Study on Performance Behavior of Compressive Sensing Measurements for Multiple Sensor System

Sangjun Park, Hwanchol Jang and Heung-no Lee
Information and Communications
Gwangju Institute of Science and Technology
Gwangju, Republic of Korea
sjpark1@gist.ac.kr, hcjang@gist.ac.kr and heungno@gist.ac.kr

***Abstract*—In this paper, we will analyze the performance limit for a multiple sensor system (MSS) based on compressive sensing. In our MSS, all of the sensors measure signals from a common source. There exists the redundancy in the measured signal because the measured signal comes from the common source. To reduce communication costs, this redundancy must be removed. For this purpose, we use compressive sensing at each sensor to obtain compressed measurements. After all of the sensors obtain compressed measurements, they transmit them to a central unit. A decoder at the central unit receives all of the transmitted signals and attempts to jointly estimate the correct support set, which is the set of indices corresponding to the locations of the non-zero coefficients of the measured signals. In order to analyze our MSS, we present a jointly typical decoder inspired by recent work [4]. We first obtain the upper bound probability that the jointly typical decoder fails to estimate the correct support set. Next, we prove that as the number of sensors increases, the compressed measurements per sensor (per-sensor measurements) can be reduced to sparsity, which is the number of non-zero coefficients in the measured signal. We present the sufficient number of sensors required with the increase in the noise variance.**

Keywords–Multiple Sensor System, Joint Typicality, Compressive Sensing, Sparse Signal, Per-Sensor Measurements

I. INTRODUCTION

We consider a multiple sensor system (MSS), where all of the sensors measure signals from a common source in a limited region. Each sensor in the MSS independently transmits the measured signal to a central unit. At the central unit of the MSS, all of the transmitted signals can be used to recognize the common source.

The distribution of several sensors is a good way to precisely knowing about a common source. However, this approach has a drawback: The coverage area for each sensor can significantly overlap with that of other sensors. As a result, there may be a high level of redundancy in the measured signal at each sensor. This redundancy must be removed to reduce communication costs before all of the sensors transmit the measured signal to the central unit. Thus, we need a compression technique to remove the redundancy.

Recently, compressive sensing [3] (CS) has been extensively studied since it has been proven that a sparse signal, which is highly dimensional and contains a small number of non-zero coefficients, can be recovered from a small number of compressed measurements with dimensions much smaller than the sparse signal. Compressive sensing can be considered to be a compression technique that exploits prior information that any signal can be sparse in a certain transform domain.

In the MSS, each sensor compresses the measured signal to remove redundancy. One way to accomplish this is to allow the sensors to communicate with each other. However, an exchange of signals incurs additional communication costs and should hence be avoided. Therefore, we need another compression technique that allows sensors to remove redundancy without any communication with other sensors.

In order for all of the sensors to compress the measured signal, we propose the use of compressive sensing. Compressive sensing can be effectively used to remove redundancy without communication of sensors, by exploiting the fact that the locations of non-zero coefficients are shared. Thus, sensors can compress their measured signals by using compressive sensing independently. Subsequently, they transmit their compressed measurements to the central unit. A decoder at the central unit jointly estimates the locations of non-zero coefficients, which are a support set, by using the fact that the locations of the non-zero coefficients are shared. Eventually, the decoder computes all of the measured signals by using the estimated support set.

Using compressive sensing in the MSS provides many advantages. First, as has been mentioned, it allows the sensors to remove the redundancy. Another advantage is the ability to control the performance of the MSS by controlling the number of compressed measurements. For example, if we increase the number of compressed measurements, the decoder at the central unit tends to accurately estimate the support set. On the other hand, if we decrease the number of compressed measurements, the decoder may fail to estimate the correct support set.

In order to investigate the performance limit for an MSS based on compressive sensing, we present a jointly typical decoder inspired by recent work [4]. Then, we will define some failure events where we consider that the jointly typical decoder fails to estimate the support set. We will also provide the probabilities for these events. Finally, we will show that as the number of sensors increases, the number of compressed measurements per sensor (PSM) converges to sparsity, which is the number of non-zero coefficients in the measured signal or the cardinality of the correct support set. We will determine the number of sensors required with the increase in noise variance as well.

The rest of this paper is organized as follows. In Section II, the system model is described. The jointly typical decoder and events are described in section III. The main results and discussion are presented in section IV, and the conclusions of this study and plans for future works are described in Section V.

II. System Model

There exist S sensors measuring a signal from a common source. Let the measured signal at each sensor be $\mathbf{x}_s \in \mathbb{R}^N$, where $s \in \{1, 2, \cdots, S\}$. The measured signal at each sensor is a K sparse signal, which implies that $\|\mathbf{x}_s\|_0 = K$, where $\|\mathbf{x}\|_0$ is the number of non-zero coefficients in $\mathbf{x}$. The notation of a support set is defined as

$$\mathcal{I} \triangleq \mathrm{supp}(\mathbf{x}) = \{i \mid x(i) \neq 0\}.$$

The support set $\mathrm{supp}(\mathbf{x})$ consists of indices corresponding to the non-zero elements of $\mathbf{x}$. Because of the redundancy, we assume that all of the support sets are the same. Thus, we have $\mathrm{supp}(\mathbf{x}_1) = \cdots = \mathrm{supp}(\mathbf{x}_S)$. The compressed measurements at each sensor are given as

$$\mathbf{y}_s = \mathbf{F}_s \mathbf{x}_s, \tag{1}$$

where $\mathbf{y}_s$ denotes the compressed measurements, and all of the elements of $\mathbf{F}_s$ are i.i.d. Gaussian random variables with zero-mean and unit variance. Then, the received signal from the s^{th} sensor is defined as

$$\mathbf{r}_s = \mathbf{y}_s + \mathbf{n}_s, \tag{2}$$

where all of the elements of $\mathbf{n}_s$ are i.i.d. Gaussian with zero-mean and variance σ^2_{noise}. We also assume that all of the noise vectors and all of the sensing matrices are mutually independent.

III. Jointly Typical Decoder and Events

The aim of the jointly typical decoder is to estimate the support set.

Definition 1: The jointly typical decoder gives us the set by employing all of the received vectors and all of the sensing matrices.

Clearly, if the output from the jointly typical decoder is different from the support set, it is a failure. Next, we introduce the notation of a joint typicality.

Definition 2: We say that an $SM \times 1$ vector $\mathbf{r} = \begin{bmatrix} \mathbf{r}_1^{\mathrm{T}} & \cdots & \mathbf{r}_S^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ and a set $\mathcal{J}$ with $|\mathcal{J}| = K$ are $\delta-$jointly typical if $\forall_s rank(\mathbf{F}_{s,\mathcal{J}}) = K$ and

$$\frac{1}{SM} \left| \sum_s \left(\left\| \mathbf{Q}_{\mathbf{F}_{s,\mathcal{J}}} \mathbf{r}_s \right\|^2 - (M-K)\sigma^2_{\text{noise}} \right) \right| < \delta, \tag{3}$$

where $\mathbf{Q}_{\mathbf{F}} = \mathbf{I} - \mathbf{F}\left(\mathbf{F}^{\mathrm{T}}\mathbf{F}\right)^{-1}\mathbf{F}^{\mathrm{T}}$ and $\mathbf{F}_{s,\mathcal{J}}$ is constructed by collecting a set of the column vectors of $\mathbf{F}_s$ corresponding to the indices of $\mathcal{J}$.

For simplicity, we denote a $\delta-$jointly typical event as $\mathrm{E}(\mathbf{r}, \mathcal{J}, \delta)$. For example, if $\mathrm{E}(\mathbf{r}, \mathcal{J}, \delta)$ occurs, then the jointly typical decoder provides $\mathcal{J}$ as the estimated support set.

Now, we define a couple of failure events. The first is when $\mathrm{E}\left(rank(\mathbf{F}_{s,\mathcal{J}}) \neq K\right)$ occurs. Clearly, for such an event, the jointly typical decoder cannot estimate the support set because we cannot evaluate (3). The second is when $\mathrm{E}(\mathbf{r}, \mathcal{J} \neq \mathcal{I}, \delta)$ occurs. In this case, the jointly typical decoder yields an incorrect support set. Thus, it must be a failure event. The last is when $\mathrm{E}(\mathbf{r}, \mathcal{I}, \delta)$ does not occur. The jointly typical decoder is not aware of the correct support set in this case.

The jointly typical decoder definitely gives an incorrect support set decision whenever any one of these three events occurs. Thus, an event where the jointly typical decoder fails to estimate the correct support set is defined as

$$\begin{aligned} \mathrm{E}(failure) = \mathrm{E}(\mathbf{r}, \mathcal{I}, \delta)^c & \bigcup_{\forall \mathcal{J} \neq \mathcal{I}, |\mathcal{J}|=K} \mathrm{E}(\mathbf{r}, \mathcal{J}, \delta) \\ & \bigcup_{\forall s, \forall \mathcal{J}, |\mathcal{J}|=K} \mathrm{E}\left(rank(\mathbf{F}_{s,\mathcal{J}}) \neq K\right) \end{aligned}. \tag{4}$$

By using the union bound approach, the probability of (4) is upper bounded. Thus, we have

$$\begin{aligned} \Pr\{\mathrm{E}(failure)\} \leq & \Pr\left\{\mathrm{E}(\mathbf{r}, \mathcal{I}, \delta)^c\right\} \\ & + \sum_{\forall \mathcal{J} \neq \mathcal{I}, |\mathcal{J}|=K} \Pr\{\mathrm{E}(\mathbf{r}, \mathcal{J}, \delta)\} \\ & + \sum_{\forall \mathcal{J}, |\mathcal{J}|=K, \forall s} \Pr\left\{\mathrm{E}\left(rank(\mathbf{F}_{s,\mathcal{J}}) \neq K\right)\right\} \end{aligned}. \tag{5}$$

Now, our aim is to obtain a couple of probabilities. In fact, the last term, i.e., $\Pr\left\{\mathrm{E}\left(rank(\mathbf{F}_{s,\mathcal{J}}) \neq K\right)\right\}$, in (5) is considered to be zero, which implies that the rank for each matrix $\mathbf{F}_{s,\mathcal{J}}$ is K with a high probability because all of the entries in matrix $\mathbf{F}_{s,\mathcal{J}}$ follow a Gaussian distribution. Thus, we can remove it in (5). Now, we consider the remaining terms. Let us consider the first term, i.e., $\mathrm{E}(\mathbf{r}, \mathcal{I}, \delta)^c$. We notice that $\mathrm{E}(\mathbf{r}, \mathcal{I}, \delta)^c$ is random because of all the noise vectors. Thus, we can obtain the probability of $\mathrm{E}(\mathbf{r}, \mathcal{I}, \delta)^c$. Lemma 1 provides the upper bound on the probability of $\mathrm{E}(\mathbf{r}, \mathcal{I}, \delta)^c$.

Lemma 1: Let $\mathcal{I}$ be the correct support set and $\forall_s rank(\mathbf{F}_{s,\mathcal{I}}) = K$. Then, for any $\delta > 0$,

$$\Pr\left\{\mathrm{E}\left(\mathbf{r},\mathcal{I},\delta\right)^c\right\} \le \exp\left(-\frac{SM}{2}\times\frac{\delta}{\sigma^2_{\text{noise}}}\right) \times\left(1+\frac{M}{M-K}\times\frac{\delta}{\sigma^2_{\text{noise}}}\right)^{S\times\frac{(M-K)}{2}}. \quad (6)$$

Similarly, we notice that $\mathrm{E}\left(\mathbf{r},\mathcal{J}\ne\mathcal{I},\delta\right)$ is random because of all of the noise vectors as well as all of the sensing matrices. Thus, we can compute its probability. Lemma 2 provides the upper bound on the probability of $\mathrm{E}\left(\mathbf{r},\mathcal{J}\ne\mathcal{I},\delta\right)$.

Lemma 2: Let $\mathcal{J}\ne\mathcal{I}$, $\forall_s rank\left(\mathbf{F}_{s,\mathcal{J}}\right)=K$, and $0\le\left|\mathcal{J}\bigcap\mathcal{I}\right|<K$. Then, for any $\delta>0$,

$$\Pr\left\{E\left(\mathbf{r},\mathcal{J},\delta\right)\right\} \le \exp\left(-\frac{S}{2\sigma^2_{\min}}\times\left(\left(M-K\right)\times\gamma+M\delta\right)\right) \times\left(\frac{\sigma^2_{\text{noise}}}{\sigma^2_{\min}}+\frac{M}{M-K}\times\frac{\delta}{\sigma^2_{\min}}\right)^{S\times\left(\frac{M-K}{2}\right)} \quad (7)$$

where $\gamma=\sigma^2_{\text{noise}}-\sigma^2_{\min}$, $\sigma^2_{s,\mathcal{J}}=\sum_{i\in\mathcal{I}\setminus\mathcal{J}} x_s\left(i\right)^2+\sigma^2_{\text{noise}}$, and $\sigma^2_{\min}=\min_s\left(\sigma^2_{1,\mathcal{J}},\cdots,\sigma^2_{s,\mathcal{J}}\right)$.

The proofs of both Lemma 1 and Lemma 2 are given in [2]. Using these two lemmas, we have the upper bound on the probability of $\mathrm{E}\left(failure\right)$. Before we introduce our theorems, we would like to emphasize the fact that all of the upper bound probabilities converge to zero as the number of sensors increases.

Proposition 1: Let $M>K$, $\forall_s rank\left(\mathbf{F}_{s,\mathcal{I}}\right)=K$, and $\delta>0$. Then, $\Pr\left\{\mathrm{E}\left(\mathbf{r},\mathcal{I},\delta\right)^c\right\}$ converges to zero as the number of sensors increases.

Propositions 2: Let $M>K$, $\forall_s rank\left(\mathbf{F}_{s,\mathcal{J}}\right)=K$, $\delta>0$, and $\sigma^2_{\text{noise}}<\min_s\sum_{i\in\mathcal{I}\setminus\mathcal{J}} x_s\left(i\right)^2$. Then, $\Pr\left\{E\left(\mathbf{r},\mathcal{J},\delta\right)\right\}$ converges to zero as the number of sensors increases.

Clearly, if $\sigma^2_{\text{noise}}\ge\min_s\sum_{i\in\mathcal{I}\setminus\mathcal{J}} x_s\left(i\right)^2$ occurs, the jointly typical decoder yields $\mathcal{J}$ as the correct support set because this decoder cannot distinguish between the noise components and the signal components. Thus, $\sigma^2_{\text{noise}}<\min_s\sum_{i\in\mathcal{I}\setminus\mathcal{J}} x_s\left(i\right)^2$ must be satisfied. The proofs of both Proposition 1 and Proposition 2 are also given in [2].

IV. Main Results

In this section, we introduce our two theorems. The first is related to the relationship between PSM and the number of sensors. The second shows the relationship between the number of sensors and the noise variance.

Theorem 1: Let $\forall_s rank\left(\mathbf{F}_{s,\mathcal{J}}\right)=K$, $\delta>0$, and $\sigma^2_{\text{noise}}<\min_s\sum_{i\in\mathcal{I}\setminus\mathcal{J}} x_s\left(i\right)^2$. Then, $\Pr\left\{\mathrm{E}\left(failure\right)\right\}$ converges to zero as the number of sensors increases till $M>K$.

Proof: The complete proof of this theorem is given in [2]; however, we would like to provide a brief description of the same here. Using Proposition 1 and Proposition 2, we know that all of the upper bound probabilities converge to zero as the number of sensors increases. When we carefully examine (5), we notice that the number of sensors does not impact the summation operation. Therefore, the upper bound on $\Pr\left\{\mathrm{E}\left(failure\right)\right\}$ converges to zero as the number of sensors increases.

Theorem 1 states that PSM converges to K as the number of sensors increases. Similar results were reported in [1][5]. In [5], it was reported that PSM converges to K under the noiseless case. In [1], it was reported that PSM converges to $2K$ under the noisy case. However, in [1], it was assumed that all of the sensing matrices were the same, i.e., $\mathbf{F}_1=\cdots=\mathbf{F}_S$. In our work, we consider the noisy case and assume that all of the sensing matrices are different.

Theorem 2: Let $\sigma^2_{1,\text{noise}}$, S_1, and $\mathrm{E}_1\left(failure\right)$ be the noise variance, number of sensors, and failure event of the 1st MSS, respectively. Let us assume that $\Pr\left\{\mathrm{E}_1\left(failure\right)\right\}\le\alpha$, where $\alpha\in\left(0,1\right)$. Let $\sigma^2_{2,\text{noise}}$, S_2, and $\mathrm{E}_2\left(failure\right)$ be the noise variance, number of sensors, and failure event of the 2nd MSS, respectively. If we suppose that $\sigma^2_{1,\text{noise}}\le\sigma^2_{2,\text{noise}}$ and $\sigma^2_{2,\text{noise}}<\min_s\sum_{i\in\mathcal{I}\setminus\mathcal{J}} x_s\left(i\right)^2$, then the sufficient condition for $\Pr\left\{\mathrm{E}_2\left(failure\right)\right\}\le\alpha$ is

$$S_2\ge\max\left(\frac{f\left(\frac{\delta}{\sigma_1^2}\times\lambda\right)}{f\left(\frac{\delta}{\sigma_2^2}\times\lambda\right)},\frac{g\left(\frac{\sigma_1^2}{\sigma^2_{\min,1}}+\frac{\delta}{\sigma^2_{\min,1}}\times\lambda\right)}{g\left(\frac{\sigma_2^2}{\sigma^2_{\min,2}}+\frac{\delta}{\sigma^2_{\min,2}}\times\lambda\right)}\right)\times S_1, \quad (8)$$

where $\lambda=\frac{M}{M-K}$, $f\left(x\right)=\log\left(1+x\right)-x$, and $g\left(x\right)=\log\left(x\right)-x+1$.

Proof: The proof is given in [2].

Theorem 2 provides the sufficient number of sensors for the jointly typical decoder to accurately estimate the correct support set even if the noise variance increases. It is an interesting theorem because it can be used to make the MSS robust against noise by sufficiently increasing the number of sensors.

V. CONCLUSIONS AND FUTURE WORKS

Our contributions in this paper are as follows. First, we showed that PSM converges to K as the number of sensors increases. This result is slightly better than the result in [5]. Duarte et al. obtained the same result, but they did not consider the presence of noise. On the other hand, we obtained our result under the noisy case. The second contribution was to providing the sufficient number of sensors for the MSS to make it robust against increases in the noise variance. This is important because it allows us to adaptively respond to the variation in the noise variance.

In future works, we would like to consider more practical models. For example, each signal uses the shared support set and an individual support set. In addition, we would like to consider our model again with the different assumption that all of the sensing matrices are the same. As already mentioned, a previous paper [1] showed that PSM converges to $2K$ under the assumption that all of the sensing matrices are the same. Thus, determining whether PSM converges to K or $2K$ would be beneficial.

ACKNOWLEDGMENT

This research was supported by the Leading Foreign Research Institute Recruitment Program through the National Research Foundation of Korea (NRF) funded by Do-Yak Research (N0.2011-0016496).

REFERENCES

[1] Gongguo Tang and Arye Nehorai, "Performance Analysis for Sparse Support Recovery," IEEE Trans. Inform. Theory, vol. 56, no. 3, March. 2010.

[2] Sangjun Park and Heung-No Lee, "Per-Sensor Measurements Required for a Noisy Distributed Compressive Sensor System," will be submitted to IEEE Trans. Inform. Theory.

[3] D. Donoho, "Compressive sensing," IEEE Trans. On Information Theory, vol. 52, pp. 1289-1306, 2006.

[4] Mehmet Akçakaya and Vahid Tarokh, "Shannon-Theoretic Limits on Noisy Compressive Sampling," IEEE Trans. Inform. Theory, VOL. 56, NO. 1, Jan. 2010.

[5] M. F. Duarte, S. Sarovotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed Compressed Sensing of Jointly sparse Signals," Proceedings of the 39th Asilomar Conference on Signals, Systems and Computation, Nov. 2005.