Off-the-Grid Estimation of Singular Measures

Paul Catala*, Mathias Hockmann*, Stefan Kunis* and Markus Wageringel*
* Institute of Mathematics, Osnabrück University
{paul.catala,mathias.hockmann,stefan.kunis,markus.wageringel}@uos.de

***Abstract*—Many problems in imaging science involve reconstructing, from partial observations, highly concentrated signals, *e.g.* pointwise sources or contour lines. We consider in this work the problem of recovering measures supported on such singular sets, given finitely many of their trigonometric moments. We introduce simple polynomial estimates, and prove pointwise and weak convergence as the number n of known moments increases. We show that the optimal weak convergence rate with respect to the Wasserstein-1-distance is inversely proportional to n, and that it is achieved by our estimate.**

I. Introduction

Signed Radon measures are a relevant model in many applications, *e.g.* single-molecule localization microscopy or X-ray crystallography, as they can capture diverse and complex structures, such as points, curves or manifolds for instance. In this work, we consider the following problem, ubiquitous in imaging science. Let $\mathcal{M}$ be the space of (signed) Radon measures on the d-dimensional torus. Given trigonometric moments of $\mu \in \mathcal{M}$ up to some order $n \in \mathbb{N}$, *i.e.*

$$\hat{\mu}(k) \overset{\text{def.}}{=} \int e^{-2\imath\pi\langle k,\, x\rangle} \mathrm{d}\mu(x), \quad k \in \{-n, \dots, n\}^d,$$

how can one recover or approximate the measure μ?

When μ is discrete, a wide variety of methods exist, that are able to retrieve its parameters from a few of its moments. Examples include subspace methods [1], [2], [3] and TV-minimization approaches [4]. The general case on the other hand is more involved. Finite rate of innovation approaches generalize subspace methods to continuous settings, but are restricted to specific curves [5], [6]. In another line of work, one may also focus on approximation rather than exact recovery. When the measure satisfies some regularity conditions, Christoffel functions provide approximations of its density, with guarantees with respect to the uniform norm [7], [8].

Similarly to [9], our approach requires no assumptions on the measure, except that it has finite total variation. We provide polynomial estimates and tight bounds on the pointwise approximation as well as with respect to the Wasserstein-1 distance. For details and proof we refer the interested reader to [10].

Given $\mu, \nu \in \mathcal{M}$ of equal mass, their Wasserstein-1 distance is

$$\mathcal{W}_1(\mu, \nu) \overset{\text{def.}}{=} \sup_{\mathrm{Lip}(f) \leqslant 1} \left| \int_{\mathbb{T}^d} f(x) \mathrm{d}(\mu - \nu)(x) \right|,$$

where the Lipschitz constant is defined with respect to the wrap-around distance $d(x, y) = \min_{k \in \mathbb{Z}^d} \|x - y + k\|_1$. The $\mathcal{W}_1$-distance is well-defined for signed measures, and metrizes weak convergence.

Given $\mu \in \mathcal{M}$ and $n \in \mathbb{N}$, the moment matrix of μ of order n is

$$T_n \overset{\text{def.}}{=} \left[\hat{\mu}(k - l)\right]_{k,l \in \{0,\dots,n\}^d}.$$

We identify a vector $p \in \mathbb{C}^{(n+1)^d}$ with the trigonometric polynomial $p(x) = \sum p_k e^{-2\imath\pi\langle k,\, x\rangle}$.

II. Main results

A. Polynomial approximations

Let $\mu \in \mathcal{M}$. We consider the following polynomial

$$p_n = F_n * \mu,$$

where $F_n(x) \overset{\text{def.}}{=} (n+1)^{-d} \prod_j \frac{\sin^2((n+1)\pi x_j)}{\sin^2(\pi x_j)}$ is the Fejér kernel. It can be computed efficiently using only Fast Fourier Transforms.

1) An upper bound: The next result establishes the weak convergence of the measure with density p_n towards μ.

Theorem 1 (Fejér approximation)**.** *For all $n \in \mathbb{N}_*$, we have*

$$\mathcal{W}_1(\mu, p_n) \leqslant \frac{d}{\pi^2} \frac{\log(n+1) + 3}{n}. \tag{1}$$

2) Sharpness: The previous rate in $O(n^{-1}\log(n))$ is matched for $\mu = \delta_0$. For an arbitrary measure, however regular it may be, a rate of order n^{-1} is in any case the best that one can expect for p_n.

Theorem 2 (Saturation)**.** *For every $\mu \in \mathcal{M}$ not being the Lebesgue measure, there exists a constant $c(\mu, d)$ such that*

$$\mathcal{W}_1(\mu, p_n) \geqslant \frac{c(\mu, d)}{n+1}.$$

3) Better estimates: Without any further assumption on the measure, the rate of n^{-1} in the previous lower bound is actually tight.

Theorem 3 (Best polynomial approximation)**.** *For any $d, n \in \mathbb{N}$ and $\mu \in \mathcal{M}$, there exists a polynomial of best approximation $p^*(\mu)$ (of degree at most n) in the Wasserstein-1 distance. It satisfies*

$$\sup_{\mu \in \mathcal{M}} \mathcal{W}_1(\mu, p^*(\mu)) \geqslant \frac{1}{4(n+1)}. \tag{2}$$

The gap between (1) and (2) can be filled, *e.g.* by considering convolution with the Jackson kernel $J_{2n} \overset{\text{def.}}{=} \alpha_n F_n^2(x)$ (where α_n is a normalizing constant). Note that these rates can still be improved, at the cost of additional regularity assumptions on the measure.

B. Polynomial interpolation

We now assume that μ is nonnegative. Let $r = \operatorname{rank} T_n$, and let $u_j^{(n)}$ be the singular vectors (seen as polynomials) of T_n. We define

$$p_{1,n} \overset{\text{def.}}{=} \frac{1}{(n+1)^d} \sum_{j=1}^{r} |u_j^{(n)}|^2.$$

Note that $p_{1,n} \leqslant 1$. If n is large enough, $p_{1,n}$ identifies the support.

Theorem 4. *Let $V \overset{\text{def.}}{=} \overline{\operatorname{Supp} \mu}^Z$ be the Zariski closure of the support. For n large enough (and $\mu \geqslant 0$), $p_{1,n}(x) = 1$ if and only if $x \in V$.*

Theorem 4 combined with the next result proves that $p_{1,n}$ converges pointwisely towards the characteristic function of the support.

Theorem 5. *Assume that $V \neq \mathbb{T}^d$. Let $y \in \mathbb{T}^d \setminus V$ and let g be a polynomial of degree m such that $g(y) \neq 0$ and g vanishes on $\operatorname{Supp} \mu$. Then, for all $n \geqslant m$,*

$$p_{1,n+m}(y) \leqslant \frac{\|g\|_{\mathrm{L}^2}^2}{|g(y)|} \frac{m(4m+2)^d}{n+1} + \frac{dm}{n+m+1}.$$

III. Conclusion

Handling measures via a limited number of moments is key to computational efficiency in recovery algorithms. Our results provide easily computable proxies for general measures on the torus, which come with sharp bounds in terms of weak and pointwise convergence.

References

[1] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoustics Speech Signal Process.*, vol. 37, no. 7, pp. 984–995, 1989.

[2] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[3] Y. Hua and T. Sarkar, "Generalized pencil-of-function method for extracting poles of an em system from its transient response," *IEEE Trans. Antennas Propagation*, vol. 37, no. 2, 1989.

[4] E. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Comm. Pure Appl. Math.*, vol. 67, no. 6, pp. 906–956, 2014.

[5] H. Pan, T. Blu, and P. Dragotti, "Sampling curves with finite rate of innovation," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 458–471, 2014.

[6] G. Ongie and M. Jacob, "Off-the-grid recovery of piecewise constant images from few fourier samples," *SIAM J. Imaging Sci.*, vol. 9, no. 3, pp. 1004–1041, 2016.

[7] E. Pauwels, M. Putinar, and J. Lasserre, "Data analysis from empirical moments and the Christoffel function," *F. Comp. Math.*, 2020.

[8] S. Marx, E. Pauwels, T. Weisser, D. Henrion, and J. Lasserre, "Tractable semi-algebraic approximation using Christoffel-darboux kernel," *Constr. Approx.*, vol. 54, pp. 391–429, 2021.

[9] H. N. Mhaskar, "Super-resolution meets machine learning: Approximation of measures," *J. Fourier Anal. Appl.*, vol. 25, no. 6, pp. 3104–3122, 2019.

[10] P. Catala, M. Hockmann, S. Kunis, and M. Wageringel, "Approximation and interpolation of singular measures by trigonometric polynomials," 2022.