Pilot-aided angle-domain channel estimation techniques for MIMO-OFDM systems

Citation for published version (APA):

Huang, L., Ho, C. K., Bergmans, J. W. M., & Willems, F. M. J. (2008). Pilot-aided angle-domain channel estimation techniques for MIMO-OFDM systems. *IEEE Transactions on Vehicular Technology*, 57(2), 906-920. <https://doi.org/10.1109/TVT.2007.905621>

DOI:

[10.1109/TVT.2007.905621](https://doi.org/10.1109/TVT.2007.905621)

Document status and date:

Published: 01/01/2008

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Pilot-Aided Angle-Domain Channel Estimation Techniques for MIMO-OFDM Systems

Li Huang, *Member, IEEE*, Chin Keong Ho, *Member, IEEE*,
J. W. M. Bergmans, *Senior Member, IEEE*, and Frans M. J. Willems, *Fellow, IEEE*

Abstract—Early multiple-input multiple-output with orthogonal frequency division multiplexing (MIMO-OFDM) channel estimation techniques treat channels as spatially uncorrelated. However, in many situations, MIMO-OFDM channels tend to be spatially correlated, for example, due to limited scattering. For such channels, estimation performance can be improved through exploitation of prior knowledge of the channel spatial correlation, for example, by means of the linear multiple mean square error (MMSE) technique. This knowledge is, however, not always available. As an alternative, we investigate techniques in the angle domain, where the MIMO-OFDM channel model lends itself to a physical interpretation. Our theoretical analysis and simulation results indicate that the proposed angle-domain approximated MMSE (AMMSE) channel estimation technique performs well in terms of the mean square error (mse) for various channel models representing different indoor environments. When a suitable threshold is chosen, we can use the angle-domain most-significant-taps selection technique instead of the angle-domain AMMSE technique to simplify the channel estimation procedure with little performance loss.

Index Terms—Channel estimation, multiple-input multiple-output (MIMO), orthogonal frequency division multiplexing (OFDM), spatial correlation.

I. INTRODUCTION

RECENT research trends have shown that combining the multiple-input multiple-output (MIMO) approach with orthogonal frequency division multiplexing (OFDM) can help to achieve spatial diversity and/or space-division multiplexing gain [1]–[4]. As coherent demodulation, which requires and utilizes the knowledge of channel coefficients, can achieve a 3-dB performance gain compared with differential demodulation [5], it is quite commonly adopted in MIMO-OFDM systems. Therefore, accurate and robust channel estimation¹ that permits

the realization of coherent demodulation is very important to ensure reliable data recovery.

Generally speaking, pilot-aided channel estimation is based on either the least squares (LS) [6] or the linear MMSE (LMMSE) technique [7], [8]. The essential difference between these two types of techniques is that the channel coefficients are treated as deterministic but unknown constants in the former, and as random variables of a stochastic process in the latter. Compared with LS-based techniques, LMMSE-based techniques yield better performance because they additionally exploit and require prior knowledge of the channel correlation. However, the channel correlation is sometimes not *a priori* known, which makes LMMSE-based techniques infeasible. To consider wider applications, we will focus on techniques that do not require prior knowledge of the channel correlation.

A typical MIMO-OFDM channel is conceived of as the unique link between the transmitted and noiseless received signals, and is referred to as the *array-domain* channel. This array-domain channel is treated as spatially uncorrelated in most previous pilot-aided channel estimation techniques for MIMO-OFDM systems (e.g., [4], [9], and [10]), possibly due to the fact that early MIMO studies assume the array-domain channel to be spatially uncorrelated (e.g., [11] and [12]). We call these techniques LS-based techniques in the array domain. However, in many realistic scenarios, the MIMO-OFDM channel tends to be spatially correlated, for example, due to antenna spacing constraints and limited scattering [13]–[15]. For these spatially correlated MIMO-OFDM systems, the LMMSE-based techniques in the array domain, which exploit and require prior knowledge of the channel spatial correlation, yield better performance than LS-based techniques in the array domain [16]–[18]. However, when the channel spatial correlation is not *a priori* known, which is the assumption made in this paper, these techniques are not applicable. In such cases, to improve the performance of conventional LS-based techniques in the array domain, we investigate techniques in the *angle domain*, where the channel model lends itself to a simple physical interpretation.

In the angle domain, beamforming patterns with different main lobes are used to characterize the physical propagation environment [19], [20]. For a MIMO system with N_t transmit and N_r receive antennas, the beamforming patterns have N_t transmit lobes and N_r receive lobes. A pair of transmit and receive lobes forms one angle-domain bin, and, thus, the angle domain is partitioned into $(N_t \times N_r)$ angle-domain bins. For example, as shown in Fig. 1, the transmit lobe 0 together with receive lobe 0 corresponds to the angle-domain bin (0, 0).

Manuscript received September 20, 2006; revised May 9, 2007 and June 17, 2007. The review of this paper was coordinated by Prof. X.-G. Xia.

L. Huang is with the Interuniversity Microelectronics Centre Nederland (IMEC-NL), Holst Centre, 5605 KN Eindhoven, The Netherlands, the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576, and the Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands (e-mail: Huang_Li@ieee.org).

C. K. Ho is with the Institute for Infocomm Research, Singapore 119613.

J. W. M. Bergmans and F. M. J. Willems are with the Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2007.905621

¹In this paper, we refer to the channel as the physical fading channel exclusive of the additive electronics noise. This notation is conventionally used in the field of channel estimation techniques.

Fig. 1. Schematic angle-domain representation of a MIMO channel with four transmit and four receive antennas.

Then, multiple unresolvable physical paths (e.g., paths 1 and 2) that occur in the angle-domain bin $(0, 0)$ can be approximately aggregated into one resolvable path, and the paths from other directions (e.g., paths 3 and 4) will have little effect on this resolvable path because they originate or end at other lobes. Consequently, different physical paths approximately contribute to different angle-domain bins, and the channel coefficients in different angle-domain bins can be assumed to be approximately spatially uncorrelated. Furthermore, when some angle-domain bins contain few physical paths due to limited scattering, the corresponding channel coefficients should approach zero. Based on these two special properties for the angle-domain channel coefficients, we introduce three novel channel estimation techniques for MIMO-OFDM systems in this paper.

Note that in MIMO-OFDM systems, we classify the signals and channels in two domain types. The first type is represented by either array or angle domain, and the latter one is represented by either time or frequency domain. The channel correlation in the two domain types is referred to as the channel spatial correlation and the channel frequency correlation, respectively. Additionally, we use the term *channel correlation* for both the channel spatial and frequency correlation in this paper. Hereinafter, we will explicitly state which representation is used for each domain type. For example, the angle-time domain means that the angle and time representations are used for the above two domain types, respectively. When we state the representation for only one domain type, we mean that both representations for the other domain type are applicable unless indicated otherwise. For example, the angle domain refers to either the angle-time or angle-frequency domain.

We focus on angle-domain channel estimation techniques in this paper. More specifically, we consider the angle-frequency and angle-time domain techniques. In the angle-frequency domain, when some angle-frequency domain bins contain few physical paths due to limited scattering, the corresponding channel coefficients should approach zero. This allows us to choose a suitable threshold for ignoring the small-valued channel taps and retaining only the most significant taps (MST) to reduce the effect of noise on the estimates, thereby improving the performance of the channel estimation technique. We call this technique of retaining channel coefficients of sufficiently large power as the MST selection technique. In the angle-time

domain, we can also use the MST selection technique. Furthermore, the channel coefficients in the angle-time domain are approximately uncorrelated. Thus, we may use the channel power instead of the channel correlation to approximately perform the angle-time domain LMMSE technique when the SNR is known or reliably estimated. We call the resulting technique the approximated MMSE (AMMSE) technique. Note that we refer to the *channel power* as the channel average power in this paper. When the channel power is not available, we can use the *channel instantaneous power* (i.e., the instantaneous power of estimated channel coefficients) to estimate the channel power. In such cases, to maintain the estimated channel power positive and the estimation reliable, a threshold is required to ignore coefficients with low instantaneous power. Furthermore, we will not consider the AMMSE technique in the angle-frequency domain because the channel correlation in the corresponding frequency domain might be too high to be reasonably replaced by the channel power.

To our best knowledge, this paper is the first to systematically study angle-domain channel estimation techniques for MIMO-OFDM systems. We note that one angle-domain channel estimation technique has been investigated for single-input multiple-output OFDM systems [23]. In this technique, the angle-time domain bins that correspond to the identical angular lobes (but with different time indexes) are grouped into one angle-time domain *beam*. Examining all the estimated channel coefficients in one beam, only the beams that contain significant peak values along the time axis are identified as the signal beams and, thus, retained. All the remaining beams are considered as noise (or interference) beams and are ignored. This technique was shown to be effective when the channel power is concentrated in a few beams. However, when the channel power is distributed over all the beams (e.g., the channel model E in [24]), this technique may hardly improve the performance. To overcome this problem and investigate the techniques for MIMO-OFDM systems, we do not group the angle-time domain bins into beams. Instead, we independently filter the noise in each angle-time domain bin based on the fact that the multipath components are approximately disjoint in the angle-time domain. Then, as introduced, we use either the MST selection technique or the AMMSE technique to estimate channels in the angle-time domain. Furthermore, we also investigate the MST selection technique in the angle-frequency domain.

The three angle-domain channel estimation techniques proposed in this paper have two main advantages. First, the achieved performance gain over the conventional LS technique does not require prior knowledge of the channel correlation and even the channel power. Therefore, the techniques are applicable to various propagation environments. Second, they can use conventional array-domain estimators as the coarse estimators and perform postprocessing in the angle domain. Thus, they are well suited for chip design because integrating these estimators will not significantly change the existing system architecture. It should be noted that materialization as a postprocessor is not a requirement for the angle-domain techniques. They can be also directly implemented in the angle domain by first transforming the transmitted and received signals from the array domain into the angle domain. However, this may modify the existing system architecture and, thus, will not be a focus in this paper. Instead, we concentrate on describing the angle-domain techniques in the framework of a postprocessor.

The major contributions and results of this paper are as follows.

- 1) We systematically develop the channel estimation techniques in the angle-time and angle-frequency domains for MIMO-OFDM systems. We find that the proposed techniques perform particularly well in the angle-time domain.
- 2) We develop a unified approach to analyze the performance of the MST and AMMSE channel estimation techniques in terms of the mse. Based on this approach, we also develop a simple way to compare the performance of different angle-domain techniques with the help of the first derivative test [25].
- 3) The performance of all the angle-domain techniques is dependent on the respective thresholds. Nevertheless, we find that setting the threshold to be two times of noise variance is sufficient for the angle-time domain MST selection and AMMSE techniques to yield better performance than the conventional LS technique at all the SNRs for various IEEE 802.11n channel models [24].
- 4) Of all the proposed angle-domain techniques, our theoretical analysis and simulation results demonstrate that the angle-time domain AMMSE technique results in the best performance and achieves up to 8-dB performance gain when the mse is 10^{-2} compared to the conventional LS technique.

Throughout this paper, we make four assumptions. First, all the transmit and receive antennas have the same polarization and radiation patterns, and the geometry of the antenna array is the uniform linear array. Second, the spacing between the antennas at the transmitter or the receiver is much smaller than the distances between the scatterers and antenna arrays so that the paths from a given scatterer to all the transmit or receive antennas are approximated to be parallel. This is a typical assumption for the analysis of MIMO-OFDM systems [20]–[22]. Third, as our main concern is the indoor propagation environment, we assume the channel to be time invariant over a given training period. Fourth, the channel correlation and even the channel power are not available to the receiver.

This paper is organized as follows. Section II describes conventional MIMO-OFDM systems. In Section III, the angle-domain representation of MIMO-OFDM systems is presented. In Section IV, three angle-domain techniques are proposed to estimate the MIMO-OFDM channels, and their performance is analyzed in Section V. We evaluate the performance of different techniques in Section VI by simulating typical IEEE 802.11n channel models. Last, we conclude this paper in Section VII.

II. MIMO-OFDM SYSTEM MODEL

In this section, we first introduce the array-domain representation of MIMO-OFDM systems. Therefore, for the sake of convenience, we refer to the time and frequency domains as the array-time and array-frequency domains, respectively, in this section. Then, in the next section, we describe the angle-domain representation of MIMO-OFDM systems.

In a typical MIMO-OFDM system with N_t transmit and N_r receive antennas, the high rate symbols to be transmitted are first grouped into blocks of N_d data symbols at the transmitter. These groups are called the frequency-domain OFDM symbols, and the n th group at the (i_t) th transmitter is represented by the vector $\mathbf{x}_{i_t}(n) = [x_{i_t}(0, n), x_{i_t}(1, n), \dots, x_{i_t}(N_d - 1, n)]^T$, where i_t and n denote the indexes of the transmitter and the OFDM symbol, respectively. Next, an inverse discrete Fourier transform (IDFT) block is applied to each OFDM symbol at each transmitter. The IDFT block at the transmitter and the discrete Fourier transform (DFT) block at the receiver serve to modulate and demodulate the data on the orthogonal subcarriers, respectively. At the IDFT output (i.e., in the time domain), a cyclic prefix (CP) of length N_g , as a copy of the last part of the current OFDM symbol, is inserted at the beginning of each symbol to avoid ISI, and its length N_g is assumed to be not shorter than the channel length. The resulting n th time-domain OFDM symbol at the (i_t) th transmitter is represented by $\mathbf{s}_{i_t}(n) = [s_{i_t}(0, n), s_{i_t}(1, n), \dots, s_{i_t}(N_d + N_g - 1, n)]^T$, where

$$s_{i_t}(m, n) = \frac{1}{N_d} \sum_{l=0}^{N_d-1} x_{i_t}(l, n) e^{j2\pi l(m-N_g)/N_d} \quad (1)$$

for $m = 0, 1, \dots, N_d + N_g - 1$.

The samples $\{s_{i_t}(m, n)\}$ are sent through a frequency-selective fading channel, which can be represented by an equivalent discrete-time linear finite duration channel impulse response (CIR) given by a sequence of channel matrices $\mathbf{C}(l)$ for $l = 0, 1, \dots, N_h - 1$, where N_h is the time span of the MIMO channel, and $\mathbf{C}(l)$ is an $N_r \times N_t$ matrix whose (i_r, i_t) th element $c_{i_r, i_t}(l)$ represents the channel coefficient from the (i_t) th transmit antenna to the (i_r) th receive antenna at delay l .² Assuming that the transmitter and the receiver are perfectly synchronized, the received OFDM symbols become free of ISI when the CP is removed from each symbol at all the

²For the sake of simple description, $\mathbf{C}(l)$ represents the channel matrix at each integer time index. This representation is obvious for the sample-spaced channels. For the nonsample-spaced channels, $\mathbf{C}(l)$ can be obtained via the interpolation, as shown in [6].

N_r receivers. By performing an N_d -point DFT on the resulting symbols at each receiver, we obtain the corresponding frequency-domain symbols. The resulting frequency-domain data sample at the k th subcarrier and (i_r) th receiver in the n th OFDM symbol is given by

$$y_{i_r}(k, n) = \sum_{i_t=0}^{N_t-1} h_{i_r, i_t}(k) x_{i_t}(k, n) + \vartheta_{i_r}(k, n) \quad (2)$$

where $h_{i_r, i_t}(k)$ is the channel transfer function (CTF) at the k th subcarrier from the (i_t) th transmitter to the (i_r) th receiver, and the noise $\vartheta_{i_r}(k, n)$ is assumed to be additive white Gaussian with variance σ_f^2 . In a compact notation, the received block of the n th OFDM symbols at the k th subcarrier can be written as

$$\mathbf{y}(k, n) = \mathbf{H}(k) \mathbf{x}(k, n) + \boldsymbol{\vartheta}(k, n) \quad (3)$$

where $\mathbf{y}(k, n) = [y_0(k, n), y_1(k, n), \dots, y_{N_r-1}(k, n)]^T$, $\mathbf{x}(k, n) = [x_0(k, n), x_1(k, n), \dots, x_{N_t-1}(k, n)]^T$, $\boldsymbol{\vartheta}(k, n) = [\vartheta_0(k, n), \vartheta_1(k, n), \dots, \vartheta_{N_r-1}(k, n)]^T$, and

$$\mathbf{H}(k) = \begin{bmatrix} h_{0,0}(k) & h_{0,1}(k) & \cdots & h_{0,N_t-1}(k) \\ h_{1,0}(k) & h_{1,1}(k) & \cdots & h_{1,N_t-1}(k) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r-1,0}(k) & h_{N_r-1,1}(k) & \cdots & h_{N_r-1,N_t-1}(k) \end{bmatrix}$$

is the CTF matrix at the k th subcarrier. Furthermore, the received n th OFDM symbol in the time domain is given by

$$\mathbf{z}(m, n) = \sum_{l=0}^{N_h-1} \mathbf{C}(l) \mathbf{s}(m-l, n) + \mathbf{u}(m, n) \quad (4)$$

where $\mathbf{s}(m, n) = [s_0(m, n), s_1(m, n), \dots, s_{N_t-1}(m, n)]^T$, $\mathbf{z}(m, n) = [z_0(m, n), z_1(m, n), \dots, z_{N_r-1}(m, n)]^T$, and $\mathbf{u}(m, n)$ are the transmitted, received, and noise vectors, respectively, at the m th sample of the n th OFDM symbol. If the columns of $\mathbf{C}(l)$ and $\mathbf{H}(k)$ are stacked into the vectors $\mathbf{c}(l)$ and $\mathbf{h}(k)$, respectively, the time- and frequency-domain variables are related by

$$\underline{\mathbf{y}}(n) = (\mathbf{F} \otimes \mathbf{I}_{N_r}) \underline{\mathbf{z}}(n) \quad (5)$$

$$\underline{\mathbf{x}}(n) = (\mathbf{F} \otimes \mathbf{I}_{N_t}) \underline{\mathbf{s}}(n) \quad (6)$$

$$\underline{\mathbf{h}} = (\mathbf{F} \otimes \mathbf{I}_{N_r \times N_t}) \underline{\mathbf{c}} \quad (7)$$

where $\underline{\mathbf{y}}(n) = [\mathbf{y}^T(0, n), \mathbf{y}^T(1, n), \dots, \mathbf{y}^T(N_d-1, n)]^T$, $\underline{\mathbf{z}}(n) = [\mathbf{z}^T(0, n), \mathbf{z}^T(1, n), \dots, \mathbf{z}^T(N_d-1, n)]^T$, $\underline{\mathbf{x}}(n) = [\mathbf{x}^T(0, n), \mathbf{x}^T(1, n), \dots, \mathbf{x}^T(N_d-1, n)]^T$, $\underline{\mathbf{s}}(n) = [\mathbf{s}^T(0, n), \mathbf{s}^T(1, n), \dots, \mathbf{s}^T(N_d-1, n)]^T$, $\underline{\mathbf{h}} = [\mathbf{h}^T(0), \mathbf{h}^T(1), \dots, \mathbf{h}^T(N_d-1)]^T$, $\underline{\mathbf{c}} = [\mathbf{c}^T(0), \mathbf{c}^T(1), \dots, \mathbf{c}^T(N_h-1), \mathbf{0}_{1 \times [(N_d-N_h)N_t N_r]}]^T$, $\otimes$ denotes the Kronecker product, and $\mathbf{F}$, $\mathbf{I}_N$, and $\mathbf{0}_{N_1 \times N_2}$ are the $N_d \times N_d$ unitary Fourier matrix, $N \times N$ identity matrix, and $N_1 \times N_2$ zero matrix, respectively.

III. ANGLE-DOMAIN MIMO-OFDM SYSTEMS

In this section, we represent MIMO-OFDM systems in the angle domain. This is an extension of the work for MIMO flat-fading systems, as shown in [20].

As the channel models for MIMO-OFDM systems are commonly introduced in the array-time domain [24], [26], we start to represent the corresponding angle-time domain MIMO-OFDM systems from (6). Suppose that there is an arbitrary number of physical paths (four paths are illustrated in Fig. 1) between the transmit and receive antennas at time l ; the i th path has an attenuation of a_i with an angle ϕ_i^t ($\Omega_i^t := \sin \phi_i^t$) and ϕ_i^r ($\Omega_i^r := \sin \phi_i^r$) for the transmit and receive antennas, respectively. Then, $\mathbf{C}(l)$ is given by

$$\mathbf{C}(l) = \sum_i a_i^b \mathbf{e}_r(\Omega_i^r) \mathbf{e}_t^H(\Omega_i^t) \quad (8)$$

where³

$$a_i^b := a_i \sqrt{N_t N_r} \exp\left(\frac{j2\pi d_i}{\lambda_c}\right) \quad (9)$$

$$\mathbf{e}_r(\Omega_i^r) := \frac{1}{\sqrt{N_r}} \begin{bmatrix} 1 \\ \exp[j(2\pi\Delta_r\Omega_i^r)] \\ \vdots \\ \exp[j(N_r-1)(2\pi\Delta_r\Omega_i^r)] \end{bmatrix} \quad (10)$$

$$\mathbf{e}_t(\Omega_i^t) := \frac{1}{\sqrt{N_t}} \begin{bmatrix} 1 \\ \exp[j(2\pi\Delta_t\Omega_i^t)] \\ \vdots \\ \exp[j(N_t-1)(2\pi\Delta_t\Omega_i^t)] \end{bmatrix} \quad (11)$$

where the superscript H denotes the Hermitian transpose; $\{N_t, \Delta_t, \mathbf{e}_t(\Omega_i^t)\}$ and $\{N_r, \Delta_r, \mathbf{e}_r(\Omega_i^r)\}$ are the number of antennas, the separation between adjacent antennas normalized by λ_c , and the array response vectors,⁴ respectively, for the transmit and receive antennas, respectively; λ_c is the carrier wavelength; and d_i is the distance between the last transmit and receive antennas along path i .

As from [20], the orthonormal bases for the angle-time transmitted and received signals are given by

$$\boldsymbol{\xi}_t := \left\{ \mathbf{e}_t(0), \mathbf{e}_t\left(\frac{1}{L_t}\right), \dots, \mathbf{e}_t\left(\frac{N_t-1}{L_t}\right) \right\} \quad (12)$$

$$\boldsymbol{\xi}_r := \left\{ \mathbf{e}_r(0), \mathbf{e}_r\left(\frac{1}{L_r}\right), \dots, \mathbf{e}_r\left(\frac{N_r-1}{L_r}\right) \right\} \quad (13)$$

respectively, where $L_t = N_t \Delta_t$ and $L_r = N_r \Delta_r$ are the normalized antenna array lengths of the transmitter and the receiver, respectively. Let $\mathbf{U}_t$ and $\mathbf{U}_r$ be the unitary matrices whose columns are the basis vectors in (12) and (13), respectively. Then, we can transform the m th samples of the n th

³For notational convenience, we ignore the time index l in these three variables. We also assume that the fractional bandwidth is small such that λ_c , $\mathbf{e}_r(\Omega_i^r)$, and $\mathbf{e}_t(\Omega_i^t)$ are approximated to be unchanged over the whole signal bandwidth.

⁴In the literature, the array response vector has also been alternatively called the array vector, the array steering vector, the array propagation vector, and the array manifold vector.

transmitted and received OFDM symbol from the array-time domain into the angle-time domain by

$$\mathbf{s}^a(m, n) := \mathbf{U}_t^H \mathbf{s}(m, n) \quad (14)$$

$$\mathbf{z}^a(m, n) := \mathbf{U}_r^H \mathbf{z}(m, n) \quad (15)$$

respectively, where the superscript “a” denotes the angle-domain variables. From (4), we obtain the angle-time domain MIMO-OFDM system equation as

$$\mathbf{z}^a(m, n) = \sum_{l=0}^{N_h-1} \mathbf{C}^a(l) \mathbf{s}^a(m-l, n) + \mathbf{u}^a(m, n) \quad (16)$$

where

$$\begin{aligned} \mathbf{C}^a(l) &:= \begin{bmatrix} c_{0,0}^a(l) & c_{0,1}^a(l) & \cdots & c_{0,N_t-1}^a(l) \\ c_{1,0}^a(l) & c_{1,1}^a(l) & \cdots & c_{1,N_t-1}^a(l) \\ \vdots & \vdots & \ddots & \vdots \\ c_{N_r-1,0}^a(l) & c_{N_r-1,1}^a(l) & \cdots & c_{N_r-1,N_t-1}^a(l) \end{bmatrix} \\ &= \mathbf{U}_r^H \mathbf{C}(l) \mathbf{U}_t \end{aligned} \quad (17)$$

is the angle-time domain channel matrix, and

$$\mathbf{u}^a(m, n) = \mathbf{U}_r^H \mathbf{u}(m, n) \quad (18)$$

is the angle-time domain noise vector that still satisfies an independent identically distributed multivariate complex normal distribution. As explained in [20], due to the finite number of antennas, multiple unresolvable physical paths can be appropriately aggregated into one resolvable path with gain $c_{k_r, k_t}^a(l)$. This gain uniquely corresponds to the channel coefficient for the (k_r, k_t) angle-domain bin at delay l . Hence, different physical paths (approximately) contribute to different elements of $\mathbf{C}^a(l)$. This means that the angle-time domain channel matrix $\mathbf{C}^a(l)$ lends itself to a physical interpretation. For example, in Fig. 1, paths 1 and 2 contribute to the element $c_{0,0}^a(l)$. Paths 3 and 4 contribute to the elements $c_{3,3}^a(l)$ and $c_{1,2}^a(l)$, respectively. All the other elements of $\mathbf{C}^a(l)$ approach zero because no paths contribute to these elements.

By analogy to the angle-time domain case, from (3), we obtain the angle-frequency domain MIMO-OFDM system equation at the k th subcarrier as

$$\mathbf{y}^a(k, n) = \mathbf{H}^a(k) \mathbf{x}^a(k, n) + \mathbf{v}^a(k, n) \quad (19)$$

where

$$\mathbf{x}^a(k, n) := \mathbf{U}_t^H \mathbf{x}(k, n) \quad (20)$$

$$\mathbf{y}^a(k, n) := \mathbf{U}_r^H \mathbf{y}(k, n) \quad (21)$$

$$\mathbf{v}^a(k, n) := \mathbf{U}_r^H \mathbf{v}(k, n) \quad (22)$$

$$\mathbf{H}^a(k) := \mathbf{U}_r^H \mathbf{H}(k) \mathbf{U}_t. \quad (23)$$

IV. ANGLE-DOMAIN CHANNEL ESTIMATION

As the physical path within one angle-time domain beam has most of its energy within this beam, the elements within

one angle-domain channel matrix $\mathbf{C}^a(l)$ or $\mathbf{H}^a(k)$ maintain low spatial correlation. Furthermore, various elements in the angle-domain channel matrices tend to approach zero due to limited scattering. Therefore, we may independently perform noise filtering in each angle-time or angle-frequency domain bin to improve the estimation performance. The angle-domain channel can be directly estimated in the angle domain [27] by first transforming the transmitted and received signal from the array-frequency domain into the angle-frequency domain as shown in (20) and (21). However, this may modify the existing system architecture and, thus, will not be a focus in this paper. Instead, without affecting the estimation performance, we concentrate on describing the angle-domain techniques in the framework of a postprocessor. This follows three steps: 1) performing the coarse channel estimation in the array domain; 2) in the postprocessor, transforming the estimated channel from the array domain into the angle-time or angle-frequency domain where the noise filtering process in each angle-domain bin is independently performed; and 3) transforming back the filtered estimated channel into the array domain.

In the literature, there exist two types of frequency-domain pilot arrangements in OFDM systems for the implementation of pilot-aided channel estimation techniques—the block-type and comb-type pilot arrangements. The first type is realized by inserting pilots into all of the subcarriers within a periodical time interval, and the estimated channel coefficients that are obtained from one period are used for further signal processing (e.g., equalization, detection, etc.) until the next period of pilots is received. This type of pilot arrangement is particularly suitable for slowly fading channels. The second type is performed by inserting pilots into a certain number of subcarriers of each OFDM symbol and using the estimated channel coefficients at these pilot subcarriers to interpolate the channel coefficients at other subcarriers. Compared with the first type, this type is more suitable for rapidly time-varying channels in that it can track time variation within each OFDM symbol. However, it may suffer an irreducible estimation error floor due to the interpolation.

In typical MIMO-OFDM systems, only one transmit antenna sends pilot symbols in a given time or frequency position [3], [10], [28]. This technique can lead to very simple MIMO-OFDM channel estimation because the received signals at a given time or frequency position correspond to one unique channel coefficient from a given transmit antenna to a given receive antenna. In such cases, the channel estimation for a MIMO-OFDM system with N_t transmit and N_r received antennas becomes the channel estimation for the total $N_t \times N_r$ single-input single-output (SISO) OFDM systems. Thus, the well-developed SISO-OFDM channel estimation techniques (e.g., [6], [7], [29], and [31]) are directly applicable to the estimation of MIMO-OFDM channels. In this paper, we use the conventional array-frequency domain LS technique [6] to coarsely estimate the array-frequency domain $h_{i_r, i_t}(k)$ because the knowledge of channel correlation is assumed to be not available. For the sake of a simple description, as shown in Fig. 2, we assume that pilots from different transmit antennas are time orthogonal to each other. Thus, only one transmit antenna is used to transmit pilots in each OFDM training

Fig. 2. Time-orthogonal pilot pattern for a MIMO-OFDM system with N_t transmit antennas.

symbol period as introduced in [3]. The proposed techniques can be easily extended to other pilot transmission schemes (e.g., [28]) but will not be covered in this paper.

A. Angle-Frequency Domain Technique

From (3) and (7) and using the assumption that the pilots from different transmit antennas are time orthogonal to each other, we obtain

$$\mathbf{Y} = \mathbf{X}\mathbf{h} + \mathbf{v} \quad (24)$$

where $\mathbf{Y} = [\mathbf{Y}^T(n), \mathbf{Y}^T(n+N_t), \dots, \mathbf{Y}^T(n+(N_c-1)N_t)]^T$ with $\mathbf{Y}(n) = [\mathbf{y}^T(0, n), \mathbf{y}^T(0, n+1), \dots, \mathbf{y}^T(0, n+N_t-1), \mathbf{y}^T(1, n), \dots, \mathbf{y}^T(1, n+N_t-1), \dots, \mathbf{y}^T(N_d-1, n+N_t-1)]^T$, $\mathbf{X} = [\mathbf{X}(n), \mathbf{X}(n+N_t), \dots, \mathbf{X}(n+(N_c-1)N_t)]^T$ with $\mathbf{X}(n) = \text{diag}[\mathbf{x}^T(0, n), \mathbf{x}^T(0, n+1), \dots, \mathbf{x}^T(0, n+N_t-1), \mathbf{x}^T(1, n), \dots, \mathbf{x}^T(1, n+N_t-1), \dots, \mathbf{x}^T(N_d-1, n+N_t-1)]$, and $\mathbf{v} = [\mathbf{v}^T(n), \mathbf{v}^T(n+N_t), \dots, \mathbf{v}^T(n+(N_c-1)N_t)]^T$ with $\mathbf{v}^T(n) = [\mathbf{v}^T(0, n), \mathbf{v}^T(0, n+1), \dots, \mathbf{v}^T(0, n+N_t-1), \mathbf{v}^T(1, n), \dots, \mathbf{v}^T(1, n+N_t-1), \dots, \mathbf{v}^T(N_d-1, n+N_t-1)]^T$ are the received signal vector, the transmitted signal matrix, and the noise vector, respectively, and N_c is the number of pilots used for each channel coefficient in the LS channel estimation. Note that $\mathbf{X}$ is a block diagonal matrix. Then, the array-frequency domain LS estimator is given by

$$\mathbf{h}_{\text{LS}} = (\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H \mathbf{Y}. \quad (25)$$

By rearranging the vector form $\mathbf{h}_{\text{LS}}$ into its matrix form, in this first step, we obtain the coarsely estimated array-frequency domain channel matrix $\tilde{\mathbf{H}}(k)$ for $k = 0, 1, \dots, N_d - 1$.

In the second step, we transform the estimated array-frequency domain channel matrix from the array-frequency domain $\tilde{\mathbf{H}}(k)$ into the angle-frequency domain $\tilde{\mathbf{H}}^a(k)$ by the use of (23) for all the subcarriers under consideration. Let $\tilde{h}_{i_r, i_t}^a(k)$ denote the (i_r, i_t) th element of $\tilde{\mathbf{H}}^a(k)$. Then, the filtered angle-frequency domain channel coefficient is given by comparing the power of $\tilde{h}_{i_r, i_t}^a(k)$ with a threshold η as follows:

$$\tilde{h}_{i_r, i_t, \text{MST}}^a(k) = \begin{cases} \tilde{h}_{i_r, i_t}^a(k), & \text{if } |\tilde{h}_{i_r, i_t}^a(k)|^2 \geq \eta \\ 0, & \text{otherwise.} \end{cases} \quad (26)$$

Last, the filtered angle-frequency domain channel matrices are transformed back to the array-frequency domain. We

call this technique the angle-frequency domain MST selection technique.

To make a fair complexity comparison among all the techniques, we assume that LS-estimated array-frequency domain channel coefficients are available beforehand. Then, from (23), obtaining the angle-frequency domain $\tilde{\mathbf{H}}^a(k)$ from the array-frequency domain $\tilde{\mathbf{H}}(k)$ requires $N_t + N_r$ complex multiplication operations for each channel coefficient. Furthermore, transforming the estimated angle-frequency domain into the array-frequency domain also requires $N_t + N_r$ complex multiplication operations for each channel coefficient. Therefore, the total required number of complex multiplication operations for each channel coefficient is $2(N_t + N_r)$.

B. Angle-Time Domain Techniques

In many cases such as the nonline-of-sight (NLOS) scenario, the mean angle of departure (AoD) and angle of arrival (AoA) of the clusters of multipath components tend to be uniformly distributed over all angles [26], [33]. Thus, when the number of clusters is relatively large (e.g., the channel model E in [24]), the above angle-frequency domain channel estimation technique may hardly improve over the conventional array-frequency domain LS technique because nearly all the elements of angle-frequency domain channel matrices may not approach zero. As the clusters of multipath components are disjoint in the angle-time domain, we may perform the noise filtering in this domain instead of the angle-frequency domain.

For the implementation of angle-time domain techniques, we first transform the estimated channel coefficients from the array-frequency domain into the array-time domain by the use of DFT [31]. Then, we transform the channel matrices into the angle-time domain by using (17). In the angle-time domain, we can select the MST in channel matrices to reduce the effect of noise on the estimates. Now, the estimated angle-time domain channel coefficient becomes

$$\tilde{c}_{i_r, i_t, \text{MST}}^a(l) = \begin{cases} \tilde{c}_{i_r, i_t}^a(l), & \text{if } |\tilde{c}_{i_r, i_t}^a(l)|^2 \geq \eta \\ 0, & \text{otherwise} \end{cases} \quad (27)$$

where $\tilde{c}_{i_r, i_t}^a(l)$ is the coarsely estimated angle-time domain channel coefficient.

Note that we presume that the channel spatial correlation is not available to the receiver in this paper. Therefore, the conventional LMMSE technique that utilizes the channel spatial correlation is not applicable here. However, as discussed, the channel coefficients in the angle domain at a given time are approximately spatially uncorrelated. Therefore, we may use the channel instantaneous power to approximate the channel correlation (here, the channel power is also assumed to be not available). As the approximated channel correlation matrix is a diagonal matrix, the LMMSE technique that jointly filters all the channel coefficients becomes the independent spatial filtering for each channel coefficient. Furthermore, the channel coefficient is uncorrelated to the noise. Therefore, we may estimate the channel instantaneous power for each coefficient as $(|\tilde{c}^a(l)|^2 - \sigma_f^2)$. The noise variance σ_f^2 is assumed to be known here. In practice, it can be estimated during periods when no

transmitted signal is detected or at virtual carriers where no data are transmitted. Then, the angle–time domain AMMSE technique is realized as

$$\tilde{c}_{i_r, i_t}^a(l) = \begin{cases} \frac{|\tilde{c}_{i_r, i_t}^a(l)|^2 - \sigma_f^2}{|\tilde{c}_{i_r, i_t}^a(l)|^2} \tilde{c}_{i_r, i_t}^a(l), & \text{if } |\tilde{c}_{i_r, i_t}^a(l)|^2 \geq \eta \\ 0, & \text{otherwise.} \end{cases} \quad (28)$$

Here, the threshold η is usually chosen to be smaller than σ_f^2 . Otherwise, the approximated channel power ($|\tilde{c}_{i_r, i_t}^a(l)|^2 - \sigma_f^2$) becomes negative. In comparison with the MST selection technique, the difference is the use of a dynamic multiplication factor instead of a constant one. The factor turns out to be crucial in improving the performance of channel estimation, as shown in the following two sections. However, the complexity is only increased by only an additional complex multiplication for each channel coefficient.

After the noise filtering in the angle–time domain either by (27) or (28), we transform back the estimated channel into the array–time domain and then into the array–frequency domain.

Similar to the angle–frequency domain MST technique, we assume that LS-estimated array–frequency domain channel coefficients are available beforehand to make a fair complexity comparison. Then, from (17), we know that the transformations between the array–time domain and the angle–time domain require a total of $2(N_t + N_r)$ complex multiplication operations for each channel coefficient. In addition, the angle–time domain techniques require the transformations between the array–frequency domain and the array–time domain. This requires a total of $2N_d$ complex multiplication operations for each channel coefficient. Typically, N_d is a power of 2. Then, using the FFT and the IFFT [35] for transformations between the array–time and array–frequency domains, the total complex multiplication operations required for each channel coefficient is reduced to $\log_2 N_d$ complex multiplication operations for each channel coefficient. From the above, the total required complex multiplication operations for each channel coefficient in the angle–time domain MST selection technique is $2(N_t + N_r) + \log_2 N_d$. For the angle–time domain AMMSE technique, an additional complex multiplication is needed. Therefore, the total required complex multiplication operations for each channel coefficient in the angle–time domain AMMSE technique is $2(N_t + N_r) + 1 + \log_2 N_d$. Compared with the angle–frequency domain MST technique, these angle–time domain techniques always have higher complexity. Nevertheless, our results show that these angle–time domain techniques always outperform the angle–frequency domain MST technique for all the channel models under consideration.

Note that the angle–time domain MST selection and AMMSE techniques are well suited for the sample-spaced channels. The performance will be degraded for the nonsample-spaced channels due to the power leakage [6]. Nevertheless, our results show that the angle–time domain techniques still outperform the array–frequency domain LS technique for all the channel models under consideration.

V. PERFORMANCE ANALYSIS OF ANGLE-DOMAIN CHANNEL ESTIMATION TECHNIQUES

For channel estimation techniques, one of the most important performance measures is the mse, which measures the average mean squared deviation of the estimator from the true value [30]. In this section, we present a unified approach to computing the mse of the angle-domain (i.e., either angle–time or angle–frequency domain) MST selection and AMMSE techniques. Note that the array–time and array–frequency domains are related by a unitary transformation. Thus, of a given estimation technique, the mse represented in either the array–time or array–frequency domain yields the same result and is given by

$$\text{mse} = \frac{1}{N_d N_t N_r} \mathbb{E} [\|\tilde{\mathbf{h}} - \hat{\mathbf{h}}\|^2] \quad (29)$$

where $\tilde{\mathbf{h}}$ is either $\mathbf{h}$ or $\mathbf{c}$, and $\hat{\mathbf{h}}$ is the estimated $\tilde{\mathbf{h}}$.

Here, we use the array–frequency domain LS technique to perform the coarse channel estimation. Then, the resulting mse of the coarse estimation is given by

$$\begin{aligned} \text{mse}_{\text{LS}} &= \frac{1}{N_d N_t N_r} \mathbb{E} [\|\mathbf{h}_{\text{LS}} - \mathbf{h}\|^2] \\ &= \frac{\sigma_f^2}{N_d N_t N_r} \text{trace} \{ \mathbb{E} [(\mathbf{X}^H \mathbf{X})^{-1}] \}. \end{aligned} \quad (30)$$

In many cases such as in the IEEE 802.11a standard [32], the powers of all pilots are unity. Therefore, we also assume that $\mathbb{E}[(\mathbf{X}^H \mathbf{X})^{-1}]$ is the identity matrix. Then, since $\mathbf{X}$ is diagonal, (30) becomes

$$\text{mse}_{\text{LS}} = \sigma_f^2. \quad (31)$$

Let $\mathbf{h}^a$ represent the stacked angle-domain channel vector. Then, from (23), we obtain

$$\mathbf{h}^a = \mathbf{B} \tilde{\mathbf{h}} \quad (32)$$

where

$$\mathbf{B} = \mathbf{I}_{N_d} \otimes [(\mathbf{I}_{N_t} \otimes \mathbf{U}_r^H) (\mathbf{U}_t^T \otimes \mathbf{I}_{N_r})] \quad (33)$$

is an $(N_d N_t N_r \times N_d N_t N_r)$ matrix, and the superscript T denotes the transpose. It is easily verified that the matrix $\mathbf{B}$ is unitary, i.e.,

$$\mathbf{B}^H \mathbf{B} = \mathbf{I}_{N_r \times N_t}. \quad (34)$$

Then, the angle-domain MST selection or AMMSE technique is given by

$$\tilde{\mathbf{h}}^a = \mathbf{M} \mathbf{h}_{\text{LS}}^a \quad (35)$$

where $\mathbf{M}$ is a diagonal $(N_d N_t N_r \times N_d N_t N_r)$ matrix that represents either the MST selection or AMMSE process, and $\mathbf{h}_{\text{LS}}^a$ is the stacked LS-estimated channel vector in the angle domain. Note that the i th diagonal element of $\mathbf{M}$ (denoted as m_i) is dependent on the channel coefficient and the noise.

Denote $\hat{r}_i$ as the instantaneous power of the i th element of $\underline{\mathbf{h}}_{\text{LS}}^{\text{a}}$. In the MST selection techniques, we have

$$m_i = \begin{cases} 1, & \text{if } \hat{r}_i \geq \eta \\ 0, & \text{otherwise.} \end{cases} \quad (36)$$

In the AMMSE technique, from (28), we have

$$m_i = \begin{cases} \frac{\hat{r}_i - \sigma_f^2}{\hat{r}_i}, & \text{if } \hat{r}_i \geq \eta \\ 0, & \text{otherwise} \end{cases} \quad (37)$$

where the threshold η is chosen not to be smaller than σ_f^2 .

From (29), the mse of the angle-domain techniques is given by

$$\text{mse} = \frac{1}{N_d N_t N_r} \mathbb{E} \left[\|\mathbf{B}^H \mathbf{M} \mathbf{B} \underline{\mathbf{h}}_{\text{LS}} - \underline{\mathbf{h}}\|^2 \right] \quad (38)$$

where $\underline{\mathbf{h}}_{\text{LS}}$ is the LS-estimated $\underline{\mathbf{h}}$. Since $\mathbf{M}$ is diagonal and real, using (34), we obtain

$$\text{mse} = \frac{1}{N_d N_t N_r} \text{trace} \left\{ \mathbb{E} \left[\mathbf{M} \underline{\mathbf{h}}^{\text{a}} (\underline{\mathbf{h}}^{\text{a}})^H \mathbf{M}^H + \underline{\mathbf{h}}^{\text{a}} (\underline{\mathbf{h}}^{\text{a}})^H - 2\mathbf{M} \underline{\mathbf{h}}^{\text{a}} (\underline{\mathbf{h}}^{\text{a}})^H + \mathbf{M} \underline{\mathbf{v}}^{\text{a}} (\underline{\mathbf{v}}^{\text{a}})^H \mathbf{M}^H \right] \right\} \quad (39)$$

where $\underline{\mathbf{v}}^{\text{a}}$ is the angle-domain noise vector. Denoting $\hat{h}_i$ and $\hat{v}_i$ as the instantaneous power of the i th element of $\underline{\mathbf{h}}^{\text{a}}$ and $\underline{\mathbf{v}}^{\text{a}}$, respectively, we rewrite (39) as

$$\text{mse} = \frac{1}{N_d N_t N_r} \sum_{i=1}^{N_d N_t N_r} \mathbb{E} \left[(m_i - 1)^2 \hat{h}_i + m_i^2 \hat{v}_i \right]. \quad (40)$$

More specifically

$$\mathbb{E} \left[(m_i - 1)^2 \hat{h}_i \right] = \int_0^\infty \hat{h}_i P_h(\hat{h}_i) \int_\eta^\infty (m_i - 1)^2 P_{r|h}(\hat{r}_i|\hat{h}_i) d\hat{r}_i d\hat{h}_i + \int_0^\eta \hat{h}_i P_h(\hat{h}_i) \int_0^\eta P_{r|h}(\hat{r}_i|\hat{h}_i) d\hat{r}_i d\hat{h}_i \quad (41)$$

is the mse part due to the filtered instantaneous power of the i th channel coefficient, and

$$\mathbb{E} \left[m_i^2 \hat{v}_i \right] = \int_0^\infty \hat{v}_i P_v(\hat{v}_i) \int_\eta^\infty m_i^2 P_{r|v}(\hat{r}_i|\hat{v}_i) d\hat{r}_i d\hat{v}_i \quad (42)$$

is the mse part due to the unfiltered noise components in the i th estimated channel coefficient, where $P_h(\hat{h}_i)$, $P_v(\hat{v}_i)$, $P_{r|h}(\hat{r}_i|\hat{h}_i)$, and $P_{r|v}(\hat{r}_i|\hat{v}_i)$ are the probability density functions of $\hat{h}_i$ and $\hat{v}_i$, and the conditional probability density functions of $\hat{r}_i$ conditioned on $\hat{h}_i$ and $\hat{v}_i$, respectively. The channel

and the noise have complex normal distributions. Thus, $P_h(\hat{h}_i)$ and $P_v(\hat{v}_i)$ have exponential distributions, and $P_{r|h}(\hat{r}_i|\hat{h}_i)$ and $P_{r|v}(\hat{r}_i|\hat{v}_i)$ have noncentral chi-square distributions [25].

Let σ_i^2 denote the channel power of the i th element of $\underline{\mathbf{h}}^{\text{a}}$. When σ_i^2 is not equal to zero, (41) and (42) become

$$\begin{aligned} \mathbb{E} \left[(m_i - 1)^2 \hat{h}_i \right] &= \int_0^\infty \frac{\hat{h}_i}{\sigma_i^2} e^{-\frac{\hat{h}_i}{\sigma_i^2}} \int_\eta^\infty (m_i - 1)^2 \frac{e^{-\frac{\hat{r}_i + \hat{h}_i}{\sigma_f^2}}}{\sigma_f^2} \\ &\quad \times J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{h}_i}}{\sigma_f^2} \right) d\hat{r}_i d\hat{h}_i + \int_0^\eta \frac{\hat{h}_i}{\sigma_i^2} e^{-\frac{\hat{h}_i}{\sigma_i^2}} \\ &\quad \times \int_0^\eta \frac{e^{-\frac{\hat{r}_i + \hat{h}_i}{\sigma_f^2}}}{\sigma_f^2} J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{h}_i}}{\sigma_f^2} \right) d\hat{r}_i d\hat{h}_i \end{aligned} \quad (43)$$

$$\begin{aligned} \mathbb{E} \left[m_i^2 \hat{v}_i \right] &= \int_0^\infty \frac{\hat{v}_i}{\sigma_f^2} e^{-\frac{\hat{v}_i}{\sigma_f^2}} \int_\eta^\infty m_i^2 \frac{e^{-\frac{\hat{r}_i + \hat{v}_i}{\sigma_i^2}}}{\sigma_i^2} \\ &\quad \times J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{v}_i}}{\sigma_i^2} \right) d\hat{r}_i d\hat{v}_i \end{aligned} \quad (44)$$

respectively, where $J_0(z)$ is the modified Bessel function of the first kind.

When σ_i^2 is equal to zero, $\hat{h}_i = 0$, and $\hat{r}_i = \hat{v}_i$, then $\mathbb{E}[(m_i - 1)^2 \hat{h}_i] = 0$, and the mse of the i th estimated channel coefficient is simplified as

$$\text{mse}_i = \mathbb{E} \left[m_i^2 \hat{v}_i \right] = \int_\eta^\infty m_i^2 \frac{\hat{v}_i}{\sigma_f^2} e^{-\frac{\hat{v}_i}{\sigma_f^2}} d\hat{v}_i. \quad (45)$$

A. Performance of the MST Selection Techniques

In this subsection, we analyze the performance of the MST selection techniques and show that the optimal strategy to minimize the mse for the MST selection techniques is to ignore all the array-frequency domain LS-estimated channel coefficients, whose corresponding channel powers are smaller than the noise variance, and retain all the remaining estimated channel coefficients.

1) *When σ_i^2 Is Not Equal to Zero:* From (43) and (44), the mse of the i th estimated channel coefficient is given by

$$\begin{aligned} \text{mse}_i &= \mathbb{E} \left[(m_i - 1)^2 \hat{h}_i \right] + \mathbb{E} \left[m_i^2 \hat{v}_i \right] \\ &= \sigma_f^2 + \int_0^\infty \frac{\hat{h}_i}{\sigma_i^2} e^{-\frac{\hat{h}_i}{\sigma_i^2}} \int_0^\eta \frac{e^{-\frac{\hat{r}_i + \hat{h}_i}{\sigma_f^2}}}{\sigma_f^2} \\ &\quad \times J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{h}_i}}{\sigma_f^2} \right) d\hat{r}_i d\hat{h}_i - \int_0^\eta \frac{\hat{v}_i}{\sigma_f^2} e^{-\frac{\hat{v}_i}{\sigma_f^2}} \\ &\quad \times \int_0^\eta \frac{e^{-\frac{\hat{r}_i + \hat{v}_i}{\sigma_i^2}}}{\sigma_i^2} J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{v}_i}}{\sigma_i^2} \right) d\hat{r}_i d\hat{v}_i. \end{aligned} \quad (46)$$

Differentiating mse_i with respect to η , we obtain

$$\frac{\partial \text{mse}_i}{\partial \eta} = \frac{(\sigma_i^2 - \sigma_f^2) e^{-\frac{\eta}{\sigma_i^2 + \sigma_f^2}}}{(\sigma_i^2 + \sigma_f^2)^2}. \quad (47)$$

From (47), we find that although mse_i is difficult to analytically calculate, its gradient is surprisingly simple.

- When $\sigma_i^2 = \sigma_f^2$, the last two terms in (46) become identical. This is reasonable since setting the estimated channel coefficient as zero or retaining the original coarse estimate yields the same mse. Therefore, mse_i will always be σ_f^2 no matter what threshold is chosen.
- When $\sigma_i^2 > \sigma_f^2$, from (47) and the first derivative test [25], we conclude that mse_i reaches its extremum when $\eta = 0$ or $\eta = \infty$.⁵ Since (47) is always positive, mse_i is monotonically increasing and has its minimum σ_f^2 at $\eta = 0$. This result indicates that the LS-estimated channel coefficient should be retained when the corresponding channel power is larger than the noise variance. It is reasonable since when we ignore the estimated channel coefficient, the resulting performance loss due to the ignorance of instantaneous power will be larger than the performance gain due to the removal of the noise.
- When $\sigma_i^2 < \sigma_f^2$, (47) is always negative. Thus, mse_i reaches its minimum σ_i^2 at $\eta = \infty$. This result indicates that the LS-estimated channel coefficient is required to be ignored when the corresponding channel power is smaller than the noise variance. It is reasonable as the performance gain surpasses the performance loss when the corresponding estimated coefficient is ignored.

2) When σ_i^2 Is Equal to Zero: In this case, (45) becomes

$$\text{mse}_i = \int_{\eta}^{\infty} \frac{\hat{v}_i}{\sigma_f^2} e^{-\frac{\hat{v}_i}{\sigma_f^2}} d\hat{v}_i = e^{-\frac{\eta}{\sigma_f^2}} (\eta + \sigma_f^2). \quad (48)$$

Differentiating mse_i with respect to η , we obtain

$$\frac{\partial \text{mse}_i}{\partial \eta} = -\frac{\eta e^{-\frac{\eta}{\sigma_f^2}}}{\sigma_f^2} \leq 0. \quad (49)$$

Similar to the discussion above, we find that mse_i reaches its minimum at $\eta = \infty$. This result indicates that the LS-estimated channel coefficient is required to be ignored when the corresponding channel power is smaller than the noise variance. It is reasonable since the corresponding estimated channel coefficient only contains the noise, and ignoring this estimated coefficient will result in no performance loss.

From the discussion above, we conclude that the choice of threshold η is dependent on σ_i^2 and σ_f^2 , as shown in Table I, because we need to balance the performance gain and loss when we ignore the estimated channel coefficients. Thus, the optimum threshold should be obtained for each channel

TABLE I
MAXIMUM AND MINIMUM MSE_i AND THE CORRESPONDING THRESHOLDS FOR THE MST SELECTION TECHNIQUES

	$\sigma_i^2 > \sigma_f^2$	$\sigma_i^2 < \sigma_f^2$	$\sigma_i^2 = \sigma_f^2$
Maximum MSE_i	σ_i^2	σ_f^2	σ_f^2
Minimum MSE_i	σ_f^2	σ_i^2	σ_f^2
Threshold for Minimum MSE_i	0	∞	arbitrary

coefficient. Since the channel is independent of noise, the average power of the LS-estimated channel coefficient is the sum of the channel power and the noise variance (i.e., $\sigma_i^2 + \sigma_f^2$). For this reason, we may set the threshold η to be $2\sigma_f^2$. When the average power of the i th LS-estimated channel coefficient exceeds this threshold $2\sigma_f^2$, it means that $\sigma_i^2 > \sigma_f^2$. Therefore, we set m_i to be 1 [see (36)] to retain the i th LS-estimated channel coefficient. This is optimum to minimize the mse_i , as we discussed above. Similarly, this threshold is optimum to the minimization of mse_i for the case when $\sigma_i^2 < \sigma_f^2$. Therefore, when the average power of the LS-estimated channel coefficient is available, we can still obtain the corresponding optimal threshold.

As σ_i^2 is assumed to be not available in this paper, the average power of the corresponding LS-estimated channel coefficient may also not be available. Then, we may use the channel instantaneous power (i.e., the instantaneous power of the estimated channel coefficient) to approximate the average power of the LS-estimated channel coefficient. Due to the monotone property of mse_i with the increase in η , Table I implies that for the given threshold $\eta = 2\sigma_f^2$, mse_i is always smaller than σ_f^2 when $\sigma_i^2 < \sigma_f^2$, and larger than σ_f^2 when $\sigma_i^2 > \sigma_f^2$. Therefore, the overall performance of the MST selection techniques is dependent on the portion⁶ of the number of channel coefficients, whose average powers are below σ_f^2 , to the total number of channel coefficients. When a majority of channel coefficients have average powers that are below σ_f^2 (such as in the angle–time domain), the MST selection technique can improve over the array–frequency domain LS technique. This is also verified in our simulation results presented in Section VI.

B. Performance of the AMMSE Technique

As m_i is dependent on the channel instantaneous power, directly analyzing the performance of the AMMSE technique results in high computational complexity. Therefore, in this subsection, we only compare the AMMSE technique with the MST selection technique to provide some general understanding on the performance trends.

1) When σ_i^2 Is Not Equal to Zero: As shown in the Appendix, the AMMSE technique performs better than the MST selection technique when σ_i^2 is not equal to zero.

⁵For $\sigma_i^2 = \sigma_f^2$, $\partial \text{mse}_i / \partial \eta$ is always zero, hence, indicating that mse_i is constant, as claimed earlier.

⁶ $\eta = 2\sigma_f^2$ may not be the optimum choice but is reasonable when no prior information of σ_i^2 is available to the receiver. When additional information (e.g., the portion) is available, we may moderately adjust the threshold to improve the estimation performance, as shown in Section VI.

Fig. 3. Relation of the clustered model and the angle-domain representation.

2) *When σ_i^2 Is Equal to Zero:* In this case, (45) becomes

$$\text{mse}_i = \int_{\eta} m_i^2 \frac{\hat{v}_i}{\sigma_f^2} e^{-\frac{\hat{v}_i}{\sigma_f^2}} d\hat{v}_i < \int_{\eta} \frac{\hat{v}_i}{\sigma_f^2} e^{-\frac{\hat{v}_i}{\sigma_f^2}} d\hat{v}_i \quad (50)$$

because m_i is always smaller than one. This implies that the AMMSE technique performs better than the MST selection technique when σ_i^2 is equal to zero.

Therefore, we conclude that the AMMSE technique always performs better than the MST selection technique. Since the angle–time domain MST selection technique yields better performance than the array–frequency domain LS technique when the threshold is $2\sigma_f^2$, as discussed in the previous subsection, the angle–time domain AMMSE technique should achieve further performance gain.

VI. SIMULATION RESULTS

Computer simulations are carried out for the IEEE 802.11n channel models [24], [34]. As similar conclusions can be drawn from both the line-of-sight and NLOS scenarios based on our simulations, we only present the results for the NLOS scenario. We evaluated different channel estimation techniques for the five IEEE 802.11n channel models that represent various indoor environments such as residential homes and small offices. Model A corresponds to the MIMO flat-fading channel with a single cluster. This simple model serves as the basis to investigate the characteristics of cluster-based channel models in the angle domain. Here, $\{\text{AoD}_m, \text{AS}_t\}$ and $\{\text{AoA}_m, \text{AS}_r\}$ refer to the mean angle of clusters and angular spread of clusters, respectively, for the transmit and receive antennas, respectively, as illustrated in Fig. 3. All these four parameters are physically determined for a given propagation scenario and will affect the relative average power for each angle–time domain beam. In the simulation, we assign different values to these parameters for the model A to represent various propagation environments. This model is usually used for stressing the detection performance. It occurs only a small percentage of time in reality for the systems under consideration [24]. Models B–E correspond to 15-, 30-, 50-, and 100-ns root-mean-square delay spread, respectively. All these four models represent nonsample-spaced channels. As we found that the performance trends in different channel estimation techniques

are similar in all these four models, we only present the results for models B and E because of their minimum and maximum degrees of freedom, respectively. In the simulations, we assign the values indicated in [24, App. C] to $\{\text{AoD}_m, \text{AS}_t\}$ and $\{\text{AoA}_m, \text{AS}_r\}$ to represent typical small environments. We also evaluated different channel estimation techniques for the ideally assumed channel model that is spatially uncorrelated and sample spaced [4], [9], [10]. This channel is set to have the exponential power delay profile with the channel length N_h being 12. Throughout the simulations, we assume that $N_d = 64$ and $N_t = N_r = 4$, and that the normalized separation between adjacent antennas $\Delta_t = \Delta_r = 0.5$. Furthermore, we assume that the channel power for each link between one transmit and one receive antenna is normalized to one throughout the simulations.

A. Channel Model A

As discussed, we assign various values to $\{\text{AoD}_m, \text{AS}_t\}$ and $\{\text{AoA}_m, \text{AS}_r\}$ in channel model A to investigate the performance of the angle-domain channel estimation techniques in the presence of a single cluster. As the angular spread is usually not smaller than 40° in indoor environments [24], we consider 40° for both AS_t and AS_r as the worst case in performing angle-domain channel estimations. We also consider the angular spread to be 2° , which is valid for outdoor environments, for both AS_t and AS_r as the best case. Furthermore, the energy of multipath components from 45° leaks into more than one angle–time domain beam, which is undesirable in the angle-domain channel estimations. Thus, we consider 45° to be the AoD_m and the AoA_m of the cluster for the worse-case consideration. We also consider 0° , from which most of the energy of multipath components concentrates on one angle–time domain beam, to be the AoD_m and the AoA_m . In the following figures, which represent the channel power for each angle–time domain beam (see Figs. 4 and 6), the areas of square and circle are proportional to the angle- and array-domain normalized average power, respectively, with respect to the corresponding maximum average power in the angle domain.

Fig. 4 corresponds to the best case where the angular spread is 2° . As expected, we see that the angle-domain channel power is concentrated in the (1, 1)th angle–time domain beam, whereas the array-domain channel power is uniformly distributed over all the array–time domain beams. These observations

Fig. 4. Normalized channel power for each angle-time domain beam of model A with $\text{AoD}_m = 0^\circ$, $\text{AS}_t = 2^\circ$, $\text{AoA}_m = 0^\circ$, and $\text{AS}_r = 2^\circ$.

Fig. 6. Normalized channel power of model A for each angle-time domain beam with $\text{AoD}_m = 45^\circ$, $\text{AS}_t = 40^\circ$, $\text{AoA}_m = 45^\circ$, and $\text{AS}_r = 40^\circ$.

Fig. 5. Performance of different channel estimation techniques for model A with $\text{AoD}_m = 0^\circ$, $\text{AS}_t = 2^\circ$, $\text{AoA}_m = 0^\circ$, and $\text{AS}_r = 2^\circ$.

Fig. 7. Performance of different channel estimation techniques for model A with $\text{AoD}_m = 45^\circ$, $\text{AS}_t = 40^\circ$, $\text{AoA}_m = 45^\circ$, and $\text{AS}_r = 40^\circ$.

indicate the effectiveness of the MST selection and AMMSE techniques. From Fig. 5, it is clear that all the angle-domain techniques improve over the array-frequency domain LS technique at all the SNRs under consideration. We also find that the angle-time domain AMMSE technique improves over the angle-time domain MST selection technique by around 3.5 dB because the former is more like an LMMSE technique. Although the MST selection is a nonlinear process, it is interesting to find that the performance of the angle-domain techniques is proportional to that of the array-frequency domain LS technique in Fig. 5 because the ignored channel instantaneous power is not significant.

We consider the worst case where the angular spread is 40° in Figs. 6 and 7. Compared to Fig. 4, the angle-domain channel power tends to leak into other angle-time domain beams. Therefore, the channel powers in all angle-time domain beams are relatively high. From Fig. 7, we find that all the angle-time domain techniques still outperform the array-frequency domain

LS technique for all the SNRs. However, the angle-frequency domain MST selection technique does not perform well because the number of channel coefficients whose channel powers are above the noise variance is relatively large. We also observe that the angle-time domain AMMSE technique achieves the best performance, as expected.

B. Typical Channel Models

We consider typical nonsample-spaced indoor channels in Figs. 8 and 9. In such cases, AoD_m and AoA_m of clusters of multipath components tend to be uniformly distributed over all the angles. Then, the technique proposed in [23] will be the same as the array-frequency LS technique because all the angle-time domain beams are identified as signal beams. For this reason, its simulation results are not presented here. Compared with Figs. 5 and 7, we find that unlike in model A, the achieved performance gain of the angle-domain

Fig. 8. Performance of different channel estimation techniques for model B.

Fig. 9. Performance of different channel estimation techniques for model E.

techniques over the array–frequency domain LS technique is not significant at high SNRs in models B and E. This is because the ignored channel instantaneous powers are always larger in the nonsample-spaced channels. Nevertheless, the angle–time domain AMMSE technique still achieves the best performance for models B and E. Therefore, we can choose the angle–time domain AMMSE technique to perform channel estimation in the IEEE-802.11-based MIMO-OFDM systems because of its superior performance as well as its robustness.

As the channel power tends to be concentrated over a short time interval in the angle–time domain,⁷ the number of channel coefficients whose corresponding channel powers are below the noise variance is relatively large. For such channel coefficients, the optimum threshold should be infinite, as discussed in Section V. Intuitively, increasing the threshold will improve the estimation performance for these channel coefficients. However, this will result in an adverse effect on the estimation

⁷The time intervals for different angle–time domain beams are not likely to be the same because the first paths falling in each beam may asynchronously arrive.

Fig. 10. Performance of the angle-domain channel estimation techniques with different thresholds for model B. The number α shown in the bracket indicates that the threshold is set to $\alpha\sigma_f^2$. Otherwise, the threshold is set to $2\sigma_f^2$.

performance for the channel coefficients whose channel powers exceed σ_f^2 . As the number of channel coefficients whose channel powers are below σ_f^2 is larger compared to that of the remaining channel coefficients, it should be interesting to investigate the overall performance of the angle–time domain techniques when the threshold is increased. Fig. 10 shows that the angle–time MST selection and AMMSE techniques are improved in terms of performance when the threshold is increased to $3\sigma_f^2$ at nearly all the SNRs. Note that the selection of threshold is a tradeoff between the performance loss due to the ignorance of channel instantaneous power and the performance gain due to the removal of noise. Therefore, we find that further increasing the threshold to $6\sigma_f^2$ degrades the performance at high SNRs because the performance loss will dominate. Note that when the threshold is relatively high, we only reserve the channel coefficients whose corresponding $|\tilde{c}_{i_r, i_t}^a(l)|^2 - \sigma_f^2 / |\tilde{c}_{i_r, i_t}^a(l)|^2$ shown in (28) approaches one. This implies that the performance of the angle–time domain MST selection technique should gradually approach that of the angle–time domain AMMSE technique with the increase in thresholds. This implication is verified in Fig. 10. Therefore, we may use the angle–time domain MST selection technique when the threshold is relatively high for typical MIMO-OFDM systems.

Note that the noise variance σ_f^2 is required to be priorly known to decide the threshold in the above simulations. However, the exact knowledge of σ_f^2 may not always be available. Therefore, for a robust estimator design, we should fix the threshold for a target range of SNRs. As illustrated in Fig. 10, increasing the threshold will even improve the performance of angle–time domain techniques particularly at low SNRs. Therefore, for a given target SNR range, we may use $2\sigma_f^2$ at the lowest SNR (because of the largest σ_f^2) as the fixed threshold. In Fig. 11, we divide the whole SNR range into five disjoint groups. Each corresponds to one target SNR range, within which the threshold is fixed. Fig. 10 indicates that the threshold should be smaller (or at most slightly larger)

Fig. 11. Performance of the angle-domain channel estimation techniques for model B. “Fixed” in the brackets indicates that the threshold is fixed for a given SNR range.

than the corresponding $6\sigma_f^2$ at low SNRs and smaller than the corresponding $3\sigma_f^2$ at high SNRs. These results imply that we could choose a larger factor of σ_f^2 (and, thus, more elements for each group) at low SNRs than at high SNRs. From Fig. 11, we observe that the performance of angle-time domain estimations is always better than that of the array-frequency domain LS technique. Furthermore, we find that the performance of the angle-time domain MST selection technique is close to that of the angle-time domain AMMSE technique at the highest SNR in each target range (see the points when SNR = 10, 15, 19, 22, and 25 dB in Fig. 11). These results are consistent with the observations in Fig. 10, which shows that increasing the threshold makes the performance of the angle-time domain MST selection technique approach that of the angle-time domain AMMSE technique. However, the angle-frequency domain MST selection technique may not perform well at the highest SNR for each target range because of the relatively large performance loss due to the channel instantaneous power ignorance. As the angle-time domain AMMSE technique performs best at all the SNRs under consideration, we may conclude that the angle-time domain AMMSE technique is suitable for the typical IEEE 802.11n MIMO-OFDM systems when the target range of SNRs is available.

Note that at higher SNRs, the threshold η approaches zero, and thus, almost all the channel coefficients will not be filtered. Therefore, the angle-time domain MST selection and angle-frequency domain MST selection techniques will perform more similarly to the array-frequency domain LS technique as the SNR increases. In addition, the angle-time AMMSE technique will also perform more similarly to the angle-time domain MST selection technique as the SNR increases because the multiplication factor shown in (28) approaches to one for each angle-time domain channel coefficient. In summary, all the channel estimation techniques will converge at a relatively high SNR, as implied in all the figures shown in this subsection. Thus, the angle-time domain

Fig. 12. Performances of angle-domain channel estimation techniques for the spatially uncorrelated channel.

AMMSE technique is more preferable at lower SNRs than at higher SNRs because of the larger achievable performance gain.

C. Spatially Uncorrelated Channel

In Sections VI-A and B, we consider IEEE 802.11n channel models that are spatially correlated and show that the angle-domain channel estimation techniques can achieve good performance. In this subsection, we consider the sample-spaced spatially uncorrelated channel. In such a channel, the angle-time domain and array-time domain channel coefficients have the same statistics. Still, as the channel coefficients in the angle-time domain are uncorrelated with each other, from Fig. 12, we find that the AMMSE technique in this domain can achieve the best performance. Therefore, compared with the array-frequency domain LS technique, we can always use the angle-domain AMMSE technique for different channels to achieve higher performance at the price of higher complexity.

VII. CONCLUSION

In this paper, the angle-domain representation is used to describe MIMO-OFDM systems. Based on this representation, we propose the angle-frequency domain MST selection technique, the angle-time domain MST selection technique, and the angle-time domain AMMSE technique. These three techniques do not require prior knowledge of the channel correlation and are shown to be effective when the angular spread of clusters of multipath components is small. More importantly, the angle-time domain techniques can improve over the array-frequency domain LS technique in all the cases under consideration even when the angular spread is large, at the price of higher complexity. Furthermore, our theoretical analysis and simulation results indicate that the angle-time domain AMMSE technique achieves the best performance and is robust to the choice of threshold and mismatch of operating an SNR. Thus, when only the target SNR range is available to the receiver,

the angle-time domain AMMSE technique is a potential candidate for typical IEEE 802.11n MIMO-OFDM systems. In addition, we also find that with a suitable threshold and a known operating SNR, the angle-time domain MST selection technique results in little performance degradation compared to the angle-time domain AMMSE technique. Therefore, in such cases, the angle-time domain MST selection technique may be used for IEEE 802.11n MIMO-OFDM systems because of its lower computational complexity compared to the angle-time domain AMMSE technique.

APPENDIX PERFORMANCE COMPARISON OF AMMSE AND MST SELECTION TECHNIQUES

In this appendix, we analytically compare the performance of the AMMSE and MST selection techniques. From (43) and (44), the mse due to the i th estimated channel coefficient is given by

$$\begin{aligned} \text{mse}_i &= \int_0^\infty \frac{\hat{v}_i}{\sigma_f^2} e^{-\frac{\hat{v}_i}{\sigma_f^2}} \int_\eta^\infty \frac{(\hat{r}_i - \sigma_f^2)^2}{\hat{r}_i^2} \frac{e^{-\frac{\hat{r}_i + \hat{v}_i}{\sigma_i^2}}}{\sigma_i^2} \\ &\quad \times J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{v}_i}}{\sigma_i^2} \right) d\hat{r}_i d\hat{v}_i + \int_0^\infty \frac{\hat{h}_i}{\sigma_i^2} e^{-\frac{\hat{h}_i}{\sigma_i^2}} \\ &\quad \times \int_0^\eta \frac{\sigma_f^4}{\hat{r}_i^2} \frac{e^{-\frac{\hat{r}_i + \hat{h}_i}{\sigma_f^2}}}{\sigma_f^2} J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{h}_i}}{\sigma_f^2} \right) d\hat{r}_i d\hat{h}_i \\ &\quad + \int_0^\infty \frac{\hat{h}_i}{\sigma_i^2} e^{-\frac{\hat{h}_i}{\sigma_i^2}} \int_0^\eta \frac{e^{-\frac{\hat{r}_i + \hat{h}_i}{\sigma_f^2}}}{\sigma_f^2} J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{h}_i}}{\sigma_f^2} \right) d\hat{r}_i d\hat{h}_i. \end{aligned} \quad (51)$$

Compared to (46), the difference of mse_i , which is contributed by the i th channel coefficient, between the AMMSE technique and the MST selection technique is given by

$$\begin{aligned} \text{DIF}_i &= \int_0^\infty \frac{\hat{v}_i}{\sigma_f^2} e^{-\frac{\hat{v}_i}{\sigma_f^2}} \int_\eta^\infty \frac{\sigma_f^4}{\hat{r}_i^2} \frac{e^{-\frac{\hat{r}_i + \hat{v}_i}{\sigma_i^2}}}{\sigma_i^2} \\ &\quad \times J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{v}_i}}{\sigma_i^2} \right) d\hat{r}_i d\hat{v}_i + \int_0^\infty \frac{\hat{h}_i}{\sigma_i^2} e^{-\frac{\hat{h}_i}{\sigma_i^2}} \\ &\quad \times \int_\eta^\infty \frac{\sigma_f^4}{\hat{r}_i^2} \frac{e^{-\frac{\hat{r}_i + \hat{h}_i}{\sigma_f^2}}}{\sigma_f^2} J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{h}_i}}{\sigma_f^2} \right) d\hat{r}_i d\hat{h}_i \\ &\quad - 2 \int_0^\infty \frac{\hat{v}_i}{\sigma_f^2} e^{-\frac{\hat{v}_i}{\sigma_f^2}} \int_\eta^\infty \frac{\sigma_f^2}{\hat{r}_i} \frac{e^{-\frac{\hat{r}_i + \hat{v}_i}{\sigma_i^2}}}{\sigma_i^2} \\ &\quad \times J_0 \left(\frac{2\sqrt{\hat{r}_i \hat{v}_i}}{\sigma_i^2} \right) d\hat{r}_i d\hat{v}_i. \end{aligned} \quad (52)$$

Differentiating DIF_i with respect to η , we obtain

$$\frac{\partial \text{DIF}_i}{\partial \eta} = \frac{\sigma_f^4 e^{-\frac{\eta}{\sigma_i^2 + \sigma_f^2}} f(\sigma_i^2, \sigma_f^2, \eta)}{(\sigma_i^2 + \sigma_f^2)^3 \eta^2} \quad (53)$$

where

$$\begin{aligned} f(\sigma_i^2, \sigma_f^2, \eta) &= 2\sigma_f^2 \eta^2 + (\sigma_i^4 + 2\sigma_i^2 \sigma_f^2 - \sigma_f^4) \eta - 2\sigma_i^2 \sigma_f^2 (\sigma_i^2 + \sigma_f^2) \\ &= 2\sigma_f^2 \left(\eta - \frac{\sigma_f^4 - 2\sigma_f^2 \sigma_i^2 - \sigma_i^4}{4\sigma_f^2} \right)^2 - 2\sigma_i^2 \sigma_f^2 (\sigma_i^2 + \sigma_f^2) \\ &\quad - \frac{(\sigma_f^4 - 2\sigma_f^2 \sigma_i^2 - \sigma_i^4)^2}{8\sigma_f^2}. \end{aligned} \quad (54)$$

As the threshold η is not smaller than σ_f^2 , which is larger than $(\sigma_f^4 - 2\sigma_f^2 \sigma_i^2 - \sigma_i^4)/(4\sigma_f^2)$, (54) reaches its minimum 0 when the threshold η is equal to σ_f^2 . Therefore, when the threshold η is larger than σ_f^2 , (54) is larger than 0 (i.e., positive), and so is (53). Therefore, DIF_i is monotonically increasing with the increase in η and reaches its maximum 0 when $\eta = \infty$. Consequently, DIF_i is always negative, which implies that the AMMSE technique performs better than the MST selection technique when σ_i^2 is not equal to zero.

ACKNOWLEDGMENT

The authors would like to thank Cellular Systems Division, Aalborg University (AAU-CSys), Aalborg, Denmark, and Computer Science Institute, The University of Namur (FUNDP-INFO), Namur, Belgium, for their parenthood on the MATLAB packages for the simulation of MIMO-OFDM channel models. The authors would also like to thank Project IST-2000-30148 I-METRA, which permitted the use of MATLAB packages of channel models.

REFERENCES

- [1] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband channels," in *Proc. Int. IEEE Veh. Technol. Conf.*, Ottawa, ON, Canada, May 1998, pp. 2232–2236.
- [2] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless system: Design, performance, and field trial results," *IEEE Commun. Mag.*, vol. 40, no. 9, pp. 143–149, Sep. 2002.
- [3] A. van Zelst and T. C. W. Schenk, "Implementation of MIMO OFDM-based wireless LAN system," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 483–494, Feb. 2004.
- [4] G. L. Stüber, J. R. Barry, S. W. McLaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, Feb. 2004.
- [5] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001, ch. 5.
- [6] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Börjesson, "On channel estimation in OFDM systems," in *Proc. IEEE Int. Veh. Technol. Conf.*, Chicago, IL, Jul. 1995, pp. 815–819.
- [7] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson, and P. O. Börjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, Jul. 1998.

- [8] Y. Li, L. J. Cimini, Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.
- [9] M. Shin, H. Lee, and C. Lee, "Enhanced channel-estimation technique for MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 261–265, Jan. 2004.
- [10] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.
- [11] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wirel. Pers. Commun.*, vol. 6, no. 3, pp. 311–335, May 1998.
- [12] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.
- [13] H. Shin, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.
- [14] X. Mestre, J. R. Fonollosa, and A. P. Zamora, "Capacity of MIMO channels: Asymptotic evaluation under correlated fading," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 829–838, Jun. 2003.
- [15] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.
- [16] M. Enescu, T. Roman, and V. Koivunen, "Channel estimation and tracking in spatially correlated MIMO OFDM systems," in *Proc. IEEE Workshop Stat. Signal Process.*, St. Louis, MO, Sep. 2003, pp. 347–350.
- [17] H. Miao and M. J. Juntti, "Spatial signature and channel estimation for wireless MIMO-OFDM systems with spatial correlation," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun.*, Lisbon, Portugal, Jul. 2004, pp. 522–526.
- [18] H. Zhang, Y. Li, A. Reid, and J. Terry, "Channel estimation for MIMO OFDM in correlated fading channel," in *Proc. IEEE ICC*, Seoul, Korea, May 2005, pp. 2626–2630.
- [19] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.
- [20] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005, ch. 7.
- [21] R. B. Ert, P. Cardieri, K. W. Sowerby, T. S. Rapport, and J. H. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Pers. Commun.*, vol. 5, no. 1, pp. 10–22, Feb. 1998.
- [22] F. Gross, *Smart Antennas for Wireless Communications With MATLAB*. New York: McGraw-Hill, 2005, ch. 4.
- [23] B. Liu, R. Jin, and Y. Fan, "Modified pre-FFT OFDM adaptive antenna array with beam-space channel estimation," *Electron. Lett.*, vol. 40, no. 5, pp. 287–288, Mar. 2004.
- [24] V. Erceg *et al.*, *TGn Channel Models*, May 2004. IEEE 802.11 document 03/940r4. [Online]. Available: <ftp://ieee:wireless@ftp.802wirelessworld.com/11/03/11-03-0940-04-000n-tgn-channel-models.doc>
- [25] *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, M. Abramowitz and I. A. Stegun, Eds. Mineola, NY: Dover, 1972, ch. 3 and 26.
- [26] C.-C. Chong, C.-M. Tan, D. I. Laurenson, S. McLaughlin, M. A. Beach, and A. R. Nix, "A new statistical wideband spatio-temporal channel model for 5-GHz band WLAN systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 139–150, Feb. 2003.
- [27] L. Huang, J. W. M. Bergmans, and F. M. J. Willems, "Low-complexity LMMSE-based MIMO-OFDM channel estimation via angle-domain processing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5668–5680, Dec. 2007.
- [28] S. Sun, I. Wiemer, C. K. Ho, and T. T. Tjhung, "Training sequence assisted channel estimation for MIMO OFDM," in *Proc. IEEE WCNC*, New Orleans, LA, Mar. 2003, pp. 38–43.
- [29] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.
- [30] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993, ch. 2 and 3.
- [31] L. Huang, G. Mathew, and J. W. M. Bergmans, "Pilot-aided channel estimation for systems with virtual carriers," in *Proc. ICC*, Istanbul, Turkey, Jun. 2006, pp. 3070–3075.
- [32] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer in the 5 GHz Band*, IEEE Std. 802.11a-1999, Sep. 1999.
- [33] Q. H. Spencer, B. D. Jeffs, M. A. Jensen, and A. L. Swindlehurst, "Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 347–360, Mar. 2000.
- [34] J. P. Kermaol, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.
- [35] J. S. Walker, *Fast Fourier Transforms*, 2nd ed. Boca Raton, FL: CRC, 1996, ch. 3.

Li Huang (S'04–M'08) received the B.Eng. degree in electronics and information engineering from Huazhong University of Science and Technology, Wuhan, China, in 2002. After 2002, he was initially enrolled in an M.Eng. program at the National University of Singapore (NUS), Singapore, under the Data Storage Institute (DSI) Scholarship and was later selected to be part of the joint Ph.D. program between NUS and Eindhoven University of Technology, Eindhoven, The Netherlands, in 2004 under the Design Technology Institute (DSI) Scholarship.

Since 2006, he has been with the Interuniversity Microelectronics Centre Nederland (IMEC-NL), Holst Centre, Eindhoven, working on physical layer design for ultralow power wireless systems. His main research interests include statistical signal processing, wireless communications, magnetics, and optical recording.

Chin Keong Ho (M'08) received the B.Eng. (first-class honors) and M.Eng. degrees from the National University of Singapore, Singapore, in 1999 and 2001, respectively. He is currently working toward the Ph.D. degree at the Eindhoven University of Technology, Eindhoven, The Netherlands.

He is also currently with the Institute for Infocomm Research, Singapore. His research interests include signal processing and wireless communications for multicarrier and space-time communications.

J. W. M. Bergmans (M'95–SM'91) received the Ingenieur and Ph.D. degrees from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 1981 and 1987, respectively.

From 1982 to 1999, he was with Philips Research Labs, Eindhoven, working on signal processing for digital transmission and recording systems. In 1988 and 1989, he was an Exchange Researcher with Hitachi Central Research Labs, Tokyo, Japan. He is currently a Professor of signal processing systems with the Eindhoven University of Technology.

Frans M. J. Willems (F'05) was born in Stein, The Netherlands, on June 26, 1954. He received the M.Sc. degree in electrical engineering from the Eindhoven University of Technology, Eindhoven, The Netherlands, and the Ph.D. degree from the Catholic University of Louvain, Louvain, Belgium, in 1979 and 1982, respectively.

From 1979 to 1982 he was a Research Assistant with the Catholic University of Louvain. Since 1982, he has been a Staff Member with the Department of Electrical Engineering, Eindhoven University of Technology.

Since 1999, he has been connected with Philips Research Laboratories as an Advisor. His research contributions are in the areas of multiuser information theory and noiseless source coding.

Dr. Willems received the Marconi Young Scientist Award in 1982 and was a corecipient of the IEEE Information Theory Society Paper Award in 1996. From 1988 to 1990, he served as an Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY. From 1998 to 2000, he was a member of the Board of Governors of the IEEE Information Theory Society. From 2001 to 2004, he served as an Associate Editor for information theory for the *European Transactions on Telecommunications*.