No. 10

OCTUBRE de 2018

quantil | Documentos de Trabajo

ISSN 1234-5790 Edición electrónica

Approximate super-resolution and truncated moment problems in all dimensions.

Investigación Ganadora Premio MAPI1[a], Categoría Estudiante

Hernán García

Camilo Hernandez

Mauricio Junca

Mauricio Velasco

[a]Primera Conferencia Colombiana de Matemáticas Aplicadas e Industriales (2018). Sociedad Colombiana de Matemáticas.

quantil | matemáticas aplicadas

Serie Documentos de Trabajo Quantil, 2018-20
Edición electrónica.

OCTUBRE de 2018

Carrera 7 # 77 - 07. Oficina 901, Bogotá, D. C., Colombia
Teléfonos: 3718132 – (310)6791459 – (320)8461236
E-mail: info@quantil.com.co
http://www.quantil.com.co

Impreso en Colombia – Printed in Colombia

APPROXIMATE SUPER-RESOLUTION AND TRUNCATED MOMENT PROBLEMS IN ALL DIMENSIONS.

HERNÁN GARCÍA, CAMILO HERNÁNDEZ, MAURICIO JUNCA, AND MAURICIO VELASCO

Abstract. We study the problem of reconstructing a discrete measure on a compact set $K \subseteq \mathbb{R}^n$ from a finite set of moments (possibly known only approximately) via convex optimization. We give new uniqueness results, new quantitative estimates for approximate recovery and a new sum-of-squares based hierarchy for approximate super-resolution on compact semi-algebraic sets.

1. Introduction

Let $K \subseteq \mathbb{R}^n$ be a compact set and let V be a finite-dimensional vector space of continuous real-valued functions on K. If $L : V \to \mathbb{R}$ is linear and μ is a finite borel measure on K then μ *represents* L in V if $L(f) = \int_K f d\mu$ for all $f \in V$. In this article we study the *discrete reconstruction problem* which, given a representable operator L, asks us to find a discrete measure $\mu^* := \sum_{i=1}^k c_i \delta_{x_i}$ with $c_i \geq 0$ and $x_i \in K$ which represents L on V.

Under very general conditions, such measures μ^* exist (see Lemma 2.1 for details). Moreover, constructing explicit solutions μ^* is useful in a wide variety of applications, for instance:

(1) Polynomial optimization: via Lasserre's method of moments [L1] one can find a representable operator L such that every representing measure is supported on minimizers of a given multivariate polynomial.
(2) Numerical integration: any representing measure μ^* gives us a cubature rule [L2] for computing integrals of functions in V with respect to the measure μ via evaluation.
(3) Optimal control theory: optimal control problems can be reformulated as problems on occupation measures as in [LHPT]. Any discrete measure representing optima gives us explicit optimal control policies.

A celebrated approach to solve the reconstruction problem goes by the name of *superresolution* [CFG, FG] or of *Beurling minimal interpolation* [dCG, AdCG] and consists of finding a minimizer μ^* of the total variation norm in the set $\mathcal{S}(K)$ of all signed Borel measures on K. More precisely, letting $\|\mu\|_{\mathrm{TV}} := \sup \int_K g d\mu$ as g runs over all continuous functions g on K with $\|g\|_\infty \leq 1$ we want to solve the problem

$$\min_{\nu \in \mathcal{S}(K)} \|\nu\|_{\mathrm{TV}} : \forall f \in V \left(\int_K f d\nu = L(f) \right) \tag{1}$$

2000 *Mathematics Subject Classification.* Primary 15A29 Secondary 15B52,52A22.
Key words and phrases. Super-resolution, Compressed sensing, truncated moment problems.

Our first result is a uniqueness theorem for recovery of discrete measures extending a result of De Castro and Gamboa [dCG, Theorem 2.1] to positive measures in dimension greater than one. Let $V_{\leq d}$ be the set of real-valued polynomials of degree $\leq d$ in the variables $x_1, \dots, x_n$. Recall that a finite set of points $X \subseteq \mathbb{R}^n$ has an ideal $I(X)$ consisting of all polynomials vanishing on X, a generator degree $g(X)$ defined as the maximum degree of a minimal generator of $I(X)$ and an interpolation degree $i(X)$ defined as the minimum degree d such that every real-valued function on X is given by the restriction to X of a polynomial of degree at most d.

Theorem 1.1. *For an integer k, let $X \subseteq \mathbb{R}^n$ be a set of k points and let $V := V_{\leq d}$. If $d \geq \max(2g(X), i(X))$ then any discrete measure μ supported on X is the unique minimizer of (1) with $L(f) := \int_K f d\mu$. Moreover $\max(2g(X), i(X)) \leq 2(k - n + 1)$ for every set X of size k which is not contained in any hyperplane in $\mathbb{R}^n$.*

In applications one is often interested in measures whose support X is not an arbitrary set of points but rather a *generic* set of points $X = \{p_1, \dots, p_k\}$, meaning that $(p_1, \dots, p_k)$ lie in the complement of a given algebraic set of $(\mathbb{R}^n)^k$ (see Section 2.3 for details). For such sets we can strengthen the upper bound in the previous Theorem and provide a lower bound on d below which uniqueness is lost.

Theorem 1.2. *If X is a generic set of k points in $\mathbb{R}^n$ then:*

(1) *The interpolation degree $i(X)$ equals the smallest integer e for which the inequality $k \leq \binom{n+e}{e}$ holds.*
(2) *The inequality $g(X) \leq i(X) + 1$ holds. In particular, if $d \geq 2(i(X) + 1)$ then any discrete measure μ supported on X is the unique minimizer of (1) with $V := V_{\leq d}$.*
(3) *For all d such that $\binom{n+d}{d} \leq k(n + 1)$ problem (1) does not have a unique solution for every measure μ supported on X.*

For this approach to be useful we need a procedure for carrying out the optimization in (1). This can be done either via discretization as in [FG, D] or via sum-of-squares hierarchies as De Castro, Gamboa, Henrion and Lasserre propose in [DCGHL]. Since we are working in the context of reconstructing positive measures (not signed measures) one can also use a simple moment relaxation which is guaranteed to work under the hypotheses of Theorem 1.1 (see Section 3.1 for details). In Section 5 we show that this procedure works well in practice via several numerical examples in dimensions $1, 2$ and 4.

In many applications of the measure reconstruction problem, however, the moments of the measure are known only approximately. More precisely, we fix a basis $\phi_1, \dots, \phi_m$ for V and would like to recover a point measure μ from a known vector y with components given by $y_i = \int_K \phi_i d\mu + \epsilon_i$ where $\epsilon := (\epsilon_1, \dots, \epsilon_m)$ is a noise term with magnitude $\|\epsilon\|_2 \leq \delta$ bounded by a known value δ. A very significant contribution in this setting is the work of Azais, De Castro and Gamboa [AdCG] who give quantitative estimates for the error when the recovery mechanism is to solve the following *Beurling Lasso* (BLASSO) optimization problem:

$$\min_{\nu \in \mathcal{S}(K)} \|\nu\|_{\mathrm{TV}} : \left\| \left(\int_K \phi_i d\nu - y_i \right)_{i=0,\dots,m} \right\|_2 \leq \delta \tag{2}$$

Our next result uses their ideas to give a quantitative localization bound for discrete measures valid in all dimensions. If Δ is a discrete measure and $z \in K$ we will write $\Delta(z)$ to mean the coefficient of δ_z in the unique decomposition of Δ as a sum of Dirac measures. We will write $d(X,z)$ for the euclidean distance between a point z and a set X and write $N(X,\delta)$ (resp. $F(X,\delta)$) for the set of points which are at distance at most (resp. at least) δ from X.

Theorem 1.3. *Let μ be any discrete measure supported on a finite set X and let $\hat{\Delta}$ be a discrete minimizer of (2) with $V = V_{\leq d}$, $y_i := \int \phi_i d\mu + \epsilon_i$ and $\|(\epsilon_i)_i\|_2 \leq \delta$. If $d \geq 2g(X)$ and $\phi_0, \ldots, \phi_m$ are an orthonormal basis for V with respect to some probability measure on K then there exist constants $C_a > 0$ and $0 < C_b < 1$ depending only on X such that if $c_0 := \sqrt{\frac{C_b}{C_a}}$ then the following statements hold:*

(1) *If $z \in K$ is such that $\hat{\Delta}(z) > \frac{2\delta}{C_b}$ then $d(X,z) \leq c_0$.*
(2) *The following inequalities hold:*

$$\sum_{z \in N(X,c_0), \hat{\Delta}(z)>0} \hat{\Delta}(z) d(X,z)^2 \leq \frac{2\delta}{C_a}$$

$$\sum_{z \in F(X,c_0), \hat{\Delta}(z)>0} \hat{\Delta}(z) \leq \frac{2\delta}{C_b}$$

$$\sum_{z : \hat{\Delta}(z)<0} |\hat{\Delta}(z)| \leq 2\delta$$

In order to apply Theorem 1.3 we must be able to solve the (infinite-dimensional) optimization problem (2). Our next Theorem recasts (2) as a finite-dimensional convex optimization problem extending the main results of De Castro, Gamboa, Henrion and Lasserre in [DCGHL] to the approximate recovery problem.

Theorem 1.4. *The optimal value of (2) coincides with the optimal value of the following finite-dimensional convex optimization problem*

$$\sup_{(\vec{a},b) \in \mathbb{R}^n \times \mathbb{R}} \left\{ \langle \vec{a}, y \rangle - b\delta : P = \sum_{i=1}^{m} a_i \phi_i, \|P\|_\infty \leq 1, \|\vec{a}\|_2 \leq b \right\} \tag{3}$$

Next we propose a hierarchy of semidefinite programs for solving (3) when K is semialgebraic and explicitly bounded and $V \subseteq \mathbb{R}[\vec{x}] := \mathbb{R}[x_1, \ldots, x_n]$. To describe the hierarchy we will need the following basic definition. If $g_1, \ldots, g_t \in \mathbb{R}[\vec{x}]$ and $e > 0$ is an integer define the quadratic module of degree e of $g_1, \ldots, g_t$ as

$$Q_e(g_1, \ldots, g_t) = \left\{ f \in \mathbb{R}[\vec{x}] : \exists (s_i)_{i=0,1,\ldots,t} \text{ such that } f = s_0 + \sum_{i=1}^{t} g_i s_i \right\}.$$

where the $s_i \in \mathbb{R}[\vec{x}]$ are sums-of-squares of polynomials of degree bounded by e. Henceforth we let $\vec{\phi} = (\phi_1, \ldots, \phi_m)$ be the vector whose components are our chosen basis for V.

Theorem 1.5. *Suppose $K = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_t(x) \geq 0\}$ for some $g_i \in \mathbb{R}[x_1, \ldots, x_n]$ and assume there exist positive integers N, e such that $N - \|x\|_2^2 \in Q_e(g_1, \ldots, g_t)$. If α_s denotes the number*

$$\alpha_s = \sup_{(\vec{a},b) \in \mathbb{R}^n \times \mathbb{R}} \left\{ \langle \vec{a}, y \rangle - b\delta : 1 - \langle \vec{a}, \vec{\phi} \rangle, 1 + \langle \vec{a}, \vec{\phi} \rangle \in Q_s(g), \|\vec{a}\|_2 \leq b \right\}.$$

then the following statements hold:

(1) *For each s the number α_s is the optimal value of a semidefinite programming problem.*
(2) *The equality $\lim_{s\to\infty} \alpha_s = \alpha$ holds where α is the optimal value of problem (3).*

In Section 5 we use Theorem 1.5 for carrying out BLASSO minimization to recover discrete measures and show that we obtain good approximations in dimensions one and two.

Finally we propose a new application of super-resolution for finding good approximate discretizations of general probability measures on K in a sense to be defined. Fix a family of polynomials $\phi_1, \dots, \phi_m$ which is orthonormal with respect to some probability measure on K and which spans $V := V_{\leq d}$.

Definition 1.6. A (δ, k)-summary of a (not necessarily discrete) measure μ on K with respect to $\phi_1, \dots, \phi_m$ is a measure Δ with at most k-atoms for which the following inequality holds

$$\left\| \left(\int_K \phi_i d\mu - \int_K \phi_i d\Delta \right)_{i=1,\dots,m} \right\|_2 \leq \delta$$

We will assume we know the exact values of the moments of a measure μ on K and that we would like to find a (δ, k) summary (for given δ and k). The following Theorem shows that if such a summary exists then it is possible to use super-resolution to approximate it. Let $i(X)$ be the number defined in Theorem 1.2 part (1).

Theorem 1.7. *Suppose Δ is a (δ, k) summary of μ supported on a set X and let $\hat{\Delta}$ be a discrete minimizer of the problem*

$$\min_{\nu \in \mathcal{S}(K)} \|\nu\|_{\mathrm{TV}} : \left\| \left(\int_K \phi_i d\nu - \int_K \phi_i d\mu \right)_{i=0,\dots,m} \right\|_2 \leq \delta. \tag{4}$$

If $d \geq 2g(X)$ then the conclusions of Theorem 1.3 hold for $\hat{\Delta}$.

Based on the previous Theorem we propose taking the k largest coefficients of a discrete minimizer $\hat{\Delta}$ of (4) if such a minimizer exists as an algorithm for summarization. In Section 5 we present numerical examples of summarization of some measures in dimensions one and two. Our examples in dimension one show that the summarization procedure recovers good approximations of the Gauss-Chebyshev quadrature rule and suggests ways to generalize it to higher dimensions.

Acknowledgements. We wish to thank Greg Blekherman and Fabrice Gamboa for very useful conversations during the completion of this project. M. Junca and M. Velasco were partially supported by the FAPA funds from Universidad de los Andes.

2. Preliminaries

2.1. **Representability via discrete measures.** Let $K \subseteq \mathbb{R}^n$ be a compact set and let V be a finite-dimensional vector subspace of the space $C(K)$ continuous real-valued functions on K. By a *measure* on K we will always mean a positive measure. We will use the term *signed measure* to refer to these. By a discrete measure on K we mean a conic combination of Dirac delta measures supported at points of K. If ν is a finite Borel measure on K let $L_\nu : C(X) \to \mathbb{R}$ be the map given by $L_\nu(f) := \int_K f d\nu$. We say that an operator $L : V \to \mathbb{R}$ is representable by a measure if there exists a finite Borel measure ν such that $L(f) = L_\nu(f)$ for every $f \in V$. The following Lemma, due to Blekherman and Fialkow [BF], explains the key role played by discrete measures in truncated moment problems. It is a generalization of results of Tchakaloff [T] and Putinar [P]. We include a proof for the reader's benefit.

Lemma 2.1. *If the functions in V have no common zeroes on K then every linear operator $L \in V^*$ representable by a measure is representable by a discrete measure with at most $\dim(V^*) + 1$ atoms.*

Proof. Let $P \subseteq V$ be the closed convex cone of functions in V which are nonnegative at all points of K. It is immediate that $P = \operatorname{Conv}(L_{\delta_x} : x \in K)^*$. By the bi-duality Theorem from convex geometry we conclude that $P^* = \overline{\operatorname{Conv}(L_{\delta_x} : x \in K)}$. Now consider the map $\phi : K \to V^*$ sending a point x to the restriction of L_{δ_x} (i.e. to the evaluation at x). This map is continuous and therefore $S := \phi(K)$ is a compact set. Since the functions in V have no points in common the convex hull of S does not contain zero and therefore the cone of discrete measures $\operatorname{Conv}(L_{\delta_x} : x \in K)$ is closed in V^*. Let $\mathcal{M}(V) \subseteq V^*$ be the cone of operators representable by a finite borel measure. Since $\operatorname{Conv}(L_{\delta_x} : x \in K) \subseteq \mathcal{M}(V) \subseteq P^*$ we conclude that $\mathcal{M}(V)$ equals the cone of discrete measures as claimed. The bound on the number of atoms follows from Caratheodory's Theorem [B]. □

2.2. **Ideals and coordinate rings of points in projective space.** Suppose $X \subseteq \mathbb{R}^n$ is a finite set of points of size k. To be able to make arguments with graded rings we will embed X in the real projective space $\mathbb{P}^n$. For basic background on graded rings and projective space the reader should refer to [CLO, Chapter 1,2,8].

We endow $\mathbb{P}^n$ with homogeneous coordinates $[X_0 : \cdots : X_n]$ and identify $\mathbb{R}^n$ with the open subset of $\mathbb{P}^n$ where $X_0 \neq 0$ via the map $\phi(x_1, \ldots, x_n) = [1 : x_1 : \cdots : x_n]$. We identify X with its image under ϕ and define the homogeneous coordinate ring of X as $A := \mathbb{R}[X_0, \ldots, X_n]/I(X)$ where $I(X)$ is the ideal generated by all homogeneous polynomials vanishing at all points of X. Since $X \subseteq \mathbb{P}^n$, the ring A is standardly graded (i.e. $A_t := \mathbb{R}[X_0, \ldots, X_n]_t/I(X)_t$) and is generated, as an algebra over $\mathbb{R}$, by elements of degree one. Denote by $HF(A, t) := \dim_{\mathbb{R}} A_t$ the Hilbert function of A. The following Lemma summarizes some key basic facts about the homogeneous coordinate ring of a set of k points in $\mathbb{P}^n$. These are well known classical results in algebraic geometry for which we provide a self-contained elementary proof (see [E, Chapter 3] for further background on ideals of points in projective space).

Lemma 2.2. *The following statements hold:*

1. *The Hilbert function of A is strictly increasing until it attains the value k and then becomes constant.*
2. *The equality $i(X) = \min\{t : HF(A, t) = k\}$ holds and $i(X) \leq k - 1$.*
3. *The degree of every minimal homogeneous generator of $I(X)$ is bounded above by $\alpha(X) := i(X) + 1$.*

Proof. (1) Let $\ell \in A$ be a linear form which does not vanish at any point of X (for instance X_0). If $F \in A$ satisfies $\ell F = 0$ then F must vanish at all points of X and therefore $F = 0$ in A. We conclude that multiplication by ℓ, $m_\ell : A_t \to A_{t+1}$ is injective for every $t \geq 0$ proving that $HF(A,t)$ is non-decreasing. Let $B := A/(\ell)$ and note that for every t we have $B_t = 0$ if and only if $HF(A,t) = HF(A,t-1)$. Since B is generated in degree one the equality $B_t = 0$ for some t implies that $B_r = 0$ for all $r \geq t$. We conclude that if t satisfies $HF(A,t) = HF(A,t-1)$ then the Hilbert function becomes constant after t, proving (1). (2) For any $t \in \mathbb{N}$ consider the linear map $\phi^* : A_t \to \mathrm{Fun}(X, \mathbb{R})$ which maps $F(X_0, \dots, X_n)$ to the polynomial function $F(1, x_1, \dots, x_n)$ of degree at most t on X. This map is always injective and is therefore surjective whenever the dimension of A_t equals the dimension k of the space of all real-valued functions on X. To prove the inequality note that $HF(A,0) = 1$ and that it increases strictly at every stage so $HF(A,k-1) \geq k$ so $i(X) \leq k-1$.

(3) Let J be the ideal generated by $I(X)_{\leq \alpha(X)}$ and let $S := \mathbb{R}[X_0, \dots, X_n]$. Since $J \subseteq I(X)$ there is a surjective homomorphism $A' := S/J \to A$ and we will show that it is an isomorphism by proving that $\dim A'_t = \dim A_t$ for all t. Define $D := A'/(\ell)$ and note that it satisfies $D_j = B_j$ for $j \leq \alpha(X)$ and in particular $D_{\alpha(X)} = 0$. Since D is generated in degree one this implies that $D_q = 0$ for all $q \geq \alpha(X)$ and therefore multiplication by ℓ is surjective on A' in all components $t \geq \alpha(X)$. We conclude that $\dim A'_t \geq \dim A'_{t+1}$ for $t \geq \alpha(X)$ and in particular $k \geq \dim A'_s$ for $s \geq \alpha(X)$. By surjectivity of $A' \to A$ we know that $\dim A'_t \geq \dim A_t = k$ for $t \geq \alpha(X)$. Putting both inequalities together we conclude that $\dim A'_t = \dim A_t$ for all t as claimed.

□

Remark 2.3. The number $\alpha(X)$ is the Castelnuovo-Mumford regularity of X, the key measure of the (cohomological) complexity of algebraic varieties [M] (see [E, Chapter 4] for details).

2.3. **Generic points.** A property of k-tuples of points $(p_1, \dots, p_k) \in (\mathbb{P}^n)^k$ holds generically if the locus of points $(p_1, \dots, p_k)$ which satisfy it contains a nonempty Zariski open set. Equivalently, the set of points where the property fails is contained in a proper Zariski closed subset of $(\mathbb{P}^n)^k$ (i.e. one defined by homogeneous polynomial equations). Following common terminology we say that *a generic set of points X of size k satisfies a property Q* to mean that property Q holds generically. If $p_1, \dots, p_k$ are an independent sample of points in $\mathbb{R}^n$ sampled from a distribution which has a density with respect to the Lebesgue measure then $p_1, \dots, p_k$ satisfies every generic property with probability one (because every proper Zariski closed set has empty interior and in particular null Lebesgue measure). Understanding generic properties should therefore be of much interest for applications.

3. A basic uniqueness result for exact super-resolution.

Proof of Theorem 1.1. Let $\mu := \sum_{x \in X} c_x \delta_x$ for some real coefficients $c_x \geq 0$ and let $h_1, \dots, h_k$ be a set of generators of the ideal $I(X)$ of polynomials vanishing on X. Define $H := \sum h_i^2$ and $M := \sup_{x \in K} H(x)$. By our assumption on d the polynomial $P := 1 - \frac{H}{M}$ belongs to $V_{\leq d}$. We will show that P is a dual certificate in the sense of Candés, Romberg and Tao [CRT].

By construction we know that $\|P\|_\infty = 1$ and that $P(z) = 1$ if and only if $z \in X$. If Δ is a feasible solution of (1) then

$$\|\mu\|_{\mathrm{TV}} = \mu(X) = \int_K P d\mu = \int_K P d\Delta \leq \|\Delta\|_{\mathrm{TV}}$$

and therefore any optimal solution Δ of (1) satisfies $\|\mu\|_{\mathrm{TV}} = \|\Delta\|_{\mathrm{TV}}$. For Δ an optimal solution of (1) we write $\Delta = \Delta_X + \Delta_X^{\perp}$ where Δ_X is supported on X and $\Delta_X^{\perp}$ in $K \setminus X$. Since $|P(z)| < 1$ outside X we conclude that $\int_K P d\Delta_X^{\perp} < \|\Delta_X^{\perp}\|_{\mathrm{TV}}$ if $\Delta_X^{\perp} \neq 0$. It follows that

$$\|\Delta\|_{\mathrm{TV}} = \int_K P d\Delta < \|\Delta_X\|_{\mathrm{TV}} + \|\Delta_X^{\perp}\|_{\mathrm{TV}} = \|\Delta\|_{\mathrm{TV}}$$

a contradiction so $\Delta_X^{\perp} = 0$ and every minimizer Δ is supported on X. Since $d \geq i(X)$ there exists, for each point $x \in X$ a polynomial q_x in $V_{\leq d}$ which takes value one in x and zero at all other points of X. Since $\int_K q_x d\mu = \int_K q_x d\Delta$ we conclude that $\Delta = \mu$ as claimed. By Lemma 2.2 part (3) we know that $\alpha(X) = 1 + i(X)$ satisfies $g(X) \leq \alpha(X)$ and therefore $\max(2g(X), i(X)) \leq 2\alpha(X)$. If X is not contained in any hyperplane then $\dim(A_1) = n + 1$ and therefore by Lemma 2.2 part (1) $\dim(A_t) \geq n + t$ for all $1 \leq t \leq i(X)$ and we conclude that $i(X) \leq k - n$ so $\alpha(X) \leq k - n + 1$, proving the claim. □

Remark 3.1. If $X \subseteq \mathbb{R}$ consists of k points then it is immediate that $g(X) = k$ and $i(X) = k + 1$ so our Theorem proves uniqueness for $d \geq 2k$, giving another proof of [dCG, Proposition 2.3]. This upper bound is sharp since the dimension of the space of discrete measures supported at k points of $\mathbb{R}$ is $2k$ (one for the location and another one for the coefficient for each support point). An identical sharp bound is obtained for any set of points $X \subseteq \mathbb{R}^2$ which lie on a conic and in particular for any set of at most 5 points in $\mathbb{R}^2$.

Proof of Theorem 1.2. (1) The Hilbert function of the homogeneous coordinate ring A of a generic set of k points in $\mathbb{P}^n$ is given by

$$HF(A,t) = \min\left(\binom{n+t}{t}, k\right)$$

which agrees with k for the smallest e with $\binom{n+e}{e} \geq k$ as claimed.(2) By Lemma 2.2 we know that $\alpha(X) = i(X) + 1$ and $\max(2g(X), i(X)) \leq 2\alpha(X)$ proving the claim. (3) Suppose the discrete measure μ is supported on X and is the unique minimizer of (1). It follows from strong duality [DCGHL, Lemma 1] that there exists a polynomial G^* which is optimal for the dual optimization problem

$$\max_{G \in V_{\leq d}} L(G)\colon \|G\|_\infty \leq 1 \text{ on } K$$

where $L : V_{\leq d} \to \mathbb{R}$ is defined by $L(G) := \int_K G d\mu$. and in particular that there exists a nonnegative polynomial $F := 1 - G^*$ of degree d vanishing at all points of X. Since F is nonnegative we conclude that F is singular (i.e. has vanishing gradient) at all points of X. Since X consists of generic points, vanishing with multiplicty at least two at all of them imposes $k(n+1)$ independent conditions and therefore such F do not exist whenever $\binom{n+d}{d} \leq k(n+1)$ as claimed. □

Remark 3.2. The maximum in the quantity $\max(2g(X), i(X))$ of Theorem 1.1 can be achieved in either side as the following examples show. If X is a complete intersection of n quadrics then $g(X) = 2$ and $i(X) = n - 1$ so the interpolation degree is larger than $4 = 2 \cdot 2$ for $n \geq 5$. If X is a generic set of $\binom{d+n}{n}$ points in $\mathbb{P}^n$ then $g(X) = i(X) = d + 1$ so $2g(X) > i(X)$. The second example shows that the inequality $\max(2g(X), i(X)) \leq 2\alpha(X)$ is sharp since it becomes an equality for this choice of points.

Remark 3.3. The degrees of all minimal generators, and more generally the structure of the minimal free resolutions of ideals of points in $\mathbb{P}^2$ are well understood (See [E, Chapter 3] for details). By contrast the minimal free resolution of even generic sets of points s in $\mathbb{P}^n$ is widely open. The conjectural answer suggested by Lorenzini [L] was later disproved in celebrated work by Eisenbud and Popescu [EP].

3.1. **A moments relaxation.** Mirroring the proof of Theorem 1.1 one can use the following problem of moments reconstruction procedure for recovering μ given its moments operator $L : V_{\leq 2d} \to \mathbb{R}$ with $L(g) := \int_K g d\mu$ on polynomials of degree at most $2d$.

(1) Finding the *support* of μ by constructing a minimizer H^* of the optimization problem $\min_H L(H)$ where H runs over the sums-of-squares of elements of $V_{\leq d}$. More explicitly if $\vec{\phi}$ is a basis for $V_{\leq d}$ then we find H^* by solving the semidefinite programming problem:

$$\min L\left(\vec{\phi}^t A\vec{\phi}\right) \text{ s.t. } A \succeq 0 \text{ and } \operatorname{tr}(A) = 1.$$

and find the support of μ by finding the zeroes of H^* in K.

(2) Finding the *coefficients* of μ by linear algebra. If $z_1, \dots, z_k$ are the zeroes of H^* we find the coefficients $c_1, \dots, c_k$ by solving the linear equations $\sum_{i=1}^k c_i f(z_i) = L(f)$ for $f \in V$.

Theorem 1.1 guarantees that the procedure works for sufficiently high d and Theorem 1.2 gives a better guarantee for measures supported on generic points. We finish the Section with two remarks about the above procedure:

(1) If H^* is any minimizer in the relative interior of the face $L(H) = 0$ of convex cone Q of the sums-of-squares of elements in $V_{\leq g(X)}$ has X as its only real zeroes since otherwise evaluation at any additional zero would define a proper face of Q containing an interior point and hence all of Q. In particular the kernel of this evaluation would contain the dual certificate constructed in the proof of Theorem 1.1 all of whose real zeroes lie on X deriving a contradiction. As a result, interior point numerical methods for solving the SDP would produce optima H^* with X as its set of zeroes, as can be seen in our numerical examples in Section 5.

(2) If X is a generic set of $k = \binom{e+n}{n}$ points in $\mathbb{P}^n$ then $g(X) = i(X) = e+1$ and Theorem 1.2 shows that there is unique recovery when $d \geq 2(e+1)$. We claim that, if the recovery is carried out with the sum-of-squares procedure above then this bound is sharp in the sense that the recovery would fail for $d < 2(e+1)$. The reason is that every sum of squares $H = \sum P_i^2$ which vanishes at the points would have summands P_i of degree less than $e+1$ and therefore be identically zero on X because $I(X)$ contains no forms of degree less than $e+1$.

4. Approximate recovery

In this section we focus on the problem of approximate recovery. The following key property was proposed by Azais, De Castro and Gamboa as central for BLASSO quantitative localization results. We modify their definition slightly since our interest is only measures and not signed measures.

Definition 4.1. (Quadratic isolation condition)[AdCG, Definition 2.2] A finite set $X \subseteq K$ satisfies a quadratic isolation condition with parameters $C_a > 0$ and

$0 < C_b < 1$ respect to V if there exists $P \in V$ satisfying $\|P\|_\infty \leq 1$ on K, $P \equiv 1$ on X and such that the following inequality holds

$$\forall z \in K \left(P(z) \leq 1 - \min\{C_a d(z,X)^2, C_b\}\right)$$

Lemma 4.2. *If $d \geq 2g(X)$ then X satisfies a quadratic isolation condition on $V_{\leq d}$.*

Proof. Let $h_1, \dots, h_k$ be a set of minimal generators of the ideal $I(X)$ of polynomials vanishing on X and define $H := \sum h_i^2$ and $M := \sup_{x \in K} H(x)$. By our assumption on d the polynomial $P := 1 - \frac{H}{M}$ is nonnegative, belongs to $V_{\leq d}$ and is identical to one on X. Since $h_1, \dots, h_k$ are generators of the ideal $I(X)$ and a reduced set of points is nonsingular the differential of the map $\mathcal{H} : \mathbb{R}^n \to \mathbb{R}^k$ given by $\mathcal{H}(x) = (h_1(x), \dots, h_k(x))$ has trivial kernel at every $x \in X$. As a result, the Hessian at $x \in X$ of the polynomial H is strictly positive definite and in particular there exist positive real numbers η_x and $(C_a)_x$ such that $1 - P(z) \leq (C_a)_x \|z - x\|^2$ for z with $\|z - x\| \leq \eta_x$. Define $\delta = \min_{x \in X} \eta_x$, $C_a := \min_{x \in X} (C_a)_x$ and let $C_b = \sup_{z : d(z,X) \geq \frac{\delta}{2}} (1 - P(z))$. We conclude that X satisfies a quadratic isolation condition with parameters C_a and C_b with $0 < C_b < 1$ and $C_a > 0$. □

The following Lemma, of interest in its own right, extracts the essence of [AdCG, Theorem 2.1]. Note that we do not require the basis ϕ_i to be orthonormal.

Lemma 4.3. *Suppose X satisfies a quadratic isolation condition on V with witness P. If $\hat{\Delta}$ is a minimizer of (2) then the inequalities*

$$0 \leq \|\hat{\Delta}\|_{\mathrm{TV}} - \int_K P d\hat{\Delta} \leq 2\delta \|a\|_2$$

hold, where a is the vector of coefficients of P in the basis ϕ_i of V.

Proof. Since the measure μ is feasible for (2) we know that $\|\mu\|_{\mathrm{TV}} \geq \|\hat{\Delta}\|_{\mathrm{TV}}$. Since P satisfies $\|P\|_\infty \leq 1$ on K we know that

$$\|\hat{\Delta}\|_{\mathrm{TV}} \geq \int_K P d\hat{\Delta} = \sum a_i \int_K \phi_i d\hat{\Delta} = \sum a_i \left(r_i + \int_K \phi_i d\mu + \epsilon_i\right)$$

where $r_i := \int_K \phi_i d\hat{\Delta} - y_i$ and $y_i = \int_K \phi_i d\mu + \epsilon_i$. Since $\sum a_i \int_K \phi_i d\mu = \int_K P d\mu = \|\mu\|_{\mathrm{TV}}$ we conclude that

$$\|\mu\|_{\mathrm{TV}} + \sum a_i(r_i + \epsilon_i) \leq \int_K P d\hat{\Delta} \leq \|\hat{\Delta}\|_{\mathrm{TV}} \leq \|\mu\|_{\mathrm{TV}}$$

So the difference between the last and first terms is an upper bound for the difference between the interior terms yielding

$$0 \leq \|\hat{\Delta}\|_{\mathrm{TV}} - \int_K P d\hat{\Delta} \leq \left|\sum a_i(r_i + \epsilon_i)\right| \leq \|a\|_2 \left(\|r\|_2 + \|\epsilon\|_2\right) \leq 2\delta \|a\|_2$$

where the last two inequalities follow from the Cauchy-Schwartz and triangle inequalities. □

We are now in a position to prove our main result on approximate recovery

Proof of Theorem 1.3. By Lemma 4.2 the set X satisfies a quadratic isolation condition with parameters $C_a > 0$ and $0 < C_b < 1$. Let P be the witness constructed in the Lemma. Since the basis ϕ is assumed to be orthonormal with respect to a probability measure on K, Parseval's equality shows that the vector a of coefficients of P in the basis ϕ_i satisfies $\|a\|_2^2$ agrees with the L_2-norm of P with respect to the probability measure that makes the ϕ_i orthonormal. It follows that this norm is bounded by one since $\|P\|_\infty \leq 1$. By Lemma 4.3 we know that any minimizer $\hat{\Delta}$ of problem (2) satisfies

$$0 \leq \|\hat{\Delta}\|_{\mathrm{TV}} - \int_K P d\hat{\Delta} \leq 2\delta\|a\|_2 \leq 2\delta$$

We compute the quantity in the middle with $\hat{\Delta} = \sum_{z\in K}\hat{\Delta}(z)\delta_z$. Separating the coefficients into three sets, negative coefficients, and two sets of positive coefficients according to which of the two terms achieves the minimum in $\min\{C_a d(z,X)^2, C_b\}$ we obtain, since $P \geq 0$, the inequality

$$2\delta \geq \sum_{\hat{\Delta}(z)<0} |\hat{\Delta}(z)| + \sum_{\hat{\Delta}(z)>0, d(z,X)^2 \leq \frac{C_b}{C_a}} \hat{\Delta}(z) C_a d(z,X)^2 + \sum_{\hat{\Delta}(z)>0, d(z,X)^2 > \frac{C_b}{C_a}} \hat{\Delta}(z) C_b$$

from which the three inequalities in part (2) of the Theorem follow immediately. □

4.1. **An algorithm for approximate super-resolution.** In this section we focus on solving problem (2). We begin by proving Theorem 1.4 which reformulates (2) as a finite-dimensional convex optimization problem amenable to computation whenever (2) has a discrete minimizer.

Proof of Theorem 1.4. During the proof we will identify problem (3) with the dual of (2) and prove that there is no duality gap. To do this we first reformulate (2) as a primal problem in standard form (as in [B, Section 7.1]). Recall that a signed Radon measure ν admits a unique Hahn decomposition as a difference of Radon measures ν_+ and ν_- and that in this decomposition the total variation is given by $\|\nu\|_{\mathrm{TV}} = \nu_-(K) + \nu_+(K)$ which is a linear function in ν_+, ν_-. The ambient vector space of our primal optimization problem will be $E = C(K)^* \times C(K)^* \times \mathbb{R}^m \times \mathbb{R}$ endowed with the weak $*$-topology. We will denote its elements by 4-tuples $(\nu_-, \nu_+, \vec{z}, w)$. Define the convex cone

$$D := \{(\nu_-, \nu_+, \vec{z}, w) : \|\vec{z}\|_2 \leq w \text{ and } \nu_+, \nu_- \in R(K)_+\}$$

where $R(K)_+$ denotes the cone of positive radon measures on K. The continuous dual of E, denoted E^* is given by $E^* := C(K) \times C(K) \times \mathbb{R}^m \times \mathbb{R}$ and we will write its elements as 4-tuples $(f_1, f_2, \vec{a}, b)$. In this notation the dual cone $D^* \subseteq E^*$ is given by:

$$D^* := \{(f_1, f_2, \vec{a}, b) : \|\vec{a}\|_2 \leq b \text{ and } f_1, f_2 \geq 0 \text{ on } K\}.$$

To simplify the notation we will write $\int_K f d\nu := \langle f, \nu\rangle$. Define the continuous linear map $A : E \to \mathbb{R}^m \times \mathbb{R}$ by the formula

$$A(\nu_-, \nu_+, \vec{z}, w) = \left((\langle \phi_i, \nu_+ - \nu_-\rangle - z_i)_{i=1,\dots,m}, w \right)$$

and note that problem (2) is equivalent to

$$\min_{(\nu_-,\nu_+,\vec{z},w)\in D} \langle 1, \nu_+ + \nu_-\rangle \text{ s.t. } A(\nu_-, \nu_+, \vec{z}, w) = (\vec{y}, \delta)$$

its dual problem is therefore given by (see [B, Section 7.1]

$$\sup_{(f_1,f_2,\vec{a},b)} \langle \vec{a}, y\rangle + \delta b \text{ s.t. } (1,1,0,0) - A^*(\vec{a}, b) \in D^*.$$

By definition of adjoint we have $A^*(\vec{a},b) = \left(\langle \vec{a},\vec{\phi}\rangle, -\langle \vec{a},\vec{\phi}\rangle, -\vec{a}, b\right)$ so the dual is equivalent to (3) after the change of variables $b \to -b$. To prove the Theorem we will show that there is no duality gap. Since the objective function is nonnegative and the domain of the problem is nonempty (because its feasible set contains the measure μ which we would like to recover) by [B, Theorem 7.1] it suffices to prove that $\hat{A}(D) \subseteq \mathbb{R}^{m+2}$ is closed where $\hat{A}: E \to \mathbb{R}\times\mathbb{R}^m\times\mathbb{R}$ is given by

$$\hat{A}(\nu_-,\nu_+,\vec{z},w) = \left(\langle 1, \nu_+ + \nu_-\rangle, A(\nu_-,\nu_+,\vec{z},w)\right).$$

Assume $\beta_j = (\nu_-^j, \nu_+^j, \vec{z}^j, w^j)$ is a sequence in D for which $\hat{A}(\beta_j)$ converges to $s \in \mathbb{R}^{m+2}$ as $j \to \infty$. We will show that there exists $\beta \in D$ such that $\hat{A}(\beta) = s$. Since $(\hat{A}(\beta_j))_j$ is a convergent sequence in $\mathbb{R}^{m+2}$ it is bounded and therefore both the total variation of the $\nu_\pm^j$ and the w^j which are the first and last components of the map $\hat{A}$ are bounded. By the Theorem of Banach-Alaoglu we know that balls in $C(K)^*$ are compact in the weak $*$ topology and therefore conclude that the points β_j lie in a compact subset of the closed cone $D \subseteq E$. As a result there is a subsequence $(\beta_{j_n})_n$ converging to a point $\beta \in D$. Since $\hat{A}$ is a continuous linear map we conclude that $\hat{A}(\beta) = s$ as claimed.

□

Remark 4.4. If we think of a signed measure as a linear operator $L \in V^*$ in the ellipsoid $\mathcal{E}$ defined by $\|(L(\phi_i) - y_i)_{i=1\dots m}\|_2 \le \delta$ then the quantity $-\|\vec{a}\|_2\delta + \sum_{i=1}^m a_i y_i$ equals $\inf_{L\in\mathcal{E}} L(P)$ where $P := \sum a_i\phi_i$ and the optimization problem above can be thought of as solving

$$\sup_{P:\|P\|_\infty\le 1, P\in V}\left(\inf_{L\in\mathcal{E}} L(P)\right)$$

This suggests a methodology for recovering an optimizer measure, given an optimal solution $(\vec{a}^*, \alpha^*)$ of (3), namely:

1. Define $P^* := \sum a_i^*\phi_i$ and find an operator L^* which is a minimizer of the second-order cone optimization problem $\inf_{L\in\mathcal{E}} L(P^*)$.
2. The values $L^*(\phi_i)$ are the moments of a measure which we can try to recover via exact superresolution as in the previous section. The moments of this measure are contained in $\mathcal{E}$ and $L^*(P^*) = \alpha^*$ so the measure has total variation α^* and is therefore a minimizer of (2).

Next we prove Theorem 1.5 which gives a semidefinite programming hierarchy for solving (3) on explicitly bounded semialgebraic sets.

Proof of Theorem 1.5. (1) A polynomial h is a sum-of-squares of polynomials of degree at most e iff there exists a PSD matrix A such that $h = \vec{m}^t A\vec{m}$ where $\vec{m}$ is the vector of monomials of degree at most e. It follows that $Q_s(g)$ is an SDR set (a linear projection of a spectrahedron) for any $s > 0$. We conclude that the set

$$\left\{(\vec{a},b) : 1 - \langle\vec{a},\vec{\phi}\rangle, 1 + \langle\vec{a},\vec{\phi}\rangle \in Q_s(g), \|\vec{a}\|_2 \le b\right\}$$

is also SDR since it is an intersection of two affine slices of SDR sets and a second-order cone constraint. Since the function $\langle\vec{a},y\rangle - b\delta$ is linear on $(\vec{a},b)$ we conclude that α_s is the optimal value of a semidefinite programming problem as claimed.

(2) Suppose that $(\vec{a}^*, b^*)$ is an optimal solution of (3). For $\epsilon > 0$ let $\vec{a}' := (1-\epsilon)\vec{a}^*$ and $b' := (1-\epsilon)b$. It is immediate that $1 - \langle\vec{a}',\vec{\phi}\rangle > 0$ and $1 + \langle\vec{a}',\vec{\phi}\rangle > 0$. Since K is

explicitly bounded Putinar's Theorem [P] implies that there exists an integer $e > 0$ such that $1 - \langle \vec{a}', \vec{\phi} \rangle, 1 + \langle \vec{a}', \vec{\phi} \rangle \in Q_e(g)$ and therefore α_s is at least the optimal value at $(\vec{a}', b)$, that is $(1-\epsilon)\alpha$. We conclude that $(1-\epsilon)\alpha \leq \alpha_e \leq \alpha$ proving the claim since $\epsilon > 0$ was arbitrary.

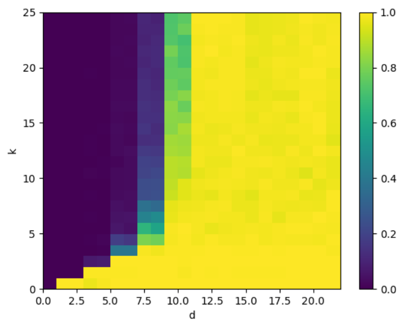
□

We are now in a position to prove the summarization Theorem 1.7.

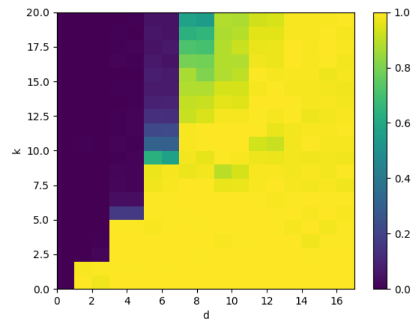
Proof of Theorem 1.7. Suppose Δ is a (δ, k) summary of μ and let $X := \mathrm{supp}(\Delta)$ Since the moments depend continuously on the location of the points we can assume, by slightly perturbing the support of Δ, if necessary, that X is a generic set of points. If we define $y_i := \int_K \phi_i d\mu$ then $y_i = \int_K \phi_i d\Delta + \epsilon_i$ with $\|(\epsilon_i)_i\|_2 \leq \delta$. Since $d \geq \max(2g(X), i(X))$ the claim follows from Theorem 1.3. □

5. Numerical Experiments

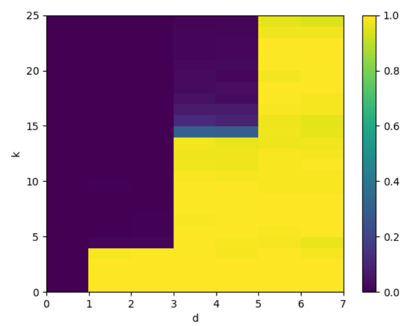
5.1. **Exact Recovery.** In this section we use the SDP procedure outlined in Section 3.1 to recover discrete measures in $K := [-1,1]^n$, for $n = 1,2,4$ with $V = V_{\leq d}$. The goal is to record the behavior of the algorithm as d and k vary for measures supported on generic points. For each pair (k,d) we generate 100 uniform discrete measures $\Delta_j = \sum_{i=1}^k \frac{1}{k}\delta_{x_i^j}$ with support $S_j := \{x_1^j, ..., x_k^j\}$ in $[-1,1]^n$ chosen uniformly at random. For each j we compute the moments with respect to the standard monomial basis of $V_{\leq d}$. To quantify the quality of the recovery we evaluate the function $q := H^*$ at the points x_i^j and report the proportion of points where this quantity is very close to zero. Figure 1 reports the average of these proportions over the 100 simulations. Figure 2 shows the function H^* for degrees $d = 2,3,4$ where Δ is a counting measure supported at four points in $[-1,1]$. Figure 3 shows the heatmap of the function $\log(H^*)$, for degrees $d = 1,2,4,6$ where Δ is a counting measure supported in four points on $K = [0,1] \times [0,1]$. As expected, location accuracy increases with degree.



(A) Recovery in dimension 1

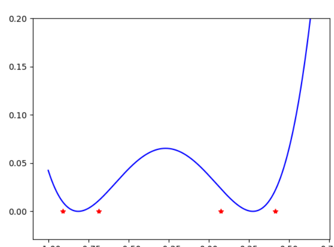


(B) Recovery in dimension 2

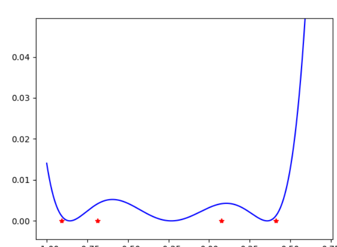


(C) Recovery in dimension 4

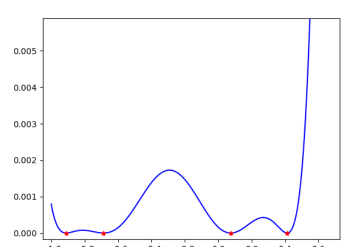
FIGURE 1



(A) d=2

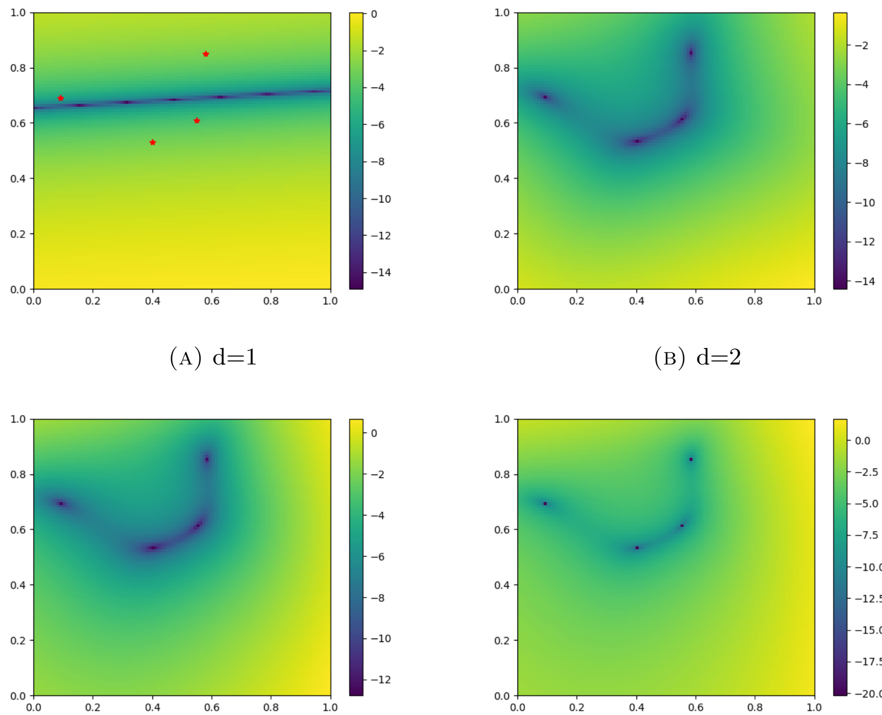


(B) d=3



(C) d=4

FIGURE 2. Polynomial H^* associated to the counting measure on 4 points for different values of degree d via the recovery procedure.



(A) d=1

(B) d=2

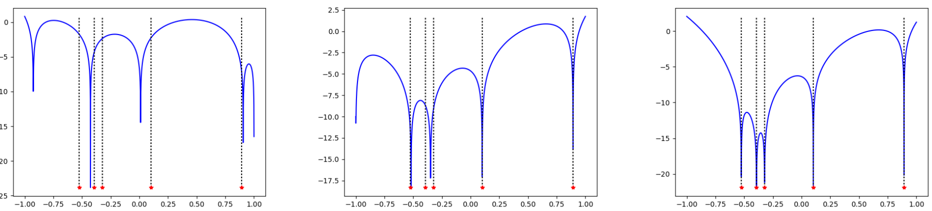
(C) d=4

(D) d=6

FIGURE 3. logarithm of H^* when μ is a counting measure supported in four points and different values of the degree d.

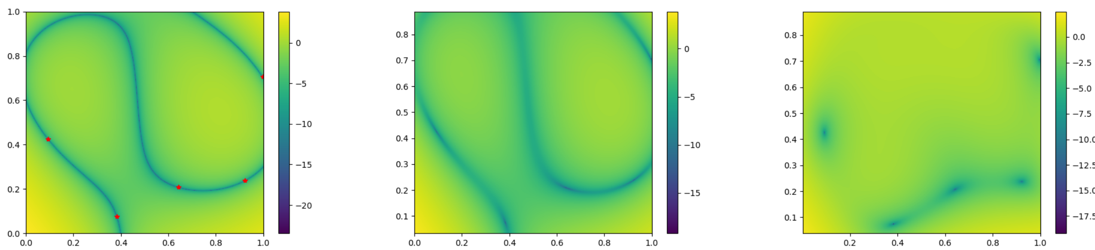
5.2. **Approximate Recovery.** We let μ be the counting measure supported on the five red points of Figures 4 and 5 (in dimensions one and two respectively). Noisy measurements $y'_j = \int \Phi_j d\mu + \epsilon_i$ are generated, where ϵ_i is a sample with distribution $\mathcal{N}(0, \epsilon)$ and $\{\Phi_1, ..., \Phi_m\}$ is the ortonormalization of the monomial basis of $V_{\leq d}$ with respect to the inner product given by the Lebesgue measure in $[-1, 1]$ and $[0, 1]^2$ for $d = 11$ and $d = 6$, respectively in dimension 1 and 2. We choose $\delta = \|(\epsilon_i)_i\|_2$ and use the hierarchy defined in 1.5 with $e = d$.

5.3. **Measure summarization.** Applying Theorem 1.7 to the Lebesgue measure on the interval $[-1, 1]$ with $V = V_{\leq d}$, we obtain a very good approximation of the $d - th$ Gauss-Legendre nodes as local minima of the optimal polynomial H^*. This is illustrated in Figure 6. The vertical lines correspond to the location of the Gauss-Legendre nodes.



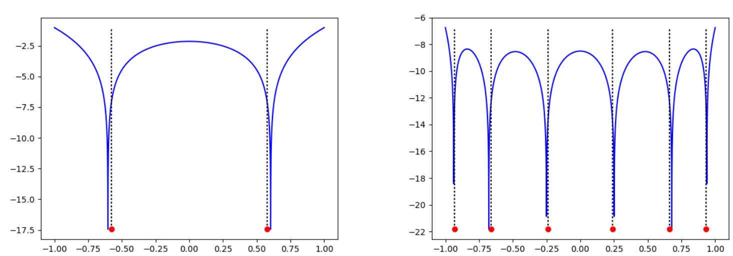
(A) $\epsilon = 10^{-1}$ (B) $\epsilon = 10^{-3}$ (C) $\epsilon = 10^{-5}$

FIGURE 4. Logarithms of optimal polynomials H^* with $d = 11$ for noisy measurements and varying ϵ.

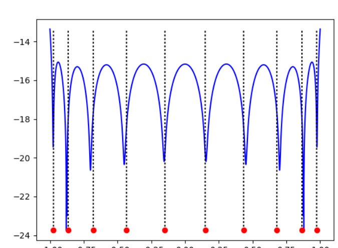


(A) $\epsilon = 10^{-3}$ (B) $\epsilon = 10^{-5}$ (C) $\epsilon = 10^{-7}$

FIGURE 5. Logarithms of optimal polynomials H^* with $d = 6$ for noisy measurements and varying ϵ.

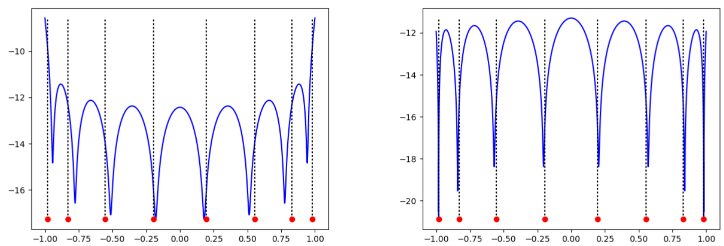


(A) $d = 2$ (B) $d = 6$



(C) $d = 10$

FIGURE 6. Summarization of the uniform measure in $[-1, 1]$



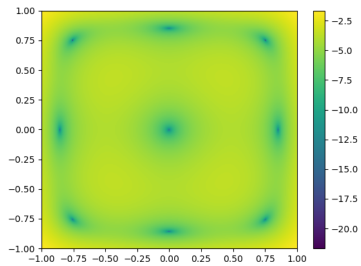
(A) Density $w_1(x)$ (B) Density $w_2(x)$

FIGURE 7

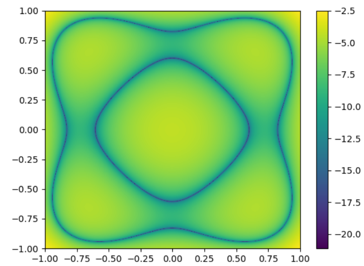
Similarly we use Theorem 1.7 to obtain discrete approximations to the measures in $[0,1]$ given by the densities $w_1(x) := \sqrt{1-x^2}$ and $w_2(x) := \frac{1}{\sqrt{1-x^2}}$. The results are shown in Figure 7. The recovered measures turn out to be supported on a set very close to the roots of Chebyshev polynomials (marked in red) of degree d, which are known to lead to the best interpolation formulas [K, Section 6.1].

Finally, in Figure 8 we apply Theorem 1.7 to the Lebesgue measure over the square $[-1,1]^2$ with $V = V_{\leq d}$ for $d = 3,4,5$. Note that when $d = 3,5$ the obtained summary is not the product measure of the one-dimensional summaries since its support contains $(0,0)$ (compare with Figure 6). When $d = 4$ the algorithm finds an H^* with infinitely many real zeroes and is therefore unable to locate the support of a discrete summary. It would be interesting to find criteria which guarantee that problem (4) has discrete minimizers.

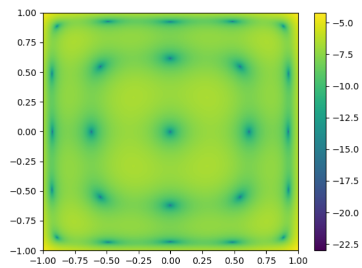
All computations in this section were made with the Julia programming language [BJEKS] using the specialized solver [MA] and the JuMP modeling language [IDaJHaML]. The code used to generate the examples in this section is freely available at `https://github.com/hernan1992garcia/super_resolution_recovery`.



(A) Nodes for $d = 3$.



(B) Nodes for $d = 4$.



(C) Nodes for $d = 5$.

Figure 8

References

[AdCG] Jean-Marc Azaïs, Yohann de Castro, and Fabrice Gamboa, *Spike detection from inaccurate samplings*, Appl. Comput. Harmon. Anal. **38** (2015), no. 2, 177–195, DOI 10.1016/j.acha.2014.03.004. MR3303671

[B] Alexander Barvinok, *A course in convexity*, Graduate Studies in Mathematics, vol. 54, American Mathematical Society, Providence, RI, 2002. MR1940576

[CLO] David A. Cox, John Little, and Donal O'Shea, *Ideals, varieties, and algorithms*, 4th ed., Undergraduate Texts in Mathematics, Springer, Cham, 2015. An introduction to computational algebraic geometry and commutative algebra. MR3330490

[BF] Greg Blekherman and Lawrence Fialkow, *The core variety and representing measures in the truncated moment problem*, https://arxiv.org/abs/1804.04276.

[CRT] Emmanuel J. Candès, Justin Romberg, and Terence Tao, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory **52** (2006), no. 2, 489–509, DOI 10.1109/TIT.2005.862083. MR2236170

[CFG] Emmanuel J. Candès and Carlos Fernandez-Granda, *Towards a mathematical theory of super-resolution*, Comm. Pure Appl. Math. **67** (2014), no. 6, 906–956, DOI 10.1002/cpa.21455. MR3193963

[CRT] Emmanuel J. Candès, Justin K. Romberg, and Terence Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math. **59** (2006), no. 8, 1207–1223, DOI 10.1002/cpa.20124. MR2230846 (2007f:94007)

[CT] Emmanuel J. Candes and Terence Tao, *Decoding by linear programming*, IEEE Trans. Inform. Theory **51** (2005), no. 12, 4203–4215, DOI 10.1109/TIT.2005.858979. MR2243152 (2007b:94313)

[dCG] Yohann de Castro and Fabrice Gamboa, *Exact reconstruction using Beurling minimal extrapolation*, J. Math. Anal. Appl. **395** (2012), no. 1, 336–354, DOI 10.1016/j.jmaa.2012.05.011. MR2943626

[DCGHL] Yohann De Castro, F. Gamboa, Didier Henrion, and J.-B. Lasserre, *Exact solutions to super resolution on semi-algebraic domains in higher dimensions*, IEEE Trans. Inform. Theory **63** (2017), no. 1, 621–630, DOI 10.1109/TIT.2016.2619368. MR3599963

[D] David L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory **52** (2006), no. 4, 1289–1306, DOI 10.1109/TIT.2006.871582. MR2241189

[E] David Eisenbud, *The geometry of syzygies*, Graduate Texts in Mathematics, vol. 229, Springer-Verlag, New York, 2005. A second course in commutative algebra and algebraic geometry. MR2103875

[FG] Carlos Fernandez-Granda, *Super-resolution of point sources via convex programming*, Inf. Inference **5** (2016), no. 3, 251–303, DOI 10.1093/imaiai/iaw005. MR3551798

[BJEKS] Bezanson J., A. Edelman, S. Karpinski, and V. Shah, *Julia: A Fresh Approach to Numerical Computing*, SIAM Review **59** (2017), no. 1, 65–98, DOI 10.1137/141000671.

[IDaJHaML] Iain Dunning and Joey Huchette and Miles Lubin, *JuMP: A Modeling Language for Mathematical Optimization*, SIAM Review **59** (2017), no. 2, 295-320, DOI 10.1137/15M1020575.

[K] D. and Cheney Kincaid W., *Numerical Analysis: Mathematics of Scientific Computing, Third Edition* (2009).

[FG] Carlos Fernandez-Granda, *Super-resolution of point sources via convex programming*, Inf. Inference **5** (2016), no. 3, 251–303, DOI 10.1093/imaiai/iaw005. MR3551798

[DHL] I. Dunning, J. Huchette, and M. Lubin, *JuMP: A Modeling Language for Mathematical Optimization*, SIAM Review **59** (2017), no. 2, 295-320, DOI 10.1137/15M1020575.

[L] Anna Lorenzini, *The minimal resolution conjecture*, J. Algebra **156** (1993), no. 1, 5–35, DOI 10.1006/jabr.1993.1060. MR1213782

[E] David Eisenbud, *Commutative algebra*, Graduate Texts in Mathematics, vol. 150, Springer-Verlag, New York, 1995. With a view toward algebraic geometry. MR1322960

[EP] David Eisenbud and Sorin Popescu, *Gale duality and free resolutions of ideals of points*, Invent. Math. **136** (1999), no. 2, 419–449, DOI 10.1007/s002220050315. MR1688433

[L1] Jean Bernard Lasserre, *Moments, positive polynomials and their applications*, Imperial College Press Optimization Series, vol. 1, Imperial College Press, London, 2010. MR2589247

[L2] Jean B. Lasserre, *The existence of Gaussian cubature formulas*, J. Approx. Theory **164** (2012), no. 5, 572–585, DOI 10.1016/j.jat.2012.01.004. MR2903118

[LHPT] Jean B. Lasserre, Didier Henrion, Christophe Prieur, and Emmanuel Trélat, *Nonlinear optimal control via occupation measures and LMI-relaxations*, SIAM J. Control Optim. **47** (2008), no. 4, 1643–1666, DOI 10.1137/070685051. MR2421324

[MA] MOSEK ApS, *The MOSEK large scale optimization solver, Version 8.1*, https://www.mosek.com/Downloads (2017).

[M] David Mumford, *Lectures on curves on an algebraic surface*, With a section by G. M. Bergman. Annals of Mathematics Studies, No. 59, Princeton University Press, Princeton, N.J., 1966. MR0209285

[P] Mihai Putinar, *A note on Tchakaloff's theorem*, Proc. Amer. Math. Soc. **125** (1997), no. 8, 2409–2414, DOI 10.1090/S0002-9939-97-03862-8. MR1389533

[T] L. Tchakaloff, *Formules générales de quadrature mécanique du type de Gauss*, Colloq. Math. **5** (1957), 69–73 (French). MR0092885

[P] Mihai Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J. **42** (1993), no. 3, 969–984, DOI 10.1512/iumj.1993.42.42045. MR1254128

Departamento de matemáticas, Universidad de los Andes, Carrera $1^{\mathrm{ra}}\#18A-12$, Bogotá, Colombia
E-mail address: jh.garcia1776@uniandes.edu.co, mj.junca20@uniandes.edu.co, mvelasco@uniandes.edu.co

Industrial Engineering and Operations Research Department, Columbia University, 500 West 120-th street, New York, NY 10027,
E-mail address: camilo.hernandez@columbia.edu

Departamento de matemáticas, Universidad de los Andes, Carrera $1^{\mathrm{ra}}\#18A-12$, Bogotá, Colombia
E-mail address: mj.junca20@uniandes.edu.co

Departamento de matemáticas, Universidad de los Andes, Carrera $1^{\mathrm{ra}}\#18A-12$, Bogotá, Colombia
E-mail address: mvelasco@uniandes.edu.co