Collaborative Framework of Algorithms for Sparse Channel Estimation in OFDM Systems

Anthony Ngozichukwuka Uwaechia and Nor Muzlifah Mahyuddin

Abstract: **For proper matrix ensembles, it has been known that the greedy pursuit (GP) algorithms are computationally efficient and fast to reconstruct sparse signals from far fewer linear measurements. In considering several parameters such as sparsity level, sparse signal ambient dimension and the number of linear measurements, the GP algorithms have been shown to perform differently in estimating sparse signals. According to data fusion principle, fusing completely the estimated support set of different reconstruction algorithms can improve signal recovery performance. It can, however, lead to the increased probability of estimating incorrect support indices, and thus degrades the signal reconstruction accuracy. In this paper, a new fusion framework, namely collaborative framework of algorithms (CoFA), is proposed to pursue accurate reconstruction of the sparse signals from far fewer linear measurements. The two main ingredients of the proposed scheme that control the estimation of incorrect support indices are pre-selection support of orthogonal matching pursuit (OMP) algorithm and Thresholding -to eliminate unpromising indices from the identied support set of any participating algorithm. Using the restricted isometry property, the theoretical analysis of the CoFA scheme and the sufficient conditions (guarantees) for realizing an improved reconstruction performance are presented. Simulation results demonstrate that the proposed scheme is effective and offer a better channel estimation performance in terms of mean-squared-error (MSE) and bit-error-rate (BER) when compared to other reconstruction algorithms, without the significant increase in computational complexity.**

Index Terms: **Compressed sensing (CS), orthogonal frequency division multiplexing (OFDM), restricted isometry property (RIP), sparse channel estimation, sparse signal reconstruction.**

I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (OFDM) system is a multicarrier modulation technique that has widely been adopted in modern communication systems [1]. This is mainly due to its immunity to frequency-selective fading channels, which support very high spectral efficiency [2]–[4]. OFDM manages to transform the frequency-selective fading channel into several flat fading subchannels with independent additive noise vectors [3], which significantly reduces the complexity of the receiver design [5]. Nonetheless, channel estimation (CE) is critical in OFDM systems [6], as accurate channel state information (CSI) can notably improve performance.

Recently, by leveraging the notion of transform coding, a new signal acquisition paradigm known as compressed sensing (CS) has emerged for sensing signals that have a sparse or compressible representation [7]–[9]. Therefore, CS relies on linear dimensionality reduction which typically recovers an inherently sparse signal from a limited number of linear measurements [9]. This has proven to be more accurate than the conventional least squares (LS) [10]. Fortunately, many high-dimensional real-world signals have coefficient vectors containing few nonzero entries compared to their ambient dimensionality [8]–[12]. However, knowledge of the locations of the nonzero signal coefficients remain unknown [10]. Therefore, the design of suitable reconstruction algorithms for identifying sparse signals still remains a much more fundamental problem in the rapidly growing field of CS [13]–[15].

Typically, solving the sparse representation problem is nondeterministic polynomial-time hard (NP-hard) [9], [15], [16], and various techniques have been proposed in the literature for solving this problem [8], [15], [17]. These techniques approximate the coefficients of the representing functions [14], employing, for instance, linear programming [18], gradient descent optimization [19] and greedy pursuit (GP) algorithms [9], [15], [17].The GP algorithms iteratively identify the support set of the unknown signal until a halting condition is met [15], [20]. Considering that the GP algorithms are typically significantly faster with much lower computational complexity than the methods of convex optimization [15], [18], [20], they have been widely used for finding sparse solutions to underdetermined, or ill-conditioned, linear systems. Some of the GP algorithms include orthogonal matching pursuit (OMP) [21], regularized OMP (ROMP) [22], stagewise OMP (StOMP), generalized OMP (gOMP) [24], multipath matching pursuit (MMP) [25], compressive sampling matching pursuit (CoSaMP) [22], [26] and subspace pursuit (SP) [27] algorithms. The summary of notations and symbols used in this paper are listed in Table 1.

The OMP algorithm proposed in [21], remains one of the most intriguing GP algorithms, wherein at each iteration, the column index of the measurement matrix that is best correlated with the modified measurements (or residual) is selected. Tropp and Gilbert [21] demonstrated that, for a given k-sparse signal $\boldsymbol{x} \in \mathbb{R}^{\mathbb{N}}$ of size N, (where k denotes the number of nonzero elements of the signal $\boldsymbol{x}$) and a random Gaussian measurement matrix $\boldsymbol{A} \in \mathbb{R}^{\mathbb{N}_{\mathbb{P}} \times \mathbb{N}}$, the OMP succeeds in reconstructing $\boldsymbol{x}$ from $\boldsymbol{y} \in \mathbb{R}^{\mathbb{N}_{\mathbb{P}}} = \boldsymbol{Ax}$ linear measurements. This is achieved in just k iterations with immense probability if the number of measurements follows $N_P \sim k\log N$ [10], [21], [22], [24]. An algo-

Manuscript received April 10, 2017; approved for publication by Yun Hee Kim, Division I, September 26, 2017.

This work was supported by the Universiti Sains Malaysia under RUI grant [1001/PELECT/814206].

The authors are with the School of Electrical and Electronic Engineering Universiti Sains Malaysia, Seri Ampangan, 14300, Nibong Tebal, Pulau Pinang, Malaysia, email: anu14_eee037@student.usm.my, eemnmuzlifah@usm.my.

N. M. Mahyuddin is the corresponding author.

Digital object identifier 10.1109/JCN.2018.000002

rithm, namely, StOMP has been proposed [10], that selects more than one highly correlated atoms that surpass a carefully designed threshold. Similarly, in [22], another technique that select multiple columns with the largest correlation in magnitude has been proposed, namely ROMP. However, the ROMP selects vectors with similar magnitude (i.e., comparable coordinates) of the correlation value before selecting the vector with the maximal energy. Another pursuit algorithm, known as gOMP has been proposed [24]. However, the gOMP targets the selection of a fixed number of atoms per iteration by modifying OMP, mainly on the identification step, for complexity reduction. Similarly, in [25], another algorithm, known as MMP, that selects multiple promising candidates at each stage of iteration, has been proposed. However, each iteration step of MMP, is associated with an increase in the number of child paths for each candidate, and hence necessarily exhibit increasing complexity. Hence, the StOMP, ROMP, gOMP and MMP algorithms consider various ways to correctly identify the component of correlation values, by modifying OMP, mainly on the identification step. However, the CoSaMP [22], [26] and SP [27] algorithms admit $2k$ and k atoms, respectively per iteration. These then require the incorporation of an extra operational step, which has the impact of pruning (i.e., backtracking), to correct erroneously estimated atoms chosen from the sensing basis.

For sparse signal recovery problem, many GP algorithms perform differently to recover or reconstruct sparse or compressible signals [28]. In the seminal work of Blanchard *et al*. [29], it is shown that if the bounds on the restricted isometry property (RIP) constant are placed on the GP algorithm RIP-based reconstruction condition, while expressing these conditions into required number of measurements, sparse signal dimension and sparsity level of the true signal, the GP algorithms will perform differently to reconstruct the unknown sparse signal. Additionally, in noisy CS measurements, Hurley and Rickard [30] observed that, the cardinality of the identified support-set, say $\hat{T}$, is often far larger than the sparsity level k of the target signal (i.e., $|\hat{T}| \gg k$, where $|\cdot|$ denotes the cardinality operator) and may require other sparse reconstruction algorithms for support identification. This consequently results in the notion of fusing different reconstruction algorithms together to enables successful signal recovery probability from a seemingly incomplete set of linear measurements [28], [31].

Ambat *et al*. [28] exploited the idea of fusing several CS reconstruction algorithms together and developed a fusion framework, namely fusion of algorithms for compressed sensing (FACS). FACS fuses completely the estimated support-sets of several viable reconstruction algorithms together, to determine the final signal estimate. Although, fusing completely the estimated support indices of algorithms may be elegant in the noiseless regime or in a considerable low level of channel noise but may lead to an increased probability of fusing several erroneously estimated indices over noisy channels. Moreover, since FACS fuses completely the estimated support sets of reconstruction algorithms, then reconstructing the signal using the restricted support (union of support set) LS estimate, may somewhat be achieved at the expense of an increased computational complexity.

In this paper, a new fusion framework, namely Collaborative Framework of Algorithms (CoFA), is proposed to pursue accurate reconstruction of the sparse signals from far fewer linear measurements. Unlike FACS, the CoFA scheme exploits both the estimated support-set and sparse coefficients of algorithms to revise the estimate of support set of the target signal, with the aim of correcting erroneously estimated atoms chosen from the sensing basis. Hence, the CoFA scheme builds on two main ingredients which include the pre-selection support of OMP algorithm and thresholding—to yield a revised estimate of the estimated support set of any participating algorithm considered in the fusion framework. Using the RIP, the theoretical analysis of the CoFA scheme and the sufficient conditions (guarantees) for realizing an improved reconstruction performance are presented. Simulation results demonstrate that an improvement in mean squared-error (MSE) and bit-error-rate (BER) performance can be achieved by using the proposed scheme when compared with the FACS scheme.

The remainder of this paper is organized as follows. Section II presents an overview of the CS theory and the formulation of the system model as a CS problem. Section III presents the proposed sparse recovery scheme as well as the derivation of the performance bounds and reconstruction guarantees of the scheme. Section IV presents the results of numerical experiments. Finally, Section V concludes this paper.

Table 1. Summary of notations and symbols.

Notation	Representation of the symbol or symbol
k	Sparsity number i.e., number of non-zero elements of a vector
$\lvert\cdot\rvert$	Cardinality operator
$\lceil\cdot\rceil$	The ceiling function
$\#\{\cdot\}$	Number of elements
L	Channel length
$\boldsymbol{h}$	Channel impulse response
N	Total number of subcarriers
$\boldsymbol{y}$	Received CS measurements
$\boldsymbol{A}$	Measurement matrix
$(\cdot)^T$	Transpose
$\mathtt{diag}\{\cdot\}$	Diagonal Matrix
$\lVert\cdot\rVert_0$	l_0-norm
$\lVert\cdot\rVert_1$	l_1-norm
$\lVert\cdot\rVert_2$	l_2-norm
$\boldsymbol{A}^\dagger$	Moore-Penrose pseudo-inverse matrix of $\boldsymbol{A}$
$\hat{T}_\triangle$	The set of indices contained in $\breve{T}$ that is pruned by the algorithm.
$\boldsymbol{I}_N$	Identity matrix of size N
$\mathbb{R}^{(\cdot)}$	Real field with dimension $(\cdot)$
$\mathcal{CN}$	Complex Gaussian distribution
$\emptyset$	Empty set
$\boldsymbol{A}_{\breve{T}}$	Column submatrix of $\boldsymbol{A}$, with column indices equal to elements in the set $\breve{T}$
$\mathrm{supp}(x)$	The support of a vector x
x_p	Sub-vector formed by the elements of vector x listed in the set p.
T^c	Complements of the set T
$\setminus$	Difference set
$\tilde{T}$, $\breve{T}$ and $\hat{T}$	Estimated support set
$\tilde{\boldsymbol{h}}$, and $\hat{\boldsymbol{h}}$	Cox-Ingersoll-Ross (CIR) estimates
$\boldsymbol{A}^*$	Conjugate transpose or Hermitian transpose of $\boldsymbol{A}$

II. CS THEORY AND SYSTEM MODEL

In this section, an overview of the fundamental theories associated with the CS scheme is presented. The detailed description

of the basic improvements in CS can be found in the study of [8], [32]–[35].

A. CS Theory

CS is a very efficient sub-Nyquist signal sampling scheme that allows, under certain assumptions, the accurate recovery of signals that are sparse or compressible with respect to some known representation basis [10], [15]. The efficiency of this scheme relies on the measurement technique employed and the reconstruction algorithm used for support identification (i.e., to identify the sparse set of representation coefficients). If $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_N]$, with $\boldsymbol{\psi}_n \in \mathbb{R}^N$, is a representation basis for a real-valued, discrete-time signal $\boldsymbol{x} \in \mathbb{R}^{\mathbb{N}}$, where N is the signal dimension, then the sparse representation assumes that the original signal $\boldsymbol{x} = \sum_{i=1}^{N} \psi_i \phi_i = \boldsymbol{\Psi}\Phi$ is compressible. Such that $\phi_i = \langle \boldsymbol{x}, \psi_i \rangle$ is the representation of $\boldsymbol{x}$ in a basis $\boldsymbol{\Psi}$ which form an $N \times 1$ vector $\Phi = [\phi_1, \phi_2, \cdots, \phi_N]^T$. Typically, since Φ is sparse, then the signal $\boldsymbol{x}$ is called k-sparse. Consequently, $\|\Phi\|_0 \leq k$ with $k \ll N$, which denotes the l_0-norm measure, counts the number of non-zero elements in the vector Φ i.e., $\|\Phi\|_0 = \#\{\phi_i \neq 0, i = 1, 2, \cdots, N\}$. In many practical applications, $\boldsymbol{x}$ is unknown, but through CS (sampling), it is measurable in the form of compressed measurements given as

$$\boldsymbol{y} = \boldsymbol{\Theta x} = \boldsymbol{\Theta\Psi}\Phi = \boldsymbol{A}\Phi, \tag{1}$$

where $\boldsymbol{\Theta} \in \mathbb{R}^{\mathbb{N}_\mathbb{P} \times \mathbb{N}}$ represents the sensing matrix, N_P the size of $\boldsymbol{y}$ with $N_P < N$, $\boldsymbol{A} \in \mathbb{R}^{\mathbb{N}_\mathbb{P} \times \mathbb{N}} = \boldsymbol{\Theta\Psi}$ represents the measurement matrix such that $k < N_P < N$. When $N_P < N$, this necessarily leads to an underdetermined system of equations with infinitely many solutions of $\boldsymbol{x}$ satisfying (1). To have a single unique solution, the measurement matrix $\boldsymbol{A}$ is required to satisfy $N_P > N$. Nonetheless, the required estimation is, however, an underdetermined system in order to reduce the amount of pilot overhead required for CE. Since $\boldsymbol{x}$ is k-sparse, the unique solution (which is necessarily the sparsest possible) can be derived by using the CS framework to solve the l_0-norm minimization problem, as long as $\texttt{spark}(\boldsymbol{A}) > 2k$ is satisfied [9]. The spark of a matrix is the smallest number of atoms of that matrix that are linearly dependent [9]. To express differently, the following minimization problem needs to be solved

$$\hat{\Phi} = \arg\min_{\Phi} \|\Phi\|_0, \quad \text{s.t.} \quad \boldsymbol{y} = \boldsymbol{A}\Phi, \tag{2}$$

where $\|\cdot\|$ is the l_0-norm, but unfortunately, this minimization problem is combinatorial (specifically, NP-hard), and cannot be solved in polynomial time [22], [23]. Alternatively, since the measurement matrix $\boldsymbol{A}$, approximately preserves the norm of $\boldsymbol{x}$ when the matrix $\boldsymbol{A}$ is occupied by independent and identically distributed (i.i.d.) random entries [10], [9], in relation to basis pursuit (BP), one can attempt to solve the l_1-norm minimization problem of the form

$$\hat{\Phi} = \arg\min_{\Phi} \|\Phi\|_1, \quad \text{s.t.} \quad \boldsymbol{y} = \boldsymbol{A}\Phi. \tag{3}$$

This problem is a convex optimization problem and can be recast as a linear programming problem [9]. A reliable accurate solution can be obtained provided that the matrix $\boldsymbol{A}$ is a random matrix and that the number of $N_P = O(k\log(N/k))$ measurements required, are satisfied [22]. For large scale problems, the BP algorithm is known to converge very slowly [36]. Fortunately, approximate solutions are available using the GP algorithms [22].

For GP algorithms, let the ith column of $\boldsymbol{A}$ be represented by c_i, where $i \in [N]$ and $[N] := \{1, 2, \cdots, N\}$. As the entries in $\boldsymbol{y}$ are a linear combinations of k columns of $\boldsymbol{A}$, solving the sparse approximation problem can be formulated to as the problem of accurately identifying the support $\{c_1, c_2, \cdots, c_k\}$ of the k-sparse signal, to then efficiently recover the signal $\boldsymbol{x}$ from its measurements $\boldsymbol{y}$.

B. System Model

Consider an OFDM system with a comb-type pilot arrangement having N subcarriers, of which N_P subcarriers are indexed with pilots, in position $P = \{p_1, p_2, \cdots, p_{N_p}\} \subseteq \{1 \leq p_1 < p_2 < \cdots < p_{N_p} \leq N\}$ which are known *a priori* to the receiver. Let $\boldsymbol{W} \in \mathbb{C}^{\mathbb{N}_\mathbb{P} \times \mathbb{N}}$ represents the pilot placement matrix. If the equivalent transmitted signal is denoted as $x(0), x(1), \cdots, x(N-1)$, then the received signal vector can be formulated as

$$\boldsymbol{y} = \boldsymbol{XH} + \boldsymbol{v} = \boldsymbol{XFh} + \boldsymbol{v}, \tag{4}$$

where $\boldsymbol{y} \triangleq [y(0), y(1), \cdots, y(N-1)]^T$ is the received channel measurements, $\boldsymbol{X}$ is an $N \times N$ diagonal matrix denoted as $\boldsymbol{X} \triangleq \text{diag}\{x(0), x(1), \cdots, x(N-1)\}$ which represents the equivalent transmitted signals on the main diagonal, $\boldsymbol{H} \triangleq [H(0), H(1), \cdots, H(N-1)]^T$ is a vector of sampled channel frequency response, $\boldsymbol{v} \triangleq [v(0), v(1), \cdots, v(N-1)]^T \sim \mathcal{CN}(0, \sigma_v^2 I_N)$ is the additive white Gaussian noise (AWGN) experienced by the channel, and $\boldsymbol{h} \triangleq [h(0), h(1), \cdots, h(L-1)]^T$ is the k-sparse received baseband CIR with length L, F is an $N \times L$ partial discrete Fourier transform (DFT) submatrix which includes only the first L columns of a standard $N \times N$ DFT submatrix whose (m, n)th element of F is given by $[F]_{m,n} = \frac{1}{\sqrt{N}} e^{-j2\pi mn/N}$, where $0 \leq m \leq N-1$ and $0 \leq n \leq L-1$. Therefore, the received signal at pilot location is expressed as

$$\boldsymbol{y}(\boldsymbol{p}) = \boldsymbol{X}(\boldsymbol{p})\boldsymbol{F}(\boldsymbol{p})\boldsymbol{h} + \boldsymbol{v}(\boldsymbol{p}) = \boldsymbol{Ah} + \boldsymbol{v}(\boldsymbol{p}), \tag{5}$$

where $\boldsymbol{y}(\boldsymbol{p}) \triangleq \boldsymbol{Wy} \triangleq [y(p_1), y(p_2), \cdots, y(p_{N_P})]^T$ represents the received channel measurements at the pilot placement set $\boldsymbol{p}$, $\boldsymbol{X}(\boldsymbol{p}) \triangleq \boldsymbol{WXW}^T \triangleq \text{diag}\{x(p_1), x(p_2), \cdots, x(p_{N_P})\}$ represents the transmitted signal matrix at the pilot placement set $\boldsymbol{p}$, $\boldsymbol{v}(\boldsymbol{p}) \triangleq \boldsymbol{Wv} \triangleq [v(p_1), v(p_2), \cdots, v(p_{N_P})]^T \sim \mathcal{CN}(0, \sigma_v^2 I_{N_P})$ represents the AWGN at pilot placement set $\boldsymbol{p}$, $\boldsymbol{F}(\boldsymbol{p}) \triangleq \boldsymbol{WF}$ represents an $N_P \times L$ DFT submatrix which includes only the rows that corresponds to the pilot placement set $\boldsymbol{p}$ and the first L columns of a standard $N \times N$ DFT submatrix and whose (m, n)th element of $\boldsymbol{F}$ is given by $[\boldsymbol{F}]_{m,n} = \frac{1}{\sqrt{N}} e^{-j2\pi mn/N}$, where $0 \leq m \leq N_P - 1$ and $0 \leq n \leq L-1$ while $\boldsymbol{A} \triangleq \boldsymbol{X}(\boldsymbol{p})\boldsymbol{F}(\boldsymbol{p})$ represents the designed $N_P \times L$ measurement matrix.

According to (5), the objective of CE is to obtain an estimate of the CIR, $\boldsymbol{h}$ by using both the measurement matrix $\boldsymbol{A}$ and the very small set of linear measurements $\boldsymbol{y}(\boldsymbol{p})$. In order to reduce the overhead of pilot transmission, the measurement matrix with less number of rows (pilot signals, N_P) than columns (channel coefficients, L) is considered i.e., $N_P < L$. Consequently, the

matrix $\boldsymbol{A}$ becomes ill-conditioned and the resulting system of linear equations becomes underdetermined. Hence, the method of LS which is capable of accurately estimating $\boldsymbol{h}$ under the condition of $N_P > L$, now becomes inaccurate. However, through CS, $\boldsymbol{h}$ is accurately estimated and the originally transmitted signal is uniquely reconstructed on a condition that the target signal is k-sparse.

In CS, robust recovery guarantee of the sparse vectors, $\boldsymbol{h}$ is possible from noisy measurements if the matrix $\boldsymbol{A}$ satisfies the RIP. This property is presented as follows

***Definition 1** (RIP [18])*: A matrix $\boldsymbol{A} \in \mathbb{R}^{\mathbb{N}_{\mathbb{P}} \times \mathbb{L}}$ satisfies RIP with restricted isometry constant (RIC) of δ_k, if

$$(1-\delta_k)\|\boldsymbol{h}\|_2^2 \leq \|\boldsymbol{A}\boldsymbol{h}\|_2^2 \leq (1+\delta_k)\|\boldsymbol{h}\|_2^2 \tag{6}$$

holds for all k-sparse vectors $\boldsymbol{h} \in \mathbb{R}^{\mathbb{L}}$ such that $\|\boldsymbol{h}\|_0 \leq k$, and $0 \leq \delta_k < 1$. Equivalently, for all k-sparse vectors $\boldsymbol{h} \in \mathbb{R}^{\mathbb{L}}$ with $\|\boldsymbol{h}\|_2 = 1$, (6) may be reformulated as $(1-\delta_k) \leq \|\boldsymbol{A}\boldsymbol{h}\|_2^2 \leq (1+\delta_k)$. Hence, the smaller the k-th RIP parameter of $\boldsymbol{A}$, i.e., $\delta_k(\boldsymbol{A})$, the better RIP is for that k value. In other words, RIP measures the "overall conditioning" of the set of $N_P \times k$ submatrices of $\boldsymbol{A}$ [37], [38]. The following Lemma shows that RIC satisfies the monotonicity property.

***Lemma 1** (Monotonicity of δ_k [27])*: If the measurement matrix satisfies the RIP of orders $2k$ and $3k$, respectively, then

$$\delta_k \leq \delta_{2k} \leq \delta_{3k} \tag{7}$$

for any $\delta_{2k} \leq \delta_{3k}$. This property is called the monotonicity of the RIC.

III. COLLABORATIVE FRAMEWORK OF ALGORITHMS

In this section, a new sparse recovery scheme, namely CoFA, is proposed to reconstruct the unknown sparse signal from far fewer linear measurements. In what first follow, is an exploratory investigation of the viability of the proposed scheme, followed by a detailed description of the proposed scheme. Lastly, the theoretical performance analysis of the proposed scheme is presented.

A. Exploratory Investigation

Consider an OFDM system with a data sequence modulated by 4QAM with $N = 256$ subcarriers, where N_P subcarriers are assumed to be pilots subcarriers. The length L of the sparse Rayleigh multipath fading channel $\boldsymbol{h}$, is modeled with $L = 50$ taps, where $k = 4$ positions are randomly selected as nonzero channel taps that are independently and identically distributed (i.i.d.) $\mathcal{CN}(0,1)$. Hence, these nonzero entries corresponds to $\mathcal{T} = \{c_1, c_2, \cdots, c_k\} \subseteq \{1 \leq c_1 < c_2 < \cdots < c_k \leq L\}$ columns of $\boldsymbol{A} \in \mathbb{R}^{\mathbb{N}_{\mathbb{P}} \times \mathbb{L}}$ which represents the true support set of the sparse signal. For the recovery of the sparse channel, two GP reconstruction algorithms were used, the OMP [21] and the ROMP [26] and averaged over 5,500 sparse channel realizations. The simulations are performed using MATLAB v8.5 (Release 2015a) on a PC Workstation equipped with Intel Core i5-4460 CPU at 3.20 GHz with 4 GB installed memory (RAM).

Table 2. Average number of correctly estimated atoms $N = 256$, $k = 4$, $L = 50$.

N_P	8	10	12	14	16
$\texttt{Avg}\lvert\hat{\mathcal{T}}_{(OMP)}\rvert$	1.8	2.4	3.2	3.8	3.9
$\texttt{Avg}\lvert\hat{\mathcal{T}}_{(ROMP)}\rvert$	1.7	2.1	2.4	3.1	3.8
$\texttt{Avg}\lvert\bar{\mathcal{T}}^{\min}\rvert$	1.9	2.6	3.4	4.0	4.4
$\texttt{Avg}\lvert\bar{\mathcal{T}}^{\max}\rvert$	**2.5**	**3.1**	**3.9**	**4.7**	**5.3**

The conceptual framework of this exploratory investigation is such that, in each sparse channel realization, each algorithm compares their estimated support set $\hat{\mathcal{T}}$ with the true support set of the sparse signal $\mathcal{T}$ (i.e., $|b| = \hat{\mathcal{T}} \cap \mathcal{T}\{\boldsymbol{x} : \boldsymbol{x} \in \hat{\mathcal{T}} \text{ and } \boldsymbol{x} \in \mathcal{T}\}$, where $|b|$ is the cardinality of the elements of the intersection $\hat{\mathcal{T}} \cap \mathcal{T}$. The number of true atoms detected are then average over $t = 5500$ sparse channel realizations i.e., $\texttt{avg}|b| = (\sum_i^t |b_i|)/t$.

Two new functions are defined that finds support of a vector, as follows

$$\texttt{supp}(\hat{\boldsymbol{h}}, \hat{T}, \beta) \triangleq \{\text{the subset of indices in set } \hat{T} \text{ that corresponds to } \beta \text{ largest magnitude of the estimated sparse coefficients in } \hat{\boldsymbol{h}}\}$$

and

$$\texttt{supp}(\hat{\boldsymbol{h}}, \hat{T}, \ddot{\beta}) \triangleq \{\text{the subset of indices in set } \hat{T} \text{ that corresponds to } \beta \text{ smallest magnitude of the estimated sparse coefficients in } \hat{\boldsymbol{h}}\}.$$

Let $\hat{\mathcal{T}}_{(OMP)}$ and $|\hat{\mathcal{T}}_{(OMP)}|$ and then $\hat{\mathcal{T}}_{(ROMP)}$ and $|\hat{\mathcal{T}}_{(ROMP)}|$ represent the estimated support-set and cardinality of OMP and ROMP, respectively. In the same simulation environment, a framework of collaboration between OMP and ROMP is investigated in two different scenarios. Let $\hat{\boldsymbol{h}}_{(ROMP)}$ represent the output vector of estimated sparse coefficients of ROMP which has k nonzero elements corresponding to $\hat{\mathcal{T}}_{(ROMP)}$. Let $\tilde{\mathcal{T}}_{(ROMP)} = \hat{\mathcal{T}}_{(ROMP)} \setminus \hat{\mathcal{T}}_{(OMP)} = \{\boldsymbol{x} : \boldsymbol{x} \in \hat{\mathcal{T}}_{(ROMP)} \textit{ and } \boldsymbol{x} \notin \hat{\mathcal{T}}_{(OMP)}\}$ represent the indices in ROMP but not in OMP. As an illustration, we consider a particular case where $\beta = 2$. Let $\tilde{\mathcal{T}}^{\max}_{(ROMP)} = \texttt{supp}(\hat{\boldsymbol{h}}_{(ROMP)}, \tilde{\mathcal{T}}_{(ROMP)}, \beta)$, and $\tilde{\mathcal{T}}^{\min}_{(ROMP)} = \texttt{supp}(\hat{\boldsymbol{h}}_{(ROMP)}, \tilde{\mathcal{T}}_{(ROMP)}, \ddot{\beta})$, represent the subset of indices in set $\tilde{\mathcal{T}}_{(ROMP)}$ that corresponds to $\beta = 2$ largest and smallest magnitudes of $\hat{\boldsymbol{h}}_{(ROMP)}$, respectively. Let $\breve{\mathcal{T}}^{\max} = \hat{\mathcal{T}}_{(OMP)} \cup \tilde{\mathcal{T}}^{\max}_{(ROMP)}$ and $\breve{\mathcal{T}}^{\min} = \hat{\mathcal{T}}_{(OMP)} \cup \tilde{\mathcal{T}}^{\min}_{(ROMP)}$ represent the framework of collaboration between OMP and ROMP based on the two different scenarios. Consequently, $|\bar{\mathcal{T}}^{\max}| = \breve{\mathcal{T}}^{\max} \cap \mathcal{T}$ and $|\bar{\mathcal{T}}^{\min}| = \breve{\mathcal{T}}^{\min} \cap \mathcal{T}$ represent the cardinality of the column indices that were correctly identified for the two different scenarios, respectively. The results of this experiment is presented in Table 2. The simulation results in Table 2 shows that, the average number of atoms correctly identified by OMP is 3.02 and by ROMP is 2.62. Additionally, the average atom identified by $\texttt{Avg}|\bar{\mathcal{T}}^{\min}| = 3.26$, and $\texttt{Avg}|\bar{\mathcal{T}}^{\max}| = 3.90$. Interestingly, while ensuring that too many incorrect indices were not merged in the union of support sets (i.e., by admitting only β number of promising indices from the participating algorithm, ROMP), the set $\breve{\mathcal{T}}^{\max} = \hat{\mathcal{T}}_{(OMP)} \cup \tilde{\mathcal{T}}^{\max}_{(ROMP)}$ dominated by its ability to find the true indices of the target signal. Hence, this collaborative framework between algorithms

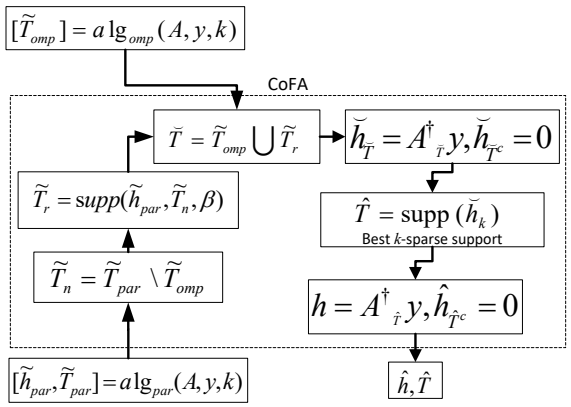

Fig. 1. Schematics block diagram of the CoFA scheme.

will be exploited in the following section and may possibly lead to accurate reconstruction of sparse signals from noisy CS measurements.

B. Algorithm Description

The CoFA scheme is proposed to pursue accurate sparse signal recovery by identifying the true support set of the target signal. Since, OMP can reliably recover any k-sparse signal of dimension N using $N_P = 0(k\ln N)$ limited number of linear measurements and have comparable performance guarantee to that of BP [23], [39], it has become an attractive alternative to BP [23]. However, OMP may fail to identify a correct index at a fixed iteration if the cumulative coherence of the measurement matrix is too similar [40], i.e. if the measurement matrix coherence is too high. Therefore, many revolutionary GP algorithms have been proposed to improve the identification step of OMP. This is also revisited, as a special case of the CoFA scheme, for accurate support recovery of the high-dimensional sparse signal. Hence, the two main ingredients of CoFA are the pre-selection wherein, the estimated support-set of OMP is selected and the thresholding–to remove the unpromising indices from the estimated support set of any participating algorithm before fusion (i.e., $\breve{T}_r = \texttt{supp}(\tilde{\boldsymbol{h}}_{par}, \breve{T}, \beta)$ such that, indices in set $\breve{T}$ that corresponds to the β largest coefficients of $\breve{\boldsymbol{h}}_{par}$ yield a revised estimate of support $\breve{T}_r$). The schematic block diagrame of the CoFA scheme is presented as Fig. 1.

The pseudo-code for the CoFA scheme is shown as Algorithm 1. Define $\tilde{T}_{omp}$ and $\tilde{T}_{par}$, as the estimated support set of OMP and any participating algorithm, respectively. Let $\beta = \lceil \alpha |\tilde{T}_{par}| \rceil = \lceil \alpha k \rceil$, where α takes values defined by $0.1 \leq \alpha \leq 1$. In other words, $\alpha \in [\Lambda]$, such that $[\Lambda] := \texttt{linspace}(0.1, 1, k)$ generates a row vector consisting of k linearly equally spaced points within the range of 0.1 and 1, in the form $[\Lambda] := \{\lambda_1, \lambda_2, \cdots, \lambda_k\}$. Here, k denotes the sparsity level of the true signal. Hence, β can take on any value within $\lceil \lambda_1 \cdot k \rceil$, $\lceil \lambda_2 \cdot k \rceil, \cdots, \lceil \lambda_k \cdot k \rceil$. Since β is dependent on parameters such as the dimension of the sparse signal, sparsity level k, measurement noise level, and the number of measurements N_P, then proper value of α must be selected to determine the thresholding parameter β.

Algorithm 1 CoFA Scheme

Input: Received signal at pilot subcarriers $\boldsymbol{y}_p$, measurement matrix $\boldsymbol{A}$, sparsity level k, parameter α;

Output: Estimated sparse channel vector $\hat{\boldsymbol{h}}$, estimated sparse support set $\hat{T}$

Initialization: $\hat{\boldsymbol{h}} = [0, 0, \cdots, 0]^T$;

1: $[\tilde{T}_{omp}] = \texttt{alg}(\boldsymbol{A}, \boldsymbol{y}_p, k)$; {Estimated support set for OMP}

2: $[\tilde{\boldsymbol{h}}_{par}, \tilde{T}_{par}] = \texttt{algPar}(\boldsymbol{A}, \boldsymbol{y}_p, k)$; {$\tilde{\boldsymbol{h}}_{par}$ and $\tilde{T}_{par}$ represent the estimated channel coefficients and the corresponding support-set of any participating algorithm, respectively}

3: $\beta = \lceil \alpha k \rceil$; {$\alpha \in [\Lambda]$, and $[\Lambda] := \texttt{linspace}(0.1, 1, k)$}

4: $\tilde{T}_n = \tilde{T}_{par} \backslash \tilde{T}_{omp}$; {$\boldsymbol{x} : \boldsymbol{x} \in \tilde{T}_{par}$ and $\boldsymbol{x} \notin \tilde{T}_{omp}$}

5: $\tilde{T}_r = \texttt{supp}(\tilde{\boldsymbol{h}}_{par}, \tilde{T}_n, \beta)$; {$\tilde{T}_r \subset \tilde{T}_n$ i.e., $\tilde{T}_r$ contains indices in $\tilde{T}_n$ corresponding to β largest magnitude elements in $\hat{\boldsymbol{h}}_{par}$ }

6: $\breve{T} = \tilde{T}_{omp} \cup \tilde{T}_r$;

7: $\breve{\boldsymbol{h}}_{\breve{T}} = \boldsymbol{A}^{\dagger}_{\breve{T}} \boldsymbol{y}_p, \breve{\boldsymbol{h}}_{\breve{T}^c} = 0$;

8: $\hat{T} = \texttt{supp}(\breve{\boldsymbol{h}}_k)$; {$\breve{\boldsymbol{h}}_k$ is the best k-sparse approximation of $\breve{\boldsymbol{h}}$}

9: $\hat{\boldsymbol{h}}_{\hat{T}} = \boldsymbol{A}^{\dagger}_{\hat{T}} \boldsymbol{y}_p, \hat{\boldsymbol{h}}_{\hat{T}^c} = 0$;

10: **return** $\hat{\boldsymbol{h}}, \hat{T}$.

In order to identify the support set of the underlying sparse signal, the estimated support sets of OMP and that of any other participating algorithm are first generated in the form of $[\tilde{T}_{omp}] = \texttt{alg}(\boldsymbol{A}, \boldsymbol{y}_p, k)$ and $[\tilde{\boldsymbol{h}}_{par}, \tilde{T}_{par}] = \texttt{algPar}(\boldsymbol{A}, \boldsymbol{y}_p, k)$, respectively. The notation $\tilde{T}_{omp}$, denotes the estimated support set of OMP, while $\tilde{\boldsymbol{h}}_{par}$ and $\tilde{T}_{par}$ denote the estimated sparse coefficients and support-set of any participating algorithm. The support set of OMP is then pre-selected for inclusion into the union of support set. Subsequently, the set difference of $\tilde{T}_{omp}$ and $\tilde{T}_{par}$, i.e., $\tilde{T}_n = \tilde{T}_{par} \backslash \tilde{T}_{omp}$ $\{\boldsymbol{x} : \boldsymbol{x} \in \tilde{T}_{par}$ and $\boldsymbol{x} \notin \tilde{T}_{omp}\}$, which represents the set of elements in $\tilde{T}_{par}$ but not in $\tilde{T}_{omp}$ is generated. And then $\tilde{T}_r = \texttt{supp}(\hat{\boldsymbol{h}}_{par}, \tilde{T}_n, \beta)$, which yields a revised estimate of the support set of the participating algorithm, where $\tilde{T}_r$ represents the indices in $\tilde{T}_n$ that corresponds to β largest magnitude of $\hat{\boldsymbol{h}}_{par}$. Subsequently, a union of indices is formed, i.e., $\breve{T} = \tilde{T}_{omp} \cup \tilde{T}_r$. Following the identification of the sparse approximation supports, $\breve{T}$, a new estimate of the CIR, $\tilde{\boldsymbol{h}}_{\breve{T}}$ is obtained. Lastly, the k largest values of the estimated channel coefficient are maintained and a new estimate of the sparse coefficient is generated, $\hat{\boldsymbol{h}}$ while the remainder are zero padded, i.e., $\hat{\boldsymbol{h}}_{\hat{T}^c} = 0$.

C. Theoretical Analysis of CoFA Scheme in the CS Framework

In this section, the theoretical analysis of the CoFA scheme is presented using the RIP of the measurement matrix $\boldsymbol{A}$. Let $|\tilde{T}_{omp}| = |\tilde{T}_{par}| = k$ be the support set cardinality of OMP and that of any participating algorithm, respectively. In the CoFA scheme, $\breve{T} = \tilde{T}_{omp} \cup \tilde{T}_r$ denotes the union set formed by OMP and the revised estimate of support set of any participating algorithm. Consequently, $|\breve{T}| = k + \lceil \alpha k \rceil = k + \beta$, since $|\tilde{T}_{omp}| = k$

and $|\tilde{T}_r| = \beta$, where $(k+\beta) \leq N_P$. Subsequently, the LS solution is then obtained $\boldsymbol{h}_{\hat{T}} = \arg\min_u \|\boldsymbol{y}_p - \boldsymbol{A}_{\hat{T}} u\|_2 = \boldsymbol{A}_{\hat{T}}^{\dagger} \boldsymbol{y}_p$ which furnishes a k-sparse signal estimate, $\boldsymbol{h}$ with a corresponding support set $\hat{T}$

Hence, since the matrix $\boldsymbol{A}$ satisfies the k-RIP with RIC of $\delta_{k+\beta}$, then the $N_P \times L$ dimensional matrix $\boldsymbol{A}$ satisfies RIP with RIC $\delta_{k+\beta}$ if

$$(1-\delta_{k+\beta})\|\boldsymbol{h}\|_2^2 \leq \|\boldsymbol{A}\boldsymbol{h}\|_2^2 \leq (1+\delta_{k+\beta})\|\boldsymbol{h}\|_2^2 \tag{8}$$

holds for all $(k+\beta) \leq N_P$ sparse vector $\boldsymbol{h}$, where $0 \leq \delta_{k+\beta} < 1$. Since matrix $\boldsymbol{A}$ has RIC $\delta_{k+\beta}$, then according to Proposition 3.1 in [15], it holds that

$$\begin{aligned}
\|\boldsymbol{A}_{\hat{T}}^{*}\boldsymbol{y}\|_2 &\leq \sqrt{1+\delta_{k+\beta}}\|\boldsymbol{y}\|_2 \\
\|\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{y}\|_2 &\leq \frac{1}{\sqrt{1-\delta_{k+\beta}}}\|\boldsymbol{y}\|_2 \\
(1-\delta_{k+\beta})\|\boldsymbol{h}\|_2 &\leq \|\boldsymbol{A}_{\hat{T}}^{*}\boldsymbol{A}_{\hat{T}}\boldsymbol{h}\|_2 \leq (1+\delta_{k+\beta})\|\boldsymbol{h}\|_2 \\
\frac{1}{1+\delta_{k+\beta}}\|\boldsymbol{h}\|_2 &\leq \|(\boldsymbol{A}_{\hat{T}}^{*}\boldsymbol{A}_{\hat{T}}\|)^{-1}\boldsymbol{h}\|_2 \leq \frac{1}{1-\delta_{k+\beta}}\|\boldsymbol{h}\|_2,
\end{aligned} \tag{9}$$

where $(\boldsymbol{A})^{\dagger} = (\boldsymbol{A}^{*}\boldsymbol{A})^{-1}\boldsymbol{A}^{*}$ (for complex $\boldsymbol{A}$ with linearly independent columns, assuming $\boldsymbol{A}$ is full-rank). Hence, from Corollary 3.3 in [15], given that $\boldsymbol{A}$ has RIC of $\delta_{k+\beta}$ with a set of indices, $\hat{T}$, and with $\boldsymbol{h}$ as a vector of CIR, it then follows that

$$\|\boldsymbol{A}_{\hat{T}}^{*}\boldsymbol{A}\boldsymbol{h}_{\hat{T}^c}\|_2 = \|\boldsymbol{A}_{\hat{T}}^{*}\boldsymbol{A}_{\hat{T}^c}\boldsymbol{h}_{\hat{T}^c}\|_2 \leq \delta_{k+\beta}\|\boldsymbol{h}_{\hat{T}^c}\|_2 \tag{10}$$

provided that $k+\beta \geq |\hat{T} \cup \mathtt{supp}(\boldsymbol{h})|$. Accordingly, the following lemma can be shown

Lemma 2: For the CS framework stated in (4), if the union of identified support-set is given as $\breve{T}$ and $\hat{\boldsymbol{h}}$ is the estimated channel coefficient of $\boldsymbol{h}$ derived via $\hat{\boldsymbol{h}}_{\breve{T}} = \boldsymbol{A}_{\breve{T}}^{\dagger}\boldsymbol{y}, \hat{\boldsymbol{h}}_{\hat{T}^c} = 0$, then the following relation holds

$$\|\boldsymbol{h} - \hat{\boldsymbol{h}}\|_2 \leq \frac{\delta_{2k+\beta}}{1-\delta_{2k+\beta}}\|\boldsymbol{h}_{\hat{T}^c}\|_2 + \frac{1}{\sqrt{1-\delta_{2k+\beta}}}\|\boldsymbol{v}\|_2, \tag{11}$$

from this it can be deduced further that

$$\|\boldsymbol{h}_{\hat{T}^c}\|_2 \leq \|\boldsymbol{h} - \hat{\boldsymbol{h}}\|_2,$$

and that

$$\|\boldsymbol{h}_{\hat{T}^c}\|_2 = \frac{1+\delta_{2k+\beta}}{1-\delta_{2k+\beta}}\|\boldsymbol{h}_{\breve{T}^c}\|_2 + \frac{2}{1-\delta_{2k+\beta}}\|\boldsymbol{v}\|_2.$$

This then provides a reconstruction guarantee of

$$\|\boldsymbol{h} - \hat{\boldsymbol{h}}\|_2 \leq \frac{1+\delta_{2k+\beta}}{(1-\delta_{2k+\beta})^2}\|\boldsymbol{h}_{\breve{T}^c}\|_2 + \frac{3}{(1-\delta_{2k+\beta})^2}\|\boldsymbol{v}\|_2.$$

Proof of Lemma 1 can be found in Appendix B. ■

Hence, according to the monotonicity property of RIC in (7), the CoFA scheme is seen to satisfy RIP with RIC $\delta_{2k+\beta}$, where $\beta = 2$. However, in the FACS scheme, the measurement matrix $\boldsymbol{A}$ satisfies RIP with RIC δ_{R+k}, where $R \triangleq |\Gamma| \leq N_P$ and $\Gamma = \cup_{i=1}^{P} \hat{T}_i$ represents the union of P number of estimated support-sets. Hence

$$\delta_{2k} \leq \delta_{2k+\beta} \leq \delta_{R+k}.$$

Finally, it is worth mentioning that if matrix $\boldsymbol{A}$ satisfies RIC of order δ_{2k}, $\delta_{2k+\beta}$, or δ_{R+k}, then it can be interpreted that, matrix $\boldsymbol{A}$ approximately preserves the distance between any pair of k-sparse vectors for the respective value of RIC. Therefore, the RIC of δ_{R+k} for FACS will indeed have significant implications regarding to its robustness to noise and therefore, FACS may not provide performance guarantee improvement in all cases.

IV. NUMERICAL EXPERIMENTS AND RESULTS

In this section, simulation experiments were performed to investigate and compare the performance of the proposed scheme with other estimation algorithms such as LS, OMP, CoSaMP, SP and the FACS scheme. The following simulation experiments were performed using MATLAB v8.5 (Release 2015a) on a PC Workstation equipped with Intel Core i5-4460 CPU at 3.20GHz with 4GB installed memory (RAM).

A. Simulation Set-up

An OFDM system with 4QAM baseband modulation, with $N = 256$ subcarriers is considered, where $N_P = 16$ subcarriers are assumed to be pilots, unless otherwise mentioned. For the recovery of sparse channel, 4,000 realizations of sparse multipath channels with $L = 50$ are employed in each settings, where $k = 6$ location of the nonzero channel taps are randomly generated and the channel attenuation is based on the independently and identically distributed (i.i.d.) $\mathcal{CN}(0,1)$. In each of the following experiments, the parameter $\alpha \in [\Lambda]$, with $[\Lambda] := \mathtt{linspace}(0.1, 1, k)$ is employed, such that $\alpha \in \{0.10, 0.28, 0.46, 0.64, 0.82, 1.00\}$. Hence, in applying the term to $\beta = \lceil \alpha |\tilde{T}_{pat}| \rceil$, a corresponding value of $\beta \in \{1,2,3,4,5,6\}$ indices of the participating algorithm will be admitted into the union of support set.

B. Performance of the CoFA Scheme

In the first experiment, the variation on the performance of the proposed scheme, namely CoFA with respect to signal-to-noise ratio (SNR) for different settings of β is studied, where $\beta = \lceil \alpha k \rceil$ and $\alpha \in [\Lambda]$. Two different randomly-generated pilot patterns were used as presented in Appendix-A with $N_p = 12$ and $N_p = 14$ number of pilot subcarriers, respectively. Fig. 2 plots the MSE versus SNR of the proposed scheme over the various settings of α for a corresponding value for β. Fig. 2 indicates how the CoFA scheme is robustness in terms of sparsity level, the number of linear measurements, measurement noise level and sparse signal dimension over various parameter settings of β. It is observed that using the different parameter settings of $\alpha \in \{0.10, 0.28, 0.46, 0.64, 0.82, 1.00\}$ strongly influences the overall system performance. However, the proposed scheme showed a better MSE performance with a setting of $\alpha = 0.28$ (i.e., $\beta = \lceil \alpha k \rceil = \lceil 0.28 \times k \rceil = 2$, where $k = 6$) for the two pilot patterns.

In the second experiment, CE performance is compared between the proposed scheme and the FACS scheme using

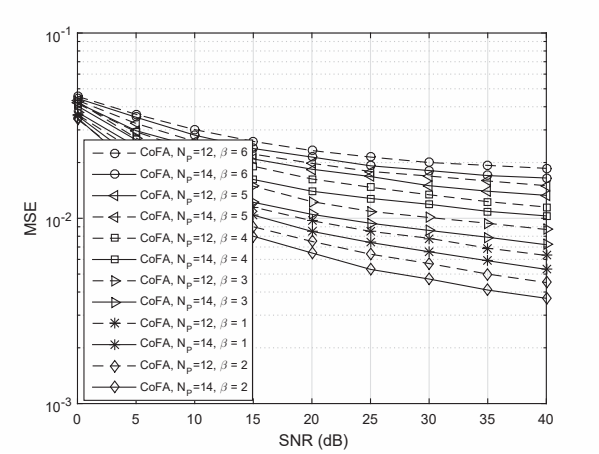

Fig. 2. MSE performance of the CoFA scheme with various number of pilots and settings of β.

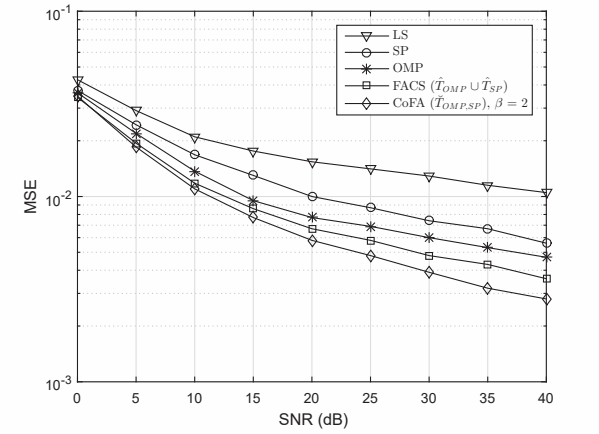

Fig. 3. MSE performance of the LS, SP, OMP algorithms and the FACS and CoFA schemes.

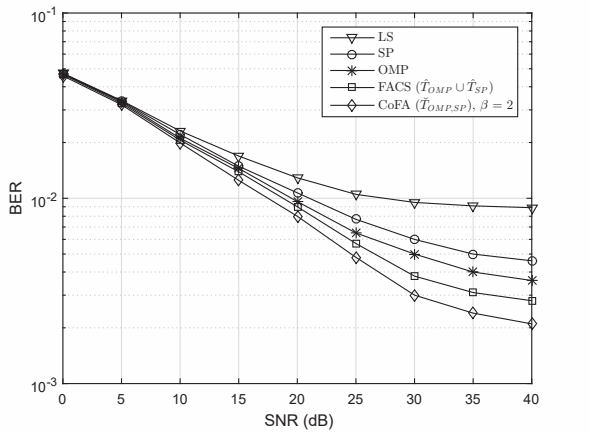

Fig. 4. BER performance of the LS, SP, OMP algorithms and the FACS and CoFA schemes.

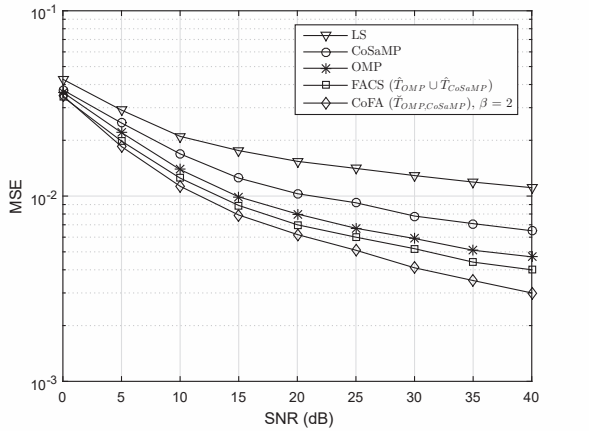

Fig. 5. MSE performance of the LS, CoSaMP, OMP algorithms and the FACS and CoFA schemes.

OMP and SP algorithms to develop CoFA($\breve{T}_{OMP,SP}$) and FACS($\hat{T}_{OMP} \cup \hat{T}_{SP}$) schemes, respectively. In this experiment, a randomly generated pilot pattern with $N_P = 16$ subcarriers was employed (see Appendix A). Additionally, the uniformly spaced pilot pattern was also used, as it proves to be the most effective approach for the conventional CE methods. The comparisons of MSE performance and the BER performance for sparse CE are shown in Figs. 3 and 4, respectively. It can be seen that the proposed scheme provides better MSE performance as well as system BER than the other methods of channel estimation including the FACS scheme. For example, at SNR = 30 dB the MSE for LS, SP, OMP, FACS and CoFA algorithms are 1.29×10^{-2}, 7.4×10^{-3}, 6.0×10^{-3}, 4.8×10^{-3}, and 3.9×10^{-3}, respectively. In other words, CoFA achieves a better MSE performance than FACS using a reduced number of fused support indices.

In the third experiment, CE performance is compared between the proposed scheme and the FACS scheme using OMP and CoSaMP algorithms to develop CoFA($\breve{T}_{OMP,CoSaMP}$) and FACS($\hat{T}_{OMP} \cup \hat{T}_{CoSaMP}$) schemes, respectively. The uniformly spaced pilot pattern is also plotted as the performance benchmark. The MSE and BER performance of CE are shown in Fig. 5 and Fig. 6, respectively. It can be seen that the proposed scheme has a significantly improved performance in terms of CE errors

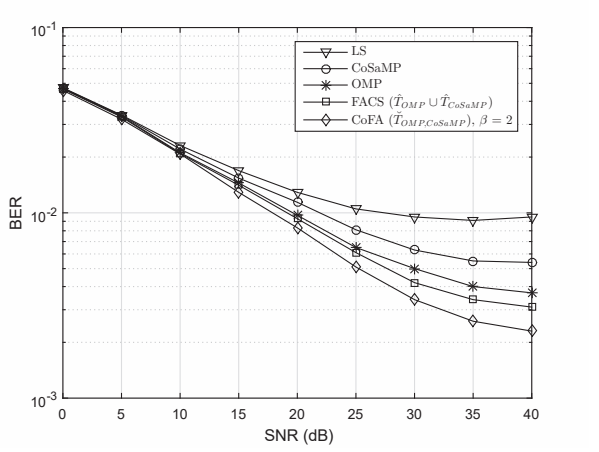

Fig. 6. BER performance of the LS, CoSaMP, OMP algorithms and the FACS and CoFA schemes.

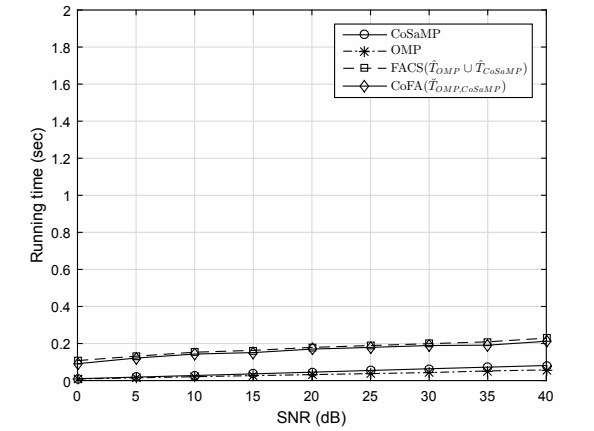

Fig. 7. CPU run time of the CoSaMP, OMP, FACS and CoFA algorithms for CE.

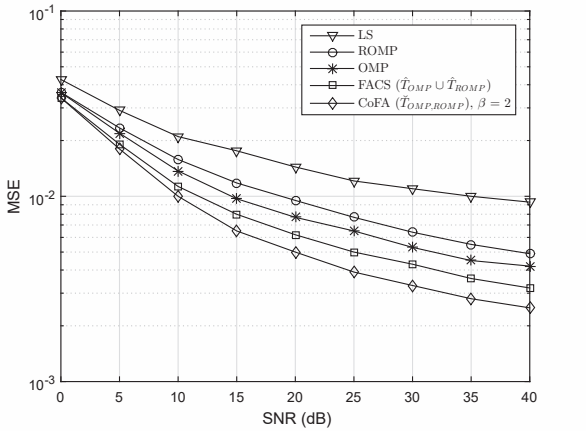

Fig. 9. MSE performance of the LS, OMP, ROMP algorithms and the FACS and CoFA schemes.

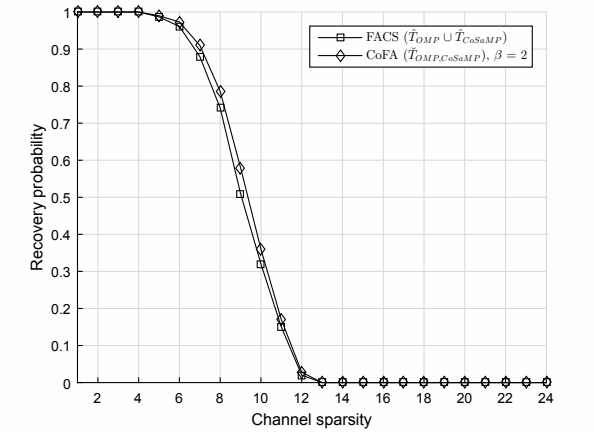

Fig. 8. Successful channel recovery percentage versus channel sparsity.

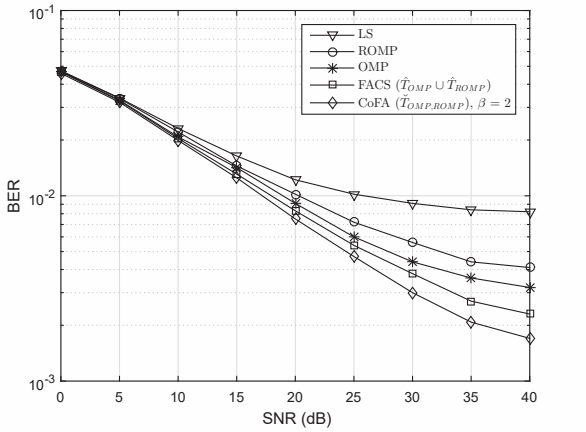

Fig. 10. BER performance of the LS, OMP, ROMP algorithms and the FACS and CoFA schemes.

compared to the FACS scheme. This improved performance by CoFA is best explained by its ability to find the true support indices of the sparse signal, while ensuring that too many incorrect indices are not fused. Additionally, Fig. 7 plots a rough estimate of the computational complexity of OMP, CoSaMP, FACS and CoFA algorithms based on the CPU run time. However, it is noteworthy to mention that, in spite of the fact that the CPU runtime is not an exact measure of complexity, it can still provide a rough estimate of the computational complexity. In Fig. 7 it is observed that, the computational complexity of CoSaMP is much higher than that of OMP. This can be further explained as the obvious effect of always selecting a fixed number of atoms, $2k$, with the largest correlation value in magnitude per iteration, unlike OMP which selects only one element from the correlation vector per iteration. Furthermore, FACS is observed to have a somewhat higher computational complexity than the CoFA scheme due to the fact that if fuses all indices to the union of support set (say Γ), and hence, requires an extra computation for the calculation of the LS approximation restricted to the union of support set via the Moore-Penrose pseudoinverse (i.e., $\tilde{\boldsymbol{h}}_{\Gamma} = \boldsymbol{A}_{\Gamma}^{\dagger}\boldsymbol{y}, \tilde{\boldsymbol{h}}_{\Gamma^c} = 0$).

The fourth experiment compares the sparse signal recovery probability with respect to the sparsity of the channel (i.e., the number of nonzero channel taps) for SNR $=$ 22 dB. Fig. 8 plots the sparse signal recovery probability with respect to the sparsity of the channel for both FACS($\hat{T}_{OMP} \cup \hat{T}_{CoSaMP}$) and CoFA($\breve{T}_{OMP,CoSaMP}$) schemes. It shows how the two schemes are robust against the sparsity of the channel. It can be observed that as the channel sparsity level k increases, the probability of successful channel recover decreases for both schemes. This is because, as the level of sparsity k increases, the number of required measurements $N_P = Ck\log(N/k)$, also increases. However, since the number of measurements (pilot subcarriers) used in this experiment, is kept fixed at $N_P = 16$ (where $N = 256$), at higher sparsity level k the available measurements becomes insufficient for perfect reconstruction of the sparse signal and results in a decrease in the successful channel recovery probability. However, in Fig. 8, it is observed that the proposed scheme is more robust and offered the best channel recovery probability against channel sparsity compared to the FACS scheme.

In the last experiment, CE performance is compared between the proposed scheme and the FACS scheme using the OMP and ROMP algorithms to develop the CoFA($\breve{T}_{OMP,ROMP}$) and the

FACS($\hat{T}_{OMP} \cup \hat{T}_{ROMP}$) schemes, respectively. The uniformly spaced pilot pattern is also plotted as the performance benchmark. The comparisons of the MSE performance and the BER performance for sparse CE are shown in Figs. 9 and 10, respectively. It can be seen from Figs. 9 and 10 that the use of the proposed scheme provides noticeable MSE and BER performance improvement over the FACS scheme.

V. CONCLUSION

In this paper, the fusion based signal recovery problem was investigated for sparsity-based CE technique in OFDM system and a new scheme, namely CoFA, was proposed to efficiently realize sparse signal reconstruction from far fewer linear measurements. The CoFA scheme uses a pair of algorithms beneath a collaborative framework, to correct erroneously estimated atoms chosen from the sensing basis, for the accurate reconstruction of the k-sparse signal. The performance guarantees of the CoFA scheme based on the RIP, has been examined. It was shown that the requirement on the RIC of the measurement matrix for guaranteeing exact/stable signal recovery from the compressed measurements is improved, compared to the FACS scheme. Additionally, the experimental results demonstrate that the CoFA scheme achieves a better MSE performance for the channel estimate, as well as the system BER when compared with the existing CE algorithms including the FACS scheme.

APPENDIX A

SPECIFIC PILOT PATTERN

The location of the randomly generated pilot pattern for $N_P = 12$ is set as $\{2, 9, 37, 68, 111, 170, 182, 205, 216, 226, 248, 255\}$. The location of the randomly generated pilot pattern for $N_P = 14$ is set as $\{1, 7, 15, 28, 35, 63, 74, 138, 199, 238, 242, 248, 250, 251\}$. The location of the randomly generated pilot pattern for $N_P = 16$ is set as $\{1, 6, 38, 76, 82, 89, 107, 128, 137, 154, 160, 189, 204, 222, 240, 255\}$. The location of the uniformly spaced pilot pattern for $N_P = 16$ is set as $\{1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256\}$

APPENDIX B

PROOF OF LEMMA 2

We note here that this lemma is as well employed in a somewhat similar form to that of Theorem 1 in [28], [15] and Appendix A in [27]. However despite these connections, we can obtain a more tighter error bounds with the CoFA scheme. To prove (11), we first consider

$$\begin{aligned}
\|\boldsymbol{h}-\hat{\boldsymbol{h}}\|_2 &\le \|\boldsymbol{h}_{\hat{T}} - \boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{y}\|_2 + \|\boldsymbol{h}_{\hat{T}^c}\|_2 \\
&= \|\boldsymbol{h}_{\hat{T}} - \boldsymbol{A}_{\hat{T}}^{\dagger}(\boldsymbol{A}_T\boldsymbol{h}_T + \boldsymbol{v})\|_2 + \|\boldsymbol{h}_{\hat{T}^c}\|_2 \\
&\le \|\boldsymbol{h}_{\hat{T}} - \boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{A}_T\boldsymbol{h}_T\|_2 + \|\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{v}\|_2 + \|\boldsymbol{h}_{\hat{T}^c}\|_2 \\
&= \|\boldsymbol{h}_{\hat{T}} - \boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{A}_{T\cap\hat{T}}\boldsymbol{h}_{T\cap\hat{T}} - \boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{A}_{\hat{T}^c}\boldsymbol{h}_{\hat{T}^c}\|_2 + \|\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{v}\|_2 \\
&\quad + \|\boldsymbol{h}_{\hat{T}^c}\|_2 \\
&= \|\boldsymbol{h}_{\hat{T}} - \boldsymbol{A}_{\hat{T}}^{\dagger}\begin{bmatrix}\boldsymbol{A}_{T\cap\hat{T}}\boldsymbol{A}_{\hat{T}-T}\end{bmatrix}\begin{bmatrix}\boldsymbol{h}_{T\cap\hat{T}} \\ \boldsymbol{0}\end{bmatrix} - \boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{A}_{\hat{T}^c}\boldsymbol{h}_{\hat{T}^c}\|_2 \\
&\quad + \|\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{v}\|_2 + \|\boldsymbol{h}_{\hat{T}^c}\|_2 \\
&= \|\boldsymbol{h}_{\hat{T}} - \boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{A}_{\hat{T}}\boldsymbol{h}_{\hat{T}} - \boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{A}_{\hat{T}^c}\boldsymbol{h}_{\hat{T}^c}\|_2 + \|\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{v}\|_2 \\
&\quad + \|\boldsymbol{h}_{\hat{T}^c}\|_2 \\
&= \|\boldsymbol{h}_{\hat{T}} - \boldsymbol{h}_{\hat{T}} - \boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{A}_{\hat{T}^c}\boldsymbol{h}_{\hat{T}^c}\|_2 + \|\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{v}\|_2 + \|\boldsymbol{h}_{\hat{T}^c}\|_2 \\
&= \|\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{A}_{\hat{T}^c}\boldsymbol{h}_{\hat{T}^c}\|_2 + \|\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{v}\|_2 + \|\boldsymbol{h}_{\hat{T}^c}\|_2. \qquad (12)
\end{aligned}$$

Note that the term $-\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{A}_{\hat{T}^c}\boldsymbol{h}_{\hat{T}^c}$ is replace with $\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{A}_{\hat{T}^c}\boldsymbol{h}_{\hat{T}^c}$ as there will be no effect in an l_2-norm. Having that $\boldsymbol{A}^{\dagger} = (\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*$, and that the cardinality of $\breve{T}$ is $k+\beta$, with $\boldsymbol{h}$ being k-sparse, then following Proposition 3.1 and Corollary 3.3 in [15], implies that

$$\begin{aligned}
\|\boldsymbol{h}-\hat{\boldsymbol{h}}\|_2 &\le \|(\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*\boldsymbol{A}_{\hat{T}^c}\boldsymbol{h}_{\hat{T}^c}\|_2 + \|(\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*\boldsymbol{v}\|_2 \\
&\quad + \|\boldsymbol{h}_{\hat{T}^c}\|_2 \\
&\le \left(1 + \frac{\delta_{(k+\beta)+k}}{1-\delta_{(k+\beta)+k}}\right)\|\boldsymbol{h}_{\hat{T}^c}\|_2 + \frac{\|\boldsymbol{v}\|_2}{\sqrt{1-\delta_{(k+\beta)+k}}} \\
&\le \frac{1}{1-\delta_{(k+\beta)+k}}\|\boldsymbol{A}^*\boldsymbol{A}_{\hat{T}^c}\boldsymbol{h}_{\hat{T}^c}\|_2 + \frac{\|\boldsymbol{v}\|_2}{\sqrt{1-\delta_{(k+\beta)+k}}} \\
&\le \frac{\delta_{(k+\beta)+k}}{1-\delta_{(k+\beta)+k}}\|\boldsymbol{h}_{\hat{T}^c}\|_2 + \frac{1}{\sqrt{1-\delta_{(k+\beta)+k}}}\|\boldsymbol{v}\|_2 \\
&\le \frac{\delta_{2k+\beta}}{1-\delta_{2k+\beta}}\|\boldsymbol{h}_{\hat{T}^c}\|_2 + \frac{1}{\sqrt{1-\delta_{2k+\beta}}}\|\boldsymbol{v}\|_2. \qquad (13)
\end{aligned}$$

Additionally, having that

$$\|\boldsymbol{h}-\hat{\boldsymbol{h}}\|_2^2 = \|\boldsymbol{h}_{\hat{T}} - \boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{y}\|_2^2 + \|\boldsymbol{h}_{\hat{T}^c}\|_2^2 \ge \|\boldsymbol{h}_{\hat{T}^c}\|_2^2$$

by further simplification, could be reduced to

$$\|\boldsymbol{h}-\hat{\boldsymbol{h}}\|_2 \ge \|\boldsymbol{h}_{\hat{T}^c}\|_2.$$

To determine an upper bound for $\|\boldsymbol{h}_{\hat{T}^c}\|_2$, we consider the following definitions for notational convenience; let $\hat{T}_{\triangle} := \breve{T} \setminus \hat{T}$ be the set of indices contained in $\breve{T}$ that is pruned by Algorithm 1. We can perceive that $\hat{T} \subset \breve{T}$, subsequently, by applying $\hat{T}^c = \breve{T}^c \cup \hat{T}_{\triangle}$ we arrive at

$$\|\boldsymbol{h}_{\hat{T}^c}\|_2 \le \|\boldsymbol{h}_{\breve{T}^c}\|_2 + \|\boldsymbol{h}_{\hat{T}_{\triangle}}\|_2. \qquad (14)$$

Furthermore, consider the following

$$\begin{aligned}
\|(\tilde{\boldsymbol{h}}_{\breve{T}})_{\hat{T}_{\triangle}}\|_2 &= \|\boldsymbol{h}_{\hat{T}_{\triangle}} + (\tilde{\boldsymbol{h}}_{\breve{T}} - \boldsymbol{h}_{\breve{T}})_{\hat{T}_{\triangle}}\|_2 \\
&\ge \|\boldsymbol{h}_{\hat{T}_{\triangle}}\|_2 - \|(\tilde{\boldsymbol{h}}_{\breve{T}} - \boldsymbol{h}_{\breve{T}})_{\hat{T}_{\triangle}}\|_2.
\end{aligned}$$

By means of re-ordering the terms, we obtain

$$\begin{aligned}\|\boldsymbol{h}_{\check{T}_\triangle}\|_2 &\le \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\hat{T}_\triangle}\|_2 + \|(\tilde{\boldsymbol{h}}_{\check{T}} - \boldsymbol{h}_{\check{T}})_{\hat{T}_\triangle}\|_2 \\ &\le \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\hat{T}_\triangle}\|_2 + \|(\tilde{\boldsymbol{h}}_{\check{T}} - \boldsymbol{h}_{\check{T}})\|_2. \end{aligned} \tag{15}$$

Since, $\|(\tilde{\boldsymbol{h}}_{\check{T}})_T\|_2^2 - \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\hat{T}}\|_2^2 \le 0$ we then have

$$\begin{aligned}\|(\tilde{\boldsymbol{h}}_{\check{T}})_{\hat{T}_\triangle}\|_2^2 &= \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\hat{T}_\triangle}\|_2^2 + \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\hat{T}}\|_2^2 - \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\hat{T}}\|_2^2 \\ &= \|\tilde{\boldsymbol{h}}_{\check{T}}\|_2^2 - \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\hat{T}}\|_2^2 \\ &= \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\check{T}\setminus T}\|_2^2 + \|(\tilde{\boldsymbol{h}}_{\check{T}})_T\|_2^2 - \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\hat{T}}\|_2^2 \\ &\le \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\check{T}\setminus T}\|_2^2. \end{aligned}$$

Applying square root property we get

$$\begin{aligned}\|(\tilde{\boldsymbol{h}}_{\check{T}})_{\hat{T}_\triangle}\|_2 &\le \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\check{T}\setminus T}\|_2 \\ &= \|(\tilde{\boldsymbol{h}}_{\check{T}})_{\check{T}\setminus T} - \boldsymbol{h}_{\check{T}\setminus T}\|_2 \\ &= \|(\tilde{\boldsymbol{h}}_{\check{T}} - \boldsymbol{h})_{\check{T}\setminus T}\|_2 \\ &\le \|\tilde{\boldsymbol{h}}_{\check{T}} - \boldsymbol{h}_{\check{T}}\|_2. \end{aligned} \tag{16}$$

Merging (16) with (15) gives

$$\|\boldsymbol{h}_{\hat{T}_\triangle}\|_2 \le 2\|(\tilde{\boldsymbol{h}}_{\check{T}} - \boldsymbol{h}_{\check{T}})\|_2. \tag{17}$$

Simplifying the RHS of (17) gives

$$\begin{aligned}\|(\tilde{\boldsymbol{h}}_{\check{T}} - \boldsymbol{h}_{\check{T}})\|_2 &= \|\boldsymbol{h}_{\check{T}} - \boldsymbol{A}_{\check{T}}^{\dagger}(\boldsymbol{A}\boldsymbol{h} + \boldsymbol{v})\|_2 \\ &= \|\boldsymbol{h}_{\check{T}} - \boldsymbol{A}_{\check{T}}^{\dagger}(\boldsymbol{A}_T\boldsymbol{h}_T + \boldsymbol{v})\|_2 \\ &\le \|\boldsymbol{h}_{\check{T}} - \boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{A}_T\boldsymbol{h}_T\|_2 + \|\boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{v}\|_2 \\ &= \|\boldsymbol{h}_{\check{T}} - \boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{A}_{T\cap\check{T}}\boldsymbol{h}_{T\cap\check{T}} - \boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{A}_{\check{T}^c}\boldsymbol{h}_{\check{T}^c}\|_2 \\ &+ \|\boldsymbol{A}_{\hat{T}}^{\dagger}\boldsymbol{v}\|_2 \\ &= \|\boldsymbol{h}_{\check{T}} - \boldsymbol{A}_{\check{T}}^{\dagger}\begin{bmatrix}\boldsymbol{A}_{T\cap\check{T}}\boldsymbol{A}_{\check{T}-T}\end{bmatrix}\begin{bmatrix}\boldsymbol{h}_{T\cap\check{T}} \\ \boldsymbol{0}\end{bmatrix} \\ &- \boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{A}_{\check{T}^c}\boldsymbol{h}_{\check{T}^c}\|_2 + \|\boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{v}\|_2 \\ &= \|\boldsymbol{h}_{\check{T}} - \boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{A}_{\check{T}}\boldsymbol{h}_{\check{T}} - \boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{A}_{\check{T}^c}\boldsymbol{h}_{\check{T}^c}\|_2 + \|\boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{v}\|_2 \\ &= \|\boldsymbol{h}_{\check{T}} - \boldsymbol{h}_{\check{T}} - \boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{A}_{\check{T}^c}\boldsymbol{h}_{\check{T}^c}\|_2 + \|\boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{v}\|_2 \\ &= \|\boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{A}_{\check{T}^c}\boldsymbol{h}_{\check{T}^c}\|_2 + \|\boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{v}\|_2 \\ &= \|(\boldsymbol{A}_{\check{T}}^{*}\boldsymbol{A}_{\check{T}})^{-1}\boldsymbol{A}_{\check{T}}^{*}\boldsymbol{A}_{\check{T}^c}\boldsymbol{h}_{\check{T}^c}\|_2 + \|\boldsymbol{A}_{\check{T}}^{\dagger}\boldsymbol{v}\|_2 \\ &\overset{(a)}{\le} \frac{\delta_{(k+\beta)+k}}{1-\delta_{k+\beta}}\|\boldsymbol{h}_{\check{T}^c}\|_2 + \frac{1}{\sqrt{1-\delta_{k+\beta}}}\|\boldsymbol{v}\|_2 \\ &\le \frac{\delta_{(k+\beta)+k}}{1-\delta_{(k+\beta)+k}}\|\boldsymbol{h}_{\check{T}^c}\|_2 + \frac{1}{1-\delta_{(k+\beta)+k}}\|\boldsymbol{v}\|_2. \end{aligned} \tag{18}$$

Hence, $\delta_{k+\beta} \le \delta_{(k+\beta)+k}$, $0 \le \delta_{k+\beta} \le 1$. Substituting (17), and (18) in (14) gives

$$\begin{aligned}\|\boldsymbol{h}_{\hat{T}^c}\|_2 &\le \left(1 + \frac{2\delta_{(k+\beta)+k}}{1-\delta_{(k+\beta)+k}}\right)\|\boldsymbol{h}_{\check{T}^c}\|_2 \\ &+ \frac{2}{1-\delta_{(k+\beta)+k}}\|\boldsymbol{v}\|_2 \\ &= \frac{1+\delta_{(k+\beta)+k}}{1-\delta_{(k+\beta)+k}}\|\boldsymbol{h}_{\check{T}^c}\|_2 + \frac{2}{1-\delta_{(k+\beta)+k}}\|\boldsymbol{v}\|_2 \\ &= \frac{1+\delta_{2k+\beta}}{1-\delta_{2k+\beta}}\|\boldsymbol{h}_{\check{T}^c}\|_2 + \frac{2}{1-\delta_{2k+\beta}}\|\boldsymbol{v}\|_2. \end{aligned} \tag{19}$$

Combining (19) in (13) yields

$$\begin{aligned}\|\boldsymbol{h} - \hat{\boldsymbol{h}}\|_2 &\le \frac{1+\delta_{2k+\beta}}{(1-\delta_{2k+\beta})^2}\|\boldsymbol{h}_{\check{T}^c}\|_2 \\ &+ \left(\frac{1}{\sqrt{1-\delta_{2k+\beta}}} + \frac{2}{(1-\delta_{2k+\beta})^2}\right)\|\boldsymbol{v}\|_2 \\ &\le \frac{1+\delta_{2k+\beta}}{(1-\delta_{2k+\beta})^2}\|\boldsymbol{h}_{\check{T}^c}\|_2 + \frac{3}{(1-\delta_{2k+\beta})^2}\|\boldsymbol{v}\|_2. \end{aligned} \tag{20}$$

REFERENCES

[1] A. Weiss, A. Yeredor, and M. Shtaif, "Iterative symbol recovery for power-efficient DC-biased optical OFDM systems," *J. Light. Technol.,* vol. 34, no. 9, pp. 2331–2338, May 2016.

[2] X. He, R. Song, and W. P. Zhu, "Pilot allocation for distributed-compressed-sensing-based sparse channel estimation in MIMO-OFDM systems," *IEEE Trans. Veh. Tech.,* vol. 65, no. 5, pp. 2990–3004, May 2016.

[3] M. N. Suresh, G. Ananthi, S. J. Thiruvengadam, and V. Abhaikumar "Timing estimation algorithm for OFDM-based wireless systems," *IET Sci., Meas. Tech.*, vol. 11, no. 2, pp. 149–154, Mar. 2016.

[4] D. Lee, J. So, and S. R. Lee, "Power allocation and subcarrier assignment for joint delivery of unicast and broadcast transmissions in OFDM systems," *J. Comm. Net.*, vol. 18, no. 3, pp. 375–386, June 2016.

[5] N. Wang, G. Gui, Z. Zhang, T. Tang, and J. Jiang, "A novel sparse channel estimation method for multipath MIMO-OFDM systems," in *Proc. IEEE VTC Fall*, 2011.

[6] Y. Zhang, R. Venkatesan, O. A. Dobre, and C. Li, "Novel compressed sensing-based channel estimation algorithm and near-optimal pilot placement scheme," *IEEE Trans. Wireless Commun.,* vol. 15, no. 4, pp. 2590–2603, Apr. 2016.

[7] R. Mohammadian, A. Amini, and B. H. Khalaj, "Compressive sensing-based pilot design for sparse channel estimation in OFDM systems," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 4–7, Jan. 2017.

[8] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.

[9] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications, Cambridge University Press*, 2012.

[10] D. L. Donoho, I. Johnstone, and A. Montanari, "Accurate prediction of phase transitions in compressed sensing via a connection to minimax denoising," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3396–3433, June 2013.

[11] X. Ma, F. Yang, W. Ding, and J. Song, "Novel approach to design time-domain training sequence for accurate sparse channel estimation," *IEEE Trans. Broadcast.*, vol. 62, no. 3, pp. 512–520, Sept. 2016.

[12] X. Cheng, M. Wang, and S. Li, "Compressive sensing-based beamforming for millimeter-wave OFDM systems," *IEEE Trans. Commun.,* vol. 65, no. 1, pp. 371–386, Jan. 2017.

[13] R. Jose, S. K. Ambat, and K. V. S. Hari, "Low complexity joint estimation of synchronization impairments in sparse channel for MIMO-OFDM system," *AEU – Int. J. Electron. Commun.*, vol. 68, no. 2, pp. 151–157, 2014.

[14] E. J. Candès and M. A. Davenport, "How well can we estimate a sparse vector?," *Appl. Comput. Harmon. Anal.*, vol. 34, no. 2, pp. 317–323, Mar. 2013.

[15] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied Comput. Harmonic Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.
[16] A. S. Bandeira, E. Dobriban, D. G. Mixon, and W. F. Sawin, "Certifying the restricted isometry property is hard," *IEEE Trans. Signal Process.,* vol. 59, no. 6, pp. 3448–3450, June 2013.
[17] 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.
[18] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
[19] S. Bahmani, P. T. Boufounos, and B. Raj, "Learning model-based sparsity via projected gradient descent," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 2092–2099, Apr. 2016.
[20] J. Choi, "Sparse index multiple access for multi-carrier systems with precoding," *J. Commun. Netw.*, vol. 18, no. 3, pp. 439–445, June 2016.
[21] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, July 2011.
[22] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 310–316, June 2010.
[23] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[24] J. Wang, S. Kwon, and B. Shim, "Generalized orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6202–6216, Dec. 2012.
[25] S. S. Kwon, J. Wang, and B. Shim, "Multipath matching pursuit," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2986–3001, May 2014.
[26] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Found. Comput. Math.*, vol. 9, no. 3, pp. 317–334, June 2009.
[27] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.
[28] S. K. Ambat, S. Chatterjee, and K. V. S. Hari, "Fusion of algorithms for compressed sensing," *IEEE Trans. Signal Process.,* vol. 61, no. 14, pp. 3699–3704, July 2013.
[29] J. D. Blanchard, C. Cartis, J. Tanner, and A. Thompson, "Phase transitions for greedy sparse approximation algorithms," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 188–203, Mar. 2011.
[30] N. Hurley and S. Rickard, "Comparing measures of sparsity," *EEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4723–4741, Oct. 2009.
[31] M. J. Fadili, J. L. Starck, and L. Boubchir, "Morphological diversity and sparse image denoising," in *Proc. IEEE ICASSP*, 2007
[32] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, "Compressive sensing: From theory to applications, a survey," *J. Commun. Netw.,* vol. 15, no. 5, pp.443–456, Oct. 2013.
[33] C. Qi, G. Yue, L. Wu, Y. Huang, and A. Nallanathan "Pilot design schemes for sparse Channel Estimation in OFDM Systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1493–1505, Apr. 2015.
[34] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: Algorithms and applications," *IEEE Access*, vol. 3, pp. 490–530, May 2015.
[35] M. F. Duarte and Y. C. Eldar"Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, Apr. 2011.
[36] E. D. Livshitz and V. N. Temlyakov, "Sparse approximation and recovery by greedy algorithms," *IEEE Trans. Inf. Theory,* vol. 60, no. 7, pp. 3989–4000, July 2014.
[37] Y. Zhang, "Theory of compressive sensing via l_1-minimization: A non-RIP analysis and extensions," *J. Oper. Res. Soc. China,* vol. 1, no. 1, pp. 79–105, Mar. 2013.
[38] N. Lee, "MAP support detection for greedy sparse signal recovery algorithms in compressive sensing," *IEEE Trans. Sig. Proc.,* vol. 64, no. 19, pp. 4987–4999, Oct. 2016.
[39] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
[40] K. Schnass and P. Vandergheynst, "Dictionary preconditioning for greedy algorithms," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1994–2002, Oct. 2008.

Anthony Ngozichukwuka Uwaechia received both the B.Eng. and M.Sc. degrees in Electrical Engineering from Ahmadu Bello University, Zaria, in the year 2006 and 2013, respectively. He received the Nigeria-sao Tome and Principe Joint Development Postgraduate Scholarship Award during his M.Sc. degree programme. He worked with Lilleker Bro. Nigeria Ltd. as an Electrical Site Engineer in the year 2007-2010. Currently, he is a Graduate Teaching Assistant in the School of Electrical and Electronic Engineering, Universiti Sains Malaysia (USM), Malaysia, since June 2015. He is also a Postgraduate Student (Ph.D.) in USM. His research interests lie in the areas of communication and signal processing and particularly focusing in the area of digital signal processing, sparse channel estimation, channel measurement, inverse problem, sparse representation, sparsity, and the OFDM system.

Nor Muzlifah Mahyuddin received B.Eng. degree in Electric-Telecommunication from Universiti Teknologi Malaysia, in 2005, and M.Sc. degree in Electronics System Design Engineering from Universiti Sains Malaysia, in 2006. In addition, she also received a Ph.D. degree in Microelectronics System Design from Newcastle University, Newcastle upon Tyne, United Kingdom, in 2011. She is currently working as a Lecturer in Universiti Sains Malaysia, starting from March 2012. She has produced several papers on the topic of a low-swing signalling scheme and split-ring resonator design. Her current research interests are in the field of RF and microwave engineering, reliability, and signal integrity. The topic of interests include the modelling design of split-ring resonator in a high-performance application, the impact of variability on the design of microstrip-based circuits, and the power integrity in the high-performance circuits. Dr Nor Muzlifah Mahyuddin is currently a Member of IEEE and involved in the Communications Society (ComSoc). Subsequently, she is also a Member of IET and Professional Member of Association for Computing Machinery (ACM). She is also registered with Board of Engineers Malaysia (BEM).