



HAL
open science

Non-uniform spline recovery from small degree polynomial approximation

Yohann de Castro, Guillaume Mijoule

► **To cite this version:**

Yohann de Castro, Guillaume Mijoule. Non-uniform spline recovery from small degree polynomial approximation. 2014. hal-00950850v1

HAL Id: hal-00950850

<https://hal.science/hal-00950850v1>

Submitted on 23 Feb 2014 (v1), last revised 25 May 2015 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NON-UNIFORM SPLINE RECOVERY FROM SMALL DEGREE POLYNOMIAL APPROXIMATION

YOHANN DE CASTRO AND GUILLAUME MIJOULE

ABSTRACT. We investigate the sparse spikes deconvolution problem onto spaces of algebraic polynomials. Our framework encompasses the measure reconstruction problem from a combination of noiseless and noisy moment measurements. We study a TV-norm regularization procedure to localize the support and estimate the weights of a target discrete measure in this frame. Furthermore, we derive quantitative bounds on the support recovery and the amplitudes errors under a Chebyshev-type minimal separation condition on its support. Finally, we study the localization of the knots of non-uniform splines from small degree polynomial approximations.

1. INTRODUCTION

1.1. Sparse spikes deconvolution onto spaces of algebraic polynomials. This paper is devoted to the extension of some recent results in spike deconvolution to the frame of algebraic polynomials. Beyond the theoretical interest, we focus on this model in order to bring tools and quantitative guarantees from the super-resolution theory [9, 7, 8] to the companion problem of the recovery of knots of non-uniform splines [3]. At first glance, this setting can be depicted as a deconvolution problem where one wants to recover the location of the support of a discrete measure from the observation of its convolution with an algebraic polynomial of given degree \mathbf{m} . Equivalently, we aim at recovering a discrete measure from the knowledge of the true $(\mathbf{d} + 1)$ first moments and a noisy version of the $(\mathbf{m} - \mathbf{d})$ next ones.

1.2. Non-uniform spline recovery. Our framework involves the recovery of non-uniform splines, i.e. a smooth polynomial function that is piecewise-defined on subintervals of different lengths. More precisely, we investigate a grid-free procedure to estimate a non-uniform spline from a polynomial approximation of small degree. Our estimation procedure can be used as a post-processing technique in various fields such as data assimilation [15], shape optimization [14] or spectral methods in PDE's [13].

For instance, one gets a polynomial approximation of the solution of a PDE when using spectral methods such as Galerkin method. In this setting, one seeks a weak solution of a PDE using bounded degree polynomials as test functions. Then, Lax-Milgram theorem grants the existence of a unique weak solution \mathbf{f} for which a polynomial approximation P can be computed. Moreover, Céa's lemma shows that the Galerkin approximation P is comparable to the best polynomial approximation $\mathbf{p}(\mathbf{f})$ of the weak solution \mathbf{f} . In this paper, this situation is depicted by (1.6). Hence, if one knows the weak solution \mathbf{f} is a non-uniform spline then our (post-processing) procedure can provide a grid-free estimate $\hat{\mathbf{f}}$ from the Galerkin

Date: February 23, 2014.

Key words and phrases. LASSO; Super-resolution; Non-Uniform Splines; Algebraic polynomials; ℓ_1 -minimization.

approximation P . Moreover, Theorem 2 shows that the recovered spline has large discontinuities near the large discontinuities of the target spline f . Hence, the location of the large enough discontinuities of the weak solution \mathbf{f} can be quantitatively and in a grid-free manner estimated from the Galerkin approximation using our algorithm.

As an example, Figure 1 illustrates how our procedure improves a polynomial approximation of a non-uniform spline. Observe that discontinuities of splines make them difficult to approximate by polynomials. Even the best polynomial approximation given by orthogonal projection (thin dashed gray line) fails in localizing the knots of splines. Consider a worse approximation (thin black line) of the spline (thick dashed gray line). It seems rather difficult to localize the discontinuities of the spline from the knowledge of this polynomial approximation and the boundary conditions.

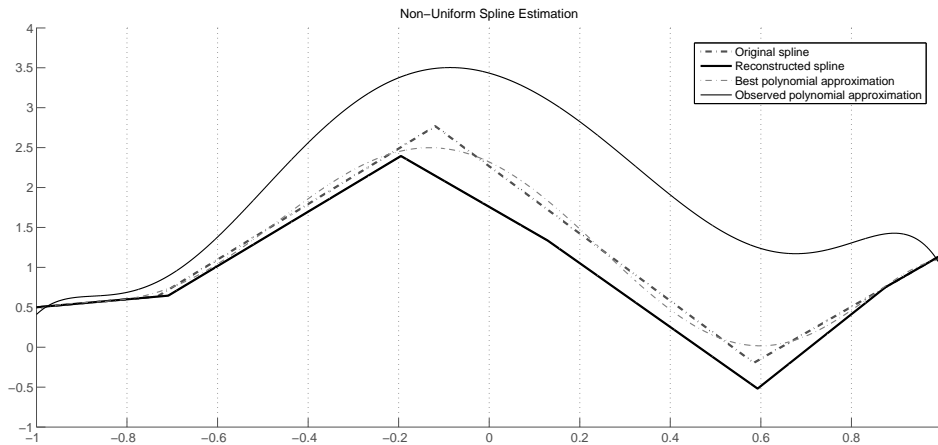


FIGURE 1. Estimated spline (thick black line) of a non-uniform spline (thick dashed gray line) and its knots from a polynomial approximation (thin black line).

Nevertheless, our procedure produces a non-uniform spline (thick black line) whose large discontinuities are close to the knots of the target spline.

1.3. Previous works. The super-resolution problem has been intensively investigated in the last years. In [8] the authors give an exact recovery condition for the noiseless problem in a general setting. In the Fourier frame, this analysis was greatly refined in [7] which shows that the exact recovery condition is satisfied for all measure satisfying a “minimum separation condition”. The recovery from noisy samplings was investigated in [6] which characterizes the reconstruction error as the resolution increases. The first result on quantitative localization was brought by [1] which gives the bounds on the support detection error in a general frame. This analysis was derived in terms of the amplitude of the target measure in [12]. In the Fourier frame, the optimal rates in prediction error have been investigated in [17]. Lastly, the behavior and the stability of ℓ_1 -norm regularization in the space of measures has been investigated in [11] when observing small noise errors.

The spline recovery problem in the noiseless case has been studied in [3] where the authors assume that one knows the orthogonal projection $\mathbf{p}(f)$ of the non-uniform spline f . Our frame extends their point of view to the noisy case where

one observes a polynomial approximation P close to the best polynomial approximation $\mathbf{p}(\mathbf{f})$. To the best of our knowledge, there is no result on a quantitative localization of the knots of non-uniform splines from noisy measurements.

1.4. General model and notation. Let $[-1, 1]$ equipped with the distance:

$$\forall u, v \in [-1, 1], \quad d(u, v) = |\arccos u - \arccos v|.$$

Let \mathbf{x} be a complex measure on $[-1, 1]$ with finite support of size s . In particular, \mathbf{x} has polar decomposition:

$$(1.1) \quad \mathbf{x} = \sum_{k=1}^s \mathbf{a}_k \delta_{\mathbf{t}_k},$$

where $\mathbf{a}_k \in \mathbb{R} \setminus \{0\}$, $\mathbf{t}_k \in [-1, 1]$, and δ_t denotes the Dirac measure at point t . Let \mathbf{m} be a positive integer and $\mathcal{F} = \{\varphi_0, \varphi_1, \dots, \varphi_{\mathbf{m}}\}$ be such that $\varphi_0 = 1$ and for $k = 1, \dots, \mathbf{m}$,

$$\varphi_k = \sqrt{2} T_k,$$

where $T_k(t) = \cos(k \arccos(t))$ is the k -th Chebyshev polynomial of the first kind. Define the k -th generalized moment of a signed measure μ on $[-1, 1]$ as:

$$c_k(\mu) = \int_{[-1, 1]} \varphi_k d\mu,$$

for $k = 0, 1, \dots, \mathbf{m}$. Assume that we observe $c_k(\mathbf{x})$ for $0 \leq k \leq \mathbf{d}$ and a noisy version of $c_k(\mathbf{x})$ for $\mathbf{d} + 1 \leq k \leq \mathbf{m}$ where possibly $\mathbf{d} = -1$. Define $y_k = c_k(\mathbf{x}) + \varepsilon_k$ such as $\varepsilon_k = 0$ for $0 \leq k \leq \mathbf{d}$ and ε_k are i.i.d. $\mathcal{N}(0, \sigma^2)$ for $\mathbf{d} + 1 \leq k \leq \mathbf{m}$. This can be written as:

$$(1.2) \quad \mathbf{y} = \mathbf{c}(\mathbf{x}) + \mathbf{e},$$

where $\mathbf{c}(\mathbf{x}) = (c_k(\mathbf{x}))_{k=0}^{\mathbf{m}}$ and $\mathbf{e} = (0, \mathbf{n})$ with $\mathbf{n} \sim \mathcal{N}(0, \sigma^2 \text{Id}_{\mathbf{m}-\mathbf{d}})$. Note we know the first true moments up to the order \mathbf{d} and a noisy version of them up to the order \mathbf{m} . Moreover, the degree \mathbf{d} is allowed to be -1 . In this case, we only observe a noisy version of the moments up to the order \mathbf{m} .

1.5. An ℓ_1 -minimization procedure. Our analysis follows recent proposal on ℓ_1 -minimization [1, 17, 11]. Denote by \mathbf{M} the set of all finite signed measures on $[-1, 1]$ which is isometrically isomorphic to the dual $\mathcal{C}([-1, 1])^*$ of the continuous function endowed with the supremum norm and by $\|\cdot\|_{TV}$ the total variation norm. We recall that for all $\mu \in \mathbf{M}$,

$$\|\mu\|_{TV} = \sup_{\mathcal{P}} \sum_{E \in \mathcal{P}} |\mu(E)|,$$

where the supremum is taken over all partitions \mathcal{P} of $[-1, 1]$ into a finite number of disjoint measurable subsets. Consider a modified version of the convex program BLASSO [1] given by:

$$(1.3) \quad \hat{\mathbf{x}} \in \arg \min_{\mu \in \mathbf{C}_{\mathbf{d}}(\mathbf{x})} \frac{1}{2} \|\mathbf{c}(\mu) - \mathbf{y}\|_2^2 + \lambda \|\mu\|_{TV},$$

where $\mathbf{C}_{\mathbf{d}}(\mathbf{x}) := \{\mu \in \mathbf{M}; \quad \forall k = 0, \dots, \mathbf{d}, c_k(\mu) = c_k(\mathbf{x})\}$ and $\lambda > 0$ is a tuning parameter. Questions immediately arise:

- How close is the recovered spike train from the target x ?
- How accurate is the localization of (1.3) in terms of the noise and the amplitude of the recovered/original spike?

To the best of our knowledge, this paper is the first to quantitatively address these questions in the frame of algebraic polynomials.

1.6. Contribution.

Definition 1.1 (Minimum separation). *Let $\mathbf{T} \subset [-1, 1]$. We define $\Delta(\mathbf{T})$, the minimum separation of \mathbf{T} , by*

$$\Delta(\mathbf{T}) = \inf_{(t,t') \in \mathbf{T}^2; t \neq t'} \min \{d(t, t'), \pi - d(t, t')\},$$

that is the minimum modulus between two points of $\arccos(T) + \pi\mathbb{Z}$.

Let $\epsilon(\mathbf{T})$ denote the distance from $\mathbf{T} \setminus \{-1, 1\}$ to the edges of $[-1, 1]$:

$$\epsilon(\mathbf{T}) = \inf \{ \min(d(t, 1), d(t, -1)); t \in \mathbf{T} \setminus \{-1, 1\} \}.$$

Theorem 1. *Assume $m \geq 128$. Let $\alpha > 0$ and set:*

$$\lambda_0 := 2\sigma[2(1 + \alpha)(\mathbf{m} - \mathbf{d}) \log(5(\mathbf{m} + \mathbf{d} + 1))]^{1/2},$$

then with probability greater than $1 - \left[\frac{1}{5(\mathbf{m} + \mathbf{d})}\right]^\alpha$ the following holds. If $\lambda \geq \lambda_0$ and

$$(1.4) \quad \min\{\Delta(\mathbf{T}), 2\epsilon(\mathbf{T})\} \geq \frac{5\pi}{\mathbf{m}}.$$

then there exists a solution $\hat{\mathbf{x}}$ to (1.3) with finite support $\hat{\mathbf{x}} = \sum_{k=1}^s \hat{\mathbf{a}}_k \delta_{\hat{\mathbf{t}}_k}$ satisfying:

(1) *Global control:*

$$\sum_{k=1}^s |\hat{\mathbf{a}}_k| \min \left\{ \mathbf{m}^2 \min_{\mathbf{t} \in \mathbf{T}} d(\mathbf{t}, \hat{\mathbf{t}}_k)^2; c_0^2 \right\} \leq c_1 \lambda,$$

(2) *Local control:*

$$\forall i = 1, \dots, s, \quad \left| \mathbf{a}_i - \hat{\mathbf{x}} \left(\left\{ t \mid d(\mathbf{t}_i, t) \leq \frac{c_0}{\mathbf{m}} \right\} \right) \right| \leq c_2 \lambda,$$

(3) *Large spike localization:*

$$\forall i = 1, \dots, s, \text{ s.t. } |a_i| > c_2 \lambda, \quad \exists \hat{\mathbf{t}} \in \text{Supp}(\hat{\mathbf{x}}) \text{ s.t. } d(\mathbf{t}_i, \hat{\mathbf{t}}) \leq \left[\frac{c_1 \lambda}{|a_i| - c_2 \lambda} \right]^{1/2} \frac{1}{\mathbf{m}},$$

where $c_0 = 1.0361$, $c_1 = 235.85$, and $c_2 = 220.72$.

A proof can be found in Appendix C.

1.7. Non-uniform spline reconstruction. In this subsection, we assume $\mathbf{d} \geq 0$. Observe that the frame investigated in this paper covers the recovery problem of non-uniform splines of order \mathbf{d} from projections onto spaces $\mathbb{R}_{\mathbf{m}-\mathbf{d}-1}[X]$ of algebraic polynomials of given degree at most $\mathbf{m} - \mathbf{d} - 1$. Indeed, consider an univariate spline \mathbf{f} of degree \mathbf{d} over the knot sequence $\mathbf{T} = \{-1, \mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_s, 1\}$ that is a continuously differentiable function \mathbf{f} of order $\mathbf{d} - 1$ piecewise-defined by:

$$\mathbf{f} = \mathbb{1}_{[-1, \mathbf{t}_1)} \mathbf{P}_0 + \mathbb{1}_{[\mathbf{t}_1, \mathbf{t}_2)} \mathbf{P}_1 + \dots + \mathbb{1}_{[\mathbf{t}_{s-1}, \mathbf{t}_s)} \mathbf{P}_{s-1} + \mathbb{1}_{[\mathbf{t}_s, 1]} \mathbf{P}_s,$$

where \mathbf{P}_k belongs to $\mathbb{R}_{\mathbf{d}}[X]$, and for all subset $E \subseteq [-1, 1]$, $\mathbb{1}_E(t)$ equals 1 if t belongs to E and 0 otherwise. Consider $\mathbf{f}^{(\mathbf{d}+1)}$, the $(\mathbf{d} + 1)$ -th distributional derivative of \mathbf{f} . Note it can be written as :

$$\mathbf{f}^{(\mathbf{d}+1)} = \sum_{k=1}^s (\mathbf{P}_k^{(\mathbf{d})} - \mathbf{P}_{k-1}^{(\mathbf{d})}) \delta_{\mathbf{t}_k},$$

where $\mathbf{P}_k^{(\mathbf{d})} \in \mathbb{R}$ is the \mathbf{d} -th derivative of \mathbf{P}_k . Using matrix notation, Appendix E shows that:

$$(1.5) \quad \mathbf{c}(\mathbf{f}^{(\mathbf{d}+1)}) = \begin{bmatrix} 0 & W_1 \\ (-1)^{\mathbf{d}+1} \text{Id}_{\mathbf{m}-\mathbf{d}} & W_2 \end{bmatrix} \begin{pmatrix} \mathbf{p}(\mathbf{f}) \\ \mathbf{b} \end{pmatrix},$$

where:

- $\mathbf{p}(\mathbf{f}) = (\langle \mathbf{f}, \varphi_{\mathbf{d}+1}^{(\mathbf{d}+1)} \rangle, \langle \mathbf{f}, \varphi_{\mathbf{d}+2}^{(\mathbf{d}+1)} \rangle, \dots, \langle \mathbf{f}, \varphi_{\mathbf{m}}^{(\mathbf{d}+1)} \rangle)$,
- $\mathbf{b} = (\mathbf{P}_0(-1), \dots, \mathbf{P}_0^{(\mathbf{d}-1)}(-1), \mathbf{P}_0^{(\mathbf{d})}, \mathbf{P}_s(1), \dots, \mathbf{P}_s^{(\mathbf{d}-1)}(1), \mathbf{P}_s^{(\mathbf{d})})$,
- and W_1, W_2 are known matrices, see (E.1), (E.2) and (E.3), whose entries belong to the set $\{-1, 1, \sqrt{2}(-1)^a \mathbf{w}_{b,c}; a \in \{0, 1\} \text{ and } b, c \in \mathbb{N}\}$.

Let \mathcal{L} denote the Lebesgue measure and observe that $\mathbf{p}(\mathbf{f})$ is entirely determined by the orthogonal projection in $L^2(\mathcal{L})$ of \mathbf{f} onto $\mathbb{R}_{\mathbf{m}-\mathbf{d}-1}[X]$, and \mathbf{b} describes the boundary conditions on \mathbf{f} . Note $\mathbf{c}(\mathbf{f}^{(\mathbf{d}+1)})$ can be obtained from the projection $\mathbf{p}(\mathbf{f})$ and \mathbf{b} . Then, our model considers a gaussian perturbation of the projection $\mathbf{p}(\mathbf{f})$.

Assumption 1 (Approximate projection of non-uniform splines). *We say that a random polynomial P with values in $\mathbb{R}_{\mathbf{m}-\mathbf{d}-1}[X]$ satisfies **Assumption 1** if and only if:*

$$(1.6) \quad \Theta(P) = \mathbf{p}(\mathbf{f}) + \mathbf{n},$$

where $\mathbf{n} \sim \mathcal{N}(0, \sigma^2 \text{Id}_{\mathbf{m}-\mathbf{d}})$ and $\Theta : P \mapsto (\langle P, \varphi_{\mathbf{d}+1}^{(\mathbf{d}+1)} \rangle, \langle P, \varphi_{\mathbf{d}+2}^{(\mathbf{d}+1)} \rangle, \dots, \langle P, \varphi_{\mathbf{m}}^{(\mathbf{d}+1)} \rangle)$.

Remark. For the sake of simplicity, we choose a Gaussian approximation error but our analysis can be extended to other types of noise. This extension can be done by bounding the ℓ_∞ -norm of the polynomial whose coefficients are given by the random vector $\mathbf{e} = (0, \mathbf{n})$, as done in Lemma 8.

Remark. Observe the mapping Θ defines an isomorphism from $\mathbb{R}_{\mathbf{m}-\mathbf{d}-1}[X]$ onto $\mathbb{R}^{\mathbf{m}-\mathbf{d}}$. Moreover, note the inverse image of $\mathbf{n} \sim \mathcal{N}(0, \sigma^2 \text{Id}_{\mathbf{m}-\mathbf{d}})$ under Θ is a random polynomial whose entries are Gaussian random variables in any basis of $\mathbb{R}_{\mathbf{m}-\mathbf{d}-1}[X]$. Therefore, Assumption (1.6) can be equivalently formulated as the knowledge of some Gaussian perturbation in any given basis of $\mathbb{R}_{\mathbf{m}-\mathbf{d}-1}[X]$.

Let $\mathbf{b} = (\mathbf{P}_0(-1), \dots, \mathbf{P}_0^{(\mathbf{d}-1)}(-1), \mathbf{P}_0^{(\mathbf{d})}, \mathbf{P}_s(1), \dots, \mathbf{P}_s^{(\mathbf{d}-1)}(1), \mathbf{P}_s^{(\mathbf{d})})$ and P be a random vector with values in $\mathbb{R}_{\mathbf{m}-\mathbf{d}-1}[X]$. Set:

$$(1.7) \quad \hat{\mathbf{x}} \in \arg \min_{\mu \in \mathbf{C}_d(\mathbf{f}^{(\mathbf{d}+1)})} \frac{1}{2} \|\mathbf{c}(\mu) - \mathbf{y}\|_2^2 + \lambda \|\mu\|_{TV}.$$

Recall $\mathbf{C}_d(\mathbf{f}^{(\mathbf{d}+1)}) := \{\mu \in \mathbf{M}; \quad \forall k = 0, \dots, \mathbf{d}, c_k(\mu) = c_k(\mathbf{f}^{(\mathbf{d}+1)})\}$, $\lambda > 0$ is a tuning parameter and

$$\mathbf{y} := \begin{bmatrix} 0 & W_1 \\ (-1)^{\mathbf{d}+1} \text{Id}_{\mathbf{m}-\mathbf{d}} & W_2 \end{bmatrix} \begin{pmatrix} \Theta(P) \\ \mathbf{b} \end{pmatrix}.$$

Note that if a discrete measure $\hat{\mathbf{x}}$ enjoys

$$(1.8) \quad \forall k = 0, \dots, \mathbf{d}, \quad c_k(\hat{\mathbf{x}}) = c_k(\mathbf{f}^{(\mathbf{d}+1)})$$

then one can explicitly construct the unique non-uniform spline $\hat{\mathbf{f}}$ with $(\mathbf{d} + 1)$ -th derivative $\hat{\mathbf{x}}$ and boundary conditions \mathbf{b} . Indeed, observe that we can uniquely construct a non-uniform spline $\hat{\mathbf{f}}$ from the knowledge of the $(\mathbf{d} + 1)$ boundary conditions at point -1 and its $(\mathbf{d} + 1)$ -th derivative. Moreover, Eq.'s (1.8), (E.2) and (E.3) show that $\hat{\mathbf{f}}$ satisfies the $(\mathbf{d} + 1)$ boundary conditions at point 1 and so the boundary conditions \mathbf{b} . Eventually, we consider the algorithm described in Table 1.

Theorem 2. *Let $\mathbf{m} > \mathbf{d} \geq 0$. Let \mathbf{f} be a non-uniform spline of degree \mathbf{d} that can be written as:*

$$\mathbf{f} = \mathbb{1}_{[-1, t_1]} \mathbf{P}_0 + \mathbb{1}_{[t_1, t_2]} \mathbf{P}_1 + \dots + \mathbb{1}_{[t_{s-1}, t_s]} \mathbf{P}_{s-1} + \mathbb{1}_{[t_s, 1]} \mathbf{P}_s,$$

Inputs: Boundary conditions \mathbf{b} , a polynomial approximation P , an upper bound σ on the noise standard deviation and $\alpha > 0$.

1. Set $\mathbf{d} = \text{Size}(\mathbf{b})/2 - 1$ and $\mathbf{m} = \deg(P) + \mathbf{d} + 1$,
2. Compute $\Theta(P) = (\langle P, \varphi_{\mathbf{d}+1}^{(\mathbf{d}+1)} \rangle, \langle P, \varphi_{\mathbf{d}+2}^{(\mathbf{d}+1)} \rangle, \dots, \langle P, \varphi_{\mathbf{m}}^{(\mathbf{d}+1)} \rangle)$,
3. Compute $\mathbf{y} = \begin{bmatrix} 0 & W_1 \\ (-1)^{\mathbf{d}+1} \text{Id}_{\mathbf{m}-\mathbf{d}} & W_2 \end{bmatrix} \begin{pmatrix} \Theta(P) \\ \mathbf{b} \end{pmatrix}$, where W_1 and W_2 are described in Appendix E.
4. Set $\lambda = 4\sigma[2(1 + \alpha)(\mathbf{m} - \mathbf{d}) \log(5(\mathbf{m} + \mathbf{d} + 1))]^{1/2}$,
5. Find a discrete solution $\hat{\mathbf{x}} = \sum_{k=1}^{\hat{s}} \hat{\mathbf{a}}_k \delta_{\hat{\mathbf{t}}_k}$ to (1.7),
6. Find the unique spline $\hat{\mathbf{f}}$ of order $\mathbf{d} - 1$ such that $\hat{\mathbf{f}}^{(\mathbf{d}+1)} = \hat{\mathbf{x}}$ and $(\hat{\mathbf{f}}, \dots, \hat{\mathbf{f}}_0^{(\mathbf{d}-1)}(-1), \hat{\mathbf{f}}_0^{(\mathbf{d})}, \hat{\mathbf{f}}_s(1), \dots, \hat{\mathbf{f}}_s^{(\mathbf{d}-1)}(1), \hat{\mathbf{f}}_s^{(\mathbf{d})}) = \mathbf{b}$.

Output: A non-uniform spline $\hat{\mathbf{f}}$.

TABLE 1. Non-uniform spline recovery algorithm.

where $\mathbf{P}_k \in \mathbb{R}_{\mathbf{d}}[X]$ and $\mathbf{T} = \{-1, \mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_s, 1\}$ enjoys:

$$\min\{\Delta(\mathbf{T}), 2\epsilon(\mathbf{T})\} \geq \frac{5\pi}{\mathbf{m}}.$$

Set $\mathbf{b} = (\mathbf{P}_0(-1), \dots, \mathbf{P}_0^{(\mathbf{d}-1)}(-1), \mathbf{P}_0^{(\mathbf{d})}, \mathbf{P}_s(1), \dots, \mathbf{P}_s^{(\mathbf{d}-1)}(1), \mathbf{P}_s^{(\mathbf{d})})$ and let P be such that **Assumption 1** holds. Let $\alpha > 0$ then, with probability greater than $1 - [\frac{1}{5(\mathbf{m}+\mathbf{d})}]^\alpha$, any output $\hat{\mathbf{f}}$ of the aforementioned algorithm enjoys:

(1) *Global control:*

$$\sum_{k=1}^{\hat{s}} |\hat{\mathbf{P}}_k^{(\mathbf{d})} - \hat{\mathbf{P}}_{k-1}^{(\mathbf{d})}| \min \left\{ \mathbf{m}^2 \min_{\mathbf{t} \in \mathbf{T}} d(\mathbf{t}, \hat{\mathbf{t}}_k)^2; c_0^2 \right\} \leq c_1 \lambda,$$

(2) *Large discontinuity localization:* $\forall i = 1, \dots, s$, s.t. $|\mathbf{P}_i^{(\mathbf{d})} - \mathbf{P}_{i-1}^{(\mathbf{d})}| > c_2 \lambda$,

$$\exists \hat{\mathbf{t}} \in \{\hat{\mathbf{t}}_1, \dots, \hat{\mathbf{t}}_s\} \text{ s.t. } d(\mathbf{t}_i, \hat{\mathbf{t}}) \leq \left[\frac{c_1 \lambda}{|\mathbf{P}_i^{(\mathbf{d})} - \mathbf{P}_{i-1}^{(\mathbf{d})}| - c_2 \lambda} \right]^{1/2} \frac{1}{\mathbf{m}},$$

where $c_0 = 1.0361$, $c_1 = 235.85$, $c_2 = 220.72$, $\lambda = 4\sigma[2(1 + \alpha)(\mathbf{m} - \mathbf{d}) \log(5(\mathbf{m} + \mathbf{d} + 1))]^{1/2}$ and $\hat{\mathbf{f}}$ is written as:

$$\hat{\mathbf{f}} = \mathbb{1}_{[-1, \hat{\mathbf{t}}_1]} \hat{\mathbf{P}}_0 + \mathbb{1}_{[\hat{\mathbf{t}}_1, \hat{\mathbf{t}}_2]} \hat{\mathbf{P}}_1 + \dots + \mathbb{1}_{[\hat{\mathbf{t}}_{s-1}, \hat{\mathbf{t}}_s]} \mathbf{P}_{s-1} + \mathbb{1}_{[\hat{\mathbf{t}}_s, 1]} \mathbf{P}_s,$$

with $\hat{\mathbf{P}}_k \in \mathbb{R}_{\mathbf{d}}[X]$.

A proof can be found in Appendix F.

2. QUANTITATIVE LOCALIZATION

2.1. Zero-noise problem. In the noiseless case, observe that $\mathbf{n} = 0$. Exact recovery from moment samples has been investigated in [1, 3] where one considers the program:

$$(2.1) \quad \mathbf{x}^0 \in \arg \min_{\mu \in \mathbf{M}} \|\mu\|_{TV} \quad \text{s.t.} \quad \int \Phi d\mu = \int \Phi d\mathbf{x},$$

whith $\Phi = (\varphi_0, \dots, \varphi_m)$ is the Chebyshev moment curve. The optimality condition of (2.1) shows that the sub-gradient of the ℓ_1 -norm vanishes at any solution point \mathbf{x}^0 . Therefore a sufficient condition for exact recovery is that \mathbf{x} satisfies the optimality condition. This is covered by the notion of “dual certificate” [8, 7] or equivalently the notion of “source condition” [5].

Definition 2.1 (Dual certificate). *We say that a polynomial $P = \sum_{k=0}^m \alpha_k \varphi_k$ is a dual certificate for the measure \mathbf{x} defined by (1.1) if and only if it satisfies the following properties:*

- *phase interpolation: $\forall k \in \{1, \dots, S\}, P(\mathbf{t}_k) = \mathbf{a}_k / |\mathbf{a}_k|$,*
- *ℓ_∞ -constraint: $\|P\|_\infty \leq 1$.*

One can prove [8] that \mathbf{x} is a solution to (2.1) if and only if \mathbf{x} has a dual certificate.

2.2. Semi-noisy moment sample model. In our model, we deal with an observation \mathbf{y} described by (1.2). In this case, the existence of a dual certificate is not sufficient to derive support localization, see [1]. One needs to strengthen this notion using the Quadratic Isolation Condition [1].

Definition 2.2 (Quadratic isolation condition). *We say that a finite set $\mathbf{T} = \{\mathbf{t}_1, \dots, \mathbf{t}_s\} \subset [-1, 1]$ satisfies the quadratic isolation condition with parameters $C_a > 0$ and $0 < C_b < 1$, denoted by $\text{QIC}(C_a, C_b)$, if and only if for all $(\theta_k)_{k=1}^s \in \mathbb{R}^s$, there exists $P \in \text{Span}(\mathcal{F})$ such that for all $k = 1, \dots, s$, $P(\mathbf{t}_k) = \exp(-i\theta_k)$, and*

$$\forall x \in [-1, 1], \quad 1 - |P(x)| \geq \min_{t \in \mathbf{T}} \{C_a \mathbf{m}^2 d(x, t)^2, C_b\}.$$

As showed by Lemma 6, if the support \mathbf{T} satisfy a minimal separation condition described in (1.4) then \mathbf{T} satisfies $\text{QIC}(C_a, C_b)$ with constants $C_a = 0.00848$ and $C_b = 0.00879$. Using Bernstein’s inequality for algebraic polynomials and the dual certificate construction of [7], we prove Theorem 1, see Appendix C.

3. SEMI-DEFINITE PROGRAMMING

Observe that the Fenchel-Legendre dual program of (1.3) is given by:

$$(3.1) \quad \hat{\alpha} \in \arg \min_{\|\sum_{k=0}^m \alpha_k \varphi_k\|_\infty \leq \lambda} \left\{ \frac{1}{2} \|\alpha\|_2^2 + \sum_{k=d+1}^m \alpha_k y_k \right\},$$

and strong duality holds, see Lemma 9. This dual program can be seen as the orthogonal projection of the vector $(0, \dots, 0, y_{d+1}, \dots, y_m)$ onto the convex set:

$$\left\{ \alpha \in \mathbb{R}^{m+1}, \quad \left\| \sum_{k=0}^m \alpha_k \varphi_k \right\|_\infty \leq \lambda \right\}.$$

Therefore, there is a unique solution to (3.1). Moreover, observe that the constraint $\|\sum_{k=0}^m \alpha_k \varphi_k\|_\infty \leq \lambda$ can be re-cast as imposing that the algebraic polynomials:

$$(3.2) \quad P_1 := \lambda + \sum_{k=0}^m \alpha_k \varphi_k \geq 0 \quad \text{and} \quad P_2 := \lambda - \sum_{k=0}^m \alpha_k \varphi_k \geq 0.$$

Considering the change of variables $\theta = \arccos(t)$, the aforementioned inequalities can be equivalently drawn for some trigonometric polynomials. Using Riesz-Fejér theorem, one can show that non-negative trigonometric polynomials are sums of squares polynomials (SOS). A standard result, see for instance [10], ensures that the convex set of sum of square polynomials (SOS) can be described as

the intersection between the set of positive hermitian semi-definite (SDP) matrices and an affine constraint.

Lemma 3. *The constraint (3.2) can be re-casted into a semi-definite constraint.*

Hence, we can compute $\hat{\alpha}$ using a SDP program. Moreover, Fenchel's duality theorem shows that the dual polynomial:

$$\hat{P} = \frac{1}{\lambda} \sum_{k=0}^{\mathbf{m}} \hat{\alpha}_k \varphi_k,$$

is a sub-gradient of the TV-norm at point $\hat{\mathbf{x}}$. In particular, the support $\hat{\mathbf{T}}$ of $\hat{\mathbf{x}}$ is included in:

$$\{t \in [-1, 1], \quad |\hat{P}| = 1\}.$$

If \hat{P} is not constant, this level set has at most $\mathbf{m} + 1$ points and it defines the support of the solution. Hence, we can find the weights of $\hat{\mathbf{x}}$ using a least-square-type estimator subject to the affine constraint given by the intersection between $\mathbf{C}_d(\mathbf{x})$ and discrete measures with support included in $\hat{\mathbf{T}}$. In this case, the solution to (1.3) is unique and can be computed using the aforementioned SDP program. If \hat{P} is constant then there always exists a solution to (1.3) with finite support. Indeed, using the fact that there is no duality gap, one can check that the solution has non-negative (resp. non-positive) weights if $\hat{P} = 1$ (resp. $\hat{P} = -1$). Therefore, Carathéodory's theorem shows that there always exists a solution with finite support¹. However, one can not use the dual program (3.1) to compute the solution to the primal program (1.3). We deduce the following lemma.

Lemma 4. *There always exists a solution to the primal problem (1.3) with a support of size at most $\mathbf{m} + 2$. Moreover, if \hat{P} is not constant, the solution to (1.3) is unique, its support is included in the level set $\{t \in [-1, 1], \quad |\hat{P}| = 1\}$ and has size at most $\mathbf{m} + 1$.*

APPENDIX A. DUAL CERTIFICATES

This section capitalizes on the recent breakthrough presented in [7] and builds an explicit dual certificate in the frame of algebraic polynomials. More precisely, we explicitly upper bound the dual certificates by a quadratic function near the support points, as done in [7].

Lemma 5. *Assume (1.4) holds. Then for all $\mathbf{t}_j \in \mathbf{T}$, there exists a polynomial $q_{\mathbf{t}_j}$ of degree \mathbf{m} such that:*

- (1) $q_{\mathbf{t}_j}(\mathbf{t}_j) = 1$,
- (2) $\forall t_l \in \mathbf{T} \setminus \{\mathbf{t}_j\}, \quad q_{\mathbf{t}_j}(t_l) = 0$,
- (3) if $d(t, \mathbf{t}_j) \leq c_0/\mathbf{m}$ then:

$$1 - C_2 \mathbf{m}^2 d(t, \mathbf{t}_j)^2 \leq q_{\mathbf{t}_j}(t) \leq 1 - C_1 \mathbf{m}^2 d(t, \mathbf{t}_j)^2,$$

- (4) if $d(t, t_l) \leq c_0/\mathbf{m}$ and $t_l \in \mathbf{T} \setminus \{\mathbf{t}_j\}$ then:

$$C_1 \mathbf{m}^2 d(t, \mathbf{t}_j)^2 \leq q_{\mathbf{t}_j}(t) \leq C_2 \mathbf{m}^2 d(t, \mathbf{t}_j)^2,$$

- (5) if $d(t, t_l) > c_0/\mathbf{m}$ for all $t_l \in \mathbf{T}$ then:

$$c_0^2 C_1 \leq q_{\mathbf{t}_j}(t) \leq 1 - c_0^2 C_1,$$

where $c_0 = 2\pi \cdot 0.1649$, $C_1 = 0.00424$, and $C_2 = 0.25$.

¹The interested reader may find a valuable reference on the geometry of the cone of non-negative measures in [16].

Proof. By symmetrizing the support, we can use existing results for real trigonometric polynomials. Let $X = \frac{1}{2\pi} (\arccos(\mathbf{T}) \cup [-\arccos(\mathbf{T})]) + \frac{1}{2}$. Note that $X \subset [0, 1]$. It is easy to check that (1.4) implies:

$$(A.1) \quad \min_{(x, x') \in X; x \neq x'} |x - x'| \geq 2.5/\mathbf{m}$$

Thus, according to Proposition 2.1 and Lemma 2.5 of [7], for all $x_j \in X$, there exists a real trigonometric polynomial of degree \mathbf{m} , $\tilde{q}_{x_j} : x \mapsto \sum_{k=-\mathbf{m}}^{\mathbf{m}} c_k e^{2i\pi kx}$, such that:

- $\tilde{q}_{x_j}(x_j) = \tilde{q}_{x_j}(-x_j) = 1$,
- $|\tilde{q}_{x_j}(x)| < 1, x \in [0, 1] \setminus X$,
- $\tilde{q}_{x_j}(x_l) = -1, x_l \in X \setminus \{x_j, -x_j\}$,
- $\forall (x, x_l) \in [0, 1] \times X, |x - x_l| \leq 0.1649/\mathbf{m}$,

$$|\tilde{q}_{x_j}(x)| \leq 1 - 0.3353 \mathbf{m}^2 (x - x_l)^2,$$

- $\forall x \in [0, 1], \forall x_l \in X, |x - x_l| > 0.1649/\mathbf{m}$,

$$|\tilde{q}_{x_j}(x)| \leq 1 - 0.3353 \cdot 0.1649^2.$$

By construction [7] the trigonometric polynomial function $p_{x_j} : x \in [-\pi, \pi] \mapsto \tilde{q}_{x_j} \left(\frac{1}{2\pi}x + \frac{1}{2} \right)$ is real and even, so we have the expansion:

$$p_{x_j}(x) = \sum_{k=0}^{\mathbf{m}} a_k \cos(kx).$$

Moreover, since $\sup_{x \in [0, 2\pi]} |p_{x_j}(x)| = 1$, Bernstein's inequality [4] implies:

$$(A.2) \quad \sup_{x \in [0, 1]} |p_{x_j}''(x)| \leq \mathbf{m}^2.$$

Let $\mathbf{t}_j \in \mathbf{T}$ and $x_j = \arccos(\mathbf{t}_j)$. We define:

$$q_{\mathbf{t}_j}(t) = \frac{1}{2} p_{x_j}(\arccos t) + \frac{1}{2} = \frac{1}{2} \sum_{k=0}^{\mathbf{m}} a_k T_k(t) + \frac{1}{2},$$

where T_k is the k -th Chebyshev polynomial of the first kind. Lemma 5 is a direct consequence of the properties verified by \tilde{q}_{x_j} and (A.2). \square

Lemma 6. Assume (1.4) holds. Then for all (v_1, \dots, v_S) such that $\forall j \in [1, S], |v_j| = 1$, there exists a polynomial q of degree \mathbf{m} such that:

- (1) $\forall j \in [1, S], q(\mathbf{t}_j) = v_j$,
- (2) if $d(t, \mathbf{t}_j) \leq c_0/\mathbf{m}$ then:

$$1 - |q(t)| \geq 2C_1 \mathbf{m}^2 d(t, \mathbf{t}_j)^2,$$

- (3) if $d(t, t_l) > 2\pi \cdot 0.1649/\mathbf{m}$ for all $t_l \in \mathbf{T}$ then:

$$1 - |q(t)| \geq 2c_0^2 C_1,$$

where $c_0 = 2\pi \cdot 0.1649$ and $C_1 = 0.00424$.

Proof. Similarly as previous lemma, if $X = \frac{1}{2\pi} (\arccos(\mathbf{T}) \cup [-\arccos(\mathbf{T})]) + \frac{1}{2}$, then we can construct a trigonometric polynomial $\tilde{q} : x \mapsto \sum_{k=-\mathbf{m}}^{\mathbf{m}} c_k e^{2i\pi kx}$, such that:

- $\tilde{q}(x_j) = \tilde{q}(-x_j) = v_j, \forall j \in [1, S]$,
- $|\tilde{q}_{x_j}(x)| < 1, \forall x \in [0, 1] \setminus X$,

- $\forall (x, x_l) \in [0, 1] \times X, |x - x_l| \leq 0.1649/\mathbf{m},$
 $|\tilde{q}_{x_l}(x)| \leq 1 - 0.3353 \mathbf{m}^2 (x - x_l)^2,$
- $\forall x \in [0, 1], \forall x_l \in X, |x - x_l| > 0.1649/\mathbf{m},$
 $|\tilde{q}_{x_l}(x)| \leq 1 - 0.3353 \cdot 0.1649^2.$

Then $p : x \in [-\pi, \pi] \mapsto \tilde{q}\left(\frac{1}{2\pi}x + \frac{1}{2}\right)$ is even, so we have the expansion $p(x) = \sum_{k=0}^{\mathbf{m}} a_k \cos(kx)$ where $a_k \in \mathbb{R}$. Putting

$$q : t \mapsto \sum_{k=0}^{\mathbf{m}} a_k \cos(k \arccos t) = \sum_{k=0}^{\mathbf{m}} a_k T_k(t),$$

we can show q verifies the needed properties. \square

APPENDIX B. RICE METHOD

Define the Gaussian process $\{X_{\mathbf{m}, \mathbf{d}}(t), t \in [-1, 1]\}$ by:

$$\forall t \in [-1, 1], \quad X_{\mathbf{m}, \mathbf{d}}(t) = \xi_{\mathbf{d}+1} \varphi_{\mathbf{d}+1}(t) + \xi_{\mathbf{d}+2} \varphi_{\mathbf{d}+2}(t) + \dots + \xi_{\mathbf{m}} \varphi_{\mathbf{m}}(t),$$

where $\xi_{\mathbf{d}+1}, \dots, \xi_{\mathbf{m}}$ are i.i.d. standard normal. Its covariance function is:

$$r(s, t) = \varphi_{\mathbf{d}+1}(t) \varphi_{\mathbf{d}+1}(s) + \varphi_{\mathbf{d}+2}(t) \varphi_{\mathbf{d}+2}(s) + \dots + \varphi_{\mathbf{m}}(t) \varphi_{\mathbf{m}}(s),$$

where the dependence in \mathbf{m} and \mathbf{d} has been omitted. Observe its maximal variance is attained at point 1 and is given by $\sigma_{\mathbf{m}, \mathbf{d}}^2 = 2(\mathbf{m} - \mathbf{d})$, and its variance function is $\sigma_{\mathbf{m}, \mathbf{d}}^2(t) = \varphi_{\mathbf{d}+1}(t)^2 + \varphi_{\mathbf{d}+2}(t)^2 + \dots + \varphi_{\mathbf{m}}(t)^2$.

Lemma 7. Let $\mathcal{M} = \max_{t \in [-1, 1]} |X_{\mathbf{m}, \mathbf{d}}(t)|$, then:

$$\forall u > \sqrt{2(\mathbf{m} - \mathbf{d})}, \quad \mathbb{P}\{\mathcal{M} > u\} \leq 5(\mathbf{m} + \mathbf{d} + 1) \exp\left[-\frac{u^2}{8(\mathbf{m} - \mathbf{d})}\right].$$

Proof. By the change of variables $t = \cos \theta$, for all $t \in [-1, 1]$:

$$X_{\mathbf{m}, \mathbf{d}}(t) = X_{\mathbf{m}, \mathbf{d}}(\cos \theta) = \sqrt{2} \xi_{\mathbf{d}+1} \cos((\mathbf{d} + 1)\theta) + \dots + \sqrt{2} \xi_{\mathbf{m}} \cos(\mathbf{m}\theta).$$

Set $T(\theta) := X_{\mathbf{m}, \mathbf{d}}(t)$. We recall that its variance function is given by:

$$\sigma_{\mathbf{m}, \mathbf{d}}^2(\theta) = 2 \cos^2((\mathbf{d} + 1)\theta) + \dots + 2 \cos^2(\mathbf{m}\theta) = \mathbf{m} - \mathbf{d} + \frac{\mathbf{D}_{\mathbf{m}}(2\theta) - \mathbf{D}_{\mathbf{d}}(2\theta)}{2},$$

where \mathbf{D}_k denotes the Dirichlet kernel of order k . Observe that:

$$\forall \theta \in \mathbb{R}, \quad \sigma_{\mathbf{m}, \mathbf{d}}^2(\theta) \leq \sigma_{\mathbf{m}, \mathbf{d}}^2(0) = 2(\mathbf{m} - \mathbf{d}),$$

By the Rice method [2], for $u > 0$:

$$\begin{aligned} \mathbb{P}\{\mathcal{M} > u\} &\leq 2\mathbb{P}\left\{\max_{\theta \in [0, \pi]} T(\theta) > u\right\}, \\ &\leq 2\mathbb{P}\{T(0) > u\} + 2\mathbb{E}[U_u([0, \pi])], \\ &= 2\left[1 - \Psi\left(\frac{u}{\sqrt{2(\mathbf{m} - \mathbf{d})}}\right)\right] + 2 \int_0^\pi \mathbb{E}((T'(\theta))^+ | T(\theta) = u) \psi_{\sigma_{\mathbf{m}, \mathbf{d}}(\theta)}(u) dt \end{aligned}$$

where U_u is the number of crossings of the level u , Ψ is the standard normal distribution, and ψ_σ is the density of the centered normal distribution with standard error σ . First, observe that for $v > 0$, $(1 - \Psi(v)) \leq (1/2) \exp(-v^2/2)$. Hence,

$$1 - \Psi\left(\frac{u}{\sqrt{2(\mathbf{m} - \mathbf{d})}}\right) \leq (1/2) \exp\left(-\frac{u^2}{2(\mathbf{m} - \mathbf{d})}\right).$$

Moreover, regression formulas implies that:

$$\begin{aligned}\mathbb{E}(T'(\theta)|T(\theta) = u) &= \frac{r_{0,1}(\theta, \theta)}{r(\theta, \theta)} u, \\ \text{Var}(T'(\theta)|T(\theta) = u) &\leq \text{Var}(T'(\theta)) = r_{1,1}(\theta, \theta),\end{aligned}$$

where, for instance, $r_{1,1}(\nu, \theta) = \frac{\partial^2 r(\nu, \theta)}{\partial \nu \partial \theta}$. We recall that the covariance function is given by:

$$\begin{aligned}r(\nu, \theta) &= 2 \cos((\mathbf{d} + 1)\nu) \cos((\mathbf{d} + 1)\theta) + \dots + 2 \cos(\mathbf{m}\nu) \cos(\mathbf{m}\theta), \\ &= \frac{1}{2} [\mathbf{D}_{\mathbf{m}}(\nu - \theta) + \mathbf{D}_{\mathbf{m}}(\nu + \theta) - \mathbf{D}_{\mathbf{d}}(\nu - \theta) - \mathbf{D}_{\mathbf{d}}(\nu + \theta)].\end{aligned}$$

Observe that:

$$\begin{aligned}r_{0,1}(\theta, \theta) &= \frac{1}{2} [\mathbf{D}'_{\mathbf{m}}(2\theta) - \mathbf{D}'_{\mathbf{d}}(2\theta)] = - \sum_{k=\mathbf{d}+1}^{\mathbf{m}} k \sin(2k\theta), \\ r_{1,1}(\theta, \theta) &= \sum_{k=\mathbf{d}+1}^{\mathbf{m}} k^2 (1 - \cos(2k\theta)).\end{aligned}$$

On the other hand, if $Z \sim \mathcal{N}(\mu, \sigma^2)$ then

$$\mathbb{E}(Z^+) = \mu \Psi\left(\frac{\mu}{\sigma}\right) + \sigma \psi\left(\frac{\mu}{\sigma}\right) \leq \mu^+ + \frac{\sigma}{\sqrt{2\pi}},$$

where ψ is the standard normal density. We get that:

$$\begin{aligned}&\int_0^\pi \mathbb{E}((T'(\theta))^+ | T(\theta) = u) \psi_{\sigma_{\mathbf{m},\mathbf{d}}(\theta)}(u) d\theta \\ &\leq \int_0^\pi \frac{[\mathbf{D}'_{\mathbf{m}}(2\theta) - \mathbf{D}'_{\mathbf{d}}(2\theta)]^+}{2 \sigma_{\mathbf{m},\mathbf{d}}^2(\theta)} u \psi_{\sigma_{\mathbf{m},\mathbf{d}}(\theta)}(u) d\theta \\ &+ \frac{1}{\sqrt{2\pi}} \int_0^\pi \left[\sum_{k=\mathbf{d}+1}^{\mathbf{m}} k^2 (1 - \cos(2k\theta)) \right]^{1/2} \psi_{\sigma_{\mathbf{m},\mathbf{d}}(\theta)}(u) d\theta, \\ &= A + B.\end{aligned}$$

We use the following straightforward relations:

- $\forall 0 < \sigma_1 < \sigma_2 < u$, $\psi_{\sigma_1}(u) \leq \psi_{\sigma_2}(u)$,
- $\forall \theta$, $[\mathbf{D}'_{\mathbf{m}}(2\theta) - \mathbf{D}'_{\mathbf{d}}(2\theta)]^+ \leq \sum_{k=\mathbf{d}+1}^{\mathbf{m}} k = \frac{(\mathbf{m}+\mathbf{d}+1)(\mathbf{m}-\mathbf{d})}{2}$,
- $\forall \theta \in [0, \pi]$,

$$\frac{u}{2 \sigma_{\mathbf{m},\mathbf{d}}^2(\theta)} \psi_{\sigma_{\mathbf{m},\mathbf{d}}(\theta)}(u) \leq \frac{1}{2\sqrt{2\pi}u^2} \frac{u^3}{\sigma_{\mathbf{m},\mathbf{d}}^3(\theta)} e^{-\frac{u^2}{4\sigma_{\mathbf{m},\mathbf{d}}^2(\theta)}} e^{-\frac{u^2}{4\sigma_{\mathbf{m},\mathbf{d}}^2(\theta)}} \leq \frac{2}{3u^2} e^{-\frac{u^2}{8(\mathbf{m}-\mathbf{d})}}.$$

Eventually, we get, for $u > \sqrt{2(\mathbf{m}-\mathbf{d})}$:

$$\begin{aligned}A &\leq \frac{\pi}{3} \frac{(\mathbf{m} + \mathbf{d} + 1)(\mathbf{m} - \mathbf{d})}{u^2} \exp\left(-\frac{u^2}{8(\mathbf{m} - \mathbf{d})}\right), \\ B &\leq \left[\frac{\pi}{12} ((2\mathbf{m} + 1)(\mathbf{m} + 1)\mathbf{m} - (2\mathbf{d} + 1)(\mathbf{d} + 1)\mathbf{d}) \right]^{1/2} \psi_{\sqrt{2(\mathbf{m}-\mathbf{d})}}(u).\end{aligned}$$

and the result follows. \square

Lemma 8. Set $\lambda_R := \sigma[8(\mathbf{m} - \mathbf{d}) \log(5(\mathbf{m} + \mathbf{d} + 1))]^{1/2}$ and $\lambda > \lambda_R$, then:

$$\mathbb{P}\left(\left\| \sum_{k=0}^{\mathbf{m}} \varepsilon_k \varphi_k \right\|_{\infty} > \lambda\right) \leq \exp\left[-\frac{\lambda^2 - \lambda_R^2}{8\sigma^2(\mathbf{m} - \mathbf{d})}\right].$$

In particular, for all $t > 0$, if

$$\lambda_0(t) := \sigma[8(1+t)(\mathbf{m} - \mathbf{d}) \log(5(\mathbf{m} + \mathbf{d} + 1))]^{1/2},$$

then

$$(B.1) \quad \mathbb{P} \left(\left\| \sum_{k=0}^{\mathbf{m}} \varepsilon_k \varphi_k \right\|_{\infty} > \lambda_0(t) \right) \leq \frac{1}{[5(\mathbf{m} + \mathbf{d} + 1)]^t}.$$

APPENDIX C. PROOF OF THEOREM 1

Assume that $\lambda \geq \lambda_0$ where λ_0 is described in Lemma 8 (the dependence in t has been omitted). Observe that the condition in Lemma 11 is met. One can prove that there exists a solution $\hat{\mathbf{x}}$ to (1.3) with finite support, see Lemma 4. Set:

$$\hat{\mathbf{x}} = \sum_{k=1}^{\hat{s}} \hat{\mathbf{a}}_k \delta_{\hat{\mathbf{t}}_k}.$$

Set $v_j = \bar{\mathbf{a}}_j / |\mathbf{a}_j|$ for $j = 1, \dots, s$ and consider $q = \sum_{k=0}^{\mathbf{m}} \beta_k \varphi_k$ the algebraic polynomial described in Lemma 6. Set:

$$\mathcal{D} := \|\hat{\mathbf{x}}\|_{TV} - \|\mathbf{x}\|_{TV} - \int_{-1}^1 q d(\hat{\mathbf{x}} - \mathbf{x}).$$

Note that $\mathcal{D} \geq 0$. Since \mathbf{x} is feasible, it holds:

$$\frac{1}{2} \|\mathbf{c}(\hat{\mathbf{x}}) - \mathbf{y}\|_2^2 + \lambda \mathcal{D} + \lambda \int_{-1}^1 q d(\hat{\mathbf{x}} - \mathbf{x}) \leq \frac{1}{2} \|\mathbf{e}\|_2^2 + \lambda \|\mathbf{x}\|_{TV}.$$

Hence,

$$\frac{1}{2} \|\mathbf{c}(\hat{\mathbf{x}}) - \mathbf{y} + \lambda \beta\|_2^2 + \lambda \mathcal{D} \leq \frac{1}{2} \|\mathbf{e}\|_2^2 + \frac{1}{2} \|\lambda \beta\|_2^2 - \lambda \langle \mathbf{e}, \mathbf{a} \rangle,$$

where $\beta := (\beta_k)_{k=0}^{\mathbf{m}}$ denotes the coefficients of q in the basis \mathcal{F} . Eventually,

$$\mathcal{D} \leq \frac{\lambda}{2} \|\beta - \frac{\mathbf{e}}{\lambda}\|_2^2.$$

Observe that:

$$\left\| \sum_{k=0}^{\mathbf{m}} (\beta_k - \varepsilon_k / \lambda) \varphi_k \right\|_{\infty} \leq 2,$$

so that:

$$(C.1) \quad \mathcal{D} \leq 2\lambda.$$

Moreover, note that:

$$(C.2) \quad \begin{aligned} \mathcal{D} &= \|\hat{\mathbf{x}}\|_{TV} - \int_{-1}^1 q d\hat{\mathbf{x}}, \\ &\geq \sum_{k=1}^{\hat{s}} |\hat{\mathbf{a}}_k| (1 - |q|)(\hat{\mathbf{t}}_k), \\ &\geq \sum_{k=1}^{\hat{s}} |\hat{\mathbf{a}}_k| \min\{2C_1 \mathbf{m}^2 \min_{\mathbf{t} \in \mathbf{T}} d(\mathbf{t}, \hat{\mathbf{t}}_k)^2; 2c_0^2 C_1\}, \end{aligned}$$

where $c_0 = 2\pi \cdot 0.1649$ and $C_1 = 0.00424$. Now, let $\mathbf{t} \in \mathbf{T}$ and consider the polynomial $q_{\mathbf{t}}$ described in Lemma 5. Using (C.2) we get that:

$$\begin{aligned}
& \left| \sum_{\{k \mid d(\mathbf{t}, \hat{\mathbf{t}}_k) > \frac{c_0}{\mathbf{m}}\}} \hat{\mathbf{a}}_k q_{\mathbf{t}}(\hat{\mathbf{t}}_k) + \sum_{\{k \mid d(\mathbf{t}, \hat{\mathbf{t}}_k) \leq \frac{c_0}{\mathbf{m}}\}} \hat{\mathbf{a}}_k (q_{\mathbf{t}}(\hat{\mathbf{t}}_k) - 1) \right| \\
& \leq \sum_{\{k \mid d(\mathbf{t}, \hat{\mathbf{t}}_k) > \frac{c_0}{\mathbf{m}}\}} |\hat{\mathbf{a}}_k| |q_{\mathbf{t}}(\hat{\mathbf{t}}_k)| + \sum_{\{k \mid d(\mathbf{t}, \hat{\mathbf{t}}_k) \leq \frac{c_0}{\mathbf{m}}\}} |\hat{\mathbf{a}}_k| |q_{\mathbf{t}} - 1|(\hat{\mathbf{t}}_k), \\
& \leq \sum_{k=1}^{\hat{s}} |\hat{\mathbf{a}}_k| \min\{C_2 \mathbf{m}^2 \min_{\mathbf{t} \in \mathbf{T}} d(\mathbf{t}, \hat{\mathbf{t}}_k)^2; 1 - c_0^2 C_1\}, \\
& \leq C' \times \sum_{k=1}^{\hat{s}} |\hat{\mathbf{a}}_k| \min\{2C_1 \mathbf{m}^2 \min_{\mathbf{t} \in \mathbf{T}} d(\mathbf{t}, \hat{\mathbf{t}}_k)^2; 2c_0^2 C_1\}, \\
\text{(C.3)} \quad & \leq 2C' \lambda.
\end{aligned}$$

where $C_2 = 0.25$ and $C' = \max\{\frac{C_2}{2C_1}; \frac{1-c_0^2 C_1}{2c_0^2 C_1}\} = 109.36$. Invoking (D.9), we deduce that for all $i = 1, \dots, s$,

$$\begin{aligned}
| \mathbf{a}_i - \hat{\mathbf{x}}(\mathbf{t}_i + \mathcal{B}(c_0/\mathbf{m})) | & \leq \left| \int q_{\mathbf{t}_i} d\mathbf{x} - \int q_{\mathbf{t}_i} d\hat{\mathbf{x}} \right. \\
& \quad \left. + \sum_{\{k \mid d(\mathbf{t}_i, \hat{\mathbf{t}}_k) > \frac{c_0}{\mathbf{m}}\}} \hat{\mathbf{a}}_k q_{\mathbf{t}_i}(\hat{\mathbf{t}}_k) + \sum_{\{k \mid d(\mathbf{t}_i, \hat{\mathbf{t}}_k) \leq \frac{c_0}{\mathbf{m}}\}} \hat{\mathbf{a}}_k (q_{\mathbf{t}_i}(\hat{\mathbf{t}}_k) - 1) \right|, \\
& \leq 2(C' + 1)\lambda,
\end{aligned}$$

where $\mathbf{t}_i + \mathcal{B}(c_0/\mathbf{m}) = \{t \mid d(\mathbf{t}_i, t) \leq c_0/\mathbf{m}\}$. Finally, observe that (2) is a consequence of the aforementioned inequalities.

APPENDIX D. FENCHEL-LEGENDRE CONJUGATE AND FIRST ORDER CONDITIONS

Lemma 9. *The program:*

$$\text{(D.1)} \quad \min_{\mu \in \mathbf{C}_d(\mathbf{x})} \frac{1}{2} \|\mathbf{c}(\mu) - \mathbf{y}\|_2^2 + \lambda \|\mu\|_{TV},$$

has dual Fenchel-Legendre dual program:

$$\text{(D.2)} \quad \kappa - \min_{\|\sum_{k=0}^{\mathbf{m}} \alpha_k \varphi_k\|_{\infty} \leq \lambda} \left\{ \frac{1}{2} \|\alpha\|_2^2 + \sum_{k=d+1}^{\mathbf{m}} \alpha_k y_k \right\},$$

where κ is a constant. Moreover, there is no duality gap.

Proof. The case $\mathbf{d} = -1$ has been treated in [1]. Assume that $\mathbf{d} \geq 0$. Program (D.1) can be viewed as:

$$\min_{\mu \in \mathbf{M}} h(\mathbf{c}(\mu)) + \psi(\mu),$$

where the function $h(c) = (1/2) \|c - \mathbf{y}\|_2^2$ has Legendre conjugate:

$$\forall \alpha \in \mathbb{R}^{\mathbf{m}+1}, \quad h^*(\alpha) = \langle \alpha, \mathbf{y} \rangle + \frac{1}{2} \|\alpha\|_2^2,$$

and:

$$\psi(\mu) = \lambda \|\mu\|_{TV} + \mathbf{u}_{\mathbf{C}_d(\mathbf{x})}(\mu),$$

with $\mathbf{u}_{\mathbf{C}_d(\mathbf{x})}(\mu) = 0$ if $\mu \in \mathbf{C}_d(\mathbf{x})$ and ∞ otherwise. We can compute the Legendre conjugate of ψ as follows. Let $f \in \mathcal{C}([-1, 1])$. The convex conjugate ψ^* at point f

is defined by:

$$(D.3) \quad \psi^*(f) = \sup_{\mu \in \mathbf{M}} \int f d\mu - \lambda \|\mu\|_{TV} - \mathbf{1}_{\mathbf{C}_d(\mathbf{x})}(\mu) = \sup_{\mu \in \mathbf{C}_d(\mathbf{x})} \int f d\mu - \lambda \|\mu\|_{TV}.$$

Observe that $[-1, 1]$ is compact. Therefore there exists a point $t \in [-1, 1]$ such that $|f(t)| = \|f\|_\infty$. Since ψ^* is symmetric, we can assume that $f(t)$ is non-negative. Set:

$$\mu_\rho(du) = \rho \delta_t(u) - \rho \left[\sum_{k=0}^d (c_k(\delta_t) - \frac{c_k(\mathbf{x})}{\rho}) \varphi_k(u) \mathcal{L}(du) \right] := \rho \delta_t(u) - \rho \theta_\rho(u) \mathcal{L}(du).$$

Notice that $\mu_\rho \in \mathbf{C}_d(\mathbf{x})$ and so:

$$(D.4) \quad \psi^*(f) \geq \rho(f(t) - \lambda) + \rho \left(\int f(u) \theta_\rho(u) du - \lambda \|\theta_\rho\|_1 \right).$$

Set:

$$\theta = \sum_{k=0}^d c_k(\delta_t) \varphi_k,$$

which can be viewed as the convolution of δ_t with the idempotent $\sum_{k=0}^d \varphi_k$. This last kernel can be seen as a Dirichlet-type kernel and we claim that, if $\mathbf{d} \geq 1$, it takes positive and negative values on $[-1, 1]$. Since f is continuous and θ has at least a change of sign, Hölder's inequality can be written as:

$$(D.5) \quad \int f \theta < \|f\|_\infty \|\theta\|_1.$$

Moreover, remark that θ_ρ converges toward θ uniformly. Henceforth, if $\|f\|_\infty > \lambda$, we deduce from (D.4) and (D.5) that $\psi^*(f) = \infty$, letting ρ go to infinity. If $\mathbf{d} = 0$ then $\theta = 1$. In this case, Eq. (D.4) shows that $\psi^*(f) = \infty$ if $\|f\|_\infty > \lambda$. Moreover, remark that:

$$\mathcal{C}([-1, 1]) \subset L^2(\Pi) = \text{Span}\{\varphi_0, \dots, \varphi_d\} \oplus \text{Span}\{\varphi_0, \dots, \varphi_d\}^\perp := V \oplus V^\perp.$$

Using this decomposition it holds that for all $\mu \in \mathbf{C}_d(\mathbf{x})$,

$$\int f d\mu = \sum_{k=0}^d c_k(\mathbf{x}) \langle f, \varphi_k \rangle + \int \Pi_{V^\perp}(f) d\mu,$$

where Π_{V^\perp} is the orthogonal projection onto V^\perp . In particular, it holds

$$(D.6) \quad \forall f \in V, \quad \psi^*(f) = \mathbf{1}_{\{\|f\|_\infty \leq \lambda\}}(f) + \sum_{k=0}^d c_k(\mathbf{x}) \langle f, \varphi_k \rangle + \kappa_d(\mathbf{x}),$$

where $\kappa_d(\mathbf{x}) = -\lambda \inf_{\mu \in \mathbf{C}_d(\mathbf{x})} \|\mu\|_{TV}$ is a constant. Observe that the dual operator \mathbf{c}^* of \mathbf{c} is given by:

$$\forall \alpha \in \mathbb{R}^{\mathbf{m}+1}, \quad \mathbf{c}^*(\alpha) = \sum_{k=0}^{\mathbf{m}} \alpha_k \varphi_k.$$

Using the definition of conjugate $h(\mathbf{c}(\mu)) = \sup_{\alpha} \langle \alpha, \mathbf{c}(\mu) \rangle - h^*(\alpha)$, notice that:

$$\min_{\mu \in \mathbf{M}} h(\mathbf{c}(\mu)) + \psi(\mu) = - \inf_{\alpha \in \mathbb{R}^{\mathbf{m}+1}} h^*(\alpha) + \psi^*(-\mathbf{c}^*(\alpha)).$$

It follows that the program (D.1) has Fenchel-Legendre dual:

$$- \inf_{\alpha \in \mathbb{R}^{\mathbf{m}+1}} h^*(\alpha) + \psi^*(-\mathbf{c}^*(\alpha)) = \kappa_d(\mathbf{x}) - \inf_{\|\sum_{k=0}^{\mathbf{m}} \alpha_k \varphi_k\|_\infty \leq \lambda} \left\{ \frac{1}{2} \|\alpha\|_2^2 + \sum_{k=d+1}^{\mathbf{m}} \alpha_k y_k \right\}.$$

Since $\mathbf{C}_d(\mathbf{x})$ is an affine space, Slater's condition shows that strong duality holds. \square

Lemma 10. *The first order conditions read :*

$$(D.7) \quad \|\hat{P}\|_\infty \leq \lambda \quad \text{and} \quad \lambda \|\hat{\mathbf{x}}\|_{TV} \leq \int_{-1}^1 \hat{P} d(\hat{\mathbf{x}}) + \lambda \inf_{\mu \in \mathbf{C}_d(\mathbf{x})} \|\mu\|_{TV},$$

where:

$$\hat{P} = \sum_{k=0}^m (y_k - c_k(\hat{\mathbf{x}})) \varphi_k.$$

Proof. Let $\mu \in \mathbf{C}_d(\mathbf{x})$ and $\gamma \in (0, 1)$. Set $\nu = \hat{\mathbf{x}} + \gamma(\mu - \hat{\mathbf{x}})$ then, by convexity:

$$\|\mu\|_{TV} - \|\hat{\mathbf{x}}\|_{TV} \geq \frac{1}{\gamma} (\|\nu\|_{TV} - \|\hat{\mathbf{x}}\|_{TV}).$$

Observe that $\nu \in \mathbf{C}_d(\mathbf{x})$, by optimality:

$$\begin{aligned} \lambda (\|\nu\|_{TV} - \|\hat{\mathbf{x}}\|_{TV}) &\geq \frac{1}{2} (\|\mathbf{c}(\hat{\mathbf{x}}) - \mathbf{y}\|_2^2 - \|\mathbf{c}(\nu) - \mathbf{y}\|_2^2), \\ &= \gamma \langle \mathbf{y} - \mathbf{c}(\hat{\mathbf{x}}), \mathbf{c}(\mu) - \mathbf{c}(\hat{\mathbf{x}}) \rangle - \frac{\gamma^2}{2} \|\mathbf{c}(\mu) - \mathbf{c}(\hat{\mathbf{x}})\|_2^2. \end{aligned}$$

Letting γ go to 0, we deduce:

$$(D.8) \quad \forall \mu \in \mathbf{C}_d(\mathbf{x}), \quad \lambda (\|\mu\|_{TV} - \|\hat{\mathbf{x}}\|_{TV}) \geq \langle \mathbf{y} - \mathbf{c}(\hat{\mathbf{x}}), \mathbf{c}(\mu) - \mathbf{c}(\hat{\mathbf{x}}) \rangle.$$

Conversely, if (D.8) holds then, for all $\mu \in \mathbf{C}_d(\mathbf{x})$:

$$\begin{aligned} \frac{1}{2} \|\mathbf{c}(\mu) - \mathbf{y}\|_2^2 + \lambda \|\mu\|_{TV} &\geq \frac{1}{2} \|\mathbf{c}(\hat{\mathbf{x}}) - \mathbf{y} + \mathbf{c}(\mu) - \mathbf{c}(\hat{\mathbf{x}})\|_2^2 \\ &\quad + \langle \mathbf{y} - \mathbf{c}(\hat{\mathbf{x}}), \mathbf{c}(\mu) - \mathbf{c}(\hat{\mathbf{x}}) \rangle + \lambda \|\hat{\mathbf{x}}\|_{TV}, \\ &= \frac{1}{2} \|\mathbf{c}(\hat{\mathbf{x}}) - \mathbf{y}\|_2^2 + \lambda \|\hat{\mathbf{x}}\|_{TV} + \frac{1}{2} \|\mathbf{c}(\mu) - \mathbf{c}(\hat{\mathbf{x}})\|_2^2, \\ &\geq \frac{1}{2} \|\mathbf{c}(\hat{\mathbf{x}}) - \mathbf{y}\|_2^2 + \lambda \|\hat{\mathbf{x}}\|_{TV}. \end{aligned}$$

Therefore, Eq. (D.8) is a necessary and sufficient condition for the measure $\hat{\mathbf{x}}$ to be a solution to (1.3). In particular, it follows:

$$\lambda \|\hat{\mathbf{x}}\|_{TV} - \langle \mathbf{y} - \mathbf{c}(\hat{\mathbf{x}}), \mathbf{c}(\hat{\mathbf{x}}) \rangle \leq \inf_{\mu \in \mathbf{C}_d(\mathbf{x})} \{ \lambda \|\mu\|_{TV} - \langle \mathbf{y} - \mathbf{c}(\hat{\mathbf{x}}), \mathbf{c}(\mu) \rangle \} = -\psi^*(\hat{P}),$$

where ψ^* is defined by (D.3) and $\hat{P} = \sum_{k=d+1}^m (y_k - c_k(\hat{\mathbf{x}})) \varphi_k$. The optimality conditions can be deduced from (D.6). \square

Lemma 11. *Let $\hat{\mathbf{x}}$ be the solution to (1.3), then the following holds:*

$$(D.9) \quad \forall P \in \text{Span}(\mathcal{F}), \quad \left| \int_{-1}^1 P d(\hat{\mathbf{x}} - \mathbf{x}) \right| \leq (\lambda + \lambda_0) \|P\|_\infty,$$

where $\lambda_0 \geq \left\| \sum_{k=0}^m \varepsilon_k \varphi_k \right\|_\infty$.

Proof. Let $(a_k)_{k=0}^m$ be the coefficients of P , namely:

$$P = \sum_{k=0}^m a_k \varphi_k.$$

It holds:

$$\begin{aligned} \int_{-1}^1 P d(\hat{\mathbf{x}} - \mathbf{x}) &= \sum_{k=0}^m a_k \int_{-1}^1 \varphi_k d(\hat{\mathbf{x}} - \mathbf{x}) = \sum_{k=0}^m a_k (c_k(\hat{\mathbf{x}}) - c_k(\mathbf{x})), \\ &= \int_{-1}^1 \left(\sum_{k=0}^m a_k \varphi_k \right) \left(\sum_{k=0}^m (c_k(\hat{\mathbf{x}}) - c_k(\mathbf{x})) \varphi_k \right) d\Pi, \end{aligned}$$

since \mathcal{F} is an orthonormal family with respect to the probability measure Π . By Hölder's inequality, we get:

$$\left| \int_{-1}^1 P d(\hat{\mathbf{x}} - \mathbf{x}) \right| \leq \left\| \sum_{k=0}^{\mathbf{m}} (c_k(\hat{\mathbf{x}}) - c_k(\mathbf{x})) \varphi_k \right\|_{\infty} \int_{-1}^1 |P| d\Pi.$$

The triangular inequality gives:

$$\left\| \sum_{k=0}^{\mathbf{m}} (c_k(\hat{\mathbf{x}}) - c_k(\mathbf{x})) \varphi_k \right\|_{\infty} \leq \left\| \sum_{k=0}^{\mathbf{m}} (c_k(\hat{\mathbf{x}}) - y_k) \varphi_k \right\|_{\infty} + \left\| \sum_{k=0}^{\mathbf{m}} (y_k - c_k(\mathbf{x})) \varphi_k \right\|_{\infty}.$$

The result follows from (D.7). \square

APPENDIX E. MATRIX FORMULATION OF THE SPLINE RECOVERY PROBLEM

By induction, for $k = 0, 1, \dots, \mathbf{m}$,

$$c_k(\mathbf{f}^{(\mathbf{d}+1)}) = \langle \mathbf{f}^{(\mathbf{d}+1)}, \varphi_k \rangle = \sum_{l=0}^{\mathbf{d}} (-1)^l [\mathbf{f}^{(\mathbf{d}-l)} \varphi_k^{(l)}]_{-1}^1 + (-1)^{\mathbf{d}+1} \langle \mathbf{f}, \varphi_k^{(\mathbf{d}+1)} \rangle.$$

Moreover, it is known that for $k \geq l$, $T_k^{(l)}(-1) = (-1)^{k+l} \mathbf{w}_{k,l}$ and $T_k^{(l)}(1) = \mathbf{w}_{k,l}$ where:

$$\mathbf{w}_{k,l} := \prod_{j=0}^{l-1} \frac{k^2 - j^2}{2j+1}.$$

Therefore, for $\mathbf{m} \geq k > \mathbf{d}$,

$$\begin{aligned} \text{(E.1)} \quad c_k(\mathbf{f}^{(\mathbf{d}+1)}) &= \sqrt{2} \sum_{l=0}^{\mathbf{d}} (-1)^l \mathbf{w}_{k,l} \mathbf{P}_s^{(\mathbf{d}-l)}(1) \\ &\quad + (-1)^{k+1} \sqrt{2} \sum_{l=0}^{\mathbf{d}} \mathbf{w}_{k,l} \mathbf{P}_0^{(\mathbf{d}-l)}(-1) + (-1)^{\mathbf{d}+1} \langle \mathbf{f}, \varphi_k^{(\mathbf{d}+1)} \rangle, \end{aligned}$$

for $\mathbf{d} \geq k \geq 1$,

$$\begin{aligned} \text{(E.2)} \quad c_k(\mathbf{f}^{(\mathbf{d}+1)}) &= \sqrt{2} \sum_{l=0}^k (-1)^l \mathbf{w}_{k,l} \mathbf{P}_s^{(\mathbf{d}-l)}(1) \\ &\quad + (-1)^{k+1} \sqrt{2} \sum_{l=0}^k \mathbf{w}_{k,l} \mathbf{P}_0^{(\mathbf{d}-l)}(-1), \end{aligned}$$

and

$$\text{(E.3)} \quad c_0(\mathbf{f}^{(\mathbf{d}+1)}) = \mathbf{P}_s^{(\mathbf{d})} - \mathbf{P}_0^{(\mathbf{d})}.$$

APPENDIX F. PROOF OF THEOREM 2

From (1.5) deduce that if P satisfies **Assumption 1** then:

$$\mathbf{y} := \begin{bmatrix} 0 & W_1 \\ (-1)^{\mathbf{d}+1} \text{Id}_{\mathbf{m}-\mathbf{d}} & W_2 \end{bmatrix} \begin{pmatrix} \Theta(P) \\ \mathbf{b} \end{pmatrix} = \mathbf{c}(\mathbf{f}^{(\mathbf{d}+1)}) + (-1)^{\mathbf{d}+1} \begin{pmatrix} 0 \\ \mathbf{n} \end{pmatrix},$$

where W_1 and W_2 are described in Appendix E. Observe the result follows from Theorem 1.

Acknowledgments. YdC would like to thank V. Duval and G. Peyré for valuable discussions on Super-Resolution.

REFERENCES

- [1] J.-M. Azaïs, Y. de Castro, and F. Gamboa. Spike detection from inaccurate samplings. *arXiv preprint arXiv:1301.5873*, 2013.
- [2] J.-M. Azaïs and M. Wschebor. A general expression for the distribution of the maximum of a gaussian field and the approximation of the tail. *Stochastic Processes and their Applications*, 118(7):1190 – 1218, 2008.
- [3] T. Bendory, S. Dekel, and A. Feuer. Exact recovery of non-uniform splines from the projection onto spaces of algebraic polynomials. 2013.
- [4] P. Borwein, T. Erdélyi, and J. Zhang. Chebyshev polynomials and markov–bernstein type inequalities for rational spaces. *Journal of the London Mathematical Society*, 50(3):501–519, 1994.
- [5] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse problems*, 20(5):1411, 2004.
- [6] E.J. Candes and C. Fernandez-Granda. Super-resolution from noisy data. *arXiv preprint arXiv:1211.0290*, 2012.
- [7] E.J. Candes and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 2012.
- [8] Y. de Castro and F. Gamboa. Exact reconstruction using beurling minimal extrapolation. *Journal of Mathematical Analysis and Applications*, 2012.
- [9] D.L. Donoho. Superresolution via sparsity constraints. *SIAM Journal on Mathematical Analysis*, 23(5):1309–1331, 1992.
- [10] B. Dumitrescu. *Positive trigonometric polynomials and signal processing applications*. Springer, 2007.
- [11] V. Duval and G. Peyré. Exact support recovery for sparse spikes deconvolution. *arXiv preprint arXiv:1306.6909*, 2013.
- [12] C. Fernandez-Granda. Support detection in super-resolution. *arXiv preprint arXiv:1302.3921*, 2013.
- [13] D. Gottlieb and S. A. Orszag. *Numerical analysis of spectral methods*, volume 2. SIAM, 1977.
- [14] A. Henrot and M. Pierre. *Variation et optimisation de formes: une analyse géométrique*, volume 48. Springer, 2006.
- [15] E. Kalnay. *Atmospheric modeling, data assimilation, and predictability*. Cambridge university press, 2003.
- [16] M. G. Krein and A. A. Nudelman. The markov moment problem and extremal problems, volume 50 of translations of mathematical monographs. *American Mathematical Society, Providence, Rhode Island*, 1977.
- [17] G. Tang, B. Narayan Bhaskar, and B. Recht. Near minimax line spectral estimation. *arXiv preprint arXiv:1303.4348*, 2013.

YDC IS WITH THE DÉPARTEMENT DE MATHÉMATIQUES (CNRS UMR 8628), BÂTIMENT 425, FACULTÉ DES SCIENCES D’ORSAY, UNIVERSITÉ PARIS-SUD 11, F-91405 ORSAY CEDEX, FRANCE.

E-mail address: `yohann.decastro@math.u-psud.fr`

URL: `www.math.u-psud.fr/~decastro`

GM IS WITH THE DÉPARTEMENT DE MATHÉMATIQUES-INFORMATIQUE, UNIVERSITÉ PARIS OUEST NANTERRE LA DÉFENSE, FRANCE.

E-mail address: `guillaume.mijoule@u-paris10.fr`