frontiers in Robotics and AI

ORIGINAL RESEARCH
published: 03 July 2015
doi: 10.3389/frobt.2015.00016

Sensitivity analysis of compressive sensing solutions

*Liyi Dai**

Computing Sciences Division, U.S. Army Research Office, Research Triangle Park, NC, USA

The compressive sensing framework finds a wide range of applications in signal processing, data analysis, and fusion. Within this framework, various methods have been proposed to find a sparse solution x from a linear measurement model $y = Ax$. In practice, the linear model is often an approximation. One basic issue is the robustness of the solution in the presence of uncertainties. In this paper, we are interested in compressive sensing solutions under a general form of measurement $y = (A + B)x + v$ in which B and v describe modeling and measurement inaccuracies, respectively. We analyze the sensitivity of solutions to infinitesimal modeling error B or measurement inaccuracy v. Exact solutions are obtained. Specifically, the existence of sensitivity is established and the equation governing the sensitivity is obtained. Worst-case sensitivity bounds are derived. The bounds indicate that sensitivity is linear to measurement inaccuracy due to the linearity of the measurement model, and roughly proportional to the solution for modeling error. An approach to sensitivity reduction is subsequently proposed.

Keywords: compressive sensing, sparse solutions, sensitivity analysis, robustness, gradient method

OPEN ACCESS

Edited by:
Soumik Sarkar,
Iowa State University, USA

Reviewed by:
Soumalya Sarkar,
Pennsylvania State University, USA
Han Guo,
Iowa State University, USA

****Correspondence:***
Liyi Dai,
Computing Sciences Division, U.S. Army Research Office, 4300 S. Miami Blvd, Research Triangle Park, NC 27703, USA
liyi.dai.civ@mail.mil

Specialty section:
This article was submitted to Sensor Fusion and Machine Perception, a section of the journal Frontiers in Robotics and AI

Received: *09 March 2015*
Accepted: *18 June 2015*
Published: *03 July 2015*

Citation:
Dai L (2015) Sensitivity analysis of compressive sensing solutions. Front. Robot. AI 2:16. doi: 10.3389/frobt.2015.00016

1. Introduction

Consider the following minimization problem from perfect measurements

$$\min ||x||_{l_0}, \text{ subject to } y = Ax, \quad (1)$$

where $y \in R^m$ is a vector of measurements, $x \in R^n$ is the vector to be solved, $A \in R^{m\times n}$ is a matrix, and l_0 denotes the l_0 norm, i.e., the number of non-zero entries. In the compressive sensing framework, the number of measurements available is far smaller than the dimension of the solution x, i.e., $m \ll n$. Because the l_0 norm is not convex, equation (1) is a combinatorial optimization problem, solving, which directly is computationally intensive and often prohibitive for problems of practical interest. Therefore, equation (1) is replaced with the following l_1 minimization problem

$$\min ||x||_{l_1}, \text{ subject to } y = Ax, \quad (2)$$

where the l_1 norm is defined as $||x||_{l_1} = \sum_i |x_i|$. Note that the matrix A in equations (1) or (2) is assumed to be known exactly and that y is free from measurement inaccuracy. In practice, the problem formulation equation (2) is often an approximation because there may exist modeling errors in A and measurement inaccuracies in y. Therefore, a realistic measurement equation would be

$$y = [A + B(\theta)]x + v(\theta), \quad (3)$$

where $\theta = [\theta_j] \in R^p$ is a vector of unknown parameters, $B(\theta) \in R^{m\times n}$ is a matrix describing modeling errors, and $v(\theta) \in R^m$ is measurement noise. One appealing feature of the measurement form equation (3) is that it can incorporate prior knowledge of the inaccuracies. For example, we may have

$$B(\theta) = \sum_j \theta_j B_j, v(\theta) = \sum_j \theta_j v_j, \quad (4)$$

where B_j and v_j are known matrices or vectors of appropriate dimensions from prior knowledge. When no prior knowledge is available, B_j and v_j can be chosen as the entry indicator matrix or vector, e.g., v_j has 1 for the j-th entry of $v(\theta)$ and 0 everywhere else. The form equation (3) can be used to describe a wide range of inaccuracies. In equation (4), θ_j is the unknown magnitude of inaccuracy. Another interpretation of the measurement form equation (3) is that it represents structured uncertainties and can be used to describe the characteristics of inaccuracies for different measurements. For example, in multi-spectrum imaging, the noise characteristics are different for different spectrum. Images taken from different views or at different times may exhibit different noise characteristics. In the setting of a sensor network, a particular sensor may produce bad data due to malfunctioning, which likely leads to different characteristics of inaccuracy compared with other functioning sensors. Without loss of generality, it is assumed that the nominal value of θ is 0, which corresponds to the case with perfect modeling and measurements. Under measurement equation (3), the problem equation (2) can be recast as

$$\min ||x||_{l_1}, \text{ subject to } y = [A + B(\theta)]x + v(\theta). \quad (5)$$

A natural question is how sensitive is the solution of equation (5) to small perturbations at θ, without loss of generality, particularly around $\theta = 0$? Because equation (2) is often only an approximation for problems of practical interest, such sensitivity characterizes the robustness of the solution to modeling or measurement inaccuracies. Among the vast literature on compressive sensing, there has been significant interest in analyzing the robustness of compressive sensing solutions in the presence of measurement inaccuracies. One widely adopted approach, such as those in Candes et al. (2006) and Candes (2008), is to reformulate the problem equation (2) as

$$\min ||x||_{l_1}, \text{ subject to } ||y - Ax||_{l_2} \le \epsilon, \quad (6)$$

in which ϵ is a tolerable bound of solution inaccuracy, and the l_2 norm is defined as $||x||_{l_2} = \left(\sum_i x_i^2\right)^{1/2}$. Existing literature on equation (6) is extensive. Interested readers are referred to Donoho (2006), Candes and Wakin (2008), and Eldar and Kutyniok (2012) for a comprehensive treatment and literature review. Of particular relevance to this paper, Donoho et al. (2011) (Donoho and Reevew, 2012) analyzed sensitivity of signal recovery solutions to the relaxation of sparsity under the Least Absolute Shrinkage and Selection Operator (LASSO) problem formulation and obtained asymptotic performance bounds in terms of underlying parameters in the methods of finding the solutions. Herman and Strohmer (2010) derived l_2 bounds between the solutions of equations (2) and (5) in terms of perturbation bounds of unknown $B(\theta)$ and $v(\theta)$. Only the magnitudes (i.e., $||.||_{l_2}$ bounds) of $B(\theta)$ and $v(\theta)$ are assumed known. Chi et al. (2011) analyzed solution bounds to modeling errors of unknown $B(\theta)$ and $v(\theta)$. The goal was to investigate the effects of basis mismatch since the matrix A is also known as the basis matrix in compressive sensing. Upper bounds of solution deviation were obtained. The treatments in those papers considered both strictly sparse signals and compressible signals, i.e., the ordered entries of the signal vector decay exponentially fast. The bounds are of the following form Upper bounds of solutions are obtained (Candes et al., 2006; Candes, 2008)

$$||x(\theta) - x||_{l_2} \le C_1||x - x_k||_{l_1} + C_2\epsilon,$$

where C_1 and C_2 are constants dependent of A, k is the sparsity of x, and x_k is x with all but the k-largest entries set to zero. Recent publications (Arias-Castro and Eldar, 2011; Davenport et al., 2012; Tang et al., 2013) analyzed the statistical bounds when measurement inaccuracies are modeled as Gaussian noises. A recent publication (Moghadam et al., 2014) after this paper was written-derived sensitivity bounds while this paper seeks exact solutions to sensitivity.

The objective of this paper is to analyze the sensitivity of compressive sensing solutions to perturbations (inaccuracies) in matrix A and measurement y, i.e., the sensitivity of solutions to equation (5) to the unknown parameter θ at $\theta = 0$. Such sensitivity characterizes the effects of infinitesimal perturbations in modeling and measurements. Deterministic problem setting is adopted, and the true solution is assumed to be k-sparse. The sensitivity is local at $\theta = 0$. Exact expressions of the sensitivity are derived. The results of this paper provide complementary insights regarding the effects of modeling and measurement inaccuracies on compressive sensing solutions.

The rest of the paper is arranged as the follows: in Section 2, we examine the continuity of the solution $x(\theta)$ at $\theta = 0$. In Section 3, we establish the existence, finiteness, and equation for the sensitivity of $x(\theta)$ at $\theta = 0$. Bounds for worst-case infinitesimal perturbations are derived in Section 4. An approach to sensitivity reduction is proposed in Section 5. A numerical example is provided in Section 6 to illustrate the sensitivity reduction algorithms. Finally, concluding remarks are provided in Section 7.

In this paper, we use lower case letters to denote column vectors, upper case letters to denote matrices. For parameterized vectors, the following notational conventions, if exist, will be adopted.

$$\nabla v(\theta) = \begin{bmatrix} \frac{\partial v(\theta)}{\partial \theta_1} & \frac{\partial v(\theta)}{\partial \theta_2} & \dots & \frac{\partial v(\theta)}{\partial \theta_p} \end{bmatrix} \in R^{m\times p},$$
$$\nabla B(\theta) = \begin{bmatrix} \frac{\partial B(\theta)}{\partial \theta_1} & \frac{\partial B(\theta)}{\partial \theta_2} & \dots & \frac{\partial B(\theta)}{\partial \theta_p} \end{bmatrix} \in R^{m\times np}.$$

Consequently, for a vector $x \in R^n$ independent of θ,

$$\nabla[B(\theta)x] = \begin{bmatrix} \frac{\partial B(\theta)}{\partial \theta_1}x & \frac{\partial B(\theta)}{\partial \theta_2}x & \dots & \frac{\partial B(\theta)}{\partial \theta_p}x \end{bmatrix} \in R^{m\times p}.$$

For notational consistence, the l_2 norm of a matrix $A = [a_{i,j}]$ is defined as the Frobenius norm, $||A||_{l_2} = \left(\sum_{i,j} a_{i,j}^2\right)^{1/2}$.

2. Solution Continuity

Let $x(\theta)$ denote the solution to equation (5). Without confusion, we reserve $x = x(0)$ to denote a solution to the unperturbed problem equation (2). In this section, we examine the continuity of $x(\theta)$ in the neighborhood of $\theta = 0$. A vector $x \in R^n$ is said to be k-sparse, $k \ll n$, if it has at most k non-zero entries (Candes et al., 2006). In the compressive sensing framework, we are interested in k-sparse solutions.

Without further assumption, the existence of a (sparse) solution to the optimization problem equation (5) is not sufficient in guaranteeing the continuity of $x(\theta)$ near $\theta = 0$. To illustrate this issue, consider the following example.

Example 2.1. Consider the problem equation (5) with

$$y = 1, A = \begin{bmatrix} 1 & 1+\theta \end{bmatrix}$$

near $\theta = 0$. Then the minimum $||x||_{l_1}$ is achieved by

$$x(\theta) = \begin{cases} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, & \text{if } \theta < 0 \\ \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (1-\alpha)\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \text{ for any } 0 \le \alpha \le 1 & \text{if } \theta = 0 \\ \begin{bmatrix} 0 \\ 1/(1+\theta) \end{bmatrix}, & \text{if } \theta > 0. \end{cases}$$

It's clear that the solution $x(\theta)$ is not continuous at $\theta = 0$.

In Example 2.1, the solution is not unique at $\theta = 0$. It turns out that the uniqueness of the solution is critical in guaranteeing the continuity of the solution near $\theta = 0$.

One popular approach to finding the solution to equation (5) is to solve the following LASSO problem.

$$\min_{x(\theta)} \left\{ \frac{1}{2}||y - [(A + B(\theta))x(\theta) + v(\theta)]||_{l_2}^2 + \tau||x(\theta)||_{l_1} \right\} \quad (7)$$

which $\tau > 0$ is a weighting parameter.

Theorem 2.1. Assume that both $B(\theta)$ and $v(\theta)$ are continuous at $\theta = 0$, $B(0) = 0$, $v(0) = 0$, AA^T is positive definite, and that the solution $x(\theta)$ to equation (7) is unique at $\theta = 0$. Then, $x(\theta)$ is continuous at $\theta = 0$, i.e.,

$$\lim_{\theta \to 0} x(\theta) = x, \quad (8)$$

where x is the solution of equation (7) at $\theta = 0$.

Proof. Consider

$$\hat{x}(\theta) = [A + B(\theta)]\left[(A + B(\theta))(A + B(\theta))^T\right]^{-1}(y - v(\theta)).$$

Then $\hat{x}(\theta)$ is the Moore–Penrose pseudoinverse solution to $y = [A + B(\theta)]\hat{x}(\theta) + v(\theta)$. Under the assumptions that AA^T is positive definite and that both $B(\theta)$ and $v(\theta)$ are continuous at $\theta = 0$ with $B(0) = 0$, $v(0) = 0$, we know that $\hat{x}(\theta)$ is continuous and uniformly bounded in a small neighborhood of $\theta = 0$, i.e., there exist $c > 0$, $\delta > 0$ such that $||\hat{x}(\theta)||_{l_1} \le c$ for all $||\theta||_{l_2} \le \delta$. Because $x(\theta)$ is the optimal solution to equation (7) while $\hat{x}(\theta)$ may not, we have

$$\tau||x(\theta)||_{l_1} \le \frac{1}{2}||y - [(A + B(\theta))x(\theta) + v(\theta)]||_{l_2}^2 + \tau||x(\theta)||_{l_1}$$
$$\le \frac{1}{2}||y - [(A + B(\theta))\hat{x}(\theta) + v(\theta)]||_{l_2}^2 + \tau||\hat{x}(\theta)||_{l_1} = \tau||\hat{x}(\theta)||_{l_1}.$$

or

$$||x(\theta)||_{l_1} \le ||\hat{x}(\theta)||_{l_1} \le c$$

for all $||\theta||_{l_2} \le \delta$. Therefore, $x(\theta)$ is uniformly bounded for all $||\theta||_{l_2} \le \delta$.

We next prove equation (8) by contradiction. Assume that equation (8) is not true. Because $x(\theta)$ is uniformly bound for all $||\theta||_{l_2} \le \delta$, there exists a sequence θ_i, $i = 1, 2, \ldots$, such that

$$\lim_{i \to \infty} \theta_i = 0 \text{ and } \lim_{i \to \infty} x(\theta_i) = \tilde{x} \ne x. \quad (9)$$

Note that $x(\theta)$ is the unique solution to equation (7). It must be that

$$\frac{1}{2}||y - [(A + B(\theta_i))x(\theta_i) + v(\theta_i)]||_{l_2}^2 + \tau||x(\theta_i)||_{l_1}$$
$$< \frac{1}{2}||y - [(A + B(\theta_i))x + v(\theta_i)]||_{l_2}^2 + \tau||x||_{l_1}.$$

By setting $i \to \infty$, also noting the continuity of $B(\theta)$ and $v(\theta)$ at $\theta = 0$, we obtain

$$\frac{1}{2}||y - A\tilde{x}||_{l_2}^2 + \tau||\tilde{x}||_{l_1} \le \frac{1}{2}||y - Ax||_{l_2}^2 + \tau||x||_{l_1}.$$

Therefore, we must have $\tilde{x} = x$ because the solution to equation (7) at $\theta = 0$ is unique, which contradicts equation (9). The contradiction establishes equation (8). ■

Note that the number of rows of A is far smaller than the number of columns, $m = n$. The positive definiteness of AA^T is equivalent to that all measurements y are not redundant, which is technical and mild. Theorem 2.1 states that the solution to equation (7) is continuous if it is unique at $\theta = 0$. Similarly, we can establish the continuity of the solution to equation (5). The proof is similar to that for Theorem 2.1 and omitted to avoid repetitiveness.

Theorem 2.2. Assume that both $B(\theta)$ and $v(\theta)$ are continuous at $\theta = 0$, $B(0) = 0$, $v(0) = 0$, AA^T is positive definite, and that the solution $x(\theta)$ to equation (5) is unique at $\theta = 0$. Then $x(\theta)$ is continuous at $\theta = 0$, i.e.,

$$\lim_{\theta \to 0} x(\theta) = x. \quad (10)$$

We introduce the following concept, which is a one-sided relaxation of the *Restricted Isometry Property (RIP)* (Candes et al., 2006).

Definition 2.1. A matrix $A \in R^{m \times n}$ is said to be k-sparse positive definite if there exists a constant $c > 0$ such that

$$||Ax||_{l_2}^2 \ge c||x||_{l_2}^2 \quad (11)$$

for any k-sparse vector $x \in R^n$.

The $2k$-sparse positive definiteness is a sufficient condition for guaranteeing the uniqueness of the optimal solution to equation (5) (Candes, 2008).

3. Solution Sensitivity

In this section, we are interested in the gradient information

$$\nabla x(\theta) = \begin{bmatrix} \frac{\partial x(\theta)}{\partial \theta_1} & \frac{\partial x(\theta)}{\partial \theta_2} & \dots & \frac{\partial x(\theta)}{\partial \theta_p} \end{bmatrix} \tag{12}$$

which, if exists, is an $n \times p$ matrix in $R^{n \times p}$.

The next theorem establishes the existence of the gradient equation (12).

Theorem 3.1. Consider problem equation (5). Assume that both $B(\theta)$ and $v(\theta)$ are differentiable at $\theta = 0$, $B(0) = 0$, $v(0) = 0$, and that there exists a $\delta > 0$ such that A is $3k$-sparse positive definite for all $||\theta||_{l_2} \leq \delta$. Then at $\theta = 0$, the gradient $\bigtriangledown x(\theta)$ exists, is finite, and satisfies

$$AZ + EX + W = 0, \tag{13}$$

where

$$Z = \nabla x(\theta)|_{\theta=0} \in R^{n \times p}, E = \nabla B(\theta)|_{\theta=0} \in R^{m \times np},$$
$$W = \nabla v(\theta)|_{\theta=0} \in R^{m \times p},$$
$$X = \begin{bmatrix} x & 0 & 0 & \dots & 0 \\ 0 & x & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x \end{bmatrix} \in R^{np \times p},$$

and x is the solution to equation (2).

Proof. For the sake of simple notations, we only prove the theorem for scalar θ. The proof can be extended component wisely to the vector case.

We first establish the existence of $\bigtriangledown x(\theta)$. Let $h(\theta) = \theta^{-1}[x(\theta) - x]$ where $x(\theta)$ and x are solutions to the optimization problems equations (5) and (2), respectively.

Under the continuity of $B(\theta)$, $A + B(\theta)$ is $3k$-sparse positive definite with a constant $\tilde{c}$ independent of θ in a small neighborhood of $\theta = 0$ if A is. Therefore,

$$\tilde{c}||x - x(\theta)||_{l_2}^2 \leq ||[A + B(\theta)][x - x(\theta)]||_{l_2}^2 \leq ||B(\theta)||_{l_2}^2 ||x||_{l_2}^2 + ||v(\theta)||_{l_2}^2. \tag{14}$$

or

$$||h(\theta)||_{l_2}^2 \leq \tilde{c}^{-1}[||\theta^{-1}B(\theta)||_{l_2}^2 ||x||_{l_2}^2 + ||\theta^{-1}v(\theta)||_{l_2}^2].$$

By assumption, $B(\theta)$ and $v(\theta)$ are differentiable at $\theta = 0$. The previous inequality shows that $h(\theta)$ is uniformly bounded in a small neighborhood of $\theta = 0$. Consequently, there exists a sequence $\theta^{(i)} \to 0$ such that

$$\hat{h} \triangleq \lim_{i \to \infty} h(\theta^{(i)}) \tag{15}$$

exists and is finite. Recall that

$$[A + B(\theta)]h(\theta) = -\theta^{-1}B(\theta)x - \theta^{-1}v(\theta)$$

and

$$[A + B(\theta^{(i)})]h(\theta^{(i)}) = -(\theta^{(i)})^{-1}B(\theta^{(i)})x - (\theta^{(i)})^{-1}v(\theta^{(i)}).$$

Their difference gives

$$A[h(\theta) - h(\theta^{(i)})] = -B(\theta)h(\theta) + B(\theta^{(i)})h(\theta^{(i)}) - \theta^{-1}B(\theta)x + (\theta^{(i)})^{-1}B(\theta^{(i)})x - \theta^{-1}v(\theta) + (\theta^{(i)})^{-1}v(\theta^{(i)})$$

or

$$||A[h(\theta) - h(\theta^{(i)})]||_{l_2}^2 \leq ||B(\theta)h(\theta)||_{l_2}^2 + ||B(\theta^{(i)})h(\theta^{(i)})||_{l_2}^2 + ||[\theta^{-1}B(\theta) - (\theta^{(i)})^{-1}B(\theta^{(i)})]x||_{l_2}^2 + ||\theta^{-1}v(\theta) - (\theta^{(i)})^{-1}v(\theta^{(i)})||_{l_2}^2.$$

Since A is $3k$-sparse positive definite, there exists a constant $c > 0$ such that

$$c||h(\theta) - h(\theta^{(i)})||_{l_2}^2 \leq ||A[h(\theta) - h(\theta^{(i)})]||_{l_2}^2.$$

Consequently,

$$c||h(\theta) - h(\theta^{(i)})||_{l_2}^2 \leq ||B(\theta)h(\theta)||_{l_2}^2 + ||B(\theta^{(i)})h(\theta^{(i)})||_{l_2}^2 + ||\theta^{-1}B(\theta) - (\theta^{(i)})^{-1}B(\theta^{(i)})||_{l_2}^2 ||x||_{l_1}^2 + ||\theta^{-1}v(\theta) - (\theta^{(i)})^{-1}v(\theta^{(i)})||_{l_2}^2. \tag{16}$$

The right hand side of equation (16) goes to zero as $\theta \to 0$, $i \to \infty$. Therefore,

$$\lim_{\theta \to 0, i \to \infty} ||h(\theta) - h(\theta^{(i)})||_{l_2}^2 = 0.$$

Or equivalently,

$$\lim_{\theta \to 0} h(\theta) = \lim_{i \to \infty} h(\theta^{(i)}) = \hat{h}, \tag{17}$$

which establishes that $\bigtriangledown x(\theta)$ exists at $\theta = 0$ and is finite.

Finally, the fact that $x(\theta)$ and x are the solutions to problems equations (5) and (2), respectively, gives

$$[A + B(\theta)]h(\theta) + \theta^{-1}B(\theta)x + \theta^{-1}v(\theta) = 0.$$

Denote $Z \triangleq \hat{h} = \lim_{\theta \to 0} h(\theta)$ which exists and is finite. By setting $\theta \to 0$ in the above equation, we obtain

$$AZ + EX + W = 0$$

which is equation (13). This completes the proof of the theorem. ■

Since the solution $x(\theta)$ is k-sparse for all θ in a small neighborhood of $\theta = 0$, by focusing on the non-zero entries of x and $x(\theta)$, the proof of Theorem 3.1 also shows that each column of Z is k-sparse. The support of Z is the same as that of x.

Equation (13) indicates that only those perturbations in E that corresponds to non-zero entries of x contribute to the sensitivity.

A note is needed regarding the requirement of the $3k$-sparse positive definiteness in Theorem 3.1. It has been known in the literature [e.g., Candes (2008) in the form of the Restricted Isometry Property] that $2k$-sparse positive definiteness is needed for guaranteeing the uniqueness of a k-sparse solution. The $3k$-sparse positive definiteness is needed for ensuring the differentiability of the solution, a higher order property. This is a sufficient condition, and its relaxation is a subject of current effort.

4. Sensitivity Analysis

Equation (13) gives an explicit expression of the sensitivity $\triangledown x(\theta)$ at $\theta = 0$. Theorem 3.1 shows that the k-sparse solution Z is unique for given x, E, and W. In this section, we consider the worst-case perturbations, i.e., the value(s) of E and W that maximizes $||Z||_{l_2}$. It's obvious that the sensitivity increases when either $||E||_{l_2}$ or $||W||_{l_2}$ increases. We analyze the sensitivity based on normalized perturbations. Specifically, we consider the following

$$\max_{E:||E||_{l_2}=1;W=0} ||Z||_{l_2}, \max_{E=0;W:||W||_{l_2}=1} ||Z||_{l_2}, \text{ and } \max_{E,W:||[E\ W]||_{l_2}=1} ||Z||_{l_2}.$$

Lemma 4.1. The sensitivity Z satisfies the following relationship

$$\max_{E:||E||_{l_2}=1;W=0} ||AZ||_{l_2} = ||x||_{l_2}, \tag{18}$$

$$\max_{E=0;W:||W||_{l_2}=1} ||AZ||_{l_2} = 1, \tag{19}$$

$$\max_{E,W:||[E\ W]||_{l_2}=1} ||AZ||_{l_2} = (||x||_{l_2}^2 + 1)^{1/2}. \tag{20}$$

Proof. According to equation (13), Z satisfies

$$AZ = -EX - W.$$

Next we consider each of the three cases.
Case I: $W = 0$. In this case,

$$||AZ||_{l_2} = ||EX||_{l_2}.$$

Recall that

$$EX = [E_1 x \quad E_2 x \quad \dots \quad E_p x].$$

By applying the Cauchy-Schwarz inequality $||Ax||_{l_2} \leq ||A||_{l_2} ||x||_{l_2}$, we obtain

$$\begin{aligned} ||EX||_{l_2}^2 &= \sum_{j=1}^{p} ||E_j x||_{l_2}^2 \leq \sum_{j=1}^{p} ||E_j||_{l_2}^2 ||x||_{l_2}^2 \\ &= \left(\sum_{j=1}^{p} ||E_j||_{l_2}^2 \right) ||x||_{l_2}^2 = ||E||_{l_2}^2 ||x||_{l_2}^2. \end{aligned}$$

Therefore,

$$\begin{aligned} \sup_{E:||E||_{l_2}=1;W=0} ||AZ||_{l_2} &= \sup_{E:||E||_{l_2}=1;W=0} ||EX||_{l_2} \\ &\leq \max_{E:||E||_{l_2}=1} ||E||_{l_2} ||x||_{l_2} = ||x||_{l_2}. \end{aligned} \tag{21}$$

Next, we prove that the upper bound is achievable for a specifically chosen E. Let

$$\hat{E} = [\hat{E}_1 \quad \hat{E}_2 \quad \dots \quad \hat{E}_p], \hat{E}_j = \alpha e_m x^T, \alpha = (mp)^{-1/2} ||x||_{l_2}^{-1},$$
$$e_m = [1 \quad 1 \quad \dots \quad 1]^T \in R^m.$$

Then

$$||\hat{E}_j||_{l_2}^2 = \text{trace}(\hat{E}_j \hat{E}_j^T) = \alpha^2 ||x||_{l_2}^2 \text{trace}(e_m e_m^T) = m\alpha^2 ||x||_{l_2}^2$$

and

$$||\hat{E}||_{l_2}^2 = \sum_{j=1}^{p} ||E_j||_{l_2}^2 = \sum_{j=1}^{p} m\alpha^2 ||x||_{l_2}^2 = \alpha^2 (mp||x||_{l_2}^2) = 1.$$

For this $E = \hat{E}$, we have

$$\begin{aligned} ||\hat{E}X||_{l_2}^2 &= \sum_{j=1}^{p} ||\hat{E}_j x||_{l_2}^2 = \sum_{j=1}^{p} ||\alpha e_m x^T x||_{l_2}^2 \\ &= \alpha^2 ||x||_{l_2}^4 mp = ||x||_{l_2}^2, \end{aligned}$$

or equivalently

$$||AZ||_{l_2} = |\hat{E}X||_{l_2} = ||x||_{l_2}. \tag{22}$$

The combination of equations (21) and (22) gives (18).
Case II: $E = 0$. In case,

$$||AZ||_{l_2} = ||W||_{l_2}.$$

from which equation (19) follows.
Case III: In general,

$$||AZ||_{l_2} = ||EX + W||_{l_2} = ||\tilde{E}\tilde{X}||_{l_2}$$

where

$$\tilde{E} = [\tilde{E}_1 \tilde{x} \quad \tilde{E}_2 \tilde{x} \quad \dots \quad \tilde{E}_p \tilde{x}], \tilde{x} = \begin{bmatrix} x \\ 1 \end{bmatrix},$$
$$\tilde{E}_j = [E_j \quad W_j], j = 1, 2, ..., p$$

and W_j is the j-th column of W. By following the proof of Case I, we know that

$$\begin{aligned} \max_{E,W:||[E\ W]||_{l_2}=1} ||AZ||_{l_2} &= \max_{\tilde{E}:||\tilde{E}||_{l_2}=1} ||\tilde{E}\tilde{X}||_{l_2} \\ &= ||\tilde{x}||_{l_2} = (||x||_{l_2}^2 + 1)^{1/2} \end{aligned}$$

which establishes equation (20). ■

Equations (18)–(20) show that the worst-case $||AZ||_{l_2}$ is a constant for measurement noise but proportional to the solution vector for modeling error E.

Theorem 4.1. Under the assumptions in Theorem 3.1, the sensitivity Z satisfies the following bounds

$$\max_{E:||E||_{l_2}=1;W=0} ||Z||_{l_2} \leq \sigma_{\min}^{-1}(A) ||x||_{l_2}, \tag{23}$$

$$\max_{E=0;:||W||_{l_2}=1} ||Z||_{l_2} \leq \sigma_{\min}^{-1}(A), \tag{24}$$

$$\max_{E:||E||_{l_2}=1;W:||W||_{l_2}=1} ||Z||_{l_2} \leq \sigma_{\min}^{-1}(A)(||x||_{l_2}^2 + 1)^{1/2}. \tag{25}$$

in which $\sigma_{\min}(A) > 0$ is the minimal singular value of A.

Proof. Without loss of generality, assume θ is a scalar. Let S be the collection of indices of possibly non-zero entries of Z, $Z_S \in \mathrm{R}^{|S|}$ be the vector of non-zero entries of Z corresponding to the indices in S, and A_S be the matrix consisting of corresponding columns of A. Then, A_S is of full column rank if it is k-sparse positive definite, $\sigma_{\min}(A_S) \geq \sigma_{\min}(A)$, and

$$AZ = A_S Z_S.$$

Consider the matrix AZ,

$$\begin{aligned}||AZ||_{l_2}^2 &= \text{trace}(Z^T A^T AZ) = \text{trace}(Z_S^T A_S^T A_S Z_S)\\ &\geq \text{trace}(Z_S^T (\sigma_{\min}(A_S))^2 Z_S)\\ &\geq (\sigma_{\min}(A))^2 ||Z_S||_{l_2}^2 = (\sigma_{\min}(A))^2 ||Z||_{l_2}^2,\end{aligned}$$

which gives

$$||Z||_{l_2} \leq \sigma_{\min}^{-1}(A)||AZ||_{l_2}. \tag{26}$$

The rest follows from combining the previous inequality with equations (18)–(20) in Lemma 4.1. ■

If $\sigma_{\min}(A)$ is achieved over k columns of A that correspond to the k-sparse signal to be recovered, then the equality in equation (26) holds and the bounds in Theorem 4.1 could be tight. The bounds are not tight in general because the k columns of A corresponding to the indices in S are unknown *a priori*.

5. Sensitivity Reduction

One natural way of reducing the sensitivity of compressive sensing solutions is to carefully select the basis matrix A so that Z is as small as possible. There has been extensive research on how to select A (Donoho et al., 2006). The general idea is to improve the incoherence of the matrix A, for example, by making A as close as possible to the identity matrix in terms of the RIP property. The problem could be difficult because selecting the correct columns of A from available choices is a combinatorial optimization problem. In this section, we consider an alternative. We assume that the matrix A is given. The objective is to find a solution x that is least sensitive to perturbations B and v. One example where this problem arises is in pattern recognition where each column of A represents a vector in a feature space. The issue is how to classify the pattern based on features of the training data so that the solution is robust to potential perturbations (i.e., noise in data).

We consider to improve the robustness of a signal recovery solution by reducing its sensitivity to model error and measurement inaccuracy. Toward that goal, we reformulate the l_1 optimization problem equation (5) by including a term to penalize high sensitivity. This approach offers an alternative to the conventional approach of basis matrix selection and could be performed after the matrix A is selected.

Solution robustness can be improved by reducing the magnitude of sensitivity. The l_∞ norm is a natural choice because it characterizes the largest entry of a vector or a matrix. Therefore, we consider the following optimization problem.

$$\min ||x||_{l_1} + \lambda||Z||_{l_\infty}, \text{ subject to } y = Ax, AZ + EX + W = 0, \tag{27}$$

where $\lambda > 0$ is a penalizing weight and the l_∞ norm is defined as $||Z||_{l_\infty} = \max_{i,j}|z_{ij}|$. Note that $Z \in R^{n\times p}$ is a matrix. We convert Z into a vector by defining

$$\tilde{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{bmatrix} \in R^{np},$$

where for each j, $1 \leq j \leq p$, z_j is the j-th column of Z. Then problem equation (27) can be re-written as

$$\min ||x||_{l_1} + \lambda||\tilde{z}||_{l_\infty}, \text{ subject to } Cx + D\tilde{z} = f, \tag{28}$$

in which

$$C = \begin{bmatrix} A \\ E_1 \\ E_2 \\ \vdots \\ E_p \end{bmatrix} \in R^{(m+1)p\times n}, D = \begin{bmatrix} 0 & 0 & \dots & 0 \\ A & 0 & \dots & 0 \\ 0 & A & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A \end{bmatrix} \in R^{(m+1)p\times np},$$

$$f = \begin{bmatrix} y \\ -w_1 \\ -w_2 \\ \vdots \\ -w_p \end{bmatrix} \in R^{(m+1)p},$$

and w_j is the j-th column of W, $1 \leq j \leq p$.

Several methods are available to solve the optimization problem equation (28), one of which is the Alternating Direction Method of Multipliers (ADMM) (Boyd et al., 2011). In fact, equation (28) is in the standard form of the ADMM problem formulation. There is a trade-off in selecting the method or algorithm to solve the optimization problem equation (28) in terms of accuracy, efficiency, and reliability. We adopt the gradient projection algorithms proposed by Figueiredo et al. (2007) for solving LASSO problems. Toward that end, we consider two modifications: The first modification is to replace the l_∞ norm with the l_ν norm defined as $||Z||_{l_\nu} = (\sum_{i,j} |z_j(i)|^\nu)^{1/\nu}$ where $z_j(i)$ is the ith entry of z_j-the jth column of Z. It is known that $\lim_{\nu\to\infty}||Z||_{l_\nu} = ||Z||_{l_\infty}$. So a large ν is chosen. To ensure the smoothness of the objective function, ν is chosen as an even number. The second modification is to reformulate equation (28) as a modified LASSO problem so that the efficient gradient projection algorithms are applicable with minimum modifications. Finally, it is noted that for each j, $z_j(i) = 0$ for all i not in the support of x. Consequently, we consider the following optimization problem for sensitivity reduction.

$$\begin{aligned}&\min \tau||x||_{l_1} + \lambda||\tilde{z}||_{l_\nu} + \frac{1}{2}||Cx + D\tilde{z} - f||_{l_2}^2 \\ &\text{subject to } S(x) = S(z_j), j = 1, 2, \dots, p\end{aligned} \tag{29}$$

where $\nu > 1$ is a large even number, $\tau > 0$ and $\lambda > 0$ are user-selected weights, $S(x)$ is the set of indices corresponding to possibly non-zero entries of a sparse vector x, i.e., the support of x.

We next provide the algorithmic details of implementing the gradient projection method. First, the problem equation (29) is equivalent to the following constrained optimization problem (Figueiredo et al., 2007).

$$\min F(u, r, \tilde{z}) \triangleq \frac{\tau}{2} e^T (u + r) + \lambda ||\tilde{z}||_{l_\nu} + \frac{1}{2} ||C(r - u) + D\tilde{z} - f||_{l_2}^2 \quad (30)$$
$$\text{subject to } u \geq 0, r \geq 0, S(r - u) = S(z_j), j = 1, 2, ..., p$$

in which e is the vector of all ones, $u = [u_j] \geq 0$ (or $r = [r_j] \geq 0$) denotes $u_j \geq 0$ (or $r_j \geq 0$, respectively) for all $j = 1, 2, \ldots, n$. The sparse solution x is given by

$$x = \frac{1}{2}(r - u),$$

and thus $S(r - u) = S(x)$ in equation (30). The gradient of the objective function $F(u, r, \tilde{z})$ in equation (30) is

$$\nabla F(u, r, \tilde{z}) = \begin{bmatrix} \nabla_u F(u, r, \tilde{z}) \\ \nabla_r F(u, r, \tilde{z}) \\ \nabla_{\tilde{z}} F(u, r, \tilde{z}) \end{bmatrix}, \quad (31)$$

where

$$\nabla_u F(u, r, \tilde{z}) = \frac{\tau}{2} e - \frac{1}{2} C^T [C(r - u) + D\tilde{z} - f],$$
$$\nabla_r F(u, r, \tilde{z}) = \frac{\tau}{2} e + \frac{1}{2} C^T [C(r - u) + D\tilde{z} - f],$$
$$\nabla_{\tilde{z}} F(u, r, \tilde{z}) = \lambda ||\tilde{z}||_{l_\nu}^{1-\nu} \tilde{z}^{\nu-1} + D^T [\frac{1}{2} C(r - u) + D\tilde{z} - f],$$

in which $\tilde{z}^{\nu-1} = [\tilde{z}_j^{\nu-1}]$ for $\tilde{z} = [\tilde{z}_j]$.

The gradient projection algorithms proposed in Figueiredo et al. (2007) find the optimal solutions iteratively. Let $u^{(i)}, r^{(i)}, \tilde{z}^{(i)}$ denote the values of $u, r, \tilde{z}$ at the ith iteration. Then, at the $(i + 1)$th iteration, $u^{(i+1)}, r^{(i+1)}, \tilde{z}^{(i+1)}$ are updated according to

$$\begin{bmatrix} u^{(i+1)} \\ r^{(i+1)} \\ \tilde{z}^{(i+1)} \end{bmatrix} = \begin{bmatrix} u^{(i)} + \eta^{(i)} (u^{(i+1:i)} - u^{(i)}) \\ r^{(i)} + \eta^{(i)} (r^{(i+1:i)} - r^{(i)}) \\ \tilde{z}^{(i)} + \eta^{(i)} (\tilde{z}^{(i+1:i)} - \tilde{z}^{(i)}) \end{bmatrix} \quad (32)$$

where $\eta^{(i)} \in [0,1]$ is a weighting factor, and the transition values $u^{(i+1:i)}, r^{(i+1:i)}, \tilde{z}^{(i+1:i)}$ are gradient projection updates

$$u^{(i+1:i)} = \Pi_{C^+} (u^{(i)} - \alpha^{(i)} \nabla_u F(u^{(i)}, r^{(i)}, \tilde{z}^{(i)})). \quad (33)$$
$$r^{(i+1:i)} = \Pi_{C^+} (r^{(i)} - \alpha^{(i)} \nabla_r F(u^{(i)}, r^{(i)}, \tilde{z}^{(i)})). \quad (34)$$
$$\tilde{z}^{(i+1:i)} = \Pi_{S(r^{(i+1)} - u^{(i+1)})} (\tilde{z}^{(i)} - \alpha^{(i)} \nabla_{\tilde{z}} F(u^{(i)}, r^{(i)}, \tilde{z}^{(i)})). \quad (35)$$

In equations (33)–(35), $\alpha^{(i)} > 0$ is the step-size of the gradient algorithm, $C^+ = \{u = [u_j] \in R^n | u_j \geq 0, j = 1,2,\ldots n\}$ is the non-negative subspace, $S(x)$ is the support of x as defined in equation (29), and $\Pi_C(u)$ is the projection operator that projects u onto the subspace C. Because $x = 1/2(r - u)$, the subspace $S[r^{(i+1)} - u^{(i+1)}]$ in equation (35) is the subspace of x at the $(i + 1)$th iteration. To increase the efficiency of the algorithm, the step-size $\alpha^{(i)}$ is selected according to the Armijo rule, i.e., $\alpha^{(i)} = \alpha_0 \beta^{i_0}$ in which α_0 is the initial step-size, $\beta \in (0,1)$, and i_0 is the first number in the sequence of $\alpha_0, \alpha\beta, \alpha\beta^2, \alpha\beta^3, \ldots$ that achieves

$$\min_i F(\Pi_{C^+} (u^{(i)} - \alpha_0 \beta^i \nabla_u F), \Pi_{C^+} (r^{(i)} - \alpha_0 \beta^i \nabla_r F), \Pi_{S(r^{(i+1)} - u^{(i+1)})} (\tilde{z}^{(i)} - \alpha_0 \beta^i \nabla_{\tilde{z}} F)),$$

in which the values of $\bigtriangledown F$ are evaluated at $(u^{(i)}, r^{(i)}, \tilde{z}^{(i)})$. Note that the projection of $\tilde{z}$ onto the support of x in equation (35) is unique to solving the optimization problem equation (30).

The value of x at the ith iteration is given by $\frac{1}{2}(r^{(i)} - u^{(i)})$.

The dimension of z is np which could be high for large p since n typically is a large number. However, the matrix multiplications in updating the gradient equation (31) could be done offline. Updating equations (32)–(35) is straightforward.

The algorithm described by equations (32)–(35) has been observed with fast convergence for the numerical example in Section 6. It is nevertheless a basic version of the gradient projection algorithms. Our purpose is to illustrate the incorporation sensitivity information in improving the robustness of compressive sensing solutions through sensitivity reduction. Readers interested in the gradient projection approach are referred to Figueiredo et al. (2007) for a comprehensive treatment and improvements.

6. A Numerical Example

The goal of this section is to illustrate the performance of the sensitivity reduction algorithm described in Section 5 through a numerical example. In this example, the matrix A and measurement y are taken from the numerical example described in the software package accompanying (Candes and Romberg, 2005). In this example, $n = 1024$, $m = 512$. The matrix A and measurement y are, respectively, a random matrix and vector drawn from uniform distributions. Sparsity factor $k = 102$, about 10% of the entries of x. The sparse solution x is shown in **Figure 1**.

We consider the sensitivity analysis and reduction for the following perturbation

$$B = \theta a_{i_0}, v = \theta \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in R^m, \quad (36)$$

where $i_0 = 959$, a_{i_0} is the i_0-th column of A. Note that θ is a scalar, thus $p = 1$. Perturbation equation (36) leads to a spike in sensitivity as shown in **Figure 2**. For pattern classification, each column of A represents a feature vector of a training data point. The choice of B implies that we are interested in the sensitivity of a bad training data point and seek the reduction of its effects to pattern classification. Other user-selected parameters are $\nu = 50$, $\tau = 0.85$, $\eta^{(i)} = 0.4$ for all i, $\alpha_0 = 0.9$, and $\beta = 0.8$. For comparison, we show the sparse solution x and the corresponding sensitivity for two cases.

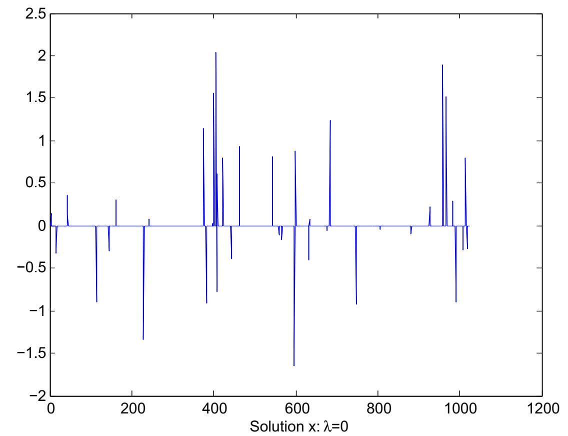

FIGURE 1 | Solution x without sensitivity reduction: $\lambda = 0$.

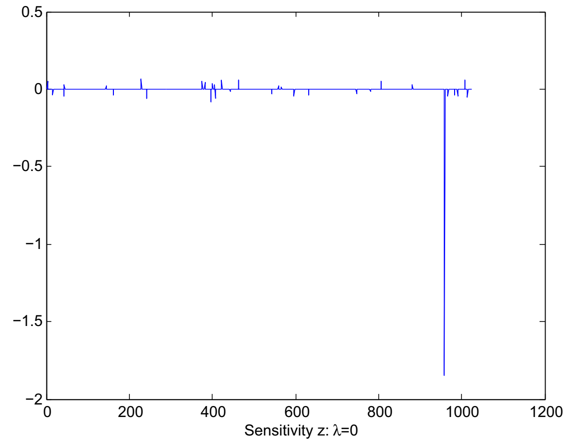

FIGURE 2 | Sensitivity Z without sensitivity reduction: $\lambda = 0$.

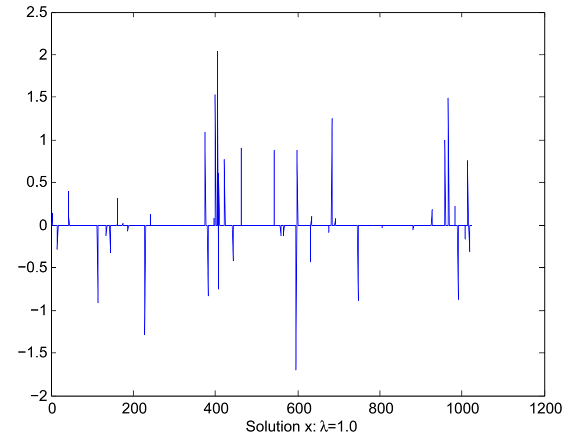

FIGURE 3 | Solution x with sensitivity reduction: $\lambda = 1.0$.

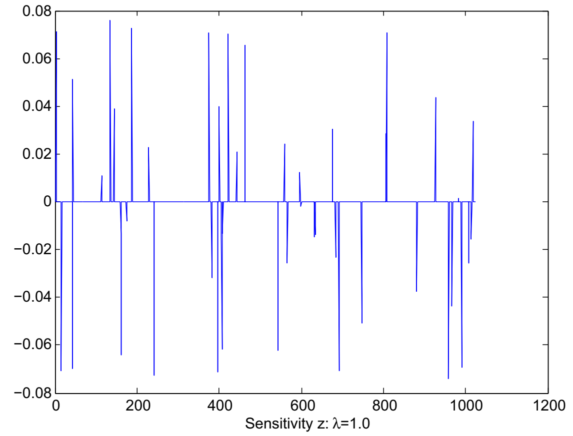

FIGURE 4 | Sensitivity Z with sensitivity reduction: $\lambda = 1.0$.

Case 1: $\lambda = 0$. This case corresponds to the sparse solution to the following LASSO optimization:

$$\min \tau ||x||_{l_1} + \frac{1}{2}||y - Ax||_{l_2}^2.$$

Figure 1 shows the values of the solution x, which clearly indicates the sparseness of the solution. **Figure 2** shows the sensitivity, Z displayed in the vector form $\tilde{z}$ ($Z = \tilde{z}$ for $p = 1$), of x with respect to the parameter θ. The largest value of sensitivity is $||\tilde{z}||_{l_\infty} = 1.8443$, which corresponds to $x(i_0) = x(959)$. The largest value of sensitivity is orders of magnitude larger than the rest of sensitivity values, which indicates the solution x is sensitive to the value of the i_0th column of A.

Case 2: $\lambda = 1.0$. This case corresponds to the sparse solution with sensitivity reduction through the gradient projection algorithm described in Section 5. The values of x and $\tilde{z}$ are shown in **Figures 3** and **4**. **Figure 3** indicates that $x(i_0)$ is now set to 0 to remove the effect of the i_0th column of A. For pattern classification, this would mean that the feature vector corresponding to the i_0 column of A is sensitive to noise and thus is removed for this particular case. Such feature vector may be resulted from biased data or incorrect choice, depending on the particular problem of concern or scenario. The largest value of sensitivity is $||Z||_{l_\infty} = 0.0761$, which represents orders of magnitude reduction from $||Z||_{l_\infty} = 1.8443$.

7. Conclusion

Robustness of compressive sensing solutions has attracted extensive interest. Existing efforts have been focused on obtaining analytical bounds between the solutions of the perturbed and unperturbed linear measurement models. The perturbation is unknown but finite with a known upper bound. In this paper, we consider solution sensitivity to infinitesimal perturbations, the "other" side of perturbation has opposed to finite perturbations. The problem formulation enables the derivation of exact solutions. The results show that solution sensitivity is linear to measurement noise and proportional to the solution. We have also demonstrated how the sensitivity information can be incorporated in problem formulation to improve solution robustness. The new problem formulation provides a trade-off (i.e., by adjusting the parameter λ) between accuracy and robustness of compressive sensing solutions. One future research direction would be using the sensitivity information to adaptive optimization of parameterized compressive sensing problems. In computer vision, it is well recognized that the performance, e.g., the probability of detection, of object recognition is sensitive to feature selection and training data. The sensitivity information may be used to reduce such sensitivity, for example, in compressive sensing-based approaches to object recognition.

Acknowledgments

This work is part of the author's staff research at the North Carolina State University. The author would like to thank Prof. Dror Baron, Jin Tan, and Nikhil Krishnan for their helpful comments and constructive suggestions. Their comments lead to the correction of an error in Lemma 3.1 in an earlier version. This work is supported in part by the U.S. Army Research Office under agreement W911NF-04-D-0003.

References

Arias-Castro, E., and Eldar, Y. C. (2011). Noise folding in compressed sensing. *IEEE Signal Process. Lett.* 18, 478–481. doi:10.1109/LSP.2011.2159837

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* 3, 1–122. doi:10.1561/2200000016

Candes, E. J. (2008). The restricted isometry property and its implications for compressed sensing. *C.R. Acad. Sci. Ser. I* 346, 589–592. doi:10.1016/j.crma.2008.03.014

Candes, E. J., and Romberg, J. (2005). l_1-MAGIC: Recovery of Sparse Signals via Convex Programming. Available at: http://users.ece.gatech.edu/~justin/l1magic/

Candes, E. J., Romberg, J., and Tao, T. (2006). Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* 59, 1207–1223. doi:10.1002/cpa.20124

Candes, E. J., and Wakin, M. B. (2008). An introduction to compressive sampling. *IEEE Signal Process. Mag.* 25, 21–30. doi:10.1109/MSP.2007.914731

Chi, Y., Scharf, L. L., Pezeshki, A., and Calderbank, A. R. (2011). Sensitivity to basis mismatch in compressed sensing. *IEEE Trans. Signal Process.* 59, 2182–2195. doi:10.1109/TSP.2011.2112650

Davenport, M. A., Laska, J. N., Treichler, J. R., and Baraniuk, R. G. (2012). The pros and cons of compressive sensing for wideband signal acquisition: noise folding versus dynamic range. *IEEE Trans. Signal Process.* 60, 4628–4642. doi:10.1109/TSP.2012.2201149

Donoho, D., and Reevew, G. (2012). "The sensitivity of compressed sensing performance to relaxation of sparsity," in *Proceedings of the 2012 IEEE International Symposium on Information Theory* (Cambridge, MA: IEEE Conference Publications), 2211–2215. doi:10.1109/ISIT.2012.6283846

Donoho, D. L. (2006). Compressed sensing. *IEEE Trans. Inf. Theory* 52, 1289–1306. doi:10.1109/TIT.2006.871582

Donoho, D. L., Elad, M., and Temlyakov, V. (2006). Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Trans. Inf. Theory* 52, 6–18. doi:10.1109/TIT.2005.860430

Donoho, D. L., Maleki, A., and Montanari, A. (2011). The noise-sensitivity phase transition in compressed sensing. *IEEE Trans. Inf. Theory* 57, 6920–6941. doi:10.1109/TIT.2011.2165823

Eldar, Y. C., and Kutyniok, G. (2012). *Compressed Sensing: Theory and Applications*. New York: Cambridge University Press.

Figueiredo, M. A. T., Nowak, R. D., and Wright, S. J. (2007). Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. *IEEE J. Sel. Top. Signal Process.* 1, 586–597. doi:10.1109/JSTSP.2007.910281

Herman, M. A., and Strohmer, T. (2010). General deviants: an analysis of perturbations in compressed sensing. *IEEE J. Sel. Top. Signal Process.* 4, 342–349. doi:10.1109/JSTSP.2009.2039170

Moghadam, A. A., Aghagolzadeh, M., and Radha, H. (2014). "Sensitivity analysis in RIPless compressed sensing," in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (Monticello, IL: IEEE Conference Publications), 881–888. doi:10.1109/ALLERTON.2014.7028547

Tang, Y., Chen, L., and Gu, Y. (2013). On the performance bound of sparse estimation with sensing matrix perturbation. *IEEE Trans. Signal Process.* 61, 4372–4386. doi:10.1109/TSP.2013.2271481

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.