

# Distributed Compressive Sensing Reconstruction Via Common Support Discovery

Wei Chen\*, Miguel R. D. Rodrigues<sup>†</sup> and Ian J. Wassell\*

\* Computer Laboratory, University of Cambridge, UK

<sup>†</sup> Instituto de Telecomunicações, Department of Computer Science, University of Porto, Portugal

**Abstract**—This paper presents a novel signal reconstruction method based on the distributed compressive sensing (DCS) framework for application to wireless sensor networks (WSN). The proposed method exploits both the intra-sensor correlation and the inter-sensor correlation to reduce the number of samples required for recovering the original signals. An innovative feature of our method is using the Fréchet mean of the signals to discover the common support of their sparse representations in some basis. Then a new greedy algorithm, called precognition matching pursuit (PMP), is proposed to further reduce the number of required samples with the knowledge of the common support. The superior reconstruction quality of the proposed method is demonstrated by both computer-generated signals and real data gathered by a WSN located in the Intel Berkeley Research lab.

## I. INTRODUCTION

Wireless sensor networks (WSN) provide a flexible way to monitor the physical parameters of the environment through the deployment of a large number of sensor nodes. There are three main challenges in WSN, i.e., network lifetime, computational ability and bandwidth constraints. The first two challenges result from the size and cost limitation of sensor nodes, while the third challenge is due to the physical character of the wireless channel. Although the actual physical environmental information such as temperature and humidity is compressible owing to the presence of temporal correlation and spatial correlation, traditional signal compression methods are not usually applied until all the data is gathered at a fusion center (FC). Compressive sensing (CS) is a sampling theory, which leverages the compressibility of the signal to reduce the number of samples required for reconstruction. Using the CS technique as the data acquisition approach in a WSN can significantly reduce the energy consumed in the process of sampling and transmission through the network, and also lower the wireless bandwidth required for communication.

The traditional approach to measure 1-D environmental information, e.g., temperature, humidity, light and sound, is to uniformly sample and then report to the FC. The sampling rate could range from milliseconds to minutes depending on the particular application. In [1], [2], CS is used to directly acquire a reduced number of samples via randomized sampling in the time domain, and so no explicit compression process is required. To further utilize the spatial correlation of the environmental information, an extension of CS, named distributed compressive sensing (DCS) [3], is applied to reconstruct the signal obtained from a group of sensors [4], [5]. The spatial signal obtained from different sensors could turn out to be

a more compressible signal in a certain basis via reordering of the samples [6]. However, to the best of our knowledge there is no efficient method to find the optimal permutation of the sensors, although most existing variations of CS to WSN requires the preknowledge of the spatial distribution of the specific signal.

In this paper, we focus on a DCS model [3] to reduce the number of samples required for reconstruction. In this DCS model, signals detected by different sensors have different innovation supports but only one common support due to some global factors and spatial correlations. We propose a method to find the common support by calculating a Fréchet mean. With the knowledge of the common support, we propose a new greedy algorithm, called precognition matching pursuit (PMP), to reconstruct a group of signals collected by a WSN. The knowledge of the common support enable the PMP to reconstruct the original signals with a significantly reduced number of samples. The advantage of our method is demonstrated by recovering not only computer-generated signals but also real data gathered by a WSN at the Intel Berkeley Research lab [7].

The following notation is used. Upper-case letters denote numbers, boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and italics upper-case letters denote supports. The superscripts  $(\cdot)^H$  and  $(\cdot)^\dagger$  denote complex conjugate-transpose and pseudoinverse, respectively.  $\{\mathbf{X}_k\}$  and  $\{\mathbf{x}_k\}$  ( $k = 1, 2, \dots, K$ ) denote a group of matrices and vectors, respectively.  $\bar{A}$  denotes the complement of set  $A$ .  $\text{supp}\mathbf{x}$  and  $\text{supp}\mathbf{x}_B$  denote the set of points of vector  $\mathbf{x}$  where its absolute value is nonzero and one of the  $B$  largest values, respectively. By  $\mathbf{x}_J = D$ , we mean that points whose indexes are in the support  $J$  of vector  $\mathbf{x}$  is set to a number  $D$ .

## II. SYSTEM DESCRIPTION

### A. Compressive Sensing (CS)

We first consider a WSN with only one sensor. We are interested in a discrete-time signal  $\mathbf{f} \in \mathbb{R}^N$  which is sparse in some sparsifying basis  $\Psi \in \mathbb{R}^{N \times N}$  such as the Fourier transform. This discrete-time vector  $\mathbf{f}$  can be written in terms of a sparse representation  $\mathbf{x} \in \mathbb{R}^N$  as follows:

$$\mathbf{f} = \Psi\mathbf{x}. \quad (1)$$

The sensor generates a vector of measurements  $\mathbf{y} \in \mathbb{R}^M$  ( $M \ll N$ ) via random projections, which is represented

by multiplying by a sensing matrix  $\Phi \in \mathbb{R}^{M \times N}$ :

$$\mathbf{y} = \Phi \mathbf{f} + \mathbf{n} = \Phi \Psi \mathbf{x} + \mathbf{n}, \quad (2)$$

where  $\mathbf{n} \in \mathbb{R}^M$  represents the measurement noise. The entries of the sensing matrix could be generated using a random distribution, for example a Gaussian or a Bernoulli distribution. However, these random projection methods consume extra sampling power and require strict time synchronization with the FC in the reconstruction process. A technique, called causal random sampling, is used to avoid those issues. The entries of the sensing matrix  $\Phi$  are all zeros except for  $M$  entries in  $M$  different columns and rows. To maintain the causality of the sampling process, the order of the  $M$  unity entries in different rows should be sorted by the column number, e.g.

$$\Phi = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 \\ & & & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Define the matrix  $\mathbf{A} = \Phi \Psi$ . Then the recovery procedure corresponds to the solution of the optimization problem:

$$\min_{\hat{\mathbf{x}}} \|\hat{\mathbf{x}}\|_{\ell_1}, \quad \text{s.t.} \quad \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|_{\ell_2} \leq \epsilon, \quad (3)$$

where  $\epsilon > 0$  is an estimate of the measurement noise level. In [8], the author established that the reconstruction error of (3) obeys

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_{\ell_2} \leq C_0 S^{-1/2} \|\mathbf{x} - \mathbf{x}_S\|_{\ell_1} + C_1 \epsilon, \quad (4)$$

where  $\mathbf{x}_S$  is an approximation of  $\mathbf{x}$  with all but the  $S$ -largest entries set to zero, and  $C_0$  and  $C_1$  are constants depending on the restricted isometry constant (RIC)  $\delta$  of matrix  $\mathbf{A}$ . We note that with a reduced number of samples, the reconstructed signal  $\Psi \hat{\mathbf{x}}$  is a good approximation to the original signal if the signal is compressible.

Many approaches and their variants have been proposed in the literature to solve (3). Most of those algorithms can be classified into two categories - convex optimization algorithms and greedy algorithms. Generally speaking, convex optimization algorithms [9], [10] provide higher reconstruction performance in terms of number of samples required than do greedy algorithms. However, greedy algorithms still attract many researcher's attention owing to their low computational complexity. In particular, some greedy algorithms such as Compressive Sampling Matching Pursuit (CoSaMP) [11] and Subspace Matching Pursuit [12] deliver the same guarantees in terms of the RIC as the best optimization-based approaches.

### B. Distributed Compressive Sensing (DCS)

Here we consider a WSN with  $K$  sensors. Each sensor monitors the environmental parameters independently which are then transmitted to the FC through the network. DCS is proposed to reduce the total number of samples required to reconstruct the environmental signals by exploiting the joint

sparsity of the monitored signals. Suppose all the signals  $\{\mathbf{f}_k\}$  ( $k = 1, \dots, K$ ) monitored by different sensors have sparse representations  $\{\mathbf{x}_k\}$  in a sparsifying basis  $\Psi$ . The sparse common component and innovations model in [3] assumes that each signal contains a common component that is sparse plus an innovation component that is also sparse:

$$\mathbf{x}_k = \mathbf{z}_c + \mathbf{z}_k, \quad k = 1, \dots, K. \quad (5)$$

In this paper, we focus on a relaxed model of the sparse common component and innovations model. In our model, each signal contains a sparse innovation  $\mathbf{z}_{d,k}$  and a sparse similar component  $\mathbf{z}_{c,k}$ , where the similar components  $\{\mathbf{z}_{c,k}\}$  share the same sparse support and sign but have different amplitudes:

$$\mathbf{x}_k = \mathbf{z}_{c,k} + \mathbf{z}_{d,k}. \quad (6)$$

This relaxed model can be seen as a generalized sparse common component and innovations model where the common component condition is relaxed to become what we will name a similar component condition.

An application that is well-suited to this model is environmental temperature monitoring by a large number of wireless sensors deployed at different locations. Owing to the smooth change of physical parameters, the space-time temperature signals  $\{\mathbf{f}_k\}$  are compressible in both the time and space domain. Global factors, e.g., the sun and prevailing winds, could result in an effect  $\mathbf{z}_{c,k}$  which affects all sensors but with a different order of severity due to the terrain and shade. Local factors, e.g., animal activities and fire, could lead to sparse innovations  $\mathbf{z}_{d,k}$ . There and also related scenarios such as relative humidity, light intensities and air pressure monitoring via a WSN.

### III. COMMON SUPPORT DISCOVERY

In this section, we aim to find the common support  $J_c$  of the similar components  $\{\mathbf{z}_{c,k}\}$ , which could be used to enhance the algorithm performance in terms of the number of samples required. A natural idea is using some form of the mean of all the sparse representations as an approximation. As the sparse similar components  $\{\mathbf{z}_{c,k}\}$  share the same support and sign, the mean result keeps the value in the common support while depresses the value in the innovation supports. However, the mean vector cannot be directly calculated without explicitly presenting the sparse representations. Instead, we exploit the Fréchet mean to derive an estimate of the common support, which can be acquired by solving an ordinary least squares problem.

The Fréchet mean  $\tilde{\mathbf{x}}$  of  $K$  sparse representations is defined as follows:

$$\tilde{\mathbf{x}} = \arg \min_{\tilde{\mathbf{x}}} \sum_{k=1}^K d^2(\tilde{\mathbf{x}}, \mathbf{y}_k, \mathbf{A}_k), \quad (7)$$

where  $\mathbf{y}_k = \Phi_k \mathbf{f}_k + \mathbf{n}_k = \mathbf{A}_k \mathbf{x}_k + \mathbf{n}_k$  is the sample vector from the  $k$ th sensor,  $\Phi_k$  is the  $k$ th sensor's sensing matrix,  $\mathbf{A}_k = \Phi_k \Psi$ , and the distance function is defined here as

$$d(\tilde{\mathbf{x}}, \mathbf{y}_k, \mathbf{A}_k) = \|\mathbf{A}_k \tilde{\mathbf{x}} - \mathbf{y}_k\|_{\ell_2}. \quad (8)$$

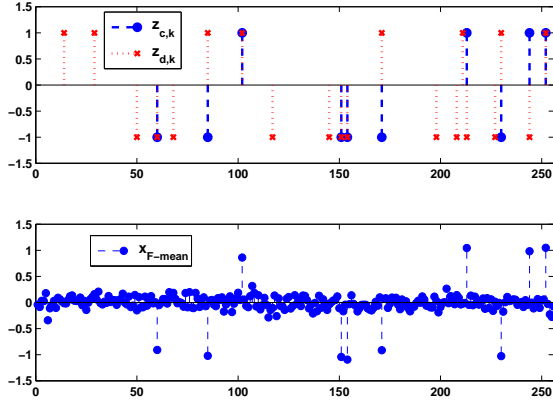


Fig. 1. Comparing the sparse representation (above) with the Fréchet mean (below). Each sparse representation ( $N=256$ ) contains 10 common components (blue) and 20 innovations (red), all of which are randomly placed  $\pm 1$  spikes. The Fréchet mean is calculated from 40 sparse representations measured by independent random Gaussian matrices ( $M=100$ ).

Then, from the equation  $\mathbf{y}_k = \mathbf{A}_k \mathbf{x}_k + \mathbf{n}_k$ , we can rewrite the distance function as

$$d(\tilde{\mathbf{x}}, \mathbf{y}_k, \mathbf{A}_k) = \|\mathbf{A}_k \tilde{\mathbf{x}} - \mathbf{A}_k \mathbf{x}_k - \mathbf{n}_k\|_{\ell_2}, \quad (9)$$

which is equal to the Euclidean distance in a low-dimensional space between  $\tilde{\mathbf{x}}$  and  $\mathbf{x}_k$ , in the absence of the noise term. The Fréchet mean can be obtained by solving the following problem:

$$\min_{\tilde{\mathbf{x}}} \left\| \hat{\mathbf{A}} \tilde{\mathbf{x}} - \hat{\mathbf{y}} \right\|_{\ell_2}^2, \quad (10)$$

where the extended sensing matrix  $\hat{\mathbf{A}}$  and the extended measurement vector  $\hat{\mathbf{y}}$  are given by

$$\hat{\mathbf{A}} = [\mathbf{A}_1^H \cdots \mathbf{A}_K^H]^H, \quad (11)$$

and

$$\hat{\mathbf{y}} = [\mathbf{y}_1^H \cdots \mathbf{y}_K^H]^H. \quad (12)$$

The Fréchet mean can be considered as the result of minimizing the sum of the squared distances between the observed set and the predicted point. If  $\text{rank}(\hat{\mathbf{A}}) = N$ , (10) turns to be an ordinary least squares problem, of which the result can be written using an explicit formula  $\tilde{\mathbf{x}} = (\hat{\mathbf{A}}^H \hat{\mathbf{A}})^{-1} \hat{\mathbf{A}}^H \hat{\mathbf{y}}$  or computed with the use of a speeding-up algorithm such as conjugate gradient (CG) [13]. Consequently, for a sequence of independently generated random matrices  $\Phi_k$  (and hence  $\mathbf{A}_k$ ), at least  $K = \lceil \frac{N}{M} \rceil$  measured signals are required to make the system overdetermined, which is satisfied in a WSN having only a low number of sensors.

Fig 1 gives a simplified example of the Fréchet mean, which is derived from 40 correlated sparse representations. As shown in Fig 1 the mean result keeps the value in the common support while depresses the value in the innovation support, then the common support can be detected by selecting absolute values above a given threshold. In addition, if the common support size is known, we can detect the support by choosing a given number of largest absolute values. The detected common support cannot be guaranteed to be exactly the same as the real

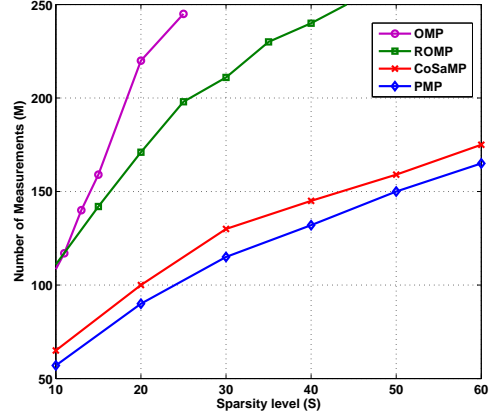


Fig. 2. The 99% reconstruction trend of different algorithms ( $N=256$ ). Each sparse representation in the simulation contains half common components and half innovations, all of which are randomly placed  $\pm 1$  spikes. For the PMP, the Fréchet mean is calculated from 40 sparse representations measured by independent random Gaussian matrices.

one. However, the proposed reconstruction algorithm put forth in the next section does not require an exactly correct common support for proper reconstruction.

#### IV. PRECOGNITION MATCHING PURSUIT (PMP)

We have developed a new greedy algorithm, namely pre-cognition matching pursuit (PMP), which aims to achieve fast reconstruction and a reduced requirement on the number of measurements compared to traditional matching pursuit (MP) and its variations [11], [14], [15]. The PMP learns from partial support information, e.g., the common support detected by calculating the Fréchet mean in the DCS model, and returns a better signal in terms of the number of samples required. The pseudo-code of the PMP for a group of signals can be described as follows:

---

##### Algorithm 1 Pre-cognition matching pursuit (PMP)

---

**Input:** A group of measurement vectors  $\mathbf{y}_k$  ( $k = 1, 2, \dots, K$ ), a group of measurement matrices  $\mathbf{A}_k$  ( $k = 1, 2, \dots, K$ ), a group of sparsity levels  $S_k$  ( $k = 1, 2, \dots, K$ ), and an estimated common support  $J_c$ ;

**Output:** An estimate group of sparse signals  $\hat{\mathbf{x}}_k$  ( $k = 1, 2, \dots, K$ );

**Process:** For  $k > 0$ , do

- 1) **Initialization:** Let  $\hat{\mathbf{x}}_k = 0$  and  $\mathbf{r} = \mathbf{y}_k$ ;
  - 2) **Repeat:**
    - $\mathbf{u} = \mathbf{A}_k^H \mathbf{r}$ ,  $W = 2 \cdot (S_k - \|J_c\|_{\ell_0})$ ;
    - $T = \text{supp} \mathbf{u}_W$ ,  $O = \text{supp} \hat{\mathbf{x}}_k$ ;
    - $J = T \cup J_c \cup O$ ;
    - $\mathbf{v} = \mathbf{A}_{k,J}^\dagger \mathbf{y}_k$ ,  $Q = \text{supp} \mathbf{v}_{S_k}$ ;
    - $\hat{\mathbf{x}}_{k,Q} = \mathbf{v}_Q$ ,  $\hat{\mathbf{x}}_{k,\bar{Q}} = 0$ ,  $\mathbf{r} = \mathbf{y}_k - \mathbf{A}_{k,J} \hat{\mathbf{x}}_k$ ;
    - **if** halting condition true **then**  
quit the iteration;
    - end if.**
-

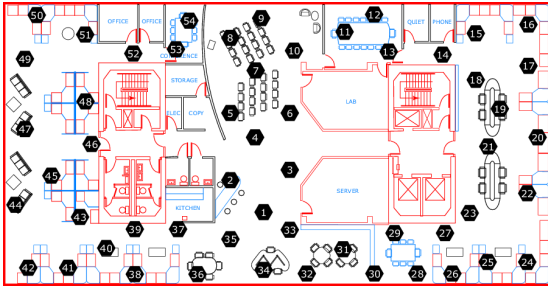


Fig. 3. Layout of the WSN of the Intel Berkeley Research lab.

The key idea of the PMP is to try to iteratively estimate the support of the sparse signal  $\mathbf{x}_k$ . The selected support  $J$  in each iteration contains the locations of  $W = 2 \cdot (S_k - \|J_c\|_{\ell_0})$  largest values of a residual vector returned in the previous iteration, the known partial support  $J_c$  of  $\mathbf{x}_k$ , and the locations of the nonzeros of previous  $\hat{\mathbf{x}}_k$ . With the estimated support  $J$ , the reconstructed signal  $\hat{\mathbf{x}}_k$  is calculated by solving the following least square problem

$$\min_{\hat{\mathbf{x}}_k} \|\mathbf{A}_{k,J} \hat{\mathbf{x}}_{k,J} - \mathbf{y}_k\|_{\ell_2}^2. \quad (13)$$

Next, all the elements in  $\hat{\mathbf{x}}_k$  are set to zero except for the  $S_k$  largest values. For each reconstructed signal in the group, iterations stop when the result satisfies some halting condition. One halting condition for sparse signals is that the residual's norm  $\mathbf{r}$  is smaller than a given threshold. For an exactly sparse signal, the threshold can be chosen as the noise level for noisy measurements or zero for noiseless measurements. For a compressible signal, to the best of our knowledge, there is no way to identify the best threshold. Therefore, one can either set a threshold based on empirical knowledge or let the algorithm stop when the improvement of the residual is smaller than a certain threshold.

We note that the PMP needs knowledge about the sparsity level  $S$ , which is also required in many other MP algorithms, such as orthogonal matching pursuit (OMP), regularized orthogonal matching pursuit (ROMP) and compressive sampling matching Pursuit (CoSaMP). As  $\mathcal{O}(S \log \frac{N}{S})$  measurements are enough for reliable reconstruction [16], the sparsity level can be roughly estimated as  $\frac{M}{\log N}$ . Alternatively, we can also test several sparsity levels and choose the one with minimal least error  $\|\mathbf{A}_k \hat{\mathbf{x}}_k - \mathbf{y}_k\|_{\ell_2}$ .

One can observe that the PMP is similar to the CoSaMP [11] except for an additional common support and a smaller size of discovered support in each iteration. This key innovation enables the PMP to return a better reconstructed signal. A comparison of different reconstruction algorithms is shown in Fig 2, which considers an exactly sparse signal and noiseless measurements.

## V. EXPERIMENTAL RESULTS

In this section, we apply the proposed method described previously to reconstruct the signals gathered by the WSN located in the Intel Berkeley Research lab [7]. In this WSN,

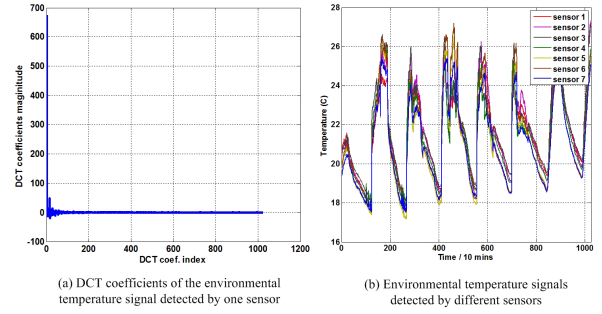


Fig. 4. Environmental temperature signals detected by the WSN.

54 Mica2Dot sensors with weather boards were deployed in one floor of the lab building as shown in Fig 3. Each sensor detected the environmental humidity, temperature and light every 31 seconds for more than one month. We use randomly selected samples of the original data for evaluating the proposed method, i.e., the sampling manner is similar to randomly sampling the original signals.

Instead of uniform sampling, each sensor independently and randomly collects a small portion of the original samples and transmits them to the FC. Note that the signals monitored by the WSN might be compressible rather than exactly sparse in some sparsifying basis, and all the measurements are disturbed by the noise. This differs from the scenario used for simulations in the previous sections. In the following evaluation of our method, we use the discrete cosine transform (DCT) as the sparsifying domain for simplicity since it avoids dealing with the complex values yielded by the Fourier transform. We note that the environmental temperature signals in the detected area have sparse DCT coefficients vectors as shown in Fig 4(a), and have high spatial correlations as shown in Fig 4(b).

All signals we study in the following have a length of

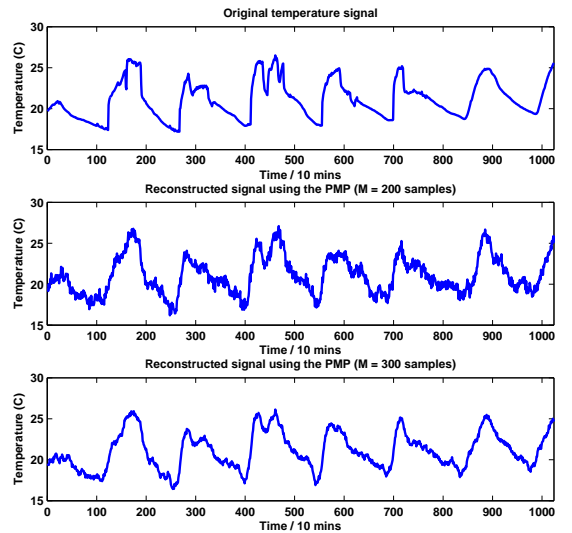


Fig. 5. Environmental temperature signals detected by the WSN and reconstruction results using the PMP algorithm.

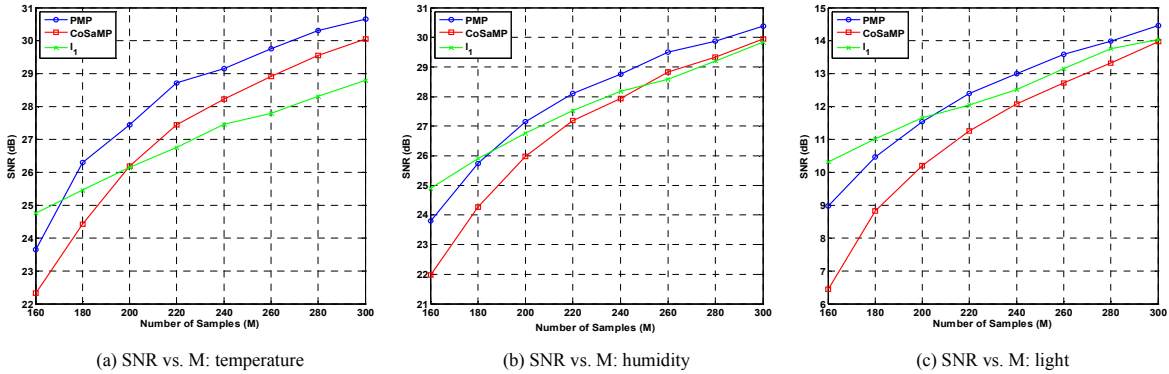


Fig. 6. Performance of three different reconstruction algorithms in terms of real signals ( $N=1024$ ) detected by the WSN.

1024. We assume the sparsity level  $S = 50$  in our simulation according to Fig 4(a), and half of the sparse components have a common support. We first calculate the Fréchet mean from the randomly sampled signal of different sensors, and then obtain the common support from the Fréchet mean, which is used as an input of the PMP to independently reconstruct each signal. In Fig 5, we demonstrate the effectiveness of our method in terms of reconstructing the temperature signal with a reduced number of samples. We achieve similar performance in reconstructing the humidity signal and the light signal, but these are not shown here due to the limited space. We define the reconstruction signal to noise ratio (SNR) to be  $\frac{\sum_{k=1}^K \|\hat{\mathbf{f}}_k\|_{\ell_2}^2}{\sum_{k=1}^K \|\mathbf{f}_k - \hat{\mathbf{f}}_k\|_{\ell_2}^2}$ , where  $\mathbf{f}_k$  and  $\hat{\mathbf{f}}_k$  are the original signal and reconstructed signal of the  $k$ th sensor respectively. In Fig 6, we show the superiority of our method in comparison with the  $\ell_1$  minimization algorithm and the CoSaMP algorithm which is one of the most efficient algorithms in the category of greedy methods. The performance of the  $\ell_1$  minimization algorithm, where we use the CVX code (<http://www.stanford.edu/~boyd/cvx/>), is better than the proposed method for smaller numbers of samples. However, the  $\ell_1$  minimization algorithm takes several minutes to reconstruct the group of signals while our methods just requires a few seconds using the same PC, which makes our method promising in terms of computational complexity.

## VI. CONCLUSIONS

In this paper, we study the problem of monitoring the physical parameters of a real environment by a WSN, and focus on the DCS framework to reduce the number of samples required to reconstruct the original signals. As the number of samples is proportional to the total energy consumed by a sensor, which includes sampling, processing and transmission. Therefore, a reduced number of samples made by each sensor would save the battery power of the sensor and prolong the life time of the WSN. We propose a method to discover the common support of a group of signals by calculating a Fréchet mean, which is shown to be effective. By using the estimated common support, we also propose a greedy algorithm, called PMP, to reconstruct a group of correlated signals. We show that the PMP algorithm requires less samples for exact reconstruction than most of the

other greedy algorithms including OMP, ROMP and CoSaMP in the DCS model. In addition, experiments demonstrate the advantage of our method in the reconstruction of actual signals collected by a practical WSN.

## REFERENCES

- [1] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.
- [2] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [3] D. Baron, M. Wakin, M. Duarte, S. Sarvotham, and R. Baraniuk, "Distributed compressed sensing," 2005, Preprint.
- [4] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressive wireless sensing," in *Proceedings of the 5th international conference on Information processing in sensor networks*. ACM, 2006.
- [5] W. Wang, M. Garofalakis, and K. Ramchandran, "Distributed sparse random projections for refinable approximation," in *6th International Symposium on Information Processing in Sensor Networks, IPSN*, 2007, pp. 331–339.
- [6] M. Mahmudimanesh, A. Khelil, and N. Suri, "Reordering for Better Compressibility: Efficient Spatial Sampling in Wireless Sensor Networks," in *2010 IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*, 2010, pp. 50–57.
- [7] "Intel berkeley lab wsn, <http://db.csail.mit.edu/labdata/labdata.html>."
- [8] E. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes rendus-Mathématique*, vol. 346, no. 9–10, pp. 589–592, 2008.
- [9] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale 1-regularized least squares," *IEEE journal of selected topics in signal processing*, vol. 1, no. 4, pp. 606–617, 2007.
- [10] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal on selected topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.
- [11] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.
- [12] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing: Closing the gap between performance and complexity," *submitted*, 2008.
- [13] J. Nocedal and S. Wright, *Numerical optimization*, 2nd ed. Springer verlag, 2006.
- [14] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, p. 4655, 2007.
- [15] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Foundations of computational mathematics*, vol. 9, no. 3, pp. 317–334, 2009.
- [16] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.