Channel Identification for Block Transmission*

Olivier Rousseaux[1], Geert Leus[1†], Petre Stoica[2] and Marc Moonen[1]
[1] K.U.Leuven - ESAT
Kasteelpark Arenberg 10, 3001 Leuven, Belgium
Email: {olivier.rousseaux, geert.leus, marc.moonen}@esat.kuleuven.ac.be
[2] Uppsala University - Department of Systems and Control
P.O. Box 337 SE-751 05, Uppsala Sweden
Email: ps@syscon.uu.se

***Abstract*— In this paper, we first review popular schemes for block transmission within the framework of affine precoding. We then briefly present existing channel identification techniques that only rely on the redundant symbols introduced by these block transmission schemes. We then present the major contribution of this paper, which is a new stochastic channel identification method suited for Non-Zero Padded (NZP) block transmission, a special case of Known Symbol Padded (KSP) block transmission. This new stochastic method is iterative and takes into account both the knowledge of the padded symbols and the expected structure of the unknown data symbols. We compare this new method with existing channel identification methods for block transmission. Experimental results show that the new method performs better than the existing ones and has a lower complexity.**

I. Introduction

The market need for high data rate communications has driven the research in telecommunication during the last years. In order to increase data rates when transmitting data over wireless channels, it is often needed to use broadband communication systems. A major impediment of such systems is that the sampling period can get smaller than the delay spread of the channel, especially in multipath scenarios. This results in ISI (Inter Symbol Interference, interference between two successive symbols of information), a phenomenon that needs to be combatted at the receiver. This is usually done using serial equalization techniques. Recently, a family of transmission techniques, known as block transmission techniques [1], [2], [3], [4] have been proposed in order to address the problem of ISI in a cheap and efficient way. A wide variety of such transmission and equalization schemes have been proposed and analyzed in the last years. All these block transmission schemes introduce redundant symbols at the transmitter in the form of a Cyclic Prefix (CP) or a padded sequence of known symbols (KSP, Known Symbol Padding) in order to avoid Inter Block Interference (IBI) and allow cheap equalization. While channel identification based on training sequences or pilot symbols in this context has attracted much attention in the past, the specific problem of identifying the channel using only the redundant symbols that are introduced in block transmission schemes to avoid IBI is the problem we focus on in this paper. We first review popular schemes for block transmission within the framework of affine precoding [5]. We then briefly present existing channel identification techniques that only rely on the redundant symbols introduced by these block transmission schemes. We next present the major contribution of this paper, which is a new stochastic channel identification method suited for Non-Zero Padded (NZP) block transmission, a special case of Known Symbol Padded (KSP) block transmission. This new stochastic method is iterative and takes into account both the knowledge of the padded symbols and the expected structure of the unknown data symbols. We compare this new method with existing channel identification methods for KSP block transmission and finally draw some conclusions.

Notation: We use upper (lower) case bold face letters to denote matrices (column vectors). $\mathbf{I}_N$ is the identity matrix of size N and $\mathbf{O}_{M\times N}$ is an all-zeros matrix of size $M \times N$. The operator * denotes the complex conjugate, ^{T} the transpose, ^{H} the complex conjugate transpose and † the pseudo-inverse of a matrix.

II. General Data Model for Block Transmission

We consider stationary Finite Impulse Response (FIR) convolutive channels of order L: $\mathbf{h} = [h[0], h[1], \cdots, h[L]]^T$. A sequence $x[n]$ of symbols is transmitted over the channel. The received sequence $y[n]$ is the linear convolution of the transmitted sequence with the channel impulse response:

$$y[n] = \sum_{i=0}^{L} h[i]x[n-i] + \eta[n], \tag{1}$$

where $\eta[n]$ is the Additive White Gaussian Noise (AWGN) at the receiver. When a sequence $s[n]$ of data symbols has to be transmitted, a straightforward solution is to transmit these symbols directly over the channel.

An alternative to this straightforward approach is known as block transmission [1], where data are organized in blocks and fed to a precoding matrix before transmission. Training symbols can be mixed with data symbols, which is equivalent to affine precoding [5].

* This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the Belgian State, Prime Minister's Office - Federal Office for Scientific, Technical and Cultural Affairs - Interuniversity Poles of Attraction Programme (2002-2007) - IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modeling') and P5/11 ('Mobile multimedia communication systems and networks'), the Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology) of the Flemish Government, Research Project FWO nr.G.0196.02 ('Design of efficient communication techniques for wireless time-dispersive multi-user MIMO systems') and was partially sponsored by IMEC (Flemish Interuniversity Microelectronics Center). The scientific responsibility is assumed by its authors.

† Postdoctoral Fellow of the F.W.O. Vlaanderen.

A block $\mathbf{s}_k(n_s \times 1)$ of data symbols is defined as a column vector of size n_s, $k = 1, \cdots, K$ being the block index: $\mathbf{s}_k = [s[(k-1)n_s + 1], \cdots, s[kn_s]]^T$. The kth transmitted block, $\mathbf{x}_k(n_x \times 1)$ is computed as follows:

$$\mathbf{x}_k = \mathbf{P}_s \mathbf{s}_k + \mathbf{P}_t \mathbf{t}_k, \tag{2}$$

where $\mathbf{t}_k = [t_k[1], \cdots, t_k[n_t]]^T$ is a block of training symbols, $\mathbf{P}_s(n_x \times n_s)$ is the precoding matrix for data symbols and $\mathbf{P}_t(n_x \times n_t)$ is the precoding matrix for training symbols. The transmitted sequence is obtained by parallel-to-serial conversion of $\mathbf{x}_k$: $\mathbf{x}_k^T = [x[n_x(k-1)+1], \cdots, x[n_x k]]$ and the received sequence $y[n]$ is derived from (1). This received sequence can be organized in blocks $\mathbf{y}_k(n_x \times 1)$ of received symbols corresponding to the blocks of transmitted symbols: $\mathbf{y}_k = [y[(k-1)n_x + 1], \cdots, y[kn_x]]^T$. Using these definitions, the transmission scheme can be expressed on a block level: the kth received block contains a contribution from the kth transmitted block plus interference from the $k-1^{th}$ transmitted block and some noise:

$$\mathbf{y}_k = \mathbf{H}_0 \mathbf{x}_k + \mathbf{H}_1 \mathbf{x}_{k-1} + \boldsymbol{\eta}_k, \tag{3}$$

where $\boldsymbol{\eta}_k$ is the AWGN vector, $\mathbf{H}_1(n_x \times n_x)$ is an upper triangular Toeplitz matrix with $[0, \cdots, 0, h[L], \cdots, h[1]]$ as first row and $\mathbf{H}_0(n_x \times n_x)$ is a lower triangular Toeplitz matrix with $[h[0], \cdots, h[L], 0, \cdots, 0]^T$ as first column.

Equalization or symbol detection will usually be done after a pre-processing step on the received block, which yields the vector $\mathbf{z}_k(n_z \times 1) : \mathbf{z}_k = \mathbf{R}\mathbf{y}_k$, where $\mathbf{R}(n_z \times n_x)$ is the pre-processing matrix at the receiver. The general scheme for block transmission can then be described by the following relation:

$$\begin{aligned}\mathbf{z}_k = {} & \mathbf{R}\mathbf{H}_0\mathbf{P}_s\mathbf{s}_k + \mathbf{R}\mathbf{H}_1\mathbf{P}_s\mathbf{s}_{k-1} \\ & + \mathbf{R}\mathbf{H}_0\mathbf{P}_t\mathbf{t}_k + \mathbf{R}\mathbf{H}_1\mathbf{P}_t\mathbf{t}_{k-1} + \mathbf{R}\boldsymbol{\eta}_k.\end{aligned} \tag{4}$$

Traditional block transmission techniques aim at avoiding Inter Block Interference (IBI, interference between successive blocks of data). Several approaches are possible towards this goal, all of which can be expressed within this general framework. In this paper, we consider two families of block transmission techniques that are used in practice: CP based transmission and KSP based transmission, which we detail next.

A. Cyclic Prefix (CP) Transmission

CP techniques avoid IBI by adding to the transmitted data blocks a cyclic prefix whose length μ is at least equal to the channel order L. This cyclic prefix simply copies the last symbols of the block to its beginning before transmission. The CP collects all the interference from the data symbols of the previous block and is simply discarded at the receiver. Appending the CP translates into precoding matrices of the form:

$$\mathbf{P}_s = \mathbf{P}_\mu \mathbf{P}'_s, \qquad \mathbf{P}_\mu = \left[\begin{array}{c|c} \mathbf{O}_{\mu \times (n_x - 2\mu)} & \mathbf{I}_\mu \\ \hline \multicolumn{2}{c}{\mathbf{I}_{n_x - \mu}} \end{array} \right], \tag{5}$$

where $\mathbf{P_s}'$ is a general precoding matrix of size $(n_x - \mu) \times n_s$. Discarding the CP at the receiver translates into pre-processing matrices of the form:

$$\mathbf{R} = \mathbf{R}'\mathbf{R}_\mu, \qquad \mathbf{R}_\mu = \left[\begin{array}{c|c} \mathbf{O}_{(n_x - \mu) \times \mu} & \mathbf{I}_{n_x - \mu} \end{array} \right]. \tag{6}$$

where $\mathbf{R}'$ is a general pre-processing matrix.

Using precoding and pre-processing matrices of this kind, it is easy to check that, when $\mu \geqslant L$, CP transmission yields two remarkable properties:

- The channel matrix effectively becomes a circulant matrix of size $(n_x - \mu) \times (n_x - \mu)$: $\mathbf{R}_\mu \mathbf{H}_0 \mathbf{P}_\mu = \mathbf{H}_{circ}$, $\mathbf{H}_{circ}$ being a Toeplitz matrix with $[h[0], \cdots, h[L], 0, \cdots, 0]^T$ on the first column and $[h[0], 0, \cdots, 0, h[L], \cdots, h[1]]$ on the first row.
- The IBI is cancelled: $\mathbf{R}_\mu \mathbf{H}_1 \mathbf{P}_\mu = \mathbf{O}_{(n_x - \mu) \times (n_x - \mu)}$.

Ignoring the training symbol terms, (4) can then be rewritten as:

$$\mathbf{z}_k = \mathbf{R}'\mathbf{H}_{circ}\mathbf{P_s}'\mathbf{s}_k + \mathbf{R}'\mathbf{R}_\mu\boldsymbol{\eta}_k. \tag{7}$$

Two popular block transmission systems fall under this category of CP systems:**CP-Only** and **OFDM** systems [1], [2], [3], [4]. None of them mixes training symbols with the transmitted signals $(n_t = 0)$.

1) Cyclic Prefix Only (CP-Only): CP-Only [1], [3] is a straightforward example of a CP transmission system:it simply adds a CP of length μ to the transmitted data symbol block of size n_s without any extra encoding ($\mathbf{P}'_s = \mathbf{I}_{n_s}$). The CP is discarded at the receiver without extra processing ($\mathbf{R}' = \mathbf{I}_{n_s}$), which results in the following transmission scheme: $\mathbf{z}_k = \mathbf{H}_{circ}\mathbf{s}_k + \mathbf{n}_k$.

A nice property of circulant matrices is that they are diagonalized by the use of DFT and IDFT operations. The channel matrix being circulant in CP-Only systems, this property can be used for implementing efficient and cheap equalization systems.

2) Orthogonal Frequency Division Modulation (OFDM): Also known as **DMT** (Discrete Multi Tone), these systems were the first block transmission techniques used in practical systems [1], [2]. They directly exploit the circulant structure of the channel matrix of CP systems. In OFDM, an IDFT (implemented as an IFFT) is performed on data blocks of size n_s before the CP is added. The receiver throws the CP away and performs a DFT (implemented as an FFT) on the received block.

OFDM system can thus be described as follows: $\mathbf{z}_k = \boldsymbol{\mathcal{F}}_{n_s}\mathbf{H}_{circ}\boldsymbol{\mathcal{I}}_{n_s}\mathbf{s}_k + \boldsymbol{\mathcal{F}}_{n_s}\boldsymbol{\eta}_k$, where $\boldsymbol{\mathcal{I}}_{n_s}$, the n_s-point IDFT matrix and $\boldsymbol{\mathcal{F}}_{n_s}$, the n_s-point DFT are respectively the pre-coding and pre-processing matrices $\mathbf{P}'_s$ and $\mathbf{R}'$.

Exploiting the diagonalization property of circulant matrices, this system can be further simplified into: $\mathbf{z}_k = \mathbf{H}_{diag}\mathbf{s}_k + \boldsymbol{\mathcal{F}}_P\mathbf{n}_k$, where $\mathbf{H}_{diag}$ is a diagonal matrix with the n_s-point DFT of the channel impulse response $\mathbf{h}$ on the main diagonal.

B. Known Symbol Padding (KSP) Transmission

In KSP transmission [1], [4], IBI is avoided by padding known symbols to each data block. The kth transmitted block is made of n_s data symbols followed by n_t known symbols: $\mathbf{x}_k = [\mathbf{s}_k^T, \mathbf{t}_k^T]^T$. This translates into pre-coding matrices of the form $\mathbf{P}_s = [\mathbf{I}_{n_s}, \mathbf{O}_{n_t \times n_s}]^T$ and $\mathbf{P}_t = [\mathbf{O}_{n_s \times n_t}, \mathbf{I}_{n_t}]^T$.

No pre-processing is done at the receiver, which translates into an identity pre-processing matrix: $\mathbf{R} = \mathbf{I}_{n_x}$. It is easy

to check that $\mathbf{H}_1\mathbf{P}_s = \mathbf{O}_{(nx\times n_s)}$, which means that there is no IBI. Using these precoding and pre-processing matrices, the total transmission scheme (4) translates into:

$$\mathbf{z}_k = \mathbf{H}_{KSP}\mathbf{s}_k + \mathbf{H}_{t0}\mathbf{t}_k + \mathbf{H}_{t1}\mathbf{t}_{k-1} + \boldsymbol{\eta}_k, \tag{8}$$

where $\mathbf{H}_{KSP} = \mathbf{H}_0\mathbf{P}_s$ is a full rank tall Toeplitz matrix, $\mathbf{H}_{t0} = \mathbf{H}_0\mathbf{P}_t$, and $\mathbf{H}_{t1} = \mathbf{H}_1\mathbf{P}_t$.

Note that when the same training sequence is used for all data blocks, i.e. $\mathbf{t}_k = \mathbf{t}\ \ \forall k$, the system can be described as: $\mathbf{z}_k = \left[\ \mathbf{H}\ \middle|\ \mathbf{H}_{t0} + \mathbf{H}_{t1}\ \right] \left[\frac{\mathbf{s}_k}{\mathbf{t}}\right] + \boldsymbol{\eta}_k = \mathbf{H}_{circ}\mathbf{x}_k + \boldsymbol{\eta}_k$.

Two KSP methods are used in practice: The first one that was proposed is called **Zero Padding (ZP)**. In this case, a sequence of zeros is appended at the end of every block ($\mathbf{t}_k = \mathbf{0}$). The alternative to ZP is known as **Non-Zero Padding (NZP)**. NZP corresponds to the case where the known symbols differ from zero ($\mathbf{t}_k \neq \mathbf{0}$).

III. Channel Identification

Block transmission techniques were presented here in the general framework of affine precoding. Many channel identification techniques have been proposed in this affine precoding framework that aim at identifying the channel on a block per block basis. These techniques always require the presence of $2L+1$ training symbols per transmitted block and are best suited for highly time varying channels. Other techniques have been developed recently that only rely on the redundant symbols inserted by the block transmission scheme to avoid IBI and consider stationary channels for several blocks of data. We briefly present here recent techniques that are suited for the specific cases of CP and KSP transmission [6], [7], [8], [9]. We finally propose and describe in more details a novel stochastic channel identification method that is suited for NZP transmission, a special case of KSP.

A. Existing Channel Identification Methods for Block Transmission

1) Channel Identification in CP Transmission: In CP transmission systems, if we are to exclude the use of explicit training sequences, the only property that can be used towards channel identification is the CP-induced cyclostationary nature of the transmitted data.

A first method was proposed in [8] directly exploits this property. That method relies on the fact that the cyclostationarity is also present in the received signal fore this property is not affected by linear filtering (i.e. the transmission channel). This method allows to identify the channel even when it is longer than the cyclic prefix.

A second method, which was proposed in [9], directly exploits the fact that a part of the transmitted sequence is repeated in the cyclic prefix. Subtracting the end of a block from its beginning will thus yield zeros in the data model that can be used to identify the channel using subspace decomposition based algorithms. Note that this method only works when the true channel order is shorter than the cyclic prefix length.

2) Channel Identification in KSP Transmission: In KSP transmission, while the percentage of redundant symbols remains roughly the same and the cyclostationarity of transmitted symbols is still present if the padded sequence stays unchanged from block to block, the knowledge of the transmitted sequence can be further exploited at the receiver to perform more efficient channel estimation.

ZP Transmission

A first method was proposed in. [7] for the special case of ZP transmission where the padded sequence is set to zero. In this case, the presence of the all-zero padded sequence in the data model allows the blind estimation of the channel using a subspace-based algorithm. This method is a deterministic one and allows to identify the channel up to a complex scaling factor.

NZP Transmission

When the padded sequence is not set to zero, the knowledge of this sequence can be exploited at the receiver to perform a better estimate of the channel. In [6], a deterministic method inspired by [7] is proposed that is suited for the case where the padded sequence is non-zero. It combines the algorithm of [7] together with an additional constraint that comes from the knowledge of the padded symbols. Note that for this method to work it is mandatory that the padded sequence remains the same from block to block.

B. A New Stochastic Channel Identification Method for NZP

In order to identify the channel, we focus on the received samples that contain a contribution from the known symbols. Define $\mathbf{u}_k$ as the vector that contains a contribution from the known symbols in the kth transmitted block:

$$\mathbf{u}_k = \left[\mathbf{z}_k(n_s+1), \cdots, \mathbf{z}_k(n_s+n_t), \mathbf{z}_{k+1}(1), \cdots, \mathbf{z}_{k+1}(L)\right]^T.$$

The vector $\mathbf{u}_k$ contains a contribution from the known symbols in the kth transmitted block/ plus an additional term that collects the contributions from both the unknown surrounding data symbols and the noise. Rephrasing the specific transmission scheme of KSP transmission (8), $\mathbf{u}_k$ can be described as follows:

$$\mathbf{u}_k = \mathbf{T}_k\mathbf{h} + \boldsymbol{\epsilon}_k, \tag{9}$$

where $\mathbf{T}_k$ is an $(n_t+L)\times(L+1)$ tall Toeplitz matrix with $\left[\mathbf{t}_k^T, 0, \cdots, 0\right]$ as first column and

$$\boldsymbol{\epsilon}_k = \underbrace{\begin{bmatrix} h_L & \cdots & h_1 & 0 & \cdots & 0 \\ 0 & \ddots & \vdots & \vdots & & \\ & \ddots & h_L & 0 & & \vdots \\ \vdots & & 0 & h_0 & \ddots & \\ & & \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & h_{L-1} & \cdots & h_0 \end{bmatrix}}_{\mathbf{H}_s(n_t+L)\times(2L)} \mathbf{s}'_k + \boldsymbol{\eta}'_k, \tag{10}$$

where $\mathbf{s}'_k = \left[s_k[n_s-L+1], \cdots, s_k[n_s], s_{k+1}[1], \cdots, s_{k+1}[L-1]\right]^T$ and $\boldsymbol{\eta}'_k$ is the corresponding AWGN term.

We can collect all these equations in a single relation. Define $\mathbf{T}$ as the collection of all the $\mathbf{T}_k$ matrices: $\mathbf{T} = \left[\mathbf{T}_1^H, \cdots, \mathbf{T}_K^H\right]^H$, $\mathbf{u}$ as the collection of all vectors $\mathbf{u}_k$: $\mathbf{u} = \left[\mathbf{u}_1^H, \cdots, \mathbf{u}_K^H\right]^H$, and $\boldsymbol{\epsilon}$ is similarly defined as the concatenation of all $\boldsymbol{\epsilon}_k$ terms. Collecting the K equations from (9) then yields:

$$\mathbf{u} = \mathbf{T}\mathbf{h} + \boldsymbol{\epsilon}. \tag{11}$$

Assuming that both the noise and the data are white and zero-mean ($E\{s[n]s^*[l]\} = E\{\eta[n]\eta^*[l]\} = 0$ for $n \neq l$ and $E\{s[n]\} = E\{\eta[n]\} = 0$) and defining the signal and noise variances as $\lambda^2 = E\{s[n]s^*[n]\}$ and $\sigma^2 = E\{\eta[n]\eta^*[n]\}$ respectively, the statistics of the term $\boldsymbol{\epsilon}_k$ are the following:

- **C1**: $E\{\boldsymbol{\epsilon}_k\} = 0$
- **C2**: $E\left\{\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^H\right\} \triangleq \mathbf{Q} = \lambda^2\mathbf{H}_s\mathbf{H}_s^H + \sigma^2\mathbf{I}$
- **C3**: $E\left\{\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_l^H\right\} = \mathbf{0}_{(n_t+L)\times(n_t+L)}$

Note that **C3** only holds if $E\left\{\mathbf{s}'_k\mathbf{s}_l^{'H}\right\} = \mathbf{0}$ for $n \neq l$, which is only the case if there is a sufficiently large number of transmitted data symbols between the blocks of known symbols, i.e. **C3 is true when** $n_s \geqslant 2L$, a relatively weak condition which we assume to be true from now on. Based on these statistics, we can now propose a new stochastic method to estimate the transmission channel $\mathbf{h}$ from the received symbols $\mathbf{u}_k$.

Step 1

We start from a LS estimator of the channel impulse response $\mathbf{h}$ for the kth block:

$$\hat{\mathbf{h}}_k^{LS} = \left(\mathbf{T}_k^H\mathbf{T}_k\right)^{-1}\mathbf{T}_k^H\mathbf{u}_k. \tag{12}$$

Due to **C1** this estimator is unbiased $\left(E\left\{\hat{\mathbf{h}}_k\right\} = \mathbf{h}\right)$. Note that the inverse in (12) always exists since $\mathbf{T}_k$ is always full column rank, independently of the length and the composition of the padded sequence. We can build a similar LS estimator that jointly uses all the transmitted padded sequences in order to estimate the channel as

$$\hat{\mathbf{h}}_{LS} = \left(\mathbf{T}^H\mathbf{T}\right)^{-1}\mathbf{T}^H\mathbf{u} = \left(\sum_{k=1}^{K}\mathbf{T}_k^H\mathbf{T}_k\right)^{-1}\sum_{k=1}^{K}\mathbf{T}_k^H\mathbf{u}_k. \tag{13}$$

Step 2

We know from **C2** that the interference plus noise term $\boldsymbol{\epsilon}_k$ present in the LS estimate proposed in the previous paragraph is not white. The accuracy of the estimate can thus be improved if we apply a weighted least squares method that first whitens the noise term. Assume for now that the auto-correlation $\mathbf{Q}$ of the noise term is known at the receiver. Define the whitened vector of received symbols $\tilde{\mathbf{u}}_k = \mathbf{Q}^{-1/2}\mathbf{u}_k$. Adapting (9) correspondingly, we get:

$$\tilde{\mathbf{u}}_k = \underbrace{\mathbf{Q}^{-1/2}\mathbf{T}_k}_{\tilde{\mathbf{T}}_k}\mathbf{h} + \underbrace{\mathbf{Q}^{-1/2}\epsilon_k}_{\tilde{\boldsymbol{\epsilon}}_k}. \tag{14}$$

It is easy to verify that $E\{\tilde{\boldsymbol{\epsilon}_k}\} = \mathbf{0}$ and that $E\left\{\tilde{\boldsymbol{\epsilon}}_k\tilde{\boldsymbol{\epsilon}}_k^H\right\} = \mathbf{I}$. This shows that the interference plus noise term in this alternative formulation of the problem is white. Adapting (12) by replacing $\mathbf{T}_k$ with $\tilde{\mathbf{T}}_k$ and $\mathbf{u}_k$ with $\tilde{\mathbf{u}}_k$ will thus produce a weighted least squares estimate of the channel which is more accurate than the initial one. This estimate can be expressed directly as a function of the original system variables:

$$\hat{\mathbf{h}}_k^{WLS} = \left(\tilde{\mathbf{T}}_k^H\tilde{\mathbf{T}}_k\right)^{-1}\tilde{\mathbf{T}}_k^H\tilde{\mathbf{u}}_k = \left(\mathbf{T}_k^H\mathbf{Q}^{-1}\mathbf{T}_k\right)^{-1}\mathbf{T}_k^H\mathbf{Q}^{-1}\mathbf{u}_k \tag{15}$$

Here again, this approach can be extended to the larger data model (11) to build a weighted LS estimator of the channel that jointly uses all the padded sequences: $\hat{\mathbf{h}}_{WLS} = \left(\mathbf{T}^H\mathbf{Q}_{tot}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^H\mathbf{Q}_{tot}^{-1}\mathbf{u}$, where $\mathbf{Q}_{tot} = E\left\{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^H\right\}$ is the autocorrelation matrix of the total noise term $\boldsymbol{\epsilon}$. Exploiting **C2** and **C3** allows us to say that $\mathbf{Q}_{tot}$ is a block-diagonal matrix with the autocorrelation matrix $\mathbf{Q}$ repeated over the main diagonal. This block-diagonal structure allows us to split the product of large matrices in a sum of products of smaller matrices. We can thus re-express the weighted LS channel estimate in a way that is computationally less demanding:

$$\hat{\mathbf{h}}_{WLS} = \left(\sum_{k=1}^{K}\mathbf{T}_k^H\mathbf{Q}^{-1}\mathbf{T}_k\right)^{-1}\sum_{k=1}^{K}\mathbf{T}_k^H\mathbf{Q}^{-1}\mathbf{u}_k \tag{16}$$

Note that $\mathbf{Q}^{-1}$ always exists and is always full rank thanks to the noise term in $\mathbf{Q}$. Because $\mathbf{Q}$ is an auto-correlation matrix, it is positive semi-definite. Since we know it is full rank, we can say it is strictly positive definite. If we additionally remember that $\mathbf{T}_k$ is always full column rank, we can say that all the terms in the summation of the left side of (16) are strictly positive definite as well. Taking into account that the sum of strictly positive definite matrices is a strictly positive definite matrix proves that the inverse in (16) always exists.

In order to produce this weighted least squares channel estimate, an estimate of the noise autocorrelation matrix $\mathbf{Q}$ is needed. The channel estimation method proposed in Step 1 can be used to produce an initial channel estimate $\hat{\mathbf{h}}^{(0)}$. Using (9) together with this channel estimate allows us to estimate the noise term $\boldsymbol{\epsilon}_k$:

$$\hat{\boldsymbol{\epsilon}}_k = \mathbf{u}_k - \mathbf{T}_k\hat{\mathbf{h}}^{(0)} \tag{17}$$

This noise estimate can be used to produce an estimate of the noise autocorrelation: $\hat{\mathbf{Q}}_k = \hat{\boldsymbol{\epsilon}}_k\hat{\boldsymbol{\epsilon}}_k^H$. We can then average over a large number of such $\hat{\mathbf{Q}}_k$ estimates to better estimate $\mathbf{Q}$:

$$\hat{\mathbf{Q}} = K^{-1}\sum_{k=1}^{K}\hat{\boldsymbol{\epsilon}}_k\hat{\boldsymbol{\epsilon}}_k^H \tag{18}$$

This noise autocorrelation matrix can be fed into (16) in order to produce a refined estimate $\hat{\mathbf{h}}^{(1)}$ of the channel.

Step 3

$\hat{\mathbf{h}}^{(1)}$ can be used to produce a better estimate $\hat{\mathbf{Q}}$ that is in turn used to produce a better channel estimate. We proceed for a few iterations to gradually refine our channel estimate.

When the final channel estimate $\hat{\mathbf{h}}$ is available, we can estimate the noise power. We see from **C2** that subtracting $\lambda^2\mathbf{H}_s\mathbf{H}_s^H$ from $\mathbf{Q}$ leaves us with a diagonal matrix with the

noise power σ^2 on the diagonal. We can reconstruct $\hat{\mathbf{H}}_s$ using the available channel estimate $\hat{\mathbf{h}}$. Additionally, we can use (17) and (18) with the final channel estimate $\hat{\mathbf{h}}$ to approximate $\mathbf{Q}$ and, assuming that λ^2 is known, produce an estimate $\hat{\sigma}^2$ of the noise power:

$$\hat{\sigma}^2 = (n_t + L)^{-1} \, tr\left(\hat{\mathbf{Q}} - \hat{\mathbf{H}}_s \hat{\mathbf{H}}_s^H\right), \tag{19}$$

where tr denotes the trace of a matrix. Note that the initial channel estimate proposed in Step 1 is equivalent to the general method of Step 2 initialized with $\mathbf{Q} = \mathbf{I}$.

A first advantage of this iterative method is its low computational cost. Another advantage is the guaranteed existence of a solution independently of the length of the padded sequence and the channel order. This method can thus safely be applied even when the true channel order is larger than the length of the padded sequence. This method could thus as well be considered in the case of PSAM transmission (Pilot Symbol Assisted Modulation), where single equispaced pilot symbols are mixed with the data symbols in order to allow optimal channel identification in a time-varying context. Finally note that when our method is used, we can freely choose either to repeat the same padded sequence $\mathbf{t}_k$ from block to block or to change it for every new block, the latter allowing a better channel identification because of a higher diversity in the transmitted known symbols.

C. Experimental Results

In this section, we analyze the performance of the proposed stochastic algorithm and compare it with existing channel identification methods for KSP block transmission. The channel identification methods suited for CP block transmission are not investigated in this section fore they perform worse than their KSP counterparts that can rely on a richer knowledge of the transmitted signal structure in order to identify the channel.

Fig 1 compares the NMSE of different methods as a function of the SNR. The experiments are done over a set of 250 different 6th order random Rayleigh-fading channels. The channel is considered to stay constant for a burstlength (total number of transmitted blocks) K of 128, the transmitted block size $n_x = n_s + n_t$ is set to 64. The length of the padded sequences is set to the channel order ($n_t = 6$). In Fig 1, we compare the NMSE of the channel as a function of the SNR expressed in dB for the proposed iterative method, the proposed initial channel estimate and other existing methods, namely the ZP method proposed in [7] and the NZP method proposed in [6]. The comparisons are done using the same channel, noise and data realizations. In the comparison the padded sequence was changed after every block for the new stochastic method, whereas it was kept constant for the method presented in [6] (which is mandatory in order to have this method work). For the ZP method we perform optimal scalar ambiguity removal, which is not feasible in practice and thus obtain an upper bound on its performance. We observe that the performance achieved by the iterative method stay constant after the first iteration, which means that the iterative procedure converges in just one step.

The experiments clearly show a saturation effect for the initial channel estimate. This saturation occurs when the unknown data symbol terms become dominant in $\boldsymbol{\epsilon}_k$, in which case the reduction of the noise has no impact on the quality of the estimate. The saturation disappears when the iterative method that takes the structure of the unknown data symbols into account is used. We even see that the iterative method performs better than the existing ones, especially at low SNR. The ZP method seems to perform slightly better at very high SNR, but we have to keep in mind that its performance is artificially boosted by the use of optimal scalar ambiguity removal.

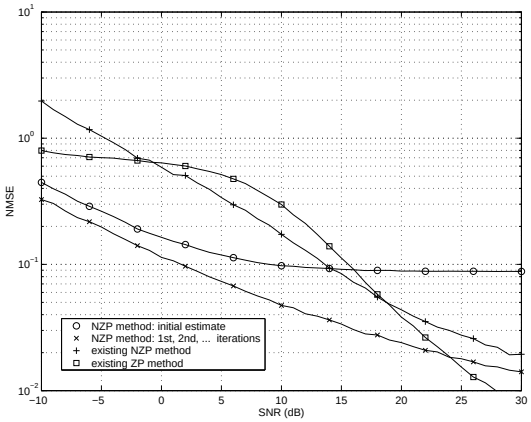

Fig. 1. NMSE vs. SNR for different Channel Estimation Techniques

IV. Conclusion

In this paper, we have presented a new channel identification method that is suited for NZP block transmission. The experiments show that this method yields more accurate results than the existing ones and this at a lower computational cost. The theoretical developments show that the method is flexible (freedom of choosing the length and composition of the padded sequence regardless of the channel) and robust (guaranteed existence of the solution).

References

[1] Z. Wang and G.B. Giannakis, "Wireless Multicarrier Communications: Where Fourier meets Shannon," *IEEE Signal Processing Magazine*, pp. 29–48, May 2000.

[2] J.A.C. Bingham, "Multicarrier Modulation for Data Transmission: An Idea whose Time has Come," *IEEE Communications Magazine*, pp. 5–14, May 1990.

[3] H. Sari, G. Karam and I. Jeanclaude, "Transmission Techniques for Digital Terrestrial TV Broadcasting," *IEEE Communicatons Magazine*, pp. 100–109, Feb. 1995.

[4] L. Deneire, B. Gyselinckx and M. Engels., "Training Sequence vs. Cyclic Prefi x: A New Look on Single Carrier Communication," in *Proc. GLOBECOM*, San Fransisco, California, November/December 2000.

[5] J. H. Manton and Y. Hua, "Affi ne Precoders for Reliable Communications," in *Proc. of ICASSP*, Istanbul, Turkey, June 2000.

[6] G. Leus and M. Moonen, "Semi-Blind Channel Estimation for Block Transmission with Non-Zero Padding," in *Proc. of the Asilomar Conference on Signals, Systems and Computers*, Pacifi c Grove, California, Nov. 4-7 2001.

[7] S. Barbarossa, A. Scaglione and G. B. Giannakis, "Performance Analysis of a Deterministic Channel Estimator for Block Transmission Systems With Null Guard Intervals," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 684–695, March 2002.

[8] R. W. Heath Jr and G. B. Giannakis, "Exploiting Input Cyclostationarity for Blind Channel Identifi cation in OFDM Systems," *IEEE Transactions on Signal Processing*, vol. 47, no. 3, pp. 848–856, March 1999.

[9] U. Tureli and H. Liu, "Blind Carrier Synchronisation and Channel Identifi cation for OFDM Communications," in *Proc. of the ICASSP Conference*, Seattle, Washington, 1998, vol. 6, pp. 3509–3512.