Identifying Bad Measurements in Compressive Sensing

H. T. Kung, Tsung-Han Lin, Dario Vlah
Harvard University
Cambridge, MA

***Abstract*—We consider the problem of identifying bad measurements in compressive sensing. These bad measurements can be present due to malicious attacks and system malfunction. Since the system of linear equations in compressive sensing is underconstrained, errors introduced by these bad measurements can result in large changes in decoded solutions. We describe methods for identifying bad measurements so that they can be removed before decoding. In a new separation-based method we separate out top nonzero variables by ranking, eliminate the remaining variables from the system of equations, and then solve the reduced overconstrained problem to identify bad measurements. Comparing to prior methods based on direct or joint ℓ_1-minimization, the separation-based method can work under a much smaller number of measurements. In analyzing the method we introduce the notion of inversions which governs the separability of large nonzero variables.**

I. Introduction

Compressive sensing has emerged as a major research area due to, among others, the surprising property that sub-Nyquist sampling can capture the information present in a sparse signal. In general, this is made possible by having each measurement be some incoherent linear combination of the signal, thus ensuring that sparse signal components contribute to the measurement with high probability.

A standard compressive sensing formulation is as follows:

$$y = \mathbf{A}x$$

where x is an N-dimensional vector representing the sparse signal being sampled, $\mathbf{A}$ is an $M \times N$ measurement matrix containing random entries, and y is a vector of M measurements which are random linear combinations of components of x. Typically, $M \ll N$, so this is an underconstrained system that does not have a unique solution for general x. Nevertheless, suppose that x is K-sparse in the sense that it can be expressed as a linear combination of K basis vectors in some basis with K being a fraction of M. Then it is possible to recover x with high probability using a random measurement matrix $\mathbf{A}$. A rich volume of literature examines this topic starting with the seminal work of Candès and Tao [1].

Clearly, given the underconstrained nature of the above formulation the M measurements in the vector y are, in some sense, information-dense. As a result, perturbations of y can severely impact the successful reconstruction of the signal x. In particular, in this paper we focus on coping with such perturbations when just a few of the elements of y exhibit arbitrary error. Such cases could occur when, for example, the compressive measurements are being collected in a distributed system and a small fraction of the nodes either malfunction or have been compromised to output incorrect values in order to interfere with the decoding process.

Note that when the errors are substantial, the application can detect their existence for the following reason. Because the decoded solution will not yield sufficiently many zero or near-zero components, the application should be able to notice the lack of expected sparsity in the decoded solution, and conclude it is not trustworthy.

However, when errors are not large enough to allow easy observation of the decoded solutions on the existence of bad measurements, detecting and identifying them can be challenging. In this case, since the compressive sensing system is underconstrained, the incorrect solution led by bad measurements can be anywhere in a $(N-(M-h))$-dimensional space, where h is the number of bad measurements. This makes the problem of identifying bad measurements difficult.

In this paper, we show that we can identify bad measurements for the case where there are only relatively few of them. (Throughout the paper, we use the terms "bad samples", "bad measurements" and "bad equations" interchangeably.) We describe a method for doing so, called the separation-based method, and compare it with prior methods based on direct or joint ℓ_1-minimization.

For performance analysis, we have identified a key parameter, called the number of inversions. We can estimate inversions via empirical sampling and use the resulting estimate to configure a separation-based method. In practice, this almost always leads to successful identification of bad measurements.

II. Background

In this section we introduce some notation, review the standard ℓ_1-minimization method, and describe two relevant kinds of error.

A. Notation

We will use the following notation to write down the standard compressive sensing formulation:

$$y = \mathbf{A}x \tag{1}$$

Here, $y \in \mathbb{R}^M$ are the compressive samples, $\mathbf{A} \in \mathbb{R}^{M,N}$ is the random measurement matrix such that $M \ll N$, and $x \in \mathbb{R}^N$ is the original sparse input to be reconstructed. We will say

This material is based on research sponsored by Air Force Research Laboratory under agreement number FA8750-10-2-0180. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory or the U.S. Government.

that x is (K, ε)-sparse if only K entries of x have a magnitude greater than ε. Thus, with $\varepsilon > 0$ we can describe vectors which are only *approximately sparse*. On the other hand, for $\varepsilon = 0$, x will have exactly K or fewer nonzero entries; for those cases we will just say that x is K-sparse or that x is *exactly sparse*. For brevity, for an exactly or approximately sparse signal x, we will often use the term *nonzeros* to refer to the K largest entries of x, even if for the latter case there are other smaller elements which are not equal to zero. Finally, given an arbitrary vector x, we will write x_K to denote a "truncated" version of x, where all except the K largest-magnitude entries have been set to 0.

We adopt the standard solution method via ℓ_1-minimization, as follows:

$$x^* = \arg\min_{y=\mathbf{A}x} \|x\|_{\ell_1} \tag{2}$$

It has been shown that the ℓ_1-reconstruction of K-sparse signals is exact with high probability if

$$M > CK \log \frac{N}{K} \tag{3}$$

for some small constant C [2]. For example, in practice, $C = 1.7$ with $\log_2$ gives a probability of decoding failure of less than 0.1%.

B. Truncation Error

Consider a (K, ε)-sparse signal x, with some $\varepsilon > 0$. Reconstructing this approximately sparse x exactly is not possible with the underconstrained system of $M < N$ measurements. This means that the reconstruction x^* will inevitably contain some error induced by the minimization in (2) treating the target vector x as K-sparse, that is, by truncating it so that some of its entries with magnitudes $\leq \varepsilon$ become zeros. Errors in x^* resulting from this truncation are bounded; indeed, from compressive sensing literature [3] we have

$$\|x^* - x\|_{\ell_2} < \frac{C_0}{\sqrt{K}} \|x - x_K\|_{\ell_1} \tag{4}$$

with C_0 being some constant dependent on properties of the measurement matrix $\mathbf{A}$. Note that use of larger M will allow us to use larger K, tightening the bound in (4) and thus reducing $\|x^* - x\|_{\ell_2}$.

C. Direct ℓ_1-minimization

We consider the case where h individual elements of y, the compressive samples, are corrupted for some small h. We then have the following formulation with the error vector e being h-sparse:

$$y = \mathbf{A}x + e \tag{5}$$

Candès et al. [4] showed that in the presence of noise, solving the linear system with solely ℓ_1-minimization can still give a reasonable solution. They show that the magnitude of the solution error is bounded proportionally to the noise level as follows:

$$\|x^* - x\|_{\ell_2} \leq \frac{C_0}{\sqrt{K}} \|x - x_K\|_{\ell_1} + C_1 \|e\|_{\ell_2} \tag{6}$$

where C_0 and C_1 are some constants. This suggests use of conventional ℓ_1-minimization, which we refer to as *direct ℓ_1-minimization* in this paper, for scenarios where small additive noise is introduced in every measurement, such as the measurement noise from the sensing devices or the quantization error. However, for other scenarios where even just few measurements are corrupted due to malfunctioning sensors, corrupted packets, or malicious attacks the upper bound in (6) may still be arbitrarily large.

D. Errors Introduced by Bad Measurements

Under direct ℓ_1-minimization, any amount of error introduced by bad measurements will cause distortion in decoded solutions. As an illustration, the red curve in Figure 1(a) shows the effect of a single nonzero entry in the error vector e, that is, a single bad measurement. We obtained the compressive sensing solutions for $N = 200, M = 50, K = 10$, with all nonzero elements in the input set to a value of 100. For the error metric we used the average of the displacements of computed nonzero element values normalized by their respective original values.

The effect of this type of error will depend on the application. If the application's goal is to recover exact values of the nonzero inputs, any amount of error will limit the precision of the recovered values. On the other hand, suppose that the objective is to merely identify the locations of the nonzero elements; for example, they could represent bad measurements (see Section III-B) or a small number of sensors in a large population detecting certain exceptional conditions. A straightforward method to perform the identification is to admit any values exceeding some threshold. The red curves in Figures 1(b) and 1(c) show how a single bad equation affects such identification success under direct ℓ_1-minimization, given same parameters as in the previous example, using an identification threshold of 50. We can see that when error approaches 50 or 100, the solution begins to fail to identify all nonzeros, or begin to identify incorrect nonzeros, respectively.

Therefore, to tackle the error problems caused by bad measurements, one should identify and remove these measurements from the system to allow solutions with better accuracy.

III. Two Methods

In this section, we will discuss two approaches to detect and recover errors stemming from bad measurements. The joint ℓ_1-minimization method utilizes the fact that because we assume that there are only few bad measurements, the error vector is *sparse*. On the other hand, the separation-based method aims to separate out sufficiently many zero variables so they can be eliminated from the measurements, resulting in an overconstrained system with which the error vector can be solved.

A. Joint ℓ_1-minimization Method

A method has been described in literature to deal with sparse errors in compressive sensing [5], based on joint ℓ_1-minimization for solving both vectors x and e simultaneously. The method works by expressing the formulation in (5) as follows:

$$y = \begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix} \begin{bmatrix} x \\ e \end{bmatrix} \tag{7}$$

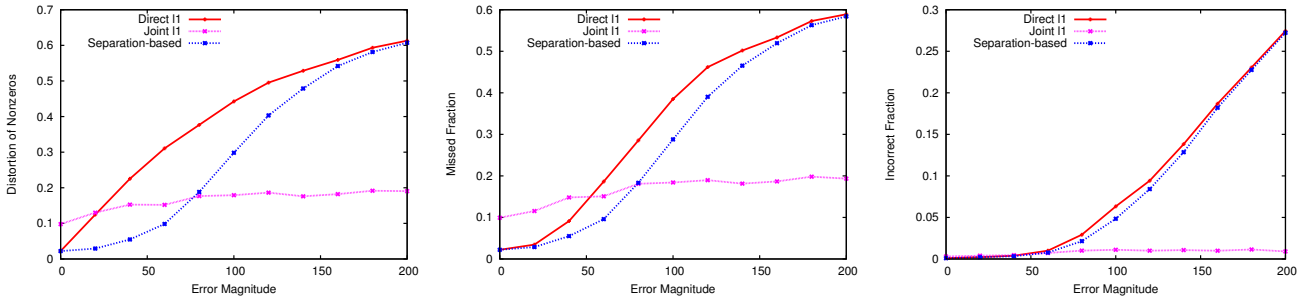

Fig. 1. Effects under increasing error magnitude of a single bad measurement: (a) relative displacement of nonzeros, (b) missed identification of nonzeros, and (c) incorrect identification of nonzeros. This example assumes that $N = 200$, $M = 50$ and $K = 10$, with the value of nonzeros equal to 100.

That is, by using the fact that e is sparse, we can treat it as part of the input, and reconstruct it using the error-free formulation in (1).

This method has been shown to work in numerical experiments [5]. Later, it was proven to work by showing that the joint measurement matrix $[\ \mathbf{A}\ \mathbf{I}\ ]$ satisfies a sufficient condition, restricted isometry property (RIP), which guarantees that the minimization in (2) succeeds with high probability, although under slightly worse bounds than a purely random matrix of same dimension [6].

B. Separation-based Method

We now describe an alternative method for dealing with bad measurements called the *separation-based method,* which works in three steps: first, it reduces the original underconstrained system to a smaller, overconstrained system by eliminating sufficiently many zero variables. Then, it transforms that system into an underconstrained system for e, and solves it using ℓ_1-minimization. Finally, once we know e, we eliminate it from the original system and solve for x. We describe these steps in detail below for K-sparse x. We will extend the method to approximately K-sparse x in Section V.

- **Step I: Separating nonzeros.** We begin with the system of measurements in (5), which is an underconstrained $M \times N$ system, and use ℓ_1-minimization to obtain x^*. From there, we extract a subset $\mathcal{L}$ of $L < M$ indices which we make sure (see discussion in Section IV) will include the K nonzeros in x with high probability; namely, we construct $\mathcal{L}$ by taking the L elements of x^* with the largest magnitudes. Note that elements of x at indices not in $\mathcal{L}$ are all zeros for K-sparse x; thus, they don't contribute any weight to the measurement vector y and thus can be removed from the system of equations. We now have a new overconstrained system

$$y = \mathbf{B}x' + e \tag{8}$$

  where $\mathbf{B}$ is a new $M \times L$ matrix composed of $\mathbf{A}$'s columns $\{a_i : i \in \mathcal{L}\}$, and, correspondingly, x' is a new vector of L unknowns $\{x_i : i \in \mathcal{L}\}$, which we refer to as "left-over" variables. The parameter L is chosen so that we have a good probability of finding the original K nonzero elements among the L largest-magnitude elements of x^*. Thus, at the very least, $L \geq K$.

- **Step II: Identifying Bad Measurements.** The overconstrained system in (8) with a sparse error vector e has been studied previously in the context of error correction [1]. The method reported there is not directly applicable to our problem because it targets general non-sparse signals and thus needs more measurements to identify errors. However, after Step I is performed and the system becomes overconstrained, we can use the method to find e. First, we construct a $(M - L) \times M$ matrix $\mathbf{F}$ whose null-space is the range of $\mathbf{B}$, so that $\mathbf{FB} = 0$. Then we multiply both sides of (8) by $\mathbf{F}$ to get

$$\mathbf{F}y = \mathbf{FB}x' + \mathbf{F}e = \mathbf{F}e \tag{9}$$

  Rewriting the left-hand side for clarity as $\tilde{y}$ we get

$$\tilde{y} = \mathbf{F}e \tag{10}$$

  which is an underconstrained system. We can view (10) as a compressive sensing of sparse vector e, and, noting $\mathbf{F}$ satisfies RIP, we solve e via ℓ_1-minimization.
- **Step III: Solving Error-free System.** Clearly, once we know e, we can eliminate it from the right-hand side of (5) and solve the resulting system using standard ℓ_1-minimization. Alternatively, we can use the locations of the nonzero elements in e to identify bad equations, remove these equations in (5), and solve the resulting system. (The latter approach is useful for the approximately sparse case of Section V.)

Figure 1(a) shows the simulation results by repeating the same experiments in Section II-C with the two methods. The separation-based method improves the performance of the direct ℓ_1 method when the error magnitude is small. When errors are large, their performance becomes similar. Note that for the separation-based method, the curves of nonzero distortion and missed fraction on nonzeros look very similar. This is due to the fact that when the method misses a nonzero variable, it will treat it as a zero variable. In this case, the method would eliminate the component by setting it to zero, causing a distortion in the decoded solution. Thus the relative distortion is the same as the missed fraction. In comparison with the joint-ℓ_1 method, we observe that the separation-based method performs better when the magnitude of errors is smaller.

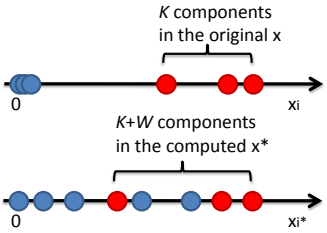

Fig. 2. Magnitudes of x's components, with computed x^* exhibiting inversion under error. Top: The original x with blue nodes denoting zero-valued variables and red nodes denoting the K nonzero variables. Bottom: When solving the system using ℓ_1-minimization, the error will displace the blue and red nodes from the correct solution, and may result in some blue nodes getting mixed into the red nodes. The number of blue nodes larger than the smallest red node is the number of inversions W.

IV. Analysis for Exactly Sparse Signal Using The Notion of Inversions

In this section, we analyze the performance of the separation-based method by introducing the notion of *inversions*.

We assume the signal x to be recovered is exactly sparse with K nonzeros. That is, components in x other than the K largest ones are all zero. The separation-based method aims to separate and eliminate some of these zero components in order to reduce the total number of unknowns. Therefore, separability of zero components using the computed x^* is a key to its performance. Figure 2 illustrates the component magnitudes of a K-sparse x vector and a computed x^* vector after solving x directly via ℓ_1-minimization. The figure shows the value of every component in x and x^*. We denote the K nonzero components as the red nodes and the other $N-K$ zero components as the blue nodes. Due to the errors from the bad equations, or insufficient good equations, the component values in x^* are displaced from their original values. As the figure shows, some blue nodes may end up with larger magnitudes than some red nodes, resulting in separation difficulty. We call the number of blue nodes that have larger magnitudes than the smallest red node *the number of inversions*, W, or simply inversions.

The value W determines how well we can separate the components in Step I. Note that all the K nonzero components need to be in the set $\mathcal{L}$ of L extracted components. Otherwise, incorrectly dropped the nonzero terms would result in wrong constraints to the system of equations. Therefore, L has to be sufficiently large to ensure

$$L \geq K + W \tag{11}$$

On the other hand, we want to turn the system of equations into an overconstrained one involving the L left-over variables, and later use ℓ_1-minimization to identify errors. Therefore we also need to ensure that (10) can be solved with ℓ_1-minimization. Since the $\mathbf{F}$ matrix has dimension $(M-L)\times M$, we know from (3) that condition (12) below ensures error recovery with high probability:

$$M - L > O(h\log(M/h)) \tag{12}$$

Combining (11) and (12) yields a lower bound on the number M of measurements required for the separation-based method:

$$M > K + W + O(h\log(M/h)) \tag{13}$$

In addition, to ensure that we can solve Step I with ℓ_1-minimization, we need to bound M by the sparsity of x:

$$M > O(K\log(N/K)) \tag{14}$$

From (13), we can see that the number of measurements required for error detection is a function of W. Fortunately W can be small and is related to how "separated" the nonzero components in x are away from zero. Here we need to introduce another parameter α, referred to as "nonzeros floor," to denote the magnitude of the smallest nonzero component in x. That is, α represents the degree of separability of x's nonzero components from zero. As long as α is large enough relative to e, separation by rank is possible.

To see this clearly, as shown in (6), the ℓ_2 distance between x and the reconstructed x^* is bounded by the error magnitude. In other words, the displacement of every component x_i in x is also bounded:

$$|x_i^* - x_i| < \|x^* - x\|_{\ell_2} < C_1\|e\|_{\ell_2} \tag{15}$$

Hence when α is sufficiently large relative to e, there will be no inversions, that is, $W=0$. In this case, those components which assume the K largest magnitudes in the computed x^* are precisely those K nonzero variables in the original x.

V. Extension to Approximately Sparse Signal

We show that the separation-based method is applicable to approximately-sparse signals. Recall that such a (K,ε)-sparse signal x may have components with small magnitudes relative to the K-leading components.

We first argue that small or zero W can be ensured so the separation is possible. We know from (6) that the displacement of components in the computed x^* is still bounded even if x is approximately sparse:

$$|x_i^* - x_i| < \|x^* - x\|_{\ell_2} < C_1\|e\|_{\ell_2} + \frac{C_0\|x_K - x\|_{\ell_1}}{\sqrt{K}} \tag{16}$$

Therefore, following the same argument as before, W is guaranteed to be zero given sufficiently large α.

However, additional errors are introduced in eliminating the separated components with small magnitudes. This is due to the fact that x is (K,ε)-sparse, not exactly sparse; so the eliminated components are small but not zero. Denoting the elimination error as e_E, we can rewrite (8) after elimination of the variables as an overconstrained system:

$$y = \mathbf{B}x' + e + e_E \tag{17}$$

and

$$e_E = \sum_{j\notin\mathcal{L}} a_j x_j \tag{18}$$

where a_j is the j-th column of the measurement matrix $\mathbf{A}$.

Similarly, we can rewrite (10) in Step II as

$$\tilde{y} = \mathbf{F}(e + e_E) = \mathbf{F}e' \tag{19}$$

Note that e' is not sparse in general. Thus at first glance it may appear that we cannot solve (19) by ℓ_1-minimization. Nevertheless, our goal is to know the locations of nonzero components in e instead of recovering e' exactly. ℓ_1-minimization may still be applicable because it can separate out nonzeros which are distinguishably larger than other components. For this, we need to assume that the entries e_i of e constitute the largest components in e' and that they are significantly larger in their magnitudes than other components. That is,

$$|e_i| > \frac{2C_0 \|e_E\|_{\ell_1}}{\sqrt{h}}, \ \forall i = 1, 2, ..., m, \ e_i \neq 0 \qquad (20)$$

The condition follows from (4), and the constant 2 is to avoid e_i canceling e_E due to sign difference. We argue that the assumption in (20) is reasonable. If e has magnitudes similar to the small-valued components in x, one will not be able to observe the effect of e in the solution and thus in this case there is no need to remove the bad equations.

VI. Contrasting the Two Methods and Application Strategy

Finally, let us compare the joint ℓ_1-minimization method and the separation-based method. The joint ℓ_1-minimization method has the advantage that it only makes assumptions on the sparsity of e. Thus it can recover any sparse e regardless of its magnitude. In contrast, the separation-based method needs to have assumptions on error magnitudes, that is, the error needs to be sufficiently small relative to the nonzeros floor α, to allow separation of variables. On the other hand, the ℓ_1-minimization method needs a larger number of measurements than the separation-based method, since the former recovers both x and e simultaneously, whereas the latter recovers only x in Steps I and III.

The above suggests a strategy of detecting and identifying bad measurements. That is, for a given set of equations which may or may not contain any bad ones, we first apply the joint ℓ_1-minimization method to solve for x and e. The procedure is deemed to be successful if the computed e^* is sparse and the computed solution of the system corrected by e^* is also sparse. Otherwise, apply the separation-based method as it can survive with fewer measurements.

We can furthermore use the properties of the two methods to design a strategy for choosing the number of measurements M. Suppose the application is such that measurements are relatively costly. Then, a straightforward strategy is to start with small M and slowly add more measurements. We note there can be three outcomes, depending on the magnitude of the error: 1) if the error is extreme the application can detect the bad measurement as an outlier, and we are finished; 2) if the error is moderate, separation may fail at small M, and joint-ℓ_1 succeed when M increases sufficiently; finally, 3) if the error is small, separation may succeed at smaller M than joint-ℓ_1.

A similar strategy may be employed for the type of applications where sets of M compressive measurements are taken periodically, e.g., to repeatedly determine the status of a set of objects. In that case, the application can, over time, adapt M so that at least one of the methods succeeds; one potential adaptation scheme may be to additively decrease, multiplicatively increase the value of M so that the system quickly leaves the state where both methods fail.

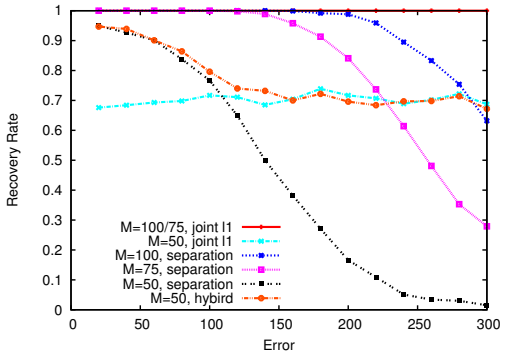

Fig. 3. The joint ℓ_1-minimization method (joint l1) can detect errors regardless of the error magnitude. However, it also requires more measurements. When there are insufficient measurements, the separation-based method (separation) performs better given sufficiently large nonzeros floor α, which is configured to be 150 in these experiments.

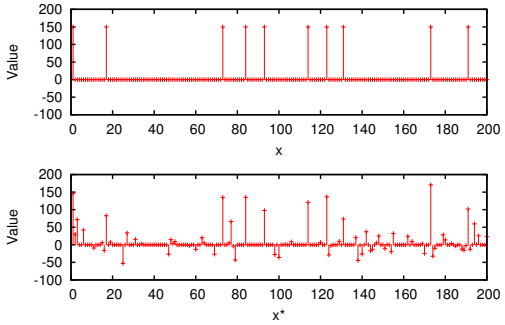

Fig. 4. Top: An exemplar test signal x with N=200, K=10, and α=150. Bottom: The decoded solution x^* after directly solving with ℓ_1-minimization with M=75. A bad measurement was introduced into the linear system $y = \mathbf{A}x$. The computed x^* still preserves the "peaks" structure of x; thus one can eliminate a large number of small-valued components to reduce the number of unknowns.

VII. Numerical Experiments

In this section we demonstrate and compare the performance of the separation-based and the joint ℓ_1-minimization methods via numerical experiments.

A. Performance of the Separation-based Method

Figure 3 shows the performance of both methods. In each experiment, we use an exactly sparse signal of length 200. The signal has 10 peaks at random locations as shown in Figure 4. For simplicity, all the peak values are set to 150 while all others are set to 0. An $M \times 200$ Gaussian matrix is used to obtain the M measurements. One of the measurements is deliberately corrupted by adding an error amount ranging from 20 to 300. For the separation-based method, the number of left-over variables L is set to its lower bound $K + W$ assuming we know W. In reality, W will not be known a priori and it can only be estimated empirically or analytically. We use this smallest possible L to establish a performance upper bound for the separation-based method.

We varied the number of measurements M from 50 to 100. The joint ℓ_1-minimization method performs very well when

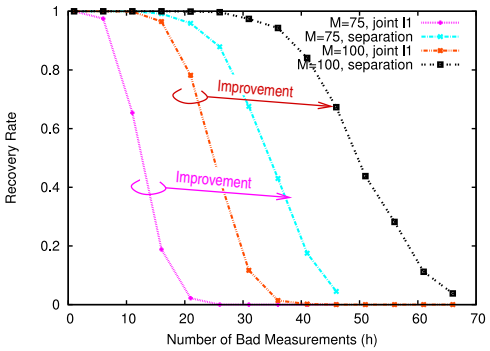

Fig. 5. The separation-based method can tolerate a higher number of bad measurements. In this experiment we fixed $\alpha = 150$, $M = 75$, $e_i = 30$, and varied the number of bad measurements.

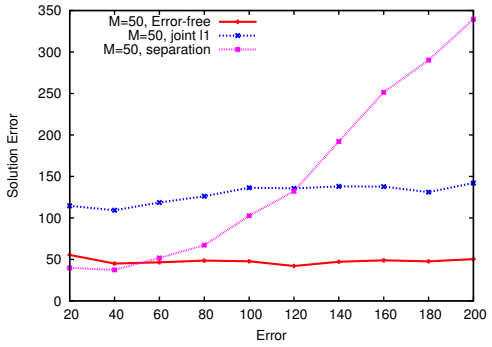

Fig. 6. The separation-based method applies to an approximately sparse signal. When the error is small, the solution error recovery rate is similar to the case that no bad measurements are introduced.

there is a sufficient amount of measurements. When M is 75 or 100, it can always recover from the corrupted measurements regardless of the error magnitudes. In contrast, although the separation-based method is also capable of recovering from the bad measurements, it cannot tolerate errors too large due to an excessive amount of inversions. For large M, e.g. $M = 100$, the separation-based method can handle larger errors because more inversions can be tolerated.

The separation-based method shows its main advantage when the measurements are less plentiful. When M is dropped to 50, it is insufficient for the joint ℓ_1-minimization method, resulting in a recovery rate around 70% for the entire error range. The separation-based method however can achieve over 90% recovery rate when the nonzeros floor α is large enough relative to the errors. This suggests, as described earlier, a hybrid method to combine the advantages of both joint ℓ_1 and the separation-based method. That is, one should use the joint ℓ_1 method first. If it fails, the separation-based method can be used instead. The results of the hybrid method are also shown in Figure 3, and the hybrid method indeed can improve the joint ℓ_1 method when the number of measurements is not sufficient.

This experiment shows that the separation-based method can recover with fewer number of measurements than the joint ℓ_1-minimization method. For similar reasons, the separation-based method can tolerate a higher number of bad measurements. Figure 5 shows the results of an experiment where we vary h to change the number of bad equations. We use the same setting as the previous experiment, and set the total number of measurements to 75 and 100. When a bad measurement is added, we always add a constant error 30 to the measured value. The results show that the separation-based method can generally achieve the same recovery rate as the joint ℓ_1-minimization method while allowing 20-30 additional bad measurements.

Next, we show that the separation-based method also applies to approximately sparse signals. In this experiment, we use the same settings except we set all the original zero components to a fixed value 1. This number is chosen to satisfy Eq (20). Note that a truncation error exists even if no bad measurements are presented. Figure 6 shows the performance of separation-based method applied to this approximately sparse signal. Note that when error introduced by the bad measurement is small (e.g., less than 60), the performance of the separation-based method is comparable to the error-free case which involves no bad equations. This shows that the method has successfully removed the bad equations. In contrast, the joint ℓ_1-minimization method has a higher solution error, due to the lack of sufficient measurements.

VIII. Conclusions

In this paper we describe a new method, called separation-based method, for detecting and also identifying bad measurements in compressive sensing. If errors introduced by bad measurements are small, then the separation-based method requires a smaller number of good measurements than the previous method based on joint ℓ_1-minimization. We note that if there are sufficient good measurements, the joint ℓ_1-minimization works independently of error magnitudes. In practice we suggest a hybrid approach. That is, we will use the joint ℓ_1-minimization method first. If the joint method fails due to an insufficient amount of good measurements, then we use the separation-based method, as it can survive with a smaller number of good measurements.

We have shown that the separation-based method works for both exactly and approximately sparse signals. In the analysis we have introduced the notion of inversions. A tight upper bound on the number of inversions is a key to ensure the effectiveness of the separation-based method. We argue that obtaining such an upper bound is empirically feasible in practice.

References

[1] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[2] E. J. Candès, J. K. 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, 2006.

[3] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus de l'Académie des sciences, Paris*, Series I, vol. 346, pp. 589, 2008.

[4] E. J. Candès, J. K., Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[5] J. Wright and Y. Ma, "Dense error correction via l1-minimization," in *ICASSP 2009*.

[6] Z. Li, F. Wu, and J. Wright, "On the systematic measurement matrix for compressed sensing in the presence of gross errors," in *2010 Data Compression Conference*, Snowbird, UT, Mar. 2010.