Distributed Compressive Sensing: Performance Analysis with Diverse Signal Ensembles

Sung-Hsien Hsieh*, Wei-Jie Liang†, Chun-Shien Lu‡ and Soo-Chang Pei§
*Inst. of Info. Science, Academia Sinica & Graduate Inst. Comm. Eng., NTU, ROC, Email: parvaty316@iis.sinica.edu.tw
†Inst. of Info. Science, Academia Sinica & Department of Mathematics, NCKU, ROC, Email: l18991037@mail.ncku.edu.tw
‡Inst. of Info. Science, Academia Sinica, Taiwan, ROC, Email: lcs@iis.sinica.edu.tw
§Graduate Inst. Comm. Eng., NTU, ROC, Email: pei@cc.ee.ntu.edu.tw

***Abstract*—Distributed compressive sensing is a framework considering jointly sparsity within signal ensembles along with multiple measurement vectors (MMVs). The current theoretical bound of performance for MMVs, however, is derived to be the same with that for single MV (SMV) because the characteristics of signal ensembles are ignored.**

In this work, we introduce a new factor called "Euclidean distances between signals" for the performance analysis of a deterministic signal model under MMVs framework. We show that, by taking the size of signal ensembles into consideration, MMVs indeed exhibit better performance than SMV. Although our concept can be broadly applied to CS algorithms with MMVs, the case study conducted on a well-known greedy solver, called simultaneous orthogonal matching pursuit (SOMP), will be explored in this paper. We show that the performance of SOMP, when incorporated with our concept by modifying the steps of support detection and signal estimations, will be improved remarkably, especially when the Euclidean distances between signals are short. The performance of modified SOMP is verified to meet our theoretical prediction.

I. INTRODUCTION

A. Background & Related Works

Compressive sensing (CS) [1], [2] of sparse signals in achieving data acquisition and compression simultaneously has been extensively studied. Conventionally, given a measurement vector, CS shows that a sparse signal can be reconstructed via different solvers such as ℓ_1-minimization [3], [4] or greedy approaches [5], [6]. To further reduce the number of measurements, distributed compressive sensing (DCS) [7], [8] is a framework considering jointly sparsity within signal ensembles along with multiple measurement vectors (MMVs).

The model of MMVs is described as follows. Let $X = [x^1, x^2, ..., x^L] \in \mathbb{R}^{N\times L}$ be the signal ensembles, where $L \geq 1$ is the size of signal ensembles, and let $\Phi^i \in \mathbb{R}^{M\times N}$ for $1 \leq i \leq L$ be a sensing matrix. X is called jointly K-sparse if $\left|\bigcup_{i=1}^{L} \text{supp}\left(x^i\right)\right| = K$, where supp $\left(x^i\right)$ returns a support set of x^i and $|\cdot|$ is the cardinality function. Then, signal sampling is conducted via:

$$y^i = \Phi^i x^i. \tag{1}$$

Another common formulation assumes $\Phi = \Phi^1 = \dots \Phi^L$. Therefore, let $Y = [y^1, y^2, ..., y^L]$, we have:

$$Y = \Phi X. \tag{2}$$

When Φ^i's for all i's are drawn from i.i.d Gaussian random variables, a difference between the above two formulations is that $\text{rank}(Y) = \min(L, M)$ in Eq. (1) but $\text{rank}(Y) = \min(\text{rank}(X), M)$ in Eq. (2) with probability one.

Conventionally, the performance analysis can be classified into two categories: stochastic signal model or deterministic signal model. Stochastic signal model often assumes that either Φ or x is random so that the theoretical result involves probability. In contrast, deterministic signal model does not require this assumption and is unrelated to probability.

Furthermore, in the stochastic signal model, DCS [8] shows the fundamental bounds on the number of noiseless measurements such that signals can be jointly recovered based on Eq. (1). In addition, DCS shows that supports can be detected correctly with probability one when $L \to \infty$. In other words, DCS cannot accurately characterize the relationship between L and the performances of solvers such as SOMP [9]. For Eq. (2), [10] focuses on the average case analysis based on random matrix. [11] studies signals with a sparse fusion frame representation and provides a probabilistic analysis. [12] uses conic geometric for the performance analysis. Basically, [8]∼[12] show that the performance is proportional to L.

In the deterministic model based on Eq. (2), most of works [13], [14], [15] show that the performance is proportional to rank(Y) with noiseless measurements. In fact, if the performance analysis in MMVs does not consider rank(Y) as a factor, the analysis will be same as that in SMV. For example, [16] shows the performance of SOMP that is irrelevant to L, which is almost the same with OMP (a special case of SOMP with $L = 1$) [17]. This is because multiple sensors sense the same source with $x^1 = x^2 = \ldots = x^L$ such that Eq. (2) is degraded into SMV due to rank$(Y) = 1$. In addition, whatever the signal ensembles are, $\text{rank}(Y) = \min(L, M)$ based on Eq. (1) always holds. However, signal ensembles actually affect the performance in Eq. (1). For example, when $x^1 = x^2 = \ldots = x^L$, the performance based on Eq. (1) still is improved [18] since Eq. (1) can be reformulated as the following SMV formulation :

$$\hat{y} = A\hat{x}, \tag{3}$$

where $\hat{y} = \left[y^1; y^2; \ldots; y^L\right] \in \mathbb{R}^{ML}$, $A = \left[\Phi^1; \Phi^2; \ldots; \Phi^L\right] \in \mathbb{R}^{ML\times N}$, and $\hat{x} = x^1 = \ldots = x^L$. Clearly, the performance of Eq. (3) will be better along with the increase of L since the number of measurements is ML.

ISBN 978-0-9928626-7-1 © EURASIP 2017

B. Motivation

The discussions so far motivate us to consider a question: how to characterize the performance of DCS based on Eq. (1), especially when signals' strengths are different but close to each other. Since rank no longer characterizes the performance accurately in the DCS framework, here, we consider "Euclidean distances between signals are nonzero". It is noted that the Euclidean distances considered in this paper include two parts: one is $\|x^i - x^*\|_2$ for all i's with $x^* = \frac{1}{L}\sum_{i=1}^{L} x^i$ and another one is $\|x^i\|_2$ for all i's. Specifically, if the assumption that either $\|x^i - x^*\|_2$ is small for any i or $\|x^i\|_2$ approaches to $\|x^j\|_2$ holds for all $i \neq j$ exists, the performance will be good based on Eq. (1). Such an assumption is practical and occurs in cooperative spectrum sensing [19], where MMVs are obtained from different sensors to observe a single signal source (spectrum). Under the circumstance, when the sensors are too close to each other, the observed signal spectra also are similar, implying that distances between signals are short.

C. Contributions

In this paper, we are interested in the performance analysis of deterministic MMVs model in Eq. (1). Compared with previous works using rank, the novel factor with "Euclidean distances between signals" will reveal that the performance is related to the size L of signal ensembles. We take SOMP as a case study here even though our concept can be generally applied to other greedy algorithms. More specifically, we present a new mechanism for support detection and derive the sufficient condition of correct support detection. We show that when the Euclidean distance between signals are short or the signals have the same sign, the new mechanism outperforms conventional method remarkably. In terms of signal estimation, individual sparse signal is conventionally estimated by its corresponding measurement vector. In our work, however, we explore a strategy of estimating an individual signal from all measurement vectors and show that this strategy is potential to make support detection possible even when $M < K \leq ML$.

II. Preliminaries

For a matrix H, we denote its transpose by H^T and its pseudo inverse matrix by $H^\dagger$. For a set V collecting indices, H_V is a submatrix formed by columns of H with indices belonging to V. $\mathcal{P}(V)$ is the power set of V. For a vector u, the i^{th} entry of u is $u[i]$. $u_V \in \mathbb{R}^{|V|}$ is a vector formed by entries of u with indices belonging to V. $\|\cdot\|_p$ denotes the ℓ_p-norm. $\text{sign}(u)$ extracts the sign of u. $|u|$ returns the absolute value of u. Denote $\Omega = \bigcup_{i=1}^{L} \text{supp}\left(x^i\right)$ as the ground truth of support set. $\mathcal{N}(0, \sigma^2)$ denotes a normal distribution with zero mean and variance σ^2.

III. Main Results

To induce the new factor "Euclidean distance between signals" into the theoretical performance analysis of MMVs and see how many advantages we can have, we take SOMP as a case study here (but keep in mind that our idea can be generally applied to other greedy algorithms). In the following procedure of SOMP, the steps of support detection and signal estimation contain the original ones ((a) and (c)) and the newly added one ((b) and (d)) as alternative of the former.

1) Initialization: $t = 1$, $S = \{\ \}$, and $r^{i,t} = y^i$ for $i = 1, \ldots, L$.
2) Support detection:
   (a) $I = \text{argmax}_i\, u[i]$ with $u = \sum_{j=1}^{L} \left|(\Phi^j)^T r^{j,t}\right|$
   (b) $I = \text{argmax}_i\, u[i]$ with $u = \left|\sum_{j=1}^{L} (\Phi^j)^T r^{j,t}\right|$.
3) Support update: $S = S \bigcup \{I\}$.
4) Signal estimation:
   (c) $\hat{x}^i = (\Phi_S^i)^\dagger y^i$ with $i = 1, \ldots, L$
   (d) $\hat{x}^i = (A_S)^\dagger \hat{y}$ with $i = 1, \ldots, L$, where $A_S = \left[\Phi_S^1; \Phi_S^2; \ldots; \Phi_S^L\right] \in \mathbb{R}^{ML \times K}$.
5) Residual update: $r^{i,t+1} = y^i - \Phi_S^i \hat{x}^i$ with $i = 1, \ldots, L$.
6) If $t = K$, stop and output $\bar{x}_S^i = (\Phi_S^i)^\dagger y^i$ with $i = 1, ..., L$; otherwise, $t = t + 1$ and go to Step 2.

In the above procedure, SOMP-(a+c) denotes the traditional SOMP by choosing (a) as support detection and (c) as signal estimation. In contrast, steps (b) and (d) are proposed as alternative of steps (a) and (c) to accommodate for the conditions that the Euclidean distances between signals are short or the signals have the same sign, as mentioned in Sec. I-B.

We first explain why we present (b) as an alternative of (a) in certain situations. In the first iteration of steps (a) and (b), we expect that $u[j]$ for $j \in \Omega$ is large enough to make support detection correct. We derive the lower bounds of $u[j]$ for $j \in \Omega$ in steps (a) and (b), respectively, as follows:

$$
\begin{aligned}
(a): u[j] &= \sum_{i=1}^{L} \left|(\Phi_j^i)^T r^{i,1}\right| = \sum_{i=1}^{L} \left|x^i[j] + \left((\Phi_j^i)^T \Phi_\Omega^i - 1\right) x_\Omega^i\right| \\
&\geq \sum_{i=1}^{L} \left|x^i[j]\right| - \sum_{i=1}^{L} \left|\left((\Phi_j^i)^T \Phi_\Omega^i - 1\right) x_\Omega^i\right| \\
(b): u[j] &= \left|\sum_{i=1}^{L} (\Phi_j^i)^T r_1^i\right| \\
&= \left|\sum_{i=1}^{L} x^i[j] + \sum_{i=1}^{L} \left((\Phi_j^i)^T \Phi_\Omega^i - 1\right) x_\Omega^i\right| \\
&\geq \left|\sum_{i=1}^{L} x^i[j]\right| - \left|\sum_{i=1}^{L} \left((\Phi_j^i)^T \Phi_\Omega^i - i\right) x_\Omega^i\right|.
\end{aligned}
$$

If $\text{sign}(x^1) = \ldots = \text{sign}(x^L)$, we have $\sum_{i=1}^{L} \left|x^i[j]\right| = \left|\sum_{i=1}^{L} x^i[j]\right|$ and $\sum_{i=1}^{L} \left|\left((\Phi_j^i)^T \Phi_\Omega^i - 1\right) x^i\right| \geq \left|\sum_{i=1}^{L} \left((\Phi_j^i)^T \Phi_\Omega^i - I\right) x^i\right|$. It is easy to derive that (b) achieves more accurate support detection than (a) under the case that all signals have the same sign. We will further integrate this assumption into our performances analysis later.

Second, we discuss why we present (d) as an alternative of (c). In this paper, steps (b+d) is equivalent to solving Eq. (3) when $x^1 = x^2 \ldots = x^L$. Compared with (c), (d) is potential to make support detection possible when $M < K \leq ML$ since the number of measurements in Eq. (3) is ML. In other words, no matter what L is, there are infinite solutions to the least

ISBN 978-0-9928626-7-1 © EURASIP 2017

square problem with $M < K$ and, thus, (c) fails to estimate the signal correctly. On the other hand, when $\|x^i - x^j\| \leq \epsilon$ for all $i \neq j$ with small ϵ, (b+d) is no longer formulated as SMV in Eq. (3). On the contrary, we show that SOMP-(b+d) conducted with Eq. (1) still works when $M < K \leq ML$.

To begin with the performance analyses of SOMP-(a+c), SOMP-(b+c), and SOMP-(b+d), we first introduce restricted isometric property (RIP) as follows.

Lemma 1. ***(RIP)***
Let $\Phi \in \mathbb{R}^{M\times N}$. Suppose that there exists a constant $\delta_{|I|}(\Phi) < 1$ such that for any $x \in \mathbb{R}^{|I|}$ and any $I \subset \Omega$,

$$(1-\delta_{|I|}(\Phi))\|x\|_2^2 \leq \|\Phi_I x\|_2^2 \leq (1+\delta_{|I|}(\Phi))\|x\|_2^2 \tag{4}$$

holds. The matrix Φ is said to satisfy the $|I|$-restricted isometry property with restricted isometry constant (RIC) $\delta_{|I|}(\Phi)$.

When Φ satisfies RIP, the following consequence always holds.

Lemma 2. ***(Consequence of RIP** [20])*
Given a matrix Φ, for $I \subset \Omega$, if $\delta_{|I|}(\Phi) < 1$, then, for any $x \in \mathbb{R}^{|I|}$, we have

$$(1-\delta_{|I|}(\Phi))\|x\|_2 \leq \|\Phi_I^T \Phi_I x\|_2 \leq (1+\delta_{|I|}(\Phi))\|x\|_2. \tag{5}$$

According to RIC, OMP recovers all K-sparse vectors provided Φ satisfies the sufficient condition that $\delta_{K+1}(\Phi) < \frac{1}{\sqrt{K}+1}$ [17], [21]. Similarity, traditional SOMP-(a+c) with signal ensembles sensed via Eq. (2) needs to satisfy $\delta_{K+1}(\Phi) < \frac{1}{\sqrt{K}+1}$ or $\delta_K(\Phi) < \frac{\sqrt{K-1}}{\sqrt{K-1}+\sqrt{K}}$ [16]. Nevertheless, as mentioned in Sec. I-A, the sufficient condition never contains L due to no assumption about signal ensembles was made. In addition, it should be noted that this sufficient condition [16] cannot be applied to SOMP-(a+c) with Eq. (1). On the other hand, DCS focuses on SOMP-(a+c) with Eq. (1) [7], [8] but it does not prove such a sufficient condition. Thus, in addition to conducting analyses for SOMP-(b+c) and SOMP-(b+d), we also provide analysis for SOMP-(a+c).

To induce L into the sufficient condition of SOMP with signal ensembles being sensed via Eq. (1), our main results are summarized as the following three theorems. Due to limited space, we only show the proof of Theorem 2 here. Please refer to our complete version in arXiv[1] for other proofs.

Theorem 1. *Suppose $x^i \in \mathbb{R}^N$ is a $K-$sparse signal sensed via Eq. (1) for $i = 1,\ldots,L$ and Φ^i's satisfy RIP. Then, the SOMP-(a+c) algorithm will perfectly reconstruct x^i's if*

$$\sum_{i=1}^{L} \frac{\epsilon_1 \delta_{K+1}^2(\Phi^i) - (\sqrt{K}+2\epsilon_1)\delta_{K+1}(\Phi^i) + \epsilon_1}{1-\delta_{K+1}(\Phi^i)} > 0, \tag{6}$$

where $\epsilon_1 = \max_{U\in\mathcal{P}(\Omega)\setminus\emptyset} \frac{\min_j \|x_U^j\|_2}{\max_j \|x_U^j\|_2}$.

Theorem 2. *Let $A = \frac{1}{\sqrt{L}}[\Phi^1;\Phi^2;\ldots;\Phi^L]$ and let $\delta_K^{max} = \max_i \delta_K(\Phi^i)$. Suppose $x^i \in \mathbb{R}^N$ is a $K-$sparse signal sensed via Eq. (1) and Φ^i satisfies RIP for $i = 1,\ldots,L$. Then, the SOMP-(b+c) algorithm will perfectly reconstruct x^i's if*

$$\left(\sqrt{K}+1\right)\delta_{K+1}(A) + \left(1+\left(\sqrt{K}+1\right)(L\epsilon_2+\eta)\right)\delta_{K+1}^{max} < 1, \tag{7}$$

where $\epsilon_2 = \max_{U\in\mathcal{P}(\Omega)\setminus\emptyset} \frac{\sum_{i=1}^L \|x_U^i - x_U^*\|_2}{L\|x_U^*\|_2}$, $x^* = \frac{1}{L}\sum_{i=1}^L x^i$, and $\eta = \delta_{K+1}^{max} - \delta_{K+1}(A)$.

Proof. Please see Appendix in Sec. VI for detailed proof. ☐

Theorem 3. *Let $A = \frac{1}{\sqrt{L}}[\Phi^1;\Phi^2;\ldots;\Phi^L]$ and let $\delta_K^{max} = \max_i \delta_K(\Phi^i)$. Suppose $x^i \in \mathbb{R}^N$ is a $K-$sparse signal sensed via Eq. (1) for $i = 1,\ldots,L$, and A satisfies RIP with $K \leq M$. Then, the SOMP-(b+d) algorithm will perfectly reconstruct x^i's with $i = 1,\ldots,L$ if*

$$\sqrt{K}(1+L^2\epsilon_3)\delta_{K+1}(A) + (1+L\epsilon_3)\delta_{K+1}^{max} < 1, \tag{8}$$

where $\epsilon_3 = \max_{U\in\mathcal{P}(\Omega)\setminus\emptyset} \frac{\sum_{i=1}^L \|x^i - x^*\|_2}{L\|x_U^*\|_2}$ with $x^* = \frac{1}{L}\sum_{i=1}^L x^i$.

In the above three theorems, ϵ_1, ϵ_2, and ϵ_3 describe the characteristics of involved signal ensembles, respectively. First, among them, Theorem 1 shows that when the entries of x^i's have the same energy (unrelated to sign(x^i)'s), we have $\epsilon_1 = 1$ and SOMP-(a+c) performs best. On the other hand, the analysis is derived for SOMP-(a+c) without considering signal ensembles as follows.

Corollary 1. *Let $\delta_K^{max} = \max_i \delta_K(\Phi^i)$. Other assumptions follow Theorem 1. Then, the SOMP-(a+c) algorithm will perfectly reconstruct x^i's with $i = 1,\ldots,L$ if*

$$\delta_{K+1}^{max} < \frac{1}{\sqrt{K}+2}.$$

In comparison with Theorem 1, the result of Corollary 1 even is worse than SMV since δ_{K+1}^{max} is the maximum among $\delta_K(\Phi^i)$'s and is increased with $L > 1$. However, when $\epsilon_1 = 1$, $\delta_{K+1}^{max} < \frac{1}{\sqrt{K}+2}$ is one of solutions to satisfy (6) in Theorem 1. In fact, Theorem 1 requires that the mean of $\delta_{K+1}(\Phi^i)$'s instead of δ_{K+1}^{max} is small.

Second, as shown in Theorem 2, ϵ_2 indicates that x^i's should be distributed around the center x^*, which should be far away from the origin. In other words, x^i's have the same sign to maximize the denominator of ϵ_2. To fairly compare Theorem 1 and Theorem 2, we need to build the relationship between $\delta_K(A)$ and $\delta_K(\Phi^i)$. In fact, $\sqrt{L}\delta_K(A) \sim \delta_{K+1}^{max}$. In addition, a random matrix is known to satisfy $\delta_{cK} < \theta$ with high probability provided one chooses $M = O(\frac{cK}{\theta^2}\log\frac{N}{K})$ [22]. Then, it is trivial to check that when $\epsilon_1 = 1$, $\epsilon_2 = 0$ (the best case for both theorems), and $L = K$, the number of measurements required in Theorem 1 is about $O(K)$ larger than that in Theorem 2.

Finally, we note that the desired signal ensembles for both the cases of ϵ_2 and ϵ_3 are the same. Since the numerator in ϵ_3 is fixed, it implies that $\epsilon_2 \leq \epsilon_3$. However, it should be noted

[1] http://arxiv.org/abs/1609.01899

ISBN 978-0-9928626-7-1 © EURASIP 2017

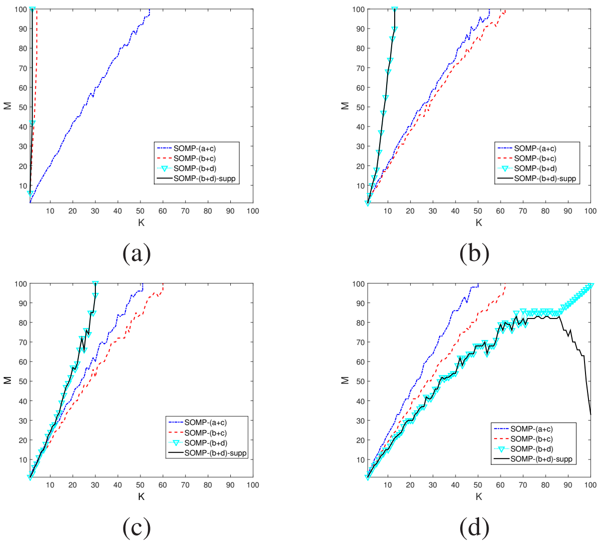

Fig. 1. Performance analysis for different types of signals: (a) Type I; (b) Type II; (c) Type III; (d) Type IV, under $L = 3$ and $N = 100$. The curve denotes the phase transition of probability of success achieving 50%. The region above the curve means the probability $\geq 50\%$.

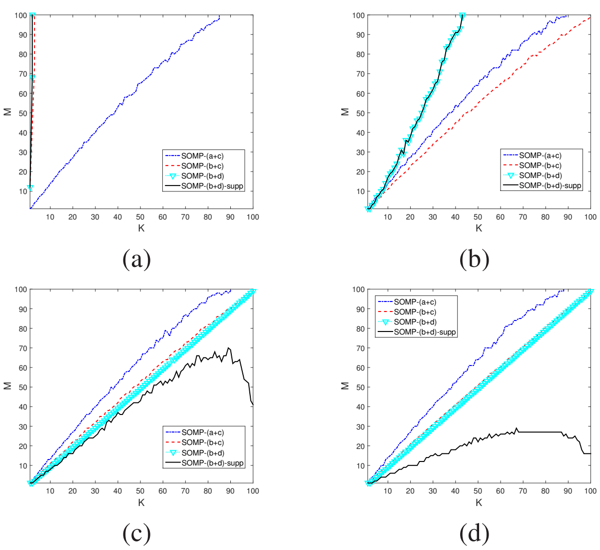

Fig. 2. Performance analysis for different types of signals: (a) Type I; (b) Type II; (c) Type III; (d) Type IV, under $L = 9$ and $N = 100$.

that, when $\epsilon_2 = \epsilon_3 = 0$, the sufficient condition in Theorem 3, compared with that in Theorem 2, is slightly relaxed. In addition, the assumption in Theorem 3 only requires that A, instead of all Φ^i's, satisfies RIP and that $K \leq M$. Thus, even though individual Φ^i does not satisfy RIP, perfect reconstruction is still possible. The following corollary shows that $K \leq M$ can be further removed for perfect support detection.

Corrollary 2. *Suppose A satisfies RIP. Then, the SOMP-(b+d) algorithm perfectly detects the support set of x^i's for $i = 1, \ldots, L$ with the same sufficient condition in Theorem 3.*

IV. Experiments

In this section, we validate our three theorems from empirical simulations. We first randomly pick a support set Ω with $|\Omega| = K$. Then, four types of signal ensembles are generated as follows. For all $j \in \Omega$, we have

I. $x^i[j] \sim \mathcal{N}(0, 1)$ with $i = 1, \ldots, L$.
II. $x^i[j] \sim |\mathcal{N}(0, 1)|$ with $i = 1, \ldots, L$.
III. $x^i[j] \sim \mathcal{N}(1, 0.25)$ with $i = 1, \ldots, L$.
IV. $x^i[j] = 1$ with $i = 1, \ldots, L$.

For $j \notin \Omega$, $x^i[j] = 0$. Then, Φ^i's are standard normal matrices. We run SOMP-(a+c), SOMP-(b+c), and SOMP-(b+d) to obtain $\bar{x}^i$'s and declare success if $\sum_{i=1}^{L} \|\bar{x}^i - x^i\| \leq 10^{-5}$. The successful probability is the number of successes divided by 100.

These types of signals present different values of ϵ_1, ϵ_2 and ϵ_3. For example, ϵ_2 and ϵ_3 are gradually decreased from Type I to Type IV. In addition, ϵ_1's in Types I and II are the same but are smaller than those in Types III and IV.

The results for Types I, II, III, and IV are shown in Figs. 1 and 2(a)-(d) with $L = 3$ and $L = 9$, respectively. It should be noted that SOMP-(b+d)-supp considers the success of "support detection" instead of signal reconstruction in SOMP-(b+d). Thus, by Corollary 2, success may happen even when $K > M$.

It is also observed from Figs. 1(a)-(b) that the curve of SOMP-(b+d) overlaps with that of SOMP-(b+d)-supp. This is because correct support detection implies perfect reconstruction for $K \leq M$. In addition, it is surprising to see from Figs. 1(c)-(d) that SOMP-(b+d)-supp exhibits higher probability of success when K approaches N. This may be due to that since the number $\binom{N}{K}$ of candidate support sets approaches 1.

SOMP-(a+c) maintains the performance from Fig. 1(a)-(d). In addition, SOMP-(b+c) outperforms SOMP-(a+c) remarkably when signals have the same sign, as shown in from Figs. 2(b)-(d). Compared with SOMP-(b+c), the assumption in SOMP-(b+d) is more sensitive to Euclidean distances between signals, implying large ϵ_3, such that its performance is worse than SOMP-(b+c). However, in terms of support detection, SOMP-(b+d) has potential to lower the number of measurements when $M < K$. In addition, when $\epsilon_2 = \epsilon_3 = 0$, SOMP-(b+d) outperforms SOMP-(b+c). Fig. 2 reaches the same conclusions with Fig. 1 but exhibits higher successful probability under $L = 9$.

In summary, SOMP-(a+c) has the weakest assumption about signal ensembles such that it can be applied to all different types of signal ensembles. Even so, for Types II-IV, its performance is not the best among the methods used for comparisons. In fact, when ϵ_1, ϵ_2, and ϵ_3 are the best choices such that the sufficient conditions are easy to satisfy in Theorems 1$\sim$3, respectively, the sufficient condition for Theorem 1 is relatively not easy to satisfy.

V. Conclusion

In this paper, we have introduced the novel concept of "Euclidean distances between signals" for the theoretical performance analysis of MMVs. Our results based on RIP clearly illustrate how the size of signal ensembles affects the performance of SOMP. In contrast, previous works based on RIP are unrelated to the size of signal ensembles and are degenerated to SMV. Finally, we have demonstrated that the performance of CS algorithms under MMVs model can be remarkably improved especially when the Euclidean distances between signals are short.

VI. Appendix

Lemma 3. *[23] Let I_1, $I_2 \subset \Omega$ be two disjoint sets ($I_1 \cap I_2 = \emptyset$). If $\delta_{|I_1|+|I_2|} < 1$, then*

$$\|(\Phi_{I_1})^T \Phi_{I_2} x\|_2 \le \delta_{|I_1|+|I_2|}(\Phi)\|x\|_2$$

holds for any x.

Proof of Theorem 2:

Proof. Here, I_1 and I_2 in Lemma 3 denote the chosen index at t-th iteration and gound truth r Ω, respectively. By the contrapositive of statement in Lemma 3, the chosen index and Ω are not disjoint; *i.e.*, support detection is correct.

Let $\Phi_\Omega^i = [\Phi_S^i \mid \Phi_U^i]$, where S denotes the support set that has been solved and U denotes the support set that has not solved yet. Let $r^{i,t}$ be the residual and let I be the chosen index at t-th iteration. For simplicity, let r^i and δ^i denote $r^{i,t}$ and $\delta(\Phi^i)$, respectively. When $I \notin \Omega$, we first derive the upper bound of $\left\|\sum_i (\Phi_I^i)^T r^i\right\|$ by Lemma 3 as:

$$\begin{aligned}&\sqrt{K}\left\|\sum_i (\Phi_I^i)^T r^i\right\| = \left\|\sum_i (\Phi_U^i)^T (I - \Phi_S^i(\Phi_S^i)^\dagger)\Phi_U^i x_U^i\right\| \\ &\ge \left\|\sum_i (\Phi_U^i)^T \Phi_U^i x_U^i\right\| - \left\|\sum_i (\Phi_U^i)^T \Phi_S^i(\Phi_S^i)^\dagger \Phi_U^i x_U^i\right\| \\ &= (I) + (II).\end{aligned}$$

Let $x_U^i = x_U^* - c^i$, where $c^i \in \mathbb{R}^{|U|}$ is any constant and $x_U^* = \dfrac{\sum_i x_U^i}{L}$, and $\epsilon_2 = \max\limits_{U \in \mathcal{P}(\Omega)\setminus\emptyset} \dfrac{\sum_i \|x_U^i - x_U^*\|}{L\|x_U^*\|}$. Then,

$$\begin{aligned}(I) \quad &\ge \quad L\|x^*\| - \left\|\sum_i \left[(\Phi_U^i)^T \Phi_U^i - I\right] x_U^*\right\| - \\ &\quad\quad \left\|\sum_i \left[(\Phi_U^i)^T \Phi_U^i - I\right] c^i\right\| \\ &\ge \quad L\|x_U^*\|\left(1 - \delta_{K+1}(A) - L\epsilon_2\delta_{K+1}^{max}\right). \\ (II) \quad &\le \quad L\frac{(\delta_{K+1}^{max})^2}{1-\delta_{K+1}^{max}}(1 + L\epsilon_2)\|x_U^*\|.\end{aligned}$$

Therefore, the lower bound of $\left\|\sum_i (\Phi_I^i)^T r^i\right\|$ is

$$\frac{L\|x^*\|}{\sqrt{K}}\left[1 - \delta_{K+1}(A) - L\epsilon_2\delta_{K+1}^{max} - \frac{(\delta_{K+1}^{max})^2}{1-\delta_{K+1}^{max}}(1 + L\epsilon_2)\right]. \quad (9)$$

By the similar technique, the upper bound is obtained by:

$$L\|x_U^*\|\left[\delta_{K+1}(A) + L\epsilon_2\delta_{K+1}^{max} + \frac{(\delta_{K+1}^{max})^2}{1-\delta_{K+1}^{max}}(1 + L\epsilon_2)\right] \quad (10)$$

Hence, the SOMP-(b+c) algorithm will choose correct support if (9) $>$ (10), which implies

$$(\sqrt{K}+1)\delta_{K+1}(A) + \left(1 + \left(\sqrt{K}+1\right)(L\epsilon_2 + \eta)\right)\delta_{K+1}^{max} < 1,$$

with $\eta = \delta_{K+1}^{max} - \delta_{K+1}(A)$. When all support are found correctly, the SOMP-(b+c) algorithm will perfectly reconstruct x^i's. ☐

Acknowledgment

This work was supported by MOST 104-2221-E-001-030-MY3 and MOST 104-2221-E-001-019-MY3 (Taiwan, ROC).

References

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[3] E. H. Candes, J. Romberg, and T. Taio, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Information Theory*, vol. 52, no. 2, pp. 489–507, 2006.

[4] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal of Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[5] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[6] D. Needella and J. A. Troppb, "Cosamp: Iterative signal recovery from incomplete and inaccurate samplesstar," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[7] D. Baron, M. Wakin, M. Duarte, S. Sarvotham, and R. Baraniuk, "Distributed compressed sensing," *Preprint*, 2005.

[8] M. F. Duarte, M. B. Wakin, D. Baron, S. Sarvotham, and R. G. Baraniuk, "Measurement bounds for sparse signal ensembles via graphical models," *IEEE Trans. on Information Theory*, vol. 59, no. 7, pp. 4280–4289, 2013.

[9] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. part i: Greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.

[10] Y. C. Eldar and H. Rauhut, "Average case analysis of multichannel sparse recovery using convex relaxation," *IEEE Trans. on Information Theory*, vol. 56, no. 1, pp. 505–519, 2010.

[11] P. Boufounos, G. Kutyniok, and H. Rauhut, "Sparse recovery from combined fusion frame measurements," *IEEE Trans. on Information Theory*, vol. 57, no. 6, pp. 3864–3876, 2011.

[12] S.-W. Hu, G.-X. Lin, S.-H. Hsieh, and C.-S. Lu, "Phase transition of joint-sparse recovery from multiple measurements via convex optimization," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2015, pp. 3576–3580.

[13] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. on Signal Processing*, vol. 54, no. 12, pp. 4634–4643, 2006.

[14] M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," *IEEE Trans. on Information Theory*, vol. 58, no. 2, pp. 1135–1146, 2012.

[15] Y. Jin and B. D. Rao, "Support recovery of sparse signals in the presence of multiple measurement vectors," *IEEE Trans. on Information Theory*, vol. 59, no. 5, pp. 3139–3157, 2013.

[16] J. F. Determe, J. Louveaux, L. Jacques, and F. Horlin, "On the exact recovery condition of simultaneous orthogonal matching pursuit," *IEEE Signal Processing Letters*, vol. 23, no. 1, pp. 164–168, 2016.

[17] Q. Mo and Y. Shen, "A remark on the restricted isometry property in orthogonal matching pursuit," *IEEE Trans. on Information Theory*, vol. 58, no. 6, pp. 3654–3656, 2012.

[18] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp, "Living on the edge: phase transitions in convex programs with random data," *Information and Inference*, vol. 3, no. 3, pp. 224–294, 2014.

[19] P. Paysarvi-Hoseini and N. C. Beaulieu, "Optimal wideband spectrum sensing framework for cognitive radio systems," *IEEE Trans. on Signal Processing*, vol. 59, no. 3, pp. 1170–1182, 2011.

[20] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[21] J. Wang and B. Shim, "On the recovery limit of sparse signals using orthogonal matching pursuit," *IEEE Trans. on Signal Processing*, vol. 60, no. 9, pp. 4973–4976, 2012.

[22] P. Jain, A. Tewari, and I. S. Dhillon, "Orthogonal matching pursuit with replacement," in *Neural Information Processing Systems*, 2011.

[23] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589 – 592, 2008.

ISBN 978-0-9928626-7-1 © EURASIP 2017