# *Sparse image super-resolution via superset selection and pruning*

# Sparse image super-resolution via superset selection and pruning

Nam Nguyen
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA 02139
Email: namnguyen@math.mit.edu

Laurent Demanet
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA 02139
Email: laurent@math.mit.edu

***Abstract*—This note extends the superset method for sparse signal recovery from bandlimited measurements to the two-dimensional case. The algorithm leverages translation-invariance of the Fourier basis functions by constructing a Hankel tensor, and identifying the signal subspace from its range space. In the noisy case, this method determines a superset which then needs to undergo pruning. The method displays reasonable robustness to noise, and unlike $\ell_1$ minimization, always succeeds in the noiseless case.**

**Acknowledgments.** This project was sponsored by the Air Force Office of Scientific Research under Arje Nachman. LD also acknowledges funding from NSF, ONR, Total SA, and the Alfred P. Sloan Foundation.

## I. Introduction

We investigate a simple 2D super-resolution problem, where the goal is to recover a sparse image from its low frequency components. We call "spike" any image consisting of a single nonzero pixel; a sparse image $X_0 \in \mathbb{R}^{n_1 \times n_2}$ is a weighted superposition of a small number of spikes. The only available data are the noisy $2D$ low-frequency components

$$Y = F_{(1)} X_0 F_{(2)}^T + E, \tag{1}$$

where $F_{(1)} \in C^{m_1 \times n_1}$ and $F_{(2)} \in \mathbb{C}^{n_2 \times m_2}$ are partial, short and wide Fourier matrices, $F_{(a)}(j,k) = e^{-i2\pi kj/n_a}$, $0 \le j < m_a$, $-n_a/2 \le k < n_a/2$, $a = 1, 2$, ($n_1, n_2$ are even), and $E$ is the noise term whose entries are assumed i.i.d. $\mathcal{N}(0, \sigma^2)$.

Though various techniques have been proposed in the literature to tackle the super-resolution problem in the $1D$ setting, few algorithms are set up for this problem in higher dimensions. Among those, the most notable contributions are the generalized matrix pencil approaches of Hua and Sarkar [**?**], Andersson et al. [1], and Potts and Tasche [3]; Fannjiang and Liao's greedy algorithm [5]; and $\ell_1$ minimization [2]. Already in one dimension, the performance of the matrix-pencil-based algorithms is poorly understood in the presence of noise. In addition, the algorithm requires to know the support size, a value that is not often available in practice. In contrast, $\ell_1$ minimization is a natural idea for super-resolution problems, but often fails in situations where cancellations occur. Candès and Fernandez-Granda [2] show that $\ell_1$ minimization can recover the signal or image faithfully from its consecutive Fourier measurements as long as locations of any two consecutive spikes are separated by a constant times the super-resolution factor, $SRF = \max\{\frac{n_1}{m_1}, \frac{n_2}{m_2}\}$.

We should note that the data model considered here has the same form as the well-known compressed sensing (CS) problem, except for the crucial difference that we observe consecutive rather than random samples of the discrete Fourier transform.

In this paper, we extend the algorithm proposed in [4] for 1D signals to solve the super-resolution problem in 2D. In 1D, this algorithm has been experimentally shown to outperform methods such as matrix pencil and $\ell_1$ minimization in situations when the distances between spikes are much smaller than the Rayleigh limit. The superset method consists of two main ingredients:

- *Selection* of a set $\Omega$ which contains the spike locations of $X_0$ (with high probabiity), by inspection of the range space of a Hankel tensor; and
- *Pruning* to tighten the superset $\Omega$ by gradually removing atoms that do not associate with the true support $T$.

The 2D superset method considers the same object as the generalized matrix pencil methods, namely the tensor of all translates of the original image $X_0$ in *both* spatial directions. This 4-index object has an invariance property along the new translation indices, hence we may call it a Hankel tensor. Where traditional matrix pencil methods consider rank-reducing numbers of shifted versions of this Hankel tensor, the selection step of the superset method is simpler and deals with membership tests of individual atoms in the tensor's range space. The resulting procedure is not entirely fragile to the presence of additive noise, though this short note does not attempt to resolve the important associated theoretical recovery questions.

In the rest of the paper, we make use of the following notations. Let $[n]$ be the set $\{0, 1, ..., n-1\}$, we denote $T \in ([n_1], [n_2])$ as the support of $X_0$. For a matrix $A$, we write $A_S$ for the restriction of $A$ on its columns in $S$. The $k$-th column of $A$ is denoted as $A_{\bullet,k}$, while the $k$-th row is $A_{k,\bullet}$. The $(j,k)$-th entry of $A$ is $A(j,k)$. We also denote the restriction of $A$ to its first $L$ rows as $A^{(L)}$. Finally, vec $(A)$ is the vector constructed from vectorizing $A$.

## II. Hankel tensor in block matrix form

In this section, we start by describing the construction of the Hankel tensor in block matrix form. Throughout the paper, we use a convenient, vectorized form of the observations in (1) with respect to the elements of $X_0$. With $\mathbf{1}_{n_a}$ the $n_a \times 1$ vector of ones, denote

$$A = \mathbf{1}_{n_2} \otimes F_{(1)} = \underbrace{[F_{(1)} \quad F_{(1)} \quad \cdots \quad F_{(1)}]}_{n_2 \text{ times}},$$

and

$$B = \begin{pmatrix} \mathbf{1}_{n_1} \otimes (F_{(2)}^T)_{1,\bullet} \\ \vdots \\ \mathbf{1}_{n_1} \otimes (F_{(2)}^T)_{n_2,\bullet} \end{pmatrix}.$$

Then (1) can be written as the product

$$Y = AZ_0B + E = Y_0 + E, \tag{2}$$

where $Z_0 \in \mathbb{R}^{n_1 n_2 \times n_1 n_2}$ is the diagonal matrix with the vectorized form of $X_0$ along its diagonal. Let $S$ be the set that associates with nonzero entries on the diagonal of $Z_0$, then the observation matrix is $Y = A_S Z_{0_S} B_S^T + E$. Notice that $S$ is simply the vectorized form of the support $T$ of $X_0$.

In the absence of Gaussian noise, it is clear from (1) that $\operatorname{rank}(Y_0) \leq |S|$. Strict inequality occurs whenever the locations of any two spikes align horizontally or vertically – a phenomenon that did not occur in 1D. Therefore, $\operatorname{Ran} Y_0 \in \operatorname{Ran} A_S$ and $\operatorname{Ran} Y_0^T \in \operatorname{Ran} B_S$. One way to definitively link $|S|$ to a "rank" is to introduce the Hankel tensor, or "enhanced matrix", $Y_{0e}$ as in [**?**]. Let $K \in [0, m_1)$ and $L \in [0, m_2)$ be numbers that can be picked by users. We define the (block matrix realization of the) Hankel tensor as the $K \times (m_1 - K + 1)$ block Hankel matrix

$$Y_{0e} \triangleq \begin{pmatrix} H_0 & H_1 & \cdots & H_{m_1-K} \\ H_1 & H_2 & \cdots & H_{m_1-K+1} \\ \vdots & \vdots & \vdots & \vdots \\ H_{K-1} & H_K & \cdots & H_{m_1-1} \end{pmatrix}. \tag{3}$$

where each block $H_j$ is a $L \times (m_2 - L + 1)$ Hankel matrix constructed from the $j$-th row of $Y_0$. The form of $H_j$ is as follows:

$$\begin{pmatrix} Y_0(j,0) & Y_0(j,1) & \cdots & Y_0(j,m_2-L) \\ Y_0(j,1) & Y_0(j,2) & \cdots & Y_0(j,m_2-L+1) \\ \vdots & \vdots & \vdots & \vdots \\ Y_0(j,L-1) & Y_0(j,L) & \cdots & Y_0(j,m_2-1) \end{pmatrix}. \tag{4}$$

From (2) and (4), each block $H_j$ can be expressed as $H_j = B^{(L)} Z_0 D^j (B^{(m_2-L)})^T$, where $D \triangleq \operatorname{diag}(A_{1,\bullet})$. Substituting the formulation of $H_j$ into (3) allows to express $Y_{0e}$ as the products of three matrices,

$$Y_{0e} = \mathcal{A} Z_0 \mathcal{B}^T, \tag{5}$$

where we use calligraphic letters $\mathcal{A}$ and $\mathcal{B}$ to denote matrices

$$\mathcal{A} = \begin{pmatrix} B^{(L)} \\ B^{(L)}D \\ \vdots \\ B^{(L)}D^{K-1} \end{pmatrix} \tag{6}$$

$$\text{and} \qquad \mathcal{B} = \begin{pmatrix} B^{(m_2-L)} \\ B^{(m_2-L)}D \\ \vdots \\ B^{(m_2-L)}D^{m_1-K} \end{pmatrix}. \tag{7}$$

If we denote $\mathcal{A}_S = B_S^{(L)} D_S$ and $\mathcal{B}_S = B_S^{(m_2-L)} D_S$, then we can write $Y_{0e} = \mathcal{A}_S Z_{0_S} \mathcal{B}_S$. The following lemma is an important observation which allows us to exploit the subspace of the enhanced matrix to find the support $S$. The proof is presented in [**?**].

**Proposition 1.** *If $K, m_1 - K + 1, L$, and $m_2 - L + 1$ are greater than $|S|$, then the rank of $Y_{0e}$ is $|S|$, and*

$$\textit{Ran}\,(Y_{0e}) = \textit{Ran}\,(\mathcal{A}_S).$$

The lemma implies that if we loop over all the candidate atoms $\mathcal{A}_{\bullet,k}$ for $0 \leq k < n_1 n_2$ and select those with

$$\angle(\mathcal{A}_{\bullet,k}, \operatorname{Ran} Y_{0e}) = 0, \tag{8}$$

then those candidate set is associated with the true support. Once the set $S$ is detected, the image is recovered by solving the determined system

$$Y_0 = A_S Z_{0_S} (B_S)^T, \qquad \operatorname{diag}(Z_{0_{S^c}}) = 0.$$

## III. Subspace identification and pruning

The strategy just presented is not quite as simple in the noisy setting, when

$$Y = AZ_0B + E, \tag{9}$$

where $E$ is the noise matrix. In this situation, Proposition 1 is no longer satisfied. However, selecting indices that associate with the smallest angles $\angle(\mathcal{A}_{\bullet,k}, \operatorname{Ran} Y_e)$ is still a useful idea.

**Proposition 2.** *Let $Y = Y_0 + E$ with $e \sim N(0, \sigma^2 I_m)$, and form the corresponding $KL \times (m_1 - L + 1)(m_2 - K + 1)$ enhanced matrix $Y_e$ as in (5). Denote the $j$-th singular values of the block symmetric matrix $Y_{0e}^{(m_1-K)}$ by $s_{0,j}$. Then there exists positive $c_1, C_1$ and $c$, such that with probability at least $1 - c_1(m_1 m_2)^{-C_1}$,*

$$\sin\angle(\mathcal{A}_{\bullet,k}, \textit{Ran}\, Y_e) \leq c\,\varepsilon_1 \tag{10}$$

*for all indices $k$ in the support set and*

$$\varepsilon_1 = \frac{|S|}{\|\mathcal{A}_{\bullet,k}\|_2} \frac{\sigma\sqrt{KL \log m_1 m_2}}{|X_{0_{\min}}|} \sqrt{\frac{|X_{0_{\max}}|}{s_{0,|S|}}}, \tag{11}$$

*where $|X_{0_{\min}}|$ and $|X_{0_{\max}}|$ are respectively the smallest and largest nonzero entries of $X_0$.*

This proposition provides a procedure that guarantees the selection of a *superset* $\Omega$ of the true support $T$, with high probability.

Even though Proposition 2 reveals a constructive approach to define the superset, there are a few unknown quantities involved in the bound (10). The support size $|S|$, though not specified, can be upper bounded by a constant such as $\min(m_1, m_2)$. $|x_{0_{\min}}|$ and $|x_{0_{\max}}|$ can be estimated if we know the dynamic range of the image's intensity. These estimations might loosen the bound (10) and make increasing the superset size. However, $T$ will be included in $\Omega$ with high probability.

A second, pruning step is required to extract $T$ from $\Omega$. The idea is the same as in the 1D case, namely that the quality of the data fit is tested for each subset $\Omega\backslash k$ of $\Omega$ where a single atom $k$ is removed.

First rewrite $Y = A_\Omega Z_{0_\Omega} B_\Omega + E = \sum_{j\in\Omega} Z_0(j,j) A_{\bullet,j} \otimes B_{j,\bullet} + E$. Then express this identity in the vectorized form $\text{vec}\,(Y) = M_\Omega \text{vec}\,(Z_{0_\Omega}) + \text{vec}\,(E)$ where the matrix $M_\Omega = [\text{vec}\,(A_{\bullet,j} \otimes B_{j,\bullet})]_{j\in\Omega}$ and $\text{vec}\,(Z_{0_\Omega})$ is the vector containing the diagonal of $Z_0$ at location $\Omega$. Observe that in the noiseless setting, any $k \notin S$ will give zero angle between $\text{vec}\,(Y)$ and $\text{Ran}\, M_{\Omega\backslash k}$ – the range of $M_\Omega$ with the $k$-th column removed. Hence $\angle(\text{vec}\,(Y), \text{Ran}\, M_{\Omega\backslash k}) \neq 0$ iff $k \in S$. This suggests using this criterion for check the membership of $k$ in $T$. We then iterate until all $k \notin S$ are removed.

In the presence of noise, instead of using $\angle(\text{vec}\,(Y), \text{Ran}\, M_{\Omega\backslash k})$, we use the slightly better variant $\angle(\mathcal{P}_\Omega \text{vec}\,(Y), \text{Ran}\, M_{\Omega\backslash k})$, or equivalently $\left\|(\mathcal{P}_\Omega - \mathcal{P}_{\Omega\backslash k})\text{vec}\,(Y)\right\|_2$, where $\mathcal{P}_\Omega$ is the orthogonal projection of $\text{vec}\,(Y)$ onto $M_\Omega$. This process helps filtering out some of the noise embedded in $\text{vec}\,(Y)$. The pruning step will loop over $k \in \Omega$ and estimate $k \in S$ only when the angle is above a certain threshold.

Algorithm 1 describes superset selection followed by pruning to estimate the solution's support.

A good choice of threshold $\epsilon_2$ for the pruning step is suggested by the following result.

**Proposition 3.** *Assume the set $\Omega$ obtained from the $t$-th iteration of the pruning step still contains the true support $S$. Also assume that*

$$\min_{k\in S} \left\|(\mathcal{P}_\Omega - \mathcal{P}_{\Omega\backslash k}) vec\,(Z_{0_S}) M_{\bullet,k}\right\|_2 \geq 3\sigma. \qquad (12)$$

*Then there exists positive constants $c_1$ and $c_2$ such that with probability at least $1 - c_1(m_1 m_2)^{-1}$,*

$$\min_{k\notin S} \left\|(\mathcal{P}_\Omega - \mathcal{P}_{\Omega\backslash k}) vec\,(Y)\right\|_2 < \epsilon_2$$

$$\textit{and} \quad \min_{k\in S} \left\|(\mathcal{P}_\Omega - \mathcal{P}_{\Omega\backslash k}) vec\,(Y)\right\|_2 \geq \epsilon_2$$

*with the choice $\epsilon_2 = c_2\sigma$.*

Proposition 3 implies that as long as the noise level is sufficiently small so that (12) is satisfied, at each iteration, the pruning step is able to remove an atom $k \notin S$.

**Algorithm 1** Superset selection and pruning

**input:** Matrices $A \in \mathbb{C}^{m_1\times n_1 n_2}$, $B \in C^{n_1 n_2 \times n_2}$, $Y = AX_0B + E$, parameter $L$ and $K$, thresholds $\varepsilon_1$ and $\varepsilon_2$.
**initialization:** Enhanced matrix $Y_e \in \mathbb{C}^{KL\times(m_1-K+1)(m_2-L+1)}$, matrices $\mathcal{A} \in \mathbb{C}^{KL\times n_1 n_2}$ and $\mathcal{B} \in \mathbb{C}^{n_1 n_2\times(m_1-K+1)(m_2-L+1)}$ defined in (6) and (7)
**support identification**
  **decompose:** $\widetilde{Q}\widetilde{R} = Y_e\widetilde{E}$, $\widetilde{Q} \in \mathbb{C}^{L\times r}$
  **project:** $a_k \leftarrow \mathcal{A}_{\bullet,k}$ ( for all $k$)
  $\gamma_k \leftarrow \left\|a_k - \widetilde{Q}\widetilde{Q}^* a_k\right\| / \|a_k\|$
  $\Omega = \{k : \gamma_k \leq \varepsilon_1\}$
**while** true **do**
  **vectorize:** $M_\Omega = [\text{vec}\,(A_{\bullet,j} \otimes B_{j,\bullet})]_{j\in\Omega}$
  **decompose:** $QR = M_\Omega E$, $Q \in \mathbb{C}^{m\times|\Omega|}$
  **remove:** $\forall k \in \Omega$: $Q_{(k)}R_{(k)} = M_{\Omega\backslash k}E_{(k)}$
  $\delta_k \leftarrow \|(Q_{(k)}Q_{(k)}^* - QQ^*)\text{vec}\,(Y)\|_2$
  $k_0 \leftarrow \text{argmin}_k\, \delta_k$
  if $\delta_{k_0} < \varepsilon_2$, $\Omega \leftarrow \Omega\backslash k_0$
  else break
**end while**
**output:** $\widehat{X} = \text{argmin}_X \|Y - A_\Omega X B_\Omega\|$

## IV. Simulations

In the first simulation, we fix the image dimension to be $n_1 = 60$ and $n_2 = 60$. Only $m_1 = 12$ and $m_2 = 12$ frequency components are observed in each each dimension. We generate an $n_1 \times n_2$ image with spikes separated by $\lfloor 2.5n_1/m_1 \rfloor$ grid points vertically and $\lfloor 2.5n_2/m_2 \rfloor$ grid points horizontally. The spike magnitudes are randomly set to $\pm 1$ with equal probability. The noise matrix $E$ is generated from a Gaussian distribution whose entries are i.i.d. $N(0, \sigma^2)$ with $\sigma = 10^{-2}$. The parameters $K$ and $L$ used to construct the Hankel tensor are set to $\lfloor 2m_1/3 \rfloor$ and $\lfloor 2m_2/3 \rfloor$, respectively, as advised by Hua and Sarkar. We also set the threshold $\epsilon_1$ in the superset selection via (11) with $|S|$ replaced by $\min(m_1, m_2)$, $|X_{0_{\min}}| = |X_{0_{\max}}| = 1$, and $s_{0,|S|}$ replaced by the smallest singular value of the block symmetric matrix $Y_e^{(m_1-K)}$. The constant $c$ in (10) is fixed $c = 1$ and the threshold in the pruning step is $\epsilon_2 = \sigma\sqrt{\log m_1 m_2}$. Figure 1 shows the recovered images via $\ell_1$ minimization (basis pursuit) and the superset method (or SSP, for superset selection and pruning). Both methods accurately recover the image, hence are able to super-locate the spikes.

The next simulation demonstrates a more challenging situation in which the spacings between the spikes are well below the Rayleigh length, $\frac{n_1}{m_1}$ horizontally and $\frac{n_2}{m_2}$ vertically. As explained in [2], $\ell_1$ minimization fails to recover the image. With the same setting as in the previous simulation and with $m_1 = 6$, $m_2 = 6$, and $\sigma = 10^{-5}$, we plot the images recovered by the two methods in figure 2. The Rayleigh box (with sidelength equal to the Rayleigh length) is also plotted for reference. It is clear from the figure that the recovery via SSP is accurate, while $\ell_1$ minimization isn't.

For the record, the SSP method fails when $\sigma = 10^{-3}$ in

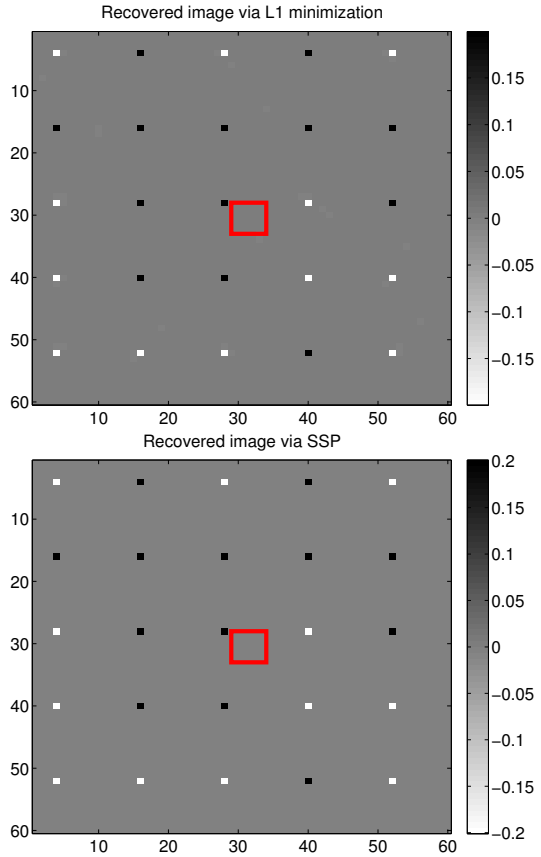


Fig. 1. An image with well-separated spikes. Top: recovered image via $\ell_1$ minimization. Bottom: recovered image via SSP method. Both methods recover the image accurately. In this example, $\sigma = 1e-2$. The "Rayleigh box" is shown for reference.

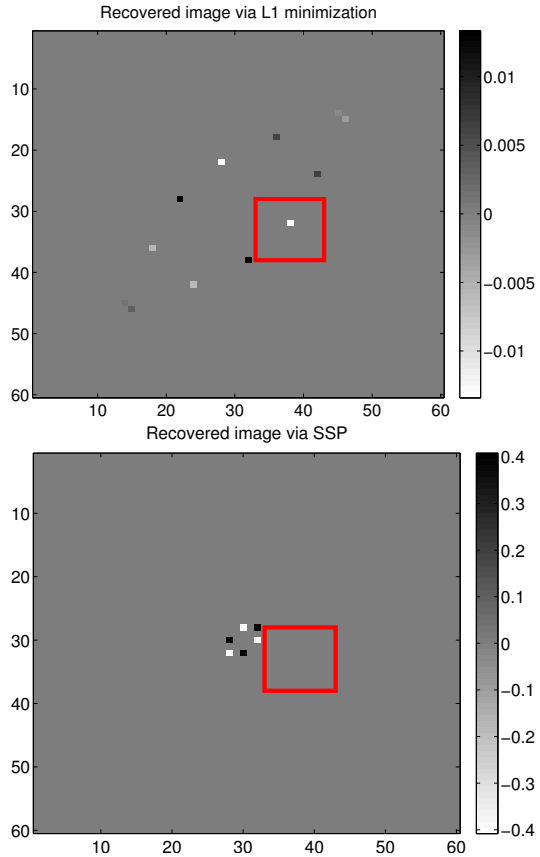


Fig. 2. An image with spacing between spikes well below the Rayleigh length. Top: image recovered via $\ell_1$ minimization. Right: image recovered via the SSP method, visually indistinguishable from the original image. In this example, $\sigma = 1e-5$. The "Rayleigh box" is shown for reference.

the configuration of the second example. This failure is due to the violation of the assumption (12) in Proposition 3. Super-resolution remains an ill-posed problem which can only be mitigated to a certain extent. The super-resolution problem in the form considered here will only be resolved once a method is exhibited for which the admissible noise level provably reaches the statistical recoverability limit.

## V. Conclusion

This paper desribes the superset method for dealing with the image super-resolution problem. The method is still competitive and not entirely unstable in the presence of noise, even when the distance between spikes is below the Rayleigh limit – a situation in which few other methods operate.

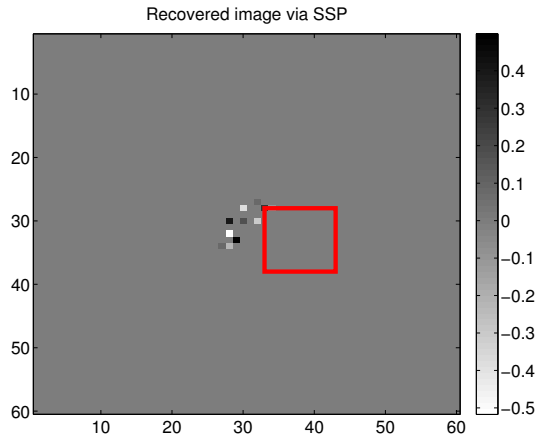


Fig. 3. With the noise level $\sigma = 1e-3$, SSP fails to recover the image shown in the bottom of Fig. 2.

## References

[1] F. Andersson, M. Carlsson, and M.V. de Hoop. Sparse approximation of functions using sums of exponentials and aak theory. *J. Approx. Theory*, 163(2):213–248, 2011.

[2] E. Candès and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *Commun. Pure Appl. Math.* To appear.

[3] Potts. D. and M. Tasche. Parameter estimation for multivariate exponential sums. *Elec. Trans. Num. Anal.*, 40:204–224, 2013.

[4] L. Demanet, D. Needell, and N. Nguyen. Super-resolution via superset selection and pruning. In *Proc. of SAMPTA*, Bremen, Germany, 2013. IEEE.

[5] A. Fannjiang and W. Liao. Coherence-pattern guided compressive sensing with unresolved grids. *IAM J. Imaging Sci.*, 5:179–202, 2012.