IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. 55, NO. 10, OCTOBER 2007

A Unified Framework for Tomlinson–Harashima Precoding in MC-CDMA and OFDMA Downlink Transmissions

Michele Morelli, Senior Member, IEEE, and Luca Sanguinetti, Member, IEEE

Abstract—We consider a unified framework comprising both multicarrier code-division multiple-access (MC-CDMA) and orthogonal frequency-division multiple-access (OFDMA), and discuss nonlinear prefiltering for downlink transmissions based on Tomlinson–Harashima precoding. The base station (BS) is equipped with multiple transmitting antennas and channel state information is assumed to be available at the transmit side. We design the prefiltering matrices so as to minimize the sum of the mean square errors at all mobile terminals when a conventional single-user data detector is employed at the receiver side. In this way, most of the computational burden is moved to the BS, where power consumption and computational resources are not critical requirements. Computer simulations are used to assess the performance of the proposed scheme under different operating scenarios. It turns out that OFDMA outperforms MC-CDMA when the system resources (subcarriers and/or spreading codes) are optimally assigned to the active users according to the channel quality.

Index Terms—Minimum mean square error (MMSE) prefiltering, multicarrier code-division multiple-access (CDMA), orthogonal frequency-division multiple-access (OFDMA), Tomlinson–Harashima precoding (THP).

I. INTRODUCTION

The demand for high data rates in wireless transmissions has led to a strong interest in multicarrier modulations due to their intrinsic robustness against frequency-selective fading [1]. Among them, multicarrier code-division multiple-access (MC-CDMA) and orthogonal frequency-division multiple-access (OFDMA) [2] are particularly promising for future broadband communications because of their high spectral efficiency and flexibility for integrated multimedia applications. These schemes can support multiple users by exploiting either code- or frequency-division multiplexing. In particular, in MC-CDMA, the data of different users are spread in the frequency domain using preassigned signature sequences while in OFDMA an exclusive set of orthogonal subcarriers is allocated to each user.

Recent publications show that MC-CDMA is particularly attractive for downlink transmissions, i.e., from the base station (BS) to the mobile terminals (MTs) [3]. In these applications, orthogonal spreading codes are usually employed to provide protection against cochannel interference. In the presence of multipath propagation, however, signals undergo frequency-selective fading, and the code orthogonality is lost. This gives rise to multiple-access interference (MAI), which strongly limits the system performance. In conventional MC-CDMA systems, MAI mitigation is accomplished at the receiver side using well-known single-user or multiuser detection techniques [4]. Alternatively, linear prefiltering based on the zero-forcing (ZF) or minimum-mean-square-error (MMSE) criteria can be employed to mitigate MAI and channel distortions [5], [6]. A possible drawback of linear prefiltering is represented by the power boosting effect [7], which occurs in the presence of deeply faded subcarriers and leads to high power consumption at the transmit side.

An effective solution to mitigate power boosting is represented by nonlinear prefiltering based on Tomlinson–Harashima precoding (THP). This technique employs modulo arithmetic and was originally proposed to combat intersymbol interference (ISI) in single-user transmissions over highly dispersive channels [8], [9]. Recently, it has been employed in multiuser systems to counteract the detrimental effect of MAI [10]–[13]. In these applications, THP can be viewed as the transmit counterpart of the vertical Bell Labs layered space–time (V-BLAST) architecture [14]. The main difference is that the latter operates at the receiver side, and can only exploit data decisions for interference cancellation, while the former is employed at the transmitter, where the true data symbols are available. This means that THP does not suffer from the error propagation problem, and accordingly, is expected to outperform V-BLAST.

So far, nonlinear prefiltering based on THP has been extensively studied for multiple-input multiple-output systems with decentralized receivers (MIMO multiuser) [10]–[12]. In particular, the scheme discussed in [10] is based on a ZF approach and exploits the QR decomposition of the MIMO channel matrix. The processing matrices are designed in [11] according to an MMSE criterion under a constraint on the overall transmit power. Unfortunately, the solution to this problem requires a large number of matrix inversions (equal to the number of active users) and its application to heavy-loaded systems may be impractical. An efficient implementation of the aforementioned algorithm is discussed in [12], where all matrix inversions are replaced by a single Cholesky factorization. Nonlinear prefiltering for MC-CDMA downlink transmissions has been investigated in [13] under a ZF constraint. Although reasonable, this scheme produces significant differences in the error rate performance of the receive terminals, which may be undesirable in commercial applications where a fair treatment of the active users is normally required.

Paper approved by R. Kohno, the Editor for Spread Spectrum Theory and Applications of the IEEE Communications Society. Manuscript received November 3, 2005; revised April 9, 2006. This work was supported by the Italian Ministry of Education under FIRB Project PRIMO, 2002.

The authors are with the Department of Information Engineering, University of Pisa, Pisa 56126, Italy (e-mail: luca.sanguinetti@iet.unipi.it; michele.morelli@iet.unipi.it).

Digital Object Identifier 10.1109/TCOMM.2007.906430
In this paper, we employ a unified framework to investigate nonlinear prefiltering in MC-CDMA and OFDMA systems equipped with multiple transmit antennas. To keep the complexity of the MTs at a reasonable level, single-user detection (SUD) strategies are adopted at the remote units, and the goal is to derive the optimal prefiltering matrices at the BS for a preassigned receiver structure. This means that we do not attempt to perform any joint transmit-receive optimization since, although powerful, this approach would require fully cooperative receivers, which is difficult to achieve in multiuser downlink transmissions [15]. The optimality criterion adopted for the design of the processing matrices is the minimization of the sum of mean square errors (MSEs) at all MTs under a constraint on the overall transmit power. The resulting scheme mitigates MAI and channel distortions and outperforms the method in [13]. Furthermore, it allows the MTs to employ conventional SUD techniques, thereby moving most of the computational burden to the BS, where power consumption and computational resources are not critical requirements.

The use of a unified framework comprising both MC-CDMA and OFDMA allows a fair comparison between these multiple-access technologies under the same operating conditions. As we shall see, in OFDMA applications, the proposed THP-based scheme is equivalent to maximum ratio transmission (MRT) over the available transmit antennas [16], thereby leading to a significant reduction of complexity as compared to MC-CDMA. It is also found that OFDMA outperforms MC-CDMA when the system resources are optimally assigned to the active users according to the actual channel realization.

The remainder of the paper is organized as follows. Next section outlines the signal model and introduces basic notation. In Section III, we design the prefiltering matrices following an MMSE criterion. Numerical results are illustrated in Section IV, while some concluding remarks are given in Section V.

II. SYSTEM MODEL

A. Transmitter Structure

We consider the downlink of a multicarrier system in which the BS is equipped with $N_T$ antennas and the total number of subcarriers $N$ is divided into smaller groups of $Q$ elements [2, p. 73]. Without loss of generality, we concentrate on a single group and denote $\{i_n; 1 \leq n \leq Q\}$ the corresponding subcarrier indices. The BS employs the $Q$ subcarriers to communicate with $K$ active users ($K \leq Q$), which can be separated in various ways. In this paper, we consider both code- and frequency-division multiplexing, corresponding to MC-CDMA and OFDMA, respectively. The symbol transmitted to the $k$th user is denoted $a_k$ and belongs to an $M$-ary quadrature-amplitude modulation (QAM) constellation with average energy $\sigma_a^2 = 2(M-1)/3$. This amounts to saying that both the real and imaginary parts of $a_k$ are taken from the set $A = \{\pm 1, \pm 3, \ldots, \pm \sqrt{M-1}\}$. For convenience, we collect the users’ data into a $K$-dimensional vector $\mathbf{a} = [a_1, a_2, \ldots, a_K]^T$, where the notation $[\cdot]^T$ denotes the transpose operation.

The previous equation may be written in matrix form as

$$\mathbf{b} = \mathbf{C}^{-1}\mathbf{P}\mathbf{a}$$

where $\mathbf{C} = \mathbf{B} + \mathbf{I}_K$ is a unit-diagonal and lower triangular matrix, i.e., $[\mathbf{C}]_{k,k} = 1$ and $[\mathbf{C}]_{k,k} = 0$ for $k < 1$. Inspection of (3) reveals that the energy of the precoded symbols depends on $\mathbf{C}$ and may be very large in the presence of deep fades, thereby leading to a significant increase of the transmit power [10]. This is a manifestation of the power boosting effect that occurs in conjunction with linear prefiltering. To overcome this problem, we adopt a THP approach, and introduce a modulo operator that acts independently over the real and imaginary parts of its input.

![Fig. 1.](image) (a) Block diagram of the transmitter. (b) Equivalent block diagram of the prefiltering unit.
according to the following rule:

$$\text{MOD}_M(x) = x - 2\sqrt{M} \cdot \left\lfloor \frac{x - \sqrt{M}}{2\sqrt{M}} \right\rfloor$$  \hfill (4)$$

where the notation $\lfloor z \rfloor$ indicates the smallest integer larger than or equal to $z$. In this way, the precoded symbols $b_k$ at the output of the nonlinear device are constrained into the square region $\mathcal{N} = \{ x^{(R)} + j x^{(I)} \mid x^{(R)}, x^{(I)} \in (-\sqrt{M}, \sqrt{M}) \}$ and the transmit power is consequently reduced as compared to conventional linear prefiltering.

Applying the nonlinearity (4) to the right-hand side (RHS) of (2) yields

$$b_k = \tilde{a}_k - \sum_{l=1}^{k-1} [B]_{k,l} b_l + \tilde{d}_k, \quad k = 1, 2, \ldots, K$$  \hfill (5)$$

where $\tilde{d}_k = 2\sqrt{M} \cdot \tilde{p}_k$ and $\tilde{p}_k$ is a complex-valued quantity whose real and imaginary components are suitable integers that constraint $b_k$ within the region $\mathcal{N}$ (clearly, a unique $\tilde{p}_k$ exists with such a property). The previous equation indicates that the modulo operator in Fig. 1(a) is equivalent to adding a vector $d = P^T \tilde{d}$ to the input data $a$, where $d = [d_1, d_2, \ldots, d_K]^T$, and we have borne in mind that $PP^T = I_K$. This results into the equivalent block diagram of Fig. 1(b), from which it follows that $b = Pv - Bb$, or equivalently,

$$b = C^{-1}Pv$$  \hfill (6)$$

where $v = a + d$ is the effective data vector [17]. Comparing (6) with (3), we see that the only difference between linear prefiltering and THP is that, in the latter, the true data vector $a$ is replaced by $v$. In this way, the transmit power is reduced without any loss of information since $a$ can easily be regenerated from $v$ using the same module device in (4).

After nonlinear precoding, vector $b$ is passed to $N_T$ matrices $U_i$ ($i = 1, 2, \ldots, N_T$), one for each antenna branch. This produces the $Q$-dimensional vectors $s_i = U_i b$, which can also be written as

$$s_i = \sum_{k=1}^K b_k u_{k,i}, \quad i = 1, 2, \ldots, N_T$$  \hfill (7)$$

where $u_{k,i}$ is the $k$th column of $U_i$. Next, each $s_i$ is frequency interleaved with the contributions of the other groups, and the resulting vector $z_i$ is finally mapped onto $N$ subcarriers using a conventional OFDM modulator. Clearly, in OFDMA transmissions, only one entry of $u_{k,i}$ is expected to be nonzero since $b_k$ is transmitted over a single subcarrier. On the other hand, in MC-CDMA, $u_{k,i}$ can be interpreted as the prefiltered spreading code of the $k$th user at the $i$th transmit branch.

Collecting vectors $s_i = U_i b$ into a single $N_T Q$-dimensional vector $s = [s_1^T s_2^T \cdots s_{N_T}^T]^T$, we have

$$s = Ub$$  \hfill (8)$$

where $U = [U_1^T U_2^T \cdots U_{N_T}^T]^T$ is a matrix of dimensions $N_T Q \times K$.

B. Receiver Structure

The signals transmitted by the BS array propagate through multipath channels and undergo frequency-selective fading. Without loss of generality, we concentrate on the $m$th MT and assume that it is equipped with a single-antenna receiver. As shown in Fig. 2, the incoming waveforms are implicitly combined by the receive antenna and passed to an OFDM demodulator. We denote by $X_m = [X_m(1), X_m(2), \ldots, X_m(Q)]^T$ the demodulator outputs corresponding to the $Q$ subcarriers of the considered group. Then, we have

$$X_m = H_m Ub + n_m$$  \hfill (9)$$

where $H_m = [H_{m,1} \, H_{m,2} \cdots H_{m,Q}]$ is a $Q \times N_T Q$ matrix in which $H_{m,i} = \text{diag}\{H_{m,i}(i_1), H_{m,i}(i_2), \ldots, H_{m,i}(i_Q)\}$ represents the channel frequency response between the $i$th transmit antenna and the $m$th MT over the $Q$ subcarriers. Also, $n_m = [n_m(1), n_m(2), \ldots, n_m(Q)]^T$ is thermal noise that is modeled as a Gaussian vector with zero mean and covariance matrix $\sigma_n^2 I_Q$.

To keep the complexity of the MT at a tolerable level, a simple SUD scheme is employed for data detection. For this purpose, the entries of $X_m$ are linearly combined to form

$$y_m = \sqrt{e} \cdot q_m^H X_m$$  \hfill (10)$$

where $q_m = [q_m(1), q_m(2), \ldots, q_m(Q)]^T$ is a unit-norm vector, $(\cdot)^H$ denotes Hermitian transpose, and $e > 0$ is a real-valued parameter that can be thought of as being a part of the automatic gain control (AGC) unit. Recalling that the nonlinear precoding is equivalent to replacing the true data symbols $a$ with $v = a + d$, in normal operating conditions, we expect that $y_m = v_m + \eta_m$, where $v_m = a_m + d_m$ while $\eta_m$ is a disturbance term that accounts for residual interference and thermal noise. Hence, in order to remove the contribution of $d_m$, we pass $y_m$ to a nonlinear device that operates according to (4). The output is finally fed to a threshold unit that delivers an estimate of $a_m$.

Different selections of $q_m$ in (10) correspond to different SUD schemes and/or multiple-access techniques. For example, in OFDMA, each subcarrier in the considered group is exclusively assigned to a different user, and in consequence, the entries of $q_m$ take the form

$$q_m(n) = \begin{cases} 1, & \text{if } n = j_m \\ 0, & \text{otherwise} \end{cases}$$  \hfill (11)$$

where $j_m$ belongs to the set $\{1, 2, \ldots, Q\}$ and satisfies the constraints $j_m \neq j_k$ for $m \neq k$ since different users must be allocated over different subcarriers. In case of dynamic subcarrier allocation, the mapping function $m \rightarrow j_m$ ($m = 1, 2, \ldots, K$) is properly designed so as to optimize the system performance.
Alternatively, we can simply set $j_m = m$ if the system operates in a rigid fashion without employing any dynamic resource allocation policy. On the other hand, in MC-CDMA, the users’ data are spread over the $Q$ subcarriers using orthogonal Walsh–Hadamard (WH) codes. This amounts to setting

$$q_m(n) = c_{j_m}(n) t_m(n), \quad n = 1, 2, \ldots, Q$$

(12)

where the coefficients $\{t_m(n)\}$ are designed according to the selected SUD strategy, while $c_{j_m} = [c_{j_m}(1), c_{j_m}(2), \ldots, c_{j_m}(Q)]^T$ is the WH spreading sequence assigned to the $m$th user, with $c_{j_m}(n) \in \{\pm 1/\sqrt{Q}\}$. Here, the mapping $m \rightarrow j_m$ accounts for the possible dynamic allocation of the signature codes among the active users. In the sequel, we consider the following SUD techniques.

1) Pure de-spreading (PD):

$$t_m(n) = 1.$$  

(13)

2) Maximum ratio combining (MRC):

$$t_m(n) = \frac{\sum_{i=1}^{Q} H_{m,i}(i_n)}{\sqrt{1/\sum_{i=1}^{Q} \sum_{i=1}^{Q} |H_{m,i}(i_n)|^2}}.$$  

(14)

3) Equal gain combining (EGC):

$$t_m(n) = \frac{\sum_{i=1}^{Q} H_{m,i}(i_n)}{\sum_{i=1}^{Q} H_{m,i}(i_n)}.$$  

(15)

Note that the PD strategy dispenses from channel knowledge at the receiver, thereby keeping the complexity of the MT at a very low level.

Substituting (9) into (10) yields

$$y_m = \sqrt{\rho} \cdot g_m^H U b + \sqrt{\nu} \cdot w_m$$  

(16)

where $g_m = H_m^H q_m$ is a $QK$-dimensional vector, while $w_m = q_m^H n_m$ is the noise contribution. Finally, by letting $y = [y_1, y_2, \ldots, y_K]^T$ and $F = \sqrt{\rho} \cdot U$, we obtain

$$y = G^H F b + \sqrt{\nu} \cdot w$$  

(17)

where $G$ is the following matrix of dimensions $N_T Q \times K$

$$G = \begin{bmatrix}
H_{1,1}^H q_1 & H_{1,2}^H q_1 & \cdots & H_{1,K}^H q_K \\
H_{1,1}^H q_1 & H_{2,2}^H q_1 & \cdots & H_{1,K}^H q_K \\
\vdots & \vdots & \ddots & \vdots \\
H_{1,1}^H q_1 & H_{2,2}^H q_1 & \cdots & H_{1,K}^H q_K \\
H_{1,N_T}^H q_1 & H_{2,N_T}^H q_1 & \cdots & H_{1,K}^H q_K
\end{bmatrix}$$  

(18)

while $w = [w_1, w_2, \ldots, w_K]^T$ is a Gaussian vector with zero mean and covariance matrix $\sigma_w^2 I_Q$.

III. DESIGN OF THE PREFILTERING MATRICES

A. Optimality Criterion

In this section, the processing matrices $P$, $C$, and $U$ are designed so as to minimize the sum of MSEs at all MTs. Since the contribution of $d$ to the received vector $y$ is removed by the modulo operator employed at each MT, the desired value for $y$ is $y = v$ instead of $y = a$. On the other hand, from (6), it follows that $v = P^T C b$ so that our optimality criterion leads to the minimization of the following objective function

$$J = E\{\| y - P^T C b \|^2 \}$$  

(19)

where $\| \cdot \|$ denotes the Euclidean norm of the enclosed vector, and the statistical expectation is computed over data symbols and thermal noise.

To maintain the same power as in the absence of any prefiltering, we impose a constraint on the overall transmit power. Also, to make the problem mathematically tractable, we assume that the precoded symbols $b_j$ are statistically independent with zero mean and the same power $\sigma_b^2$ as the user data. Although not rigorously true, this assumption is reasonable for large $M$-QAM constellations with size $M \geq 16$ [17]. In the aforementioned hypothesis, the power constraint can be formulated as $tr\{U^H U\} = K$, or equivalently,

$$tr\{F^H F\} = e \cdot K$$  

(20)

where $tr\{\cdot\}$ denotes the trace of the enclosed matrix and we have used the identity $F = \sqrt{e} \cdot U$.

B. Design of the Backward and Forward Matrices

In minimizing the objective function (19) under the constraint (20), we can follow the line of reasoning employed in [18], or alternatively, the more conventional technique presented in [12]. Since, in both cases, we arrive at the same final result, in the sequel, we adopt the approach of [18] that is mathematically simpler. We begin by substituting (17) into (19) and computing the statistical expectation over $b$ and $w$. This produces

$$J = \sigma_a^2 \cdot tr\{ (G^H F - P^T C)(G^H F - P^T C)^H \} + e \cdot K \sigma_b^2$$  

(21)

where we have borne in mind that $b$ and $w$ are statistically independent with zero mean and covariance matrices $\sigma_a^2 I_Q$ and $\sigma_w^2 I_Q$, respectively. Next, substituting (20) into (21) and letting $\rho = \sigma_a^2 / \sigma_b^2$, we obtain

$$J = \sigma_a^2 \cdot tr\{ (G^H F - P^T C)(G^H F - P^T C)^H + \rho \cdot F^H F \}.$$  

(22)

Our goal is to determine the matrices $F$, $C$, and $P$ that minimize the RHS of (22). For this purpose, we keep $C$ and $P$ fixed and set to zero the gradient of $J$ with respect to $F$. This yields

$$F = \hat{G}(G^H \hat{G} + \rho I_K)^{-1} C$$  

(23)

where $\hat{G} = GP^T$, and we have used the identity $P^T P = I_K$. Next, we substitute (23) into (22) to obtain

$$J = \sigma_a^2 \cdot tr\{ C^H (G^H \hat{G} + \rho I_K)^{-1} C \}$$  

(24)

and look for the unit-diagonal and lower-triangular matrix $C$ that minimizes the RHS of (24). This problem is addressed in Appendix A and the solution is given by

$$C = \tilde{L} \tilde{D}$$  

(25)
where \( \bar{D} \bar{L}^H \) is the Cholesky factorization of \( \bar{G}^H \bar{G} + \rho I_K \), i.e.,

\[
\bar{G}^H \bar{G} + \rho I_K = \bar{L} \bar{L}^H
\]

while \( \bar{D} \) is a \( K \times K \) diagonal matrix that scales the elements on the main diagonal of \( \bar{C} \) to unity and reads

\[
\bar{D} = \text{diag}\{1/|\bar{L}|_{k,k}; k = 1, 2, \ldots, K\}. \tag{27}
\]

Substituting (25) and (26) into (23) and recalling that \( U = (1/\sqrt{e}) \cdot F \) yields

\[
U = \frac{1}{\sqrt{e}} \cdot \bar{G}(\bar{L}^{-1})^H \bar{D}
\]

where \( e \) is found from (20) after replacing \( F \) with \( \bar{G}(\bar{L}^{-1})^H \bar{D} \). This produces

\[
e = \frac{\text{tr}\{\bar{D}^H \bar{D} - \rho \bar{D}^H (\bar{L}^H \bar{L})^{-1} \bar{D}\}}{K}\tag{29}
\]

where we have borne in mind that \( \bar{G}^H \bar{G} = \bar{L} \bar{L}^H - \rho I_K \), as indicated in (26).

Finally, collecting (24)–(27), it follows that

\[
J = \sigma_n^2 \sum_{k=1}^{K} \frac{1}{|\bar{L}|_{k,k}^2}.
\]

At this stage, we should find the permutation matrix \( P \) that minimizes the RHS of (30), where the dependence on \( P \) is hidden in the quantities \( |\bar{L}|_{k,k} \). Unfortunately, the optimal solution can only be found through an exhaustive search over all \( K! \) permutations of \( a \), which becomes prohibitive even in the presence of a few active users. An interesting alternative to the exhaustive search is the best-first-ordering strategy, which was originally proposed in [14] for V-BLAST and has recently been extended to THP in [12]. This method operates on the rows of the channel matrix, and in many cases, achieves the optimal order with a significant reduction of complexity with respect to the exhaustive search. The best first strategy of [12] can also be applied to our scheme. However, since we are interested in the ultimate performance achievable by the investigated system, in this paper, the exhaustive search is used to find the optimal \( P \). Although computationally demanding, the resulting procedure is still affordable since \( K \leq Q \), and \( Q \) is typically small in practical applications [2, p. 72].

In the sequel, the precoding scheme based on (25) and (27)–(29) is referred to as the nonlinear transmit Wiener filter (NL-TWF) [19]. Calling \( \bar{L}_{\text{opt}} \), the matrix \( \bar{L} \) that corresponds to the optimal choice of \( P \), from (30), it follows that the minimum of \( J \) is given by

\[
J_{\text{min}} = \sigma_n^2 \sum_{k=1}^{K} \frac{1}{|\bar{L}_{\text{opt}}|_{k,k}^2}. \tag{31}
\]

C. Remarks

1) Setting \( B = 0 \) (corresponding to \( C = I_K \) ) yields the linear transmit Wiener filter (L-TWF). In such a case, from (23), we obtain

\[
U = \frac{1}{\sqrt{e}} \cdot \bar{G}(\bar{G}^H \bar{G} + \rho I_K)^{-1}
\]

where we have used the identity \( U = (1/\sqrt{e}) \cdot F \). The scalar \( e \) is computed from (20) and reads

\[
e = \frac{\text{tr}\{\bar{G}(\bar{G}^H \bar{G} + \rho I_K)^{-2} \bar{G}^H \}}{K} \tag{33}
\]

while the sum of MSEs in (24) becomes

\[
J' = \sigma_n^2 \cdot \text{tr}\{(\bar{G}^H \bar{G} + \rho I_K)^{-1}\}
\]

and can also be rewritten as

\[
J' = \sigma_n^2 \sum_{k=1}^{K} \frac{1}{|\bar{L}|_{k,k}^2} + \sigma_n^2 \sum_{k=2}^{K-1} |\bar{L}|_{k,k}^2. \tag{35}
\]

Comparing (35) with (30) indicates that \( J' \geq J \), meaning that L-TWF cannot perform better than NL-TWF. Interestingly, letting \( \bar{G} = GP^T \) into (34) and recalling that \( P^T P = I_K \) yields

\[
J' = \sigma_n^2 \cdot \text{tr}\{P(\bar{G}^H \bar{G} + \rho I_K)^{-1} P^T \}
\]

from which, by using the identity \( \text{tr}\{AB\} = \text{tr}\{BA\} \) with \( A = P \) and \( B = (\bar{G}^H \bar{G} + \rho I_K)^{-1} P^T \), it follows that the performance of L-TWF is independent of the permutation matrix \( P \).

2) In OFDMA, the combining coefficients \( q_m(n) \) \( (n = 1, 2, \ldots, Q) \) at the \( m \)th MT are selected according to (11). Then, from (18), it turns out that \( \bar{G}^H \bar{G} + \rho I_K \) is diagonal and reads

\[
\bar{G}^H \bar{G} + \rho I_K = \text{diag}\{\rho + \sum_{\ell=1}^{N_T} |H_{k,\ell}(i)\|^2; k=1,2,\ldots,K\}. \tag{36}
\]

The previous fact, together with the identity \( \bar{G}^H \bar{G} + \rho I_K = P(\bar{G}^H \bar{G} + \rho I_K)P^T \), indicates that \( \bar{G}^H \bar{G} + \rho I_K \) is diagonal as well and the same occurs with \( L \). Hence, from (27), we see that \( \bar{D} = \bar{L}^{-1} \) so that (25) becomes \( C = I_K \) and \( B \) reduces to the null matrix. This means that NL-TWF boils down to L-TWF when applied to OFDMA transmissions. The physical reason is that separating different users at a subcarrier level eliminates the MAI, thereby avoiding the need for nonlinear precoding. As discussed previously, in this case, the objective function \( J \) becomes independent of \( P \). Thus, letting \( P = I_K \) for simplicity, we have \( \bar{G} = G \) and the forward matrix \( U \) in (32) takes the form

\[
U = \frac{1}{\sqrt{e}} \cdot \bar{G}(\bar{G}^H \bar{G} + \rho I_K)^{-1}. \tag{37}
\]

Substituting (37) into (8) and recalling that \( b = a \) yields

\[
s = \frac{1}{\sqrt{e}} \cdot \bar{G}(\bar{G}^H \bar{G} + \rho I_K)^{-1} a. \tag{38}
\]

Finally, collecting (18), (36), and (38), it follows that the \( s \)th entry of \( s_i \) is given by

\[
s_i(n) = \frac{H_{m,i}(i_n)}{\sqrt{e} \cdot \left[ \rho + \sum_{\ell=1}^{N_T} |H_{m,\ell}(i_n)|^2 \right]^{1/2}}, \quad i = 1, 2, \ldots, N_T
\]

where \( m \) is the index of the user allocated over the \( i_n \)th subcarrier, i.e., \( n = j_m \). The previous equation indicates that, in OFDMA systems, the proposed scheme reduces to MRT over the available transmit antennas, thereby leading to a significant reduction of complexity as compared to MC-CDMA.
The minimum of $J$ is found from (24) after setting $C = I_K$ and $\tilde{G} = G$. Then, from (36), we have

$$J_{\min} = \sigma_n^2 \cdot \sum_{k=1}^K \left[ \rho + \sum_{\ell=1}^{N_T} |H_k,\ell(i_{j_\ell})|^2 \right]^{-1}. \quad (40)$$

3) A possible drawback of NL-TWF is that it requires knowledge of $\rho$. This parameter is related to the noise power $\sigma_n^2$ and is not available at the BS. A simple solution to this problem is found by assuming high signal-to-noise ratio (SNR) values and setting $\rho = 0$. In this case, the processing matrices are still given in (25)–(28) with

$$e = \text{tr} \left\{ \tilde{D}^H \tilde{D} \right\}$$

while $\tilde{L}^H$ is the Cholesky factorization of $\tilde{G}^H \tilde{G}$ instead of $G^H G + \rho I_K$. The previous solution corresponds to a nonlinear transmit zero-forcing (NL-TZF) scheme. This can be seen by substituting (25)–(28) into (17) to obtain

$$y = v + \sqrt{e} \cdot w$$

where we have borne in mind that $F = \sqrt{e} \cdot U$ and $\tilde{G}^H \tilde{G} = \tilde{L}^H \tilde{L}$. Inspection of (42) reveals that there is no MAI at the input of the decision device. Also, recalling that $v = a + d$, and assuming for simplicity that the modulo-operator at the receiver side can suppress the contribution of $d$ without modifying the thermal noise, the SNR at the $m$th MT is given by $\text{SNR}_m = \sigma_n^2 / (e \cdot \sigma_n^2)$, or using (41) and (27)

$$\text{SNR}_m = \sigma_n^2 \cdot \frac{1}{K} \left[ \sum_{k=1}^K \frac{1}{|H_k|^2} \right]^{-1}. \quad (43)$$

The previous equation indicates that $\text{SNR}_m$ is independent of $m$, meaning that NL-TZF ensures the same SNR at all MTs. For this reason, it is particularly suited for commercial applications, where fair treatment of the active users is typically recommended.

4) At this stage, we introduce the nonlinear precoding scheme discussed by Cosovic, Sand, and Raulefs (CSR) in [13], where $C = D L$, $U = G (L^{-1})^H$, and $\tilde{L}^H = G^H \tilde{G}$. Comparing these results with (25) and (28), we see that CSR and NL-TZF only differ in the presence of the scalar factor $1/\sqrt{e}$ and in the different position of $\tilde{D}$ into the expression of the backward and forward matrices. However, selecting $C$ and $U$ as indicated in [13] leads to

$$y_m = \sqrt{e} \cdot [L]_{m,m} v_m + \sqrt{e} \cdot w_m$$

from which it follows that

$$\text{SNR}_m = \frac{\sigma_n^2}{\sigma_n^2} \cdot |L|_{m,m}^2. \quad (45)$$

From the previous equation, we see that the SNRs are user dependent and may be significantly different depending on the actual channel responses. Interestingly, the minimum of $J$ is the same with both CSR and NL-TZF, as given in (31). However, since the error rate performance is dominated by the user with the smallest SNR, significant differences are expected between these schemes in terms of average bit-error-rate (BER).

IV. SIMULATION RESULTS

A. System Parameters

Computer simulations have been run to assess the performance of the proposed THP-based scheme. The simulated system is inspired by the HiperLAN/II standardization project [20] and employs a total of $N = 64$ subcarriers. The latter are divided into 16 groups, each containing $Q = 4$ elements. To better exploit the channel frequency diversity, the subcarriers belonging to the considered group are uniformly distributed over the signal bandwidth with indices $i_n = 16(n-1)$ for $1 \leq n \leq 4$. The signal bandwidth is $B = 20$ MHz, so that the useful part of each block has duration $N/B = 3.2 \mu$s, while the sampling period is $1/B = 5 \times 10^{-7}$ $\mu$s. The channel frequency response between the $i$th transmit antenna and the $m$th MT is expressed by

$$H_{m,i}(\ell) = \sum_{\ell=0}^{L-1} h_{m,i}(\ell) e^{-j 2 \pi \ell i_n / N}, \quad n = 1, 2, \ldots, Q$$

where $h_{m,i} = [h_{m,i}(0), h_{m,i}(1), \ldots, h_{m,i}(L-1)]^T$ represents the corresponding discrete-time channel impulse response (CIR) of length $L = 8$. The entries of $h_{m,i}$ are modeled as independent Gaussian random variables with zero mean and power $E\{|h_{m,i}(\ell)|^2\} = \lambda \cdot \exp(-\ell/4)$ ($\ell = 0, 1, \ldots, 7$), where $\lambda$ is chosen such that the average energy of $h_{m,i}$ is normalized to unity, i.e., $E\{|h_{m,i}|^2\} = 1$. The CIRs are kept fixed over the downlink time slot (slow fading), but vary independently from slot to slot. The transmit antennas are adequately separated so as to make vectors $h_{m,i}$ statistically independent for different values of $i$ and $m$.

We assume an uncoded transmission in which the information bits are mapped onto 16-QAM symbols using a Gray map. The number of active users in the considered group of subcarriers is $K = 4$, corresponding to a fully loaded system.

B. Performance Assessment

Fig. 3 illustrates the objective function $J$ vs. $1/\sigma_n^2$ for an MC-CDMA system employing NL-TWF. The results have been obtained through computer simulations by averaging the RHS of (30) and (31) with respect to the channel statistics. The number of transmit antennas is either $N_T = 1$ or $2$ and only a PD operation is accomplished at the receiver. The curve labeled optimal ordering strategy (OOS) has been obtained by selecting the permutation matrix $P$ that leads to the minimum of $J$ in (31). As mentioned previously, in doing so, we have employed an exhaustive search over all possible permutations, although more practical techniques like the best first strategy [12] can be used without incurring significant degradations with respect to the exhaustive search. The optimal code assignment (OCA) indicates that the spreading codes are dynamically assigned to the active users at each new simulation run. In practice, calling $c_{j_m}$ the code assigned to the $m$th user, we select the indices $\{j_1, j_2, \ldots, j_K\}$ in the set $\{1, 2, \ldots, Q\}$ so as to minimize the RHS of (30). Clearly, in order to avoid that a given code be assigned to more than one user, we must ensure that $j_m \neq j_k$ for $m \neq k$. Again, the optimal mapping $m \rightarrow j_m$ is found through an exhaustive
search over all \( N_a = Q!/(Q-K)! \) possible assignments. The curve labeled OOS&OCA is obtained by looking for the joint optimum ordering and code allocation strategies, and requires a search over \( N_p \cdot N_a = K!Q!/(Q-K)! \) candidates at each new channel realization. Albeit too complex for practical implementation, OOS&OCA is used as a benchmark to assess the ultimate performance achievable by the investigated system. Finally, the curve labeled nonadaptive ordering and code assignment (NOCA) refers to a system where \( P = I_k \) and \( j_k = k \) for \( k = 1,2,\ldots,K \), i.e., the users’ data are not optimally reordered before precoding nor the spreading codes are adaptively assigned to users. The results of Fig. 3 indicate that both OOS&OCA lead to a significant mean square error (MSE) reduction in case of a single transmit antenna, while they only produce marginal gains when \( N_T = 2 \). As expected, increasing \( N_T \) provides the BS with more degrees of freedom to mitigate MAI and channel distortions, thereby leading to a significant improvement of the system performance. In any case, it turns out that OCA practically achieves the same performance of OOS&OCA, and is more beneficial than using OOS singularly. A possible explanation is that dynamic assignment of users’ codes modifies the statistics of the eigenvalues of \( G^H G + \rho I_K \) and provides the system with some form of multiuser diversity [21]. In particular, it is found by simulation that OCA corresponds to a situation in which the condition number (i.e., the ratio between the largest and the smallest eigenvalues) of \( G^H G + \rho I_K \) achieves a minimum among all possible code assignments. On the other hand, the ordering strategy does not provide any multiuser diversity gain since the eigenvalues of \( \tilde{G}^H \tilde{G} + \rho I_K \) are independent of \( P \).

The BER of an MC-CDMA system equipped with NL-TWF is shown in Fig. 4 as a function of \( E_T/N_0 \), where \( E_T \) is the transmitted energy per bit and \( N_0 \) is the two-sided noise power spectral density. The operating conditions are the same as in Fig. 3, i.e., the receiver employs a simple PD detection strategy and the number of transmit antennas is either \( N_T = 1 \) or 2. Again, we see that OCA provides better results than OOS. In particular, for \( N_T = 2 \), OCA has virtually the same performance of OOS&OCA, while for \( N_T = 1 \), the loss is limited to 1 dB at an error rate of \( 10^{-3} \). Note that significant performance degradations occur when NL-TWF operates with NOCA and \( N_T = 1 \).

Fig. 5 compares the BER performance of NL-TWF, NL-TZF, L-TWF, and the CSR scheme discussed in [13] when applied to an MC-CDMA system employing a PD operation at the receiver side. For a fair comparison with L-TWF, results for NL-TWF, NL-TZF, and CSR have been obtained without any OCA or OOS. We see that NL-TWF outperforms the other schemes,
especially in case of a single transmit antenna. However, NL-TZF is a promising alternative to NL-TWF when $N_T = 2$ as it achieves similar performance without requiring knowledge of $\rho$ at the transmit side.

Fig. 6 illustrates the impact of various SUD strategies on the BER performance of an MC-CDMA system equipped with NL-TWF. In order to assess the ultimate performance achievable by the considered scheme, we only present results obtained with OOS & OCA. It turns out that MRC is the best strategy, while PD performs poorly since it does not exploit any channel state information (CSI) at the receiver side. In particular, at an error rate of $10^{-3}$, the loss of PD with respect to MRC is approximately 1 dB with $N_T = 2$, and increases to 2 dB when $N_T = 1$.

We now assess the performance of the proposed prefiltering scheme when used in conjunction with OFDMA. In this case, the sum of MSEs is computed by averaging the RHS of (40) with respect to the channel statistics. The results obtained by means of computer simulations are shown in Fig. 7 as a function of $1/\sigma_n^2$ for either $N_T = 1$ or 2. Here, the optimal subcarrier assignment (OSA) denotes a system in which the available subcarriers are dynamically assigned to users according to the actual channel realization. In this case, the $n$th user transmits on the $j_m$th subcarrier, where the indices $\{j_1, j_2, \ldots, j_K\}$ are selected in the set $\{1, 2, \ldots, Q\}$ so as to minimize the RHS of (40). In doing so, we resort to an exhaustive search over all possible subcarrier assignments, bearing in mind that $j_m \neq j_k$ for $m \neq k$ since each subcarrier must be exclusively assigned to only one user. The curve labeled nonadaptive subcarrier assignment (NSA) refers to a system without OSA where $j_m = m$ for $m = 1, 2, \ldots, K$. As mentioned previously, in OFDMA transmissions, NL-TWF reduces to a simple MRT scheme over the available transmit antennas. In this case, there is no advantage in adopting a particular ordering strategy for the transmitted data symbols, and we may arbitrarily set $\mathbf{P} = \mathbf{I}_K$. The results of Fig. 7 indicate that OSA can significantly reduce the sum of MSEs at the MTs. In particular, OSA with $N_T = 1$ has virtually the same performance of NSA with $N_T = 2$.

Fig. 8 illustrates BER results for an OFDMA system obtained in the same operating conditions of Fig. 7. Again, we see that OSA dramatically improves the system performance with both $N_T = 1$ and 2. In particular, it is worth noting that curves obtained with OSA are asymptotically steeper than those with NSA. This is a manifestation of the multiuser diversity gain that occurs in OFDMA systems in the presence of dynamic subcarrier assignment [21]. Interestingly, we see that employing
OSA with $N_T = 1$ provides a much lower BER than using NSA with $N_T = 2$, although both configurations are characterized by similar performance in terms of MSE (see Fig. 7). A possible explanation is that the BER is strictly related to the statistical distribution of the disturbance term that affects the decision statistic, and accordingly, is not exclusively dictated by the MSE, which only represents the power of the disturbance. This means that systems characterized by the same MSE may exhibit different BER performance.

Fig. 9 illustrates the ultimate BER achievable by the proposed prefiltering technique when applied to MC-CDMA and OFDMA. In particular, OFDMA operates with OSA while MC-CDMA employs OOS & OCA in conjunction with MRC as a detection strategy. We see that OFDMA outperforms MC-CDMA and achieves a gain of approximately 1 dB with both $N_T = 1$ and 2.

C. Computational Complexity

Table I illustrates the overall number of complex operations (products plus additions) required by the investigated schemes at both ends of the wireless link and for each data symbol. In writing these figures, we have considered the computation of the combining coefficients $t_m(n)$ in (16) and (17) as well as the complexity of the receive DFT unit, but not the operations involved with OOS & OCA. Compared to the other schemes, we see that OFDMA is simpler to implement. Recalling the results of Fig. 9, it turns out that OFDMA has some potential advantages over MC-CDMA, as it achieves better error rate performance with a reduced complexity. Actually, the computational load involved with OFDMA is rather limited for the following reasons. First, no data reordering is required at the transmit side. Second, the decision statistic at the $m$th MT is simply given by the $i_m$th output of the DFT unit, whereas in MC-CDMA, it is obtained by linearly combining the DFT outputs and passing the result through a modulo device. As discussed previously, data reordering at the transmitter can be avoided even with MC-CDMA without incurring any significant performance penalty (provided that OCA is employed). However, signal precoding and data detection still remain more complex than in OFDMA.

V. CONCLUSION

We have considered a unified framework for nonlinear prefiltering in MC-CDMA and OFDMA downlink transmissions. The proposed scheme exploits channel knowledge at the transmit side and aims at minimizing the sum of mean square errors at all MTs. In this way, MAI and channel distortions are precompensated at the transmitter and low complex SUD schemes can be employed at the receiver, thereby keeping the power consumption at the MT at a tolerable level. Specifically, the following results have been found.

1) The performance of MC-CDMA is greatly improved if the spreading codes are properly assigned to the active users (OCA). Adopting an optimal precoding order (OOS) in addition to OCA provides marginal gains at the price of a significant increase of complexity.

2) Among the considered SUD techniques, MRC gives the best results in a prefiltered MC-CDMA system operating with OCA and/or OOS.

3) When applied to MC-CDMA, NL-TWF outperforms conventional linear prefiltering techniques as well as other existing nonlinear schemes based on the zero-forcing criterion.

4) In OFDMA applications, the proposed prefiltering scheme reduces to MRT over the available antennas, and the decision statistics at the receiver are directly found at the DFT output without any further processing. This leads to a significant reduction of complexity with respect to MC-CDMA.

5) OSA among the active users produces dramatic improvements in the error rate performance of an OFDMA system.

6) A comparison between OFDMA (with OSA) and MC-CDMA (with OCA and/or OOS) indicates that the former achieves better performance with a reduced complexity.

APPENDIX I

In this appendix, we highlight the major steps leading to (25). Our goal is to find the unit-diagonal and lower triangular matrix $C$ that minimizes the RHS of (24). For this purpose, we consider the Cholesky factorization of $\tilde{G}^H \tilde{G} + \rho I_K$, as given in (26), where $L$ is a $K \times K$ lower triangular matrix with positive real-valued elements $[L]_{k,k}$ ($k = 1, 2, \ldots, K$) on its main diagonal.
Substituting (26) into (24) produces
\[ J = \sigma_n^2 \sum_{k=1}^{K} \frac{1}{|L|^2_{k,k}} + \sigma_n^2 \sum_{k=1}^{K} \sum_{i=1}^{K} |\tilde{L}^{-1}C|_{k,i}^2 \] (48)
where we have borne in mind that \(|C|_{k,k} = 1\) for \(k = 1, 2, \ldots, K\). From (48), it follows that the minimum of \(J\) is achieved when \(\tilde{L}^{-1}C\) is diagonal, i.e.,
\[ \tilde{L}^{-1}C = \tilde{D}. \] (49)
Finally, premultiplying both sides of (49) by \(\tilde{L}\) produces the result (25) in the text.

REFERENCES
