Received August 28, 2021, accepted September 10, 2021, date of publication September 16, 2021,
date of current version September 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3113391

Matrix Element-Based Theory of Compressive Sensing and Its Application to Electromagnetic Imaging

PRATIK SHAH[1] AND MAHTA MOGHADDAM[2], (Fellow, IEEE)
[1]Acutus Medical Inc., Carlsbad, CA 92008, USA
[2]Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA
Corresponding author: Pratik Shah (pratik0828@gmail.com)

ABSTRACT We present a simple and general theory of compressive sensing (CS) that relies on elements of the sensing matrix rather than on the number of measurements. We prove the exact recovery using a dual certificate by showing that the sensing matrix satisfies an incoherence property and isotropy property if the sparsity level is kept lower than the reciprocal of the largest element of a matrix created from the sensing matrix. Unlike the CS literature, this unconventional approach does not require a linear relationship between the sparsity and the number of measurements and at the same time, can easily be evaluated. This adaptability captures anisotropic measurements appropriately as with anisotropic measurements, adding more measurements does not really imply that a signal with more nonzero elements will be recovered exactly. As an illustration, we demonstrate the theory's ability to accurately handle the anisotropic (Green's function-based sensing matrix) measurements and also its similarity to the existing CS literature for isotropic (Fourier) measurements. Further, we show the usefulness of the theory in comparing different sensing matrices and in generating dielectric images. The dielectric images are perfectly recovered even when there is only a single transmitter.

INDEX TERMS Compressed sensing, electromagnetic propagation, inverse problems, inverse scattering, sparse recovery.

I. INTRODUCTION

Many problems in science and engineering require solving an inverse problem, where parameters of interest, $\mathbf{x}$, are estimated from a set of linear measurements, $\mathbf{y}$. The set of measurements is not always orthogonal and sometimes the set is small. As such, the problem is ill-posed but if the desired parameters are known to be sparse (nonzero at only a few locations), it is possible to recover the parameters exactly through CS framework [1], [2], where $\mathbf{x}$ is estimated by solving the following ℓ_1 minimization problem:

$$\min \|\mathbf{x}\|_1, \quad s.t \ \|\mathbf{A}\,\mathbf{x} - \mathbf{y}\|_2^2. \tag{1}$$

To have exact recovery, the sensing matrix, $\mathbf{A}$ has to satisfy conditions derived broadly under the CS framework [1]–[12]. One of the most popular conditions in the CS framework is restricted isometry property(RIP) [2], which is satisfied with the Restricted Isometry Constant δ_s if

$$(1 - \delta_s)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1 + \delta_s)\|\mathbf{x}\|_2^2 \tag{2}$$

holds for all $\mathbf{x}$ with sparsity less than or equal to s. Proving RIP for a given matrix $\mathbf{A}$ is NP-hard [13]. However, the RIP has been satisfied for a variety of matrices, where $\mathbf{A}$ is created by randomly sampling i.i.d. entries from symmetric Bernoulli distribution [11], frequencies from Fourier Transform [1], i.i.d entries from a Gaussian distribution with zero mean [14], and from an orthogonal matrix [15]. For these matrices, it has been shown that if the number of measurements, m is about the order of $\mathcal{O}(s \log n)$, where n is the number of unknowns, the RIP holds. It is believed that the same order also requires to satisfy the RIP for a general matrix $\mathbf{A}$. Another popular condition is null space property (NSP), which is satisfied with a constant $\alpha_k < \frac{1}{2}$ if

$$\|\mathbf{v}_\mathbf{k}\|_1 \leq \alpha_k \ \|\mathbf{v}\|_1, \tag{3}$$

The associate editor coordinating the review of this manuscript and approving it for publication was Gerardo Di Martino.

where $\mathbf{v}$ is in the null space of $\mathbf{A}$ ($\mathbf{A}\,\mathbf{v} = 0$) and $\mathbf{v_k}$ denotes the k largest absolute values of $\mathbf{v}$. NSP depends only on the kernel of $\mathbf{A}$, provides sufficient and necessary conditions for sparse recovery, whereas RIP is only sufficient. Again, it is NP-hard to compute α_k [13]; therefore several relaxation approaches have been used to estimate bounds on α_k. For Gaussian distribution, the number of measurements is about the order of $\mathcal{O}(s \log n)$ [16].

To address the difficulty in verifying the RIP and extend the use of CS theory in noisy scenarios and for compressible signals, Candes *et al.* in [17] have developed a probabilistic approach. The approach relies on the sensing matrix confining local isotropy and incoherence (μ) properties that are shown to be valid for orthogonal matrices. This approach is extended to anisotropic measurements in [18] (anisotropic measurements are defined as $\mathbf{A_\Omega}^\mathsf{T}\mathbf{A_\Omega} \neq I$) with an additional condition that $\mathbf{A_\Omega}^\mathsf{T}\mathbf{A_\Omega}$ is invertible, where $\mathbf{A_\Omega}$ is a superset of measurements from which m measurements are chosen in $\mathbf{A}$. In both formulations, m is about the order of $\mathcal{O}(\mu\, s \log n)$ (For anisotropic case, $\mathcal{O}$ contains constant proportional to the conditional number of $\mathbf{A_\Omega}^\mathsf{T}\mathbf{A_\Omega}$). Another notable work on anisotropic measurements is carried out within the original RIP framework in [19], where the authors show m is at the same order $\mathcal{O}$ as others except it has an additional $\log^3$ term.

Electromagnetic imaging systems are used in biomedical and subsurface sensing/imaging [20]–[22], where usually Poisson equation or inhomogeneous Helmholtz equation is solved using Green's function. These systems do not possess a system matrix that has been evaluated for satisfying CS theory. In addition, the system has anisotropic measurements, and many times, the number of measurements is constrained by practical limitations such as accessibility, available technology, and cost. Under this scenario, evaluating if CS theory holds or not for a given system can not be examined by existing CS literature because either the conditions are NP-hard to evaluate or the conditions can only be applied when m is pretty big ($\approx$ 18000 in [18]), or more importantly the number of measurements, m, can not remain proportional to the sparsity, i.e. $\mathcal{O}(s)$. The rationale for the last point is that each additional anisotropic measurement may not provide the same level of information as the previous measurement. Therefore, the rate of required additional measurements can not remain linear with an increasing number of nonzeros in $\mathbf{x}$. In an extreme case, adding new measurements may not bring new information. For example, in electromagnetic imaging, electromagnetic fields scattered by scatterers in the far field region can be reconstructed accurately at a fixed radial distance using only a small set of samples/measurements [23]. Any additional measurement won't provide new information about the scatterers. Therefore, a higher number of measurements cannot provide an exact recovery with a higher number of nonzeros in $\mathbf{x}$.

To address the challenges described above, we propose a fundamentally different approach to establish the conditions for the exact recovery, which can efficiently incorporate the nonlinear relationship between sparsity and the number of measurements as well as can be applicable even when m is small. We establish a relationship between elements of the sensing matrix and sparsity level and prove that $\mathbf{x}$ can be recovered exactly given that its sparsity and dual variable are bounded from below. In other words, the theory can provide a lower estimate of maximum sparsity that can be recovered. The main advantage of our element-based approach over other CS theories is that it can represent anisotropy accurately and applies to any sensing matrix.

In this paper, we present the sensing matrix's elements-based sparse recovery theory and demonstrate the theory's usage for anisotropic measurements. This paper is organized as follows: in Section II, background and related work are summarized. In Section III, element based CS is introduced and corresponding theorems are proved. In Section IV, the theory is evaluated for isotropic and anisotropic measurements. In Section V, the numerical results are presented for anisotropic measurements. In Section VI, applications in the electromagnetic imaging regime are presented. Finally, we summarize the paper and discuss future work in Section VII.

A. NOTATION

We use the following notations throughout the paper. $\mathbf{x}$ is a column vector with n elements. $\mathbf{x_K}$ denotes a part of $\mathbf{x}$ containing entries indexed by K, where a subset $K \subset \{1, \ldots, n\}$. Similarly, for an $m \times n$ matrix $\mathbf{A}$, $\mathbf{A_K}$ denotes an $m \times |K|$ submatrix with column indices are in K. $\mathbf{A_{\{i\}}}$ is the i^{th} column of $\mathbf{A}$. $\|\mathbf{A}\|_2$ is the ℓ_2 norm of $\mathbf{A}$. $\|\mathbf{A}\|_1$ is the ℓ_1 norm of $\mathbf{A}$, which is defined as the maximum absolute column sum of $\mathbf{A}$. $\|\mathbf{A}\|_\infty$ is the ℓ_∞ norm of $\mathbf{A}$, which is defined as the maximum absolute row sum of $\mathbf{A}$. K^c denotes the complement of the set K. sgn(α) is the sign of α for $\alpha \neq 0$ and zero otherwise. For a vector, sgn($\mathbf{x}$) is an element based operation. $\mathbf{A}^\mathsf{T}$ denote the transpose of $\mathbf{A}$ if $\mathbf{A}$ is real and complex conjugate transpose if $\mathbf{A}$ is complex.

II. BACKGROUND AND RELATED WORK

We are interested in estimating $\mathbf{x}$ that has at most s nonzero (s-sparse) elements, for a given measurement vector $\mathbf{y}$ of length m ($m < n$), where $\mathbf{x}$ and $\mathbf{y}$ are related by a linear operator $\mathbf{A}$. Assuming that the rank of $\mathbf{A}$ is m, we investigate the conditions under which the following minimization problem,

$$\min \|\mathbf{x}\|_1, \quad s.t\ \mathbf{A}\,\mathbf{x} = \mathbf{y}, \tag{4}$$

can provide the exact recovery of $\mathbf{x}$.

We use relaxed isometry and incoherence properties of $\mathbf{A}$ and *dual certificate* approach to derive the conditions. The relaxed isometry property requires isometry property ($\mathbf{A}^\mathsf{T}\mathbf{A} = I$) holds only locally - only for the columns of $\mathbf{A}$, where $\mathbf{x}$ is nonzero. Let K be the set of locations of nonzero within $\mathbf{x}$, the relaxed isometry property is defined as

$$\|\mathbf{A_K}^\mathsf{T}\mathbf{A_K} - \mathbf{I}\|_2 \leq \delta, \tag{5}$$

where $\mathbf{I}$ is the identity matrix and δ is a constant. The incoherence property ensures that any column of $\mathbf{A}$ in K^c should

not be well approximated by any column of $\mathbf{A}$ in K, which requires

$$\max_{i\in K^c} \|\mathbf{A_K}^\mathsf{T}\mathbf{A}_{\{i\}}\|_2 \leq 1 \tag{6}$$

holds. In the *dual certificate* approach, the dual problem of the given primal problem (for example (4)) is formulated and a dual feasible point (dual variable) is found under certain conditions to certify that the primal variable, $\mathbf{x}$, is optimal. In other words, if *dual certificate* - the dual variable that certifies optimality of $\mathbf{x}$ exists, the recovery of $\mathbf{x}$ is exact.

In [17], the authors use relaxed isometry and incoherence properties to establish exact recovery even for nearly sparse signal in a probabilistic way using *dual certificate*. To get exact recovery with probability at least $1 - e^{-w}$, for any scalar $w \geq 1$, δ in (5) should be less than $1/2$ for noiseless measurements and less than $1/4$ for noisy measurements; a dual variable $\mathbf{v} \in \mathbb{R}^n$ should satisfy

$$\|\mathbf{v_K} - \text{sgn}(\mathbf{x_K})\|_2 \leq \frac{1}{4} \quad \text{and} \quad \|\mathbf{v_{K^c}}\|_\infty \leq \frac{1}{4}; \tag{7}$$

and m should be greater than $C\mu s \log n$, where C is a constant and μ is incoherence parameter ($\max_{i\in 1,..,n} |\mathbf{A}_{\{i\}}^\mathsf{T}\mathbf{A}_{\{i\}}|^2 \leq \mu$). The authors proved this using the Matrix Bernstein and Vector Bernstein inequalities in the golfing scheme.

For anisotropic measurements, [18] extends the results of [17] by using the same constraints with an additional assumption that the distribution has to be complete in the sense that $\mathbf{A}_\Omega{}^\mathsf{T}\mathbf{A}_\Omega$ is invertible. Under these conditions, the authors generalize the theoretical applicability of sparse recovery to any invertible linear system. For successful recovery the required number of measurements should be greater than $18044\,\kappa_s\,\mu s \log n$, where κ_s is the $s-$sparse condition number [18] of $\mathbf{A}_\Omega{}^\mathsf{T}\mathbf{A}_\Omega$.

III. ELEMENT BASED COMPRESSIVE SENSING

Element based CS is a generalized CS theory that can be used for isotropic as well as for anisotropic measurements. The theory has two main theorems: (1) exact recovery - the solution to the optimization problem (4) equals the unknown vector $\mathbf{x}$, and (2) under what conditions, the exact recovery is possible? We use dual certificate approach to show the exact recovery and then formulate conditions using $\|\mathbf{A}\|_\infty$ and $\|\mathbf{A}\|_1$ norms to show the existence of dual certificate. Since these norms operate at an individual element of the matrix, different than most commonly used $\|\mathbf{A}\|_2$ norm, we term our approach as Element based CS.

For the following theorems, we assume that $(\mathbf{A_K}^\mathsf{T}\mathbf{A_K})^{-1}$ exists and $\mathbf{A}$ is normalized. The normalization can be performed different ways: (1) most used approach in CS literature - divide each row by $\sqrt{m}$, (2) set $\|\mathbf{A_K}^\mathsf{T}\mathbf{A_K}\|_2$ close to c, (3) check a range of values for normalization and pick the one that maximize a certain property of $\mathbf{A}$.

Theorem 1 (Exact Recovery): Let's assume that $\|c\,\mathbf{I} - \mathbf{A_K}^\mathsf{T}\mathbf{A_K}\|_2 \leq \delta$, where $0 \leq \delta < \frac{c}{c+1}$, $c > 1$ and $\max_{i\in K^c} \|\mathbf{A_K}^\mathsf{T}\mathbf{A}_{\{i\}}\|_2 \leq 1$. Suppose there exists $\mathbf{v}$ in the row space of $\mathbf{A}$ obeying $\mathbf{v_K} = c\,\text{sgn}(\mathbf{x_K})$ and $\|\mathbf{v_{K^c}}\|_\infty < 1$, then $\mathbf{x}$ is the unique ℓ_1 minimizer to (4).

This theorem states that $\mathbf{x}$ can be recovered exactly, if dual certificate $\mathbf{v}$ satisfies certain conditions and $\mathbf{A_K}$ holds certain norm constraints. The proof of the theorem is given in Appendix A.

Next, we will prove that such dual certificate exists. To prove this, we use element entries of a matrix. Let $\mathbf{B} := c\,\mathbf{I} - \mathbf{A}^\mathsf{T}\mathbf{A}$ and $|B_{pq}| \leq \zeta,\ 1 \leq p, q \leq n$, where ζ is defined as a value that is higher than or equal to all elements of the matrix $\mathbf{B}$. In other words, the elements of $\mathbf{B}$ are bounded above by ζ. The normalization step on $\mathbf{A}$ ensures that $\zeta \leq 1$.

Theorem 2 (Existence of a Dual Certificate): If $s \leq \frac{\delta}{\zeta}$ then the constraints on $\mathbf{A}$ in Theorem 1 hold and a vector $\mathbf{v}$ with the properties required for Theorem 1 exists.

The proof of the theorem is detailed in Appendix B.

Comparing these conditions to the conditions of the CS, the following key differences are observed.

1) Element based CS conditions can be used to evaluate any system matrix, whereas except for [18], none of the other CS theories can be used to evaluate an anisotropic measurements system because either it is NP-hard or the assumptions do not hold.
2) We do not establish a direct relation between m and s because it can change depending on the type of the measurements. This varying relationship can be captured by ζ. We will discuss it in the next section. For the CS theory, it is always kept linear even for the anisotropic measurements such as in [18].
3) Our approach is deterministic in the sense that it works for all sparse vectors, whereas this is not the case for [17] and [18].
4) Evaluating the conditions for Element based CS on any system matrix require (a) construct the system matrix, (b) choose a value for c, (c) select value of δ between 0 and $\frac{c}{c+1}$, (d) normalize $\mathbf{A}$, and finally (e) calculate ζ and then s. This is different than using conventional CS formulation: $m = \mathcal{O}(s \log n)$, where s is estimated using m and n even though for many systems the multiplication factor in $\mathcal{O}$ notation is unknown or is too big.

It should be clear that (1) we have not optimized the constraints in the theory to make them optimal, i.e, the bounds on conditions are not tight, and (2) the conditions do not give universal results in the sense that they are applicable only to the given set of measurements or the given sensing matrix but not to all possible choices of the sensing matrix for the given measurement system. For example, if the conditions hold for a given set of Fourier measurements, the exact recovery is guaranteed for a system matrix that uses measurements only at those frequencies, not for arbitrarily selected frequency measurements.

A. SELECTING VALUE FOR CONSTANTS

The value for the constants, δ and c are selected such that all conditions in Theorems are satisfied. Since $s = \frac{\delta}{\zeta}$, setting the

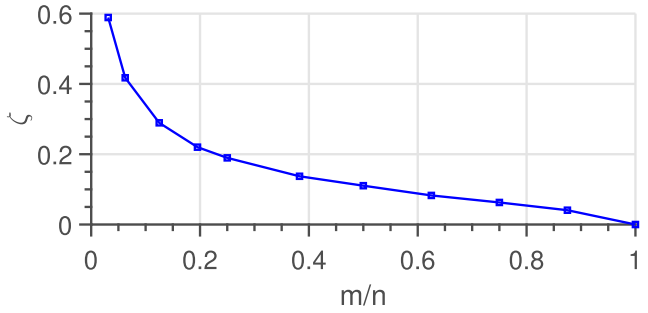

FIGURE 1. Isotropic measurements: effect of a number of measurements (m) on ζ (inversely proportional to sparsity) for a fixed number of unknowns (n).

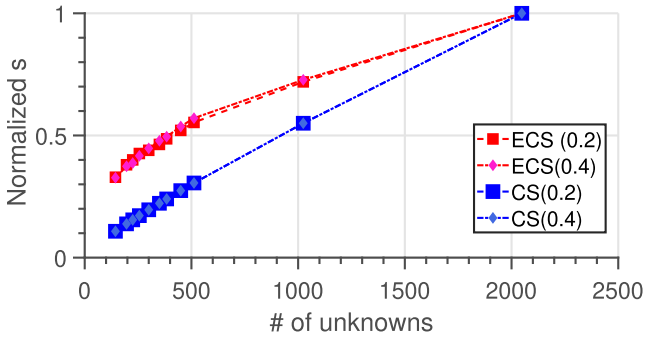

FIGURE 2. Comparative analysis for isotropic measurements: for two set of measurements- 0.2n and 0.4n, normalized sparsity (s) is evaluated against number of unknowns for Element based CS (ECS) and CS literature (CS).

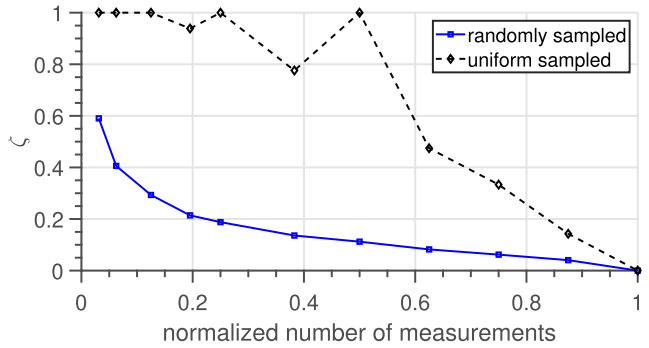

FIGURE 3. Isotropic measurements: ζ for randomly sampled measurements and uniformly sampled measurements across a set of measurements. The number of measurements are normalized to the number of unknowns.

higher value of δ gives larger recoverable sparsity. However, δ cannot be higher than $\frac{c}{c+1}$. So, we can set δ slightly smaller than the upper bound, such as $0.999(\frac{c}{c+1})$.

Constant c appears in the definition of $\mathbf{B}$. Larger c can increase ζ value through higher value in either off-diagonal elements of $\mathbf{B}$ or diagonal elements of $\mathbf{B}$ depending on the normalization step used for $\mathbf{A}$. Therefore, c cannot be set very high and can be set between 1 and 1.5. The optimal c value that can provide maximum recoverable sparsity can be obtained by brute-force search within this range.

IV. ANALYSIS OF CANONICAL CASES

The conditions presented in Theorem 2 do not have direct relation between the number of measurements, m, and the number of unknowns, n, rather it has ζ that is derived from $\mathbf{A}$. To evaluate the effectiveness of ζ in capturing the changes in measurements, we will discuss both types of measurements setup in canonical cases: A. isotropic measurements and B. anisotropic measurements.

A. ISOTROPIC MEASUREMENTS

The isotropic measurements are carried out synthetically in the Discrete Fourier Transform (DFT) domain to assess the relationship of ζ to the number of measurements m and the number of unknowns n. The DFT matrix is normalized by dividing each row by $\sqrt{m}$. For the relationship between ζ and m, we randomly select frequency components (measurements), keep n constant ($n = 512$), vary the number of measurements (m) from 30 to 512, and calculate ζ for each m. We repeat the same experiment 50 times. The resultant average value of ζ is plotted against m/n as shown in Fig. 1. As can be seen, the value of ζ monotonically decreases as m increases. This behavior is inline with the literature on CS because ζ is inversely proposal to the sparsity, and in the literature, sparsity increases with increasing m. Also, note that for $m = n$, ζ goes to 0, which indicates that the exact recovery of the largest sparse object ($s = n$) is guaranteed.

For the relation between ζ and the number of unknowns, we set m as a fraction of the number of unknowns and vary the number of unknowns from 256 to 2048. For each value of n, sparsity is calculated using Eq. (24) and the formula used in CS ($m/\log(n)$) for $m = 0.2n$ and $m = 0.4n$. Sparsity, normalized by its maximum value is plotted in Fig. 2. As can be seen, Element based CS has a monotonically increasing relation between sparsity and the number of unknowns, similar to the CS literature. It is also interesting to note that for the Element based CS, even though s is estimated from two different DFT matrices for two different value m, the rate of the change in sparsity remains the same, comparably to the behavior of CS.

Next, we evaluate the effect of random measurements, the fundamental component of the CS theory, on ζ. We compare randomly selected measurements to the uniformly sampled measurements. The estimated value of ζ for each value of m is shown in Fig. 3. For all values of m, ζ remains higher for uniformly sampled measurements than for randomly selected measurements. These higher values of ζs indicate that the uniform sampling performs worst than randomly sampling for the DFT matrix. This observation is in harmony with preference of random measurements in the CS theory.

B. ANISOTROPIC MEASUREMENTS

For anisotropic measurements, a two-dimensional electromagnetic imaging problem is considered, where a single transmitting antenna is used to excite a region of interest and multiple receiving antenna are used to collect the reflected and scattered waves. This imaging problem has an interesting fact that only a handful number of measurements are needed to capture the scattered waves at an arbitrary but fixed radial

distance [23]. We investigate below whether ζ can capture this bandlimited nature or not.

The electric field $E_s(\mathbf{r})$ radiated by the current source $J(\mathbf{r}')$ lying in region of interest, D, at a position $\mathbf{r}'$ can be written through Green's function $G(\mathbf{r}, \mathbf{r}')$ integral formula as

$$E_s(\mathbf{r}) = k_b^2 \int_D G(\mathbf{r}, \mathbf{r}') J(\mathbf{r}')\, dr', \quad \mathbf{r} \in S,\ \mathbf{r}' \in D, \tag{8}$$

where k_b is the background wavenumber, $\mathbf{r}$ is the position of a receiver, and S is the region outside of D. In a discrete form, $\mathbf{b} = c\,\mathbf{G}\,\mathbf{z}$, where $\mathbf{b}$ contains m measurements, an element of $\mathbf{G}$, $\mathbf{G}(\mathbf{r}, \mathbf{r}') = H_0^{(2)}(k_b|\mathbf{r} - \mathbf{r}'|)$, $\mathbf{z}$ contains n discretized points of D, and c is constant. $H_0^{(2)}$ is the 0^{th} order Hankel function of the second kind. Clearly, $\mathbf{G}$ does not comprise of isotropic measurements.

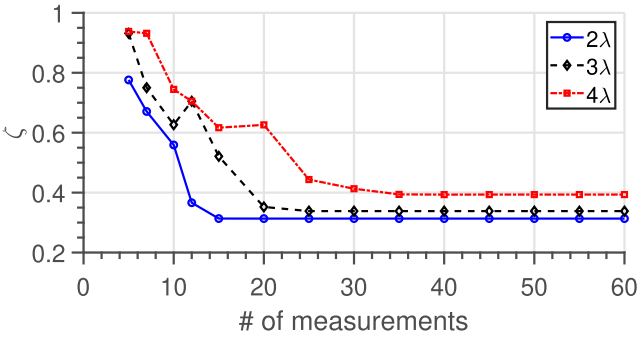

FIGURE 4. Anisotropic measurements: effect of a number of measurements (m) on ζ for three different size of region of interest (n). (compare it to the results for isotropic measurements in (Fig. 1).

Let's consider a typical imaging setup and evaluate the value of ζ in the same manner as we did for the isotropic measurements. For the relation between ζ and m, we vary the number of measurements from 5 to 60, which are sampled uniformly at a 3λ radial distance and evaluate ζ for the square region, D, having width - 2λ, 3λ, and 4λ, where λ is the wavelength in the background medium. $\mathbf{G}$ is normalized by $\sqrt{m}$. The results are reported in Fig. 4 in terms of ζ as a function of m for three different width of D. It shows that ζ decreases until a certain number of measurements and then it becomes constant; i.e. additional measurements do not provide new information. This ζ response is totally different than the ζ response for the isotropic measurements but perfectly in line with the existing theory on the bandlimited nature of scattered fields.

ζ's constant response for higher value of m is not specific to these examples. In fact, it can also be derived analytically as follows.

An element of $\mathbf{F} := \frac{1}{m}\mathbf{G}^\intercal\mathbf{G}$, F_{pq} is expressed as

$$\begin{aligned} F_{pq} &= \frac{1}{m}\mathbf{G}_{\{p\}}{}^\intercal\mathbf{G}_{\{q\}} \\ &= \frac{1}{m}\sum_{l=1}^{m} H_0^{(2)}(k_b|\mathbf{r_l} - \mathbf{r'_p}|)^\intercal H_0^{(2)}(k_b|\mathbf{r_l} - \mathbf{r'_q}|). \end{aligned} \tag{9}$$

Applying the far-field approximation [24], we get

$$F_{pq} \approx \frac{2}{\pi k_b m}\sum_{l=1}^{m}\frac{1}{r_l}\exp\left[jk_b\left(r'_q\cos\left(\theta_{lq}\right) - r'_p\cos\left(\theta_{lp}\right)\right)\right], \tag{10}$$

where $r_{\{.\}}$ is the magnitude of the quantity and $\theta_{\{.\}}$ is the angle between measurement point $\mathbf{r_l}$ and a point in D. Using the Bessel function property [25] and with some arithmetic, we get,

$$\begin{aligned} F_{pq} \approx \frac{2}{\pi k_b m}\sum_{s=-\infty}^{\infty}\sum_{t=-\infty}^{\infty}\sum_{l=1}^{m} j^s J_s(k_b r'_q)(-1)^t j^t J_t(k_b r'_p) \\ \times \frac{1}{r_l}\exp\left[j(s\theta_{lq} - t\theta_{lp})\right]. \end{aligned} \tag{11}$$

Scattered fields on a constant radius, $r_l = a$, can be optimally sampled using uniform sampling scheme [26]. Using uniform sampling for measurements, each angle can be calculated as $\theta_{lp} = \frac{2\pi m_l}{m} - \theta_p$, $m_l \in \{0, 1, .., m\}$, assuming the center of the region of interest is the reference point (origin of the coordinate system). Incorporating these, we get

$$\begin{aligned} F_{pq} \approx \frac{2}{\pi k_b m a}\sum_{s=-\infty}^{\infty}\sum_{t=-\infty}^{\infty}\sum_{l=1}^{m} j^s J_s(k_b r'_q)(-1)^t j^t J_t(k_b r'_p) \\ \times \exp\left[j\left(\frac{2\pi m_l}{m}(s-t) + (t\theta_p - s\theta_q)\right)\right]. \end{aligned} \tag{12}$$

Using delta function formula, we get

$$\begin{aligned} F_{pq} \approx \frac{2}{\pi k_b a}\sum_{s=-\infty}^{\infty}\sum_{t=-\infty}^{\infty} j^s J_s(k_b r'_q)(-1)^t j^t J_t(k_b r'_p) \\ \times \exp\left[j(t\theta_p - s\theta_q)\right]\sum_{v=-\infty}^{\infty}\delta[s - t + vm]]. \end{aligned} \tag{13}$$

It can be further simplified as

$$\begin{aligned} F_{pq} &\approx \frac{2}{\pi k_b a}\sum_{v,t=-\infty}^{\infty} J_{t-vm}(k_b r'_q) J_t(k_b r'_p) e^{j\phi}, \\ \text{where } \phi &= (t\theta_p - (t - vm)\theta_q - \pi vm/2). \end{aligned} \tag{14}$$

Using $J_{-n}(x) = (-1)^n J_n(x)$, we get

$$\begin{aligned} F_{pq} \approx \frac{2}{\pi k_b a}\sum_{v,t=0}^{\infty}\Big[J_{t-vm}(k_b r'_q)\left(e^{j\phi} + e^{-j\phi}(-1)^{vm}\right) \\ + J_{t+vm}(k_b r'_q)\left(e^{j\phi_1} + e^{-j\phi_1}(-1)^{vm}\right)\Big] J_t(k_b r'_p), \end{aligned} \tag{15}$$

where $\phi_1 = (t\theta_p - (t + vm)\theta_q + \pi vm/2)$.
Now setting

$$\zeta = \max_{1\le p,\, q\le n}|I - F_{pq}|, \tag{16}$$

where $I = 1$ if $p = q$, else $I = 0$, we get the equation, which shows that ζ is the sum of the order of the Bessel functions. Bessel function has a property that $J_n(x) \approx 0$ if $|n| > 2\lceil x\rceil$ [27]. Applying this condition on the order of three Bessel functions, we get $F_{pq} \not\approx 0$, only if

$$0 \le t < 2\lceil k_b r'_m\rceil \quad \&$$

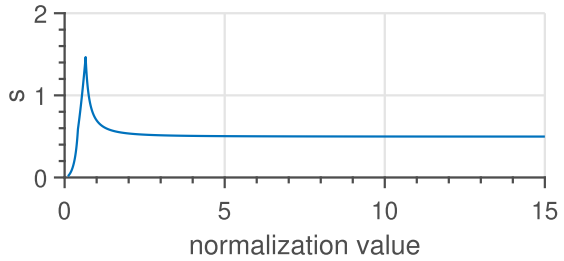

FIGURE 5. A normalization step: evaluating recoverable sparsity s for a range of values and picking one value as the normalization value that gives highest recoverable sparsity.

$$\frac{-t}{m} \leq v < \frac{2\lceil k_b r'_m \rceil - t}{m} \quad \text{or}$$
$$\frac{t - 2\lceil k_b r'_m \rceil}{m} < v \leq \frac{t}{m}, \tag{17}$$

where r'_m is the maximum magnitude of $r'_{\{.\}}$. As m increases, $\frac{t-2\lceil k_b r'_m \rceil}{m}$ and $\frac{2\lceil k_b r'_m \rceil - t}{m}$ become smaller and range of values that v can take decreases. After $m > 2\lceil k_b r'_m \rceil$, only possible value that v can have is 0. Plugging this value of v gives

$$\zeta = \max_{1\leq p,\, q\leq n} \left| I - \frac{8}{\pi k_b a} \sum_{t=0}^{2\lceil k_b r'_m \rceil} J_t(k_b r'_q) J_t(k_b r'_p) e^{jt(\theta_p - \theta_q)} \right|, \tag{18}$$

which shows that ζ is independent of m and the location of receivers for all $m > 2\lceil k_b r'_m \rceil$. Fig. 4 highlights the same point - the flatness of the curve happens at $m \approx 2\lceil 2\sqrt{2}\pi\, b \rceil$, where b is the half-width of the square region. In other words, any additional measurement above this m does not provide additional information. It can be reconstructed from the existed measurements. It is interesting to note that the condition on m obtained here matches with the value in [23], while also providing a link between measurements and recovery - a sparse object that can be recovered exactly.

V. NUMERICAL RESULTS

In the previous section, we showed that the element based CS theory captures additions of new measurements correctly for both isotropic and anisotropic measurements. In this section, we will evaluate the applicability of the Element based CS conditions in performing sparse recovery.

We consider a typical imaging setup, where receiving antennas are located on the circumference of a circular domain of 4λ diameter and the investigation domain is within a square of 2λ length with pixel size $\frac{\lambda}{4}$. We normalize $\mathbf{G}$ by the value that gives the minimum ζ value, i.e. the maximum sparsity that can be recovered. Fig. 5 shows an example for $m = 2\lceil 2\sqrt{2}\pi \rceil = 18$ that 0.656 can be used as the normalization value as this value provides the maximum sparsity, 1.47, that can be recovered. Since this is a small size problem, we have numerically validated that when $s < 1.47$, conditions on $\mathbf{A}$ as well as $\mathbf{v}$ hold.

Next, we perform sparse recovery for $\mathbf{z}$ that can have sparsity between $s = 1$ and $s = 6$. At each sparsity value, 50 samples of $\mathbf{z}$ are created. Out of 50 samples, across all sparsity value, 15 samples have fixed but arbitrary contrast value and nonzero positions, 10 samples have fixed but arbitrary only contrast value, and 15 samples have only fixed positions. The remaining 10 samples have random positions and random contrast values. The number of measurements used in this numerical experiments is 8, 10, 12, 15, 20, and 40. For these measurements, the conditions in theorems are validated and using the above normalization step, corresponding maximum value for the sparsity is 1.1, 1.1, 1.13, 1.46, 1.47, and 1.47. The sparse recovery is obtained using SPGL software [28], which uses Spectral Projected Gradient algorithm and is applicable for a complex valued problem. The results are evaluated using the weighted absolute percentage error (WAPE) metric, where WAPE is defined as the sum of absolute differences divided by the sum of actual values (WAPE $= \frac{\sum_{i=0}^{i=n} |x_i - \hat{x}_i|}{/} \sum_{i=0}^{i=n} |x_i|$). If the WAPE value is less than 10^{-3}, we consider it as a successful recovery. For each sparsity value, the recovery rate for different values of m is shown in Fig. 6.

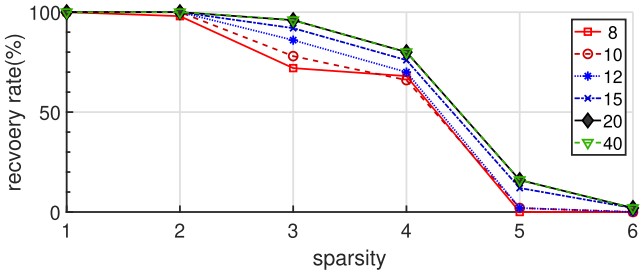

FIGURE 6. Sparse recovery for 50 samples at each sparsity value for different number of measurements. Recovery rate 100% shows that all 50 samples are successfully recovered.

As can be seen, the recovery rate is 100% for all sparsity values lower than estimated maximum sparsity value for a given m and it decreases as s increases. For $s = 6$, the recovery rate is 0% for all measurements. The recovery rates increases as m increases from 8 to 20. Above 20, the recovery rate remains identical, as we expect for all $m > 18$ for this setup.

VI. APPLICATIONS

In this section, we present applications of the Element based CS to Electromagnetic imaging. Specifically, we discuss comparing different sensing matrices and retrieving dielectric images from electromagnetic imaging.

A. COMPARING SENSING MATRICES

Let's assume we have multiple choices for a sensing matrix and we are interested in resolving which sensing matrix can provide accurate sparse signal recovery with the same number of measurements. This scenario can occur when there is flexibility in placing measurement recorders or when more than one type of sensing matrices are available. For flexibility in recorders' placement, we can optimize the location of recorders such that ζ is minimum. This problem is typically known as maximizing the sensing capacity and it is out of the scope for this paper. Instead, we focus on two types of the sensing matrix: (1) matrix defined by Green's function

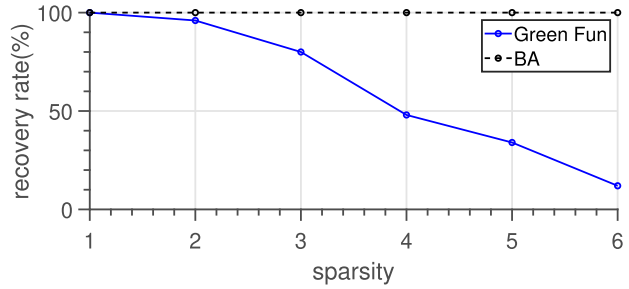

FIGURE 7. Comparing sensing matrices: sparse recovery for 50 samples at each sparsity value for two sensing matrices. In legend, BA stands for the sensing matrix defined by the Born approximation of (8) with two transmitters; and Green Fun stands for the Green function based sensing matrix.

integral formula (8) and (2) Born approximation of (8) [29]. In Born approximation, the total field is approximated by the incident field. The total field becomes a part of the integral because the current source is expressed as the product of complex permittivity and the total field.

For the simulation setup, we use the same setup that we used in the previous section ($m = 18$) and the system is configured for two different sensing matrices - (1) the Green's function, and (2) the Born approximation with two transmitters. Both systems are evaluated for $s = 1$ and $s = 6$ for 50 samples having random contrast values and locations at each sparsity level. For both types of measurements, we use same c and δ, which gives $\zeta = 0.34$ for Green's function based system and 0.27 for Born approximation based system. Based on the ζ value, we expect that the sensing matrix based on Born approximation will have a better recovery rate. Fig. 7 shows the recovery rate for both sensing matrices across six sparsity level. As can be seen, the recovery rate remains 100% for all sparsity level for Born approximation based system, while Green's function based recovery rate drops significantly and reduces to 12% for $s = 6$. It can be inferred from this simulation study that ζ value can be used to do comparative analysis between two systems and lower ζ value indicates better recovery rate.

B. GENERATING DIELECTRIC IMAGES

Generating dielectric images accurately from scattered electric fields has applications in through-the-wall imaging [30] or medical therapy/monitoring [31], [32]. This problem is addressed many times in literature and also when the images are sparse. For sparse imaging, the problem is formulated in various forms and a large set of solvers with different characteristics are used to improve recovery. Details can be in found in review articles written by Oliveri *et al.* [33] and Massa *et al.* [34]. In this section, we present an efficient process to generate sparse dielectric images. This process is particularly valuable, when there are practical only a limited number of sources are practically feasible, and fast detection or identification is one of the requirements.

We propose to use a two steps process to generate dielectric images. In the first step, we estimate current sources from the scattered electric fields similarly to our method described in

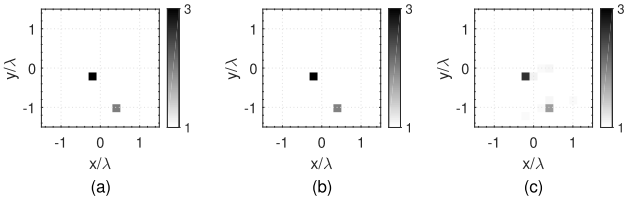

FIGURE 8. (a) true dielectric 2 points object; (b) reconstructed object, where reconstruction and measurement grids are at the same resolution (λ/5); (c) reconstructed object, where reconstruction and measurement grids are at the different resolution (λ/5 vs λ/10). Colorbar shows relative permittivity.

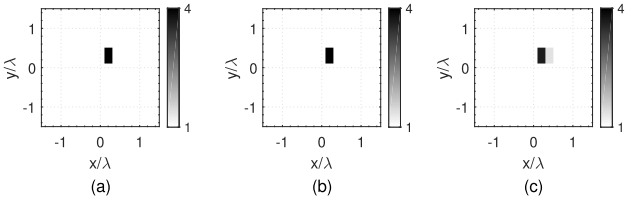

FIGURE 9. (a) true dielectric box object; (b) reconstructed object, where reconstruction and measurement grids are at the same resolution (λ/5); (c) reconstructed object, where reconstruction and measurement grids are at the different resolution (λ/5 vs λ/10). Colorbar shows relative permittivity.

Section IV-B and Section V. In the second step, the dielectric values, $\mathbf{x}$, are estimated from the current sources, $\mathbf{z}$, using the following equation:

$$x_i = \frac{z_i E_i^{\mathsf{T}}}{\|E_i\|^2}, \quad \forall i \in \{1, .., n\}, \tag{19}$$

where E_i is the electric field inside the domain. It is calculated using the following discretized version of the electric field volume integral equation [29]:

$$E_i = E_{in,i} + \sum_{j=1}^{n} C_{ij}\, G_{ij}\, z_j, \quad \forall i \in \{1, .., n\}, \tag{20}$$

where $E_{in,i}$ is the incident electric field, $E_i\, x_i = z_i$, and C_{ij} is a constant estimated similarly to [35]. Since z_i is already estimated, (20) does not require any matrix inversion and (19) requires only the element wise operation (Hadamard product). Thus, this step is computationally very efficient. In addition, getting dielectric values do not require an iterative approach such as Born Iterative Method [36] or an approximation such as Born approximation, rather they are estimated directly without any approximation.

To test the method, we place a single source (frequency = 20 MHz) and 30 receiving antennas uniformly distributed on the circumference of a circular domain of radius 2.5λ and the investigation domain is within a square of 3λ length. The background medium is considered to be air. Two objects that are electrically larger than required by Born approximation ([29]) are considered here - the 2 points object and the box object (see Fig. 8(a) and Fig. 9(a)). These objects are reconstructed at λ/5 spatial resolution, while "measured" scattered field is estimated at both λ/5 and λ/10. Corresponding reconstructed images are shown in Fig. 8 and Fig. 9. As can be seen, objects reconstructed in two different scenarios are visually very similar to the true object. When

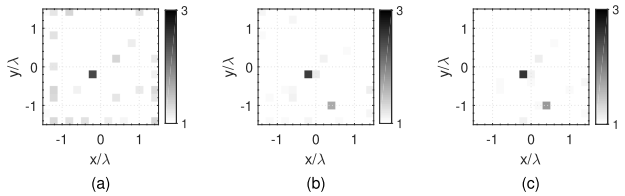

FIGURE 10. Reconstructed the 2 points object at different SNRs. (a) 10 dB; (b) 20 dB; (c) 30 dB. Colorbar shows relative permittivity.

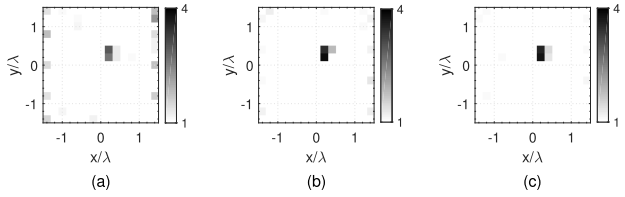

FIGURE 11. Reconstructed the box object at different SNRs. (a) 10 dB; (b) 20 dB; (c) 30 dB. Colorbar shows relative permittivity.

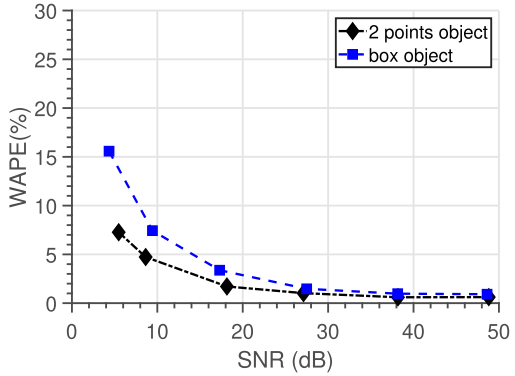

FIGURE 12. Noise analysis of two different objects reconstructed at the different resolution $\lambda/5$ than measurement grid resolution $\lambda/10$.

'inverse crime' (reconstruction and measurement grids are the same) is committed, the reconstruction results are perfect. Quantitatively, for the 2 points WAPE is $0.25 \times 10^{-7}\%$ when the grids are the same and 0.61%when they are different. For the box object, WAPE is $0.49 \times 10^{-7}\%$ when the grids are the same and 0.88% when they are different.

We have shown that sparse recovery with a single transmitter is possible theoretically and numerically. Our results do not contradict with the well-accepted point that the multiple solutions exist for (8) as the point is pertinent to any number of nonzeros in $\mathbf{z}$ (see [37], [38]), whereas we assume to have limited number of nonzeros in $\mathbf{z}$. We actually show exact recovery for the subset of $\mathbf{z}$ is feasible, which was never investigated separately and as a part of the full set of $\mathbf{z}$, the subset was thought to have multiple solutions or did not have the appropriate physical solution [39].

Next, we evaluate the sparse recovery under measurement noise, where we add random Gaussian noise to the measurements. We perform a sequence of simulations by varying SNR from 5 dB to 50 dB for both type of objects mentioned above. All parameters are kept the same and the objects are reconstructed at $\lambda/5$ resolution, whereas measurements are simulated at $\lambda/10$ resolution. The recovered objects at 10 dB, 20 dB, and 30 dB SNRs are shown in Fig. 10 and Fig. 11 for the 2 points object and the box object, respectively. WAPEs for all SNRs are shown in Fig. 12 for both type of objects. As can be seen, WAPE remains $< 5\%$ for SNR $> 20\%$.

VII. CONCLUSION

We presented Element based CS theory that can be applicable to any sensing matrix. We showed that the sparse recovery would be exact, if:

1) $(\mathbf{A_K}^{\mathsf{T}}\mathbf{A_K})^{-1}$ exists, and
2) $s \leq \frac{\delta}{\zeta}$.

We evaluated this theory on two types of measurements. (1) Fourier measurements as isotropic measurements and (2) anisotropic measurements using Green's function.

For Fourier measurements, the Element based CS theory showed similar dependencies of sparsity on the number of unknowns and the number of measurements as they have in the published literature of CS. For measurements involving Green's function, the Element based CS theory showed the nonlinear relationship between sparsity and number of measurements and also an upper bound on the usable number of measurements, whereas according to the CS literature, the relationship is linear or it cannot be evaluated. The upper bound on the usable number of measurements exists in the electromagnetic imaging domain but here, we were able to incorporate it with the recovery problem.

Next, we demonstrated that in addition to providing a theory for exact recovery, the Element based CS theory had applications in generating dielectric images and in comparing different measurement systems. Although the analysis and results were for Green's function based formulation, similar steps can be carried out for any other sensing matrix. For generating dielectric images, the following aspects were inferred.

1) The recovery process is very simple and does not need any forward problem's matrix inversion. Hence, it has very low computational complexity.
2) It is possible to have the exact recovery for very sparse objects. The finding rectifies the conventional understanding of the existence of non-unique solutions for the single-source setup in the inverse scattering domain.
3) The finding was numerically validated, which showed that the dielectric values were recovered perfectly when 'inverse crime' was committed and had relative low error under both noisy and noiseless scenarios, when 'inverse crime' was avoided.

The future work includes extending the theory when $\mathbf{x}$ is not exactly sparse and the measurement set is noisy. The other direction that we will further study is how ζ can help maximizing the sensing capacity, where the sensing matrix is designed by optimizing sensors' locations and the system's other variables.

APPENDIX A
PROOF OF THEOREM 1

The proof of the theorem is as follows. It is by contradiction similar to the proof in [17].

Proof: Let's assume that $\hat{\mathbf{x}} = \mathbf{x} + \mathbf{h}$ be the solution to (4), which states that $\|\hat{\mathbf{x}}\|_1 \leq \|\mathbf{x}\|_1$ and $\mathbf{A}\,\mathbf{h} = \mathbf{0}$ since both $\mathbf{x}$ and $\hat{\mathbf{x}}$ are feasible.

Now, $\|\hat{\mathbf{x}}\|_1$ can be expanded in the sets K and K^c as,

$$\begin{aligned} c\,\|\hat{\mathbf{x}}\|_1 &= c\,\|\mathbf{x}_{\mathbf{K}} + \mathbf{h}_{\mathbf{K}}\|_1 + c\,\|\mathbf{h}_{\mathbf{K}^c}\|_1 \\ &\geq c\,\|\mathbf{x}_{\mathbf{K}}\|_1 + \langle c\,\mathrm{sgn}(\mathbf{x}_{\mathbf{K}}), \mathbf{h}_{\mathbf{K}}\rangle + c\,\|\mathbf{h}_{\mathbf{K}^c}\|_1. \end{aligned} \tag{21}$$

Next, $|\langle c\,\mathrm{sgn}(\mathbf{x}_{\mathbf{K}}), \mathbf{h}_{\mathbf{K}}\rangle|$ can be bounded above as

$$\begin{aligned} |\langle c\,\mathrm{sgn}(\mathbf{x}_{\mathbf{K}}), \mathbf{h}_{\mathbf{K}}\rangle| &= |\langle c\,\mathrm{sgn}(\mathbf{x}_{\mathbf{K}}) - \mathbf{v}_{\mathbf{K}}, \mathbf{h}_{\mathbf{K}}\rangle + \langle \mathbf{v}_{\mathbf{K}}, \mathbf{h}_{\mathbf{K}}\rangle| \\ &= |\langle c\,\mathrm{sgn}(\mathbf{x}_{\mathbf{K}}) - \mathbf{v}_{\mathbf{K}}, \mathbf{h}_{\mathbf{K}}\rangle - \langle \mathbf{v}_{\mathbf{K}^c}, \mathbf{h}_{\mathbf{K}^c}\rangle| \\ &= |-\langle \mathbf{v}_{\mathbf{K}^c}, \mathbf{h}_{\mathbf{K}^c}\rangle| \\ &\leq \|\mathbf{v}_{\mathbf{K}^c}\|_\infty \|\mathbf{h}_{\mathbf{K}^c}\|_1 \;<\; \|\mathbf{h}_{\mathbf{K}^c}\|_1. \end{aligned} \tag{22}$$

Here we used that $\mathbf{v}_{\mathbf{K}} = c\,\mathrm{sgn}(\mathbf{x}_{\mathbf{K}})$ and $\mathbf{v} = \mathbf{A}^\intercal\mathbf{w}$ is a dual vector. For the dual vector, we get $\langle \mathbf{v}, \mathbf{h}\rangle = \langle \mathbf{w}, \mathbf{A}\mathbf{h}\rangle = \mathbf{0}$. Thus, $\langle \mathbf{v}_{\mathbf{K}}, \mathbf{h}_{\mathbf{K}}\rangle = \langle \mathbf{v}, \mathbf{h}\rangle - \langle \mathbf{v}_{\mathbf{K}^c}, \mathbf{h}_{\mathbf{K}^c}\rangle = -\langle \mathbf{v}_{\mathbf{K}^c}, \mathbf{h}_{\mathbf{K}^c}\rangle$. Therefore from (21), we get

$$c\,\|\hat{\mathbf{x}}\|_1 \geq c\,\|\mathbf{x}\|_1 + (c-1)\|\mathbf{h}_{\mathbf{K}^c}\|_1. \tag{23}$$

Since $c > 1$, the multiplier of $\mathbf{h}_{\mathbf{K}^c}$ is greater than zero. This implies $\mathbf{h}_{\mathbf{K}^c} = \mathbf{0}$ given that $\hat{\mathbf{x}}$ is the solution to (4). Also Since $\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}}$ has full rank, $(\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}})\mathbf{h}_{\mathbf{K}} = \mathbf{0}$ only if $\mathbf{h}_{\mathbf{K}} = \mathbf{0}$. So $\mathbf{h}$ has to be zero, which states $\hat{\mathbf{x}} = \mathbf{x}$ and $\mathbf{x}$ as the unique ℓ_1 minimizer to (4). Thus proves the theorem. □

APPENDIX B
PROOF OF THEOREM 2

Proof: Since $|B_{pq}| < \zeta$, $1 \leq p, q \leq n$, we have $\|\mathbf{B}_{\mathbf{K}}\|_\infty < s\zeta$.

If the sparsity in $\mathbf{x}$ satisfy

$$s \leq \delta/\zeta, \tag{24}$$

then $\|c\,\mathbf{I} - \mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}}\|_2 = \|\mathbf{B}_{\mathbf{K}}\|_2 \leq \sqrt{\|\mathbf{B}_{\mathbf{K}}\|_\infty\|\mathbf{B}_{\mathbf{K}}\|_1} \leq \zeta s \leq \delta$.

For the second condition on A,

$$\begin{aligned} \max_{i\in K^c}\|\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\{i\}}\|_2 &= \max_{i\in K^c}\sqrt{\sum_{l\in T}(\mathbf{A}_{\{l\}}{}^\intercal\mathbf{A}_{\{i\}})^2} \\ &= \max_{i\in K^c}\sqrt{\sum_{l\in T} B_{li}^2} \\ &\leq \sqrt{s}\zeta \leq \frac{\delta}{\sqrt{s}} \leq 1. \end{aligned} \tag{25}$$

where we used the fact that $\delta < 1$ and $s \geq 1$. Also, since $\mathbf{A}$ is normalized, $\zeta \leq 1 \implies \sqrt{s}\zeta \leq \sqrt{\delta\zeta} \leq 1$.

To get the dual certificate, we chose to define $\mathbf{v}$ as

$$\mathbf{v} = c\,\mathbf{A}^\intercal\mathbf{A}_{\mathbf{K}}(\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}})^{-1}\,\mathrm{sgn}(\mathbf{x}_{\mathbf{K}}). \tag{26}$$

For the entries of $\mathbf{v}$ in the support set K, we get

$$\begin{aligned} \mathbf{v}_{\mathbf{K}} &= c\,\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}}(\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}})^{-1}\,\mathrm{sgn}(\mathbf{x}_{\mathbf{K}}) \\ &= c\,\mathrm{sgn}(\mathbf{x}_{\mathbf{K}}), \end{aligned} \tag{27}$$

where we use the definition of right inverse [40]. This proves condition for $\mathbf{v}_{\mathbf{K}}$.

For $\mathbf{v}_{\mathbf{K}^c}$,

$$\begin{aligned} \|\mathbf{v}_{\mathbf{K}^c}\|_\infty &= \|c\,\mathbf{A}_{\mathbf{K}^c}{}^\intercal\mathbf{A}_{\mathbf{K}}(\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}})^{-1}\,\mathrm{sgn}(\mathbf{x}_{\mathbf{K}})\|_\infty \\ &\leq \|c\,\mathbf{A}_{\mathbf{K}^c}{}^\intercal\mathbf{A}_{\mathbf{K}}(\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}})^{-1}\|_\infty \\ &\leq c\,s\,\zeta\,\|(\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}})^{-1}\|_\infty. \end{aligned} \tag{28}$$

where we use the defintion of ζ on the last line. Next, since $\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}}$ is nonsingular, using Neumann Series, we get

$$c(\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}})^{-1} = \sum_{n=0}^{\infty}\frac{1}{c^n}(c\,I - \mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}})^n = \sum_{n=0}^{\infty}\frac{1}{c^n}\mathbf{B}_{\mathbf{K}}{}^n. \tag{29}$$

From Eq. (28), we get

$$\begin{aligned} c\,s\,\zeta\,\|(\mathbf{A}_{\mathbf{K}}{}^\intercal\mathbf{A}_{\mathbf{K}})^{-1}\|_\infty &= s\,\zeta\,\|\sum_{n=0}^{\infty}\frac{1}{c^n}\mathbf{B}_{\mathbf{K}}{}^n\|_\infty \\ &\leq \delta\sum_{n=0}^{\infty}\|\frac{1}{c^n}\mathbf{B}_K^n\|_\infty \\ &\leq \delta\left(1 + \frac{\delta}{c} + \left(\frac{\delta}{c}\right)^2 + \cdots\right) \\ &\leq \delta\left(\frac{1}{1-\frac{\delta}{c}}\right) \\ &\leq \frac{c\,\delta}{c-\delta} < 1, \end{aligned} \tag{30}$$

where for the last line of Eq. (30), we use the following algebra.

$$\begin{aligned} \delta &< \frac{c}{c+1} = \frac{1}{1+1/c} \\ \therefore\ 1 + \frac{1}{c} &< \frac{1}{\delta} \implies 1 < \frac{c-\delta}{c\,\delta} \\ \therefore\ \frac{c\,\delta}{c-\delta} &< 1. \end{aligned} \tag{31}$$

□

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. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
[3] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
[4] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, p. 969, 2007.
[5] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.
[6] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[7] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.
[8] Y. Zhang, "Theory of compressive sensing via ℓ_1-minimization: A non-RIP analysis and extensions," *J. Oper. Res. Soc. China*, vol. 1, no. 1, pp. 79–105, Mar. 2013.
[9] 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.
[10] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k-term approximation," *J. Amer. Math. Soc.*, vol. 22, no. 1, pp. 211–231, 2009.

[11] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
[12] E. J. Candés, 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.
[13] A. M. Tillmann and M. E. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1248–1259, Feb. 2014.
[14] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Construct. Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.
[15] J. Romberg, "Compressive sensing by random convolution," *SIAM J. Imag. Sci.*, vol. 2, no. 4, pp. 1098–1128, Nov. 2009.
[16] M. Stojnic, W. Xu, and B. Hassibi, "Compressed sensing-probabilistic analysis of a null-space characterization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 3377–3380.
[17] E. J. Candès and Y. Plan, "A probabilistic and RIPless theory of compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7235–7254, Nov. 2011.
[18] R. Kueng and D. Gross, "RIPless compressed sensing from anisotropic measurements," *Linear Algebra Appl.*, vol. 441, pp. 110–123, Jan. 2014.
[19] M. Rudelson and S. Zhou, "Reconstruction from anisotropic random measurements," in *Proc. Conf. Learn. Theory*, 2012, pp. 1–10.
[20] B. He and D. Wu, "Imaging and visualization of 3-D cardiac electric activity," *IEEE Trans. Inf. Technol. Biomed.*, vol. 5, no. 3, pp. 181–186, Sep. 2001.
[21] L.-P. Song, C. Yu, and Q. H. Liu, "Through-wall imaging (TWI) by radar: 2-D tomographic results and analyses," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2793–2798, Dec. 2005.
[22] M. Pastorino, "Recent inversion procedures for microwave imaging in biomedical, subsurface detection and nondestructive evaluation applications," *Measurement*, vol. 36, nos. 3–4, pp. 257–269, Oct. 2004.
[23] O. M. Bucci and T. Isernia, "Electromagnetic inverse scattering: Retrievable information and measurement strategies," *Radio Sci.*, vol. 32, no. 6, pp. 2123–2137, Nov./Dec. 1997.
[24] C. A. Balanis, *Advanced Engineering Electromagnetics*. Hoboken, NJ, USA: Wiley, 1999.
[25] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. Washington, DC, USA: US Government Printing office, 1948, vol. 55.
[26] O. M. Bucci and G. Franceschetti, "On the degrees of freedom of scattered fields," *IEEE Trans. Antennas Propag.*, vol. 37, no. 7, pp. 918–926, Jul. 1989.
[27] U. K. Khankhoje and K. Shah, "Spatial bandlimitedness of scattered electromagnetic fields," 2015, *arXiv:1505.00886*. [Online]. Available: http://arxiv.org/abs/1505.00886
[28] E. van den Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM J. Sci. Comput.*, vol. 31, no. 2, pp. 890–912, 2008.
[29] W. C. Chew, *Waves Fields Inhomogeneous Media*. Piscataway, NJ, USA: IEEE Press, 1995.
[30] G. Gennarelli, I. Catapano, and F. Soldovieri, "RF/microwave imaging of sparse targets in urban areas," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, pp. 643–646, 2013.
[31] G. Chen, J. Stang, M. Haynes, E. Leuthardt, and M. Moghaddam, "Real-time three-dimensional microwave monitoring of interstitial thermal therapy," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 3, pp. 528–538, Mar. 2018.
[32] W. Shao and T. McCollough, "Advances in microwave near-field imaging: Prototypes, systems, and applications," *IEEE Microw. Mag.*, vol. 21, no. 5, pp. 94–119, May 2020.
[33] M. Salucci, N. Anselmi, and A. Massa, "Compressive sensing as applied to inverse problems for imaging: Theory, applications, current trends, and open challenges," *IEEE Antennas Propag. Mag.*, vol. 59, no. 5, pp. 34–46, Oct. 2017.
[34] A. Massa, P. Rocca, and G. Oliveri, "Compressive sensing in electromagnetics-A review," *IEEE Antennas Propag. Mag.*, vol. 57, no. 1, pp. 224–238, Jan. 2015.
[35] J. Richmond, "Scattering by a dielectric cylinder of arbitrary cross section shape," *IEEE Trans. Antennas Propag.*, vol. AP-13, no. 3, pp. 334–341, May 1965.
[36] M. Moghaddam and W. C. Chew, "Study of some practical issues in inversion with the Born iterative method using time-domain data," *IEEE Trans. Antennas Propag.*, vol. 41, no. 2, pp. 177–184, Feb. 1993.
[37] A. J. Devaney and G. Sherman, "Nonuniqueness in inverse source and scattering problems," *IEEE Trans. Antennas Propag.*, vol. AP-30, no. 5, pp. 1034–1037, Sep. 1982.
[38] N. Bleistein and J. K. Cohen, "Nonuniqueness in the inverse source problem in acoustics and electromagnetics," *J. Math. Phys.*, vol. 18, no. 2, pp. 194–201, 1977.
[39] T. M. Habashy, M. L. Oristaglio, and A. T. de Hoop, "Simultaneous nonlinear reconstruction of two-dimensional permittivity and conductivity," *Radio Sci.*, vol. 29, no. 4, pp. 1101–1118, 1994.
[40] R. Penrose, "A generalized inverse for matrices," *Math. Proc. Cambridge Phil. Soc.*, vol. 51, no. 3, pp. 406–413, 1955.

PRATIK SHAH received the B.Tech. degree in electronics and communication engineering from Nirma University, Ahmedabad, India, in 2007, the M.Tech. degree from the Indian Institute of Technology Technology, Kharagpur, India, in 2009, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, USA, in 2017 and 2018, respectively. He was a Scientist with GE Global Research, Bengaluru, India, from 2009 to 2012, and a Postdoctoral Research Associate with the University of Southern California, in 2018. He is currently a Principal Scientist at Acutus Medical Inc., Carlsbad, CA, USA. His current research interests include inverse problems, inverse scattering, optimization methods, compressive sensing, image and signal processing, and machine learning and their applications in medical imaging and sensing.

MAHTA MOGHADDAM (Fellow, IEEE) received the B.S. degree (Hons.) in electrical and computer engineering from the University of Kansas, Lawrence, Kansas, in 1986, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana–Champaign, in 1989 and 1991, respectively. She was a System Engineer for the Cassini radar. She is currently a Professor of electrical engineering with the University of Southern California, Los Angeles, CA, USA. Prior to that, she was at the University of Michigan, from 2003 to 2011, and NASA Jet Propulsion Laboratory (JPL), from 1991 to 2003. She has introduced new approaches for quantitative interpretation of multichannel radar imagery based on analytical inverse scattering techniques applied to complex and random media. Her most recent research interests include the development of new radar instrument and measurement technologies for subsurface and subcanopy characterization, development of forward and inverse scattering techniques for layered random media, especially for soil moisture and permafrost applications, geophysical retrievals using signal-of-opportunity reflectometry, and transforming concepts of radar remote sensing to medical imaging and therapy systems. She is a member of the Soil Moisture Active and Passive (SMAP) Mission Science Team, the Cyclones Global Navigation Satellite System (CYGNSS) Science Team, and the National Academy of Engineering. She was the Principal Investigator of the AirMOSS NASA Earth Ventures-1 Mission. She served as the Science Chair for the JPL Team X (Advanced Mission Studies Team).