Receiver Cancellation Technique for Nonlinear Power Amplifier Distortion in SDMA–OFDM Systems

Fernando Gregorio, *Student Member, IEEE*, Stefan Werner, *Member, IEEE*, Timo I. Laakso, *Senior Member, IEEE*, and Juan Cousseau, *Senior Member, IEEE*

***Abstract*—Space-division multiple access (SDMA) and orthogonal frequency-division multiplexing (OFDM) can be combined to design a robust communications system with increased spectral efficiency and system capacity. This combination is one of the most promising candidates for future wireless local area network implementations. However, one drawback of OFDM systems is the high peak-to-average power ratio, which imposes strong requirements on the linearity of power amplifiers (PAs). Such linearity requirements translate into high back-off that results in low power efficiency. In order to improve power efficiency, a *PA nonlinearity cancellation* (PANC) technique is introduced in this paper. This technique reduces the nonlinear distortion effects on the received signal. The performance of the new technique is evaluated with simulations, which show significant power efficiency improvements. To obtain meaningful results for comparison purposes, we derive a theoretical upper bound on the bit error rate performance of an SDMA-OFDM system subject to PA nonlinearities. In addition, a novel channel estimation technique that combines frequency- and time-domain channel estimation with PANC is also presented. Simulation results show the robustness of the cancellation method also when channel estimation is included.**

***Index Terms*—Multiuser, orthogonal frequency-division multiplexing (OFDM), power amplifier (PA) nonlinearities cancellation, space-division multiple access (SDMA).**

I. Introduction

SPACE-DIVISION multiple access (SDMA)–orthogonal frequency-division multiplexing (OFDM) can be applied in wireless local area network (WLAN) [1] systems to increase the data rate and the system capacity. Accurate estimation of the uplink user channels enables multiuser detection techniques for user separation. Systems that have low or medium channel mobility, which is the case for WLAN systems, are well suited for that situation since the channel estimation process does not require high complexity.

The combination of several signals with different phases and frequencies that are typical for OFDM systems causes a large peak-to-average power ratio (PAPR) [2]. This can result in considerable distortion effects when the composite signal is amplified by a power amplifier (PA), which typically has nonlinear (NL) characteristics. The high PAPR of OFDM signals require linear PAs with a high dynamic range. However, linear amplifiers tend to have low power efficiency [3], which leads to a reduced battery life, which is a critical resource in mobile systems.

Several techniques that combat NL effects at the transmitter side in OFDM systems have been proposed in the literature, e.g., PAPR reduction via mapping or coding [4], [5], linear scaling [6], clustered OFDM for low PAPR implementations [7], allocation methods that minimize intermodulation products [8], and beamforming designs that employ PAPR constraints [9]. Techniques that are applicable at the transmitter side also include predistorters (PDs) [10] that enable power-efficient amplifiers with reduced out-of-band emission and low waveform distortion. However, the rather high computational complexity that is associated with PDs may prohibit their use in small mobile transceiver structures where low power consumption is required.

Two important aspects for reducing NL PA effects are out-of-band and in-band distortion. The in-band distortion degrades the own bit error rate (BER) performance, whereas the out-of-band distortion affects users that are located in the adjacent frequency bands.

The effect of NL PA on the bit error probability in an SDMA–OFDM system is an important issue that must be considered in a realistic system design. This paper derives an upper bound on the BER for a least squares (LS) detector in an SDMA–OFDM system subject to PA nonlinearities. Our results are not restricted to a particular PA model. Furthermore, we consider the realistic case of low clipping levels, i.e., several clipping events during an OFDM symbol. The assumption of low clipping levels enables us to model the NL distortion as an additive Gaussian noise (see [11]). We also verify the validity of this assumption for our setup through simulations. The solution obtained also provides an upper bound for the achievable BER of more advanced receiver structures, like the minimum mean square error (MMSE) and NL detectors [12], where the derivation of BER expressions is far more complicated.

Manuscript received March 27, 2006; revised August 19, 2006 and December 9, 2006. This work was supported in part by ALβAN, European Union Programme of High Level Scholarships for Latin America, under Contract E03D19254AR and in part by the Nokia Foundation. The review of this paper was coordinated by Prof. L. Lampe.

F. Gregorio, S. Werner, and T. I. Laakso are with the Signal Processing Laboratory, Smart and Novel Radios (SMARAD) Centre of Excellence, Helsinki University of Technology, 02015 Espoo, Finland (e-mail: fernando.gregorio@tkk.fi; stefan.werner@tkk.fi; timo.laakso@tkk.fi).

J. Cousseau is with the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)–Department of Electrical and Computer Engineering, Universidad Nacional del Sur, Bahía Blanca 8000, Argentina (e-mail: jcousseau@uns.edu.ar).

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

Digital Object Identifier 10.1109/TVT.2007.899965

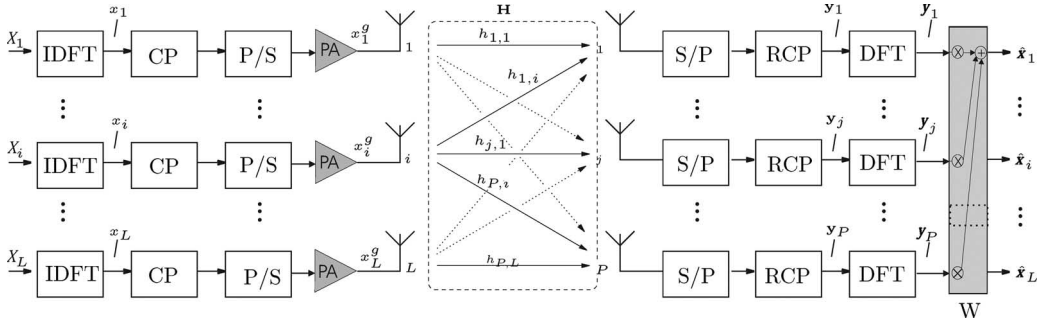


Fig. 1. Block diagram of an SDMA–OFDM system that consists of one base station equipped with P antennas and L mobile users, each with a single antenna element.

The BER upper bound also motivates us to introduce a new cancellation technique that reduces the NL effects. To reduce the NL distortion, we propose the combination of multiuser detection and NL distortion mitigation techniques at the receiver side. The basic idea of the proposed PA nonlinearity cancellation (PANC) technique is given as follows. With an initial estimate of the user symbols, the distortion effects can be estimated if the PA model is known. After that, the NL distortion can be removed from the received signal, and new and improved symbol estimates can be obtained. This procedure can be repeated in an iterative manner to obtain almost undistorted estimates in two or three iterations. This conceptual idea was used in [13] in a single-user wireline system using adaptive OFDM with a large number of carriers. In practice, PA model parameters need to be estimated and sent to the receiver. Our simulations confirm the good performance of the PANC technique for the cases of known and estimated PA model parameters. In our simulation studies, the nonlinearity is modeled as a simple static nonlinearity. However, we stress that the PANC technique can be used with more general models that include memory effects, e.g., Wiener models [10], [14], [15].

Another important issue, which is by far not addressed in the literature, is channel estimation in systems that are affected by nonlinearities. In addition to PANC, we propose a multiuser channel estimation technique that gives accurate results when PA nonlinearities are considered. The novel channel estimation method is combined with PANC for quasi-stationary and time-varying channels. We show that the combination gives mean square error (MSE) levels for the channel estimates that approach the values obtained with linear amplifiers. The evaluation of PANC for low mobility systems and systems with moderate mobility is included to illustrate the proposed methods in a practical context.

In summary, the contributions of the paper are listed as follows:

1) a theoretical BER upper bound for SDMA–OFDM systems with PA nonlinearities;
2) extension of the PANC technique in [13] to the case of mobile SDMA–OFDM systems;
3) a new frequency- and time-domain (FD–TD) channel estimation scheme compatible with the PANC technique.

The paper is organized as follows. Section II reviews concepts of multiuser SDMA–OFDM systems and multiuser detection techniques and introduces the notation to be used in the remaining parts of the article. The theoretical BER bound for SDMA–OFDM systems with PA nonlinearities is derived in Section III. The new PANC technique for the multiuser case is introduced in Section IV. Section V presents a new FD–TD channel estimation approach tailored for the proposed PANC technique. Section VI considers some practical issues that are related to the implementation of the proposed method. Section VII provides simulations of system efficiency, out-of-band distortion analysis, MSE channel estimation, and BER. Finally, conclusions are drawn in Section VIII.

Notation: In this paper, small boldface letters are used to denote vectors, and capital boldface letters are used to denote matrices. In addition, the standard font (Times Roman) is used for time-domain (TD) variables, and calligraphic letters are used to denote frequency-domain (FD) variables. For example, $\mathbf{H}$, $\mathbf{h}$, and h denote a matrix, a vector, and a scalar variable in the TD, respectively. Their corresponding notations in the FD are $\boldsymbol{H}$, $\boldsymbol{h}$, and h, respectively.

II. Multiuser Multiple-Input–Multiple-Output (MIMO)–OFDM Systems

This section presents the SDMA–OFDM system model that takes into account PA effects. Thereafter, we briefly review the most common receiver techniques for separating the user signals at the base station. Of particular interest is the LS detector, which is used in the BER and capacity studies in Sections III and IV, respectively.

A. System Model

The multiuser SDMA–OFDM system under consideration has N subcarriers and consists of one base station that is equipped with P antennas and L mobile users with a single transmit antenna. That results in a $P \times L$ MIMO–OFDM system. It is assumed that all users are simultaneously transmitting independent signals on all N subcarriers.

A block diagram of the system is shown in Fig. 1, where the cyclic prefix insertion/elimination blocks for combatting intersymbol interference are denoted CP and RCP, respectively.

The transmitted signal from user j at time instant n is given by

$$\mathbf{x}_j(n) = \mathbf{G}_{\mathrm{cp}}\bar{\mathbf{x}}_j(n) = \mathbf{G}_{\mathrm{cp}}\mathbf{Q}_N\boldsymbol{x}_j(n) \tag{1}$$

where $\mathbf{G}_{\rm cp}$ is the $(N+v)\times N$ *cyclic prefix insertion matrix* [16], v is the length of the cyclic prefix, $N+v$ is the total length of the OFDM symbol, $\bar{\mathbf{x}}(n)$ is the inverse discrete Fourier transform (IDFT) of the modulated symbols $\boldsymbol{x}_j(n)\in\mathbb{C}^{N\times 1}$ without the cyclic prefix, and $\mathbf{Q}_N$ is the $N\times N$ IDFT matrix.

The multicarrier signal after passing the NL PA $g[\cdot]$ can be written as

$$\mathbf{x}_j^g(n)=g\left[\mathbf{x}_j(n)\right]=K_L\mathbf{x}_j(n)+\mathbf{d}_j(n) \tag{2}$$

where the first term $\mathbf{x}_j(n)$ is the distortion-free discrete-time input signal vector of (1), and K_L is the gain of the linear part. The second term $\mathbf{d}_j(n)$ is the NL distortion, which is a function of the modulated symbol vector $\boldsymbol{x}_j(n)$ and the PA transfer function $g[\cdot]$ (details of some common NL PA models are presented in Appendix A). The value of K_L approaches unity for clipping levels higher than 6 dB [17], i.e., when the constellation scaling is insignificant.

The received signal at antenna i after removing the cyclic prefix $\mathbf{y}_i(n)$ is formed by the superposition of the independently faded signals that are associated with the L users sharing the same space–frequency resource. The received signal, which is assumed to be corrupted by circular complex Gaussian noise at the array elements, is given by

$$\mathbf{y}_i(n)=\sum_{j=1}^{L}\left[K_L\mathbf{H}_{i,j}(n)\bar{\mathbf{x}}_j(n)+\mathbf{H}_{i,j}(n)\mathbf{d}_j(n)\right]+\mathbf{n}_i(n) \tag{3}$$

where $\mathbf{H}_{i,j}(n)$ is an $N\times N$ circulant TD channel matrix at time instant n, which is formed by the channel response vector $\mathbf{h}_{i,j}(n)$ for the link between user j and base station antenna i. The FD expression of the received signal is obtained by taking the discrete Fourier transform (DFT) of (3).

Let $\boldsymbol{y}(n,k)=[y_1(n,k),\ldots,y_P(n,k)]^T$ denote the vector of received signals at each antenna on subcarrier k. Then, the received signal vector for each subcarrier can be written as

$$\boldsymbol{y}(n,k)=K_L\boldsymbol{H}(n,k)\boldsymbol{x}(n,k)+\boldsymbol{H}(n,k)\boldsymbol{d}(n,k)+\boldsymbol{n}(n,k) \tag{4}$$

where $\boldsymbol{H}(n,k)\in\mathbb{C}^{P\times L}$ is the channel transfer matrix, $\boldsymbol{x}(n,k)=[\boldsymbol{x}_1(n,k),\ldots,\boldsymbol{x}_L(n,k)]^T$ is the vector that contains transmitted signals from each user, $\boldsymbol{d}(n,k)=[d_1(n,k),\ldots,d_L(n,k)]^T$ is the vector that contains the NL distortion of each user on subcarrier k, and $\boldsymbol{n}\in\mathbb{C}^{P\times 1}$ is the additive noise assumed to be circular complex Gaussian with $E[\boldsymbol{n}\boldsymbol{n}^H]=\sigma_n^2\mathbf{I}$.

The FD channel transfer matrix $\boldsymbol{H}(n,k)$ in (4) is given by [12]

$$\boldsymbol{H}(n,k)=\begin{bmatrix}h_{1,1}(n,k) & h_{1,2}(n,k) & \cdots & h_{1,L}(n,k)\\ h_{2,1}(n,k) & h_{2,2}(n,k) & \cdots & h_{2,L}(n,k)\\ \vdots & \vdots & \ddots & \vdots\\ h_{P,1}(n,k) & h_{P,2}(n,k) & \cdots & h_{P,L}(n,k)\end{bmatrix}$$

where $h_{i,j}(n,k)$ denotes the channel response on subcarrier k at time n between antenna element i of the base station and user j.

B. FD Detectors

An estimate $\hat{\boldsymbol{x}}(n,k)$ of the L user transmitted signals $\boldsymbol{x}(n,k)$ can be obtained by linearly combining the signals at the P receive antennas using a weight matrix $\boldsymbol{W}\in\mathbb{C}^{P\times L}$ as follows:

$$\hat{\boldsymbol{x}}(n,k)=\boldsymbol{W}^H\boldsymbol{y}(n,k). \tag{5}$$

The standard LS combiner $\boldsymbol{W}=\boldsymbol{W}_{\rm LS}$ is given by [18]

$$\boldsymbol{W}_{\rm LS}=\frac{1}{K_L}\boldsymbol{H}(n,k)\left[\boldsymbol{H}^H(n,k)\boldsymbol{H}(n,k)\right]^{-1}. \tag{6}$$

Alternatively, an MMSE detector can be used. This alternative exploits the available statistical knowledge of the noise and makes a tradeoff between the multiuser interference (MUI) and the measurement noise. The optimal MMSE weight is obtained as follows:

$$\boldsymbol{W}_{\rm MMSE}=\frac{1}{K_L}\left[\boldsymbol{H}(n,k)\mathbf{P}_\rho\boldsymbol{H}^H(n,k)+\sigma_n^2\mathbf{I}\right]^{-1}\boldsymbol{H}(n,k)\mathbf{P}_\rho \tag{7}$$

where matrix $\mathbf{P}_\rho$ is an $(L\times L)$ diagonal matrix of the form $\mathtt{diag}(\rho_1,\ldots,\rho_L)$, with $\rho_j=\sigma_j^2/\sigma_n^2$ being the signal-to-noise ratio (SNR) of user j.

Other options for user separation include NL techniques, e.g., parallel interference cancellation (PIC) and successive interference cancellation (SIC) [12]. This paper will only consider the use of the linear LS and MMSE receivers.

III. BER Upper Bound in an SDMA–OFDM System With an NL PA

In this section, a theoretical BER analysis is carried out for the case when an LS receiver is used to separate the transmitted multiuser signal. The obtained results provide an upper bound for the performance of different receiver structures. In particular, the LS performance will approach that of the MMSE detector for high SNRs.

In the analysis, we use the following assumptions.

- A1) We have low and medium clipping levels, i.e., the input saturation voltage of the PA is fixed at a low level, creating several clipping events during an OFDM symbol.
- A2) The FD channel coefficients $\boldsymbol{h}_{i,j}(n,k)$ are assumed to be independent stationary zero-mean unit variance circular complex Gaussian-distributed processes.
- A3) The LS detector defined in (6) is used for user separation.
- A4) A low channel noise approximation is used.

Assumption A1) enables us to consider the distortion term in (3) to be additive Gaussian noise with a variance that is equal to σ_d^2 (see [11] for details).

To validate assumption A1), Fig. 2 shows a Gaussian complementary cumulative distribution function (CCDF) with an adequate variance, and the simulation results for an OFDM system with $N=512$ subcarriers and QPSK modulation. As can be observed, the curves show good agreement when

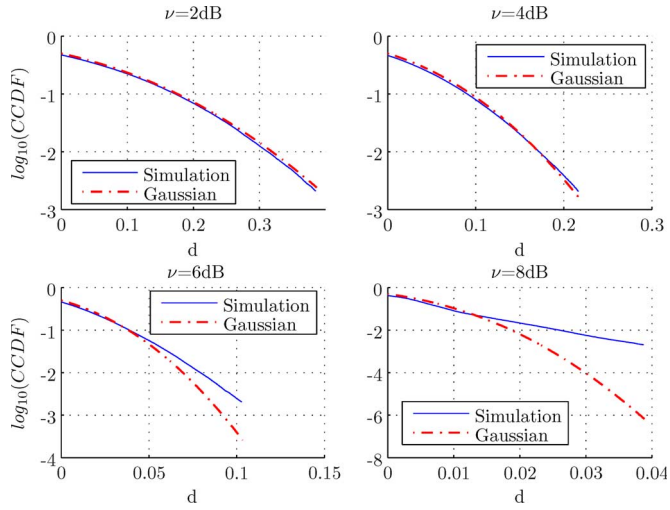


Fig. 2. CCDF of distortion noise for the SSPA model with $p = 2$.

clipping levels are lower than 6 dB. The clipping level ν is defined as

$$\nu = \frac{A_s}{\sqrt{E\left\{|x(n)|^2\right\}}} \tag{8}$$

where A_s is the amplifier input saturation, $\sqrt{E\{|x(n)|^2\}}$ is the root-mean-square value of the OFDM signal. From the above, we conclude that the distortion noise can be considered as additive Gaussian noise if clipping levels are lower than 6 dB. Higher clipping levels can be analyzed using the impulsive noise model proposed in [19].

Based on assumptions A1)–A4), the derivation of a BER upper bound follows an approach that is similar to that presented in [20], in which the authors present an analysis of the error probability for a system with antenna diversity.

Applying the LS detector to the signal model of (4), the following estimate of the transmitted signal is obtained:

$$\begin{aligned}\hat{\boldsymbol{x}}_{\mathrm{LS}}(n,k) &= \boldsymbol{W}_{\mathrm{LS}}^{H}\left[\boldsymbol{H}(n,k)\boldsymbol{x}(n,k)\right. \\ &\qquad \left. + \boldsymbol{H}(n,k)\boldsymbol{d}(n,k) + \boldsymbol{n}(n,k)\right] \\ &= \boldsymbol{x}(n,k) + \frac{\boldsymbol{d}(n,k)}{K_L} + \boldsymbol{W}_{\mathrm{LS}}^{H}\boldsymbol{n}(n,k). \end{aligned} \tag{9}$$

The LS estimate $\hat{\boldsymbol{x}}_{\mathrm{LS}}(n,k)$ is a noisy estimate of the original vector $\boldsymbol{x}(n,k)$ and can be modeled by a Gaussian distribution with mean value μ and covariance matrix $\mathbf{R}_{\mathrm{LS}}$, where $\mu = E_x\{\hat{\boldsymbol{x}}_{\mathrm{LS}}(n,k)\} = \boldsymbol{x}(n,k)$, and $\mathbf{R}_{\mathrm{LS}} = E_x\{[\boldsymbol{W}_{\mathrm{LS}}^{H}\boldsymbol{n}(n,k)] + \boldsymbol{d}(n,k)/K_L]\,[\boldsymbol{W}_{\mathrm{LS}}^{H}\boldsymbol{n}\,(n,k) + \boldsymbol{d}(n,k)/K_L]^{H}\} = \sigma_n^2[\boldsymbol{H}^{H}(n,k)\boldsymbol{H}(n,k)]^{-1}/K_L^2 + \sigma_d^2/K_L^2$ (E_x denotes conditional expectation with respect to x).

In the following equations, time index n and frequency index k are dropped to simplify the notation. Under assumption A2), the MSE of the LS receiver for user j is given by [12]

$$\xi_j = \frac{\left[\boldsymbol{H}^{H}(n,k)\boldsymbol{H}(n,k)\right]_{j,j}^{-1}\sigma_n^2 + \sigma_d^2}{K_L^2} \tag{10}$$

where $1/[\boldsymbol{H}^{H}\boldsymbol{H}]_{j,j}^{-1}$, which corresponds to the inverse of element (j,j) of matrix $[\boldsymbol{H}^{H}\boldsymbol{H}]^{-1}$, is a scalar random variable with a chi-square *distribution* with $2(P-L+1)$ degrees of freedom [20]. Defining $\gamma = (1/[\boldsymbol{H}^{H}\boldsymbol{H}]_{j,j}^{-1})/\sigma_n^2$ as the instantaneous SNR, the MSE can be written in a compact form as

$$\xi_j = \frac{\sigma_a^2}{K_L^2\gamma} + \frac{\sigma_d^2}{K_L^2} \tag{11}$$

where σ_a^2 is the energy of the transmitted complex-valued data symbols.

The conditional error probability $P_e(E|\gamma)$ and the MSE can be related by [20]

$$P_e(E|\gamma) \le \exp\left(-\frac{1}{\xi_j}\right) \tag{12}$$

where the approximation is valid for low noise levels [see Assumption A4)]. Combining (11) and (12) results in the following upper bound on the conditional error probability:

$$P_e(E|\gamma) \le \exp\left(-\frac{K_L^2\gamma}{\sigma_d^2\gamma + \sigma_a^2}\right). \tag{13}$$

To determine the average bit error probability, the conditional probability of error must be averaged over the fading channel statistics as

$$\bar{P}_e \le \int_0^{\infty} P_e(E|\gamma)P(\gamma)d\gamma \tag{14}$$

where $P(\gamma)$ is the statistics of the channel, which is a chi-square probability density function (pdf) with $2(P-L+1)$ degrees of freedom given by

$$P(\gamma) = \frac{1}{(P-L)!\bar{\gamma}^{P-L+1}}\gamma^{P-L}\exp\left(-\frac{\gamma}{\bar{\gamma}}\right) \tag{15}$$

and $\bar{\gamma} = E[\gamma]$ is the average SNR. Substituting the pdf (15) and the conditional bit error probability (13) in the integral (14) gives

$$\begin{aligned}\bar{P}_e \le{}& \frac{1}{(P-L)!\bar{\gamma}^{P-L+1}} \\ &\times \int_0^{\infty} \gamma^{P-L}\exp\left(-\frac{K_L^2\gamma}{\sigma_d^2\gamma + \sigma_a^2} - \frac{\gamma}{\bar{\gamma}}\right)d\gamma. \end{aligned} \tag{16}$$

Using a Taylor expansion of the exponential function in (16), the error probability can be expressed as

$$\begin{aligned}\bar{P}_e \le{}& \frac{1}{(P-L)!\bar{\gamma}^{P-L+1}}\sum_{m=0}^{\infty}\frac{(-1)^m(m+P-L)!}{m!} \\ &\times \left(\frac{1}{\sigma_d^2}\right)^{m+P-L+1} K_L^{2m}U_{P-L+m+1,\frac{P-L+2}{2}}\left(\frac{\sigma_a^2}{\sigma_d^2\bar{\gamma}}\right) \end{aligned} \tag{17}$$

where $U_{P-L+m+1,(P-L+2/2)}(\sigma_a^2/\sigma_d^2\bar{\gamma})$ is the *Confluent Hypergeometric Function* [21] defined as

$$U_{a,b}(z) = \frac{1}{\Gamma(a)} \int_0^\infty \exp(-zt) t^{a-1}(1+t)^{b-a-1} dt. \quad (18)$$

Extensive simulations show that a close fit to the infinite sum of the original expression (see Appendix B) is obtained by employing only the first 20–30 terms.

IV. PANC Technique

In this section, we propose a PANC technique to combat the in-band distortion that is created by an NL PA.

We start by studying the effect of PA-induced NL distortion on the system capacity. A closed-form capacity expression is obtained for the LS receiver. After that, the novel PANC technique is developed for use in an LS receiver structure.

A. NL PA Effects on System Capacity

Let us drop the time and frequency indices in (4), and let $\boldsymbol{w}$ denote the additive distortion, i.e., $\boldsymbol{w} = \boldsymbol{Hd} + \boldsymbol{n}$. The interference vector $\boldsymbol{w}$ is uncorrelated with the transmitted signal. The channel capacity for the system in (4) can be written as [22]

$$C = E_{\boldsymbol{H}} \left[\log_2 \left\{ \frac{\det(\mathbf{R}_w + \boldsymbol{H}\mathbf{R}_x\boldsymbol{H}^H)}{\det(\mathbf{R}_w)} \right\} \right] \quad (19)$$

where the expectation operation $\mathbf{E}_{\boldsymbol{H}}[\cdot]$ is over the random channel matrix $\boldsymbol{H}$, $\mathbf{R}_x$ is the correlation matrix of the transmitted signal vector given by $\mathbf{R}_x = K_L^2\sigma_j^2\mathbf{I}_L$. The correlation matrix $\mathbf{R}_w$ of the uncorrelated interference terms, namely, channel noise and NL distortion, can be expressed as $\mathbf{R}_w = E_{\boldsymbol{H}}[(\boldsymbol{Hd}+\boldsymbol{n})(\boldsymbol{Hd}+\boldsymbol{n})^H] = \sigma_d^2\boldsymbol{HH}^H + \sigma_n^2\mathbf{I}_P$. Therefore, substitution in (19) and using singular value decomposition for the product $\boldsymbol{HH}^H$ [23], the capacity becomes

$$\begin{aligned} C &= E\left[\log_2 \left\{ \prod_{i=1}^{\operatorname{rank}(\boldsymbol{HH}^H)} \left(1 + \frac{K_L^2\sigma_j^2\lambda_i}{\sigma_n^2 + \sigma_d^2\lambda_i} \right) \right\} \right] \\ &= E\left[\sum_{i=1}^{\operatorname{rank}(\boldsymbol{HH}^H)} \log_2 \left\{ 1 + \frac{K_L^2\sigma_j^2\lambda_i}{\sigma_n^2 + \sigma_d^2\lambda_i} \right\} \right] \end{aligned} \quad (20)$$

where λ_i is the ith eigenvalue of $\boldsymbol{HH}^H$, $E[\cdot]$ is the expectation over the eigenvalues λ_i, and $(\cdot)$ denotes the rank of the matrix.

From (20), we see how the capacity is reduced when the NL distortion is considered. This is because λ_i is positive, and the logarithmic function is monotonically increasing. Note also that when σ_d^2 increases, K_L decreases.

In case a least squares receiver is employed, the capacity will be affected by noise enhancement. Following the same steps as above, the capacity for user j, assuming an LS equalizer, can be derived from (9) as follows:

$$\begin{aligned} C_{\mathrm{LS}} &= E\left[\log_2 \left\{ 1 + \frac{\sigma_j^2}{\sigma_n^2[\boldsymbol{H}^H\boldsymbol{H}]_{j,j}^{-1} + \sigma_d^2/K_L^2} \right\} \right] \\ &= E\left[\log_2 \left\{ 1 + \frac{\sigma_j^2\gamma}{\sigma_a^2 + \sigma_d^2\gamma/K_L^2} \right\} \right] \end{aligned} \quad (21)$$

where the instantaneous SNR $\gamma = (1/[\boldsymbol{H}^H\boldsymbol{H}]_{j,j}^{-1})/\sigma_n^2$ is the same chi-square-*distributed* variable with $2(P-L+1)$ degrees of freedom that was used in the BER derivation of Section III. The capacity is obtained by integrating (21) over the chi-square *distribution* defined in (15). This integral can be numerically evaluated. However, assuming a high SNR approximation, the capacity can be solved for in closed form as (see Appendix C)

$$\begin{aligned} C_{\mathrm{LS}} \approx{}& \log_2\left(1 + \frac{1}{\sigma_{d'}^2}\right) - \frac{\sigma_a^2}{\ln(2)} \\ &\times \left[\ln(\beta) + \gamma_e - \exp(\beta)\left([1+\beta u(P-L)] \right.\right. \\ &\left.\left. \times \mathrm{E}i(-\beta) + \sum_{k=2}^{P-L} \frac{\beta^k \Gamma_I(1-k,\beta)}{k} \right) \right] \end{aligned} \quad (22)$$

where $\beta = (\sigma_a^2/(\sigma_d^2/K_L^2))\bar{\gamma}$, γ_e is the Euler–Mascheroni constant ($\gamma_e = 0.5772\ldots$) [24], $\mathrm{E}i(\cdot)$ is the exponential integral function, and $\Gamma_I(\cdot)$ is the incomplete gamma function [21]. The function $u(P-L)$ is the unit step function defined as $u(x) = 1\ \forall x > 0$ and $u(x) = 0\ \forall x \le 0$.

Equation (22) was evaluated for $L = 1$ and $L = 2$ users and $P = 4$ for clipping levels of $\nu = 1$ dB, $\nu = 2$ dB, and $\nu = 4$ dB, and the results are illustrated in Fig. 3. The exact capacity values were obtained using numerical evaluation. We see that (22) provides a tight approximation for SNR levels higher than 5 dB. In order to evaluate the effect of NL distortion, the linear case was also included. The asymptotic capacity is also plotted in Fig. 3 and is given by

$$C_{\mathrm{LS}}^{\infty} = \lim_{\bar{\gamma}\to\infty} C_{\mathrm{LS}} = \log_2\left(1 + \frac{K_L^2}{\sigma_d^2}\right). \quad (23)$$

It is clear from the expression for C_{LS}^{∞} that the capacity will be bounded due to the NL distortion.

B. PANC Technique

The previous basic analysis shows that system capacity is strongly affected by NL distortion. The result motivates the introduction of a new cancellation technique in order to keep system capacity close to that of the linear case.

Assuming that the transmitter nonlinearity is known at the receiver, the receiver can compute and estimate $\hat{\boldsymbol{d}}(n)$ from the received vector $\boldsymbol{x}(n,k)$. An initial estimate of vector $\hat{\boldsymbol{x}}(n,k)$ can be used to calculate the distortion vector $\boldsymbol{d}(n,k)$. The estimated distortion vector is removed from the original received vector, and a new estimation of $\boldsymbol{x}(n,k)$ can be obtained.

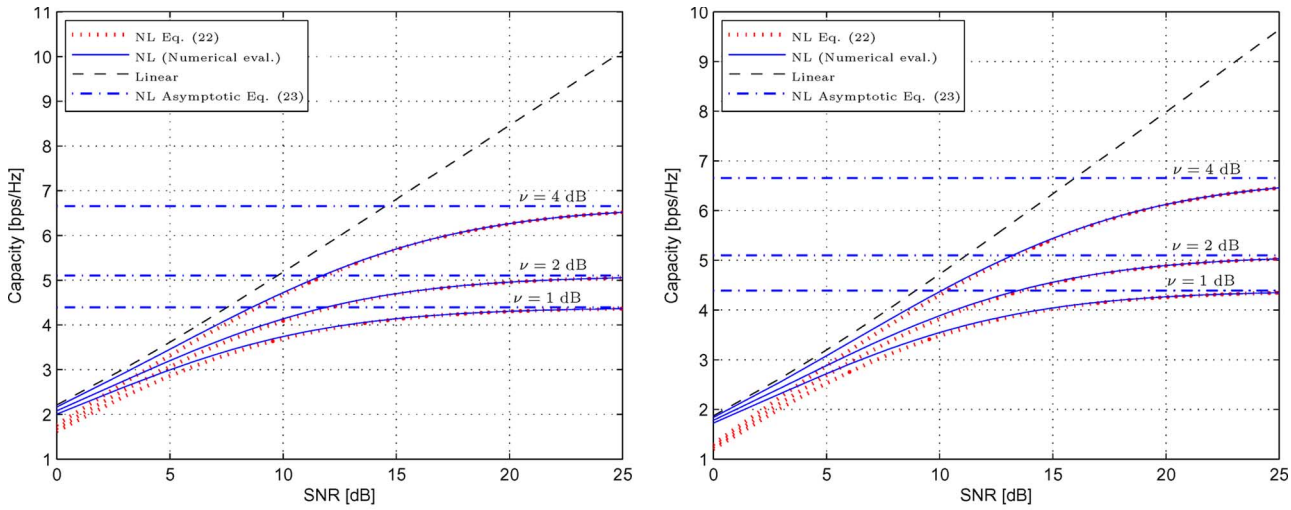


Fig. 3. Capacity evaluation for an LS receiver with $L = 1$ and $L = 2$ users, $P = 4$ antennas, and clipping levels of $\nu = 1$ dB, $\nu = 2$ dB, and $\nu = 4$ dB. The exact value of capacity is obtained using numerical integration, the high SNR approximation is computed using (22), and the linear case was also included for comparison.

This second estimate $\hat{\boldsymbol{x}}(n,k)$ can be used to re-estimate the distortion vector. This process can be iteratively performed until some specified bound is reached.

Using the LS detector of (6), the estimation of the jth user signal is given by

$$\hat{\boldsymbol{x}}_j(n,k) = \frac{1}{K_L}\left[\boldsymbol{H}(n,k)\mathrm{col}_j \times \left\{\left[\boldsymbol{H}(n,k)^H \boldsymbol{H}(n,k)\right]^{-1}\right\}\right]^H \boldsymbol{y}(n,k) \quad (24)$$

where $\mathrm{col}_j\{\cdot\}$ denotes the jth column vector of the corresponding matrix. Using $\hat{\boldsymbol{x}}_j$, an estimate $\boldsymbol{z}_j$ of the original transmitted constellation $\boldsymbol{x}_j$ is obtained by applying hard decoding (or more efficient methods if coding is employed). This process is carried out for all active carriers. Using the recovered symbols, the TD signal is reproduced via IDFT as $\hat{\mathbf{x}}_j(n) = \mathbf{Q}_N \boldsymbol{z}_j(n)$.

Assuming that the NL model of the PA, i.e., $g[\cdot]$, is known at the receiver (details of $g[\cdot]$ are discussed in Section VI and Appendix A), the distortion term can be obtained using (2) as follows:

$$\hat{\mathbf{d}}_j(n) = g\left[\hat{\mathbf{x}}_j^{(m)}(n)\right] - K_L \hat{\mathbf{x}}_j^{(m)}(n) \quad (25)$$

where $\hat{\mathbf{x}}_j^{(m)}(n)$ is the TD representation of the recovered signal for user j at iteration m. The FD distortion term is obtained applying the DFT operator

$$\hat{\boldsymbol{d}}_j(n) = \mathbf{Q}_N^H \left\{g\left[\hat{\mathbf{x}}_j^{(m)}(n)\right] - K_L \hat{\mathbf{x}}_j^{(m)}(n)\right\}. \quad (26)$$

The distortion $\hat{d}_j(n) = [\hat{d}_j(n,1)\cdots\hat{d}_j(n,N)]^T$ is subtracted from the estimated signal $\hat{\boldsymbol{x}}_j^{(m)}$. Using this result, the transmitted constellation is re-estimated in a new decoding/distortion cancellation step. The process can be carried out iteratively. Our simulation study, which is presented in Section VI, suggests that two or three iterations are usually sufficient.

The PANC technique for an SDMA–OFDM system with an LS receiver is summarized in Table I. Alternatively, the PANC technique can be combined with an MMSE receiver, which estimates user signal j using (7).

V. PANC With Channel Estimation

The PANC technique proposed in Section IV assumes perfect knowledge of the channel. In practice, the channel needs to be estimated, and estimation errors will degrade the overall performance. There are two fundamental approaches for channel estimation. One approach assumes that a whole OFDM symbol is periodically transmitted, i.e., pilots are transmitted on all subcarriers. The second approach uses a reduced set of dedicated subcarriers for pilot data, and the channel estimation is performed using interpolation. The second approach, where each user has its own dedicated pilot subcarriers, is addressed in this paper.

Channel estimation for the stationary case is executed in the initialization. The obtained estimates can be used for relatively long time until the channel conditions are modified. In the case of time-varying channels, tracking capability must be considered, and the channel estimation process needs to be carried out for each OFDM symbol.

In the following, we discuss issues that are related to the choice of the training symbols and the number of pilot subcarriers employed for systems with NL PA. The objective is to define an adequate number of pilot subcarriers, which reduces the NL PA effects on channel estimates. Finally, a channel estimation technique with tracking capabilities to combat NL distortion and MUI is introduced.

The novel combination of FD and TD channel estimation in addition to PANC is developed in Section V-B. The main motivation behind a combined TD and FD scheme is to improve the estimation accuracy for a normal overhead (i.e., pilot carriers). If only FD channel estimation was used, the performance of the system will be severally degraded when a normal overhead

TABLE I
NL CANCELLATION TECHNIQUE (PANC) ALGORITHM

$\boldsymbol{H}$: Frequency domain channel

$g[\cdot]$: PA transfer function (see Appendix A)

For $j = 1$ **to** L **number of users**

For $m = 1$ **to number of iterations**

For $k = 1$ **to** N

Estimate the jth user signal using LS receiver:

$$\hat{X}_j(n,k) = \frac{1}{K_L}\left[\boldsymbol{H}(n,k)\mathrm{col}_j\{[\boldsymbol{H}(n,k)^{\mathrm{H}}\boldsymbol{H}(n,k)]^{-1}\}\right]^{\mathrm{H}}\boldsymbol{y}(n,k)$$

if $m = 1$

Demodulate the jth user symbol:

$$Z_j^{(m)}(n,k) = \left\langle \hat{X}_j(n,k)\right\rangle^{1}$$

else

Demodulate the jth user symbol:

$$Z_j^{(m)}(n,k) = \left\langle \hat{X}_j(n,k) - \hat{d}_j^{(m)}(n,k)\right\rangle^{1}$$

end

end

Transmitted symbols in time domain:

$$\mathbf{x}_j^{(m)}(n) = \mathbf{Q}_N \boldsymbol{Z}_j^{(m)}(n)$$

Calculate the distortion term:

$$\hat{\mathbf{d}}_j(n) = g[\mathbf{x}_j^{(m)}(n)] - K_L\mathbf{x}_j^{(m)}(n)$$

Obtain the distortion in frequency domain:

$$\hat{\boldsymbol{d}}_j^{(m)}(n) = \mathbf{Q}^{\mathrm{H}}\hat{\mathbf{d}}_j(n)$$

end

end

[1] $\langle\ ,\rangle$ denotes the symbol decision process in which the decoder chooses the value of $Z_j^{(m)}(n,k)$ that is closest in Euclidean distance to $\hat{X}_j(n,k)$.

is used. In order to improve the accuracy, more pilot tones are required, which reduces the effective data rate of the system. The proposed combination of FD and TD channel estimation techniques that are presented in Section V-B outperforms the results that are obtained with conventional FD techniques, as shown in [25] and [26].

A. Channel Estimation Considerations

It is known that PA nonlinearities have a strong impact on the channel estimation process [27]. Therefore, it is important that training symbols are not affected by nonlinearities in order to obtain good channel estimates. One option, which is applied in HIPERLAN II, is to design training symbols with low PAPR.

In the case of an SDMA–OFDM system, a group of equally spaced subcarriers can be assigned to each user. With N nonzero subcarriers and L users, a maximum of N/L subcarriers can be allocated to each user, i.e., all users transmit their training symbols on nonoverlapping subcarriers. The advantage of this method is that each user has only a group of N/L active subcarriers during the training period, which reduces the PAPR of the OFDM signal [7]. This is explained by the fact that the pdf of the OFDM signal with a large number of carriers can be well approximated with a Gaussian distribution $\mathcal{N}(N\mu, N\sigma^2)$. Therefore, if only N/L carriers are active, the variance of the OFDM signal is reduced, and the PA mostly operates in the linear region. Thus, the effect of nonlinearities is reduced in the channel estimation process.

However, there is a tradeoff between the number of pilot carriers (maximum N/L), the channel estimation accuracy, and the level of NL distortion. The accuracy of the channel estimation is related to the number of subcarriers that are used for training, i.e., the interpolation error is reduced if the number of subcarriers is increased. Therefore, from a channel estimation point of view, the more pilot carriers used, the better. On the other hand, increasing the number of subcarriers leads to an increased PAPR, i.e., increased NL distortion. Therefore, a total of $T \times L < N$ subcarriers are reserved for pilot data, where T is chosen to trade off the estimation accuracy and distortion level.

Fig. 4 illustrates the pilot carrier allocation strategy for a system with $L = 4$ users. In the figure, both stationary and time-varying channels are considered. For the case of a stationary channel, the OFDM symbol includes only pilot carriers (shown

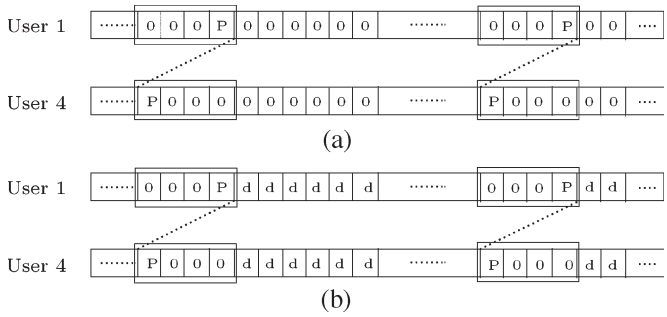


Fig. 4. Pilot carrier allocation. P: Pilot subcarriers. d: Data subcarriers. (a) Stationary channel. (b) Time-varying channel.

with P in the figure). In this case, as previously commented, the channel estimation process is performed during the initialization. For time-varying channels, where the channel estimation is performed for each OFDM symbol, these symbols include pilot subcarriers and data subcarriers (shown with d in the figure). The level of overhead is then equal to $O = (TL/N)$.

Our simulation results suggest that $T = 32$ pilot subcarriers per active user is a good value for a system with $L = 4$ users and $N = 512$ subcarriers, considering low clipping levels in stationary channels. In the case of time-varying channels, where tracking capabilities need to be considered, the number of pilot subcarriers is reduced to $T = 16$ in order to obtain a reduced level of overhead.

B. FD–TD Channel Estimation With PANC

In this section, an iterative channel estimation technique that combines FD–TD channel estimation with PANC is discussed. The technique is applicable to estimation of both stationary (e.g., WLAN) and nonstationary channels (e.g., mobile systems). The basic idea is to use the equalized signals from the FD processing to remove the NL distortion (and MUI in case of nonstationary channels) and improve the channel estimate in TD.

The combined FD–TD channel estimation with PANC can be summarized in the following two steps:

Step 1) *FD channel estimation with PANC.* For the FD channel estimation, we will assume that each user is given a set of T dedicated pilot carriers according to Fig. 4. For stationary channels, pilot carriers are used only during the initialization. Let $\mathcal{T}_j(k_{j,1}, \ldots, k_{j,T})$ denote the set that specifies the T pilot carriers of user j. The channel frequency response coefficients can be estimated over these subcarriers as in a single-user system. In the case of a stationary channel, only pilot data are transmitted (see Fig. 4). This allows us to exactly reproduce the transmitted TD signal and estimate the NL distortion using (21). Thereafter, channel estimation may be carried out on the pilot subcarriers ($k \in \mathcal{T}_j$) using the following expression:

$$\hat{h}_{i,j}(n,k) = \frac{y_i(n,k)}{K_L x_j(n,k) + \hat{d}_j(n,k)} \tag{27}$$

where $y_i(n,k)$ is the received signal at antenna i on subcarrier $k \in \mathcal{T}_j$, and $x_j(n,k)$ is the training symbol that is transmitted by user j at the corresponding subcarrier.

Collecting the channel estimates obtained from the pilot carriers for user j in vector $\hat{\boldsymbol{h}}^{c}_{i,j}(n) = [\hat{h}_{i,j}(n,k_{j,1}), \ldots, \hat{h}_{i,j}(n,k_{j,T})]^T$, the whole channel FD response can be obtained through interpolation using truncated DFT matrices [28], i.e.,

$$\hat{\boldsymbol{h}}_{i,j}(n) = \mathbf{Q}_N \left[\mathbf{Q}^H_{\mathcal{T}_j}\mathbf{Q}_{\mathcal{T}_j}\right]^{-1} \mathbf{Q}^H_{\mathcal{T}_j}\hat{\boldsymbol{h}}^{c}_{i,j}(n) \tag{28}$$

where $\mathbf{Q}_{\mathcal{T}_j}$ is formed by the T columns of the IDFT matrix $\mathbf{Q}_N$ associated with the initially estimated subcarriers and the L_h rows associated with the nonzero TD channel taps.

In case of a nonstationary channel, pilot symbols are tone multiplexed with useful data (see Fig. 4). This means that the user symbols on the data carriers need to be estimated before we can construct the TD signal that is required for estimating $d_j(n,k)$ through (21). Therefore, an *a priori* channel estimate that enables us to detect each user signal is required (here, we use the estimate $\hat{h}_{i,j}(n-1,k)$). After acquiring an estimate of the NL distortion, (27) is used to obtain the new channel estimates over the pilots $k \in \mathcal{T}_j$ at time n. Finally, (28) provides the channel for all subcarriers. Using the FD channel estimate, each user's signal is detected and decoded to obtain $\boldsymbol{z}_j(n) \in \mathbb{C}^N$, $j = 1, \ldots, L$.

Step 2) *TD channel estimation with PANC.* In this step, TD channel estimation is performed. With NL distortion estimates $\{\hat{\boldsymbol{d}}_j(n)\}_{j=1}^{L}$ and detected user symbols $\{\boldsymbol{z}_j(n)\}_{j=1}^{L}$ (or known pilots in case of stationary channel) obtained from Step 1), we can now cancel their effect from the TD received signal $\mathbf{y}_i(n)$. Hence, we have

$$\bar{\mathbf{y}}_i(n) = \mathbf{y}_i(n) - \mathbf{Q}_N \times \sum_{j=1;i\neq j}^{L} \Lambda_{i,j}(n)[\boldsymbol{z}_j(n) + \hat{\boldsymbol{d}}_j(n)] \tag{29}$$

where $\Lambda_{i,j}(n) = \mathtt{diag}[\hat{\boldsymbol{h}}_{i,j}(n)]$ is a diagonal matrix that contains the FD channel estimates obtained in the FD step.

After removal of the MUI and NL distortion, TD channel estimation can now be performed on the single-user signal $\bar{\mathbf{y}}_i(n)$ using an LS approach [26], [29], i.e.,

$$\hat{\mathbf{h}}_{i,j}(n) = \left[\mathbf{C}^H_j(n)\mathbf{C}_j(n)\right]^{-1}\mathbf{C}^H_j(n)\bar{\mathbf{y}}_i(n) \tag{30}$$

where $\mathbf{C}_j$ is the $N \times L_h$ circulant matrix formed by the TD transmitted vector $\mathbf{x}_j$ (i.e., each column of $\mathbf{C}_j$ equals the previous column rotated downward by one element [23]). A new iteration can be performed in TD in order to improve the channel estimation accuracy [29].

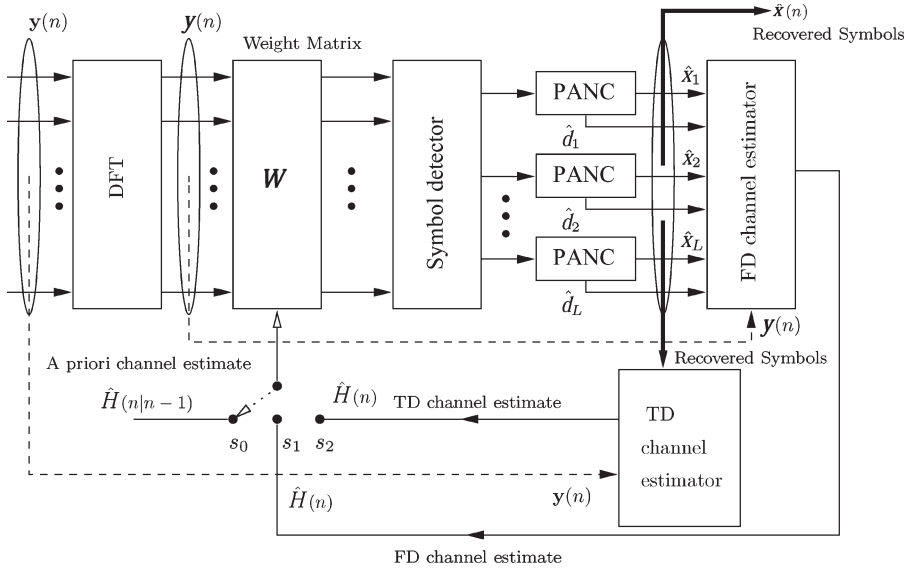


Fig. 5. FD–TD channel estimation with PANC. s_0: Transmitted symbols are recovered using the previous channel estimate. s_1: Transmitted symbols are recovered using the new FD channel estimate obtained after NL distortion cancellation (after PANC). s_2: Transmitted symbols are recovered with a refined TD channel estimate.

A block diagram of the FD–TD channel estimation with PANC is illustrated in Fig. 5.

Finally, some remarks about the application of the aforementioned channel estimation technique in stationary and nonstationary channels are given as follows. In a stationary channel, the estimation of the NL distortion is not effected by MUI (due to known pilots). Furthermore, the level of NL distortion is lower during the initial channel estimation process because only T out of N carriers are used, resulting in a reduced PAPR (see Section V-A). For the case of nonstationary channels, the pilot carriers are tone multiplexed with data carriers. This increases the NL distortion component in the channel estimation (due to larger PAPR). Furthermore, an estimate of the MUI on the data carriers is required in order to reconstruct the TD signal used for estimating the NL distortion. Therefore, as the *a priori* channel estimate is based on the previous time instant, it is expected that the decision errors will increase when the channel variation increases (higher mobile speeds). Our simulation results presented in Section VI confirm that the combination of FD–TD channel estimation and PANC is feasible for stationary and moderately time-varying channels. The FD–TD channel estimation with PANC is summarized in Table II.

VI. Practical Considerations

An attractive feature of the PANC technique is that it is implemented at the base station receiver where more resources are available than in the mobile terminal. Furthermore, the handset hardware does not need to be modified.

The PANC technique assumes the PA model of the user to be known at the base station. The transmission of PA parameters must be included in the system initialization jointly with the channel estimation and synchronization. As discussed in Appendix A, only one parameter (clipping level) must be known for the limiter model. The clipping level and the smoothness factor are required for the solid-state power amplifier (SSPA) model. These parameters can be estimated in the initialization process before the channel estimation. A more realistic alternative is to use a polynomial approach to model the nonlinearity, i.e.,

$$\hat{g}[x(t)] = \sum_{j=0}^{P} a_j |x(t)|^j \tag{31}$$

where $x(t)$ is the input signal, $\hat{g}[x(t)]$ is the output signal of the PA model, and $\{a_j\}_{j=0}^{P}$ represents the polynomial coefficients, with P being the polynomial order. By exciting the PA with a power-swept single-tone signal, an LS estimate of the coefficients a_j can be obtained (see [30] for details). In order to have a complete description of the PA at the base station, all coefficients of the polynomial model are required. We stress that the PANC technique can be easily implemented with more general PA models, including memory effects. However, this is beyond the scope of this paper.

VII. Simulation Results

In Sections VII-A–C, a complete performance study is presented, where the channel estimation technique with PANC is evaluated in terms of MSE, BER, and efficiency for stationary and nonstationary channel environments. The applicability of PANC in WLAN systems is also tested. Finally, we compare the BER results of PANC with a solution that uses a PD in the transmitter. In addition, we evaluate the performance improvement achieved with a solution that combines PANC at the receiver and PD at the transmitter.

In our simulations, we assume $P = 4$ receive antennas and $L = 1$–4 users with a single transmit antenna that is equipped with an SSPA with smoothness factor $p = 2$. The number of subcarriers employed is $N = 512$, and the length of the cyclic

TABLE II
FD–TD CHANNEL ESTIMATION WITH THE PANC ALGORITHM

$\boldsymbol{H}$: Initial FD channel estimate

$g[\cdot]$: PA transfer function (see Appendix A)

For $m = 1$ **to number of iterations**

For $k = 1$ **to** N

Calculate the weight matrix:

$$\boldsymbol{W}_{LS} = \frac{1}{K_L}\hat{\boldsymbol{H}}(n,k)\left[\hat{\boldsymbol{H}}^{\mathrm{H}}(n,k)\hat{\boldsymbol{H}}(n,k)\right]^{-1}$$

Estimate the users signal using LS detector:

$$\hat{\boldsymbol{x}}(n,k) = \boldsymbol{W}_{LS}^{\mathrm{H}}\boldsymbol{y}(n,k)$$

end

FD channel estimation

For $j = 1$ **to** L **number of users**

For $i = 1$ **to** P **number of receive antennas**

Calculate the distortion term:

$$\hat{\mathbf{d}}_j(n) = g[\mathbf{x}_j^{(m)}(n)] - K_L\mathbf{x}_j^{(m)}(n)$$

Obtain the distortion in frequency domain:

$$\hat{\boldsymbol{d}}_j(n) = \mathbf{Q}^{\mathrm{H}}\hat{\mathbf{d}}_j(n)$$

Calculate a new channel estimation on the pilot carriers:

$$\hat{h}_{i,j}(n,k) = \frac{y_i(n,k)}{K_L X_j(n,k) + \hat{d}_j(n,k)}$$

Obtain the whole channel FD response:

$$\hat{\boldsymbol{h}}_{i,j}(n) = \mathbf{Q}_N\left[\mathbf{Q}_T^{\mathrm{H}}\mathbf{Q}_T\right]^{-1}\mathbf{Q}_T^{\mathrm{H}}\hat{\boldsymbol{h}}^{c}_{i,j}(n)$$

end

end

TD channel estimation

Calculate a new weight matrix and detect each user symbol

Calculate the distortion term with the detected symbols

For j=1 to L **number of users**

For i=1 to P **number of receive antennas**

Remove the MUI and the distortion from the received signal:

$$\bar{\mathbf{y}}_i(n) = \mathbf{y}_i(n) - \mathbf{Q}_N\sum_{j=1;j\neq i}^{L}\mathtt{diag}\left[\hat{\boldsymbol{h}}_{i,j}(n)\right]\left[\boldsymbol{z}_j(n) + \hat{\boldsymbol{d}}_j(n)\right]$$

Obtain the time domain channel response:

$$\hat{\mathbf{h}}_{i,j}(n) = \left[\mathbf{C}^{\mathrm{H}}(n)\mathbf{C}(n)\right]^{-1}\mathbf{C}^{\mathrm{H}}(n)\bar{\mathbf{y}}_i(n)$$

end

end

end

prefix is $v = 8$. The modulation employed is either QPSK or 16-quadratic-amplitude modulation (16-QAM). The channel is Rayleigh fading, with independent propagation paths, each generated according to Jakes' Doppler spectrum, with an exponential delay profile. The subcarrier frequency is 5 GHz, and a bandwidth of 20 MHz is used. For the stationary channel, the mobile speed is set to $v(t) = 5$ km/h, and $v(t) = 30$ km/h for nonstationary channels. The channel estimation for stationary channels uses $T = 32$ subcarriers per user in the initialization process, and $T = 16$ subcarriers per user in the case of nonstationary channels. In all simulations, an LS detector is used for separating the user signals.

To characterize NL distortion and efficiency, the evaluations that are performed take into account the following parameters:

1) *Input back-off (IBO)* is defined as the ratio of the average power at the PA input and the input saturation power. It can be represented in decibels as $\mathrm{IBO_{dB}} = 10\log A_s^2 - 10\log\epsilon^2\sigma_s^2$, where σ_s^2, is the average power at the PA input, and A_s^2 is the input saturation power. Parameter

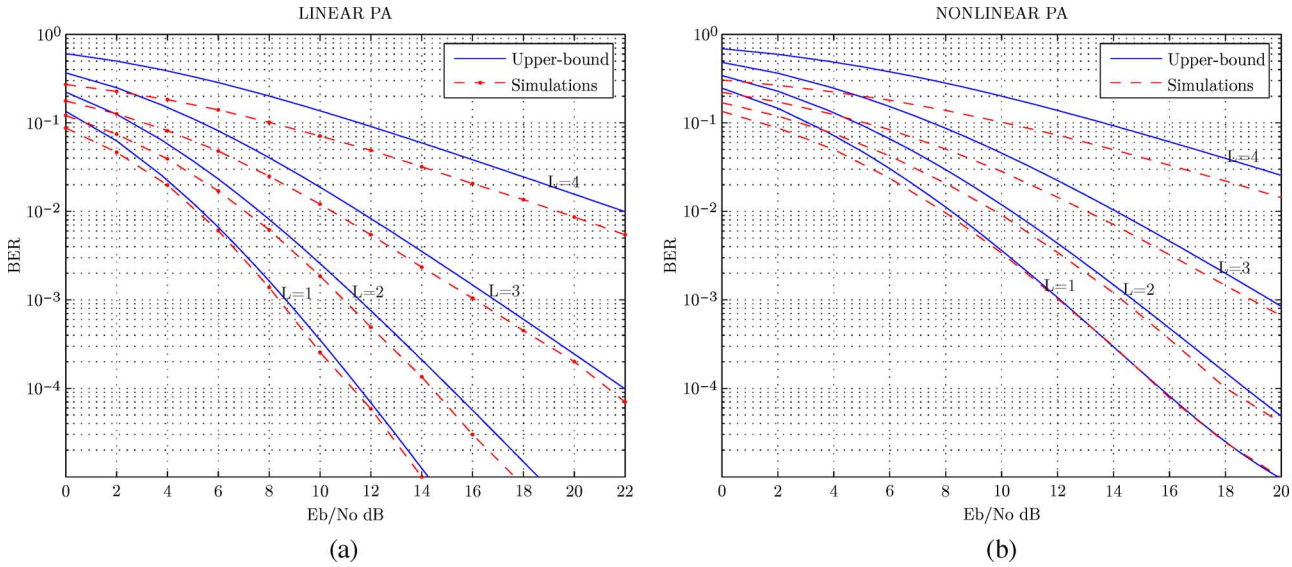


Fig. 6. BER versus E_b/N_0 of a $P \times L$ SDMA–OFDM system with QPSK modulation for the case of (a) linear PA and (b) NL PA with clipping level $\nu = 1$ dB and an LS detector. The mobile speed is set to $v(t) = 5$ km/h. The curves are plotted for $P = 4$ receive antennas, and $L = 1$, 2, 3, and 4. The upper bound of (17) is valid for all SNRs.

$\epsilon < 1$ is used for scaling the symbols in order to reduce the clipping probability and limit the in-band distortion. The IBO employed determines the power that is driven by the PA. Large values of IBO reduce the clipping probability and decrease the BER degradation but also reduce the power efficiency (the dynamic range is decreased). In the simulations (Fig. 7), we use a class A PA that is typical for OFDM systems, whose power efficiency curve is given in [31] and [32].

2) *Total degradation* Ψ is a measure for the balance between the level of degradation and power efficiency, which is defined as $\Psi_{\rm dB} = {\rm SNR}^{NL}_{\rm dB} - {\rm SNR}^{L}_{\rm dB} + {\rm IBO}_{\rm dB}$, where ${\rm SNR}^{NL}_{\rm dB} = E_b/N_0$ is the SNR required to obtain a fixed BER target in the presence of PA nonlinearities with a fixed IBO, and ${\rm SNR}^{L}_{\rm dB}$ expresses the SNR required in the case of linear PA. The BER target is typically set to 10^{-4}.

A. Verification of the BER Analysis

Fig. 6 illustrates the BER results of the LS detector obtained by theory and simulations. QPSK modulation is employed, and the clipping level of the SSPA model is set to $\nu = 1$ dB. The BER curve for a linear PA is included for comparison purposes in Fig. 6(a).

As expected, the simulation results show a lower BER than the theoretical upper bound. For high SNR values, the BER bound is tighter than in the low-SNR region.

Note that if the upper bound derived in [20] were used (assuming linear PA), the predicted BER will be too optimistic and cannot be reached in practice. In that case, the system design would be based on an erroneous limit. We believe that the upper bounds derived in Section III are more realistic for a practical system design.

B. Evaluation of PANC for Stationary Channels

The system considered in this section uses 16-QAM modulation on each carrier. We present the results obtained with PANC, assuming perfect channel state information (CSI) and using the FD–TD channel estimation approach proposed in Section V.

The BER versus SNR curves are shown in Fig. 7. As can be seen, a significant improvement is obtained with PANC. The robustness of the PANC technique to channel estimation errors is evident, showing BER results that are close to those obtained with perfect channel knowledge.

Total degradation curves are shown in Fig. 8 for $L = 1$ and $L = 2$ active users. From these curves, it is possible to conclude that the best operation points for a system with PANC and one without PANC are given as follows.

- Without PANC, IBO $= 3.75$ dB ($\eta = 19\%$) with $\Psi = 4.5$ dB for $L = 1$ active user, and IBO $= 4.4$ dB ($\eta = 16.2\%$) with $\Psi = 4.4$ dB for $L = 2$ active users.
- With PANC, IBO $= 1.2$ dB ($\eta = 29.5\%$) with $\Psi = 2.55$ dB for $L = 1$ active user, and IBO $= 1.9$ dB ($\eta = 26.4\%$) with $\Psi = 2.58$ dB for $L = 2$ users.

Note that for the PANC case, if IBO is fixed at 0 dB, Ψ raises to 3.5 dB for $L = 1$ and to 3.2 dB for $L = 2$. This small increment in Ψ gives an important improvement in the power efficiency, which raises to 35%. We can then conclude that the optimal operation point for the PA can be moved to IBO $= 0$ dB, instead of the previous value (i.e., IBO $= 1.2$ dB for $L = 1$ and IBO $= 1.9$ dB for $L = 2$). This is explained by the specific shape of the total degradation curve which is almost flat when PANC is used in the low IBO region.

In order to verify the out-of-band distortion for the aforementioned operation points, we consider a WLAN system with 512 carriers and an oversampling factor that is equal to 8. The PSD of the SSPA model is compared with the WLAN

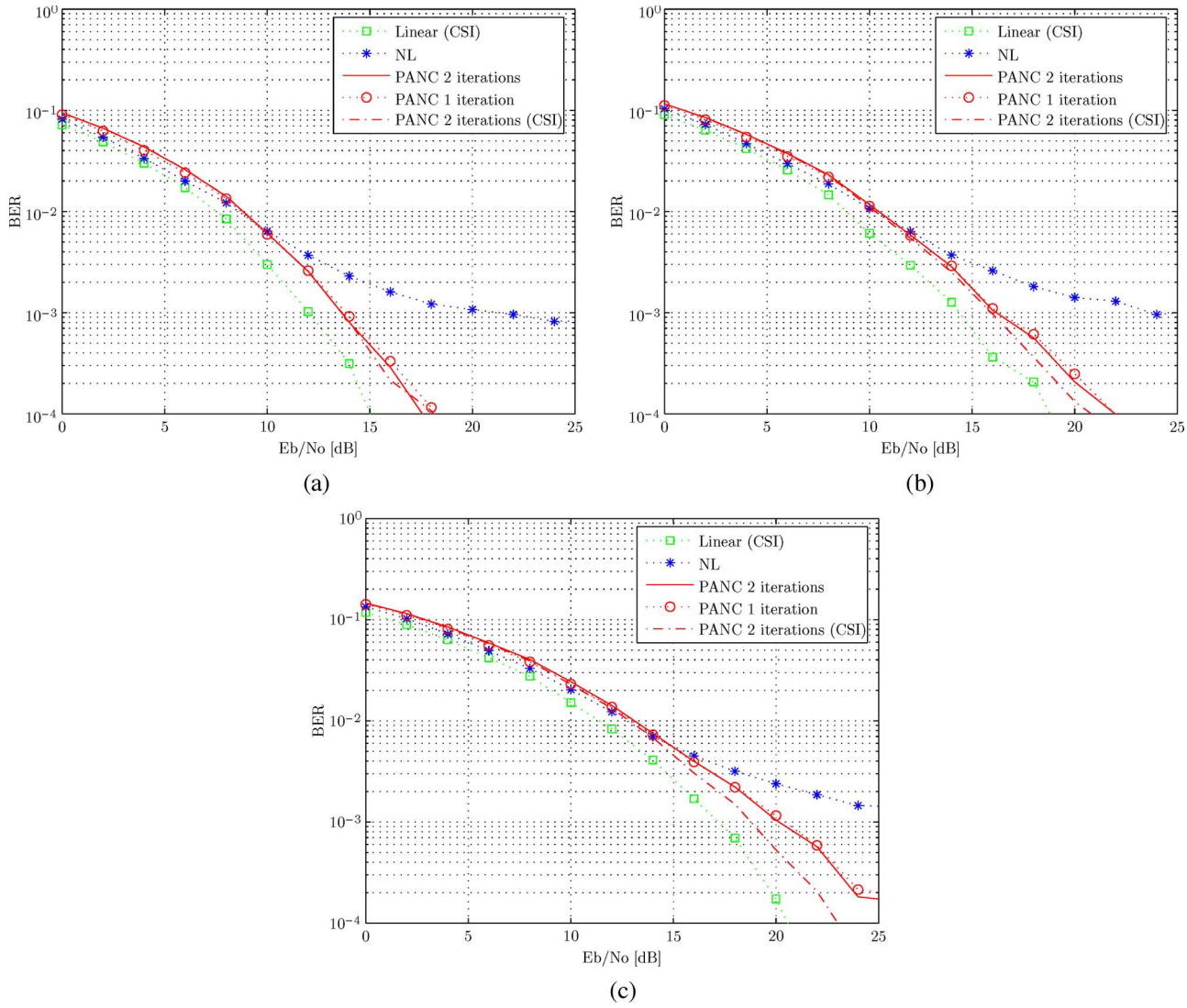


Fig. 7. BER versus E_b/N_0 of a $P \times L$ SDMA–OFDM system ($P = 4$) with 16-QAM for PANC with FD–TD channel estimation ($T = 32$ subcarriers per user), $v(t) = 5$ km/h, and using perfect CSI and LS detector, SSPA model with clipping $\nu = 4$ dB. The curves are plotted for (a) $L = 1$ user, (b) $L = 2$ users, and (c) $L = 3$ users. Results obtained with linear PA and NL PA without PANC are included for reference.

transmit spectrum mask of a WLAN IEEE 802.11a system [33] in Fig. 9. The PSD of a linear PA is included as a reference. The employed clipping level is $\nu = 4$ dB, and the back-off values are 0, 1.2, and 1.9 dB. As can be seen in the figure, the resulting out-of-band distortion from an SSPA that works in the optimal IBO points meets the transmit spectrum mask. Consequently, the PANC technique will give a reduced in-band distortion that works in the optimal work points with an adequate distortion over adjacent bands, as defined in the standard.

C. PANC in Time-Varying Channels

The system considered in this section uses 16-QAM on each carrier. We present the results obtained with PANC, assuming perfect CSI and using the FD–TD channel estimation approach proposed in Section V. The SSPA model uses a smoothing factor $p = 2$ and $\nu = 4$ dB. The FD–TD channel estimation employs one iteration (i.e., one FD–TD cycle). An increase in the number of iterations gives a small improvement in the estimation procedure, unnecessarily increasing the implementation complexity and delay. To compare the performance of the novel technique with other methods that combat NL PA effects, the LS-based PD in [30] was implemented. The PD was modeled with a fifth-order polynomial.

Fig. 10 shows the MSE of the channel estimates versus SNR for different numbers of users. The alternatives evaluated include the folowing.

1) *FD–TD channel estimation for linear and NL PA models:* These MSE curves give upper and lower bounds for the system performance.
2) *FD–TD channel estimation using a seventh-order polynomial model of the PA model in the receiver:* In this evaluation, we show the robustness of the channel estimation methodology to PA modeling errors.
3) *FD–TD channel estimation with PANC:* In this case, we can evaluate the effect of MUI on the channel estimation

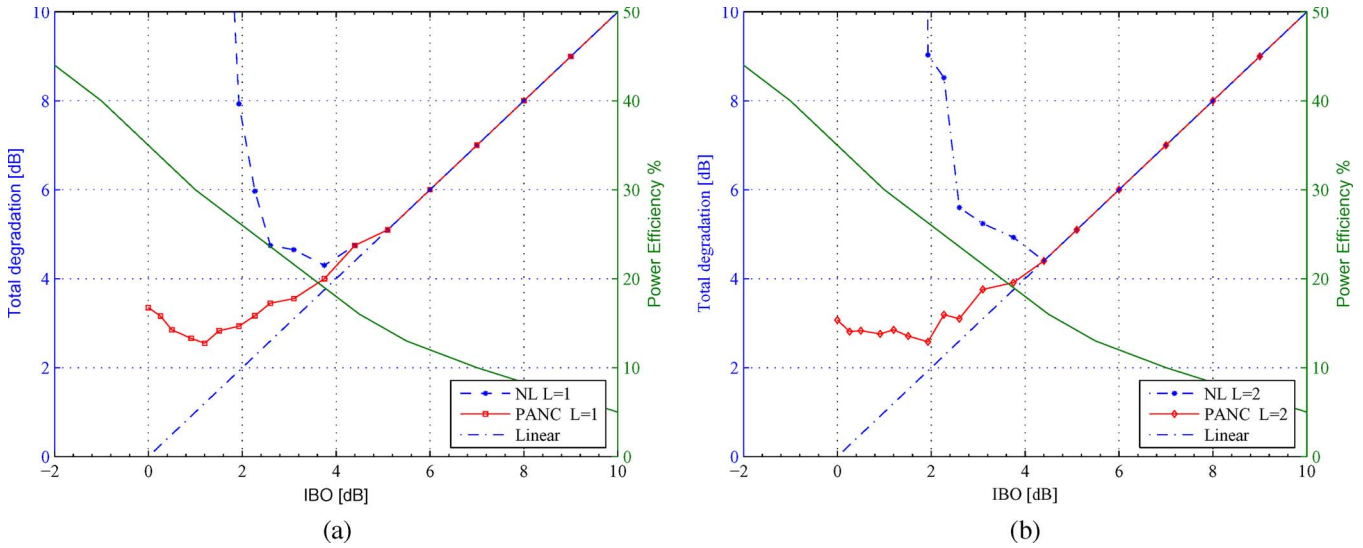


Fig. 8. Total degradation versus IBO of a $P \times L$ SDMA–OFDM system ($P = 4$) with 16-QAM for a receiver with and without PANC, $v(t) = 5$ km/h, LS detector with FD–TD channel estimation ($T = 32$ subcarriers per user), $v(t) = 5$ km/h, SSPA model with clipping $\nu = 4$ dB, and smoothness factor $p = 2$. The curves are plotted for $P = 4$ receive antennas, and $L = 1$ and $L = 2$ users. The power efficiency curve is obtained from [31].

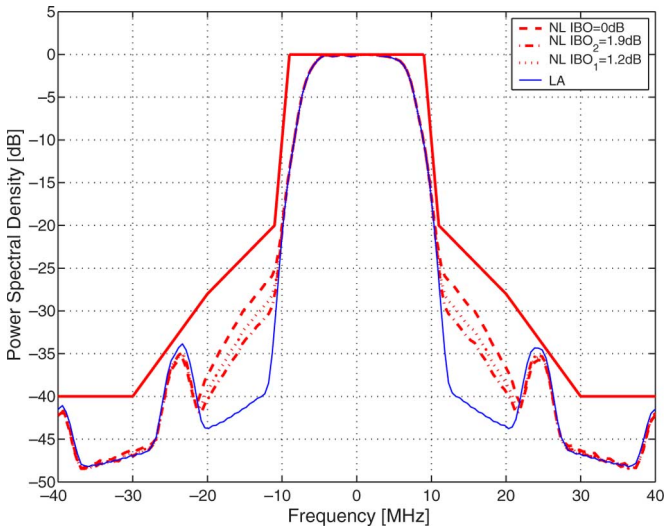


Fig. 9. Out-of-band distortion of an OFDM system with 16-QAM, for different values of IBO and SSPA model with $p = 2$ and $\nu = 4$ dB. The results are plotted for optimal back-off values of $\text{IBO}_1 = 1.2$ dB for $L = 1$ user and $\text{IBO}_2 = 1.9$ dB for $L = 2$ users. The curves are seen to meet the requirements that are set by the WLAN transmit spectrum mask.

accuracy. This effect is clearly evident in these figures, given the lower MSE when the single-user case $L = 1$ is considered. However, for a higher number of users, the BER performance is still suitable. These curves show results that approach the values obtained when a linear PA is used.

4) *FD–TD channel estimation with PD:* This evaluation demonstrates the performance of the PD without PANC. As can be observed, the results obtained with PANC (instead of PD) introduce improvements on the order of 20 (for one user) to 10 dB (for four users).
5) *FD–TD channel estimation with PD and PANC:* This evaluation demonstrates that PD and PANC techniques are complementary techniques and can be combined.

The curves in Fig. 10 show the advantage of using FD–TD channel estimation with the PANC technique, even when combined with a PD.

The BER curves are found in Fig. 11. It can be observed that an SNR gain that is larger than 2 dB is obtained for FD–TD channel estimation with PANC when compared with a PD with FD–TD channel estimation, for $L = 1$. An even larger SNR gain is obtained for a higher number of users.

The performance degradation of the PD is related with the channel estimation errors, and the results are even worse when the number of users is increased. On the other hand, in an NL system with PANC, the BER performance is slightly degraded when the number of users is increased.

Another issue that is shown in the BER results is that PANC is robust to PA modeling errors. The BER curves for PANC with PA known and PANC with estimated PA parameters are almost identical.

The results obtained by FD–TD channel estimation with PANC, and with the addition of PD, do not show a significant improvement in the BER. However, this combination can be useful because the PD provides a reduction in the out-of-band distortion and PANC is useful for the in-band region.

VIII. Conclusion

A novel PANC technique in an SDMA–OFDM system is presented, drastically reducing the harmful effects of PA nonlinearities in the performance of the system. Significant levels of power efficiency improvements with reduced SNR degradation are obtained in SDMA–OFDM systems.

An upper bound for the BER for an SDMA–OFDM system subject to NL PA distortion was also derived. Simulation results agree well with the analytical derivation. The theoretical upper bound obtained can be used for a realistic system design, whereas the upper bound derived in [20], assuming a linear amplifier, may yield very optimistic BER values.

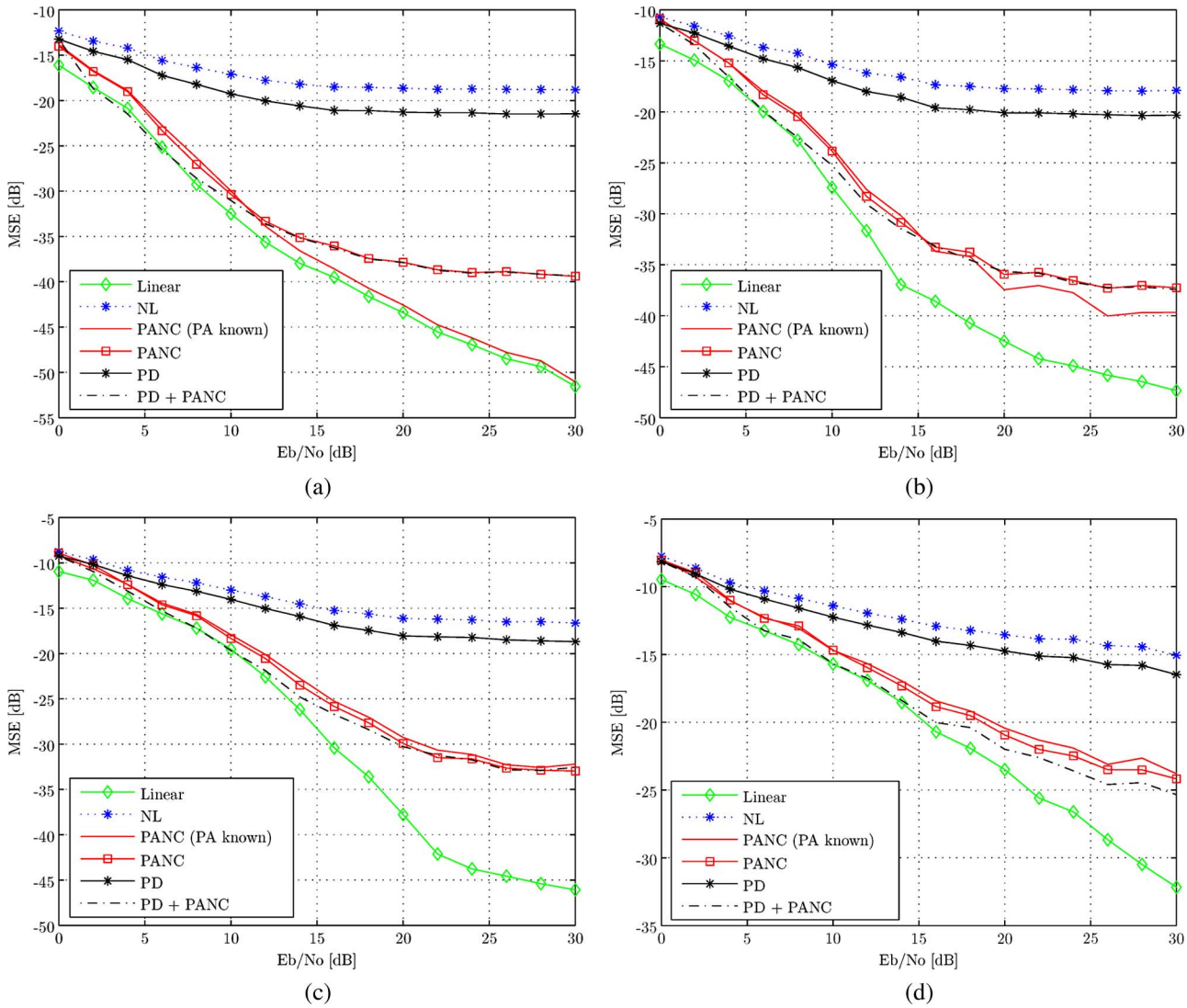


Fig. 10. MSE versus E_b/N_0 of a $P \times L$ SDMA–OFDM system ($P = 4$) with 16-QAM for PANC and PD with FD–TD channel estimation ($T = 16$ subcarriers per user) and $v(t) = 30$ km/h, and LS detector, SSPA model with $p = 2$ and clipping $\nu = 4$ dB. The curves are plotted for (a) $L = 1$ user, (b) $L = 2$ users, (c) $L = 3$, and (d) $L = 4$ users. Results obtained with perfectly known PA model and estimated PA parameters are included. The linear PA and NL PA without PANC are included for reference.

Finally, a channel estimation strategy that incorporates the channel estimation into the PANC technique was proposed. The channel estimation algorithm operates in both the FD and TD. It was verified that, when incorporating the new channel estimation strategy into the PANC technique, the performance remains close to that of a perfectly known channel, thus preserving the significant improvement in BER levels that can be obtained when compared with a solution that ignores the NL PA effects. The robustness of the technique to PA modeling errors was also verified.

To compare the performance of the novel technique with other methods that combat the PA nonlinearity effects, a PD was also implemented. Simulation results show that PANC reaches better results than using a PD. The combination of PANC and PD was also evaluated, and the results show that this combination can be useful, giving suitable BER levels and a reduced out-of-band distortion.

An interesting feature of PANC is that it implements the computationally intensive processing at the receiver. This is an attractive scheme for uplink processing where the computational burden is concentrated at the base station.

APPENDIX A
LIMITER AND SSPA MODELS

The PA input signal can be represented in polar coordinates as $x = \beta e^{j\phi}$, and the output of the PA can be written as

$$g[x] = M[\beta] \exp\left(j\left(\phi + P[\beta]\right)\right) \tag{32}$$

where $M[\beta]$ represents the AM/AM conversion, and $P[\beta]$ represents the AM/PM conversion characteristics of the PA [34]. The output of the PA can be written as

$$g[x] = K_L x + d \tag{33}$$

where K_L is a complex scalar that defines the linear gain, and d is an uncorrelated distortion term. Thus, we have

$$E[xd^*] = E\left[x\left(g[x] - K_L x\right)^*\right] = 0 \tag{34}$$

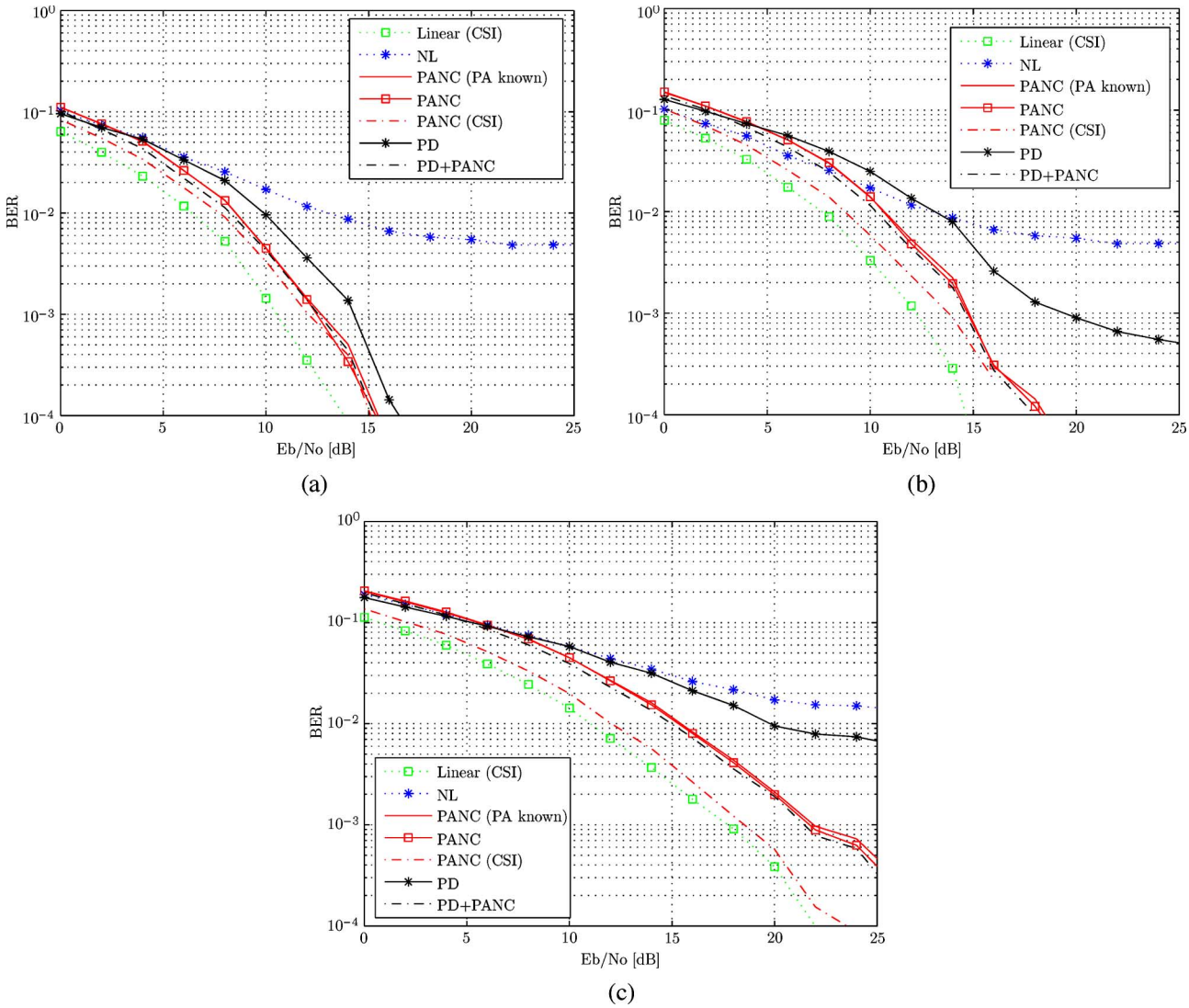


Fig. 11. BER versus E_b/N_0 of a $P \times L$ SDMA–OFDM system ($P = 4$) with 16-QAM for PANC with FD–TD channel estimation ($T = 16$ subcarriers per user), $v(t) = 30$ km/h, and using perfect CSI and LS detector, SSPA model with clipping $\nu = 4$ dB. The curves are plotted for (a) $L = 1$ user, (b) $L = 2$ users, and (c) $L = 3$ users. Results using perfectly known PA model and estimated PA model using a seventh-order polynomial, and linear PA and NL PA without PANC are included for reference.

from which we obtain the gain as

$$K_L = \frac{E\left[xg^*[x]\right]}{E[xx^*]}. \tag{35}$$

The NL distortion term can be calculated as

$$\sigma_d^2 = \frac{E\left[|g[x]|^2\right] - |K_L|^2}{E\left[|x|^2\right]} = \frac{\int_0^{+\infty} |g[u]|^2 p(u)du - |K_L|^2}{\sigma^2} \tag{36}$$

where $g[u] = M_S(\beta)$ is the PA transfer function, and $p(u)$ is the pdf of the OFDM signal, which, for large N, takes the form $p(u) = (1/\sqrt{2\pi}\sigma)\exp(-(u^2/2\sigma^2))$.

In case of the SSPA, the transfer characteristic is modeled by

$$M_S(\beta) = \frac{\beta}{\left[1 + (\beta/A_s)^{2p}\right]^{1/2p}} \tag{37}$$

where the parameter p adjusts the smoothness of the transition from the linear region to the saturation region. In the case of the SSPA model, it is not possible to obtain a closed-form expression for the K_L and σ_d^2 values. For values of p larger than 20, the SSPA model approaches the limiter model given by

$$M_L(\beta) = \begin{cases} \beta, & \beta < A_s \\ A_s, & \beta > A_s \end{cases} \tag{38}$$

where A_s is the amplifier input saturation voltage.

Using the limiter model, it is possible to obtain closed-form solutions for the K_L and σ_d^2 values. From (35), and using the PA model given by (38), we arrive at the following expression for K_L [6]:

$$K_L = 1 - \exp(-\nu^2) + \frac{1}{2}\sqrt{\pi}\nu\,\mathrm{erfc}(\nu) \tag{39}$$

where ν is the clipping level. The power of the distortion noise σ_d^2 is obtained using (36), and it can be written as

$$\begin{aligned}\sigma_d^2 &= \frac{\int_0^{+\infty} |g[u]|^2 p(u)du - K_L^2}{\sigma^2} \\ &= \frac{\int_0^{A_s} |u|^2 p(u)du + \int_{A_s}^{+\infty} |A_s|^2 p(u)du - K_L^2}{\sigma^2} \\ &= \sigma^2 \left[1 - \exp(-\nu^2) - K_L^2\right]. \end{aligned} \tag{40}$$

Appendix B
BER Upper Bound Derivation

Equation (16) can be rewritten as

$$\begin{aligned}\bar{P}_e &\le \frac{1}{(P-L)!\bar{\gamma}^{P-L+1}} \\ &\quad \times \int_0^{\infty} \gamma^{P-L} \exp\left(-\frac{\gamma}{\bar{\gamma}}\right) \exp\left(-\gamma \frac{K_L^2}{\sigma_a^2 + \sigma_d^2 \gamma}\right) d\gamma \\ &= \frac{1}{(P-L)!\bar{\gamma}^{P-L+1}} I \end{aligned} \tag{41}$$

where

$$\begin{aligned} I &= \int_0^{\infty} \gamma^k \exp\left(-\frac{\gamma}{\bar{\gamma}}\right) f(\gamma) d\gamma, \\ k &= P - L, \\ f(\gamma) &= \exp\left(-\gamma \frac{K_L^2}{\sigma_a^2 + \sigma_d^2 \gamma}\right). \end{aligned} \tag{42}$$

To proceed, we use a Taylor expansion for f, i.e.,

$$f(\gamma) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \left(\frac{\gamma K_L^2}{\sigma_a^2 + \sigma_d^2 \gamma}\right)^m. \tag{43}$$

Now, integral I attains a sum form, where we need to evaluate the integral

$$I_m = \int_0^{\infty} \gamma^k e^{-\frac{\gamma}{\bar{\gamma}}} \left(\frac{\gamma K_L^2}{\sigma_a^2 + \sigma_d^2 \gamma}\right)^m d\gamma. \tag{44}$$

Using the substitution $t = \sigma_d^2 \gamma$, we find that

$$I_m = \left(\frac{1}{\sigma_d^2}\right)^{m+k+1} K_L^{2m} \bar{\gamma} \int_0^{\infty} \frac{t^{m+k}}{(\sigma_a^2 + t)^m} e^{-t/\sigma_d^2 \bar{\gamma}}. \tag{45}$$

We can use [21, eq. (8), p. 941] as follows:

$$\begin{aligned}&\int_0^{\infty} x^{\nu-1}(x+\beta)^{-\rho} e^{-\mu x} dx \\ &\qquad = \beta^{\frac{\nu-\rho-1}{2}} \mu^{\frac{\rho-\nu-1}{2}} e^{\frac{\beta\mu}{2}} \Gamma(\nu) W_{\frac{1-\nu-\rho}{2}, \frac{\nu-\rho}{2}}(\beta\mu) \end{aligned} \tag{46}$$

where $\Gamma(\cdot)$ is the Gamma function, and $W(\cdot)$ is the Whittaker function defined as

$$\begin{aligned} W_{\lambda,\mu}(z) &= \frac{z^{\mu+1/2} e^{-z/2}}{\Gamma(\mu - \lambda + 1/2)} \\ &\quad \times \int_0^{\infty} \exp(-zt) t^{\mu-\lambda-1/2} (1+t)^{\lambda+\mu-1/2} dt. \end{aligned} \tag{47}$$

By combining these equations, we can formulate the original integral as an expansion in Whittaker functions

$$\begin{aligned} I_m &= \left(\frac{1}{\sigma_d^2}\right)^{m+k+1} K_L^{2m} \left(\frac{1}{\sigma_d^2 \bar{\gamma}}\right)^k \\ &\quad \times \exp\left(\frac{1}{2\sigma_d^2 \bar{\gamma}}\right) (m+k)! W_{\frac{-k-2m}{2}, \frac{k+1}{2}} \left(\frac{\sigma_a^2}{\sigma_d^2 \bar{\gamma}}\right). \end{aligned} \tag{48}$$

Finally, substituting (48) in (41), the error probability can be expressed as

$$\begin{aligned} \bar{P}_e &\le \frac{1}{(P-L)!\bar{\gamma}^{P-L+1}} \sum_{m=0}^{\infty} I_m \\ &= \frac{1}{(P-L)!\bar{\gamma}^{P-L+1}} \left(\frac{1}{\sigma_d^2 \bar{\gamma}}\right)^{P-L} \exp\left(\frac{1}{2\sigma_d^2 \bar{\gamma}}\right) \\ &\quad \times \sum_{m=0}^{\infty} \frac{(-1)^m (m+P-L)!}{m!} \left(\frac{1}{\sigma_d^2}\right)^{m+P-L+1} \\ &\quad \times K_L^{2m} W_{\frac{-(P-L)-2m}{2}, \frac{P-L+1}{2}} \left(\frac{\sigma_a^2}{\sigma_d^2 \bar{\gamma}}\right). \end{aligned} \tag{49}$$

The Whittaker function can be calculated in *Mathematica Symbolic Software* using the *Confluent Hypergeometric Function*, which is defined in (18). Equation (17) follows from the relation between the Whittaker function and the *Confluent Hypergeometric* functions. Hence

$$W_{\lambda,\mu}(z) = \exp^{-z/2} z^{1/2+\mu} U_{1/2+\mu-\lambda, 1+2\mu}(z). \tag{50}$$

Appendix C
Capacity for LS Receivers

The capacity is obtained by integrating (21) over the chi-square *distribution* as

$$\begin{aligned} C_{\mathrm{LS}} &= \int_0^{\infty} \log_2\left(1 + \frac{\sigma_j^2 \gamma}{\sigma_a^2 + \sigma_d^2 \gamma / K_L^2}\right) \\ &\quad \times \frac{1}{(P-L)!\bar{\gamma}^{P-L+1}} \gamma^{P-L} \exp\left(-\frac{\gamma}{\bar{\gamma}}\right) d\gamma. \end{aligned} \tag{51}$$

Applying integration by parts and expressing the primitive function of the chi-square distribution as ($\Gamma_I(P-L+1,$

$\gamma/\bar{\gamma})/(P-L)!)$ [24, eq. (26.4.19)], where $\Gamma_I(\cdot,\cdot)$ is the incomplete gamma function, the integral can be written as

$$\begin{aligned} C_{\mathrm{LS}} &= \log_2\left(1+\frac{\sigma_j^2\gamma}{\sigma_a^2+\sigma_{d'}^2\gamma,}\right)\frac{\Gamma_I(P-L+1,\gamma/\bar{\gamma})}{(P-L)!}\Bigg|_0^\infty \\ &\quad -\frac{1}{\ln(2)}\int_0^\infty \frac{\Gamma_I(P-L+1,\gamma/\bar{\gamma})}{(P-L)!}g(\gamma)d\gamma \\ &= T_1 - T_2 \end{aligned} \tag{52}$$

where $\sigma_{d'}^2=\sigma_d^2/K_L^2$, $\sigma_j^2=1$, and $g(\gamma)=(\sigma_a^2/[\sigma_a^2+(\sigma_{d'}^2+1)\gamma][\sigma_a^2+\sigma_{d'}^2\gamma])$.

The first term of (52) is equal to

$$T_1=\log_2\left(1+\frac{1}{\sigma_{d'}^2}\right). \tag{53}$$

In order to solve the second term T_2, we assume that $\sigma_{d'}^2 \ll 1$ and $\gamma \gg \sigma_a$. The assumptions are for high SNR levels.

Now, $g(\gamma)$ can be accurately approximated as

$$g(\gamma)=\frac{\sigma_a^2}{(\sigma_a^2+(\sigma_{d'}^2+1)\,\gamma)\,(\sigma_a^2+\sigma_{d'}^2\gamma)} \approx \frac{\sigma_a^2}{\gamma\,(\sigma_a^2+\sigma_{d'}^2\gamma)}. \tag{54}$$

Then, we express the incomplete gamma function $\Gamma_I(\cdot,\cdot)$ using a series expansion [35, Sec. 8.352, eq. (1)] as

$$\Gamma_I(P-L+1,\gamma/\bar{\gamma})=(P-L)!\left[1-\exp\left(\frac{\gamma}{\bar{\gamma}}\right)\sum_{k=0}^{P-L}\frac{(\gamma/\bar{\gamma})^k}{k!}\right]. \tag{55}$$

Now, the term T_2 can be written as

$$\begin{aligned} T_2 &= \frac{1}{\ln(2)}\int_0^\infty\left[1-\exp\left(\frac{\gamma}{\bar{\gamma}}\right)\sum_{k=0}^{P-L}\frac{(\gamma/\bar{\gamma})^k}{k!}\right]\frac{\sigma_a^2}{\gamma\,(\sigma_a^2+\sigma_{d'}^2\gamma)}d\gamma \\ &= \Bigg[1-\exp\left(\frac{\gamma}{\bar{\gamma}}\right)-\exp\left(\frac{\gamma}{\bar{\gamma}}\right)\frac{\gamma}{\bar{\gamma}}u(P-L) \\ &\quad -\exp\left(\frac{\gamma}{\bar{\gamma}}\right)\sum_{k=2}^{P-L}\frac{(\gamma/\bar{\gamma})^k}{k!}\Bigg]\frac{\sigma_a^2}{\gamma\,(\sigma_a^2+\sigma_{d'}^2\gamma)})d\gamma \\ &= I_1+I_2+I_3 \end{aligned} \tag{56}$$

where $u(P-L)$ is the unit step function defined as $u(x)=1$ for $x>0$, and $u(x)=0$ for $x\le 0$. After some manipulation and using [35], integral I_1 is obtained as

$$\begin{aligned} I_1 &= \frac{1}{\ln(2)}\int_0^\infty\left[1-\exp\left(\frac{\gamma}{\bar{\gamma}}\right)\right]\frac{\sigma_a^2}{\gamma\,(\sigma_a^2+\sigma_{d'}^2\gamma)}d\gamma \\ &= \frac{1}{\ln(2)}\left[\ln\left(\frac{\sigma_a^2}{\sigma_{d'}^2\bar{\gamma}}\right)-\exp\left(\frac{\sigma_a^2}{\sigma_{d'}^2\bar{\gamma}}\right)\mathrm{E}i\left(-\frac{\sigma_a^2}{\sigma_{d'}^2\bar{\gamma}}\right)+\gamma_e\right] \end{aligned} \tag{57}$$

where γ_e is the Euler–Mascheroni constant ($\gamma_e=0.5772\ldots$) [24], and $\mathrm{E}i(\cdot)$ is the exponential integral function [35]. This integral was solved using a series expansion for the exponential integral function [35, Sec. 8.214, eq. (4)] in order to evaluate the integration limit in 0.

Integral I_2 is obtained using [35, Sec. 3.352, eq. (4)]

$$\begin{aligned} I_2 &= \frac{1}{\ln(2)}\int_0^\infty \exp\left(\frac{\gamma}{\bar{\gamma}}\right)\frac{1}{\bar{\gamma}}\frac{\sigma_a^2}{(\sigma_a^2+\sigma_{d'}^2\gamma)}u(P-L)d\gamma \\ &= \frac{1}{\ln(2)}\frac{\sigma_a^2}{\sigma_{d'}^2\bar{\gamma}}\exp\left(\frac{\sigma_a^2}{\sigma_{d'}^2\bar{\gamma}}\right)\mathrm{E}i\left(-\frac{\sigma_a^2}{\sigma_{d'}^2\bar{\gamma}}\right)u(P-L). \end{aligned} \tag{58}$$

Integral I_3 is solved using [35, Sec. 3.384, eq. (10)]

$$\begin{aligned} I_3 &= \frac{1}{\ln(2)}\int_0^\infty \exp\left(\frac{\gamma}{\bar{\gamma}}\right)\sum_{k=2}^{P-L}\frac{\gamma^{k-1}}{\bar{\gamma}^k k!}\frac{\sigma_a^2}{(\sigma_a^2+\sigma_{d'}^2\gamma)}d\gamma \\ &= \frac{1}{\ln(2)}\exp\left(\frac{\sigma_a^2}{\sigma_{d'}^2\bar{\gamma}}\right)\sum_{k=2}^{P-L}\frac{\sigma_a^{2k}}{\bar{\gamma}^k k\sigma_{d'}^{2k}}\Gamma_I\left(1-k,\frac{\sigma_a^2}{\sigma_{d'}^2\bar{\gamma}}\right). \end{aligned} \tag{59}$$

Finally, the expression for the capacity can be expressed as

$$\begin{aligned} C_{\mathrm{LS}} \approx{}& \log_2\left(1+\frac{1}{\sigma_{d'}^2}\right)-\frac{\sigma_a^2}{\ln(2)} \\ &\times\Bigg[\ln(\beta)+\gamma_e-\exp(\beta) \\ &\quad\times\Bigg([1+\beta u(P-L)]\,\mathrm{E}i(-\beta) \\ &\qquad +\sum_{k=2}^{P-L}\frac{\beta^k\Gamma_I(1-k,\beta)}{k}\Bigg)\Bigg] \end{aligned} \tag{60}$$

where $\beta=(\sigma_a^2/\sigma_{d'}^2\bar{\gamma})$.

ACKNOWLEDGMENT

The authors would like to thank Dr. R. Wichman and Dr. J. Hämäläinen for their helpful suggestions.

REFERENCES

[1] J. Heiskala and J. Terry, *OFDM Wireless LANs: A Theoretical and Practical Guide*. Indianapolis, IN: Sams, 2002.

[2] D. Wulich, "Definition of efficient PAPR in OFDM," *IEEE Commun. Lett.*, vol. 9, no. 9, pp. 832–834, Sep. 2005.

[3] S. Cripps, *Advanced Techniques in RF Power Amplifiers Design*. Norwood, MA: Artech House, 2002.

[4] C. Tellambura, "A coding technique for reducing peak to average power ratio in OFDM," in *Proc. IEEE GLOBECOM Conf.*, Nov. 1998, vol. 5, pp. 2783–2785.

[5] R. Baxley and G. Zhou, "Power savings analysis of peak-to-average power ratio in OFDM," *IEEE Trans. Consum. Electron.*, vol. 50, no. 3, pp. 792–798, Aug. 2004.

[6] H. Ochiai, "Performance analysis of peak power and band-limited OFDM system with linear scaling," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 1055–1065, Sep. 2003.

[7] F. Gregorio, T. Laakso, and R. Wichman, "Comparison of clustered and antenna interleaving OFDM in multiple-antenna systems with amplifier nonlinearities," in *Proc. IEEE SPCOM*, Dec. 2004, pp. 111–115.

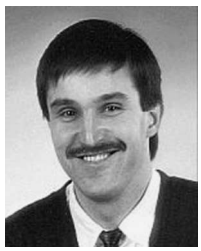
[8] F. Gregorio and T. Laakso, "A novel carrier allocation for multi-user OFDM system with amplifier nonlinearities," in *Proc. EUSIPCO*, Antalya, Turkey, Sep. 2005.
[9] D. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint TX–RX beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.
[10] R. Raich, H. Qian, and G. T. Zhou, "Orthogonal polynomials for power amplifier modeling and predistorter design," *IEEE Trans. Veh. Technol.*, vol. 53, no. 5, pp. 1468–1479, Sep. 2004.
[11] D. Dardari, V. Tralli, and A. Vaccari, "A theoretical characterization of nonlinear distortion effects in OFDM systems," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1755–1764, Oct. 2000.
[12] L. Hanzo, M. Munster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-User Communication, WLANs and Broadcasting*. Hoboken, NJ: Wiley, 2003.
[13] J. Tellado, L. Hoo, and J. Cioffi, "Maximum-likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 218–228, Feb. 2003.
[14] H. Kang, Y. Cho, and D. Youn, "On compensating nonlinear distortions of an OFDM system using an efficient adaptive predistorter," *IEEE Trans. Commun.*, vol. 47, no. 4, pp. 522–526, Apr. 1999.
[15] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. Hoboken, NJ: Wiley, 1980.
[16] T. Roman, M. Enescu, and V. Koivunen, "Joint time-domain tracking of channel and frequency offsets for MIMO OFDM systems," *Wirel. Pers. Commun.*, vol. 31, no. 3/4, pp. 181–200, Dec. 2004.
[17] M. Friese, "On the degradation of OFDM-signals due to peak-clipping in optimally predistorted power amplifiers," in *Proc. IEEE GLOBECOM Conf.*, Nov. 1998, vol. 2, pp. 939–944.
[18] S. Verdu, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[19] A. Bahai, M. Singh, A. Goldsmith, and B. Saltzberg, "A new approach for evaluating clipping distortion in multicarrier systems," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1037–1046, Jun. 2002.
[20] J. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, no. 2–4, pp. 1740–1751, Feb.–Apr. 1994.
[21] I. Gradshteyn and I. Ryzhik, *Table of Integrals and Products*. New York: Academic, 1965.
[22] S. Haykin and M. Moher, *Modern Wireless Communications*. Englewood Cliffs, NJ: Prentice-Hall, 2005.
[23] G. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins Univ. Press, 1993.
[24] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. Norwich, NY: Knovel, 2002.
[25] F. Gregorio, S. Werner, J. Cousseau, and T. Laakso, "Channel estimation for multiuser OFDM systems in the presence of power amplifier nonlinearities," in *Proc. IEEE Int. Symp. PIMRC*, Sep. 2006, vol. 1, pp. 1–5.
[26] S. Werner, M. Enescu, and V. Koivunen, "Combined frequency and time domain channel estimation in mobile MIMO–OFDM systems," in *Proc. IEEE ICASSP*, May 2006, p. IV.
[27] B. Wu, S. Cheng, and H. Wang, "Clipping effects on channel estimation and signal detection in OFDM," in *Proc. IEEE Int. Symp. PIMRC*, Sep. 2003, vol. 1, pp. 531–534.
[28] L. Deneire, P. Vandenameele, L. van der Perre, B. Gyselinckx, and M. Engels, "A low-complexity ML channel estimator for OFDM," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 135–140, Feb. 2003.
[29] S. Werner, M. Enescu, and V. Koivunen, "Low-complexity time-domain channel estimators for mobile wireless OFDM systems," in *Proc. IEEE Workshop SIPS*, Athens, Greece, Nov. 2005, vol. 1, pp. 245–250.
[30] Y. Guo and J. Cavallaro, "Enhanced power efficiency of mobile OFDM radio using predistortion and post-compensation," in *Proc. IEEE VTC—Fall*, Sep. 2002, vol. 1, pp. 214–218.
[31] J. Tubbax, L. V. D. Perre, M. Engels, H. D. Man, and M. Moonen, "OFDM versus single carrier: A realistic multi-antenna comparison," *EURASIP J. Appl. Signal Process.*, vol. 9, pp. 1275–1287, Sep. 2004.
[32] X. Zhang, L. Larson, and P. Asbeck, *Design of Linear RF Outphasing Power Amplifiers*. Norwood, MA: Artech House, 2003.
[33] IEEE, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High Speed Physical Layer in the 5 GHz Band, IEEE Std. 802.11a-1999, 1999.
[34] P. Kenington, *High-Linearity RF Amplifier Design*. Norwood, MA: Artech House, 2000.
[35] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series and Products*. New York: Academic, 1994.

Fernando Gregorio (S'04) received the B.Sc. degree from the Universidad Tecnologica Nacional (UTN), Bahía Blanca, Argentina, and the M.Sc. degree in electrical engineering from the Universidad Nacional del Sur (UNS), Bahía Blanca. He is currently working toward the Ph.D. degree with the Signal Processing Laboratory, Smart and Novel Radios (SMARAD) Centre of Excellence, Helsinki University of Technology, Espoo, Finland.

His research interests include power amplifier nonlinearities in MIMO–OFDM systems and multiuser communications.

Stefan Werner (S'99–A'02–M'03) received the M.Sc. degree in electrical engineering from the Royal Institute of Technology, Stockholm, Sweden, in 1998 and the D.Sc. degree (with honors) in electrical engineering from the Helsinki University of Technology (HUT), Espoo, Finland, in 2002.

He is currently a Senior Researcher with the Signal Processing Laboratory, Smart and Novel Radios (SMARAD) Center of Excellence, HUT. His research interests are in multiuser communications and adaptive filtering.



Timo I. Laakso (SM'95) was born in Vantaa, Finland, on February 1, 1961. He received the M.Sc.(tech.), Lic.Sc.(tech.), and D.Sc.(tech.) degrees in electrical engineering from the Helsinki University of Technology (HUT), Espoo, Finland, in 1987, 1990, and 1991, respectively.

From 1992 to 1994, he was with the Nokia Research Center, Helsinki, Finland, where he researched third-generation mobile communication systems. During 1994–1996, he was a Lecturer at the University of Westminster, London, U.K. During 1996–2006, he was a Professor of signal processing in telecommunications with the Department of Electrical and Communications Engineering, HUT. He is currently with the Signal Processing Laboratory, Smart and Novel Radios (SMARAD) Centre of Excellence, HUT. He is the author of approximately 130 journal and conference publications. He is the holder of 16 patents.

Juan Cousseau (S'90–M'92–SM'00) received the B.Sc. degree in electrical engineering from the Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina, in 1983 and the M.Sc. and Ph.D. degrees in electrical engineering from the Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, in 1989 and 1993, respectively.

Since 1984, he has been with the Department of Electrical and Computer Engineering and the Departamento de Estudios de Postgrado y Educación Contínua, UNS. He is currently a Senior Researcher with the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Bahía Blanca, Argentina. He has been involved in scientific and industry projects with research groups from Argentina, Brazil, Spain, and the USA. He is the Coordinator of the Signal Processing and Communication Laboratory (LaPSyC), UNS. He was a Visiting Professor at the University of California, Irvine, in 1999 and at the Signal Processing Laboratory, Helsinki University of Technology, Espoo, Finland, in 2004 and 2006. His research interests are in the areas of adaptive algorithms and statistical signal processing addressed to high-speed digital communications.

Dr. Cousseau was the Chair of the IEEE Circuits and Systems Society (CASS), Argentine Chapter, from 1997 to 2000 and a member of the Executive Committee of the IEEE CASS during 2000/2001. He was part of the IEEE Signal Processing Society Distinguished Lecturer Program in 2006.