On the Relationship Between Compressive Sensing and Random Sensor Arrays

Lawrence Carin
Department of Electrical & Computer Engineering
Duke University
Durham, NC
lcarin@ece.duke.edu

Abstract

Random sensor arrays are examined from a compressive sensing (CS) perspective. It is demonstrated that the natural random-array projections manifested by the media Green's function are consistent with the projection-type measurements associated with CS. This linkage allows the use of existing CS theory to quantify the performance of random arrays, of interest for array design. The analysis demonstrates that the CS theory is applicable to arrays in vacuum as well as in the presence of a surrounding media; further, the presence of a surrounding media with known properties may be used to improve array performance.

I. Introduction

Most existing sensor arrays are sampled uniformly with inter-element spacing less than or equal to $\lambda/2$, where λ is the system wavelength. This inter-element spacing enhances performance by reducing side lobes [1], at the cost of array resolution. Specifically, it is well known that the resolution with which an array may focus is dictated by the size of the array aperture [1]; if one has a budget on the number of array elements that may be used, the $\lambda/2$ spacing also implies an associated aperture size. This limitation has motivated the development of arrays with inter-element spacing greater than $\lambda/2$. Further, to mitigate the "grating lobes" that are manifested by such a sub-sampled array, it is desirable to constitute non-uniform inter-element spacing. This has motivated the development of randomly spaced elements [2], [3]. Non-uniform arrays have been constituted for similar reasons in interferometric sensing [4]. We therefore note that the main motivation for the use of random and non-uniform arrays has typically been the goal of achieving high-resolution sensing while reducing sensing costs (relative to constituting the same array aperture with uniformly spaced elements at $\lambda/2$).

While the use of non-uniform and random arrays constitutes an old problem, the analysis applied to date is unsatisfactory, from multiple standpoints. For the case of non-uniform arrays [1], there are limited general theoretical developments; each array is generally designed from "scratch" to achieve a particular design goal. The theory associated with random arrays is more developed, as a result of statistical analyses [1]. However, this theory is largely unsatisfying, in that it constitutes statistical properties of side lobes, as averaged across many randomly constituted arrays. It does not explicitly define relationships on the accuracy one may expect when estimating the sources responsible for the signal on the sensor array, as a function of the number of angle-dependent sources and as a function of the noise level.

More recently, the new field of compressive sensing (CS) has been developed [5]–[9]. This theory was constituted in a more-general setting than the aforementioned random arrays, but there are also clear relationships. In CS one is interested in measuring a signal $\boldsymbol{u} \in \Re^n$, and it is assumed that $\boldsymbol{u}$ is compressible in an orthonormal basis represented by the columns of $\boldsymbol{\Psi} \in \Re^{n \times n}$ (for simplicity we assume real signals, but CS theory is applicable to the complex data generally of interest to array processing). Specifically, for transform coefficients $\boldsymbol{x}$, we have $\boldsymbol{u} = \boldsymbol{\Psi}\boldsymbol{x}$; if $\boldsymbol{x}_s$ represents $\boldsymbol{x}$ with the smallest $n - s$ components set to zero, then $\boldsymbol{x}$ is compressible in the sense that $\|\boldsymbol{x} - \boldsymbol{x}_s\|_{\ell_2}/\|\boldsymbol{x}\|_{\ell_2}$ is negligibly small for $s \ll n$. In CS, rather than measuring $\boldsymbol{u}$ directly, one performs a set of measurements $\boldsymbol{y} \in \Re^m$ with $\boldsymbol{y} = \boldsymbol{\Sigma}\boldsymbol{u}$, where $\boldsymbol{\Sigma} \in \Re^{m \times n}$, with $m < n$. We therefore have $\boldsymbol{y} = \boldsymbol{\Sigma}\boldsymbol{\Psi}\boldsymbol{x}_s + \boldsymbol{\Sigma}\boldsymbol{\Psi}(\boldsymbol{x} - \boldsymbol{x}_s) = \boldsymbol{\Phi}\boldsymbol{x}_s + \boldsymbol{z}$, with $\boldsymbol{\Phi} = \boldsymbol{\Sigma}\boldsymbol{\Psi}$ and $\boldsymbol{z}$ representing "noise" manifested by discarding the small transform coefficients. Compressive sensing theory [10] has demonstrated that there are explicit designs for $\boldsymbol{\Sigma}$ and hence $\boldsymbol{\Phi}$ by which one may recover $\boldsymbol{x}_s$ accurately, using a relatively simple ℓ_1-based inversion algorithm. The inversion problem for $\boldsymbol{x}_s$ based on measured $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}_s + \boldsymbol{z}$ constitutes a well-known linear-regression problem under the constraint that $\boldsymbol{x}_s$ must be sparse [11].

We note that the motivation for CS is related to that associated with random arrays. Specifically, it is known that $\hat{\boldsymbol{u}} = \boldsymbol{\Psi}\boldsymbol{x}_s$ is a good approximation to $\boldsymbol{u}$, and since $\boldsymbol{x}_s$ is sparse, it is hoped that the number of *projection* measurements m that may be performed satisfies $m \ll n$, where again n constitutes the number of samples in $\boldsymbol{u}$ that one may measure conventionally (hence, n defines the resolution with which $\boldsymbol{u}$ is represented). Therefore, both random arrays and CS are manifested by the goal of realizing high-resolution data via a relatively small number of measurements. The projection measurements in CS correspond to the rows of $\mathbf{\Sigma}$, and here we demonstrate how such projections are manifested in array-based measurements.

The principal contribution of CS concerns explicit theorems for the design of $\mathbf{\Phi}$ that assure that sparseness-constrained inversion algorithms of the form discussed above will perform reliably (even *perfectly* [10] under specific circumstances). This design procedure provides the important linkage to random sensor arrays. In particular, it has been demonstrated [5]–[9] that $\mathbf{\Sigma}$ and hence $\mathbf{\Phi}$ may be designed *randomly*, with specific constructions. There are several different random designs one may consider, with these closely linked to the embeddings associated with the Johnson-Lindenstrauss Lemma [12]. A contribution of this paper is to demonstrate that one of these designs is consistent with the type of projection measurements performed implicitly by sensor arrays with appropriate randomly designed inter-element spacing. The random nature of CS measurements provides the explicit link to random sensor arrays. Further, we also demonstrate how the sparseness associated with CS plays an important role in the performance of random arrays (to our knowledge, this linkage to sparseness has not been recognized previously within the array-processing literature).

We also make the connection between existing array processing algorithms, such as CLEAN [4], which were developed decades ago for random arrays, and new algorithms such as OMP and STOMP [13], [14] which have been applied and developed much more recently for CS. We demonstrate that these algorithms, as well as RELAX from array processing [15], while developed independently, are highly related to one another (in fact, OMP and CLEAN are essentially the same algorithm).

This paper makes the explicit connection between decades-old random sensor arrays and the much newer CS, demonstrating that the former is a special case of the latter. It is therefore not surprising that the aforementioned independently developed algorithms are highly inter-related. Further, using CS theory, we are able to make explicit statements about the performance of random arrays for sensing multiple angle-dependent sources. In particular, the accuracy of algorithms of the type discussed above is quantified as a function of the number of array elements, number of sources, and as a function of the additive noise (without requiring explicit statements about the noise statistics).

The remainder of the paper is organized as follows. In Section II we review CS theory of relevance for random sensor arrays. It is demonstrated in Section III that measurements of the type required for CS are implemented naturally in random arrays via the medium Green's function; this is true for general array constructions and general linear, isotropic media. Algorithms used for random arrays and for CS are summarized in Section IV, where it is demonstrated that the different research communities have developed highly related algorithms. Having made the connection between CS and random sensor arrays, in Section V we demonstrate how the former may be used to provide explicit quantitative statements about the performance of the latter, of importance for random-array design. Conclusions and directions for future research are provided in Section VI.

II. Relevant Compressive-Sensing Theory

A. Restricted isometry property and noisy-signal recovery

A brief summary of compressive sensing (CS) is provided based on the theory presented in [10], with a focus on the application of interest here. The discussion assumes that the signals of interest are real, although in the array-processing application considered in Section III the data are complex. Compressive sensing was first considered for a special class of complex measurements [5]–[7], and all of the theory presented below may be extended to complex signals, with added complexity that is unnecessary for current purposes; all results for real signals are retained, with small modifications to the final constants associated with the results.

Consider measured data $\boldsymbol{y} \in \Re^m$ that may be expressed in the form

$$\boldsymbol{y} = \boldsymbol{\Phi} \boldsymbol{x} + \boldsymbol{z} \tag{1}$$

where $\boldsymbol{x} \in \Re^n$ with $n > m$, and $\boldsymbol{z} \in \Re^m$ represents additive noise. Our objective is to recover $\boldsymbol{x}$ from $\boldsymbol{y}$, which is an ill-posed problem without further restrictions on $\boldsymbol{x}$. We assume that $\boldsymbol{x}$ is sparse, which means that only a small set of its components are nonzero. We are interested in understanding how large m must be to assure reliable recovery of $\boldsymbol{x}$, as well as the form $\mathbf{\Phi} \in \Re^{m\times n}$ must obey.

For each integer $s = 1, 2, ...$ Candès [10] defines the isometry constant δ_s of matrix $\mathbf{\Phi}$ as the smallest number such that

$$(1-\delta_s)\|\boldsymbol{x}\|_{\ell_2}^2 \leq \|\mathbf{\Phi}\boldsymbol{x}\|_{\ell_2}^2 \leq (1+\delta_s)\|\boldsymbol{x}\|_{\ell_2}^2 \tag{2}$$

holds for all s-sparse vectors $\boldsymbol{x}$. For small δ_s this is called a *Restricted Isometry Property* (RIP) because the near isometry is restricted to s-sparse signals. Consider the following solution to (1):

$$\min_{\tilde{\boldsymbol{x}}\in\Re^n} \|\tilde{\boldsymbol{x}}\|_{\ell_1} \quad \text{subject to} \quad \|\boldsymbol{y} - \Phi\tilde{\boldsymbol{x}}\|_{\ell_2} \leq \epsilon \tag{3}$$

where ϵ is an upper bound on the energy in $\boldsymbol{z}$. Candès proves the following [10]: Assume that $\delta_{2s} < \sqrt{2} - 1$ and $\|\boldsymbol{z}\|_{\ell_2} \leq \epsilon$, then the solution $\boldsymbol{x}^*$ to (3) obeys

$$\|\boldsymbol{x}^* - \boldsymbol{x}\|_{\ell_2} \leq C_0 s^{-1/2}\|\boldsymbol{x} - \boldsymbol{x}_s\|_{\ell_1} + C_1\epsilon \tag{4}$$

for explicit and small constants C_0 and C_1, where $\boldsymbol{x}_s$ is the same as $\boldsymbol{x}$ with the $n - s$ smallest components set to zero. Note that if $\boldsymbol{x}$ is s-sparse (the number of non-zero components in $\boldsymbol{x}$ is less than or equal to s), then $\|\boldsymbol{x} - \boldsymbol{x}_s\|_{\ell_1} = 0$, and (4) implies that the CS reconstruction error is then proportional to the energy in the "noise" $\boldsymbol{z}$.

B. Projection design

To design $\mathbf{\Phi} \in \Re^{m\times n}$ for measurement of a sparse signal $\boldsymbol{x} \in \Re^n$, consider a matrix $\mathbf{U} \in \Re^{n\times n}$, defined by orthonormal rows. One way to design $\mathbf{\Phi}$ is to select m rows of $\mathbf{U}$ uniformly at random, and then normalize the associated columns to have unit norm. With overwhelming probability, a $\mathbf{\Phi}$ matrix so designed yields [16] $\delta_{2s} < \sqrt{2} - 1$ if the number of projections m satisfies

$$m \geq C_3 \cdot s \cdot \mu^2 \cdot (log\ n)^4 \tag{5}$$

where $\mu = \sqrt{n} \cdot \max_{i,j} |\mathbf{U}_{i,j}|$, and there is an explicit form for the (small) constant C_3.

Reconsidering normalization of the columns of Φ, if we assume that the normalization constants are approximately equal for each of the columns, then without normalization the RIP is satisfied to within this multiplicative constant, which does not impact the solution of (3). We therefore henceforth ignore column normalization, but note that without normalization the constants on the right side of (4) are now multiplied by the normalization constant $\sqrt{n/m}$.

The signal of interest $\boldsymbol{u} \in \Re^n$ may not be sparse, although its transform coefficients may be sparse (or compressible) in an appropriate orthonormal basis; the vector $\boldsymbol{x}$ is now defined by the transform coefficients. Specifically, let $\boldsymbol{\Psi} \in \Re^{n \times n}$ have columns that represent an orthonormal basis, and $\boldsymbol{u} = \boldsymbol{\Psi}\boldsymbol{x} + \boldsymbol{\nu}$, where $\boldsymbol{x}$ is s-sparse and $\boldsymbol{\nu}$ is additive "noise" defined by the error manifested by setting the $n-s$ transform coefficients exactly to zero. We again consider a matrix $\boldsymbol{\Sigma}_o \in \Re^{n \times n}$ with orthonormal rows. We note that the matrix product $\mathbf{U} = \boldsymbol{\Sigma}_o \Psi$ corresponds to projecting the rows of $\boldsymbol{\Sigma}_o$ onto the column space of $\boldsymbol{\Psi}$, and therefore the rows of $\mathbf{U}$ are also orthonormal. For fixed $\boldsymbol{\Psi}$, random selection of rows of $\mathbf{U}$ corresponds to random selection of rows of $\boldsymbol{\Sigma}_o$. The projection measurement $\boldsymbol{y}$ is defined by randomly selecting m rows of $\boldsymbol{\Sigma}_o$, with the m selected rows defining the projection matrix $\boldsymbol{\Sigma} \in \Re^{m \times n}$, and

$$\boldsymbol{y} = \boldsymbol{\Sigma}\boldsymbol{u} = \boldsymbol{\Sigma}\boldsymbol{\Psi}\boldsymbol{x} + \boldsymbol{\Sigma}\boldsymbol{\nu} = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{z} \tag{6}$$

which yields the expression in (1). The required number of CS measurements of this type for s-sparse $\boldsymbol{x}$ is defined by (5), and therefore we desire small mutual coherence μ. Recognizing that now $\mathbf{U} = \boldsymbol{\Sigma}_o\boldsymbol{\Psi}$, μ is minimized by selecting $\boldsymbol{\Sigma}_o$ such that the mutual coherence between the row space of $\boldsymbol{\Sigma}_o$ and column space of $\boldsymbol{\Psi}$ is as small as possible. This implies that we desire the rows of $\boldsymbol{\Sigma}_o$ to be "spread out" in the column space of $\boldsymbol{\Psi}$. The number of required CS measurements m is also proportional to the number of significant components in $\boldsymbol{x}$, and therefore it is also desirable to choose $\boldsymbol{\Psi}$ such that $\boldsymbol{x}$ is as sparse as possible.

C. Summary

If a signal $\boldsymbol{u} \in \Re^n$ is compressible in the orthonormal basis $\boldsymbol{\Psi}$, and therefore $\boldsymbol{u} = \boldsymbol{\Psi x} + \boldsymbol{\nu}$ for sparse $\boldsymbol{x}$ and small error $\boldsymbol{\nu}$, one may perform m projection measurements (constituting $\boldsymbol{y} \in \Re^m$) of the form $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{z}$, where $\boldsymbol{\Phi} \in \Re^{m \times n}$ is constructed as $\boldsymbol{\Sigma \Psi}$, with the rows of $\boldsymbol{\Sigma}$ defined by m randomly selected orthonormal vectors. If the mutual coherence between the rows of $\boldsymbol{\Sigma}$ and columns of $\boldsymbol{\Psi}$ is small, then one may recover $\boldsymbol{x}$ accurately based on $m < n$ measurements $\boldsymbol{y}$, with error proportional to the energy in $\boldsymbol{z}$. It is also desirable to choose $\boldsymbol{\Psi}$ such that $\boldsymbol{x}$ is as sparse as possible, to minimize the required m.

Before proceeding, we note that there are many other constructions one may use for $\boldsymbol{\Phi}$ [16], but as discussed below the design of $\boldsymbol{\Phi}$ considered above is of most interest to array signal processing. Similar types of CS projections have also been used successfully in MRI applications [17].

III. Relationship to Random Arrays

A. Array measurements as nearly-orthogonal projections

Consider a current $\boldsymbol{J}(r = R, \theta = \pi/2, \phi)$ for large R and $\phi \in [0, 2\pi]$; this is a ring of current at radius $r = R$, and we assume these currents are responsible for the angle-dependent fields observed on a sensor array situated in the $\theta = \pi/2$ plane. There are typically out-of-plane sources, but when performing imaging with an array assumed to reside within a plane, all sources are imaged into the array plane. For the analysis that follows a finite sensor system ($e.g.$, array) is assumed located in the vicinity of the coordinate origin $r = 0$ (with array dimensions infinitesimal relative to R). To simplify notation, the far-zone source current $\boldsymbol{J}(r = R, \theta = \pi/2, \phi)$ is henceforth represented as $\boldsymbol{J}(\phi)$, to emphasize that it is only a function of the angle ϕ. When performing array processing our objective is to infer $\boldsymbol{J}(\phi)$, this representing the angle-dependent sources responsible for the measured radiation.

The environment in which the sensor exists is arbitrary, and the characteristics of the sensor array are general, as long as the media is isotropic and linear. For simplicity we assume the antennas are point (isotropic) radiators and receivers, but the theory may be generalized to consider the patterns of real antennas. The electric field due to $\boldsymbol{J}(\phi)$ as observed at the ith receiver, positioned at $\boldsymbol{r}_i$, may be expressed as

$$\boldsymbol{E}(\boldsymbol{r}_i) = \int_0^{2\pi} d\phi \boldsymbol{G}(\boldsymbol{r}_i; r = R, \theta = \pi/2, \phi) \cdot \boldsymbol{J}(\phi) \tag{7}$$

Therefore, the field $\boldsymbol{E}(\boldsymbol{r}_i)$ is a linear combination (projection) of the source current $\boldsymbol{J}(\phi)$ with the dyadic Green's function $\boldsymbol{G}(\boldsymbol{r}_i; r = R, \theta = \pi/2, \phi)$. Recognizing that the source current we wish to infer always exists at $r = R$ and $\theta = \pi/2$, we also simplify the representation of the Green's function, with $\boldsymbol{G}(\boldsymbol{r}_i; r = R, \theta = \pi/2, \phi)$ henceforth represented as $\boldsymbol{G}(\boldsymbol{r}_i; \phi)$. Concerning notation, the Green's function $\boldsymbol{G}(\boldsymbol{r}; \phi)$ is in general a dyadic, of the form $\boldsymbol{G}(\boldsymbol{r}; \phi) = \sum_{i=1}^{3} \sum_{j=1}^{3} \hat{\boldsymbol{a}}_i \hat{\boldsymbol{a}}_j g_{ij}(\boldsymbol{r}; \phi)$, where $\hat{\boldsymbol{a}}_i$ is a unit vector in the ith orthogonal direction, $g_{ij}(\boldsymbol{r}; \phi)$ is a scalar function, and $\boldsymbol{J}$ is a vector; hence $\boldsymbol{G}(\boldsymbol{r}_i; \phi) \cdot \boldsymbol{J}(\phi)$ is a dyadic-vector dot product, which yields a vector ($\boldsymbol{E}(\boldsymbol{r}_i)$ is of course a vector).

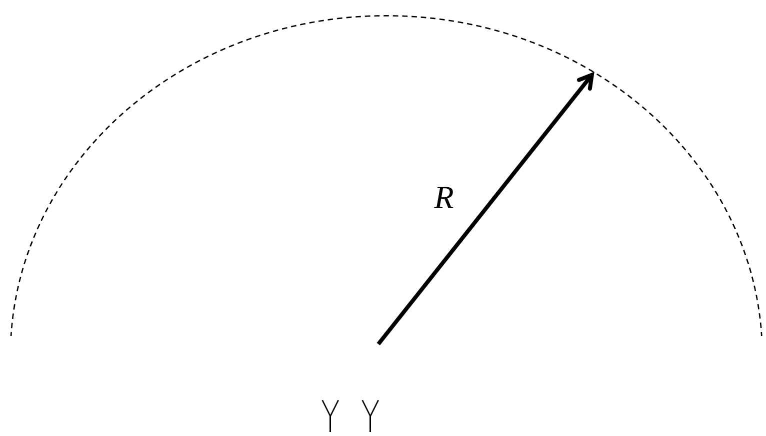

Fig. 1. Geometry of sensor array.

The observed electric fields at $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$ (two antennas on the array) are manifested by projecting $\boldsymbol{J}(\phi)$ onto $\boldsymbol{G}(\boldsymbol{r}_i; \phi)$ and $\boldsymbol{G}(\boldsymbol{r}_j; \phi)$, respectively. It is therefore of interest to examine the relationship between these two projections. For this, we recall the reciprocity theorem [18]. Assume that $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ represent two general currents, responsible for generating electric field $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$, respectively. The reciprocity theorem states

$$\int dr^3 \boldsymbol{E}_1 \cdot \boldsymbol{J}_2 = \int dr^3 \boldsymbol{E}_2 \cdot \boldsymbol{J}_1 \tag{8}$$

Let $\boldsymbol{J}_1$ be generated by an $\hat{\boldsymbol{a}}_i$-directed point source situated at $\boldsymbol{r}_i$, and let $\boldsymbol{J}_2$ be generated by $\boldsymbol{J}_{PC}(\phi)$ for $\phi \in [0, 2\pi]$, where $\boldsymbol{J}_{PC}(\phi)$ represents the phase-conjugated electric fields at $r = R$ and $\theta = \pi/2$ for $0 \leq \phi \leq 2\pi$, due to a $\hat{\boldsymbol{a}}_j$-directed point source at $\boldsymbol{r}_j$; stated explicitly, $\boldsymbol{J}_1 = \hat{\boldsymbol{a}}_i \delta(\boldsymbol{r} - \boldsymbol{r}_i)$ and $\boldsymbol{J}_2 = \boldsymbol{G}^*(\boldsymbol{r}_j; \phi) \cdot \hat{\boldsymbol{a}}_j$. Using the reciprocity relationship this yields

$$\boldsymbol{E}_{PC}(\boldsymbol{r}_{obs} = \boldsymbol{r}_i; \boldsymbol{r}_{source} = \boldsymbol{r}_j) \cdot \hat{\boldsymbol{a}}_i = \int_0^{2\pi} d\phi [\boldsymbol{G}(\boldsymbol{r}_i; \phi) \cdot \hat{\boldsymbol{a}}_i] \cdot [\boldsymbol{G}^*(\boldsymbol{r}_j; \phi) \cdot \hat{\boldsymbol{a}}_j] \tag{9}$$

where $\boldsymbol{E}_{PC}(\boldsymbol{r}_{obs} = \boldsymbol{r}_i; \boldsymbol{r}_{source} = \boldsymbol{r}_j)$ is the radiated field observed at $\boldsymbol{r}_i$ due to $\boldsymbol{J}_2 = \boldsymbol{J}_{PC}(\phi) = \boldsymbol{G}^*(\boldsymbol{r}_j; \phi)$. Because of the fact that the current $\boldsymbol{J}_{PC}(\phi)$ resides for all $\phi \in [0, 2\pi]$, and using insights from phase conjugation or time-reversal [19], the $\boldsymbol{E}_{PC}(\boldsymbol{r}_{obs} = \boldsymbol{r}_i; \boldsymbol{r}_{source} = \boldsymbol{r}_j)$ will be strongly peaked at $\boldsymbol{r}_i = \boldsymbol{r}_j$, and will become very small for $\|\boldsymbol{r}_i - \boldsymbol{r}_j\|_{\ell_2} > \lambda/2$, where λ is the wavelength. This demonstrates that if $\|\boldsymbol{r}_i - \boldsymbol{r}_j\|_{\ell_2} \geq \lambda/2$ the projections manifested by $\boldsymbol{G}(\boldsymbol{r}_i; \phi) \cdot \hat{\boldsymbol{a}}_i$ and $\boldsymbol{G}(\boldsymbol{r}_j; \phi) \cdot \hat{\boldsymbol{a}}_j$ are approximately orthogonal:

$$\int_0^{2\pi} d\phi [\boldsymbol{G}(\boldsymbol{r}_i; \phi) \cdot \hat{\boldsymbol{a}}_i] \cdot [\boldsymbol{G}^*(\boldsymbol{r}_j; \phi) \cdot \hat{\boldsymbol{a}}_j] \approx 0 \quad \text{for} \quad \|\boldsymbol{r}_i - \boldsymbol{r}_j\|_{\ell_2} \geq \lambda/2 \tag{10}$$

The $\lambda/2$ resolution is dictated by the fact that, for the large R of interest, the current $\boldsymbol{J}_{PC}(\phi)$ loses evanescent-field information about the point source at $\boldsymbol{r}_j$ [18].

We may therefore generally view, for any antenna positions, the observed fields as projections of far-field sources at $\phi \in [0, 2\pi]$ onto the corresponding Green's function, and the projections are orthogonal over $\phi \in [0, 2\pi]$ as long as the antennas are separated by a half wavelength or more. This is important, as the compressive-sensing theory summarized in Section II dictates that orthogonal projections are desirable, and they are here naturally manifested by the wave physics. Note that the CS theory desires *orthonormal* projections, not simply orthogonal. Since the source current $\boldsymbol{J}(\phi)$ is assumed to reside in the far zone of the sensor array, the Green's-function projections all approximately have the same amplitude, and hence the projections are orthonormal to within a constant – this constant does not impact the form of (3). Nevertheless, the Green's functions may be normalized to have unit energy if desired, this simply manifesting a scaling of the source current $\boldsymbol{J}(\phi)$.

B. Random projections and compressive sensing

Consider an array of m isotropic antennas in the plane $\theta = \pi/2$, with the antennas situated near the origin $r = 0$. While it is not required, we assume that all antennas are polarized in the same way: $\hat{\boldsymbol{a}}_i = \hat{\boldsymbol{a}}$ for all antennas i. Let e_i represent the (complex) signal measured on antenna i, and $\boldsymbol{e} = [e_1, e_2, ..., e_m]^T$ represents the data observed on the m antennas. Further, let $\boldsymbol{j} = [J(\phi = 0), J(\phi = \Delta), J(\phi = 2\Delta), ..., J(\phi = 2\pi)]^T$ represent the (unknown) radiating currents as a function of angle (with current direction consistent with the polarization of the antennas). In this representation the current is discretized at an appropriate (fine) angular rate Δ, yielding an n-dimensional vector $\boldsymbol{j}$. The current represented by $\boldsymbol{j}$ is also assumed to reside in the plane $\theta = \pi/2$ of the sensor array.

In matrix form we have

$$\boldsymbol{e} = \boldsymbol{\Sigma}\boldsymbol{j} \tag{11}$$

where $\boldsymbol{\Sigma}$ is an $m \times n$ matrix (with $m < n$, and ideally $m \ll n$). The ith row of $\boldsymbol{\Sigma}$ represents $\boldsymbol{G}(\boldsymbol{r}_i; \phi) \cdot \hat{\boldsymbol{a}}$ discretized with respect to ϕ; if all antennas are separated by at least $\lambda/2$, the rows of $\boldsymbol{\Sigma}$ are approximately orthogonal. The vector $\boldsymbol{j}$ is typically compressible in an appropriate orthonormal basis, which we represent by the $n \times n$ matrix $\boldsymbol{\Psi}$; we therefore have $\boldsymbol{j} = \boldsymbol{\Psi}\boldsymbol{x} + \boldsymbol{\nu}$, with $\boldsymbol{x}$ assumed to be sparse and $\boldsymbol{\nu}$ absorbing the error manifested by setting the small transform coefficients exactly to zero. As an example, if the source $\boldsymbol{j}$ is characterized by a relatively small number of radiators distributed over $\phi \in [0, 2\pi]$, with additive noise, then we may let $\boldsymbol{\Psi}$ be the identity matrix. Summarizing, we have

$$\boldsymbol{e} = \boldsymbol{\Sigma}\boldsymbol{\Psi}\boldsymbol{x} + \boldsymbol{\Sigma}\nu = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{z} \tag{12}$$

The key observation is that the randomly situated array elements effectively constitute projections onto the desired ϕ-dependent sources responsible for the radiation, and these projections may be designed to be approximately orthogonal if $\|\boldsymbol{r}_i - \boldsymbol{r}_j\|_{\ell_2} \geq \lambda/2 \;\; \forall\, i$ and j. The most compact sensor array with which the m near-orthogonal projections may be manifested corresponds to a uniform array with $\lambda/2$ spacing. However, uniform sampling is not required, or even desired. The enhanced aperture size manifested by wider than $\lambda/2$ separation provides value concerning resolution, as discussed further below. Additionally, the random spacing mitigates grating lobes [1], which we discuss below from a CS perspective. Finally, we note that the choice of $\mathbf{\Psi} = \mathbf{I}^{n\times n}$ is advantageous from the standpoint of the mutual coherence μ discussed in (5), as the Green's function (rows of $\mathbf{\Sigma}$) are nearly uniformly weighted across angle (for the case of an array in vacuum).

Random arrays have been used previously in the context of more-conventional array-processing techniques [1], [3]. The distinction provided by CS is that the inversion for $\boldsymbol{J}(\phi)$ takes advantage of additional information not exploited by conventional techniques. Specifically, CS exploits knowledge that the unknown current is compressible in an appropriate transform relative to angle ϕ (possibly after typical averaging of array measurements to reduce sensor noise).

Before proceeding, it is important to emphasize that the orthogonality of the Green's-function-produced projections is important, but it is *not* in itself sufficient for CS. For example, the rows of the identity matrix $\mathbf{I}^{n\times n}$ are also orthonormal, but these are often a poor choice for CS projections. Hence, to examine the appropriateness of projections, one must pay careful attention to the mutual coherence μ in (5). To minimize μ, and hence the number of projection measurements m, the orthonormal projection vectors that constitute the rows of $\mathbf{\Sigma}$ should be "spread out" when expanded in the column space of $\mathbf{\Psi}$. If we consider a current $\boldsymbol{J}(\phi) = \sum_{k=1}^{s} w_k \delta(\phi - \phi_k^*) + \boldsymbol{\nu}(\phi)$, where $\boldsymbol{\nu}(\phi)$ is arbitrary noise with energy less than ϵ, and ϕ_k^* represent possible source angles selected at random, with random weight w_k and $|w_k|^2$ large relative to the average (across angle) noise energy, then in this case the appropriate orthonormal basis in which to represent $\boldsymbol{j}$ is $\mathbf{\Psi} = \mathbf{I}^{n\times n}$ (the noise-free current $\sum_{k=1}^{s} w_k \delta(\phi - \phi_k^*)$ is sparse in this basis, and for sufficiently high SNR $\boldsymbol{J}(\phi) = \sum_{k=1}^{s} w_k \delta(\phi - \phi_k^*) + \boldsymbol{\nu}(\phi)$ is compressible in $\mathbf{\Psi} = \mathbf{I}^{n\times n}$). For $\mathbf{\Psi} = \mathbf{I}^{n\times n}$, mutual coherence is minimized by projections (rows of $\mathbf{\Sigma}$) that have energy uniformly distributed across angle ϕ.

We now consider the special but important case of an antenna array in vacuum (free-space). For large observation distances R and time dependence $exp(j\omega t)$, the free-space Green's function is of the form $exp(-j\beta R)exp(-j\beta\rho cos(\phi))/R$, where ρ is the small (relative to R) distance from the origin to the corresponding antenna element on the array, $\beta = 2\pi/\lambda$, and $j = \sqrt{-1}$. Such a projection has *constant* amplitude with respect to ϕ, and therefore is maximally spread out in the column space of $\mathbf{\Psi} = \mathbf{I}^{n\times n}$. Therefore, Green's-function-constituted projections for random arrays in vacuum are *optimal* for (typical) sources of the type mentioned above, for which $\mathbf{\Psi} = \mathbf{I}^{n\times n}$. More generally, if the sources are more diffuse with respect to ϕ, and therefore not sparse in $\mathbf{\Psi} = \mathbf{I}^{n\times n}$, a basis such as wavelets is appropriate for $\mathbf{\Psi}$, since such localized basis functions will still provide a relatively "spread out" representation of the Green's function, which has constant amplitude as a function of angle ϕ.

C. CS perspective on aperture size and angular resolution

Let $\hat{\boldsymbol{j}}$ represent the CS-computed approximation to $\boldsymbol{j}$, based on the m measurements $\boldsymbol{e}$. Compressive-sensing theory places bounds on the accuracy of the error $\|\boldsymbol{j} - \hat{\boldsymbol{j}}\|_{\ell_2}$. However, based on the above discussion this error is computed over all angles $\phi \in [0, 2\pi]$, and therefore it does not address the accuracy of the reconstruction over narrower angular support ($i.e.$, resolution). For that, we now assume that $\boldsymbol{J}(\phi)$ is only non-zero over a contiguous extent of angles S_ϕ, and S_ϕ may be made arbitrarily small; we desire CS reconstruction guarantees for currents that reside over the support of S_ϕ. The angular extent (in radians) of S_ϕ is denoted $|S_\phi|$.

Under these circumstances, the field observed at antenna $\boldsymbol{r}_i$ may be expressed as

$$e_i = \int_{\phi \in S_\phi} d\phi \ \hat{\boldsymbol{a}}_i \cdot \boldsymbol{G}(\boldsymbol{r}_i; \phi) \cdot \boldsymbol{J}(\phi) \tag{13}$$

As above, we wish to examine relationships between $\boldsymbol{G}(\boldsymbol{r}_i; \phi)$ for different $\boldsymbol{r}_i$, now under the restriction that we are only interested in $\phi \in S_\phi$ (see Figure 2). We again consider reciprocity, and now $\boldsymbol{J}_{PC}(\phi) = \boldsymbol{G}^*(\boldsymbol{r}_j; \phi) \cdot \hat{\boldsymbol{a}}_j$ for $\phi \in S_\phi$ and $\boldsymbol{J}_{PC}(\phi) = 0$ for $\phi \notin S_\phi$. This yields

$$\boldsymbol{E}_{PC}(\boldsymbol{r}_{obs} = \boldsymbol{r}_i; \boldsymbol{r}_{source} = \boldsymbol{r}_j, S_\phi) \cdot \hat{\boldsymbol{a}}_i = \int_{\phi \in S_\phi} d\phi [\boldsymbol{G}(\boldsymbol{r}_i; \phi) \cdot \hat{\boldsymbol{a}}_i] \cdot [\boldsymbol{G}^*(\boldsymbol{r}_j; \phi) \cdot \hat{\boldsymbol{a}}_j] \tag{14}$$

where $\boldsymbol{E}_{PC}(\boldsymbol{r}_{obs} = \boldsymbol{r}_i; \boldsymbol{r}_{source} = \boldsymbol{r}_j, S_\phi)$ represents the fields due to a phase-conjugated source that only exists over $\phi \in S_\phi$. From aperture theory [1], for small $|S_\phi|$ and for an array in vacuum, to achieve $\boldsymbol{E}_{PC}(\boldsymbol{r}_{obs} = \boldsymbol{r}_i; \boldsymbol{r}_{source} = \boldsymbol{r}_j, S_\phi) \approx 0$ we require $\|\boldsymbol{r}_i - \boldsymbol{r}_j\|_{\ell_2} \geq \lambda/|S_\phi|$, which demonstrates that to achieve near-orthogonal projections over $\phi \in S_\phi$, the inter-element spacing must be increased inversely proportional to the support of S_ϕ. Hence, to achieve m near-orthogonal CS projections over small angular support S_ϕ, which will allow CS reconstruction-error guarantees over this support (addressing resolution), the size of the antenna aperture must be increased to assure that $\|\boldsymbol{r}_i - \boldsymbol{r}_j\|_{\ell_2} \geq \lambda/|S_\phi|$. This constitutes the CS connection to the well-known relationship between the array aperture size and system resolution (this is viewed from another CS-related standpoint in the next subsection).

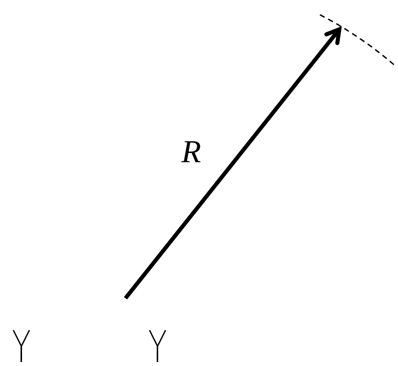

Fig. 2. Consideration of CS inversion accuracy over a small range of source angles, addressing the issue of resolution. To achieve orthogonal projections over a narrow support of angles, and hence CS reconstruction accuracy, the antennas must be separated further, implying a larger array. This maps the conventional issue of array resolution and aperture size onto a CS perspective.

The above discussion implies that the size of the aperture must be very large to achieve accurate CS results over narrow S_ϕ ($i.e.$, resolution). However, recall from the CS theory in Section II that the number of CS measurements m is related to the signal sparseness over the projection support. As $|S_\phi| \rightarrow 0$ the signal $\boldsymbol{J}(\phi)$ is expected to be very smooth and hence highly compressible. Therefore, the number of near-orthogonal CS projections required over that support is likely to be small. Consequently only a relatively small set of projections are anticipated for inter-element spacing $\lambda/|S_\phi|$. Hence, on a sensor array, only the relatively few widely separated elements (which yield near-orthogonal projections over narrow $|S_\phi|$) are expected to play an important role in the reconstruction accuracy over narrow $|S_\phi|$.

D. Restricted orthornomality and resolution

Consider a linear sensor array with *uniform* inter-element spacing Δ_x; the array is assumed to reside in vacuum. We consider the orthonormal basis $\mathbf{\Psi} = \mathbf{I}^{n\times n}$, and examine the properties of the columns of $\mathbf{\Phi}$. Each column of $\mathbf{\Phi}$ corresponds to the Green's function from a discrete source angle $\phi \in [0, 2\pi]$ to the m uniformly spaced antennas that constitute the array. Apart from a constant, the components that constitute the ith column of $\mathbf{\Phi}$ may be expressed as

$$[\mathbf{\Phi}_{1,i}, \mathbf{\Phi}_{2,i}, \cdots, \mathbf{\Phi}_{m,i}]^T = [1, \omega_i, \omega_i^2, \cdots, \omega_i^{m-1}]^T \tag{15}$$

where $\omega_i = exp[-j2\pi\frac{\Delta_x}{\lambda}cos(\phi_i)]$, and ϕ_i corresponds to the ith angular bin, with $j = \sqrt{-1}$. We observe that each column of $\mathbf{\Phi}$ corresponds to a sampled Fourier basis function, at angular frequency ω_i, truncated over m samples. By setting $\Delta_x = \lambda/2$ we achieve the desired near-orthogonal projections on the source currents (Section III), while also removing ambiguities in the angular frequencies; the latter manifest "grating lobes" in the array response [1].

Now consider any two columns of $\mathbf{\Phi}$, corresponding to angles ϕ_1 and ϕ_2; the degree to which these two columns are orthogonal is dictated by the difference $\|\phi_1 - \phi_2\|_{\ell_2}$, and by the size of the array, dictated by m. Each column of $\mathbf{\Phi}$ is associated with the contribution to the array response from one source angle, and it appears reasonable that the degree to which any two columns are orthogonal will impact the resolution with which one may infer the desired source current $\boldsymbol{J}(\phi)$. This implies the well-known relationship between angular resolution and aperture size (size of m), discussed in Section III-C. However, there is not a *precise* relationship from above concerning the near-orthogonality of the columns of $\mathbf{\Phi}$ and the accuracy with which $\boldsymbol{J}(\phi)$ may be recovered. Further, the above discussion was relegated to conventional uniformly sampled linear arrays in vacuum, with inter-element spacing $\Delta_x \leq \lambda/2$. Compressive sensing affords the opportunity to *explicitly* link the relationship between the near-orthogonality of the columns of $\mathbf{\Phi}$ to the accuracy with which $\boldsymbol{J}(\phi)$ may be measured; further, the theory is applicable to general random arrays situated in arbitrary environments (not necessarily vacuum). As is known from conventional array theory [1], the random array locations mitigate grating lobes; this is a product of the random projections proscribed by CS theory, although it was not an explicit *a priori* objective.

Consider a $\mathbf{\Phi}$ matrix, with the kth row defined by the ϕ-dependent Green's function for array element situated at $\boldsymbol{r}_k$ (we again assume $\mathbf{\Psi} = \mathbf{I}^{n \times n}$). The array may be situated in an arbitrary linear, isotropic environment, and the array elements are situated randomly, with the restriction that the rows are nearly orthogonal (separated by greater than or equal to $\lambda/2$, as discussed in Section III-A). It is important to emphasize that the *rows* of $\mathbf{\Phi}$ are orthogonal, and we wish to examine the properties of the columns; the columns of $\mathbf{\Phi}$ define the vectors in which the angle-dependent source $\boldsymbol{J}(\phi)$ is expanded to constitute the observed signal across the arrays.

Recall the Restricted Isometry Property (RIP) from Section II-A, and the associated parameter δ_s defined there. Consider two vectors $\boldsymbol{x} \in \Re^n$ and $\boldsymbol{x}' \in \Re^n$, and these vectors are supported on *disjoint* subsets T and T', with $T, T' \subseteq \{1, 2, \cdots, n\}$, with $|T| \leq s$ and $|T'| \leq s'$. The following relationship holds [10]:

$$|<\mathbf{\Phi}\boldsymbol{x}, \mathbf{\Phi}\boldsymbol{x}'>| \leq \delta_{s+s'}|\boldsymbol{x}||\boldsymbol{x}'| \tag{16}$$

Recall that if δ_{2s} is small there are guarantees concerning the accuracy of the reconstruction in (3), and (5) provides an explicit expression for the required number of array elements m. As discussed further in Section V, the integer s may be linked to the number of discrete sources responsible for the observed fields. From RIP and (16), when interested in accurately observing s sources (from s discrete angles), the matrix $\mathbf{\Phi}$ should be designed such that if one randomly selects up to $2s$ columns of $\mathbf{\Phi}$, these columns should be nearly orthonormal, in the sense defined by (2) and (16). Thus, CS makes explicit the relationship between the degree of orthonormality in the columns of $\mathbf{\Phi}$ (linked to the number of array elements m) and the resolution with which one may distinguish multiple angle-dependent discrete sources.

IV. Reconstruction Algorithms

In this section we demonstrate that many of the algorithms developed independently in the CS and array-processing communities are highly related to each other; this is not surprising given the inter-relationships between these fields, as discussed in Sections II and III. We discretize the unknown current $\boldsymbol{j}$ in (11) in terms of n uniform bins over 2π, where the angular support of the bins Δ is small relative to the array resolution. The current is represented as $\boldsymbol{j} = \boldsymbol{\Psi}\boldsymbol{x} + \boldsymbol{\nu}$, where $\boldsymbol{x}$ is s-sparse. In this example we assume that $\boldsymbol{\Psi}$ is the $n \times n$ identity matrix, with $\boldsymbol{\nu}$ representing stationary (relative to angle) sensor noise. The measured data is therefore $\boldsymbol{e} = \boldsymbol{\Sigma}\boldsymbol{x} + \boldsymbol{\Sigma}\boldsymbol{\nu} = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{z}$; in this case the non-zero components of $\boldsymbol{x}$ correspond to discrete sources situated at s angular bins distributed across 2π. It is assumed by construction that the energy in $\boldsymbol{\Phi}\boldsymbol{x}$ is large relative to that in $\boldsymbol{z}$.

If we ignore the noise $\boldsymbol{z}$ for now, we have $\boldsymbol{e} = \boldsymbol{\Phi}\boldsymbol{x}$. Since $\boldsymbol{x}$ is s-sparse, this implies that $\boldsymbol{e}$, which corresponds to the measured signals across the m sensors, may be expanded in terms of a small set of the columns of $\boldsymbol{\Phi}$. Each column corresponds to the observed signal across the m receivers due to a source at one of the angular bins; if the array is uniformly sampled, the columns of $\boldsymbol{\Phi}$ represent n Fourier bins [15], indexed by the source angle relative to the array (see (15)). However, as discussed in the previous subsection, this near-orthogonality of the columns of $\boldsymbol{\Phi}$ is applicable to appropriately designed random arrays in general media (those with greater than $\lambda/2$ inter-element spacing).

Methods have been developed in the array-processing community for inferring $\boldsymbol{x}$ based on measured $\boldsymbol{e} = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{z}$, using greedy or near-greedy constructions. The CLEAN algorithm was first developed for sparsely sampled antenna arrays in radio astronomy [4], and was later applied in other radar applications [20], where it was again employed for sparsely sampled antenna arrays. The RELAX algorithm [15] is closely related to CLEAN, with modifications employed to remove some of its greedy nature. Further, RELAX was developed based on the assumption of uniformly sampled arrays in vacuum; under this assumption the columns of $\boldsymbol{\Phi}$ correspond to discrete Fourier components (recall (15)), allowing use of the FFT to accelerate computations.

The CLEAN algorithm is highly related to Orthogonal Matching Pursuits (OMP) [21], [22], which has been applied recently in the CS literature [13]. The OMP algorithm has also motivated new CS reconstruction algorithms, such as STOMP [14]

These algorithms are all very similar, despite the fact that they have been developed independently in disparate but, as demonstrated above, highly related fields. They sequentially search the n columns of $\mathbf{\Phi}$, adding the new column that most minimizes the mean-square error $\|\mathbf{\Phi}_\Omega \boldsymbol{x}_\Omega - \boldsymbol{e}\|_{\ell_2}$, where Ω is the subset of columns of $\mathbf{\Phi}$ selected up to a particular point of the iterative solution, and $\boldsymbol{x}_\Omega$ represents the $|\Omega|$ components of $\boldsymbol{x}$ that correspond to the selected columns of $\mathbf{\Phi}$. At each iteration Ω is expanded by one vector, selecting one of the remaining columns of $\mathbf{\Phi}$ not employed thus far. The algorithm terminates when $\|\mathbf{\Phi}_\Omega \boldsymbol{x}_\Omega - \boldsymbol{e}\|_{\ell_2}$ stabilizes with expanding Ω, implying the algorithm is starting to reconstruct the noise $\boldsymbol{z}$.

V. CS PERSPECTIVE ON PERFORMANCE OF RANDOM ARRAYS

From the above discussion there are strong linkages between the motivations for random arrays and for compressive sensing, as well as in the state-of-the-art algorithms applied to CS and array processing. Concerning the latter, the algorithmic similarities are manifested because appropriately designed random arrays implicitly perform the class of projection measurements (across angle) that are required for CS. This connection may be used to infer fundamental relationships of random arrays, which CS theory now places on a firm mathematical footing.

Assume an angle-dependent source current $\boldsymbol{J}(\phi)$ at large radius R is responsible for the fields observed at an array in the $\theta = \pi/2$ plane and situated near the origin. The current is discretized into n angular bins, yielding the vector $\boldsymbol{j}$, with n sufficiently large such that the angular bins are small relative to the array resolution. The array has randomly constituted inter-element spacing, with separation greater than $\lambda/2$. Based on the analysis in Section III and existing CS theory, the following statements may be made:

- Assume that the vector $\boldsymbol{x}$ defines the source strength from each of the n angular bins, with $\boldsymbol{x}$ an arbitrary s-sparse signal (s discrete sources distributed arbitrarily across the n angular bins); in this case $\mathbf{\Psi} = \mathbf{I}^{n\times n}$. In the noise-free case, the ℓ_1-based inversion algorithm in (3), with $\epsilon = 0$, yields an *exact* reconstruction of $\boldsymbol{j}$ with overwhelming probability if

$$m \geq C_3 \cdot \mu^2 \cdot s \cdot (log\ n)^4 \tag{17}$$

for small constant C_3 defined in [16]. The constant μ is the maximum element of $\boldsymbol{\Sigma}$, the latter defined by $G(\boldsymbol{r}_i;\phi_k)$; this corresponds to the far-zone Green's function between array element $\boldsymbol{r}_i$ and discretized angular bin ϕ_k, with the normalization $\sum_{k=1}^{n}|G(\boldsymbol{r}_i;\phi_k)|^2 = 1$.

- Now considering the case for which there *is* additive noise, let $\boldsymbol{j} = \boldsymbol{x} + \boldsymbol{z}$, with $\|\boldsymbol{z}\|_{\ell_2} \leq \epsilon$; $\boldsymbol{x}$ again represents the strengths of the sources distributed across the n angular bins ($i.e.$, $\boldsymbol{\Psi} = \mathbf{I}^{n\times n}$). Let $\boldsymbol{\Sigma} \in \Re^{m\times n}$ represent the Green's function defined projection matrix, as above. If $m \geq C_3 \cdot s \cdot \mu^2 \cdot (log\ n)^4$ then $\delta_{2s} < \sqrt{2} - 1$ for this matrix, then the solution $\boldsymbol{x}^*$ to (3) satisfies (4) [10]. The mutual coherence μ is defined on $\boldsymbol{\Sigma}$, as above.

- More generally, $\boldsymbol{j} = \boldsymbol{\Psi x}$, where $\boldsymbol{\Psi} \in \Re^{n\times n}$ is defined by orthonormal columns in which $\boldsymbol{x}$ is compressible. In this case the bins of $\boldsymbol{x}$ do not represent angle-dependent sources, but rather represent the source current $\boldsymbol{j}$ in the basis $\boldsymbol{\Psi}$ (in which $\boldsymbol{x}$ is compressible). The above relationships still hold, but now the mutual coherence μ is defined with respect to the matrix $\boldsymbol{\Sigma\Psi}$, with $\boldsymbol{\Sigma}$ defined as above.

- The above relationships hold for reconstruction accuracy over the entire angular support $\phi \in [0, 2\pi]$. Similar relationships may be constituted for accuracy over narrower angular support ($i.e.$, resolution), using ideas discussed in Section III-C. In this case, to achieve the required orthogonal projections associated with the CS theory, the inter-element spacing must be larger than $\lambda/2$.

VI. Conclusions and Future Directions

The principal focus of this paper has been to explicitly make the connection between random sensor arrays and compressive sensing (CS). The former has been investigated for many decades, typically motivated by the idea of constituting sufficient angular resolution (large aperture) with a relatively small number of array elements. By contrast, the field of compressive sensing is only a few years old; it has also been motivated by the goal of reducing sensing costs, for general sensing missions. We have demonstrated that the types of measurements employed in arrays is consistent with the projection measurements associated with CS. In fact, random arrays may be viewed as a special case of CS measurements. The inversion algorithms widely employed in CS are based on regularized inversion, with the regularization manifested by minimizing the ℓ_1 norm of the transform coefficients for the signal of interest. While array signal processing algorithms have not been explicitly motivated by this goal, we have demonstrated that two widely used array-processing algorithms, CLEAN and RELAX, are approximate means of realizing the CS inversion. Further, OMP and STOMP, two recently developed algorithms applicable to CS inversion, are very closely linked to CLEAN and RELAX.

With this strong linkage of the motivations and algorithms associated with CS and sensor arrays, one may ask what new is provided by CS, and where does CS point concerning future research directions for array-processing applications. Concerning what is new, CS provides an explicit, quantitative theory for design of random sensor arrays. Specifically, it quantifies how many array elements m are required to reliably recover s angle-dependent sources embedded in noise. It provides guarantees on algorithm performance, which may be useful in designing random arrays for particular applications. Further, most previous random array theory was only applicable to structures in vacuum, while the CS-based theory has been demonstrated here to be appropriate for any linear, isotropic medium (the theory demonstrating the near-orthogonality of Green's-function based projections did not assume the array was in vacuum).

Concerning future research, there are several directions of interest. For example, one may consider design of new array concepts that *exploit* the properties of relatively complex propagation media placed in the presence of the sensor array. Specifically, to realize a large effective aperture while still maintaining a relatively small number of proximate antennas, one may place the array antennas in the presence of scatterers, as in Figure 3. The *effective* large aperture is manifested through scattering from the media placed in the presence of the antennas; incident waves at a diversity of incident angles impinge the scatterers, and ultimately make their way to the receiver antennas. Concerning these scatterers placed in the presence of the array antennas, it is desirable that they not break out-of-plane symmetry, to preserve polarization purity in the observed fields. This suggests placing all randomly placed antennas in the same plane, with the surrounding scatterers defined by spheres, with sphere equators residing in this same plane.

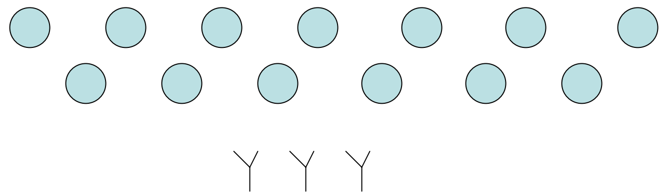

Fig. 3. Antennas in presence of scatterers.

The measured data are represented as $\boldsymbol{e} = \boldsymbol{\Phi}\boldsymbol{x}+\boldsymbol{\nu}$, where $\boldsymbol{x}$ represents the sparse set of significant transform coefficients of the angle-dependent far-zone source currents. We note that the same noise mitigation manifested by averaging multiple measurements may also be manifested through multiple measurements (averaging) of $\boldsymbol{e}$. A challenge for CS inversion involves a requirement for knowledge of $\boldsymbol{\Phi}$, which is sensitive to the exact placement of the antennas and of the surrounding scattering media (if a design like that in Figure 3 is employed). However, a given structure may be "calibrated" to infer $\boldsymbol{\Phi}$, by simply performing far-field measurements of $\boldsymbol{e}$, with a source antenna placed at large R within the aforementioned plane; the array response $\boldsymbol{e}$ is then measured as a function of the source angle ϕ. By performing this *one-time* calibration for a sufficient set of angles $\phi \in [0, 2\pi]$ (or desired subset of angles), one may directly measure $\boldsymbol{\Phi}$. Once this one-time measurement of $\boldsymbol{\Phi}$ is performed, "standard" CS inversion algorithms may be used to recover $\boldsymbol{x}$ for an arbitrary source $\boldsymbol{J}(\phi)$.

In addition to new sensor designs, there is interest in new algorithms for recovery of the source current $\boldsymbol{j}$. It has been demonstrated that the CLEAN [4], [20] and RELAX [15] algorithms applied in array processing are performing a greedy or near-greedy ℓ_1-type inversion, analogous to algorithms developed for CS [13], [14]. Compressive sensing is a very active field, motivating a large set of new algorithms that are generally superior to these early approaches. For example, algorithms have been developed to provide a full posterior density function on the sparse coefficients, using fast Bayesian techniques [23], [24]. This same class of algorithms may also be adapted for improved processing of data from random sensor arrays. In fact, almost all of the inversion algorithms developed in the CS community recently may be directly applied to random-array processing.

Another promising direction in compressive sensing involves development of adaptive projections

[23]–[25]. In this context one may constitute a sensor array in which the elements may be adaptively selected, which based on the discussion in Section III corresponds to adaptive projections. It has been demonstrated that this approach provides significant value, particularly in noise [25]. The success of this framework for general CS applications suggests its use for the specific application of sensor arrays.

We close by noting that some sensor arrays are synthetically constituted, for example for synthetic aperture radar (SAR) [26] and synthetic aperture sonar (SAS) [27]. In such systems the sensor platform most often be carefully designed to achieve as uniform sampling as possible along the array length. Further, the data are typically measured to achieve $\lambda/2$ sampling. The compressive sensing theory indicates that random sampling is desirable, and that the samples may be situated greater than $\lambda/2$ apart. Random samples are likely naturally manifested by the flight of the platform, and sampling more coarsely than $\lambda/2$ offers advantages for reduced data storage and processing. The CS theory also provides fundamental bounds on reconstruction accuracy, as a function of sample rate and noise level.

Acknowledgements

The author wishes to thank Prof. Emmanuel Candès of Caltech for a careful reading of an early version of this paper, which resulted in important improvements. The research reported here was funded by the DARPA Mathematical Time Reversal Program and by the ONR Basic Research Challenge program.

References

[1] B. Steinberg, *Principles of Aperture and Array System Design Including Random and Adaptive Arrays.* Wiley Press, 1976.

[2] Y. Lo, "A probabalistic approach to the problem of large antenna arrays," *Radio Science*, vol. 68D, pp. 1011–1019, 1964.

[3] ——, "A mathematical theory of antenna arrays with randomly spaced elements," *IEEE Trans. Antennas and Propagation*, vol. 12, p. 257268, 1964.

[4] J. Hogbom, "Aperture synthesis with a nonregular distribution of interferometer baselines," *Astron. Astrophys. Supplements*, vol. 15, pp. 417–426, 1974.

[5] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Inform. Theory*, vol. 52, pp. 489–502, 2006.

[6] E. Candès and T. Tao, "Near-optimal signal recovery from random projections and universal encoding strategies," *IEEE Trans. on Inform. Theory*, vol. 52, pp. 5402–5425, 2006.

[7] E. Candès, J. Romberg, and T. Tao, "Signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, pp. 1207–1223, 2005.

[8] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. on Inform. Theory*, vol. 51, pp. 4203–4215, 2005.

[9] D. L. Donoho, "Compressed sensing," *IEEE Trans. Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[10] E. J. Candès, "The restricted isometry propert and its implications for compressed sensing," *Academic des Sciences*, 2008.

[11] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal. Statist. Soc B.*, vol. 58, no. 1, pp. 267–288, 1996.

[12] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, pp. 253–263, 2008.

[13] J. A. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," 2005.

[14] D. L. Donoho, Y. Tsaig, I. Drori, and J.-C. Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit," 2006.

[15] J. Li and P. Stoica, "Efficient mixed-spectrum estimation with applications to target feature extraction," *IEEE Trans. Signal Processing*, vol. 44, pp. 281–295, 1996.

[16] E. Candès, "Compressive sampling," *Proceedings of the International Congress of Mathematics*, vol. 3, pp. 1433–1452, 2006.

[17] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, pp. 1182–1195, 2007.

[18] C. Balanis, *Advanced Engineering Electromagnetics*. Wiley Press, 1989.

[19] P. Blomberg, G. Papanicolaou, and H. Zhao, "Super-resolution in time-reversal acoustics," *J. Acoust. Soc. of Am.*, vol. 111, pp. 230–248, 2002.

[20] J. Tsao and B. Steinberg, "Reduction of sidelobe and speckle artifacts in microwave imaging: The CLEAN technique," *IEEE Trans. Antennas Propagat.*, vol. 36, pp. 543–556, 1988.

[21] G. Davis, S. Mallat, and M. Avellaneda, "Greedy adaptive approximation," *J. Constr. Approx.*, vol. 13, pp. 57–98, 1997.

[22] Y. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition."

[23] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. on Signal Processing*, pp. 2346–2356, 2008.

[24] M. Seeger and H. Nickisch, "Compressed sensing and bayesian experimental design," in *ICML '08: Proceedings of the 25th international conference on Machine learning*, 2008.

[25] R. Castro, J. Haupt, R. Nowak, and G. Raz, "Finding needles in noisy haystacks," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2008.

[26] J. Curlander and R. McDonough, *Synthetic aperture radar: Systems and signal processing*. Wiley Press, 1991.

[27] L. Cutrona, "Comparison of sonar system performance achievable using synthetic-aperture techniques with the performance achievable by more conventional means," *J. Acoustical Soc. Am.*, vol. 58, pp. 336–348, 1975.