Compressed Acquisition of Correlated Signals

J.V.Satyanarayana, A.G .Ramakrishnan
jvsat29@yahoo.co.in, ramkiag@ee.iisc.ernet.in
Indian Institute of Science
Bangalore, India

Abstract

Distributed compressed sensing exploits information redundancy, inbuilt in multi-signal ensembles with inter- as well as intra-signal correlations, to reconstruct undersampled signals. In this paper we revisit this problem, albeit from a different perspective, of taking streaming data, from several correlated sources, as input to a real time system which, without any a priori information, incrementally learns and admits each source into the system.

Index Terms—Distributed compressed sensing, Joint sparsity models, Karhunen–Loueve Transform, Multiple measurement vector problem

1. Introduction

Distributed Compressed Sensing (DCS), where several correlated sources are sampled as multiple measurement vectors (MMV), has been dealt from different perspectives by researchers in recent years. DCS invariably leverages upon the information redundancy owing to the intra- as well as inter- signal correlations which contribute to what is known as joint sparsity of the signal ensemble. Such problem formulations have exciting applications in wireless multisensor data acquisition as well as biomedical signal capture. Chen and Huo [1] have given a thorough elucidation of the multiple measurement vector (MMV) problem of recovering the source matrix, given a matrix whose columns are measurement vectors that are projections of the source vectors on an over-complete dictionary. The authors extend the classical single measurement vector (SMV) case to MMV, with respect to recovery algorithms like convex optimization and orthogonal matching pursuit. Zhang and Rao [2] have demonstrated their approach based on a block sparse Bayesian learning framework, in cases where the source vectors are temporally correlated, with good results. However, they make a key assumption of common sparsity amongst measurement vectors (MV), which is valid only if the number of MVs is small. Reconstruction of two correlated sequences, in which the second sequence is the sum of the first sequence and an additional innovation, has been handled deftly in [3] by making use of the expectation maximization (EM) algorithm. The method depends on the a priori knowledge of the spectral characteristics of the innovation component. In this work, we propose a scheme in which the system learns the structure in the signal ensemble during the process of acquisition without any prior information. We deal with the real-time acquisition of streaming data from N correlated sources, with only $M < N$ measurement channels available. Since we wish to reconstruct the signal sources at every sampling instant, we cannot make use of an MMV formulation. We therefore, treat the problem as an SMV one. We do not make any specific assumption on the sparsity of the individual signals in the ensemble and rely on the underlying structure manifested as inter-signal correlation. Our undersampling scheme is very simple, so we avoid using any explicit measurement matrix on to which the signals are projected. The next section introduces the SMV problem.

2. Signal Correlation Model

The signals used in our work are based on the Joint Sparsity Models (JSM-1, JSM-2 and JSM-3) proposed in [4]-[5]. In these papers, the authors have explained these models in detail along with examples of practical real world signals. We introduce JSM-4 that eliminates some assumptions in the original JSMs to incorporate additional generality.

2.1 JSM-4

Let there be N correlated signals, $\boldsymbol{s}_n = \boldsymbol{c} + \boldsymbol{i}_n, n = \{0,1,2,..N-1\}$ where $\boldsymbol{c} = \boldsymbol{\mathcal{F}\alpha}$ is the common component and $\boldsymbol{i}_n = \boldsymbol{\mathcal{F}\alpha}_n$ is an innovation component unique to each signal where $\boldsymbol{\mathcal{F}}$ is a dictionary with $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}_n$ being the corresponding coefficient vectors for the common and the innovation components respectively. We choose the Fourier basis as the dictionary $\boldsymbol{\mathcal{F}}$. Like in JSM-3 proposed in [3], we do not restrict $\boldsymbol{\alpha}$ to be sparse. In addition, we remove the sparsity constraint on each $\boldsymbol{\alpha}_n$. Thus, $\gamma = \|\boldsymbol{\alpha}\|_0$ and $\gamma_n = \|\boldsymbol{\alpha_n}\|_0$, $n = \{0,1,2,..N-1\}$ need not necessarily be small compared to the dimensionality of the basis $\boldsymbol{\mathcal{F}}$. The frequencies and the corresponding amplitudes of the sinusoids in the common as well as innovation components are chosen at random, from an interval of real numbers. For example, the n^{th} innovation component $\boldsymbol{i}_n$, $n = 0 \dots N-1$ has γ_n frequency components, $f_i^{(n)}$, $1 \le i \le \gamma_n$ where $f_i^{(n)} \in \mathbb{R}$, in the band $[0, F]$. Similarly, the common component $\boldsymbol{c}$ has γ constituent frequencies, $f_j, 1 \le j \le \gamma$ in the same band. The Nyquist sampling frequency of the system, $F_S = 2F$. To neutralize the effect of incorporating additional generality in JSM-4, we restrict the amplitudes of the constituent sinusoids in all the innovation components to less than ten percent of the amplitudes of the sinusoids in the common component,

$$\frac{\|A_n\|_\infty}{\|A\|_\infty} < 0.1,\ \forall n \quad (1)$$

where$\|\,.\,\|_\infty$, indicating the maximum norm of a vector $\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{D}\times\boldsymbol{1}}$,is defined as $\|\boldsymbol{x}\|_\infty \stackrel{\text{def}}{=} max(|x_1|, |x_2|, \dots, |x_j|, \dots, |x_D|)$. $\boldsymbol{A}$ and $\boldsymbol{A}_n$ are the vectors of amplitudes of the common and innovation components.

Multiple correlated measurements of a physical process polluted with measurement noise at the individual sensors is a real world situation which can fit into this model. We show empirically, that signals based on JSM-4 can be reconstructed with reasonable level of accuracy.

3. Signal Measurement and Reconstruction

It is necessary to identify a tool which can exploit the common structure shared by the correlated sources. Karhunen Loueve Transform (KLT) is a very efficient means of sparsifying a set of correlated signals. Consider the signal matrix $\boldsymbol{S} \in \mathbb{R}^{\boldsymbol{K}\times\boldsymbol{N}}$, the rows of which are indexed by K successive time instants and the columns are indexed by the N signal sources. The covariance matrix of $\boldsymbol{S}$, denoted by $\boldsymbol{\Sigma}_S$, is symmetric for real valued $\boldsymbol{S}$ and its eigen vectors, $\boldsymbol{\psi}_n$, are orthogonal. Consequently, one can construct an orthogonal matrix, $\boldsymbol{\Psi} \triangleq [\boldsymbol{\psi}_0, \boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_{N-1}]$ such that $\boldsymbol{\Sigma}_s\boldsymbol{\Psi} = \boldsymbol{\Psi\Lambda}$ where $\boldsymbol{\Lambda}$ is a diagonal matrix consisting of the corresponding eigen values. The KLT is given by

$$\boldsymbol{\Gamma} = \boldsymbol{\Psi}^{\boldsymbol{T}} \quad (2)$$

The measurement vector, $\boldsymbol{g} \in \mathbb{R}^{N\times 1}$ at any instant, comprising the samples from all the N sources can be transformed into the N dimensional basis spanned by the eigen vectors, $\boldsymbol{\psi}_n$ as

$$\boldsymbol{G} = \boldsymbol{\Gamma g} \quad (3)$$

If the original signal sources are correlated then $\boldsymbol{\Gamma}$ can serve as a sparsity inducing basis and $\boldsymbol{G}$ is a sparse vector. This opens up the possibility of compressed acquisition and reconstruction [6]-[7] of the signals. Essentially this means that once the system learns the correlation structure in the signals, after several measurement cycles, and the KLT, $(\boldsymbol{\Gamma})$ converges, all the sources need not be sampled in the subsequent sampling cycles. From an undersampled measurement vector, which has only $M < N$ elements, the full $\widehat{\boldsymbol{g}}$ can be reconstructed using CS. Since $\boldsymbol{\Gamma}$ is unitary, from (3) we have

$$\boldsymbol{\Gamma}^{-1}\boldsymbol{G} = \boldsymbol{g} \quad (4)$$

where from (2)

$$\boldsymbol{\Gamma}^{-1} = \boldsymbol{\Psi} \quad (5)$$

is the inverse-KLT. If we measure only M out of N signal sources, then we have a vector $\boldsymbol{f} = \boldsymbol{\phi g}$ that is an undersampled version of $\boldsymbol{g}$. Here $\boldsymbol{\phi} \in \mathbb{R}^{M\times N}$ is the measurement matrix derived by choosing the M rows of $\mathbf{I_N}$, the identity matrix of dimension N which correspond to the indices of the measured signal sources in $\boldsymbol{g}$. Multiplying (4) on both sides by $\boldsymbol{\phi}$,

$$\boldsymbol{\phi\Gamma}^{-1}\boldsymbol{G} = \boldsymbol{\phi g} = \boldsymbol{f} \quad (6)$$

The vector, $\boldsymbol{f} \in \mathbb{R}^{M\times 1}$ is the undersampled version of $\boldsymbol{g}$ in which only M out of the N sources have been sampled. Given that $\boldsymbol{G}$ is sparse, the undetermined set of equations (6) can be solved under a CS framework, using the well known Basis Pursuit (BP) algorithm [8] which is basically a convex optimization (CO).

$$\widehat{\boldsymbol{G}} = \text{argmin}\|\boldsymbol{h}\|_1 \text{ subject to } \boldsymbol{\phi\Gamma}^{-1}\boldsymbol{h} = \boldsymbol{f} \quad (7)$$

(7) and (4) together give us $\widehat{\boldsymbol{g}}$.

With all the tools required for an undersampling and reconstruction identified, we are ready to devise a scheme for acquiring continuous streaming data from multiple signal sources. Description of such a scheme is the content of the next section.

4. Acquisition and Reconstruction of Correlated Signals

In this paper we propose Acquisition and Reconstruction of Correlated Signals (ARCS), a scheme for compressed acquisition of streaming data from several correlated signal sources of the JSM-4 class, and reconstruction of the same, in real time. Given an ensemble of N correlated signal sources and a maximum of $M < N$ measurement channels available in the system, all N channels are incrementally learnt. To start with, M signal sources are sampled at each sampling instant the sampling rate being F_S. The inverse-KLT (5) is calculated for the M signals until it converges. Inverse-KLT convergence is detected, when the RMS difference between the vectors formed out of the corresponding eigen values of the associated KLT matrix at successive time instants is below a small threshold τ. At this point, the system will have learnt the first M sources, which we include in the set of *familiar* sources. From the next sampling cycle onwards, at every instant, one source from the M *familiar* sources is randomly chosen to be omitted and only $M - 1$ out of the remaining are sampled. CS reconstruction (7) , recovers the omitted source. The one free measurement channel is made to sample a new source, which we call *stranger*, out of the remaining $N - M$ *strangers*. Inverse-KLT is again calculated for the $M + 1$ sources consisting of M *familiar* sources and one *stranger* source. After convergence, the *stranger* is added to the set of *familiar* sources, as before. Subsequently, out of the M + *1* familiar sources, two are randomly picked to be left out and the vector of $M + 1$ *familiar* signals is reconstructed (7) from only M-1 measurements. The process continues until Inverse-KLT converges for the $M + 2$ sources consisting of $M + 1$ *familiar* sources and one *stranger* source. This goes on until the set of *strangers* is empty and all the N channels join the *familiar* set.

4.1 ARCS Algorithm

A brief description of the algorithm, listed in Figure 1, is given here. The pre-specified algorithm parameters are N-the number of correlated signal sources, M-the number of available measurement channels and T_S –the sampling

```
Algorithm ARCS
Input:
  N               :  number of correlated signal sources
  M               :  number of available measurement channels
  T_s             :  time period of acquisition cycle
  Signal Sources

Output:
  S               :  Signal matrix

Var:
  fam             :  set of familiar signals
  str             :  set of stranger signals
  f               :  column vector with M or M − 1 elements
  Γ^{-1}(n)       :  Inverse KLT matrix of dimension n
  τ               :  threshold for KLT convergence
  φ               :  Measurement matrix
  k               :  acquisition cycle count
  time            :  current acquisition time
  acquisition_time:  total time of acquisition
  sel             :  set of randomly chosen channels
  eigval          :  vector of eigen values of the covariance matrix of !
  −Inf(n)         :  n dimensional vector, each element equal to −∞
  d               :  l2 norm between successive vectors of eigen value
  n               :  index count for the signal sources

Method:
 1. n ← M, k ← 1, str ← {1,2..M}
 2. while time ≤ acquisition_time do
 3.   if n = M then
 4.     sampling of channels in str into k^{th} row of S
 5.   else
 6.     if n < N then
 7.       sel ← Randomly select M − 1 numbers from fam
 8.     else
 9.       sel ← Randomly select M numbers from fam
10.     end if
11.     f ← sampling of channels in sel
12.     Ĝ ← CONVEXOPT(φ, Γ^{-1}(n − 1), f, (n − 1))
        //Ĝ ∈ ℝ^{n−1×1}
13.     k^{th} row of S ← (Γ^{-1}(n − 1)g')^T
14.   end if
15.   if n < N then
16.     if n > M then
17.       S(k, n) ← sampling of the channel number in str
18.     end if
19.     if k > 1 then
20.       (Γ^{-1}(n), eigval) ← INVERSEKLT(S)
21.       d ← ||eigval_prev − eigval||_2
22.       if d < τ then
23.         k ← 0, n ← n + 1, eigval_prev ← −Inf(n)
24.         S ← NULL,  fam ← fam ∪ str,  str ← {n}
25.         n_F ← size of fam
26.       else
27.         eigval_prev ← eigval
28.       end if
29.     end if
30.   end if
31.   time ← time + T_s,  k ← k + 1
32. end while
33. return
end ARCS


Subroutine CONVEXOPT(φ, ψ, f, p)
Var:
    h ∈ ℝ^{p×1}
Method:
 1. G ← argmin(||h||_1) subject to φψh = f
 2. return G
end CONVEXOPT

Subroutine INVERSEKLT(array)
Method:
 1. C ← Covariance Matrix of array
 2. [V E] ← Eigen vectors and Eigen values of C
 3. return (V, E)
end INVERSEKLT
```

Fig. 1 ARCS algorithm

period. N analog signal sources are available at the input of the system. The output is derived, in real time, from the signal matrix $\boldsymbol{S}$ whose rows correspond to the measurement vectors at successive time instants and whose columns represent the signal sources. The number of rows in $\boldsymbol{S}$, indexed by k, increases with each successive time instant and the number of columns, indexed by n, increases with each new signal source admitted into the system. Initially n, initialized to M, signals are sampled and stored in successive rows of $\boldsymbol{S}$. The inverse-KLT is calculated from the second sampling instant onwards (lines 19 and 20). Once convergence is achieved (line 22), the stranger set, str is merged into the set fam of familiar signals. Further, the new signal $n+1$, is admitted into the set str. The matrix $\boldsymbol{S}$ is initialized to *NULL* after its content is sent to the output,. The variable k is reset to zero and later incremented by one (line 31). In the subsequent sampling cycles, $M-1$ sources are randomly chosen (line 7) from fam and sampled (line 11). The numbers of the selected signals is stored in the set sel. The vector $\widehat{\boldsymbol{G}}$ (line 12) of size $n-1$ is obtained by invoking the convex optimization subroutine. The approximation $\widehat{\boldsymbol{g}}$ of the signal vector is obtained using (4) and assigned to the k^{th} row of $\boldsymbol{S}$ (line 13). The convex optimization routine takes as input the measurement matrix $\boldsymbol{\phi}$ and the inverse KLT matrix of order $n-1$. $\boldsymbol{\phi}$ is equal to the downsized identity matrix, consisting of only those rows whose indices are in sel. It is to be noted here, that the convex optimization is of order $n-1$, due to the fact that after the previous convergence, n had been incremented. After the signals in fam have been reconstructed, the stranger signal is sampled and stored (line 17) into a new column n of the k^{th} row. After several sampling cycles the inverse KLT of order n converges. The process continues until all the signals have been admitted into the system, that is $n > N$,. Subsequently, the convergence step is not required and lines 16-30 do not execute. M instead of $M-1$ signals are randomly chosen (line 9), at every sampling instant, from the set fam which is of size N.

5. Simulation and Results

5.1 Signals of JSM-4 class

Input Signals: We take as input a set of $N = 10$ signals that are captured with the help of a maximum available measurement channels, $M = 5$. The parameters of the *JSM-4* model, discussed in Section II, upon which the signals are built, are as follows:

$F_l = 5\ Hz,\ F_u = 10\ Hz,\ F_s = 20\ Hz,\ \gamma_n = 50, \gamma = 50.$

The amplitudes of the sinusoids are randomly chosen such that (1) holds true.

The acquisition and reconstruction are simulated for a duration of 100 seconds. Snapshot of the reconstruction,

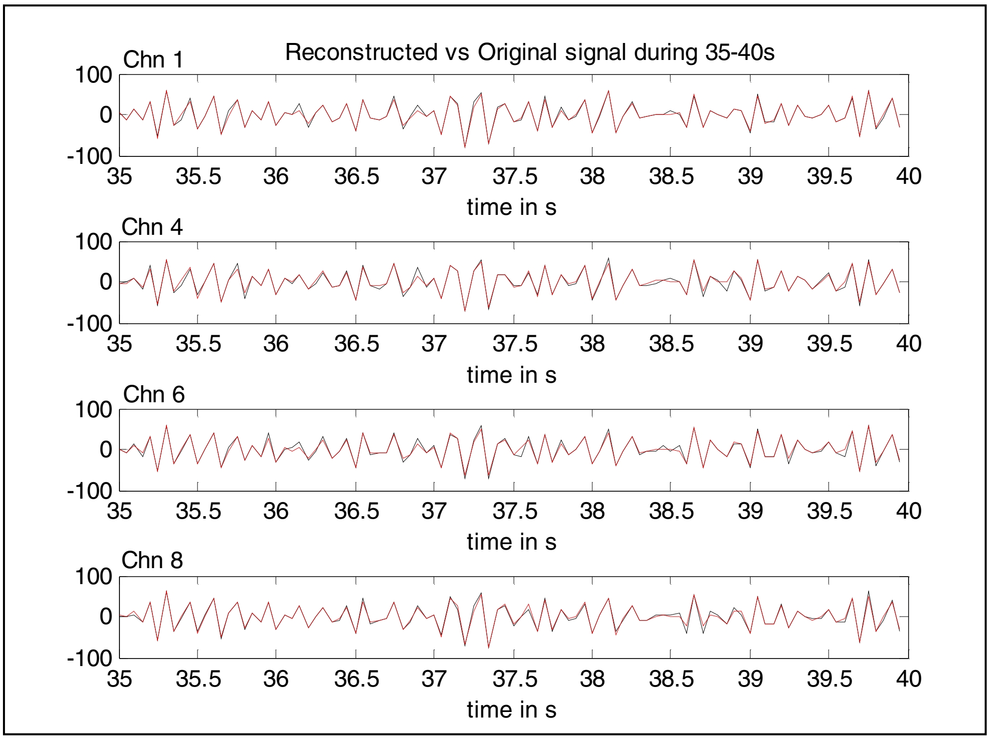

Fig. 2. Reconstructed signal (shown in red color) is plotted against the original signal for four channels, 1, 4, 6 and 8 during the interval 35-40sec. The corresponding PSNR values in dB are: 25.6, 23.3, 22.6 and 24.0.

TABLE I
RECONSTRUCTION PSNR VALUES (IN DB) AFTER REINITIALIZING COMMON COMPONENT

	Chnl 1	Chnl 4	Chnl 6	Chnl 8
After 60 sec	22.7	21.8	24.3	23.5
After 80 sec	25.8	22.7	21.1	21.5

plotted in red color, is shown against the original signal in figure 2 for channels, 1, 4, 6 and 8 in the time interval 35-40 sec. The choice of these channel numbers, for illustration, is arbitrary. As can be seen in the figure, the deviation between the original and the reconstructed signals is marginal. The learning curve of the system is shown in figure 3. All ten signals are learnt by the system within 40 sec, subsequent to which they are acquired with only the available five channels. After this point, the acquisition and reconstruction continue without the KLT convergence step and there is no more learning taking place.

5.2 Non-stationary signals of JSM-4 class

To probe into the robustness of ARCS, the constituent frequencies and amplitudes of the common component $\boldsymbol{c}$ have been altered in the middle of the simulation, at sampling time instants of 60sec and 80sec. In spite of this non-stationarity introduced, it is observed in the simulation, subsequent to the change in the signal frequency content, the reconstruction matches the original signal fairly well, as is indicated by the corresponding PSNR values for the channels 1, 4, 6 and 8 in Table I. This is because the correlation properties of the signals do not change even after the change in the frequency content of the common component. Since the correlation structure is already learnt by the system before the instant the no-stationarity is introduced, the reconstruction proceeds fairly well.

6. Conclusion

In this paper, we have suggested an approach for capturing and reconstructing a group of signals with joint correlation properties. The scheme, which we have called ARCS acquires streaming data from multiple signal sources in real time using fewer acquisition channels. We conclude the paper with some of the advantages offered by ARCS.

Practical measurement matrix: Under ARCS, the measurement matrix $\boldsymbol{\phi}$ is implicit in the sampling operation. The measurement vector is not a projection of the signal onto $\boldsymbol{\phi}$ and is obtained from the original vector, consisting of samples from the different sources, by simple undersampling, wherein some of the signal sources are not sampled. Therefore, $\boldsymbol{\phi}$ is just the downsized identity matrix, in which, the rows corresponding to the indices of the elements not sampled, are absent. With this arrangement, the sampling architecture is quite simple and

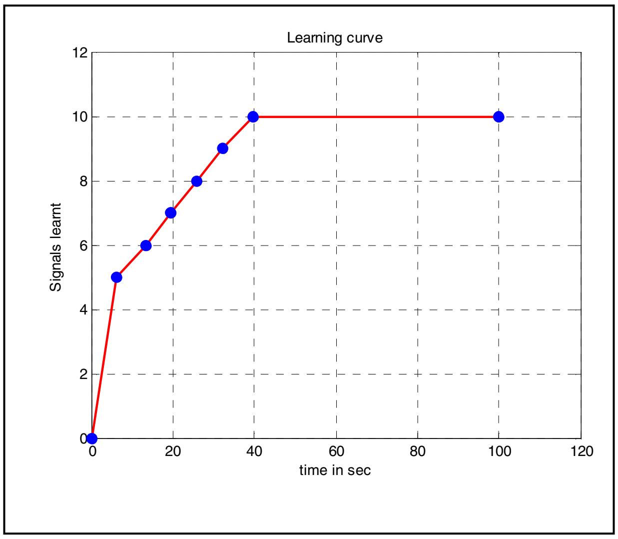

Fig. 3. The curve shows how the correlation structure is learnt by the system. The blue circles indicate the points at which a new signal is added.

general, thereby having the potential to support a variety of applications.

Support for non-sparse signals: Empirically, we have been able to demonstrate that even signals in which neither the common nor the innovation components are sparse, can be reconstructed to a reasonable degree ofaccuracy as long as the inter-signal correlation structure exists.

Robustness with respect to non-stationarity: ARCS is able to reconstruct the signal even after the common component in the signals is reinitialized to include a different set of randomly chosen frequencies, once the correlation structure of the ensemble has been learnt.

Invariance to CS recovery algorithm: ARCS does not depend on any particular compressed sensing recovery algorithm used. In this paper we have employed basis pursuit approach which is a convex optimization. However, the CS recovery can equally well be done by other methods like matching pursuit [9] which could be faster but possibly less accurate. Essentially, the execution time of ARCS depends on the CS recovery method used. Thus, a faster algorithm can equally well be plugged into ARCS, to speed up its execution.

Finally, a candidate application we envisage, where ARCS could be employed, deserves mention. Consider a configuration of M electrodes on the human scalp in an EEG trial. Once the correlation structure of the signals is learnt by ARCS, one of the electrodes could be moved to a different location. Through CS, ARCS can reconstruct the signal at the location from where the electrode is removed, at each sampling instant. Simultaneously, ARCS starts learning the new configuration of $M + 1$ signals which includes the location from where electrode is removed as well as the new location where it is placed. This process continues until ARCS learns about $N > M$ locations and reconstructs N signals using only Melectrodes.

References

[1] J. Chen and X. Huo, "Theoretical Results on Sparse representation for multiple measurement vectors (MMV) in an over complete dictionary". *Signal Processing, IEEE Transactions on*, vol. 54, no.12, pp.4634-46430, Dec. 2006

[2] Z. Zhang and B.D. Rao, "Sparse Signal Recovery with Temporally Correlated Source Vectors using Sparse Bayesian Learning*", Selected Topics in Signal Processing, IEEE Jour. of,* vol.5, no.5, pp.912-926, 2011

[3] Garcia-Frias, J. and Esnaola, I, "Distributed compression of correlated real sequences using random projections", *Data Compression Conference, Proceedings of,* 25-27 March 2008.,pp 322-331

[4] M.F. Duarte, S. Sarvotham, D.Baron, M.B. Wakin, R.G. Baraniuk, "Joint Sparsity Models for Distributed Compressed Sensing", *Computer (2005)*

[5] M.F.Duarte, S.Sarvotham, D.Baron, M.B.Wakin, R.G.Baraniuk, "Distributed Compressed Sensing of Jointly Sparse Signals", *Signals, Systems and Computers, 2005. Conference Record of the Thirty-Ninth Asilomar Conference on*, vol., no., pp1537-1541, October 28 – November 1, 2005.

[6] D.L. Donoho, "Compressed Sensing", *Information Theory, IEEE Transactions on,* vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[7] E. Candes and M. Wakin, "An introduction to compressive sampling," IEEE Signal Proc. Magazine, vol. 25, no. 2, pp. 21–30, March 2008.

[8] Scott Shaobing Chen, David L. Donoho, Michael A.Saunders, "Atomic Decomposition by basis pursuit", *SIAM Journal on Scientific Computing,* vol.20, no.1,pp. 33-61, December 1998.

[9] Joel Tropp and Anna Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit", *Information Theory, IEEE Transactions on,* vol. 53, no.12, pp. 4655-4666, Dec 2007.