

Parallel-Interference-Cancellation-Assisted Decision-Directed Channel Estimation for OFDM Systems Using Multiple Transmit Antennas

Matthias MÜNSTER and Lajos HANZO, *Fellow, IEEE*

Abstract—The number of transmit antennas that can be employed in the context of least-squares (LS) channel estimation contrived for orthogonal frequency division multiplexing (OFDM) systems employing multiple transmit antennas is limited by the ratio of the number of subcarriers and the number of significant channel impulse response (CIR)-related taps. In order to allow for more complex scenarios in terms of the number of transmit antennas and users supported, CIR-related tap prediction-filtering-based parallel interference cancellation (PIC)-assisted decision-directed channel estimation (DDCE) is investigated. New explicit expressions are derived for the estimator's mean-square error (MSE), and a new iterative procedure is devised for the offline optimization of the CIR-related tap predictor coefficients. These new expressions are capable of accounting for the estimator's novel recursive structure. In the context of our performance results, it is demonstrated, for example, that the estimator is capable of supporting $L = 16$ transmit antennas, when assuming $K = 512$ subcarriers and $K_0 = 64$ significant CIR taps, while LS-optimized DDCE would be limited to employing $L = 8$ transmit antennas.

Index Terms—Decision-directed channel estimation (DDCE), multiple transmit antennas, orthogonal frequency division multiplexing (OFDM), parallel interference cancellation (PIC).

I. MOTIVATION

IN RECENT years, the family of single- and multiuser orthogonal frequency division multiplexing (OFDM) schemes [1] using time-domain, frequency-domain, as well as spatial-domain spreading [2] has enjoyed a renaissance. Hence, OFDM has found its way into numerous wireless systems that require accurate channel estimation. Accordingly, the topic of decision-directed channel estimation (DDCE) has been addressed in a variety of contributions, notably, for example, in the detailed discussions of [3]–[8], in the context of single-user single-transmit antenna OFDM environments. The basic idea

is to equalize the channel transfer function experienced by an OFDM symbol during the current transmission period by capitalizing on that encountered during the previous OFDM-symbol period [1].¹ This implies assuming quasi-invariance of the channel's transfer function between the two consecutive OFDM symbols' transmission intervals. An improved channel transfer-function estimate can then be obtained for detecting the most recently received OFDM symbol upon dividing the complex symbol received in each subcarrier by the sliced and remodulated information symbol hosted by a subcarrier. The updated channel estimate is then employed again as an initial channel estimate during the next OFDM symbol's transmission period.

By contrast, channel estimation for multiuser OFDM has been considered in [9]–[15]. In the context of the multiuser OFDM scenario to be outlined in Section II, the signal received by each antenna is constituted by the superposition of the signal contributions associated with the different users or transmit antennas. Note that, in terms of the multiple-input multiple-output (MIMO) structure of the channel, the multiuser single-transmit-antenna scenario is equivalent, for example, to a single-user space-time coded (STC) scenario using multiple transmit antennas, provided that the channels associated with the different transmit antennas are perfectly uncorrelated. Hence, channel-estimation techniques designed for multiuser OFDM can be employed for single-user STC and vice versa. For the latter scenario, a least-squares (LS) error channel estimator was proposed by Li *et al.* [9], which aims at recovering the different transmit antennas' channel transfer functions on the basis of the output signal of a specific reception antenna element and by also capitalizing on the remodulated received symbols associated with the different users. The performance of this estimator was found to be limited in terms of the mean-square estimation error in scenarios where the product of the number of transmit antennas and the number of channel impulse response (CIR) taps to be estimated per transmit antenna approaches the total number of subcarriers hosted by an OFDM symbol.

In [10], a DDCE was proposed by Jeon *et al.* for an STC OFDM scenario of two transmit antennas and two receive

Manuscript received July 3, 2002; revised October 16, 2003 and June 13, 2005; accepted June 15, 2005. The editor coordinating the review of this paper and approving it for publication is S. Roy. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), U.K., and the European Commission.

The authors are with the Department of Electronics and Computer Science (ECS), University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

Digital Object Identifier 10.1109/TWC.2005.855793

¹For sample chapters, please refer to www-mobile.ecs.soton.ac.uk.

antennas. Specifically, the channel transfer function² associated with each transmit–receive antenna pair was estimated in the frequency domain on the basis of the output signal of the specific receive antenna upon subtracting the estimated interfering signal contributions associated with the remaining transmit antennas.

By contrast, in [11] and [13], a similar technique was proposed by Li with the aim of simplifying the DDCE approach of [9], which operates in the time domain. A prerequisite for the operation of this parallel interference cancellation (PIC)-assisted DDCE is the availability of a reliable estimate of the various channel transfer functions for the current OFDM symbol, which are employed in the cancellation process in order to obtain updated channel transfer-function estimates for the demodulation of the next OFDM symbol. In order to compensate for the channel’s variation as a function of the OFDM-symbol index, linear prediction techniques can be employed, as it was also proposed, for example, in [11]. However, due to the estimator’s recursive structure, determining the optimum predictor coefficients is not as straightforward as for a transversal finite-impulse response (FIR) filter-assisted predictor such as that proposed in [4]. This will be the topic of our following discourse.

In contrast to the above backcloth, the paper proposes rank-reduced DDCE for multiuser OFDM systems supporting a high number of users. To set the scene more explicitly, channel estimation for single-user OFDM was discussed, for example, by Li *et al.* [9], as well as by Edfors *et al.* [3]. To elaborate further, the rank-reduced filtering approach, the linear predictive procedure along the time direction, and the PIC process have been used before in various different contexts. However, in addition to supporting a high number of multiple users, a further novel aspect of our scheme is the specific iterative method proposed for calculating the predictor’s coefficients, which can be invoked offline. This important aspect was not treated by Li in [13], since single-tap prediction filtering was considered. The predictor’s stability analysis is also new.

The outline of the papers is as follows. In Section II, we portray the signal model associated with the spatial division multiple access (SDMA) uplink transmission scenario. Note again that in terms of the MIMO structure of the channel, this SDMA system is equivalent, for example, to a single-user STC scenario employing multiple transmit antennas. Hence, the algorithms discussed here are amenable to a wide range of applications involving multiple transmit antennas. In Section III, we will then focus on the PIC-assisted DDCE employing prediction filtering along the time axis in the CIR-related domain. Expressions are derived for the estimator’s mean-square error (MSE) and an iteration-based novel approach is devised for evaluating the optimum CIR-related tap predictor coefficients. This is followed in Section IV by an extensive performance assessment under both sample-spaced and nonsample-spaced CIR conditions. Finally, in Section V, the PIC-assisted DDCE’s complexity will be studied. Our conclusions will be offered in

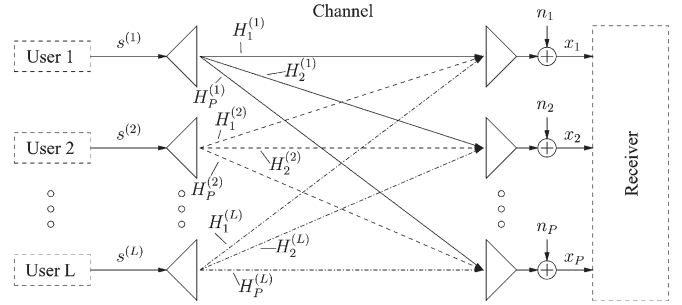


Fig. 1. Schematic of an SDMA uplink scenario, as observed on an OFDM subcarrier basis, where each of the L users is equipped with a single transmit antenna and the BS’s receiver is assisted by a P -element antenna front end. For comparison, in an STC scenario, the L transmit antennas are used for providing L th-order transmit diversity for a single user. Note that, for the sake of simplicity, we have omitted the subcarrier index k .

Section VI. In the Appendix, the conditions for the estimator’s stability are presented.

II. THE SDMA SIGNAL MODEL ON A SUBCARRIER BASIS

In Fig. 1, we have portrayed an SDMA uplink transmission scenario, where each of the P simultaneous users is equipped with a single transmission antenna, while the receiver capitalizes on a P -element antenna front end. The set of complex signals, $x_p[n, k]$, $p = 1, \dots, P$ received by the P -element antenna array in the k th subcarrier of the n th OFDM symbol is constituted by the superposition of the independently faded frequency-domain signals associated with the L users sharing the same space-frequency resource. The received signal was corrupted by the Gaussian noise at the array elements. Regarding the statistical properties of the different signal components depicted in Fig. 1, we assume that the complex data signal $s^{(l)}$ transmitted by the l th user has zero mean and a variance of σ_l^2 . The additive white Gaussian noise (AWGN) noise process n_p at any antenna-array element p also exhibits zero mean and a variance of σ_n^2 , which is identical for all array elements. The frequency-domain channel transfer factors $H_p^{(l)}$ of the different array elements $p = 1, \dots, P$ or users $l = 1, \dots, L$ are independent stationary complex Gaussian distributed random variables with zero mean and unit variance, namely, $\sigma_H^2 = 1$.

III. ANALYTICAL DESCRIPTION OF FREQUENCY-DOMAIN PIC-ASSISTED DDCE³

The specific structure of Section III is as follows. Our portrayal of the frequency-domain PIC-assisted DDCE commences in Section III-A, where we introduce the channel-estimation algorithm. Furthermore, in Section III-B, an iterative algorithm is developed for the offline calculation of the predictor coefficients. Since normally the exact knowledge of the channel’s statistics in the form of the space–time space–frequency correlation function is unavailable, in Section III-C, we discuss potential strategies for providing estimates of the statistics required.

²In the context of the OFDM system, the set of K different subcarriers’ channel transfer factors is referred to as the channel transfer function, or simply as the channel.

³Unless otherwise stated, the channel transfer-function estimator is simply referred to as the channel estimator.

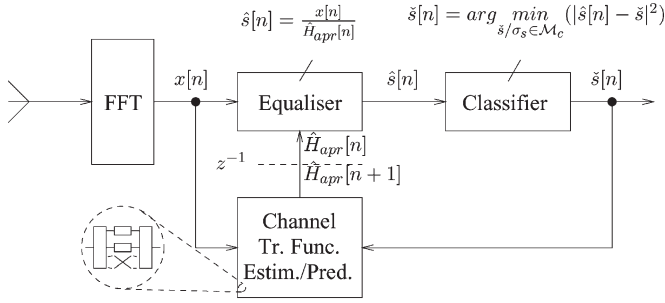


Fig. 2. DDCE-aided OFDM receiver.

A. Structure of the PIC-Assisted DDCE

In Section III-A1, we discuss the structure of the PIC unit. Expressions are provided both for the *a posteriori* channel transfer-factor estimates arrived at after the PIC, as well as for the *a priori* channel transfer-factor estimates upon taking into account the effects of the CIR-related tap prediction filter. The specific structure of the predictor arrangement is detailed in Section III-A2.

1) *Design of the PIC Unit:* The structure of the OFDM receiver is outlined in Fig. 2 and our discussions are related to the DDCE block seen at the bottom of Fig. 2, which will be discussed in more detail in the context of Fig. 3. To elaborate a little further, following fast Fourier transform (FFT) processing, the complex frequency-domain signal $x[n, k]$ associated with each of the K subcarriers, $k = 0, \dots, K - 1$, is equalized based on the *a priori* channel transfer-factor estimates generated during the previous OFDM-symbol period for employment during the current period. As a result, the linear estimates $\hat{s}[n, k]$ of the transmitted signals $s[n, k]$ are obtained. These estimates are classified, yielding the complex symbols, $\check{s}[n, k]$ that are most likely to have been transmitted. The classified symbols $\check{s}[n, k]$ are then employed together with the received subcarrier signals $x[n, k]$ for generating *a priori* channel transfer-factor estimates for employment during the $(n + 1)$ th OFDM-symbol period. The specific structure of the DDCE scheme, which is indicated by the stylized illustration at the bottom left corner, will be detailed in the paper.

Explicitly, the complex output signal $x_p[n, k]$ of the p th receiver-antenna element in the k th subcarrier of the n th OFDM symbol is given by

$$x_p[n, k] = \sum_{i=1}^L H_p^{(i)}[n, k] s^{(i)}[n, k] + n_p[n, k] \quad (1)$$

where the different variables have been defined in Section II. Upon invoking vector notation, (1) can be rewritten as

$$\mathbf{x}_p[n] = \sum_{i=1}^L \mathbf{S}^{(i)}[n] \mathbf{H}_p^{(i)}[n] + \mathbf{n}_p[n] \quad (2)$$

where $\mathbf{x}_p[n] \in \mathbb{C}^{K \times 1}$, $\mathbf{H}_p^{(i)}[n] \in \mathbb{C}^{K \times 1}$, and $\mathbf{n}_p[n] \in \mathbb{C}^{K \times 1}$ are column vectors hosting the subcarrier-related variables $x_p[n, k]$, $H_p^{(i)}[n, k]$, and $n_p[n, k]$, respectively, and $\mathbf{S}^{(i)}[n] \in \mathbb{C}^{K \times K}$ is a diagonal matrix having elements given by $s^{(i)}[n, k]$,

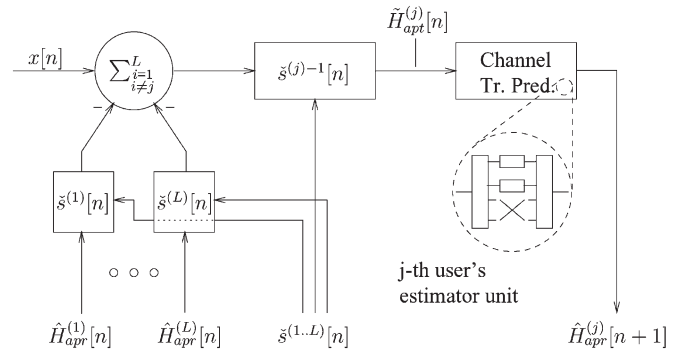


Fig. 3. Illustration of the PIC-assisted channel transfer-function estimation or prediction block, associated with the j th user and any of the P receiver-antenna elements. The PIC process is described by (3). The structure of the channel transfer-function predictor follows the concepts of low-rank DDCE investigated in [3], [4], and [16] for the single-user scenario.

where $k = 0, \dots, K - 1$. An *a posteriori* (apt) estimate $\tilde{\mathbf{H}}_{\text{apt}}^{(j)}[n] \in \mathbb{C}^{K \times 1}$ of the vector $\mathbf{H}^{(j)}[n]$ of “true” channel transfer factors between the j th user’s single transmit antenna and the p th receiver antenna can be obtained by subtracting all the $(L - 1)$ vectors of interfering users’ estimated signal contributions from the vector $\mathbf{x}_p[n]$ of composite received signals of the L users, followed by normalization with the j th user’s diagonal matrix of detected complex symbols $\check{\mathbf{S}}^{(j)}[n]$, yielding

$$\tilde{\mathbf{H}}_{\text{apt}}^{(j)}[n] = \check{\mathbf{S}}^{(j)-1}[n] \left(\mathbf{x}[n] - \sum_{\substack{i=1 \\ i \neq j}}^L \check{\mathbf{S}}^{(i)}[n] \hat{\mathbf{H}}_{\text{apr}}^{(i)}[n] \right) \quad (3)$$

where, for simplicity’s sake, we have omitted the receiver antenna’s index p .⁴ The PIC process based on (3) has been further illustrated in Fig. 3. In (3), $\hat{\mathbf{H}}_{\text{apr}}^{(i)}[n] \in \mathbb{C}^{K \times 1}$ denotes the i th user’s vector of complex *a priori* (apr) channel transfer-factor estimates predicted during the $(n - 1)$ th OFDM-symbol period for the n th OFDM symbol, as a function of the vectors of *a posteriori* channel transfer-factor estimates $\tilde{\mathbf{H}}_{\text{apt}}^{(i)}[n - \hat{n}]$ associated with the previous $N_{\text{tap}}^{[i]}$ number of OFDM symbols, which is formulated as

$$\hat{\mathbf{H}}_{\text{apr}}^{(i)}[n] = f \left(\tilde{\mathbf{H}}_{\text{apt}}^{(i)}[n - 1], \dots, \tilde{\mathbf{H}}_{\text{apt}}^{(i)}[n - N_{\text{tap}}^{[i]}] \right). \quad (4)$$

We will further elaborate on the specific structure of the predictor in the next section.

2) *Design of the Predictor Unit:* The channel transfer-function prediction along the time direction follows the philosophy of the two-dimensional (2-D) minimum MSE (MMSE) channel transfer-function-estimation approach proposed by Li *et al.* [4], which in turn is based on the rank-reduction-assisted one-dimensional (1-D) MMSE channel-estimation scheme proposed by Edfors and co-workers [3], [16]. A prerequisite for the optimality of these techniques was stated by Li *et al.* [4], arguing that the channel’s space–time

⁴This is also justified by the observation that the channel-estimation processes of different receiver antennas are decoupled.

space–frequency correlation function $r_H(\Delta t, \Delta f) \in \mathbb{C}$ is separable:

$$r_H(\Delta t, \Delta f) = E \{ H(t_1, f_1) H^*(t_2, f_2) \} \quad (5)$$

$$= r_H(\Delta t) \cdot r_H(\Delta f) \quad (6)$$

where $H(t, f) \in \mathbb{C}$ denotes the channel's frequency-domain transfer factor at time–frequency position (t, f) and $r_H(\Delta t) \in \mathbb{C}$ is the channel's space–time correlation function, while $r_H(\Delta f) \in \mathbb{C}$, the space-frequency correlation function. It was demonstrated by Li *et al.* [4] that this separation is valid under the assumption of uncorrelated scattering and upon further assuming that identical normalized space–time correlation functions are associated with the different paths. For further related notes, please refer to [17, Ch. 2]. Let us now briefly portray the unitary transform-based channel transfer-function predictor associated with the i th user.

- 1) In the first step, in order to obtain the i th user's vector of *a priori* channel transfer-factor estimates for the n th OFDM-symbol period during the $(n-1)$ th OFDM-symbol period, which is denoted by $\tilde{\mathbf{H}}_{\text{apr}}^{(i)}[n]$, the vector of *a posteriori* channel transfer-factor estimates $\tilde{\mathbf{H}}_{\text{apt}}^{(i)}[n-1]$ is subjected to a unitary linear inverse transform $\tilde{\mathbf{U}}^{[f](i)\text{H}} \in \mathbb{C}^{K \times K}$, yielding the vector $\tilde{\mathbf{h}}_{\text{apt}}^{(i)}[n-1] \in \mathbb{C}^{K \times 1}$ of CIR-related *a posteriori* tap values:

$$\tilde{\mathbf{h}}_{\text{apt}}^{(i)}[n-1] = \tilde{\mathbf{U}}^{[f](i)\text{H}} \tilde{\mathbf{H}}_{\text{apt}}^{(i)}[n-1]. \quad (7)$$

From a statistical point of view, the optimum unitary transform to be employed is the Karhunen–Loeve transform (KLT) [3], [18] with respect to the Hermitian space-frequency correlation matrix of *a posteriori* channel transfer-factor estimates, which is given by $\mathbf{R}_{\text{apt}}^{[f](i)} = E \{ \tilde{\mathbf{H}}_{\text{apt}}^{(i)} \tilde{\mathbf{H}}_{\text{apt}}^{(i)\text{H}} \}$, when assuming the wide-sense stationarity of $\tilde{\mathbf{H}}_{\text{apt}}^{(i)}[n]$. The matrix $\mathbf{R}_{\text{apt}}^{[f](i)} \in \mathbb{C}^{K \times K}$ can be decomposed as $\mathbf{R}_{\text{apt}}^{[f](i)} = \mathbf{U}_{\text{apt}}^{[f](i)} \mathbf{\Lambda}_{\text{apt}}^{[f](i)} \mathbf{U}_{\text{apt}}^{[f](i)\text{H}}$, where $\mathbf{U}_{\text{apt}}^{[f](i)} \in \mathbb{C}^{K \times K}$ is the unitary KLT matrix of eigenvectors, and $\mathbf{\Lambda}_{\text{apt}}^{[f](i)} \in \mathbb{R}^{K \times K}$ exhibits the diagonal form of $\mathbf{\Lambda}_{\text{apt}}^{[f](i)} = \text{diag}(\lambda_{\text{apt},0}^{[f](i)}, \dots, \lambda_{\text{apt},K-1}^{[f](i)})$. The diagonal elements of $\mathbf{\Lambda}_{\text{apt}}^{[f](i)}$ are referred to as the eigenvalues of $\mathbf{R}_{\text{apt}}^{[f](i)}$ [19]. Similarly, the desired channel's “true” space-frequency correlation matrix $\mathbf{R}^{[f](i)} = E \{ \mathbf{H}^{[i]} \mathbf{H}^{[i]\text{H}} \}$ can be decomposed as $\mathbf{R}^{[f](i)} = \mathbf{U}^{[f](i)} \mathbf{\Lambda}^{[f](i)} \mathbf{U}^{[f](i)\text{H}}$. At this stage, we note that the error components contaminating the vector $\tilde{\mathbf{H}}_{\text{apt}}^{(i)}[n-1]$ estimating the vector $\mathbf{H}^{(i)}[n-1]$ of “true” channel transfer factors are uncorrelated due to the statistical independence of the AWGN and that of the modulated symbols transmitted in the different subcarriers. Hence, both $\mathbf{R}_{\text{apt}}^{[f](i)}$ and $\mathbf{R}^{[f](i)}$ share the same eigenvectors [18], which implies that we have $\mathbf{U}_{\text{apt}}^{[f](i)} = \mathbf{U}^{[f](i)}$. In reality, however, the explicit knowledge of the channel's space-frequency correlation

matrix $\mathbf{R}^{[f](i)}$ and that of its unitary KLT matrix $\mathbf{U}^{[f](i)}$ is typically unavailable. Instead, an estimate $\tilde{\mathbf{R}}^{[f](i)}$ and its associated unitary KLT matrix $\tilde{\mathbf{U}}^{[f](i)}$ has to be employed, which—in contrast to the optimum KLT matrix $\mathbf{U}^{[f](i)}$ —results in an imperfect decorrelation of the *a posteriori* channel transfer-factor estimates.

- 2) In the second step, linear $N_{\text{tap}}^{[t]}$ -tap filtering is performed in the time direction separately for those K_0 number of CIR-related components of $\tilde{\mathbf{h}}_{\text{apt}}^{(i)}$, for which the variance is significant. This is achieved by capitalizing on the current vector $\tilde{\mathbf{h}}_{\text{apt}}^{(i)}[n-1]$ and the vectors $\tilde{\mathbf{h}}_{\text{apt}}^{(i)}[n-\hat{n}]$, $\hat{n} = 2, \dots, N_{\text{tap}}^{[t]}$ of the previous $(N_{\text{tap}}^{[t]} - 1)$ number of OFDM symbols. As a result, in the case of estimation filtering [4], an improved estimate $\hat{\mathbf{h}}_{\text{apt}}^{(i)}[n-1]$ of $\mathbf{h}^{(i)}[n-1]$ is obtained, although this technique was not employed here. By contrast, in the case of the prediction filtering employed here, an *a priori* estimate $\hat{\mathbf{h}}_{\text{apr}}^{(i)}[n] \in \mathbb{C}^{K \times 1}$ of $\mathbf{h}^{(i)}[n]$ is obtained. In mathematical terms, this can be formulated as

$$\hat{\mathbf{h}}_{\text{apr}}^{(i)}[n] = \mathbf{I}_{K_0}^{(i)} \sum_{\hat{n}=1}^{N_{\text{tap}}^{[t]}} \tilde{c}_{\text{pre}}^{(i)}[\hat{n}-1] \tilde{\mathbf{h}}_{\text{apt}}^{(i)}[n-\hat{n}] \quad (8)$$

where $\mathbf{I}_{K_0}^{(i)} \in \mathbb{C}^{K \times K}$ denotes a sparse unity matrix having unity entries only at those K_0 number of diagonal positions, for which the variance of the associated components of $\tilde{\mathbf{h}}_{\text{apt}}^{(i)}$ is significant. Furthermore, in (8), the variable $\tilde{c}_{\text{pre}}^{(i)}[\hat{n}-1] \in \mathbb{C}$ denotes the $(\hat{n}-1)$ th CIR-related tap prediction-filter coefficient. Note that, for simplicity, here we employ the same coefficient $\tilde{c}_{\text{pre}}^{(i)}[\hat{n}-1]$ for filtering each of the different K_0 number of taps of the specific \hat{n} th CIR-related vector $\tilde{\mathbf{h}}_{\text{apt}}^{(i)}[n-\hat{n}]$, which follows the concepts of robust channel estimation advocated by Li *et al.* [4].

- 3) In the last step, the vector of CIR-related *a priori* tap estimates $\hat{\mathbf{h}}_{\text{apr}}^{(i)}[n]$ is transformed back to the OFDM frequency domain with the aid of the unitary KLT matrix $\tilde{\mathbf{U}}^{[f](i)}$, yielding the vector of *a priori* channel transfer-factor estimates $\hat{\mathbf{H}}_{\text{apr}}^{(i)}[n]$ for the n th OFDM-symbol period:

$$\hat{\mathbf{H}}_{\text{apr}}^{(i)}[n] = \tilde{\mathbf{U}}^{[f](i)} \hat{\mathbf{h}}_{\text{apr}}^{(i)}[n]. \quad (9)$$

This vector of *a priori* channel transfer-factor estimates is in turn employed in the detection stage during the n th OFDM-symbol period. Upon substituting (7) into (8), and by substituting the result into (9), we obtain the following relation between the vector of *a priori* channel transfer-factor estimates derived for the n th OFDM symbol and the vectors of *a posteriori* channel transfer-factor estimates of the past $N_{\text{tap}}^{[t]}$ number of OFDM symbols:

$$\hat{\mathbf{H}}_{\text{apr}}^{(i)}[n] = \mathbf{T}_{K_0}^{(i)} \sum_{\hat{n}=1}^{N_{\text{tap}}^{[t]}} \tilde{c}_{\text{pre}}^{(i)}[\hat{n}-1] \tilde{\mathbf{H}}_{\text{apt}}^{(i)}[n-\hat{n}] \quad (10)$$

TABLE I

SUMMARY OF PROCESSING STEPS ASSOCIATED WITH PIC-ASSISTED DDCE FOR MULTIPLE TRANSMIT ANTENNAS DURING THE n TH OFDM-SYMBOL PERIOD. THE LINEAR FILTERING ALONG THE TIME DIRECTION FOLLOWS THE PRINCIPLES OF A ROBUST DDCE FOR SINGLE-USER OFDM AS PROPOSED, e.g., BY LI *et al.* [9]. THE CALCULATION OF THE FILTER COEFFICIENTS $\tilde{c}_{\text{pre}}^{(i)}[\hat{n} - 1]$, $\hat{n} = 1, \dots, N_{\text{tap}}^{[t]}$ IS CONDUCTED WITH THE AID OF THE ITERATIVE APPROACH DEvised IN SECTION III-B.2, WHICH HAS BEEN SUMMARIZED AGAIN, IN TABLE II

Description	Operation
PIC-Step	$\tilde{\mathbf{H}}_{\text{apt}}^{(j)}[n] = \check{\mathbf{S}}^{(j)-1}[n] \left(\mathbf{x}[n] - \sum_{i=1, i \neq j}^L \check{\mathbf{S}}^{(i)}[n] \hat{\mathbf{H}}_{\text{apr}}^{(i)}[n] \right)$, $j = 1, \dots, L$ of Eq. 3
Filtering	$\hat{\mathbf{H}}_{\text{apr}}^{(i)}[n] = \mathbf{T}_{K_0}^{(i)} \sum_{\hat{n}=1}^{N_{\text{tap}}^{[t]}} \tilde{c}_{\text{pre}}^{(i)}[\hat{n} - 1] \tilde{\mathbf{H}}_{\text{apt}}^{(i)}[n - \hat{n}]$, $i = 1, \dots, L$ of Eq. 10 with $\mathbf{T}_{K_0}^{(i)} = \tilde{\mathbf{U}}^{[f](i)} \mathbf{I}_{K_0}^{(i)} \tilde{\mathbf{U}}^{[f](i)H}$ of Eq. 11

where $\mathbf{T}_{K_0, p}^{(i)} \in \mathbb{C}^{K \times K}$ is given by

$$\mathbf{T}_{K_0}^{(i)} = \tilde{\mathbf{U}}^{[f](i)} \mathbf{I}_{K_0}^{(i)} \tilde{\mathbf{U}}^{[f](i)H}. \quad (11)$$

Again, the different steps of the PIC-assisted DDCE have been illustrated in Fig. 3, as well as in Table I.

After having described the process of generating the vectors of *a posteriori* and *a priori* channel transfer-factor estimates in Sections III-A1 and III-A2, we will embark in Section III-B1 on an analytical evaluation of the associated *a priori* estimation MSE.

B. Derivation of an Iterative Approach for Determining the Set of Optimum Predictor Coefficients

Our discussions commence in Section III-B1 with the derivation of an expression for the average *a priori* channel-estimation MSE as a function of the corresponding estimation MSEs associated with the previous $N_{\text{tap}}^{[t]}$ number of OFDM symbols. This expression is then employed in Section III-B2—under the assumption that the system is in its steady-state condition—for generating the different users' vectors of optimum predictor coefficients, again, as a function of the predictor-coefficient-dependent *a priori* estimation MSEs. Since the recursive structure of the channel transfer-function estimator does not allow for an algebraic solution to be generated for the desired predictor coefficients, an iterative approach is applied, which exploits the contractive properties of the system equations. This approach was proposed earlier by Rashid-Farrokhi *et al.* [20] in the context of simultaneously optimizing the transmit power allocation and base-station (BS) antenna-array weights in wireless networks.

1) *A Priori Channel-Estimation MSE*: Let us commence our discussions in this section by developing an expression for the vector of *a priori* channel transfer-factor estimation errors associated with the j th user during the n th OFDM-symbol period as a function of the vectors of *a priori* channel transfer-factor estimation errors of the $(L - 1)$ remaining users during the $N_{\text{tap}}^{[t]}$ number of previous OFDM-symbol periods. Assuming error-free symbol decisions, we have $\check{\mathbf{S}}^{(j)}[n - \hat{n}] = \mathbf{S}^{(j)}[n - \hat{n}]$, $j = 1, \dots, L$, $\hat{n} = 1, \dots, N_{\text{tap}}^{[t]}$. Substituting (2) into (3), and then substituting the result into (10), yields an expression for the vector of channel transfer-

factor estimation errors $\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}[n] \in \mathbb{C}^{K \times 1}$ in the following form:

$$\begin{aligned} \Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}[n] = & - \mathbf{T}_{K_0}^{(j)} \sum_{\hat{n}=1}^{N_{\text{tap}}^{[t]}} \tilde{c}_{\text{pre}}^{(j)}[\hat{n} - 1] \mathbf{S}^{(j)-1}[n - \hat{n}] \\ & \times \sum_{i=1, i \neq j}^L \mathbf{S}^{(i)}[n - \hat{n}] \Delta \hat{\mathbf{H}}_{\text{apr}}^{(i)}[n - \hat{n}] \\ & - \mathbf{T}_{K_0}^{(j)} \sum_{\hat{n}=1}^{N_{\text{tap}}^{[t]}} \tilde{c}_{\text{pre}}^{(j)}[\hat{n} - 1] \mathbf{S}^{(j)-1}[n - \hat{n}] \mathbf{n}[n - \hat{n}] \\ & + \mathbf{H}^{(j)}[n] - \mathbf{T}_{K_0}^{(j)} \sum_{\hat{n}=1}^{N_{\text{tap}}^{[t]}} \tilde{c}_{\text{pre}}^{(j)}[\hat{n} - 1] \mathbf{H}^{(j)}[n - \hat{n}] \end{aligned} \quad (12)$$

where

$$\Delta \hat{\mathbf{H}}_{\text{apr}}^j[n] = \mathbf{H}^j[n] - \hat{\mathbf{H}}_{\text{apr}}^j[n]. \quad (13)$$

Please observe that, for the sake of avoiding notational confusion, the variable i of (10) has been substituted by the variable j . The vector of *a priori* channel transfer-factor estimation errors given by (12) is constituted by three components. Specifically, the first term of (12) is due to the effects of the *a priori* prediction errors of the $N_{\text{tap}}^{[t]}$ number of past OFDM symbols, the second term is attributed to the contaminating effect of the AWGN, and the third term is due to the lack of “perfect predictability” of the channel transfer factors by the $N_{\text{tap}}^{[t]}$ -order predictor. In other words, the last term is due to the channel transfer function's decorrelation with time.

The average variance of the j th user's vector of *a priori* channel transfer-factor estimation errors, or equivalently, the average mean-square *a priori* estimation error, can be expressed in mathematical terms as

$$\overline{\text{MSE}}_{\text{apr}}^{(j)}[n] = \frac{1}{K} \text{Trace} \left(\mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n] \right) \quad (14)$$

where $\mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n] \in \mathbb{C}^{K \times K}$ denotes the autocorrelation matrix of the vector $\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}[n]$ of *a priori* channel transfer-factor estimation errors. The computation of $\overline{\text{MSE}}_{\text{apr}}^{(j)}[n]$ of the j th user's vector of *a priori* estimation errors associated with the n th OFDM-symbol period, as given by (12), will be carried out in two steps.

In the first step, let us evaluate the autocorrelation matrix $\mathbf{R}_{\Delta\hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n]$. This is achieved by substituting (12) into

$$\begin{aligned} & \mathbf{R}_{\Delta\hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n] \\ &= E \left\{ \Delta\hat{\mathbf{H}}_{\text{apr}}^{(j)}[n] \Delta\hat{\mathbf{H}}_{\text{apr}}^{(j)\text{H}}[n] \right\} \\ &= \frac{\alpha_j}{\sigma_j^2} \mathbf{T}_{K_0}^{(j)} \left(\sum_{\hat{n}=1}^{N_{\text{tap}}^{[t]}} \left| \tilde{\mathbf{c}}_{\text{pre}}^{(j)}[\hat{n}-1] \right|^2 \sum_{\substack{i=1 \\ i \neq j}}^L \sigma_i^2 \mathbf{R}_{\Delta\hat{\mathbf{H}}_{\text{apr}}^{(i)}}[n-\hat{n}]_{\text{Diag}} \right) \\ & \quad \times \mathbf{T}_{K_0}^{(j)\text{H}} + \frac{\alpha_j}{\sigma_j^2} \sigma_n^2 \sum_{\hat{n}=1}^{N_{\text{tap}}^{[t]}} \left| \tilde{\mathbf{c}}_{\text{pre}}^{(j)}[\hat{n}-1] \right|^2 \mathbf{T}_{K_0}^{(j)} \mathbf{T}_{K_0}^{(j)\text{H}} + \mathbf{R}_{\mathbf{H}_{\text{dec}}^{(j)}} \end{aligned} \quad (15)$$

where we introduced a new definition, namely, that of the channel transfer-function decorrelation-related matrix $\mathbf{R}_{\mathbf{H}_{\text{dec}}^{(j)}} \in \mathbb{C}^{K \times K}$, which is given by

$$\begin{aligned} \mathbf{R}_{\mathbf{H}_{\text{dec}}^{(j)}} &= \mathbf{R}^{[f](j)} - \mathbf{T}_{K_0}^{(j)} \mathbf{R}^{[f](j)\text{H}} \cdot \left(\tilde{\mathbf{c}}_{\text{pre}}^{(j)\text{H}} \mathbf{R}^{[t](j)} \right) \\ & \quad - \mathbf{R}^{[f](j)} \mathbf{T}_{K_0}^{(j)\text{H}} \cdot \left(\tilde{\mathbf{c}}_{\text{pre}}^{(j)T} \mathbf{R}^{[t](j)*} \right) \\ & \quad + \mathbf{T}_{K_0}^{(j)} \mathbf{R}^{[f](j)} \mathbf{T}_{K_0}^{(j)\text{H}} \cdot \left(\tilde{\mathbf{c}}_{\text{pre}}^{(j)\text{H}} \mathbf{R}^{[t](j)} \tilde{\mathbf{c}}_{\text{pre}}^{(j)} \right). \end{aligned} \quad (17)$$

In the context of (16), we have exploited the fact that the three additive components of the vector $\Delta\hat{\mathbf{H}}_{\text{apr}}^{(j)}[n]$ of *a priori* channel transfer-factor estimation errors in (12) are uncorrelated. The uncorrelated nature of these three terms accrues from the statistical independence of the complex AWGN process and that of the complex valued process describing the channel transfer function's evolution versus frequency and time. We have also exploited the fact that the complex symbols transmitted in different subcarriers of a specific user's signal during a specific OFDM-symbol period, as well as the symbols transmitted by the same user in different OFDM-symbol periods and the symbols transmitted by different users, are statistically independent, which also implies that they are uncorrelated. Still considering (16), the variable α_j denotes the so-called "modulation noise-enhancement factor" [3], [21] defined as $\alpha_j = E\{|s^{(j)}[n, k]|^2\} E\{|1/s^{(j)}[n, k]|^2\}$. For M -ary phase shift keying (MPSK)-based modulation schemes, such as quaternary PSK (QPSK), we have $\alpha = 1$, while for higher order quadrature amplitude modulation (QAM) schemes, we have $\alpha > 1$ [3], [21]. Note that here we have implicitly assumed that the same modulation scheme is employed on different subcarriers of a specific user's transmitted signal. To elaborate further, the variables to be defined in (17) are the space-time correlation-function-related autocorrelation vector $\mathbf{r}^{[t](j)} \in \mathbb{C}^{N_{\text{tap}}^{[t]} \times 1}$ of the channel transfer function, where the \hat{n} th element is given by $\mathbf{r}^{[t](j)}|_{\hat{n}} = E\{H^{(j)*}[n, k] H^{(j)}[n-\hat{n}, k]\}$, and the space-time correlation-function-related autocorrelation matrix $\mathbf{R}^{[t](j)} \in \mathbb{C}^{N_{\text{tap}}^{[t]} \times N_{\text{tap}}^{[t]}}$ of the channel transfer function, with the element (\hat{n}_1, \hat{n}_2) given by $\mathbf{R}^{[t](j)}|_{\hat{n}_1, \hat{n}_2} = E\{H^{(j)}[n-\hat{n}_1, k] H^{(j)*}[n-\hat{n}_2, k]\}$. Furthermore, $\tilde{\mathbf{c}}_{\text{pre}}^{(j)} \in \mathbb{C}^{N_{\text{tap}}^{[t]} \times 1}$ is the vector of conjugate complex CIR-related tap prediction-filter

coefficients with its \hat{n} th element given by $\tilde{\mathbf{c}}_{\text{pre}}^{(j)}|_{\hat{n}} = \tilde{\mathbf{c}}_{\text{pre}}^{(j)*}[\hat{n}]$. The channel's space-frequency correlation matrix $\mathbf{R}^{[f](j)}$ was defined earlier in Section III-A.2. Let us now return to our original objective, namely, that of developing an expression for the average *a priori* channel transfer-factor estimation MSE during the n th OFDM-symbol period.

In the second step, (16) is invoked in conjunction with (14) for obtaining an expression for the j th user's average *a priori* channel transfer-factor estimation MSE as a function of the remaining users' *a priori* estimation MSEs associated with the $N_{\text{tap}}^{[t]}$ number of previous OFDM-symbol periods:

$$\begin{aligned} \overline{\text{MSE}}_{\text{apr}}^{(j)}[n] &= \frac{K_0}{K} \frac{\alpha_j}{\sigma_j^2} \sum_{\hat{n}=1}^{N_{\text{tap}}^{[t]}} \left| \tilde{\mathbf{c}}_{\text{pre}}^{(j)}[\hat{n}-1] \right|^2 \sum_{\substack{i=1 \\ i \neq j}}^L \sigma_i^2 \overline{\text{MSE}}_{\text{apr}}^{(i)}[n-\hat{n}] \\ & \quad + \frac{K_0}{K} \frac{\alpha_j}{\sigma_j^2} \left\| \tilde{\mathbf{c}}_{\text{pre}}^{(j)} \right\|_2^2 \sigma_n^2 + \overline{\text{MSE}}_{\text{dec}}^{(j)} \end{aligned} \quad (18)$$

where we have

$$\begin{aligned} \overline{\text{MSE}}_{\text{dec}}^{(j)} &= \frac{1}{K} \text{Trace} \left(\mathbf{R}_{\mathbf{H}_{\text{dec}}^{(j)}} \right) \\ &= \sigma_H^2 - \frac{1}{K} \text{Trace} \left(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)} \right) \\ & \quad \times \left(2\Re \left\{ \tilde{\mathbf{c}}_{\text{pre}}^{(j)\text{H}} \mathbf{r}^{[t](j)} \right\} - \tilde{\mathbf{c}}_{\text{pre}}^{(j)\text{H}} \mathbf{R}^{[t](j)} \tilde{\mathbf{c}}_{\text{pre}}^{(j)} \right). \end{aligned} \quad (20)$$

In the context of deriving (18), we have capitalized on the relations $\text{Trace}(\mathbf{A} + \mathbf{B}) = \text{Trace}(\mathbf{A}) + \text{Trace}(\mathbf{B})$, as well as on $\text{Trace}(\mathbf{U}\mathbf{A}\mathbf{U}^{\text{H}}) = \text{Trace}(\mathbf{A})$, which are valid for a unitary matrix \mathbf{U} [18], [22]. Furthermore, in the context of deriving the first additive term in (18), we exploited the fact that $(1/K)\text{Trace}(\mathbf{T}_{K_0}^{(j)} \mathbf{R}_{\Delta\hat{\mathbf{H}}_{\text{apr}}^{(i)}}[n-\hat{n}]_{\text{Diag}} \mathbf{T}_{K_0}^{(j)\text{H}}) = (K_0/K)\overline{\text{MSE}}_{\text{apr}}^{(i)}[n-\hat{n}]$, which is only valid for a unitary transform matrix $\tilde{\mathbf{U}}^{[f](j)}$ having elements of unity magnitude. This is the case, for example, when employing the discrete FT (DFT) matrix \mathbf{W} as the unitary transform matrix. The second additive term in (18) is based on exploiting the relationship of $(1/K)\text{Trace}(\mathbf{T}_{K_0}^{(j)} \mathbf{T}_{K_0}^{(j)\text{H}}) = K_0/K$. We also note in this context that $\mathbf{T}_{K_0}^{(j)\text{H}} = \mathbf{T}_{K_0}^{(j)}$, and that $\mathbf{T}_{K_0}^{(j)} \mathbf{T}_{K_0}^{(j)\text{H}} = \mathbf{T}_{K_0}^{(j)}$. Furthermore, in (20), the matrix $\Upsilon^{[f](j)} \in \mathbb{C}^{K \times K}$ denotes the decomposition of the j th user's channel's space-frequency correlation matrix $\mathbf{R}^{[f](j)}$ with respect to the unitary transform matrix $\tilde{\mathbf{U}}^{[f](j)}$, which is expressed as $\Upsilon^{[f](j)} = \tilde{\mathbf{U}}^{[f](j)\text{H}} \mathbf{R}^{[f](j)} \tilde{\mathbf{U}}^{[f](j)}$. Note that in contrast to $\Lambda^{[f](j)}$ associated with the decomposition of $\mathbf{R}^{[f](j)}$ with respect to $\mathbf{U}^{[f](j)}$, the matrix $\Upsilon^{[f](j)}$ is not necessarily of diagonal shape, constrained to having real-valued elements only.

2) *Iterative Calculation of the CIR-Related Tap Predictor Coefficients*: In the steady-state condition, we can assume that the specific user's *a priori* and *a posteriori* estimation MSEs are identical for different OFDM symbols, which is expressed as

$$\overline{\text{MSE}}_{\text{apr}}^{(i)} = \overline{\text{MSE}}_{\text{apr}}^{(i)}[n-\hat{n}] \quad (21)$$

where $i = 1, \dots, L$ and $\hat{n} = 0, \dots, N_{\text{tap}}^{[t]}$. Hence, (18) simplifies to

$$\overline{\text{MSE}}_{\text{apr}}^{(j)} = \frac{K_0}{K} \frac{\alpha_j}{\sigma_j^2} \left\| \tilde{\mathbf{c}}_{\text{pre}}^{(j)} \right\|_2^2 \times \left(\sum_{\substack{i=1 \\ i \neq j}}^L \sigma_i^2 \overline{\text{MSE}}_{\text{apr}}^{(i)} + \sigma_n^2 \right) + \overline{\text{MSE}}_{\text{dec}}^{(j)}. \quad (22)$$

Upon invoking (22), the j th user's vector of CIR-related tap predictor coefficients $\tilde{\mathbf{c}}_{\text{pre}}^{(j)}$ can be evaluated conditioned on the remaining $(L-1)$ number of users' *a priori* estimation MSEs, namely on $\overline{\text{MSE}}_{\text{apr}}^{(i)}$, $i = 1, \dots, L, i \neq j$, which ensues by calculating the gradient of $\overline{\text{MSE}}_{\text{apr}}^{(j)}$ with respect to the j th user's coefficients, yielding

$$\nabla^{(j)} \overline{\text{MSE}}_{\text{apr}}^{(j)} = \frac{K_0}{K} \frac{\alpha_j}{\sigma_j^2} \tilde{\mathbf{c}}_{\text{pre}}^{(j)} \left(\sum_{\substack{i=1 \\ i \neq j}}^L \sigma_i^2 \overline{\text{MSE}}_{\text{apr}}^{(i)} + \sigma_n^2 \right) - \frac{1}{K} \text{Trace} \left(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)} \right) \times \left(\mathbf{r}^{[t](j)} - \mathbf{R}^{[t](j)} \tilde{\mathbf{c}}_{\text{pre}}^{(j)} \right) \quad (23)$$

where $\mathbf{R}^{[t](j)}$ and $\mathbf{r}^{[t](j)}$ were defined in the context of (17). The gradient vector, with respect to the j th user's coefficients, is defined here as $\nabla^{(j)} = \partial / (\partial \tilde{\mathbf{c}}_{\text{pre}}^{(j)*})$. In the context of (23), we have exploited the fact that $\nabla^{(j)} \tilde{\mathbf{c}}_{\text{pre}}^{(j)\text{H}} = \mathbf{I}$, as well as the fact that $\nabla^{(j)} \tilde{\mathbf{c}}_{\text{pre}}^{(j)\text{T}} = \mathbf{0}$ and $\nabla^{(j)} (\|\tilde{\mathbf{c}}_{\text{pre}}^{(j)}\|_2^2) = \tilde{\mathbf{c}}_{\text{pre}}^{(j)}$ [19].

In the optimum point of operation, we have $\nabla^{(j)} \overline{\text{MSE}}_{\text{apr}}^{(j)} = \mathbf{0}$ and hence, (23) can be solved for the j th user's vector of predictor coefficients, resulting in the Wiener-filter-related solution of⁵

$$\tilde{\mathbf{c}}_{\text{pre}}^{(j)}|_{\text{opt}} = \left[\mathbf{R}^{[t](j)} + \frac{K_0}{\text{Trace} \left(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)} \right)} \frac{\alpha_j}{\sigma_j^2} \times \left(\sum_{\substack{i=1 \\ i \neq j}}^L \sigma_i^2 \overline{\text{MSE}}_{\text{apr}}^{(i)} + \sigma_n^2 \right) \mathbf{I} \right]^{-1} \cdot \mathbf{r}^{[t](j)}. \quad (24)$$

Based on (22) and (24), a fixed-point iteration algorithm [19] can be devised for obtaining the different users' vectors of

predictor coefficients under the constraint of minimizing the sum of the different users' *a priori* estimation MSEs. This approach was proposed earlier by Rashid-Farrokh *et al.* [20] in the context of simultaneously optimizing both the transmit power allocation and the base-station antenna-array weights in wireless networks, leading to formulas similar to (22) and (24). In our forthcoming discourse, we will briefly present the steps of the algorithm with respect to our specific optimization problem, but for a formal proof of the algorithm's convergence and that of the uniqueness of the solution, we refer to [20]. Note that, in the context of our description of the algorithm, the iteration index—and not the OFDM-symbol index—is given in the square brackets.

- 1) Initialize the different users' *a priori* estimation MSEs, for example, by setting $\overline{\text{MSE}}_{\text{apr}}^{(j)}[0] = 0$ for $j = 1, \dots, L$.
- 2) For the n th iteration: Conditioned on the *a priori* estimation MSE values obtained during the $(n-1)$ th iteration, namely, on $\overline{\text{MSE}}_{\text{apr}}^{(j)}[n-1]$, $j = 1, \dots, L$, calculate the different users' vectors of optimum predictor coefficients for the n th iteration, namely, $\tilde{\mathbf{c}}_{\text{pre}}^{(j)}[n]|_{\text{opt}}$, $j = 1, \dots, L$, with the aid of (24).
- 3) Conditioned on the n th iteration's predictor-coefficient vectors $\tilde{\mathbf{c}}_{\text{pre}}^{(j)}[n]|_{\text{opt}}$, $j = 1, \dots, L$, obtained in step 2) and also conditioned on the $(n-1)$ th iteration's *a priori* estimation MSE values, namely on $\overline{\text{MSE}}_{\text{apr}}^{(j)}[n-1]$, $j = 1, \dots, L$, calculate the n th iteration's *a priori* estimation MSE values of $\overline{\text{MSE}}_{\text{apr}}^{(j)}[n]$, $j = 1, \dots, L$, with the aid of (22).⁶
- 4) Start a new iteration by returning to step 2).

The essential equations of this optimization procedure have been summarized again in Table II. Note that, instead of invoking (22) separately for each user, the different users' *a priori* estimation MSEs can also be calculated in parallel with the aid of (33), as a result of which, an even faster convergence is achieved. The price to be paid is a higher computational complexity, since an explicit matrix inversion is required in (33).

In the next section, we will address the problem of a potential lack of knowledge about the channel's exact statistics, namely, that of the space-time space-frequency correlation function.

C. Channel Statistics

As it was observed in (22) and (24), a prerequisite for determining the different users' vectors of optimum CIR-related tap predictor coefficients is the knowledge of the users' space-time

⁵Note that in the context of identical transmit powers, modulation modes, and channel statistics, (24) is significantly simplified, namely, we obtain the same vector of predictor coefficients given by $\tilde{\mathbf{c}}_{\text{pre}}|_{\text{opt}}^{\text{SIMPLE}} = [\mathbf{R}^{[t]} + (K_0/\text{Trace}(\Upsilon^{[f]} \mathbf{I}_{K_0})) \alpha((L-1)\overline{\text{MSE}}_{\text{apr}}^{\text{SIMPLE}} + (\sigma_n^2/\sigma_s^2))\mathbf{I}]^{-1} \cdot \mathbf{r}^{[t]}$, as well as the same average estimation MSE for the different users. Furthermore, note that upon removing the $(L-1)$ number of contributions, which are related to the PIC process, we then obtain the expressions for the estimation MSE and the vector of coefficients associated with a transversal predictor, which can be expressed as $\tilde{\mathbf{c}}_{\text{pre},\text{FIR}}|_{\text{opt}}^{\text{SIMPLE}} = [\mathbf{R}^{[t]} + (K_0/\text{Trace}(\Upsilon^{[f]} \mathbf{I}_{K_0}))\alpha(\sigma_n^2/\sigma_s^2)\mathbf{I}]^{-1} \cdot \mathbf{r}^{[t]}$.

⁶Recall that the representation of the j th user's estimation MSE was valid only when assuming that the unitary transform matrix $\tilde{\mathbf{U}}^{[f]}$ has elements of unity magnitude, which is the case, for example, for the DFT matrix \mathbf{W} . However, in the more general scenario of employing an arbitrary unitary transform matrix, which could be, for example, one of the robust transforms proposed in [23], this condition is not fulfilled. In this case, the matrix $\mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n]$ has to be explicitly iterated with the aid of (16), upon assuming $\mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n] = \mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n - \hat{n}]$, followed by the application of (14), in order to obtain the estimation MSE required by step 2).

TABLE II
SUMMARY OF PROCESSING STEPS ASSOCIATED WITH THE OPTIMIZATION OF THE PIC-ASSISTED DDCE'S PREDICTOR COEFFICIENTS

Description	Operation
0-th iter.	Init. $\overline{\text{MSE}}_{\text{appr}}^{(j)}[0] = 0$ for $j = 1, \dots, L$
n -th iter.	1.) $\tilde{\mathbf{c}}_{\text{pre}}^{(j)}[n] = \left[\mathbf{R}^{[t](j)} + \frac{K_0}{\text{Trace}(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)})} \frac{\alpha_j}{\sigma_j^2} \left(\sum_{i=1, i \neq j}^L \sigma_i^2 \overline{\text{MSE}}_{\text{appr}}^{(i)}[n-1] + \sigma_n^2 \right) \mathbf{I} \right]^{-1} \cdot \mathbf{r}^{[t](j)}$ 2.) $\overline{\text{MSE}}_{\text{appr}}^{(j)}[n] = \frac{K_0}{K} \frac{\alpha_j}{\sigma_j^2} \ \tilde{\mathbf{c}}_{\text{pre}}^{(j)}[n]\ _2^2 \left(\sum_{i=1, i \neq j}^L \sigma_i^2 \overline{\text{MSE}}_{\text{appr}}^{(i)}[n-1] + \sigma_n^2 \right) + \overline{\text{MSE}}_{\text{dec}}^{(j)}$ with $\overline{\text{MSE}}_{\text{dec}}^{(j)} = \sigma_H^2 - \frac{1}{K} \text{Trace}(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)}) \cdot \left(2\Re\{\tilde{\mathbf{c}}_{\text{pre}}^{(j)H} \mathbf{r}^{[t](j)}\} - \tilde{\mathbf{c}}_{\text{pre}}^{(j)H} \mathbf{R}^{[t](j)} \tilde{\mathbf{c}}_{\text{pre}}^{(j)} \right)$ for $j = 1, \dots, L$ of Eq. 24, 22 and 20, respectively.

channel transfer-factor correlation functions $r_H^{[t](j)}[\Delta n]$, $j = 1, \dots, L$, defined by

$$r_H^{[t](j)}[\Delta n] = E \left\{ H^{(j)}[n, k] \cdot H^{(j)*}[n - \Delta n, k] \right\}. \quad (25)$$

These are required for evaluating the autocorrelation matrices $\mathbf{R}^{[t](j)}$ and cross-correlation vectors $\mathbf{r}^{[t](j)}$, for $j = 1, \dots, L$. Assuming Jakes' fading model [24] for example, the channel correlation along the time direction is given by [4]

$$r_{H,J}^{[t](j)}[\Delta n] = J_0 \left(\Delta n \cdot \omega_D^{(j)} \right) \quad (26)$$

$$\approx 1 - \frac{1}{4} \left(\Delta n \cdot \omega_D^{(j)} \right)^2, \quad \Delta n \cdot \omega_D^{(j)} \ll 1 \quad (27)$$

where $J_0(\cdot)$ denotes the zero-order Bessel function of the first kind and $\omega_D^{(j)} = 2\pi T_f f_D^{(j)}$, and T_f being the OFDM-symbol duration including the guard period time, while $f_D^{(j)}$ denotes the channel's Doppler frequency. Since usually the exact Doppler frequency $f_D^{(j)}$ is not known, it was demonstrated in [4], in the context of a transversal-type estimator, that the MSE performance degradation incurred due to a mismatch of the channel statistics is only marginal, if a uniform ideally support-limited Doppler power spectrum associated with $\tilde{f}_D^{(j)} \geq f_D^{(j)}$ is assumed for the calculation of the correlation coefficients of (25). The associated space-time correlation function is given as the inverse FT of the uniform Doppler power spectrum, which leads to

$$\tilde{r}_{H,\text{unif}}^{[t](j)}[\Delta n] = \frac{\sin \left(\Delta n \cdot \tilde{\omega}_D^{(j)} \right)}{\Delta n \cdot \tilde{\omega}_D^{(j)}}. \quad (28)$$

Furthermore, the calculation of the vectors of CIR-related tap predictor coefficients according to (24) also requires the evaluation of the expression $\text{Trace}(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)})$. More explicitly, we recall from Section III-A2 that $\Upsilon^{[f](j)}$ is the decomposition of the j th user's channel's space-frequency correlation matrix $\mathbf{R}^{[f](j)}$ with respect to the unitary transform matrix $\tilde{\mathbf{U}}^{[f](j)}$, which is formulated as $\Upsilon^{[f](j)} = \tilde{\mathbf{U}}^{[f](j)H} \mathbf{R}^{[f](j)} \tilde{\mathbf{U}}^{[f](j)}$, and $\mathbf{I}_{K_0}^{(j)}$ is a sparse identity matrix having unity entries only at

those K_0 number of positions, which are associated with a significant value of $\Upsilon^{[f](j)}$. Hence, we note that the evaluation of $\text{Trace}(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)})$ requires the knowledge of $\mathbf{R}^{[f](j)}$, which is not directly available in practice.

A viable approach is that of obtaining an "average" value of $\text{Trace}(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)})$ by employing the space-frequency correlation matrix $\tilde{\mathbf{R}}^{[f](j)}$ based on the space-frequency correlation function associated with a uniform ideally support-limited multipath intensity profile.⁷ The sparse identity matrix $\mathbf{I}_{K_0}^{(j)}$ could be designed for retaining the first K_0 CIR-related coefficients of $\Upsilon^{[f](j)}$ —rather than the K_0 largest one—or alternatively, for retaining the first K_0^I and the last K_0^{II} CIR-related coefficients of $\Upsilon^{[f](j)}$, where $K_0 = K_0^I + K_0^{II}$. This was suggested by van de Beek *et al.* [25] in the context of DFT-based channel transfer-function estimation employed designed for single-user OFDM systems.

IV. PERFORMANCE

With the exception of the results to be presented in Section IV-B, our investigations were conducted for an SDMA uplink scenario supporting four simultaneous equal-power OFDM users, each equipped with one transmit antenna. At the BS, four reception antennas were assumed. Furthermore, for the sake of simplicity, the different users were assumed to employ the same modulation scheme, and the channels between the different transmit antennas and each receiver antenna were assumed to have the same Doppler power spectrum.⁸ Unless otherwise stated, the specific channel statistics invoked were that of the channel's space-time correlation function provided by the Jakes model, as given by (26), having an OFDM-symbol-normalized Doppler frequency⁹ of $F_D = 0.007$, which corresponds to a vehicular speed of 50 km/h, or equivalently,

⁷The continuous unit-energy uniform power delay profile is given by $r_{h,\text{unif}}(\tau) = (1/T_m) \text{rect}(\tau - \tau_{\text{shift}}/T_m)$, while its FT, namely the space-frequency correlation function, is given by $r_{H,\text{unif}}(\Delta f) = \sin(\pi T_m \Delta f) \cdot e^{-j2\pi \tau_{\text{shift}} \Delta f}$.

⁸Recall that in this scenario, (24) and (33), associated with the iterative optimization of the predictor coefficients, as proposed in Section III-B2, are significantly simplified.

⁹The OFDM-symbol-normalized Doppler frequency is defined as $F_D = f_D T_f$, where f_D is the Doppler frequency and T_f denotes the OFDM symbol duration.

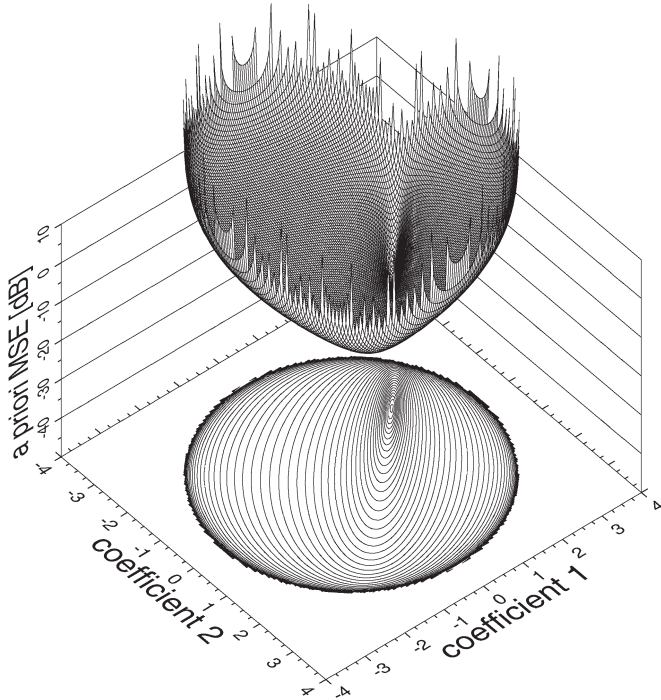


Fig. 4. Evolution of the *a priori* channel-estimation MSE according to (33), as a function of the real-valued coefficients of the two-tap CIR-related tap predictor employed in this particular example. The number of subcarriers was $K = 512$, while the number of significant CIR-related taps was $K_0 = 16$, in the context of a sample-spaced CIR. Furthermore, the number of users was $L = 4$ and the OFDM-symbol-normalized Doppler frequency was $F_D = 0.1$. The space-time channel correlation function of (28), associated with a uniform ideally support-limited Doppler power spectrum, was invoked. The SNR at the reception antenna was equal to 20 dB.

31.25 mi/h in the context of the indoor wireless asynchronous transfer mode (WATM) system’s parameters [1], [26] invoked here.¹⁰ Furthermore, we considered “frame-invariant” fading, where the fading envelope of each CIR-related tap has been kept constant during each OFDM symbol’s transmission period. This avoided the obfuscating effects of intersubcarrier interference and hence, enabled us to study the various channel transfer-function-estimation effects in isolation. Furthermore, apart from our investigations in Section IV-G, error-free symbol decisions were assumed in the generation of the remodulated reference used for DDCE.

A. Evolution of the A Priori Channel Estimation MSE in a 2-Tap CIR-Related Tap Prediction Scenario

In Fig. 4, we have exemplified the evolution of the average *a priori* channel transfer-factor estimation MSE according to (33), as a function of the CIR-related tap predictor coefficients’ associated values, where we employed a two-tap predictor, since for a higher number of predictor taps, a visualization

¹⁰Note that associated with the indoor WATM channel is a sample-spaced CIR having a multipath spread of $T_m = 12T_s$. In the context of our simulations, which employed the indoor WATM channel’s CIR, the matrix $\mathbf{I}_{K_0}^{(j)}$ was designed such as to retain the first K_0 number of CIR-related taps. Hence, in case that $K_0 > 12$, we have $\text{Trace}(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)}) = K \sigma_H^2$.

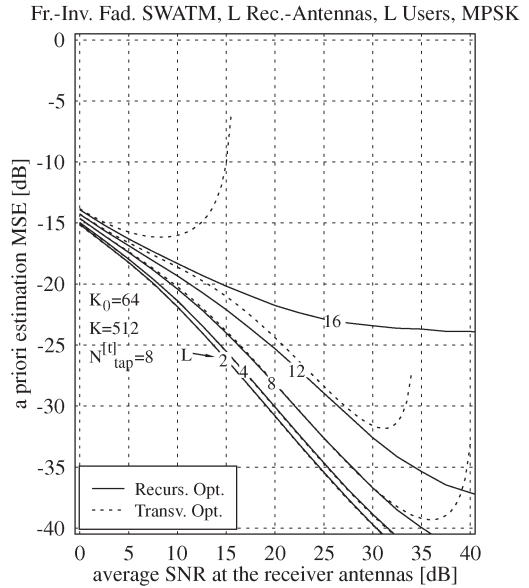


Fig. 5. *A priori* channel-estimation MSE versus SNR performance exhibited by the PIC-assisted DDCE of Fig. 3, using the optimum recursive predictor coefficients evaluated with the aid of the iterative approach of Section III-B2. As a benchmarker, we have plotted the *a priori* channel-estimation MSE performance achieved with the aid of the suboptimum transversal predictor coefficients. Each of the SDMA scenario’s independently faded channels is characterized by the indoor WATM channel parameters of [1] and [26].

is less convenient. Also, note that the predictor coefficients are real valued due to the employment of the real-valued space-time channel correlation function of (26). In our particular example, the *a priori* channel-estimation MSE evaluated from (33) is minimized for a coefficient vector of $\tilde{\mathbf{c}}_{\text{pre}}|_{\text{opt}} \approx (1.771, -0.898)^T$. By contrast, for coefficient pairs outside the circle having a radius of $\sqrt{K/K_0 \alpha(L-1)} \approx 3.27$, centered around the origin of the \mathbb{R}^2 space, the channel estimator is unstable, which is evidenced by an excessive MSE.

B. Influence of the Number of Simultaneous Users on the A Priori Channel-Estimation MSE

In this section, we will demonstrate that the PIC-assisted approach advocated here is capable of supporting scenarios of a higher complexity in terms of the L number of simultaneous users, than a maximum of $L_{\text{max}} = K/K_0$, as supported by the LS-assisted DDCE of [9]. Hence, in Fig. 5, we have plotted the average *a priori* channel transfer-factor estimation MSE of the PIC-assisted DDCE as a function of the L number of simultaneous users, assuming an eight-tap CIR-related tap prediction filter and a fixed number of $K_0 = 64$ significant CIR-related taps. This corresponds to 12.5% of the duration of a 512-subcarrier OFDM symbol’s time-domain representation, which may be viewed as the relative upper bound of the CIR length in a well-designed OFDM system. Here, we capitalized again on the idealistic assumption of encountering error-free symbol decisions. We observe in Fig. 5 that the *a priori* channel-estimation MSE performance is degraded upon increasing the number of users supported. This is

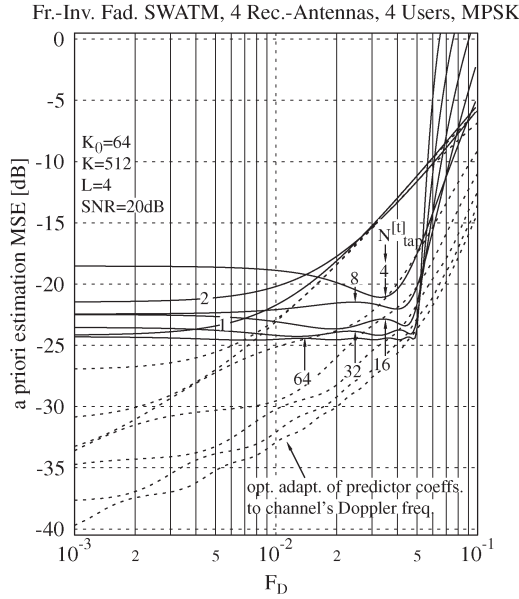


Fig. 6. *A priori* channel-estimation MSE versus OFDM-symbol-normalized Doppler-frequency performance exhibited by the PIC-assisted DDCE of Fig. 3, using optimum recursive predictor coefficients. The predictor coefficients were optimized for $\tilde{F}_D = 0.05$, using the iterative approach of Section III-B2. As in previous graphs, a Jakes's spectrum-related space-time correlation function obeying (26) was associated with the channel. Each of the SDMA scenario's independently faded channels is characterized by the indoor WATM channel parameters of [1] and [26].

because more multiuser interference-related noise is inflicted by the *a posteriori* channel estimates during the PIC process, which is then injected into the *a priori* channel estimates' prediction process. However, these effects can be mitigated by increasing the predictor's range. Again, we observe that, in the context of the suboptimum transversal predictor coefficients of footnote 5, the PIC-assisted DDCE tends to become unstable at higher signal-to-noise ratios (SNRs).

C. Influence of a Mismatch of the OFDM-Symbol-Normalized Doppler Frequency

In Fig. 6, we have portrayed the average *a priori* channel transfer-factor estimation MSE versus OFDM-symbol-normalized Doppler-frequency performance of the recursive estimator of Fig. 3, in the context of employing a uniform ideally support-limited Doppler power spectrum having a space-time correlation function obeying (28) in the calculation of the CIR-related tap predictor coefficients. Furthermore, a Doppler power spectrum obeying Jakes's model [24] and having a space-time correlation function given by (26) was associated with the channel. In our particular example, the predictor coefficients were calculated upon invoking once again the iterative approach of Section III-B2 for an OFDM-symbol-normalized Doppler frequency of $\tilde{F}_D = 0.05$. Furthermore, as a reference, we have also plotted the *a priori* channel-estimation MSE performance in the context of predictor coefficients, which were optimized for the channel's specific Doppler frequency. As reported in [4] and also observed in Fig. 6, upon increasing the number of predictor taps, the *a priori* channel-estimation MSE is ren-

dered quasi-invariant for OFDM-symbol-normalized Doppler frequencies, which are lower than that assumed in the calculation of the CIR-related tap predictor coefficients, namely $\tilde{F}_D = 0.05$. By contrast, for higher Doppler frequencies, a rapid degradation of the MSE is observed in Fig. 6. This "robustness" is achieved at the cost of a potentially significant loss in performance compared to the case of optimally adapted predictor coefficients. To give an example, it is seen in Fig. 6 that for an SNR of 20 dB and for 64 predictor coefficients, the *a priori* channel-estimation MSE performance loss is as high as 10 dB at an OFDM-symbol-normalized Doppler frequency of $\tilde{F}_D = 0.007$, when the predictor coefficients were designed for $\tilde{F}_D = 0.05$.

D. Effects of Correlated Domain Leakage Upon Assuming a Uniform Nonsample-Spaced CIR¹¹

Our analytical evaluations in the previous sections were conducted so far under the assumption of a sample-spaced CIR. Using a sample-spaced CIR facilitates the recovery of almost all the energy of the channel's output, upon invoking a finite number of $K_0 < K$ significant taps. By contrast, in the context of the more realistic scenario of a nonsample-spaced CIR, the energy conveyed by the channel is distributed over a higher number of CIR-related taps, i.e., it potentially "leaks" to all CIR-related taps.

In order to demonstrate the effects of leakage, in Fig. 7, we have plotted the normalized variance of the diagonal elements of the decomposition of $\mathbf{R}_{\text{unif}}^{[f]}$ based on the space-frequency correlation function $r_{H,\text{unif}}(\Delta f)$ associated with the uniform multipath intensity profile,¹² when employing the DFT matrix \mathbf{W} as the unitary transform matrix $\tilde{\mathbf{U}}^{[f]}$, which is expressed mathematically as $\Upsilon_{\text{unif}}^{[f]} = \tilde{\mathbf{U}}^{[f]H} \mathbf{R}_{\text{unif}}^{[f]} \tilde{\mathbf{U}}^{[f]}$. The "u"-shaped evolution of the tap variances seen in Fig. 7 for tap indices in excess of $T_m/T_s = 16$ is, again, a result of the leakage incurred. By contrast, in the context of decomposing the matrix $\mathbf{R}_{\text{unif}}^{[f]}$ with the aid of the optimum KLT matrix, namely $\Lambda_{\text{unif}}^{[f]} = \mathbf{U}_{\text{unif}}^{[f]H} \mathbf{R}_{\text{unif}}^{[f]} \mathbf{U}_{\text{unif}}^{[f]}$, the channel's energy is concentrated on a number of CIR-related taps, which is only slightly higher [3] than the multipath spread T_m normalized to the sampling period duration T_s , as observed in Fig. 7.

E. A Priori Channel-Estimation MSE for a Nonsample-Spaced Uniform CIR¹³

The corresponding *a priori* channel-estimation MSE curves, which were evaluated with the aid of the iterative approach of Section III-B2, are plotted in Fig. 8, as a function of the K_0 number of significant CIR-related taps. The factor $(K_0/\text{Trace}(\Upsilon_{\text{unif}}^{[f](j)} \mathbf{I}_{K_0}^{(j)}))$ of (24) was evaluated upon selecting

¹¹The normalized multipath spread was set equal to one eighth of the $K = 512$ subcarriers assumed here, namely, to $T_m/T_s = 64$.

¹²The continuous unit-energy uniform power delay profile is given by $r_{h,\text{unif}}(\tau) = (1/T_m)\text{rect}(\tau - \tau_{\text{shift}}/T_m)$, while its Fourier transform, namely the space-frequency correlation function, is given by $r_{H,\text{unif}}(\Delta f) = \sin(\pi T_m \Delta f) \cdot e^{-j2\pi \tau_{\text{shift}} \Delta f}$.

¹³The normalized multipath spread was set equal to one eighth of the $K = 512$ subcarriers assumed here, namely, to $T_m/T_s = 64$.

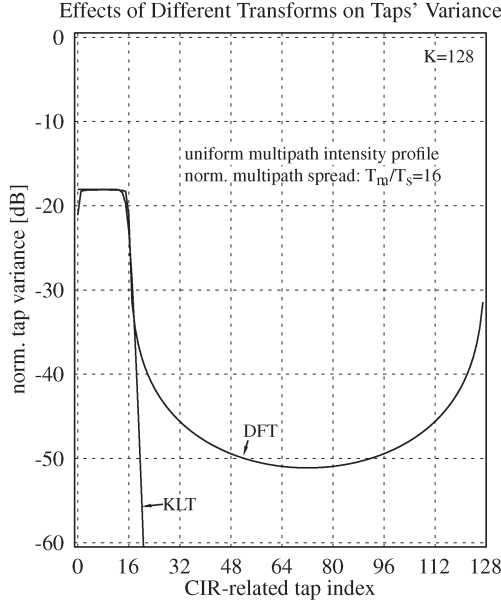


Fig. 7. Illustration of the normalized variance associated with the diagonal elements of the decomposition $\Upsilon^{[f]} = \tilde{\mathbf{U}}^{[f]H} \mathbf{R}^{[f]} \tilde{\mathbf{U}}^{[f]}$ of $\mathbf{R}^{[f]}$ with respect to $\tilde{\mathbf{U}}^{[f]} = \mathbf{W}$ and of the decomposition $\Lambda^{[f]} = \mathbf{U}^{[f]H} \mathbf{R}^{[f]} \mathbf{U}^{[f]}$ with respect to the optimum KLT matrix $\mathbf{U}^{[f]}$, in the context of a uniform multipath intensity profile, having a normalized multipath spread of $T_m/T_s = 16$. The normalization of the diagonal elements of $\Upsilon^{[f]}$ was carried out with respect to the $K = 128$ number of subcarriers.

the K_0 largest tap variances from the decomposition $\Upsilon_{\text{unif}}^{[f](j)} = \tilde{\mathbf{U}}^{[f](j)H} \mathbf{R}_{\text{unif}}^{[f](j)} \tilde{\mathbf{U}}^{[f](j)}$ of the specific space–frequency correlation matrices of the channel with respect to $\tilde{\mathbf{U}}^{[f](j)} = \mathbf{W}$.¹⁴ The curves are also parameterized with the $N_{\text{tap}}^{[t]}$ number of predictor taps. A rapid improvement of the estimator’s MSE is observed upon increasing the K_0 number of significant CIR-related taps up to a certain optimum K_0 value, which is a consequence of retaining more of the channel’s energy. At the same time, more of the undesired noise is retained, since a gradually decreasing fraction of the CIR-related taps are discarded. Upon increasing the K_0 number of significant taps beyond the optimum point seen in Fig. 8, the opposite behavior is observed, namely, that the MSE is degraded again. This is because for these taps, the benefit of extracting more of the channel’s energy is lower than the penalty incurred due to retaining more of the undesired noise. Note that this behavior is a result of employing the same set of predictor coefficients for the filtering of each of the different CIR-related taps. By contrast, in the context of a predictor arrangement employing individually optimized sets of coefficients for the prediction of each of the different CIR-related taps, a “leveling out” of the *a priori* estimation MSE performance would be observed, instead of the explicit degradation seen in Fig. 8. This is because for the low-energy CIR-related taps suffering from a low channel-

¹⁴Note that while for the sample-spaced CIR we have $\text{Trace}(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)}) = K \sigma_H^2$ for $K_0 < K$ upon appropriately selecting $\mathbf{I}_{K_0}^{(j)}$, in the context of the nonsample-spaced CIR, we potentially have $\text{Trace}(\Upsilon^{[f](j)} \mathbf{I}_{K_0}^{(j)}) < K \sigma_H^2$ for $K_0 < K$, which results from the leakage.

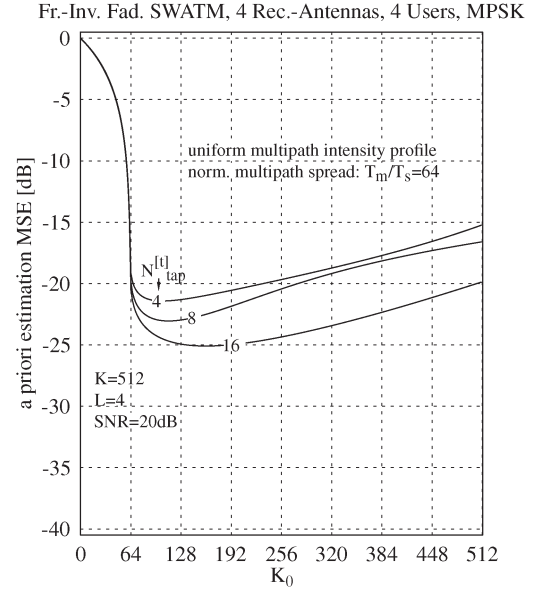


Fig. 8. *A priori* channel-estimation MSE performance of the PIC-assisted DDCE of Fig. 3, using optimum recursive predictor coefficients, versus the K_0 number of significant CIR-related taps retained in the context of a uniform multipath intensity profile, having a normalized multipath spread of $T_m/T_s = 64$. The DFT matrix $\tilde{\mathbf{U}}^{[f]} = \mathbf{W}$ was employed as a transform basis.

related signal-component-to-noise ratio, the noise would be more mitigated.

F. A Priori Channel Transfer-Factor Estimation MSE for a Nonsample-Spaced CIR on a Subcarrier Basis

The specific distribution of the *a priori* channel transfer-factor estimation MSE across the different subcarriers can also be obtained using the approach outlined in Section III-B2 for jointly optimizing the average *a priori* channel-estimation MSE and the predictor coefficients.

However, this involves invoking (16) instead of (22) in the algorithm outlined above. Again, in the context of a stable operation, as defined in the Appendix, we assume that the estimator’s statistics recorded in the form of the *a priori* channel transfer-factor estimation errors’ correlation matrix $\mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n] = \mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n - \hat{n}]$ is invariant for $\hat{n} = 1, \dots, N_{\text{tap}}^{[t]}$, yielding

$$\mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n] = \frac{\alpha_j}{\sigma_j^2} \left\| \tilde{\mathbf{c}}_{\text{pre}}^{(j)} \right\|_2^2 \mathbf{T}_{K_0}^{(j)} \times \left(\sum_{\substack{i=1 \\ i \neq j}}^L \sigma_i^2 \mathcal{D} \text{diag} \left\{ \mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(i)}}[n] \right\} + \sigma_n^2 \right) \mathbf{T}_{K_0}^{(j)H} + \mathbf{R}_{\mathbf{H}_{\text{dec}}^{(j)}}. \quad (29)$$

Recall that the desired subcarrier-based *a priori* channel transfer-factor estimation MSE variances are found on the main diagonal of the matrix $\mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n]$ of (29). The iteration commences with an initial assignment for the matrices $\mathbf{R}_{\Delta \hat{\mathbf{H}}_{\text{apr}}^{(j)}}[n]$, $j = 1, \dots, L$, potentially constrained by the condition that the matrices are supposed to be Hermitian.

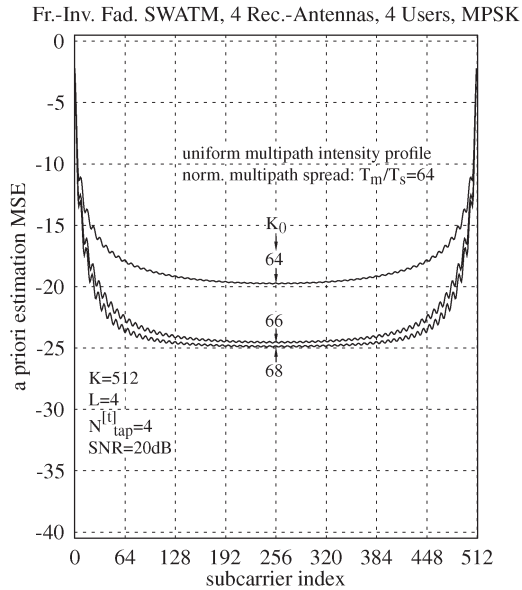


Fig. 9. *A priori* channel-estimation MSE performance versus the subcarrier index exhibited by the PIC-assisted DDCE of Fig. 3, using optimum recursive predictor coefficients. The DFT matrix $\tilde{\mathbf{U}}^{[f]} = \mathbf{W}$ was employed as a transform basis.

The j th user's *a priori* channel transfer-factor estimation-error correlation matrix is then updated with the aid of (29), on the basis of the remaining users' error correlation matrices' diagonals, denoted by $\text{Diag}\{\mathbf{R}_{\Delta\hat{\mathbf{h}}_{\text{app}}^{(j)}}[n]\}$, employing the remaining users' associated current vectors of predictor coefficients. After updating all users' error correlation matrices, the vectors of predictor coefficients are updated with the aid of (24). This involves evaluating first the average *a priori* channel transfer-factor estimation MSEs with the aid of (14), on the basis of the updated error correlation matrices. The iteration continues by updating the error correlation matrices, again, upon invoking the updated vectors of predictor coefficients.

Our analytical performance evaluations have been carried out for the uniform multipath intensity profile, again, in conjunction with a normalized multipath spread of $T_m/T_s = 64$ and for $K = 512$ subcarriers. The number of predictor taps was $N_{\text{tap}}^{[t]} = 4$. Our simulation results are portrayed in Fig. 9 for an SNR of 20 dB recorded at the reception antennas. The curves are further parameterized with the K_0 number of significant CIR-related taps. As also evidenced by the simulation results of Fig. 8, the value of K_0 should be in excess of $T_m/T_s = 64$ in order to be able to extract all the significant taps and hence, to prevent an excessive degradation of the MSE. The most important observation drawn from Fig. 9 is that, as a result of the effects of leakage imposed by the uniform multipath intensity profile, the estimation MSE is substantially degraded for the outer subcarriers of the frequency-domain OFDM symbol. Estimation MSEs as high as -5 dB are observed. Based on the relatively high MSE associated with the outer subcarriers, we also expect, for these subcarriers, a significantly deteriorated bit error rate (BER) performance, compared to the subcarriers at the center of the OFDM symbol.

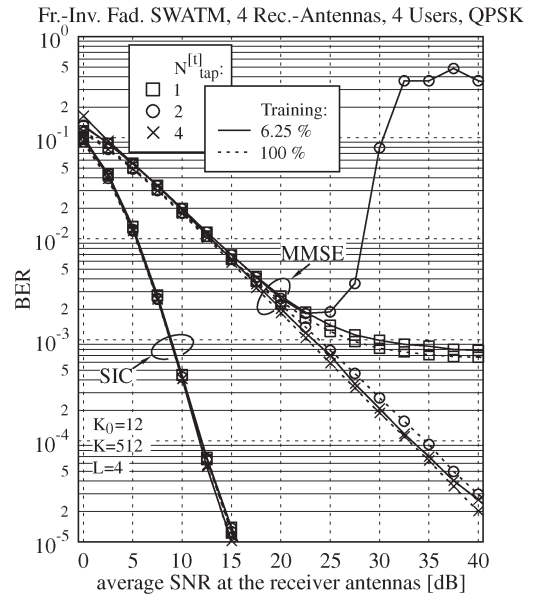


Fig. 10. BER versus SNR performance of an uncoded system employing the PIC-assisted DDCE of Fig. 3, using optimum recursive predictor coefficients in conjunction with both MMSE and M-SIC ($M = 2$)-based detection at the receiver. The fraction of training overhead imposed was either 6.25% or 100%, which corresponds to transmitting one dedicated training OFDM symbol per every block of 16 OFDM symbols, while the latter denotes the idealistic case of an error-free reference. Each of the SDMA scenario's independently faded channels is characterized by the indoor WATM channel parameters of [1] and [26].

G. System BER in the Context of Imperfect Error-Contaminated Symbol Decisions Assuming a Sample-Spaced CIR

So far, in this section, we have capitalized on the idealistic assumption of error-free symbol decisions. By contrast, in a realistic scenario, the channel-estimation process is impaired by erroneous symbol decisions. These effects will be further highlighted during our forthcoming discussions.

In Fig. 10, we have plotted the system's QPSK-related BER as a function of the SNR measured at the receiver antennas upon invoking both MMSE detection and the more effective, but also more complex, successive interference cancellation (SIC) detection technique, where $M = 2$ tentative symbol decisions were tracked from each detection node, which follows a similar philosophy as that of [27]. The curves are further parameterized with the number of CIR-related predictor taps, and the fraction of training overhead incorporated. While a training overhead of 6.25% corresponds to transmitting one dedicated training OFDM symbol in every block of 16 OFDM symbols, a training overhead of 100% indicates here the scenario where an error-free remodulated reference was made available to the DDCE for benchmarking. An interesting phenomenon is observed in the context of MMSE detection, when using two predictor taps. Due to an "excessive" number of erroneous subcarrier symbol decisions encountered in a specific OFDM symbol—which may be potentially induced by a deep fade on one of the channels of the multiple users—the *a priori* channel-estimation MSE encountered during the next OFDM symbol is severely degraded, which in turn may trigger an avalanche of errors. This may lead to the system's instability. In the context of

one-tap CIR-related tap prediction filtering, error-propagation events exceeding the length of a training period duration are prevented by periodically transmitting dedicated training OFDM symbols. By contrast, in the case of a higher number of predictor taps, the OFDM training block length should ideally be identical to the number of predictor taps in order to eliminate the possibility of error propagation across the training OFDM symbols. Note furthermore that these effects are not observed for the four-tap predictor—at least not in the range of SNRs considered—since the effects of errors imposed by a single OFDM symbol are more efficiently mitigated. Fig. 10 is, again, an evidence of the SIC combiner's more powerful detection capability.

V. COMPLEXITY

In this section, let us briefly analyze the algorithm's complexity. Since the optimization of the predictor coefficients following the concepts of Section III-B2 could be conducted offline, here, we will only consider the complexity of the PIC and the predictor units, as portrayed in Section III-A. It can be shown that the normalized complexity, expressed in terms of the number of complex multiplications $\mathcal{C}_0^{\text{C}\times\text{C}}$ and complex additions $\mathcal{C}_0^{\text{C}+\text{C}}$, is given by

$$\mathcal{C}_0^{\text{C}\times\text{C}} = 2 + \log_2 K + \frac{K_0}{K} N_{\text{tap}}^{[t]} \quad (30)$$

$$\mathcal{C}_0^{\text{C}+\text{C}} = (L - 1) + 2 \log_2 K + \frac{K_0}{K} N_{\text{tap}}^{[t]} \quad (31)$$

where the normalization was conducted with respect to the K , L , and P number of subcarriers, users, and receiver antennas, respectively. Note in (30) and (31) that the first additive term is related to the PIC operation of Section III-A1, while the second and third terms, to the transform and time-direction filtering operations of Section III-A2.

VI. CONCLUSION

In this paper, PIC-assisted DDCE, designed for OFDM scenarios supporting multiple users employing multiple transmit antennas at the BS, was investigated. Our mathematical analysis provided in Section III characterized the PIC-assisted DDCE's structure in Section III-A, separately for the PIC unit in Section III-A1, and for the predictor unit in Section III-A2, respectively. Furthermore, in Section III-B, an iterative approach was devised for the offline optimization of the estimator's CIR-related tap predictor coefficients. These discussions included the derivation of the predictor's *a priori* channel-estimation MSE in Section III-B1, based on which, the iterative optimization approach was derived later on in Section III-B2. In the context of the MSE and BER performance analysis conducted in Section IV, we found in Section IV-B that the estimator is capable of supporting a higher number of users employing more transmit antennas than $L = K/K_0$, which was found to be a limitation in the context of LS-assisted DDCE of [9]. Furthermore, we demonstrated in Section IV-C that the principles of robustness, with respect to

the specific shape of the channel's Doppler power spectrum, can be applied similarly to the case of DDCE designed for single-user OFDM, as investigated by Li *et al.* in [4]. The investigations of Sections IV-D–F were conducted with respect to the employment of the PIC-assisted DDCE in the context of nonsample-spaced CIRs, which resulted in the effect of CIR-related domain leakage, when employing the DFT matrix \mathbf{W} as the unitary transform matrix $\tilde{\mathbf{U}}^{[f]}$ associated with the estimator. Furthermore, we observed that the associated estimation MSE was nonuniformly distributed across the different subcarrier, where a significant degradation was observed at the outer subcarriers. These DFT-related effects could be mitigated upon employing the improved transforms proposed by Li in [13], instead of the DFT matrix \mathbf{W} . Our analysis of the system's BER recorded in the context of the sample-spaced indoor WATM CIR, and upon assuming imperfect potentially error-contaminated symbol decisions, showed that in order to guarantee the estimator's stability, the symbol decisions should be as reliable as possible, which was achieved with the advent of the powerful SIC detection scheme. A promising further research area is constituted by the OFDM-symbol-by-symbol-based adaptation of the CIR-related tap predictor's coefficients.

APPENDIX

A. Stability Analysis of the Recursive Channel Estimator

Recall that (22), which establishes a relation between the j th user's *a priori* estimation MSE and the remaining $L - 1$ users' *a priori* estimation MSEs, can be viewed as a system of equations for different values of $j = 1, \dots, L$, namely for the different users. It can be shown that (22) can be represented in a compact vectorial notation as

$$\mathbf{MSE}_{\text{apr}} = \mathbf{C}_{\text{pre}} \cdot \mathbf{P}_{\text{s}}^{-1} \cdot \mathbf{F} \cdot \mathbf{P}_{\text{s}} \cdot \mathbf{MSE}_{\text{apr}} + \mathbf{C}_{\text{pre}} \cdot \mathbf{P}_{\text{s}}^{-1} \cdot \mathbf{p}_{\text{n}} + \mathbf{MSE}_{\text{dec}} \quad (32)$$

where $\mathbf{MSE}_{\text{apr}} \in \mathbb{R}^{L \times 1}$ hosts the different users' *a priori* estimation MSEs denoted by $\overline{\text{MSE}}_{\text{apr}}^{(j)}$, $j = 1, \dots, L$, and the diagonal matrix $\mathbf{C}_{\text{pre}} \in \mathbb{R}^{L \times L}$ hosts the different users' CIR-related tap prediction-coefficient-related terms of $(K_0/K)\alpha_j \|\tilde{\mathbf{c}}_{\text{pre}}^{(j)}\|_2^2$, $j = 1, \dots, L$. A characteristic component is the feedback matrix $\mathbf{F} \in \mathbb{R}^{L \times L}$, where the elements on all side diagonals are of unit value, except for the main diagonal, whose elements are zero. The relation to the PIC process is that for the estimation of the j th user's channel transfer function, the cochannel interference imposed by the $L - 1$ remaining users has to be removed. Note in this context that the j th row of matrix \mathbf{F} is associated with the estimation process of the j th user's channel. Furthermore, the diagonal matrix $\mathbf{P}_{\text{s}} \in \mathbb{R}^{L \times L}$ hosts the different users' signal variances σ_i^2 , $i = 1, \dots, L$, while the vector $\mathbf{p}_{\text{n}} \in \mathbb{R}^{L \times 1}$ exhibits identical elements equal to the AWGN noise variance σ_{n}^2 . Finally, the matrix $\mathbf{MSE}_{\text{dec}} \in \mathbb{R}^{L \times 1}$ hosts the different users' residual channel decorrelation-related MSEs values, given by $(1/K)\text{Trace}(\mathbf{R}_{\mathbf{H}_{\text{dec}}}^{(j)})$, $j = 1, \dots, L$, which are also a function of the individual users' CIR-related

tap predictor coefficients, as evidenced by (20). In order to proceed further, (32) can be solved for the vector of *a priori* estimation MSEs, conditioned on the knowledge of the vectors $\tilde{\mathbf{c}}_{\text{pre}}^{(j)}$, $j = 1, \dots, L$ of predictor coefficients, yielding¹⁵

$$\mathbf{MSE}_{\text{apr}} = (\mathbf{I} - \mathbf{C}_{\text{pre}} \cdot \mathbf{P}_s^{-1} \cdot \mathbf{F} \cdot \mathbf{P}_s)^{-1} \cdot (\mathbf{C}_{\text{pre}} \cdot \mathbf{P}_s^{-1} \cdot \mathbf{p}_n + \mathbf{MSE}_{\text{dec}}). \quad (33)$$

Per definition, the elements of $\mathbf{MSE}_{\text{apr}}$ or, equivalently, the different users' *a priori* estimation MSEs must have a finite positive value. This is coupled to the following conditions:

- 1) existence of $(\mathbf{I} - \hat{\mathbf{F}})^{-1}$, where $\hat{\mathbf{F}} = \mathbf{C}_{\text{pre}} \cdot \mathbf{P}_s^{-1} \cdot \mathbf{F} \cdot \mathbf{P}_s$;
- 2) all elements of $(\mathbf{I} - \hat{\mathbf{F}})^{-1}$ must be positive.

It can be demonstrated that these two conditions are fulfilled, if the spectral radius¹⁶ $\rho(\hat{\mathbf{F}})$ of the matrix $\hat{\mathbf{F}} = \mathbf{C}_{\text{pre}} \cdot \mathbf{P}_s^{-1} \cdot \mathbf{F} \cdot \mathbf{P}_s$ is less than unity [28]. An upper bound estimate of the spectral radius is given by the largest Euclidean distance measured from the origin in \mathbb{C} , exhibited by a point contained in the union $G(\hat{\mathbf{F}})$ of Gershgorin disks of $\hat{\mathbf{F}}$. We note here that with the aid of the Gershgorin circle theorem [19], explicit bounds can be placed on the regions in \mathbb{C} , which host the eigenvalues of a matrix $\mathbf{A} \in \mathbb{C}^{m \times m}$. The i th Gershgorin disk is defined as: $R_i(\mathbf{A}) = \{x \in \mathbb{C} : |x - a_{ii}| \leq \sum_{j \neq i}^m |a_{ij}|\}$, where a_{ij} is the element of the matrix \mathbf{A} associated with its i th row and j th column. The eigenvalues of the matrix \mathbf{A} reside within the union of Gershgorin disks of \mathbf{A} , which is formulated in a compact form as $\lambda(\mathbf{A}) \subset \bigcup_{i=1}^m R_i(\mathbf{A}) = G(\mathbf{A})$ [19]. Hence, provided that we have

$$\max_{j=1, \dots, L} \left(\frac{K_0}{K} \alpha_j \left\| \tilde{\mathbf{c}}_{\text{pre}}^{(j)} \right\|_2^2 \sum_{\substack{i=1 \\ i \neq j}}^L \frac{\sigma_i^2}{\sigma_j^2} \right) < 1 \quad (34)$$

it can be shown that $(\mathbf{I} - \hat{\mathbf{F}})$ is invertible. By contrast, if this condition is not fulfilled, no immediate conclusion can be drawn with respect to the invertibility of $(\mathbf{I} - \hat{\mathbf{F}})$.

A further criterion for the existence of the matrix inverse $(\mathbf{I} - \hat{\mathbf{F}})^{-1}$ is coupled to the condition that the determinant of $(\mathbf{I} - \hat{\mathbf{F}})$ is nonzero, namely, that we have $\det(\mathbf{I} - \hat{\mathbf{F}}) \neq 0$. It can be shown that for all elements of this specific matrix inverse to be positive, as stipulated in (2), we have to satisfy the condition of $\det(\mathbf{I} - \hat{\mathbf{F}}) > 0$. Furthermore, it can be shown that $\det(\mathbf{I} - \hat{\mathbf{F}}) = \det(\mathbf{I} - \mathbf{C}_{\text{pre}} \cdot \mathbf{F})$, which implies that the channel estimator's stability is only a function of the estimator coefficients to be determined. Even if the channel conditions are subjected to variations, the estimator remains stable for

¹⁵Note that in the context of identical transmit powers, modulation schemes, and channel statistics, (33) is significantly simplified, namely, we obtain the same average estimation MSE given by $\overline{\mathbf{MSE}}_{\text{apr}}^{\text{SIMPLE}} = (\chi(\sigma_n^2/\sigma_s^2) \|\tilde{\mathbf{c}}_{\text{pre}}\|_2^2 + (1/K)\text{Trace}(\mathbf{R}_{\mathbf{H}_{\text{dec}}})) / (1 - \chi(L-1) \|\tilde{\mathbf{c}}_{\text{pre}}\|_2^2)$, where $\chi = (K_0/K)\alpha$, as well as the same predictor coefficients for the different users [1].

¹⁶Recall that the spectral radius of a matrix is the smallest radius of a circle centered around the origin of \mathbb{C}_{pre} that contains all the matrix's eigenvalues [19].

a "stable" set of coefficients—provided that correct symbol decisions are performed.

REFERENCES

- [1] L. Hanzo, M. Münster, B.-J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs and Broadcasting*. New York: Wiley, 2003.
- [2] L. Hanzo, L. L. Yang, E. L. Kuan, and K. Yen, *Single- and Multi-Carrier DS-SS-CDMA*. New York: Wiley, 2003.
- [3] 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.
- [4] Y. Li, L. J. Cimini, 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.
- [5] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 467–479, Mar. 2001.
- [6] B. Yang, Z. Cao, and K. B. Letaief, "Analysis of low-complexity windowed DFT-based MMSE channel estimator for OFDM systems," *IEEE Trans. Commun.*, vol. 49, no. 11, pp. 1977–1987, Nov. 2001.
- [7] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Pilot-based channel estimation for OFDM systems by tracking the delay-subspace," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 315–325, Jan. 2004.
- [8] R. Lin and A. Petropulu, "Linear precoding assisted blind channel estimation for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 54, no. 3, pp. 983–995, May 2005.
- [9] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 461–471, Mar. 1999.
- [10] W. G. Jeon, K. H. Paik, and Y. S. Cho, "An efficient channel estimation technique for OFDM systems with transmitter diversity," in *IEEE Proc. Int. Symp. Personal, Indoor and Mobile Radio Communications*, London, U.K., Sep. 18–21, 2000, vol. 2, pp. 1246–1250.
- [11] Y. Li, "Optimum training sequences for OFDM systems with multiple transmit antennas," in *IEEE Proc. Global Telecommunications Conf.*, San Francisco, CA, Nov. 27–Dec. 1, 2000, vol. 3, pp. 1478–1482.
- [12] Y. Xie and C. N. Georghiades, "An EM-based channel estimation algorithm for OFDM with transmitter diversity," in *IEEE Proc. Global Telecommunications Conf.*, San Antonio, TX, Nov. 25–29, 2001, vol. 2, pp. 871–875.
- [13] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.
- [14] H. Bölcskei, J. R. W. Heath, and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multi-antenna systems," *IEEE Trans. Signal Process.*, vol. 50, no. 1, pp. 96–109, Jan. 2002.
- [15] H. Minn, D. I. Kim, and V. K. Bhargava, "A reduced complexity channel estimation for OFDM systems with transmit diversity in mobile wireless channels," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 799–807, May 2002.
- [16] O. Edfors, M. Sandell, J. J. Van de Beek, S. K. Wilson, and P. O. Börjesson, "OFDM channel estimation by singular value decomposition," in *IEEE Proc. Vehicular Technology Conf.*, Atlanta, GA, Apr. 28–May 1, 1996, vol. 2, pp. 923–927.
- [17] R. Steele and L. Hanzo, *Mobile Radio Communications*. Chichester, U.K.: Wiley, 1999.
- [18] M. Sandell, "Design and analysis of estimators for multicarrier modulation and ultrasonic imaging," Ph.D. thesis, Lulea Univ. Technol., Lulea, Sweden, 1996.
- [19] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Upper Saddle River, NJ: Prentice-Hall, 2000.
- [20] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1313–1324, Oct. 1998.
- [21] A. Chini, Y. Wu, M. El-Tanany, and S. Mahmoud, "Filtered decision feedback channel estimation for OFDM-based DTV terrestrial broadcasting system," *IEEE Trans. Broadcast.*, vol. 44, no. 1, pp. 2–11, Mar. 1998.
- [22] G. Strang, *Linear Algebra and Its Applications*, 2nd ed. New York: Academic, 1980.
- [23] Y. Li and N. R. Sollenberger, "Clustered OFDM with channel estimation for high rate wireless data," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2071–2076, Dec. 2001.
- [24] J. W. C. Jakes, *Mobile Microwave Communications*. New York: Wiley, 1974.

- [25] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Börjesson, "On channel estimation in OFDM systems," in *IEEE Proc. Vehicular Technology Conf.*, Chicago, IL, Jul. 1995, vol. 2, pp. 815–819.
- [26] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: A convenient framework for time–frequency processing in wireless communications," *Proc. IEEE*, vol. 88, no. 5, pp. 611–642, May 2000.
- [27] P. Vandenameele, L. V. D. Perre, M. G. E. Engels, B. Gyselinckx, and H. J. D. Man, "A combined OFDM/SDMA approach," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2312–2321, Nov. 2000.
- [28] F. R. Gantmacher, *The Theory of Matrices*, vol. 2. New York: Chelsea, 1990.

Matthias Münster received the Dipl. Ing. degree from the Rheinisch-Westfälische Technische Hochschule, Aachen, Germany, in 1998. After graduation, he embarked on postgraduate research at the University of Southampton, Southampton, U.K., where he received the Ph.D. degree in mobile communications in 2002.

His areas of interest include adaptive multiuser orthogonal frequency division multiplexing (OFDM) transmission, wideband channel estimation, multiuser detection, and a range of related signal-processing aspects. During his Ph.D. research, he contributed over a dozen various research papers and following the completion of his thesis, he returned to his native Germany, where he is currently involved in the development of sophisticated signal-processing algorithms.

Lajos Hanzo (M'91–SM'92–F'04) received the degree in electronics in 1976, the doctorate degree in 1983, and the D.Sci. degree in 2004, all from the University of Southampton, Southampton, U.K.

During his career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the Department of Electronics and Computer Science, University of Southampton, where he holds the Chair in Telecommunications. He coauthored 11 books totaling 9000 pages on mobile radio communications, published in excess of 500 research papers, organized and chaired conference sessions, and presented overview lectures. Currently, he heads an academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by the industry, the Engineering and Physical Sciences Research Council (EPSRC) U.K., the European Information Society Technologies (IST) Programme, and the Mobile Virtual Centre of Excellence (VCE), U.K. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses.

Dr. Hanzo has been awarded a number of distinctions. He is also an IEEE Distinguished Lecturer of both the Communications as well as the Vehicular Technology Society and a Fellow of the Institution of Electrical Engineers, U.K. (IEE). He is a Fellow of the Royal Academy of Engineering, a Governor of the IEEE Vehicular Technology Society and an Editor of the PROCEEDINGS OF THE IEEE.