# Per-Sensor Measurements Behavior of Compressive Sensing System for Multiple Measurements

Sangjun Park, Junho Lee and Heung-no Lee
School of Information and Communications
Gwangju Institute of Science and Technology
Gwangju, Republic of Korea
sjpark1@gist.ac.kr, ljh8097@gmail.com and heungno@gist.ac.kr

***Abstract*— We aim to mathematically obtain the probability of recovery failure in distributed sensing and sampling for a Multiple Sensor System (MSS). In this system, sensors take samples while compressing the signal with linear projection operations using the idea of compressive sensing (CS) [1]. In particular, we show that a bound for per-sensor measurements (PSM), the number of compressed measurements required at each sensor for good signal recovery. Our focus is to see how PSM behaves as the number of sensors increases based on the failure probability. Using the idea of Joint Typicality [2][3], we show that PSM converges to the sparsity as the number of sensors increases.**

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

## I. INTRODUCTION

Multiple Sensor System (MSS) occurs in many application domains as the system in which many sensors are densely distributed in a limited region. In MSS, each sensor takes the samples on the relevant signals such as temperature/humidity/wind variations. The resolution of MSS is highly related with the number of sensors. It means that if we use many sensors, resolution increases. However, the downside may happen. It is that the coverage area for each sensor may be significantly overlapped and as the results, there may exist a huge amount of redundancy in the sensed data. Thus, there exists a trade off as we have to choose between the resolution and redundancy.

Meanwhile, the Compressive Sensing (CS) is an emerging field. It offers the new paradigm. CS is capable of combining the sampling and the compression operation into a single step [1]. CS exploits the prior knowledge that a naturally occurring signal has a sparse representation in some transform domain and is compressible. In the CS framework, the signal can be reconstructed by linear programming methods such as simplex method, interior point method and so on with the number of signal samples taken at a rate smaller than the Nyquist. This number is about $O\left(K\log\left(N/K\right)\right)$, where $K$ is the number of nonzero elements in the sensed signal and $N$ is the signal length.

The CS approach provides the following advantage. The time required to obtaining signal samples and the amount of computations required for the compression operation of stored signal samples can be significantly reduced since sampling and compression are performed simultaneously.

Applying this CS framework to MSS may offer a good solution in terms of power, processing time and so on. In the context of MSS, the central receiver gathers all compressed measurements and applies a joint recovery algorithm to reconstruct each signal. The central receiver may obtain benefits since theses compressed measurements taken from different sensors may exhibit *inter-sensor correlation*. With the exploitation of the inter-sensor correlation, per-sensor measurements (PSM), the number of compressed measurements required at each sensor for good signal recovery, can be reduced. We aim to investigate how PSM changes as the number of sensors increases. For this goal, we propose to use an information theoretic approach.

We achieve our goal by deriving the probability of recovery failure as the function of PSM, the number of sensors, and the sparsity. An information theoretic idea, called *the joint typicality* [3], is used in Akçakaya and Tarokh [2] for the compressive sensing context, whose work is for single sensor problem. We use the joint typicality to analyze MSS, and show that PSM converges to the sparsity as the number of sensors increases.

The rest of this paper is organized as follows. In Section 2, the system model is described. The main result is described in Section 3; numerical result is described in Section 4; conclusion and future works are described in Section 5.

## II. SYSTEM MODEL

In MSS, there are $S$ sensors distributed in the limited region, each observing a signal of its own. Let the sensed signal at each sensor be $\mathbf{x}_s \in \mathbb{R}^N$ where $s$ denotes the sensor index, i.e., $s \in \{1, 2, \cdots, S\}$. The signal at each sensor is a $K$ sparse signal with the sparsity equaling $K$.

We introduce the notation of *support set* which is defined as

$$\mathcal{I} \triangleq \operatorname{supp}\left(\mathbf{x}_s\right) = \{i \mid \mathbf{x}_s\left(i\right) \neq 0\},$$

where $\operatorname{supp}\left(\mathbf{x}_s\right) = \mathcal{I} \subset \{1, 2, ..., N\}$ is the set of indices corresponding to non-zero elements of $\mathbf{x}_s$. We assume that all support sets are the same. It means that indices corresponding to nonzero elements of signal are the same at each sensor and all sensors have the same sparsity. This is a reasonable assumption when each sensor independently observes a single sparse signal representing a globally occurring original signal pattern. An independent observation obtained at each sensor deployed at different locations within the site makes the values of the sparse signal different attenuations, caused by different signal propagation paths. Compressed measurements made at each sensor are then sent to the joint receiver. The joint receiver makes a noisy measurement of the compressed signal $\mathbf{y}_s$ from the $s^{th}$ sensor which can be expressed as

$$\mathbf{r}_s = \mathbf{y}_s + \mathbf{n}_s, \quad (1)$$

where $\mathbf{y}_s$ denotes the clean measurements, i.e., $\mathbf{y}_s = \mathbf{F}_s\mathbf{x}_s$ and $\mathbf{n}_s \in \mathbb{R}^M$ is the additive observation noise, elements of the $\mathbf{n}_s$ are i.i.d Gaussian with zero-mean and variance $\sigma_n^2$. The elements of $\mathbf{F}_s$ are i.i.d. Gaussian samples with zero-mean and a fixed variance with one. Each signal has the same sparsity $K$, i.e., $\|\mathbf{x}_s\|_0 = K$. The noise $\mathbf{n}_s$ and the sensing matrix $\mathbf{F}_s$ are mutually independent. Now, we consider the performance metric for the estimation.

**Definition of Proposed Error Metric:** Let $\hat{\mathbf{x}}_s$ be the recovered signal for sensor $s$. Then, the proposed error-metric (PEM) is defined:

$$p(\hat{\mathbf{x}}, x, S) = \mathbb{I}\left(\sum_{s=1}^{S} \frac{|\{i \mid \hat{\mathbf{x}}_s(i) \neq 0\} \bigcap \mathcal{I}|}{|\mathcal{I}|} = S\right), \quad (2)$$

where $\mathbb{I}(\bullet)$ is the indicator function. It is useful to note that PEM is 1 when all the recovered signals are correct. Otherwise it is zero.

**Definition of Decoder:** For PEM, we consider the average probability of error for all $\mathbf{F}_s$ matrices whose elements follow the normal Gaussian distribution.

$$\Pr(\textit{failure}) = \mathbf{E}_A\left(\Pr\left(p(\hat{\mathbf{x}}, x, S) = 0\right)\right)$$

The average probability of error is expectation over error event for PEM. Now, we say the decoder can recover signal reliably if $\Pr(\textit{failure}) \to 0$ as $S \to \infty$ if $M > K$.

We also define an utility index set $\mathcal{J}$, i.e., $\mathcal{J} \subset \{1, 2, ..., N\}$. Similar to $\mathcal{I}$, the maximum size of $\mathcal{J}$ is $K$. The number of elements in the intersection of $\mathcal{I}$ and $\mathcal{J}$ varies from 0 to $K-1$. We define $\mathbf{f}_s(i)$ the $i^{th}$ column of $\mathbf{F}_s$. We define two sub-matrices based on $\mathbf{F}_s$. A sub-matrix $\mathbf{F}_{s,\mathcal{I}}$ is constructed by collecting the set of column vectors of $\mathbf{F}_s$ corresponding to the indices of $\mathcal{I}$. Likewise, a sub-matrix $\mathbf{F}_{s,\mathcal{J}}$ is obtained from collecting the column vectors of $\mathbf{F}_s$ corresponding to the indices of $\mathcal{J}$. Finally, we define $\mathbf{O}_\mathbf{A}$ as the *orthogonal projection* matrix $\mathbf{O}_\mathbf{A} = \mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$ and $\mathbf{Q}_\mathbf{A}$ as the *projection matrix* which is defined as $\mathbf{Q}_\mathbf{A} = \mathbf{I} - \mathbf{O}_\mathbf{A}$.

**Definition of Jointly Typical Event:** We define a $\delta-$*jointly typical event* $E(\mathbf{r}, \mathcal{J})$, i.e., $E(\mathbf{r}, \mathcal{J}) \triangleq \{\,\mathbf{r}$ and $\mathcal{J}$ are $\delta-$ jointly typical$\}$ for all $\mathcal{J} \subset \{1, 2, ..., N\}$ with $|\mathcal{J}| = K$, which is given by

$$\left\{\left(\{\mathbf{F}_s\}, \mathcal{J}\right) \left| \; \left| \frac{1}{S}\sum_s \frac{1}{M}\left(\left\|\mathbf{Q}_{F_{s,\mathcal{J}}}\mathbf{r}_s\right\|^2 - (M-K)\sigma_n^2\right)\right| < \delta \right.\right\}. \quad (3)$$

**Probability of Jointly Typical Event:** If all column vectors of $\mathbf{F}_{s,\mathcal{J}}$ are mutually independent for all sensors, all noises are mutually independent, and for some $\delta > 0$, we can find $0 < \varepsilon < 1$ as we increase $S$ and $M$ that satisfies the following inequality

$$\Pr\{E(\mathbf{r}, \mathcal{J})\} \geq 1 - \varepsilon. \quad (4)$$

Before we introduce our main result, we provide two probabilities.

The received signal $\mathbf{r}$ is jointly typical with the correct index set $\mathcal{I}$ : Let all column vectors of $\mathbf{F}_{s,\mathcal{I}}$ be linearly independent. Then, $\mathbf{r}$ and $\mathcal{I}$ are $\delta-$ jointly typical with probability

$$\Pr\{E(\mathbf{r}, \mathcal{I})^c\} \leq \exp\left(-SM\frac{\delta}{2\sigma_n^2}\right)\left(\frac{M-K+M\frac{\delta}{\sigma_n^2}}{M-K}\right)^{\frac{S(M-K)}{2}} \quad \text{if } 0 < M\frac{\delta}{\sigma_n^2}$$

or

$$\Pr\{E(\mathbf{r}, \mathcal{I})^c\} \leq \exp\left(-SM\frac{\delta}{2\sigma_n^2}\right)\left(\frac{M-K+M\frac{\delta}{\sigma_n^2}}{M-K}\right)^{\frac{S(M-K)}{2}} +$$

$$\exp\left(SM\frac{\delta}{2\sigma_n^2}\right)\left(\frac{M-K-M\frac{\delta}{\sigma_n^2}}{M-K}\right)^{\frac{S(M-K)}{2}} \quad \text{if } 0 < M\frac{\delta}{\sigma_n^2} < S(M-K)$$

where $\delta_1 := \dfrac{\delta}{\sigma_n^2}$.

This probability is the failure probability with the support set. Similar, we define another failure probability with the incorrect set.

The received signal $\mathbf{r}$ is jointly typical with an incorrect index set $\mathcal{J}$ : Let all column vectors of $\mathbf{F}_{s,\mathcal{J}}$ for all sensors be linearly independent. Then, $\mathbf{r}$ and $\mathcal{J}$ are $\delta-$ jointly typical with probability

$$\Pr\{E(\mathbf{r},\mathcal{J})\} \le \exp\left(-\frac{S\left((M-K)\left(\sigma_n^2-\sigma_{\min}^2\right)+M\delta\right)}{2\sigma_{\min}^2}\right)\left(\frac{(M-K)\sigma_n^2+M\delta}{(M-K)\sigma_{\min}^2}\right)^{\frac{S(M-K)}{2}}$$

$$\text{if } S\left((M-K)\sigma_n^2-M\delta\right) \ge \mathbf{E}\left[\sum_s \left\|\mathbf{Q}_{F_{s,\mathcal{J}}}\mathbf{r}_s\right\|^2\right]$$

$$\Pr\{E(\mathbf{r},\mathcal{J})\} \le \exp\left(-\frac{S\left((M-K)\left(\sigma_n^2-\sigma_{\min}^2\right)+M\delta\right)}{2\sigma_{\min}^2}\right)\left(\frac{(M-K)\sigma_n^2+M\delta}{(M-K)\sigma_{\min}^2}\right)^{\frac{S(M-K)}{2}} -$$

$$\exp\left(-\frac{S}{2\sigma_{\min}^2}\left((M-K)\left(\sigma_n^2-\sigma_{\min}^2\right)-M\delta\right)\right)\left(\frac{(M-K)\sigma_n^2-M\delta}{\sigma_{\min}^2(M-K)}\right)^{\frac{S(M-K)}{2}} \text{ if } o.w.$$

where $\sigma_{s,\mathcal{J}}^2 := \sum_{i\in\mathcal{I}\setminus\mathcal{J}} x_s(i)^2+\sigma_n^2$ and $\sigma_{\min}^2 = \min\left(\sigma_{1,\mathcal{J}}^2,\cdots,\sigma_{S,\mathcal{J}}^2\right)$.

Detail proofs and explanations of them are described in [4].

## III. Main Result

**Theorem**: For any $\frac{(M-K)\sigma_n^2}{M} > \delta > 0$ and Gaussian measurement matrix $\mathbf{F}_s$ and with PEM given as (2), $\Pr(failure) \to 0$ as $S \to \infty$ if $M > K$.

*Proof*: The proof is described in [4].

This Theorem states that PSM can be reduced to $K$+1 when the number of sensors increases.

## IV. Numerical Result

In this section, we provide a numerical result. It shows that PSM can be reduced to $K$+1. According to a Figure 1, we easily notice that PSM is getting smaller while the number of sensors increases.

## V. Conclusions And Future works

We have shown that PSM converges to the sparsity of the sensed signal when the number of sensors is sufficiently large. In addition, we have observed that the required SNR decreases as the number of sensors increases for fixed PSM, which will be discussed in details [4]. For further work, it would be interesting to mathematically express the required PSM for a fixed $S$ and design a possible recovery algorithm approaching the analysis results of this paper. Furthermore, we aim to consider other models such that each sensor may have different support sets.


## Acknowledgment

This work was supported by the National Research Foundation of KOREA (NRF) grant funded by the Korea Government (MEST) (N0.2010-0017944).


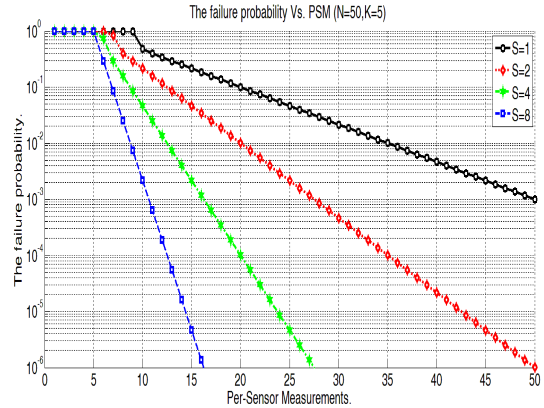


**Figure 1. Parameters are as follows. The signal length is 50; the number of nonzero elements in each signal is 5; *S* denotes the number of sensors; the noise variance is 0.0001. PSM decreases as the number of sensors increases.**

## References


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

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

[3] Thomas M. Cover and Joy A. Thomas, *Elemens of Information Theory, Second Edition*, Wiley Interscience, 2006.

[4] Sangjun Park, Junho Lee and Heug-no Lee, "Typical Analysis for Compressive Sensing Based Multi Sensor Systems", In preparation