Recoverability Analysis for Modified Compressive Sensing with Partially Known Support

Jun Zhang¹, Yuanqing Li^{2*}, Zhenghui Gu², Zhu Liang Yu²

¹ College of Information Engineering, Guangdong University of Technology, Guangzhou, People's Republic of China, ² Center for Brain Computer Interfaces and Brain Information Processing, South China University of Technology, Guangzhou, People's Republic of China

Abstract

The recently proposed modified-compressive sensing (modified-CS), which utilizes the partially known support as prior knowledge, significantly improves the performance of recovering sparse signals. However, modified-CS depends heavily on the reliability of the known support. An important problem, which must be studied further, is the recoverability of modified-CS when the known support contains a number of errors. In this letter, we analyze the recoverability of modified-CS in a stochastic framework. A sufficient and necessary condition is established for exact recovery of a sparse signal. Utilizing this condition, the recovery probability that reflects the recoverability of modified-CS can be computed explicitly for a sparse signal with ℓ nonzero entries. Simulation experiments have been carried out to validate our theoretical results.

Citation: Zhang J, Li Y, Gu Z, Yu ZL (2014) Recoverability Analysis for Modified Compressive Sensing with Partially Known Support. PLoS ONE 9(2): e87985. doi:10.1371/journal.pone.0087985

Editor: Holger Fröhlich, University of Bonn, Bonn-Aachen International Center for IT, Germany

Received: August 7, 2013; **Accepted:** January 2, 2014; **Published:** February 10, 2014

Copyright: © 2014 Zhang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the National High-tech R&D Program of China (863 Program) under grant 2012AA011601, the National Natural Science Foundation of China under grants 91120305, 61105121 and 61175114, the Natural Science Foundation of Guangdong under grant S2012020010945 and the Excellent Youth Development Project of Universities in Guangdong Province under grant 2012LYM 0057. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: auyqli@scut.edu.cn

Introduction

A central problem in CS is the following: given an $m \times n$ matrix $\mathbf{A}$ ($m < n$), and a measurement vector $\mathbf{y} = \mathbf{A}\mathbf{x}^*$, recover $\mathbf{x}^*$. To deal with this problem, the most extensively studied recovery method is the ℓ_1 -minimization approach (Basis Pursuit) [1–5]

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad s.t. \quad \mathbf{y} = \mathbf{A}\mathbf{x} \quad (1)$$

This convex problem can be solved efficiently; moreover, $\mathcal{O}(\ell \log(n/\ell))$ probabilistic measurements are sufficient for it to recover a ℓ -sparse vector $\mathbf{x}^*$ (i.e., all but at most ℓ entries are zero) exactly.

Recently, Vaswani and Lu [6–9], Miosso [10,11], Wang and Yin [12,13], Friedlander et.al [14], Jacques [15] have shown that exact recovery based on fewer measurements than those needed for the ℓ_1 -minimization approach is possible when the support of $\mathbf{x}^*$ is partially known. The recovery is implemented by solving the optimization problem.

$$\min_{\mathbf{x}} \|\mathbf{x}_{\mathbf{T}^c}\|_1 \quad s.t. \quad \mathbf{y} = \mathbf{A}\mathbf{x} \quad (2)$$

where $\mathbf{T}$ denotes the “known” part of support, $\mathbf{T}^c = [1, \dots, n] \setminus \mathbf{T}$, $\mathbf{x}_{\mathbf{T}^c}$ is a column vector composed of the entries of $\mathbf{x}$ with their indices being in $\mathbf{T}^c$. This method is named modified-CS [6] or truncated ℓ_1 minimization [12]. One application of the modified-CS is the recovery of (time) sequences of sparse signals, such as dynamic magnetic resonance imaging (MRI) [8,9]. Since the support evolve slowly over time, the previously recovered support can be used as known part for later reconstruction.

As an important performance index of modified-CS, its recoverability, i.e., when is the solution of (2) equal to $\mathbf{x}^*$, has been discussed in several papers. In [6], a sufficient condition on the recoverability was obtained based on restricted isometry property. From the view of t -null space property, another sufficient condition to recover ℓ -sparse vectors was proposed in [12]. However, there always exist some signals that do not satisfy these conditions but still can be recovered. Specifically, in real-world applications, the known support often contains some errors. The existing sufficient conditions can not reflect accurately the recoverability of modified-CS in many cases. Therefore, it is necessary to develop alternative techniques for analyzing the recoverability of modified-CS.

In this paper, a sufficient and necessary condition (SNC) on the recoverability of modified-CS is derived. Then, we discuss the recoverability of modified-CS in a probabilistic way. The main advantage of our work is that, for a randomly given vector $\mathbf{x}^*$ with ℓ nonzero entries, the exact recovery percentage of modified-CS can be computed explicitly under a given matrix $\mathbf{A}$ and a randomly given $\mathbf{T}$ that satisfied $|\mathbf{T}|=p$ but includes p_1 errors, where $|\mathbf{T}|$ denotes the size of the known support $\mathbf{T}$. Hence, this paper provides a quantitative index to measure the reliability of modified-CS in real-world applications. Simulation experiments validate our results.

Materials and Methods

1 A Sufficient and Necessary Condition for Exact Recovery

In this subsection, a SNC on the recoverability of modified-CS is derived. Firstly, we give some notations in the follows. The

Figure 1. Probabilities curves obtained in example 1. The horizontal axis represents the sampling numbers. The vertical axis represents the probabilities $\mathbf{P}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A})$ obtained by (6). The three curves from the top to the bottom correspond to $(m, n, \ell, p, p_1) = (7, 9, 4, 2, 1)$, $(48, 128, 20, 8, 2)$ and $(182, 1280, 60, 32, 4)$ respectively. doi:10.1371/journal.pone.0087985.g001

support of vector $\mathbf{x}^* = (x_1^*, \dots, x_n^*)^T$ is denoted by $\mathbf{N}$, i.e. $\mathbf{N} \triangleq \{j | x_j^* \neq 0\}$. Suppose $\mathbf{N}$ can be split as $\mathbf{N} = \mathbf{T} \cup \Delta \Delta_e$, where $\Delta \triangleq \mathbf{N} \setminus \mathbf{T}$ is the unknown part of the support and $\Delta_e \triangleq \mathbf{T} \setminus \mathbf{N}$ is set of errors in the known part support $\mathbf{T}$. The set operations $\cup$ and $\setminus$ stand for set union and set difference respectively. Let $\mathbf{x}^{(1)}$ denote the solution of the model in (2) and $\mathbf{F}$ denote the set of all subsets of Δ . A SNC on the recoverability of modified-CS is given in the following theorem, which is an extension of a result in [16].

Theorem 1 For a given vector $\mathbf{x}^*$, $\mathbf{x}^{(1)} = \mathbf{x}^*$, if and only if $\forall \mathbf{I} \in \mathbf{F}$, the optimal value of the objective function of the following optimization problem is greater than zero, provided that this optimization problem is solvable:

$$\begin{aligned} \min_{\delta} \quad & \left(\sum_{k \in (\mathbf{T}^c \setminus \mathbf{I})} |\delta_k| - \sum_{k \in \mathbf{I}} |\delta_k| \right), \text{ s.t.} \\ & \mathbf{A}\delta = \mathbf{0}, \|\delta\|_1 = 1 \\ & \delta_k x_k^* > 0 \text{ for } k \in \mathbf{I} \\ & \delta_k x_k^* \leq 0 \text{ for } k \in \Delta \setminus \mathbf{I} \end{aligned} \quad (3)$$

where $\delta = (\delta_1, \dots, \delta_n)^T \in \mathbb{R}^n$.

The proof of this theorem is given in Appendix S1.

Remark 1: For a given measurement matrix $\mathbf{A}$, the recoverability of the sparse vector $\mathbf{x}^*$ based on the model in (2) depends only on the index set of nonzeros of $\mathbf{x}^*$ in $\mathbf{T}^c$ and the signs of these nonzeros. In other words, the recoverability relies only on the sign pattern of $\mathbf{x}^*$ in $\mathbf{T}^c$ instead of the magnitudes of these nonzeros.

Remark 2: It follows from the proof of Theorem 1 that, even if $\mathbf{T}$ contains several errors, Theorem 1 still holds.

Remark 3: Recently, the recoverability analysis of the modified-CS were reported in [6] and [12]. However, we establish a sufficient and necessary condition for the modified-CS to exactly reconstructs a sparse vector, which differs from the sufficient conditions proposed in these works.

2 Probability Estimation on Recoverability of Modified-CS

In this subsection, we utilize Theorem 1 to estimate the probability that the vector $\mathbf{x}^*$ can be recovered by modified-CS, i.e., the conditional probability $\mathbf{P}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A})$, where $\|\mathbf{x}^*\|_0$ is defined as the number of nonzero entries of $\mathbf{x}^*$, $|\mathbf{T}|$ and $|\Delta_e|$ denote the size of $\mathbf{T}$ and Δ_e respectively. This probability reflects the recoverability of modified-CS, and is hereafter named as recovery probability.

Figure 2. Comparison of theoretical results (solid curves) and simulation results (dotted curves) on recovery probability. Figure (a) shows the experimental results in part I) and figure (b) shows the ones in part II). In both figures, the three pairs of solid and dotted curves from the top to the bottom correspond to $|\Delta_e|=0, 1, 2$ respectively.
doi:10.1371/journal.pone.0087985.g002

Let $\mathbf{G}$ denote the index set $\{1, 2, \dots, n\}$, it is easy to know that there are $C_n^\ell = \frac{n!}{\ell!(n-\ell)!}$ index subsets of $\mathbf{G}$ with size ℓ . We denote these subsets as $\mathbf{G}_j^{(\ell)}$, $j=1, \dots, C_n^\ell$. For each $\mathbf{G}_j^{(\ell)}$, there are $C_\ell^{p_2}$ subsets with size $p_2 = (p - p_1)$. We denote these subsets as $\mathbf{N}_s^{(p_2)}$, $s=1, \dots, C_\ell^{p_2}$. At the same time, for the set $\mathbf{G} \setminus \mathbf{N}$ (the index set of the zero entries of $\mathbf{x}^*$), there are $C_{n-\ell}^{p_1}$ subsets with size p_1 . These subsets are denoted as $\mathbf{H}_i^{(p_1)}$, $i=1, \dots, C_{n-\ell}^{p_1}$. Firstly, we discuss the estimation of the recovery probability under the following assumption.

Assumption 1 The index set $\mathbf{N}$ of the ℓ nonzero entries of $\mathbf{x}^*$ can be one of the C_n^ℓ index sets $\mathbf{G}_j^{(\ell)}$, $j=1, \dots, C_n^\ell$, with equal probability. The index set Δ_e of p_1 errors in known support can be one of the $C_{n-\ell}^{p_1}$ index sets $\mathbf{H}_i^{(p_1)}$, $i=1, \dots, C_{n-\ell}^{p_1}$, with equal probability. The index set $\mathbf{T} \setminus \Delta_e$ of p_2 nonzero entries can be one of the $C_\ell^{p_2}$ index sets $\mathbf{N}_s^{(p_2)}$, $s=1, \dots, C_\ell^{p_2}$, with equal probability. All the nonzero entries of the vector $\mathbf{x}^*$ take either positive or negative sign with equal probability.

For a given vector $\mathbf{x}^*$ and the known support $\mathbf{T}$, there is a sign column vector $\mathbf{t} = \text{sign}(\mathbf{x}_{\mathbf{T}^c}^*) \in \mathbb{R}^{n-p}$ in $\mathbf{T}^c$. The recoverability of the vector $\mathbf{x}^*$ only relates with the sign column vector $\mathbf{t}$ (see Remark 1). Under the conditions that the index set of the nonzero entries of $\mathbf{x}^*$ is $\mathbf{G}_j^{(\ell)}$ and the known support $\mathbf{T}$ is $\mathbf{N}_s^{(p_2)} \cup \mathbf{H}_i^{(p_1)}$, it is easy to derive that $\mathbf{T}^c$ contains $\ell - p_2$ indexes of the nonzeros of $\mathbf{x}^*$, where $p_2 = (p - p_1)$. Then there are $2^{\ell-p_2}$ sign column vectors. Among these sign column vectors, suppose that $w_{s,i}^j$ sign column vectors

can be recovered, then $\frac{w_{s,i}^j}{2^{\ell-p_2}}$ is the probability of a vector $\mathbf{x}^*$ being recovered by solving the modified-CS. Hence, following Assumption 1, the recovery probability is calculated by

$$\begin{aligned} \mathbf{P}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A}) \\ = \sum_{j=1}^{C_n^\ell} \frac{1}{C_n^\ell} \sum_{s=1}^{C_\ell^{p_2}} \frac{1}{C_\ell^{p_2}} \sum_{i=1}^{C_{n-\ell}^{p_1}} \frac{1}{C_{n-\ell}^{p_1}} \frac{w_{s,i}^j}{2^{\ell-p_2}} \end{aligned} \quad (4)$$

where $\ell=1, \dots, m$, $p=0, \dots, \ell$ and $p_1=0, \dots, p$. Because the measurement matrix $\mathbf{A}$ is known, we can determine $w_{s,i}^j$ in (4) by checking whether the SNC (3) is satisfied for all the $2^{\ell-p_2}$ sign column vectors corresponding to the index set $\mathbf{G}_j^{(\ell)}$, $\mathbf{H}_i^{(p_1)}$ and $\mathbf{N}_s^{(p_2)}$.

Because many practical situations such as Electroencephalogram (EEG) signals in wavelet domain do not completely satisfy those assumptions in “Assumption 1”, we further extend our analysis to more general case. Without loss of generality, we have the following assumption

Assumption 2 The index set $\mathbf{N}$ of the ℓ nonzero entries of $\mathbf{x}^*$ can be one of the C_n^ℓ index sets $\mathbf{G}_j^{(\ell)}$, $j=1, \dots, C_n^\ell$, with probability P_j . The index set Δ_e of p_1 errors in known support can be one of the $C_{n-\ell}^{p_1}$ index sets $\mathbf{H}_i^{(p_1)}$, $i=1, \dots, C_{n-\ell}^{p_1}$, with probability P_i . The index set $\mathbf{T} \setminus \Delta_e$ of p_2 nonzero entries can be one of the $C_\ell^{p_2}$ index sets $\mathbf{N}_s^{(p_2)}$, $s=1, \dots, C_\ell^{p_2}$, with probability

Figure 3. An original segment of record No. 100 in the MIT-BIH arrhythmia database and the reconstructed one in time and wavelet domain. Figures (a) and (b) show the original and the reconstructed one respectively.
doi:10.1371/journal.pone.0087985.g003

P_s . All the nonzero entries of the vector $\mathbf{x}^*$ take either positive or negative sign with probability P_+ or P_- respectively.

Similarly, suppose the index set of the nonzero entries of $\mathbf{x}^*$ is $\mathbf{G}_j^{(\ell)}$ and the known support is $\mathbf{N}_s^{(p_2)} \cup \mathbf{H}_i^{(p_1)}$, there are $2^{\ell-p_2}$ sign column vectors. Since all the nonzero entries of the vector $\mathbf{x}^*$ take either positive or negative sign with probability P_+ or P_- respectively, the probability of the sign pattern of vector $\mathbf{x}^*$ equals one of $2^{\ell-p_2}$ sign column vectors is $(P_-)^k (P_+)^{\ell-p_2-k}$, where k denotes the number of negative signs in this sign column vector and $0 \leq k \leq \ell-p_2$. Obviously, there are $C_{\ell-p_2}^k$ sign column vectors that has k negative signs. Among these vectors, suppose that $w_{s,i,k}^j$ sign column vectors can be recovered, then $\sum_{k=0}^{\ell-p_2} ((P_-)^k (P_+)^{\ell-p_2-k} \cdot w_{s,i,k}^j)$ is the probability of the vector $\mathbf{x}^*$ being recovered by solving the modified-CS. Hence, under the Assumption 2, the recovery probability is calculated by

$$\mathbf{P}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A}) = \sum_{j=1}^{C_n^\ell} P_j \sum_{s=1}^{C_\ell^{p_2}} P_s \sum_{i=1}^{C_{n-\ell}^{p_1}} P_i \left(\sum_{k=0}^{\ell-p_2} ((P_-)^k (P_+)^{\ell-p_2-k} \cdot w_{s,i,k}^j) \right) \quad (5)$$

where $\ell = 1, \dots, m$, $p = 0, \dots, \ell$ and $p_1 = 0, \dots, p$. Because the measurement matrix $\mathbf{A}$ is known, we can determine $w_{s,i,k}^j$ in (5), $0 \leq k \leq \ell-p_2$, by checking whether the SNC (3) is satisfied for all the $C_{\ell-p_2}^k$ sign column vectors corresponding to the index set $\mathbf{G}_j^{(\ell)}$, $\mathbf{H}_i^{(p_1)}$ and $\mathbf{N}_s^{(p_2)}$.

Remark 4: Equation (4) is a special case of equation (5) under the equal probability assumption.

However, the computational burden to calculate (5) increases exponentially as the problem dimensions increase. For each sign column vector $\mathbf{t}_\tau$ and the corresponding index set $\mathbf{G}_j^{(\ell)}$, $\mathbf{H}_i^{(p_1)}$ and $\mathbf{N}_s^{(p_2)}$, we denote the quads $[\mathbf{G}_j^{(\ell)}, \mathbf{N}_s^{(p_2)}, \mathbf{H}_i^{(p_1)}, \mathbf{t}_\tau]$, where $j = 1, \dots, C_n^\ell$, $s = 1, \dots, C_\ell^{p_2}$, $i = 1, \dots, C_{n-\ell}^{p_1}$ and $\tau = 1, \dots, 2^{\ell-p_2}$. Suppose $\mathbf{Z}$ is a set composed by all the quads, there are $C_n^\ell C_\ell^{p_2} C_{n-\ell}^{p_1} 2^{\ell-p_2}$ elements in set $\mathbf{Z}$. For each element of set $\mathbf{Z}$, if the sign column vector $\mathbf{t}_\tau$ can be recovered by modified-CS with a given measurement matrix $\mathbf{A}$ and known support $\mathbf{T} = \mathbf{N}_s^{(p_2)} \cup \mathbf{H}_i^{(p_1)}$, we call the quad can be recovered. In (5), the estimation of recover probability need to check the total number of quads in set $\mathbf{Z}$. When n increases, the computational burden will increase exponentially. To avoid the computational burden problem, we state the following Theorem.

Theorem 2 Suppose that M quads are randomly taken from set $\mathbf{Z}$, where M is a large positive integer ($M \ll C_n^\ell C_\ell^{p_2} C_{n-\ell}^{p_1} 2^{\ell-p_2}$), and K of the M quads can be recovered by solving modified-CS. Then

$$\mathbf{P}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A}) \simeq \frac{K}{M} \quad (6)$$

The proof of this theorem is given in Appendix S2.

Remark 5: In real-world applications, by sampling randomly M sign vectors with ℓ nonzero entries, we can check the number of the vectors that can be exact recovered by modified-CS with a random known support $\mathbf{T}$ whose size is p but contains p_1 errors. Suppose K sign vectors can be recovered, the recovery probability

Figure 4. Probabilities curves obtained in example 3. The horizontal axis of each subfig represents the sampling numbers. The vertical axis represents the recovery probabilities. Red curves are the theoretic probability curves; Blue curves are the practical probability curves.
doi:10.1371/journal.pone.0087985.g004

$\mathbf{P}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A})$ can be computed approximately through calculating the ratio of K/M .

Remark 6: It is well-known that certifying the restricted isometry property is hard, while based on the proposed method, the recoverability probability that reflects the recoverability of modified-CS can be computed explicitly.

From the proof of Theorem 2, the sampling numbers M , which controls the precision in the approximation of (6), is related to the two-point distribution of v_k other than the size of $\mathbf{Z}$. Thus, there is no need for M increasing exponentially as n increases.

Results and Discussion

In this section, simulation examples on both synthesis data and real-world data have been conducted to demonstrate the validity of our theoretical results.

Example 1: In this example, the conclusion in Theorem 2 are demonstrated.

According to the uniform distribution in $[-0.5, 0.5]$, we randomly generate three matrices $\mathbf{A}_i \in \mathbb{R}^{m \times n}$ ($i = 1, 2, 3$) with $(m, n) = (7, 9), (48, 128)$ and $(182, 1280)$ respectively. For matrices $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{A}_3$, we set $(\ell, p, p_1) = (4, 2, 1), (20, 8, 2)$ and $(60, 32, 4)$ respectively. As n increases in their three cases, the number of sign vectors increases exponentially. For example, for $(m, n, \ell, p, p_1) = (7, 9, 4, 2, 1)$, $(48, 128, 20, 8, 2)$, the set $\mathbf{Z}$ contains approximately 2×10^4 and 7×10^{36} elements respectively. Hence, for their three cases, we estimate the probabilities

$\mathbf{P}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A})$ by the sampling method. For each case, we sample $M = 100, 500, 1000, 5000, 10000$ respectively. The resultant probability estimates depicted in Fig. 1 indicate that 1) the estimation precision of the sampling method is stable in our experiments with different samplings. Therefore, we only need a very few samplings to obtain the satisfied estimation precision in real-world applications; 2) as n increases in three cases, the sampling M don't increase exponentially.

Example 2: Suppose $\mathbf{A} \in \mathbb{R}^{7 \times 9}$ was taken according to the uniform distribution in $[-0.5, 0.5]$. This example contains two parts in which the recovery probability estimates (4) and (5) are considered in simulation, respectively.

- (I) All nonzero entries of the sparse vector $\mathbf{x}^*$ were drawn from a uniform distribution valued in the range $[-1, +1]$. Without loss of generality, we set $p = 2$. For a vector $\mathbf{x}^*$ with ℓ nonzero entries, where $\ell = 2, 3, \dots, 7$, we calculated the recovery probabilities by (4), where $p_1 = 0, 1, 2$ respectively. For every ℓ ($\ell = 2, \dots, 7$) nonzero entries, we also sampled 1000 vectors with random indices. For each vector, we solved the modified-CS with a randomly given $\mathbf{T}$, whose size equals to p but contains p_1 errors, and checked whether the solution is equal to the true vector. Suppose that n_p^ℓ vectors can be recovered, we calculated the ratio $p_p^\ell = \frac{n_p^\ell}{1000}$ as the recovery probability $\hat{\mathbf{P}}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A})$. The experimental results are presented in

Fig. 2(a). Therein, solid curves denote the theoretic recovery probability estimated by (4). Dotted curves denote probabilities $\hat{\mathbf{P}}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A})$. Experimental results show that the theoretical estimates fit the simulated values very well.

- (II) Now we consider the probability estimate (5). We suppose that all nonzero entries of the sparse vector $\mathbf{x}^*$ were drawn from a uniform distribution valued in the range $[-0.5, +1]$. Obviously, the nonzero entries of the vector $\mathbf{x}^*$ take the positive sign with probability $2/3$ or the negative sign with probability $1/3$. Similarly, we set $p = 2$. For a vector $\mathbf{x}^*$ with ℓ nonzero entries, where $\ell = 2, 3, \dots, 7$, we randomly generate the probabilities P_j where $j = 1, \dots, C_n^{\ell}$. For an index set $\mathbf{T}$ whose size equals to p but contains p_1 errors, we randomly generate the probabilities P_i where $i = 1, \dots, C_{n-\ell}^{p_1}$ and P_s where $s = 1, \dots, C_{\ell}^{p_2}$. The recovery probabilities are calculated by (5), where $p_1 = 0, 1, 2$ respectively. For every ℓ ($\ell = 2, \dots, 7$) nonzero entries, we also sampled 1000 vectors $\mathbf{x}^*$ and $\mathbf{T}$ that satisfy the assumption 2 with the above-generated probabilities. For each vector and $\mathbf{T}$, we solved the modified-CS and checked whether the solution is equal to the true vector. Finally, the $\hat{\mathbf{P}}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A})$ can be calculated with the same way in the part I. We present the experimental results in Fig. 2(b). Therein, solid curves denote the theoretic recovery probability estimated by (5). Dotted curves denote probabilities $\hat{\mathbf{P}}(\mathbf{x}^{(1)} = \mathbf{x}^*; \|\mathbf{x}^*\|_0 = \ell, |\mathbf{T}| = p, |\Delta_e| = p_1, \mathbf{A})$. Experimental results show that in the general case, the theoretical estimates also fit the simulated values very well.

Example 3: In this example, we test on real-world ECG reconstruction to demonstrate the accuracy of the probability estimation by (5).

Firstly, eight ECG data have been chosen from the MIT-BIH arrhythmia database [17] as the test signals. Each data file includes two-channel ambulatory ECG recordings, and each channel contains 650000 binary data instances in a 16-bits data format, including the index and amplitude. In our simulation, ECG vector s is extracted from the original data at the window size $n = 256$. A random sparse binary matrix [18] is used as our sensing matrix Φ and we use D6 Daubechies wavelet dictionary Ψ to represent ECG segment, i.e.,

$$y = \Phi s = \Phi \Psi x = \Lambda x \quad (7)$$

It is well-known that vector $\mathbf{x}$ is not strictly sparse, but can be approximated by ℓ -sparse vector. Therefore, to obtain the ℓ -sparse approximation $\tilde{\mathbf{x}}$ of vector $\mathbf{x}$, we calculate the standard derivation σ of the high-frequency coefficients in vector $\mathbf{x}$ and shrink the coefficients whose magnitudes are less than 3σ to zero. We define the theoretic recovery of ECG segment $\mathbf{x}$ as the SNC (3) can be satisfied for the sign pattern $\text{sign}(\tilde{\mathbf{x}})$. On the other hand, for the recovery of a compressible vector $\mathbf{x}$, the best one can expect is that

the solution of modified-CS and $\mathbf{x}$ have their nonzero components at the same locations [19]. Considering the noise contamination, we think the ECG segment $\mathbf{x}$ is recovered in practice if the solution of modified-CS and $\mathbf{x}$ have the overwhelming majority (e.g. 95%) of their nonzero components at the same locations.

Hence, we randomly extracted 100 segments from each ECG data. According to the Theorem 2, we can estimate the recovery probability of modified-CS through calculating how many segments can be recovered, i.e., the recovery ratio. In our experiment, we suppose $\mathbf{T}$ is the index set of low-frequency coefficients in vector $\mathbf{x}$. On the one hand, we check the SNC (3) for all the sign patterns $\text{sign}(\tilde{\mathbf{x}})$ of these segments to obtain the theoretic recovery ratio; on the other hand, we obtain the practical recovery ratio by checking whether these ECG segments can be recovered in practice. For illustration, an original segment of record No. 100 in the MIT-BIH arrhythmia database and its wavelet coefficients are plotted in Fig. 3(a). At the same time, the reconstructed ones in time and wavelet domain are shown in Fig. 3(b). We present the experimental results of eight ECG data in Fig. 4. Therein, red curves denote the theoretic recovery probabilities. Blue curves denote practical recovery probabilities. Experimental results show that the proposed probability estimation is very accurate.

Conclusion

In this letter we study the recoverability of the modified-CS in a stochastic framework. A sufficient and necessary condition on the recoverability is presented. Based on this condition, the recovery probability of the modified-CS can be estimated explicitly. It is worth mentioning that Theorem 1 can be easy to extend to weighted- ℓ_1 minimization approach that was proposed in [20] for nonuniform sparse model. Moreover, the recovery probability estimation provides alternative way to find (numerically) the optimal set of weights in weighted- ℓ_1 minimization approach, which has the largest recovery probability to recover the signals.

Supporting Information

Appendix S1 Proof of Theorem 1.
(PDF)

Appendix S2 Proof of Theorem 2.
(PDF)

Acknowledgments

The authors would like to thank anonymous reviewers and Academic Editor for the insightful and constructive suggestions.

Author Contributions

Conceived and designed the experiments: JZ YL. Performed the experiments: JZ ZG. Analyzed the data: JZ ZY. Contributed reagents/materials/analysis tools: JZ YL ZG ZY. Wrote the paper: JZ YL.

References

- Donoho DL (2006) Compressed sensing. *IEEE Trans Inf Theory* 52: 1289–1306.
- Candès EJ, Romberg J, Tao T (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans Inf Theory* 52: 489–509.
- Candès EJ, Tao T (2006) Near-optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans Inf Theory* 52: 5406–5425.
- Chen SS, Donoho DL, Saunders MA (2001) Atomic decomposition by basis pursuit. *SIAM REV* 43: 129–159.
- Candès E, Tao T (2005) Decoding by linear programming. *IEEE Trans Inf Theory* 51: 4203–4215.
- Vaswani N, Lu W (2010) Modified-CS: modifying compressive sensing for problems with partially known support. *IEEE Trans Signal Process* 58: 4595–4607.
- Vaswani N, Lu W (2009) Modified-CS: modifying compressive sensing for problems with partially known support. In: *Proc IEEE Int Symp Inf Theory (ISIT)*. pp. 488–492.
- Lu W, Vaswani N (2009) Modified compressive sensing for real-time dynamic mr imaging. In: *Proc IEEE Int Conf Image Proc (ICIP)*. pp. 3009–3012.

9. Qiu C, Lu W, Vaswani N (2009) Real-time dynamic mri reconstruction using kalman filtered cs. In: Proc Int Conf Acoustics, Speech, Signal Processing(I-CASSP). pp. 393–396.
10. Miosso C, von Borries R, Argez M, Velazquez L, Quintero C, et al. (2009) Compressive sensing reconstruction with prior information by iteratively reweighted least-squares. *IEEE Trans Signal Process* 57: 2424–2431.
11. von Borries R, Miosso C, Potes C (2007) Compressed sensing using prior information. In: Computational Advances in Multi-Sensor Adaptive Processing, 2007. CAMPSAP 2007. 2nd IEEE International Workshop on. pp. 121–124.
12. Wang Y, Yin W (2010) sparse signal reconstruction via iterative support detection. *SIAM J Imag Sci* 3: 462–491.
13. Guo W, Yin W (2010) Edgcs: edge guided compressive sensing reconstruction. Technical Report TR10-02, Department of Computational and Applied Mathematics, Rice University, Houston, TX.
14. Friedlander M, Mansour H, Saab R, Yilmaz O (2010) Recovering compressively sampled signals using partial support information. to appear in the *IEEE Trans Inf Theory*.
15. Jacques L (2010) A short note on compressed sensing with partially known signal support. *Signal Processing* 90: 3308–3312.
16. Li Y, Amari S, Cichocki A, Guan C (2006) Probability estimation for recoverability analysis of blind source separation based on sparse representation. *IEEE Trans Inf Theory* 52: 3139–3152.
17. Moody G, Mark R (2001) The impact of the mit-bih arrhythmia database. *IEEE Eng Med Biol Mag* 20: 45–50.
18. Mamaghanian H, Khaled N, Atienza D, Vanderghenst P (2011) Compressed sensing for realtime energy-efficient ecg compression on wireless body sensor nodes. *IEEE Trans Biomed Eng* 58: 2456–2466.
19. Fuchs JJ (2005) Recovery of exact sparse representations in the presence of bounded noise. *IEEE Trans Inf Theory* 51: 3601–3608.
20. Khajehnejad M, Xu W, Avestimehr S, Hassibi B (2011) Analyzing weighted ℓ_1 minimization for sparse recovery with nonuniform sparse models. *IEEE Trans Signal Process* 59: 1985–2001.