On the Error Probability of Support Recovery for Orthogonal Matching Pursuit With a Random Measurement Matrix

YONGGU LEE[ID]**[1], (Member, IEEE), JINHO CHOI**[ID]**[2], (Senior Member, IEEE), AND EUISEOK HWANG**[ID]**[3], (Member, IEEE)**

[1]Research Institute for Solar and Sustainable Energies, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea
[2]School of Information Technology, Deakin University, Geelong, VIC 3220, Australia
[3]School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea

Corresponding author: Euiseok Hwang (euiseokh@gist.ac.kr)

ABSTRACT In this paper, an asymptotic bound on the recovery error probability of a sparse signal is derived for the orthogonal matching pursuit algorithm. The proposed bound is based on the support recovery analysis with a random measurement matrix, which gets closer to the empirical bound tightly in a large system and high signal-to-noise ratio regime. During recovery, all signal associated parameters introduced in the existing analysis are considered together. Furthermore, the necessary conditions for the conventional bound derivation such as the minimum value limit of non-zero coefficients in the sparse signal can be relaxed in our proposed approach. Through numerical evaluations, our theoretical performance bounds are demonstrated to be close to the simulated results, notably closer than those obtained previously.

INDEX TERMS Compressive sensing, orthogonal matching pursuit, support recovery.

I. INTRODUCTION

In general, signals are sampled at least at a Nyquist sampling rate, corresponding to twice the maximum frequency of interest for signal reconstruction. However, it may be inefficient to sample at a Nyquist rate for a signal, when the signal can be sparsely represented in a transformed domain. For this reason, compressive sensing (CS) has received a considerable attention in that a sparse signal can be recovered from fewer samples than the Nyquist requirement. CS has been applied to various applications such as wireless communications, and image compressions [1], [2].

In most applications of CS, a noisy measurement can be expressed as

$$\mathbf{y} = \Phi \mathbf{s} + \mathbf{n}, \qquad (1)$$

where $\Phi \in \mathbb{R}^{L \times N}$ is a measurement matrix, and $\mathbf{s} \in \mathbb{R}^{N \times 1}$ is a Q-sparse vector (i.e., $\|\mathbf{s}\|_0 = Q$) with a support set denoted by $\Lambda = \{i | \mathbf{s}_i \neq 0\}$ whose indices indicate

The associate editor coordinating the review of this manuscript and approving it for publication was Nianqiang Li[ID].

the positions of non-zero coefficients of $\mathbf{s}$. Here, $\mathbf{s}_i$ stands for the i-th element of $\mathbf{s}$. The recovery problem in CS is underdetermined because there are fewer measurements than unknowns. However, the problem in CS can be solved by taking advantage of the sparsity of recovering signal. Suppose that $\mathbf{s}$ belongs to the set obtained by the union of all the $\binom{N}{Q} = \frac{N!}{Q!(N-Q)!}$ Q-dimensional subspaces in $\mathbb{R}^{N \times 1}$. In the recovery problem, although an l_0-optimization is good to seek the sparsest solution, it requires a combinatorial search. Alternatively, the recovery problem of finding the coefficients of Q basis vectors in the signal decomposition can be formulated as an l_1-optimization problem [3]. Besides, there are low-complexity greedy algorithms to recover a sparse signal such as orthogonal matching pursuit (OMP) and compressive sampling matching pursuit (CoSaMP) [4].

OMP is an iterative algorithm to find out the best matching projection with low-complexity, which is frequently used for sparse signal recovery in CS because of its efficiency and effectiveness. For the OMP, the best column index can be chosen by searching for the most closely correlated column in the measurement matrix with the current residual vector at

each step. At each step, the index of the selected column is included in the estimated support set and the residual vector is updated by projecting the observation onto the linear subspace spanned by the columns in the set. Detailed procedures are provided in Algorithm 1.

Algorithm 1 OMP Algorithm

Input: $\mathbf{y}$, Φ, and stopping criterion
Output: $\hat{\mathbf{s}} \in \mathbb{R}^{N \times 1}$, and $\hat{\Lambda}$

Notation : q is an iteration index. $\hat{\mathbf{s}}$ and $\hat{\Lambda}$ denote estimated sparse signal and estimated support set, respectively. In addition, $\mathbf{r}_{(q)}$ is the q-th residual vector.
Initialization : $q = 0$, $\hat{\Lambda}_{(0)} = \emptyset$, and $\mathbf{r}_{(0)} = \mathbf{y}$

1: **while** Stopping criterion (e.g., $q < Q$) **do**
2:     $q = q + 1$
3:     $\phi_{(q)} = \arg\max_{1 \geq j \geq N} |\langle \mathbf{r}_{(q-1)}, \Phi_j \rangle|$
4:     $\hat{\Lambda}_{(q)} = \hat{\Lambda}_{(q-1)} \cup \{\phi_{(q)}\}$
5:     $\hat{\mathbf{s}}_{(q)} = \arg\min_{\mathbf{z} \in \mathbb{E}^{|\hat{\Lambda}_{(q)}|}} \|\mathbf{y} - \Phi_{\hat{\Lambda}_{(q)}} \mathbf{z}\|$
6:     $\mathbf{r}_{(q)} = \mathbf{y} - \Phi_{\hat{\Lambda}_{(q)}} \hat{\mathbf{s}}_{(q)}$
7: **end while**
8: **return** $\hat{\mathbf{s}} = \hat{\mathbf{s}}_{(q)}$, and $\hat{\Lambda} = \hat{\Lambda}_{(q)}$

Despite the weaknesses such as poor reconstruction performance in a low signal-to-noise ratio (SNR), OMP has been widely used with other strengths owing to its simple procedures, such easy and fast implementability with low complexity. In particular, it is suitable for CS recovery in wireless communications and data compressions for Internet of things (IoT) because IoT devices have limited hardware resources (e.g., small memory, low battery) with low processing power [5], [6]. Thus, it is important to investigate theoretical performance predictions of OMP for applications, which is a primary motivation of this study.

The accuracy of the support recovery is one of the most important issues in CS. Once the support set is correctly reconstructed, (1) is well-posed and provides an accurate estimate of $\mathbf{s}$ using the least square approach [7]. Therefore, in some applications, the estimation of a sparse signal refers to the recovery of the support set. For example, distributed detection problem can be solved using CS and the estimation performance is directly related to the accuracy of support recovery [8]. Furthermore, statistical performance of support recovery with a random measurement matrix is required in some applications. For example, in CS encryption, a measurement matrix used as a secret key can be randomly changed to enhance the security (e.g., one-time pad) [9]–[11]. To evaluate the performance of these applications, a statistical recovery performance for ensemble should be analyzed.

Two features of a measurement matrix are used to analyze the support recovery of OMP. One is the restricted isometry property (RIP) and the other is the mutual coherence. For the RIP, the restricted isometry constant (RIC) of order Q can be defined as the smallest constant of δ_Q such that

$$(1 - \delta_Q)\|\mathbf{s}\|_2^2 \leq \|\Phi\mathbf{s}\|_2^2 \leq (1 + \delta_Q)\|\mathbf{s}\|_2^2. \quad (2)$$
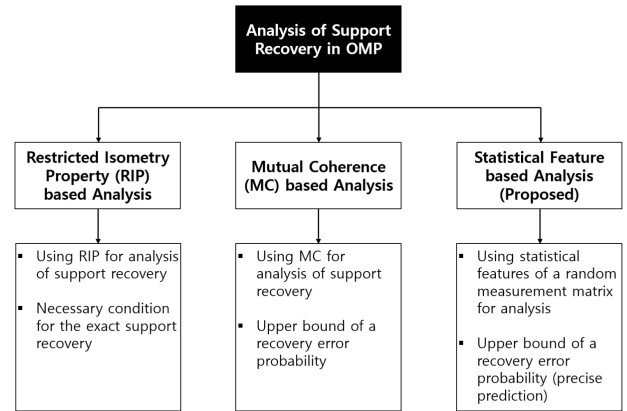

FIGURE 1. Comparison of analyses for support recovery in OMP.

There exist several studies using the RIP to see the performance of the support recovery of OMP in [12]–[20]. Most of these provide sufficient conditions for the exact support recovery of OMP theoretically. In [14], a sufficient condition of the RIC for the exact recovery in a noiseless case is given by $\delta_{Q+1} < \frac{1}{3\sqrt{Q}}$. In addition, in a noisy case, when $\delta_{Q+1} < \frac{1}{\sqrt{Q+1}}$, a sufficient condition of the minimum value of non-zero coefficients in $\mathbf{s}$ for the exact support recovery is provided as follows:

$$s_{min} \geq \frac{2}{1 - (\sqrt{Q} + 1)\delta_{Q+1}} \left(1 + \sqrt{\frac{(1 + \delta_2)Q}{1 - \delta_{Q+1}}}\right) \epsilon, \quad (3)$$

where $\|\Phi^{\mathsf{T}}\mathbf{n}\|_\infty \leq \epsilon$ in [20]. However, as shown in [21], determining whether or not a measurement matrix satisfies the RIP is NP-hard. In addition, most of the analyses are focused on derivations of sufficient conditions for the support recovery. On the other hand, mutual coherence-based analyses have been performed to derive the upper bounds on the error probability of the support recovery in OMP in [22], [23], where the mutual coherence can be defined by

$$\mu_{max} = \max_{1 \leq i \neq j \leq N} |\langle \Phi_i, \Phi_j \rangle|. \quad (4)$$

However, they have a non-negligible gap between the derived theoretical bounds and empirical results. Besides, the information-theoretic limits of the sparsity recovery problem are investigated in [24].

In this paper, we derive an asymptotic upper bound on the recovery error probability in the high SNR regime and large system for a random ensemble in OMP. The proposed analysis is not affected by a single realization of a measurement matrix; however, it depends on statistical features of the random measurement matrix. We are motivated to provide an acceptable performance guarantee for identifying a support set of a sparse signal. To this end, our analysis is based on the statistical characteristics of a random measurement matrix in CS. The derivation can be used for some applications such as CS encryption [9]–[11] and wireless communications [25]–[27], which can directly exploit sparsity with an acceptable prediction of the support recovery.

Through numerical evaluations, it is demonstrated that the proposed theoretical analysis closely agrees with the empirically obtained results compared to the conventional mutual coherence-based analysis [23] in terms of the recovery error probability. Furthermore, it is shown that the proposed analysis is more valid in a large system and high-SNR regime.

Notation: Upper-case and lower-case boldface letters are used for matrices and vectors, respectively. $\mathbf{X}^T$ denotes the transpose of $\mathbf{X}$. For a matrix $\mathbf{X}$ (a vector $\mathbf{x}$), $\mathbf{X}_m$ ($\mathbf{x}_m$) and $\mathbf{X}_{n,m}$ represent the m-th row (element, resp.) and the m-th row and the n-th column element of $\mathbf{X}$, respectively. The p-norm of a vector $\mathbf{a}$ is denoted by $\|\mathbf{a}\|_p$ (If $p = 2$, the norm is denoted by $\|\mathbf{a}\|$ without the subscript). $\mathcal{N}(\boldsymbol{\mu}, \mathbf{R})$ and $U(a, b)$ represent the distribution of Gaussian random vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$ and the uniform distribution with an interval $[a, b]$, respectively. $\langle \mathbf{x}, \mathbf{y} \rangle$ denotes the inner product between $\mathbf{x}$ and $\mathbf{y}$.

II. SUPPORT RECOVERY ANALYSIS IN OMP

In this section, we first introduce an existing recovery analysis for a support set that utilizes the mutual coherence of a measurement matrix in OMP. Then, the proposed support recovery analysis is presented as the main contribution in this paper and compared with a conventional approach in [23].

A. CONVENTIONAL ANALYSIS OF SUPPORT RECOVERY

In previous studies [22], [23], the support recovery in OMP is analyzed with the mutual coherence of a measurement matrix, denoted by the maximum absolute cross-correlation between columns of a measurement matrix. Then, recent results on support recovery in [23] can be abstracted by the following theorem, providing good insights on coherence-based analysis.

Theorem 1 (Miandji et al. [23]): Consider a noisy CS measurement in (1) with $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$. Let $\lambda = \Pr\{|\langle \tilde{\boldsymbol{\Phi}}_j, \mathbf{n} \rangle| \leq \beta\}$, for some constant $\beta \geq 0$ and $j \in \{1, 2, \cdots, N\}$. If $s_{min} \geq \beta$, then an upper bound on the recovery error probability is given by

$$P_e < 1 - \lambda\left(1 - 2N\exp\left(\frac{-N(s_{min}/2-\beta)^2}{2Q^2\gamma^2+2N\gamma(s_{min}/2-\beta)/3}\right)\right), \tag{5}$$

where $\gamma = \mu_{max}s_{max}$. Here, s_{min} and s_{max} are the minimum and maximum values of non-zero coefficients in $\mathbf{s}$, respectively. Moreover, λ is lower bounded by

$$\lambda \geq 1 - N\sqrt{\frac{2}{\pi}}\frac{\sigma}{\beta}\exp\left(\frac{-\beta^2}{2\sigma^2}\right). \tag{6}$$

In Theorem 1, the mutual coherence, μ_{max}, is used to derive the correlation between a column of a measurement matrix and a noisy measurement (for the proof of this theorem, see [23] for details). Theorem 1 is an updated version of the support recovery analysis derived in [22] to take into account various parameters (L, N, Q, s_{min}, s_{max}, σ^2 and μ_{max}) and to improve the performance bound. However, it requires to meet a certain condition, $s_{min} \geq \beta$. In addition, there is a significant

difference between theoretical and empirical results, making it difficult to apply the analysis with (5) to actual applications. In particular, if a Gaussian measurement matrix having a relatively high mutual coherence (e.g., $\mathbb{E}[\mu_{max}] = 0.3899$ with $L = 128$ and $N = 256$) is used, the upper bound on the error probability in (5) approaches one even though the other parameters are set for favorable conditions. Motivated by these problems, we present a new analysis of support recovery in OMP, as follows.

B. PROPOSED ANALYSIS OF SUPPORT RECOVERY

In this subsection, we derive an asymptotic upper bound on the recovery error probability in high-SNR regime and large system with a random measurement matrix in OMP.

Theorem 2: Consider a noisy CS measurement in (1), where a measurement matrix, $\boldsymbol{\Phi}$ is a random matrix whose elements are independent and identically distributed random variables (e.g., Gaussian, Bernoulli) of zero mean and $1/L$ variance, $\mathbf{s}_{\Lambda(q)} \sim U(s_{min}, s_{max})$, and $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$. Here, $\Lambda(q)$ denotes an index of the q-th non-zero element in $\mathbf{s}$. Then, an asymptotic upper bound on the recovery error probability in the high-SNR regime and large system is given by

$$
\begin{aligned}
P_e \\
< \int_0^\infty \frac{\sqrt{2}(N-Q)}{\sqrt{\pi}\tilde{\sigma}_\kappa^2}\exp\left(-\frac{z^2}{2\tilde{\sigma}_\kappa^2}\right)\left[\text{erf}\left(\frac{z}{\tilde{\sigma}_\kappa\sqrt{2}}\right)\right]^{N-Q-1} \\
\times \left(1 - \left[1 - \frac{1}{2}\left(\text{erf}\left(\frac{z+\tilde{\mu}_\eta}{\tilde{\sigma}_\eta\sqrt{2}}\right)+\text{erf}\left(\frac{z-\tilde{\mu}_\eta}{\tilde{\sigma}_\eta\sqrt{2}}\right)\right)\right]^Q\right)dz,
\end{aligned}
\tag{7}
$$

where $\tilde{\mu}_\eta = s_{min}$, $\tilde{\sigma}_\eta^2 = \frac{s_{min}^2(Q-1)}{L}+\sigma^2$, $\tilde{\sigma}_\kappa^2 = \frac{Q}{L}((\frac{s_{max}+s_{min}}{2})^2+\frac{1}{12}(s_{max}-s_{min})^2)+\sigma^2$, and $\text{erf}(\cdot)$ denotes the error function.

Proof of Theorem 2: The following two lemmas will be used to prove Theorem 2 and the proofs of lemmas are presented at Appendices A and B, respectively.

Lemma 1: Let $\Gamma_{k,\Lambda} = \langle \tilde{\boldsymbol{\Phi}}_k, \boldsymbol{\Phi}_\Lambda\mathbf{s}_\Lambda + \mathbf{n} \rangle$ and $\Gamma_{j,\Lambda^\circ} = \langle \tilde{\boldsymbol{\Phi}}_j, \boldsymbol{\Phi}_{\Lambda^\circ}\underline{\mathbf{s}}_{\Lambda^\circ} + \mathbf{n} \rangle$, where $\tilde{\boldsymbol{\Phi}}_j = \frac{\boldsymbol{\Phi}_j}{\|\boldsymbol{\Phi}_j\|^2}$, $\tilde{\boldsymbol{\Phi}}_k = \frac{\boldsymbol{\Phi}_k}{\|\boldsymbol{\Phi}_k\|^2}$, $\underline{\mathbf{s}}_\Lambda = s_{min} \cdot \mathbf{1}$, and $\Lambda^\circ = \Lambda\setminus\{j\}$. Then, the distributions of $\Gamma_{k,\Lambda}$ and Γ_{j,Λ° become $\mathcal{N}(0, \tilde{\sigma}_\kappa^2)$ and $\mathcal{N}(0, \tilde{\sigma}_\eta^2)$, respectively, in a large system.

Lemma 2: $\Pr(\min_{j\in\Lambda}|\langle \tilde{\boldsymbol{\Phi}}_j, \boldsymbol{\Phi}_\Lambda\underline{\mathbf{s}}_\Lambda + \mathbf{n} \rangle| > \min_{j\in\Lambda}|\langle \tilde{\boldsymbol{\Phi}}_j, \boldsymbol{\Phi}_\Lambda\mathbf{s}_\Lambda + \mathbf{n} \rangle|)$ approaches zero in a large-system and high-SNR regime.

In OMP, a support set cannot be recovered exactly when

$$\left(\max_{j\in\Lambda}\left|\left\langle \tilde{\boldsymbol{\Phi}}_j, \mathbf{r}_{(q)} \right\rangle\right| < \max_{k\notin\Lambda}\left|\left\langle \tilde{\boldsymbol{\Phi}}_k, \mathbf{r}_{(q)} \right\rangle\right|\right), \tag{8}$$

where $\mathbf{r}_{(q)}$ is the residual vector in the q-th iteration and $1 \leq k \leq N$ hereinafter. Here, $\mathbf{r}_{(1)} = \mathbf{y}$ and $\mathbf{r}_{(q)} = \mathbf{y} - \hat{\boldsymbol{\Phi}}_{(q-1)}\hat{\mathbf{s}}_{(q-1)}$, where $\hat{\boldsymbol{\Phi}}_{(q-1)}$ and $\hat{\mathbf{s}}_{(q-1)}$ are the measurement matrix with the estimated support set and the estimated sparse vector in the $(q-1)$-th iteration, respectively. However, it is difficult to derive the probability of an event due to the recursively computed residual vector. Alternatively, in [22], [23],

a recovery error event can be approximated as

$$\min_{j \in \Lambda} |\langle \tilde{\Phi}_j, \Phi_\Lambda \mathbf{s}_\Lambda + \mathbf{n} \rangle| \le \max_{k \notin \Lambda} |\langle \tilde{\Phi}_k, \Phi_\Lambda \mathbf{s}_\Lambda + \mathbf{n} \rangle|. \quad (9)$$

The error event of (9) is defined using a measurement vector, $\mathbf{y}$ in [22], [23], while the indices of a support set can be found using the residual vectors in OMP. It is unfavorable for the support recovery to use a measurement vector instead of residual vectors, which can reduce the interferences caused by the superposition with other atoms for the support set in $\mathbf{y}$. However, the probability of (9) can be used as an asymptotic upper bound on a recovery error probability (i.e., a probability of an error event of (8) for Q iterations) because the interferences become negligibly small with a large L. In addition, since $\Pr(\min_{j \in \Lambda} |\langle \tilde{\Phi}_j, \Phi_\Lambda \underline{\mathbf{s}}_\Lambda + \mathbf{n} \rangle| > \min_{j \in \Lambda} |\langle \tilde{\Phi}_j, \Phi_\Lambda \mathbf{s}_\Lambda + \mathbf{n} \rangle|)$ approaches zero with a large L and a high SNR (from Lemma 2), the recovery error event in (9) can be modified to easily derive an upper bound on the recovery error probability in a large-system and high-SNR regime as follows:

$$\min_{j \in \Lambda} |\langle \tilde{\Phi}_j, \Phi_\Lambda \underline{\mathbf{s}}_\Lambda + \mathbf{n} \rangle| \le \max_{k \notin \Lambda} |\langle \tilde{\Phi}_k, \Phi_\Lambda \underline{\mathbf{s}}_\Lambda + \mathbf{n} \rangle|. \quad (10)$$

The term on the left-hand side of (10) can be rewritten as

$$\min_{j \in \Lambda} |\langle \tilde{\Phi}_j, \Phi_\Lambda \underline{\mathbf{s}}_\Lambda + \mathbf{n} \rangle| = \min_{j \in \Lambda} |\underline{\mathbf{s}}_j + \langle \tilde{\Phi}_j, \Phi_{\Lambda^\circ} \underline{\mathbf{s}}_{\Lambda^\circ} + \mathbf{n} \rangle|. \quad (11)$$

Then, from (10), (11) and Lemma 1, we can rewrite the recovery error event as follows:

$$\min_{j \in \Lambda} |s_{min} + \Gamma_{j,\Lambda^\circ}| - \max_{k \notin \Lambda} |\Gamma_{k,\Lambda}| \le 0. \quad (12)$$

For convenience, let $\eta = |s_{min} + \Gamma_{j,\Lambda^\circ}|$, $\kappa = |\Gamma_{k,\Lambda}|$. Thus, to obtain a recovery error probability, the distributions of η and κ have to be derived. From Lemma 1, the distribution of $\Gamma_{k,\Lambda}$ becomes $\mathcal{N}(0, \tilde{\sigma}_\kappa^2)$ in a large-system. Thus, κ follows the half-normal distribution, where the probability density function and cumulative distribution function of κ can be approximated by

$$f_\kappa(z) \approx \frac{\sqrt{2}}{\sqrt{\pi \tilde{\sigma}_\kappa^2}} \exp\left(-\frac{z^2}{2\tilde{\sigma}_\kappa^2}\right) \quad (13)$$

and

$$F_\kappa(z) \approx \mathrm{erf}\left(\frac{z}{\tilde{\sigma}_\kappa \sqrt{2}}\right). \quad (14)$$

Then, the mean and variance of κ are given by, respectively,

$$\mu_\kappa = \tilde{\sigma}_\kappa \sqrt{\frac{2}{\pi}} \quad (15)$$

and

$$\sigma_\kappa^2 = \tilde{\sigma}_\kappa^2 \left(1 - \frac{2}{\pi}\right). \quad (16)$$

In the same manner, from Lemma 1, the distribution of $s_{min} + \Gamma_{j,\Lambda^\circ}$ becomes a Gaussian distribution with a mean of s_{min} and a variance of $\tilde{\sigma}_\eta^2$ in a large system (i.e., with a large L). Thus, η follows the folded normal distribution,

where the probability density function and cumulative distribution function of η can be approximated by

$$f_\eta(x) \approx \frac{1}{\sqrt{2\pi \tilde{\sigma}_\eta^2}} \exp\left(-\frac{(x - \tilde{\mu}_\eta)^2}{2\tilde{\sigma}_\eta^2}\right)$$
$$+ \frac{1}{\sqrt{2\pi \tilde{\sigma}_\eta^2}} \exp\left(-\frac{(x + \tilde{\mu}_\eta)^2}{2\tilde{\sigma}_\eta^2}\right) \quad (17)$$

and

$$F_\eta(x) \approx \frac{1}{2}\left(\mathrm{erf}\left(\frac{x + \tilde{\mu}_\eta}{\tilde{\sigma}_\eta \sqrt{2}}\right) + \mathrm{erf}\left(\frac{x - \tilde{\mu}_\eta}{\tilde{\sigma}_\eta \sqrt{2}}\right)\right). \quad (18)$$

Then, the mean and variance of η are given by, respectively,

$$\mu_\eta = \tilde{\sigma}_\eta \sqrt{\frac{2}{\pi}} \exp\left(-\frac{\tilde{\mu}_\eta^2}{2\tilde{\sigma}_\eta^2}\right) + \tilde{\mu}_\eta\left(-\mathrm{erf}\left(\frac{-\tilde{\mu}_\eta}{\sqrt{2}\tilde{\sigma}_\eta}\right)\right) \quad (19)$$

and

$$\sigma_\eta^2 = s_{min} + \frac{s_{min}^2(Q-1)}{L} + \sigma^2 - \mu_\eta^2. \quad (20)$$

Let $\Omega = \min_{j \in \Lambda} \eta - \max_{k \notin \Lambda} \kappa$. Then, according to order statistics [28], the probability density function of $\max_{k \notin \Lambda} \kappa$ becomes $(N - Q)f_\kappa(z)[F_\kappa(z)]^{N-Q-1}$, where $|\{1, 2, \cdots, N\} \setminus \Lambda| = N - Q$. In addition, the cumulative distribution function of $\min_{j \in \Lambda} \eta$ is given by $1 - [1 - F_\eta(x)]^Q$, where $|\Lambda| = Q$. From the distributions, an asymptotic upper bound on the recovery error probability in the high-SNR regime can be approximated as

$$P_e < \Pr(\Omega \le 0)$$
$$\approx \int_0^\infty (N-Q)f_\kappa(z)[F_\kappa(z)]^{N-Q-1}\left(1 - [1 - F_\eta(z)]^Q\right) dz. \quad (21)$$

Then, our proof can be completed by substituting (13), (14), and (18) into (21).
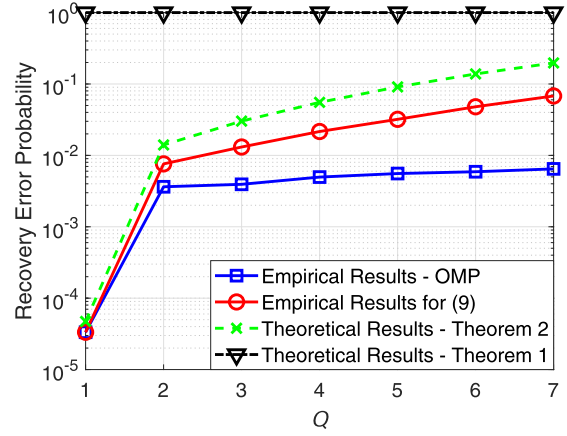
C. DISCUSSIONS ON ANALYSIS

In this subsection, detailed descriptions for the analysis of the support recovery are presented with a particular focus on the features of Theorem 2.
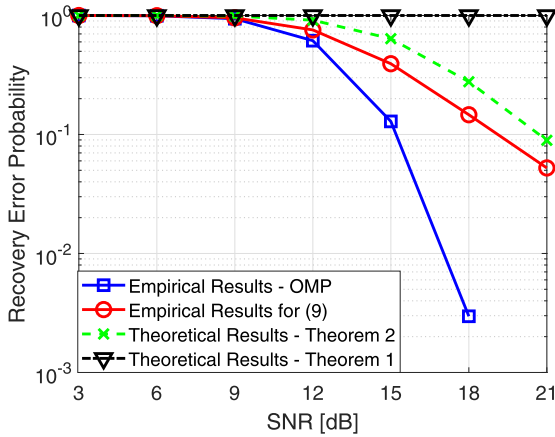
Correlation is an important factor in the support recovery of OMP since an index associated with the highest correlation coefficient between a measurement vector, $\mathbf{y}$, and a column vector of a measurement matrix, Φ, is iteratively selected. Note that $\frac{\eta}{\|\mathbf{y}\|}$ and $\frac{\kappa}{\|\mathbf{y}\|}$ are the correlation coefficients for the correct and incorrect supports associated with $j \in \Lambda$ and $k \notin \Lambda$, respectively. Then, from (13) and (17), it is shown that η follows a different distribution from κ. In detail, η is an absolute value of a Gaussian random variable with the mean of s_{min} and variance of $\frac{s_{min}^2(Q-1)}{L} + \sigma^2$. In addition, κ is an absolute value of a Gaussian random variable with zero mean and variance of $\frac{Q}{L}\left(\left(\frac{s_{max}+s_{min}}{2}\right)^2 + \frac{1}{12}(s_{max} - s_{min})^2\right) + \sigma^2$. Then, we can intuitively know that as s_{min} and L increase (or Q and σ^2 decrease), the recovery error probability decreases
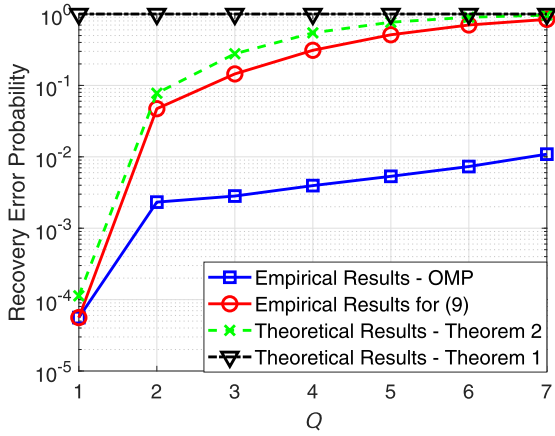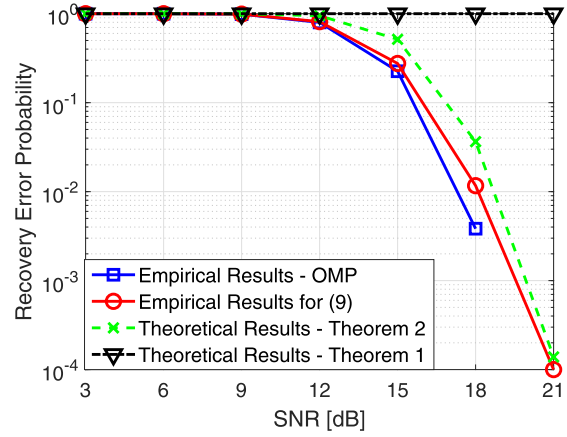
(a) $L = 128$, $N = 2L$, SNR= 18dB, and $\zeta = 0.7$

(b) $L = 128$, $N = 2L$, $Q = 3$, and $\zeta = 0.7$

FIGURE 2. Performances for a recovery error probability over various parameters, where $L = 128$, $N = 2L$, and $\zeta = 0.7$.

(a) $L = 1024$, $N = 2L$, SNR= 18dB, and $\zeta = 0.7$

(b) $L = 1024$ $N = 2L$ $Q = 3$, and $\zeta = 0 7$

FIGURE 3. Performances for a recovery error probability over various parameters, where $L = 1024$, $N = 2L$, and $\zeta = 0.7$.

because $\Pr(\eta > \kappa)$ becomes high with a large $|\mu_\eta - \mu_\kappa|$ and small σ_η^2 and σ_κ^2.

From (21), an upper bound on the recovery error probability which is an integral expression is derived. Although it is difficult to transform the upper bound to a closed-form expression due to an error function, it is possible to ensure the stability of the integral expression. In (21), if $z \to \infty$, a term in the integral approaches zero due to $f_\kappa(\infty) = 0$, $F_\kappa(\infty) = 1$ and $F_\eta(\infty) = 1$. Therefore, with a sufficiently large number for the integral upper limit, a numerical technique can be used to perform the integral in (21).

Let us compare Theorem 1 and 2 theoretically. A recovery error in CS is defined as $\hat{\Lambda} \neq \Lambda$, where $\hat{\Lambda}$ is an estimated support set. Instead of $\hat{\Lambda} \neq \Lambda$, a recovery error event in (9) is alternatively used to derive an upper bound on the recovery error probability in both analyses. In addition, the same parameters (i.e., L, N, Q, s_{min}, s_{max}, σ^2) are considered in both the analyses. However, statistical features of a random measurement matrix are used to derive an upper bound on the recovery error probability in the proposed derivation,

while the conventional derivation in [23] is based on mutual coherence of a measurement matrix (i.e., μ_{max}). Above all, Theorem 2 can ensure a significantly improved bound, compared to the conventional bound (i.e., Theorem 1) (Detailed performance evaluation will be presented in Section III).

III. NUMERICAL RESULTS

In this section, we present the numerical results to compare the theoretical results of Theorem 1 and 2 with the empirical results. For simulations, it is assumed that s_{min} and s_{max} are included in a non-zero coefficient set of **s** and the other coefficients follow a uniform distribution on the interval $[s_{min} \; s_{max}]$ and elements of a measurement matrix follow the Gaussian distribution (i.e., $\Phi_{l,n} \sim \mathcal{N}(0, 1/L)$). The SNR is defined as $\frac{\mathbb{E}[\|\mathbf{s}_j\|^2]}{\sigma^2}$ for $j \in \Lambda$. The empirical results for the recovery error probability are divided into two cases: 1) empirical results for $\hat{\Lambda} \neq \Lambda$ obtained by OMP algorithm, 2) empirical results for the alternative recovery error event in (9) used to derive Theorems 1 and 2. As mentioned in Section II, the second case has a larger recovery error probability than that of the first case.
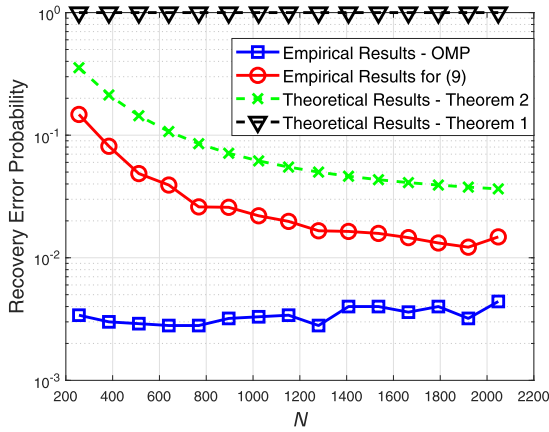
(a) $Q = 3$, $L = N/2$, SNR= 18dB, and $\zeta = 0.7$
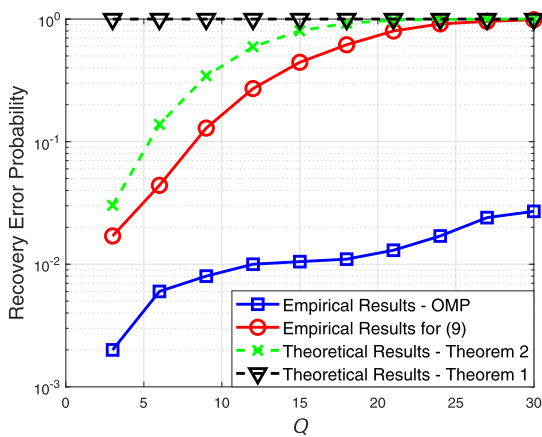
(b) $L = 1024$ $N = 2L$, SNR= 18dB, and $\zeta = 0.7$

FIGURE 4. Performances for a recovery error probability with large *L, N,* and *Q*, where SNR= 18dB, and $\zeta = 0.7$.
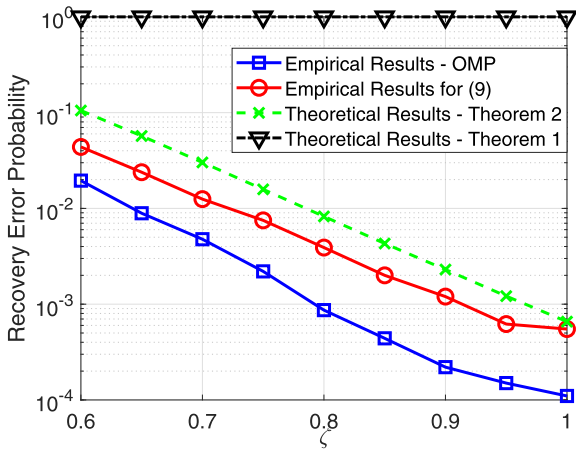

FIGURE 5. Performances for a recovery error probability over various ζ, where *L* = 1024, SNR = 18dB, *N* = 2*L*, and *Q* = 3.

Meanwhile, since the mutual coherence of a random measurement matrix is relatively high (e.g., $\mathbb{E}[\mu_{max}] = 0.3899$ when $L = 128$), all theoretical results of Theorem 1 converge to one in the following simulation environment. Even if $\sigma^2 \to 0$ (which results in SNR $\to \infty$), an upper bound on the recovery
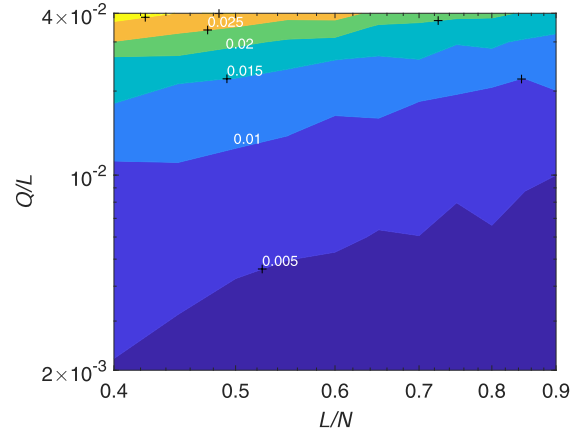
(a) Empirical error probability
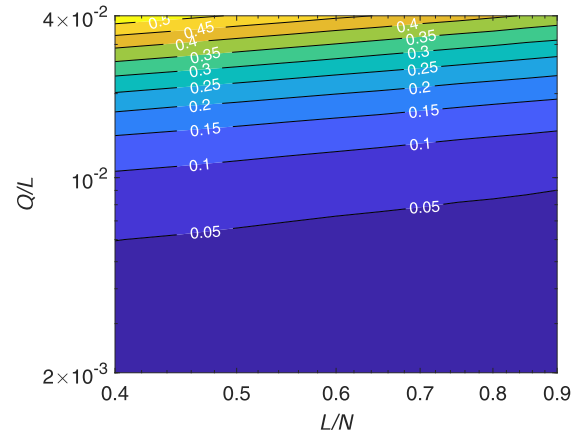
(b) The gap between the theoretical and empirical error probabilities

FIGURE 6. Performances over various undersampling and relative sparsity ratios, where $\zeta = 0.7$ and SNR = 18dB.

error probability in Theorem 1 remains one with $\mathbb{E}[\mu_{max}] = 0.3899$.

From the following figures, we can find an improved performance guarantee in Theorem 2 over various parameters (i.e., $L, N, Q, \sigma^2, s_{min}$ and s_{max}) in terms of the recovery error probability. Fig. 2 and Fig. 3 show that as the SNR and the sparsity, defined as $\varphi = \frac{N-Q}{Q}$, increase (i.e., Q decreases), the recovery error probability decreases. In addition, by comparing Fig. 2 with Fig. 3, it is shown that Theorem 2 can yield theoretical results that are much closer to the empirically obtained results with a large L because it is easier to discriminate between Φ_j and Φ_k with a large dimension of a column vector of Φ in a recovery error event defined in (9). From Fig. 4(a), we can clearly find that the prediction accuracy of Theorem 2 is significantly improved in a large system (i.e., large L and N). Furthermore, as seen in Fig. 4(b), the theoretical results in Theorem 2 are sufficiently close to empirical results with a high sparsity (i.e., small Q) because the superposition interference of a measurement vector in (9) becomes small with a small Q. A ratio of s_{min} to s_{max}, i.e., $\zeta = \frac{s_{min}}{s_{max}}$, is also a crucial parameter that influences on the recovery error probability. If ζ is small (e.g., $0.1 \sim 0.5$),

the recovery performance of OMP deteriorates significantly. As shown in (3), it is difficult to obtain a moderate recovery performance with a low ζ. So, we deal with the recovery performance for a high ζ. From Fig. 5, we can find that a large ζ, which is maintained to obtain a reasonable recovery performance in CS is favorable for theoretical results in Theorem 2. Furthermore, in Fig. 6, we show the gap between the theoretical and empirical results (i.e., accuracy of Theorem 2) over various undersampling (i.e., L/N) and relative sparsity ratios (i.e., Q/L) when $\zeta = 0.7$ and SNR = 18dB. From the figure, it is found that Theorem 2 is valid with large undersampling and small relative sparsity ratios (i.e., a large system).

IV. CONCLUSION

In this paper, we analyzed support recovery with a random measurement matrix in OMP. An asymptotic upper bound on the recovery error probability for the random ensemble was derived using statistical features of a random measurement matrix. The upper bound can ensure a more precise performance prediction than the conventional analysis without any required condition. From the numerical results, we showed that the upper bound is sufficiently close to the empirical results which is desirable for its applicability. Furthermore, for a large system (i.e., large L and N), a gap between the theoretical and empirical results becomes negligibly small.

APPENDIX

A. PROOF OF LEMMA 1

Expanding $\Gamma_{k,\Lambda}$, it is shown that

$$
\Gamma_{k,\Lambda} = \sum_{q=1}^{Q} \tilde{\Phi}_k^H \Phi_{\Lambda(q)} s_{\Lambda(q)} + \tilde{\Phi}_k^H \mathbf{n},
$$

$$
= \sum_{q=1}^{Q} \left(\sum_{l=1}^{L} \tilde{\Phi}_{l,k}^* \Phi_{l,\Lambda(q)} s_{\Lambda(q)} \right) + \sum_{l=1}^{L} \tilde{\Phi}_{l,k}^* \mathbf{n}_l. \quad (22)
$$

Here, $\text{Var}[\tilde{\Phi}_{l,k}^* \Phi_{l,\Lambda(q)} s_{\Lambda(q)}] = \frac{1}{L^2}((\frac{s_{max}+s_{min}}{2})^2 + \frac{1}{12}(s_{max} - s_{min})^2)$ and $\text{Var}[\tilde{\Phi}_{l,k}^* \mathbf{n}_l] = \frac{\sigma^2}{L}$, because $\Phi_{l,\Lambda(q)} \sim \mathcal{N}\left(0, \frac{1}{L}\right)$, $s_{\Lambda(q)} \sim U(s_{min}, s_{max})$, and $\mathbf{n}_l \sim \mathcal{N}\left(0, \sigma^2\right)$ are independent (Note that $\text{Var}(AB) = \text{Var}(A)\text{Var}(B) + \text{Var}(A)(\mathbb{E}[B])^2 + \text{Var}(B)(\mathbb{E}[A])^2$ if A and B are independent). Then, according to the central limit theorem (CLT) for large L and Q (i.e., a large system), the probability density functions of $\sum_{q=1}^{Q} \left(\sum_{l=1}^{L} \tilde{\Phi}_{l,k}^* \Phi_{l,\Lambda(q)} s_{\Lambda(q)} \right)$ and $\sum_{l=1}^{L} \tilde{\Phi}_{l,k}^* \mathbf{n}_l$ become $\mathcal{N}\left(0, \frac{Q}{L}((\frac{s_{max}+s_{min}}{2})^2 + \frac{1}{12}(s_{max} - s_{min})^2)\right)$ and $\mathcal{N}\left(0, \sigma^2\right)$, respectively. From them, the distribution of $\Gamma_{k,\Lambda}$ becomes $\mathcal{N}\left(0, \tilde{\sigma}_\kappa^2\right)$, where $\tilde{\sigma}_\kappa^2 = \frac{Q}{L}((\frac{s_{max}+s_{min}}{2})^2 + \frac{1}{12}(s_{max} - s_{min})^2) + \sigma^2$.

In the same manner, it can be shown that

$$
\Gamma_{j,\Lambda^\circ} = \sum_{q=1}^{Q-1} \tilde{\Phi}_j^H \Phi_{\Lambda^\circ(q)} \underline{s}_{\Lambda^\circ(q)} + \tilde{\Phi}_j^H \mathbf{n},
$$

$$
= s_{min} \sum_{q=1}^{Q-1} \left(\sum_{l=1}^{L} \tilde{\Phi}_{l,j}^* \Phi_{l,\Lambda^\circ(q)} \right) + \sum_{l=1}^{L} \tilde{\Phi}_{l,j}^* \mathbf{n}_l. \quad (23)
$$

Note that Φ_{Λ° is supported on Λ° which has $Q - 1$ indices excluding j in Λ. Then, as $\Gamma_{k,\Lambda}$, the distribution of Γ_{j,Λ° becomes $\mathcal{N}\left(0, \tilde{\sigma}_\eta^2\right)$ due to the CLT for large L and Q, where $\tilde{\sigma}_\eta^2 = \frac{s_{min}^2(Q-1)}{L} + \sigma^2$.

B. PROOF OF LEMMA 2

To show that $\text{Pr}(\min_{j\in\Lambda} |\langle \tilde{\Phi}_j, \Phi_\Lambda s_\Lambda + \mathbf{n}\rangle| < \min_{j\in\Lambda} |\langle \tilde{\Phi}_j, \Phi_\Lambda \underline{s}_\Lambda + \mathbf{n}\rangle|)$ approaches zero in a large-system and high-SNR regime, $\text{Pr}(|\langle \tilde{\Phi}_j, \Phi_\Lambda s_\Lambda + \mathbf{n}\rangle| < |\langle \tilde{\Phi}_j, \Phi_\Lambda \underline{s}_\Lambda + \mathbf{n}\rangle|)$ is alternatively analyzed in this subsection. To this end, consider

$$
\langle \tilde{\Phi}_j, \Phi_\Lambda s_\Lambda + \mathbf{n}\rangle = s_j + \tilde{\Phi}_j^T \Phi_{\Lambda^\circ} s_{\Lambda^\circ} + \tilde{\Phi}_j^T \mathbf{n},
$$

$$
\langle \tilde{\Phi}_j, \Phi_\Lambda \underline{s}_\Lambda + \mathbf{n}\rangle = s_{min} + \Gamma_{j,\Lambda^\circ}. \quad (24)
$$

For convenience, let $\tilde{\Gamma}_{j,\Lambda^\circ} = \tilde{\Phi}_j^T \Phi_{\Lambda^\circ} s_{\Lambda^\circ} + \tilde{\Phi}_j^T \mathbf{n}$. From Lemma 1, it can be found that the distribution of $s_{min} + \Gamma_{j,\Lambda^\circ}$ becomes $\mathcal{N}(s_{min}, \tilde{\sigma}_\eta^2)$ in a large system. Then, $\text{Pr}(s_{min} + \Gamma_{j,\Lambda^\circ} < 0)$ approaches zero in a large-system and high-SNR regime because $s_{min} > 0$ and $\tilde{\sigma}_\eta^2(= \frac{s_{min}^2(Q-1)}{L} + \sigma^2) \to 0$. In the same manner, as $\Gamma_{k,\Lambda}$ and Γ_{j,Λ°, the distribution of $\tilde{\Gamma}_{j,\Lambda^\circ}$ becomes to $\mathcal{N}(0, \sigma_{\tilde{\Gamma}}^2)$ with a large L, where $\sigma_{\tilde{\Gamma}}^2 = \frac{Q-1}{L}((\frac{s_{max}+s_{min}}{2})^2 + \frac{1}{12}(s_{max} - s_{min})^2) + \sigma^2$. From this, the distribution of $s_{min} + \tilde{\Gamma}_{j,\Lambda^\circ}$ also becomes $\mathcal{N}(s_{min}, \sigma_{\tilde{\Gamma}}^2)$. Then, in a large-system and high-SNR regime, $\text{Pr}(\mathbf{s}_j + \tilde{\Gamma}_{j,\Lambda^\circ} < 0)$ approaches zero, because $\text{Pr}(\mathbf{s}_j + \tilde{\Gamma}_{j,\Lambda^\circ} < 0) < \text{Pr}(s_{min} + \tilde{\Gamma}_{j,\Lambda^\circ} < 0)$ and $\sigma_{\tilde{\Gamma}}^2 \to 0$ when $L \to \infty$ and $\sigma^2 \to 0$. Then, it can be shown that

$$
\text{Pr}(|\langle \tilde{\Phi}_j, \Phi_\Lambda s_\Lambda + \mathbf{n}\rangle| < |\langle \tilde{\Phi}_j, \Phi_\Lambda \underline{s}_\Lambda + \mathbf{n}\rangle|)
$$
$$
\to \text{Pr}(s_j + \tilde{\Gamma}_{j,\Lambda^\circ} < s_{min} + \Gamma_{j,\Lambda^\circ}),
$$
$$
= \text{Pr}(\tilde{s}_j + \tilde{\Phi}_j^T \Phi_{\Lambda^\circ} \tilde{s}_{\Lambda^\circ} < 0), \quad (25)
$$

where $\tilde{s}_j(= s_j - s_{min})$ and an element of $\tilde{s}_{\Lambda^\circ}(= s_{\Lambda^\circ} - s_{min})$ follow a uniform distribution with an interval $[0, s_{max} - s_{min}]$. Note that the mean and variance of $\tilde{s}_j$ are given by $\mathbb{E}[\tilde{s}_j] = \frac{1}{s_{max}-s_{min}}$ and $\text{Var}[\tilde{s}_j] = (\frac{s_{max}-s_{min}}{2})^2 + \frac{1}{12}(s_{max}-s_{min})^2$, respectively. Then, as $\tilde{\Gamma}_{j,\Lambda^\circ}$, the distribution of $\tilde{\Phi}_j^T \Phi_{\Lambda^\circ} \tilde{s}_{\Lambda^\circ}$ becomes $\mathcal{N}(0, \sigma_{\tilde{s}}^2)$ because of the CLT for large L and Q, where $\sigma_{\tilde{s}}^2 = \frac{Q-1}{L}((\frac{s_{max}-s_{min}}{2})^2 + \frac{1}{12}(s_{max} - s_{min})^2)$. Thus, by using the joint distribution of $\tilde{s}_j$ and $\tilde{\Phi}_j^T \Phi_{\Lambda^\circ} \tilde{s}_{\Lambda^\circ}$, the probability is given by

$$
\text{Pr}(\tilde{s}_j + \tilde{\Phi}_j^T \Phi_{\Lambda^\circ} \tilde{s}_{\Lambda^\circ} < 0)
$$
$$
= \int_{-\infty}^{0} \int_{0}^{\alpha} \frac{1}{\alpha} \frac{1}{\sqrt{2\pi}\sigma_{\tilde{s}}} \exp\left(\frac{-(x-u)^2}{2\sigma_{\tilde{s}}^2}\right) du\, dx
$$
$$
= \int_{-\infty}^{0} \frac{1}{2\alpha} \left(\left[1 + \text{erf}\left(\frac{u-x}{\sigma_{\tilde{s}}\sqrt{2}}\right)\right] \Big|_0^\alpha \right) dx
$$
$$
= \int_{-\infty}^{0} \frac{1}{2\alpha} \left[\text{erf}\left(\frac{\alpha-x}{\sigma_{\tilde{s}}\sqrt{2}}\right) - \text{erf}\left(\frac{-x}{\sigma_{\tilde{s}}\sqrt{2}}\right)\right] dx, \quad (26)
$$

where $\alpha = s_{max} - s_{min}$. For convenience, let $a = \frac{\alpha}{\sigma_{\tilde{s}}\sqrt{2}}$ and $b = -\frac{1}{\sigma_{\tilde{s}}\sqrt{2}}$. Then, from $\int \text{erf}(a + bx)dx = \frac{(a+bx)\text{erf}(a+bx)}{b} +$

$\frac{\exp(-(a+bx)^2)}{b\sqrt{\pi}}$, it can be shown that

$$\Pr(\tilde{s}_j + \tilde{\Phi}_j^T \Phi_{\Lambda \circ} \tilde{s}_{\Lambda \circ} < 0)$$

$$= \frac{1}{2\alpha} \int_{-\infty}^{0} \mathrm{erf}(a+bx) - \mathrm{erf}(bx)dx,$$

$$= \frac{1}{2\alpha} \left(\left[\frac{(a+bx)\mathrm{erf}(a+bx)}{b} + \frac{\exp(-(a+bx)^2)}{b\sqrt{\pi}} \right. \right.$$

$$\left. \left. - \frac{bx\,\mathrm{erf}(bx)}{b} - \frac{\exp(-(bx)^2)}{b\sqrt{\pi}} \right] \Big|_{-\infty}^{0} \right),$$

$$= \frac{1}{2\alpha} \left(\left[\frac{a\mathrm{erf}(a+bx)}{b} + \frac{1}{b\sqrt{\pi}} \left(\exp(-(a+bx)^2) \right. \right. \right.$$

$$\left. \left. \left. - \exp(-(bx)^2) \right) + x(\mathrm{erf}(a+bx) - \mathrm{erf}(bx)) \right] \Big|_{-\infty}^{0} \right),$$

$$= \frac{1}{2\alpha} \left(\frac{a\,\mathrm{erf}(a)}{b} + \frac{1}{b\sqrt{\pi}}(\exp(-a^2)-1) - \frac{a}{b} \right),$$

$$= \frac{1}{2} \left(1 - \mathrm{erf}\left(\frac{s_{max} - s_{min}}{\sigma_{\tilde{s}}\sqrt{2}} \right) \right) + \frac{\sigma_{\tilde{s}}}{\sqrt{2\pi}(s_{max} - s_{min})}$$

$$\times \left(1 - \exp\left(-\left(\frac{s_{max} - s_{min}}{\sigma_{\tilde{s}}\sqrt{2}} \right)^2 \right) \right). \tag{27}$$

When $L \to \infty$, $\sigma_{\tilde{s}}^2$ approaches zero. Furthermore, $a = \infty$, $b = -\infty$, $\mathrm{erf}(a) = 1$ and $\exp(-a^2) = 0$, when $\sigma_{\tilde{s}}^2 = 0$. Then, from (27), $\Pr(\tilde{s}_j + \tilde{\Phi}_j^T \Phi_{\Lambda \circ} \tilde{s}_{\Lambda \circ} < 0)$ approaches zero in a large-system and high-SNR regime.

REFERENCES

[1] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[2] M. Rani, S. B. Dhok, and R. B. Deshmukh, "A systematic review of compressive sensing: Concepts, implementations and applications," *IEEE Access*, vol. 6, pp. 4875–4894, 2018.

[3] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comp. Rendus Mathematique*, vol. 346, nos. 9–10, pp. 589–592, May 2008.

[4] D. Needell, J. Tropp, and R. Vershynin, "Greedy signal recovery review," in *Proc. 42nd Asilomar Conf. Signals, Syst. Comput.*, Oct. 2008, pp. 1048–1050.

[5] H. Djelouat, A. Amira, and F. Bensaali, "Compressive sensing-based IoT applications: A review," *J. Sensor Actuator Netw.*, vol. 7, no. 4, p. 45, Oct. 2018.

[6] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications: Useful tips and tricks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1527–1550, 3rd Quart., 2017.

[7] S. Park, N. Yul Yu, and H.-N. Lee, "An information-theoretic study for joint sparsity pattern recovery with different sensing matrices," *IEEE Trans. Inf. Theory*, vol. 63, no. 9, pp. 5559–5571, Sep. 2017.

[8] Q. Ling and Z. Tian, "Decentralized sparse signal recovery for compressive sleeping wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3816–3827, Jul. 2010.

[9] Y. Rachlin and D. Baron, "The secrecy of compressed sensing measurements," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2008, pp. 813–817.

[10] J. Choi, "Secure transmissions via compressive sensing in multicarrier systems," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1315–1319, Oct. 2016.

[11] Y. Zhang, L. Y. Zhang, J. Zhou, L. Liu, F. Chen, and X. He, "A review of compressive sensing in information security field," *IEEE Access*, vol. 4, pp. 2507–2519, 2016.

[12] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. New York, NY, USA: Springer, 2013.

[13] 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, Dec. 2008.

[14] M. A. Davenport and M. B. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4395–4401, Sep. 2010.

[15] L.-H. Chang and J.-Y. Wu, "An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5702–5715, Sep. 2014.

[16] M. Yang and F. de Hoog, "New coherence and rip analysis for weak orthogonal matching pursuit," in *Proc. IEEE Workshop Stat. Signal Process. (SSP)*, Jun. 2014, pp. 376–379.

[17] H. Li, G. Liu, and J. Zou, "Improved RIP-based performance guarantee for sparse signal recovery via A*OMP," *Electron. Lett.*, vol. 54, no. 21, pp. 1216–1218, Oct. 2018.

[18] X. Cai, Z. Zhou, Y. Yang, and Y. Wang, "Improved sufficient conditions for support recovery of sparse signals via orthogonal matching pursuit," *IEEE Access*, vol. 6, pp. 30437–30443, 2018.

[19] J. Chen, F. Zhu, and Q. Liang, "Information theoretic bounds for sparse reconstruction in random noise," *IEEE Access*, vol. 7, pp. 102304–102312, 2019.

[20] J. Wen, Z. Zhou, J. Wang, X. Tang, and Q. Mo, "A sharp condition for exact support recovery with orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 65, no. 6, pp. 1370–1382, Mar. 2017.

[21] A. S. Bandeira, E. Dobriban, D. G. Mixon, and W. F. Sawin, "Certifying the restricted isometry property is hard," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3448–3450, Jun. 2013.

[22] Z. Ben-Haim, Y. C. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5030–5043, Oct. 2010.

[23] E. Miandji, M. Emadi, J. Unger, and E. Afshari, "On probability of support recovery for orthogonal matching pursuit using mutual coherence," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1646–1650, Nov. 2017.

[24] M. J. Wainwright, "Information-theoretic limits on sparsity recovery in the high-dimensional and noisy setting," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5728–5741, Dec. 2009.

[25] T. Wimalajeewa and P. K. Varshney, "Application of compressive sensing techniques in distributed sensor networks: A survey," 2017, *arXiv:1709.10401*. [Online]. Available: http://arxiv.org/abs/1709.10401

[26] C. Li, G. Li, B. Kailkhura, and P. K. Varshney, "Secure distributed detection of sparse signals via falsification of local compressive measurements," *IEEE Trans. Signal Process.*, vol. 67, no. 18, pp. 4696–4706, Sep. 2019.

[27] W. Lu, Y. Wang, X. Wen, X. Hua, S. Peng, and L. Zhong, "Compressive downlink channel estimation for FDD massive MIMO using weighted l_p minimization," *IEEE Access*, vol. 7, pp. 86964–86978, 2019.

[28] H. David and H. Nagaraja, *Order Statistics*. New York, NY, USA: Wiley, 2004.

YONGGU LEE (Member, IEEE) received the B.E. degree in electronics engineering from Kyungpook National University, Daegu, South Korea, in 2013, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the Gwangju Institute of Science and Technology (GIST), Gwangju, South Korea, in 2016 and 2019, respectively. He is currently a Senior Researcher with the Research Institute for Solar and Sustainable Energies, GIST. His research interests include the areas of communications and signal processing, with emphasis on physical layer security, 5G communication systems, and machine type communications.

JINHO CHOI (Senior Member, IEEE) was born in Seoul, South Korea. He received the B.E. *(magna cum laude)* degree in electronics engineering from Sogang University, Seoul, in 1989, and the M.S.E. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), in 1991 and 1994, respectively. He was a Professor/Chair in wireless with Swansea University, U.K. He was also a Professor with the Gwangju Institute of Science and Technology (GIST), South Korea. He is currently a Professor with the School of Information Technology, Deakin University, Burwood, NSW, Australia. He authored two books published by Cambridge University Press, in 2006 and 2010. His research interests include the Internet of Things (IoT), wireless communications, and statistical signal processing. He received the 1999 Best Paper Award for Signal Processing from EURASIP and the 2009 Best Paper Award from WPMC (Conference). He is also an Editor of the IEEE TRANSACTIONS COMMUNICATIONS and the IEEE WIRELESS COMMUNICATIONS LETTERS. He is a Division Editor of the *Journal of Communications and Networks* (JCN). He also had served as an Associate Editor or Editor for other journals, including the IEEE COMMUNICATIONS LETTERS, the *Journal of Communications and Networks*, the IEEE TRANSACTIONS VEHICULAR TECHNOLOGY, and *ETRI Journal*.

EUISEOK HWANG (Member, IEEE) received the B.S. and M.S. degrees from the School of Engineering, Seoul National University, Seoul, South Korea, in 1998 and 2000, respectively, and the M.S. and Ph.D. degrees in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2010 and 2011, respectively. From 2000 to 2006, he was with the Digital Media Research Center, Daewoo Electronics Company Ltd., South Korea. From 2011 to 2014, he was with the Channel Architecture Group, LSI Corporation (now Broadcom), San Jose, CA, USA. Since 2015, he has been an Assistant/Associate Professor with the School of Mechatronics/Electrical Engineering and Computer Science/Artificial Intelligence, Gwangju Institute of Science and Technology (GIST), South Korea. He has published over 100 journal articles and conference papers in information and signal processing area. He holds over 20 granted U.S. patents. His research interests include data channel signal processing and coding, energy informatics, intelligence implementations for smart grid, and information processing for system intelligence in emerging the ICT/IoT applications.

• • •