AN ITERATIVE METHOD FOR IMPROVED TRAINING-BASED ESTIMATION OF DOUBLY SELECTIVE CHANNELS†

Olivier Rousseaux[1] *and Geert Leus*[2]

[1] K.U.Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium
Email: {olivier.rousseaux,marc.moonen}@esat.kuleuven.ac.be

[2] T.U.Delft, Mekelweg 4, 2628CD Delft, The Netherlands
Email: leus@cas.et.tudelft.nl

ABSTRACT

A new approach has been proposed recently to describe doubly-selective channels (i.e. time- and frequency-selective channels) with a limited number of parameters referred to as the Basis Expansion Model (BEM). In the BEM, the true channel coefficients are approximated with a high accuracy using a limited number of complex exponentials. In this paper, we propose a new method in order to identify the BEM coefficients of the transmission channel. We consider a transmission scheme where several short training sequences (i.e. their length is comparable to the channel order) are inserted in the stream of data symbols. We propose an iterative method that exploits all the received symbols that contain contributions from the training sequences and blindly filters out the contribution of the unknown surrounding data symbols. The proposed method has a low computational complexity and outperforms existing methods proposed in a similar context.

1. INTRODUCTION

In order to increase data rates when transmitting data over wireless channels, it is often needed to use broadband communication systems. The sampling period can than get smaller than the delay spread of the channel, especially in multipath scenarios, which gives rise to frequency-selective channels. High user mobility combined with high carrier frequencies causes the transmission channel to change rapidly in time, which is referred to as time-selectivity of the channel. Doubly-selective channels that are encountered in high mobility broadband communications with high carrier frequencies thus exhibit both time- and frequency-selectivity.

Many techniques have been proposed to model such channels (e.g. piece-wise constant models, linear interpolation, polynomial interpolation, Bessel functions, etc...). In this paper, we focus on a recently proposed technique for the modeling of doubly-selective channels: the Basis Expansion Model (BEM) [1], [2] and [3]. This new model has attracted a lot of attention recently for it allows an accurate representation of doubly-selective channels with a limited number of parameters and allows cheap and efficient channel equalization [4], [5] [2].

The problem of identifying the BEM parameters of the transmission channel through training has already been discussed in [6] and [7], where in [6] optimal training sequences are presented. These optimal training sequences consist of $2Q + 1$ equispaced bursts of $2L + 1$ pilot symbols with a single non-zero element placed in the middle ($2Q + 1$ represents the number of complex exponentials in the BEM and $L+1$ represents the channel length). However, these methods assume that the period of the BEM is equal to the interval over which we want to identify the channel, which generally leads to a large modeling error at the edges of the interval. Moreveor, these methods only exploit the channel output samples that solely contain contributions from the training symbols. In this paper, we develop a new channel estimation technique that also takes into account the channel output samples that contain contributions from both the training symbols and the unknown surrounding data symbols. The method is independent of the period of the BEM. Moreover, it works for any structure and composition of the available training sequences.

Notation: We use upper (lower) case bold face letters to denote matrices (column vectors). $\mathbf{I}_N$ is the identity matrix of size $N \times N$ and $\mathbf{0}_{M\times N}$ is the all-zero matrix of size $M \times N$; the subscripts are omitted when the dimension of the matrices is clear from the context. The operator $(.)^*$ denotes the complex conjugate, $(.)^T$ the transpose of a matrix, $(.)^H$ its complex conjugate transpose, $(.)^{1/2}$ represents its square root and $tr(.)$ its trace. Finally, $diag(\mathbf{v})$ is a diagonal matrix with the elements of the vector $\mathbf{v}$ placed on its main diagonal.

2. CHANNEL MODEL

2.1. Time-Varying Channels

We propose the following model to describe the transmission of data symbols over a doubly selective channel: Let $x[n]$ be the sequence of transmitted data symbols. Sampling the receive antenna at the symbol rate, the sequence of received data symbols can with-

†The authors would like to acknowledge Marc Moonen for his scientific contribution. They consider him to be a co-author of this paper, however he cannot be listed due to conference restrictions. 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'). The scientific responsibility is assumed by its authors.

out loss of generality be described by:

$$y[n] = \sum_{\nu=-\infty}^{+\infty} h[n;\nu]x[n-\nu] + w[n],$$

where $h[n;\nu]$ accounts for the effects of the transmission channel and the transmit and receive filters ($h[n;\nu]$ is thus the complex multiplicative channel coefficient that accounts for the contribution of the $(n-\nu)^{th}$ transmitted data symbol into the n^{th} received sample), and $w[n]$ is the additive noise, that we will consider to be Gaussian distributed. The large number of independent coefficients of this model ($N(L+1)$ independent coefficients for a channel of order L) makes its use for channel identification or equalization purposes quite unpractical.

2.2. Physical Channel Model

In practical situations, the channel parameters introduced do not vary randomly as they are linked to the physical properties of the transmission channel. The proposed physical channel model allows to parametrize the channel coefficients as a function of the physical transmission channel. Consider a multipath propagation channel where c clusters each consisting in r reflected or scattered rays arrive at the receiver. Considering that the transmission interval is short enough such that the number of rays and clusters does not change during the transmission, and the time-variation of the channel is negligible during the time-span of the receive filter, the transmission channel can be described as:

$$h[n;\nu] = \sum_c \psi(\nu T_s - \tau_c) \sum_r G_{c,r} e^{j2\pi f_{c,r} n T_s}, \qquad (1)$$

where T_s is the symbol period and $\psi(t)$ is the total impulse response of the transmit and receive filters, τ_c is the delay of the c^{th} cluster, $G_{c,r}$ and $f_{c,r}$ are respectively the complex gain and the frequency offset of the r^{th} ray of the c^{th} cluster. The frequency offset is caused by the relative motion between the receiver and the scatterer and is the source of the time-variation of the channel coefficients. The Jakes model [8], which is often proposed to simulate time-varying transmission channels, is a special case of the presented physical channel model.

The Doppler spread f_{max} of the channel is the maximum of all these frequency offsets. When $NT_s > 1/f_{max}$, the channel coefficients undergo significant changes during the transmission of $x[n]$ and the channel is labeled as time-varying, which is the situation we will consider here. The physical channel model presented here, though very handy for simulating realistic time-varying transmission channels, still contains many parameters which makes it impractical to use for channel estimation/equalization applications.

2.3. Basis Expansion Model (BEM)

The Basis Expansion Model (BEM), which has been proposed recently, models time-varying channels with a limited number of parameters and allows low-complexity equalization of these channels. The BEM approximates the actual channel with a limited number of complex exponentials. Assuming that the channel impulse response length is constant and limited to $L+1$, the true channel $h[n;\nu]$ can be approximated over the interval $n = 1 \cdots N$ by its BEM model:

$$h[n;\nu] = \sum_{l=1}^{L} \delta[\nu - l] \sum_{q=-Q}^{Q} h_{q,l} e^{j2\pi qn/N_{mod}} \qquad (2)$$

Each channel tap is modeled as the sum of $2Q+1$ complex exponentials and the whole channel is described with a limited number of $(2Q+1)(L+1)$ parameters, namely the $h_{q,l}$ coefficients. The parameters Q and N_{mod} should be selected carefully in order to allow an accurate approximation of the true channel. The Doppler spread of the BEM channel model (which is its highest frequency component) is equal to $Q/(N_{mod}T_s)$. Q and N_{mod} should be chosen such that the BEM Doppler spread is approximately equal to the Doppler spread of the true channel. Furthermore, the BEM is periodic with a period N_{mod}. Therefore, as the true channel is not periodic, N_{mod} should at least be as large as N; the match of the BEM to the true channel gets tighter as N_{mod} increases. However, increasing N_{mod} forces us to increase Q in order to fulfill the Doppler spread requirement. A good empirical rule for most practical cases is to choose $N_{mod} = 3N$ and then choose Q according to the Doppler spread rule: $Q = \lceil f_{max} N_{mod} T_s \rceil$, which yields a very tight match of the BEM with a limited number of parameters. When the channel varies slowly and $1/(3NT_s) \gg f_{max}$, the above procedure yields $Q = 1$ but the Doppler Spread of the BEM will be significantly larger than the true Doppler spread, yielding a poor match of the BEM. In this case, increasing N_{mod} in order to make the true Doppler spread equal to the BEM Doppler spread largely improves the accuracy of the BEM: $N_{mod} = \lceil 1/(T_s f_{max}) \rceil$.

Using the BEM, the input-output relationship of the transmission channel over the interval $n = 1 \cdots N$ is written as:

$$y[n] = \sum_{l=0}^{L} \sum_{q=-Q}^{Q} h_{q,l} e^{j2\pi qn/N_{mod}} x[n-l] + w[n]. \qquad (3)$$

3. IDENTIFICATION OF THE BEM COEFFICIENTS

In this section, we analyze how a time-varying channel can be identified at the receiver based on the knowledge of training symbols inserted in the stream of transmitted data symbols. We consider a transmission scheme where K equispaced short clusters of training symbols of length n_t are inserted in the stream of data symbols, which is a natural placement of the training symbols towards the identification of time-varying channels [6]. Note however that this hypothesis of equi-spaced training bursts of constant length is not mandatory for the proposed method to work. We adopt it only for the clarity of the presentation but it is straightforward to adapt the method to the more general situation where the length of the training sequences and the spacing between them varies. We aim at identifying the BEM coefficients that provide the best match to the true channel taking into account all the channel output samples that contain contributions from the training symbols, including those who contain contributions from both the training symbols and the unknown surrounding data symbols.

3.1. Data Model

Let $\mathbf{t}_k = [t_k[1], \cdots, t_k[n_t]]^T, \quad k = 1 \cdots K$ be the k^{th} training sequence inserted into the stream of data symbols. Let $\mathbf{s}_k = \left[s_k[1], \cdots, s_k[n_s]^T\right]^T$ be the block of data symbols placed after the k^{th} training sequence. The transmitted burstcan then be written as: $\mathbf{x} = [x[1], \cdots, x[N]]^T = [\mathbf{t}_1^T, \mathbf{s}_1^T, \cdots \mathbf{t}_K^T, \mathbf{s}_K^T]^T$ (note that $N = K(n_t) + n_s$). Existing methods for training-based estimation of doubly-selective channels [6], [7] only exploit the channel output samples that solely contain contributions from the training sequence $\mathbf{t}_k$ and discard all the channel output samples that contain contributions from the unknown data symbols $\mathbf{s}_k$. The method we shall present here exploits all the channel output samples that contain contributions from $\mathbf{t}_k$, including those who contain contributions from both unknown data symbols and training symbols.

Let $\mathbf{u}_k$ be the $(n_t+L\times 1)$ vector of the channel output samples that contain contributions from $\mathbf{t}_k$: $\mathbf{u}_k = [y[(k-1)(n_s+n_t)+1], \cdots, y[(k-1)(n_s+n_t)+n_t+L]]^T$, the characterization of which can be obtained using (3) if the channel can be described accurately by the BEM (2).

Developing the BEM expression of the channel parameters and re-arranging the resulting expression, we obtain the following data model that is well-suited for the identification of the channel's BEM coefficients:

$$\mathbf{u}_k = \boldsymbol{\mathcal{T}}_k \mathbf{h}_{BEM} + \boldsymbol{\epsilon}_k. \tag{4}$$

• The first term, $\boldsymbol{\mathcal{T}}_k \mathbf{h}_{BEM}$ is a deterministic term where $\mathbf{h}_{BEM}$ is the $(2Q+1)(L+1)$-wide vector of the channel's BEM coefficients: $\mathbf{h}_{BEM} = [h_{-Q,0}, \cdots, h_{Q,0}, \cdots, h_{Q,L}]^T$ and $\boldsymbol{\mathcal{T}}_k$ is an $(n_t+L) \times ((2Q+1)(L+1))$ matrix accounting for the contributions of the complex exponentials of the BEM and the training sequences, which has the following structure:

$$\boldsymbol{\mathcal{T}}_k = \begin{bmatrix} \boxed{\mathbf{T}_{k,0}} & \boxed{\mathbf{T}_{k,1}} & \ddots & \boxed{\mathbf{T}_{k,L}} \end{bmatrix},$$

with $\mathbf{T}_{k,l} = diag(\mathbf{t}_k)\mathbf{C}_{k,l}$, where $\mathbf{C}_{k,l}$ accounts for the BEM's complex exponentials multiplying the $h_{q,l}$ coefficients: $\mathbf{C}_{k,l}[x,y] = e^{j2\pi(y-Q-1)\frac{(k-1)(n_s+n_t)+l+x}{N_{mod}}}$.

• The second term, $\boldsymbol{\epsilon}_k$ is stochastic and represents the contributions of the unknown surrounding data symbols and the AWGN:

$$\boldsymbol{\epsilon}_k = \underbrace{\left[\ \mathbf{H}_{s,k}^L \ \middle|\ \mathbf{H}_{s,k}^R\ \right]}_{\mathbf{H}_{s,k}} \mathbf{s}'_k + \mathbf{w}_k, \tag{5}$$

where $\mathbf{w}_k$ is the AWGN vector, $\mathbf{s}'_k = [s_{k-1}[n_s-L+1], \cdots, [s_{k-1}[n_s], s_k[1], \cdots, s_k[L]]^T$ is the vector of the unknown data symbols contributing to $\mathbf{u}_k$ (assuming $n_s \geq L$). $\mathbf{H}_{s,k}$ is an $(n_t+L) \times 2L$ matrix gathering the channel coefficients that multiply these data symbols. It is the concatenation of two matrices:

$$\mathbf{H}_{s,k}^L = \begin{bmatrix} h[n_{k,1};L] & \cdots & h[n_{k,1};1] \\ & \ddots & \vdots \\ \mathbf{0} & & h[n_{k,L};L] \\ \hline & \mathbf{0}_{(n_t \times L)} & \end{bmatrix},$$

$$\mathbf{H}_{s,k}^R = \begin{bmatrix} & \mathbf{0}_{(n_t \times L)} & \\ \hline h[n_{k,n_t};0] & & \mathbf{0} \\ \vdots & \ddots & \\ h[n_{k,n_t+L};L-1] & \cdots & h[n_{k,n_t+L};0] \end{bmatrix},$$

where $n_{k,l}$ is a shorthand notation for the index of the l^{th} element of $\mathbf{u}_k$: $n_{k,l} = (k-1)(n_s+n_t)+l$.

3.2. Proposed Algorithms

Assuming that the noise and the data are white and zero-mean with variance σ^2 for the noise samples and λ^2 for the data symbols (i.e. $\mathrm{E}\{\mathbf{s}_k\} = \mathbf{0}$ and $\mathrm{E}\{\mathbf{w}_k\} = \mathbf{0}$, $\mathrm{E}\{\mathbf{s}_k\mathbf{s}_k^H\} = \lambda^2\mathbf{I}$, $\mathrm{E}\{\mathbf{w}_k\mathbf{w}_k^H\} = \sigma^2\mathbf{I}$, $\forall k$, $\mathrm{E}\{\mathbf{s}_k\mathbf{s}_l^H\} = \mathbf{0}$ and $\mathrm{E}\{\mathbf{w}_k\mathbf{w}_l^H\} = \mathbf{0}$, $\forall k,l; l \neq k$), it is straightforward to derive the first- and second order statistics of $\boldsymbol{\epsilon}_k$ (assuming also $n_s \geq 2L$):

$$\begin{aligned} &\mathrm{E}\{\boldsymbol{\epsilon}_k\} = \mathbf{0}, \\ &\mathrm{E}\left\{\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_l^H\right\} = \delta_{k,l}\mathbf{Q}_k, \ \ \forall k,l, \\ &\mathbf{Q}_k = \lambda^2\mathbf{H}_{s,k}\mathbf{H}_{s,k}^H + \sigma^2\mathbf{I}. \end{aligned} \tag{6}$$

LS Channel Estimate

Relying on the first-order statistics of $\boldsymbol{\epsilon}_k$, a simple Least Squares (LS) approach provides us with an unbiased estimator of $\mathbf{h}_{BEM}$:

$$\hat{\mathbf{h}}_{LS} = \left(\sum_{k=1}^{K} \boldsymbol{\mathcal{T}}_k^H \boldsymbol{\mathcal{T}}_k\right)^{-1} \sum_{k=1}^{K} \boldsymbol{\mathcal{T}}_k^H \mathbf{u}_k. \tag{7}$$

Because of the presence of the complex exponentials, the inverse of the sum will always exist as soon as $K(n_t+L) \geq (2Q+1)(L+1)$.

WLS Channel Estimate

Since $\boldsymbol{\epsilon}_k$ is not white, the LS approach is not optimal. A Weighted Least Squares (WLS) approach taking into account the color of $\boldsymbol{\epsilon}_k$ would yield an improved estimate of the channel parameters. *Assuming that all the* $\mathbf{Q}_k$*'s are known* (see also next paragraph), the WLS estimate of $\mathbf{h}_{BEM}$ can be computed as:

$$\hat{\mathbf{h}}_{WLS} = \left(\sum_{k=1}^{K} \boldsymbol{\mathcal{T}}_k^H \mathbf{Q}_k^{-1} \boldsymbol{\mathcal{T}}_k\right)^{-1} \sum_{k=1}^{K} \boldsymbol{\mathcal{T}}_k^H \mathbf{Q}_k^{-1} \mathbf{u}_k. \tag{8}$$

The presence of the AWGN term in $\mathbf{Q}_k$ ensures the existence of its inverse and the inverse of the sum exists under the same conditions as for the LS estimate.

Iterative WLS Channel Estimate

Unfortunately, $\mathbf{Q}_k$ is not known at the receiver for it depends on the sought channel. The WLS approach can thus not be straightforwardly adopted. We propose below an iterative approach that allows to cope with the dependence of $\mathbf{Q}_k$ on the channel.

Assume a channel estimate $\hat{\mathbf{h}}_{BEM}^{(i)}$ is available at the receiver (i^{th} iteration). Exploiting (2) and the definition of $\mathbf{H}_{s,k}$, it is possible to construct its estimate $\hat{\mathbf{H}}_{s,k}^{(i)}$ from $\hat{\mathbf{h}}_{BEM}^{(i)}$. Relying on the parametric definition of $\mathbf{Q}_k$ and assuming that σ^2 is known, we construct the estimate $\hat{\mathbf{Q}}_k^{(i)}$ of the color of $\boldsymbol{\epsilon}_k$. This estimate is used to produce a refined estimate $\hat{\mathbf{h}}_{BEM}^{(i+1)}$ of the channel model with a WLS approach:

$$\hat{\mathbf{h}}_{BEM}^{(i+1)} = \left(\sum_{k=1}^{K} \boldsymbol{\mathcal{T}}_k^H \hat{\mathbf{Q}}_k^{(i)-1} \boldsymbol{\mathcal{T}}_k\right)^{-1} \sum_{k=1}^{K} \boldsymbol{\mathcal{T}}_k^H \hat{\mathbf{Q}}_k^{(i)-1} \mathbf{u}_k.$$

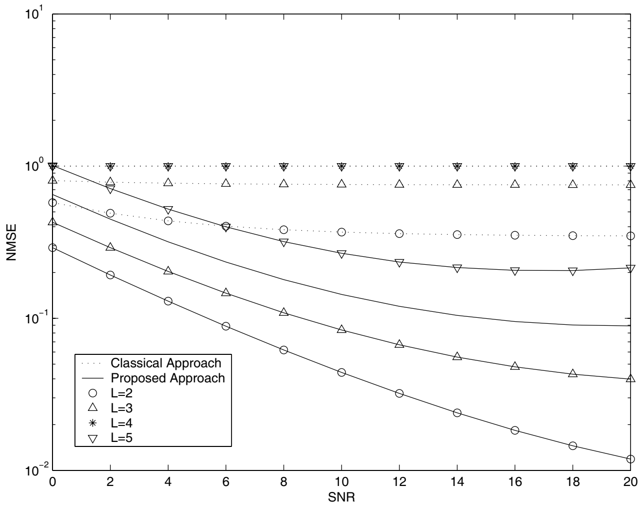

Fig. 1. NMSE vs.SNR for increasing channel orders L when training sequences of length 4 are inserted in the transmitted data

The iterative procedure is stopped when there is no significant difference between two consecutive channel estimates. If the starting point is sufficiently accurate, this iterative procedure converges to a solution which is close to the true WLS estimate. It is possible to show that the convergence point of this iterative procedure is the gaussian ML channel estimate (i.e. the ML channel estimate when the surrounding data symbols are assumed to be Gaussian distributed).

The iterative procedure can be initialized with the LS channel estimate of (7): $\hat{\mathbf{h}}^{(0)}_{BEM} = \hat{\mathbf{h}}_{LS}$, which is equivalent to choosing $\hat{\mathbf{Q}}^{(0)}_k = \mathbf{I}$, $\forall k$. Experimental results show that this choice yields good convergence properties of the iterative procedure.

4. EXPERIMENTAL RESULTS

We compare the proposed method with the method proposed in [6] and [7] (note that we generalize these methods to handle arbitrary BEM periods). The channels that are used for the simulations are obtained using a physical channel model with 10 clusters of 100 rays each.

When identifying the BEM of a physical channel, there are two sources of mismatch between the resulting channel model and the actual channel. The first source of error is the BEM-induced modelling error: if all the channel coefficients $h[n;\nu]$ were known, the best possible BEM (with fixed parameters Q and N_{mod}) would not match exactly the physical channel. Moreover, there is a difference between the best possible BEM and the estimated BEM obtained through the proposed identification procedure. This results in an identification error, which is the second source of error.

The performance of the proposed identification method could thus be assessed either by the identification error (the difference between the obtained BEM and the best possible BEM) or by the total error (the difference between the obtained BEM and the actual channel). We adopt here the second possibility as it assesses the total performance of the system. The performance metric is thus the normalized mean square error (NMSE) between the obtained channel BEM and the true channel. The results are averaged over 100 different channels, performing 100 runs for each channel. A run consists in the transmission of 64 blocks, each containing 4 training symbols and 16 data symbols. The Doppler spread and the noise power are assumed to be known at the receiver. The training sequences are constant-modulus symbols with a uniform phase distribution. Given the chosen Doppler spread and burst length ($N = 1024$), the parameters of the BEM obtained using the procedure described in this paper are the following: $Q = 2$ and $N_{mod} = 3072$. We perform the simulations using the same setup for different channel orders, namely $L = 2, 3, 4$ and 5. The results are presented in Fig.1. The proposed method clearly outperforms the existing one. Note that the existing method is unable to cope with long channels ($L = 4, 5, \cdots$), whilst the proposed method keeps generating accurate channel estimates.

5. CONCLUSIONS

In this paper, we have introduced a new training-based method that allows to identify the BEM model of doubly-selective channels. The method is able to cope with training sequences of various lengths and compositions. Taking into account the channel output samples that contain contributions from both the training symbols and the unknown surrounding data symbols allows the proposed method to clearly outperform existing methods.

6. REFERENCES

[1] M. K. Tsatsanis and G. B. Giannakis, "Modeling and Equalization of Rapidly Fading Channels," *International Journal of Adaptive Control and Signal Processing*, vol. 10, pp. 159–176, 1996.

[2] G. B. Giannakis and C. Tepedelioglo, "Basis Expansion Model and Diversity Techniques for Blind Identification and Equalization of Time Varying Channels," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1969–1986, Oct 1998.

[3] A. M. Sayeed and B. Aazhang, "Joint Multipath-Doppler Diversity in Mobile Wireless Communications," *IEEE Trans. on Communications*, vol. 47, pp. 123–132, Jan 1999.

[4] I. Barhumi, G. Leus and M. Moonen, "Time-Varying FIR Equalization of Doubly Selective Channels," *IEEE Transactions on Wireless Communications*, 2004, to appear.

[5] G. Leus, I. Barhumi, M. Moonen, "MMSE Time-Varying FIR Equalization of Doubly-Selective Channels," in *Proc. of the International Conference on Acoustics, Speech, and Signal Processing (ICCASP 2003)*, Hong Kong, April 2003, pp. IV–485–IV–488.

[6] G. B. Giannakis X. Ma and S. Ohno, "Optimal Training for Block Transmission Over Doubly Selective Wireless Channels," *IEEE Transaction on Signal Processing*, vol. 51, no. 5, pp. 1351–1366, May 2003.

[7] P. Schniter, "Low-Complexity Estimation of Doubly-Selective Channels," in *proc. of Signal Processing Advances in Wireless Communications*, Rome, Italy, June 2003.

[8] W. C. Jakes, *Microwave Mobile Communications*, chapter 3, Wiley, 1974.