

SUPER-RESOLUTION FROM SHORT-TIME FOURIER TRANSFORM MEASUREMENTS

Céline Aubel, David Stotz, and Helmut Bölcskei

Dept. IT&EE, ETH Zurich, Switzerland
{aubelc, dstotz, boelcskei}@nari.ee.ethz.ch

ABSTRACT

While spike trains are obviously not band-limited, the theory of super-resolution tells us that perfect recovery of unknown spike locations and weights from low-pass Fourier transform measurements is possible provided that the minimum spacing, Δ , between spikes is not too small. Specifically, for a cutoff frequency of f_c , Donoho [2] shows that exact recovery is possible if $\Delta > 1/f_c$, but does not specify a corresponding recovery method. On the other hand, Candès and Fernandez-Granda [3] provide a recovery method based on convex optimization, which provably succeeds as long as $\Delta > 2/f_c$. In practical applications one often has access to windowed Fourier transform measurements, i.e., short-time Fourier transform (STFT) measurements, only. In this paper, we develop a theory of super-resolution from STFT measurements, and we propose a method that provably succeeds in recovering spike trains from STFT measurements provided that $\Delta > 1/f_c$.

Index Terms— Super-resolution, inverse problems in measure spaces, short-time Fourier transform.

1. INTRODUCTION

The recovery of spike trains with unknown spike locations and weights from low-pass Fourier measurements, commonly referred to as super-resolution, has been a topic of long-standing interest [4–11], with recent focus on ℓ^1 -minimization-based recovery techniques [3, 12, 13]. It was recognized in [2, 3, 14–16] that a measure-theoretic formulation of the super-resolution problem in continuous-time not only leads to a clean mathematical framework, but also to results that are “grid-free” [17]. This theory is inspired by Beurling’s seminal work on the “balayage” of measures in Fourier transforms [6, 7], and on interpolation using entire functions of exponential type [8]. Specifically, Castro and Gamboa [15], and Candès and Fernandez-Granda [3], propose to solve a total variation minimization problem for recovering a discrete measure (modeling the spike train) from low-pass Fourier measurements. Despite its infinite-dimensional nature, this optimization problem can be solved

exploiting Fenchel duality, as described in [3, 14, 16]. Concretely, it is shown in [3, 15] that the analysis of the Fenchel dual problem leads to an interpolation problem, which can be solved explicitly as long as the elements in the support of the measure to be recovered are separated by at least $2/f_c$, where f_c represents the cutoff frequency of the low-pass measurements. Donoho [2] proves that a separation of $1/f_c$ is sufficient for perfect recovery, but does not provide a concrete method for reconstructing the measure. Kahane [1] shows that recovery is possible if the minimum separation between spikes is at least $\frac{5}{f_c} \sqrt{\log(1/f_c)}$, but this result suffers from a log-factor penalty.

Contributions: In practical applications one often has access to windowed Fourier transform, i.e., short-time Fourier transform (STFT), measurements only. It is therefore of interest to develop a theory of super-resolution from STFT measurements. This is precisely the goal of the present paper. Inspired by [2, 3, 14–16, 18], we use a measure-theoretic formulation and consider the continuous-time case. Our main result shows that exact recovery through convex optimization is possible, for a Gaussian window function, provided that the minimum separation between points in the measure support exceeds $1/f_c$.

Notation and preparatory material: The complex conjugate of $z \in \mathbb{C}$ is denoted by \bar{z} . The derivative of the function φ is designated by φ' . The sinc function is defined as $\text{sinc}(t) := \sin(t)/t$ for all $t \neq 0$ and $\text{sinc}(0) = 1$. Uppercase boldface letters stand for matrices. The entry in the k th row and ℓ th column of the matrix \mathbf{M} is $m_{k,\ell}$. The superscript H stands for Hermitian transposition. For matrices $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{M \times N}$, we write $\langle \mathbf{X}, \mathbf{Y} \rangle := \text{Re} \{ \text{Tr}(\mathbf{Y}^H \mathbf{X}) \}$ for their real inner product. General linear operators are designated by uppercase calligraphic letters. If X and Y are topological vector spaces, and X^* and Y^* their topological duals, the adjoint of the linear operator $\mathcal{L}: X \rightarrow Y$ is denoted by $\mathcal{L}^*: Y^* \rightarrow X^*$. The set of all solutions of an optimization problem (P) is denoted by $\text{Sol}\{\text{P}\}$. For a measure space (X, Σ, μ) and a measurable function $\varphi: X \rightarrow \mathbb{C}$, we write $\int_X \varphi(x) \mu(x)$ for the integration of φ with respect to μ , where we set $dx := \lambda(x)$ if λ is the Lebesgue measure. For $p \in [1, \infty)$, $L^p(X, \Sigma, \mu)$ denotes the space of all functions $\varphi: X \rightarrow \mathbb{C}$ such that $\|\varphi\|_{L^p} := (\int_X |\varphi(x)|^p \mu(x))^{1/p} < \infty$. The space $L^\infty(X, \Sigma, \mu)$ contains all functions $\varphi: X \rightarrow \mathbb{C}$ such that $\|\varphi\|_{L^\infty} := \inf\{C > 0: |\varphi(x)| \leq C \text{ for } \mu\text{-almost all } x \in X\}$.

The authors would like to thank J.-P. Kahane for answering questions on results in [1], M. Lerjen for his help with simulations, and C. Chenot for inspiring discussions.

$X\} < \infty$. For functions $\varphi \in L^p(X, \Sigma, \mu)$ and $\psi \in L^q(X, \Sigma, \mu)$ with $p, q \in [1, \infty]$ satisfying $1/p + 1/q = 1$, we set $\langle \varphi, \psi \rangle := \operatorname{Re} \left\{ \int_X \varphi(x) \overline{\psi(x)} \mu(x) \right\}$. For a separable locally compact metric abelian group G (e.g., the additive group \mathbb{R} or the torus $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ endowed with the natural topology), we write $L^p(G)$ in the particular case where Σ is the Borel σ -algebra of G and μ the Haar measure on G . We denote by $\mathcal{B}(G)$ the Borel σ -algebra of G and by $\mathcal{M}(G)$ the space of all complex Radon measures on $(G, \mathcal{B}(G))$. For $t \in G$, $\delta_t \in \mathcal{M}(G)$ designates the Dirac measure at t , which for $B \in \mathcal{B}(G)$ is given by $\delta_t(B) = 1$ if $t \in B$ and $\delta_t(B) = 0$ otherwise. The support $\operatorname{supp}(\mu)$ of a complex Radon measure $\mu \in \mathcal{M}(G)$ is the largest closed set $C \subseteq G$ such that for every open set $B \in \mathcal{B}(G)$ satisfying $B \cap C \neq \emptyset$, it holds that $\mu(B \cap C) \neq 0$. For every $\mu \in \mathcal{M}(G)$, the total variation (TV) of μ is defined as the measure $|\mu|$ satisfying

$$\forall B \in \mathcal{B}(G), \quad |\mu|(B) := \sup_{\pi \in \Pi(B)} \sum_{A \in \pi} |\mu(A)|,$$

where $\Pi(B)$ denotes the set of all partitions of B . The space $\mathcal{M}(G)$ can be equipped with the TV norm $\|\mu\|_{\text{TV}} := |\mu|(G)$. Based on the Riesz representation theorem [19, Thm. 6.19], $\mathcal{M}(G)$ can be characterized as the dual space of $C_0(G)$. With $C_c(G)$ the space of all complex-valued continuous functions on G whose support is compact, $C_0(G)$ is the completion of $C_c(G)$ relative to the metric defined by the supremum norm $\|\varphi\|_\infty = \sup_{t \in G} |\varphi(t)|$. By analogy with the real inner product in $L^2(G)$, we define the real dual pairing of the measure $\mu \in \mathcal{M}(G)$ and the function $\varphi \in C_0(G)$ as $\langle \mu, \varphi \rangle := \operatorname{Re} \left\{ \int_G \overline{\varphi(t)} \mu(t) \right\}$. We endow $\mathcal{M}(G)$ with the weak-* topology [20], i.e., the coarsest topology on $\mathcal{M}(G)$ for which every linear functional $\mathcal{L}_\varphi: \mathcal{M}(G) \rightarrow \mathbb{R}$ defined by $\mu \mapsto \langle \mu, \varphi \rangle = \langle \mu, \varphi \rangle$, with $\varphi \in C_0(G)$, is continuous.

2. STATEMENT OF THE PROBLEM

We consider a complex Radon measure on \mathbb{R} of the form

$$\mu = \sum_{\ell \in \Omega} a_\ell \delta_{t_\ell}. \quad (1)$$

The measure models a spike train, and is supported on the closed discrete set $T := \{t_\ell\}_{\ell \in \Omega} \subseteq \mathbb{R}$ with complex mass $a_\ell \neq 0$ attached to the point t_ℓ . The locations t_ℓ and the associated a_ℓ are assumed unknown throughout the paper. We do, however, assume that the measure μ is known to be discrete. The set Ω is assumed to be countable, and we require that $\sum_{\ell \in \Omega} |a_\ell| < \infty$ for μ to be in $\mathcal{M}(\mathbb{R})$. Throughout the paper, μ designates exclusively the measure defined in (1).

Suppose we obtain measurements of μ in the time-frequency domain in the form:

$$y(\tau, f) = (\mathcal{V}_g \mu)(\tau, f)$$

for $\tau \in \mathbb{R}$ and $f \in B_{f_c} := \{f \in \mathbb{R} : |f| \leq f_c\}$, where f_c is the cutoff frequency and

$$(\mathcal{V}_g \mu)(\tau, f) := \int_{\mathbb{R}} \overline{g(t - \tau)} e^{-2\pi i f t} \mu(t) \quad (2)$$

denotes the STFT [21] of μ with respect to the window function g . Although our theory applies to more general window functions taken from the Schwartz space of rapidly decaying functions, for concreteness, we choose g to be Gaussian, i.e.,

$$\forall t \in \mathbb{R}, \quad g(t) = \frac{1}{\sqrt{\sigma}} \exp\left(-\frac{t^2}{2\sigma^2}\right),$$

where $\sigma > 0$ is a parameter controlling the width of the window. We wish to recover the measure μ from the measurements y by solving the following optimization problem:

$$\text{(SR)} \quad \underset{\nu \in \mathcal{M}(\mathbb{R})}{\text{minimize}} \quad \|\nu\|_{\text{TV}} \quad \text{subject to } y = \mathcal{A}_g \nu,$$

where $\mathcal{A}_g: \mathcal{M}(\mathbb{R}) \rightarrow L^1(\mathbb{R}^2)$ maps $\nu \in \mathcal{M}(\mathbb{R})$ to the function $\varphi \in L^1(\mathbb{R}^2)$ such that

$$\forall(\tau, f) \in \mathbb{R}^2, \quad \varphi(\tau, f) = \begin{cases} (\mathcal{V}_g \nu)(\tau, f), & f \in B_{f_c} \\ 0, & \text{otherwise.} \end{cases}$$

The motivation for considering time-frequency measurements is twofold. First, signals are often partitioned into short time segments and windowed for acquisition. Second, the frequency characteristics of the signal modeled by the measure μ often vary over time, i.e., the $t_\ell, \ell \in \Omega$, can be more packed in certain intervals, so that time-localized spectral information about μ will lead to improved reconstruction quality for the same frequency limitation.

3. RECONSTRUCTION FROM COMPLETE MEASUREMENTS

Before embarking on the problem of reconstructing μ by solving (SR) for a given $f_c < \infty$, we need to convince ourselves that reconstruction is possible from complete measurements, i.e., for $f_c = \infty$. This should be evident as the STFT, defined for functions in $L^2(\mathbb{R})$, is invertible [21]. The measure-theoretic STFT considered here is, however, non-standard. While one can show that the STFT of a measure still determines the underlying measure uniquely in the sense that, for $\nu \in \mathcal{M}(\mathbb{R})$, $\mathcal{V}_g \nu = 0$ implies $\nu = 0$, to the best of our knowledge, no general inversion formula is available for complex Radon measures. Since μ is discrete, all we need to find is its support $T = \{t_\ell\}_{\ell \in \Omega}$ and the corresponding complex masses $\{a_\ell\}_{\ell \in \Omega}$. Specifically, it can be shown that μ can be recovered from its STFT $\mathcal{V}_g \mu$ according to

$$\begin{aligned} \lim_{F \rightarrow \infty} \frac{1}{2F} \int_{-F}^F \int_{\mathbb{R}} (\mathcal{V}_g \mu)(\tau, f) g(t - \tau) e^{2\pi i f t} d\tau df \\ = \begin{cases} a_\ell, & \text{if } t = t_\ell \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

4. RECONSTRUCTION FROM PARTIAL MEASUREMENTS

Now, we consider the reconstruction of μ from band-limited STFT measurements using (SR). Since the space $\mathcal{M}(\mathbb{R})$ is

infinite-dimensional, the existence of a solution of (SR) is delicate. It turns out, however, that relying on the convexity of the TV norm $\|\cdot\|_{\text{TV}}$ and on compactness of the unit ball $\{\nu \in \mathcal{M}(G) : \|\nu\|_{\text{TV}} \leq 1\}$ with respect to the weak-* topology, the result provided in [20, Cor. 3.20] ensures the existence of a solution of (SR). Next, with the help of Fenchel duality theory [22, Chap. 4], we derive necessary and sufficient conditions for μ to be the unique solution of (SR).

Theorem 1 (Fenchel predual) *The Fenchel predual problem of (SR) is*

$$(PD\text{-}SR) \quad \underset{c \in L^\infty(\mathbb{R}^2)}{\text{maximize}} \langle c, y \rangle \quad \text{subject to} \quad \|\mathcal{A}_g^* c\|_\infty \leq 1.$$

In addition, the following equality holds

$$\begin{aligned} \min \left\{ \|\nu\|_{\text{TV}} : \mathcal{A}_g \nu = y, \nu \in \mathcal{M}(\mathbb{R}) \right\} \\ = \sup \left\{ \langle c, y \rangle : \|\mathcal{A}_g^* c\|_\infty \leq 1, c \in L^\infty(\mathbb{R}^2) \right\}. \end{aligned} \quad (3)$$

Moreover, if (PD-SR) has a solution $c_0 \in L^\infty(\mathbb{R}^2)$, then

$$\bigcup_{\nu_0 \in \text{Sol}\{(\text{SR})\}} \text{supp}(\nu_0) \subseteq \{t \in \mathbb{R} : |(\mathcal{A}_g^* c_0)(t)| = 1\}. \quad (4)$$

Theorem 1 follows by application of [22, Thms. 4.4.2 and 4.4.3] similarly to what was done in [14, Prop. 2]. We emphasize that (PD-SR) is the *predual* problem of (SR), meaning that (SR) is the dual problem of (PD-SR). The dual problem of (SR) is, however, not (PD-SR) as the space $L^1(\mathbb{R}^2)$ is not reflexive.

The consequences of Theorem 1 are the following: Assuming that (PD-SR) has a solution, which we denote by $c_0 \in L^\infty(\mathbb{R}^2)$, the support $T = \{t_\ell\}_{\ell \in \Omega}$ of the measure μ to be recovered must satisfy $|(\mathcal{A}_g^* c_0)(t_\ell)| = 1$ if μ is to be in the set of solutions of (SR). Furthermore, if (PD-SR) has a solution $c_0 \in L^\infty(\mathbb{R}^2)$ such that $|\mathcal{A}_g^* c_0|$ is not identically 1 on \mathbb{R} , then every solution to (SR) is a discrete measure despite the fact that (SR) is an optimization problem over the space of *all* complex Radon measures. This can be seen as follows: Both the window function g and $\mathcal{A}_g^* c_0$ can be extended to entire functions which we also denote by g and $\mathcal{A}_g^* c_0$, i.e.,

$$\begin{aligned} \forall z \in \mathbb{C}, \quad g(z) &= \frac{1}{\sqrt{\sigma}} \exp\left(-\frac{\pi z^2}{2\sigma^2}\right) \quad \text{and} \\ (\mathcal{A}_g^* c_0)(z) &= \int_{-f_c}^{f_c} \int_{\mathbb{R}} c_0(\tau, f) g(z - \tau) e^{2\pi i f z} d\tau df. \end{aligned} \quad (5)$$

We can then define the function

$$\forall z \in \mathbb{C}, \quad h(z) := 1 - (\mathcal{A}_g^* c_0)(z) \overline{(\mathcal{A}_g^* c_0)(\bar{z})}.$$

Since $|\mathcal{A}_g^* c_0|$ is not identically 1 on \mathbb{R} , h is not identically zero. Consequently, on the basis of [19, Thm. 10.18] the set $\{z \in \mathbb{C} : h(z) = 0\}$, and a fortiori the set $\{t \in \mathbb{R} : |\mathcal{A}_g^* c_0| = 1\}$, are at most countable and have no limit points. But since (4) holds, this implies that any solution ν_0 to (SR) must have

discrete support, and therefore, ν_0 is necessarily a discrete measure.

Similarly to [16, Sec. 2.4], the following theorem provides a necessary and sufficient condition for μ to be a solution of (SR).

Theorem 2 (Optimality conditions) *The measure μ to be recovered is in the set of solutions of (SR) if and only if there exists $c_0 \in L^\infty(\mathbb{R}^2)$ such that*

$$\|\mathcal{A}_g^* c_0\|_\infty \leq 1 \quad \text{and} \quad \forall \ell \in \Omega, \quad (\mathcal{A}_g^* c_0)(t_\ell) = \frac{a_\ell}{|a_\ell|}.$$

It is important to note that (PD-SR) has at least one solution if the measure μ to be recovered is a solution of (SR), as the sup in (3) is attained for c_0 . Indeed, the following holds:

$$\begin{aligned} \langle c_0, y \rangle &= \langle c_0, \mathcal{A}_g \mu \rangle = \langle \mathcal{A}_g^* c_0, \mu \rangle = \int_G \overline{(\mathcal{A}_g^* c_0)(t)} \mu(t) \\ &= \sum_{\ell \in \Omega} a_\ell \overline{(\mathcal{A}_g^* c_0)(t_\ell)} = \sum_{\ell \in \Omega} |a_\ell| = \|\mu\|_{\text{TV}}. \end{aligned}$$

Since $\|\mathcal{A}_g^* c_0\|_\infty \leq 1$, c_0 is a solution of (PD-SR).

Theorem 2 provides conditions on μ to be a solution of (SR). However, we hope for more, namely, we want conditions on μ to be the *unique* solution of (SR). Such conditions are given in the following theorem, which is a straightforward adaptation of [3, App. A].

Theorem 3 (Uniqueness) *If for every sequence $\varepsilon = \{\varepsilon_\ell\}_{\ell \in \Omega}$ of unit magnitude complex numbers, there exists a function $c_0 \in L^\infty(\mathbb{R}^2)$ obeying*

$$\forall \ell \in \Omega, \quad (\mathcal{A}_g^* c_0)(t_\ell) = \varepsilon_\ell \quad (6)$$

$$\forall t \in \mathbb{R} \setminus T, \quad |(\mathcal{A}_g^* c_0)(t)| < 1, \quad (7)$$

then μ is the unique solution of (SR).

Verifying the conditions of Theorem 3 requires solving constrained interpolation problems associated with the support set $T = \{t_\ell\}_{\ell \in \Omega}$ of μ , as specified by (6) and (7). The following theorem provides conditions ensuring that these solutions can be given in explicit form.

Theorem 4 (Exact recovery) *Let $\sigma = \frac{1}{4f_c}$. If the minimum distance Δ between any two points of T ,*

$$\Delta = \inf_{\substack{\ell, \ell' \in \Omega \\ \ell \neq \ell'}} |t_\ell - t_{\ell'}|,$$

satisfies $\Delta > 1/f_c$, then the conditions of Theorem 3 are met, and, hence, μ is the unique solution of (SR).

We next briefly describe the main ingredients of the proof of Theorem 4, which is mostly inspired by [3, Sec. 2, pp. 15–27]. The proof is accomplished by verifying the conditions of Theorem 3. To this end, we fix a sequence $\varepsilon = \{\varepsilon_\ell\}_{\ell \in \Omega}$ of

complex unit-magnitude numbers, and take a function $c_0 \in L^\infty(\mathbb{R}^2)$ of the form

$$c_0(\tau, f) := \sum_{\ell \in \Omega} [\alpha_\ell g(t_\ell - \tau) e^{-2\pi i f t_\ell} + \beta_\ell g'(t_\ell - \tau) e^{-2\pi i f t_\ell}]$$

for $(\tau, f) \in \mathbb{R}^2$, where $\alpha = \{\alpha_\ell\}_{\ell \in \Omega}$ and $\beta = \{\beta_\ell\}_{\ell \in \Omega}$ are sequences in $\ell^\infty(\Omega)$. The function $\mathcal{A}_g^* c_0$ is thus given by

$$\begin{aligned} \forall t \in \mathbb{R}, \quad (\mathcal{A}_g^* c_0)(t) &= \sum_{\ell \in \Omega} \alpha_\ell G(t - t_\ell) \text{sinc}(2\pi f_c(t - t_\ell)) \\ &+ \sum_{\ell \in \Omega} \beta_\ell G'(t_\ell - t) \text{sinc}(2\pi f_c(t - t_\ell)), \end{aligned} \quad (8)$$

where G designates the autocorrelation of g , that is,

$$\forall t \in \mathbb{R}, \quad G(t) = \int_{\mathbb{R}} g(\tau) g(t + \tau) d\tau = \exp\left(-\frac{\pi t^2}{4\sigma^2}\right).$$

Next, we determine sequences α and β such that the interpolation conditions $(\mathcal{A}_g^* c_0)(t_\ell) = \varepsilon_\ell$ are satisfied and $\mathcal{A}_g^* c_0$ has a local extremum at every t_ℓ , $\ell \in \Omega$. The coefficients α_ℓ play the main role in interpolating the points $(t_\ell, \varepsilon_\ell)$. The coefficients β_ℓ add a correction term to ensure that $\mathcal{A}_g^* c_0$ does not exceed 1 on \mathbb{R} . To conclude the proof, we finally show that (7) holds.

We can see from (8) how the windowing in the STFT helps to improve the recovery guarantee in Theorem 4 compared to recovery based on unwindowed Fourier measurements, as considered in [3, Sec. 2]. Specifically, in the case of unwindowed Fourier measurements, the interpolation function $\mathcal{A}^* c_0$ must be a Paley-Wiener function [19, Thm. 19.3], while here $\mathcal{A}_g^* c_0$ is clearly not band-limited due to Gaussian windowing and therefore has better time-localization. This, in turn, allows the minimum separation Δ to be smaller.

5. SIMULATIONS

For the simulation results we consider the recovery of the discrete complex measure $\mu = \sum_{\ell \in \Omega} a_\ell \delta_{t_\ell} \in \mathcal{M}(\mathbb{T})$ over the torus \mathbb{T} (the set Ω is then finite). The Gaussian window function is periodized so that for all $t \in \mathbb{R}$,

$$g(t) = \sum_{n \in \mathbb{Z}} \frac{1}{\sqrt{\sigma}} \exp\left(-\frac{\pi(t+n)^2}{2\sigma^2}\right) = \sum_{n \in \mathbb{Z}} g_n e^{2\pi i n t}, \quad (9)$$

with $g_n := \sqrt{2\pi} \exp(-2\pi\sigma^2 n^2)$, $n \in \mathbb{Z}$. The dual group of \mathbb{T} is \mathbb{Z} . The corresponding STFT measurements of μ are given by the sequence $\{y_k\}_{k=-f_c}^{f_c}$ of functions

$$\forall \tau \in \mathbb{T}, \quad y_k(\tau) := \sum_{\ell \in \Omega} a_\ell g(t_\ell - \tau) e^{-2\pi i k t_\ell} = \sum_{n \in \mathbb{Z}} y_{k,n} e^{2\pi i n \tau},$$

where $y_{k,n} := \sum_{\ell \in \Omega} a_\ell g_n e^{-2\pi i(n+k)t_\ell}$ is the n th Fourier coefficient of y_k . Using Parseval's theorem, the objective function for (PD-SR) can be rewritten as

$$\langle c, y \rangle = \sum_{k=-f_c}^{f_c} \sum_{n \in \mathbb{Z}} c_{k,n} y_{k,n},$$

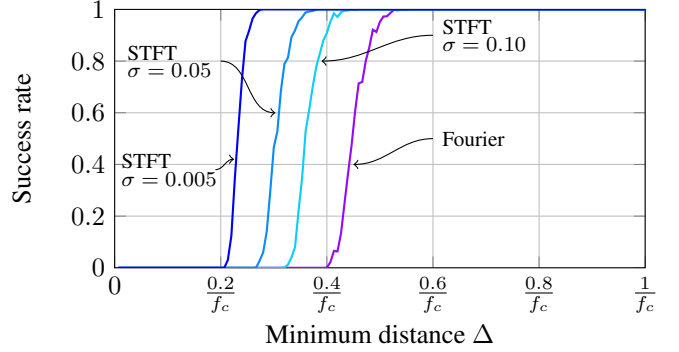


Fig. 1: Success rate for support recovery from (unwindowed) Fourier measurements and from STFT measurements with $f_c = 50$ and $N = 50$.

where $c_{k,n}$ denotes the n th Fourier coefficient of $c_k \in L^\infty(\mathbb{T})$ for $|k| \leq f_c$. The function $\mathcal{A}_g^* c$ can be expressed as

$$\forall t \in \mathbb{T}, \quad (\mathcal{A}_g^* c)(t) = \sum_{k=-f_c}^{f_c} \sum_{n \in \mathbb{Z}} g_n c_{k,n} e^{2\pi i(k+n)t}.$$

In order to render the problem (PD-SR) finite-dimensional, we approximate the function g in (9) by keeping only $2N + 1$ Fourier coefficients, that is, we replace g by

$$\forall t \in \mathbb{T}, \quad \tilde{g}(t) = \sum_{n=-N}^N g_n e^{2\pi i n t},$$

where N is chosen large enough for the coefficients g_n , $|n| \geq N$, to be small. The objective of (SR) then is $\langle \mathbf{C}, \mathbf{Y} \rangle$, where $\mathbf{C} := (c_{k,n})_{|k| \leq f_c, |n| \leq N}$ and $\mathbf{Y} := (y_{k,n})_{|k| \leq f_c, |n| \leq N}$. The function $\mathcal{A}_g^* c$ becomes a trigonometric polynomial which can be expressed as

$$\sum_{m=-(f_c+N)}^{f_c+N} x_m e^{2\pi i m t} \quad \text{with} \quad x_m = \sum_{n=n_{\min}}^{n_{\max}} g_m c_{m-n,n},$$

where $n_{\min} := \max\{-N, m - f_c\}$ and $n_{\max} := \min\{N, f_c + m\}$. We can now apply a procedure similar to the one developed in [3, Sec. 4, pp. 31–36] to solve (PD-SR) and to reconstruct the corresponding solution of (SR).

To assess the performance of our recovery procedure, we run 1500 trials. For each Δ , we construct a discrete Radon measure μ supported on the set $T = \{t_\ell\}_{\ell=0}^S$ with $S = \lfloor 1/(2\Delta) \rfloor$ and $t_\ell = 2\ell\Delta + r_\ell$, where r_ℓ is chosen uniformly at random in $[0, \Delta]$. Thus, the minimum distance between two distinct points of T is at least Δ . The complex amplitudes are obtained by choosing their real and imaginary parts uniformly at random in $[0, 1000]$. If the reconstructed measure $\hat{\mu}$ has support $\hat{T} = \{\hat{t}_\ell\}_{\ell \in \Omega}$ satisfying $\|\hat{T} - T\|_{\ell^2} / \|T\|_{\ell^2} \leq 10^{-3}$, we declare success. The corresponding results are depicted in Fig. 1. As predicted by our theoretical results, we, indeed, observe a factor-of-two improvement in the case of recovery from STFT measurements relative to recovery from unwindowed Fourier measurements as in [3, Sec. 2]. Note, however, that time-frequency measurements provide more information than frequency-only measurements as considered in [3].

References

- [1] J.-P. Kahane, “Analyse et synthèse harmoniques,” in *Histoires de mathématiques (Journées X-UPS 2011)*, École Polytechnique, Palaiseau, France, May 2011, pp. 17–53.
- [2] D. L. Donoho, “Super-resolution via sparsity constraint,” *SIAM Journal on Mathematical Analysis*, vol. 23, no. 5, pp. 1303–1331, Sep. 1992.
- [3] E. J. Candès and C. Fernandez-Granda, “Towards a mathematical theory of super-resolution,” *Communications on Pure and Applied Mathematics*, Apr. 2013, online issue.
- [4] B. F. Logan, “Properties of high-pass signals,” Ph.D. dissertation, Columbia University, New York, NY, USA, 1965.
- [5] —, “Bandlimited functions bounded below over an interval,” *Notices of the American Mathematical Society*, vol. 24, p. A331, 1977.
- [6] A. Beurling, “Local harmonic analysis with some applications to differential operators,” in *The Collected Works of Arne Beurling: Volume 2, Harmonic Analysis*, L. Carleson, P. Malliavin, J. Neuberger, and J. Werner, Eds. Boston, MA, USA: Birkhäuser, 1966, pp. 299–315.
- [7] —, “Balayage of Fourier-Stieltjes transforms,” in *The Collected Works of Arne Beurling: Volume 2, Harmonic Analysis*, L. Carleson, P. Malliavin, J. Neuberger, and J. Werner, Eds. Boston, MA, USA: Birkhäuser, 1989, pp. 341–350.
- [8] —, “V. Interpolation for an interval on \mathbb{R}^1 . 1. A density theorem. Mittag-Leffler Lectures on Harmonic Analysis,” in *The Collected Works of Arne Beurling: Volume 2, Harmonic Analysis*, L. Carleson, P. Malliavin, J. Neuberger, and J. Werner, Eds. Boston, MA, USA: Birkhäuser, 1989, pp. 351–359.
- [9] D. L. Donoho and B. F. Logan, “Signal recovery and the large sieve,” *SIAM Journal on Applied Mathematics*, vol. 52, no. 2, pp. 577–591, Apr. 1992.
- [10] M. Vetterli, P. Marziliano, and T. Blu, “Sampling signals with finite rate of innovation,” *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002.
- [11] P. L. Dragotti, M. Vetterli, and T. Blu, “Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix,” *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1741–1757, May 2007.
- [12] S. Levy and P. K. Fullagar, “Reconstruction of a sparse spike train from a portion of its spectrum, and application to high-resolution deconvolution,” *Geophysics*, vol. 46, no. 9, pp. 1235–1243, Sep. 1981.
- [13] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM Review*, vol. 43, no. 1, pp. 129–159, Feb. 2001.
- [14] K. Bredies and H. K. Pikkarainen, “Inverse problems in spaces of measures,” *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 19, no. 1, pp. 190–218, Jan. 2013.
- [15] Y. de Castro and F. Gamboa, “Exact reconstruction using Beurling minimal extrapolation,” *Journal of Mathematical Analysis and Applications*, vol. 395, no. 1, pp. 336–354, Nov. 2012.
- [16] V. Duval and G. Peyré, “Exact support recovery for sparse spikes deconvolution,” CNRS and Université Paris-Dauphine, Paris, France, Tech. Rep., July 2013. [Online]. Available: <http://hal.archives-ouvertes.fr/docs/00/83/96/35/PDF/DuvalPeyre-SparseSpikes.pdf>
- [17] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, “Compressed sensing off the grid,” in *50th Annual Allerton Conference on Communication, Control, and Computing*, Oct. 2012, pp. 778–785.
- [18] E. Au-Yeung and J. J. Benedetto, “Balayage and short-time Fourier transform frames,” in *10th International Conference on Sampling Theory and Applications*, July 2013. [Online]. Available: <http://arxiv.org/abs/1309.0539>
- [19] W. Rudin, *Real and Complex Analysis*, 3rd ed., P. R. Devine, Ed. New York, NY, USA: McGraw-Hill, 1987.
- [20] H. Brezis, *Analyse fonctionnelle – Théorie et applications*, ser. Collection Mathématiques appliquées pour la maîtrise, P. G. Ciarlet and J. L. Lions, Eds. Paris, France: Masson, 1983.
- [21] K. Gröchenig, *Foundations of Time-Frequency Analysis*, ser. Appl. Numer. Harmonic Anal., J. J. Benedetto, Ed. Boston, MA, USA: Birkhäuser, 2000.
- [22] J. Borwein and Q. Zhu, *Techniques of Variational Analysis*, ser. CMS Books in Mathematics, J. Borwein and K. Dilcher, Eds. New York, NY, USA: Springer, 2005.