COMPRESSED REMOTE SENSING OF SPARSE OBJECTS

ALBERT C. FANNJIANG [*], PENGCHONG YAN [†], AND THOMAS STROHMER [‡]

Abstract. The linear inverse source and scattering problems are studied from the perspective of compressed sensing. By introducing the sensor as well as target ensembles, the maximum number of recoverable targets is proved to be at least proportional to the number of measurement data modulo a log-square factor with overwhelming probability.

Important contributions include the discoveries of the threshold aperture, consistent with the classical Rayleigh criterion, and the incoherence effect induced by random antenna locations.

The prediction of theorems are confirmed by numerical simulations.

1. Introduction. We consider the imaging problem in the form of inverse source or scattering problem which has wide-range applications such as radar, sonar and computed tomography. The imaging problem is typically plagued by nonuniqueness and instability and hence mathematically challenging. Traditional methods such as matched field processing [35] are limited in the number of targets that can be reliably recovered at high resolution. They often fail to detect a substantial number of targets, while at the same time they tend to produce artifacts obscuring the real target images. These limitations are due to the presence of noise and the fact that the imaging problem is in practice underdetermined. The standard regularization methods can handle to some extent the problem with noise but are inadequate to remedy the issue of nonuniqueness of the solution.

In this paper we utilize the fact that in many imaging applications the targets are sparse in the sense that they typically occupy a small fraction of the overall region of interest (the target domain). This sparsity assumption suggests to approach the imaging problem by using the framework of *compressed sensing*.

At the core of compressed sensing lies the following problem (here we focus, as is common in the compressed sensing community, on the discrete setting). Assume $X \in \mathbb{C}^m$ is a signal that is sparse, i.e., the number of its non-zero components (measured by the ℓ_0-quasinorm $\|X\|_0$ which is simply the number of non-zero entries of X) satisfies $s := \|X\|_0 \ll m$. Let $Y \in \mathbb{C}^n$ be the measurement data vector. We explore in this paper the *linear* inverse problem which can be formulated as $Y = \mathbf{\Phi}X$ where $\mathbf{\Phi}$ is an $n \times m$ matrix with $n \ll m$. The goal is to recover X, given the data vector Y and the sensing matrix $\mathbf{\Phi}$ of full rank. As $n \ll m$, $\mathbf{\Phi}X = Y$ is severely underdetermined and unique reconstruction of X is in general impossible. The number $\|X\|_0$ of nonzero elements in X is called the sparsity of X.

However, due to the sparsity of X one can compute X by solving the optimization problem

$$\min \|X\|_0 \qquad \text{s.t. } \mathbf{\Phi}X = Y. \tag{L0}$$

Since (L0) is NP-hard and thus computationally infeasible, we consider instead its convex relaxation, also known as Basis Pursuit (BP),

$$\min \|X\|_1 \qquad \text{s.t. } \mathbf{\Phi}X = Y \tag{L1}$$

which can be solved by linear and quadratic programming techniques. The amazing discovery due to David Donoho was that under certain conditions on the matrix $\mathbf{\Phi}$ and the sparsity of X, both (L1) and (L0) have the same unique solution [17]. One such condition is the *Restricted Isometry Property* (see 2.10 below) due to Candès and Tao [9], which requires essentially that any $n \times s$ submatrix of $\mathbf{\Phi}$ is an approximate isometry. This property is satisfied by a number of matrices such as Gaussian random matrices or random partial Fourier matrices [9, 7, 34]. In that case, as long as $s \leq \mathcal{O}(n/\log(m))$, with high probability the solution of (L1) will indeed coincide with the solution of (L0). Another condition for which equivalence between (L0) and (L1) can be proven is based on the *incoherence* of the columns of $\mathbf{\Phi}$, which refers to the property that the inner product of any two columns of $\mathbf{\Phi}$ is small [15, 24, 36, 38]. Moreover, the performance of BP is

stable w.r.t. the presence of noise and error [8, 16, 37]. Finally the computational complexity of BP can be significantly reduced by using the various greedy algorithms in place of the linear programming technique [12, 30, 31, 36, 37]. The most basic greedy algorithm relevant here is Orthogonal Matching Pursuit (OMP) which has been thoroughly analyzed in [36].

For the imaging problem, the sensing matrix $\mathbf{\Phi}$ represents a physical process (typically wave propagation) and thus its entries cannot be arbitrarily chosen at our convenience. Therefore we cannot simply assume that $\mathbf{\Phi}$ satisfies any of the conditions that make compressed sensing work. The few physical parameters that we have control over are the wavelength λ of the probe wave, the locations and number n of sensors and the aperture A of the probe array. This is one of the reasons that make the practical realization of compressed sensing a challenging task.

The paper is organized as follows. In Section 2 we describe the physical setup, formulate the imaging problem in the framework of compressed sensing and state our main results. In Section 3 we prove the results for the source and monostatic synthetic aperture imaging. In Section 4 and Appendix A, we prove the main result for the multistatic response imaging. In Section 5 and Appendix B, we discuss the numerical methods and present simulations that confirm the predictions of our theorems.

2. Problem formulations and main results. In this paper, we study the inverse source and scattering problems in the linear regime. For simplicity and definiteness we consider the three dimensional space and assume that all targets are in the transverse plane $\{z = z_0\}$ and all sensors are in another transverse plane $\{z = 0\}$. The exact Green function for the Helmholtz equation which governs the monochromatic wave propagation is

$$G(\mathbf{r}, \mathbf{a}) = \frac{e^{i\omega|\mathbf{r}-\mathbf{a}|}}{4\pi|\mathbf{r}-\mathbf{a}|}, \quad \mathbf{r} = (x, y, z_0), \quad \mathbf{a} = (\xi, \eta, 0). \tag{2.1}$$

We assume that the phase speed $c = 1$ so that the frequency ω equals the wavenumber.

We consider the Fresnel diffraction for which the Green function (2.1) can be approximated by the paraxial Green function

$$\frac{e^{i\omega z_0}}{4\pi z_0} e^{i\omega|x-\xi|^2/(2z_0)} e^{i\omega|y-\eta|^2/(2z_0)}. \tag{2.2}$$

The paraxial approximation (2.2) is valid under, for instance the condition

$$\frac{(A+L)^4}{\lambda z_0^3} \ll 1 \tag{2.3}$$

where L is the linear dimension of the target domain (see [4] for a more general condition).

The formula (2.2) follows from truncating the Taylor expansion of the function $\omega|\mathbf{r} - \mathbf{a}|$

$$\omega|\mathbf{r} - \mathbf{a}| \approx \omega \left(z_0 + \frac{|x-\xi|^2}{2z_0} + \frac{|y-\eta|^2}{2z_0} \right)$$

and neglecting the higher order terms which are of the same order of magnitude as the left hand side of (2.3).

For notational simplicity later on, let us define

$$G_p(\mathbf{r}, \mathbf{a}) = e^{i\omega|x-\xi|^2/(2z_0)} e^{i\omega|y-\eta|^2/(2z_0)} \tag{2.4}$$

without the factor $e^{i\omega z_0}/(4\pi z_0)$.

In the case of the inverse source problem, the corresponding sensing matrix $\mathbf{\Phi}$ is essentially made of the paraxial Green function for various points in the sensor array and the target domain. In this set-up, the entries (2.2) of the paraxial sensing matrix have the same magnitude so without loss of generality the column vectors of $\mathbf{\Phi}$ are assumed to have unit ℓ^2-norm.
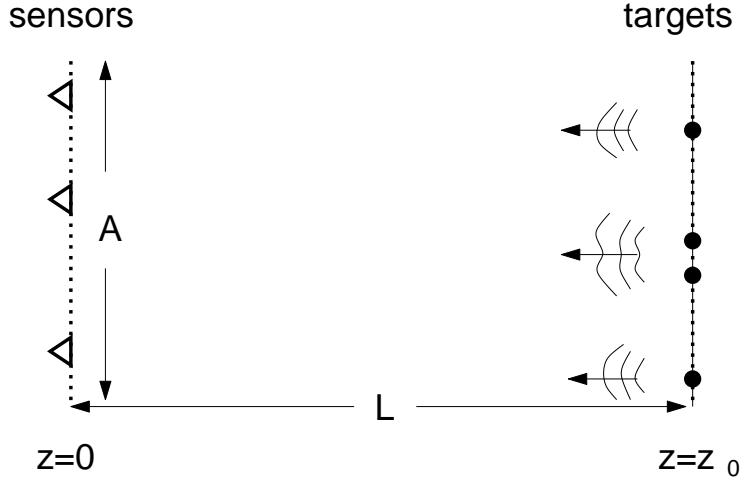
FIG. 2.1. *The imaging geometry*

A key idea in our construction of a suitable sensing matrix is to randomize the locations $\mathbf{a}_j = (0, \xi_j, \eta_j), j = 1, ..., n$ of the n sensors within a fixed aperture which is the linear size of the area (possibly) occupied by sensors (a square of size A for example). Indeed, we assume ξ_j, η_j are independent uniformly distributed in $[0, A]$. We assume that the antenna elements are independently uniformly distributed in a square array $[0, A] \times [0, A]$ in the plane $\{z = 0\}$, see Figure 2.1. Define the *sensor ensemble* to be the sample space of n *i.i.d. uniformly* distributed points in $[0, A]^2$.

We consider the idealized situation where the locations of the targets are a subset of a finite square lattice $\mathcal{M}$ of spacing ℓ

$$\mathcal{M} = \{\mathbf{r}_l : l = 1, ..., m\} = \left\{ (i\ell, j\ell, z_0) : i, j = 1, ..., \sqrt{m} \right\}, \quad l = (i-1)\sqrt{m} + j$$

in the transverse plane $\{z = z_0\}$. Hence the total number of grid points m is a perfect square. We defer the discussion on extended targets to the concluding section.

Let $\mathcal{S} = \{\mathbf{r}_{j_l} : l = 1, ..., s\}$ be the set of target locations and $\tau_{j_l}, l = 1, ..., s$ be the (source or scattering) amplitudes of the targets. Set $\tau_i = 0, i \notin \{j_1, ..., j_s\}$.

For source inversion each source point emits a signal described by (2.1) times the target amplitude which is recorded by the sensors. After proper normalization, the data vector Y can be written as

$$Y = \mathbf{\Phi} X + E \tag{2.5}$$

with the sensing matrix $\mathbf{\Phi} = n^{-1/2} \mathbf{\Phi}^{\mathrm{SI}}$ where

$$\Phi_{ij}^{\mathrm{SI}} = G_p(\mathbf{a}_i, \mathbf{r}_j), \quad \forall i = 1, ..., n, \quad j = 1, ..., m \tag{2.6}$$

and the error term E includes the mismatch between the exact Green function (2.1) and the paraxial Green function (2.2). We can write

$$E = \mathbf{\Phi}' X + E' \tag{2.7}$$

where the matrix $\mathbf{\Phi}'$ has the elements

$$\Phi_{ij}' = n^{-1/2} \left(G(\mathbf{a}_i, \mathbf{r}_j) - G_p(\mathbf{a}_i, \mathbf{r}_j) \right), \quad \forall i = 1, ..., n, \quad j = 1, ..., m \tag{2.8}$$

and E' stands for the external noise as well as the off-grid error when the targets do not exactly sit on the grid points (see Figure 5.5 and the related discussion in Section 5). When the target is sufficiently sparse and (2.3) is sufficiently enforced, $E \approx E'$.

3

For the noisy measurement model (2.5) instead of (L1) we consider the constrained minimization principle called the basis pursuit denoising (BPDN) [5]

$$\min_{Z \in \mathbb{C}^m} \|Z\|_1, \quad \text{s.t. } \|Y - \mathbf{\Phi}Z\|_2 \le \epsilon \tag{2.9}$$

where ϵ is the size of error, i.e. $\|E\|_2 \le \epsilon$. Denote the BPDN solution by $\hat{X}$.

The main reason that we consider the paraxial regime and the paraxial sensing matrix (2.6) is that the paraxial sensing matrix satisfies the restricted isometry property (RIP) which plays a crucial role in the current formulation of compressed sensing.

Precisely, let the sparsity s of a vector be the number of nonzero components and define the restricted isometry constant (RIC) $\delta_s < 1$ to be the smallest positive number such that the inequality

$$(1 - \delta_s)\|Z\|_2^2 \le \|\mathbf{\Phi}Z\|_2^2 \le (1 + \delta_s)\|Z\|_2^2$$

holds for all $Z \in \mathbb{C}^m$ of sparsity at most s. A sensing matrix is said to satisfy the $2s$-RIP if

$$\delta_{2s} < \sqrt{2} - 1 \tag{2.10}$$

[6]. Hence if the paraxial sensing matrix satisfies (2.10) and the paraxial condition (2.3) is sufficiently enforced such that

$$\|\mathbf{\Phi}'\|_2^2 < \sqrt{2} - 1 - \delta_{2s}$$

where $\|\mathbf{\Phi}'\|_2$ is the spectral norm of $\mathbf{\Phi}'$ then the exact sensing matrix also satisfies the $2s$-RIP. In this case, we can use the exact sensing matrix for inversion (see [27] for a different approach to handling multiplicative noise).

However, we do not know if the $2s$-RIP holds for the exact sensing matrix for a discrete grid in a non-paraxial regime. In the Littlewood-Paley basis the exact sensing matrix with a novel sampling scheme can be shown to satisfy the $2s$-RIP in any regime [19, 21]. But the Littlewood-Paley basis is not suitable to represent localized targets.

We consider also non-sparse, but *compressible*, targets. For the target vector X let $X^{(s)}$ denote the best s-sparse approximation of X in the sense of L^1-norm, i.e.

$$X^{(s)} = \operatorname{argmin} \|Z - X\|_1, \quad \text{s.t. } \|Z\|_0 \le s$$

where $\|Z\|_0$ denotes the number of nonzero components, called the sparsity, of Z. Clearly, $X^{(s)}$ consists of the s largest components of X.

The first main result proved in this paper is this.

THEOREM 2.1. *Let* $\mathbf{\Phi} = n^{-1/2}\mathbf{\Phi}^{\text{SI}}$. *If*

$$\frac{\ell A}{\lambda z_0} \equiv \frac{1}{\rho} \in \mathbb{N} \tag{2.11}$$

and

$$\frac{n}{\ln n} \ge C s \ln^2 s \ln m \ln \frac{1}{\gamma} \tag{2.12}$$

for $\gamma \in (0, 1)$ *and some absolute constant* C, *then*

$$\|\hat{X} - X\|_2 \le C_1 s^{-1/2}\|X - X^{(s)}\|_1 + C_2 \epsilon \tag{2.13}$$

for some constants C_1 *and* C_2 *with probability at least* $1 - \gamma$ *of the sensor ensemble.* The relation (2.11) indicates the existence of the threshold, optimal aperture given by $\lambda z_0/\ell$ corresponding to $\rho = 1$ (see

4

Remark 2 for more discussion on this point). Since the meshsize ℓ has the meaning of resolution, $\rho = 1$ is consistent with the classical Rayleigh criterion [4]

$$\ell \geq \frac{\lambda z_0}{A} \tag{2.14}$$

whose right hand side has the meaning of the spot size determined by the sensor array. The difference between the current framework and the classical resolution theory is that our sensor array is *sparse* and contains only $\mathcal{O}(s)$ sensors (up to a logarithmic factor).

Next we consider imaging methods for targets that are scatterers instead of sources. For point scatterers of amplitudes τ_{j_l} located at $\mathbf{r}_{j_l}, l = 1, 2, 3, ...s$, the resulting Green function $\tilde{G}(\mathbf{r}, \mathbf{a})$ between the sensor plane $\mathbf{a} \in \{z = 0\}$ and the target plane $\mathbf{r} \in \{z = z_0\}$, including the multiple scattering effect, obeys the Lippmann-Schwinger equation

$$\tilde{G}(\mathbf{r}, \mathbf{a}_i) = G(\mathbf{r}, \mathbf{a}_i) + \sum_{l=1}^{s} \tau_{j_l} G(\mathbf{r}, \mathbf{r}_{j_l}) \tilde{G}(\mathbf{r}_{j_l}, \mathbf{a}_i), \quad i = 1, ..., n,$$

for any $\mathbf{r} \neq \mathbf{r}_{j_l}, l = 1, ..., s$. The exciting field $\tilde{G}(\mathbf{r}_{j_l}, \mathbf{a})$ is part of the unknown and can be obtained from the so called Foldy-Lax equation (see e.g. [29, 22] for details). In the paraxial approximation, G above is to be replaced by its paraxial counterpart.

Hence, the inverse scattering problem is intrinsically nonlinear. However, often a linear scattering model is a good approximation and widely used in, e.g. radar imaging in the regimes of physical optics and geometric optics [3, 11] (see [39] for a precise formulation of the condition).

One such model is the Born approximation (also known as Rayleigh-Gans scattering in optics) in which the unknown exciting field is replaced by the incident field resulting in

$$\tilde{G}(\mathbf{r}, \mathbf{a}_i) - G(\mathbf{r}, \mathbf{a}_i) = \sum_{l=1}^{s} \tau_{j_l} G(\mathbf{r}, \mathbf{r}_{j_l}) G(\mathbf{r}_{j_l}, \mathbf{a}_i), \quad i = 1, ..., n. \tag{2.15}$$

The left hand side of (2.15) is the scattered field in the Born approximation when the incident field is emitted from a point source at $\mathbf{a}_i$. The Born approximation linearizes the relation between the scatterers and the scattered field. The goal of inverse scattering is to reconstruct the targets given the measurements of the scattered field.

In the simplest setting, called the (monostatic) synthetic aperture (SA) imaging, the real, physical array consists of only one receive/transmit antenna. The imaging aperture is synthesized by the antenna taking different transmit-receive positions $\mathbf{a}_i, i = 1, ..., n$ [23].

The SA imaging considered here is motivated by synthetic aperture radar (SAR) imaging. Monostatic SAR is a technique where a substantial aperture can be synthesized by moving a transmit-receive antenna along a trajectory and repeatedly interrogating a search area by firing repeated pulses from the antenna and measuring the responses. This can greatly leverage a limited probe resource and has many applications in remote sensing. The image formation is typically obtained via the matched filter processing (MFP) [11]. In Appendix B, we outline the rudiments of the linear processor of MFP.

First we consider a simplified set-up, neglecting the Doppler effect associated with the relative motion between the antenna and targets. Define the target vector X to be $X = (\tau_j) \in \mathbb{C}^m$. In this case, the sensing matrix $\boldsymbol{\Phi} = n^{-1/2} \boldsymbol{\Phi}^{\mathrm{SA}}$ with

$$\Phi_{ij}^{\mathrm{SA}} = G_p^2(\mathbf{a}_i, \mathbf{r}_j), \quad i = 1, ..., n, \quad j = 1, ..., m \tag{2.16}$$

after normalization. In other words, $\Phi_{ij}^{\mathrm{SA}} = \Phi_{lj}^{\mathrm{MR}}$ with $l = n(i-1) + i$. A crucial observation about SA imaging is that

$$G_p^2(\mathbf{a}_i, \mathbf{r}_j; \omega) = G_p(\mathbf{a}_i, \mathbf{r}_j; 2\omega), \tag{2.17}$$

cf. (2.4). In other words, this version of monostatic SA imaging is mathematically equivalent to the source inversion with twice the frequency. Therefore, the following theorem is an immediate consequence of Theorem 2.1.

THEOREM 2.2. Let $\boldsymbol{\Phi} = n^{-1/2}\boldsymbol{\Phi}^{\mathrm{SA}}$. Under the conditions

$$\frac{2}{\rho} \in \mathbb{N} \tag{2.18}$$

and (2.12), the estimate (2.13) holds with probability at least $1 - \gamma$ of the sensor ensemble. As a result of the SA aperture condition (2.18), the corresponding optimal aperture is half of that for the inverse source and MR imaging. In other words, SA can produce the qualitatively optimal performance with half of the aperture. This two-fold enhancement of resolving power in SA imaging has been previously established for the matched-field technique [23].

Next, we consider a multistatic imaging set-up in which we use the real array aperture as in the inverse source problem discussed above and moreover the array is also the source of n probe waves. One by one, each antenna of the array emits an impulse and the entire array receives the echo. Each transmitter-receiver pair gives rise to a datum and there are altogether n^2 data forming a datum matrix called the multistatic response matrix. This is called the *multistatic response* (MR) imaging [22, 23].

Since the full response matrix is symmetric, we use only $n(n + 1)/2$ of its elements (say, the upper triangular part including the diagonals) to form the data vector. Then the j-th column of the sensing matrix corresponds to the responses due to the presence of a scatterer at the j-th grid point of the computation lattice. By (2.15) in the paraxial approximation the corresponding sensing matrix is $\boldsymbol{\Phi} = \sqrt{\frac{2}{n(n+1)}}\boldsymbol{\Phi}^{\mathrm{MR}}$ with

$$\Phi_{lj}^{\mathrm{MR}} = G_p(\mathbf{a}_i, \mathbf{r}_j)G_p(\mathbf{r}_j, \mathbf{a}_k), \quad l = 1, .., n(n+1)/2, \quad j = 1, ..., m \tag{2.19}$$

after proper normalization where l is related to i, k as

$$l = \frac{(2n + 2 - i)(i - 1)}{2} + k - i + 1, \quad k \geq i$$

provided that the data vector is given by the upper triangular part (including the diagonals) of the multistatic response matrix.

Now that we do not know if the sensing matrix (2.19) satisfies the $2s$-RIP or not we will use the alternative approach to compressed sensing based on the notion of mutual coherence. For any matrix $\boldsymbol{\Phi}$ the mutual coherence of the matrix is defined by

$$\mu(\boldsymbol{\Phi}) = \max_{i \neq j} \frac{|\sum_k \Phi_{ik}^* \Phi_{kj}|}{\sqrt{\sum_k |\Phi_{ki}|^2}\sqrt{\sum_k |\Phi_{kj}|^2}}.$$

Our main result for the MR imaging is this.

THEOREM 2.3. *Let* $\boldsymbol{\Phi} = \sqrt{\frac{2}{n(n+1)}}\boldsymbol{\Phi}^{\mathrm{MR}}$. *Assume the aperture condition (2.11) and*

$$m \leq \frac{\delta}{2}e^{K^2/2} \tag{2.20}$$

for some positive constants δ and K.

(i). *Consider only the sensor ensemble. Then for any real-valued target vector $X \in \mathbb{R}^m$ of sparsity up to*

$$s < \frac{1}{4}\left(1 + \frac{n}{2K^2}\right)$$

the minimizer $\hat{X}$ of (2.9) obeys the error bound

$$\|\hat{X} - X\|_2 \leq \frac{2\epsilon}{\sqrt{1 - \theta}}, \quad \theta = \frac{2K^2(4s - 1)}{n} < 1$$

6

with probability at least $1 - 2\delta$.

(ii). *In addition to the sensor ensemble, consider also the* target ensemble *consisting of target vectors with at most s non-zero entries whose phases are independently, uniformly distributed in* $[0, 2\pi]$ *and whose support indices are independently and randomly selected from the index set* $\{1, 2, ..., m\}$.

Suppose $E = 0$ in (2.5). Then the target vector of sparsity

$$s \le \frac{n^2}{32K^4 \ln \frac{m}{\gamma}}, \tag{2.21}$$

where $\gamma > 0$ such that

$$K^4 \ln \frac{m}{\gamma} \ge e^{1/4}, \tag{2.22}$$

can be recovered exactly by BP with probability at least

$$\left[1 - 2\delta - \frac{\rho n^{5/2}(n+1)^{5/2}}{\pi 2^{5/2} m^{1/2}}\right] \times \left[1 - 2\gamma - 2^p s^{-p}\right], \quad p = \frac{\ln m - \ln \gamma}{144\sqrt{e} \ln s}. \tag{2.23}$$

The proof is given in Section 4. Note that the probability bound (2.23) is not nearly as good as that in Theorems 2.1 and 2.2. In particular, for (2.23) to be close to unity, it is required that $m \gg n^{10}$. In other words, only extremely sparse random target vectors are proved to be exactly recoverable with high probability and a nearly optimal number $\mathcal{O}(\sqrt{s})$ of sensors. In practice, however, the lower bound (2.23) appears to grossly underestimate the actual success probability. The other drawback of Theorem 2.3(ii) is the error-free assumption. Strictly speaking this assumption is only valid for imaging with *exact* sensing matrix in the case of Born scattering. Since the main purpose of Theorem 2.3(ii) is to illustrate the numerically observed $\mathcal{O}(n^2)$ behavior of the recoverable sparsity (2.21) (cf. Figure 5.3), we will briefly comment on how to extend Theorem 2.3(ii) to the case of exact sensing matrix in the paraxial regime (2.3) (see Remark 3).

We end this section by relating our works to previous results in the literature. In [2, 25] radar imaging is formulated in the space-time domain and the compressed sensing methodology is applied to *processing* of wide-band radar images where the incoherent sensing matrix arises from random samplings and projections of received signals. Detailed numerical and experimental studies of several sampling schemes are given in [25]. The first rigorous compressed sensing analysis of radar imaging is carried out in [26] for one spatial dimension where the incoherent sensing matrix arises from an incoherent probe signal. In contrast, in our formulation the signals are monochromatic and the incoherent sensing matrix arises from the random locations of sensors. In other words, the compression takes place in the sensor array and there is no processing of the measurement data besides the L^1-based inversion. The main advantage of our approach is, of course, reduction of array complexity (i.e. our sensor array is sparse). The single-frequency feature of our set-up is analogous to the one-pixel feature of Rice's single-pixel camera [10, 18]. Needless to say, the imaging performance can be enhanced by using multiple frequencies properly. For example, it is shown in [21] that multiple lower frequencies can be combined to achieve the same performance of a single higher frequency. From another perspective, the present work can be viewed as complementary to [26] in the sense that while the latter deals with the range information we address the directional aspect of imaging.

3. Source inversion and monostatic SA imaging: RIP. In this section, we show that the source and monostatic SA imaging schemes satisfy the 2s-RIP.

First we state the main result of the RIP approach [6].

PROPOSITION 3.1. *[6] Suppose the restricted isometry constant of* $\mathbf{\Phi}$ *satisfies the inequality*

$$\delta_{2s} < \sqrt{2} - 1. \tag{3.1}$$

Then the solution $\hat{X}$ of BPDN satisfies

$$\|\hat{X} - X\|_2 \le C_1 s^{-1/2} \|X - X^{(s)}\|_1 + C_2 \epsilon \tag{3.2}$$

where

$$C_1 = \frac{2(1 + (\sqrt{2} - 1)\delta_{2s})}{1 - (\sqrt{2} + 1)\delta_{2s}} \tag{3.3}$$

$$C_2 = \frac{4\sqrt{1 + \delta_{2s}}}{1 - (\sqrt{2} + 1)\delta_{2s}}. \tag{3.4}$$

Now we show that the sensing matrix for source inversion satisfies the $2s$-RIP. This can be readily seen by rewriting the paraxial Green function (2.2)

$$\frac{e^{i\omega z_0}}{4\pi z_0} e^{i\omega(x^2 + y^2)/(2z_0)} e^{-i\omega x\xi/z_0} e^{-i\omega y\eta/z_0} e^{i\omega(\xi^2 + \eta^2)/(2z_0)},$$

for $\mathbf{r} = (x, y, z_0), \mathbf{a} = (\xi, \eta, 0)$.

Now the sensing matrix (2.6) can be written as the product of three matrices

$$\mathbf{\Phi} = \mathbf{D}_1 \mathbf{\Psi} \mathbf{D}_2$$

where

$$\mathbf{D}_1 = \text{diag}(e^{i\omega(\xi_j^2 + \eta_j^2)/(2z_0)}), \quad \mathbf{D}_2 = \text{diag}(e^{i\omega(x_l^2 + y_l^2)/(2z_0)})$$

are unitary and

$$\mathbf{\Psi} = \frac{1}{\sqrt{n}} \left[e^{-i\omega\xi_j x_l/z_0} e^{-i\omega\eta_j y_l/z_0} \right]. \tag{3.5}$$

Assume without loss of generality that $(x_l, y_l), l = 1, ..., m$ are the grid points of the square lattice of spacing ℓ and suppose that $(\xi_j, \eta_j), j = 1, ..., n$ are independently and uniformly distributed in $[0, A] \times [0, A]$ with

$$\frac{1}{\rho} \equiv \frac{A\ell}{\lambda z_0} = 1 \tag{3.6}$$

which has the obvious generalization (2.11).

The result essential for us is due to Rauhut [33].

PROPOSITION 3.2. *[33] If*

$$\frac{n}{\ln n} \geq C\delta^{-2} s \ln^2 s \ln m \ln \frac{1}{\gamma} \tag{3.7}$$

for $\gamma \in (0, 1)$ and some absolute constant C, then with probability at least $1 - \gamma$ the matrix $\mathbf{\Psi}$ defined by (3.5)-(3.6) satisfies the RIC bound

$$\delta_s \leq \delta. \tag{3.8}$$

See [7, 34] for similar results for sensors located in a particular discrete subset of $[0, A] \times [0, A]$.

Since $\mathbf{D}_1$ and $\mathbf{D}_2$ are diagonal and unitary, $\mathbf{\Phi}$ satisfies (3.1) if and only if $\mathbf{\Psi}$ satisfies the same condition. To finish the proof of Theorem 2.1 we can choose any $\delta < \sqrt{2} - 1$ and apply Propositions 3.2 and 3.1.

From the relationships (2.16), (2.17) it follows immediately that $\mathbf{\Phi}^{\text{SA}}$ also satisfies (3.1) if

$$\rho = 2, \tag{3.9}$$

which can be generalized to (2.18).

The superiority of the RIP approach, when it works, is that the recovery performance is guaranteed for *all* compressible targets under a nearly optimal sparsity constraint. And the target ensemble needs not be introduced.

However, for the sensing matrix for the multistatic response imaging does not seem amenable to the RIP approach. Below we resort to the incoherence approach to compressed sensing.

4. Proof of Theorem 2.3: MR imaging. For the coherence calculation, we have

$$\sum_{l=1}^{n^2} \Phi_{li}^{\mathrm{MR}*} \Phi_{lj}^{\mathrm{MR}} = \sum_{p,q=1}^{n} G_p(\mathbf{a}_p, \mathbf{r}_j) G_p(\mathbf{r}_j, \mathbf{a}_q) G_p^*(\mathbf{a}_p, \mathbf{r}_i) G_p^*(\mathbf{r}_i, \mathbf{a}_q) \tag{4.1}$$

$$= \left[\sum_{p=1}^{n} G_p(\mathbf{a}_p, \mathbf{r}_j) G_p^*(\mathbf{a}_p, \mathbf{r}_i) \right]^2$$

and thus

$$\mu\left(\mathbf{\Phi}^{\mathrm{MR}}\right) = \mu^2(\mathbf{\Phi}^{\mathrm{SI}}). \tag{4.2}$$

THEOREM 4.1. *Under (2.11) and (2.20) the coherence of $\mathbf{\Phi}^{\mathrm{SI}}$ satisfies*

$$\mu(\mathbf{\Phi}^{\mathrm{SI}}) \leq \sqrt{2} K/\sqrt{n} \tag{4.3}$$

with probability greater than $(1-\delta)^2$. The proof of Theorem 4.1 is given in Section 4.3.

REMARK 1.

The general lower bound [13, 40]

$$\sqrt{\frac{m-n}{n(m-1)}} \leq \mu(\mathbf{\Phi}) \leq 1$$

for the mutual coherence of any $n \times m$ matrix $\mathbf{\Phi}$ implies that the coherence bound (4.3) is optimal asymptotically ($m \gg n$) modulo a constant factor.

REMARK 2. *Since the coherence of the sensing matrix should decrease as the aperture increases and since the coherence is of the same order of magnitude as $n^{-1/2}$ whenever (2.11) holds, simple interpolation leads to the conclusion that the coherence should be roughly constant for*

$$A \geq \frac{\lambda z_0}{\ell} \tag{4.4}$$

corresponding to $\rho \leq 1$. The right hand side of (4.4), corresponding to $\rho = 1$, defines the optimal aperture. See Figure 5.1 for numerical evidence.

COROLLARY 1. *Under the assumptions (2.18) and (2.20) the coherence of $\mathbf{\Phi}^{\mathrm{SA}}$ satisfies*

$$\mu(\mathbf{\Phi}^{\mathrm{SA}}) \leq \sqrt{2} K/\sqrt{n}$$

with probability greater than $1 - 2\delta$. This follows immediately from (2.16)-(2.17).

COROLLARY 2. *Under the assumptions (2.11) and (2.20) the coherence of $\mathbf{\Phi}^{\mathrm{MR}}$ satisfies*

$$\mu(\mathbf{\Phi}^{\mathrm{MR}}) \leq 2K^2/n$$

with probability greater than $1 - 2\delta$. This follows from (4.2).

4.1. Proof of Part (i). The following result is due to Donoho, Elad and Temlyakov (Theorem 3.1, [16])

PROPOSITION 4.2. *Suppose $\mathbf{\Phi}$ has all unit columns. Suppose the sparsity s of the target vector X satisfies*

$$s < \frac{1}{4}\left(1 + \frac{1}{\mu(\mathbf{\Phi})}\right).$$

Then the minimizer $\hat{X}$ of (2.9) obeys the error bound

$$\|\hat{X} - X\|_2 \leq \frac{2\epsilon}{\sqrt{1 - (4s-1)\mu(\mathbf{\Phi})}}.$$

Part (i) of Theorem 2.3 follows immediately from Proposition 4.2 and Corollary 2.

9

4.2. Proof of Part (ii). The following theorem is a reformulation of results due to Tropp [38] and the basis for our analysis of the MR imaging scheme. See [26, 32] for other applications.

PROPOSITION 4.3. *Let $\mathbf{\Phi} \in \mathbb{C}^{N,m}$ have all unit columns. Let X be drawn from the target ensemble. Assume that*

$$s \leq \left(8\mu^2(\mathbf{\Phi}) \ln \frac{m}{\gamma}\right)^{-1}, \quad \gamma \in (0,1) \tag{4.5}$$

and that for $p \geq 1$

$$3\left(\frac{p \ln s}{2 \ln \frac{m}{\gamma}}\right)^{1/2} + \frac{s}{m}\|\mathbf{\Phi}\|_2^2 \leq \frac{1}{4e^{1/4}}. \tag{4.6}$$

Then X is the unique solution of BP with probability $1 - 2\gamma - 2^p s^{-p}$. Here $\|\mathbf{\Phi}\|_2$ denotes the spectral norm of $\mathbf{\Phi}$. We explain the connection of the above statement with [38] in Appendix A.

To apply it to the MR imaging we set $\mathbf{\Phi} = N^{-1/2}\mathbf{\Phi}^{\mathrm{MR}}$ with $N = n(n+1)/2$.

The proof of Theorem 2.3 hinges on Corollary 2 and the following theorem.

THEOREM 4.4. *The matrix $\mathbf{\Phi}^{\mathrm{MR}}$ has full rank and obeys the spectral norm bound*

$$\|\mathbf{\Phi}^{\mathrm{MR}}\|_2^2 \leq 2m \tag{4.7}$$

with probability at least

$$1 - \frac{\rho n^{5/2}(n+1)^{5/2}}{\pi 2^{5/2} m^{1/2}}, \quad \rho = \frac{\lambda z_0}{\ell A}.$$

The proof of Theorem 4.4 is given in Section 4.4.

Conditions (2.21) and (2.20) allow the choice of K such that

$$2 \ln \frac{2m}{\delta} < K^2 < \frac{n}{\sqrt{32s \ln \frac{m}{\gamma}}}. \tag{4.8}$$

As a consequence (4.5) with $\mathbf{\Phi} = \sqrt{\frac{2}{n(n+1)}}\mathbf{\Phi}^{\mathrm{MR}}$ is satisfied with probability greater than $1 - 2\delta$ by Corollary 2.

Now the norm bound (4.7) implies (4.6) if

$$3\left(\frac{p \ln s}{2 \ln \frac{m}{\gamma}}\right)^{1/2} + \frac{4s}{n(n+1)} \leq \frac{1}{4e^{1/4}}, \quad p > 1, \tag{4.9}$$

which in turn follows from (2.21) and the condition

$$K^4 \ln \frac{m}{\gamma} \geq \frac{1}{8}\left(\frac{1}{4e^{1/4}} - 3\left(\frac{p \ln s}{2 \ln \frac{m}{\gamma}}\right)^{1/2}\right)^{-1}.$$

Hence under (2.22) we can choose p in (4.6) to be

$$p = \frac{\ln m - \ln \gamma}{144\sqrt{e} \ln s}.$$

Since Theorems 4.1 and 4.4 hold with probability greater than

$$1 - 2\delta - \frac{\rho n^{5/2}(n+1)^{5/2}}{\pi 2^{5/2} m^{1/2}}$$

10

and since the target ensemble is independent of the sensor ensemble we have the bound (2.23) for the probability of exact recovery.

REMARK 3. *When the paraxial condition (2.3) is sufficiently enforced, Theorem 2.3 (ii) can be generalized to the case of exact sensing matrix with*

$$\Phi_{lj}^{\mathrm{MR}} = G(\mathbf{a}_i, \mathbf{r}_j) G(\mathbf{r}_j, \mathbf{a}_k)$$

instead of the paraxial version (2.19). This is because the spectral norm of the difference matrix $\mathbf{\Phi}'$ can be made arbitrarily small by enforcing (2.3). As a consequence, both the mutual coherence and the spectral norm of the exact sensing matrix can be made arbitrarily close to those of the paraxial sensing matrix. The same analysis as above then leads to the desired result (cf. Figure 5.3).

4.3. Proof of Theorem 4.1: coherence bound. *Proof.* Summing over $\mathbf{a}_l, l = 1, ..., n$ we obtain

$$\sum_{l=1}^{n} \Phi_{li}^* \Phi_{lj} = e^{i\omega(x_j^2 + y_j^2 - x_i^2 - y_i^2)/(2z_0)} \frac{1}{n} \sum_{l=1}^{n} e^{i\xi_l \omega(x_i - x_j)/z_0} e^{i\eta_l \omega(y_i - y_j)/z_0}. \tag{4.10}$$

Define the random variables $X_l, Y_l, l = 1, ..., n$, as

$$X_l = \cos\left[(\xi_l(x_i - x_j) + \eta_l(y_i - y_j))\omega/z_0\right] \tag{4.11}$$
$$Y_l = \sin\left[(\xi_l(x_i - x_j) + \eta_l(y_i - y_j))\omega/z_0\right] \tag{4.12}$$

and their respective sums

$$S_n = \sum_{l=1}^{n} X_l, \quad T_n = \sum_{l=1}^{n} Y_l.$$

Then the absolute value of the right hand side of (4.10) is bounded by

$$\frac{1}{n}|S_n + iT_n| \le \frac{1}{n}\left(\sqrt{|S_n - \mathbb{E}S_n|^2 + |T_n - \mathbb{E}T_n|^2} + |\mathbb{E}(S_n + iT_n)|\right). \tag{4.13}$$

To estimate the right hand side of (4.13), we recall the Hoeffding inequality [28].

PROPOSITION 4.5. *Let $X_1, ..., X_n$ be independent random variables. Assume that $X_l \in [a_l, b_l], l = 1, ..., n$ almost surely. Then we have*

$$\mathbb{P}\left[|S_n - \mathbb{E}S_n| \ge nt\right] \le 2\exp\left[-\frac{2n^2t^2}{\sum_{l=1}^{n}(b_l - a_l)^2}\right] \tag{4.14}$$

for all positive values of t. We apply the Hoeffding inequality to both S_n and T_n. To this end, we have $a_l = -1, b_l = 1, \forall l$ and set

$$t = K/\sqrt{n}, \quad K > 0.$$

Then we obtain

$$\mathbb{P}\left[n^{-1}|S_n - \mathbb{E}S_n| \ge K/\sqrt{n}\right] \le 2e^{-K^2/2} \tag{4.15}$$
$$\mathbb{P}\left[n^{-1}|T_n - \mathbb{E}T_n| \ge K/\sqrt{n}\right] \le 2e^{-K^2/2}. \tag{4.16}$$

Note that the quantities S_n, T_n depend on $x_i - x_j, y_i - y_j$, i.e.

$$S_n = S_n(x_i - x_j, y_i - y_j), \quad T_n = T_n(x_i - x_j, y_i - y_j).$$

Note that the quantities S_n, T_n depend on $(x_i - x_j, y_i - y_j)$ but note the symmetry: $S_n(x_i - x_j, y_i - y_j) = S_n(x_j - x_i, y_j - y_i), T_n(x_i - x_j, y_i - y_j) = -T_p(x_j - x_i, y_j - y_i)$. Furthermore, a moment of reflection reveals

11

that thanks to the square symmetry of the lattice there are at most $m - 1$ different values $|S_n|$ and $|T_n|$ among the $m(m-1)/2$ pairs of points.

We use (4.15)-(4.16) and the union bound to obtain

$$\mathbb{P}\left[\max_{i \neq j} n^{-1} |S_n(x_i - x_j, y_i - y_j) - \mathbb{E}S_n(x_i - x_j, y_i - y_j)| \geq K/\sqrt{n}\right]$$

$$\leq 2(m-1)e^{-K^2/2}$$

$$\mathbb{P}\left[\max_{i \neq j} n^{-1} |T_n(x_i - x_j, y_i - y_j) - \mathbb{E}T_n(x_i - x_j, y_i - y_j)| \geq K/\sqrt{n}\right]$$

$$\leq 2(m-1)e^{-K^2/2}$$

and hence

$$\mathbb{P}\left[\max_{i \neq j} n^{-1} |S_n + iT_n - \mathbb{E}(S_n + iT_n)| \geq K\sqrt{\frac{2}{n}}\right] \leq 4(m-1)e^{-K^2/2} < 2\delta. \tag{4.17}$$

The third term on the right hand side of (4.13) can be calculated as follows. By the mutual independence of ξ_l and η_l we have

$$\frac{1}{n}|\mathbb{E}(S_n + iT_n)| = \frac{1}{n}\left|\sum_{l=1}^{n} \mathbb{E}(X_l + iY_l)\right|$$

$$= \frac{1}{n}\left|\sum_{l=1}^{n} \mathbb{E}\left(e^{i\xi_l\omega(x_i - x_j)/z_0}\right) \mathbb{E}\left(e^{i\eta_l\omega(y_i - y_j)/z_0}\right)\right|$$

$$= \left|\mathbb{E}\left(e^{i\xi_l\omega(x_i - x_j)/z_0}\right) \mathbb{E}\left(e^{i\eta_l\omega(y_i - y_j)/z_0}\right)\right|$$

since $\xi_l, \eta_l, l = 1, ..., n$ are independently identically distributed.

Simple calculation with the uniform distribution on $[0, A] \times [0, A]$ yields

$$\left|\mathbb{E}\left(e^{i\xi_l\omega(x_i - x_j)/z_0}\right) \mathbb{E}\left(e^{i\eta_l\omega(y_i - y_j)/z_0}\right)\right| = \left|\frac{e^{i\phi_{ij}} - 1}{\phi_{ij}}\right|\left|\frac{e^{i\psi_{ij}} - 1}{\psi_{ij}}\right|$$

$$= 4\left|\frac{\sin\frac{\phi_{ij}}{2}}{\phi_{ij}}\right|\left|\frac{\sin\frac{\psi_{ij}}{2}}{\psi_{ij}}\right| \tag{4.18}$$

with

$$\phi_{ij} = A\omega(x_i - x_j)/z_0, \quad \psi_{ij} = A\omega(y_i - y_j)/z_0.$$

The optimal condition is to choose A such that

$$\phi_{ij} = \psi_{ij} \in 2\pi\mathbb{Z}, \tag{4.19}$$

under which (4.18) vanishes. Condition (4.19) can be fulfilled for an equally spaced grid as is assumed here. Let

$$\ell = \min_{i \neq j} |x_i - x_j| = \min_{i \neq j} |y_i - y_j|.$$

The smallest ℓ satisfying condition (4.19) is given by

$$\ell = \frac{z_0\lambda}{A}, \quad \lambda = 2\pi/\omega \tag{4.20}$$

12

which can be interpreted as the resolution of the imaging system and is equivalent to the classical Rayleigh criterion.

In this case, $\mathbb{E}(S_n + iT_n) = 0$ and hence

$$\mu(\mathbf{\Phi}^{\text{SI}}) \leq \sqrt{2}K/\sqrt{n}$$

with probability greater than $1 - 2\delta$.
□

4.4. Proof of Theorem 4.4: spectral norm bound. For the proof, it suffices to show that the matrix $\mathbf{\Phi}$ satisfies

$$\|\frac{1}{m}\mathbf{\Phi}^{\text{MR}}\mathbf{\Phi}^{\text{MR}*} - \mathbf{I}_{n(n+1)/2}\|_2 < 1 \tag{4.21}$$

where $\mathbf{I}_{n(n+1)/2}$ is the $n(n+1)/2 \times n(n+1)/2$ identity matrix with the corresponding probability bound. By the Gershgorin circle theorem, (4.21) would in turn follow from

$$\mu\left(\mathbf{\Phi}^{\text{MR}*}\right) < \frac{1}{n(n+1)/2 - 1} \tag{4.22}$$

since the diagonal elements of $\frac{1}{m}\mathbf{\Phi}^{\text{MR}}\mathbf{\Phi}^{\text{MR}*}$ are unity.

For the MR setting, the pairwise coherence has the form

$$\frac{1}{m}\sum_{j=1}^{m}\Phi_{lj}^{\text{MR}}\Phi_{l'j}^{\text{MR}*} = \frac{1}{m}e^{i\omega(\xi_k^2 + \eta_k^2 + \xi_i^2 + \eta_i^2 - \xi_{k'}^2 - \eta_{k'}^2 - \xi_{i'}^2 - \eta_{i'}^2)/(2z_0)}$$

$$\times \frac{e^{i\omega(\xi_i + \xi_k - \xi_{i'} - \xi_{k'})(x_1 + \sqrt{m}\ell)/z_0} - e^{i\omega(\xi_i + \xi_k - \xi_{i'} - \xi_{k'})x_1/z_0}}{1 - e^{i\omega(\xi_i + \xi_k - \xi_{i'} - \xi_{k'})\ell/z_0}}$$

$$\times \frac{e^{i\omega(\eta_i + \eta_k - \eta_{i'} - \eta_{k'})(y_1 + \sqrt{m}\ell)/z_0} - e^{i\omega(\eta_i + \eta_k - \eta_{i'} - \eta_{k'})y_1/z_0}}{1 - e^{i\omega(\eta_i + \eta_k - \eta_{i'} - \eta_{k'})\ell/z_0}} \tag{4.23}$$

where

$$l = \frac{(2n + 2 - i)(i - 1)}{2} + k - i + 1, \quad k \geq i \tag{4.24}$$

$$l' = \frac{(2n + 2 - i')(i' - 1)}{2} + k' - i' + 1, \quad k' \geq i'. \tag{4.25}$$

Thus,

$$\frac{1}{m}\left|\sum_{j=1}^{m}\Phi_{lj}^{\text{MR}}\Phi_{l'j}^{\text{MR}*}\right| \leq \frac{1}{m}\left|\frac{\sin\frac{\omega\ell\sqrt{m}(\xi_i + \xi_k - \xi_{i'} - \xi_{k'})}{2z_0}}{\sin\frac{\omega\ell(\xi_i + \xi_k - \xi_{i'} - \xi_{k'})}{2z_0}}\right| \cdot \left|\frac{\sin\frac{\omega\ell\sqrt{m}(\eta_i + \eta_k - \eta_{i'} - \eta_{k'})}{2z_0}}{\sin\frac{\omega\ell(\eta_i + \eta_k - \eta_{i'} - \eta_{k'})}{2z_0}}\right| \tag{4.26}$$

where we have used the identity

$$\left|1 - e^{i\theta}\right| = 2\left|\sin\frac{\theta}{2}\right|. \tag{4.27}$$

Let

$$\kappa = \min_{l \neq l'}\min_{j \in \mathbb{Z}}\left\{\left|\frac{\ell(\xi_i + \xi_k - \xi_{i'} - \xi_{k'})}{\lambda z_0} - j\right|, \left|\frac{\ell(\eta_i + \eta_k - \eta_{i'} - \eta_{k'})}{\lambda z_0} - j\right|\right\} \tag{4.28}$$

where l and l' are related to (i, k) and (i', k'), respectively, by (4.24) and (4.25). Clearly κ is nonzero with probability one. For $l \neq l'$ the probability density functions (PDF) for the random variables

$$\frac{\ell(\xi_i + \xi_k - \xi_{i'} - \xi_{k'})}{\lambda z_0}, \quad \frac{\ell(\eta_i + \eta_k - \eta_{i'} - \eta_{k'})}{\lambda z_0}$$

13

are either the symmetric triangular distribution or its self-convolution supported on $[-2\rho^{-1}, 2\rho^{-1}]$. In either case, their PDFs are bounded by ρ. Hence the probability that $\{\kappa > \alpha\}$ for small $\alpha > 0$ is larger than

$$(1 - 2\rho\alpha)^{n(n+1)(n^2+n-2)/8} > 1 - \alpha\rho n(n+1)(n^2+n-2)/4$$

where the power counts the number of unordered different pairs (l, l'). Hence the desirable bound (4.22) follows from

$$\mu\left(\boldsymbol{\Phi}^{\mathrm{MR}*}\right) < \frac{1}{\pi^2\alpha^2 m}$$

with the choice

$$\alpha = \frac{1}{\pi}\sqrt{\frac{n(n+1)/2 - 1}{m}}$$

which holds true with probability larger than

$$1 - \alpha\rho n(n+1)(n^2+n-2)/4 > 1 - \frac{\rho n^{5/2}(n+1)^{5/2}}{\pi 2^{5/2} m^{1/2}}.$$

5. Numerical simulations . In the simulations, we set $z_0 = 10000$ and for the most part $\lambda = 0.1$ for the search domain $[-250, 250]^2$ with $\ell = 10$. For the SA setting, we also use $\lambda = 0.2$. The targets are i.i.d. uniform random variables of the grid with target amplitudes in the range $[1, 2]$. We randomly select antenna locations from $[-50, 50]^2$ with the aperture $A = 100$ which, according to Theorem 4.1, is the threshold, optimal aperture.

With these parameters

$$\frac{(A + L)^4}{\lambda z_0^3} \approx 1.3$$

the condition (2.3) is barely satisfied. On the other hand the Fresnel number is

$$\frac{(A + L)^2}{\lambda z_0} = 360 \gg 1$$

indicating that this is not the Fraunhofer diffraction regime [4] (see [20, 21] for compressed sensing analysis of inverse scattering in the Fraunhofer regime).

We use the true Green function (2.1) in the computation of scattered waves and in the inversion step the exact Green function and its paraxial approximation to construct the sensing matrix (for comparison). In other words, we allow model mismatch between the propagation and inversion steps. The degradation in performance can be seen in the figures but is still manageable.

In our simulations we have used the Matlab codes YALL1 (acronym for *Your ALgorithms for L1*, available at `http://www.caam.rice.edu/ optimization/L1/YALL1/`) and SP (acronym for *Subspace Pursuit*, available at `http://igorcarron.googlepages.com/cscodes`). YALL1 is a L1-minimization solver based on the Alternating Direction Method [41] while SP is a greedy algorithm which has a significantly lower computational complexity and, under the condition $\delta_{3s} < 0.06$, is guaranteed to exactly recover X of sparsity at most s via a finite number of iterations [12].

Our experiences indicate that YALL1 has a slightly better performance when the sensing matrix is based on the exact Green function. For the paraxial sensing matrix, however, the error tolerance (i.e. $\|E\|_2$ in (2.5)) needs to be pre-estimated in order to optimize the performance. In comparison, the SP code does not require input of error tolerance and its performance with the paraxial sensing matrix is close to that with the exact sensing matrix. Hence we will show the numerical results with SP only.

In Figure 5.1 (left), the coherence is calculated with aperture $A \in [10, 200]$ and $n = 100$ for the sensing matrices with the exact Green function (red-solid curve) as entries and its paraxial approximation (black-asterisk curve). The coherence of the exact sensing matrix at the borderline of the paraxial regime with
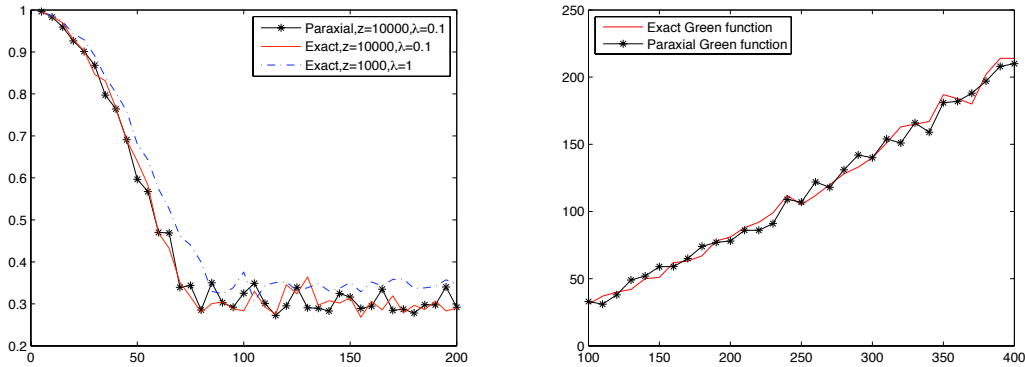
FIG. 5.1. *(left) The red-solid and black-asterisk curves are, respectively, the coherence for the exact and paraxial sensing matrices for $z_0 = 10000, \lambda = 0.1, n = 100$ as a function of aperture. The blue-dashed curve is the coherence for the exact sensing matrix for $z_0 = 1000, \lambda = 1, n = 100$; (right) the empirical, maximum number of recoverable sources v.s. the number n of antennas by using the paraxial (black-asterisk) and exact (red-solid) sensing matrices.*

$z_0 = 1000, \lambda = 1$ is also calculated (blue-dashed curve). All three curves track one another closely and flatten near and beyond $A = 100$ in agreement with the theory (Theorem 4.1), indicating the validity of the optimal aperture throughout the paraxial regime.

Figure 5.1 (right) displays the numerically found maximum number of recoverable source points as a function of n by using the exact (red-solid curve) and paraxial (black-asterisk curve) sensing matrices. The maximum number of recoverable targets (MNRT) is in principle a random variable as our theory is formulated in terms of the target and sensor ensembles. To compute MNRT, we start with one target point and apply the sensing scheme. If the recovery is (nearly) perfect a new target vector with one additional support is randomly drawn and the sensing scheme is rerun. We iterate this process until the sensing scheme fails to recover the targets and then we record the target support in the previous iterate as MNRT. The criterion for recovery is up to 5% relative error unless otherwise specified. *This is an one-trial test and no averaging is applied.* The linear profile in Figure 5.1 (right) is consistent with the prediction (3.7) of Theorem 2.1. Of course, the performance for the Born scatterers is still expected to be better than that with the Foldy-Lax scatterers and it is indeed the case.

For the purpose of demonstrating the degrading effect of multiple scattering on the sensing matrix based on the Born approximation, we use both the Foldy-Lax scheme and the Born approximation in the scattering simulation [22]. We shall refer the former targets as the Foldy-Lax scatterers and the latter as the Born scatterers. Hence there may be up to two mismatches (the paraxial approximation and the Born approximation) in the process depending on the simulations.

In Figure 5.2 (left) the compressed sensing image for source inversion is shown. For comparison, the image obtained by the linear processor of MFP is shown in Figure 5.2 (right). While it is clear that the compressed sensing image is much clearer, it also demands a higher computational complexity since a large matrix needs to be generated and stored. See Appendix B for the definition of the linear processor.

In Figure 5.3, the numerically found maximum number of recoverable Born scatterers is depicted as a function of the number of antennas for both SA and MR imaging by using the paraxial and exact sensing matrices. The result for the MR imaging shows clearly the $\mathcal{O}(n^2)$ behavior of the recoverable targets as predicted by Theorem 2.3(ii). The SA results for the Foldy-Lax scatterers in two dynamic ranges are depicted in Figure 5.4.

Recoveries for the Foldy-Lax scatterers in the dynamic range $\tau \in [1, 2]$ mostly fail as the mismatch is too great for the compressed sensing techniques to handle. Indeed, the performances with the Foldy-Lax scatterers in the range $[0.01, 0.02]$ are significantly degraded by the multiple scattering effects (Figure 5.4, left, versus, Figure 5.3, left). However, if the dynamic range is decreased to, say $[0.001, 0.002]$, then multiple scattering is negligible and the recovery performance is essentially restored, Figure 5.4 (right). Note that in
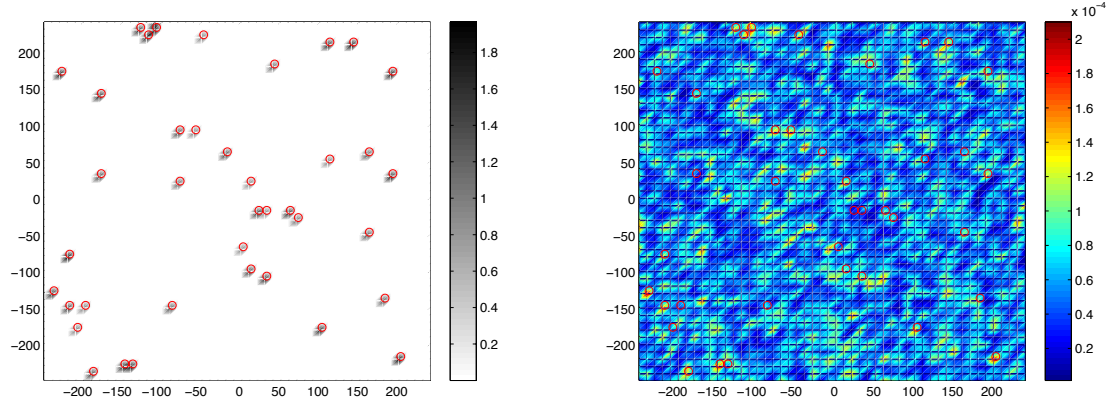
FIG. 5.2. *(left) Source inversion with the paraxial sensing matrix* 40 *source points and* 121 *antennas. The resulting error is* 0.0164 *while the error with exact Green function is* 7×10^{-16} *(not shown). (right) MFP image produced on the same grid. The red circles represent the true locations of the targets in both plots.*
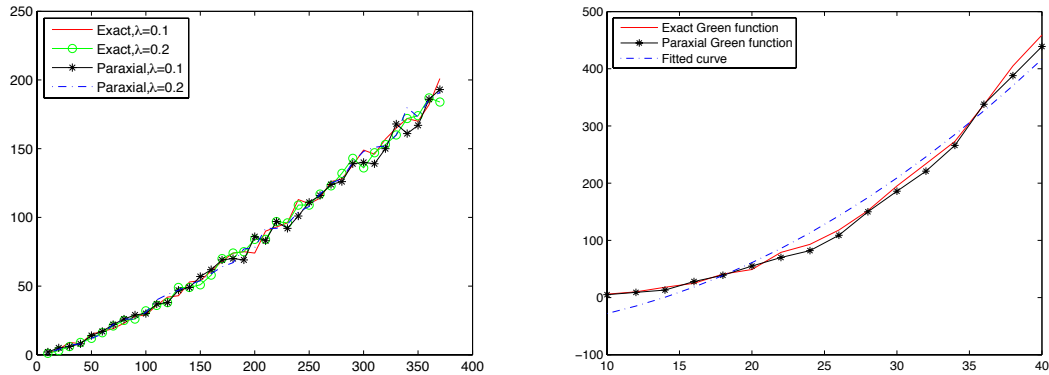

FIG. 5.3. *The empirical maximum number of recoverable* Born *scatterers (left for SA, right for MR) with* $\tau \in [1, 2]$ *v.s. the number n of antennas (or antenna locations). The data for* $n \in [10, 30]$ *in the MR plot is fitted with the parabola (blue-dashed curve):* $-57.5400 + 0.2964n^2$ *The wavelength is* 0.1 *for the MR case. The SA plot depicts the number of recoverable scatterers in four settings: paraxial sensing matrix with* $\lambda = 0.1$ *(black-asterisk), paraxial sensing matrix with* $\lambda = 0.2$ *(blue-dashed), exact sensing matrix with* $\lambda = 0.1$ *(red-solid) and exact sensing matrix with* $\lambda = 0.2$ *(green-circled)*

the intermediate range $\tau \in [0.01, 0.02]$ the performance with $\lambda = 0.1$ is significantly worse than that with $\lambda = 0.2$. This is because the multiple scattering effects are more pronounced for higher frequencies. Hence when the multiple scattering effects are significant, one should choose the longest wavelength allowed by the condition (2.18).

Finally, we consider the off-the-grid error by randomly perturbing the target locations by the uniform random variables in $[-f\ell/2, f\ell/2]^2, f \in [0, 1]$. We calculate the relative errors in source inversion as a function of f over 500 realizations of source locations, amplitudes and perturbations and take average. The results are shown in Figure 5.5. Note that the relative error decreases with the number of sensors and increases with the sparsity. This effect is qualitatively captured in the bounds (3.7), which decreases roughly as $n^{-1/2}$, and (3.4), which decreases to the limiting value 4 as n increases. Likewise, the bounds (3.7) increases roughly as $s^{1/2}$, and hence (3.4) increases as s increases.

6. Conclusions. In this paper, we have studied the imaging problem from the perspective of compressed sensing, in particular the idea that sufficient incoherence and sparsity guarantee uniqueness of the
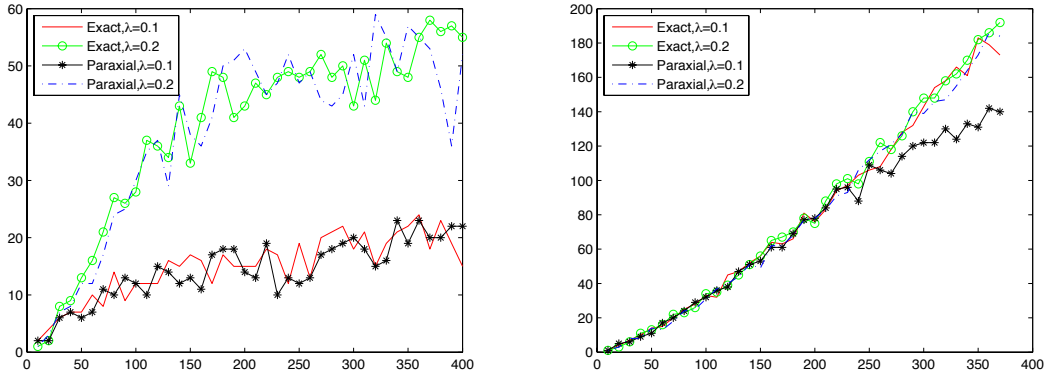
FIG. 5.4. *The empirical maximum number of* Foldy-Lax *scatterers recoverable by SA versus the number of antenna locations for the dynamic range* $\tau \in [0.01, 0.02]$ *(left) and the dynamic range* $\tau \in [0.001, 0.002]$ *(right).*
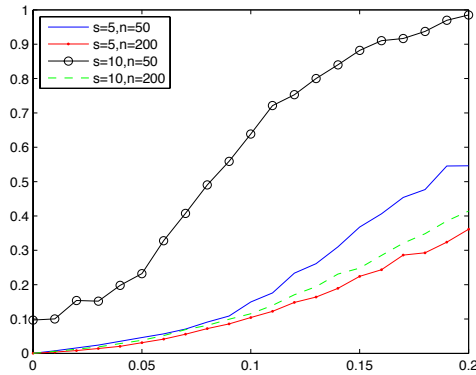

FIG. 5.5. *The stability performance of source inversion in the presence of off-grid errors with* $n = 50, 200, s = 5, 10$. *The vertical axis is for the relative error of recovery and the horizontal axis is for the range of the random perturbations in the unit of the grid spacing* ℓ.

solution. Moreover, by adopting the target ensemble following [38] and the sensor ensemble, the maximum number of recoverable targets is proved to be at least proportional to the number of measurement data modulo a log-square factor with overwhelming probability.

We have analyzed three imaging settings: the source inversion, the inverse scattering with the monostatic synthetic aperture and with the multistatic responses. Important contributions of our analysis include the discoveries of the incoherence effect induced by random antenna locations and the threshold aperture defined by $\rho = 1$ for source and MR imaging and $\rho = 1/2$ for SA imaging where $\rho = \lambda z_0/(A\ell)$.

In this paper we have considered the localization of point targets and the determination of their amplitudes. A natural next step is to consider extended targets. However our approach does not extend in a straightforward manner to imaging of extended targets, as can be easily seen. Assume that we model an extended target approximately as an ensemble of point targets that are spaced very close together. Clearly, this requires the grid spacing ℓ to be so small as to render $\rho \gg 1$. To apply our theorems would then require that the aperture and the number of antennas increase without bound. Clearly this is not a feasible way to image extended targets. Two different approaches to imaging extended targets by using a wavelet basis and an interpolation scheme are presented in [21].

Appendix A. Proof of Proposition 4.3. Proposition 4.3 is an easy consequence from the following

17

two theorems due to Tropp [38].

PROPOSITION A.1. *[38] Let $\boldsymbol{\Phi}$ be a $N \times m$ matrix with full rank and all unit columns. Let $\boldsymbol{\Phi}_S$ be a submatrix generated by randomly selecting columns of $\boldsymbol{\Phi}$ indexed by the set $S, \#S = s$. The condition*

$$6 \left(p\mu^2(\boldsymbol{\Phi})s\ln(1+s/2)\right)^{1/2} + \frac{s}{m}\|\boldsymbol{\Phi}\|_2^2 \leq \frac{\alpha}{2e^{1/4}}, \quad p \geq 1 \tag{A.1}$$

implies that

$$\mathbb{P}\left(\|\boldsymbol{\Phi}_S^*\boldsymbol{\Phi}_S - \mathbf{I}_s\|_2 < \alpha\right) \geq 1 - \left(\frac{2}{s}\right)^p. \tag{A.2}$$

PROPOSITION A.2. *[38] Let X be drawn from the target ensemble. If*

$$\mu^2 s \leq \left(8\ln\frac{m}{\gamma}\right)^{-1}, \quad \gamma \in (0,1) \tag{A.3}$$

and if the least singular value

$$\tau_{\min}(\boldsymbol{\Phi}_S) \geq 2^{-1/2}, \quad \#S = s \tag{A.4}$$

then X is the unique solution of BP (L1), except with probability 2γ.

First of all, (4.5) and (4.6) together imply (A.1) and (A.3) with $\alpha = 1/2$. Moreover, by Proposition A.1 (A.4) holds with probability greater than or equal to the right hand side of (A.2). Hence we need only to derive the claimed bound for the probability of the event E that X is the unique solution of BP. This follows from the estimate

$$\begin{aligned}
\mathbb{P}(E) &\geq \mathbb{P}(E \,\big|\, \|\boldsymbol{\Phi}_S^*\boldsymbol{\Phi}_S - \mathbf{I}_s\|_2 < 2^{-1})\mathbb{P}(\|\boldsymbol{\Phi}_S^*\boldsymbol{\Phi}_S - \mathbf{I}_s\|_2 < 2^{-1}) \\
&\geq (1 - 2\gamma)(1 - (2/s)^p) \\
&\geq 1 - 2\gamma - (2/s)^p.
\end{aligned}$$

Appendix B. Matched field processing.

Matched field processing (MFP) has been used extensively for source localization in underwater acoustics and is closely related to the matched filter in signal processing.

Let Y be the n-dimensional signals received by the array of n sensors. Let $G(\mathbf{r}, \mathbf{a})$ be the Green function of the time-invariant medium and let $\mathbf{G}$ be the Green vector

$$\mathbf{G}(\mathbf{r}) = [G(\mathbf{r}, \mathbf{a}_1), G(\mathbf{r}, \mathbf{a}_2), ..., G(\mathbf{r}, \mathbf{a}_n)]^t \tag{B.1}$$

where t denotes transpose. The conventional MFP uses the Bartlett processor with the ambiguity surface

$$B(\mathbf{r}) = \frac{\mathbf{G}^*(\mathbf{r})YY^*\mathbf{G}(\mathbf{r})}{\|\mathbf{G}(\mathbf{r})\|_2^2} \tag{B.2}$$

[35]. The Bartlett processor is motivated by the following optimization problem: Maximize the quantity

$$W^*YY^*W \tag{B.3}$$

subject to the constraint:

$$W^*W = 1.$$

The solution

$$W = Y/\|Y\|_2$$

18

is the weight vector for the matched filter. In the case of one point source of amplitude τ_1 located at $\mathbf{x}_1$,

$$Y = \tau_1 \mathbf{G}(\mathbf{r}_1)$$

hence

$$W = \frac{\tau_1 \mathbf{G}(\mathbf{r}_1)}{|\tau_1| \|\mathbf{G}(\mathbf{r}_1)\|_2}. \tag{B.4}$$

Extending (B.4) to an arbitrary field point $\mathbf{r}$ by substituting $\mathbf{r}$ for $\mathbf{r}_1$ we obtain the Bartlett processor from (B.3).

REFERENCES

[1] A.B. Baggeroer, W.A. Kuperman and P.N. Mikhalevsky, "An overview of matched field methods in ocean acoustics", *IEEE J. Oceanic Eng.* **18** (1993), 401-424.

[2] R. Baraniuk, P. Steeghs, Compressive radar imaging, *IEEE Radar Conference*, pp. 128-133, April, 2007.

[3] B. Borden, *Radar Imaging of Airborne Targets.* Institute of Physics Publishing, Bristol, 1999.

[4] M. Born and E. Wolf, *Principles of Optics*, 7-th edition, Cambridge University Press, 1999.

[5] A.M. Bruckstein, D.L. Donoho and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals," *SIAM Rev.* **51** (2009), 34-81.

[6] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences, Paris, Serie I.* **346** (2008) 589-592.

[7] E. Candès, J. Romberg and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory* **52** (2006), 489-509.

[8] E.J. Candès, J. Romberg and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.* **59** (2006), 120723.

[9] E. J. Candès and T. Tao, " Decoding by linear programming," *IEEE Trans. Inform. Theory* **51** (2005), 42034215.

[10] W.L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Appl. Phys. Lett.* **93** (2008) 121105.

[11] J. C. Curlander and R.N. McDonough, *Synthetic Aperture Radar: Systems and Signal Processing,* Wiley-Intersience, 1991.

[12] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing and reconstruction," *IEEE Trans. Inf. Theory* **55(5)** (2009), 2230-2249.

[13] P. Delsarte, J. M. Goethals, and J. J. Seidel, Bounds for systems of lines and Jacobi poynomials, *Philips Res. Repts.* **30:3** pp. 91105, 1975, issue in honour of C.J. Bouwkamp.

[14] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory* **52** (2006) 1289-1306.

[15] D.L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proc. Nat. Acad. Sci.* **100** (2003) 2197-2202.

[16] D.L. Donoho, M. Elad and V.N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inform. Theory* **52** (2006) 6-18.

[17] D.L. Donoho and X. Huo, "Uncertainty principle and ideal atomic decomposition, " *IEEE Trans. Inform. Theory* **47** (2001), 2845-2862.

[18] M.F. Duarte, M.A. Davenport, D. Takhar, J.N. Laska, Ting Sun, K.F. Kelly and R. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Sign. Proc. Mag.*, March, 2008, pp. 83-91.

[19] A.C. Fannjiang, "Compressive imaging of subwavelength structures," *SIAM J. Imag. Sci.* **2** (2009), 1277-1291.

[20] A.C. Fannjiang, "Compressive inverse scattering I. High-frequency SIMO/MISO and MIMO measurements," *Inverse Problems* **26** (2010), 035008.

[21] A.C. Fannjiang, "Compressive inverse scattering II. multi-shot SISO measurements with Born scatterers," *Inverse Problems* **26** (2010), 035009

[22] A.C. Fannjiang and P. Yan, "Multi-frequency imaging of multiple targets in Rician fading channels: stability and resolution," *Inverse Problems* **23** (2007) 1801-1819.

[23] A.C. Fannjiang, K. Solna and P. Yan, "Synthetic aperture imaging of multiple point targets in Rician fading media," *SIAM J. Imaging Sci.* **2** (2009), 344-366.

[24] R. Gribonval and M. Nielsen, "Sparse representation in unions of bases," *IEEE Trans. Inform. Theory* **49** (2003), 3320-3325.

[25] A. C. Gurbuz, J. H. McClellan, W. R. Scott, "Compressive sensing for subsurface imaging using ground penetrating radar", *Sign. Proc.* **89** (2009), 1959-1972.

[26] M. Herman and T. Strohmer, " High-resolution radar via compressed sensing," *IEEE Trans. Sign. Proc.* **57** (2009), 2275-2284.

[27] M. Herman and T. Strohmer, "General deviants: an analysis of perturbations in compressed sensing," to appear in IEEE J. Sel. Topics Signal Proc., Special Issue on Compressive Sensing.

[28] W. Hoeffding, "Probability inequalities for sums of bounded random variables", *J. Amer. Stat. Assoc.* **58** (1963) 1330.

[29] M. Lax, " Multiple scattering of waves," *Rev. Mod. Phys.* **23** (1951), 287-310.

[30] D. Needell, J. A. Tropp, and R. Vershynin, "Greedy signal recovery review," *Proc. 42nd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 2008.

[31] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit", *Found. Comput. Math.* **9** (2009), 317-334.

[32] G.E. Pfander, H. Rauhut and J. Tanner, " Identification of matrices having a sparse representation" *IEEE Trans. Signal Proc.* **56** (2008) 5376-5388.

[33] H. Rauhut, "Stability results for random sampling of sparse trigonometric polynomials," *IEEE Trans. Inform. Th.* **54** (2008), 5661-5670.

[34] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Comm. Pure Appl. Math.* **111** (2008) 1025-1045.

[35] A. Tolstoy, *Matched Field Processing in Underwater Acoustics*, World Scientific, Singapore, 1993.

[36] J.A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. Inform. Theory* **50** (2004), 2231-2242.

[37] J.A. Tropp, " Just relax: convex programming methods for identifying sparse signals in noise," *IEEE Trans. Inform. Theory* **52** (2006), 1030-1051. "Corrigendum" *IEEE Trans. Inform. Theory* (2008).

[38] J.A. Tropp, "On the conditioning of random subdictionaries," *Appl. Comput. Harmon. Anal.* **25** (2008), 124.

[39] H.C. van de Hulst, *Light Scattering by Small Particles.* Dover Publications, New York, 1981.

[40] L. Welch, Lower bounds on the maximum cross-correlation of signals, *IEEE Trans. on Information Theory,* **20** (1974), pp. 397399.

[41] J. Yang and Y. Zhang, "Alternating direction algorithms for ℓ_1-problems in compressive sensing", preprint, 2009.