

Least Squares Channel Estimation for Analog Network Coding Over Frequency Selective Fading Channels

Pedro Ivo da Cruz and Murilo Bellezoni Loiola

Abstract—Two-way relay networks can have their throughputs improved by adopting physical-layer network coding. To work properly, these systems need to know the channel impulse responses involved. Most previous works on channel estimation for physical-layer network coding systems consider flat fading or frequency-selective fading together with orthogonal frequency division multiplexing modulation. However, for single carrier systems under frequency-selective channels, the estimators proposed in the literature cannot be applied directly. In order to solve this problem, a least squares channel estimator for these scenarios is proposed here. Simulations are performed to evaluate the performance of this channel estimator and the results show the effectiveness of the proposed technique.

Index Terms—physical layer network coding, channel estimation, least squares, analog network coding, amplify-and-forward, two-way relay network

I. INTRODUCTION

The first work on communication systems that use two-way channels for exchanging information between two nodes was presented by Shannon in 1961 [1]. In his work, Shannon discussed several channels through which it was possible to exchange information between two nodes in both directions at the same time. One of these channels would allow the two nodes to send their information at the same time and frequency, and both nodes would receive the XOR mapping of the bits sent by each user. Shannon did not show how such a system could be implemented but developed the mathematics necessary to understand them. This subject has gained a renewed interest when a technique called Physical-layer Network Coding (PNC) was proposed in 2006 [2], [3]. PNC adopts a Two-way Relay Channel (TWRC) network and allows two nodes to send data simultaneously to the relay, improving the throughput of the network [3].

The PNC systems take advantage from the electromagnetic interference between the signals sent by two or more nodes at the same time and frequency. Specifically in the TWRC, two nodes send information simultaneously to the relay node, that must perform the PNC mapping [2] in order to transform its received signals originated from both users into one that can

be recognized by the end nodes. It is shown in [4] that we can consider the signal interference itself as a PNC mapping, and, therefore, the relay only needs to amplify the received signal and send it to the end nodes. The nodes must, therefore, be able to extract the information of interest from this signal. This protocol is known as Amplify-and-Forward (AF), and the system is said to use Analog Network Coding (ANC), since the relay does not need to perform any detection, and the mapping occurs in the analog domain.

Most works on PNC systems consider that the channel impulse responses (CIR) of each channel are perfectly known at the relay and at the end nodes. However, in practical situations, these CIR are not known a priori. Hence, channel estimation techniques are essential for practical deployment of PNC systems. Because of the easy implementation of the ANC systems, most works on channel estimation focus on these systems. Furthermore, it can be shown that for ANC systems, channel estimation is only necessary for detection at the end nodes, i. e., only the impulse responses of the cascaded channels needs to be estimated, while knowledge of the individual channels is not required [5]–[10].

Therefore, the literature on channel estimation techniques focus on the TWRC network and ANC systems. The work in [6] derives the maximum likelihood (ML) estimator and the Linear Maximum Signal-to-Noise Ratio (LMSNR) estimator, considering flat fading channels. It also proposes the optimal training sequence that reduces the mean squared error (MSE) associated with those estimators. In [7], the authors developed Least Squares (LS) estimators for Orthogonal Frequency Division Multiplexing (OFDM) systems, by using either one OFDM symbol as training sequence, or by using pilot sub-carriers. The Zadoff-Chu training sequence for TWRC systems was proposed in [8] for the LS estimator and OFDM modulated systems, in order to reduce the peak-to-average power ratio (PAPR). The Linear Minimum Mean Squared Error (LMMSE) estimator for OFDM systems was addressed in [9], showing that it has superior performance than the LS estimator, in terms of MSE. The work in [10] proposes a different channel estimation technique: it uses the cyclic time shifting property of the Discrete Fourier Transform to make the channels separable at the relay. However, while previous works estimated the channels at the end nodes, in [10] they are estimated at the relay, and the end nodes are assumed to receive the estimated CIRs through a perfect feedback channel.

The majority of works on channel estimation for PNC systems, included the ones cited previously, considers only

Pedro Ivo da Cruz and Murilo Bellezoni Loiola are with the Engineering, Modeling and Applied Social Sciences Center, Federal University of ABC, Santo André, SP, Brazil, E-mails: pedro.cruz@ufabc.edu.br, murilo.loiola@ufabc.edu.br. This work was partially supported by CAPES and by FAPESP (2013/25977-7). This work was presented in part at the XXXIV Brazilian Symposium on Telecommunication and Signal Processing (SBTr2016).

flat fading environments [5], [6], or frequency-selective fading environments combined with OFDM modulation [7]–[10], since OFDM is a robust technique against the frequency selectivity of wireless channels.

However, in single carrier (SC) communications systems, such as GSM [11], and satellite networks with CDMA [12]–[14], PNC technique could also be applied. In the same way, several communications protocols and standards for IoT and Wireless Sensor Networks make use of SC modulation and their importance has increased in recent years. For instance, the binary-phase shift keying (BPSK) modulation is used in several open standards such as 6LoWPAN and ZigBee, and the differential BPSK in SigFox [15]. This systems are also subject to the effects of the wireless channel. Therefore, this work proposes an LS channel estimation technique for ANC systems using an SC modulation over frequency-selective fading channels. The works in [6], [7], [9] also develop specific training sequences to minimize the MSE of the estimator. However, it is shown that there is no significant reduction on the MSE when compared to a random training sequence.

This work improves the work in [16] presenting a detailed system modeling and derivation of the estimator, its MSE expression, and the design of an optimal training sequence. Furthermore, it also presents results with some modifications in the scenarios presented, and a more detailed discussion and insights on the obtained the results.

The remainder of this paper is organized as follows: Section II presents the system model that is used in this work; Section III develops the proposed least squares channel estimator, while simulation results are presented in Section IV; finally, conclusions are made in Section V.

II. SYSTEM MODEL

A simple TWRN typically has two source nodes (1 and 2), also referred to as user nodes, and one relay node (R), as shown in Figure 1, where node 1 sends information to node 2, and vice-versa, using node R to assist their communications.

The traditional network using Time-Division Multiple Access (TDMA) would need four transmission stages to exchange data between nodes 1 and 2, as shown in Figure 1: the first stage is used to send data from node 1 to node R; the second, to send this data from node R to node 2; the third, to send data from node 2 to node R; and the fourth one to send data from node R to node 1.

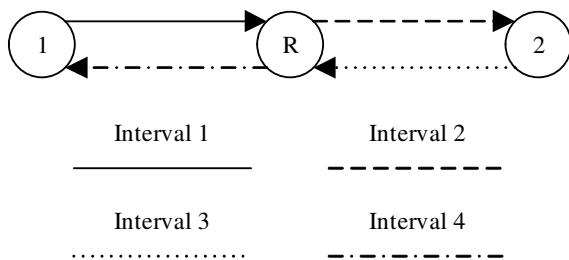


Fig. 1. Transmission stages in a TWRN with TDMA.

By applying the Network Coding (NC) concepts [17], it is possible to reduce to three the number of time intervals needed

to exchange data between nodes 1 and 2, as shown in Figure 2. The relay node detects the signals from both user nodes separately, and performs the NC of the bits of each one of them before transmission.

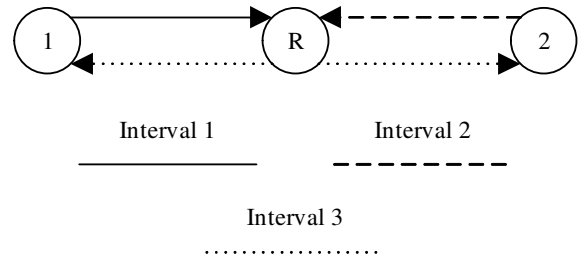


Fig. 2. Transmission stages in a TWRN with NC.

To illustrate how NC could be implemented in a TWRN consider, for instance, that node 1 transmits bit s_1 and the node 2 transmits bit s_2 . The simplest way to do NC is to perform a bitwise XOR operation $s_R = s_1 \oplus s_2$ at the relay. It will then transmit the bit s_R to both users, each of which will perform this same operation on the received bit and the bit itself transmitted. At node 1, for example, the information transmitted by node 2 can be estimated as:

$$\begin{aligned} \hat{s}_2 &= s_1 \oplus s_R \\ &= s_1 \oplus s_1 \oplus s_2 \\ &= s_2. \end{aligned} \tag{1}$$

Therefore, node 1 is able to recover the bit sent by node 2 from the bit s_R sent by the relay.

Finally, when applying the PNC technique, it only takes two transmission stages to exchange data between the two user nodes, as shown in Figure 3. In this case, nodes 1 and 2 send their information simultaneously to node R. This stage is called Multiple Access (MAC). In the next stage, R amplifies the signal by a factor α and forwards the resulting signal to both nodes 1 and 2. This factor can be used to adjust the power of transmitted signal at the relay. This stage is called Broadcast (BC).

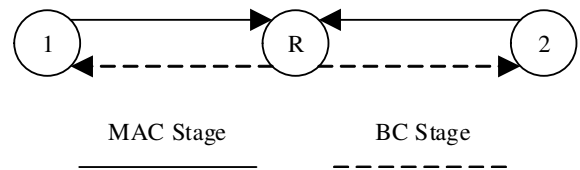


Fig. 3. Transmission stages in a TWRN with PNC.

The relay protocol presented above is called AF, and the user nodes just need to remove their self-information to obtain the information sent by the other user. The AF protocol is also called ANC, as discussed before. This works will focus specifically on this protocol.

The signal model used in this work is shown in Figure 4. Let $x_i(n)$ be the baseband symbol sent by the node i at the discrete time instant n , $h_{iR}(n)$ be the CIR between the nodes i and R, and $w_R(n)$ be the additive white Gaussian noise (AWGN) with distribution $\mathcal{N}(0, \sigma_w^2)$ at node R, and $*$ denotes the convolution

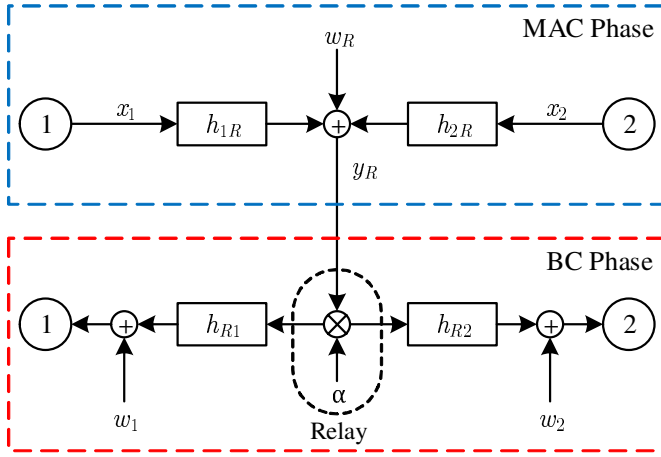


Fig. 4. Scheme for the signal model of a PNC system.

operation. The discrete-time baseband signal received at R can be written as:

$$y_R(n) = x_1(n) * h_{1R}(n) + x_2(n) * h_{2R}(n) + w_R(n). \quad (2)$$

Due to the symmetry of this network, the performance can be analyzed just at node 1, since it is equivalent for the node 2. Let $h_{Ri}(n)$ be the CIR between nodes R and i , and $w_1(n)$ be AWGN at the node 1, also with distribution $\mathcal{N}(0, \sigma_w^2)$. The signal received at the node 1 at the BC stage can be written as:

$$y_1(n) = \alpha y_R(n) * h_{R1}(n) + w_1(n). \quad (3)$$

The cascaded channels can be defined as $a(n) = h_{1R}(n) * h_{R1}(n)$, and $b(n) = h_{2R}(n) * h_{R1}(n)$, and the equivalent noise as $w(n) = \alpha w_R(n) * h_{R1}(n) + w_1(n)$, which will also be a random variable with Gaussian distribution with zero mean and variance that will depend upon the channel power profile of $h_{1R}(n)$ and the variances of $w_R(N)$ and $w_1(n)$. Then, substituting (2) into (3), the signal received at node 1 at the BC stage can be rewritten as:

$$y_1(n) = \alpha x_1(n) * a(n) + \alpha x_2(n) * b(n) + w(n). \quad (4)$$

Detection at node 1 can be performed as [18]:

$$\hat{x}_2(n) = \arg \min_{x_2(n)} |y_1(n) - \alpha x_1(n) * a(n) - \alpha x_2(n) * b(n)|^2. \quad (5)$$

It means that, given $y_1(n)$, $x_1(n)$, $a(n)$ and $b(n)$, it is possible to obtain an estimate of $x_2(n)$ through maximum likelihood data detection, which, in this case, turns out to be the value of x_2 that minimizes the squared of the absolute difference between the received signal $y_1(n)$ and a version of the transmitted signal $\alpha y_R(n)$ without noise.

A practical way of performing this at node 1 is to first extract its self-information from $y_1(n)$, i. e:

$$\begin{aligned} \tilde{x}_2(n) &= y_1(n) - \alpha x_1(n) * a(n) \\ &= \alpha x_2(n) * b(n) + w(n). \end{aligned} \quad (6)$$

Then, the signal \tilde{x}_2 needs to be equalized. Linear equalizers such as Zero-forcing (ZF) or algorithms such as the Maximum Likelihood Sequence Estimation (MLSE) can be used for this purpose [11].

The model described above shows the importance of estimating the cascaded CIR $a(n)$ and $b(n)$ at the end nodes for the self-information extraction and equalization. It also shows one advantage of the ANC scheme: it only needs the knowledge of the CIR at the end nodes, and not at the relay node. Thus, the next section is devoted to the development of a channel estimator for PNC systems over frequency-selective fading channels.

III. LEAST SQUARES CHANNEL ESTIMATION

As mentioned in Section I, most works on channel estimation for PNC systems consider frequency-selective channels with OFDM modulation. In this work, an LS estimator in time domain is proposed to estimate frequency-selective and time invariant CIR for PNC systems using SC transmission.

Assuming that the nodes 1 and 2 send N training symbols $x_1(n)$ and $x_2(n)$, respectively, (4) can be interpreted as a linear model that can be written in matrix form. To do so, a matrix \mathbf{X} can be defined as:

$$\mathbf{X} = \alpha [\mathbf{X}_1 \ \mathbf{X}_2], \quad (7)$$

where \mathbf{X}_1 and \mathbf{X}_2 are convolution matrices, which have the structure shown in (8), containing the training symbols sent by nodes 1 and 2, respectively.

$$\mathbf{X}_i = \begin{bmatrix} x_i(0) & 0 & 0 & \dots & 0 \\ x_i(1) & x_i(0) & 0 & \dots & 0 \\ x_i(2) & x_i(1) & x_i(0) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_i(N-1) & x_i(N-2) & x_i(N-3) & \dots & x_i(N-L) \end{bmatrix}. \quad (8)$$

In this matrix, the element $x_i(n)$ is the symbol sent by user i at the discrete time n . Considering that the channels have length N_{CH} , the cascaded channels $a(n)$ and $b(n)$ have length $L = 2N_{CH} - 1$. Therefore, \mathbf{X} has dimension $N \times 2L$. Although it has been considered that all channels have the same length, the estimator can be easily generalized for channels with different sizes.

Let \mathbf{h} be the column vector that contains the coefficients of both concatenated channels as:

$$\mathbf{h} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \quad (9)$$

where $\mathbf{a} = [a_0 \ a_1 \ \dots \ a_{L-1}]^T$ is a vector that contains the L coefficients of $a(n)$, and $\mathbf{b} = [b_0 \ b_1 \ \dots \ b_{L-1}]^T$, is the vector that contains the L coefficients of $b(n)$. Furthermore, let $\mathbf{w} = [w_0 \ w_1 \ \dots \ w_{N-1}]^T$ be a vector containing N samples of $w(n)$, and $\mathbf{y} = [y_1(0) \ y_1(1) \ \dots \ y_1(N-1)]^T$ be a vector of N samples from the received signal at node 1, where $y_i(n)$ represents the received sample at instant n at node i .

Then, it is possible to write (4) in matrix form as:

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{w}. \quad (10)$$

The LS estimate of \mathbf{h} can be obtained from [18]:

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} |\mathbf{y} - \mathbf{X}\mathbf{h}|^2. \quad (11)$$

Thus, the LS channel estimator may be obtained by minimizing:

$$J(\mathbf{h}) = \|\mathbf{y} - \mathbf{X}\mathbf{h}\|^2 = (\mathbf{y} - \mathbf{X}\mathbf{h})^H(\mathbf{y} - \mathbf{X}\mathbf{h}). \quad (12)$$

By applying the distributive property and Hermitian product of matrices [18], the last equation can be rewritten as:

$$\begin{aligned} J(\mathbf{h}) &= \mathbf{y}^H\mathbf{y} - \mathbf{y}^H\mathbf{X}\mathbf{h} - \mathbf{h}^H\mathbf{X}^H\mathbf{y} + \mathbf{h}^H\mathbf{X}^H\mathbf{X}\mathbf{h} \\ &= \mathbf{y}^H\mathbf{y} - 2\mathbf{y}^H\mathbf{X}\mathbf{h} + \mathbf{h}^H\mathbf{X}^H\mathbf{X}\mathbf{h}. \end{aligned} \quad (13)$$

Differentiating (13) with respect to \mathbf{h} and setting it equal to zero yields:

$$\frac{\partial J(\mathbf{h})}{\partial \mathbf{h}} = -2\mathbf{X}^H\mathbf{y} + 2\mathbf{X}^H\mathbf{X}\mathbf{h} = 0. \quad (14)$$

Thus, the solution to (11) is given by [18]:

$$\hat{\mathbf{h}} = \mathbf{X}^\dagger \mathbf{y}, \quad (15)$$

where $\hat{\mathbf{h}}$ contains the estimates of \mathbf{a} and \mathbf{b} and \mathbf{X}^\dagger denotes the pseudoinverse matrix of \mathbf{X} , which is given by [19]

$$\mathbf{X}^\dagger = (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H. \quad (16)$$

It is worth noting that the estimated cascaded channels in (15) will be used for self-extraction and equalization at the end nodes.

It is also important to highlight that, although the estimator (15) is based on classic equations, some adaptations in the model, which were not previously reported in the literature, were necessary to employ it with ANC system.

The relationship between the estimate $\hat{\mathbf{h}}$ and the real channel \mathbf{h} can be found by substituting (10) into (15), resulting in:

$$\hat{\mathbf{h}} = (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{X}\mathbf{h} + (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{w}, \quad (17)$$

where $(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{X} = \mathbf{I}$, being \mathbf{I} an identity matrix with appropriated dimensions. Then, it is possible to write:

$$\hat{\mathbf{h}} = \mathbf{h} + (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{w}. \quad (18)$$

Defining the error vector as $\mathbf{e} = \hat{\mathbf{h}} - \mathbf{h}$, the covariance matrix of the estimation error is given by $\text{cov}(\mathbf{e}) = \mathbb{E}[(\hat{\mathbf{h}} - \mathbf{h})(\hat{\mathbf{h}} - \mathbf{h})^H]$. Substituting (18) into this equation gives:

$$\text{cov}(\mathbf{e}) = \mathbb{E}[(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{w}\mathbf{w}^H\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}]. \quad (19)$$

The noise component is given by $\mathbf{w} = \alpha\mathbf{H}_{R1}\mathbf{w}_R + \mathbf{w}_1$ for node 1, where \mathbf{H}_{R1} is convolution matrix associated with channel $h_{R1}(n)$, \mathbf{w}_R and \mathbf{w}_1 are vectors containing the samples from the noises received at nodes R and 1, respectively. As the noises \mathbf{w}_R and \mathbf{w}_1 are independent, the error covariance matrix becomes:

$$\begin{aligned} \text{cov}(\mathbf{e}) &= \alpha^2 \mathbb{E}[(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{H}_{R1}\mathbf{w}_R\mathbf{w}_R^H\mathbf{H}_{R1}^H\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}] \\ &\quad + \mathbb{E}[(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{w}_1\mathbf{w}_1^H\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}]. \end{aligned} \quad (20)$$

As \mathbf{X} , \mathbf{H}_{R1} , \mathbf{w}_R and \mathbf{w}_1 are all independent, (20) can be simplified to:

$$\begin{aligned} \text{cov}(\mathbf{e}) &= \alpha^2 \sigma_w^2 (\mathbf{X}^H\mathbf{X})^{-1} \mathbf{X}^H \beta_{R1} \mathbf{X} (\mathbf{X}^H\mathbf{X})^{-1} \\ &\quad + \sigma_w^2 (\mathbf{X}^H\mathbf{X})^{-1} \mathbf{X}^H \mathbf{X} (\mathbf{X}^H\mathbf{X})^{-1}, \end{aligned} \quad (21)$$

where:

$$\beta_{R1} = \sum_{l=0}^{N_{CH}-1} \sigma_{R1,l}^2 \quad (22)$$

with $\sigma_{R1,l}^2$ being the variance of the l -th coefficient of the channel between nodes R and 1.

Rearranging the terms in (21), the covariance matrix of the estimation error is given by:

$$\text{cov}(\mathbf{e}) = \sigma_w^2 (\mathbf{X}^H\mathbf{X})^{-1} \mathbf{X}^H (\alpha^2 \beta_{R1} + 1) \mathbf{X} (\mathbf{X}^H\mathbf{X})^{-1}. \quad (23)$$

The MSE is given by $\text{MSE} = \text{tr}\{\text{cov}(\mathbf{e})\}$, where tr denotes the trace operator, thus resulting in:

$$\text{MSE} = \sigma_w^2 (\alpha^2 \beta_{R1} + 1) \text{tr}\{(\mathbf{X}^H\mathbf{X})^{-1}\}. \quad (24)$$

Therefore, it is clear that the MSE depends on the noise variance, the channel power profile, and $\text{tr}\{(\mathbf{X}^H\mathbf{X})^{-1}\}$. So, it is possible to design a training sequence to minimize this last term, which is the only term the designers can control, and thus, minimizing the MSE.

Assuming a power constraint \mathcal{P} on the transmitted signals, i.e:

$$\|\mathbf{X}\|_F^2 = \mathcal{P}, \quad (25)$$

where $\|\cdot\|_F^2$ denotes the Frobenius norm, the optimal training sequence can be obtained by solving the following:

$$\arg \min_{\mathbf{X}} \text{tr}\{(\mathbf{X}^H\mathbf{X})^{-1}\} \text{ subject to } \|\mathbf{X}\|_F^2 = \mathcal{P}. \quad (26)$$

It is possible to show that the minimum is achieved when [20]:

$$\mathbf{X}^H\mathbf{X} = \frac{\mathcal{P}}{2L} \mathbf{I}. \quad (27)$$

In other words, $\mathbf{X}^H\mathbf{X}$ must be a scaled identity matrix. This is achieved for any matrix \mathbf{X} with orthogonal lines, and with norm $\sqrt{\mathcal{P}/2L}$. There might be additional constraints on the form of the training matrix \mathbf{X}_i or on \mathcal{P} . For instance, in OFDM modulated systems [7], the training matrix must be circulant and it is also desirable to choose a training matrix that reduces the peak-to-average power ratio (PAPR) [8]. Since this work considers an SC modulated system operating in frequency-selective fading channels, the training sequence is restricted to the form shown in (8).

To satisfy these requirements, the following training sequence is proposed:

$$x_i(n) = \sqrt{\frac{\mathcal{P}}{2L}} \delta(n - (i-1)L), \quad i = 1, 2, \quad (28)$$

recalling that i denotes the number of the user node and $\delta(n)$ denotes the Kronecker delta function.

In order to properly compare the performance of this training sequence against that of a sequence containing randomly generated BPSK symbols, it is necessary to adjust \mathcal{P} so that the power of both sequences are equal. This can be done by setting:

$$\mathcal{P} = 2LN, \quad (29)$$

and, by doing so, the factor that defines amplitude of the symbols in the optimal training sequence becomes:

$$\sqrt{\frac{\mathcal{P}}{2L}} = \sqrt{N}. \quad (30)$$

Although these sequences minimize the estimator MSE, they do not improve significantly the MSE compared to that of a

random training sequence. Consequently, the BER will not have a perceptible reduction, as will be shown in section IV. This result agrees with others already presented in the literature for optimal training sequence design for OFDM modulated PNC systems.

IV. SIMULATION RESULTS AND DISCUSSION

To evaluate the performance of the proposed LS channel estimator, an ANC system is simulated using BPSK modulation.

The gain α is set in a way that the power of the received signal at the relay is equal to the power received at nodes 1 and 2, so that the signal-to-noise ratio (SNR) is the same at all channels. To do that, P_R is set to $P_R = P_1 + P_2$, being P_R the power of the signal transmitted by the relay, P_1 and P_2 the powers of the signals transmitted by nodes 1 and 2, respectively. Thus, the gain α can be set to:

$$\alpha = \sqrt{\frac{P_R}{\beta_{1R}P_1 + \beta_{2R}P_2 + \sigma_w^2}}, \quad (31)$$

where

$$\beta_{iR} = \sum_{l=0}^{N_{CH}-1} \sigma_{iR,l}^2, \quad (32)$$

with $\sigma_{iR,l}^2$ being the variance of the l -th coefficient of the channel between the nodes i and R . Then, the SNR can be defined as $\text{SNR} = P_R/\sigma_w^2$, where σ_w^2 is the noise power.

In the simulations, whose results are presented in the sequel, the channels have lengths $N_{CH} = 5$ and their coefficients are randomly generated in each realization. The coefficients are independent and identically distributed zero mean and unit variance complex Gaussian random variables. In other words, the channel magnitudes follow a Rayleigh distribution and the channel phases are uniformly distributed between 0 and 2π . The cascaded channels \mathbf{a} and \mathbf{b} have length $L = 9$.

The MSE between the original and estimated channels $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, and the bit error rate (BER) when the system uses the estimated CIR and when it has perfect knowledge of the CIR are computed through Monte Carlo simulations for different scenarios. Each realization simulates the transmission of 10^3 information bits. For each SNR, a total of 10^4 realizations are performed for averaging. The training sequence, that consists of N random BPSK symbols, is concatenated in the beginning of the block of information bits. The estimate $\hat{\mathbf{a}}$ is used to compute the self-extraction operation shown in (6). Then, the estimate $\hat{\mathbf{b}}$ is used to perform the equalization through an MLSE equalizer. Both estimates were obtained from (15). The estimation MSE is computed in each realization by:

$$\text{MSE}(\hat{\mathbf{a}}) = \frac{1}{L} \sum_{n=0}^{L-1} |a_n - \hat{a}_n|^2, \quad (33)$$

and, at the end of the simulation, it is averaged by the number of trials. A similar expression can be used to compute the MSE for channel $\hat{\mathbf{b}}$.

In the first scenario the MSE and BER performances of the proposed LS estimator are evaluated for different lengths of the training sequence. Figure 5 shows the MSE performance.

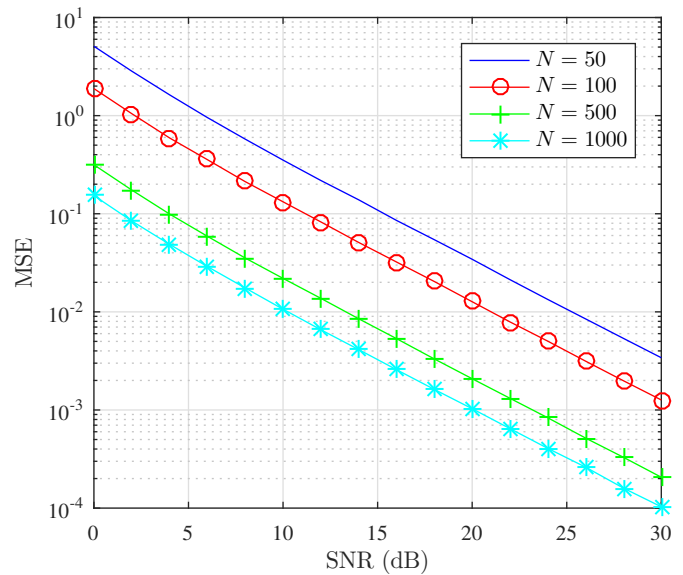


Fig. 5. MSE of $\hat{\mathbf{a}}$ for different training sequence lengths.

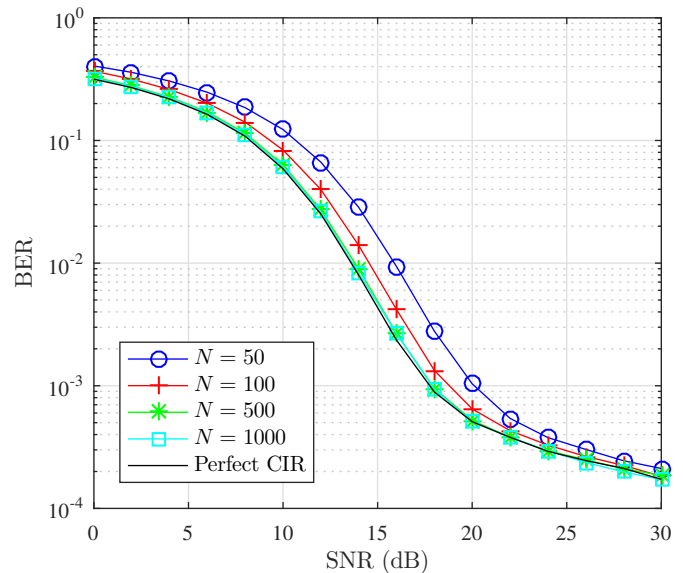


Fig. 6. BER for different training sequence lengths at the LS estimator.

As the MSE for both concatenated channels are equal, only the MSE for channel $\hat{\mathbf{a}}$ is shown.

As expected, the estimator performance degrades as the length of the training sequence decreases. For an SNR of 10 dB, the training sequence of length $N = 50$ has an MSE performance of only 0.378, while for a length of $N = 100$ training symbols it is 0.132. For $N = 500$, the MSE is 0.022, and for $N = 1000$ the system achieves an MSE of 0.011.

Figure 6 shows the BER for the same training sequence lengths simulated in Figure 5. As observed in Figures 5 and 6, although the MSE is better for a longer training sequence, it does not always improve significantly the BER. For instance, for $N = 500$, $N = 1000$ and perfect CIR, the BER are practically the same, and the differences can be due to errors in numerical simulation. For $N = 50$, $N = 100$ and $N = 500$, it is possible

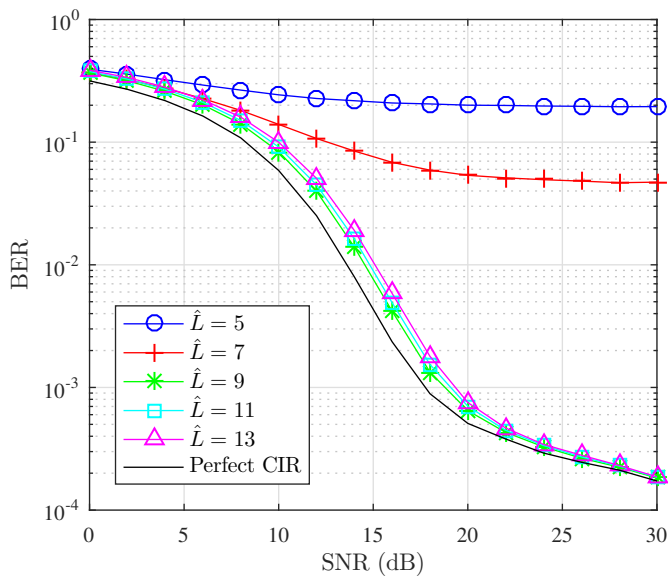


Fig. 7. BER considering different channel lengths at the estimator.

to see an improvement in the BER. However, using $N = 500$ instead of $N = 100$ saves less than 1 dB in SNR to achieve almost the same BER. For an MSE of 10^{-2} there is a difference of approximately 12 dB between the PNC system that uses a $N = 1000$ (the block length) and the PNC system that uses $N = 100$, i. e. 10 % of the block length. However, this difference does not impact significantly the BER. Therefore, the proposed estimator can be deployed without the need of a high number of training symbols. Furthermore, it is possible to see in Figure 6 that an error floor happens near the 20 dB SNR region even for the perfect knowledge of CSI. This floor is due to the depth of the traceback of the Viterbi algorithm used to implement the MLSE equalizer, that was set to $5L$. As the length of transmitted symbols block is 10^3 , the algorithm could not remove all the intersymbol interference (ISI). In order to reduce this BER floor, the length of the block of transmitted symbols can be increased.

The second scenario evaluates the impact of erroneous channel length \hat{L} at the estimator, i.e., when it considers a channel length smaller (or greater) than the actual channel length L . For this simulation, it was used a training sequence of length $N = 100$. Figure 7 shows that the BER does not change significantly when $\hat{L} > L$, since the extra coefficients given by the estimator are nearly zero, but decreases considerably when $\hat{L} < L$. For $\hat{L} = 5$, there are almost no reduction in BER as the SNR increases. This happens because the ISI generated by the frequency selective wireless channel are not totally mitigated. First, the self removal given by (6) are not performed correctly because the estimate $\hat{\mathbf{a}}$ does not consider all the channel coefficients. Second, as the equalizer uses the information $\hat{\mathbf{b}}$, that was obtained considering less coefficients than the actual cascaded channel, its output will have residual ISI, which degrades the overall performance of the system.

In the third scenario, a comparison between the proposed estimation technique with the LS estimator proposed in [7] for OFDM-based PNC is done. Although it is an ANC based on

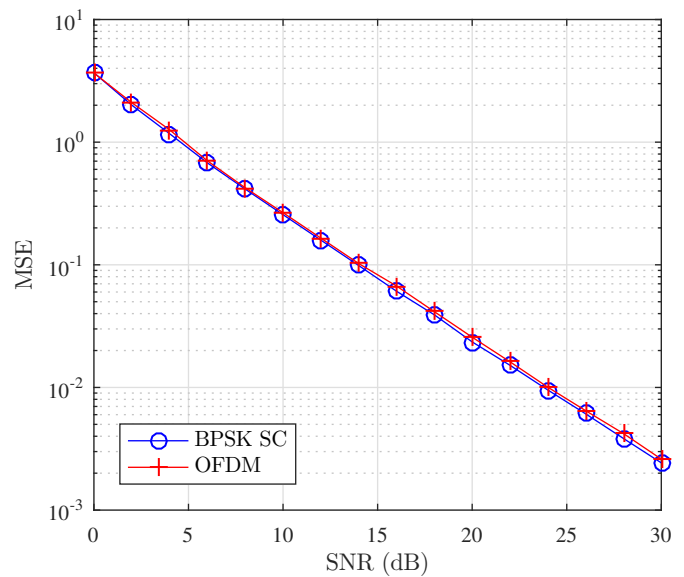


Fig. 8. MSE comparison between the LS estimator for OFDM system and the LS estimator for SC system.

OFDM modulation, the channel estimator estimates the CIR after removing the Cyclic Prefix (CP) and before computing the Discrete Fourier Transform, i. e., it works in the time domain, making it a fair comparison for the MSE. The OFDM system uses 64 carriers and CP length of 16. In this scenario, the proposed technique uses $N = 64$ training symbols to match with the length of the sequence used for the OFDM system, once it uses one OFDM block as training sequence, i. e., 64 training symbols.

It can be seen from Figure 8 that the performances of both estimators are equivalent, and minor differences are due to numerical errors in simulation. Although these estimator are based in the same criterion, they are distinct estimators, since the construction of the matrix given by (7) is different. This is because the estimator proposed here deals with SC modulation, rather than OFDM as in [7].

Finally, the MSE obtained by using the optimal training sequence, derived in this paper, is compared to the MSE obtained by using a random training sequence in Figure 9. This figure shows that the optimal training sequence provides a better MSE. However, it can be seen that for $N = 1000$, the impact is smaller than for $N = 100$. For an MSE of 10^{-2} , a random sequence of length $N = 100$ needs an SNR of around 21 dB, while the optimal one needs around 19.7 dB. It thus represents an 1.3 dB difference. However, for a training sequence length of $N = 1000$, changing from the random sequence to optimal sequence reduces the required MSE from around 10.3 dB to near 9.7 dB, which consists in a reduction of less than 1 dB.

The better performance of the optimal training sequence happens because the squared Frobenius norm $\|\mathbf{X}\|_F^2 = \text{tr}\{\mathbf{X}^H\mathbf{X}\}$ is smaller for the random sequence than for the optimal sequence, although they have the same power. This happens due to the structure of the convolution matrix in (8). As the Frobenius norm is smaller for the random sequence,

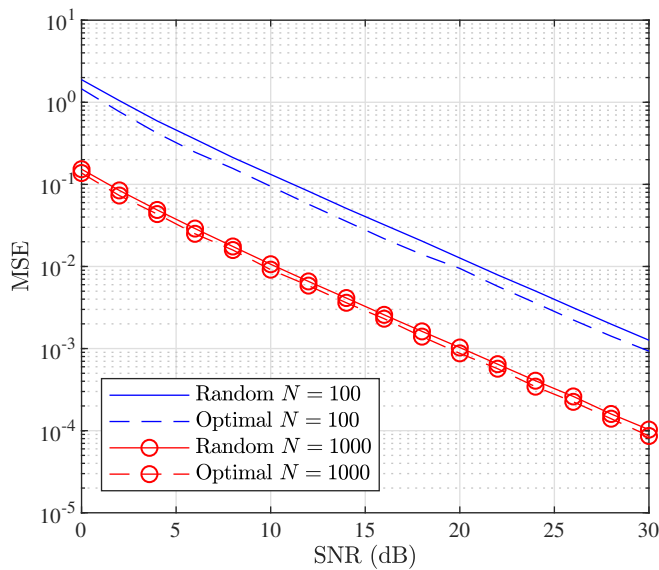


Fig. 9. MSE comparison between the random and the optimal training sequences for different lengths.

$\text{tr}\{(\mathbf{X}^H\mathbf{X})^{-1}\}$ will be larger and, therefore, the MSE given by (24) will be larger. For longer training sequences, this difference is smaller and, therefore, the impact of the optimal sequence in reducing the MSE is also smaller.

V. CONCLUSIONS

In this paper, a Least Squares channel estimation technique is proposed for PNC communication systems that uses single carrier modulation and operates under frequency-selective fading channels.

It is shown that a longer training sequence yields a better performance in terms of MSE. However, it is possible to see that a highly accurate estimation does not necessary imply a better BER performance and, therefore, shorter training sequences can be used. It can be seen from the results shown in section IV that with a training sequence of length $N = 100$, the proposed estimator performs, in terms of BER, nearly as well as a system considering the perfect knowledge of the CIR.

Furthermore, simulations considering channel lengths at the estimator different from the real one shows that using a longer channel length at the estimator does not bring any penalty to the system performance when compared to the real length, keeping an equivalent BER. On the other hand, by considering a channel length smaller than the real one, the BER performance degrades considerably. Hence, it is important to develop an accurate estimator for the channel length at the end nodes.

The proposed estimator was also compared to the estimator presented in [7]. Although the latter considers an OFDM system and has different mathematical constructions, this estimator works in time domain, so the MSE comparison is fair. It is shown that both estimators have equivalent MSE performance.

Moreover, the design of an optimal training sequence, in terms of the channel estimator MSE, is presented. It is shown that it provides a better performance compared to a random training sequence. However, for larger training

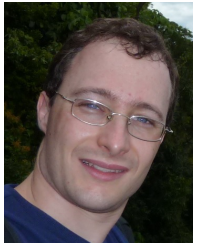
sequence lengths, the improvement is smaller than for shorter sequences.

REFERENCES

- [1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Statist. Probab.*, 1961.
- [2] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: Physical-layer network coding," in *Proc. MobiCom '06*, 2006, doi:10.1145/1161089.1161129.
- [3] P. Popovski and H. Yomo, "The anti-packets can increase the achievable throughput of a wireless multi-hop network," in *IEEE Proc. ICC '06*, IEEE, Jun. 2006, doi:10.1109/ICC.2006.255688.
- [4] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: analog network coding," in *Proc. SIGCOMM '07*, 2007, doi:10.1145/1282380.1282425.
- [5] F. Gao, R. Zhang, and Y.-C. Liang, "On channel estimation for amplify-and-forward two-way relay networks," in *Proc IEEE Globecom '08*, IEEE, Nov. 2008, doi:10.1109/GLOCOM.2008.ECP.705.
- [6] —, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, Oct. 2009, doi:10.1109/TCOMM.2009.10.080169.
- [7] —, "Channel estimation for OFDM modulated two-way relay networks," *IEEE Trans. Signal Process.*, vol. 57, Nov. 2009, doi:10.1109/TSP.2009.2026537.
- [8] W. Yang, Y. Cai, J. Hu, and W. Yang, "Channel estimation for two-way relay OFDM networks," *Eurasip Journal on Wireless Communications and Networking*, vol. 2010, Jul. 2010, doi:10.1155/2010/186182.
- [9] Jin Soo Wang, Minh Tam Tran, Dong Ryul Shin, and Yun Hee Kim, "Linear channel estimation and training for frequency-selective fading channels in two-way relay networks," in *The Sixth International Workshop on Signal Design and Its Applications in Communications*. Tokyo: IEEE, Oct. 2013, doi:10.1109/IWSDA.2013.6849092.
- [10] H. Gacanin, T. Sjödin, and F. Adachi, "On channel estimation for analog network coding in a frequency-selective fading channel," *Eurasip Journal on Wireless Communications and Networking*, vol. 2011, Jan. 2011, doi:10.1155/2011/980430.
- [11] J. Proakis and M. Salehi, *Communication Systems Engineering*. Prentice Hall, 2002.
- [12] R. Ziemer and R. Peterson, *Introduction to digital communication*. Macmillan Pub. Co., 1992.
- [13] F. Rossetto, "A comparison of different physical layer network coding techniques for the satellite environment," in *2010 5th Adv. Satell. Multimed. Syst. Conf. 11th Signal Process. Sp. Commun. Work.* IEEE, Sep. 2010, pp. 25–30.
- [14] S. D. Buch and B. S. Rajan, "Performance analysis of physical layer network coding for two-way relaying over non-regenerative communication satellites," in *2017 IEEE 86th Veh. Technol. Conf.* IEEE, Sep. 2017, pp. 1–7, doi:10.1109/VTCTFall.2017.8288244.
- [15] S. Al-Sarawi, M. Anbar, K. Alieyan, and M. Alzubaidi, "Internet of things (iot) communication protocols: Review," in *2017 8th International Conference on Information Technology (ICIT)*, May 2017, pp. 685–690, doi:10.1109/ICITECH.2017.8079928.
- [16] P. I. Cruz and M. B. Loiola, "Channel estimation for two-way relay networks over frequency selective fading channels," in *XXXIV SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES - SBrT2016*, Santarém, 2016, pp. 989–992.
- [17] R. Ahlswede, Ning Cai, S.-Y. Li, and R. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000, doi:10.1109/18.850663.
- [18] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.
- [19] L. Trefethen and D. Bau, *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, 1997.
- [20] M. Biguesh and A. Gershman, "Training-based mimo channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006, doi:10.1109/TSP.2005.863008.



Pedro Ivo da Cruz is currently pursuing his PhD in Information Engineering at Federal University of ABC (UFABC). He received the MSc and the BSc degrees in Information Engineering in 2017 and 2014, respectively, and the BSc in Science and Technology in 2013, from the same university. His research interests include wireless communications, adaptive and statistical signal processing, Wireless Physical-layer Network Coding systems and Wireless Physical-layer Security.



Murilo Bellezoni Loiola received the titles of Electrical Engineer (2002), Master in Electrical Engineering (2005) and Doctor in Electrical Engineering (2009) from the University of Campinas (UNICAMP), Brazil. Currently, he is an Associate Professor at Federal University of ABC (UFABC). His main research interests lie in the areas of adaptive and statistical signal processing, wireless communications, wireless physical-layer security, and machine learning.