IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 1, 2019, accepted March 9, 2019, date of publication March 12, 2019, date of current version April 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2904596

Spectrum-Efficient Distributed Compressed Sensing Based Channel Estimation for OFDM Systems Over Doubly Selective Channels

ANTHONY NGOZICHUKWUKA UWAECHIA AND NOR MUZLIFAH MAHYUDDIN, (Member, IEEE)
School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Nibong Tebal 14300, Malaysia

Corresponding author: Nor Muzlifah Mahyuddin (eemnmuzlifah@usm.my)

This work was supported in part by the Postdoctoral Fellowship Program of School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Malaysia, and in part by the Ministry of Education Malaysia through the Fundamental Research Grant Scheme under Grant FRGS 203/PELECT/6071373.

ABSTRACT In this paper, we deal with channel estimation (CE) for high-mobility orthogonal frequency division multiplexing (OFDM) systems. To make the numerous (unknown) estimation for the high-mobility OFDM systems practicable, the channels are assumed to be time- and frequency-selective—or doubly selective (DS) and approximated by a basis expansion model (BEM). As the DS channel requires the distributed acquisition of multiple correlated signals in the delay-Doppler channel domain, we proceed to estimate jointly sparse BEM coefficient vectors over a DS channel as against numerous channel coefficients. On account of channel time-variation, the resulting channel matrix in the frequency domain exhibits (approximately banded) pseudo-circular structure, which gives rise to a diagonally dominant yet full matrix rather than a diagonal matrix and thus induces inter-channel interference (ICI). On the premise of this observation, we propose a new pilot design scheme that identifies the optimal pilot placement and values for each pilot cluster to combat ICI. Furthermore, to obtain a channel estimator consistent with the jointly sparse delay-Doppler [i.e., two dimensional (2D)] channel model, an algorithm namely, distributed compressed sensing (DCS)-based stage determined matching pursuit (DCS-SdMP), is proposed. Our claims are supported by simulation results, which are obtained considering Jakes' channels with fairly high Doppler spreads, which show the superiority of the proposed schemes over other different methods of CE.

INDEX TERMS Channel estimation (CE), doubly selective (DS) channel, orthogonal frequency division multiplexing (OFDM), distributed compressed sensing (DCS), basis expansion model (BEM), pilot design, wireless communication.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) has emerged as an effective multicarrier transmission technique for wireless communications systems over frequency-selective fading channels [1]–[3]. Conventionally [4], OFDM has been used in several wireless communication system applications and adopted by various standards where channels are assumed to experience static (or- quasi-static) fading while effectively ignoring time selectivity. However, future high-mobility OFDM system communications [5], [6], such as express railroad trains [3], [7], digital video broadcasting (DVB), underwater acoustic (UWA) communication [8] and millimeter-wave (mmWave) communications [9] for fifth-generation (5G) wireless networks manage a fading operating environment of time- and frequency-selective—or doubly selective (DS)— fading. The effect of DS fading is due to high Doppler frequency spread (DFS) (i.e., frequency dispersion —time selectivity) and multipath effects (i.e., time dispersion —frequency selectivity) in wireless links [10]. For an OFDM system operating over DS channels, the resulting time-variation of channel impulse response (CIR) induces a large number of channel coefficients to be estimated —due to DFS [11], [12], compared to static or quasi-static channel

The associate editor coordinating the review of this manuscript and approving it for publication was Md Fazlul Kader.

scenarios. Therefore, the time-variation of the channel during one OFDM symbol duration leads to a loss of subcarrier orthogonality [13], [14], and results in power leakage among subcarriers, referred to as inter-carrier interference (ICI). In practice, this requires much more pilots for reliable CE than with an entirely frequency selective channel [5], [11], [13], [14]. Therefore, making the CE task over DS channel extremely challenging for coherent detection.

Several signal processing schemes have been proposed to effectively mitigate ICI and lower the error floor of the channel. Examples of such schemes include precoding [15], time-domain equalization [16], and time-frequency localization [17]. Recently, empirical investigations have demonstrated that high-mobility OFDM system channels encountered in practice tend to exhibit a sparse structure [5], especially at high signal space dimension −particularly the delay-Doppler domain [11], [18], [19]. Therefore, the sparse channel can be characterized with significantly fewer parameters than those required by the Nyquist theorem and can effectively reduce the overhead for CE. Consequently, to exploit the inherently low-dimensionality of the sparse channels [5], [11], [12], the application of compressed sensing (CS) in DS channels, were studied. Several basis expansion models (BEMs) have been proposed for the channel time variations within one OFDM symbol [20], to approximate and simplify the time-varying channel parameters to be estimated. Besides several other candidates basis functions, such as cosine functions [21] and polynomials (i.e., Legendre and Chebyshev polynomials) [22], the set of complex exponentials (Fourier) basis functions [23] is best suited to hardware implementation since the estimation of the Fourier coefficients is a fast Fourier transform (FFT)-based. While the theory and application of CS have been well-developed for individual signals, several CS applications require distributed acquisition of multiple correlated signals. Thus, the so-called distributed compressed sensing (DCS), which contributes to what is known as the joint sparsity of the signals ensemble (i.e., sparse signal vectors sharing the same support). Therefore, taking benefit from the complex-exponential basis expansion model (CE-BEM) and the channel sparsity in the delay-Doppler domain [18], [19], [24], we move forward to estimate jointly sparse CE-BEM coefficient vectors over a DS channel than those of the delay-Doppler channel tap gains −''numerous'' channel coefficients. The reason is that since the time-varying nature of the delay-Doppler channel is well characterized in the CE-BEM by (known) Fourier basis functions, the time-variations of the (unknown) BEM coefficients are likely far slower than those of the delay-Doppler channel tap gains, and hence more suitable to track [3], [19]. Consequently, we can use the DCS method to directly estimate the two-dimensional (2D) channel coefficients (i.e., the delay-Doppler channel), where several jointly sparse (unknown) BEM coefficient vectors become the estimation goal.

Furthermore, we focus on the pilot pattern design in sparse CE over DS channels. In [3] and [7], pilot selection schemes have been proposed that considers the design of pilot symbols and their placement for the high-mobility OFDM systems over the DS channels. However, these methods do not exploit the inherent sparsity of the channel that is related to the fact that DS channels tend to be dominated by a relatively small number of clusters of significant propagation paths [24], [25].

The recently explored CS technique to decrease the transmission overhead of pilot subcarriers, exploits the inherent sparsity of the wireless fading channel [1], [2], [5], [11], [12], whose merit is that can increase spectrum efficiency through reducing the number of pilot symbols that have to be transmitted. In [5], the original DS channel is first approximated using CE-BEM, where the channel becomes transformed into 2D channel model, subsequently, a DCS-based simultaneous orthogonal matching pursuit (DCS-SOMP) algorithm is proposed to estimate the several jointly sparse (unknown) BEM coefficient vectors. However, rather than jointly optimized for pattern and power of pilots as a solution, the proposed technique in [5] optimized for pilot placements while assuming equal power allocation of pilot symbols. This may not lead to a reliable estimate of the jointly sparse unknown) BEM coefficient vectors over a DS channel. In [12], a technique that assumes equal power allocation assignment of pilot symbols and uses the discrete stochastic optimization (DSO) to obtain the optimal pilot placement is proposed, which then applied a structured DCS namely, block-based simultaneous orthogonal matching pursuit (BSOMP) algorithm for the joint recovery of coefficient vectors. In [11], an equal-power pilot design that optimizes pilot placements for DCS estimation scheme based on structured CS (SCS) for the time-frequency training OFDM systems, is proposed. Moreover, in [11] an adaptive support-aware block orthogonal matching pursuit (ASA-BOMP) greedy based algorithm, is also proposed for DS channel estimation −to recover the jointly sparse BEM coefficient vectors. In other words, Cheng *et al.* [5], Ma *et al.*, Qin *et al.* [12] considered the problem of sparse CE for OFDM systems over a DS channel while exploiting the joint sparsity property of the CE-BEM coefficient vectors and proposed the design of the pilot pattern by assuming equal power allocation of pilot symbols which may not be optimal.

While CS estimation of OFDM systems is well-explored [1], [2], [5], [11], [12], fewer works have addressed the design of pilot symbols and their placements for the high-mobility OFDM systems in the presence of DS fading. In [14], a joint pilot symbol and pilot design pattern is proposed for the high-mobility OFDM systems over DS channel to improve CE accuracy. The scheme assumes that there is two pilot symbol power corresponding to the placement set and employs the modified discrete stochastic approximation to optimize the pilot placement via maintaining an occupation probability vector, where the algorithm converges to the state which has the largest occupation probability. However, this

assumption may not necessarily lead to an optimal pilot design.

In this paper, we study the clustered pilot design problem for sparse CE of OFDM systems over a DS channel, where the resulting frequency domain channel matrix is approximately banded and exhibits a pseudo-circular structure, which gives rise to a diagonally dominant yet full matrix rather than a diagonal matrix and thus ICI. On the premise of this observation, we propose a new pilot design scheme that identifies the optimal pilot placement and values for each pilot cluster to combat ICI. The work will show that by taking the frequency domain channel matrix ICI into account in the proposed clustered pilot design, can reduce the rapid degradation of the pilot power from the diagonal to the margin, and thus mitigates ICI. Specifically, under the pilot power constraint, the proposed approach decomposes the problem of jointly optimizing the pilot sequences over both the pilot placement and values allocable to the pilot clusters into identifiable sequential conception. Hence, we applied the second-order cone programming (SOCP) for the optimal pilot power design and the stochastic search for the optimal pilot placement. Simulation results confirms that the proposed pilot design scheme can provide optimal pilot sequence allocable to the pilot cluster which can significantly improve the estimation accuracy of the OFDM system over DS channel. To ensure a satisfactory estimation, we have further proposed a new DCS recovery algorithm based on stage determined matching pursuit (SdMP) namely, DCS-SdMP that possesses the advantage of multiple common-support indices selections per iteration and then adds a backtracking process (at a later iteration) to subsequently refine the previously selected common support indices. Hence, the proposed algorithm enables the existing SdMP algorithm [26] of our prior work to deal with the problem of the distributed recovery of jointly sparse signals that share the same support. Simulation results show that the proposed pilot design and CE scheme, achieve better performance than existing approaches.

The remainder of this paper is organized as follows. In Section II, a review of CS and DCS theories and the OFDM system model over a DS channel are presented. In Section III, description of the new pilot design scheme is presented. Section IV introduces the proposed joint reconstruction algorithm for DCS that allows for joint recovery of signals ensembles with common-sparse supports. Section V presents the simulation results and performance evaluation. Finally, Section VI concludes the paper.

Notations: We use boldface lowercase and uppercase letters to denote vectors and matrices, respectively. For a given matrix $\mathbf{A}$, $\mathbf{A}^{-1}$, $\mathbf{A}^{\dagger}$, $\mathbf{A}^{T}$ and $\mathbf{A}^{H}$ represent its inverse, pseudo inverse, transpose and conjugate transpose, respectively. $\mathbf{0}_{M\times N}$ and $\mathbf{0}_{M}$ denote the $M \times N$ all zero matrix and the zero vector with dimension M, respectively. $\mathcal{P}$ denotes a set. $[\mathbf{A}]_{\mathcal{P}}$ represents the selected rows of $\mathbf{A}$, whose indices correspond to the set $\mathcal{P}$. $\mathbb{C}^{M\times N}$ represents the set of $M \times N$ matrices in complex field. $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix. $\boldsymbol{I}_N^{\langle q\rangle}$ denotes the permutation matrix which is obtained from $\boldsymbol{I}_N$ by shifting its column circularly $|q|$-times to the right for $q < 0$ and to the left otherwise. $\mathcal{CN}(0, \sigma^2)$ represents the complex normal distribution with independent real and imaginary parts each with mean zero and variance $\sigma^2/2$. $\otimes$ denotes the Kronecker product, $[\mathbf{A}]_{a,b}$ denotes a submatrix with row indices a and columns indices b, and diag($\boldsymbol{a}$) changes a vector $\boldsymbol{a}$ into a diagonal matrix.

II. CS AND DSC THEORIES AND SYSTEM MODEL

Here, we first present the CS and DCS theories, followed by a detailed description of the OFDM system model over a DS channel and then we introduce the sparse CE-BEM channel model within multiple OFDM symbols.

A. CS AND DCS THEORIES

1) CS THEORY

The goal of CS is to recover (or approximately recover) the original signal $\boldsymbol{x} \in \mathbb{C}^N$, from its measurements $\boldsymbol{y} \in \mathbb{C}^M$ [2], [26]. The CS of a signal $\boldsymbol{x} \in \mathbb{C}^N$ is summarized as two processes: The first is termed the measured process and expressed as

$$\boldsymbol{y} = \boldsymbol{\Psi}\boldsymbol{x} + \boldsymbol{v} \tag{1}$$

where $\boldsymbol{y} \in \mathbb{C}^M$ denotes the measured signal or compressed measurements, $\boldsymbol{\Psi} \in \mathbb{C}^{M\times N}$ denotes the measurement matrix with $M \ll N$ and $\boldsymbol{v} \sim \mathcal{CN}(\mathbf{0}_M, \sigma_v^2\boldsymbol{I}_M) \in \mathbb{C}^M$ denotes the noise term. Generally, the original signal $\boldsymbol{x}$ is assumed to have a sparse representation in a certain basis $\boldsymbol{\Phi} = \{\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_N\}$, since the path delay difference from the similar scatterer is far less than the system sampling period [2], [27]–[29]. This can be denoted by the following equation:

$$\boldsymbol{x} = \sum_{i=0}^{N-1} \boldsymbol{\phi}_i h_i = \boldsymbol{\Phi}\boldsymbol{h} \tag{2}$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{N\times N}$ represents a matrix with elements $\boldsymbol{\phi}_i \in \mathbb{C}^N$, for $i \in [0, N-1]$ and $\boldsymbol{h} \in \mathbb{C}^N$ represents a vector. We say that $\boldsymbol{h} \in \mathbb{C}^N$ is S-sparse in the basis $\boldsymbol{\Phi}$ if $\|\boldsymbol{h}\|_{\ell_0} \leq S$ [27], where $\|\cdot\|_{\ell_0}$ denotes the so-called ℓ_0-norm of a vector that measures the number of nonzero components of the vector. Now, using (1) and (2) we can write

$$\boldsymbol{y} = \boldsymbol{\Psi}\boldsymbol{\Phi}\boldsymbol{h} + \boldsymbol{v} = \mathbf{A}\boldsymbol{h} + \boldsymbol{v}, \tag{3}$$

where $\mathbf{A} = \boldsymbol{\Psi}\boldsymbol{\Phi} \in \mathbb{C}^{M\times N}$ denotes the measurement matrix, which is the generally employed notation in the CS literature and will be used in the rest of this paper.

The second process is termed the reconstruction process and deals with the recovery of the signal from its compressed measurement. The following reconstruction error bound is one of the basic results of CS (See [27], [30], and [31], and for further results and a discussion). In (3), since $M \ll N$, the recovery problem is ill-posed. However, the restrict isometry property (RIP) demonstrates that the sparse vector $\boldsymbol{h}$ in (3) can be recovered exactly from the noiseless measurement $\boldsymbol{y}$ (i.e., in the noise-free case $\boldsymbol{v} = \mathbf{0}_M$) with high probability, if matrix $\mathbf{A}$ satisfies the RIP [27]. The $M \times N$

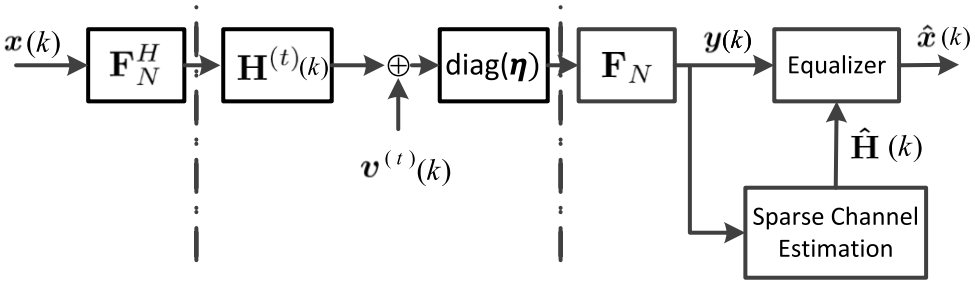

FIGURE 1. OFDM system transceiver block diagram.

matrix $\mathbf{A}$ in (3) is said to satisfy the RIP of order S if there exists a constant $\delta_S \in (0, 1)$ such that

$$(1 - \delta_S)\|\boldsymbol{h}\|_{\ell_2}^2 \leq \|\mathbf{A}\boldsymbol{h}\|_{\ell_2}^2 \leq (1 + \delta_S)\|\boldsymbol{h}\|_{\ell_2}^2, \tag{4}$$

is satisfied for all vectors with $\|\boldsymbol{h}\|_{\ell_0} \leq S$, where the ℓ_0-norm measures the nonzero elements in $\boldsymbol{h}$ [27], [28], [31], [32]. We have written $\|\cdot\|_{\ell_0}$ and $\|\cdot\|_{\ell_2}$ for the ℓ_0-norm vector and ℓ_2-norm vector, respectively. δ_S denotes the restricted isometry constant (RIC) which is the infimum of all possible δ for a given $\mathbf{A} \in \mathbb{C}^{M\times N}$ which can be expressed mathematically as

$$\delta_S := \arg\min_{\delta\geq 0}\{\delta : (1 - \delta)\|\boldsymbol{h}\|_{\ell_2}^2 \leq \|\mathbf{A}\boldsymbol{h}\|_{\ell_2}^2 \leq (1 + \delta)\|\boldsymbol{h}\|_{\ell_2}^2, \quad \forall \boldsymbol{h} \in \Sigma_S\}.$$

Unfortunately, verifying (4) for a given $\mathbf{A}$, S and $\delta_S \in (0, 1)$ has been shown to be NP-hard [27], [33] in general from the perspective of computational complexity. Fortunately, the most obvious application for the l_1-norm is to replace the l_0-norm problem [2]. Hence, computationally efficient l_1-norm algorithms are well-developed to solve a relaxed version of the l_0-norm problem. One of these algorithms is basis pursuit. However, the l_1-norm approach is often still computationally intensive and challenging to implement in real-time systems. Alternatively, a widely deployed framework for sparse signal recovery is the mutual incoherence property (MIP) presented in [29] and [34]. Using deterministic sensing matrices, if $S < \frac{1}{2}(1 + \frac{1}{\mu(\mathbf{A})})$ [30], [31], where

$$\mu(\mathbf{A}) = \max_{m\neq n} \frac{|\mathbf{a}_m^H \mathbf{a}_n|}{\|\mathbf{a}_m\|_{\ell_2}\|\mathbf{a}_n\|_{\ell_2}}, \quad \text{for } 1 \leq m, n \leq N \tag{5}$$

with $\mathbf{a}_m$ denoting the m-th column of $\mathbf{A}$, the same sparse reconstruction performance can be achieved from (3) via greedy pursuit (GP) algorithms like orthogonal matching pursuit (OMP) [28] with significantly less complexity. Therefore, in recent years, much attention has been focused on computationally efficient approaches. Further exciting results on CS can be found in [2], [26], and [33].

2) DCS THEORY

In DCS problem, a set of Q sparse vectors $\boldsymbol{x}_q$, for $q \in [0, Q-1]$ that assume a sparse representation $\{\boldsymbol{h}_q\}_{q=0}^{D-1} \in \mathbb{C}^N$ in some orthonormal basis $\boldsymbol{\Phi}$, i.e., $\boldsymbol{x}_q = \boldsymbol{\Phi}\boldsymbol{h}_q$, $q \in [0, Q-1]$ is to be jointly recovered from a set of Q measurement vectors

$$\begin{aligned}\boldsymbol{y}_q &= \boldsymbol{\Psi}\boldsymbol{\Phi}\boldsymbol{h}_q + \boldsymbol{v}_q \quad \forall q \in [0, Q-1] \\ &= \boldsymbol{A}\boldsymbol{h}_q + \boldsymbol{v}_q,\end{aligned} \tag{6}$$

where $\boldsymbol{y}_q \in \mathbb{C}^M$ is a measurement vector, $\mathbf{A} \in \mathbb{C}^{M\times N}$ a measurement matrix, $\boldsymbol{v}_q \in \mathbb{C}^N$ is the noise term, $M \ll N$. In (6), each sparse vector $\boldsymbol{h}_q$, $q \in [0, Q-1]$ is supported only on the same $\Lambda \subset \{1, \ldots, N\}$ with $|\Lambda| = S$. Thus, the signals ensemble i.e., $\{\boldsymbol{x}_q\}_{i=0}^{Q-1} \in \mathbb{C}^N$ are S-sparse and are constructed from the same S elements of the orthonormal basis $\boldsymbol{\Phi}$ [35], but with arbitrarily different coefficients. Assuming that the Q sparse vectors and the Q measurement vectors are arranged as columns of matrices $\mathfrak{H} = (\boldsymbol{h}_0, \ldots, \boldsymbol{h}_{Q-1}) \in \mathbb{C}^{N\times Q}$ and $\boldsymbol{Y} = (\boldsymbol{s}_0, \ldots, \boldsymbol{s}_{Q-1}) \in \mathbb{C}^{M\times Q}$, respectively. In DCS problem, $\mathfrak{H}$ is to be reconstructed given $\boldsymbol{Y}$ [36]

$$\mathbf{Y} = \mathbf{A}\mathfrak{H} + \mathbf{V}, \tag{7}$$

where $\mathbf{V} = [\boldsymbol{v}_0, \ldots, \boldsymbol{v}_{Q-1}] \in \mathbb{C}^{M\times Q}$ is an ensemble of noise terms.

B. OFDM SYSTEM MODEL OVER A DS CHANNEL

Consider an OFDM system with N subcarriers, as shown in Fig.1. Let the transmit signal at the n-th subcarrier during the k-th OFDM symbol be represented as $\boldsymbol{x}_n(k)$ for $n \in [0, N-1]$ and $k \in [0, K-1]$, with $[\boldsymbol{x}(k)]_n \triangleq [X_n(k)]$ for $n \in [0, N-1]$. Without loss of generality, the transmit signal is composed of $\boldsymbol{x}(k) \triangleq [X_0(k), ..., X_{N-1}(k)]^T$, for $k \in [0, K-1]$. Suppose that the set of the pilot subcarriers is denoted as $\breve{\mathcal{P}} = \{p_0, \ldots, p_{M-1}\}$, without losing generality, we assume that $\breve{\mathcal{P}} \subset \{1, \ldots, N\}$. Therefore, we use $[\boldsymbol{x}(k)]_{\breve{\mathcal{P}}}$ and $[\boldsymbol{x}(k)]_{\mathcal{D}}$ to denote the pilots of the k-th OFDM symbol and the corresponding data, respectively; such that $\breve{\mathcal{P}}(|\breve{\mathcal{P}}| = M)$ and $\mathcal{D}(|\mathcal{D}| = N - M)$ denote the set of pilots- and data- subcarrier indices, respectively, and $\breve{\mathcal{P}} \cap \mathcal{D} = \emptyset$ (where $\cap$ denotes the intersection operator). The transmit signal $\boldsymbol{x}(k)$ for $k \in [0, K-1]$ is used to modulate N carriers as $\boldsymbol{x}^{(t)}(k) = \mathbf{F}_N^H \boldsymbol{x}(k)$ for $k \in [0, K-1]$ in the time domain, where $\mathbf{F}_N \in \mathbb{C}^{N\times N}$ is the N-point normalized discrete Fourier transform (DFT) matrix with $[\mathbf{F}_N]_{a,b} = 1/\sqrt{N}\exp(-j2\pi ab/N)$ for $a, b \in [0, N-1]$. Let us now define $h_{n,l}^{(t)}(k)$ as the l-th channel tap at n-th time instant within the k-th OFDM

symbol, and let us assume that $h_{n,l}^{(t)}(k)$ have finite order L, i.e., $h_{n,l}^{(t)}(k) = 0$ for $l < 0$ or $l > L-1$ within the k-th OFDM symbol. Consequently, for each time domain channel tap l, we define

$$\boldsymbol{h}_l^{(t)}(k) \triangleq [h_{0,l}^{(t)}(k), \ldots, h_{N-1,l}^{(t)}(k)], \ l \in [0, L-1] \quad (8)$$

as a vector that stacks the time variations of the channel taps during the k-th OFDM symbol of the channel. The vector $\boldsymbol{h}^{(t)}(k)$ within the k-th OFDM symbol is assumed to be S-sparse (or approximately sparse) in some transform domain, since the number of dominant channel taps is no more than S in $\boldsymbol{h}^{(t)}(k)$. Hence, $\|\boldsymbol{h}^{(t)}(k)\|_{\ell_0} \leq S \ll L$ where $\|\cdot\|_{\ell_0}$ denotes the ℓ_0-norm of a vector that counts the number of nonzero entries of a vector.

After inserting the cyclic prefix (CP), the serially converted signal vector of the k-th OFDM symbol is transmitted through the DS channel $h_{n,l}^{(t)}(k)$. After removal of CP and reshaped by a windowing filter, the received signal vector in the time domain of the k-th OFDM symbol can be expressed as

$$\boldsymbol{y}^{(t)}(k) = \text{diag}(\boldsymbol{\eta})\mathbf{H}^{(t)}(k)\mathbf{F}_N^H\boldsymbol{x}(k) + \text{diag}(\boldsymbol{\eta})\boldsymbol{v}^{(t)}(k), \quad (9)$$

where $\boldsymbol{\eta} = (\eta_0, \ldots, \eta_{N-1})^T$ denotes the time domain window, while $\mathbf{H}^{(t)}(k)$ and $\boldsymbol{v}^{(t)}(k)$ denote the time domain channel matrix and noise term, respectively, without windowing. Herein, by the term without windowing we mean, a banded approximation of $\mathbf{H}^{(t)}(k)$ is to be determined, which does not require an appropriate window $\boldsymbol{\eta}$ that provides the smoothness [37] needed in OFDM symbol; rather than determining $\text{diag}(\boldsymbol{\eta})\mathbf{H}^{(t)}(k)$ that requires an appropriate window $\boldsymbol{\eta}$. Subsequently, the time domain signal samples $\boldsymbol{y}^{(t)}(k)$ is then demodulated by the N-point normalized DFT matrix $\mathbf{F}_N \in \mathbb{C}^{N\times N}$ in the receiver, and transformed to the frequency domain as

$$\begin{aligned}\boldsymbol{y}(k) = \mathbf{F}_N\boldsymbol{y}^{(t)}(k) &= \mathbf{F}_N\text{diag}(\boldsymbol{\eta})\mathbf{H}^{(t)}(k)\mathbf{F}_N^H\boldsymbol{x}(k) \quad (10)\\ &\quad + \mathbf{F}_N\text{diag}(\boldsymbol{\eta})\boldsymbol{v}^{(t)}(k)\\ &= \mathbf{F}_N\tilde{\mathbf{H}}^{(t)}(k)\mathbf{F}_N^H\boldsymbol{x}(k) + \mathbf{F}_N\tilde{\boldsymbol{v}}^{(t)}(k)\\ &= \mathbf{H}(k)\boldsymbol{x}(k) + \boldsymbol{v}(k), \quad (11)\end{aligned}$$

for $\forall k \in [0, K-1]$ with no appropriate window required for $\tilde{\mathbf{H}}^{(t)}(k) := \text{diag}(\boldsymbol{\eta})\mathbf{H}^{(t)}(k)$ and $\tilde{\boldsymbol{v}}^{(t)}(k) := \text{diag}(\boldsymbol{\eta})\boldsymbol{v}^{(t)}(k)$, where $\boldsymbol{y}(k) = [Y_0(k), \ldots, Y_{N-1}(k)]^T$ denotes the frequency domain received signal samples over all subcarriers during the k-th OFDM symbol, $\mathbf{H}(k) := \mathbf{F}_N\tilde{\mathbf{H}}^{(t)}(k)\mathbf{F}_N^H \in \mathbb{C}^{N\times N}$ denotes the frequency domain channel matrix and $\boldsymbol{v}(k) = [V_0(k), \ldots, V_{N-1}(k)]^T$ denotes the frequency domain noise term.

If we then assume that the length of the cyclic prefix L_{cp} satisfies $L_{cp} \geq L$ and being that the channel is time-variant, then matrix $\tilde{\mathbf{H}}^{(t)}(k) \in \mathbb{C}^{N\times N}$ is "pseudocirculant" [5], [6], [11], [38]. Consequently, matrix $\mathbf{H}(k)$ becomes practically a diagonally dominant yet full matrix rather than a diagonal matrix. Which essentially means that the received frequency-domain samples rely on both the pilot signals, known to the receiver and the unknown data signals. Mathematically, the time-domain $N \times N$ matrix $\tilde{\mathbf{H}}^{(t)}(k)$ for $k \in [0, K-1]$, which constitutes the main attribute of the channel matrix $\mathbf{H}(k)$ for $k \in [0, K-1]$ can be expressed as follows [5]

$$[\tilde{\mathbf{H}}^{(t)}(k)]_{a,b} = h^{(t)}_{(L_{cp}+a,\ \text{mod}\ (a-b,N))k}, \quad (12)$$

where $\text{mod}(a-b, N)$ denotes the remainder of $(a-b)$ divided by N where $(a-b)$ and N are integers. To be specific, the frequency domain channel matrix $\mathbf{H}(k)$ for $k \in [0, K-1]$ can be expressed as

$$[\mathbf{H}(k)]_{a+d,a} = \frac{1}{N}\sum_{n=0}^{N-1}\sum_{l=0}^{L-1} h^{(t)}_{k(L_{cp}+n,l)}e^{-j2\pi(al+dn)}, \quad (13)$$

Thus, in order to eliminate ICI at the receiver for K consecutive OFDM symbol, KNL channel coefficients of $h^{(t)}_{k(L_{cp}+n,l)}$ for $k \in [0, K-1]$, $n \in [0, N-1]$ and $l \in [0, L-1]$ needs to be identified. This, unfortunately, incurs prohibitively large number of channel coefficients and thus increased estimation complexity.

Our objective is to design a scheme for estimating the unknown channel coefficients from a reduced number of the observed sequence of channel coefficients, and thus reduced estimation complexity since a reduced number of pilot overhead will be required for reliable CE. In the following subsection, we adopt a model to reduce the number of channel coefficients to be estimated. In a later section, a new pilot design scheme that identifies the optimal pilot placement and values is proposed to provide a reliable estimate of the reduced number of the unknown channel coefficients. Subsequently, to obtain a channel estimator consistent with the jointly sparse two dimensional (2D) (i.e., delay-Doppler) channel model, a channel estimator consistent with the jointly sparse channel model is proposed.

C. REMODELING OF THE CIR WITH SPARSE CE-BEM

We provide in this section an attempt to develop an accurate time-domain sparse channel model of $h^{(t)}_{k(L_{cp}+n,l)}$ for $k \in [0, K-1]$, $n \in [0, N-1]$ and $l \in [0, L-1]$, using a BEM. Fortunately, there exists some correlation among the unknown channel coefficients [12], $h^{(t)}_{k(L_{cp}+n,l)}$ in the time index, n as the time variation of the l-th channel tap is rather smooth.[1] Thus, by exploiting the time correlations of the physical channel, the number of channel coefficients prior to CE can be reduced. On this account, the BEM is introduced as the CIR basis to exploit and capture the channel temporal correlations that exist over the physical channel. In this work, we model the rapidly time-varying channel, $h^{(t)}_{k(L_{cp}+n,l)}$ by a superposition of deterministic time-varying basis functions (e.g., complex exponentials (CE)) with time-invariant coefficients.

[1] Since channel parameters cannot change in a sudden way in the time domain, one may expect correlation among the time-varying channel parameters [39]. To model using the BEM, we assume that the time variation of the l-th channel tap is smooth, and thus $h^{(t)}_{k(L_{cp}+n,l)}$ is correlated in the time index n. Then, we can accurately model K consecutive samples of the l-th channel tap.

Now, collecting the time-varying function of the l-th channel tap $\boldsymbol{h}_l^{(t)}(k) := (h^{(t)}_{k(N+L_{cp})+L_{cp},l}, \ldots, h^{(t)}_{(k+1)(N+L_{cp})-1,l})^T \in \mathbb{C}^{N\times 1}$ for $l \in [0, L-1]$ of the k-th OFDM sysmbol, we can express $\boldsymbol{h}_l^{(t)}(k)$ as

$$\boldsymbol{h}_l^{(t)}(k) = \boldsymbol{\mathcal{B}}\boldsymbol{c}_l(k) + \boldsymbol{\xi}_l(k), \quad \forall l \in [0, L-1] \tag{14}$$

where $\boldsymbol{\mathcal{B}} := [\boldsymbol{b}_0, \ldots, \boldsymbol{b}_{Q-1}] \in \mathbb{C}^{N\times Q}$ stacks $Q(Q \ll N)$ orthonormal basis functions $\boldsymbol{b}_q \in \mathbb{C}^{N\times 1}$ for $q \in [0, Q-1]$ which are complex exponential as columns, $\boldsymbol{c}_l(k) := [c_{0,l}(k), \ldots, c_{Q-1,l}(k)]^T \in \mathbb{C}^{Q\times 1}$ represents the corresponding BEM coefficients for the l-th distinct channel delay, and $\boldsymbol{\xi}_l(k) := [\xi_{0,l}(k), \ldots, \xi_{N-1,l}(k)]^T \in \mathbb{C}^{N\times 1}$ represents the BEM modelling error. Here, a BEM with complex exponentials (CE-BEM) is used to model the time-varying unwindowed channel since it relies on a banded approximation of the frequency domain channel matrix. Notably, by utilizing the CE-BEM and exploiting the delay-Doppler domain channel sparsity within multiple OFDM symbols, we turn to estimate the jointly sparse CE-BEM coefficients $c_{q,l}(k)$ for $k \in [0, K-1], q \in [0, Q-1]$ and $l \in [0, L-1]$ which requires the estimation of KQL unknown channel coefficients with the aid of pilots. Rather than estimating the numerous channel coefficients, $h^{(t)}_{n,l}(k)$ for $k \in [0, K-1]$, $n \in [0, N-1]$ and $l \in [0, L-1]$ i.e., KNL,

Definition 1: Suppose that $\mathcal{L} = \{l : |h_{n,l}(k)| > S\}$ for $l \in [0, L-1]$ denotes the position index set of dominant channel taps for a fixed n related to the k-th OFDM symbol, then we have $\boldsymbol{h}_l^{(t)}(k) = 0$ for $k \in [0, K-1]$ and $l \neq \mathcal{L}$.

Therefore, since one can show that $c_{0,l}(k) = \cdots = c_{Q-1,l}(k) = 0$ for $l \neq \mathcal{L}$, then the CE-BEM coefficient vector $\boldsymbol{c}_q(k)$ for $q \in [0, Q-1]$ and $k \in [0, K-1]$ is as well jointly sparse and share the same sparse support with that of $\boldsymbol{h}_l^{(t)}(k)$, $\forall l$ (i.e., $(\boldsymbol{c}_{0,l}(k), \ldots, \boldsymbol{c}_{Q-1,l}(k))^T = \boldsymbol{\mathcal{B}}^\dagger \boldsymbol{h}_l^{(t)}(k)$, $\forall l \in [0, L-1]$).

By stacking all the channel taps within the k-th block in a single vector

$$h^{(t)}(k) := (h^{(t)}_{k(N+L_{cp})+L_{cp},0}, \ldots, h^{(t)}_{k(N+L_{cp})+L_{cp},L-1}, \ldots, h^{(t)}_{(k+1)(N+L_{cp})-1,0}, \ldots, h^{(t)}_{(k+1)(N+L_{cp})-1,L-1})^T$$

the BEM permits to express this vector as

$$\boldsymbol{h}^{(t)}(k) = (\boldsymbol{\mathcal{B}} \otimes \boldsymbol{I}_L)\boldsymbol{c}(k) + \boldsymbol{\xi}(k), \tag{15}$$

where $\boldsymbol{I}_L$ denotes the identity matrix of $L \times L$ dimension, $\boldsymbol{c}(k) := (c_{0,0}(k), \ldots, c_{0,L-1}(k), \ldots, c_{Q,0}(k), \ldots, c_{Q,L-1}(k))^T$ denotes the BEM stacking coefficient vector of all the channel taps within the k-th OFDM system; $\boldsymbol{\xi}(k) := (\xi_{0,0}(k), \ldots, \xi_{0,L-1}(k), \ldots, \xi_{N-1,0}(k) \ldots, \xi_{N-1,L-1}(k))^T$.

Consequently, (10) can be expressed in terms of CE-BEM as

$$\boldsymbol{y}(k) = \mathbf{H}(k)\boldsymbol{x}(k) + \boldsymbol{v}(k) \tag{16}$$

$$= \left(\sum_{q=0}^{Q-1} \boldsymbol{\mathcal{B}}_q^{CE} \mathbf{C}_q^{CE}(k)\right)\boldsymbol{x}(k) + \boldsymbol{v}(k), \tag{17}$$

with $\mathbf{H}(k) = \sum_{q=0}^{Q-1} \boldsymbol{\mathcal{B}}_q^{CE} \mathbf{C}_q^{CE}(k)$, where $\boldsymbol{\mathcal{B}}_q^{CE} := \sqrt{N}\mathbf{F}_N \mathrm{diag}(\boldsymbol{b}_q^{CE}\mathbf{F}_N^H = \sqrt{N}\boldsymbol{I}_N^{\langle q-\frac{Q-1}{2}\rangle} \in \mathbb{C}^{N\times N}$ denotes the circulant matrix in the frequency domain whose q-th basis is expressed as $\boldsymbol{b}_q^{CE} = \left(1, \ldots, e^{j\frac{2\pi}{N}n(q-\frac{Q-1}{2})}, \ldots, e^{j\frac{2\pi}{N}(N-1)(q-\frac{Q-1}{2})}\right)^T$, with $\boldsymbol{I}_N^{\langle q\rangle}$ denoting the permutation matrix which is obtained from $\boldsymbol{I}_N$ by shifting its column circularly $|q|$-times to the right for $q < 0$ and to the left otherwise, and $\mathbf{C}_q^{CE}(k)\boldsymbol{x}(k) = \mathrm{diag}(\mathbf{F}_N((\boldsymbol{c}_q^{CE}(k))^T, \mathbf{0}_{1\times(N-L)})^T)\boldsymbol{x}(k) = \mathrm{diag}(\boldsymbol{x}(k))\cdot \mathbf{F}'_N\boldsymbol{c}_q^{CE}(k)$ for $q \in [0, Q-1]$ denotes the corresponding CE-BEM coefficients, while $\mathbf{F}'_N$ denotes the discrete Fourier transform (DFT) submatrix that extracts the first L columns of $\mathbf{F}_N \in \mathbb{C}^{N\times N}$ which can be written as

$$\mathbf{F}'_N = \begin{pmatrix} 1 & 1 & & \\ cdots & 1 & & \\ 1 & \omega & \cdots & \omega^{L-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \cdots & \omega^{(N-1)(L-1)} \end{pmatrix}_{N\times L}, \quad \omega = e^{-j\frac{2\pi}{N}}, \tag{18}$$

Consequently, (16) can be rewritten as

$$\boldsymbol{y}(k) = \sum_{q=0}^{Q-1} \boldsymbol{I}_N^{\langle q-\frac{Q-1}{2}\rangle} \mathrm{diag}(\boldsymbol{x}(k))\mathbf{F}'_N\boldsymbol{c}_q(k) + \boldsymbol{v}(k). \tag{19}$$

Then we can express the K consecutive OFDM received symbols in terms of sparse CE-BEM as

$$\boldsymbol{y} = \sum_{q=0}^{Q-1} \boldsymbol{I}_{KN}^{\langle q-\frac{Q-1}{2}\rangle} \mathrm{diag}(\boldsymbol{x})(\boldsymbol{I}_K \otimes \mathbf{F}'_N) \begin{pmatrix} \boldsymbol{c}_q^{CE}(0) \\ \vdots \\ \boldsymbol{c}_q^{CE}(K-1) \end{pmatrix} + \boldsymbol{v}, \tag{20}$$

We estimate the CE-BEM coefficients, $c_{q,l}(k)$ for $k \in [0, K-1]$, $q \in [0, Q-1]$ and $l \in [0, L-1]$ with the aid of pilots $\breve{\boldsymbol{p}}^{\mathrm{eff}} = [\boldsymbol{x}_k]_{\breve{\mathcal{P}}^{\mathrm{eff}}} := [X_0(k), \ldots, X_{M-1}(k)]^T$, for $k \in [0, K-1]$ with positions represented by $\breve{\mathcal{P}}^{\mathrm{eff}} = \{p_0, \ldots, p_{M-1}\}$. Therefore, to render the equivalent channel matrix circulant and perfectly diagonalizable by the DFT, we add $(Q-1)$ guard symbols before and after each effective pilot symbol. Hence, both interblock interference (IBI) between successive transmitted blocks and ICI between the frequency subchannels in each block in the desired link, can be mitigated. We then assume that there are M pilot clusters within the k-th OFDM symbol of length $L_p = (2Q-1)$ (as will be shown later), represented as $\acute{\boldsymbol{x}}_k^{(PG)}(m)$ for $m \in [0, M-1]$ and $k \in [0, K-1]$ for each effective pilot position as illustrated in Fig. 2. Stacking together all the transmitted signal (i.e., pilot symbols and guard symbols) within the k-th OFDM symbol yields the pilot vector denoted as

$$\acute{\boldsymbol{x}}^{(PG)}(k) := (\acute{\boldsymbol{x}}_0^{(PG)T}(k), \ldots, \acute{\boldsymbol{x}}_{M-1}^{(PG)T}(k))^T \tag{21}$$

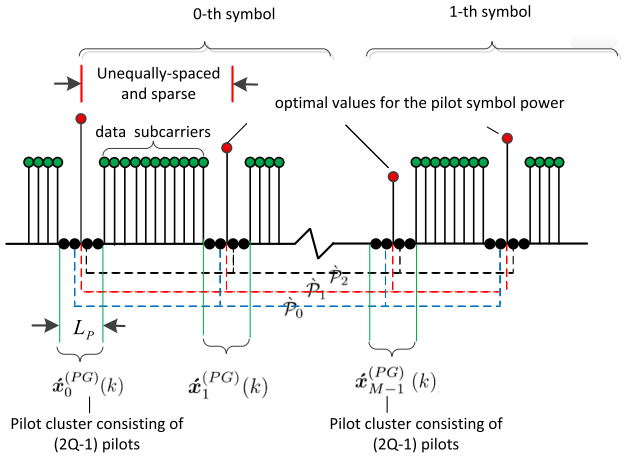

FIGURE 2. Placement of the optimal pilot symbols. Note that the distance between each pilot cluster is unequal.

where we denote $\boldsymbol{p}^{PG} := \grave{\boldsymbol{x}}^{(PG)}(k)$, where each of its elements $\{\grave{\boldsymbol{x}}_m^{(PG)}(k)\}_{m=0}^{M-1}$ is a cluster of pilot subcarriers of length L_p. The term $\boldsymbol{p}^{PG}$ indicates the total transmitted pilot symbols and guard symbols within the k-th OFDM symbol, with its corresponding location index set denoted as $\grave{\mathcal{P}}^{PG}$. From (21), the m pilot cluster of length L_p within the k-th OFDM symbol is defined as $\grave{\boldsymbol{x}}_m^{(PG)}(k) = (\boldsymbol{x}_{b_m}(k), \ldots, \boldsymbol{x}_{b_m+L_p-1}(k))^T$ for $m \in [0, M-1]$, where b_m denotes the begin position in the cluster. For example, with $m = 0$, represents the first cluster $\grave{\boldsymbol{x}}_0^{(PG)}(k) = (\boldsymbol{x}(k), \ldots, \boldsymbol{x}_{L_p-1}(k))^T$, with $b_m = 0$ denoting its begin position.

Since each cluster consists of $(2Q-2)$ guard pilot symbols $\boldsymbol{p}^G$, plus one effective pilot symbol, then a total of $(2Q-2)+1 = 2Q-1$ guard symbols and pilot symbol are transmitted per cluster. Thus, a total of $(2Q-2)M$ guard symbols is utilized within the k-th OFDM symbol, such that $|p_i - p_j| \geq 2Q-1, i \neq j$ is maintained to prevent overlapping. Consequently, the guard pilot index set can mathematically be expressed as: $\mathcal{P}^G = \cup\{t-Q-1, \ldots, t-1, t+1, \ldots, t+Q-1\}$ where $t \in \check{\mathcal{P}}$. It is worth noting that the symbol values of the guard pilots $\boldsymbol{p}^G$ are set to zero.

Now, in order to track the CE-BEM coefficients corresponding to the q-th basis function, we re-divide the effective pilot indices $\check{\mathcal{P}}^{\text{eff}}$ into Q subsets as

$$\grave{\mathcal{P}}_q = \check{\mathcal{P}}^{\text{eff}} - \left(\frac{Q-1}{2} - q\right), \quad q \in [0, Q-1] \tag{22}$$

as illustrated in Fig. 2 (with the assumption that $Q = 3$), and the number of guard pilots before and after each effective pilot is $Q-1$. Next, let us define the q-th pilot selection matrix for K consecutive OFDM symbols as $\mathbf{Z}_q =: [\boldsymbol{I}_{KN}]_{\grave{\mathcal{P}}_q} \in \mathbb{C}^{M \times KN}$ for $q \in [0, Q-1]$. It then follows from (20) that

$$[\boldsymbol{y}]_{\grave{\mathcal{P}}_q} = \mathbf{Z}_q \boldsymbol{y} = \sum_{\varrho=0}^{Q-1} \mathbf{Z}_q \boldsymbol{I}_{KN}^{\langle \varrho - \frac{Q-1}{2} \rangle} \text{diag}(\boldsymbol{x}) \boldsymbol{\Omega} \begin{pmatrix} c_\varrho^{CE}(0) \\ \vdots \\ \boldsymbol{c}_\varrho^{CE}(K-1) \end{pmatrix} + \tilde{\boldsymbol{v}}, \tag{23}$$

where $\tilde{\boldsymbol{v}} = \mathbf{Z}_q \boldsymbol{v}$ and $\boldsymbol{\Omega} \triangleq [\boldsymbol{I}_K \otimes \mathbf{F}'_N]$. From (23), since $\mathbf{Z}_q \boldsymbol{I}_{KN}^{\langle \varrho - \frac{Q-1}{2} \rangle} = \mathbf{Z}_{q-\varrho+\frac{Q-1}{2}}$, one can easily show that

$$[\boldsymbol{y}]_{\grave{\mathcal{P}}_q} = \sum_{\varrho=0}^{Q-1} \mathbf{Z}_{q-\varrho+\frac{Q-1}{2}} \text{diag}(\boldsymbol{x}) \boldsymbol{\Omega} \tilde{\boldsymbol{c}}_\varrho^{CE} + \tilde{\boldsymbol{v}}, \tag{24}$$

where $\tilde{\boldsymbol{c}}_\varrho^{CE} \triangleq ((\boldsymbol{c}_\varrho^{CE}(0))^T, \ldots, (\boldsymbol{c}_\varrho^{CE}(K-1))^T) \in \mathbb{C}^{KL \times 1}$. Without loss of generality, we assume that Q is an odd number, then the pilot subcarriers can be represented as

$$[\boldsymbol{x}]_{\grave{\mathcal{P}}_q} = \begin{cases} \mathbf{0}, & q \neq \dfrac{Q-1}{2} \\ \check{\boldsymbol{p}}^{\text{eff}}, & q = \dfrac{Q-1}{2} \end{cases} \tag{25}$$

Hence, the nonzero values from $\boldsymbol{x}$ can only be extracted if $\varrho = q$. Hence, (24) can be re-expressed as

$$[\boldsymbol{y}]_{\grave{\mathcal{P}}_q} = \text{diag}(\check{\boldsymbol{p}}^{\text{eff}}) \boldsymbol{\Omega}_{\check{\mathcal{P}}^{\text{eff}}} \tilde{\boldsymbol{c}}_q^{CE} + \tilde{\boldsymbol{v}} \quad \forall q \in [0, Q-1] \tag{26}$$

where $\tilde{\boldsymbol{c}}_q^{CE} \triangleq ((\boldsymbol{c}_q^{CE}(0))^T, \ldots, (\boldsymbol{c}_q^{CE}(K-1))^T) \in \mathbb{C}^{KL \times 1}$.

D. THE DCS FRAMEWORK FOR K CONSECUTIVE OFDM SYMBOLS

As the ensemble of signals shares the same measurement matrix, $\mathbf{A}$, it is, thus, considered inefficient to recover each sparse vector $\tilde{\boldsymbol{c}}_q^{CE}$ for $q \in [0, Q-1]$ independent. From (22), note that if $q = (Q-1)/2$ implies

$$\begin{aligned} \grave{\mathcal{P}}_{\frac{Q-1}{2}} &= \check{\mathcal{P}}^{\text{eff}} - \left(\frac{Q-1}{2} - \frac{Q-1}{2}\right) \\ &= \check{\mathcal{P}}^{\text{eff}}, \end{aligned} \tag{27}$$

then one can decompose $\boldsymbol{\Omega}_{\check{\mathcal{P}}^{\text{eff}}}$ as $\boldsymbol{\Omega}_{\grave{\mathcal{P}}_{\frac{Q-1}{2}}} \mho_q$, where

$$\mho_q := \text{diag}(1, \omega^{q-\frac{Q-1}{2}}, \ldots, \omega^{(q-\frac{Q-1}{2})(L-1)}) \tag{28}$$

is a diagonal matrix. Subsequently, (26) can be rearranged as

$$[\boldsymbol{y}]_{\grave{\mathcal{P}}_q} = (\text{diag}(\check{\boldsymbol{p}}^{\text{eff}}) \boldsymbol{\Omega}_{\grave{\mathcal{P}}_{\frac{Q-1}{2}}})(\mho_q \tilde{\boldsymbol{c}}_q^{CE}) + \tilde{\boldsymbol{v}}, \tag{29}$$

for $\forall q \in [0, Q-1]$ and then

$$[\boldsymbol{y}]_{\grave{\mathcal{P}}_q} = \mathbf{A}(\mho_q \tilde{\boldsymbol{c}}_q^{CE}) + \tilde{\boldsymbol{v}}_q, \tag{30}$$

where $\mathbf{A} \triangleq \text{diag}(\check{\boldsymbol{p}}^{\text{eff}}) \boldsymbol{\Omega}_{\grave{\mathcal{P}}_{\frac{Q-1}{2}}}$ denotes the measurement matrix which is the same for all signal ensemble. Since the pilot placement and symbol values are important in the measurement matrix design, then one can design the effective pilots by exploiting (27) (i.e., $q = (Q-1)/2$). Hence, (30) can be rewritten as

$$[\boldsymbol{y}]_{\grave{\mathcal{P}}_{\grave{\mathcal{P}}_{\frac{Q-1}{2}}}} = \mathbf{A}(\mho_{\grave{\mathcal{P}}_{\frac{Q-1}{2}}} \tilde{\boldsymbol{c}}^{CE}_{\grave{\mathcal{P}}_{\frac{Q-1}{2}}}) + \tilde{\boldsymbol{v}}_{\grave{\mathcal{P}}_{\frac{Q-1}{2}}}, \tag{31}$$

where $\Gamma_{\grave{\mathcal{P}}_{\frac{Q-1}{2}}} = \boldsymbol{I}_L$. In the sequel, we will refer to $\grave{\mathcal{P}}_{\frac{Q-1}{2}}$ as $\grave{\mathcal{P}}^*$ for brevity and the corresponding pilot symbols as $\boldsymbol{u}$.

For a given pilot pattern $\grave{\mathcal{P}}^* = \{p_0, \ldots, p_{M-1}\}$, where $0 \leq p_0 < p_1 < \cdots < p_{M-1} \leq KN-1$, we express the coherence of $\mu\{\mathbf{A}\}$ for our DCS framework according to (5) as

$$\mu\{\mathbf{A}\} = \max_{0\leq m<n\leq L-1} \frac{|\sum_{i=0}^{M-1} |x(\grave{\mathcal{P}}_i^*)|^2 e^{-j\frac{2\pi}{N}\grave{\mathcal{P}}_i^*(n-m)}|}{\sum_{i=0}^{M-1} |x(\grave{\mathcal{P}}_i^*)|^2}. \quad (32)$$

$$\text{s.t} |\grave{p}_i^* - \grave{p}_j^*| \geq 2Q-1, \quad \forall i,j,\ i \neq j \quad (32a)$$

We observe in (32) that the contribution of pilot values in the coherence measure is through their magnitudes $u(i) = |x(\grave{\mathcal{P}}_i^*)|^2$. Let $r = n-m$ and $\mathcal{L} = \{1, \ldots, L-1\}$. Then, (32) can be rewritten as

$$\mu(\grave{\mathcal{P}}^*, \boldsymbol{u}) = \max_{r\in\mathcal{L}} \frac{|\sum_{i=0}^{M-1} |x(\grave{\mathcal{P}}_i^*)|^2 e^{-j\frac{2\pi}{N}\grave{\mathcal{P}}_i^* r}|}{\sum_{i=0}^{M-1} |x(\grave{\mathcal{P}}_i^*)|^2}, \quad (33)$$

$$\text{s.t} |\grave{p}_i^* - \grave{p}_j^*| \geq 2Q-1, \quad \forall i,j,\ i \neq j \quad (33a)$$

where $\mu(\grave{\mathcal{P}}^*, \boldsymbol{u}) \triangleq \mu\{\mathbf{A}\}$. Our objective is to minimize $\mu(\grave{\mathcal{P}}^*, \boldsymbol{u})$ by selecting the optimal pilot symbols $\boldsymbol{u} = \{x(\grave{\mathcal{P}}_i^*)\}_{i=0}^{M-1} \in \mathbb{C}^M$ and their placement $\grave{\mathcal{P}}^* = \{p_0, \ldots, p_{M-1}\} \in \mathbb{C}^M$. We, however, observe in (33) that the contribution of the pilot values in $\mu(\grave{\mathcal{P}}^*, \boldsymbol{u})$ is through their magnitudes $u(i) = |x(\grave{\mathcal{P}}_i^*)|^2$ for $i \in [0, M-1]$. Therefore, the optimal pilot design problem is simplified as

$$\mho_{opt} = \arg \min_{\grave{\mathcal{P}}^*, \grave{\boldsymbol{p}}^*} \max_{r\in\mathcal{L}} \frac{|\sum_{i=0}^{M-1} |x(\grave{\mathcal{P}}_i^*)|^2 e^{-j\frac{2\pi}{N}\grave{\mathcal{P}}_i^* r}|}{\sum_{i=0}^{M-1} |x(\grave{\mathcal{P}}_i^*)|^2}. \quad (34)$$

$$\text{s.t} |\grave{p}_i^* - \grave{p}_j^*| \geq 2Q-1, \quad \forall i,j,\ i \neq j \quad (34a)$$

Here, $\mho \triangleq \mu(\grave{\mathcal{P}}^*, \boldsymbol{u})$, where $\arg\min(\cdot)$ denotes the minimizing argument to achieve $\mho_{opt}$. Rather than the equi-powered pilot symbol assumption which may not be optimal, our problem is to determine the optimal $\boldsymbol{u}$ and the corresponding index set of pilot symbols $\grave{\mathcal{P}}^*$ that can minimize the normalised mean squared error (NMSE) of CE over the DS channel allocable to the pilot clusters. For any prescribed energy to be applied for CE, we normalize the sum of pilot power (i.e symbol values). Suppose the sum power of all pilot symbols is

$$\sum_{i=0}^{M-1} |x(\grave{\mathcal{P}}_i^*)|^2 = \sum_{i=0}^{M-1} u(i) = \lambda_{\text{sum}}, \quad (35)$$

where $\boldsymbol{u} =: [u(0), \ldots, u(M-1)]$ denotes the pilot symbol[2] values that are to be transmitted on M pilot sub-channels with the additional constraint that $\lambda_{\text{Th}} < u(i) \leq \lambda_{\text{peak}}$ for $i \in [0, M-1]$, where λ_{Th} and λ_{peak} denote the threshold power and peak power constraint,[3] respectively. With these constraints, the pilot pattern design can be formulated as

$$\mho = \min_{\grave{\mathcal{P}}^*, \boldsymbol{u}} \mu(\grave{\mathcal{P}}^*, \boldsymbol{u}), \quad (36)$$

$$\text{s.t} |\grave{p}_i^* - \grave{p}_j^*| \geq 2Q-1, \quad \forall i,j,\ i \neq j \quad (36a)$$

$$\sum_{i=0}^{M-1} u(i) = \lambda_{\text{sum}}, \quad \lambda_{\text{Th}} < u(i) \leq \lambda_{\text{peak}} \quad (36b)$$

which involves the joint optimization of the pilot placement $\grave{\mathcal{P}}^*$ and pilot power $\boldsymbol{u}$ that can maximize the effective signal-to-noise ratio (SNR) for the jointly sparse CE-BEM coefficient vector estimation for OFDM Systems operating over DS channels.

For the pilot design problem of (36), it is empirically intractable to get a solution. For example, the exhaustive search over $\binom{512}{24}$ (i.e., for $N = 512$ and $M = 24$) typically requires the generation of all 9.817×10^{40} possible pilot placements. As the exhaustive search method incurs prohibitive computational costs, we decouple (36) in the sequel and then propose a low-cost suboptimal solution. Hence, (36) is decoupled in the following formulations:

For the given set of pilot pattern $\grave{\mathcal{P}}^* = \{p_0, \ldots, p_{M-1}\}$ for $\grave{\mathcal{P}}^* \subset \{1, \ldots, N\}$, we start by realizing

$$\mu_u(\grave{\mathcal{P}}^*) \triangleq \min_{\boldsymbol{u}} \mu_u(\grave{\mathcal{P}}^*, \boldsymbol{u}) \quad (37)$$

$$\text{s.t} \sum_{i=0}^{M-1} u(i) = \lambda_{\text{sum}}, \quad \lambda_{\text{Th}} < u(i) \leq \lambda_{\text{peak}} \quad (37a)$$

and then we determine

$$\min_{\grave{\mathcal{P}}^*} \mu_u(\grave{\mathcal{P}}^*) \quad (38)$$

$$\text{s.t} |\grave{p}_i^* - \grave{p}_j^*| \geq 2Q-1, \quad \forall i,j,\ i \neq j \quad (38a)$$

to achieve an optimal pilot placement set $\grave{\mathcal{P}}^*$. At the same time, the corresponding pilot symbol values $\boldsymbol{u}$ is additionally obtained.

III. OPTIMIZATION OF PILOT PLACEMENT AND VALUES

This section presents the proposed pilot placement and value optimization for reliable transmission over DS wireless channels.

A. PROPOSED PILOT PLACEMENT AND VALUE DESIGN SCHEME

To assign pilot powers to a given set of pilot locations $\grave{\mathcal{P}}_m^* = \{p_0, \ldots, p_m\}$ in the channel model, we begin from $m = 1$ and progressively move towards $m = M$. By defining

$$z(r) \triangleq \frac{\sum_{i=0}^{m-1} |x(\grave{\mathcal{P}}_i^*)|^2 e^{-j\frac{2\pi}{N}\grave{\mathcal{P}}_i^* r}}{\sum_{i=0}^{m-1} |x(\grave{\mathcal{P}}_i^*)|^2}, \quad r \in \mathcal{L}$$

$$\triangleq \frac{1}{\lambda_{\text{sum}}} \sum_{i=0}^{m-1} |x(\grave{\mathcal{P}}_i^*)|^2 e^{-j\frac{2\pi}{N}\grave{\mathcal{P}}_i^* r}, \quad r \in \mathcal{L}$$

$$\overset{a}{\triangleq} \frac{1}{\lambda_{\text{sum}}} \sum_{i=0}^{m-1} u(i) e^{-j\frac{2\pi}{N}\grave{\mathcal{P}}_i^* r}, \quad r \in \mathcal{L} \quad (39)$$

[2]Since we need to normalize the sum of pilot power for any prescribed energy to be utilized for CE, in the simulation, we set $\lambda_{\text{sum}} = 1$, which implies that the sum power of pilot symbol values is normalized.

[3]However, the power allocation among pilot symbols is not uniformly constrained but set with a peak power constraint of λ_{peak} and with a power threshold of λ_{Th} for the design and placement of pilot symbols. We, therefore, set values such that the pilot symbol values does not vary to extremely.

where (a) follows from (35), then the problem of optimal power assignment translates into

$$\mho_{opt} = \min_{\grave{\mathcal{P}}_m, \grave{\boldsymbol{p}}_m^*} \max_{r\in\mathcal{L}} |z(r)|, \tag{40}$$

$$\text{s.t } |\grave{p}_i^* - \grave{p}_j^*| \geq 2Q-1, \quad \forall i,j,\ i\neq j \tag{40a}$$

$$\sum_{i=0}^{M-1} u(i) = \lambda_{\text{sum}}, \quad \lambda_{\text{Th}} < u(i) \leq \lambda_{\text{peak}} \tag{40b}$$

and then

$$\mho_{opt} = \min_{\grave{\mathcal{P}}_m^*, \grave{\boldsymbol{p}}_m^*} \|\mathbf{z}\|_{\ell_\infty} \tag{41}$$

$$\text{s.t } |\grave{p}_i^* - \grave{p}_j^*| \geq 2Q-1, \quad \forall i,j,\ i\neq j \tag{41a}$$

$$\sum_{i=0}^{M-1} u(i) = \lambda_{\text{sum}}, \quad \lambda_{\text{Th}} < u(i) \leq \lambda_{\text{peak}}. \tag{41b}$$

We have written $\|\cdot\|_{\ell_\infty}$ for the ℓ_∞ vector norm. Subsequently, using (35) and (39), (41) can be rewritten as

$$\mho_{opt} = \min_{\grave{\mathcal{P}}_m^*, \grave{\boldsymbol{p}}_m^*} \frac{1}{\lambda_{\text{sum}}} \|\mathbf{W}_{\grave{\mathcal{P}}^*}\, \boldsymbol{u}\|_{\ell_\infty} \tag{42}$$

$$\text{s.t } |\grave{p}_i^* - \grave{p}_j^*| \geq 2Q-1, \quad \forall i,j,\ i\neq j \tag{42a}$$

$$\sum_{i=0}^{M-1} u(i) = \lambda_{\text{sum}}, \quad \lambda_{\text{Th}} < u(i) \leq \lambda_{\text{peak}}, \tag{42b}$$

where $\mathbf{W}_{\grave{\mathcal{P}}^*} := \sum_{i=0}^{m-1} e^{-j\frac{2\pi}{N}\grave{\mathcal{P}}_i^* r}$ with $m = M$ is a complex-valued submatrix. Let $g(j)$ represent the j-th row of matrix $\mathbf{W}_{\grave{\mathcal{P}}^*} \in \mathbb{C}^{L-1\times S}$ for $j = [1, L-1]$. We can moreover express

$$\mathbf{G}(j) = \left[(\mathbb{R}\text{e}\{g(j)\})^T, (\mathbb{I}\text{m}\{g(j)\})^T\right]^T \quad j = [1, L-1] \tag{43}$$

where $\mathbb{R}\text{e}\{g(j)\}$ and $\mathbb{I}\text{m}\{g(j)\}$ represent the real and imaginary parts of $g(j)$ with dimensions $1 \times S$, respectively, and $\mathbf{G}(j)$ is a $2 \times S$ real valued matrix. Then, (42) can be recast into an equivalent real-valued optimization problem in the form

$$\min_d d \tag{44}$$

$$\text{s.t } \|\mathbf{G}(j)\boldsymbol{u}\|_{\ell_2} \leq d \quad j \in [1, L-1] \tag{44a}$$

$$\sum_{i=0}^{M-1} u(i) = \lambda_{\text{sum}}, \quad \lambda_{\text{Th}} < u(i) \leq \lambda_{\text{peak}} \tag{44b}$$

containing the coefficients for the linear and second order cone constraints as an SOCP optimization problem. Due to the strengths of the linear and conic optimizers in MOSEK solver [40], MOSEK solver can be applied to solve the SOCP optimization problem for the pilot power assignment. Consequently, solutions in the form $\tilde{d}$ and $\tilde{\boldsymbol{u}}$ can be realized from (44), such that $\mu_u(\grave{\mathcal{P}}^*) := \tilde{d}$ denotes the designed pilot symbol values.

The pseudocode for the optimal pilot power design and placement scheme is shown as Algorithm 1. The input parameters to Algorithm 1 include the number of subcarriers per OFDM symbol N, the number of pilot subcarriers M, the number of non-zero taps of the channel S, the number of outer-loop iterations n_{out} and the number of inner-loop iterations n_{in}. For Algorithm 1, some notations should be further detailed. First, $\mathbf{0}_{n_{in}\times M}$ and $\mathbf{0}_{n_{in}}$ denote $n_{in} \times M$ all zero matrix and the zero vector with dimension n_{in}, respectively. Second, the pilot placement set obtained in the previous inner loop iteration is temporarily stored in $\acute{\boldsymbol{p}}_{\text{store}}$, where $\acute{\boldsymbol{p}}_{\text{store}}$ denotes the auxiliary pilot vector. Here we further explain the main steps in Algorithm 1 as follows.

For each inner-loop iteration (i.e., n_{in}), we require to store the result of optimized pilot placements. Therefore, a $n_{in} \times M$

Algorithm 1 Optimal Pilot Power Design and Placement Scheme

Input: Number of subcarriers N, number of pilot subcarriers M, number of non-zero taps of the channel S, number of outer-loop iterations n_{out}, number of inner-loop iterations n_{in} and α.

Output: Designed pilot placement $\grave{\mathcal{P}}^*_{\text{soln.}}$ with optimal pilot power u. leftmargin=1mm

- **Step 1** (*Initialization*) Store the results of optimized pilot placement within each iteration of the inner loop in $W \Leftarrow \mathbf{0}_{n_{in}\times S}$, Store the results of best optimized pilot placement of W in $s \Leftarrow 0_{n_{in}}$
- **Step 2** (*Determine the Optimal Pilot Placement*)

1: **for** $i = 0, 1 \ldots, n_{out} - 1$ **do**
2: Randomly generate a $\grave{\mathcal{P}}^* \subset \{1, \ldots, N\}$ satisfying $|\grave{p}_i^* - \grave{p}_j^*| \geq 2Q-1$
3: $\acute{\boldsymbol{p}}_{\text{store}} \Leftarrow 0_M$ {auxiliary pilot vector $\acute{\boldsymbol{p}}_{\text{store}}$ is reset to zero}
4: **for** $j = 0, 1, \ldots, n_{in} - 1$ **do**
5: **if** C.1 does not hold i.e., $\grave{\mathcal{P}}^* \neq \acute{\boldsymbol{p}}_{\text{store}}$ **then**
6: $\acute{\boldsymbol{p}}_{\text{store}} \Leftarrow \grave{\mathcal{P}}^*$
7: **for** $t = 0, 1, \ldots, \beta$ **do**
8: Obtain $\hat{\grave{\mathcal{P}}}^*_{p^*,t}$ based on (45) {determine the optimal pilot placement in the current iteration stage}
9: $\grave{\mathcal{P}}^* \Leftarrow \hat{\grave{\mathcal{P}}}^*_{p^*,t}$ {allocate the designed pilot placement for subsequent storage}
10: **end for**
11: **else**
12: **break**
13: **end if**
14: **end for**
15: $W(i) \Leftarrow \grave{\mathcal{P}}^*$ {store the designed pilot placement as rows in matrix W }
16: $s(i) \Leftarrow \mu_u(\grave{\mathcal{P}}^*)$ {solve (38)}
17: **end for**
18: $r = \arg \min_{b=0,1,\ldots,n_{out}-1} s(b)$ {determine the indices corresponding to the minimum argument to the solutions in *step* 2.16. }
19: **return** $\grave{\mathcal{P}}^*_{\text{soln.}} = W(r)$. {Designed Pilot Placement } leftmargin=1mm

- **Step 3** (*Determine the Optimal Pilot Power*) Substitute the designed pilot placement $\grave{\mathcal{P}}^*_{\text{soln.}}$ into (44) to obtain the optimal pilot symbols.

dimensional all-zero matrix W is initialized for this purpose i.e., $W \Leftarrow \mathbf{0}_{n_{in}\times M}$. Each outer loop iteration involves n_{in} number of inner loop iterations which gives rise to n_{in} number of pilot patterns stored as rows in W; and the optimal pilot placement for that inner loop iterations is subsequently stored in s, determined by the minimizing argument for the optimization problem given in (38). Therefore, s is also initialized as an all-zero vector i.e., $s \Leftarrow \mathbf{0}_{n_{in}}$. First, for **step 2.1**$\sim$ **2.16** the proposed algorithm does the following: 1) obtains the optimized pilot placement and stores in each row of W as indicated by **step 2.14** and 2) solves the corresponding objective problem using (38) and stores in s as indicated by **step 2.15** in Algorithm 1. Second, in **step 2.5**$\sim$ **2.10**, the proposed algorithm updates the solution of the pilot placement for every inner loop iteration if the newly generated pilot locations is checked to see whether it is not a previously generated pilot location. This is checked to make sure the following condition does not hold:

C.1 $\forall x[x \in \grave{\mathcal{P}}^* \longleftrightarrow x \in \acute{\boldsymbol{p}}_{\text{store}}]$, states that $\grave{\mathcal{P}}^* = \acute{\boldsymbol{p}}_{\text{store}}$ i.e., the current solution is the same as the previously generated solution.

If C.1 does not hold then $\grave{\mathcal{P}}^*$ is updated in a set of α entries invariably, where $\alpha < M$ and $\beta = M/\alpha$ is a positive integer. Hence, in the simulation, value will assigned for α. It is worth mentioning that, α influences the number of ''**for**-loop'' (i.e., **step 7**$\sim$ **step 10**) iterations for ever outer loop iteration. Hence, higher values of α means a reduced number of group entry update (i.e., $\beta = M/\alpha$ value will be smaller) but leads to an increased complexity of the algorithm which will be shown shortly in the algorithm complexity analysis. Finally, **step 2.11** $\sim$ **2.13** indicate the absence of new pilot pattern since the currently generated pilot pattern $\grave{\mathcal{P}}^*$ in **step 2.2** is entirely the same as $\acute{\boldsymbol{p}}_{\text{store}}$ and thus executes a break statement inside the inner loop. This is achieved using an auxiliary pilot vector $\acute{\boldsymbol{p}}_{\text{store}}$ which is initialized at the beginning of each iteration of the outer-loop as indicated by **step 2.3**. Taking the expression of **step 2.7**$\sim$**2.10** into consideration, the objective function for the t-th pilot selection can be represented as

$$\hat{\grave{\mathcal{P}}}^*_{p^*,t} = \arg \min_{\substack{\grave{\mathcal{P}}^* \\ \grave{\mathcal{P}}^*(j)=p(j),\ j=0,\ldots,M-1,\ \ j\notin\kappa \\ \{\grave{\mathcal{P}}^*(j),\ j\in\lambda\}\subset\boldsymbol{\varphi}}} \mu_u(\grave{\mathcal{P}}^*) \tag{45}$$

where $\boldsymbol{\varphi} = \{\boldsymbol{\psi}|\boldsymbol{\psi} \subseteq \boldsymbol{\gamma},\ \ \|\boldsymbol{\psi}\|_{\ell_0} = \alpha\}$, such that $\gamma \triangleq N \setminus \{p(j)|j = 0, \ldots, M-1,\ \ j \notin \kappa\}$ and $\kappa \triangleq \{t\alpha - \alpha + 1, t\alpha - \alpha + 2, \ldots, t\alpha\}$.

B. DISCUSSIONS ON COMPUTATIONAL COMPLEXITY

In this section, we evaluate the computational complexity of our proposed pilot design scheme over the exhaustive search. The computational complexity of our proposed scheme, given in Algorithm 1, is primarily dominated to solving both the SOCP optimization problem in (44) for the pilot power assignment in terms of realizing feasible solutions of $\tilde{d}$ and $\tilde{\boldsymbol{u}}$ and the pilot placement objective function given in (45) in terms of the number of iterations. For a given set of pilot subcarriers, **step 2.7**$\sim$ **step 2.10** requires $O\binom{N-M+\alpha}{\alpha}$ operations, to compute the pilot placement, where the binomial coefficient $\binom{N-M+\alpha}{\alpha}$ is the number of ways of picking α unordered outcomes from $(N-M+\alpha)$ possibilities and given explicitly by $\binom{N-M+\alpha}{\alpha} \equiv {}^{(N-M+\alpha)}C_{\alpha} \equiv \frac{(N-M+\alpha)!}{(N-M+\alpha-\alpha)!\alpha!}$. Since the number of iterations in **step 2.7**$\sim$ **step 2.10** is β as given in **step 2.7**, then the effect of the two outer loop iterations is intuitively equivalent to $n_{in} = n_{out} = \beta$ and thus, requires $O(\beta\binom{N-M+\alpha}{\alpha})$ operations. Moreover, by setting $\alpha = 1$, Algorithm 1 becomes efficient in the reduction of computational complexity with **step 2.7**$\sim$ **step 2.10** reduced to $O(N-M+1)$ operations. Consequently, the computational complexity of obtaining the the optimal pilot pattern in the entire **step 2** results to $O(M(N-M+1))$ operations, with $\beta = M/\alpha = M$ where $\alpha = 1$. Subsequently, after obtaining the optimal pilot placement $\hat{\grave{\mathcal{P}}}^*_{\text{soln.}}$ requires the substitution of $\hat{\grave{\mathcal{P}}}^*_{\text{soln.}}$ into (44) for the optimal power design. Hence, (44) requires the search on L columns over M pilot symbols. Hence, its complexity is given as $O((M-1)^{3.5}(L-1)^{2.5})$ operations. Thus, the overall computational complexity becomes $O\big((M-1)^{3.5}(L-1)^{2.5}M(N-M+1)n_{in}n_{out}\big)$

Exhaustive search scheme involves optimization of the pilot power and the search of all possible pilot placement. The complexity of the exhaustive search will take $O\binom{N}{M}$ operations for the seach of pilot placement (where M of N subcarriers are dedicated to transmission of pilots). Hence, the overall computational complexity becomes $O\big((M-1)^{3.5}(L-1)^{2.5}M\binom{N}{M}n_{in}n_{out}\big)$, which is relatively higher compared to our proposed scheme.

IV. DCS-BASED SDMP ALGORITHM WITH APPLICATION TO SPARSE CHANNEL ESTIMATION

In this section, we propose a novel DCS-based stage determined matching pursuit (DSC-SdMP) algorithm for sparse CE over a DS channel, capable of accelerating the reconstruction process of an ensemble of jointly sparse signals from an incomplete and inaccurate estimate of the common support-set.

A. PROPOSED DCS-BASED SDMP (DSC-SDMP) ALGORITHM

Exploiting the joint (common) sparsity structure of an ensemble of Q signals enables conventional greedy pursuit algorithms for joint signal recovery. Hence, an algorithm in our previous work namely, stage determined matching pursuit (SdMP) [26], can be easily applied to our DCS sparse model. Based on SdMP, an algorithm namely, DCS-based SdMP (DCS-SdMP) is proposed in Algorithm 2, which can efficiently capture the joint structure among the signals to improve the performance of DCS recovery for jointly sparse signal ensembles. DCS-SdMP selects $Z(< S)$ multiple support indices per iteration, and adds a pruning step at the end of some latter iterations, after satisfying a sparsity level condition to refine the selected set. Notably, this can reduce

Algorithm 2 Distributed Compressed Sensing-based SdMP (DCS-SdMP) Algorithm

Input: Received signal at pilot subcarriers $[\mathbf{Y}]_{\check{\mathcal{P}}} = ([y]_{\check{\mathcal{P}}_0}, \ldots, [y]_{\check{\mathcal{P}}_{Q-1}})$, measurement matrix $\mathbf{A}$, sparsity level S, number of selected atoms at each iteration $Z \leq S$.

Output: Estimated sparse channel $\{\tilde{\boldsymbol{c}}_q^{CE}\}_{q=0}^{Q-1}$, estimated sparse support set $\boldsymbol{\Phi}$ leftmargin=1mm

- **Step 1** (*Initialization*)
  $\boldsymbol{r}_0 = \mathbf{Y}$ { Initial residue}
  $\Gamma = \emptyset$ {Estimated support set }
  $[M, L]$ =size($\mathbf{A}$)
  $\mathcal{I} = Z$ { Number of indices for each selection $Z(Z \leq S$ and $Z \leq M/S)$}
  $k = 0$ { Iteration index}
  $\boldsymbol{\Phi}_0 = \emptyset_{KL\times Q}$ { Estimated sparse vector}
- **Step 2** (*Procedures*)

while $\|\boldsymbol{r}_k\|_{\ell_2} > \epsilon$ and $k < \min\{K, M/\mathcal{I}\}$ **do**
  $k = k + 1$
  $z_k = \arg\max(\sum_{q=0}^{Q-1} |\mathbf{A}^T[\boldsymbol{r}_{k-1}]_{:,q}|, \mathcal{I})$ { Select indices corresponding to $\mathcal{I}$ largest entries (in magnitude), where $[\boldsymbol{r}_{k-1}]_{:,q}$ denotes the q-th column of $\boldsymbol{r}_{k-1}$}
  $\Lambda_k = \Gamma_{k-1} \cup z_k$ {Merge support}
  if $|\Lambda_k| \neq |\Gamma_{k-1}|$ **then**
    if $|\Lambda_k| \leq K$ **then**
      $\acute{\mho}_{\Lambda_k}\acute{\tilde{\boldsymbol{c}}}_{\Lambda_k}^{CE} = \sum_{q=0}^{Q-1} \mathbf{A}_{\Lambda_k}^{\dagger}[\mathbf{Y}]_{:,q}$ {Channel estimation by LS }
      $\Gamma_k = \Lambda_k$
      $\boldsymbol{r}_k = \sum_{q=0}^{Q-1} \mathbf{P}_{\Lambda_k}^{\perp}[\mathbf{Y}]_{:,q}$
    else
      $\acute{\mho}_{\Lambda_k}\acute{\tilde{\boldsymbol{c}}}_{\Lambda_k}^{CE} = \sum_{q=0}^{Q-1} \mathbf{A}_{\Lambda_k}^{\dagger}[\mathbf{Y}]_{:,q}$ {Channel estimation by LS }
      $\Gamma_k = \mathtt{supp}(\mho_{\Lambda_k}^{S}\tilde{\boldsymbol{c}}_{\Lambda_k}^{S_{CE}})$ {$\mho_{\Lambda_k}^{S}\tilde{\boldsymbol{c}}_{\Lambda_k}^{S_{CE}}$ is the best S-sparse approximation of $\mho_{\Lambda_k}\tilde{\boldsymbol{c}}_{\Lambda_k}^{CE}$ }
      $\acute{\mho}_{\Gamma_k}\acute{\tilde{\boldsymbol{c}}}_{\Gamma_k}^{CE} = \sum_{q=0}^{Q-1} \mathbf{A}_{\Gamma_k}^{\dagger}[\mathbf{Y}]_{:,q}$ {Channel estimation by least squares (LS) }
      $\boldsymbol{r}_k = \sum_{q=0}^{Q-1} \mathbf{P}_{\Gamma_k}^{\perp}[\mathbf{Y}]_{:,q}$
    end if
  else
    break
  end if
end while
$\hat{\mho}_{\Gamma_k}\hat{\tilde{\boldsymbol{c}}}_{\Gamma_k}^{CE} = \acute{\mho}_{\Gamma_k}\acute{\tilde{\boldsymbol{c}}}_{\Gamma_k}^{CE}$, $\hat{\mho}_{\Gamma_k}\hat{\tilde{\boldsymbol{c}}}_{\Gamma_k}^{CE}|_{\Gamma_k^c} = 0$ {Γ_k^c is the complementary set of indices to Γ_k. }
obtain $\hat{\tilde{\boldsymbol{c}}}_{\Gamma_k}^{CE}$ from $\hat{\mho}_{\Gamma_k}\hat{\tilde{\boldsymbol{c}}}_{\Gamma_k}^{CE}$
return $\hat{\tilde{\boldsymbol{c}}}_{\Gamma_k}^{CE}, \Gamma_k$.

the probability of estimating incorrect support elements and can maximally satisfy the l_2-norm of the residual error vector based stopping criteria with fewer iterations. Thereby, increasing the probability of further accelerating the algorithm.

The pseudocode of the DCS-SdMP algorithm is shown in Algorithm 2. For simplicity, we address the essential part of the proposed algorithm without giving the obvious, redundant illustration. The following two stages are applied by the algorithm which it iterates until convergence:

1) FIRST STAGE NON-PRUNING TO EXPAND THE ESTIMATED COMMON SUPPORT SET

In every iteration, $Z(\leq S)$ maximally correlated common sparse support indices with the vector residual are selected $z_k = \arg\max(\sum_{q=0}^{Q-1} |\mathbf{A}^T[\boldsymbol{r}_{k-1}]_{:,q}|, \mathcal{I})$ and appended to the list of common support set $\Lambda_k = \Gamma_{k-1} \cup z_k$, where $\arg\max(\cdot)$ denotes the maximizing argument. Notably, the choice of Z is to be constrained to $Z = \{n_i \neq 1, i = 1, 2, \ldots, S-1\}$, where S is the level of joint sparsity. Subsequently, after adding the selected common support indices to the list of identified common support Λ_k the ensemble of signals is estimated by solving the LS problem which finds a unique solution to $\acute{\mho}_{\Lambda_k}\acute{\tilde{\boldsymbol{c}}}_{\Lambda_k}^{CE} = \sum_{q=0}^{Q-1} \mathbf{A}_{\Lambda_k}^{\dagger}[\mathbf{Y}]_{:,q}$, where $\mathbf{A}^{\dagger} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ denotes the pseudoinverse of matrix $\mathbf{A}$. Then, the residual of $\mathbf{Y}$ which is $\boldsymbol{r}^k = \mathbf{Y} - \mathbf{A}_{\Lambda_k}(\acute{\mho}_{\Lambda_k}\acute{\tilde{\boldsymbol{c}}}_{\Lambda_k}^{CE}) = (1 - \mathbf{A}_{\Lambda_k}\mathbf{A}_{\Lambda_k}^{\dagger})\mathbf{Y} = \sum_{q=0}^{Q-1} \mathbf{P}_{\Lambda_k}^{\perp}[\mathbf{Y}]_{:,q}$ (where, $\mathbf{P}_{\Lambda_k} = \mathbf{A}_{\Lambda_k}\mathbf{A}_{\Lambda_k}^{\dagger}$ and $\mathbf{P}_{\Lambda_k}^{\perp} = 1 - \mathbf{P}_{\Lambda_k}$) is updated. Using stage switching, the algorithm combine the selection of Z multiple columns per iteration with a pruning step, to manage a S element support throughout the rest of the iterations.

2) SECOND STAGE PRUNING TO ELIMINATE UNPROMISING INDICES TO MAINTAIN THE REQUIRED SPARSITY LEVEL

Thus, in every k-th iteration, if the decision $|\Lambda_k| > K$ holds true, then $|\Lambda_k| \leq K + Z$ columns of matrix A will be tested to subsequently select S reliable common support indices after pruning. The reliability property is captured by finding the subspace projection coefficients $\acute{\mho}_{\Gamma_k}\acute{\tilde{\boldsymbol{c}}}_{\Gamma_k}^{CE} = \sum_{q=0}^{Q-1} \mathbf{A}_{\Gamma_k}^{\dagger}[\mathbf{Y}]_{:,q}$ (otherwise called the LS solution) and selecting S common support indices with the largest projection coefficient. For this purpose we introduce the $\mathtt{supp}(\cdot)$ function, where $\Gamma_k = \mathtt{supp}(\mho_{\Lambda_k}^{S}\tilde{\boldsymbol{c}}_{\Lambda_k}^{S_{CE}})$ denotes the the set Γ_k which contains the best S support indices of $\mho_{\Lambda_k}^{S}\tilde{\boldsymbol{c}}_{\Lambda_k}^{S_{CE}}$. The residual error is then calculated as $\boldsymbol{r}_k = \sum_{q=0}^{Q-1} \mathbf{P}_{\Gamma_k}^{\perp}[\mathbf{Y}]_{:,q}$. This operations are performed until the l_2-norm of the residual falls below a prespecified threshold i.e., $\|\boldsymbol{r}_k\|_{\ell_2} \leq \epsilon$.

B. DISCUSSIONS ON COMPUTATIONAL COMPLEXITY

In this section, we briefly discuss the computational complexity of our proposed scheme in terms of floating-point operations (FLOPs). Notably, a complex summation requires two real summations, and a complex multiplication requires four real multiplications and two real summations, we, however, count each operation as one FLOP. Following the analysis in [26], the operation of each step can be summarized as

Correlation vector: The DCS-SdMP algorithm perfomaes a matrix multiplication $\mathbf{A}^T[r_{k-1}]_{:,q}$, which requires $(2M - 1)L$ to be determined Q times. Hence, requires a total of $(2M - 1)LQ$ flops (i.e., $O(MLQ)$).

Identification: $\mathbf{A}^T[r_{k-1}]_{:,q}$ needs to be sorted to obtain the $\mathcal{S}$ best indices, and requires performing $LS - S(S - 1)/2$ operations a total of Q times, given a complexity of

$QLS - QS(S-1)/2$ flops (i.e., $O(QLS)$).

Channel estimation: To obtain the least square solution for $\sum_{q=0}^{Q-1} \mathbf{A}_{\Lambda_k}^{\dagger}[\mathbf{Y}]_{:,q}$ requires performing the **QR** factorization of $\mathbf{A}_{\Lambda_k}^{\dagger}$ ($\mathbf{A}_{\Lambda_k}^{\dagger} = \mathbf{QR}$), hence, the solution $\acute{\mho}_{\Lambda_k}\acute{\tilde{\boldsymbol{c}}}_{\Lambda_k}^{CE}$ leads to a cost of $O(k^2MQ)$ [26].

Pruning: The pruning is similar to the identification step in terms of complexity. The estimated coefficients are revised in decreasing order of magnitude and the algorithm selects the first K at a cost of $O(QK \log K)$

Residual update: To obtain the residual $\boldsymbol{r}^k = \mathbf{Y} - \mathbf{A}_{\Lambda_k}(\acute{\mho}_{\Lambda_k}\acute{\tilde{\boldsymbol{c}}}_{\Lambda_k}^{CE}) = (1 - \mathbf{A}_{\Lambda_k}\mathbf{A}_{\Lambda_k}^{\dagger})\mathbf{Y} = \sum_{q=0}^{Q-1} \mathbf{P}_{\Lambda_k}^{\perp}[\mathbf{Y}]_{:,q}$ (where, $\mathbf{P}_{\Lambda_k} = \mathbf{A}_{\Lambda_k}\mathbf{A}_{\Lambda_k}^{\dagger}$ and $\mathbf{P}_{\Lambda_k}^{\perp} = 1 - \mathbf{P}_{\Lambda_k}$) requires matrix-vector multiplication $\mathbf{A}_{\Lambda_k}(\acute{\mho}_{\Lambda_k}\acute{\tilde{\boldsymbol{c}}}_{\Lambda_k}^{CE})$ with a complexity of $(2Sk-1)MQ$ flops followed by the subtraction which requires MQ flops i.e., $2SkMQ - MQ - MQ = 2SkMQ$. Hence, $O(SkMQ)$.

In other words, the complexity of the overall k-th iteration requires $MLQ + QLS + k^2MQ + QK \log K + SkMQ$. If the algorithm finishes in K iteration the we have $KMLQ + KQLS + k^2MQK + QK^2 \log K + KSkMQ = O(KMLQ)$, since $K < M < L < Q$.

V. SIMULATION RESULTS

In this section, we demonstrate the effectiveness of our proposed DCS based CE scheme and the proposed pilot design scheme with the conventional schemes operating over DS channel through simulation studies. For the simulations, we use the OFDM system parameters that conform with the LTE standard [41]. The entire experiments were performed using MATLAB v9.5 (Release 2018b) 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 a data sequence modulated by quadrature phase shift keying (QPSK) is considered with $N = 512$ subcarriers, cyclic prefix length $L_{CP} = N/8 = 64$, and subcarrier interval $\bigtriangleup f = 15$ KHz. In the simulations, we normalize the total transmission power of each OFDM frame by setting $\lambda_{\text{sum}} = 1$. Specifically, the parameters involved in the pilot design are set as follows. The outer loop iterations n_{out} and the inner loop iterations n_{in} are set to 1500 and 20, respectively. Unless otherwise mentioned, we set $\alpha = 2$ for the set of group update entries. We set $\lambda_{\text{sum}} = 1$, which means that the sum power of pilot subcarriers is normalized. However, the power allocation among pilot symbols is not uniformly constrained but set with a peak power constraint of $\lambda_{\text{peak}} = 0.14$ and with a power threshold of $\lambda_{\text{Th}} = 0.04$ for the design and placement of pilot symbols. The length L of the sparse DS channel $h_{n,l}$, is modeled with $L = 64$ taps, where S-positions are nonzero (dominant) taps of the channel whose elements are randomly distributed according to $\mathcal{CN}(0, \frac{1}{S})$. We compare the different pilot design schemes for DS channels that conform with Jakes' Doppler profile which assumes the uniform-angle-of-arrival assumption, while taking the CE-BEM order of $Q = 3$ (and $Q = 7$). Hence, $\nu_{\max} \leq (Q-1)/2$ must be satisfied to guarantee that the DS channel taps are well captured in the CE-BEM by the (known) Fourier basis functions. Here, $\nu_{\max}$ denotes the normalized maximum Doppler shift. Specifically, assuming a maximum carrier frequency $f_c = 3$ GHz, and two different vehicular speed of around $v = 350$ km/h and $v = 500$ km/h (unless otherwise stated), we obtain $\nu_{\max} = \frac{f_c v}{c\bigtriangleup f} \approx 0.065$ and $\nu_{\max} = \frac{f_c v}{c\bigtriangleup f} \approx 0.093$ normalized maximum Doppler shift, respectively, with $c = 3.0 \times 10^8$ m/s being the speed of light. Noting that: 1 km/m = 0.278 m/s and that 0.065 and 0.093 both satisfy the $\nu_{\max} \leq (Q-1)/2$ condition, with $Q = 3$ (unless otherwise stated). We set the number of multiple OFDM symbols that are jointly estimated to be $K = 3$. We set the average number of pilot signals for the k OFDM symbol as $M/K = 72/3 = 24$ pilot signals, where $M = 72$. We use two different pilot set based on the CE-BEM order Q. First, for a CE-BEM order of $Q = 3$, the average number of the pilot subcarrier (i.e., for both pilot signals and guard pilots) in our proposed joint multi-symbol estimation scheme is $P = (2Q-1)M/K = 5 \times (72/3) = 120$ (where guard pilots uses $(2Q-2)M/K = 4 \times (72/3) = 96$ subcarriers out of the 120 average number of pilot subcarrier). Second for a CE-BEM order of $Q = 7$, the average number of the pilot subcarrier (i.e., for both pilot signals and guard pilots) in our proposed joint multi-symbol estimation scheme is $P = (2Q-1)M/K = 13 \times (72/3) = 312$ (where guard pilots uses $(2Q-2)M/K = 12 \times (72/3) = 288$ subcarriers out of the 312 average number of pilot subcarrier). For fair comparison, we assume the same pilot overheads in each estimation scheme under evaluation, except for the cubic-spline-interpolation based LS CE higher pilots. For the estimation of the sparse channels, an Oracle estimator which has prior knowledge of the location of the non-zero channel taps is simulated as a reference and employs pilots solely to deduce the channel tap values by LS estimation. As the exact recovery of sparse signals from compressed measurements would be highly impracticable in the noisy settings, we employ the normalized mean square error (NMSE) as a performance evaluation metric defined as

$$\text{NMSE}_{\boldsymbol{h}[n]}(\text{dB}) = 10\log_{10}\left(\frac{\sum_{n=1}^{N}\|\boldsymbol{h}[n] - \hat{\boldsymbol{h}}[n]\|_{\ell_2}^2}{\sum_{n=1}^{N}\|\boldsymbol{h}[n]\|_{\ell_2}^2}\right), \quad (46)$$

where $\hat{\boldsymbol{h}}[n]$ denotes the estimate of the sparse vector $\boldsymbol{h}[n]$ at instant n. The NMSE performance of the signal recovery of the proposed algorithm is presented as a function of the SNR, where the SNR (given in dB) is defined as

$$\text{SNR(dB)} = 10\log_{10}\frac{\|\boldsymbol{A}\boldsymbol{h}\|_{\ell_2}^2}{\|\boldsymbol{V}\|_{\ell_2}^2}, \quad (47)$$

since the system model is expressed as $\boldsymbol{Y} = \mathbf{A}\boldsymbol{h} + \mathbf{V}$, where $\boldsymbol{V}$ is the noise vector whose elements are generated from Gaussian $\mathcal{N}(\mathbf{0}, 10^{-\text{SNR}/10})$. In order to recover the sparse channels, we used the the DCS-SOMP [5], ASA-BOMP [11],

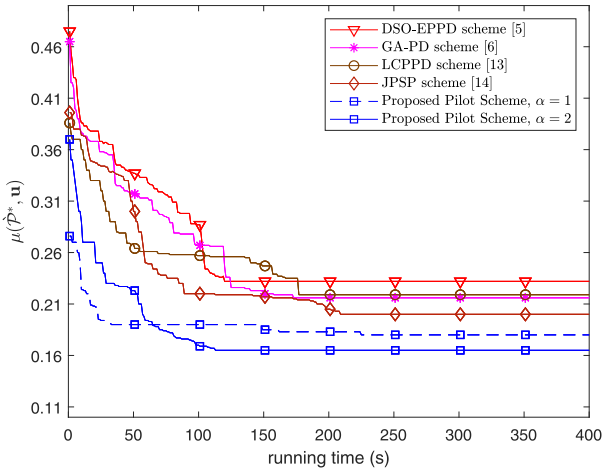

FIGURE 3. Comparison of $\mu(\grave{\mathcal{P}}^*, u)$ performance of the proposed pilot design algorithm and the conventional pilot design schemes with $Q = 3$ and $\nu_{\max} = 0.065$ (where $f_c = 3$ GHz and $\Delta f = 15$ KHz) at SNR= 20 dB.

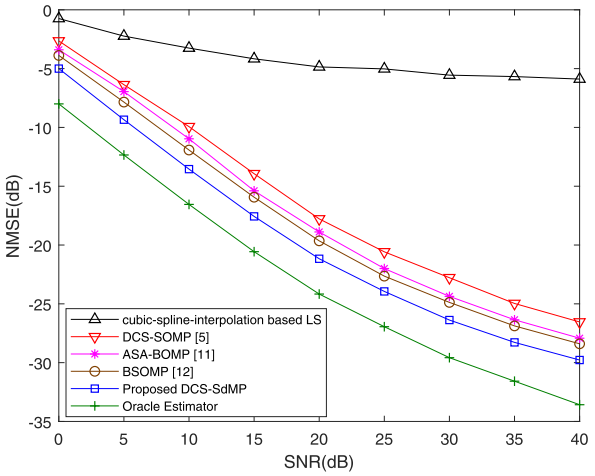

FIGURE 4. Comparison of the NMSE performance of the proposed DSC-SdMP algorithm and the conventional reconstruction algorithms with $S = 6$, $Q = 3$, $\nu_{\max} = 0.065$ (where $f_c = 3$ GHz and $\Delta f = 15$ KHz).

BSOMP [12] and the proposed DCS-SdMP algorithms and evaluate the average performance over 9500 independent trials. Additionally, we also evaluate the performance of our proposed pilot design scheme with the joint pilot symbol and pilot pattern design algorithm in [14].

B. PARAMETER SETTINGS FOR THE PILOT DESIGN SCHEME

In the first simulation experiment depicted as Fig.3, we investigate the impact of various settings of the parameter $\alpha \in [1, 2]$ for our pilot design scheme presented in section III-A with the corresponding objective function $\mu(\grave{\mathcal{P}}^*, \boldsymbol{u})$ (i.e., mutual coherence of the designed measurement matrix), based on the CPU runtime.[4] Based on the proposed pilot design scheme, one can indeed, discursively deduce its computational complexity according to Section III-B as one that increases with increasing values of α settings−used to tune the set of group update entries. In this experiment, we consider five pilot design schemes, i.e., discrete stochastic optimization (DSO)-based effective pilot pattern design scheme in [5] −we denote as DSO-EPPD scheme, the generic algorithm pilot design scheme in [6] (denoted as GA-PD scheme), the low coherence pilot pattern design scheme in [13] (denoted as LCPPD scheme), the joint pilot symbol and pilot-design-pattern scheme in [14] (denoted as JPSP scheme) and our proposed pilot design scheme presented as Algorithm 1. Fig. 3, plots the objective function $\mu(\grave{\mathcal{P}}^*, \boldsymbol{u})$ as a function of the CPU runtime, under the number of the strong paths $S = 6$, the CE-BEM order $Q = 3$, the normalized Doppler $\nu_{\max} = 0.065$ (i.e., at the speed $v = 350$ km/h) with $f_c = 3$ GHz and $\Delta f = 15$ KHz, and the average number of the pilot subcarrier for the joint multi-symbol estimation $P = 120$. It is worth mentioning that the mutual coherence is a more practical metric for assessing the DCS recovery properties of designed measurement matrices. Since, the DCS reconstruction performance improves with lower values of $\mu(\grave{\mathcal{P}}^*, \boldsymbol{u})$, in terms of NMSE [33], [42]. In Fig. 3, we observe that the proposed pilot design scheme with group update settings of $\alpha = 2$ can minimize $\mu(\grave{\mathcal{P}}^*, \boldsymbol{u})$ than that with group update settings $\alpha = 1$, but at the cost of increased computational complexity (i.e., it requires a running time of approximately 120 s for an optimal design). For power-efficient devices, using the setting of $\alpha = 1$ may be considered a better option as it can minimize $\mu(\grave{\mathcal{P}}^*, \boldsymbol{u})$ for the low algorithm runtime (say below 30 s) than with a setting of $\alpha = 2$, but at the cost of a decreased reconstruction performance. The reason is that using the setting of $\alpha = 2$ provides a better minimization of $\mu(\grave{\mathcal{P}}^*, \boldsymbol{u})$, albeit at greater computational expense. Nonetheless, since we usually offline train the joint pilot placement and symbol before the transmission, using the setting of $\alpha = 2$ will be a better option and will not bring any additional complexity to the CE at the receiver. Additionally, it is clearly seen that our proposed algorithm at both settings of $\alpha \in [1, 2]$ is superior to other pilot design schemes, based on the minimum mutual coherence criteria. Thus, in the remaining experiments, a setting of $\alpha = 2$ is used for our proposed algorithm; and all the pilot design algorithms under investigation are given sufficient time to training (> 220 s) for fairness.

[4]The running time is used rather than the number of iterations for comparison, as the complexity in each iteration varies differently for the algorithms. Nonetheless, it is worth mentioning that different hardware configurations may result in different running time measurements.

C. PERFORMANCE OF THE PROPOSED DCS-SDMP ALGORITHM

In our second experiment, depicted in Fig. 4, we show the NMSE comparison of the proposed DSC-SdMP algorithm with the DCS-SOMP, ASA-BOMP and BSOMP algorithms for the joint recovery of signals ensembles. This is achieved using an equal-power, non-uniformly spaced pilot pattern, under the number of the strong paths $S = 6$, the CE-BEM

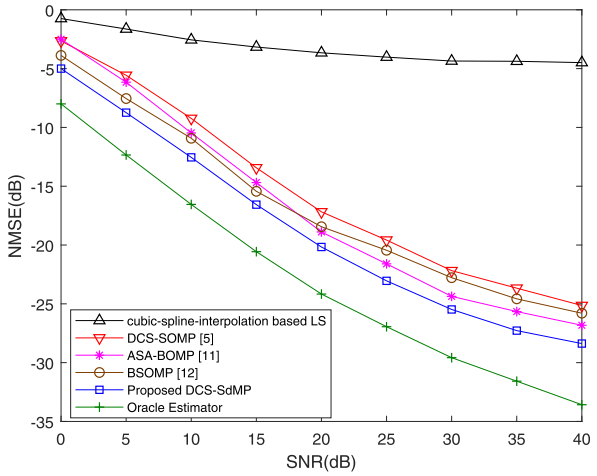

FIGURE 5. Comparison of the NMSE performance of the proposed DSC-SdMP algorithm and the conventional reconstruction algorithms with $S = 6$, $Q = 3$, $\nu_{\max} = 0.093$ (where $f_c = 3$ GHz and $\Delta f = 15$ KHz).

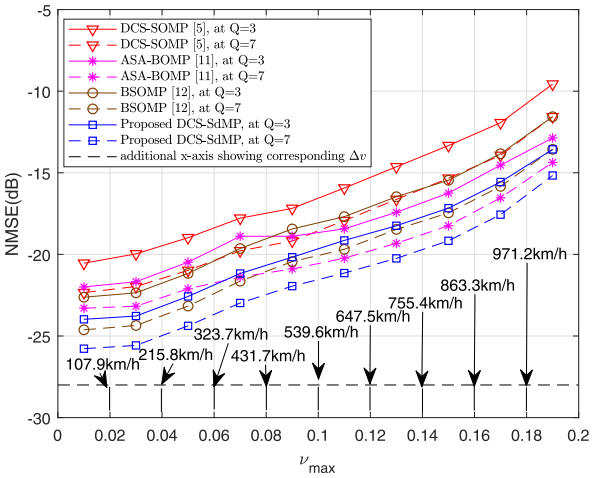

FIGURE 6. NMSE performance versus $\nu_{\max}$ at SNR= 20 dB with $S = 6$, $f_c = 3$ GHz, $\Delta f = 15$ KHz, $Q = 3$ and $Q = 7$ for the proposed algorithm and the conventional reconstruction algorithms.

order $Q = 3$, $f_c = 3$ GHz, $\Delta f = 15$ KHz, the normalized Doppler $\nu_{\max} = 0.065$ (i.e., at the speed $v = 350$ km/h) and the average number of the pilot subcarrier for the joint multi-symbol estimation of $P = 120$. In this experiment, we also simulate the cubic-spline-interpolation based LS CE with $M = 24$ pilots in each OFDM symbol, as the LS proves to be the best for traditional equispaced CE. From Fig. 4, we can see that, the DCS algorithms yield better NMSE performance than the solution obtained by using cubic-spline-interpolation based LS. However, DCS-SdMP achieves a better NMSE performance compared to DCS-SOMP [5], ASA-BOMP [11], and BSOMP [12]. For example, at SNR= 20 dB, the NMSE for cubic-spline-interpolation based LS, DCS-SOMP, BSOMP, ASA-BOMP, and the proposed DCS-SdMP algorithms are -4.86 dB, -17.78dB, -18.89dB, -19.64dB, and -21.17dB, respectively. It is clear that the proposed reconstruction algorithm can acquire very exact CIR estimation when compared to the other different methods of CE under investigation.

In our third experiment, depicted in Fig. 5, we carry out a similar comparison. We plot the NMSE versus SNR for all the reconstruction algorithms under investigation, for higher normalized Doppler shift of $\nu_{\max} = 0.093$ (i.e., at the speed $v = 500$ km/h). According to results in Fig. 5, a similar relative superiority of the proposed DCS-SdMP reconstruction algorithm can be observed. However, the performances of all the curves are degraded relative to the scenario for $\nu_{\max} = 0.065$ in Fig. 4. This is due to the larger modeling error of the CE-BEM, associated with the higher modeling error when Doppler frequency increases. Moreover, given the same sparsity level, the NMSE of DCS-SdMP is closer to the Oracle estimator than that of the other algorithms under investigation.

In our fourth experiment, Fig. 6 depicts a comparison of the NMSE versus the normalized Doppler shift, $\nu_{\max}$ under $S = 6$, $f_c = 3$ GHz, $\Delta f = 15$ KHz, $Q = 3$ at SNR $= 20$ dB, in order to examine how Doppler shift impacts the performance of CE. To further investigate the relationship between NMSE and model order Q (i.e., the number of basis functions), we have included in the same simulation environment a similar comparison of the entire algorithms under investigation under the same settings but with $Q = 7$ in Fig. 6. From Fig. 6, we can see that the resulting NMSE degrades in the high $\nu_{\max}$ regime. Hence, we could better comprehend why normalized Doppler shift $\nu_{\max}$ values are so critical in the comparative analysis of NMSE of the CE-BEM for DS channel since a large $\nu_{\max}$ leads to a large channel modeling error as a result of failure to capture the channel time-variation. However, using the proposed algorithm is shown to be more accurate in estimating the channel. According to the simulation results presented herein, increasing the BEM order from $Q = 3$ to $Q = 7$ can improve the NMSE performance of CE for all algorithms. Nonetheless, a similar relative superiority of our proposed reconstruction algorithm can be observed among algorithms under investigation. The reason is that with a high order of the basis functions Q, more mutually orthogonal elements (or basis functions) are employed to fit the DS channel, which leads to better NMSE performance. However, higher-order basis functions Q achieves better CE performance at the expense of substantially increased computational complexity. This is however, obvious since the complexity of the proposed algorithm is $O(KMLQ)$. Thus, a proper selection of the value of the BEM order Q is required to balance the reliability and transmission efficiency of practical wireless communication systems over DS channel.

In our fifty experiment depicted as Fig. 7, a rough estimate of the computational complexity of the proposed reconstruction algorithm based on the CPU runtime is presented. Fig. 7 plots the running time of each reconstruction algorithm as a function of sparsity under $\nu_{\max} = 0.065$ (i.e., for speed $v = 350$ km/h), $f_c = 3$ GHz, $\Delta f = 15$ KHz and $Q = 3$ at SNR= 20 dB. Generally, we observe that BSOMP and ASA-BOMP have the highest running time and DCS-SOMP and

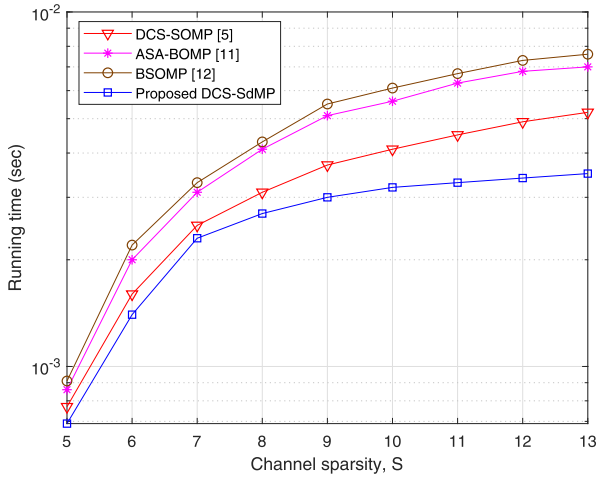

FIGURE 7. Running time as a function of sparsity S of the proposed DCS-SdMP algorithm and the conventional conventional reconstruction algorithms with $Q = 3$ and $\nu_{\max} = 0.065$ (where $f_c = 3$ GHz and $\Delta f = 15$ KHz) at SNR= 20 dB.

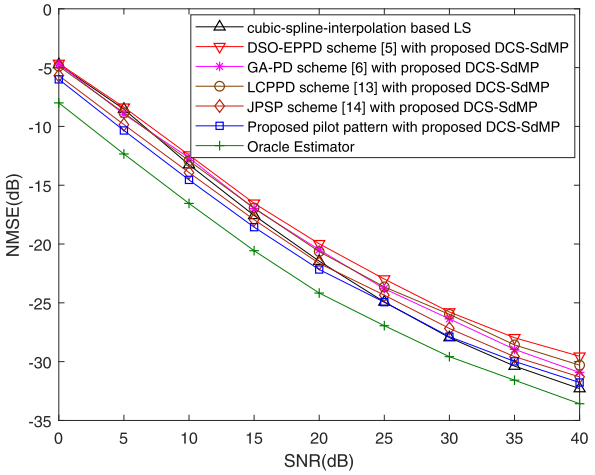

FIGURE 8. Comparison of the NMSE performance of the proposed pilot design scheme and the conventional pilot placement schemes with $S = 6$, $Q = 3$ and $\nu_{\max} = 0.065$ (where $f_c = 3$ GHz and $\Delta f = 15$ KHz) $\alpha = 2$.

DCS-SdMP have the lowest running time among algorithms under investigation. It is worth noting that the DCS-SdMP algorithm selects $Z(\leq S)$ maximally correlated common sparse support indices per iteration and incorporates a pruning step in some subsequent iterations. This increases the probability of selecting the correct common support indices per iteration, which has the benefit of maximally reducing the ℓ_2-norm of the residual error vector based stopping criteria to require fewer iterations and thus a lower running time. Notably, there are K total iteration in $O(KMLQ)$ for the DCS-SdMP algorithm. Thus, we can discursively deduce the superiority of the DCS-SdMP algorithm by examining the results of Fig. 7 as one that requires fewer iterations and, thus, a better total computational complexity. When compared to other reconstruction algorithms, we observe that DCS-SdMP exhibits better complexity for high S.

D. PERFORMANCE OF THE PROPOSED PILOT POWER AND PLACEMENT SCHEME

In our sixth experiment, we consider five pilot design schemes, i.e., the DSO-EPPD scheme [5], the GA-PD scheme [6], the LCPPD scheme [13], PSP scheme [14] and our proposed pilot design scheme presented in Algorithm 1. We also simulate the cubic-spline-interpolation based LS CE with $M = 256$ pilots in each OFDM symbol. It is worth noting that the pilot pattern design schemes in [5], [6], and [13] estimates the jointly sparse CE-BEM coefficient vectors over the DS channel based on the assumption that the pilot symbols are equally-powered. While the pilot scheme in [14] estimates the jointly sparse CE-BEM coefficient vectors over the DS channel using the joint optimal pilot placement and pilot symbol allocation. Our approach jointly optimized the pilot sequences over both the pilot symbol values and their placement allocable to the pilot clusters for the estimation of the jointly sparse CE-BEM coefficient vectors over the DS channel. In Fig. 8, we plot the NMSE versus SNR for all different pilot design schemes, under the number of the strong paths $S = 6$, $\alpha = 2$, the CE-BEM order $Q = 3$, the normalized Doppler $\nu_{\max} = 0.065$ (i.e., at the speed $v = 350$ km/h) with $f_c = 3$ GHz and $\Delta f = 15$ KHz, and the average number of the pilot subcarrier for the joint multi-symbol estimation $P = 120$. In order to recover the jointly sparse CE-BEM coefficient vectors over the DS channel, we used the proposed DCS-SdMP algorithms and evaluate the average performance over 9500 independent trials. To get an idea of the potential NMSE gain realized by exploiting the proposed pilot design scheme and the proposed DCS-SdMP reconstruction algorithm, we have included the Oracle estimator as a reference, which has prior knowledge of the location of the nonzero channel taps and employs pilots solely to deduce the channel tap values by LS estimation. We observe that Algorithm 1 (i.e., our proposed pilot design scheme) and the JPSP scheme in [14] outperforms the methods of [5], [6], and [13]. Nonetheless, our proposed pilot design scheme maintains its superiority over the JPSP scheme in [14] under the low and high SNR regime, which verifies the effectiveness of our proposed method. Moreover, it is observed that the performance of LS which utilizes 256 equally-spaced pilot signals reaches the performance of the proposed scheme at SNR 25 dB. Hence, the proposed scheme can conserve $256 - P = 136$ pilot subcarriers under the same CE performance (where $P = 120$), leading to $((256 - P)/N) \times 100 = (136/512) \times 100 = 26.56\%$ improvement in spectrum efficiency (where P and N denote the number of pilot and guard pilot subcarriers, and total number of OFDM subcarriers, respectively).

VI. CONCLUSION

In this paper, we study the pilot cluster scheme in the frequency domain for DS channel estimation and proposes a new joint design pilot scheme and subsequently introduce a DCS-SdMP algorithm for pursuing efficiency in reconstructing jointly sparse signals ensemble for OFDM systems. Taking benefit from the CE-BEM and the channel sparsity in

the delay-Doppler domain (i.e., two-dimensional (2D) channel coefficients), the proposed pilot design scheme identifies the optimal pilot placement and values for each pilot cluster and can lead to an accurate estimation of the jointly sparse CE-BEM coefficient vectors. Specifically, this approach decomposes the problem of jointly optimizing the pilot sequences over both the pilot placement and values allocable to the pilot clusters into identifiable sequential formulations. It is shown through performance analysis and simulation results that the proposed pilot design scheme can provide optimal pilot signal allocable to the pilot cluster which can significantly improve the estimation accuracy of the OFDM system over DS channel. To ensure a satisfactory estimation, we have further proposed a new DCS recovery algorithm based on stage determined matching pursuit (SdMP) namely, DCS-SdMP that possesses the advantage of multiple common-support indices selections per iteration and then adds a backtracking process (at a later iteration) to subsequently refine the previously selected common support indices. The simulation results demonstrate that the proposed DCS-SdMP algorithm with the new joint pilot design scheme presents a better NMSE performance for the CE over the DS channel when compared with existing schemes. Additionally, the proposed scheme can attain 26.56% improvement in spectrum efficiency with similar CE performance when compared with the conventional LS CE method that does not exploit sparsity.

REFERENCES

[1] A. N. Uwaechia and N. M. Mahyuddin, "A review on sparse channel estimation in ofdm system using compressed sensing," *IETE Tech. Rev.*, vol. 34, no. 5, pp. 514–531, 2017.

[2] A. N. Uwaechia and N. M. Mahyuddin, "Collaborative framework of algorithms for sparse channel estimation in OFDM systems," *J. Commun. Netw.*, vol. 20, no. 1, pp. 9–19, Feb. 2018.

[3] Z. Sheng, H. D. Tuan, H. H. Nguyen, and Y. Fang, "Pilot optimization for estimation of high-mobility OFDM channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 8795–8806, Oct. 2017.

[4] T. Cooklev, H. Dogan, R. J. Cintra, and H. Yildiz, "A generalized prefix construction for OFDM systems over quasi-static channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 3684–3693, Oct. 2011.

[5] P. Cheng *et al.*, "Channel estimation for OFDM systems over doubly selective channels: A distributed compressive sensing based approach," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4173–4185, Oct. 2013.

[6] X. Ma, F. Yang, S. Liu, J. Song, and Z. Han, "Sparse channel estimation for MIMO-OFDM systems in high-mobility situations," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6113–6124, Jul. 2018.

[7] Y. Choi and J. H. Lee, "Pilot structure design to increase wireless channel capacity for high-speed railway," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2017, pp. 1–5.

[8] X. Kuai, H. Sun, S. Zhou, and E. Cheng, "Impulsive noise mitigation in underwater acoustic OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8190–8202, Oct. 2016.

[9] X. Lin, S. Wu, C. Jiang, L. Kuang, J. Yan, and L. Hanzo, "Estimation of broadband multiuser millimeter wave massive MIMO-OFDM channels by exploiting their sparse structure," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3959–3973, Jun. 2018.

[10] H. Nguyen-Le and T. Le-Ngoc, "Pilot-aided joint CFO and doubly-selective channel estimation for OFDM transmissions," *IEEE Trans. Broadcast.*, vol. 56, no. 4, pp. 514–522, Dec. 2010.

[11] X. Ma, F. Yang, S. Liu, W. Ding, and J. Song, "Structured compressive sensing-based channel estimation for time frequency training OFDM systems over doubly selective channel," *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 266–269, Apr. 2017.

[12] Q. Qin, L. Gui, B. Gong, X. Ren, and W. Chen, "Structured distributed compressive channel estimation over doubly selective channels," *IEEE Trans. Broadcast.*, vol. 62, no. 3, pp. 521–531, Sep. 2016.

[13] X. Ren, M. Tao, and W. Chen, "Compressed channel estimation with position-based ICI elimination for high-mobility SIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6204–6216, Aug. 2016.

[14] X. Ren, X. Shao, M. Tao, and W. Chen, "Compressed channel estimation for high-mobility OFDM systems: Pilot symbol and pilot pattern design," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 4553–4557.

[15] S. Gao, M. Zhang, and X. Cheng, "Precoded index modulation for multi-input multi-output OFDM," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 17–28, Jan. 2018.

[16] X. Liu, H.-H. Chen, X. Wang, and W. Meng, "Time domain precoding for OFDM/OFDMA systems without cyclic prefix," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 5510–5514, Jun. 2018.

[17] L. Luo, B. Li, Y. Yu, X. Xu, K. Soga, and J. Yan, "Time and frequency localized pulse shape for resolution enhancement in STFT-BOTDR," *J. Sensors*, vol. 2016, Dec. 2015, Art. no. 3204130.

[18] N. Aboutorab, W. Hardjawana, and B. Vucetic, "A new iterative Doppler-assisted channel estimation joint with parallel ICI cancellation for high-mobility MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 61, no. 4, pp. 1577–1589, May 2012.

[19] J. K. Tugnait, S. He, and H. Kim, "Doubly selective channel estimation using exponential basis models and subblock tracking," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1275–1289, Mar. 2010.

[20] L. Deng, Z. Chen, and Y. Zhao, "Basis expansion model for channel estimation in LTE-R communication system," *Digit. Commun. Netw.*, vol. 2, no. 2, pp. 92–96, 2016.

[21] M. K. Tsatsanis and G. B. Giannakis, "Subspace methods for blind estimation of time-varying FIR channels," *IEEE Trans. Signal Process.*, vol. 45, no. 12, pp. 3084–3093, Dec. 1997.

[22] D. I. Shuman, P. Vandergheynst, D. Kressner, and P. Frossard, "Distributed signal processing via Chebyshev polynomial approximation," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 4, pp. 736–751, Dec. 2018.

[23] H. A. Cirpan and M. K. Tsatsanis, "Maximum likelihood blind channel estimation in the presence of Doppler shifts," *IEEE Trans. Signal Process.*, vol. 47, no. 6, pp. 1559–1569, Jun. 1999.

[24] G. Tauböck, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 255–271, Apr. 2010.

[25] V. Raghavan, G. Hariharan, and A. M. Sayeed, "Capacity of sparse multipath channels in the ultra-wideband regime," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 3, pp. 357–371, Oct. 2007.

[26] A. N. Uwaechia and N. M. Mahyuddin, "Stage-determined matching pursuit for sparse channel estimation in OFDM systems," *IEEE Syst. J.*, to be published. doi: 10.1109/JSYST.2018.2837353.

[27] T. T. Cai, L. Wang, and G. Xu, "New bounds for restricted isometry constants," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4388–4394, Aug. 2010.

[28] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, Jul. 2011.

[29] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[30] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[31] E. J. Candés, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[32] R. Obermeier and J. A. Martinez-Lorenzo, "Sensing matrix design via mutual coherence minimization for electromagnetic compressive imaging applications," *IEEE Trans. Comput. Imaging*, vol. 3, no. 2, pp. 217–229, Jun. 2017.

[33] A. N. Uwaechia and N. M. Mahyuddin, "Joint pilot placement and symbol design scheme for sparse channel estimation in OFDM systems," *Phys. Commun.*, vol. 26, pp. 71–80, Feb. 2018.

[34] A. N. Uwaechia and N. M. Mahyuddin, "New pilot allocation design schemes for sparse channel estimation in OFDM system," in *Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA)*, Sep. 2017, pp. 66–71.

[35] J. Liu, K. Huang, and G. Zhang, "An efficient distributed compressed sensing algorithm for decentralized sensor network," *Sensors*, vol. 17, no. 4, p. 907, 2017.

[36] H. Palangi, R. Ward, and L. Deng, "Distributed compressive sensing: A deep learning approach," *IEEE Trans. Signal Process.*, vol. 64, no. 17, pp. 4504–4518, Sep. 2016.
[37] S. M. R. Islam and K. S. Kwak, "Two-stage channel estimation with estimated windowing for MB-OFDM UWB system," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 272–275, Feb. 2016.
[38] Y. Guan, T. Xu, R. van Leuken, and M. Qian, "Parallel channel estimator and equalizer for mobile OFDM systems," *Circuits, Syst., Signal Process.*, vol. 33, no. 3, pp. 839–861, 2014.
[39] G. Leus, Z. Tang, and P. Banelli, "Estimation of time-varying channels-a block approach," in *Wireless Communications Over Rapidly Time-Varying Channels*, G. Matz and F. Hlawatsch. Ed. New York, NY, USA: Academic, 2011, pp. 155–197.
[40] M. ApS. (2015). *The MOSEK Optimization Toolbox for MATLAB Manual*. Vers. 7.1 (rev. 28). [Online]. Available: http://docs.mosek.com/7.1/toolbox/index.html
[41] D. Astely, E. Dahlman, A. Furuskär, Y. Jading, M. Lindström, and S. Parkvall, "LTE: The evolution of mobile broadband," *IEEE Commun. Mag.*, vol. 47, no. 4, pp. 44–51, Apr. 2009.
[42] Y. C. Eldar and G. Kutyniok, Eds., *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

ANTHONY NGOZICHUKWUKA UWAECHIA received the B.Eng. and M.Sc. degrees in electrical and electronic engineering from Ahmadu Bello University, Zaria, Nigeria, in 2006 and 2013, respectively, and the Ph.D. degree in wireless and mobile systems from Universiti Sains Malaysia (USM), Malaysia, in 2018.

He has been a Postdoctoral Research Fellow with the School of Electrical and Electronic Engineering, USM, since 2019. His research interests include signal processing for communications and particularly focusing in the area of digital signal processing, sparse channel estimation, channel measurement, inverse problem, sparse representation, sparsity, and the OFDM systems.

Dr. Uwaechia was a recipient of the Graduate on Time Award for the completion of the Ph.D. degree in USM. He was also a recipient of the Institute of Postgraduate Studies, Graduate Assistant Scheme Award in USM, Penang, Malaysia, during the course of his Ph.D. studies. He was also a recipient of the Nigeria-Sao Tome and Principe Joint Development Postgraduate Scholarship Award during his M.Sc. degree programme.

NOR MUZLIFAH MAHYUDDIN received the B.Eng. degree from Universiti Teknologi Malaysia, Malaysia, in 2005, the M.Sc. degree from Universiti Sains Malaysia, Malaysia, in 2006, and the Ph.D. degree from Newcastle University, Newcastle upon Tyne, U.K., in 2011.

She is currently a Senior Lecturer and a Doctoral Supervisor with Universiti Sains Malaysia, Malaysia. Her research interests include RF and microwave engineering, reliability, signal integrity, sparse channel estimation, and the OFDM systems.

Dr. Mahyuddin is involved in the Communications Society, a member of IET and a Professional Member of the Association for Computing Machinery. She is also registered with the Board of Engineers Malaysia.