

BLOCK-BASED TRANSCEIVERS FOR FREQUENCY SELECTIVE CHANNELS WITH REDUCED REDUNDANCY

L. REY VEGA[†] and C.G. GALARZA[†]

[†]CONICET and Facultad de Ingeniería, U.B.A., Paseo Colón 850, Buenos Aires, Argentina
lrey@fi.uba.ar, cgalar@fi.uba.ar

Abstract— We propose a transceiver structure for a frequency selective channel that allows the introduction of reduced redundancy. We optimize jointly the transmitter and receiver in this structure to maximize the information rate. The simulation results show that the proposed design has good performance while increasing the bandwidth efficiency.

Keywords— Block transceivers, inter-symbol interference (ISI), redundancy, inter-block interference (IBI), mutual information.

I. INTRODUCTION

Time invariant single-input-single-output (SISO) communication channels have been studied for a while already. However, the problem of achieving high transmission rates under limited computational resources has not been completely solved yet. For instance, block-by-block communication systems have shown to be efficient schemes for transmitting over inter-symbol interference (ISI) channels (Lechleider, 1990; Scaglione *et al.*, 1999b; Lin and Phoong, 2000). Examples of block-by-block communication systems are orthogonal frequency division multiplexing (OFDM) and discrete multitone modulation (DMT) (Chow *et al.*, 1991). In a block-based system, transmitter and receiver act on a block-by-block basis and redundancy is added to each block in order to remove inter-block interference (IBI). However, the introduction of redundancy leads to a worse bandwidth efficiency. Although this concern is immaterial for large symbol block sizes, small-sized blocks are still worthy to avoid large decoding delays and computational complexity. In the past, ISI free multirate FIR filterbank transceivers with reduced redundancy have been obtained by Lin and Phoong (2002). However, these solutions may not lead to a block-by-block communication system. Although necessary conditions for the existence of a parameterization for all block-by-block ISI free transceivers with reduced redundancy have been derived by Lin and Phoong (2002), a sufficient condition is still lacking.

In this paper we propose a family of transceivers with reduced redundancy that is obtained by controlling the IBI. Also, we derive all the transceivers that

maximize the information rate subject to a constraint of maximum transmit power.

The notation is as follows: boldfaced lower-case letters represent vectors and boldfaced upper-case letters are reserved for matrices. The notation \mathbf{A}^T represents the transpose of \mathbf{A} . The notations $\mathbf{I}_{M \times M}$ and $\mathbf{0}_{M \times N}$ represent the $M \times M$ identity matrix and $M \times N$ null matrix respectively. Also $|\mathbf{A}|$ and $\|\mathbf{A}\|_F$ denote the determinant and the Frobenius norm of matrix \mathbf{A} . The notation $\lceil \cdot \rceil$ stands for the ceiling operation on real numbers and $I(x; y)$ is the mutual information between random variables x and y .

II. PROBLEM STATEMENT

Let h_n be the discrete time impulse response of the SISO channel. We assume that h_n is a linear time invariant (LTI) system with finite impulse response (FIR) of length $L + 1$. We analyze the problem of sending and receiving blocks of N samples. For that, the vectors $\mathbf{s}(n) = [s_{nN} \ s_{nN-1} \ \dots \ s_{nN-N+1}]^T$, and $\mathbf{y}(n) = [y_{nN} \ y_{nN-1} \ \dots \ y_{nN-N+1}]^T$ represent the n -th blocks sent and received through the channel. Then, assuming that $L + 1 < N$, the IBI is limited to two consecutive blocks only and $\mathbf{y}(n)$ is expressed as

$$\mathbf{y}(n) = \mathbf{H}_1 \mathbf{s}(n) + \mathbf{H}_2 \mathbf{s}(n-1) + \mathbf{v}(n), \quad (1)$$

where $\mathbf{v}(n) \in \mathfrak{R}^N$ is the n -th block of additive noise at the output of the channel, $\mathbf{H}_1 \in \mathfrak{R}^{N \times N}$, and $\mathbf{H}_2 \in \mathfrak{R}^{N \times N}$. In particular, \mathbf{H}_1 is upper triangular Toeplitz whose first row is $[h_0 \ \dots \ h_L \ \dots \ 0] \in \mathfrak{R}^N$. Similarly, \mathbf{H}_2 is lower triangular Toeplitz whose first column is $[0 \ \dots \ h_L \ \dots \ h_1]^T \in \mathfrak{R}^N$. Finally, we consider

$$\mathbf{s}(n) = \mathbf{A} \mathbf{x}(n) \quad \text{and} \quad \hat{\mathbf{x}}(n) = \mathbf{B} \mathbf{y}(n), \quad (2)$$

where $\mathbf{x}(n) = [x_{nM} \ x_{nM-1} \ \dots \ x_{nM-M+1}]^T$ is the n -th transmitted block formed by M symbols, $\hat{\mathbf{x}}(n)$ is the n -th estimated block, $\mathbf{A} \in \mathfrak{R}^{N \times M}$ is the transmit matrix, and $\mathbf{B} \in \mathfrak{R}^{M \times N}$ is the receive matrix.

Well-known transceiver structures as the zero-padding (ZP) (Lin and Phoong, 2002; Lin and Phoong, 2001; Scaglione *et al.*, 1999a,b) and the cyclic prefix (CP) technique (Chow *et al.*, 1991) may be formulated under this framework. For instance, when using ZP with leading zeros, $N = M + L$ and the first L

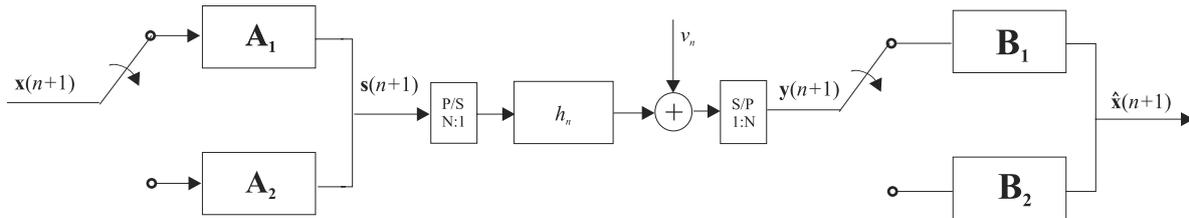


Figure 1: Transceiver structure: the blocks in the transmitter are converted to a serial stream before entering the channel; at the receiver, the serial stream is converted into blocks of N samples. In a synchronous way, the switch at the transmitter acts on each block of M symbols, and the switch at the receiver acts on each block of N samples.

rows of the transmit matrix \mathbf{A} are zero. It is also possible to define matrices \mathbf{A} and \mathbf{B} to include a prefix of length L when using the CP technique (Lin and Phoong, 2000). In general, $N > M$ and $Q := N - M$ is the redundancy that helps to remove the IBI inherent to any block-based transmission system. Although redundancy may be beneficial (Scaglione *et al.*, 1999b), it uses more bandwidth because more symbols have to be transmitted in a given period of time. If T is the signaling period, then the actual transmission rate is $(M + Q)/MT$, leading to an excess bandwidth of Q/MT due to the redundancy. Our goal is to build a structured redundancy to improve the bandwidth use without decreasing the amount of information transmitted per block, i.e., to improve the bandwidth efficiency. We replace (2) into (1) to obtain

$$\hat{\mathbf{x}}(n) = \mathbf{B}\mathbf{H}_1\mathbf{A}\mathbf{x}(n) + \mathbf{B}\mathbf{H}_2\mathbf{A}\mathbf{x}(n-1) + \mathbf{B}\mathbf{v}(n). \quad (3)$$

We see that for dispersive channels IBI between successive blocks cannot be avoided if no particular structure is imposed onto the transceiver. For example, the structure of a ZP transceiver is such that the product $\mathbf{H}_2\mathbf{A}$ in (3) vanishes and the interfering term disappears. Notice that techniques such as ZP or CP utilize a redundancy length $Q = L$. It is clear that large values of L lead to a significant decrease in the information rate that may be compensated by augmenting the block length. However, by increasing the block length we are also increasing the decoding delay and the computational load. Therefore, a tradeoff between computational requirements and bandwidth efficiency is set up. In next section, we propose to transmit a redundancy smaller than L by controlling the IBI, not by suppressing it. In particular, we pose the structure of a transceiver and we design its component so that the average mutual information per sample between the transmitted symbols and the received ones, or the information rate, is maximized.

III. TRANSCIEVERS WITH REDUCED REDUNDANCY

We assume that complete channel state information (CSI) is available at the transmitter and receiver and that the channel remains invariant at least for the

duration of a block. On a first approach, the symbols x_n are zero-mean, uncorrelated continuous random variables, i.e. coming from a constellation with infinitesimal granularity. We defer for later the case of symbols taken from finite alphabets. Under these assumptions, $E[\mathbf{x}(n)\mathbf{x}^T(n)] = \sigma_x^2\mathbf{I}_{M \times M}$. The noise v_n is assumed to be Gaussian and zero-mean. The noise autocorrelation matrix is given by $E[\mathbf{v}(n)\mathbf{v}^T(m)] = \mathbf{R}_{vv}\delta_{nm}$, where δ_{nm} is the Kronecker function. Finally, the noise and the symbols are uncorrelated, i.e. $E[\mathbf{x}(n)\mathbf{v}^T(m)] = \mathbf{0}_{M \times N}$, $\forall n, m$, where $N = M + Q$.

A. Structure of the transceivers

Suppose that two transceivers, $(\mathbf{A}_1, \mathbf{B}_1)$ and $(\mathbf{A}_2, \mathbf{B}_2)$, are available. Then, we will use them in an alternating way on two successive blocks as follows:

$$\begin{aligned} \hat{\mathbf{x}}(n+1) &= \mathbf{B}_1\mathbf{H}_1\mathbf{A}_1\mathbf{x}(n+1) + \mathbf{B}_1\mathbf{H}_2\mathbf{A}_2\mathbf{x}(n) + \\ &\quad \mathbf{B}_1\mathbf{v}(n+1), \\ \hat{\mathbf{x}}(n) &= \mathbf{B}_2\mathbf{H}_1\mathbf{A}_2\mathbf{x}(n) + \mathbf{B}_2\mathbf{H}_2\mathbf{A}_1\mathbf{x}(n-1) + \\ &\quad \mathbf{B}_2\mathbf{v}(n). \end{aligned} \quad (4)$$

This structure may be thought of as a periodic block transceiver where the two pairs of transmitters and receivers act in an alternating way. Figure 1 shows the operation of this structure. The block of data $\mathbf{x}(n)$ is processed by \mathbf{A}_2 before being sent through the channel. At the receiver, the signal block $\mathbf{y}(n)$ is multiplied by the receive matrix \mathbf{B}_2 . Then, next block of data $\mathbf{x}(n+1)$ is processed by the transmit matrix \mathbf{A}_1 , and the corresponding block $\mathbf{y}(n+1)$ is received by \mathbf{B}_1 . The same process is repeated on the following block using \mathbf{A}_2 and \mathbf{B}_2 , and so on.

Clearly, model (4) does not lead to an optimum solution if one would want to maximize the average mutual information. An optimum structure would consider joint spectral shaping and decoding of successive blocks of symbols to account for the IBI terms. However that structure would have large computational complexity. Instead, model (4) entails block by block coding and decoding operations. Still, the effect of the IBI terms can be controlled by carefully selecting the pairs $(\mathbf{A}_1, \mathbf{B}_1)$ and $(\mathbf{A}_2, \mathbf{B}_2)$.

Following the line of thought of Lin and Phoong (2003), it can be proved that if the redundancy length

satisfies that $\lceil L/2 \rceil \leq Q \leq L$, it is possible to eliminate the IBI term $\mathbf{B}_1 \mathbf{H}_2 \mathbf{A}_2 \mathbf{x}(n)$. Select \mathbf{B}_1 and \mathbf{A}_2 as

$$\mathbf{B}_1 = [\mathbf{B}'_1 \ \mathbf{0}_{M \times Q}], \quad \mathbf{A}_2 = \begin{bmatrix} \mathbf{0}_{Q \times M} \\ \mathbf{A}'_2 \end{bmatrix}, \quad (5)$$

where \mathbf{B}'_1 and \mathbf{A}'_2 are $M \times M$ matrices, and (5) corresponds to ZP for $\hat{\mathbf{x}}(n+1)$ using $Q < L$ zeros. Then,

$$\begin{aligned} \hat{\mathbf{x}}(n+1) &= \mathbf{B}'_1 \mathbf{H}_{1w} \mathbf{A}_1 \mathbf{x}(n+1) + \mathbf{B}'_1 \mathbf{v}'(n+1) \\ \hat{\mathbf{x}}(n) &= \mathbf{B}_2 \mathbf{H}_{1t} \mathbf{A}'_2 \mathbf{x}(n) + \mathbf{B}_2 \mathbf{H}_2 \mathbf{A}_1 \mathbf{x}(n-1) + \mathbf{B}_2 \mathbf{v}(n), \end{aligned} \quad (6)$$

where $\mathbf{v}'(n)$ is a length M vector obtained from $\mathbf{v}(n)$ by removing its last Q entries, \mathbf{H}_{1w} is a $M \times N$ matrix obtained by removing the last Q rows from \mathbf{H}_1 , and \mathbf{H}_{1t} is a $N \times M$ matrix obtained by removing the first Q columns from \mathbf{H}_1 . Notice that by using (5), the IBI term on $\hat{\mathbf{x}}(n+1)$ is eliminated. Moreover, the energy of the IBI term on $\hat{\mathbf{x}}(n)$, namely $\mathbf{B}_2 \mathbf{H}_2 \mathbf{A}_1 \mathbf{x}(n-1)$, may be controlled with a judicious choice of \mathbf{B}_2 and \mathbf{A}_1 . In the next subsection we will maximize the information rate for this choice of transceiver structure.

B. Optimization of the transceivers

We want to maximize the information rate for two consecutive blocks of symbols when a bound is imposed on the total power transmitted by these two blocks. For that, define the average mutual information per sample as

$$F(\mathbf{A}_1, \mathbf{A}'_2, \mathbf{B}'_1, \mathbf{B}_2) := \frac{I(\hat{\mathbf{x}}(n+1), \hat{\mathbf{x}}(n); \mathbf{x}(n+1), \mathbf{x}(n))}{2N}, \quad (7)$$

where we have emphasized the dependence of the objective function on the transceivers. We are now in conditions to set up the optimization problem:

$$\begin{aligned} \max_{\mathbf{A}_1, \mathbf{A}'_2, \mathbf{B}'_1, \mathbf{B}_2} & F(\mathbf{A}_1, \mathbf{A}'_2, \mathbf{B}'_1, \mathbf{B}_2) \\ \text{s.t.} & \sigma_x^2 \text{tr}(\mathbf{A}_1 \mathbf{A}_1^T) + \sigma_x^2 \text{tr}(\mathbf{A}'_2 \mathbf{A}'_2{}^T) \leq P. \end{aligned} \quad (8)$$

The restriction on the total power of $\mathbf{s}(n+1)$ and $\mathbf{s}(n)$ allows for an unequal power distribution between consecutive blocks. In principle this is better than assigning equal power to both output blocks, since they are not facing the same channel, as it is shown in (6).

Before solving problem (8), let us analyze the structure of the mutual information $I(\hat{\mathbf{x}}(n+1), \hat{\mathbf{x}}(n); \mathbf{x}(n+1), \mathbf{x}(n))$ when (5) is used. First, recall that $\mathbf{v}'(n+1)$ and $\mathbf{v}(n)$ are uncorrelated vectors, and the components of $\mathbf{x}(n)$ are uncorrelated in time and uncorrelated with $\mathbf{v}(n)$. Now, using the definition of mutual information (Cover and Thomas, 1991) we can show that:

$$\begin{aligned} I(\hat{\mathbf{x}}(n+1), \hat{\mathbf{x}}(n); \mathbf{x}(n+1), \mathbf{x}(n)) &= \\ I(\hat{\mathbf{x}}(n+1); \mathbf{x}(n+1)) &+ I(\hat{\mathbf{x}}(n); \mathbf{x}(n)). \end{aligned} \quad (9)$$

Let us define the equivalent noise terms in (6) as $\mathbf{z}_1(n) := \mathbf{B}'_1 \mathbf{v}'(n)$ and $\mathbf{z}_2(n) := \mathbf{B}_2 \mathbf{v}(n) + \mathbf{B}_2 \mathbf{H}_2 \mathbf{A}_1 \mathbf{x}(n-1)$, and their autocorrelation matrices

as $\mathbf{R}_{z_1 z_1}$ and $\mathbf{R}_{z_2 z_2}$. Because the noise v_n is assumed to be Gaussian and independent of the input symbols x_n , and assuming that x_n are zero-mean and uncorrelated, the mutual information terms in the RHS of (9) are maximized when $\mathbf{x}(n+1)$ and $\mathbf{x}(n)$ are Gaussian random vectors with autocorrelation matrix $\sigma_x^2 \mathbf{I}_{M \times M}$ (Cover, 1991). Then, it can be shown that

$$\begin{aligned} I(\hat{\mathbf{x}}(n+1); \mathbf{x}(n+1)) &= \frac{1}{2} \log_2 |\mathbf{I}_{M \times M} + \sigma_x^2 \mathbf{A}'_1{}^T \mathbf{H}_{1w}^T \cdot \\ &\quad \mathbf{B}'_1{}^T \mathbf{R}_{z_1 z_1}^{-1} \mathbf{B}'_1 \mathbf{H}_{1w} \mathbf{A}_1| \end{aligned} \quad (10)$$

$$\begin{aligned} I(\hat{\mathbf{x}}(n); \mathbf{x}(n)) &= \frac{1}{2} \log_2 |\mathbf{I}_{M \times M} + \sigma_x^2 \mathbf{A}'_2{}^T \mathbf{H}_{1t}^T \mathbf{B}_2^T \cdot \\ &\quad \mathbf{R}_{z_2 z_2}^{-1} \mathbf{B}_2 \mathbf{H}_{1t} \mathbf{A}'_2|. \end{aligned} \quad (11)$$

Notice that equations (10) and (11) are effectively coupled by $\mathbf{R}_{z_2 z_2}$ that is parameterized by \mathbf{A}_1 and \mathbf{B}_2 . This coupling represents the effect of the IBI on the information rate of the block system. Block $\mathbf{x}(n-1)$, which is spectrally shaped by the matrix \mathbf{A}_1 , acts as an interference source for block $\mathbf{x}(n)$. As this interference is included in $\mathbf{z}_2(n)$ and hence in its autocorrelation $\mathbf{R}_{z_2 z_2}$, equations (10) and (11) cannot be solved separately.

Hadamard's inequality (Cover, 1991) implies that (7) is maximized when the arguments of the determinants in (10) and (11) are simultaneously diagonalized. Thanks to the structure assumed in (5), simultaneous diagonalization is possible. Otherwise the optimization problem would be more difficult. Now, following to Scaglione *et al.* (1999a) we choose \mathbf{A}_1 and \mathbf{B}'_1 in the following manner

$$\mathbf{A}_1 = \mathbf{V}_1 \begin{bmatrix} \boldsymbol{\Sigma}_{A_1} \\ \mathbf{0}_{Q \times M} \end{bmatrix}, \quad \mathbf{B}'_1 = \mathbf{W}_1 \mathbf{Q}_1^T \mathbf{R}_{v'v'}^{-1/2} \quad (12)$$

to diagonalize the argument of (10). In the previous equation, \mathbf{W}_1 is an arbitrary $M \times M$ invertible matrix, $\mathbf{R}_{v'v'}$ is the autocorrelation matrix of $\mathbf{v}'(n+1)$, $\boldsymbol{\Sigma}_{A_1} := \text{diag}(\sigma_{A_1}^{ii})$ is a $M \times M$ diagonal matrix with positive entries, and \mathbf{V}_1 and \mathbf{Q}_1 are $N \times N$ and $M \times M$ orthonormal matrices given by the following singular value decomposition:

$$\mathbf{R}_{v'v'}^{-1/2} \mathbf{H}_{1w} = \mathbf{Q}_1 [\mathbf{A}_1 \ \mathbf{0}_{M \times Q}] \mathbf{V}_1^T. \quad (13)$$

Here, $\mathbf{A}_1 := \text{diag}(\lambda_1^{ii})$ has dimension $M \times M$ and its entries along the diagonal are all positive. Similarly, the argument of (11) is diagonalized when

$$\mathbf{A}'_2 = \mathbf{V}_2 \boldsymbol{\Sigma}_{A_2}, \quad \mathbf{B}_2 = [\mathbf{W}_2 \ \mathbf{0}_{M \times Q}] \mathbf{Q}_2^T \mathbf{R}_{uu}^{-1/2}. \quad (14)$$

As before, \mathbf{W}_2 is an arbitrary and invertible $M \times M$ matrix, and $\boldsymbol{\Sigma}_{A_2} := \text{diag}(\sigma_{A_2}^{ii})$ is a diagonal matrix with positive entries. Now, $\mathbf{R}_{uu} := \sigma_x^2 \mathbf{H}_2 \mathbf{A}_1 \mathbf{A}_1^T \mathbf{H}_2 + \mathbf{R}_{vv}$, and \mathbf{V}_2 and \mathbf{Q}_2 are $M \times M$ and $N \times N$ orthonormal matrices given by the following singular value decomposition:

$$\mathbf{R}_{uu}^{-1/2} \mathbf{H}_{1t} = \mathbf{Q}_2 \begin{bmatrix} \boldsymbol{\Lambda}_2 \\ \mathbf{0}_{Q \times M} \end{bmatrix} \mathbf{V}_2^T. \quad (15)$$

Once again, the diagonal matrix $\mathbf{\Lambda}_2 := \text{diag}(\lambda_2^{ii})$ has dimension $M \times M$ and its entries along the diagonal are positive. Using the expression of \mathbf{A}_1 in (12) and the definition of \mathbf{R}_{uu} we can express (15) as

$$\begin{aligned} \left(\sigma_x^2 \mathbf{H}_2 \mathbf{V}_1 \begin{bmatrix} \mathbf{\Sigma}_{A_1}^2 & \mathbf{0}_{M \times Q} \\ \mathbf{0}_{Q \times M} & \mathbf{0}_{Q \times Q} \end{bmatrix} \mathbf{V}_1^T \mathbf{H}_2 + \mathbf{R}_{vv} \right)^{-1/2} \mathbf{H}_{1t} \\ = \mathbf{Q}_2 \begin{bmatrix} \mathbf{\Lambda}_2 \\ \mathbf{0}_{Q \times M} \end{bmatrix} \mathbf{V}_2^T. \end{aligned} \quad (16)$$

Then, replacing (12) and (14) into (8) the optimization problem is re-formulated as

$$\begin{aligned} \max_{\sigma_{A_1}^{ii}, \sigma_{A_2}^{ii}} \frac{\sum_{i=1}^M \frac{1}{2} \log_2 [1 + D_i] + \sum_{i=1}^M \frac{1}{2} \log_2 [1 + E_i]}{2N} \\ \text{s.t. } \sigma_x^2 \sum_{i=1}^M (\sigma_{A_1}^{ii})^2 + \sigma_x^2 \sum_{i=1}^M (\sigma_{A_2}^{ii})^2 \leq P \end{aligned} \quad (17)$$

where $D_i = \sigma_x^2 (\sigma_{A_1}^{ii})^2 (\lambda_1^{ii})^2$ and $E_i = \sigma_x^2 (\sigma_{A_2}^{ii})^2 (\lambda_2^{ii})^2$. Equation (17) shows that given the parameters $\{\lambda_1^{ii}, \lambda_2^{ii}, i = 1 \dots M\}$ the set of solutions $\{\sigma_{A_1}^{ii}, \sigma_{A_2}^{ii}, i = 1 \dots M\}$, is obtained by the classical water-filling solution (Cover, 1991) for $2M$ parallel Gaussian channels. However this solution implicitly assumes that λ_1^{ii} and λ_2^{ii} do not depend on $\{\sigma_{A_1}^{jj}, \sigma_{A_2}^{jj}, j = 1 \dots M\}$. While this is true for λ_1^{ii} it is not true for λ_2^{ii} , which is a function of $\{\sigma_{A_1}^{jj}, j = 1 \dots M\}$ as it can be deduced from (16). Also from (13) and (16) it can be seen that \mathbf{V}_1 and \mathbf{Q}_1 are independent of $\{\sigma_{A_1}^{jj}, \sigma_{A_2}^{jj}, j = 1 \dots M\}$ while \mathbf{V}_2 and \mathbf{Q}_2 depend on $\{\sigma_{A_1}^{jj}, j = 1 \dots M\}$. To obtain a close form solution for problem (17) we need to formulate the functional dependence among the elements λ_2^{ii} and $\sigma_{A_1}^{jj}$, $i, j = 1 \dots M$. Unfortunately, this is not possible in general.

For that reason we have to turn to an algorithmic approach. For instance, using the channel knowledge, we solve (13) to obtain \mathbf{V}_1 , $\mathbf{\Lambda}_1$ and \mathbf{Q}_1 . Then, using λ_1^{ii} , we formulate the following problem:

$$\max_{\sigma_{A_1}^{ii}} \frac{\sum_{i=1}^M \frac{1}{2} \log_2 [1 + D_i]}{N}, \quad \text{s.t. } \sigma_x^2 \sum_{i=1}^M (\sigma_{A_1}^{ii})^2 \leq P/2. \quad (18)$$

The solution to this problem is obtained by using well-known algorithms (Scaglione *et al.*, 1999a). In next step we use \mathbf{V}_1 and $\mathbf{\Sigma}_{A_1} = \text{diag}(\sigma_{A_1}^{ii})$ to solve (16) for \mathbf{V}_2 , $\mathbf{\Lambda}_2$ and \mathbf{Q}_2 . Now, with the values of $\lambda_{A_1}^{ii}$ and $\lambda_{A_2}^{ii}$ already obtained, we solve problem (17). The result is a new set $\{\sigma_{A_1}^{ii}, \sigma_{A_2}^{ii}, i = 1 \dots M\}$. We iterate on the procedure by inserting the new values of $\sigma_{A_1}^{ii}$ into (16), computing a new singular value decomposition, and solving the water-filling problem (17). The procedure is repeated until a stop condition is met. The algorithm is summarized below:

Algorithm 1

1. Initialization: set $k = 0$. Compute \mathbf{V}_1 , $\mathbf{\Lambda}_1$ and \mathbf{Q}_1 using (13). Solve problem (18). Generate $\mathbf{\Sigma}_{A_1}^0$. Set $k = 1$.

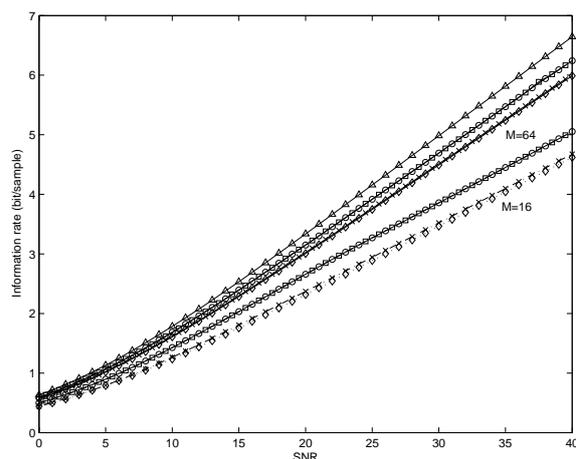


Figure 2: Channel capacity and information rates in bits per sample as functions of SNR for $M = 16$ and $M = 64$. Legend: \triangle : channel capacity; \circ : scheme proposed (algorithm 1); \square : scheme proposed (algorithm 2); \times : scheme proposed by Scaglione *et al.* (1999a); \diamond : DMT.

2. Compute \mathbf{V}_2^k , $\mathbf{\Lambda}_2^k$ and \mathbf{Q}_2^k using (16) and $\mathbf{\Sigma}_{A_1}^{k-1}$.
3. With $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2^k$ solve the water-filling problem (17). Construct $\mathbf{\Sigma}_{A_1}^k$ and $\mathbf{\Sigma}_{A_2}^k$.
4. If $\|\mathbf{\Sigma}_{A_1}^k - \mathbf{\Sigma}_{A_1}^{k-1}\|_F^2 + \|\mathbf{\Sigma}_{A_2}^k - \mathbf{\Sigma}_{A_2}^{k-1}\|_F^2 < \epsilon$, where $\epsilon > 0$ is a predefined threshold, take $\mathbf{\Sigma}_{A_1} = \mathbf{\Sigma}_{A_1}^k$, $\mathbf{\Sigma}_{A_2} = \mathbf{\Sigma}_{A_2}^k$, $\mathbf{V}_2 = \mathbf{V}_2^k$, $\mathbf{Q}_2 = \mathbf{Q}_2^k$, and exit. Otherwise, set $k = k + 1$ and go to step 2).

The convergence of the algorithm outlined above was tested with numerical experiments. It was interesting to note that for a large number of channels and thresholds, one or two iterations were enough for convergence.

IV. DISCUSSION

The derivation presented above contains certain details that are worth discussing further. For instance, it is possible to formulate a simplified but suboptimal version of the optimization problem (8). For that, the condition on the total power, would be replaced by an equal bound on the output power on both blocks of symbols $\mathbf{x}(n+1)$ and $\mathbf{x}(n)$, i.e., $\sigma_x^2 \text{tr}(\mathbf{A}_1 \mathbf{A}_1^T) \leq P/2$ and $\sigma_x^2 \text{tr}(\mathbf{A}_2' \mathbf{A}_2'^T) \leq P/2$. This formulation may be relevant during implementation because it would guarantee that the average output power is always below a specified maximum output power. In addition, an iterative procedure for obtaining $(\mathbf{A}_1, \mathbf{B}_1')$ and $(\mathbf{A}_2', \mathbf{B}_2)$ is no longer needed, and we can set the following procedure:

Algorithm 2

1. Compute \mathbf{V}_1 , $\mathbf{\Lambda}_1$ and \mathbf{Q}_1 using (13). Solve problem (18). Generate $\mathbf{\Sigma}_{A_1}$.
2. Use $\mathbf{\Sigma}_{A_1}$ in (16) for determining \mathbf{V}_2 , $\mathbf{\Lambda}_2$ and \mathbf{Q}_2 .

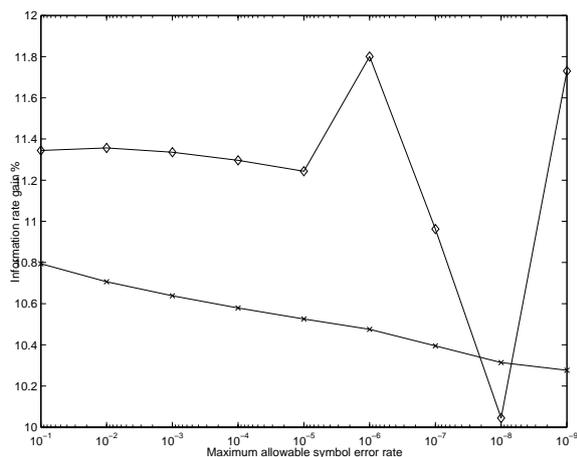


Figure 3: Information rate gain for the proposed scheme over the DMT scheme and the one proposed by Scaglione *et al.* (1999a) as function of the maximum allowable symbol error rate. SNR=10 dB and $M = 16$. Legend: ◇: gain of the proposed scheme over DMT; ×: gain if the proposed over the one by Scaglione *et al.* (1999a).

3. Obtain the water filling solution for $\sigma_{A_2}^{ii}$, which is achieved by solving problem (18) for $\sigma_{A_2}^{ii}$, with E_i replaced by D_i .

On a different item, recall that after imposing structure (5), we had to obtain the pairs $(\mathbf{A}_1, \mathbf{B}'_1)$ and $(\mathbf{A}'_2, \mathbf{B}_2)$ that solve problem (8). From (12) we know that the pair $(\mathbf{A}_1, \mathbf{B}'_1)$ is determined up to the choice of an invertible matrix \mathbf{W}_1 . Similarly, we need to select \mathbf{W}_2 to define the second pair $(\mathbf{A}'_2, \mathbf{B}_2)$. It is clear that \mathbf{W}_1 and \mathbf{W}_2 can be exploited to impose additional properties to the transceivers. For example, following the ideas by Scaglione *et al.* (1999a), we can obtain Zero-Forcing (ZF) and Minimum Mean Square Error (MMSE) receivers. It can be proved that these two choices lead to a transmission over M parallel independent channels, i.e., symbol by symbol detection within a block is optimal.

Finally, notice that the results given above are valid for Gaussian signaling. However, when the symbols x_n come from a finite alphabet we can use the gap approximation (Cioffi *et al.*, 1995). The gap Γ is defined so that the number of bits to be transmitted by the i -th channel is:

$$b_i = \frac{1}{2} \log_2 \left(1 + \frac{SNR_i}{\Gamma} \right). \quad (19)$$

Here, $SNR_i = D_i$ or $SNR_i = E_i$ depending on which transceiver pair we are considering. The choice of Γ depends on the maximum allowable symbol error rate and on the particular constellation for the i -th sub-channel. For example if we assume that all sub-channels use a PAM constellation, then the margin can be chosen as:

$$\Gamma = \frac{2}{3} [\text{erfc}^{-1}(P_e^{max})]^2, \quad (20)$$

where P_e^{max} is the maximum symbol error rate permitted in any of the sub-channels (we could also set different symbol error rates and margins for each sub-channel) and $\text{erfc}^{-1}(\cdot)$ is the inverse of complementary error function. Obviously the size of the PAM constellations for each sub-channel will be different depending on SNR_i . As pointed out by Scaglione *et al.* (1999a), the gap can be reduced relying on codification on each sub-channel.

V. SIMULATION RESULTS

We have performed numerical experiments with the algorithms introduced in previous sections. In this section, we present the results obtained by comparing our scheme with the ZP with leading zeros approach presented by Scaglione *et al.* (1999a) and with the DMT scheme. We define

$$SNR = 10 \log_{10} \left(\frac{P}{2M\sigma_v^2} \right), \quad (21)$$

where P is the average output power for two blocks of symbols, M is the number of symbols per block and σ_v^2 is the noise variance. The noise is assumed to be white, i.e. $\mathbf{R}_{vv} = \sigma_v^2 \mathbf{I}_{N \times N}$ and $\mathbf{R}_{v'v'} = \sigma_v^2 \mathbf{I}_{M \times M}$. In the case of the DMT scheme and the one by Scaglione *et al.* (1999a) this definition of SNR is valid if we think $P/2$ as the average output power per block of symbols.

First we consider the following impulse response channel: $[1, -0.3, 0.5, -0.4, 0.1, -0.02, 0.3, -0.1]$. In this case $L = 7$ and we will consider for our scheme $Q = 4$. For the scheme by Scaglione *et al.* (1999a) and the DMT scheme, the redundancy included is $Q = 7$.

In Fig. 2 we show the information rates (in bits per samples) as functions of SNR. Our scheme outperforms the one presented by Scaglione *et al.* (1999a) as well as DMT for small and large block sizes ($M = 16$ and $M = 64$). However, it is clear that for larger block sizes the gain in performance is reduced. We also show the performance of the suboptimal procedure given in algorithm 2. It is clear that this suboptimal solution performs as well as the optimal one. In Fig. 3 we show the information rate gain of the scheme proposed in this paper over the DMT scheme and the one by Scaglione *et al.* (1999a) as function of the maximum allowable symbol error rate. In this case we use equation (19) for assigning the number of bits $b_i \forall i = 1, \dots, M$. In this case the channel used was $[0.1659, 0.3045, -0.1159, -0.0773, -0.0015]$ which was drawn from an ADSL environment (Lin and Phoong, 2002). Here, $L = 4$, and $Q = 2$ for the proposed scheme and $Q = L = 4$ for the DMT scheme and the one by Scaglione *et al.* (1999a). We have used algorithm 2 for deriving our transceivers. We see that the gain in information rate for the proposed scheme is about 10 – 12% over the DMT and the scheme by Scaglione *et al.* (1999a).

As pointed out by Scaglione *et al.* (1999b) a block transmission system can be thought of as a multi-

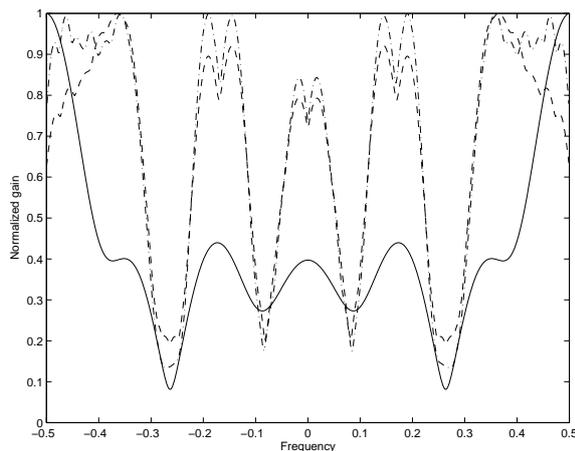


Figure 4: Average spectral magnitudes for transmit filters. SNR=10 dB, $M = 16$ and $P_e^{max} = 10^{-6}$. Legend: Dashed-dotted: Average spectral magnitudes for \mathbf{A}_1 ; Dotted: Average spectral magnitudes for \mathbf{A}_2 ; Solid: Channel transfer function magnitude $|H(f)|$.

rate filterbank system. Bearing this idea in mind we can think \mathbf{A}_1 and \mathbf{A}_2 as being banks of M filters, where the impulse responses of each of these filters are the columns of the matrices \mathbf{A}_1 and \mathbf{A}_2 . Following Scaglione *et al.* (1999b), the average spectral magnitude of the transmit filter is defined as:

$$|\mathcal{A}(f)| = \frac{1}{M} \sum_{i=0}^M |A_i(f)|, \quad (22)$$

where $A_i(f)$ is the Fourier transform of the i -th column of matrix \mathbf{A} . This average spectral magnitude gives an idea on how the transmit power is distributed in frequency.

In Fig. 4 we show the average spectral transmit magnitudes for transmit filters in \mathbf{A}_1 and \mathbf{A}_2 for the channel $[1, -0.3, 0.5, -0.4, 0.1, -0.02, 0.3, -0.1]$, SNR=10 dB, $M = 16$ and $P_e^{max} = 10^{-6}$. The other parameters are as in Fig. 2. For deriving \mathbf{A}_1 and \mathbf{A}_2 we have used the suboptimal procedure, i.e. equal average power in both successive blocks of symbols. We can see that the power distribution is not the same for \mathbf{A}_1 and \mathbf{A}_2 . This is due to the fact that the symbols passed through \mathbf{A}_2 have to deal with the noise $\mathbf{v}(n)$ on the channel plus a source of interference due to IBI from the last block of symbols spectrally shaped with \mathbf{A}_1 . In contrast, the symbols passed through \mathbf{A}_1 have to deal only with the noise $\mathbf{v}'(n+1)$.

VI. CONCLUSIONS

Assuming block transmissions for an ISI channel we have presented a structure of transceivers that operates with reduces redundancy. This structure consists of two different transceivers that act in an alternating way, i.e. one after the other. For further improvement of the spectral efficiency we have optimized the transmitters of both transceivers to achieve the maximum

information rate. We have shown that the proposed scheme outperforms other approaches presented in the literature.

ACKNOWLEDGEMENTS

This work was funded by grants PIP-6344, FONCyT PAV-000127, and UBACyT I005

REFERENCES

- Chow, J., J. Tu and J. Cioffi, "A discrete multi-tone transceiver system for HDSL applications," *IEEE J. Select. Areas Communications*, **9**, 895-908 (1991).
- Cioffi, J., G. P. Dudevoir, M. V. Eyuboglu and G. D. Froney, Jr., "MMSE decision-feedback equalizers and coding-Parts I and II," *IEEE Trans. on Communications*, **43**, 2582-2604 (1995).
- Cover, T.M and J. A. Thomas, *Elements of Information Theory*, New York, Wiley (1991).
- Lechleider, J. W., "The optimum combination of block codes and receivers for arbitrary channels," *IEEE Trans. on Communications*, **38**, 615-621 (1990).
- Lin, Y. and S. Phoong, "Perfect discrete multi-tone modulation with optimal transceivers," *IEEE Trans. on Signal Processing*, **48**, 1702-1711 (2000).
- Lin, Y. and S. Phoong, "Optimal ISI-free DMT transceivers for distorted channels with colored noise," *IEEE Trans. on Signal Processing*, **49**, 2702-2712 (2001).
- Lin, Y. and S. Phoong, "Minimum redundancy for ISI free FIR filterbank transceivers," *IEEE Trans. on Signal Processing*, **50**, 842-853 (2002).
- Scaglione, A., S. Barbarossa and G. B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. on Information Theory*, **45**, 1019-1032 (1999a).
- Scaglione, A., G. B. Giannakis and S. Barbarossa, "Redundant filterbank precoders and equalizers-Parts I and II," *IEEE Trans. on Signal Processing*, **47**, 1988-2022 (1999b).

Received: October 18, 2007.

Accepted: July 2, 2008.

Recommended by Guest Editors D. Alonso, J. Figueroa, E. Paolini and J. Solsona.