Sensing Matrix Design via Mutual Coherence Minimization for Electromagnetic Compressive Imaging Applications

Richard Obermeier and Jose Angel Martinez-Lorenzo

***Abstract*—Compressive sensing (CS) theory states that sparse signals can be recovered from a small number of linear measurements $y = Ax$ using ℓ_1- norm minimization techniques, provided that the sensing matrix satisfies a restricted isometry property (RIP). Unfortunately, the RIP is difficult to verify in electromagnetic imaging applications, where the sensing matrix is computed deterministically. Although it provides weaker reconstruction guarantees than the RIP, the mutual coherence is a more practical metric for assessing the CS recovery properties of deterministic matrices. In this paper, we describe a method for minimizing the mutual coherence of sensing matrices in electromagnetic imaging applications. Numerical results for the design method are presented for a simple multiple monostatic imaging application, in which the sensor positions for each measurement serve as the design variables. These results demonstrate the algorithm's ability to both decrease the coherence and to generate sensing matrices with improved CS recovery capabilities.**

***Index Terms*—Compressive sensing (CS), nonconvex optimization, sensing matrix design.**

I. Introduction

In ELECTROMAGNETIC imaging applications, one attempts to reconstruct the constitutive parameters or reflectivity profile of an object of interest from a set of scattered electric field measurements. Reconstruction accuracy is largely dependent upon the fidelity of the measurements, that is, the total number of measurements and their degree of independence. When classical imaging algorithms are employed, the reconstruction accuracy can be improved by increasing the number of measurements, for example by recording measurements at more locations or by exciting the object of interest using more frequencies. If performed naively, however, this approach can be expensive, both in terms of the hardware complexity of the imaging system and in terms of the information efficiency of each measurement. As more measurements are added to the system, the amount of additional information obtained by each new measurement can decrease dramatically.

Compressive Sensing (CS) [1]–[3] is a novel signal processing paradigm, which states that sparse signals can be recovered from a small number of linear measurements by solving a convex and computationally tractable $\ell_1 -$ norm minimization problem. Accurate recovery is only guaranteed when the sensing matrix is "well-behaved" according to a set of performance measures that relate the independence of the columns of the sensing matrix. CS has been applied to electromagnetic imaging applications with some success [4]–[6]. In the context of electromagnetic inverse problems, CS theory allows one to produce high resolutions images of the object of interest using undersampled data. However, CS has largely been applied as a heuristic in these applications, due to the fact that the Restricted Isometry Property (RIP), the most powerful performance measure for guaranteeing recovery using CS techniques, is very difficult to verify for matrices that are deterministically designed. In contrast, the mutual coherence is a computationally tractable measure that has been used to assess the CS recovery capabilities of sensing matrices [7], albeit with weaker reconstruction guarantees than the RIP. These properties make the coherence a suitable design measure for assessing the recovery capabilities of an electromagnetic imaging system. In this paper, we describe an algorithm for minimizing the mutual coherence of deterministically created sensing matrices. Although our motivation for this algorithm is derived from electromagnetic imaging, the method can be applied to any CS application that utilizes deterministic sensing matrices.

The remainder of this paper is organized as follows. In Section II, we discuss the fundamental theory of compressive sensing, and we describe the motivation for using the mutual coherence as the sensing matrix design metric for electromagnetic imaging applications. In this process, we describe how the mutual coherence relates to the maximum sidelobe level of the imaging system's point spread function (PSF). In Section III, we describe a first-order algorithm that can be used to minimize the coherence of a sensing matrix. Our method is more general than existing methods, e.g. [8]–[10], in that it can minimize the coherence of sensing matrices that are computed through a nonlinear, but differentiable relationship, as opposed to the

Manuscript received August 1, 2016; revised November 28, 2016 and January 27, 2017; accepted February 3, 2017. Date of publication February 17, 2017; date of current version May 8, 2017. This work has been partiality funded by the NSF CAREER program (Award no. 1653671), and the DHS (2013-ST-061-ED0001). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jong Chul Ye.

The authors are with the Department of Electrical and Computer Engineering and the Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115 USA (e-mail: jmartinez@coe.neu.edu; obermeier.ri@husky.neu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCI.2017.2671398

simple linear relationships that govern dictionary-based representations. In Section IV, we present results of the optimization technique applied to a multiple monostatic imaging system, in which the design variables are the position of the transmitting antenna in each measurement. These results demonstrate the algorithm's ability to reduce the mutual coherence of the sensing matrix, thereby improving its CS recovery capabilities. Finally, in Section V we conclude the paper by discussing several other imaging applications in which the coherence minimization can be applied.

II. Motivation

Consider a typical electromagnetic imaging problem, in which one seeks to recover the unknown reflectivity vector $x \in \mathbb{C}^N$ from a set of $y \in \mathbb{C}^M$ noisy measurements. Ignoring the mutual coupling between scattering elements in the imaging region, the relationship between x and y can be approximately linearized as $y = Ax + n$, where $n \in \mathbb{C}^M$ represents the noise in the measurements. When $M < N$, there exist an infinite number of solutions x satisfying $y = Ax$, and so regularization techniques must be employed in order to induce a unique solution. A particularly interesting scenario arises when the vector x is sparse, that is the number of non-zero elements S in x is much smaller than the total number of elements N. In this case, CS theory [1]–[3] establishes that the sparse vector can be stably recovered from an incomplete set of measurements by solving the following convex optimization program:

$$\begin{aligned} &\underset{x}{\text{minimize}} \quad \|x\|_{\ell_1} \\ &\text{subject to} \quad \|Ax - y\|_{\ell_2} \leq \eta \end{aligned} \tag{1}$$

provided that the sensing matrix A is "well behaved." Several methods have been used in the literature in order to measure the recovery properties of sensing matrices. The Restricted Isometry Property (RIP), arguably the most popular measure, uses the concept of restricted isometry constants in order to establish the tightest performance guarantees currently known. For a fixed sparsity level S, the restricted isometry constant δ_S is the smallest positive constant such that

$$(1 - \delta_S)\|x\|_{\ell_2}^2 \leq \|Ax\|_{\ell_2}^2 \leq (1 + \delta_S)\|x\|_{\ell_2}^2 \tag{2}$$

is satisfied for all vectors with $\|x\|_{\ell_0} \leq S$, where the "ℓ_0- norm" measures the number of non-zero elements in the vector. In other words, the restricted isometry constant δ_S establishes bounds for the singular values of submatrices obtained by selecting any S columns from the complete sensing matrix A. It is important to note that this formulation assumes that the sensing matrix A has ℓ_2- normalized columns.

The RIP has been expressed in many forms over the years. One of these forms, developed by Candès [11], states that distance between the optimal solution x^* to (1) and the true sparse vector x_t is bounded according to

$$\|x^* - x_t\|_{\ell_2} \leq C\eta \tag{3}$$

provided that the restricted isometry constant $\delta_{2S} < \sqrt{2} - 1$. The positive constant C in (3) depends only on the value of δ_{2S}; see [11] for details. The RIP establishes a very powerful result; for the noiseless case where $\eta = 0$, the RIP guarantees exact recovery of sparse vectors. In practice, it is very difficult to compute the restricted isometry constants of a given sensing matrix A, as one must perform a search over all combinations of S columns that can be extracted from A. This fact has led researches to utilize random matrices, where each element is drawn from an i.i.d. sub-Gaussian distribution, in order to construct sensing matrices that satisfy the RIP with high probability. Unfortunately, this approach cannot be taken in electromagnetic imaging applications, where the sensing matrix is constrained by a number of practical factors, such as the positions of the transmitting and receiving antennas and the excitation frequencies.

The mutual coherence [12] is a more practical measure for deterministic matrices, such as those used in electromagnetic imaging applications. The mutual coherence of a sensing matrix is defined as follows:

$$\mu(A) = \max_{1 \leq i \neq j \leq N} \frac{|a_i^H a_j|}{\|a_i\|_{\ell_2} \|a_j\|_{\ell_2}} \tag{4}$$

where a_i is the $i-$th column of A. The coherence has several properties that make it an intriguing measure for assessing the CS recovery capabilities of the deterministic sensing matrices. First, it can be shown that the coherence is fundamentally related to the restricted isometry constants by the following two relations [13]:

$$\delta_2 = \mu \tag{5}$$

$$\delta_S \leq (S - 1)\mu \tag{6}$$

Therefore, an upper bound on the restricted isometry constants can be established using the $O(N^2)$ computations required to evaluate the coherence, which represents a significant improvement over the combinatorial complexity of the restricted isometry constants. Second, the theoretical limitations of the coherence are also well established in the literature. It is well known that the coherence is bounded by [14] :

$$\sqrt{\frac{N - M}{M(N - 1)}} \leq \mu \leq 1 \tag{7}$$

The lower bound in particular is useful for "grading" the designed sensing matrices. Third, the coherence is intimately related to the point spread function (PSF) commonly used to assess the focusing capabilities of imaging systems. The Gramian matrix $A^H A$ can be seen as a focusing operation, in which each column represents the PSF achieved when focusing at a specific position. The mutual coherence therefore measures the absolute value of the largest sidelobe over all focusing positions. Minimizing the mutual coherence ensures that the maximum sidelobe level in the PSF is minimized. These three properties make the mutual coherence a desirable and practical metric for designing sensing matrices in electromagnetic imaging applications.

III. Coherence Minimization Algorithm

The coherence minimization algorithm described in this section is related to the prior works [8]–[10] and the references

therein. The method described in [9] uses a specialized formulation for the sensing matrix and utilizes a coherence metric based upon the ℓ_2-norm. The method described in [10] requires the design variables to be the elements of the sensing matrix and utilizes a coherence metric based on the ℓ_2-norm. The method described in [8] minimizes the mutual coherence as it is defined in (4), but requires the sensing matrix to be represented as a projected dictionary, i.e. $A = PD$. In contrast, our method minimizes the mutual coherence of (4) in applications where the sensing matrix is some nonlinear function of the design variables, i.e. $A = f(p)$. To this end, the algorithm described herein is an extension to the prior works and can used in any CS application, not just electromagnetic imaging applications. A specific instance of the coherence minimization problem for electromagnetic imaging is presented in Section IV.

Consider the general design problem, in which the sensing matrix $A \in \mathbb{C}^{M\times N}$ is related to the $p \in \mathbb{C}^L$ design variables according to the differentiable and nonlinear relationship $A = f(p)$. The design algorithm then seeks the minimizer of the following optimization program:

$$\begin{aligned} \underset{p,A}{\text{minimize}} \quad & \mu(A) \\ \text{subject to} \quad & A = f(p) \\ & p \in Q_p \end{aligned} \tag{8}$$

where Q_p defines the set of feasible values that the design variables p can take. Equation (8) is a nonlinear, non-convex, non-differentiable optimization program that cannot be easily solved in its current form. To simplify the problem, we shall make the following modifications to the optimization program. First, we consider the Gramian matrix $\hat{U} = \hat{A}^H \hat{A}$, where $\hat{A}$ is equivalent to the sensing matrix A but has normalized columns. The matrix $\hat{U}$ is Hermitian, and so it can be expressed in terms of $N(N+1)/2$ variables as follows:

$$\hat{U} = \begin{pmatrix} u_{1,1} & u_{1,2} & u_{13} & \dots & u_{1,N-2} & u_{1,N-1} & u_{1,N} \\ u_{1,2}^* & u_{2,2} & u_{2,3} & \dots & u_{2,N-2} & u_{2,N-1} & u_{2,N} \\ u_{1,3}^* & u_{2,3}^* & u_{3,3} & \dots & u_{3,N-2} & u_{3,N-1} & u_{3,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ u_{1,N-2}^* & u_{2,N-2}^* & u_{3,N-2}^* & \dots & u_{N-2,N-2} & u_{N-2,N-1} & u_{N-2,N} \\ u_{1,N-1}^* & u_{2,N-1}^* & u_{3,N-1}^* & \dots & u_{N-2,N-1}^* & u_{N-1,N-1} & u_{N-1,N} \\ u_{1,N}^* & u_{2,N}^* & u_{3,N}^* & \dots & u_{N-2,N}^* & u_{N-1,N}^* & u_{N,N} \end{pmatrix} \tag{9}$$

Only the upper triangular components of $\hat{U}$ that do not lie on the main diagonal are required in order to compute the mutual coherence, so we introduce the vector $u = (u_{1,2}, u_{1,3}, \dots, u_{1,N}, u_{2,3}, \dots, u_{N-1,N}) \in \mathbb{C}^{N(N-1)/2}$ to compactly represent these elements. Consequently, we shall use the notation $u_{i,j}$ to represent the element in the vector u that corresponds to the value of $\hat{U}$ at the location (i,j). In the second modification, we expand the definition of $f(p)$ as follows:

$$f(p) = \big(f_1(p),\ f_2(p),\ \dots,\ f_N(p) \big) \tag{10}$$

that is, $f_i(p)$ computes the $i-$th column of the sensing matrix A.

With these two modifications, the coherence minimization problem can be expressed in terms of p and u as follows:

$$\begin{aligned} \underset{p,u}{\text{minimize}} \quad & \|u\|_{\ell_\infty} \\ \text{subject to} \quad & u_{i,j} = \frac{f_i(p)^H f_j(p)}{\|f_i(p)\|_{\ell_2}\|f_j(p)\|_{\ell_2}}\ ,\ 1 \le i \ne j \le N \\ & p \in Q_p \end{aligned} \tag{11}$$

Note that (8) and (11) are equivalent; no approximations were made in formulating (11). The Augmented Lagrangian method [15] can be used in order to represent this constrained optimization problem in an unconstrained form. Using the indicator function, $I_{Q_p}(p) = \{ {}^{0}_{\infty} {}^{p\in Q_p}_{p\notin Q_p}$, for the feasible set Q_p, the Augmented Lagrangian for this problem can be expressed as follows:

$$\begin{aligned} \mathcal{L}_{\mathcal{A}}(p,u,\beta;\rho) = {} & \|u\|_{\ell_\infty} + I_{Q_p}(p) \\ & + \sum_{1\le i\ne j\le N} \beta_{i,j}\left(u_{i,j} - \frac{f_i(p)^H f_j(p)}{\|f_i(p)\|_{\ell_2}\|f_j(p)\|_{\ell_2}}\right) \\ & + \sum_{1\le i\ne j\le N} \frac{\rho}{2}\left|u_{i,j} - \frac{f_i(p)^H f_j(p)}{\|f_i(p)\|_{\ell_2}\|f_j(p)\|_{\ell_2}}\right|^2 \end{aligned} \tag{12}$$

where ρ is a positive scaling factor and $\beta = (\beta_{1,2}, \beta_{1,3}, \dots, \beta_{1,N}, \beta_{2,3}, \dots, \beta_{N-1,N}) \in \mathbb{C}^{N(N-1)/2}$ are the Lagrange multipliers. Often times, it is convenient to express the Augmented Lagrangian in the following equivalent form [16]:

$$\begin{aligned} \mathcal{L}_{\mathcal{A}}(p,u,\beta;\rho) = {} & \|u\|_{\ell_\infty} + I_{Q_p}(p) \\ & + \sum_{1\le i\ne j\le N} \frac{\rho}{2}\left|u_{i,j} - \frac{f_i(p)^H f_j(p)}{\|f_i(p)\|_{\ell_2}\|f_j(p)\|_{\ell_2}} + \beta_{i,j}/\rho\right|^2 \end{aligned} \tag{13}$$

The Augmented Lagrangian method finds a local optimal solution to (11) by solving a sequence of unconstrained subproblems of the form of (13), where β is held fixed. These subproblems can be solved using an alternating minimization procedure, which is described in detail in the Appendix.

When a given instance of (13) is solved, the Lagrange multipliers are updated using the following equation:

$$\beta_{i,j}^{(k+1)} = \beta_{i,j}^{(k)} + \rho^{(k)}\left(u_{i,j}^{(k)} - \frac{f_i(p^{(k)})^H f_j(p^{(k)})}{\|f_i(p^{(k)})\|_{\ell_2}\|f_j(p^{(k)})\|_{\ell_2}}\right) \tag{14}$$

where the superscripts indicate the iteration number, i.e. the Lagrange multiplier $\beta_{i,j}^{(k)}$ is used on the $k-$th instance of (13). To ensure that the algorithm converges to a stationary point, it is often necessary to increase ρ at each iteration. Our design method utilizes the update approach described in [15]. The coherence minimization optimization procedure is summarized in Algorithm 1.

Algorithm 1: Summary of the Augmented Lagrangian update procedure for the coherence minimization problem of Eq.(11)

Choose the initial values for $p^{(0)}$, $\rho^{(1)}$;

Set $u_{i,j}^{(0)} = \frac{f_i(p^{(0)})^H f_j(p^{(0)})}{\|f_i(p^{(0)})\|_{\ell_2}\|f_j(p^{(0)})\|_{\ell_2}}$, $\beta_{i,j}^{(1)} = 0$;

for k = *1,2,3...* **do**

Solve the unconstrained subproblem

$$\left(p^{(k)}, u^{(k)}\right) = \underset{p,u}{\operatorname{argmin}}\ \mathcal{L}_{\mathcal{A}}\left(p, u, \beta^{(k)}; \rho^{(k)}\right)$$

Update the dual variables

$$\beta_{i,j}^{(k+1)} = \beta_{i,j}^{(k)} + \rho^{(k)}\left(u_{i,j}^{(k)} - \frac{f_i(p^{(k)})^H f_j(p^{(k)})}{\|f_i(p^{(k)})\|_{\ell_2}\|f_j(p^{(k)})\|_{\ell_2}}\right)$$

Compute $\rho^{(k+1)}$ using the method described in [15]

IV. Numerical Results

In this section, we present numerical results for the coherence minimization technique applied to an electromagnetic imaging application. Consider a scenario in which a single transmitting and receiving antenna is used to excite a region of interest with a single frequency. The discretized measurement process for this system can be modeled as follows:

$$y_m = \sum_{n=1}^{N} x_n e^{-j2k\|r_m - r_n\|_{\ell_2}} = \sum_{n=1}^{N} A_{mn} x_n \qquad (15)$$

where y_m is the measured scattered field, r_m is the position of the antenna, r_n is a position in the imaging region, k is the wavenumber, and x_n is the reflectivity at the position in the imaging region. This linear sensing model can be derived by applying the Born approximation to a transmitter that radiates uniformly over the imaging region, i.e. $s(k, r_n) = Qe^{-jk\|r_m - r_n\|_{\ell_2}}$. When the signal interaction with the scatterers is linearized, the signal received by the antenna has an additional phase term $e^{-jk\|r_m - r_n\|_{\ell_2}}$ due to the Green's function of the homogeneous background medium. Combining these terms and absorbing the constant scale factors into the reflectivity vector leads to the sensing matrix formulation of (15).

Keeping the wavenumber fixed, the objective is to select the antenna positions r_m such that the coherence is minimized. Assuming that $r_m \neq r_n$, the sensing matrix relationship $A = f(r_1, r_2, \ldots, r_M)$ is differentiable, and so the method described in Section III and the Appendix can be applied. We consider two different instances of this problem, one in which the positions of the antenna are restricted to lie along a circle centered about the imaging region, and one in which the positions of the antenna are restricted to lie on a two-dimensional plane parallel to the imaging region.

A. Circular Configuration

Table I summarizes the design parameters and constraints imposed on the circular optimization problem. The antenna was constrained to operate at $M = 60$ positions along the circle of radius $20\,\lambda$ relative to the center of the imaging region. The initial positions were selected by distributing the points uniformly over the circle. The sensing matrix was computed using (15) at $N = 121$ imaging points, uniformly located on a 2D $5\,\lambda$ by $5\,\lambda$ grid centered at the origin. This corresponds to a grid size of approximately $0.5\,\lambda$.

TABLE I
Summary of Design Parameters and Constraints for the Circular Optimization Problem

Design Parameters and Constraints		
Parameter	Baseline Value	Constraint
M	60	–
N	121	–
r_n	$5\,\lambda$ by $5\,\lambda$ grid centered at origin	–
r_m	Uniformly spaced over circle	$\|r_m\|_{\ell_2} = 20\,\lambda$

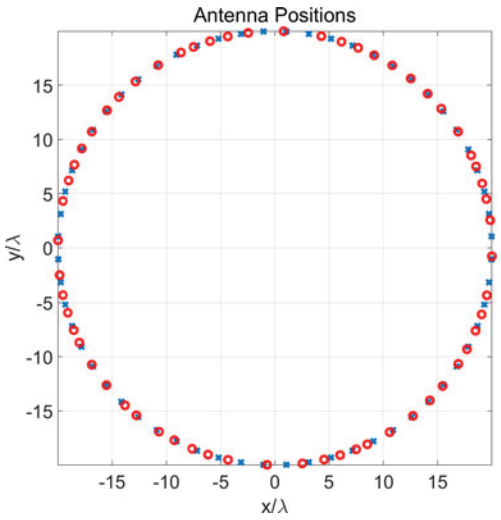

Fig. 1. Antenna positions of the baseline (blue) and optimized (red) designs in the circular configuration.

Fig. 1 displays the baseline antenna positions in blue and the optimized antenna positions in red. The optimized antenna positions are approximately symmetric about the line $y = x$ and have clearly been shifted relative to the baseline positions. Figs. 2 and 3 display the magnitude of the Gramian matrices $G = A^H A$ of the normalized sensing matrices obtained using (15) with the baseline and optimized antenna positions. Note that the red ellipses in these figures highlight the regions where the off-diagonal elements differ significantly between the two configurations. It is clear from these plots that the maximum off-diagonal element has been decreased in the optimized design. Specifically, the optimized design has a mutual coherence of approximately 0.1943, which is still an improvement over the 0.3232 coherence of the baseline design even though it does not achieve the minimum coherence of 0.0920 defined by (7). Given the restrictive nature of the design constraints in this problem, it is not surprising that the minimum coherence was not achieved. In problems where the designer has more control over the individual elements of the sensing matrices, such as the

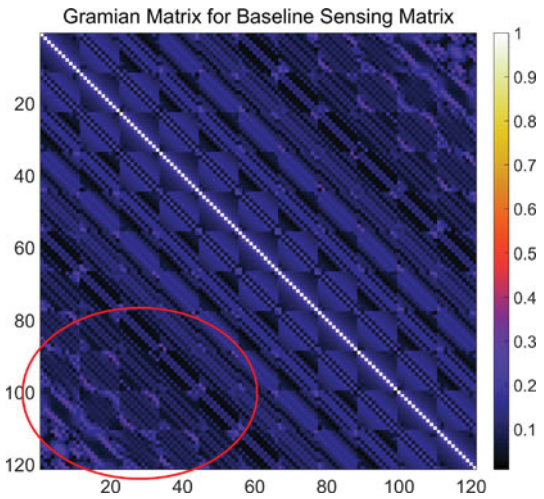

Fig. 2. Gramian matrix of the baseline design sensing matrix for the circular configuration. The mutual coherence, given by the maximum off-diagonal element, is approximately 0.3232.

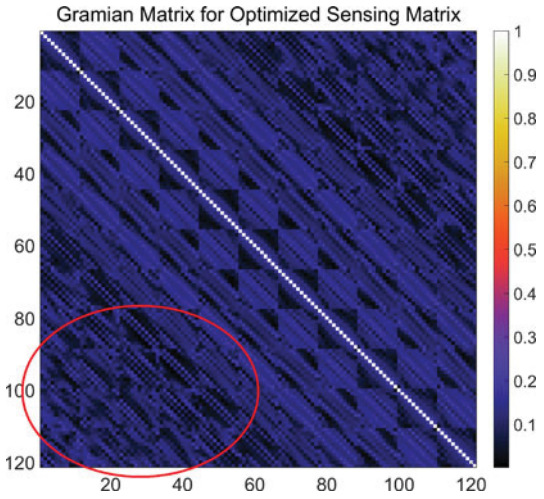

Fig. 3. Gramian matrix of the optimized design sensing matrix for the circular configuration. The mutual coherence, given by the maximum off-diagonal element, is approximately 0.1943.

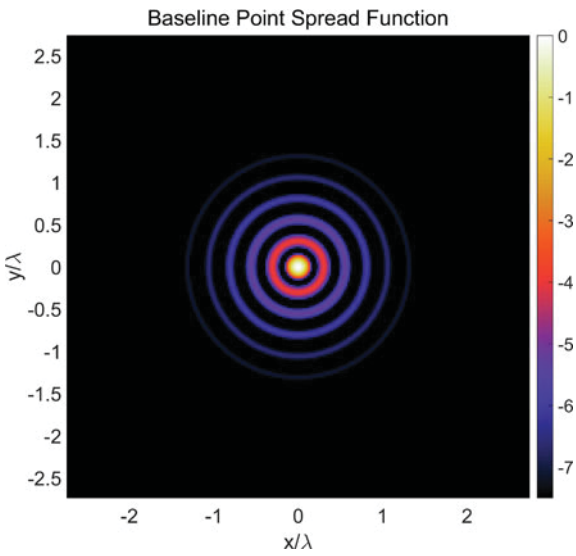

Fig. 4. Point spread function of the baseline design for the circular configuration focused at $(0, 0)$ in dB.

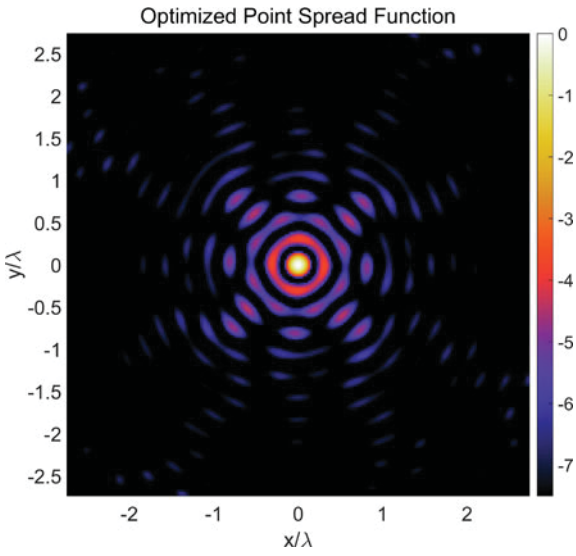

Fig. 5. Point spread function of the optimized design for the circular configuration focused at $(0, 0)$ in dB.

linear problems addressed in [8], the design method is likely to achieve solutions closer to the Welch bound.

The improvement in mutual coherence can be more greatly appreciated by inspecting the point spread functions. Figs. 4 and 5 display the PSF's of the baseline and optimized designs when focused at the origin. These plots are equivalent to the center columns of the Gramian matrices displayed in Figs. 2 and 3. For improved visibility, the PSF's are displayed in dB over the range $(-7.5, 0)$. The main lobes of the baseline and optimized designs are approximately equal in size. However, the two designs differ significantly in their sidelobe distribution. The baseline design PSF is largely concentrated near the center, and its sidelobes tend to decrease below the -7.5 dB threshold near the edges of the imaging region. The optimized design PSF has more sidelobes that exceed the lower-limit of -7.5 dB used in the plots, but it has a smaller maximum amplitude overall. This result is to be expected, since the coherence minimizes the maximum sidelobe level.

Figs. 6 and 7 display the PSF's of the baseline and optimized designs when focused at the point $(-1.9\,\lambda, -1.9\,\lambda)$. The baseline design PSF has a significant number of large amplitude sidelobes along an arc centered at the focusing point. This phenomenon commonly occurs in systems that utilize uniformly spaced sensors. In contrast, the optimized design PSF has sidelobes that are comparable to the PSF obtained when focused at the origin.

In order to further assess the improved CS recovery capabilities of the optimized design, we compared the reconstruction accuracies of two different algorithms, Orthogonal Matching Pursuit (OMP) [17] and Basis Pursuit (BP, equality-constrained

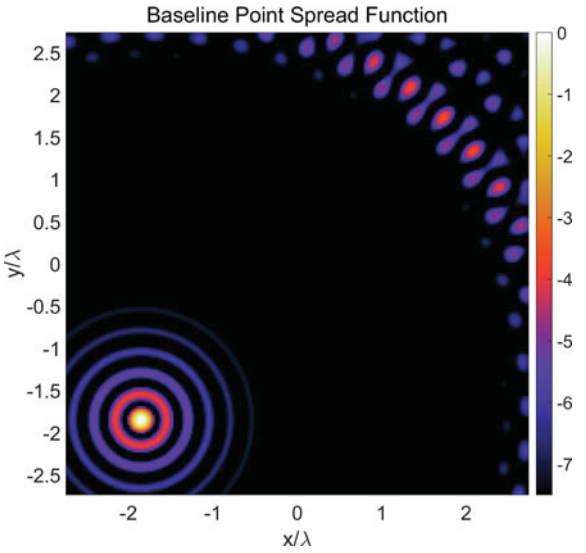

Fig. 6. Point spread function of the baseline design for the circular configuration focused at $(-1.9\,\lambda, -1.9\,\lambda)$ in dB.

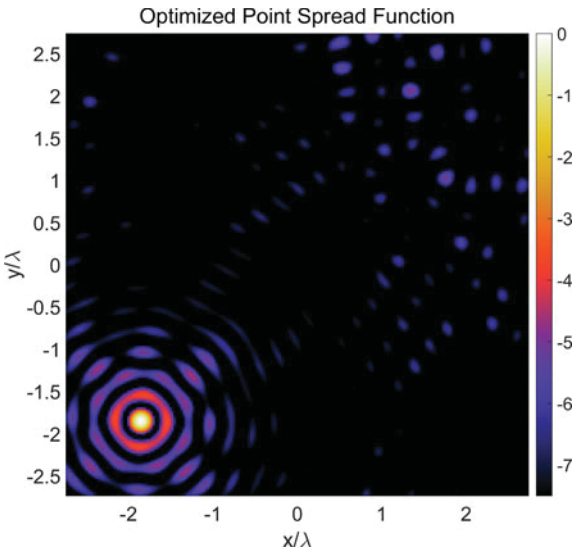

Fig. 7. Point spread function of the optimized design for the circular configuration focused at $(-1.9\,\lambda, -1.9\,\lambda)$ in dB.

ℓ_1-norm minimization). OMP is a greedy algorithm that has a lower computational complexity than BP. In practice, BP tends to outperform OMP in terms of reconstruction accuracy. Nevertheless, OMP has the same reconstruction guarantees as BP when the mutual coherence is used as the performance metric [13]. OMP and BP were tested using three different sensing matrices: that of the baseline design, that of the optimized design, and that of a randomized design where the sensor locations were randomly distributed over the feasible set. The random sensing matrix was considered in order to demonstrate that a random sensing matrix is not guaranteed to have good CS reconstruction capabilities when its elements are highly correlated to each other.

For the analysis, the noiseless measurements were generated directly from the linearized relationship $y = Ax$ in order to

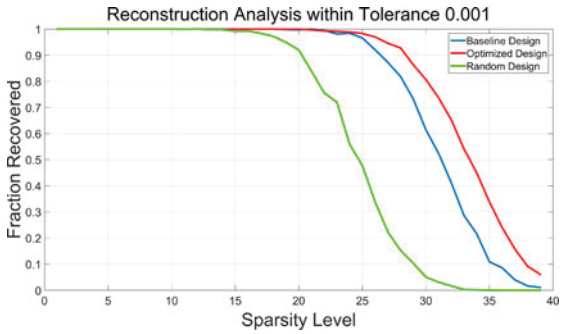

Fig. 8. Numerical comparison of the reconstruction accuracies of OMP using the baseline design (blue), optimized design (red), and random design (green) in the circular configuration.

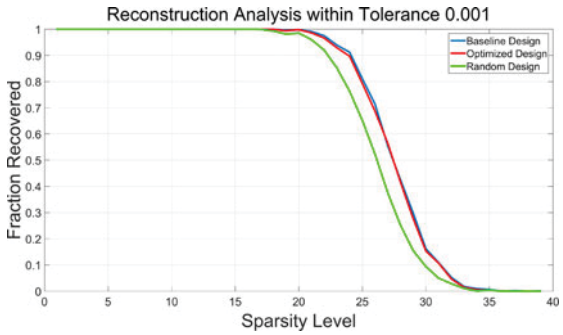

Fig. 9. Numerical comparison of the reconstruction accuracies of BP using the baseline design (blue), optimized design (red), and random design (green) in the circular configuration.

assess the matrices under ideal conditions. As a result, this analysis does not consider the effects of either measurement noise or modeling error (e.g. Born approximation) on the reconstruction accuracy. Fig. 8 displays the probability of recovering a solution vector with a normalized error $\|x_e - x_t\|_{\ell_2} / \|x_t\|_{\ell_2} \leq 0.001$ using OMP with the baseline sensing matrix (blue), the optimized sensing matrix (red), and the random sensing matrix (green), whose coherence of 0.3257 is comparable to that of the baseline sensing matrix. The probability at each sparsity level was estimated by generating 1000 vectors at each sparsity level and computing the fraction of vectors that were recovered within the tolerance criteria. For low sparsity levels, OMP performed equally well with all three sensing matrices. However, the optimized sensing matrix clearly outperformed both the baseline and random sensing matrices at higher sparsity levels. Interestingly, the random sensing matrix performed the worst.

Fig. 9 displays the reconstruction probabilities obtained using BP instead of OMP. The CVX optimization toolbox [18] was used in order to solve the BP problems. In this case, the optimized and baseline sensing matrices provided comparable reconstruction accuracy, and the random sensing matrix performed slightly worse than the other two. Interestingly, OMP outperformed BP when the baseline and optimized sensing matrices were used, and BP outperformed OMP when the random sensing matrix was used. Note that the same random number seed was used to randomly generate the sparse vectors in all four tests, so that each test operated with the same set of ground-truth vectors.

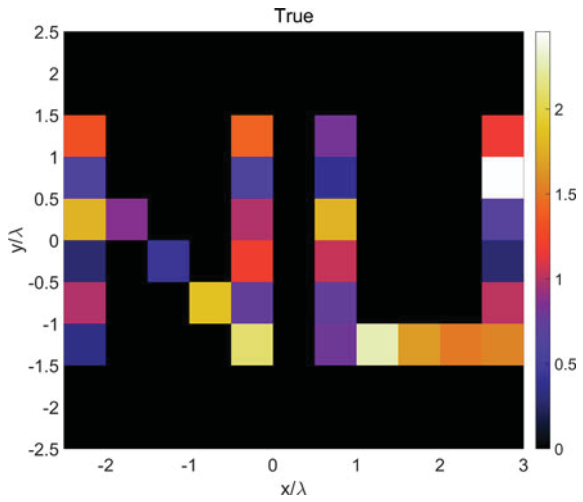

Fig. 10. Magnitude of the true reflectivity (S = 30).

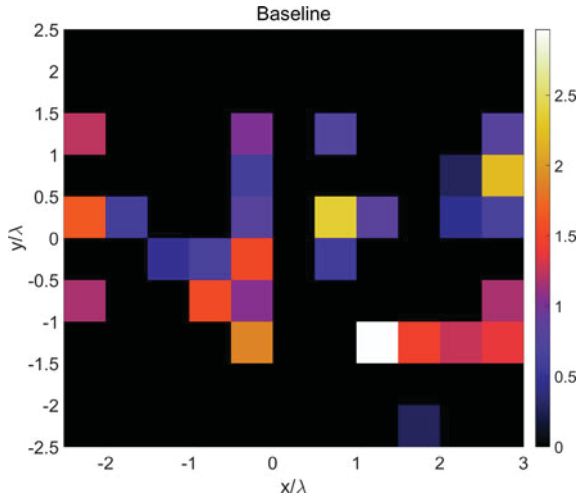

Fig. 11. Magnitude of the estimated reflectivity recovered by OMP using the baseline circular antenna configuration. The reconstruction has a normalized error $\|x_e - x_t\|_{\ell_2} / \|x_t\|_{\ell_2} = 0.375$.

The difference in reconstruction accuracy between the baseline, optimized, and random designs can be further appreciated through visual inspection of the reconstructed vectors. Fig. 10 displays the magnitude of the true scalar reflectivity vector with $S = 30$ non-zero values, and Figs. 11–13 display the reconstructions obtained by OMP using the baseline, optimized, and random antenna configurations respectively. While the baseline and random reconstructions suffer from significant errors in both the magnitudes of the scalar reflectivity and in the positions of the scatterers, the optimized design recovers the sparse vector exactly. This result is to be expected on average, given the sparsity level of the true reflectivity and the reconstruction accuracy results of Fig. 8.

B. Planar Configuration

Table II summarizes the design parameters and constraints imposed on the planar optimization problem. The antenna was constrained to operate at $M = 64$ positions within a $5\,\lambda$ by $5\,\lambda$ grid centered at $x = y = 0$ on the $z = 5\,\lambda$ plane. The initial positions were selected by distributing the points uniformly over the rectangular grid, and the sensing matrix was computed at the same positions as in the circular example.

Fig. 14 displays the baseline antenna positions in blue and the optimized antenna positions in red. Once again, the optimized design distributes the antenna positions symmetrically.

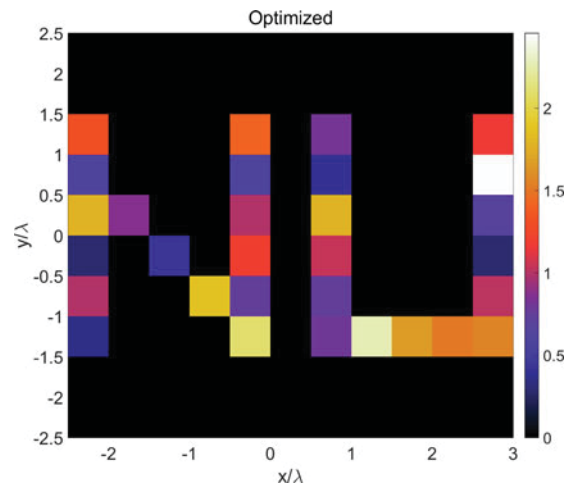

Fig. 12. Magnitude of the estimated reflectivity recovered by OMP using the optimized circular antenna configuration. The reflectivity is recovered exactly.

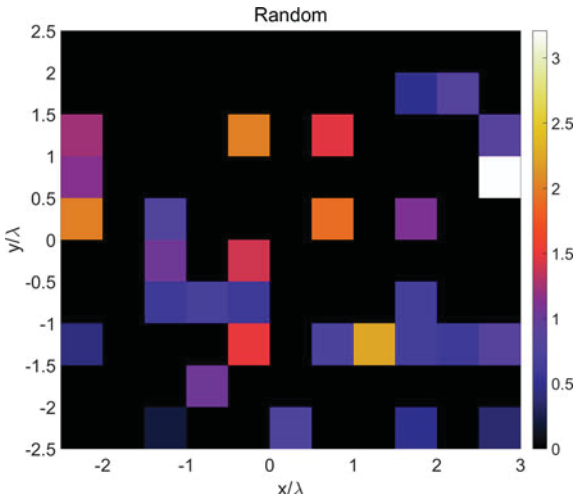

Fig. 13. Magnitude of the estimated reflectivity recovered by OMP using the random circular antenna configuration. The reconstruction has a normalized error $\|x_e - x_t\|_{\ell_2} / \|x_t\|_{\ell_2} = 0.678$.

TABLE II
SUMMARY OF DESIGN PARAMETERS AND CONSTRAINTS IN THE PLANAR OPTIMIZATION PROBLEM

Design Parameters and Constraints		
Parameter	Baseline Value	Constraint
M	64	–
N	121	–
r_n	$5\,\lambda$ by $5\,\lambda$ grid centered at origin	–
r_m	Uniformly spaced over $5\,\lambda$ by $5\,\lambda$ grid at $z = 5\,\lambda$	$\|x_m\| \leq 2.5\,\lambda$ $\|y_m\| \leq 2.5\,\lambda$ $z_m = 5\,\lambda$

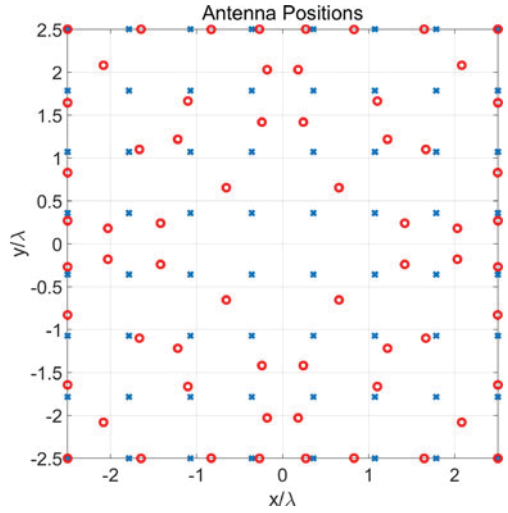

Fig. 14. Antenna positions of the baseline (blue) and optimized (red) designs in the planar configuration.

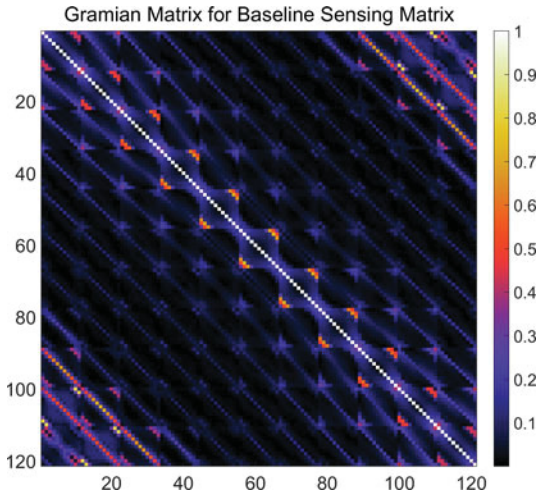

Fig. 15. Gramian matrix of the baseline design sensing matrix for the planar configuration. The mutual coherence, given by the maximum off-diagonal element, is approximately 0.8300.

Figs. 15 and 16 display the magnitude of the Gramian matrices computed from the baseline and optimized sensing matrices. In this example, the optimized design has a mutual coherence of approximately 0.2252, which is a significant improvement over the 0.8300 coherence of the baseline design even though it does not achieve the minimum coherence of 0.0862.

Like in the previous example, the improvement in mutual coherence can be appreciated by inspecting the point spread functions. Figs. 17 and 18 display the PSF's of the baseline and optimized designs when focused at the origin, and Figs. 19 and 20 display the PSF's of the baseline and optimized designs when focused at the point $(-1.9\,\lambda, -1.9\,\lambda)$. Once again, the baseline design suffers from significant sidelobe levels when focused off-center, while the optimized design provides a more uniform performance over all focusing positions.

To further assess the reconstruction capabilities of the optimized planar design, we repeated the reconstruction analysis of the previous subsection using the baseline, optimized, and

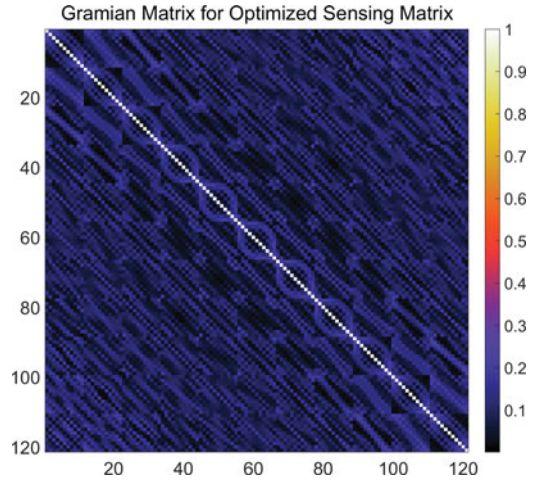

Fig. 16. Gramian matrix of the optimized design sensing matrix for the planar configuration. The mutual coherence, given by the maximum off-diagonal element, is approximately 0.2252.

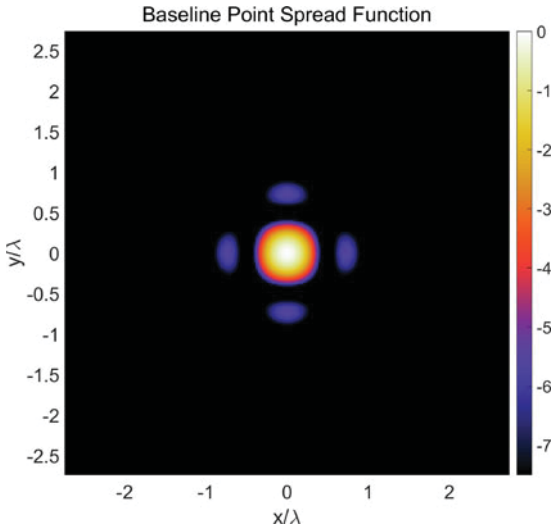

Fig. 17. Point spread function of the baseline design for the planar configuration focused at $(0, 0)$ in dB.

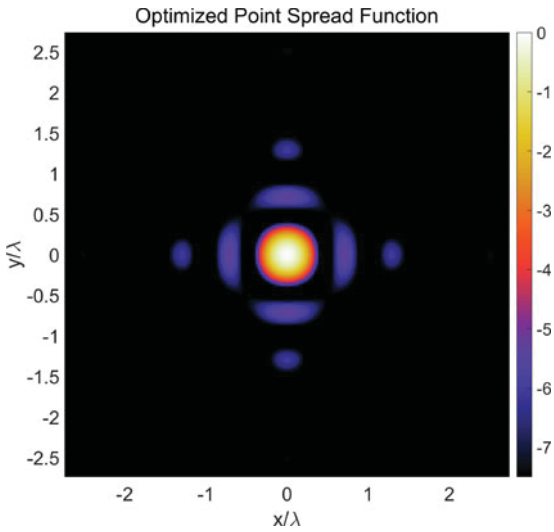

Fig. 18. Point spread function of the optimized design for the planar configuration focused at $(0, 0)$ in dB.

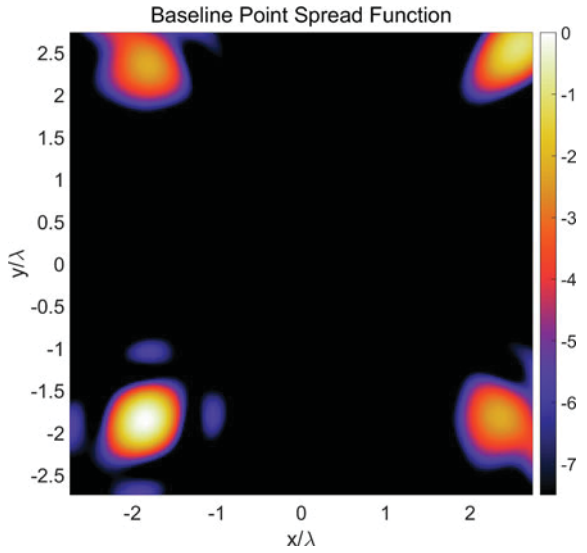

Fig. 19. Point spread function of the baseline design for the planar configuration focused at $(-1.9\,\lambda, -1.9\,\lambda)$ in dB.

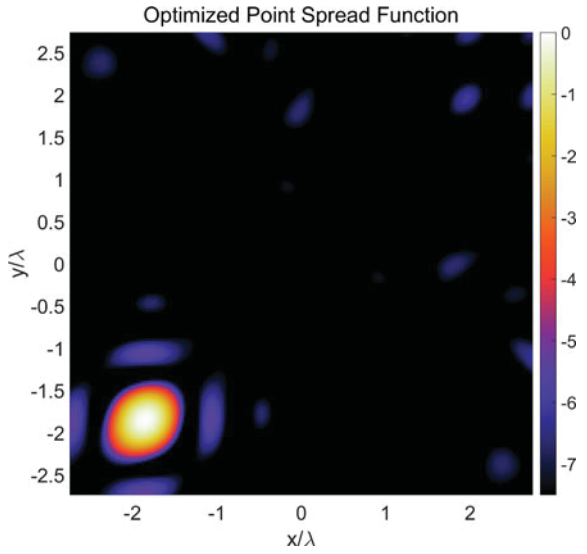

Fig. 20. Point spread function of the optimized design for the planar configuration focused at $(-1.9\,\lambda, -1.9\,\lambda)$ in dB.

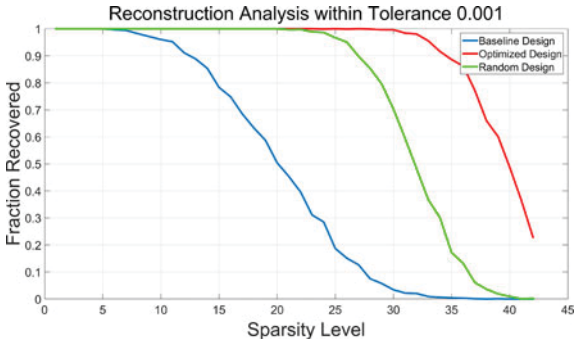

Fig. 21. Numerical comparison of the reconstruction accuracies of OMP using the baseline design (blue), optimized design (red), and random design (green) in the planar configuration.

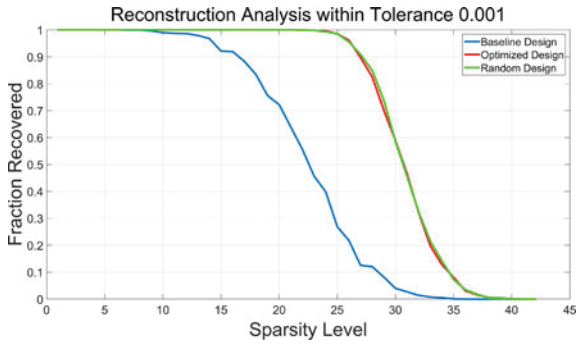

Fig. 22. Numerical comparison of the reconstruction accuracies of BP using the baseline design (blue), optimized design (red), and random design (green) in the planar configuration.

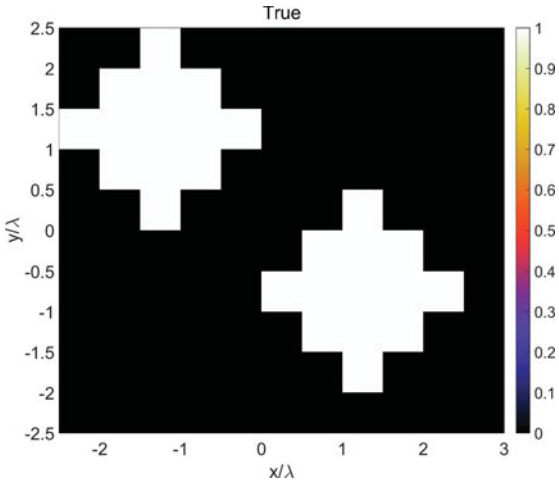

Fig. 23. Magnitude of the true reflectivity (S = 26).

random sensing matrices for the planar configuration. The random sensing matrix for this example had a coherence of 0.33564, which is much better than that of the baseline sensing matrix and worse than that of the optimized sensing matrix. Fig. 21 displays the probability of recovering a solution vector with a normalized error $\|x_e - x_t\|_{\ell_2} / \|x_t\|_{\ell_2} \leq 0.001$ using OMP with the baseline sensing matrix (blue), the optimized sensing matrix (red), and the random sensing matrix (green). The probability at each sparsity level was again estimated by generating 1000 vectors at each sparsity level and computing the fraction of vectors that were recovered within the tolerance criteria. The optimized sensing matrix for this problem significantly outperformed both the baseline and random sensing matrices, with the baseline sensing matrix performing the worst. Fig. 22 displays the reconstruction probabilities obtained using BP instead of OMP. While the optimized sensing matrix outperformed the baseline sensing matrix, it performed about the same as the random sensing matrix.

To get an idea of how noise affects the reconstruction accuracies, OMP was used to reconstruct the ground truth geometry displayed in Fig. 23 using measurements corrupted by additive noise. The elements of the additive noise vector were drawn from i.i.d. Gaussian distributions so that the measurement signal to noise ratio (SNR) was equal to 20 dB. The reconstructions for the baseline, optimized, and random sensing matrices are

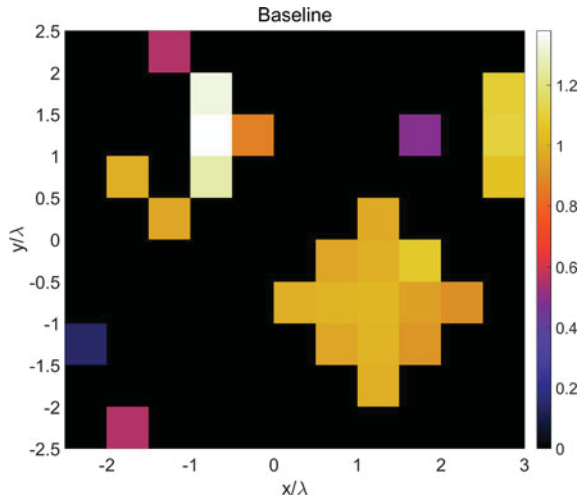

Fig. 24. Magnitude of the estimated reflectivity recovered by OMP using the baseline planar antenna configuration.

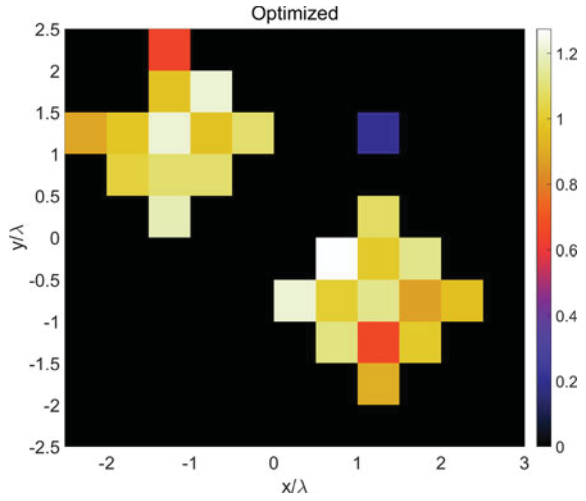

Fig. 25. Magnitude of the estimated reflectivity recovered by OMP using the optimized planar antenna configuration.

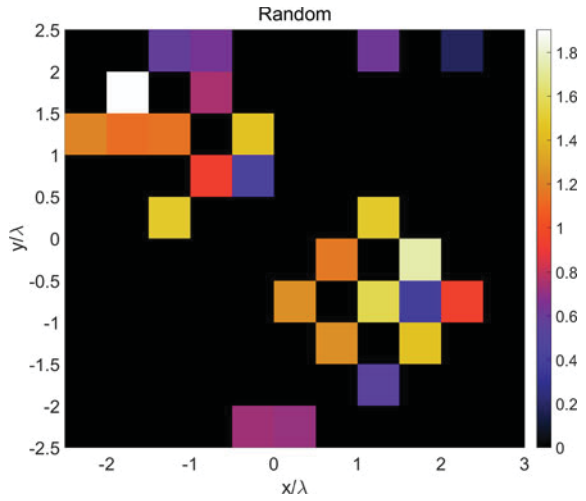

Fig. 26. Magnitude of the estimated reflectivity recovered by OMP using the random planar antenna configuration.

displayed in Figs. 24–26 respectively. These figures clearly show that the optimized configuration generated more accurate reconstructions than the baseline and random configurations.

V. Conclusion

This paper describes a novel technique for deterministically designing sensing matrices used in electromagnetic imaging applications. By minimizing the mutual coherence, the design technique directly enhances the reconstruction capabilities of the imaging system when compressive sensing techniques are employed. Although this paper only analyzed a simple multiple monostatic imaging application, there are several other scenarios where this technique can be employed. Dictionary learning techniques, similar to those studied in [8], can be applied to imaging problems. Given a dictionary of possible measurements, the method selects a subset of them such that the mutual coherence is minimized. The method can also be applied to energy allocation problems. Given several transmitting antenna and a single receiving sensor, the method designs the magnitude and phase of the currents driving each of the transmitting antennas such that the coherence is minimized. The design technique can also be applied to related imaging applications, such as the ultrasonic imaging of cargo containers [19]. Finally, the method can be used to design novel compressive antennas [20], [21], in which case the constitutive parameters of the scattering elements added to a baseline reflector antenna are optimized such that the coherence of the sensing matrix is minimized. The use of the coherence minimization technique in these applications and in others will be investigated in our future research.

Appendix A
Augmented Lagrangian Subproblem

In Section III, we described an Augmented Lagrangian method for minimizing the mutual coherence of a sensing matrix. One of the key steps in this procedure solves the unconstrained subproblem of (13), which is repeated here for convenience:

$$\mathcal{L}_{\mathcal{A}}(p,u,\beta;\rho) = \|u\|_{\ell_\infty} + I_{Q_p}(p) + \sum_{1\le i\ne j\le N} \frac{\rho}{2}\left| u_{i,j} - \frac{f_i(p)^H f_j(p)}{\|f_i(p)\|_{\ell_2}\|f_j(p)\|_{\ell_2}} + \beta_{i,j}/\rho \right|^2 \tag{13}$$

This subproblem can be solved using an alternating minimization procedure, in which u is updated while p is held fixed and vice versa. The alternating minimization procedure is summarized in Algorithm 2. Note that the subscript (m) denotes the value of the variable at the m-th iteration of the unconstrained subproblem, while the superscript (k) denotes the value of the variable at the k-th iteration of the outer loop defined in Algorithm 1. Consequently, $u_{(m)}^{(k)}$ denotes the value of u at the m-th inner iteration of the k-th outer iteration, and by convention, $u^{(k)} = u_{(m_k)}^{(k)}$, where m_k is the last iteration of the k-th subproblem. The u and p update steps displayed in Algorithm 2

Algorithm 2: Summary of the alternating minimization procedure for solving the Augmented Lagrangian subproblem of Eq.(13)

Given $p_{(0)}^{(k)}$, $u_{(0)}^{(k)}$, $\beta^{(k)}$, $\rho^{(k)}$;
for *m = 0,1,2,...* **do**
  Update u while holding p fixed

$$u_{(m+1)}^{(k)} = \underset{u}{\operatorname{argmin}}\ \mathcal{L}_{\mathcal{A}}\left(p_{(m)}^{(k)}, u, \beta^{(k)}; \rho^{(k)}\right)$$

  Update p while holding u fixed

$$p_{(m+1)}^{(k)} = \underset{p}{\operatorname{argmin}}\ \mathcal{L}_{\mathcal{A}}\left(p, u_{(m+1)}^{(k)}, \beta^{(k)}; \rho^{(k)}\right)$$

are described in detail in Appendix VI and Appendix C respectively.

APPENDIX B
u UPDATE STEP

In the u update step, we seek the minimizer of (13) with respect to u while keeping p, ρ, and β constant. By introducing the auxiliary variable $z_{i,j} = \frac{f_i(p)^H f_j(p)}{\|f_i(p)\|_{\ell_2}\|f_j(p)\|_{\ell_2}} - \beta_{i,j}/\rho$, this subproblem can be reduced to the proximal operator for the ℓ_∞-norm, i.e.:

$$\underset{u}{\text{minimize}}\ \ \|u\|_{\ell_\infty} + \frac{\rho}{2}\|u - z\|_{\ell_2}^2 \tag{16}$$

This proximal operator can be efficiently solved by introducing the following simplifications. First, we note that the optimal solution u^* to (16) has the same sign pattern as z, i.e. $\text{sign}(u^*) = \text{sign}(z)$. To demonstrate this, consider the following scenario. Suppose that an oracle has told us the ℓ_∞-norm of the optimal solution u^* to (16), which we denote by t. Given the value of t, minimizing (16) requires one to solve several scalar problems of the form:

$$\begin{aligned} \underset{u_i}{\text{minimize}} \quad & \frac{\rho}{2}|u_i - z_i|^2 \\ \text{subject to} \quad & |u_i| \le t \end{aligned} \tag{17}$$

Clearly, if $|z_i| \le t$, then the minimizer of (17) is necessarily $u_i^* = z_i$. If $|z_i| > t$, then one can express u_i and z_i in polar coordinates, i.e. $u_i = te^{j\phi_i}$ and $z_i = |z_i|e^{j\theta_i}$, and minimize the quadratic term $|te^{j\phi_i} - |z_i|e^{j\theta_i}|^2$. By expanding this term out

$$|te^{j\phi_i} - |z_i|e^{j\theta_i}|^2 = t^2 + |z_i|^2 - 2t|z_i|\cos(\phi_i - \theta_i) \tag{18}$$

we see that the solution is minimized if $\phi_i = \theta_i$, or equivalently, if $\text{sign}(u_i^*) = \text{sign}(z_i^*)$. As a result, (16) can be solved using strictly real and positive variables by making the change of variables $w = |z|$ and $q = |u|$.

With this change of variables, the problem can be recast as follows:

$$\begin{aligned} \underset{q}{\text{minimize}} \quad & \|q\|_{\ell_\infty} + \frac{\rho}{2}\|q - w\|_{\ell_2}^2 \\ \text{subject to} \quad & q \succeq 0 \end{aligned} \tag{19}$$

Algorithm 3: Summary of the procedure used to solve Eq. (16)

Compute $z_{i,j} = \frac{f_i(p)^H f_j(p)}{\|f_i(p)\|_{\ell_2}\|f_j(p)\|_{\ell_2}} - \beta_{i,j}/\rho$;
Compute $w = |z|$;
Sort w in descending order ;
Find t^* by finding the smallest value of m such that Eq. 27 is satisfied ;
Compute $q^* = \max(w, t^*)$;
Unsort q^* so that the elements are arranged in their original order ;
Compute $u^* = \text{diag}\left(\text{sign}(z)\right) q^*$;

where q and w are real vectors. In order to remove the non-smooth ℓ_∞-norm from the objective function, we can convert the objective into its epigraph form [22] by introducing the auxiliary variable t and modifying the constraints as follows:

$$\begin{aligned} \underset{q,t}{\text{minimize}} \quad & t + \frac{\rho}{2}\|q - w\|_{\ell_2}^2 \\ \text{subject to} \quad & -\ t \preceq q \preceq t \\ & t \ge 0 \end{aligned} \tag{20}$$

Considering that the elements of w are strictly positive, the constraints $t \ge 0$ and $-t \preceq q$ can be ignored, and so the simplified Lagrangian can be written in terms of a single dual variable α:

$$\mathcal{L}(q, t, \alpha) = t + \frac{\rho}{2}\|q - w\|_{\ell_2}^2 + \sum_i \alpha_i(q_i - t) \tag{21}$$

The Karush Kuhn Tucker (KKT) conditions [23] mandate that the following conditions are satisfied at the optimal point q^*, t^*, α^*:

$$q^* = w - \alpha^*/\rho \tag{22}$$

$$\sum_i \alpha_i^* = 1 \tag{23}$$

$$\alpha_i^*(q_i^* - t^*) = 0\ \forall i \tag{24}$$

$$\alpha^* \ge 0 \tag{25}$$

It is clear from (22) and (24) that the Lagrange multipliers satisfy the following relationship:

$$\alpha_i^* = \max(\rho(w_i - t^*), 0) \tag{26}$$

Therefore, one only needs to find the m non-zero elements of α where the constraints are active. Assuming that w is sorted in descending order, we can combine (26) and (23) to reveal the following condition:

$$t^* = -1/\rho + \sum_{i=1}^{m} w_i > w_{m+1} \tag{27}$$

One therefore needs only to find the smallest value of p that satisfies (27) in order to solve (20). In total, the steps for solving (16) are summarized in Algorithm 3.

APPENDIX C
p UPDATE

In the p update step, we seek the minimizer of (13) with respect to p while keeping u, ρ, and β constant. By introducing the auxiliary variable $z_{i,j} = u_{i,j} + \beta_{i,j}/\rho$, this subproblem can be reduced to the following non-convex optimization program:

$$\underset{p}{\text{minimize}}\; I_{Q_p}(p) + \sum_{1\le i\ne j\le N} \frac{\rho}{2}\left| \frac{f_i(p)^H f_j(p)}{\|f_i(p)\|_{\ell_2}\|f_j(p)\|_{\ell_2}} - z_{i,j}\right|^2 \tag{28}$$

The optimal solution to (28) is difficult to compute, even when $f_i(\cdot)$ takes a simple form. Instead, we consider an approximate update using the proximal gradient. Technically, the right-most term is not differentiable with respect to p when $f_i(p) = 0$ or $f_j(p) = 0$. However, this scenario should not occur in practical applications. Indeed, if a feasible value $p \in Q_p$ can lead to $f_i(p) = 0$ for some value of i, then the designer should seriously reconsider the choice of the design function and the feasible set. After-all, $f_i(p)$ computes the $i-$th column of the sensing matrix, and a column of all zeros is utterly useless for any sensing application, even when compressive sensing is not employed. Proceeding under the assumption that $\|f_i(p)\|_{\ell_2} > 0\,\forall i$, the gradient of the quadratic term in (28) can be computed as follows. First, the chain rule can be applied in order to compute the partial derivative $\frac{\partial \hat{f}_i(p)}{\partial p_l} = \frac{\partial}{\partial p_l}\left(\frac{f_i(p)}{\|f_i(p)\|_{\ell_2}}\right)$:

$$\begin{aligned}\frac{\partial \hat{f}_i(p)}{\partial p_l} &= \frac{\partial}{\partial p_l}\left(\frac{f_i(p)}{\|f_i(p)\|_{\ell_2}}\right)\\ &= \|f_i(p)\|_{\ell_2}^{-1}\left(I - \frac{f_i(p)f_i^H(p)}{\|f_i(p)\|_{\ell_2}^2}\right)\frac{\partial f_i(p)}{\partial p_l}\end{aligned} \tag{29}$$

Second, the chain rule can be applied once again in order to compute the partial derivative $\frac{\partial}{\partial p_l}(\hat{f}_i^H(p)\hat{f}_j(p))$:

$$\frac{\partial}{\partial p_l}\left(\hat{f}_i^H(p)\hat{f}_j(p)\right) = \frac{\partial \hat{f}_i^H(p)}{\partial p_l}\hat{f}_j(p) + \hat{f}_i^H(p)\frac{\partial \hat{f}_j(p)}{\partial p_l} \tag{30}$$

Finally, this result can be used in order to compute the l-th element of the gradient, which we denote by the vector g:

$$\begin{aligned}g_l = \sum_{1\le i\ne j\le N} &\rho\left(\frac{\partial}{\partial p_l}\left(\hat{f}_i^H(p)\hat{f}_j(p)\right)\right)\\ &\left(\frac{f_i(p)^H f_j(p)}{\|f_i(p)\|_{\ell_2}\|f_j(p)\|_{\ell_2}} - z_{i,j}\right)\end{aligned} \tag{31}$$

Given the gradient g, the proximal gradient method updates p by solving the following optimization program:

$$\underset{p}{\text{minimize}}\; I_{Q_p}(p) + \frac{1}{2\alpha}\|p - (z - \alpha g)\|_{\ell_2}^2 \tag{32}$$

where the step size α is computed using an inexact line search method. Our implementation of the design method uses the inexact proximal gradient line search method described in [24]. This line search method ensures that the objective function of (13) decreases on each iteration.

Equation (32) is simply the proximal operator for the feasible set Q_p, which will differ from problem to problem. For the design problem presented in Section IV-A, the proximal operator simply projects the updated sensor location onto the circle of radius $20\,\lambda$, i.e.:

$$\begin{aligned}\underset{r}{\text{minimize}}\;\; &\frac{1}{2\alpha}\|r - (z - \alpha g)\|_{\ell_2}^2\\ \text{subject to}\;\; &\|r\|_{\ell_2} = 20\,\lambda\end{aligned} \tag{33}$$

This problem has the following convenient closed-form solution:

$$r^* = \frac{20\,\lambda\,(z - \alpha g)}{\|\,(z - \alpha g)\,\|_{\ell_2}} \tag{34}$$

For the design problem presented in Section IV-B, the proximal operator simply projects the x and y co-ordinates onto the $5\,\lambda$ by $5\,\lambda$ box centered at the origin and enforces $z = 5\,\lambda$. Setting $h_l = (-2.5, -2.5, 5)^T\;\lambda$ and $h_u = (2.5, 2.5, 5)^T\;\lambda$, this proximal operator can be expressed as follows:

$$\begin{aligned}\underset{r}{\text{minimize}}\;\; &\frac{1}{2\alpha}\|r - (z - \alpha g)\|_{\ell_2}^2\\ \text{subject to}\;\; &h_l \preceq r \preceq h_u\end{aligned} \tag{35}$$

This problem has the following closed-form solution:

$$r^* = \max\left(\min\left(z - \alpha g, h_u\right), h_l\right) \tag{36}$$

where $\max$ and $\min$ operate on the individual vector components.

APPENDIX D
COMPLETE SUMMARY OF THE DESIGN ALGORITHM

A complete summary of the design algorithm is displayed in Algorithm 4.

Algorithm 4: Complete summary of the Augmented Lagrangian update procedure for the coherence minimization problem of Eq. (11)

Choose the initial values for $p^{(0)}$, $\rho^{(1)}$;
Set $u_{i,j}^{(0)} = \frac{f_i(p^{(0)})^H f_j(p^{(0)})}{\|f_i(p^{(0)})\|_{\ell_2}\|f_j(p^{(0)})\|_{\ell_2}}$, $\beta_{i,j}^{(1)} = 0$;
for *k = 1,2,3...* **do**
  Set $u_{(0)}^{(k)} = u^{(k-1)}$, $p_{(0)}^{(k)} = p^{(k-1)}$;
  for *m = 0,1,2,...* **do**
    Update u while holding p fixed

$$u_{(m+1)}^{(k)} = \underset{u}{\text{argmin}}\;\mathcal{L}_{\mathcal{A}}\left(p_{(m)}^{(k)}, u, \beta^{(k)}; \rho^{(k)}\right)$$

    Update p while holding u fixed

$$p_{(m+1)}^{(k)} = \underset{p}{\text{argmin}}\;\mathcal{L}_{\mathcal{A}}\left(p, u_{(m+1)}^{(k)}, \beta^{(k)}; \rho^{(k)}\right)$$

  Update the dual variables

$$\beta_{i,j}^{(k+1)} = \beta_{i,j}^{(k)} + \rho^{(k)}\left(u_{i,j}^{(k)} - \frac{f_i(p^{(k)})^H f_j(p^{(k)})}{\|f_i(p^{(k)})\|_{\ell_2}\|f_j(p^{(k)})\|_{\ell_2}}\right)$$

  Compute $\rho^{(k+1)}$ using the method described in [15]

REFERENCES

[1] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[2] E. J. Candes, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[3] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[4] M. D. Migliore and D. Pinchera, "Compressed sensing in electromagnetics: Theory, applications and perspectives," in *Proc. 5th Eur. Conf. Antennas Propag.*, 2011, pp. 1969–1973.

[5] Y. Rodriguez-Vaqueiro, Y. Alvarez-Lopez, B. Gonzalez-Valdes, J. Martinez-Lorenzo, F. Las-Heras, and C. Rappaport, "On the use of compressed sensing techniques for improving multistatic millimeter-wave portal-based personnel screening," *IEEE Trans. Antennas Propag.*, vol. 62, no. 1, pp. 494–499, 2014.

[6] A. Massa, P. Rocca, and G. Oliveri, "Compressive sensing in electromagnetics—A review," *IEEE Antennas Propag. Mag.*, vol. 57, no. 1, pp. 224–238, Feb. 2015.

[7] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.

[8] Z. Lin, C. Lu, and H. Li, "Optimized projections for compressed sensing via direct mutual coherence minimization," *arXiv:1508.03117*, 2015.

[9] Q. Bao, C. Jiang, Y. Lin, W. Tan, Z. Wang, and W. Hong, "Measurement matrix optimization and mismatch problem compensation for DLSLA 3-D SAR cross-track reconstruction," *Sensors*, vol. 16, no. 8, 2016, Art. no. Es1333.

[10] V. Abolghasemi, S. Ferdowsi, B. Makkiabadi, and S. Sanei, "On optimization of the measurement matrix for compressive sensing," in *Proc. 18th Eur. Signal Process. Conf.*, 2010, pp. 427–431.

[11] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, 2008.

[12] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[13] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Berlin, Germany: Springer-Verlag, vol. 3, no. 1, 2013.

[14] L. R. Welch, "Lower bounds on the maximum cross correlation of signals (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 3, pp. 397–399, May 1974.

[15] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed.. New York, NY, USA: Springer, 2006.

[16] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jul. 2011.

[17] Y. C. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. Conf. Rec. 27th Asilomar Conf. Signals, Syst. Comput.*, 1993, pp. 40–44.

[18] M. Grant and S. Boyd, "CVX: MATLAB software for disciplined convex programming, version 2.1," Mar. 2014. [Online]. Available: http://cvxr.com/cvx

[19] Y. Álvarez-López and J. A. Martínez-Lorenzo, "On the use of ultrasound-based technology for cargo inspection," *J. Vib. Acoust.*, vol. 138, no. 3, 2016, Art. no. 031009.

[20] J. A. Martinez Lorenzo, J. Heredia-Juesas, and W. Blackwell, "A single-transceiver compressive reflector antenna for high-sensing-capacity imaging," *IEEE Antennas Wireless Propag. Lett.*, vol. 15, pp. 968–971, 2015.

[21] J. Heredia-Juesas, G. Allan, A. Molaei, L. Tirado, W. Blackwell, and J. Martinez-Lorenzo, "Consensus-based imaging using ADMM for a compressive reflector antenna," in *IEEE Int. Symp. Antennas Propag. USNC/URSI National Radio Sci. Meet.*, Jul. 2015, pp. 1304–1305.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[23] H. Kuhn and A. Tucker, "Nonlinear programming" in *Proc. 2nd Berkeley Symp. Math. Statist. Probab*, 1950, pp. 481–492.

[24] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

Authors' photograph and biography not available at the time of publication.