Adaptive Pilot Allocation for Estimating Sparse Uplink MU-MIMO-OFDM Channels

Taoyong Li, Nele Noels, *Senior Member, IEEE*, Kamil Yavuz Kapusuz, *Member, IEEE*, Sam Lemey, *Member, IEEE*, Hendrik Rogier, *Senior Member, IEEE*, and Heidi Steendam, *Senior Member, IEEE*

Abstract—We consider uplink multiuser multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) communication. The transmit (Tx) side of the envisaged system consists of several single-antenna users or/and several multiple-antenna users. At the receive side, a multiple-antenna access point employs compressive sensing techniques to estimate the channel impulse response from the preamble portion of the observed packets. The traditional approach is that of orthogonal pilot allocation: during a short training period, each OFDM subcarrier is assigned exclusively to a single Tx antenna. In this case, the channel state information can conveniently be acquired on a per Tx antenna basis. To the best of our knowledge, all related research imposes that all Tx antennas are allocated the same amount of pilots (which must then be tailored for the most extreme channel conditions). However, in the considered system, Tx antennas may experience totally different channel conditions. Under these circumstances, the use of a fixed number of pilots per Tx antenna results in a lot of unnecessary overhead. To tackle this problem, our work addresses the design of efficient algorithms for adaptive orthogonal pilot allocation. The following design principles are applied: orthogonal pilot allocation, constant-modulus modulation, minimum measurement matrix mutual coherence optimization, and the condition that the number of pilot subcarriers allocated to each Tx antenna is adjusted to the channel conditions experienced by that Tx antenna. The paper tackles the problem of determining the optimal number of pilot subcarriers as well as the optimal positions of the pilots. To facilitate adaptive operation, we propose a reduced-complexity method to determine the optimal pilot positions. The performance of our algorithms is demonstrated by means of computer simulations, using both theoretical channel models and results from our own channel measurement campaign.

Index Terms—OFDM, MIMO, sparse channel estimation, compressive sensing, pilot allocation.

Taoyong Li is with the Information and Navigation School, Air Force Engineering University, Xi'an 710043, China, and also with the Collaborative Innovation Center of Information Sensing and Understanding, Xi'an 710071, China (e-mail: litaoyong0927@163.com).

Nele Noels and Heidi Steendam are with the Telecommunications and Information Processing Department, Ghent University/IMEC, 9000 Ghent, Belgium.

Kamil Yavuz Kapusuz, Sam Lemey, and Hendrik Rogier are with the IDLab-Electromagnetics Group, Department of Information Technology, Ghent University-imec, 9052 Ghent, Belgium.

algorithms is demonstrated by means of computer simulations, using both theoretical channel models and results from our own channel measurement campaign.

I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (ODFM) is ubiquitous in wireless communication networks, where it is often combined with other wireless technologies such as multiuser (MU) multiple-input multiple-output (MIMO). In an uplink MU-MIMO-OFDM system, multiple user nodes (UN) use the same physical resources to transmit their data simultaneously to a multiple-antenna access point (AP). For successful data reconstruction, it is imperative that the AP has accurate channel state information (CSI) [1]. Enabling the AP to recover the channel prior to data detection, it is customary for UNs to send training data in the form of a priori known pilots. For the sake of spectral efficiency, some MU-MIMO-OFDM systems opt for a scenario where all active UN antennas simultaneously send pilots over the same OFDM subcarriers [2], [3]. However, the more traditional approach is that of orthogonal pilot allocation: during a short training period, each OFDM subcarrier is assigned exclusively to a single UN antenna [4]–[6]. In this case, the CSI can conveniently be acquired on a per UN antenna basis.

Wireless communication channels often behave as linear filters with long impulse responses with most of the energy concentrated in a few short time intervals [7]. To limit the amount of pilots that are needed for the accurate estimate of such 'sparse' channels, compressive sensing (CS) based channel estimation can be used. The accuracy that can be achieved with such methods depends on (i) the channel conditions, (ii) the CS reconstruction method and (iii) the pilot allocation [8], [9]. In [10]–[12], several algorithms are proposed to effectively reconstruct a sparse signal. Any of these algorithms can be applied to estimate sparse channels. Further, some CS-based channel estimation methods have been specifically developed for the MIMO-OFDM system [13]–[15]. Accurate channel estimates can be obtained with these methods provided that a suitable pilot allocation is adopted. However, finding a good pilot allocation for estimating sparse OFDM channels is not trivial. In [4]–[6], [16]–[18], the issue was
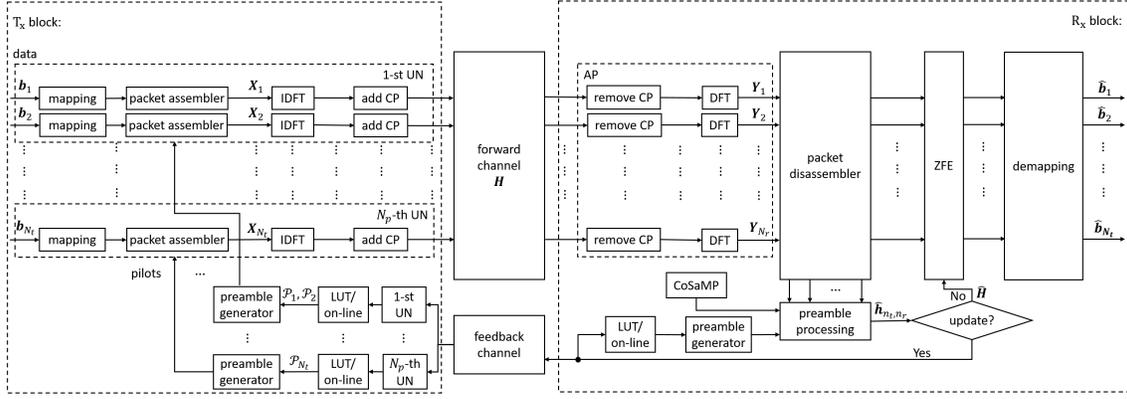
Fig. 1.   Block diagram of the considered MIMO-OFDM communication system.

studied for single-input single-output (SISO) OFDM systems. A wide range of optimization procedures is proposed; many of them are characterized by a high complexity and a slow convergence. Also, the number of data and pilot subcarriers per OFDM symbol is always fixed. The latter becomes problematic under dynamic channel conditions, where it is appropriate to instantaneously adjust the pilot allocation. A low-complexity procedure to estimate the instantaneous length and 'sparsity' of a SISO channel was proposed in [19], for OFDM symbols that are entirely allocated to pilots. So far, it has not been investigated whether the algorithm remains useful when only a portion of OFDM subcarriers is allocated to pilots. In addition, it has never been studied how the number of pilots per OFDM symbol can be adapted to match the effective channel length and sparsity. Orthogonal pilot allocation for MIMO-OFDM was considered in [4]–[6]. In all cases, all UN antennas in the system are assigned by default the same number of pilot subcarriers. In distributed systems with multiple UNs and a multi-antenna AP receiver, UNs can exhibit substantially different channel statistics. Moreover, if UNs can move independently, the overall coherence time reduces. In such cases, a fixed and unified worst-case scenario pilot allocation (always accommodating for the most extreme channel conditions) can cause a lot of unnecessary overhead. Adaptive, antenna-specific pilot allocation is then more appropriate [20].

This paper considers adaptive antenna-specific orthogonal pilot allocation for uplink MU-MIMO-OFDM. The main contributions of this paper are:

- A novel three-step procedure to continuously update the channel estimates and suitably adapt the pilot allocation in the preamble, prompting the system to accommodate the needs of mobile users.
- A novel efficient algorithm to *simultaneously* estimate the channel length and sparsity of all the channel impulse responses (CIRs). The algorithm accomodates for the fact that each antenna is allocated *only a portion of the preamble subcarriers*.
- A novel strategy to properly divide the total number of pilot subcarriers over the different UN antennas.
- An improved pilot allocation algorithm. Smart updating conditions are applied to accelerate the convergence.

Further, the concept of the 'shrinking potential' is employed, which is new in the context of orthogonal pilot allocation.

The paper is organized as follows. Section II introduces the system model and Section III reviews CS channel reconstruction. In Section IV-A, we first show how a receiver can retrieve and track the length and sparsity of all MIMO subchannels simultaneously, in parallel. Then, in Section IV-B, we propose a way to continuously adapt the number of pilot subcarriers per transmit antenna. Finally, in Section IV-C, we derive a novel orthogonal pilot allocation method. The method owes its efficiency to the introduction of a new metric, which we refer to as the 'measurement matrix mutual coherence shrinking potential'. Both theoretical channel models and results from our own channel measurement campaign are used to demonstrate the effectiveness of the algorithms. Section V describes the channel measurement set-up and Section VI discusses the numerical results. Conclusions are provided in Section VII.

Notation: Boldface italic letters denote vectors and boldface letters denote matrices. $\boldsymbol{h}_{n_t,n_r}$ and $\boldsymbol{\Phi}_{n_t,n_r}$ are the impulse response and the measurement matrix of the channel between the n_t-th Tx antenna and the n_r-th Rx antenna. The length and sparsity of $\boldsymbol{h}_{n_t,n_r}$ are denoted as L_{n_t,n_r} and K_{n_t,n_r}, respectively. There are N_t Tx antennas and N_r Rx antennas. N denotes the number of subcarriers. $\mathcal{P}_{n_t}$ denotes the set of pilot positions for the n_t-th Tx antenna, and $p_{n_t}(s)$ is the s-th element in $\mathcal{P}_{n_t}$. Further, N_{n_t} is the number of elements in $\mathcal{P}_{n_t}$. $\boldsymbol{Y}(l)$ denotes the observations on the l-th subcarrier at all Rx antennas, $\boldsymbol{Y}_{n_r}$ denotes the observations on all subcarriers at the n_r-th Rx antenna and $\boldsymbol{Y}_{n_t,n_r}$ denotes the observations on all subcarriers in $\mathcal{P}_{n_t}$ at the n_r-th Rx antenna. $\boldsymbol{\Phi}_{n_t}$ denotes the relevant measurement matrix for the n_t-th Tx antenna. $\mu\{\boldsymbol{\Phi}\}$ stands for mutual coherence of $\boldsymbol{\Phi}$. E_{po,n_t} is the 'shrinking potential' of $\mu\{\boldsymbol{\Phi}_{n_t}\}$. $\mu\{\boldsymbol{\Phi}_{n_t}\}_{\min}$ denotes Welch bound on $\mu\{\boldsymbol{\Phi}_{n_t}\}$. $\Pr[\mathcal{P}_{n_t}]$ is the probability of selecting $\mathcal{P}_{n_t}$ and δ_{n_t} is the scaling factor to adjust N_{n_t}.

II. GENERAL SYSTEM DESCRIPTION

An elementary block diagram of the considered system is shown in Fig. 1. N_p users with a total of $N_t \geq N_p$ transmit (Tx) antennas communicate with an AP that has N_r receive (Rx) antennas. The Tx antennas are not necessarily

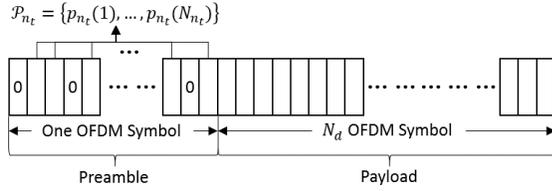$$\mathcal{P}_{n_t} = \{p_{n_t}(1), \ldots, p_{n_t}(N_{n_t})\}$$

Fig. 2. Packet structure of n_t-th Tx antenna.

physically co-located, and can therefore experience unequal channel statistics. At the Rx side a joint detection strategy is adopted. We further assume that all Tx and Rx antennas are perfectly synchronized in time and frequency.[1] After initial synchronization, the Tx antennas start sending data packets. As shown in Fig. 2, we assume that all Tx antennas employ the same packet structure: a payload of N_d data-modulated OFDM symbols is preceded by a short antenna-specific preamble with a mixture of pilot-modulated subcarriers and zero subcarriers. The pilots in the preamble are intended for channel estimation, which is the main subject of this paper. For clarity, we will limit our exposition to the case where the preamble consists of a single OFDM symbol and the CIR is estimated once per packet.[2] Each OFDM symbol consists of N subcarriers. The information bits modulating the payload subcarriers are mapped to a complex-valued constellation $\Omega = \{\omega_1, \omega_2, \ldots, \omega_{2^m}\}$ with $\frac{1}{2^m}\sum_{i=1}^{2^m}(\|\omega_i\|_2)^2 = 1$, where $\|\boldsymbol{a}\|_p$ denotes the p-norm of a vector $\boldsymbol{a}$. For simplicity, the pilots are assumed to be selected from a constant-modulus constellation $\Omega_p = \{\omega_{p,1}, \omega_{p,2}, \ldots, \omega_{p,2^{m_p}}\}$ with $\|\omega_{p,i}\|_2 = 1$. Each point in $\Omega\,(\Omega_p)$ corresponds to a unique $m(m_p)$-bit sequence. At the n_t-th Tx antenna, the l-th subcarrier is modulated by the complex value $X_{n_t}(l)$, and the corresponding bit sequence is $\boldsymbol{b}_{n_t}(l)$. The transmitted waveform is constructed by taking the inverse discrete Fourier transform (IDFT), followed by adding a cyclic prefix (CP). At the Rx side, the CP is removed and the resulting packet is sent to a discrete Fourier transform (DFT) unit for demodulation. The CP is assumed longer than the CIR of all MIMO subchannels, so that the l-th element of the DFT output at the n_r-th Rx antenna can be modeled as

$$Y_{n_r}(l) = \sum_{n_t=1}^{N_t} X_{n_t}(l)\, H_{n_t,n_r}(l) + W_{n_r}(l), \tag{1}$$

where $W_{n_r}(l)$, for $n_r = 1, 2, \ldots, N_r$ and $l = 1, 2, \ldots, N$, are independent zero-mean complex-valued circular-symmetric Gaussian noise variables with variance N_0, and

$$H_{n_t,n_r}(l) = \sum_{k=1}^{N} F(l,k)\, h_{n_t,n_r}(k), \tag{2}$$

with $F(l,k)$ the size-N DFT kernel

$$F(l,k) = e^{-j2\pi \frac{(k-1)(l-1)}{N}} \tag{3}$$

and $h_{n_t,n_r}(k)$, $k = 1, 2, \ldots$, the sample spaced CIR of the subchannel between the n_t-th Tx antenna and the n_r-th Rx antenna. The coherence time of the channel is assumed to be several times larger than the maximum duration of a packet, so that the channel can be modeled as constant over the duration of a packet transmission.[3] It is assumed that $h_{n_t,n_r}(k)$ is zero for all $k > L_{n_t,n_r}$ and the first L_{n_t,n_r} values of $h_{n_t,n_r}(k)$ are grouped in a vector $\boldsymbol{h}_{n_t,n_r} = [h_{n_t,n_r}(1), \ldots, h_{n_t,n_r}(L_{n_t,n_r})]^T$. The number of non-zero elements in $\boldsymbol{h}_{n_t,n_r}$ is denoted as K_{n_t,n_r}. The parameters L_{n_t,n_r} and K_{n_t,n_r} are termed channel length and channel sparsity, respectively. Collecting the observations on the l-th subcarrier at the various Rx antennas in a single vector, we have:

$$\boldsymbol{Y}(l) = \sqrt{E_s}\mathbf{H}^T(l)\,\boldsymbol{X}(l) + \boldsymbol{W}(l), \tag{4}$$

where E_s denotes the received symbol energy, $\boldsymbol{X}(l) = [X_1(l), X_2(l), \ldots, X_{N_t}(l)]^T$, $\boldsymbol{Y}(l) = [Y_1(l), Y_2(l), \ldots, Y_{N_r}(l)]^T$ and $\mathbf{H}(l)$ is the $N_t \times N_r$ matrix with elements $H_{n_t,n_r}(l)$ from (2). If $\mathbf{H}(l)$ is known, the receiver can estimate $\boldsymbol{X}(l)$ from $\boldsymbol{Y}(l)$, for $l = 1, 2, \ldots, N$. In this work, zero-forcing equalization (ZFE) will be employed for MIMO data detection. From the resulting symbol estimates $\hat{X}_{n_t}(l)$, estimates $\hat{\boldsymbol{b}}_{n_t}(l)$ of the corresponding m-bit sequences can be deduced. In practical systems, the quantities $\mathbf{H}(l)$ are not a priori known at the receiver and estimated values $\hat{\mathbf{H}}(l)$ are used instead. The accuracy of the estimates $\hat{\mathbf{H}}(l)$ determines the error performance of a given MIMO detector, at a given signal-to-noise ratio (SNR) E_s/N_0. The normalized mean squared error (NMSE) of the estimated CIRs $\hat{\boldsymbol{h}}_{n_t,n_r} = [\hat{h}_{n_t,n_r}(1), \hat{h}_{n_t,n_r}(2), \ldots, \hat{h}_{n_t,n_r}(L_{n_t,n_r})]^T$,

$$\hat{h}_{n_t,n_r}(k) = \frac{1}{N}\sum_{l=1}^{N} e^{j2\pi \frac{(l-1)(k-1)}{N}} \hat{H}_{n_t,n_r}(l), \tag{5}$$

serves as a first metric to evaluate the channel estimation performance. It is defined as

$$\text{NMSE} = \frac{1}{N_t N_r}\sum_{n_t=1}^{N_t}\sum_{n_r=1}^{N_r} \frac{\mathbb{E}\left[\left\|\left(\hat{\boldsymbol{h}}_{n_t,n_r} - \boldsymbol{h}_{n_t,n_r}\right)\right\|_2^2\right]}{\mathbb{E}\left[\|(\boldsymbol{h}_{n_t,n_r})\|_2^2\right]}, \tag{6}$$

where $\mathbb{E}[\cdot]$ is the statistical expectation with respect to the joint distribution of $(\mathbf{H}(l), \boldsymbol{W}(l))$ for all l. A good measure for the overall system performance is the bit error rate (BER),

$$\text{BER} = \mathbb{E}\left[\sum_{n_t=1}^{N_t}\sum_{l=1}^{N} \frac{1}{mN} d_H\left(\hat{\boldsymbol{b}}_{n_t}(l), \boldsymbol{b}_{n_t}(l)\right)\right], \tag{7}$$

with $d_H(\boldsymbol{b}, \dot{\boldsymbol{b}})$ the Hamming distance between $\boldsymbol{b}$ and $\dot{\boldsymbol{b}}$, and where $\mathbb{E}[\cdot]$ is the statistical expectation with respect to the joint distribution of $\mathbf{H}(l)$, $\boldsymbol{W}(l)$, $\boldsymbol{b}_{n_t}(l)$ for all l and all n_t.

III. CHANNEL ESTIMATION

In the remainder, it is assumed that if a Tx antenna modulates a preamble subcarrier with a pilot, then all other Tx antennas set that subcarrier to zero (orthogonal pilot allocation). For convenience, we specify the set of pilot-modulated preamble subcarriers at the n_t-th Tx antenna as $\mathcal{P}_{n_t} = \{p_{n_t}(1), p_{n_t}(2), \ldots, p_{n_t}(N_{n_t})\}$ (see Fig. 2). We have that $\cup_{j=1}^{N_t+1} \mathcal{P}_j = \mathcal{P} = \{1, 2, \ldots, N\}$, with $\mathcal{P}_i \cap \mathcal{P}_j = \varnothing$ for all $i \neq j$, and $\mathcal{P}_{N_t+1}$ the set that collects the subcarriers unused for channel estimation (possible empty). Following (1)-(3), the vector $\boldsymbol{Y}_{n_t,n_r} = [Y_{n_r}(p_{n_t}(1)), \ldots, Y_{n_r}(p_{n_t}(N_{n_t}))]^T$, collecting the DFT components with index l in $\mathcal{P}_{n_t}$ that are observed at the n_r-th Rx antenna can be expressed as

$$\boldsymbol{Y}_{n_t,n_r} = \sqrt{E_s}\boldsymbol{\Phi}_{n_t,n_r}\boldsymbol{h}_{n_t,n_r} + \boldsymbol{W}_{n_t,n_r}, \qquad (8)$$

where $\boldsymbol{W}_{n_t,n_r} = [W_{n_r}(p_{n_t}(1)), \ldots, W_{n_r}(p_{n_t}(N_{n_t}))]^T$, and $\boldsymbol{\Phi}_{n_t,n_r}$ is the $N_{n_t} \times L_{n_t,n_r}$ measurement matrix with elements

$$(\boldsymbol{\Phi}_{n_t,n_r})_{s,l} = X_{n_t}(p_{n_t}(s))\, e^{-j2\pi \frac{(p_{n_t}(s)-1)(l-1)}{N}}. \qquad (9)$$

As $\boldsymbol{\Phi}_{n_t,n_r}$ is known, the observation $\boldsymbol{Y}_{n_t,n_r}$ from (8) is suitable for estimating $\boldsymbol{h}_{n_t,n_r}$. We assume that a CS-based estimation method is employed. Several CS-based methods can be used to obtain an estimate $\hat{\boldsymbol{h}}_{n_t,n_r}$ of $\boldsymbol{h}_{n_t,n_r}$ from $\boldsymbol{Y}_{n_t,n_r}$ by solving some related problems. We can for example apply orthogonal matching pursuit (OMP) [10] to find the solution to the optimization problem

$$\min_{\boldsymbol{h}_{n_t,n_r}} \quad \|\boldsymbol{h}_{n_t,n_r}\|_0$$
$$\text{s.t. } \|\boldsymbol{Y}_{n_t,n_r} - \sqrt{E_s}\boldsymbol{\Phi}_{n_t,n_r}\boldsymbol{h}_{n_t,n_r}\|_2 \leq \epsilon, \qquad (10)$$

where ϵ is the error tolerance. Alternatively, we can use compressive sampling matching pursuit (CoSaMP) [11] to seek the estimate yielding minimum estimation error for a given (maximum) sparsity level $K_{n_t,n_r} \geq 1$

$$\min_{\boldsymbol{h}_{n_t,n_r}} \quad \|\boldsymbol{Y}_{n_t,n_r} - \sqrt{E_s}\,\boldsymbol{\Phi}_{n_t,n_r}\boldsymbol{h}_{n_t,n_r}\|_2$$
$$\text{s.t. } \|\boldsymbol{h}_{n_t,n_r}\|_0 \leq K_{n_t,n_r}. \qquad (11)$$

Moreover, expectation-maximization Gaussian-mixture approximate message passing (EM-GM-GAMP) [12] can be used to obtain an estimate $\hat{\boldsymbol{h}}_{n_t,n_r}$ by solving the following convex optimization problem

$$\hat{\boldsymbol{h}}_{n_t,n_r} = \arg \min_{\boldsymbol{h}_{n_t,n_r}} \{\|\boldsymbol{Y}_{n_t,n_r} - \sqrt{E_s}\boldsymbol{\Phi}_{n_t,n_r}\boldsymbol{h}_{n_t,n_r}\|_2^2$$
$$+ \lambda\|\boldsymbol{h}_{n_t,n_r}\|_1\}, \qquad (12)$$

where the scalar λ controls the relative importance applied to the Euclidian error and the sparseness term (the first and second expressions, respectively, inside the brackets in (12)).

The focus of this paper is, however, not on designing the CS-based channel recovery algorithm itself, but rather on designing the pilot allocation and therefore on designing the measurement matrix $\boldsymbol{\Phi}_{n_t,n_r}$ from (9). It is well-established [8], [9] that accurate CS-based reconstruction of $\boldsymbol{h}_{n_t,n_r}$ requires that:

1) The matrix $\boldsymbol{\Phi}_{n_t,n_r}$ can be constructed, which requires knowledge of $\mathcal{P}_{n_t}$, $X_{n_t}(p_{n_t}(s))$, $s = 1, 2, \ldots, N_{n_t}$ and the number L_{n_t,n_r} of columns in $\boldsymbol{\Phi}_{n_t,n_r}$.

2) The number N_{n_t} of rows in $\boldsymbol{\Phi}_{n_t,n_r}$, i.e., the number of pilot subcarriers for the n_t-th Tx antenna, is sufficiently large. Typically, it is required that [22]

$$N_{n_t} \geq \lceil K_{n_t,n_r} \log_2(L_{n_t,n_r}/K_{n_t,n_r}) \rceil. \qquad (13)$$

3) The matrix $\boldsymbol{\Phi}_{n_t,n_r}$ is an almost orthogonal matrix. A suitable measure for the orthogonality of $\boldsymbol{\Phi}_{n_t,n_r}$ is the measurement matrix mutual coherence (MMMC) $\mu\{\boldsymbol{\Phi}_{n_t,n_r}\}$ [9], with

$$\mu\{\boldsymbol{\Phi}\} = \max_{1 \leq u,v \leq L} |\langle \boldsymbol{\phi}_u, \boldsymbol{\phi}_v \rangle| / \|\boldsymbol{\phi}_u\|_2 \|\boldsymbol{\phi}_v\|_2, \quad (14)$$

where L is the number of columns in $\boldsymbol{\Phi}$, $\boldsymbol{\phi}_u$ is the u-th column of $\boldsymbol{\Phi}$, and $\langle \cdot \rangle$ is the scalar product operation.

It follows from (13) that, as opposed to conventional estimation methods, the number of required pilot subcarriers is (mainly) determined by K_{n_t,n_r} rather than L_{n_t,n_r}. Substituting (9) into (14), and taking into account that $\|X_{n_t}(p_{n_t}(s))\|_2 = 1$, yields

$$\mu\{\boldsymbol{\Phi}_{n_t,n_r}\} = \max_{1 \leq u,v \leq L_{n_t,n_r}} \frac{1}{N_{n_t}} \left| \sum_{s=1}^{N_{n_t}} e^{-j2\pi \frac{(p_{n_t}(s)-1)(u-v)}{N}} \right|,$$
$$(15)$$

which, for given N_{n_t}, only depends on $\mathcal{P}_{n_t}$. As far as CS-based estimation of $\boldsymbol{h}_{n_t,n_r}$ is concerned, minimum MMMC is a generally accepted criterion for designing the pilot allocation [4]–[6], [16]–[18], [20], [23]. The best $\mathcal{P}_{n_t}$, for given N_{n_t}, is the one that minimizes the MMMC $\mu\{\boldsymbol{\Phi}_{n_t,n_r}\}$ from (15). Due to the orthogonality principle, the positions of the pilots in the various per-Tx-antenna preambles cannot be chosen independently of each other. It follows that finding an appropriate partition $\{\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_{N_t+1}\}$ of $\mathcal{P}$ is a complex combinatorial optimization problem, which depends on the length and the sparsity of the individual CIRs.

In Section VI, it is established that with OMP, CoSaMP and EM-GM-GAMP based channel estimation virtually the same overall BER performance can be achieved. Because of space constraints we will focus on the CoSaMP approach to describe and evaluate the proposed adaptive orthogonal pilot allocation method. Since the proposed pilot allocation procedure is not particularly tailored to CoSaMP, similar performance advantages can be expected when used in combination with OMP or EM-GM-GAMP. As opposed to OMP and EM-GM-GAMP, CoSaMP requires that the value of L_{n_t,n_r} carefully matches the effective duration of $\boldsymbol{h}_{n_t,n_r}$ and that the effective sparsity K_{n_t,n_r} of $\boldsymbol{h}_{n_t,n_r}$ is also known. However, in the next section, we will show that accurate instantaneous information about L_{n_t,n_r} and K_{n_t,n_r} is indispensable anyway in practical systems with adaptive per-user pilot allocation.

IV. ADAPTIVE ORTHOGONAL PILOT ALLOCATION

We derive a practical procedure to adaptively allocate the pilot subcarriers. Using minimum NMSE as a criterion is intractable because the true channel is a priori unknown at

the Rx side. Using minimum BER as a criterion is also impractical because it requires a long acquisition period each time the pilot allocation needs updating. As an alternative, we propose a three-step procedure with a decoupled design of $\{(L_{n_t,n_r}, K_{n_t,n_r})\}$, $\{N_{n_t}\}$ and $\{\mathcal{P}_{n_t}\}$. The approach is ideally suited for adaptive pilot allocation. The first step serves to acquire and track the values of L_{n_t,n_r} and K_{n_t,n_r}, for all n_t and n_r. Acquisition is required only once, at the start of the transmission. To get the procedure started, the first packet of a transmission uses predefined initial per-Tx-antenna preambles. In the second step, the knowledge of $\{(L_{n_t,n_r}, K_{n_t,n_r})\}$ is used to determine N_{n_t}. The criterion used in this step is the condition imposed by (13). Step 1 and Step 2 are further detailed in Section IV-A and Section IV-B, respectively. Finally, in Step 3, $\{\mathcal{P}_{n_t}\}$ is selected based on the minimum MMMC criterion and subject to the condition that $\mathcal{P}_{n_t}$ should have cardinality N_{n_t}, for all n_t. The appropriate partition $\{\mathcal{P}_{n_t}\}$ for given values of N_{n_t} must be searched in large pre-computed look-up tables (LUT), or calculated on-line. A novel and more efficient subcarrier partitioning procedure (with a better accuracy/complexity trade-off than the state-of-the-art in [5]) is proposed in Section IV-C. If Step 1 does not result in an update of $\{(L_{n_t,n_r}, K_{n_t,n_r})\}$, the receiver passes the packet to the data detector for further processing; otherwise, the receiver locally updates the pilot allocation and informs the UNs at the TX side to do the same using feedback channel signaling (see Fig. 1). Upon reception of this feedback, the per-antenna-preambles are updated and the payload is resend with the novel preambles. We note that the 'LUT/on-line processing' and 'preamble generator' blocks at the Tx side need to be present in and run by every UN. An alternative could be that the updated pilots are only computed at the receiver and simply fed back to the transmit side, but this would require significantly more feedback signaling overhead. The 'preamble processing' block serves to estimate L_{n_t,n_r} and K_{n_t,n_r} of all channels (step 1) and to determine the appropriate number of pilots per Tx antenna N_{n_t} (step 2).

A. Length and Sparsity of All Subchannels

Upon reception of a novel preamble (with some given $\{\mathcal{P}_{n_t}\}$), we estimate the instantaneous values of $\{(L_{n_t,n_r}, K_{n_t,n_r})\}$. One option is to employ a channel reconstruction algorithm that does not require prior knowledge of L_{n_t,n_r} and K_{n_t,n_r}; estimates of L_{n_t,n_r} and K_{n_t,n_r} can then be derived from the estimated CIRs. However, in this work, we opt for CoSaMP channel reconstruction, in which case we compute CIR estimates for different trial values of L_{n_t,n_r} and K_{n_t,n_r}, and select the best option using minimum BER (in the preamble) as a criterion. This approach was originally proposed in [19] for a significantly different scenario, namely with $L_{n_t,n_r} > N = N_{n_t} > K_{n_t,n_r}, n_t = N_t = 1, n_r = N_r = 1$. In the following, it is shown that the algorithm from [19] can not be applied directly to the situation at hand. To ensure proper operation, important adjustments are necessary.

For given $\mathcal{P}_{n_t}$, let $\#\mathrm{err}_{n_t,n_r}(\tilde{K}, \tilde{L})$ denote the number of bit errors that are detected in $\{X_{n_t}(p_{n_t}(s)); p_{n_t}(s) \in \mathcal{P}_{n_t}\}$ after ZFE and demapping of $\{Y_{n_r}(p_{n_t}(s)); p_{n_t}(s) \in \mathcal{P}_{n_t}\}$,

when the CoSaMP CIR reconstruction unit assumes that $\tilde{K}$ and $\tilde{L}$ are the true sparsity and length of $\boldsymbol{h}_{n_t,n_r}$. For the scenario considered in [19], it was observed that (at high SNR) for any value of $\tilde{K}$ larger than the true sparsity K_{n_t,n_r} the average of $\#\mathrm{err}_{n_t,n_r}(\tilde{K}, \tilde{L})$ achieves its minimum in $\tilde{L} = L_{n_t,n_r}$. Therefore, the following two-stage optimization procedure was proposed. First, L_{n_t,n_r} was estimated to minimize $\#\mathrm{err}_{n_t,n_r}(N_{n_t}, \tilde{L})$ (taking into account that N_{n_t} is an upper bound on K_{n_t,n_r} according to (13)). Then, K_{n_t,n_r} is estimated by keeping in $\#\mathrm{err}_{n_t,n_r}(\tilde{K}, \tilde{L})$, the value of $\tilde{L}$ fixed and equal to $\hat{L}_{n_t,n_r}$. Unfortunately, straightforwardly applying this procedure to the scenario at hand (where $N > L_{n_t,n_r} > N_{n_t} > K_{n_t,n_r}$) does not yield satisfactory results. This is a consequence of the fact that the objective function $\#\mathrm{err}_{n_t,n_r}(\tilde{K}, \tilde{L})$ behaves differently if $[N > L_{n_t,n_r} > N_{n_t} > K_{n_t,n_r}]$ than if $[L_{n_t,n_r} > N = N_{n_t} > K_{n_t,n_r}]$. Simulations show that, in the former case, a reversed two-stage optimization procedure, where K_{n_t,n_r} is estimated prior to L_{n_t,n_r}, works better. The full discussion can be found in Section VI-A. Based on these observations we propose to estimate $(K_{n_t,n_r}, L_{n_t,n_r})$ from the observed preamble as

$$\hat{K}_{n_t,n_r} = \arg\min_{\tilde{K}} \#\mathrm{err}_{n_t,n_r}(\tilde{K}, L_{n_t,n_r}^{(\max)}), \quad (16)$$

$$\hat{L}_{n_t,n_r} = \arg\min_{\tilde{L}} \#\mathrm{err}_{n_t,n_r}(\hat{K}_{n_t,n_r}, \tilde{L}), \quad (17)$$

with $L_{n_t,n_r}^{(\max)}$ an upper bound on the channel length (e.g., $L_{n_t,n_r}^{(\max)} = N$). The minima in (16)-(17) are found efficiently using the dynamic window search (also used in [19]) with initial search interval $[1, N_{n_t}]$ for (16), and $[1, L_{n_t,n_r}^{(\max)}]$ for (17). Moreover, (16)-(17) can be performed simultaneously, in parallel, for all $N_t \times N_r$ subchannels. Estimating K_{n_t,n_r} (16) prior to L_{n_t,n_r} (17) significantly outperforms the original (reverse) approach from [19] in terms of estimation accuracy.

B. Number of Pilot Subcarriers Per Antenna

In distributed MIMO systems, K_{n_t,n_r} and L_{n_t,n_r} can vary significantly with (n_t, n_r). This implies that the required number of pilot subcarriers to estimate the subchannel $\boldsymbol{h}_{n_t,n_r}$ also varies significantly with (n_t, n_r). For each $n_t \in \{1, 2, \ldots, N_t\}$, the condition (13) imposes N_r different constraints on N_{n_t,n_r} (one per Rx antenna). However, subchannels $\boldsymbol{h}_{n_t,n_r}, n_r = 1, 2, \ldots, N_r$ share the same $\mathcal{P}_{n_t}$. To ensure accurate estimation of $h_{n_t,n_r}, n_r = 1, 2, \ldots, N_r$, the number N_{n_t} of elements in $\mathcal{P}_{n_t}$ must comply with (13) for that particular value of n_t and every possible value of n_r. We therefore propose to set N_{n_t} as

$$N_{n_t} = \max_{n_r} \left\lceil \delta_{n_t} \cdot \hat{K}_{n_t,n_r} \log_2 \left(\hat{L}_{n_t,n_r} / \hat{K}_{n_t,n_r} \right) \right\rceil, \quad (18)$$

with $(\hat{K}_{n_t,n_r}, \hat{L}_{n_t,n_r})$ an estimate of $(K_{n_t,n_r}, L_{n_t,n_r})$ and $\delta_{n_t} \geq 1$ a scaling factor that can be adapted to meet the system requirements (total number of pilots, desired accuracy for certain channels, etc.). Starting from an initial pilot allocation, estimates $(\hat{K}_{n_t,n_r}, \hat{L}_{n_t,n_r})$ can be obtained by applying the approach outlined in Section IV-A. These estimates can then be used to compute the appropriate number of pilot

subcarriers per Tx antenna in (18). Subsequently, a new pilot allocation with the appropriate subset sizes can be selected. In its turn, this new pilot allocation can be used to re-estimate $(K_{n_t,n_r}, L_{n_t,n_r})$. If the novel $(\hat{K}_{n_t,n_r}, \hat{L}_{n_t,n_r})$ do not (significantly) differ from the previous values, the system can continue to use the current pilot allocation. Otherwise, the procedure of updating $\{N_{n_t}\}$ and selecting a novel pilot allocation needs to be run again. A numerical example of this iterative updating procedure is provided in Section VI.

C. Orthogonal Pilot Allocation

We describe an efficient orthogonal pilot allocation procedure, termed 'extended simulated annealing' or ESA. As a smaller MMMC results in a more accurate estimate of the channel [4]–[6], [16]–[18], the MMMC will be adopted as the cost function. The ultimate goal is to find a good approximate solution to the following complex multi-objective optimization problem. Find a partition of $\mathcal{P}$ into subsets $\mathcal{P}_1,\ldots,\mathcal{P}_{N_t+1}$ with respective cardinality $N_1,\ldots,N_{N_t+1}$, so that

$$\mathcal{P}_{n_t} = \arg\min \mu\{\boldsymbol{\Phi}_{n_t,n_r}\}, \quad \begin{array}{l} n_t = 1,\ldots,N_t \\ n_r = 1,\ldots,N_r. \end{array} \quad (19)$$

From (15), it follows that $\mu\{\boldsymbol{\Phi}_{n_t,n_r}\}$ depends on $\mathcal{P}_{n_t}$ and L_{n_t,n_r}. For given $\mathcal{P}_{n_t}$, it is easily derived that $\mu\{\boldsymbol{\Phi}_{n_t,n_r}\} \leq \mu\{\boldsymbol{\Phi}_{n_t,n_r'}\}$ if $L_{n_t,n_r} \leq L_{n_t,n_r'}$. Hence, (19) simplifies as follows:

$$\mathcal{P}_{n_t} = \arg\min \mu\{\boldsymbol{\Phi}_{n_t}\}, \quad n_t = 1,\ldots,N_t, \quad (20)$$

where $\boldsymbol{\Phi}_{n_t} = \boldsymbol{\Phi}_{n_t,f(n_t)}$, with

$$f(n_t) = \arg\max_{n_r=1,2,\ldots,N_r} L_{n_t,n_r}. \quad (21)$$

In the literature, various researchers proposed ways to further simplify (20) by transforming it into a single-objective optimization problem. In [2], [6], for example, the per-Tx-antenna pilot allocations are assumed to be shifted versions of each other. In that case, solving (20) for all n_t boils down to solving (20) for $n_t = 1$ only, which significantly reduces the complexity of the problem. However, the restriction that the per-Tx-antenna pilot allocations need to be shifted versions of each other limits the performance that can be achieved. Moreover, this approach does not allow to allocate a different number of subcarriers to different Tx antennas, which may be suboptimal in distributed MIMO scenarios. In this work, we will rather follow the approach taken in [5] and use the minimum sum MMMC criterion, i.e., find a partition of the set of subcarriers $\mathcal{P}$ into subsets $\mathcal{P}_1,\ldots,\mathcal{P}_{N_t+1}$, with respective cardinality $N_1,\ldots,N_{N_t+1}$, so that

$$\{\mathcal{P}_1,\mathcal{P}_2,\ldots,\mathcal{P}_{N_t}\} = \arg\min \mu\{\boldsymbol{\Phi}\}, \quad (22)$$

with $\mu\{\boldsymbol{\Phi}\}$ the sum MMMC, given by

$$\mu\{\boldsymbol{\Phi}\} = \sum_{n_t=1}^{N_t} \mu\{\boldsymbol{\Phi}_{n_t}\}. \quad (23)$$

Similar to in [18] for SISO systems, simulated annealing (SA) can be used to efficiently find an approximate solution of (22).

Algorithm 1 ESA

1: Input: N and $\{N_{n_t}, \mu\{\boldsymbol{\Phi}_{n_t}\}_{min}; n_t = 1,2,\ldots,N_t\}$
2: Set $T_{init}, T_{rate}, T_{stop}, T_{iter}$. Set $T = T_{init}$.
3: Randomly choose a pilot allocation and calculate $\mu\{\boldsymbol{\Phi}_{n_t}\}$, $n_t = 1,\ldots,N_t$.
4: **while** $T > T_{stop}$ **do**
5:     **for** $l = 1 : T_{iter}$ **do**
6:         Using (24) and (26)-(27), compute $n_{\max}$ and $\{\Pr[\mathcal{P}_{n_t}], n_t \neq n_{\max}\}$.
7:         Randomly select an index $n_t' \neq n_{\max}$; the probability that n_t' is selected is $\Pr[\mathcal{P}_{n_t'}]$.
8:         Uniformly select a value s from $\{1,2,\ldots,N_{n_{\max}}\}$ and a value s' from $\{1,2,\ldots,N_{n_t'}\}$.
9:         Exchange $p_{n_{\max}}(s)$ and $p_{n_t'}(s')$ to form new partition $\{\mathcal{P}_1^*, \mathcal{P}_2^*,\ldots,\mathcal{P}_{N_t+1}^*\}$.
10:        Calculate the corresponding $\mu\{\boldsymbol{\Phi}_{n_{\max}}^*\}$, $\mu\{\boldsymbol{\Phi}_{n_t'}^*\}$, and the sum $\mu\{\boldsymbol{\Phi}^*\}$;
11:        **if** $((\mu\{\boldsymbol{\Phi}_{n_{\max}}^*\} - \mu\{\boldsymbol{\Phi}_{n_{\max}}\} < 0)$ and $(\mu\{\boldsymbol{\Phi}_{n_t'}^*\} - \mu\{\boldsymbol{\Phi}_{n_t'}\} < 0))$ or $((\exp(-(\mu\{\boldsymbol{\Phi}_{n_{\max}}^*\} - \mu\{\boldsymbol{\Phi}_{n_{\max}}\})/T) > rand())$ and $(\exp(-(\mu\{\boldsymbol{\Phi}_{n_t'}^*\} - \mu\{\boldsymbol{\Phi}_{n_t'}\})/T) > rand())$ and $(\mu\{\boldsymbol{\Phi}^*\} < \mu\{\boldsymbol{\Phi}\}))$ **then**
12:           $\{\mathcal{P}_1,\mathcal{P}_2,\ldots,\mathcal{P}_{N_t+1}\} \leftarrow \{\mathcal{P}_1^*,\mathcal{P}_2^*,\ldots,\mathcal{P}_{N_t+1}^*\}$;
13:           $\mu\{\boldsymbol{\Phi}_{n_{\max}}\} \leftarrow \mu\{\boldsymbol{\Phi}_{n_{\max}}^*\}$; $\mu\{\boldsymbol{\Phi}_{n_t'}\} \leftarrow \mu\{\boldsymbol{\Phi}_{n_t'}^*\}$; $\mu\{\boldsymbol{\Phi}\} \leftarrow \mu\{\boldsymbol{\Phi}^*\}$;
14:        **end if**
15:     **end for**
16:     $T \leftarrow T \cdot T_{rate}$;
17: **end while**
18: Output: $\{\mathcal{P}_1,\mathcal{P}_2,\ldots,\mathcal{P}_{N_t+1}\}$.

To further speed up convergence, we propose a smart implementation of the SA updating rules. The resulting ESA algorithm is outlined in Algorithm 1; the main novelties are the following.

a) Updating probability: The probability that a subset $\mathcal{P}_{n_t}$ gets updated, depends on the shrinking potential of its MMMC. For $n_t = 1,\ldots,N_t$, the shrinking potential E_{po,n_t} of $\mu\{\boldsymbol{\Phi}_{n_t}\}$ is defined as the difference between $\mu\{\boldsymbol{\Phi}_{n_t}\}$ and its lower Welch bound $\mu\{\boldsymbol{\Phi}_{n_t}\}_{\min}$:

$$E_{po,n_t} = \mu\{\boldsymbol{\Phi}_{n_t}\} - \mu\{\boldsymbol{\Phi}_{n_t}\}_{\min}, \quad (24)$$

where [24]

$$\mu\{\boldsymbol{\Phi}_{n_t}\}_{\min} = \sqrt{\frac{(L_{n_t,f(n_t)} - N_{n_t})}{(N_{n_t}(L_{n_t,f(n_t)} - 1))}}, \quad (25)$$

with $f(n_t)$ from (21) and N_{n_t} from (18). The shrinking potential $\mu\{\boldsymbol{\Phi}_{N_t+1}\}$ is set as the average potential of the other subsets. The subset corresponding to the largest potential, i.e., $\mathcal{P}_{n_{\max}}$, with

$$n_{\max} = \arg\max_{n_t=1,2,\ldots,N_t+1} E_{po,n_t} \quad (26)$$

is always one of the two updated subsets. The other updated subset is randomly selected, with the probability $\Pr[\mathcal{P}_{n_t}]$ of

selecting $\mathcal{P}_{n_t}$ with $n_t \neq n_{\max}$ proportional to E_{po,n_t}, i.e.,

$$\Pr[\mathcal{P}_{n_t}] = E_{po,n_t} / \sum_{n=1}^{N_t+1} (E_{po,n} - E_{po,n_{\max}}). \qquad (27)$$

b) Updating conditions: For updating the partition, ESA looks at the individual MMMC of the subsets that were selected for updating, rather than at the sum MMMC only. The old partition is replaced by the new one, only if (i) at least one of the individual MMMCs decreases and (ii) the sum MMMC decreases. If both MMMCs decrease, the new partition is always accepted. If only one of the MMMCs decreases and the sum MMMC also decreases, the new partition is accepted with a (small) probability that depends on the annealing temperature T.

The use of the shrinking potential to rationalize the SA updating probabilities and the updating conditions are innovative. They constitute essential parts of the design and they are decisive to boost the accuracy/complexity trade-off of the procedure in a MIMO scenario. Let us now analyze the complexity of ESA. Evaluating a single $\mu\{\Phi_{n_t}\}$ requires $\text{NOR}_{\text{MMMC}}(n_t) = 2N_{n_t}(L_{n_t,f(n_t)}-1)$ elementary operations on real quantities (NOR), with $f(n_t)$ defined in (21). Because ESA is a non-deterministic method, the total NOR performed with ESA (NOR_{ESA}) is a random variable and a comprehensible closed-form expression is hard to derive. As an alternative, a simple upper bound $\text{NOR}_{\text{ESA,up}}$ can be computed:

$$\begin{aligned}
\text{NOR}_{\text{ESA}} &\leq \text{NOR}_{\text{ESA,up}} \\
&= 2T_{iter} \log_{T_{rate}}(T_{stop}/T_{init}) \, \text{NOR}_{\text{MMMC,max}}
\end{aligned}$$
$$(28)$$

with $\text{NOR}_{\text{MMMC,max}} = \max_{n_t=1,2,\ldots,N_t} \text{NOR}_{\text{MMMC}}(n_t)$. The computational complexity of ESA is mainly determined by the MMMC evaluations in line 10 of Algorithm 1. In (28), $T_{iter} \log_{T_{rate}}(\frac{T_{stop}}{T_{init}})$ indicates the total number of iterations. At most two Tx antennas need to recalculate their MMMCs during each iteration. It is interesting to contrast (28) to the computational complexity associated with the SSS algorithm from [5], where in each iteration $\mu\{\Phi_{n_t}\}$ is evaluated $2(N - N_t)N_t$ times for all n_t (for more details see Section VI-D). The total NOR performed after T_{SSS} SSS iterations is

$$\begin{aligned}
\text{NOR}_{\text{SSS}} &= 2T_{\text{SSS}} \sum_{n_t=1}^{N_t} \text{NOR}_{\text{MMMC}}(n_t)(N - N_{n_t})N_{n_t} \\
&\leq 2T_{\text{SSS}} \text{NOR}_{\text{MMMC,max}} \sum_{n_t=1}^{N_t} (N - N_{n_t})N_{n_t}. \quad (29)
\end{aligned}$$

It can be concluded that for large N, SSS is significantly more complex per iteration than ESA. Moreover, previous work shows that SA converges faster than SSS [18]. It can be expected that the same holds for ESA. In the numerical results section the complexity of ESA will be discussed in more detail and further contrasted to other existing pilot allocation algorithms.
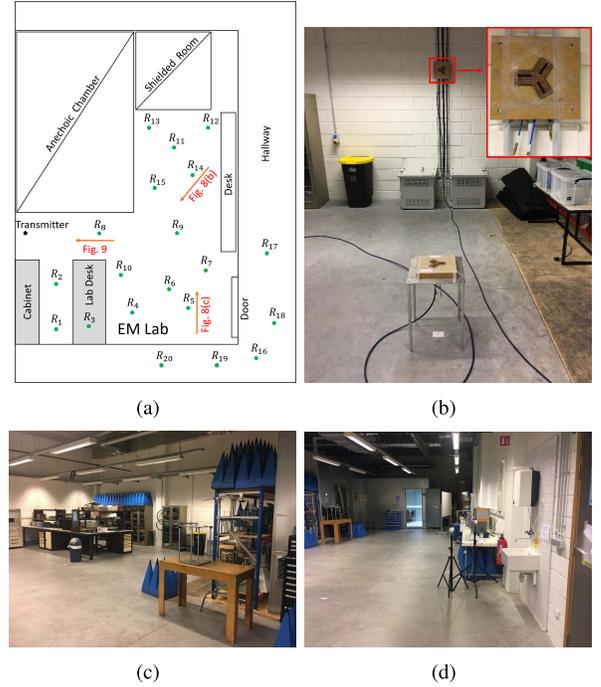

Fig. 3. (a) Floor plan of the environment, in which the channel sounding experiment was performed. (b) Integrated three-element UWB SIW cavity-backed slot antenna arrays at transmitter and receiver sides, were deployed approximately 3.1 m and 0.55 m above ground level, respectively. Measurement environment: (c) photograph of one side, (d) photograph of the other side.

V. MEASUREMENT SETUP

Measurements were conducted in an indoor laboratory environment at UGent. Fig. 3(a) depicts the laboratory plan with brick walls and reinforced concrete floors. The laboratory is approximately 19.2 m by 11.2 m. It contains two electromagnetically shielded rooms. The larger room is a rectangle of size 9.1 m by 4.2 m whereas the smaller room, is approximately 5 m long and 3.9 m wide. The rest of the laboratory is mostly equipped with metallic cabinets, desks, and hardware equipment (see Fig. 3(c) and Fig. 3(d)). The indoor radio channel with respect to a fixed transmitter was measured at 20 spatially distinct receiver positions. The Tx and Rx positions are outlined in the floor plan of Fig. 3(a). Half of the Rx positions (R_1-R_{10}) can be considered as line-of-sight (LoS) scenarios. A quarter of the Rx positions (R_{11}-R_{15}) can be considered as obstructed-LoS (OLoS) scenarios, where the free-space path from the transmitter to the receiver undergoes a reflection and/or a diffraction (walls). The last 5 Rx positions (R_{16}-R_{20}) can be considered as non-LoS (NLoS) scenarios, where the path from the transmitter to the receiver undergoes a transmission through a medium (wall, door). During the measurements, the door was kept closed and the hallway was empty.

Ultra wideband (UWB) channel sounding measurements were carried out at each of the 20 Rx positions. UWB cavity-backed slot antenna arrays in substrate integrated waveguide (SIW) technology [25] were employed at both ends of the measurement system. This type of antenna technology offers the benefit that the antenna does not suffer from mutual

coupling, which would otherwise distort the radio channel measurements. Each antenna array consists of three identical antenna elements, arranged in such a way that the array exhibits threefold rotational symmetry, resulting in an angle of $60°$ between subsequent antenna elements. The Tx and Rx antennas were placed approximately 3.1 m and 0.55 m above ground level, respectively. The measurement scenario with the receiver in position R_8 is pictured as an example in Fig. 3(b). A performance network analyzer (PNA) of type Agilent E8364B was used to probe the indoor radio channel ranging from 4 GHz to 7 GHz, this being the UWB frequency band. The PNA was used to measure the complex gain between each Tx antenna-element and each Rx antenna-element, individually. The feeder cables for the Tx and the Rx antennas were included in the PNA calibration in order to exclude them from the measurement data. All measurements were performed outside regular working hours.

VI. NUMERICAL RESULTS

In this section, simulations are conducted to demonstrate the effectiveness of the proposed algorithms. A QPSK scheme is adopted for both data and pilots. The number of subcarriers per OFDM symbol is $N = 1024$ and a bandwidth of 528 MHz is assumed, yielding a sample period of 1.89 nanoseconds. For the channel, two scenarios are considered:

1) A random scenario, where each of the MIMO subchannels is modeled in correspondence to one of the four communication environments and propagation scenarios proposed for the IEEE 803.15.3a WPAN standard. The corresponding channel models (CMs) are referred to as CM1, CM2, CM3 and CM4. CM1 and CM2 model the LoS and NLoS channel environments, for ranges smaller than 4 m. For larger ranges, the NLoS models CM3 and CM4 are used, with emphasis on a strong delay dispersion in the case of CM4 [7]. In our simulations, the channel length L and sparsity K are obtained as follows: (a) generate a channel realization according to a specific CM using the code provided by the IEEE 802.15.3a standard group, with a sampling period of 1/6 nsec; (b) resample the obtained CIR to obtain the sample spaced CIR; (c) find the amount of channel samples that suffice to cover 90% of the total power (this yields L); (d) count how many of these L channel samples have an amplitude larger than 10^{-4} (this yields K). Hence, the channel length and sparsity are determined to be $L = 180, 200, 350, 506$ and $K = 42, 56, 77, 121$ for CM1 to CM4, respectively.

2) A deterministic scenario, where the channels are taken from our measurements, performed in an indoor laboratory environment at UGent, Belgium.

Two metrics are used for performance evaluation: the NMSE (6) and the BER (7). These metrics are measured at a set of equidistant SNR points and determined by Monte-Carlo simulation. We average over 500 packets, with each packet consisting of a 1 OFDM symbol preamble and a 500 OFDM symbols payload. In the case of a random channel, a different channel realization is employed for each packet.

A. Priority of the Parameters L_{n_t,n_r} and K_{n_t,n_r}

This subsection analyzes the effect of L_{n_t,n_r} and K_{n_t,n_r} on the pilot BER, i.e.,

$$\text{BER}^{(\text{pilot})} \left(\tilde{K}_{n_t,n_r}, \tilde{L}_{n_t,n_r} \right)$$
$$= \mathbb{E} \left[\frac{\#\text{err}_{n_t,n_r} \left(\tilde{K}_{n_t,n_r}, \tilde{L}_{n_t,n_r} \right)}{m_p \cdot N_{n_t}} \right], \quad (30)$$

where $m_p = 2$ (QPSK), $\#\text{err}_{n_t,n_r}(\tilde{K}, \tilde{L})$ is defined in Section IV-A, and $\mathbb{E}[\cdot]$ denotes averaging over the noise, the channel, and the preamble of the n_t-th Tx antenna.

Fig. 4 shows the pilot BER (30) obtained for CM1 from [7], $N_{n_t} = 128$ and random pilot allocation, as a function of $\tilde{K}_{n_t,n_r}$ and $\tilde{L}_{n_t,n_r}$. In contrast to what was the case in [19], we observe from Fig. 4(b) that the pilot BER is relatively independent of $\tilde{L}_{n_t,n_r}$ for values of $\tilde{K}_{n_t,n_r}$ that significantly deviate from the actual channel sparsity $K_{n_t,n_r} = 42$. Therefore, it is almost impossible to estimate L_{n_t,n_r} prior to the estimation of K_{n_t,n_r}. On the other hand, Fig. 4(a) reveals that the pilot BER achieves a minimum at the optimum $\tilde{K}_{n_t,n_r} = K_{n_t,n_r}$ found in Fig. 4(b), independent of the value of $\tilde{L}_{n_t,n_r}$, provided that $\tilde{L}_{n_t,n_r}$ is larger than or equal to the true L_{n_t,n_r} (see the value of $\tilde{L}_{n_t,n_r}$ achieving the minimum pilot BER in Fig. 4(b) when $\tilde{K}_{n_t,n_r} = 42$ or 60). Hence, we propose to reverse the estimation order as compared to [19]. First, an estimate $\hat{K}_{n_t,n_r}$ of K_{n_t,n_r} is generated by keeping $\tilde{L}_{n_t,n_r}$ fixed and equal to some large value $L_{n_t,n_r}^{(\text{max})}$, with, e.g., $L_{n_t,n_r}^{(\text{max})} = N.^4$ Then, L_{n_t,n_r} is estimated, by minimizing $\#\text{err}_{n_t,n_r} \left(\tilde{K}_{n_t,n_r}, \tilde{L}_{n_t,n_r} \right)$ with $\tilde{K}_{n_t,n_r} = \hat{K}_{n_t,n_r}$. If $\#\text{err}_{n_t,n_r} \left(\tilde{K}_{n_t,n_r}, L_{n_t,n_r}^{(\text{max})} \right)$ is minimum for more than one trial value $\tilde{K}_{n_t,n_r}$, we choose $\hat{K}_{n_t,n_r}$ equal to the smallest one. Similarly, if $\#\text{err}_{n_t,n_r} \left(\hat{K}_{n_t,n_r}, \tilde{L}_{n_t,n_r} \right)$ is minimum for more than one value $\tilde{L}_{n_t,n_r}$, we choose $\hat{L}_{n_t,n_r}$ equal to the smallest one. Since the evaluation of $\#\text{err}_{n_t,n_r} \left(\tilde{K}_{n_t,n_r}, \tilde{L}_{n_t,n_r} \right)$ involves the reconstruction of $\boldsymbol{h}_{n_t,n_r}$, the described procedure not only estimates the duration and sparsity of the CIR, but also produces an estimate $\hat{\boldsymbol{h}}_{n_t,n_r}$ of the CIR itself. It should be noted, however, that this $\hat{\boldsymbol{h}}_{n_t,n_r}$ is not necessarily a reliable estimate of $\boldsymbol{h}_{n_t,n_r}$; this is usually only the case if (13) holds and the MMMC (14) is sufficiently small for $\hat{K}_{n_t,n_r}, \hat{L}_{n_t,n_r}$ and the considered pilot allocation. The above results can be summarized as follows. In the MU-MIMO-OFDM system, the optimal estimate of the channel sparsity in terms of minimum BER does neither depend on assumed channel length nor on the SNR provided that the assumed channel length is sufficiently large. On the other hand, no accurate estimate of the channel length can be obtained if the assumed channel sparsity is not close to the true channel sparsity. Therefore, the channel sparsity should be estimated prior to the channel length.

---

[4]Indeed, it is a valid assumption that the true value of L is smaller than N since it is assumed to be smaller than the size of the CP.
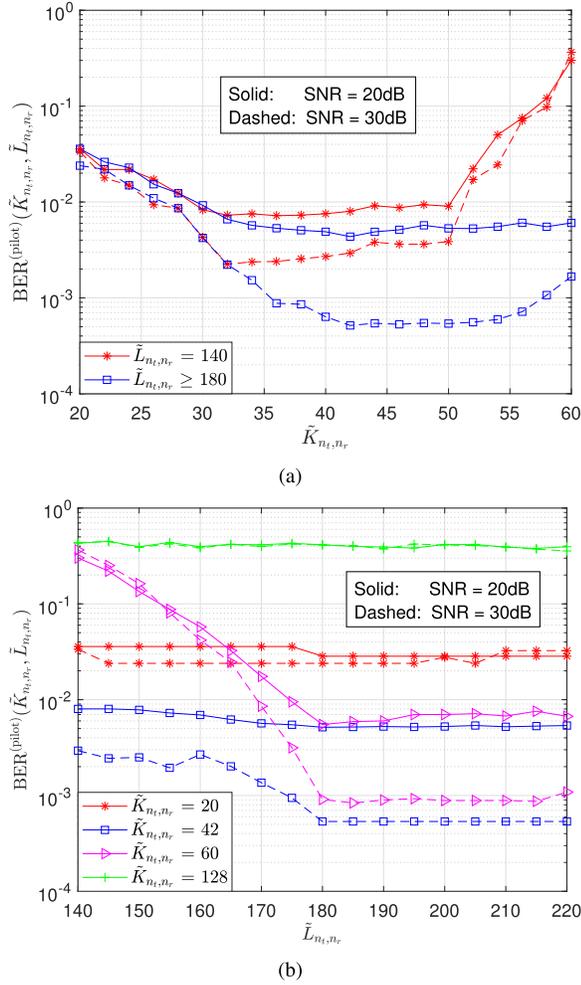
(a)

(b)

Fig. 4. Pilot BER for $(N, N_{n_t}, K_{n_t,n_r}, L_{n_t,n_r}) = (1024, 128, 42, 180)$, QPSK and $E_s/N_0 \in \{20\,\text{dB}, 30\,\text{dB}\}$; (a) as function of the assumed sparsity $\tilde{K}_{n_t,n_r}$, for different values of the assumed channel length $\tilde{L}_{n_t,n_r}$, and (b) as a function of $\tilde{L}_{n_t,n_r}$, for different values of $\tilde{K}_{n_t,n_r}$.

B. Channel Length and Sparsity Estimation

We now assess the performance of CoSaMP when the true sparsity K_{n_t,n_r} and the true length L_{n_t,n_r} of all the channels is replaced by estimates obtained using (16)-(17). We evaluate the NMSE and the BER that results from a ZFE receiver. We benchmark these performance metrics against those of a system with (1) a ZFE receiver with CoSaMP CIR reconstruction without prior knowledge of the channel statistics and assuming $L_{n_t,n_r} = N$ and $K_{n_t,n_r} = N_{n_t}$ for all (n_t, n_r), (2) a ZFE receiver with OMP CIR reconstruction [10], (3) a ZFE receiver with EM-GM-GAMP CIR reconstruction [12] and (4) a ZFE receiver with perfect channel knowledge. By means of example, a 4×4 uplink MU-MIMO system is considered with a 4 antenna AP and 4 single antenna UNs. A random channel is assumed, with CM1 for $\boldsymbol{h}_{1,n_r}$, CM2 for $\boldsymbol{h}_{2,n_r}$, CM3 for $\boldsymbol{h}_{3,n_r}$ and CM4 for $\boldsymbol{h}_{4,n_r}$, and $n_r = 1, 2, 3, 4$. Further, a random orthogonal pilot allocation is assumed, with $N_1 = 128$, $N_2 = 160$, $N_3 = 288$ and $N_4 = 448$. All subcarriers are allocated so $N_5 = 0$. Each Monte Carlo

simulation, a new orthogonal pilot allocation $\{\mathcal{P}_1, \ldots, \mathcal{P}_5\}$ with the mentioned subset sizes is randomly generated. Note that, for each (n_t, n_r) pair, the amount of pilots allocated to the n_t-th Tx antenna is smaller than the number of channel taps $h_{n_t,n_r}(k)$ that has to be estimated: $N_1 = 128 < L_{1,n_r} = 180$, $N_2 = 160 < L_{2,n_r} = 200$, $N_3 = 288 < L_{3,n_r} = 350$, $N_4 = 448 < L_{4,n_r} = 506$ (see CM specification at the beginning of this section). Fig. 5 (a) and (b) show that a CoSaMP channel estimator that assumes $L_{n_t,n_r} = N$ and $K_{n_t,n_r} = N_{n_t}$ underperforms. The NMSE is extremely high and the BER is close to 50%. This is because the CoSaMP algorithm is unable to converge if the assumed channel sparsity and/or the assumed channel length is significantly larger than the true value [26]. In contrast, it can be observed that if we first estimate K_{n_t,n_r}, then estimate L_{n_t,n_r}, and subsequently reconstruct the CIR using these estimates, the system achieves a prominent NMSE and BER improvement. CoSaMP with estimated L_{n_t,n_r} and K_{n_t,n_r} achieves virtually the same performance as OMP and EM-GM-GAMP. For the considered system set-up, the residual BER performance degradation with respect to the case that the channel is perfectly known at the receiver, is limited to a loss of about 2.5 dB in SNR. We note that this loss can be attributed to the random selection of the pilot allocation. In the following sections, we will show that this gap can be reduced further by optimizing the pilot allocation. Fig. 5 (a) and (b) also show the NMSE and BER that results when conventional LS or MMSE channel estimation (as opposed to CS-based channel estimation) is employed. In this case, the receiver first produces LS or MMSE estimates of the frequency-domain channel response samples at the pilot subcarriers, and then spline interpolation is performed to reconstruct the intermediate samples. The interpolation step is known to cause an error floor at high SNR. The MMSE estimator is assumed to follow a LS estimator and uses $\hat{\boldsymbol{h}}_{LS}\hat{\boldsymbol{h}}_{LS}^H$ instead of the true channel covariance matrix, with $\hat{\boldsymbol{h}}_{LS}$ the LS estimator output. We observe that, in the SNR range of interest, the CS-based CoSaMP algorithm performs more than an order of magnitude better than the conventional LS and MMSE channel estimation methods, both in terms of NMSE and BER. In general, proper operation of any conventional channel estimation method requires a number of pilots that is larger than or equal to the length of the unknown channel, while in the considered scenario this is not the case. Finally, we wish to note that the use of CS techniques (while reducing the required number of pilots) does not come at the expense of increased complexity. Our previous work [19] shows that the complexity of the procedure to estimate L_{n_t,n_r}, K_{n_t,n_r} and $\boldsymbol{h}_{n_t,n_r}$ using CoSaMP is not significantly larger than that of reconstructing $\boldsymbol{h}_{n_t,n_r}$ using frequency-domain MMSE estimation followed by spline interpolation. If N_{n_t} is small, the two approaches have a comparable complexity, whereas for large N_{n_t}, the CoSaMP approach even becomes the least complex. From the above results, we conclude that (in the considered MIMO context) a CoSaMP channel estimator, using previously obtained estimates of the channel length and sparsity, performs similar to other CS-based methods that do not require prior knowledge of L_{n_t,n_r} and K_{n_t,n_r}. For the
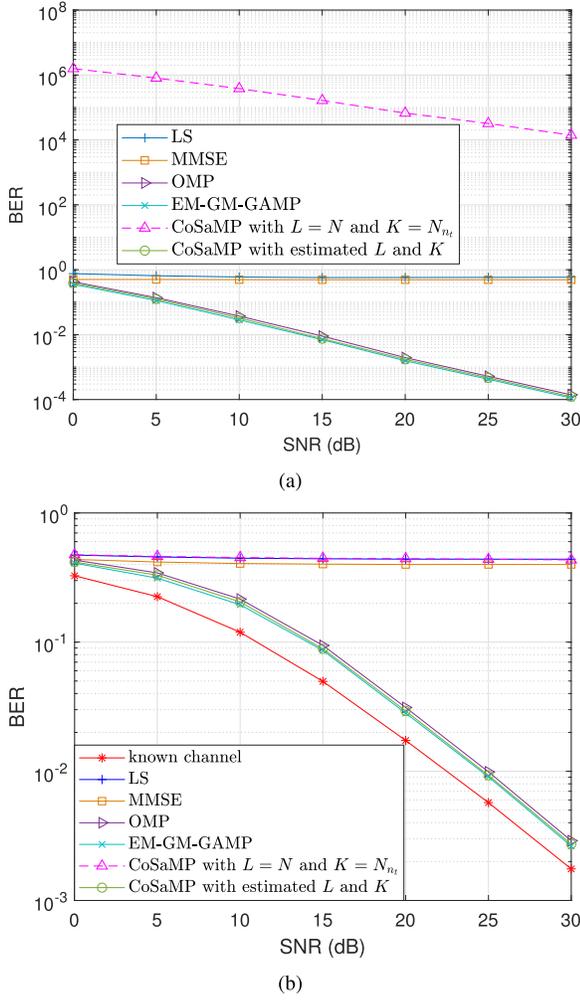
Fig. 5.   Parallel estimation of length and sparsity of 4×4 MIMO channels, with all channels from the i-th Tx antenna modeled using CMi, i=1,2,3,4. Performance comparison in terms of (a) NMSE and (b) BER.

considered scenario where L_{n_t,n_r} is larger than N_{n_t} (number of pilot subcarriers available for estimation), CS-based methods significantly outperform the conventional LS or MMSE techniques.

C. Number of Pilots Per Tx Antenna

We now demonstrate how a receiver can acquire (and track) the appropriate values for the number of pilots N_{n_t} in each per-Tx-antenna preamble. For illustration purposes, we consider an 8×8 uplink MU-MIMO system with random channels; for all $n_r = 1, 2, \ldots, 8$, we assume CM1 for $\boldsymbol{h}_{1,n_r}$, $\boldsymbol{h}_{2,n_r}$ and $\boldsymbol{h}_{3,n_r}$, CM2 for $\boldsymbol{h}_{4,n_r}$, $\boldsymbol{h}_{5,n_r}$ and $\boldsymbol{h}_{6,n_r}$, CM3 for $\boldsymbol{h}_{7,n_r}$, and CM4 for $\boldsymbol{h}_{8,n_r}$. The minimum number of pilot subcarriers N_{n_t} that should be allocated to Tx antenna $n_t = 1, 2, \ldots, 8$, is $N_{1,\min} = N_{2,\min} = N_{3,\min} = 89$, $N_{4,\min} = N_{5,\min} = N_{6,\min} = 103$, $N_{7,\min} = 169$, and $N_{8,\min} = 250$, respectively (see (13) and the CM specification at the start of this section). Summing all these values yields 995, which indicates that it should indeed be possible to accurately estimate all 8×8 MIMO channels using a single OFDM symbol preamble with 1024 subcarriers.

In practice, the true channel lengths and sparsities are not a priori known. In that case, the values N_{n_t} can be derived by following the iterative procedure described in Section IV-B. In Table I, we show how the appropriate values of $N_1, N_2, \ldots, N_8$ can be acquired in two iterations, for the considered MIMO system, when operating at an SNR of 30 dB.

- *Initialization:* No information about the channels is available, so we assign all Tx antennas the same number of pilot subcarriers ($N_1^{(0)} = N_2^{(0)} = \ldots = N_8^{(0)} = 128$) and randomly select an orthogonal pilot allocation with these equal subset sizes.
- *Iteration 1:* We estimate the sparsity K_{n_t,n_r} and channel duration L_{n_t,n_r} for each Tx-Rx pair based on the initial preamble. From these estimates, the minimum number of required pilots is determined based on (18) with $\delta_{n_t} = 1$. This results in $N_1^{(1)} = N_2^{(1)} = N_3^{(1)} = 89$, $N_4^{(1)} = N_5^{(1)} = N_6^{(1)} = 103$, $N_7^{(1)} = 182$ and $N_8^{(1)} = 266$. Note that the sum $N_1^{(1)} + N_2^{(1)} + \ldots + N_8^{(1)} = 1024$ implies that all subcarriers are allocated.
- *Iteration 2:* We use the novel pilot allocation to re-estimate the sparsity and channel length for the different Tx-Rx antenna pairs. The estimated sparsity and channel length of Tx antennas 1-6 does not change. This was to be expected as, in Iteration 1, $N_1^{(1)}, N_2^{(1)}, \ldots, N_6^{(1)}$ are smaller than 128, indicating that the sparsity and channel length of the channels from Tx antennas 1-6 can be estimated properly with the 128 pilot carriers per Tx antenna that were available in the initial preamble used in Iteration 1. However, as $N_7^{(1)}$ and $N_8^{(1)}$ are both larger than 128, the length and sparsity of the channels from Tx antennas 7-8 could not be estimated accurately in Iteration 1. Using the new estimates of K_{n_t,n_r} and L_{n_t,n_r} to compute $N_7^{(2)}$ and $N_8^{(2)}$ with (18) and $\delta_{n_t} = 1$, yields 169 and 250, respectively. Note that these values agree with the predicted minimum number of pilots, given in the first paragraph of this section. As $3 \cdot 89 + 3 \cdot 103 + 169 + 250 = 995$, there are 29 excess carriers left to judiciously distribute over the different Tx antennas. This can be done by choosing appropriate scaling factors δ_{n_t}, with a value larger than 1. Because antennas 7 and 8 correspond to the largest channel lengths and sparsities, these antennas are assigned a slightly larger scaling factor. We obtain $N_1^{(2)} = N_2^{(2)} = N_3^{(2)} = 90$, $N_4^{(2)} = N_5^{(2)} = N_6^{(2)} = 104$, $N_7^{(2)} = 176$ and $N_8^{(2)} = 266$ (summing up to 1024). A novel random orthogonal pilot allocation with these novel subset sizes needs to be selected. Since for all Tx antennas, $N_{n_t,n_r}^{(2)}$ is smaller than or equal to $N_{n_t}^{(1)}$, it is expected that the sparsity and channel length estimates obtained in Iteration 2 are accurate and further iterations are not required as long as the true channel length and sparsity values remain the same. Results not presented here confirm this presumption.

Fig. 6 illustrates the importance of allocating an appropriate number of pilots to each Tx antenna by comparing the BER performance of the considered 8×8 MIMO system when the receiver (1) has perfect channel knowledge, (2) uses

TABLE I

PROCEDURE TO DETERMINE THE APPROPRIATE NUMBER OF PILOT SUBCARRIERS IN EACH PER-ANTENNA PREAMBLE FOR AN 8×8 MIMO SYSTEM, OPERATING AT AN SNR OF 30 dB AND INITIALIZED WITH AN ARBITRARY INITIAL SUBCARRIER PARTITION WITH $N_1^{(0)} = N_2^{(0)} = \ldots = N_8^{(0)} = N/8 = 128$

n_t	n_r	$N_{n_t}^{(0)}$	min. $N_{n_t,n_r}^{(1)}(\hat{K}_{n_t,n_r}, \hat{L}_{n_t,n_r})$	$N_{n_t}^{(1)}(\delta_{n_t})$	min. $N_{n_t,n_r}^{(2)}(\hat{K}_{n_t,n_r}, \hat{L}_{n_t,n_r})$	$N_{n_t}^{(2)}(\delta_{n_t})$
1 to 3	1 to 8	128	89 (42,180)	89 (1)	89 (42,180)	90 (1.01)
4 to 6	1 to 8	128	103 (56,200)	103 (1)	103 (56,200)	104 (1.01)
7	1	128	177 (80,368)	182 (1)	169 (77,350)	176 (1.045)
	2		176 (79,367)		169 (77,350)	
	3		182 (83,379)		169 (77,350)	
	4		180 (81,375)		169 (77,350)	
	5		169 (77,352)		169 (77,350)	
	6		167 (76,348)		169 (77,350)	
	7		172 (79,356)		169 (77,350)	
	8		179 (82,372)		169 (77,350)	
8	1	128	253 (123,511)	266 (1)	250 (121,506)	266 (1.065)
	2		244 (118,494)		250 (121,506)	
	3		266 (129,538)		250 (121,506)	
	4		252 (122,509)		250 (121,506)	
	5		250 (121,506)		250 (121,506)	
	6		249 (120,505)		250 (121,506)	
	7		258 (125,521)		250 (121,506)	
	8		250 (120,506)		250 (121,506)	

CoSaMP channel reconstruction and a random preamble with $N_1 = N_2 = \ldots = N_8 = 128$ (Iteration 0), (3) uses CoSaMP channel reconstruction and a random preamble with $N_1 = N_2 = N_3 = 89$, $N_4 = N_5 = N_6 = 103$, $N_7 = 182$ and $N_8 = 266$ (Iteration 1), or (4) uses CoSaMP channel reconstruction and a random preamble with $N_1 = N_2 = N_3 = 90$, $N_4 = N_5 = N_6 = 104$, $N_7 = 176$ and $N_8 = 266$ (Iteration 2). The preamble with an equal number of pilots for every Tx antenna results in a very high BER. This is mainly because accurate estimates of subchannels of type CM3 and CM4 cannot be obtained in this case. In contrast, the preamble redesigned after Iteration 1 ensures accurate channel estimation for all subchannels and therefore results in a much lower BER. Although the pilot distribution in the preamble is further optimized in Iteration 2, we observe that the BER performance is comparable to that of Iteration 1. This is no surprise since in both iterations, the obtained number of pilot carriers to be assigned to each Tx antenna exceeds the minimum number of pilot subcarriers required to accurately estimate the channel.

Our results demonstrate that the strategy proposed in subsection IV-B is effective in appropriately distributing the pilot subcarriers over the TX antennas. At the start of the transmission, the amount of pilot subcarriers that is allocated to each Tx antenna does not yet optimally match that the channel conditions as seen by the different antennas. This has a negative impact on the system performance. During the acquisition period, the BER gradually decreases. After a few packets, a steady state is achieved. The corresponding BER is still significanty larger than in the case where the channel is perfectly known. The reason for this is that only the amount of subcarriers per TX antenna is optimized, not the actual allocation (which subcarriers can be used by which Tx antenna).

D. Joint Pilot Position Allocation for All Antenna Preambles

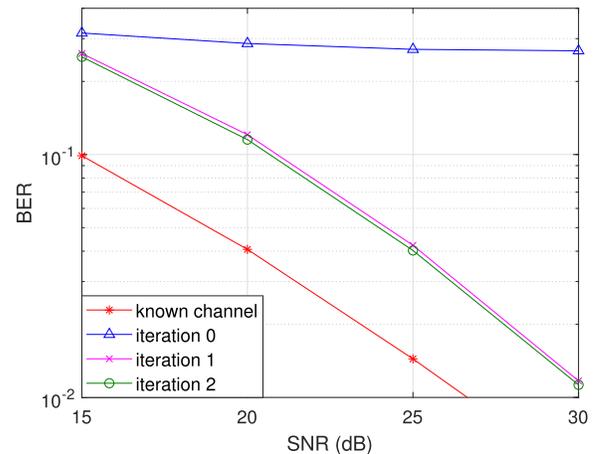In this subsection, we discuss the complexity, the convergence and the performance of the ESA procedure

Fig. 6. The measured channel between transmitter and R_8 receiver with tap spacing of $t = 1$ ns.

from Section IV-C to partition the available preamble subcarriers into per-Tx-antenna subsets $\mathcal{P}_{n_t} = \{p_{n_t}(1), p_{n_t}(2), \ldots, p_{n_t}(N_{n_t})\}$ of appropriate sizes N_{n_t}. The obtained results will be contrasted to those of the stochastic sequential search (SSS) procedure from [5], with one outer and T_{SSS} inner iterations. In each SSS inner iteration, each item $p_{n_t}(s)$ with $n_t = 1, 2, \ldots, N_t$ and $s = 1, 2, \ldots, N_{n_t}$ is successively swapped with each item of $\mathcal{P} \backslash \mathcal{P}_{n_t}$. The MMMCs of the partitions obtained after every swap are recorded, and the partition with the smallest sum MMMC is selected as the updated partition. Given the large number of MMMC evaluations involved in one iteration, it is clear that the complexity of SSS is very significant. For simplicity, the complexity of ESA and SSS is compared based on the upper bound expressions (28) and (29) derived at the end of Section IV-C, or equivalently on the maximum amount of MMMC evaluations performed, i.e., $N_{\mathrm{ESA}} = 2T_{iter} \log_{T_{rate}} \left(\frac{T_{stop}}{T_{init}} \right)$ for ESA and $N_{\mathrm{SSS}} = 2T_{\mathrm{SSS}} \sum_{n_t=1}^{N_t} N_{n_t}(N - N_{n_t})$ for SSS. It is important to note that both ESA and SSS will ultimately converge to the

TABLE II
PERFORMANCE COMPARISONS OF SSS AND ESA FOR $N = 1024$ AND $N_t = 8$

	Type	ID	Complexity	$\mu\{\Phi\}$	$[\mu\{\Phi_1\}, \mu\{\Phi_2\}, \mu\{\Phi_3\}, \mu\{\Phi_4\}, \mu\{\Phi_5\}, \mu\{\Phi_6\}, \mu\{\Phi_7\}, \mu\{\Phi_8\}]$
	initial PA	P0	\	1.7731	[0.2182, 0.2265, 0.2465, 0.2184, 0.2578, 0.2485, 0.1676, 0.1896]
SSS	$T_{\mathrm{SSS}} = 5$	SSS-1	8891060	1.1357	[0.1505, 0.1578, 0.1479, 0.1459, 0.1596, 0.1452, 0.1238, 0.1050]
	$T_{\mathrm{SSS}} = 10$	SSS-2	17782120	1.1004	[0.1483, 0.1555, 0.1616, 0.1374, 0.1419, 0.1446, 0.1151, 0.0960]
	$T_{\mathrm{SSS}} = 20$	SSS-3	35564240	1.0892	[0.1569, 0.1480, 0.1414, 0.1390, 0.1451, 0.1479, 0.1147, 0.0963]
ESA	$T_{rate} = 0.95$	ESA-1	27000	1.1073	[0.1467, 0.1518, 0.1450, 0.1437, 0.1434, 0.1446, 0.1249, 0.1073]
	$T_{rate} = 0.98$	ESA-2	68400	1.0831	[0.1497, 0.1494, 0.1499, 0.1380, 0.1397, 0.1392, 0.1186, 0.0987]
	$T_{rate} = 0.99$	ESA-3	137500	1.0396	[0.1368, 0.1454, 0.1427, 0.1298, 0.1324, 0.1348, 0.1177, 0.1001]

optimum pilot allocation (as will for example also a simple but stupid random search algorithm). As a consequence, what we are aiming for is not so much to show that ESA can outperform SSS in terms of BER, but rather to show that ESA is capable of achieving a slightly better BER performance with a significantly lower computational effort.

A first set of results is obtained by considering the same 8×8 random MIMO channel as in the previous subsection. For this channel, it is appropriate to allocate $N_1 = N_2 = N_3 = 90$, $N_4 = N_5 = N_6 = 104$, $N_7 = 176$ and $N_8 = 266$ pilot subcarriers to Tx antennas 1-3, 4-6, 7 and 8, respectively. Table II provides a comparison of SSS and ESA in terms of complexity (maximum number of MMMC evaluations performed) and performance (minimum value of sum MMMC and per-Tx-antenna MMMCs achieved). For a fair comparison, SSS and ESA are bootstrapped with the same initial randomly generated subcarrier partition P0. The sum MMMC and the per-Tx-antenna MMMCs of P0 are provided in the first row of Table II. For SSS, results are presented after T_{SSS} equal to 5, 10 and 20 iterations. For ESA, the design parameters T_{init}, T_{stop} and T_{iter} are fixed to 10^{-2}, 10^{-8} and 50, respectively; for the design parameter T_{rate} the values 0.95, 0.98 and 0.99 are employed.

We make the following observations from Table II:

- As could be expected, the performance of SSS and ESA improves when more potential subcarrier partitions are tested. It can be observed that, when T_{SSS} of SSS increases from 5 to 20, or T_{rate} of ESA increases from 0.95 to 0.99, the number of MMMC evaluations (complexity in Table II) also increases, resulting in a smaller value of sum MMMC $\mu\{\Phi\}$.
- In general, SSS converges much slower than ESA. With SSS, the smallest MMMC achieved after more than $3 \cdot 10^7$ MMMC evaluations is 1.0892; with ESA, the smallest MMMC achieved after less than $1.5 \cdot 10^5$ MMMC evaluations is already as small as 1.0396.
- As opposed to SSS, ESA not only concentrates on decreasing the sum MMMC, but also tries to make the per-Tx-antenna MMMC as small as possible.

Fig. 7 shows the BER performance of a ZFE receiver using preamble-based CoSaMP channel reconstruction when the preamble is the one with ID P0, SSS-3 or ESA-3 in Table II. The BER of a ZFE receiver with perfect channel knowledge is also shown. We observe that the BER follows the MMMC, i.e., the smaller the MMMC, the smaller the BER.

A second set of results is obtained for a deterministic 3×3 MIMO channel realization, for which we use measurements
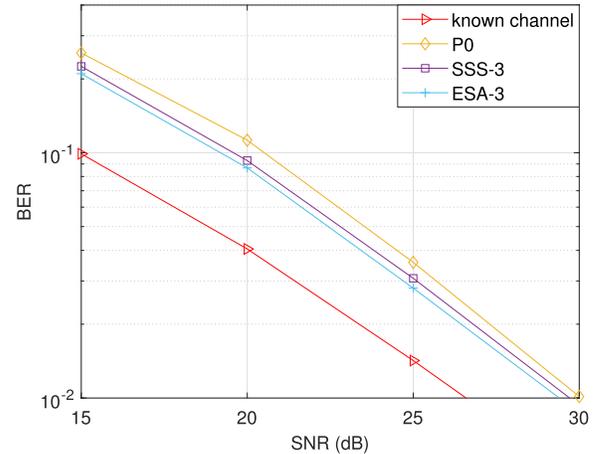

Fig. 7. Impact of pilot allocation procedure on BER performance.

from the channel sounding experiment described in Section V. We consider a set-up whereby the Tx antennas are sparsely distributed in space while the Rx antennas are co-located. The channel between Tx antennas 1, 2 and 3, and the 3-antenna receiver is modeled using measurements from one of the antenna-elements at positions R_8, R_{11} and R_{18}, respectively (see Fig. 3(a) in Section V). The true length and sparsity of the employed channels are listed in Table III (as a reference); these values were determined using the method described in [27]. While the channel length is more or less the same for all Tx antennas, the (LoS) channel from Tx antenna 1 is significantly more sparse than the (OLoS) channel from Tx antenna 2, which in turn is more sparse than the (NLoS) channel from Tx antenna 3.

Table III also shows the computation of $\{N_1, N_2, N_3\}$ and lists the measured length and sparsity for each subchannel. Initial estimates of L_{n_t,n_r} and K_{n_t,n_r} are obtained using an initial preamble with a random partition of the available subcarriers into three sets of almost equal size ($N_1 = 342$, $N_2 = 341$ and $N_3 = 341$). From these estimates, it is found that at least 165, 181 and 195 pilots have to be allocated to Tx antennas 1, 2 and 3, respectively. Since these values are significantly smaller than the values employed to estimate L_{n_t,n_r} and K_{n_t,n_r}, it can be assumed that convergence is achieved immediately, after only one iteration. To show the impact of the number of pilots on the overall system performance, BER results are computed not only for $(N_1, N_2, N_3, N_4) = (165, 181, 195, 483)$, but also for $(N_1, N_2, N_3, N_4) = (198, 217, 234, 375)$, $(247, 271, 292, 214)$

TABLE III

PARAMETER ESTIMATION AND NUMBER OF PILOT SUBCARRIERS PER TX ANTENNA FOR A 3×3 MIMO SYSTEM WITH SPATIALLY DISTRIBUTED TX ANTENNAS AND MEASURED CHANNELS

n_t	n_r	K_{n_t,n_r}	L_{n_t,n_r}	it. 0		it. 1			
				N_{n_t}	min. $N_{n_t,n_r}(\hat{L}_{n_t,n_r}, \hat{K}_{n_t,n_r})$	$N_{n_t}(\delta_{n_t})$	$N_{n_t}(\delta_{n_t})$	$N_{n_t}(\delta_{n_t})$	$N_{n_t}(\delta_{n_t})$
	1	64	339		154 (339,64)				
1	2	65	376	342	165 (376,64)	165 (1)	198 (1.2)	247 (1.5)	312 (1.9)
	3	62	365		159 (365,62)				
	1	82	366		177 (366,82)				
2	2	79	359	341	173 (359,79)	181 (1)	217 (1.2)	271 (1.5)	343 (1.9)
	3	87	366		181 (366,87)				
	1	111	371		194 (371,111)				
3	2	110	374	341	195 (374,110)	195 (1)	234 (1.2)	292 (1.5)	369 (1.9)
	3	107	366		190 (366,107)				

and (312,343,369,0), where N_4 is the number of pilot subcarriers unused for channel estimation. Each of these corresponds to choosing a different value of δ_{n_t} in (18), as specified in Table III. For each of the δ_{n_t} values, an initial preamble with a random partition of the available subcarriers into sets of size N_1, N_2 and N_3, respectively, is further optimized using ESA with $T_{init} = 10^{-2}$, $T_{stop} = 10^{-8}$, $T_{rate} = 0.99$, and $T_{iter} = 50$. For $\delta_{n_t} = 1$, 1.2, 1.5 and 1.9, the sum MMMC of the initial preamble is 0.5397, 0.4909, 0.4266 and 0.3697, respectively; after ESA this reduces to 0.3251, 0.2928, 0.2478 and 0.2140, respectively. Using the obtained preamble, CoSaMP channel reconstruction and ZFE is applied. The resulting BER is shown in Fig. 8(a). With $\delta_{n_t} = 1$, a BER below 0.2 cannot be achieved for SNR values below 30 dB. The BER performance improves significantly if δ_{n_t} increases from 1 to 1.2 and further above 1.5, the BER performance remains more or less constant. Compared to a scenario where the channel is perfectly known at the receiver, the BER degradation with $\delta_{n_t} = 1.5$ or 1.9 is about 1 dB in SNR, which is very acceptable in practice.

So far, we have shown that the proposed three-step procedure for adaptive pilot allocation is much more effective in finding an appropriate orthogonal pilot allocation than the state-of-the-art approach from [5] and that this results in a good overall BER performance. For the sake of completeness, we further compare the ESA algorithm with the deterministic pilot design (DPD) algorithm from [2] and the genetic algorithm (GA) from [6]. For this comparison, we use the same measured 3×3 MIMO channel as in Fig. 8(a) and the parameters of ESA and SSS are again set as $T_{init} = 10^{-2}$, $T_{stop} = 10^{-8}$, $T_{rate} = 0.99$, $T_{iter} = 50$ and $T_{\text{SSS}} = 20$. DPD basically employs a modified version of SSS for the SISO case, adding M_d outer iterations. In our simulations, we take $M_d = 100$. For GA, we use the design parameter values employed in [6]. All algorithms are initialized with the same preamble. As opposed to ESA and SSS, DPD and GA always allocate an equal amount of pilot subcarriers to each Tx antenna. As explained in the introduction, this is not optimum in scenarios where some Tx antennas have much better channel state conditions than others. The per-TX-antenna preambles designed by DPD or GA are, in fact, shifted versions of each other; in particular, $p_{n_t}(i + n_t - 1) = p_1(i)$, for $n_t = 2, 3, \ldots, N_t$. The orthogonality condition then implies that the subcarriers in $\mathcal{P}_1$ are spaced by at least N_t. Further, because equidistant pilot allocation is known to
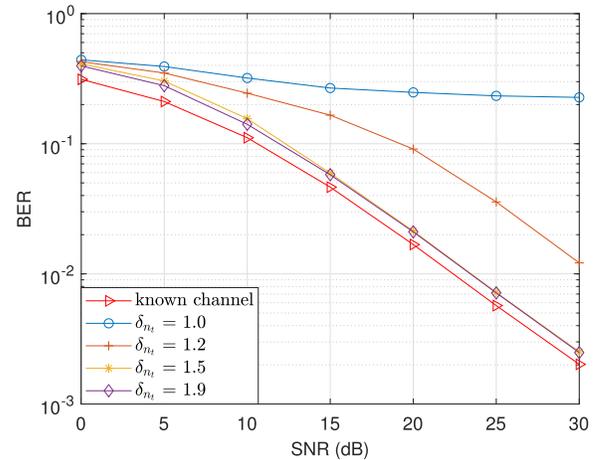
(a)
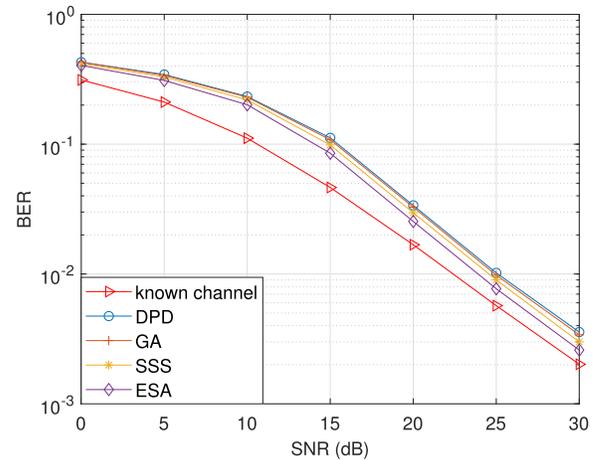
(b)

Fig. 8. A virtual 3×3 MIMO system, with 3 LoS, 3 OLoS and 3 NLoS measured subchannels: (a) BER performance with ESA preamble design if a varying fraction of the 1024 available preamble subcarriers are employed for CoSaMP channel reconstruction ($N_4 = 483$, 375, 214 and 0 for $\delta_{n_t} = 1$, 1.2, 1.5 and 1.9, respectively), (b) BER performance comparison of DPD, GA, SSS and ESA when $(N_1, N_2, N_3, N_4) = (250, 250, 250, 274)$.

yield a very high MMMC, meaningful DPD and GA designs are only possible if a significant amount of pilot subcarriers remain unused. In our simulation, we have imposed that every Tx antenna is allocated 250 pilot subcarriers, while 274 subcarriers remain unused. As can be observed from the BER comparison results shown in Fig. 8(b), ESA has a better

BER performance than DPD, GA and SSS. At a BER of $4 \cdot 10^{-3}$, the gain of ESA in terms of SNR is about 2 dB with respect to DPD and GA, and about 1 dB with respect to SSS. The sum MMMCs achieved with DPD, GA, SSS and ESA are 0.3561, 0.3336, 0.2834 and 0.2735, respectively. In terms of complexity ESA also outperforms the other algorithms. The maximum number of MMMC evaluations involved in ESA and SSS is $N_{\mathrm{ESA}} = 1.4 \cdot 10^5$ and $N_{\mathrm{SSS}} = 3.6 \cdot 10^7$, respectively (Table II). The maximum amount of MMMC calculations in DPD is $N_{\mathrm{DPD}} = M_d T_{\mathrm{SSS}} N_1 (N - 3N_1) = 1.4 \cdot 10^8$. Finally, GA produces 90 new pilot allocations per iteration and performs 5000 iterations, making a total of $N_{\mathrm{GA}} = 4.5 \cdot 10^5$ MMMC evaluations. We conclude that the proposed ESA algorithm, not only is better suited for scenarios where some Tx antennas have much better channel state conditions than others, but also provides a better performance complexity trade-off than DPD and GA.

VII. CONCLUSION

In this paper, we proposed an algorithm to accurately estimate the channel length and sparsity of all channels in a distributed MIMO-OFDM spatial-multiplexing system. Due to the distributed nature of the system, the channel length and sparsity of the diverse channels may vary significantly. As a result, the number of pilot subcarriers needed for the estimation of each channel is different. Therefore, we also proposed a method to determine the number of pilot subcarriers in each per-Tx-antenna preamble. Finally, we developed a novel low-complexity algorithm to select appropriate positions for these pilot subcarriers while adhering to the orthogonality principle. Numerical performance results were presented for both theoretical channel models and measured channels and show that compared to the state-of-the-art method, our approach not only demonstrates much faster converge, but also improves the system performance in terms of NMSE and BER.

REFERENCES

[1] Y. Zeng, A. R. Leyman, and T.-S. Ng, "Joint semiblind frequency offset and channel estimation for multiuser MIMO-OFDM uplink," *IEEE Trans. Commun.*, vol. 55, no. 12, pp. 2270–2278, Dec. 2007.

[2] R. Mohammadian, A. Amini, and B. H. Khalaj, "Deterministic pilot design for sparse channel estimation in MISO/multi-user OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 129–140, Jan. 2017.

[3] 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 Commun. Lett.*, vol. 4, no. 4, pp. 345–348, Aug. 2015.

[4] 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, Sep. 2013.

[5] C. Qi, G. Yue, L. Wu, Y. Huang, and A. Nallanathan, "Pilot design schemes for sparse channel estimation in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1493–1505, Apr. 2015.

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

[7] J. Foerster, *Channel Modeling Sub-Committee Report Final*, document IEEE P802. 15-02/490r1-SG3a, 15 Working Group for Wireless Personal Area Networks (WPANs), 2003.

[8] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[9] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.

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

[11] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, Jul. 2009.

[12] J. P. Vila and P. Schniter, "Expectation-maximization Gaussian-mixture approximate message passing," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4658–4672, Oct. 2013.

[13] R. Prasad, C. R. Murthy, and B. D. Rao, "Joint channel estimation and data detection in MIMO-OFDM systems: A sparse Bayesian learning approach," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5369–5382, Oct. 2015.

[14] X. Kuai, L. Chen, X. Yuan, and A. Liu, "Structured turbo compressed sensing for downlink massive MIMO-OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 3813–3826, Aug. 2019.

[15] S. Srivastava, M. S. Kumar, A. Mishra, S. Chopra, A. K. Jagannatham, and L. Hanzo, "Sparse doubly-selective channel estimation techniques for OSTBC MIMO-OFDM systems: A hierarchical Bayesian Kalman filter based approach," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4844–4858, Aug. 2020.

[16] Y. Nie, "Deterministic pilot pattern placement optimization for OFDM sparse channel estimation," *IEEE Access*, vol. 8, pp. 124783–124791, 2020.

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

[18] T. Li, N. Noels, and H. Steendam, "Efficient pilot allocation for sparse channel estimation in UWB OFDM systems," *Signal Process.*, vol. 175, Oct. 2020, Art. no. 107666, doi: 10.1016/j.sigpro.2020.107666.

[19] T. Li, B. Hanssens, W. Joseph, and H. Steendam, "Parameter estimation and channel reconstruction based on compressive sensing for ultra-wideband MB-OFDM systems," *Signal Process.*, vol. 167, Feb. 2020, Art. no. 107318, doi: 10.1016/j.sigpro.2019.107318.

[20] R. M. Rao, V. Marojevic, and J. H. Reed, "Adaptive pilot patterns for CA-OFDM systems in nonstationary wireless channels," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1231–1244, Feb. 2018.

[21] M. A. Matin, *Ultra Wideband Communications: Novel Trends-Antennas and Propagation*. Norderstedt, Germany: BoD–Books on Demand, 2011.

[22] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.

[23] R. Mohammadian, A. Amini, B. H. Khalaj, and N. Omidvar, "MIMO-OFDM pilot symbol design for sparse channel estimation," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA)*, Dec. 2015, pp. 975–980.

[24] L. R. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 3, pp. 397–399, May 1974.

[25] S. Lemey *et al.*, "Threefold rotationally symmetric SIW antenna array for ultra-short-range MIMO communication," *IEEE Trans. Antennas Propag.*, vol. 64, no. 5, pp. 1689–1699, May 2016.

[26] D. K. Pal and O. J. Mengshoel, "Stochastic CoSaMP: Randomizing greedy pursuit for sparse signal recovery," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Riva del Garda, Italy: Springer, 2016, pp. 761–776.

[27] B. Hanssens *et al.*, "Measurement-based analysis of specular and dense multipath components at 94 GHz in an indoor environment," *IET Microw., Antennas Propag.*, vol. 12, no. 4, pp. 509–515, 2018.

Taoyong Li was born in Liancheng, Fujian, China, in 1991. He received the B.Sc. degree in computer science and technology from Tsinghua University, Beijing, China, in 2014, and the Ph.D. degree in electrical engineering from Ghent University, Ghent, Belgium, in 2020. He is currently a Lecturer with the Information and Navigation Scholl, Air Force Engineering University, Xi'an, China, and also a Researcher with the Collaborative Innovation Center of Information Sensing and Understanding, Xi'an. His current research interests include wireless communications, ultra-wideband communication, orthogonal frequency division multiplexing, MIMO systems, channel estimation, and compressive sensing signal processing.

Nele Noels (Senior Member, IEEE) received the Diploma of Electrical Engineering and Ph.D. degrees in electrical engineering from Ghent University, Ghent, Belgium, in 2001 and 2009, respectively.

She is currently a Professor at the Department of Telecommunications and Information Processing (TELIN), Ghent University. She is the (co)author of over 60 academic articles in international journals and conference proceedings. Her main research interests are in statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, massive MIMO, optical OFDM, satellite, and mobile communication.

Dr. Noels was a recipient of several scientific awards. In 2010, she received the Scientific Award Alcatel Lucent Bell for the Best Belgian Thesis concerning an original study of information and communication technology, concepts, and/or applications. She is currently an Editor of *Communication Theory and Systems* (Division I) and the *Journal of Communications and Networks*.

Kamil Yavuz Kapusuz (Member, IEEE) received the M.Sc. degree in electrical engineering from Atilim University, Ankara, Turkey, in 2013, and the Ph.D. degree in electrical engineering from Ghent University, Ghent, Belgium, in 2021.

From 2014 to 2016, he was a Senior Antenna Engineer with Remote Sensing Technologies, Ankara. He is currently a Post-Doctoral Fellow with the Ghent University-imec, Ghent, and a Visiting Researcher with the French National Center for Scientific Research (CNRS), Institut d'Electronique et des Technologies du numéRique (IETR), Rennes, France. His research interests include the analysis and design of passive RF components, leaky-wave antennas, periodic structures, multifunction and reconfigurable antennas, active/passive phased arrays, millimeter-wave antennas, and near-field focusing techniques.

Dr. Kapusuz received the URSI Young Scientist Award at the 2021 URSI General Assembly.

Sam Lemey (Member, IEEE) received the M.Sc. degree in electronic engineering from Howest, University College West Flanders, Kortrijk, Belgium, in 2012, and the Ph.D. degree in electrical engineering from Ghent University, Ghent, Belgium, in 2016. From January 2018 to March 2018, he was a Visiting Scientist with the Terahertz Photonics Group, Institute of Electronics, Microelectronics and Nanotechnology (IEMN), University Lille Nord de France, Lille, France. He is currently an Assistant Professor with the Department of Information Technology (INTEC), Ghent University-imec. His research interests include antenna systems for wearable applications, active antenna design for the Internet of Things and (beyond) 5G applications, (opto-electronic) millimeter-wave multi-antenna systems, impulse-radio ultra-wideband antenna systems for centimeter-precision localization, and full-wave/circuit co-optimization frameworks to realize (opto-electronic) active (multi-)antenna systems. He received the URSI Young Scientist Award at the 2020 URSI General Assembly and the Best Paper Award at the 2016 IEEE MTT-S Topical Conference on Wireless Sensors and Sensor Networks. He was a co-recipient of the 2015 Best Paper Award at the 22nd IEEE Symposium on Communications and Vehicular Technology in the Benelux and of the 2019 ECOC Best Demo Award.

Hendrik Rogier (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from Ghent University, Ghent, Belgium, in 1994 and 1999, respectively.

From 2003 to 2004, he was a Visiting Scientist with the Mobile Communications Group, Vienna University of Technology, Vienna, Austria. He is currently a Senior Full Professor with the Department of Information Technology, Ghent University, and a Guest Professor with the Interuniversity Microelectronics Centre, Ghent. He has authored or coauthored over 190 articles in international journals and over 210 contributions in conference proceedings. His current research interests include antenna systems, radio wave propagation, body-centric communication, numerical electromagnetics, electromagnetic compatibility, and power/signal integrity.

Dr. Rogier is a member of the MTT-26 RFID, Wireless Sensor and IoT Committee, and he acts as the URSI Commission B Representative for Belgium. He was a recipient of the URSI Young Scientist Award (twice) at the 2001 URSI Symposium on Electromagnetic Theory and at the 2002 URSI General Assembly, the 2014 Premium Award for Best Paper in the *Electronics Letters* (IET), the Best Paper Award First Place in the 2016 IEEE MTT-S Topical Conference on Wireless Sensors and Sensor Networks, the Best Poster Paper Award at the 2012 IEEE Electrical Design of Advanced Packaging and Systems Symposium, the Best Paper Award at the 2013 IEEE Workshop on Signal and Power Integrity, and the Joseph Morrissey Memorial Award for the First Best Scientific Paper at BioEM 2013. From 2017 until 2019, he was an Associate Editor of IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES. Currently, he is an Associate Editor of *Electronics Letters* (IET) and *IET Microwaves, Antennas and Propagation*.

Heidi Steendam (Senior Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree in applied sciences from Ghent University, Ghent, Belgium, in 1995 and 2000, respectively.

Since September 1995, she has been with the Digital Communications (DIGCOM) Research Group, Department of Telecommunications and Information Processing (TELIN), Faculty of Engineering, Ghent University, first in the framework of various research projects, and since October 2002, as a Professor in the area of digital communications. In 2015, she was a Visiting Professor at Monash University. She is the author of more than 150 scientific articles in international journals and conference proceedings, for which several best paper awards were received. Her main research interests are in statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, cognitive radio and cooperative networks, radar sensing, positioning, and visible light communication. Since 2002, she has been an Executive Committee Member of the IEEE Communications and Vehicular Technology Society Joint Chapter, Benelux Section, since 2012, the Vice Chair, and since 2017, the Chair. She was active in various international conferences as the Technical Program Committee Chair/Member and the Session Chair. In 2004, 2011, and 2018, she was the Conference Chair of the IEEE Symposium on Communications and Vehicular Technology, Benelux. From 2012 till 2017, she was an Associate Editor of IEEE TRANSACTIONS ON COMMUNICATIONS and *EURASIP Journal on Wireless Communications and Networking*.