# THRESHOLDING-BASED RECONSTRUCTION OF COMPRESSED CORRELATED SIGNALS

*Alhussein Fawzi, Tamara Tošić and Pascal Frossard*

École Polytechnique Fédérale de Lausanne (EPFL)
Signal Processing Laboratory (LTS4), Lausanne, 1015-Switzerland
E-mail: {alhussein.fawzi,tamara.tosic,pascal.frossard}@epfl.ch.

**ABSTRACT**

We consider the problem of recovering a set of correlated signals (*e.g.,* images from different viewpoints) from a few linear measurements per signal. We assume that each sensor in a network acquires a compressed signal in the form of linear measurements and sends it to a joint decoder for reconstruction. We propose a novel joint reconstruction algorithm that exploits correlation among underlying signals. Our correlation model considers geometrical transformations between the supports of the different signals. The proposed joint decoder estimates the correlation and reconstructs the signals using a simple thresholding algorithm. We give both theoretical and experimental evidence to show that our method largely outperforms independent decoding in terms of support recovery and reconstruction quality.

***Index Terms***— Distributed compressed sensing, Thresholding algorithm, Sparse signals, Overcomplete dictionary.

## 1. INTRODUCTION

The growing number of distributed systems in recent years has led to an important body of work on the efficient representation of signals captured by multiple sensors. Recently, ideas based on Compressed Sensing (CS) [1, 2] have been applied to distributed reconstruction problems [3] in order to recover signals from a few measurements per sensor. When signals are correlated, a joint decoder that properly exploits the inter-sensor dependencies is expected to outperform independent decoding in terms of reconstruction quality. Very often, the correlation model restricts the unknown signals to share a common support. Using this correlation model, the authors in [3] and [4] propose decoding algorithms and show analytically that joint reconstruction outperforms independent reconstructions. In many applications, this correlation model is however too restrictive. For example, in the case of a network of neighbouring cameras capturing one scene or seismic signals captured via different sismometers, the supports of the signals are quite different even if components are linked by simple transformations.

In this paper, we adopt a more general correlation model and build a joint decoder that recovers the unknown signals from a few measurements per sensor. We assume that the unknown signals are sparse in a *redundant dictionary* $\mathcal{D}$, and not necessarily in an orthonormal basis [5, 6]. We denote the components of the dictionary as *atoms*. We assume that the support of each view $j$ is related to the support of a reference view by a transformation $T_j^*$. The transformation $T_j^*$ could be for example a translation function. Using the given correlation model, we build a joint decoder based on the thresholding algorithm [5] and prove theoretically that it outperforms an independent decoding method in terms of recovery rate. Moreover, we show experimentally that the proposed algorithm leads to better reconstruction quality.

## 2. PROBLEM FORMULATION

We consider a sensor network of $J$ nodes. Each sensor $j$ acquires $M$ linear measurements of the unknown signal $y_j \in \mathbb{R}^N$ ($M < N$) and sends it to a central decoder. The role of the decoder is to estimate the unknown signals $\mathcal{Y} = \{y_j\}_{j=1}^J$. By denoting $\mathcal{S} = \{s_j\}_{j=1}^J$ the set of compressed signals acquired by the sensors and $\mathcal{A} = \{A_j\}_{j=1}^J$ the sensing matrices, we have:

$$\underbrace{s_j}_{M\times 1} = \underbrace{A_j}_{M\times N}\underbrace{y_j}_{N\times 1}. \quad (1)$$

In the rest of this paper, we use independent sensing matrices with Gaussian i.i.d. entries. Specifically, $\sqrt{M}(A_j)_{m,n}$ follows a standard Gaussian distribution, for any $m, n, j$.

We assume that the unknown signals $y_j \in \mathbb{R}^N$ are sparse in some dictionary $\mathcal{D}$ that consists of $K$ atoms and denote by $\Phi$ its matrix representation with dimension $N \times K$. Formally, we have $y_j = \Phi c_j$, where $c_j$ is a vector of length $K$ with at most $S$ non zero components and $S < N$. By denoting the support of $y_j$ with $\Delta_j^*$ (*i.e.,* the set of $S$ atoms corresponding to the non zero entries of $c_j$), $y_j$ can be written as follows:

$$\underbrace{y_j}_{N\times 1} = \underbrace{\Phi_{\Delta_j^*}}_{N\times S}\underbrace{x_j}_{S\times 1}, \quad (2)$$

where $\Phi_{\Delta_j^*}$ is the restriction of $\Phi$ to $\Delta_j^*$ and $x_j$ corresponds to the non zero entries of $c_j$.

We adopt the following correlation model: for any $j \in \{1, \ldots, J\}$, the set of atoms in $\Delta_j^*$ can be obtained from $\Delta_1^*$ by applying a transformation $T_j^* : \mathcal{D} \to \mathcal{D}$. This can be written as: $T_j^*(\Delta_1) = \Delta_j^*$ (we consider that $T_1^*$ is the identity). In our problem, the vector of transformations $T^* = \{T_j^*\}_{j=1}^J$ is unknown. However, we assume that we are given a set $\mathcal{T}$ of candidate vector of transforms and that the correct vector $T^*$ belongs to $\mathcal{T}$.

Considering the above correlation model, we adress the following problem: Given the compressed signals $\mathcal{S}$, the sensing matrices $\mathcal{A}$, the sparsity $S$, the dictionary $\Phi$, and the set of candidate transformations vectors $\mathcal{T}$, estimate the unknown signals $\mathcal{Y}$ (*i.e.,* supports $\{\Delta_j^*\}_{j=1}^J$ and coefficients $\{x_j\}_{j=1}^J$) using a small number of measurements per sensor $M$.

## 3. JOINT THRESHOLDING ALGORITHM

We propose a solution to the problem formulated in the previous section. Our proposed decoder extends the simple thresholding algorithm [5] to multiple signals. This choice is motivated by the low complexity of thresholding algorithm with respect to other decoding methods [7]. Our joint decoder represents an efficient alternative when the signals are simple (*i.e.,* they have very sparse representations in the dictionary) and the number of sensors is fairly large, so that other decoding methods become computationally intractable.

The Joint Thresholding (JT) decoder exploits the information diversity brought by the different signals to reduce the number of measurements per sensor required for accurate signals reconstruction. It groups the measurements obtained from each individual signal and precisely estimates the unknowns $(\Delta_1^*, T^*)$ (or equivalently all the supports $\{\Delta_j^*\}_{j=1}^J$).

JT obtains an estimate $(\widehat{\Delta_1}, \widehat{T})$ of $(\Delta_1^*, T^*)$ by maximizing the following objective function, which is called the score function:

$$\Psi_s(\Delta_1, T) = \sum_{j=1}^{J} \sum_{\varphi \in T_j(\Delta_1)} s_j \cdot A_j \varphi, \tag{3}$$

where $\Delta_1$ and $T$ denote respectively the support of the reference signal and the vector of transformations variables and the dot operator ‘ $\cdot$ ’ denotes the canonical inner product. The use of $\Psi_s$ as the objective function is justified by:

$$\begin{aligned}(\widehat{\Delta_1}, \widehat{T}) &= \underset{(\Delta_1, T)}{\operatorname{argmax}}\, \Psi_s(\Delta_1, T)\\ &\approx \underset{(\Delta_1, T)}{\operatorname{argmax}}\, \Psi_y(\Delta_1, T)\\ &= (\Delta_1^*, T^*),\end{aligned}$$

where $\Psi_y(\Delta_1, T) = \sum_{j=1}^{J} \sum_{\varphi \in T_j(\Delta_1)} y_j \cdot \varphi = \mathbb{E}\Psi_s$ ($\mathbb{E}$ denotes the expected value of a random variable). Indeed, if both assumptions given in Eq.(4) and Eq.(5) hold[1], $\Psi_y(\Delta_1, T)$ is maximal for $\Delta_1 = \Delta_1^*$ and $T = T^*$. Besides, for large values of $M$, $\Psi_s(\Delta_1, T)$ concentrates around its average value $\Psi_y(\Delta_1, T)$ [8, Lemma 4.1]. The description of JT algorithm is given in Algorithm 1.

In words, the JT algorithm calculates for each transformation vector $T \in \mathcal{T}$ the vector $d_T$, given by $d_T[i] = \sum_{j=1}^{J} s_j \cdot A_j T_j(\varphi_i)$, for $1 \leq i \leq K$. Then, the largest $S$ elements in $d_T$ are summed and assigned to $\Psi_s(\Delta_1, T)$. Estimated quantities $\{\widehat{\Delta_1}, \widehat{T}\}$ are updated if $\Psi_s(\Delta_1, T)$ achieves a higher score. Knowing the set of supports, we deduce coefficients $\widehat{x_j}$ by computing the least squares solution to equation $A_j \Phi_{\widehat{\Delta_j}} \widehat{x_j} = s_j$.

**Algorithm 1** Joint Thresholding (JT) Algorithm

**Input:** compressed signals $\{s_j\}$, sensing matrices $\{A_j\}$, sparsity $S$, dictionary $\Phi$, candidate vectors of transformations $\mathcal{T}$.

**Output:** estimated signals $\{\widehat{y_j}\}$, support $\widehat{\Delta_1}$ and the vector of transformations $\widehat{T}$.

**1**. Initialization: $(\widehat{\Delta_1}, \widehat{T}, \widehat{\Psi}) \leftarrow (\varnothing, \varnothing, -\infty)$

**2**. For every $T \in \mathcal{T}$

**2.1** Build the vector $d_T$ of length $K$ in the following way:

$$d_T = \sum_{j=1}^{J} (A_j T_j(\Phi))^{\mathrm{T}} s_j,$$

where $T_j(\Phi) = \left[T_j(\varphi_1) \dots T_j(\varphi_K)\right]$ and $(\cdot)^{\mathrm{T}}$ denotes the matrix transpose.

**2.2** Keep the largest $S$ entries in $d_T$ and set the other entries to zero. The positions of the non zero entries in $d_T$ give the indices of the estimated suppport $\Delta_1$ of the first signal.

**2.3** Calculate the *score* $\Psi_s(\Delta_1, T)$ by summing the $S$ non zero entries of $d_T$.

**2.4** If $\Psi_s(\Delta_1, T)$ exceeds $\widehat{\Psi}$: update $(\widehat{\Delta_1}, \widehat{T}, \widehat{\Psi}) \leftarrow (\Delta_1, T, \Psi_s(\Delta_1, T))$

**3**. Build the coefficients vector for all $j \in \{1, \dots, J\}$:

$$\widehat{x_j} = \left(A_j \Phi_{\widehat{\Delta_j}}\right)^{+} s_j,$$

where $(\cdot)^{+}$ denotes the pseudo-inverse operator. Note that $\widehat{\Delta_j}$ is obtained using the correlation model: $\widehat{\Delta_j} = \widehat{T_j}(\widehat{\Delta_1})$, for $j \geq 2$.

**4**. Obtain signals estimates:

$$\widehat{y_j} = \Phi_{\widehat{\Delta_j}} \widehat{x_j}.$$

## 4. THEORETICAL ANALYSIS

Our theoretical analysis focuses on the performance of JT in finding the correct supports. Hence, we will not address the quality of the estimated coefficients $\{\widehat{x_j}\}_{j=1}^J$. In particular, we focus on the analysis of the *recovery rate* $R$ defined as the total number of correctly recovered atoms (in all signals combined) divided by the total number of atoms (which is equal $SJ$).

We assume the following:

- For all $j \in \{1, \dots, J\}$, $\Delta_j^*$ can be recovered entirely by applying the thresholding algorithm on $y_j$. Formally, there exists $\eta > 0$ verifying:

$$\inf_{\varphi \in \Delta_j^*} \left| \frac{y_j}{\|y_j\|_2} \cdot \varphi \right| > \sup_{\varphi \in \overline{\Delta_j^*}} \left| \frac{y_j}{\|y_j\|_2} \cdot \varphi \right| + \eta, \tag{4}$$

  where $\overline{\Delta_j^*}$ is the complement of $\Delta_j^*$ in $\mathcal{D}$. As this condition is practically hard to verify, a sufficient condition involving the coherence of the dictionary is given in [5, Eq.(3.2)].

- All the atoms in the supports have positive inner products with the corresponding signal:

$$\forall \varphi \in \Delta_j^*, y_j \cdot \varphi \geq 0 \tag{5}$$

The assumption in Eq.(4) is reasonable since we cannot hope to recover the supports using JT unless the thresholding algorithm correctly recovers the supports when applied on the full signals $y_j$. Assumption in Eq.(5) is a technical one and is used in the proof of our main theorem. This assumption can be achieved by adding the inverse of the atom in the dictionary ($\varphi \to -\varphi$) when the inner product is negative.

Our main theoretical result is given in the following theorem:

**Theorem 4.1 (Recovery rate of JT)** *Let $R$ be the recovery rate of JT defined by:*

$$R = \frac{\sum_{j=1}^{J} |\Delta_j^* \cap \widehat{\Delta_j}|}{SJ}.$$

*Then, for any* $0 < \alpha \leq 1$*:*

$$\mathbb{P}\left(R \geq 1 - \alpha\right) \geq 1 - 4SJK|\mathcal{T}| \exp\left(-CMJ\eta^2\alpha^2 \frac{m_y^2}{M_y^2}\right), \tag{6}$$

*where* $m_y = \min_j \|y_j\|_2$, $M_y = \max_j \|y_j\|_2$, $C = \left(\frac{32e}{\sqrt{6\pi}} + 4e\sqrt{2}\right)^{-1}$.

[1] The assumptions are discussed in the next section.

**Main ideas of the proof [8]**

The first step of the proof is to show that $\frac{1}{J}\sum_{j=1}^{J} A_j u_j \cdot A_j v_j$ concentrates around its average value $\frac{1}{J}\sum_{j=1}^{J} u_j \cdot v_j$ for any set of vectors $\{u_j\}_{j=1}^{J}$ and $\{v_j\}_{j=1}^{J}$ of length $N$. As in [5], we use Bernstein's inequality to establish such a result and show that the concentration inequality is tighter when $J$ increases.

Then, we bound the probability that the recovered supports have in total more than $h$ incorrectly estimated atoms. If the estimated supports have indeed more than $h$ erroneous atoms, combining Eq.(4) and Eq.(5) leads to the following inequality: $\Psi_y(\Delta_1^*, T^*) - \Psi_y(\widehat{\Delta_1}, \widehat{T}) > \eta h m_y$, where $\Psi_y$ is defined in section 3. In words, this means that the gap between the expected value of the correct and estimated scores exceeds $\eta h m_y$. We bound the probability that this gap vanishes in the compressed domain. This happens when scores $\Psi_s$ are not enough concentrated around their expected value $\Psi_y$. We thus obtain the desired bound by using the concentration inequality established in the first step. Finally, we obtain Eq.(6) by assigning $h \leftarrow \alpha S J$ which represents a part of the total number of atoms. ■

For simplicity, we consider the common case where all the signals have the same energy ($m_y = M_y$). Theorem 4.1 shows that for sufficiently high values of $J$, the recovery rate is mainly governed by $MJ$, $\eta$ and $|\mathcal{T}|$. The dependence on $MJ$ (*i.e.,* total number of measurements) follows our intuition as JT combines the measurements of the different sensors to perform the joint decoding. Increasing the total number of measurements leads to a better recovery rate. The quantity $\eta$ hides the dependence of $R$ on the signal characteristics and model. We give in [8] a lower bound on $\eta$ in terms of sparsity, coherence of the dictionary and ratio between the largest and the lowest coefficients. Another key parameter is the number of candidate vectors of transformations $|\mathcal{T}|$ which grows with $J$. In the following corollary, we provide a lower bound on the number of measurements needed per sensor to reach asymptotically a perfect recovery rate in the following two cases: (1) $|\mathcal{T}|$ grows slowly with $J$, (2) $|\mathcal{T}|$ grows exponentially with $J$.

**Corollary 4.1 (Asymptotic behaviour of $R$)** *Let* $0 < \alpha \leq 1$.

1. *If $|\mathcal{T}|$ is a subexponential function of J, then, as long as $M \geq 1$, $\mathbb{P}(R \geq 1-\alpha)$ converges to 1 as $J \to +\infty$.*
2. *If there exists $\beta > 0$ such that $|\mathcal{T}| \sim e^{\beta J}$ , then, as long as $M > \frac{\beta}{C\eta^2\alpha^2}\frac{M_y^2}{m_y^2}$, $\mathbb{P}(R \geq 1-\alpha)$ converges to 1 as $J \to +\infty$.*

The growth of $|\mathcal{T}|$ is related to the degree of uncertainty on the correct transformation vector $T^*$. For example, if $T^*$ is known in advance, then $|\mathcal{T}| = 1$ and Corrolary 4.1 guarantees an arbitrary high recovery rate with only one measurement per sensor when $J \to +\infty$. This result remains valid as long as $|\mathcal{T}| \ll e^{\beta J}$ for all $\beta > 0$. However, if $T^*$ is completely unknown and transforms between pairs of signals are independent, $|\mathcal{T}|$ grows exponentially with $J$ and we will need more measurements per sensor in order to recover the correct support estimates (consider the example where (a) the number of candidate transformations between each sensor $j \geq 2$ and the reference signal is equal to $K$; (b) $T_j$ is independent of $T_{j-1}$, then: $|\mathcal{T}| = K^{J-1}$).

Unlike independent thresholding which has a constant recovery rate in function of $J$, previous results show that the recovery rate of JT *increases* by augmenting the number of sensors $J$. Thus, in large networks, JT requires less measurements per sensor than independent thresholding for a fixed target recovery rate provided that $|\mathcal{T}|$ has a controlled growth.

## 5. EXPERIMENTAL RESULTS

### 5.1. Greedy JT

The JT algorithm, as described in Section 3, performs the search over all candidate vectors in $\mathcal{T}$. This can be very costly in terms of the computational efficiency, especially for a large number of correlated signals. Thus, instead of performing a full search, we greedily look for the relevant transformation vector. We gradually build $\widehat{T}$ by selecting at each level $V \in \{2, \ldots, J\}$ the transformation that maximizes the score calculated till the signal $V$. That is, for every possible transformation $T_V$ relating the reference signal (*i.e.,* signal 1) to $V$, the algorithm calculates the vector $d_T = \sum_{j=1}^{V} (A_j T_j(\Phi))^{\mathrm{T}} s_j$, where $T = [\widehat{T}, T_V]$ ($T_V$ appended to the vector of transforms $\widehat{T}$). Summing the largest $S$ entries in $d_T$ provides the intermediate score calculated until the signal $V$. The best estimates $\{\widehat{\Delta_1}, \widehat{T}\}$ are updated if the intermediate score exceeds the current optimal one. This procedure repeats for every $V \in \{2, \ldots, J\}$ and leads to the final support estimates $\{\widehat{\Delta_j}\}_{j=1}^{J}$. Steps 3 and 4 of Algorithm 1 are unchanged. For the detailed algorithm description, refer to [8].

Even though this algorithm has a lower complexity than JT, the price to pay is a less robust transformation estimation process: in the early stages of the algorithm ($V \ll J$), the selection of the transformations is based on a small number of signals $V$. If in addition the value of $M$ is small, this may lead to uncorrect estimation of the transformations and thus wrong support estimates. However, performance penalty is small in practice [8]. In the following, we examine the performance of Greedy JT on synthetic images and seismic signals.

### 5.2. Synthetic images

We construct a parametric dictionary where a generating function undergoes rotation, scaling and translation operations to generate the different atoms in the dictionary $\mathcal{D}$. We use the Gaussian $g(x,y) = e^{-x^2-y^2}$ as the generating function. The atoms in the dictionary are characterized by the rotation angle $\theta$, scales $s_x$ and $s_y$ and translations $t_x$ and $t_y$. If $(X,Y)$ denotes the transformed coordinate system:

$$X = \frac{\cos\theta(x-t_x) - \sin\theta(y-t_y)}{s_x}$$
$$Y = \frac{\cos\theta(y-t_y) + \sin\theta(x-t_x)}{s_y},$$

the atom $g_p$ with parameters $p = (\theta, s_x, s_y, t_x, t_y)$ is given by:

$$g_p(x,y) = K \cdot g(X,Y),$$

where $K$ is the normalization constant.

The dictionary is generated for images of size $N = 32 \times 32 = 1024$, with the following parameters: $\theta \in [0 : \frac{\pi}{6} : \pi], s_x = \{2, 4\}, s_y = \{1/2, 1\}$. Every atom is shifted in pixels of odd coordinates, so the full dictionary contains 6144 atoms.

The sparse support of the reference image is randomly selected in $\mathcal{D}$ and coefficients are chosen such that the conditions in Eq.(4) and Eq.(5) are verified. The remaining images have been obtained

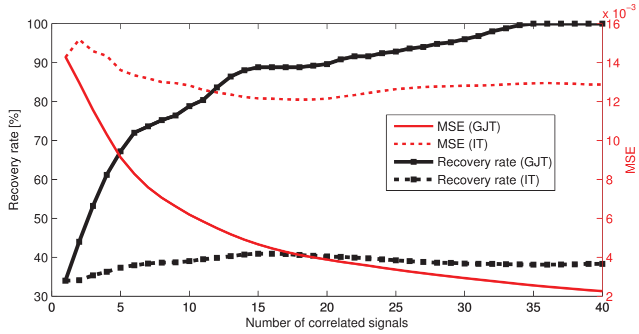


**Fig. 1**. Recovery rate and Mean Squared Error of Greedy JT (GJT) and Independent thresholding (IT) as a function of $J$. Simulation setup: 10 independent trials, $S = 5$, $M = 150$, $N = 1024$, Gaussian sensing matrices.

by applying global translations on the atoms of the reference image, under the constraint that the support of every image belongs to $\mathcal{D}$. We assume that the transformations are independent from one another and that there are 9 candidate transforms for any image. Thus, $|\mathcal{T}| = 9^{J-1}$. The recovery rate and MSE of Greedy JT and independent thresholding are shown in Fig.1 as a function of the number of correlated images. Recovery rate is defined in Section 4. For a given $J$, the calculated MSE represents the averaged MSE calculated over signals $\{1, \ldots, J\}$. We see that Greedy JT outperforms independent thresholding in terms of recovery rate and image quality, especially for high values of $J$ ($J \geq 20$). Thus, although $|\mathcal{T}|$ grows rapidly with $J$, our joint decoding approach is significantly better in practice in terms of support recovery.

### 5.3. 1D seismic signals

Seismic signals captured at neighbouring locations typically follow the correlation model proposed in this paper. Fig.2 (a), (b) represent two seismic signals that are obviously correlated as the second signal is approximately a shifted version toward the front of the first signal. We use the following sparsifying dictionary, which consists of Gaussians modulated with sinusoids:

$$g_{(t,s,\omega)}(x) = K \exp\left(-\frac{(x-t)^2}{s^2}\right) \cos\left(\omega \frac{x-t}{s}\right),$$

where $K$ is a normalization constant. The translations $t$ are chosen uniformly from 1 to $N$ with step size 10 so that the coherence of the dictionary is not too large. Scales $s$ take values in $\{4, 8, 16\}$ and $\omega$ varies from 2 to 10 with step 2. For each set of parameters $(t, s, \omega)$, $g_{(t,s,\omega)}$ and $-g_{(t,s,\omega)}$ are included in the dictionary. Fig.2 (c) and (d) illustrate the estimations of signal number 2 obtained with only $15\%$ of the measurements respectively using independent thresholding and JT algorithm. Note that as $J = 2$ in this example, Greedy JT and JT are equivalent. Due to space limitations, we omit the results on signal number 1. Visual inspection and calculated MSEs confirm the superiority of joint decoding using JT algorithm over independent thresholding in terms of reconstruction quality. This experiment shows that JT provides significantly better quality signals even when the number of correlated signals is low ($J = 2$).

## 6. CONCLUSION

In this paper, we have proposed an efficient approach for the joint recovery of correlated signals that have been compressed independently. Our solution is novel with respect to the state of the art work due to the particular geometrical correlation model based on the transformations of the sparse signal components. Mathematical analysis and experimental results demonstrate the superiority of our recovery algorithm over independent thresholding. JT is namely applicable for decoding simple multiview images, seismic signals or any other set of correlated signals satisfying the geometric correlation model. A promising future direction is to use JT for correlation estimation along with a more sophisticated recovery algorithm for the reconstruction.

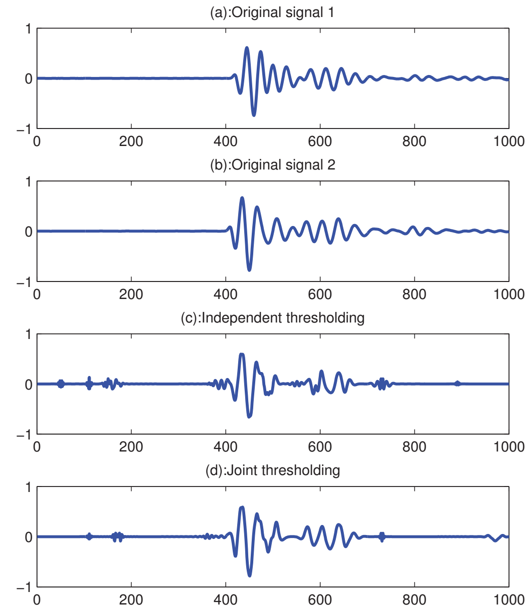


**Fig. 2**. Seismic signals (a) $y_1$ and (b) $y_2$ captured at two neighbouring locations. Estimation of $y_2$ using (c) independent thresholding and (d) JT. Simulation setup: $J = 2, N = 1000, M = 150, S = 50, |\mathcal{T}| = 3$, Gaussian sensing matrices. This experiment was conducted 200 times and we obtained $\text{MSE}_{\text{IT}} = 0.0031$ and $\text{MSE}_{\text{JT}} = 0.0025$.

## 7. ACKNOWLEDGMENTS

A.F. would like to thank Omar Fawzi for the fruitful discussions.

## 8. REFERENCES


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

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

[3] M. F. Duarte, S. Sarvotham, M. B. Wakin, D. Baron, and R. G. Baraniuk, "Joint sparsity models for distributed compressed sensing," .

[4] M. Golbabaee and P. Vandergheynst, "Distributed compressed sensing for sensor networks using thresholding," *Proc. SPIE*, 2009.

[5] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *Information Theory, IEEE Transactions on*, vol. 54, no. 5, pp. 2210–2219, 2008.

[6] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. and Comp. Harm. Anal.*, 2011.

[7] X. Chen and P. Frossard, "Joint reconstruction of compressed multi-view images," in *ICASSP 2009*. IEEE, 2009, pp. 1005–1008.

[8] A. Fawzi, T. Tosic, and P. Frossard, "Thresholding-based reconstruction of compressed correlated signals," *Technical report, arxiv preprint arXiv:1109.5938*, 2011.