Distributed Compressed Sensing for Sensor Networks, Using p-thresholding

Mohammad Golbabaee, Pierre Vandergheynst

Distributed Compressed Sensing for Sensor Networks, Using p-thresholding

Mohammad Golbabaee and *Pierre Vandergheynst*
Signal Processing Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
E-mail:{mohammad.golbabaei, pierre.vandergheynst}@epfl.ch

Problem statement

Distributed compressed sensing is the extension of compressed sampling (CS) to sensor networks. The idea is to design a *CS joint decoding* scheme at the base station which exploits the inter-sensor correlations, in order to recover the whole observations from very few number of random measurements per node. Here, the questions are about modeling the correlations, design of the joint recovery algorithms, analysis of those algorithms, the comparison between the performance of the joint and separate decoder and finally determining how optimal they are.

Originality of the work

Baraniuk et al.[1] address the first two questions by defining a joint sparsity model (i.e. a common sparsity support throughout the channels plus innovations) and proposing a simple joint recovery algorithm (OSGA) which shows a massive improvement by increasing the number of channels. However, since their analysis is limited to infinitely large networks and does not have a close form, it is not clear how effectively their algorithm take advantage of the available *diversity* from various channels. In this paper, for the same signal model, we propose a joint recovery algorithm (p-thresholding), which is more general than OSGA. Moreover, using concentration of measure techniques, we bound its performance by a closed form expression that is valid in both asymptotical and finite size configurations. This enables us to see whether our joint decoder achieves the full diversity from different channels and how near optimal it performs.

New results

In particular, we study two types of compressing matrices with i.i.d. Gaussian/Bernoulli entries. However, thanks to the generality of concentration arguments, our results are valid for a large class of random matrices including the subgaussian ones. We have proved that by increasing the number of channels, the recovery probability improves exponentially. As a consequence, we show that by choosing $p = 1$, p-thresholding exploits the *full diversity* from all channels, to reconstruct the common support from the minimum number of the measurements per node that one could hope.

1 Signal model

As mentioned above, we consider a sensor network having K nodes with sparse observations at each node k, $x(k) \in \mathbb{R}^N$. According to JSM2 in [1], observations of all nodes have the same sparsity support Λ (set of nonzero locations) of cardinality S. Each node uses its own random compressing matrix $\Phi(k)_{M\times N}$ to collect M measurements $y(k) \in \mathbb{R}^M$ as follows:

$$\begin{aligned} y(k) &= \Phi(k)x(k) && (1) \\ &= \Phi_\Lambda(k)\theta(k) \quad k = 1, ..., K. && (2) \end{aligned}$$

Where $\Phi_\Lambda(.)$ and $\theta(.)$ (nonzero coefficients) are just restrictions of $\Phi(.)$ and $x(.)$ to the joint support.

2 p-thresholding algorithm

Our algorithm is mainly inspired by *Thresholding* [2] (also known as Maximum correlation detection in [3]), which is the simplest greedy algorithm for sparse recovery and its extension to simultaneous sparse recovery, *p-thresholding* [4]. While the latter deals with joint recovery of sparse signals in a unique and deterministic dictionary, here, we modify it for distributed CS applications, where dictionaries are random and also independent at each node. The steps can be briefly summarized as follows:

1. For each node, compute $\rho \in \mathbb{R}^N$ which gives the correlations between its measurements and all columns of the corresponding sensing matrix.

$$\rho(k) = \Phi^*(k)y(k). \tag{3}$$

2. Built the vector of statistics ($f \in \mathbb{R}^N$) by p-norming the raws of the correlation matrix $\Gamma \in \mathbb{R}^{\mathbb{N}\times\mathbb{K}}$ which contains all $\rho(.)$s as its columns,

$$f_i = \left(\sum_{k=1}^{K} |\rho_i(k)|^p \right)^{1/p} \qquad i = 1, ..., N. \tag{4}$$

3. Sort out the S largest elements of the statistics vector. Their indices indicate the recovered support set $\hat{\Lambda}$.

4. For each channel, compute the nonzero coefficients by projecting its measurement vector to the corresponding recovered subspace, $\Phi_{\hat{\Lambda}}(.)$.

$$\hat{\theta}(k) = \Phi^\dagger_{\hat{\Lambda}}(k)y(k), \tag{5}$$

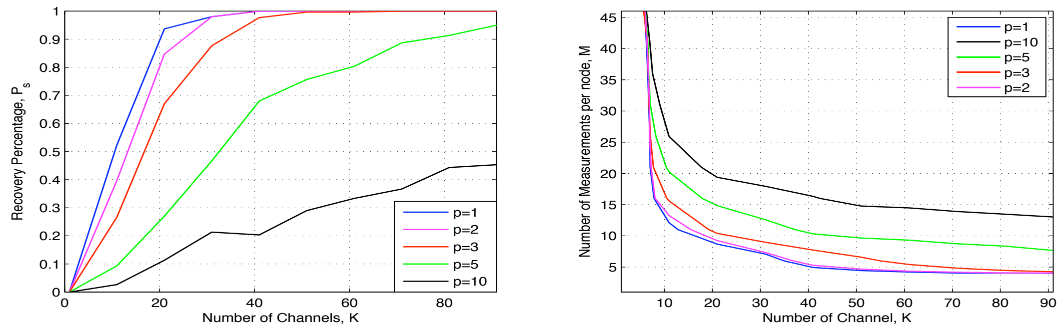
where by $A^\dagger$ we denote the pseudo inverse of the matrix A.

As we can see, for the case $p = 2$ this algorithm will be the same as OSGA in [1].

3 Main results and analysis

In this paper, we mainly focus on the support recovery, since success in this phase leads to the correct nonzero coefficients and hence, full recovery. Without getting too much into the details, for both i.i.d Gaussian and Bernoulli compressing matrices $\Phi(k)$, we upper bound the probability that the algorithm fails to recover the correct support by,

$$P_{fail} \leq CN \exp\left(-c\ R^2 \frac{MK^{1/p}}{S} \right), \tag{6}$$



(a)

(b)

Figure 1: Recovering the joint support, using p-thresholding. Simulation setup: Gaussian compressing matrices, Signal length $N = 50$, Sparsity $S = 5$, Dynamic range $R = 1$ and 300 independent trials. (a) Recovery rate vs. Number of channels for $M = 15$ and several p. (b) Recovery region for $P_s = 0.6$ and several p. The region above each line has recovery probability greater than P_s.

where c and C are some constants and R is the dynamic range of the nonzero coefficients, $R = \frac{\min_{k,i} |\theta_i(k)|}{\max_{k',i'} |\theta_{i'}(k')|}$. The proof will be presented in the original paper and is mainly based on concentration of measure arguments. We can see from both (6) and the simulation results in Figure (1a) that, by increasing the number of sensors, thanks to the higher diversity from more channels, our algorithm becomes exponentially unlikely to fail. Further, using (6) we can deduce that, a reliable recovery (i.e. $P_{fail} \leq \delta \ll 1$) requires the number of random measurements per node be greater than,

$$M \gtrsim O(S/K^{1/p} \log(N/\delta)). \tag{7}$$

This shows that, by growing the network size, each node needs to send much less measurements to our joint decoder, for a correct support estimation. Unlike in [1], this bound is not restricted only to the asymptotical cases ($K \to \infty$), but in addition, it points out clearly the tradeoff between K, M, S and N, for a reliable recovery, even in finite size sensor networks.

More importantly, (6) and (7) both imply that p-thresholding exploits the maximal (full) diversity of various channels for $p = 1$, whereas for higher p the channel diversity decreases such that, when p tends to infinity, there will be no difference between a joint or a separate decoder. This is along with Figure (1b), where simulation results indicate a greater recovery region (specially along K axis) for 1-thresholding, comparing to OSGA or any higher p.

References

[1] Dror Baron, Michael B.Wakin, Marco F.Duarte, Shriram Sarvotham, and Richard G.Baraniuk, "Distributed compressed sensing for jointly sparse signals," in *Thirty-ninth Asilomar conference on signals, systems and computers*, 2005, vol. 24, pp. 1537–1541.

[2] Karin Schnass and Pierre Vandergheynst, "Average Performance Analysis for Thresholding," *IEEE Signal Processing Letters*, vol. 14, no. 11, pp. 828–831, 2007.

[3] Alyson K. Fletcher, Sundeep Rangan, and Vivek K. Goyal, "Necessary and sufficient conditions on sparsity pattern recovery," *arXiv:0804.1839v1*, April 2008.

[4] Rémi Gribonval, Holger Rauhut, Karin Schnass, and Pierre Vandergheynst, "Atoms of all channels, unite! average case analysis of multi-channel sparse recovery using greedy algorithms," *Journal of Fourier Analysis and Applications*, 2008, Accepted to the special issue on Sparsity.