HAL
open science

The Non Degenerate Source Condition: Support Robustness for Discrete and Continuous Sparse Deconvolution

Vincent Duval, Gabriel Peyré

▶ To cite this version:

HAL Id: hal-01169371

https://inria.hal.science/hal-01169371

Submitted on 29 Jun 2015

The Non Degenerate Source Condition: Support Robustness for Discrete and Continuous Sparse Deconvolution

Vincent Duval
INRIA Rocquencourt (MOKAPLAN)
Domaine de Voluceau, Rocquencourt BP 105
78153 Le Chesnay Cedex - FRANCE
Email: vincent.duval@inria.fr

Gabriel Peyré
CNRS & Université Paris-Dauphine (CEREMADE)
Place du Maréchal De Lattre De Tassigny
75775 Paris cedex - FRANCE
Email: gabriel.peyre@ceremade.dauphine.fr

Abstract—In this paper, we study discrete and continuous versions of the LASSO in order to solve the deconvolution problem. We shed light on the Non Degenerate Source Condition, a property which yields support robustness for both the continuous and discrete problems. More precisely, we show that this property yields exact support recovery in the continuous case and the estimation of twice the number of spikes in the discrete case.

I. INTRODUCTION

A. Sparsity-based approaches for deconvolution

The deconvolution problem, which consists in removing the acquisition blur, is central in signal and image processing. The use of ℓ^1-based techniques for deconvolution can be traced back to pioneering works in Geophysics [1] for seismic prospecting purpose. On the theoretical side, the study of ℓ^1 regularization was initiated by Donoho [2] and Tibshirani [3]. As the authors remarked, that type of regularization has the remarkable property that it retains both the features of subset selection (by setting to zero some coefficients, hence favoring sparse solutions) and ridge regression (by shrinking the other coefficients).

Whereas those methods typically rely on the use of a grid for the reconstruction, several recent works [4, 5, 6, 7] have investigated the possibility to work in a continuous framework, by replacing the ℓ^1 regularization with the *total variation*, or the *atomic norm* regularization. The signal to recover is not a finite-dimensional vector anymore but a Radon measure.

An important issue is the stability of deconvolution methods w.r.t noise. We shall focus on support recovery properties, as studied by Fuchs in [8] in the discrete case. Let us mention however, that the analysis developped in [8] does not hold for the deconvolution problem where the atoms of the dictionary are highly correlated: there is no exact support recovery (on thin grids). In the continuous case, several works have studied the robustness of the model [7, 9, 10], but to our knowledge only [11] deals with the problem of exact support recovery.

Our goal is to highlight the role of an important property, the *Non Degenerate Source Condition*, which governs the support of the solutions of the LASSO at low noise, whether in the discrete or continuous models. We draw a connection with a recent paper of Tang and Recht [12] which shows that this property holds for signals that are made of sufficiently separated spikes.

For the sake of simplicity, we shall work in a one-dimensional framework, the observation being the convolution of a real signal with some kernel φ (plus possibly some noise). Typical examples of kernels include (for $\sigma > 0$)

- the ideal lowpass filter $\varphi(t) = \sqrt{\frac{1}{\pi\sigma}} \frac{\sin(t/\sigma)}{t/\sigma}$,

- the Gaussian filter $\varphi(t) = \frac{1}{\pi^{1/4}\sqrt{\sigma}} \exp\left(-\frac{t^2}{2\sigma^2}\right)$,

- the Cauchy filter $\varphi(t) = \sqrt{\frac{2}{\pi\sigma}} \frac{1}{1+(t/\sigma)^2}$.

B. Notations

We denote by $\mathcal{M}_b(\mathbb{R})$ the space of finite Radon measures on $\mathbb{R}$, it is the dual of the space $\mathscr{C}_0(\mathbb{R})$ of continuous functions which vanish at infinity endowed with the uniform norm. Given a smooth function $\varphi \in \mathscr{C}_0(\mathbb{R}) \cap L^2(\mathbb{R})$, we define the bounded linear operator $\Phi : \mathcal{M}_b(\mathbb{R}) \to L^2(\mathbb{R})$ defined by

$$\forall x \in \mathbb{R}, \quad (\Phi m)(x) = \int_{\mathbb{R}} \varphi(x-y)\mathrm{d}m(y), \tag{1}$$

and we denote by Φ' the operator obtained similarly by replacing φ with $-\varphi'$. Given $x_0 \in \mathbb{R}^N$, we let $\Phi_{x_0} : \mathbb{R}^N \to L^2(\mathbb{R})$ be the operator such that

$$\forall a_0 \in \mathbb{R}^N, \quad \Phi_{x_0}a_0 = \sum_{i=1}^{N} a_{0,i}\varphi(\cdot - x_{0,i}). \tag{2}$$

We also define $\Gamma_{x_0} : \mathbb{R}^N \times \mathbb{R}^N \to L^2(\mathbb{R}^N)$ by $\Gamma_{x_0} \stackrel{\text{def.}}{=} (\Phi_{x_0}, \Phi'_{x_0})$, *i.e.* $\Gamma_{x_0}\left(\begin{smallmatrix} a_0 \\ b_0 \end{smallmatrix}\right) = \sum_{i=1}^{N} (a_{0,i}\varphi(\cdot - x_{0,i}) - b_{0,i}\varphi'(\cdot - x_{0,i}))$.

The adjoint of Φ is $\Phi^* : L^2(\mathbb{R}) \to \mathscr{C}_0(\mathbb{R})$ is such that for all $p \in L^2(\mathbb{R})$,

$$\forall y \in \mathbb{R}, \quad (\Phi^*p)(y) = \int_{\mathbb{R}} \varphi(x-y)p(x)\mathrm{d}x. \tag{3}$$

It turns out that $(\Phi_{x_0})^*p = ((\Phi^*p)(x_{0,i}))_{1 \leqslant i \leqslant N}$.

C. Considered problem

Given $N \in \mathbb{N}^*$, we consider a measure $m_0 \stackrel{\text{def.}}{=} \sum_{i=1}^N a_{0,i} \delta_{x_{0,i}}$ where $a_0 \in \mathbb{R}^N$, and $x_0 \in \mathbb{R}^N$ has pairwise distinct components. Given a blurry observation $y_0 = \Phi m_0$ (resp. blurry and noisy $y_0 + w$, where $w \in L^2(\mathbb{R})$), we try to recover m_0 by applying the *atomic norm decomposition* or Beurling LASSO [4, 5, 6, 7], using

$$\min_{m \in \mathcal{M}_b(\mathbb{R})} |m|(\mathbb{R}) \quad \text{such that} \quad \Phi m = y_0, \tag{$\mathcal{P}_0^\infty(y_0)$}$$

$$(\text{resp.}) \quad \min_{m \in \mathcal{M}_b(\mathbb{R})} \frac{1}{2} \|y_0 + w - \Phi m\|^2 + \lambda |m|(\mathbb{R}), \tag{$\mathcal{P}_\lambda^\infty(y_0 + w)$}$$

where $|m|(\mathbb{R})$ refers to the total variation of the measure m

$$|m|(\mathbb{R}) \stackrel{\text{def.}}{=} \sup \left\{ \int_{\mathbb{R}} \psi(x) \mathrm{d}m(x) \; ; \; \psi \in \mathscr{C}_0(\mathbb{R}) \text{ and } \|\psi\|_\infty \leqslant 1 \right\}. \tag{4}$$

II. THE NON DEGENERATE SOURCE CONDITION

A. Duality

The dual problem to $(\mathcal{P}_0^\infty(y_0))$ is given by

$$\sup_{p \in C^\infty} \langle y_0, p \rangle_{L^2(\mathbb{R})}, \tag{$\mathcal{D}_0^\infty(y_0)$}$$

$$\text{where} \quad C^\infty \stackrel{\text{def.}}{=} \left\{ p \in L^2(\mathbb{R}) \; ; \; \|\Phi^* p\|_\infty \leqslant 1 \right\}.$$

The solution to $(\mathcal{D}_0^\infty(y_0))$, if it exists, is not unique in general. An important property is that if $\eta \stackrel{\text{def.}}{=} \Phi^* p$ where p is a solution to $(\mathcal{D}_0^\infty(y_0))$ and if m is a solution to $(\mathcal{P}_0^\infty(y_0))$, then

$$\begin{aligned} \operatorname{supp} m_+ &\subset \{x \in \mathbb{R} \; ; \; \eta(x) = +1\}, \\ \operatorname{supp} m_- &\subset \{x \in \mathbb{R} \; ; \; \eta(x) = -1\}, \end{aligned} \tag{5}$$

where $m = m_+ - m_-$ is the Hahn decomposition of m. Furthermore, if $p \in C^\infty$, if m is such that $\Phi m = \Phi m_0$ and if (5) holds, then m is a solution to $(\mathcal{P}_0^\infty(y_0))$ (and p is a solution to $(\mathcal{D}_0^\infty(y_0))$). In that case we say that η is a *certificate* for m.

That property is used in [4] to prove that positive measures are identifiable using $(\mathcal{P}_0^\infty(y_0))$. It is also at the core of a striking result by Candès and Fernandez-Granda [6]: there exists a positive constant $C \leqslant 2$ such that, if φ is the ideal lowpass filter with cutoff frequency f_c, every measure $m_0 = \sum_{i=1}^N a_{0,i} \delta_{x_{0,i}}$ defined on the torus $\mathbb{T} \stackrel{\text{def.}}{=} \mathbb{R}/\mathbb{Z}$ such that the minimum distance $\Delta(m_0)$ between each spike is larger than C/f_c is the unique solution to $(\mathcal{P}_0^\infty(y_0))$. An extension of that theorem to the real line has been proposed in [12], which enables to deal with the Gaussian or the Cauchy kernels, and this approach has been generalized in [13] to bivariate signals.

B. The minimal norm certificate

On the other hand, the dual problem to $(\mathcal{P}_\lambda^\infty(y_0 + w))$ is given by

$$\sup_{p \in C^\infty} \langle y_0 + w, p \rangle_{L^2(\mathbb{R})} - \frac{\lambda}{2} \|p\|_{L^2(\mathbb{R})}, \tag{$\mathcal{D}_\lambda^\infty(y_0 + w)$}$$

There is always a unique solution to $(\mathcal{D}_\lambda^\infty(y_0 + w))$, and if m_λ is a solution to $(\mathcal{P}_\lambda^\infty(y_0 + w))$ and p_λ is the solution to $(\mathcal{D}_\lambda^\infty(y_0 + w))$, they must satisfy (5). If a solution

to $(\mathcal{D}_0^\infty(y_0))$ exists, the unique solution p_λ to $\mathcal{P}_\lambda^\infty(y_0)$ converges as $\lambda \to 0^+$ to the solution of $(\mathcal{D}_0^\infty(y_0))$ with minimal norm (see [11])

$$p_0 \stackrel{\text{def.}}{=} \operatorname{argmin} \left\{ \|p\|_{L^2(\mathbb{R})} \; ; \; p \text{ is a solution to } (\mathcal{D}_0^\infty(y_0)) \right\}. \tag{6}$$

We define the *minimal norm certificate* as $\eta_0 \stackrel{\text{def.}}{=} \Phi^* p_0$. As we shall see in Section III-A, η_0 governs the recovery properties of $(\mathcal{P}_\lambda^\infty(y_0 + w))$ at low noise.

In particular the following assumption on η_0 is crucial.

Definition 1 (Non Degenerate Source Condition). *Let* $m_0 = \sum_{i=1}^N a_{0,i} \delta_{x_{0,i}}$. *We say that* m_0 *satisfies the Non Degenerate Source Condition if there exists a solution to* $(\mathcal{D}_0^\infty(y_0))$ *and if the minimal norm certificate* η_0 *satisfies*

$$\forall i \in \{1, \ldots, N\}, \; \eta_0(x_{0,i}) = \operatorname{sign}(a_{0,i}),$$
$$\forall t \in \mathbb{R} \setminus \{x_{0,1}, \ldots, x_{0,N}\}, \; |\eta_0(t)| < 1,$$
$$\forall i \in \{1, \ldots, N\}, \; \eta_0''(x_{0,i}) \neq 0.$$

It implies in particular that m_0 *is a solution to* $(\mathcal{P}_0^\infty(y_0))$.

C. The vanishing derivatives precertificate

Computing $\eta_0 = \Phi^* p_0$ from (6) is a difficult task. Observing that $\eta_0(x_{0,i}) = \operatorname{sign}(a_{0,i})$ and $\eta_0'(x_{0,i}) = 0$, we defined in [11] a "good candidate" for η_0, the *vanishing derivatives precertificate* $\eta_V \stackrel{\text{def.}}{=} \Phi^* p_V$, where

$$p_V \stackrel{\text{def.}}{=} \operatorname{argmin}\{\|p\|_{L^2(\mathbb{R})} \; ; \; (\Phi^* p)(x_{0,i}) = \operatorname{sign}(a_{0,i}), \\ (\Phi^* p)'(x_{0,i}) = 0 \text{ for } 1 \leqslant i \leqslant N\}. \tag{7}$$

If Γ_{x_0} has full column rank, p_V is well defined and may be computed in practice as $p_V = \Gamma_x^{+,*} \begin{pmatrix} \operatorname{sign}(a_0) \\ 0_N \end{pmatrix}$ where $\Gamma_{x_0}^{+,*} \stackrel{\text{def.}}{=} \Gamma_{x_0}(\Gamma_{x_0}^* \Gamma_{x_0})^{-1}$. The connection between η_V and η_0 is made clear by the following proposition.

Proposition 1 ([11]). *Assume that* Γ_{x_0} *has full rank. Then* m_0 *satisfies the Non Degenerate Source Condition if and only if*

$$\forall t \in \mathbb{R} \setminus \{x_{0,1}, \ldots, x_{0,N}\}, \; |\eta_V(t)| < 1,$$
$$\forall i \in \{1, \ldots, N\}, \; \eta_V''(x_{0,i}) \neq 0.$$

Moreover, in that case $\eta_V = \eta_0$.

The function η_V draws a surprising connection with the work of Tang and Recht. In [12], the authors give a sufficient condition for m_0 to be the unique solution of $(\mathcal{P}_0^\infty(y_0))$. To this end, under sufficient conditions, they build a certificate η for m_0 and they prove that $|\eta|(t) < 1$ for $t \in \mathbb{R} \setminus \{x_{0,1}, \ldots, x_{0,N}\}$, and $\eta''(x_{0,i}) \neq 0$. It turns out that the function η they construct is precisely η_V. Therefore we obtain

Theorem 1 ([12]). *Let* φ *be the Gaussian or the Cauchy kernel. Then there exists a constant* $C > 0$ *such that, if* $\Delta(m_0) > C\sigma$ *then the Non Degenerate Source Condition holds for* m_0.

In fact, the result proved in [12] is more general. A series of conditions which should be satisfied by the filter φ is given, and the example of Gaussian filters is studied thoroughly. One may check that those conditions also hold for the Cauchy kernel. We skip the general conditions for brevity and we refer the reader to [12].

(a) m_0 and η_0

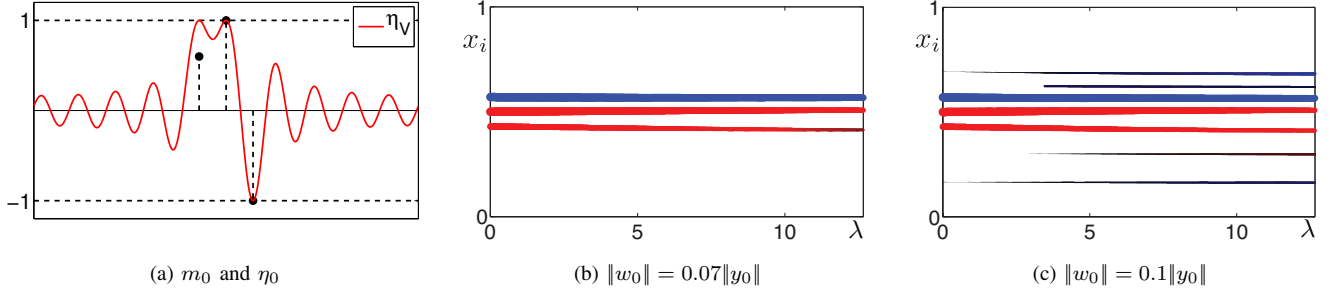(b) $\|w_0\| = 0.07\|y_0\|$

(c) $\|w_0\| = 0.1\|y_0\|$

Fig. 1: Left: original measure and its minimal norm certificate. Middle and right: positions of the reconstructed spikes (the amplitude is the width of each pattern) as λ varies, we choose w proportional to λ, *i.e.* $w = \lambda w_0$.

III. APPLICATIONS

A. Support recovery for the Beurling LASSO

A first consequence of the Non Degenerate Source Condition is the support recovery using $(\mathcal{P}_\lambda^\infty(y_0 + w))$ at small noise. For $\alpha > 0$, $\lambda_0 > 0$, let us define the small noise regime as

$$D_{\alpha,\lambda_0} \stackrel{\text{def.}}{=} \{(\lambda, w) \in \mathbb{R}_+ \times L^2(\mathbb{R}) \,;\, 0 \leqslant \lambda \leqslant \lambda_0 \quad \text{and} \quad \|w\|_{L^2(\mathbb{R})} \leqslant \alpha\lambda\}. \tag{8}$$

Theorem 2 ([11]). *Let m_0 be a discrete measure which satisfies the Non Degenerate Source Condition. Then there exist $\alpha > 0$, $\lambda_0 > 0$ such that for $(\lambda, w) \in D_{\alpha,\lambda_0}$, the solution to $(\mathcal{P}_\lambda^\infty(y_0 + w))$ is unique and it has exactly N spikes, $m_\lambda = \sum_{i=1}^N a_{\lambda,i} \delta_{x_{\lambda,i}}$. Moreover, the mapping $D_{\alpha,\lambda_0} \ni (\lambda, w) \mapsto (a_\lambda, x_\lambda) \in \mathbb{R}^N \times \mathbb{R}^N$ is $\mathscr{C}^{k-1}$ whenever $\varphi \in \mathscr{C}^k(\mathbb{R})$.*

Theorem 2 is illustrated on Figure 1, where the positions and amplitudes of the reconstructed spikes are displayed. In the first case, the ratio $\|w\|/\lambda$ is small: we are in the low noise regime and the exact number of spikes is recovered. In the second one, $\|w\|/\lambda$ is too large: parasitic spikes appear.

B. Doubling of the support on thin grids

Another application of the Non Degenerate Source condition is the low noise behavior of the discrete LASSO on thin grids. Discrete versions of $(\mathcal{P}_0^\infty(y_0))$ and $(\mathcal{P}_\lambda^\infty(y_0 + w))$, which date back to the pioneering works [3, 2], are a standard way to tackle the deconvolution problem. Given $h > 0$, we consider a grid $h\mathbb{Z} \subset \mathbb{R}$ and we look for a measure of the form $\sum_{k\in\mathbb{Z}} b_k \delta_{kh}$ by solving

$$\min_{b\in\mathbb{R}^\mathbb{Z}} \frac{1}{2}\|y_0 + w - \Phi b\|^2 + \lambda\|b\|_1, \quad (\mathcal{P}_\lambda^h(y_0 + w))$$

and $\quad \min_{b\in\mathbb{R}^\mathbb{Z}} \|b\|_1 \quad \text{such that} \quad \Phi b = y_0, \quad (\mathcal{P}_0^h(y_0))$

where we denote $\sum_{k\in\mathbb{Z}} |b_k|$ by $\|b\|_1$ and (with a slight abuse) $\Phi(\sum_{k\in\mathbb{Z}} b_k \delta_{kh})$ by Φb. The sequence the minimizers of $(\mathcal{P}_0^h(y_0))$ (resp. $(\mathcal{P}_\lambda^h(y_0 + w)))$ has limit points and every such limit point is a minimizer of $(\mathcal{P}_0^\infty(y_0))$ (resp. $(\mathcal{P}_\lambda^\infty(y_0 + w)))$, as proved in [14]. In fact, the sequence of discrete problems Γ-converge towards their continuous counterpart (see [15]).

The dual problems $(\mathcal{D}_0^h(y_0))$ and $(\mathcal{D}_\lambda^h(y_0 + w))$ are defined similarly to $(\mathcal{D}_0^\infty(y_0))$ and $(\mathcal{D}_\lambda^\infty(y_0 + w))$, replacing C^∞ with

$$C^h \stackrel{\text{def.}}{=} \left\{ p \in L^2(\mathbb{R}) \,;\, \sup_{k\in\mathbb{Z}} |(\Phi^* p)(kh)| \leqslant 1 \right\}. \tag{9}$$

We consider a sequence of refining grids, $h_n \searrow 0$ such that $\text{supp}\, m_0 \subset h_n\mathbb{Z}$ for all $n \in \mathbb{N}$ large enough (for instance $h_n = 1/2^n$ and x_0 has dyadic components). As observed in [14], the reconstructed spikes tend to cluster around the solution of $(\mathcal{P}_0^\infty(y_0))$. The following theorem describes precisely that phenomenon at low noise.

Let us define the *natural shift* of the measure as

$$\rho \stackrel{\text{def.}}{=} (\Phi_{x_0}'^* \Pi \Phi_{x_0}')^{-1} \Phi_{x_0}'^* \Phi_{x_0}^{+,*} \text{sign}(a_0) \in \mathbb{R}^N, \tag{10}$$

where Π is the orthogonal projector onto $(\text{Im}\, \Phi_{x_0}^*)^\perp$.

Theorem 3 ([11, 15]). *Assume that m_0 satisfies the Non Degenerate Source Condition. Then for each $n \in \mathbb{N}$ large enough, there exist $\alpha_n > 0$, $\lambda_{0,n} > 0$, and $\varepsilon_n \in \{-1, 0, 1\}^N$ such that for all $(\lambda, w) \in D_{\alpha_n,\lambda_{0,n}}$, the solution $m_{\lambda,n} = \sum_{k\in\mathbb{Z}} b_k \delta_{kh_n}$ to $(\mathcal{P}_\lambda)^{h_n}(y_0 + w)$ has support $\bigcup_{i=1}^N \{x_{0,i}, x_{0,i} + \varepsilon_n h_n\}$, with $\text{sign}(m_{\lambda,n}(\{x_{0,i} + \varepsilon_n h_n\})) = \text{sign}(m_{\lambda,n}(\{x_{0,i}\})) = \text{sign}(a_{0,i})$.*

If, moreover, Γ_{x_0} has full rank, and if all the components of ρ are nonzero, then $\varepsilon_n = \varepsilon$ is independent of n and

$$\varepsilon = \text{sign}\left(\text{diag}(\text{sign}(a_0)\rho)\right).$$

The involved constants can be chosen as

$$\alpha_n = O(1) \quad \text{and} \quad \lambda_{0,n} = O(h_n). \tag{11}$$

The practical consequence of Theorem (3) is that, on thin grids and at low noise, the LASSO estimates twice the number of spikes: each original spike and one of its immediate neighbors. This illustrated on Figure 2, where we have displayed the amplitudes of the reconstructed signal at two points of x_0 and their immediate neighbors. As $\lambda \to 0^+$, the immediate neighbor has nonzero amplitude but this amplitude tends to zero.

IV. CONCLUSION

The Non Degenerate Source Condition (NDSC) is an important property which governs the low noise behavior of both the LASSO for measures and the discrete LASSO : in the first case, the model recovers exactly the same number of
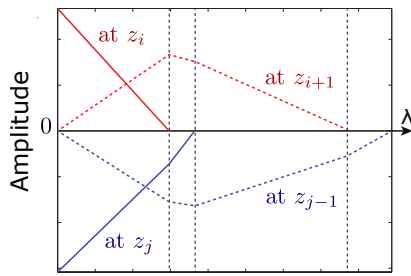
Fig. 2: The solution path (as a function of λ) for some discrete measure m_0 (the noise w is set to zero). This shows the amplitudes of the coefficients at $z_i = ih$, resp. $z_j = jh$, (continuous line) and at the next, resp. previous, point of the grid (dashed line) as λ varies.

Dirac masses, and they converge smoothly towards the original ones as the noise vanishes, in second case the model recovers pairs of Dirac masses, corresponding to the original ones and one of their immediate neighbors. By proving an identifiability result in [12], Tang and Recht have incidentally proved that the NDSC holds for measures with spikes that are sufficiently separated.

REFERENCES

[1] J. Claerbout and F. Muir, "Robust modeling with erratic data," *Geophysics*, vol. 38, no. 5, pp. 826–844, 1973.

[2] D. D. S.S. Chen and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[3] R. Tibshirani, "Regression Shrinkage and Selection Via the Lasso," in *Journal of the Royal Statistical Society, Series B*, vol. 58, 1994, pp. 267–288.

[4] 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, 2012.

[5] B. Bhaskar and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *In 2011 49th Annual Allerton Conference on Communication, Control, and Computing*, pp. 261–268, September 2011.

[6] E. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014. [Online]. Available: http://dx.doi.org/10.1002/cpa.21455

[7] K. Bredies and H. Pikkarainen, "Inverse problems in spaces of measures," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 19, no. 1, pp. 190–218, 2013.

[8] J. Fuchs, "On sparse representations in arbitrary redundant bases," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1341–1344, 2004.

[9] E. Candès and C. Fernandez-Granda, "Super-resolution from noisy data," *Journal of Fourier Analysis and Applications*, vol. 19, no. 6, pp. 1229–1254, 2013. [Online]. Available: http://dx.doi.org/10.1007/s00041-013-9292-3

[10] Y. d. C. J.-M. Azaïs and F. Gamboa, "Spike detection from inaccurate samplings," *Applied and Computational Harmonic Analysis*, vol. 38, no. 2, pp. 177–195, 2015.

[11] V. Duval and G. Peyré, "Exact support recovery for sparse spikes deconvolution," *Foundation and Computational Mathematics*, 2014. [Online]. Available: http://hal.archives-ouvertes.fr/hal-00839635/

[12] G. Tang and B. Recht, "Atomic decomposition of mixtures of translation-invariant signals," in *IEEE CAMSAP*, 2013.

[13] T. Bendory, S. Dekel, and A. Feuer, "Robust recovery of stream of pulses using convex optimization," *CoRR*, vol. abs/1412.3262, 2014. [Online]. Available: http://arxiv.org/abs/1412.3262

[14] G. Tang, B. Bhaskar, and B. Recht, "Sparse recovery over continuous dictionaries-just discretize," in *Signals, Systems and Computers, 2013 Asilomar Conference on*, Nov 2013, pp. 1043–1047.

[15] V. Duval and G. Peyré, "Sparse Spikes Deconvolution on Thin Grids," Mar. 2015. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01135200