

Deterministic Pilot Design For Sparse Channel Estimation in MISO/Multi-user OFDM Systems

Roozbeh Mohammadian, Arash Amini, and Babak Hossein Khalaj

Abstract—We study the pilot design problem for sparse channel estimation in OFDM systems where multiple channels are estimated at a single antenna receiver. Such design is applicable to downlink of massive-MIMO systems and also to scenarios where multiple users transmit to a base station at the same carrier frequency. In our design, we deviate from the conventional orthogonal pilot arrangements by assigning the same pilot subcarriers to all transmitters. In the proposed setting, the achieved improvement in spectral efficiency (by reducing pilot overhead) may come at the expense of a more challenging channel estimation block at the receiver. To address this challenge and distinguish between different signals that are arriving at the receiver at the same subcarrier, we propose to select pilot subcarriers through minimizing the coherence of the associated Fourier submatrix, as well as properly assigning different pilot values (complex numbers) to each individual transmitter. We demonstrate that if the channels are sparse enough in time domain, there are simple sparse recovery techniques to simultaneously estimate all the channels, although all transmitters share the same pilot subcarriers. Simulation results demonstrate that the proposed design outperforms existing methods in terms of both mean-square channel estimation error and bit error rate.

Index Terms—Coherence, Compressive sensing, massive-MIMO, Multi-user, OFDM, Pilot design, Sparse channel estimation.

I. INTRODUCTION

THE increasing number of users in communications systems and the emergence of new technologies such as massive-MIMO have made the spectral efficiency of communications systems of high concern. Commonly, a part of the available spectrum is dedicated to transmitting pilot signals enabling the receivers to estimate the corresponding channels.

In applications where simultaneous estimation of a number of channels is required, orthogonal pilots that occupy nonoverlapping parts of the spectrum are traditionally used. Time-division-duplexing (TDD) massive-MIMO is a typical example where different transmitters employ orthogonal pilots to enable the base station (BTS) to estimate all the channels by assuming the reciprocity of the uplink and downlink channels. However, the number of orthogonal pilots is limited due to the confined coherence time and frequency bandwidth of the channels. Consequently, in multi-user cellular OFDM systems,

transmitters in different cells may be forced to reuse the same pilots. This leads to interfering pilots at the receivers (widely known as pilot contamination) and severely degrades the performance of the TDD communications systems [1]–[3].

In frequency-division-duplexing (FDD) massive-MIMO systems, downlink channels between different antennas and a receiver should be estimated. The orthogonal pilots may correspond to OFDM subcarriers that either represent different frequencies or correspond to different OFDM symbols. Naturally, in the latter case, the obtained channel estimates are valid only if the channels remain approximately unchanged over a number of adjacent OFDM symbols, restricting the applicability of such pilot arrangement [4].

Inspired by the fact that some wireless channels have sparse structures, many researchers have applied the theory of compressive sensing (CS) [5], [6] to sparse channel estimation in single-user and MIMO OFDM systems [7]–[9]. It has been shown that CS based methods can provide a better estimate of sparse channels than conventional approaches and even with less pilot overhead [7]–[9]. Therefore, when the channel is inherently sparse, CS based techniques are promising tools to reduce the number of pilots, which in turn results in mitigating the pilot contamination problem and reducing the pilot overhead in TDD and FDD massive-MIMO OFDM systems.

The performance of CS based channel estimation methods depends on both the pilot pattern (through the measurement matrix that will be defined in the sequel) and the recovery method. Specifically, in OFDM systems, the measurement matrix design involves selection of pilot tone locations and their corresponding complex values, which constitutes the main focus of this paper.

A. Related Work

In [7]–[10], under certain conditions, it is shown that uniformly at random pilot locations guarantee perfect channel reconstruction with high probability. Unfortunately, random pilots are not applicable in practical systems [11]. Therefore, deterministic pilot designs are proposed for SISO OFDM systems [11]–[14]. The design approach is often based on minimizing the coherence of the measurement matrix [11]–[14]. While it is common to restrict the design to selection of pilot locations, as shown in [13] and [14], joint optimization of pilot pattern and pilot power results in a measurement matrix with lower coherence value.

Deterministic pilot design for sparse channel estimation in MIMO OFDM systems is considered in [12], [15] and [16].

The authors are with the Department of Electrical Engineering and Advanced Communication Research Institute (ACRI), Sharif University of Technology, Tehran, Iran. B. H. Khalaj is also with the School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran 19538-33511, Iran (e-mails: rmohammadian@ee.sharif.ir, {aamini, khalaj}@sharif.ir).

The work is supported in part by the Institute for Research in Fundamental Sciences (IPM).

Manuscript received January 23, 2016; revised xxx xx, 201x.

Such approaches aim at minimizing the coherence of the measurement matrix for generating orthogonal pilot patterns for different transmit antennas.

Due to proximity of antenna elements in a MIMO platform, the effective channels between various antenna elements and a given transmitter/receiver, may have very similar path delays. Therefore, these channels exhibit a common sparsity pattern, which is referred to as spatial common sparsity. In addition, in slowly varying channels, the sparsity pattern remains almost unchanged in multiple adjacent OFDM symbols, which in turn is called temporal common sparsity. The common spatial and temporal sparsity patterns can be utilized to achieve better channel estimation performance and superior spectral efficiency.

In order to take advantage of the common spatial sparsity among different channels in a MIMO OFDM system, a deterministic pilot allocation scheme is proposed in [17] that generates orthogonal pilot sequences. For this purpose, the concept of coherence for the Fourier submatrices involved in the SISO OFDM channel estimation is generalized to accommodate for multiple Fourier submatrices associated with the MIMO channel. By formulating the pilot selection task as a minimization problem, the authors applied a genetic algorithm to find the minimizer.

The spatial common sparsity is further exploited in [4] and [18] in (massive) MIMO systems based on FDD and TDD, respectively. In both of such designs, instead of an orthogonal pilot arrangement, the pilot subcarriers are shared among all transmit antennas in order to reduce the pilot overhead. Nevertheless, distinct transmitters are distinguished by their corresponding pilot phase values set by i.i.d. realizations of a Bernoulli distribution over (± 1) . In [4], the pilots are distributed equidistantly in the spectrum while [18] uses random pilot locations.

B. Contributions

In this paper, we consider the pilot design problem for sparse channel estimation in MISO or multi-user OFDM systems. Similar to [4] and [18], our goal is to share pilot locations among all transmitters. However, we deviate by designing both the pilot subcarriers and pilot values (complex numbers) in a *deterministic* fashion. We also avoid employing the spatial or temporal common sparsity properties of the channel to make our design applicable even in cases where such assumptions do not hold. Specifically, we allocate pilots by minimizing the coherence of the associated Fourier submatrix. As this problem is non-convex, we propose two different relaxations to make the problem tractable. In fact, one of our key contributions is proposing a new scheme for selection of pilot values for various transmitters such that the coherence of the Fourier submatrix is further reduced. In addition, such approach enables the receiver to directly and simply estimate all required channels with no ambiguity. To the best of our knowledge, this is the first fully deterministic pilot design in a MISO/multi-user scenario with shared pilot subcarriers. Moreover, our approach is modular in the sense that our pilot design schemes can employ any coherence-based SISO pilot design algorithm.

C. Outline

The remainder of the paper is organized as follows: in section II, we describe the pilot design problem for a two-transmitter system as a generic form of a multi-transmitter setting. Next, we propose two different paradigms for designing the pilot pattern in section III. We first focus on a two-transmitter setting and then, demonstrate how the design could be scaled to accommodate more transmitters. We also evaluate the performance of the proposed methods in Section IV with MATLAB simulations. Finally, we conclude the paper in Section V.

D. Notations

We represent the matrices and vectors by boldface upper case and lower case letters, respectively; for instance \mathbf{v} represents a vector while $\mathbf{\Phi}$ refers to a matrix. In addition, $v(i)$, ϕ_j and $\Phi(i, j)$ stand for the i -th element in \mathbf{v} , j -th column in $\mathbf{\Phi}$ and (i, j) -th element of $\mathbf{\Phi}$, respectively. We denote the all-one vector by $\mathbf{1}$, where the size is understood from the context. Similarly, $\mathbf{0}$ denotes the all-zero vector. The notation $\|\mathbf{v}\|_p$ refers to the ℓ_p -norm of the vector \mathbf{v} defined by $(\sum_i |v(i)|^p)^{1/p}$. We reserve the notations $\mathbf{\Phi}^T$, $\mathbf{\Phi}^H$, and $\mathbf{\Phi}^\dagger$ for the transpose, conjugate transpose, and pseudo-inverse of the matrix $\mathbf{\Phi}$, respectively. For a vector \mathbf{v} , the diagonal matrix $\text{diag}\{\mathbf{v}\}$ is the one that contains the elements of \mathbf{v} on its main diagonal. The inner-product between \mathbf{v} and \mathbf{u} is denoted by $\langle \mathbf{v}, \mathbf{u} \rangle$ and equals $\mathbf{v}^H \mathbf{u}$, given that the two vectors have equal length. By $\mathbf{v} > 0$ we imply that all elements of \mathbf{v} are positive. Finally, we denote the empty set by \emptyset , and \setminus represents the set exclusion.

II. MISO OFDM CHANNEL ESTIMATION

For the sake of simplicity, we consider a scenario in which two transmitters are communicating with a receiver (which can also be located at BTS) using the OFDM modulation in the same frequency band. Let N denote the total number of subcarriers in each OFDM symbol, and let $1 \leq p_1 < p_2 < \dots < p_{N_P} \leq N$ be the index of pilot subcarriers (N_P pilots). In fact, we are assuming that both transmitters use the same subcarriers for pilots. However, for differentiation purposes, they are allowed to use different pilot values; hence, transmitter $t \in \{1, 2\}$ assigns the value $x_t(p_i)$ to its i -th pilot subcarrier, where $1 \leq i \leq N_P$. Note that the considered scenario can be applied to both TDD uplink and FDD downlink systems. We also assume that the length of the cyclic prefix is L ; *i.e.*, after taking IDFT of the sequence assigned to all subcarriers, the ending L time-domain samples are repeated in the beginning before transmitting the main block. In this paper, we assume that L is large enough to avoid Inter-Symbol-Interference (ISI).

A. Mathematical Formulation

If we define $\mathbf{X}_t = \text{diag}\{x_t(p_1), \dots, x_t(p_{N_P})\}$ with $t \in \{1, 2\}$, we can write

$$\mathbf{y} = \mathbf{X}_1 \cdot \mathbf{F} \cdot \mathbf{h}_1 + \mathbf{X}_2 \cdot \mathbf{F} \cdot \mathbf{h}_2 + \mathbf{n}, \quad (1)$$

where $\mathbf{y} = [y(p_1), y(p_2), \dots, y(p_{N_p})]^T$ represents the vector of received signal values at the pilot subcarriers, \mathbf{F} is an $N_p \times N$ matrix formed by selecting the p_1, \dots, p_{N_p} rows of the $N \times N$ DFT matrix, namely $\mathbf{F}(i, l) = e^{-j\frac{2\pi}{N}(p_i-1)(l-1)}$, and $\mathbf{n} = [n(p_1), n(p_2), \dots, n(p_{N_p})]^T$ stands for the observed additive Gaussian noise. Here, $\mathbf{h}_t = [h_t(1), \dots, h_t(N)]^T$ corresponds to the impulse response of the channel between the t -th transmitter and the receiver. By representing the channels using the vectors \mathbf{h}_1 and \mathbf{h}_2 , we are implicitly assuming that the effective taps of the channel coincide with the grid used for sampling the received continuous signal at the receiver. Although this assumption might not hold in general, for practical scenarios it is often observed that one can obtain a fair approximation of the original channel by restricting the taps to lie on a grid.

By defining

$$\Phi_1 = \mathbf{X}_1 \cdot \mathbf{F}, \quad \Phi_2 = \mathbf{X}_2 \cdot \mathbf{F}, \quad \Phi = [\Phi_1 \ \Phi_2], \quad (2)$$

we can rewrite (1) in the form

$$\mathbf{y} = \Phi \underbrace{\begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix}}_{\mathbf{h}} + \mathbf{n} = \Phi \cdot \mathbf{h} + \mathbf{n}. \quad (3)$$

Thus, to estimate the channels \mathbf{h}_1 and \mathbf{h}_2 we need to solve the linear inverse problem (3) for \mathbf{h} . Here, we have N_p observations and $2N$ unknowns, which characterizes a highly underdetermined scenario. We can reduce the number of unknowns by recalling the fact that the delay spread of the channel is no larger than L . Hence, the elements of \mathbf{h}_1 and \mathbf{h}_2 are zero except for possibly the first L elements. If $\tilde{\mathbf{h}}_1$ and $\tilde{\mathbf{h}}_2$ represent $L \times 1$ versions of \mathbf{h}_1 and \mathbf{h}_2 by removing the ending $N - L$ zeros, respectively, and $\tilde{\mathbf{F}}$ corresponds to the $N_p \times L$ DFT submatrix by discarding the last $N - L$ columns of \mathbf{F} , we arrive at

$$\mathbf{y} = \tilde{\Phi} \underbrace{\begin{bmatrix} \tilde{\mathbf{h}}_1 \\ \tilde{\mathbf{h}}_2 \end{bmatrix}}_{\tilde{\mathbf{h}}} + \mathbf{n} = \tilde{\Phi} \cdot \tilde{\mathbf{h}} + \mathbf{n}, \quad (4)$$

where

$$\tilde{\Phi}_1 = \mathbf{X}_1 \cdot \tilde{\mathbf{F}}, \quad \tilde{\Phi}_2 = \mathbf{X}_2 \cdot \tilde{\mathbf{F}}, \quad \tilde{\Phi} = [\tilde{\Phi}_1 \ \tilde{\Phi}_2]. \quad (5)$$

Consequently, the number of unknowns in (4) is reduced to $2L$ from $2N$ in (3). Nevertheless, we are still likely to face an underdetermined system of equations as N_p rarely exceeds $\frac{N}{10}$.

Motivated by possible sparse structure of wireless channels, we assume $\tilde{\mathbf{h}}_1$ and $\tilde{\mathbf{h}}_2$ are k -sparse vectors with $k \ll L$; *i.e.*, each of these length L vectors contains at most k nonzero elements. Therefore, $\tilde{\mathbf{h}}$ in (4) is also $2k$ -sparse, making its recovery possible through use of sparse reconstruction techniques proposed in the context of compressed sensing (CS) theory. It is important to note that the matrix $\tilde{\Phi}$ directly affects the possibility and quality of the reconstruction of $\tilde{\mathbf{h}}$ from \mathbf{y} in (4). To illustrate this point, let us examine the case where both transmitters use the same pilot values, or simply, $\mathbf{X}_1 = \mathbf{X}_2 = \mathbf{X}$. In this case, we will have:

$$\mathbf{y} = \mathbf{X}\tilde{\mathbf{F}}\tilde{\mathbf{h}}_1 + \mathbf{X}\tilde{\mathbf{F}}\tilde{\mathbf{h}}_2 + \mathbf{n} = \mathbf{X}\tilde{\mathbf{F}}(\tilde{\mathbf{h}}_1 + \tilde{\mathbf{h}}_2) + \mathbf{n}.$$

Consequently, only $\tilde{\mathbf{h}}_1 + \tilde{\mathbf{h}}_2$ and not individual components of $\tilde{\mathbf{h}}_1$ and $\tilde{\mathbf{h}}_2$ can be recovered.

B. Pilot Design Requirements

One of the well-known sufficient conditions on the measurement matrix $\tilde{\Phi}$ to guarantee stable recovery of sparse vectors is the so-called Restricted Isometry Property (RIP). Since it is very likely that a matrix formed by randomly selecting the rows of a unitary matrix (such as DFT) satisfies RIP, it is proposed in [7]–[10] to select the pilots uniformly at random among all available subcarriers in OFDM systems. Since it is computationally infeasible to check whether a random matrix satisfies the RIP condition [19], it is common to replace RIP with the coherence measure [11]–[17]. The coherence of a generic matrix $\mathbf{A}_{m \times n}$ denoted by $\mu_{\mathbf{A}}$ is defined as

$$\mu_{\mathbf{A}} = \max_{\substack{1 \leq i, l \leq n \\ i \neq l}} \frac{|\langle \mathbf{a}_i, \mathbf{a}_l \rangle|}{\|\mathbf{a}_i\|_2 \cdot \|\mathbf{a}_l\|_2}. \quad (6)$$

It is shown that if $\mu_{\mathbf{A}} < \frac{1}{2k-1}$, then, any k -sparse vector \mathbf{x} can be uniquely recovered from $\mathbf{A}\mathbf{x}$ using a wide range of reconstruction techniques including the simple orthogonal matching pursuit (OMP) method [20].

Let us examine the coherence value of the measurement matrix $\tilde{\Phi}$ (or Φ) in our channel estimation problem. Since $\tilde{\Phi}$ is obtained through concatenation of $\tilde{\Phi}_1$ and $\tilde{\Phi}_2$, its coherence can be written as

$$\mu_{\tilde{\Phi}} = \max \{ \mu_{\tilde{\Phi}_1}, \mu_{\tilde{\Phi}_2}, \mu_{\tilde{\Phi}_{1,2}} \}, \quad (7)$$

where

$$\begin{aligned} \mu_{\tilde{\Phi}_1} &= \frac{\max_{1 \leq r \leq L-1} \left| \sum_{i=1}^{N_p} |x_1(p_i)|^2 e^{-j\frac{2\pi}{N}(p_i-1)r} \right|}{\sum_{i=1}^{N_p} |x_1(p_i)|^2}, \\ \mu_{\tilde{\Phi}_2} &= \frac{\max_{1 \leq r \leq L-1} \left| \sum_{i=1}^{N_p} |x_2(p_i)|^2 e^{-j\frac{2\pi}{N}(p_i-1)r} \right|}{\sum_{i=1}^{N_p} |x_2(p_i)|^2}, \\ \mu_{\tilde{\Phi}_{1,2}} &= \frac{\max_{|r| \leq L-1} \left| \sum_{i=1}^{N_p} x_1(p_i) x_2^*(p_i) e^{-j\frac{2\pi}{N}(p_i-1)r} \right|}{\sqrt{\sum_{i=1}^{N_p} |x_1(p_i)|^2} \sqrt{\sum_{i=1}^{N_p} |x_2(p_i)|^2}}. \end{aligned} \quad (8)$$

In (8), we took advantage of the circular property of DFT submatrices, which implies that the inner-product between the i -th and $i+r$ -th columns has the same magnitude as the inner-product between the l -th and $l+r$ -th columns.

The deterministic pilot design challenge in a two-transmitter (MISO or multi-user) OFDM system can be summarized as minimizing the coherence of the matrix $\tilde{\Phi}$, or $\mu_{\tilde{\Phi}}$, by selecting proper pilot subcarriers and their corresponding values. Equivalently, we aim at solving the minimization problem

$$\Omega_{\text{opt}} = \arg \min_{\Omega} \mu_{\tilde{\Phi}}, \quad (9)$$

where the feasible set $\Omega = \{\mathcal{P}, \mathcal{X}_1, \mathcal{X}_2\}$ consists of

$$\begin{aligned} \mathcal{P} &= \{ \{p_1, \dots, p_{N_p}\} \subset \{1, \dots, N\} \}, \\ \mathcal{X}_{t=1,2} &= \{ \mathbf{x}_t = [x_t(p_i)]_i \in \mathbb{C}^{N_p}, \mathbf{x}_t \neq \mathbf{0} \}. \end{aligned} \quad (10)$$

III. MISO PILOT DESIGN

In this section, we propose a new strategy in designing the pilot patterns in a multi-transmitter setting. Following the scenario in the previous section, we first assume two transmitters transmitting at the same time. Our approach is to adopt a deterministic pilot pattern designed for a SISO system for the first transmitter and subsequently add the second transmitter by tuning its pilot values, using the same pilot locations. As will be shown in the sequel, such approach can be naturally extended to accommodate more transmitters.

A. Two-Transmitter Pilot Design

Recalling (3), we can express the received signal at i -th pilot subcarrier as

$$y(p_i) = x_1(p_i) \sum_{l=1}^N e^{-j\frac{2\pi}{N}(p_i-1)(l-1)} h_1(l) + x_2(p_i) \sum_{l=1}^N e^{-j\frac{2\pi}{N}(p_i-1)(l-1)} h_2(l) + n(p_i). \quad (11)$$

As demonstrated in Section II-A, identical pilot values (*i.e.*, $x_1(p_i) = x_2(p_i)$ for all i) results in an ill-posed system of linear equations. Thus, it is mandatory to set distinct values for the pilots of different transmitters. Here, we propose to use

$$x_2(p_i) = x_1(p_i) e^{-j\frac{2\pi}{N}(p_i-1)\alpha}, \quad i = 1, \dots, N_p, \quad (12)$$

where α is an integer to be later specified. By substituting (12) in (11) we obtain

$$y(p_i) - n(p_i) = x_1(p_i) \sum_{l=1}^N e^{-j\frac{2\pi}{N}(p_i-1)(l-1)} h_1(l) + x_1(p_i) \sum_{l=1}^N e^{-j\frac{2\pi}{N}(p_i-1)(l+\alpha-1)} h_2(l). \quad (13)$$

The contribution of h_2 in (13) can be reformulated as

$$\begin{aligned} \sum_{l=1}^N e^{-j\frac{2\pi}{N}(p_i-1)(l+\alpha-1)} h_2(l) &= \sum_{l=1+\alpha}^{N+\alpha} e^{-j\frac{2\pi}{N}(p_i-1)(l-1)} h_2(l-\alpha) \\ &= \sum_{l=1}^N e^{-j\frac{2\pi}{N}(p_i-1)(l-1)} \widehat{h}_2(l), \end{aligned} \quad (14)$$

where \widehat{h}_2 represents the circularly shifted version of h_2 by α units. Consequently, (13) can be written as

$$y(p_i) - n(p_i) = x_1(p_i) \sum_{l=1}^N e^{-j\frac{2\pi}{N}(p_i-1)(l-1)} (h_1(l) + \widehat{h}_2(l)). \quad (15)$$

To illustrate the latter result, we consider the channel impulse responses \mathbf{h}_1 and \mathbf{h}_2 as depicted in the top two plots of Figure 1. Here, we assume $N = 32$ and $L = 8$. The bottom plot in Figure 1 shows $\widehat{\mathbf{h}}_2$ where $\alpha = 12$. It is evident that \mathbf{h}_1 and $\widehat{\mathbf{h}}_2$ have non-overlapping supports (non-zero taps). In general, we are guaranteed that \mathbf{h}_1 and $\widehat{\mathbf{h}}_2$ have disjoint supports if $L \leq \alpha \leq N - L$. Naturally, the main result of such choice of

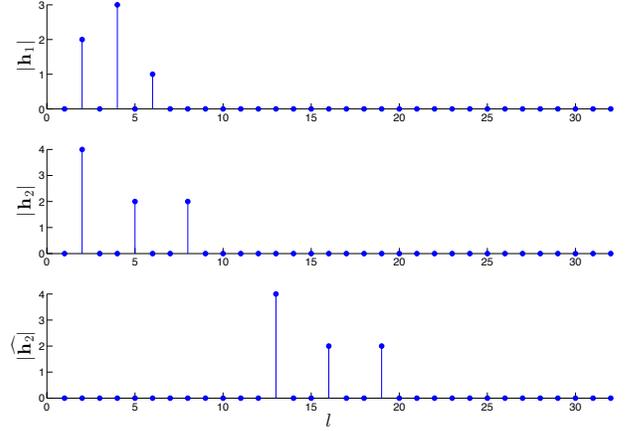


Fig. 1. Impulse responses of \mathbf{h}_1 , \mathbf{h}_2 and $\widehat{\mathbf{h}}_2$ after circularly shifting \mathbf{h}_2 by 12 units.

pilot phase values and the corresponding circular shift is that \mathbf{h}_1 and \mathbf{h}_2 can now be individually extracted from their sum.

Before further proceeding to computation of α , a number of advantages of the proposed design can be highlighted as follows:

- The complexity of pilot design based on (12) essentially consists of designing the pilots in a SISO setup, in addition to determining the single parameter α . Thus, the proposed method has almost negligible extra cost compared to ordinary SISO OFDM systems.
- Since $|x_1(p_i)|^2 = |x_2(p_i)|^2 = v(i)$ for all $1 \leq i \leq N_p$, we will have

$$\begin{aligned} \mu_{\widetilde{\Phi}_1} = \mu_{\widetilde{\Phi}_2} &= \frac{\max_{1 \leq r \leq L-1} \left| \sum_{i=1}^{N_p} v(i) e^{-j\frac{2\pi}{N}(p_i-1)r} \right|}{\sum_{i=1}^{N_p} v(i)}, \\ \mu_{\widetilde{\Phi}_{1,2}} &= \frac{\max_{|r-\alpha| \leq L-1} \left| \sum_{i=1}^{N_p} v(i) e^{-j\frac{2\pi}{N}(p_i-1)r} \right|}{\sum_{i=1}^{N_p} v(i)}. \end{aligned} \quad (16)$$

Therefore,

$$\mu_{\widetilde{\Phi}} = \max \{ \mu_{\widetilde{\Phi}_1}, \mu_{\widetilde{\Phi}_{1,2}} \}. \quad (17)$$

Consequently, we have further reduced the complexity of minimizing $\mu_{\widetilde{\Phi}}$.

As (17) suggests, the pilot design problem simplifies to minimizing $\mu_{\widetilde{\Phi}_1}$ and $\mu_{\widetilde{\Phi}_{1,2}}$. Note that minimizing $\mu_{\widetilde{\Phi}_1}$ is equivalent to designing a pilot pattern in a SISO setup. Furthermore, if the pilot pattern of the first transmitter is already known, we can minimize $\mu_{\widetilde{\Phi}_{1,2}}$ only by tuning α . In the sequel, we propose two techniques for minimizing $\mu_{\widetilde{\Phi}}$. The first and possibly the simpler approach is to apply a two-step minimization that first minimizes $\mu_{\widetilde{\Phi}_1}$ and then tunes α by keeping the pilot pattern of the first transmitter unchanged. The second approach considers simultaneous minimization of $\mu_{\widetilde{\Phi}_1}$ and $\mu_{\widetilde{\Phi}_{1,2}}$ where the pilot pattern of the first transmitter and the value of α are determined simultaneously.

B. Method I

As explained earlier, our first method extends a SISO pilot design to a MISO/multi-user design. Let the pilot locations p_1, \dots, p_{N_P} and pilot values $x(p_1), \dots, x(p_{N_P})$ be given by a SISO pilot design algorithm, such that the matrix

$$\Phi = \text{diag}\{x(p_1), \dots, x(p_{N_P})\} \mathbf{F}_{N_p \times N} \quad (18)$$

has a small coherence value. Note that $\mathbf{F}_{N_p \times N}$ represents a DFT submatrix with full column range (rather than being limited to the first L columns). There are various techniques for obtaining such a SISO pilot pattern including [11]–[14]; here, we assume availability of such a design without going into further details. In our MISO/multi-user scenario, we simply let the first transmitter proceed with the SISO settings. This implies transmitting the pilot value $x(p_i)$ at pilot subcarrier p_i . Hence, the effective $\tilde{\Phi}_1$ is equal to the first L columns of Φ . By using the same pilot locations for the second transmitter and setting the pilot values according to (12), $\tilde{\Phi}_2$ shall also be a submatrix of Φ consisting of L circularly consecutive columns starting with the $(\alpha + 1)$ -th column. Since the design of Φ is such that μ_Φ is small, the coherence of its submatrices $\tilde{\Phi}_1$ and $\tilde{\Phi}_2$ are also small (in fact $\mu_{\tilde{\Phi}_1} = \mu_{\tilde{\Phi}_2} \leq \mu_\Phi$). Our goal is to determine α such that $\mu_{\tilde{\Phi}_{1,2}}$ becomes as small as possible. To this end, we first define:

Definition 1. For an integer $0 \leq r \leq N - 1$, we define

$$C_V^P(r) = \frac{\left| \sum_{i=1}^{N_P} v(i) e^{-j \frac{2\pi}{N} (p_i - 1)r} \right|}{\sum_{i=1}^{N_P} v(i)}, \quad (19)$$

where $P = \{p_i\}_{i=1}^{N_P}$ and $V = \{v(i)\}_{i=1}^{N_P}$ represent the pilot locations and values, respectively.

It is straightforward to rewrite (16) in the form

$$\begin{aligned} \mu_{\tilde{\Phi}_1} = \mu_{\tilde{\Phi}_2} &= \max_{1 \leq r \leq L-1} C_V^P(r) \leq \max_{1 \leq r \leq N-1} C_V^P(r) = \mu_\Phi, \\ \mu_{\tilde{\Phi}_{1,2}} &= \max_{|\alpha| \leq L-1} C_V^P(r). \end{aligned} \quad (20)$$

Note that $C_V^P(0) = 1$. Therefore, to set α we should avoid $|\alpha| \leq L - 1$; otherwise, the columns of $\tilde{\Phi}_1$ and $\tilde{\Phi}_2$ shall have a nontrivial intersection which results in $\mu_{\tilde{\Phi}_{1,2}} = 1$. Consequently α should be set in the range $L, L+1, \dots, N-L$. It is worth mentioning that for any choice of α in this range, the resulting $\mu_{\tilde{\Phi}_{1,2}}$ is still upper bounded by μ_Φ , which is small by assumption. However, based on our simulation results, certain values of α may yield smaller coherence values. To find the optimum value of α , we propose the following procedure:

- **Step 1:** Evaluate $C_V^P(r)$ for all $r \in \{1, \dots, N - 1\}$.
- **Step 2:** For each choice of $\alpha \in \{L, L + 1, \dots, N - L\}$ find the maximum value of $C_V^P(r)$ for $\alpha - L < r < \alpha + L$:

$$M_\alpha = \max_{|\alpha| < L} C_V^P(r). \quad (21)$$

This is essentially implemented by sliding a window of length $2L - 1$ over the values of $C_V^P(r)$ and taking the maximum value.

- **Step 3:** Find the minimum of M_α in the considered range and declare the associated α as the optimum value. In case of a tie, any candidate can be used.

C. Method II

In the previous method, we extended an already existing SISO pilot pattern (pilot locations and values) to a two-user design by optimizing parameter α . Equivalently, we solved the problem

$$\alpha_{\text{opt}} = \arg \min_{\alpha} \max_{|r-\alpha| < L} C_V^P(r). \quad (22)$$

The selection of α affects $\mu_{\tilde{\Phi}}$ in (17) only through $\mu_{\tilde{\Phi}_{1,2}}$. When $\mu_{\tilde{\Phi}_{1,2}}$ exceeds $\mu_{\tilde{\Phi}_1}$ for all values of α , it is of interest to investigate whether another SISO pilot pattern (with possibly larger $\mu_{\tilde{\Phi}_1}$) can result in a smaller overall coherence measure $\mu_{\tilde{\Phi}}$. To this end, we generalize (22) as

$$\{P, V, \alpha\}_{\text{opt}} = \arg \min_{P, V, \alpha} \max_{r \in \mathcal{A}_\alpha} C_V^P(r), \quad (23)$$

where

$$\mathcal{A}_\alpha = \{\alpha + i\}_{i=-L+1}^{L-1} \cup \{1, 2, \dots, L-1\}. \quad (24)$$

Depending on the choice of α , the size of the set \mathcal{A}_α varies from $2L - 1$ to $3L - 2$. Although we have so far assumed α is an integer, it can be verified that (23) and (24) remain valid even for fractional values of α . However, with a fractional-valued α , the resulting $\tilde{\Phi}_2$ is no longer a DFT submatrix, while $\mu_{\tilde{\Phi}_1} = \mu_{\tilde{\Phi}_2}$ still holds.

Since it is difficult to directly address the minimization in (23), we divide it into two phases of setting α and determining P and V . Let $\Lambda_\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_M\}$ denote the finite set of possible choices for α . For each α_j , we try to find

$$\{P_{\alpha_j}, V_{\alpha_j}\} = \arg \min_{P, V} \max_{r \in \mathcal{A}_{\alpha_j}} C_V^P(r), \quad (25)$$

and set $\mu_{\tilde{\Phi}(\alpha_j)} = \max_{r \in \mathcal{A}_{\alpha_j}} C_V^{P_{\alpha_j}}(r)$. Finally, we choose j^* such that $\mu_{\tilde{\Phi}(\alpha_{j^*})}$ attains its minimum, and set $\alpha = \alpha_{j^*}$, $P = P_{\alpha_{j^*}}$, $V = V_{\alpha_{j^*}}$.

Finding the optimal $P_{\alpha_j}, V_{\alpha_j}$ in (25) is very similar to designing a pilot pattern in SISO OFDM systems. Hence, most of the available designs such as [11]–[14] could be potentially employed. In the following, we present an adaptation of Algorithm 1 proposed in [14].

The algorithm consists of outer and inner loops with T_{out} and T_{in} number of iterations, respectively. At each outer iteration, we initialize P with a random subset of the subcarriers (uniformly at random) and proceed with the inner iterations. Eventually, we shall pick the best pilot pattern among T_{out} possibilities. In fact, the greedy procedure to set P is generally sensitive to the initialization. Therefore, by considering T_{out} outer iterations, we increase the chance of having a suitable initialization point. The inner iterations are designed to sequentially modify the pilot locations P and the associated pilot powers $V_{\alpha_j, P}$.

Before describing the inner iterations, we should mention that the optimal pilot powers $V_{\alpha_j, P}$ when P and α_j are given, could be obtained via

$$\begin{aligned} V_{\alpha_j, P} &= \underset{V}{\text{argmin}} \max_{r \in \mathcal{A}_{\alpha_j}} C_V^P(r), \\ \text{s.t. } &\begin{cases} V_L \leq v(i) \leq V_H, & i = 1, \dots, N_T, \\ \sum_{i=1}^{N_P} v(i) \leq V_T. \end{cases} \end{aligned} \quad (26)$$

TABLE I
METHOD II: AN ADAPTATION FROM [14].

Input: $N, N_P, L, T_{\text{out}}, T_{\text{in}}, \Lambda_\alpha$.
Initialization: $\mathbf{m} \leftarrow \mathbf{0}_{M \times 1}, \mathbf{P} \leftarrow \mathbf{0}_{M \times N_P}$.
1: for $j = 1, \dots, M$
2: $\alpha \leftarrow \alpha_j$.
3: Set $\mathbf{D} \leftarrow \mathbf{0}_{T_{\text{out}} \times N_P}, \mathbf{c} \leftarrow \mathbf{0}_{T_{\text{out}} \times 1}$.
4: for $l = 1, \dots, T_{\text{out}}$
5: Randomly select $P \subset \{1, \dots, N\}, \bar{P} \leftarrow \emptyset$.
6: for $n = 1, \dots, T_{\text{in}}$
7: if $\bar{P} = P$
8: break.
9: end if
10: $\bar{P} \leftarrow P$.
11: for $i = 1, \dots, N_P$
12: Remove p_i from P to get \tilde{P} .
13: Obtain P_{new} according to (28), $P \leftarrow P_{\text{new}}$.
14: end for (i)
15: end for (n)
16: $\mathbf{D}[l] \leftarrow P, \mathbf{c}[l] \leftarrow \mu_{\tilde{\Phi}(\alpha_j, P, V_{\alpha_j, P})}$.
17: end for (l)
18: $s_1 = \arg \min_{l=1, \dots, T_{\text{out}}} \mathbf{c}[l]$.
19: $\mathbf{m}[j] \leftarrow \mathbf{c}(s_1), \mathbf{P}[j] \leftarrow \mathbf{D}[s_1]$.
20: end for (j)
21: $s_2 = \arg \min_{j=1, \dots, M} \mathbf{m}[j], \alpha \leftarrow \alpha_{s_2}$.
22: $P_\alpha \leftarrow \mathbf{P}[\alpha_{s_2}]$, and obtain V_{α, P_α} according to (26).

which yields

$$\mu_{\tilde{\Phi}(\alpha_j, P, V_{\alpha_j, P})} = \max_{r \in \mathcal{A}_{\alpha_j}} C_{V_{\alpha_j, P}}^P(r). \quad (27)$$

In (26), V_L, V_H and V_T represent constraints on minimum, maximum and total budget of pilot power, respectively. As shown in [14], (26) can be cast as a second-order cone programming, which could be solved using optimization packages such as MOSEK [21].

At each of the inner iterations, we first remove one of the elements from the current $P_{\text{cur}} = \{p_1, \dots, p_{N_P}\}$ to get \tilde{P} ; naturally, there are N_P ways to do this. One possible strategy is to remove an element which has not been removed in the past $N_P - 1$ inner iterations. Next, we consider all $N - N_P + 1$ possibilities of adding an element from $\mathcal{U}_{\tilde{P}} = \{1, \dots, N\} \setminus \tilde{P}$ to \tilde{P} . For each of these $N - N_P + 1$ cases, we form $\mu_{\tilde{\Phi}(\alpha_j, P, V_{\alpha_j, P})}$, and choose the one with the minimum coherence value:

$$P_{\text{new}} = \arg \min_{\substack{P = \tilde{P} \cup u \\ u \in \mathcal{U}_{\tilde{P}}}} \mu_{\tilde{\Phi}(\alpha_j, P, V_{\alpha_j, P})}. \quad (28)$$

In summary, we are improving P by replacing just one of its elements. The inner loop stops either after T_{in} iterations or when P no longer changes (optimality up to one element replacement). Table I summarizes the algorithm.

D. Multi-Transmitter Pilot Design

The design principles introduced in the two-transmitter setup of Section III-A can be extended to more than two transmitters. The key step is to generalize (12) and to simplify the overall coherence bound similar to (17).

To consider a generic MISO/Multi-user setup, let N_T represent the number of transmitters. We further assume that maximum delay spread of the channels between the transmitters

and the receiver denoted by L does not exceed $\frac{N}{N_T}$. This bound is due to the fact that our technique concatenates all the N_T channels into a single one that should be contained within the bounds of an OFDM symbol. Such constraint imposes an upperbound on the number of transmitters. We should mention that based on the received data at finite pilot locations it is not possible to uniquely estimate infinitely many channels. However, we do not claim optimality for the maximum number of allowed transmitters in our approach.

For the case of N_T transmitters, we generalize (12) as

$$x_t(p_i) = x_1(p_i) e^{-j \frac{2\pi}{N} (p_i - 1) \alpha_t}, \quad \begin{cases} i = 1, \dots, N_P \\ t = 1, \dots, N_T \end{cases}, \quad (29)$$

where $x_t(p_i)$ is the value of the i -th pilot of the t -th transmitter, and $\alpha_t s$ are integers that determine the circular shifts, with the convention that $\alpha_1 = 0$. With $\tilde{\Phi}_t = \mathbf{X}_t \tilde{\mathbf{F}}$ denoting the measurement matrix associated with the t -th transmitter, we can define the overall sensing matrix as

$$\tilde{\Phi} = [\tilde{\Phi}_1 \tilde{\Phi}_2 \dots \tilde{\Phi}_{N_T}]. \quad (30)$$

Consequently, the channel estimation task will be to recover the concatenated kN_T -sparse channel vector

$$\tilde{\mathbf{h}}_{(LN_T) \times 1} = [\tilde{\mathbf{h}}_1^T \tilde{\mathbf{h}}_2^T \dots \tilde{\mathbf{h}}_{N_T}^T]^T$$

from the equations $\mathbf{y} = \tilde{\Phi} \tilde{\mathbf{h}} + \mathbf{n}$. The coherence of $\tilde{\Phi}$ can now be written as

$$\mu_{\tilde{\Phi}} = \max_{r \in \mathcal{A}_\alpha} C_V^P(r), \quad (31)$$

where

$$\mathcal{A}_\alpha = \bigcup_{\substack{1 \leq m, n \leq N_T \\ m \neq n}} \{|\alpha_m - \alpha_n| + i\}_{i=-L+1}^{L-1} \cup \{i\}_{i=1}^{L-1}. \quad (32)$$

To avoid $0 \in \mathcal{A}_\alpha$, we need to set $\{\alpha_t\}_{t=1}^{N_T}$ such that $|\alpha_m - \alpha_n| \geq L$, for all $m \neq n$. One simple choice is

$$\alpha_t = (t - 1)L, \quad t = 1, \dots, N_T. \quad (33)$$

By assuming (33), we can see that $\mathcal{A}_\alpha \subset \{1, \dots, N - 1\}$; therefore, the coherence $\mu_{\tilde{\Phi}}$ is upper bounded by the coherence of the matrix Φ introduced in (18). In other words, we can start from a SISO design and extend it to a MISO/multi-user pattern by tuning $\{\alpha_t\}_{t=1}^{N_T}$. It should be mentioned that by increasing N_T , the optimization problem in both Method I and Method II becomes harder. For large values of N_T , we propose to directly adopt (33) instead of performing any optimization.

In summary, Method I could be extended to a multi-transmitter/multi-user setup by applying the following steps:

- **Step 1:** Adopt a SISO pilot design that minimizes the coherence of Φ (e.g., one of the designs in [11]–[14]) to obtain $P = \{p_1, \dots, p_{N_P}\}$ and $\{x(p_1), \dots, x(p_{N_P})\}$.
- **Step 2:** Assign $\{x(p_1), \dots, x(p_{N_P})\}$ to the first user/transmitter.
- **Step 3:** Set $\{\alpha_t\}_{t=2}^{N_T}$ according to Table II that generalizes the procedure in subsection III-B ($\alpha_1 = 0$ by default). Note that when $L \times N_T > \frac{N}{2}$, it is simpler to directly set $\{\alpha_t\}_{t=1}^{N_T}$ according to (33).

TABLE II
FINDING $\{\alpha_t\}_{t=2}^{N_T}$ FOR METHOD I.

Input: N, N_T, L, μ_{Φ} .
Initialization: $t \leftarrow 2, \mathcal{D}_t \leftarrow \{L, \dots, N-L\}$.
1: while $t \leq N_T$
2: For each $\alpha \in \mathcal{D}_t$ evaluate M_α defined in (21).
3: $\alpha_t \leftarrow \arg \min_{\alpha} M_\alpha$.
4: if ($M_{\alpha_t} = \mu_{\Phi}$ or $\mathcal{D}_t = \emptyset$).
5: Set $\{\alpha_t\}_{t=1}^{N_T}$ according to (33).
6: break.
7: end if
8: $\mathcal{D}_{t+1} \leftarrow \mathcal{D}_t \setminus \{\alpha_t + i\}_{i=-L+1}^{L-1}$.
9: $t \leftarrow t + 1$.
10: end while

- **Step 4:** Apply (29) on the obtained $\{\alpha_t\}_{t=1}^{N_T}$ and $\{x(p_i)\}_{i=1}^{N_P}$ to define the rest of the pilot values.

In a similar way, Method II can be generalized to a multi-transmitter/multi-user setup:

- **Step 1:** Form all N_T -tuples $\alpha \in \{0, \dots, N-1\}^{N_T}$ for which $0 \notin \mathcal{A}_\alpha$ (see (32) for the definition of \mathcal{A}_α).
- **Step 2:** For each feasible α minimize (25) to obtain P and V . This minimization is essentially similar to a SISO design and any of the techniques in [11]–[14] could be used. Hence, for each feasible α we obtain a coherence value $\mu_{\tilde{\Phi}}$ from (31).
- **Step 3:** Select one of the feasible N_T -tuples α that achieves the minimum coherence value $\mu_{\tilde{\Phi}}$. In turn, the choice of α determines P and V (**Step 2**).
- **Step 4:** Define the pilot pattern for the first transmitter/user by $\{x(p_1), \dots, x(p_{N_P})\}$. Design other pilot patterns based on (29).

As Methods I and II rest on a SISO pilot design, their actual computational complexities are directly related to the choice of the SISO design technique. However, Method II generally involves more computations in return for potentially lower coherence values (Method II never performs worse than Method I). In particular, the SISO design technique is called only once in Method I, while in Method II it is called for each feasible α (**Step 2**). For instance, with the system parameters ($N = 256, N_T = 2, L = 30$) and ($N = 256, N_T = 2, L = 60$) the SISO design technique is called 197 and 137 times in Method II, respectively. For a fixed N , larger values of $L \times N_T$ result in fewer feasible α s. Especially, the maximal case of $L \times N_T = N$ allows for only one feasible α given by (33). In this case, Method I and II are equivalent.

E. Discussion

The main ingredient of Methods I and II that simplifies the pilot design beyond the first transmitter/user, is the linear phase factor introduced in (12). We recall that our goal in this paper is to use the same pilot locations for all transmitters/users; however, we have the freedom to set the pilot values. Hence, the choice of (12) that differentiates transmitters/users only

TABLE III
COHERENCE MEASURES ACHIEVED BY DIFFERENT METHODS.

Method	Channel Length	$\mu_{\tilde{\Phi}}$
Method I	$L = 30$	0.3445
Method II	$L = 30$	0.2473
Random	$L = 30$	0.3635
Method I	$L = 60$	0.3473
Method II	$L = 60$	0.2966
Random	$L = 60$	0.3857

by a linear phase factor might seem too restrictive. Below, we discuss the benefits of this choice, as well as its suboptimality.

For an OFDM system with N_P pilots among N total subcarriers, let us assume that an optimal SISO pilot design results in coherence values μ_{Φ_L} and μ_{Φ} for channel lengths L and N (full-column DFT submatrix), respectively. Now let $\mu_{\tilde{\Phi}}$ be the achieved coherence of a given multi-transmitter/multi-user pilot design (not necessarily the proposed methods in this paper) for the same OFDM system considering channel length L . By ignoring the transmitters/users beyond the first, the multi-transmitter/multi-user design simplifies to a SISO design that cannot have a coherence value higher than $\mu_{\tilde{\Phi}}$. Because of our optimality assumption regarding the SISO design resulting in μ_{Φ_L} , we conclude that $\mu_{\Phi_L} \leq \mu_{\tilde{\Phi}}$. In other words, μ_{Φ_L} could be considered as a universal lower-bound for the coherence of all multi-transmitter/multi-user designs.

Interestingly, if the multi-transmitter/multi-user design is formed by applying (12) to the optimal SISO design, even without optimally tuning α_t s we shall have $\mu_{\tilde{\Phi}} \leq \mu_{\Phi}$. Hence, μ_{Φ} could be considered as an upper-bound or a worst-case performance of using (12). Moreover, in cases where μ_{Φ} is equal or slightly higher than μ_{Φ_L} (e.g., $L \geq \frac{1}{4}N$), the seemingly restricting choice of (12) is either optimal or suboptimal.

For further clarification, we consider the following experiment: for $N = 256, N_P = 16$ and $L = 30, 60$ we apply the SISO design of [14] to obtain the pilot pattern of the first transmitter/user; the achieved SISO coherence values are 0.2142 and 0.2525 for $L = 30$ and $L = 60$, respectively. Then, we extend the design to the second transmitter/user once by applying (12) in Methods I/II, and once by randomly setting the pilot values ($Unif(0, 1)$ for the magnitude and $Unif(0, 2\pi)$ for the phase). In Table III, we report the overall coherence values of Methods I and II, and the least achieved coherence values in the random setting among 10^6 realizations. The results in III confirm superiority of proposed methods compared to the best-case random design that avoids (12).

Here we list some of the advantages/properties of Methods I/II that use (12):

- 1) The proposed methods generate deterministic pilot patterns which is essential for practical applications.
- 2) The methods are based on minimizing the coherence measure. Unlike SPARK, NSP or RIP properties of a matrix which are NP-hard to evaluate, the coherence measure is easy to calculate ($\mathcal{O}(mn^2)$ for an $m \times n$ matrix). In addition, a matrix with coherence value μ is guaranteed to satisfy RIP of order $k < \mu^{-1} + 1$ with

TABLE IV
COHERENCE MEASURES ACHIEVED BY METHOD I.

Channel Length	μ_{Φ}	$\mu_{\tilde{\Phi}_1}$	$\mu_{\tilde{\Phi}_{1,2}}$	$\mu_{\tilde{\Phi}}$	α
$L = 30$	0.3473	0.3445	0.3405	0.3445	69
$L = 60$	0.3473	0.3473	0.3473	0.3473	60

constant $\delta_k = (k - 1)\mu$ [16]. It is also known that the performance of the greedy sparse recovery methods such as OMP is mainly influenced by the coherence measure of the sensing matrix [11].

- 3) As the proposed designs build upon a SISO pilot design and add transmitters/users, there is no need to modify the current transmitters/users in case of an exit or a new entry among the transmitters/users.
- 4) Method I is easily scalable as long as $N_T \times L \leq N$. This method simply adds a new feasible α to the existing set, which automatically defines a new pilot pattern
- 5) Independent of the number of transmitters/users, an upper-bound for the worst-case coherence measure of these methods (the coherence of the full-column DFT submatrix, μ_{Φ}) is available beforehand. This in turn defines an upper-bound on the MSE of channel estimation.
- 6) The proposed methods are transparent to the choice of the SISO pilot design, so long as it yields small coherence measures.

IV. SIMULATION RESULTS

In this section, we provide simulations to validate the performance of the proposed pilot design methods. We first study the performance of Method I and Method II, and then compare these two methods with the existing pilot designs of [4] and [18]. We should emphasize that the pilot design and channel estimation tasks in these two works are based on extra assumptions such as spatial-temporal common sparsity and block sparsity, which are missing in our setup.

As mentioned earlier, Method I and Method II essentially extend a SISO pilot design to a MISO/Multi-user pilot design. In our simulations, we adopted the SISO design of [14] (Algorithm 1). Such design minimizes the coherence measure by optimizing over both pilot locations and pilot values. We further compare their performance with the methods of [4] and [18] in Section IV-B. In order to come up with a fair comparison, we only optimize the coherence over the pilot locations and consider equal pilot powers as

$$v(i) = 1/N_P, \quad \text{for } i = 1, \dots, N_P. \quad (34)$$

The parameters involved in SISO pilot design are set similar to [14]. Specifically, the outer loop iterations T_{out} and the inner loop iterations T_{in} are set to 1000 and 15, respectively.

A. Performance Evaluation of Method I and Method II

In the first simulation, we investigate the performance of Method I and II for $N_T = 2$. We consider two cases with channel lengths $L = 30$ and $L = 60$. In either case, we assume $N = 256$ and $N_P = 16$.

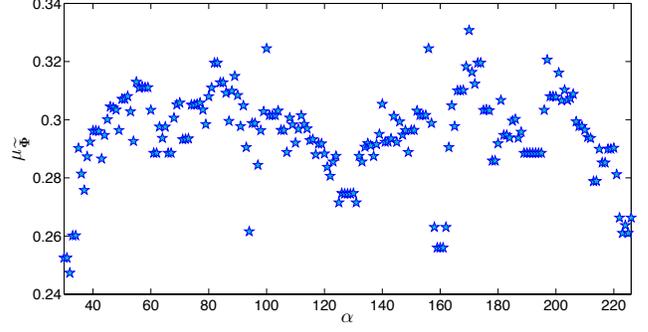


Fig. 2. The overall coherence measure for $L = 30$ at various integer values of α .

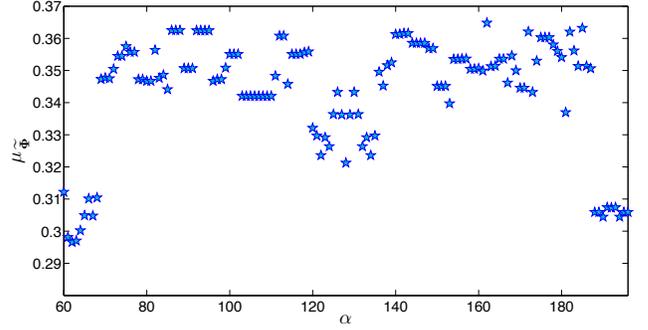


Fig. 3. The overall coherence measure for $L = 60$ at various integer values of α .

The coherence values achieved with Method I are summarized in Table IV. The value of μ_{Φ} in Table IV represents the coherence of the full-column DFT submatrix Φ with N_P rows; obviously, this number does not depend on L . The maximum cross correlation $\mu_{\Phi} = 0.3473$ among the columns of Φ corresponds to $r \in \{39, 121, 135, 217\}$ (Definition 1). By removing the $L + 1, \dots, N$ columns in Φ we obtain $\tilde{\Phi}_1$, which has a smaller coherence value $\mu_{\tilde{\Phi}_1}$. Next, we search for the optimum α that minimizes $\tilde{\Phi}_{1,2}$. For $L = 30$, we find $\alpha = 69$ that results in $\mu_{\tilde{\Phi}_{1,2}} = 0.3405 < \mu_{\tilde{\Phi}_1}$. Thus, the overall coherence measure $\mu_{\tilde{\Phi}}$ is determined by $\tilde{\Phi}_1$. In contrast, for $L = 60$ it turns out that for all values of α within the feasible range, we will have $\mu_{\tilde{\Phi}_{1,2}} = \mu_{\Phi}$. Hence, we choose $\alpha = 60$ (although other choices are equally good) and obtain $\mu_{\tilde{\Phi}} = \mu_{\Phi}$.

The coherence values achieved in Method II for different integer values of α are illustrated in Figure 2 for $L = 30$ and in Figure 3 for $L = 60$. We observe that the value of α considerably impacts the overall coherence values. This is more evident in Figure 2 for $L = 30$. Further, the minimum coherence values for $L = 30$ and $L = 60$ are 0.2473 and 0.2966 achieved at $\alpha = 32$ and $\alpha = 62$, respectively.

In principle, Method II, that is computationally more costly, is expected to emerge as the superior design as also confirmed by the results shown in Table IV and Figures 2, 3. However, the superiority of Method II is more highlighted in the case of $L = 30$. Hence, we conclude that Method II outperforms

TABLE V
THE PILOT PATTERN DESIGNED BY METHOD II FOR $L = 30$, $N_T = 2$,
 $N_P = 26$, AND $N = 256$.

$\mu_{\mathbf{p}}$	α	\mathcal{P}_{N_P}
0.2032	31	13, 25, 31, 34, 37, 49, 56, 59, 66, 70 79, 86, 94, 105, 143, 146, 157, 172 184, 187, 204, 221, 237, 240, 243, 247

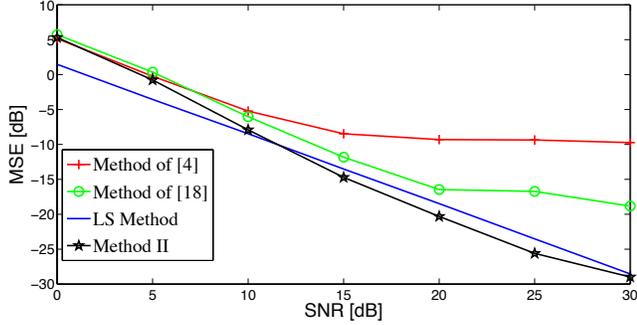


Fig. 4. The MSE of channel estimation under various pilot design schemes using OMP reconstruction technique (except for LS method). The delay spread of the channel is bounded by $L = 30$ and two transmitters are considered.

Method I in this setting (lower coherence measure).

B. Performance of Pilot Design Schemes for Multi-Transmitter Systems

In this part, we compare the performance of our proposed pilot design with that of [4] and [18] for various scenarios. It should be noted that the methods proposed in [4] and [18] exploit the common sparsity between different channels, where in our approach, such assumption is not made. We also include the least square (LS) channel estimation method as a reference in our comparisons. For applying the LS method, we consider a setting with orthogonal equidistant pilot patterns for all transmitters. Moreover, to achieve a fair comparison, we consider equal pilot powers as mentioned in (34) in all settings.

We compare the performance of various methods in terms of both the mean-square error (MSE) in estimating the sparse channel and the resulting bit error rate (BER). The former provides a measure to compare the efficiency of the pilot pattern itself, while the latter demonstrates the overall impact of the pilot pattern in the overall performance.

To simplify the communication setup, we use BPSK modulation at the transmitters and employ the maximum likelihood slicer (demodulator) at the receiver. As for the sparse channel, we set the non-zero taps uniformly at random within the available window of length L . The tap values are generated by realizations of i.i.d. complex Gaussian random variables with zero-mean and unit variance. The random channel is regenerated at every run of the simulation. Furthermore, to accommodate for the randomness involved in the designs of [4] and [18], in each run we randomly set the pilot values from $\{\pm \frac{1}{\sqrt{N_P}}\}$, while we use uniformly at random pilot locations for the method of [18]. For each of the curve points in

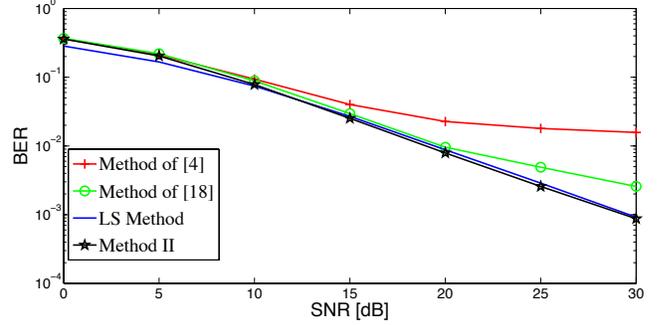


Fig. 5. Overall BER of the MISO OFDM system under various pilot design schemes. The delay spread of the channel is bounded by $L = 30$ and two transmitters are considered.

TABLE VI
OMP ALGORITHM [20].

Input: k .
Initialization: $\mathbf{z}^{(0)} \leftarrow \mathbf{0}_{k \times 1}$, $\mathbf{r}^{(0)} \leftarrow \mathbf{y}$, and $\Lambda^{(0)} \leftarrow \emptyset$.
1: for $i = 1, \dots, k$
2: $j^* \leftarrow \arg \max_j \frac{\langle \mathbf{a}_j, \mathbf{r}^{(i-1)} \rangle}{\ \mathbf{a}_j\ _2}$.
3: $\Lambda^{(i)} \leftarrow \Lambda^{(i-1)} \cup \{j^*\}$.
4: $\mathbf{z}^{(i)} \leftarrow (\mathbf{A}_{\Lambda^{(i)}})^\dagger \mathbf{y}$.
5: $\mathbf{r}^{(i)} = \mathbf{y} - \mathbf{A}\mathbf{z}^{(i)}$.
6: end for (i)

the corresponding figures we run the simulations 2000 times and report the average in the plots. To estimate the channel at the receiver, except for the LS method, we employ the orthogonal matching pursuit (OMP) method [20]. OMP is a greedy method which iteratively identifies the non-zero values of a sparse vector. Table VI summarizes how OMP recovers a sparse vector \mathbf{z} by having access to $\mathbf{y} = \mathbf{A}\mathbf{z}$ (\mathbf{A} is possibly a fat matrix). In Table VI, Λ is the set containing the indices of non-zero elements, $\mathbf{A}_{\Lambda^{(i)}}$ denotes the column submatrix of \mathbf{A} indexed by $\Lambda^{(i)}$, and \mathbf{r} is the residual vector.

In our first scenario, we consider a two-transmitter OFDM system ($N_T = 2$) with $N = 256$ subcarriers, $N_p = 26$ of which are used as pilots (except for the simulation of LS method). The channel has also $k = 3$ non-zero taps with the delay spread of $L = 30$. Since the LS method does not exploit the sparsity of the channel, we have increased the number of pilots to 64 for this method to obtain comparable results. By applying Method II to this setting, we have obtained the pilot locations shown in Table V. In all the simulation runs, we keep the same pilot pattern associated with Method II. Figures 4 and 5 present the MSE and BER performance of various pilot designs as a function of the signal to noise ratio (SNR), respectively. We observe that Method II outperforms other designs (except LS), specially at high SNR regimes. For instance, the MSE of estimating the channel by Method II is almost 10dB less than that of the method in [18] at $SNR = 30$ dB. It should be noted that the orthogonal arrangement of 64 pilots for the LS method in a two-transmitter setting occupies a total of 128 subcarriers, *i.e.*, 50% of the available spectrum. In contrast, the pilot overhead in Method II depends only on the

TABLE VII
THE PILOT PATTERN DESIGNED BY METHOD II FOR $L = 60$, $N_T = 2$,
 $N_P = 40$, AND $N = 256$.

$\mu_{\bar{\Phi}}$	α	\mathcal{P}_{N_P}
0.1889	123	2, 3, 6, 17, 38, 45, 59, 63, 80, 81, 88, 92, 93 97, 109, 125, 143, 147, 148, 152, 160, 161, 165 173, 177, 181, 185, 190, 202, 210, 211, 212, 218 219, 221, 236, 238, 243, 249, 252

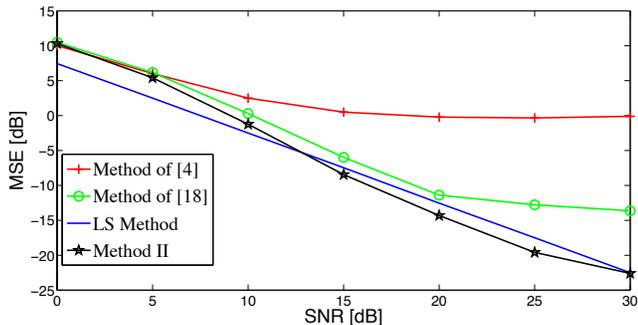


Fig. 6. The MSE of channel estimation under various pilot design schemes using OMP reconstruction technique (except for LS method). The delay spread of the channel is bounded by $L = 60$ and two transmitters are considered.

number of pilots (rather than the number of transmitters) and amounts to $26/256 \approx 10.2\%$. Therefore, a 39.8% gain in the spectral efficiency is obtained compared to conventional pilot arrangement schemes.

In our second scenario, we increase the delay spread of the channel to $L = 60$ and the number of non-zero channel taps to $k = 6$. In order to enable the receiver to estimate such channels, we increase the number of pilots to $N_P = 40$ for pilot design methods, and leave it unchanged at $N_P = 64$ for the LS method. The pilot pattern designed by Method II is provided in Table VII. Similarly, Figures 6 and 7 depict the MSE and BER performance of the considered pilot arrangements. These figures again confirm superiority of Method II compared to other approaches.

In our third scenario, we consider a MISO OFDM system with four transmitters ($N_T = 4$). We set $N = 256$, and include $N_P = 52$ pilot subcarriers for pilot design methods, while we use $N_P = 64$ for the LS method. In fact, because of the orthogonal pilot arrangement, $N_P = 64$ is the maximum number of pilots that can be assigned to each transmitter, and corresponds to a setting where all the spectrum is devoted to sending pilot subcarriers. We simulate channels with $k = 4$ non-zero taps that spread over a window of length $L = 60$. As explained earlier, when $L \times N_T$ is close to N we apply Method I and determine the values of $\{\alpha_t\}_{t=1}^4$ according to (33). The pilot locations found by this method are given in Table VIII. As in previous simulations, the pilot pattern associated with Method I does not change in different simulation runs. The MSE and BER performance curves of the considered methods are plotted in Figures 8 and 9, respectively. In line with previous results, we again observe that the proposed design outperforms the existing pilot arrangements. Furthermore, the pilot overhead of $52/256 \approx 20.3\%$ in Method I corresponds

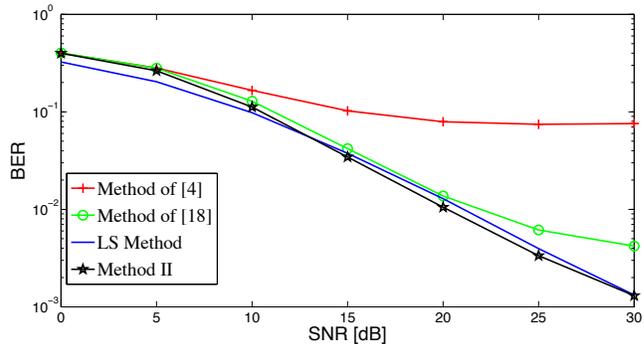


Fig. 7. Overall BER of the MISO OFDM system under various pilot design schemes. The delay spread of the channel is bounded by $L = 60$ and two transmitters are considered.

TABLE VIII
THE PILOT PATTERN DESIGNED BY METHOD II FOR $L = 60$, $N_T = 4$,
 $N_P = 52$, AND $N = 256$.

$\mu_{\bar{\Phi}}$	\mathcal{P}_{N_P}
0.1733	6, 8, 14, 15, 18, 29, 30, 33, 45, 49, 57, 61, 62, 67, 69, 71 77, 82, 84, 90, 99, 102, 103, 107, 116, 121, 123, 125, 132 135, 139, 142, 145, 146, 166, 168, 169, 171, 174, 175 180, 182, 183, 185, 191, 193, 202, 206, 211, 212, 236, 240

to a significant 79.7% spectral efficiency gain compared to the conventional orthogonal pilot allocation.

C. Performance of Pilot Design Schemes for Large Number of Transmitters

In this part, we compare the performance of our proposed pilot design with that of [4] and [18] when the number of transmitters/users grows large. For this purpose we consider four different cases as:

- **Case I:** $N = 1024$, $N_T = 8$, $L = 64$, and $N_P = 140$.
- **Case II:** $N = 1024$, $N_T = 16$, $L = 64$, and $N_P = 280$.
- **Case III:** $N = 4096$, $N_T = 32$, $L = 64$, and $N_P = 512$.
- **Case IV:** $N = 4096$, $N_T = 64$, $L = 60$, and $N_P = 1000$.

where in all cases each channel is considered to have 7 non-zero taps. As explained earlier, we consider equal pilot power distribution as in (34) to provide fairness in the comparisons. The performance of various methods is compared in terms of MSE in estimating the sparse channels. We also generate the sparse channels as described in subsection IV-B and recover them using the OMP method. Similarly, to accommodate the randomness involved in the designs of [4] and [18], in each run we randomly set the pilot values from $\{\pm \frac{1}{\sqrt{N_P}}\}$, while we use uniformly at random pilot locations for the method of [18]. For each of the curve points in the corresponding figures we average over 2000 runs. As in previous simulations, the pilot patterns associated with (proposed) deterministic methods remain unchanged in different simulation runs.

In **Case I** there are $\binom{519}{7} \approx 1.9 \times 10^{15}$ feasible 8-tuples for α ; therefore, we choose to set $\{\alpha_t\}_{t=1}^8$ of Method II according to (33) instead of inspecting the whole feasible set. The MSE performance curves of the considered methods are

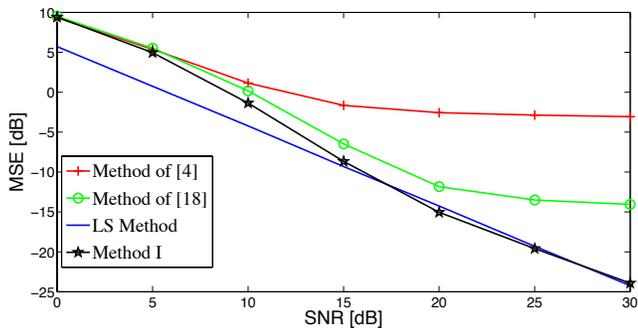


Fig. 8. The MSE of channel estimation under various pilot design schemes using OMP reconstruction technique (except for LS method). The delay spread of the channel is bounded by $L = 60$ and four transmitters are considered.

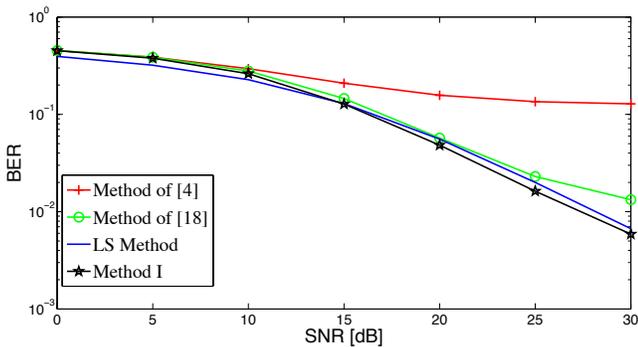


Fig. 9. Overall BER of the MISO OFDM system under various pilot design schemes. The delay spread of the channel is bounded by $L = 60$ and four transmitters are considered.

plotted in Figure 10. We observe that Method II outperforms other designs specially at high SNR regimes. For instance, the MSE of estimating the channels by Method II is almost 7dB less than that of the method of [4] at $SNR = 30$ dB.

Since $L \times N_T = N$ in **Case II**, we employ Method I. The MSE performance curves of the considered methods are plotted in Figure 11. This figure confirms superiority of Method I compared to other approaches.

In **Case III**, similar to **Case I** we apply Method II by setting $\{\alpha_t\}_{t=1}^{32}$ according to (33); again inspection of $\binom{2079}{31} \approx 6.9 \times 10^{68}$ feasible choices is computationally unaffordable. The MSE performance curves of the considered methods are plotted in Figure 12. In line with previous results, we observe that the proposed design outperforms the random pilot arrangements.

In **Case IV**, similar to **Case II**, we apply Method I. Figure 13 depicts the MSE performance of the considered pilot arrangements. Figure 13 demonstrates that the proposed method has better results than random pilots of [4], [18]. Particularly the MSE of Method I is almost 7.5dB less than the other two methods at $SNR = 30$ dB.

V. CONCLUSION

We considered the pilot design problem for sparse channel estimation in MISO or multi-transmitter OFDM systems. The goal is to design multiple deterministic and nonorthogonal

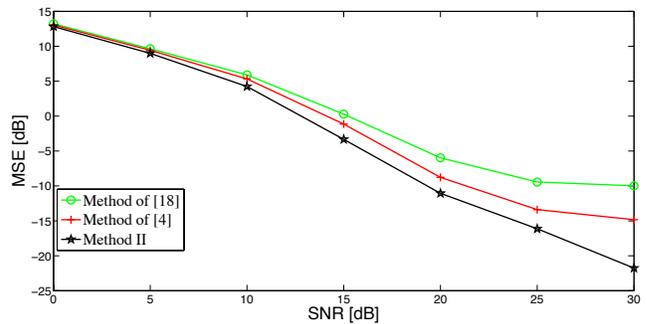


Fig. 10. The MSE of channel estimation under various pilot design schemes for **Case I**.

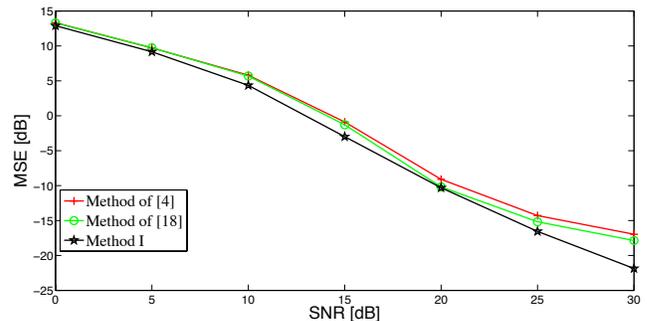


Fig. 11. The MSE of channel estimation under various pilot design schemes for **Case II**.

pilot sequences, which can improve the performance of sparse channel estimation methods. Based on minimizing the coherence of the measurement matrix, we identified for pilot patterns where all pilot sequences occupy exactly the same subcarriers. Specifically, we introduced two generic design techniques that extend arbitrary single user pilot designs to MISO settings. Simulation results confirm that the introduced methods, while imposing very low pilot overhead, outperform existing schemes in terms of MSE of the channel estimate and overall BER of the communications system.

REFERENCES

- [1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Comm.*, vol. 9, no. 11, pp. 3590–3600, November 2010.
- [2] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, January 2013.
- [3] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Comm.*, vol. 10, no. 8, pp. 2640–2651, August 2011.
- [4] Z. Gao, L. Dai, and Z. Wang, "Structured compressive sensing based superimposed pilot design in downlink large-scale MIMO systems," *Electronics Letters*, vol. 50, no. 12, pp. 896–898, June 2014.
- [5] E. J. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, August 2006.
- [6] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.
- [7] G. Taubock and F. Hlawatsch, "A compressed sensing technique for OFDM channel estimation in mobile environments: Exploiting channel sparsity for reducing pilots," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, March 2008, pp. 2885–2888.

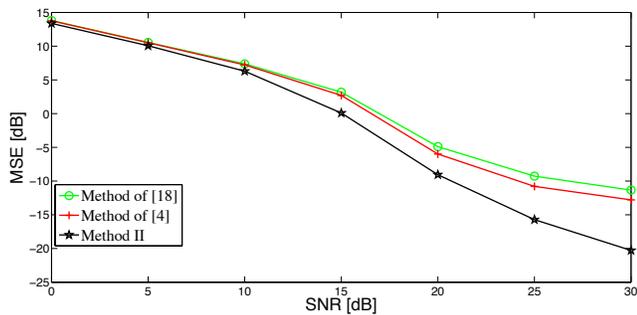


Fig. 12. The MSE of channel estimation under various pilot design schemes for **Case III**.

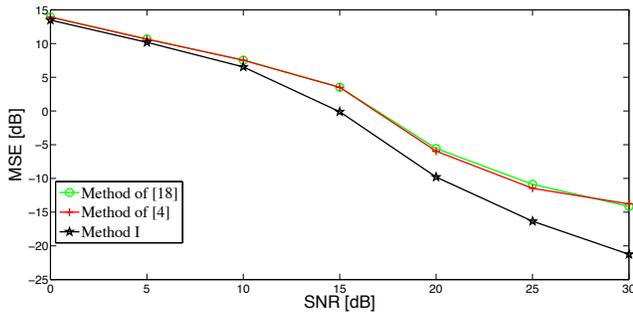


Fig. 13. The MSE of channel estimation under various pilot design schemes for **Case IV**.

the restricted isometry property is hard," *IEEE Trans. Inform. Theory*, vol. 59, no. 6, pp. 3448–3450, June 2013.

- [20] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2231–2242, October 2004.
- [21] MOSEK, *The MOSEK optimization toolbox for MATLAB manual Version 7.1 (Revision 28)*, 2015. [Online]. Available: <http://docs.mosek.com/7.1/toolbox/index.html>

- [8] G. Taubock, F. Hlawatsch, D. Eiwien, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE J. Sel. Top. Signal Process.*, vol. 4, no. 2, pp. 255–271, April 2010.
- [9] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.
- [10] C. Qi and L. Wu, "A hybrid compressed sensing algorithm for sparse channel estimation in MIMO OFDM systems," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, May 2011, pp. 3488–3491.
- [11] P. Pakrooh, A. Amini, and F. Marvasti, "OFDM pilot allocation for sparse channel estimation," *EURASIP J. Adv. Signal Process.*, vol. 59, 2012.
- [12] C. Qi, G. Yue, L. Wu, Y. Huang, and A. Nallanathan, "Pilot design schemes for sparse channel estimation in OFDM systems," *IEEE Trans. Veh. Tech.*, vol. 64, no. 4, pp. 1493–1505, April 2015.
- [13] M. Khosravi and S. Mashhadi, "Joint pilot power and pattern design for compressive OFDM channel estimation," *IEEE Comm. Lett.*, vol. 19, no. 1, pp. 50–53, January 2015.
- [14] C. Qi, L. Wu, Y. Huang, and A. Nallanathan, "Joint design of pilot power and pilot pattern for sparse cognitive radio systems," *IEEE Trans. Veh. Tech.*, vol. 64, no. 11, pp. 5384–5390, November 2015.
- [15] X. He, R. Song, and W. P. Zhu, "Pilot allocation for sparse channel estimation in MIMO-OFDM systems," *IEEE Trans. Circuits Syst. II Exp. Briefs*, vol. 60, no. 9, pp. 612–616, September 2013.
- [16] R. Mohammadian, A. Amini, B. H. Khalaj, and N. Omidvar, "MIMO-OFDM pilot symbol design for sparse channel estimation," in *Proc. Asian Pacific Signal and Inform. Process. Ass. Annual Summit and Conf.*, December 2015.
- [17] X. He, R. Song, and W. P. Zhu, "Pilot allocation for distributed compressed sensing based sparse channel estimation in MIMO-OFDM systems," to appear in *IEEE Trans. Veh. Tech.*, 2016.
- [18] Y. Nan, L. Zhang, and X. Sun, "Efficient downlink channel estimation scheme based on block-structured compressive sensing for TDD massive MU-MIMO systems," *IEEE Wireless Comm. Lett.*, vol. 4, no. 4, pp. 345–348, August 2015.
- [19] A. S. Bandeira, E. Dobriban, D. Mixon, and W. F. Sawin, "Certifying