in Proc. IEEE GLOBECOM 2000
San Francisco (CA), Nov. 2000,
pp. 1031–1035.

EFFICIENT POCS ALGORITHMS FOR DETERMINISTIC BLIND EQUALIZATION OF TIME-VARYING CHANNELS

Harold Artés, Franz Hlawatsch, and Gerald Matz

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria
phone: +43 1 58801 38915, fax: +43 1 58801 38999, email: fhlawats@email.tuwien.ac.at
web: http://www.nt.tuwien.ac.at/dspgroup/time.html

ABSTRACT

We present two POCS (projections onto convex sets) algorithms for deterministic blind equalization of linear time-varying (LTV) channels. Our approach is based on a multichannel LTI representation of LTV channels. We prove a theorem on unique reconstruction/equalization, formulate a POCS algorithm for the resolution of a matrix ambiguity, and show how to cope with unequal subchannel lengths. We also present an alternative POCS algorithm for equalization that avoids a singular value decomposition. Both algorithms are guaranteed to converge to the desired solution.

1. INTRODUCTION

This paper presents POCS (projections onto convex sets) algorithms for deterministic blind equalization or symbol estimation for linear time-varying (LTV) channels. The overall method is based on a discrete representation of LTV channels in terms of uniformly discretized Doppler shifts and LTI subchannels. As opposed to [1], we do not require estimation of the Doppler shifts of individual scatterers. The subchannels are allowed to have different lengths.

The paper is organized as follows. The channel representation is discussed in Section 2. The overall framework for the equalization method is explained in Section 3. Section 4 presents a uniqueness theorem and a POCS algorithm for resolving a matrix ambiguity occurring in the initial formulation of the method. Section 5 presents an improved equalization method that is entirely based on a POCS algorithm. Finally, simulation results are provided in Section 6.

2. CHANNEL REPRESENTATION

We consider a discrete-time signal $s[n]$ that is transmitted over a discrete-time LTV channel with impulse response $h[n,m]$. The channel's input-output relation is

$$x[n] = \sum_{m=0}^{L-1} h[n,m]\, s[n-m]\,, \qquad n \in [0, N-1]\,, \tag{1}$$

where $L-1$ is the channel's maximum time delay and $[0, N-1]$ is the interval over which the output signal $x[n]$ is observed (later we shall allow for a time offset n_0). In a communications application, the input signal $s[n]$ will typically be a sequence of data symbols (the transmit and receive filters are incorporated in the channel [2]).

Funding by FWF grants P12228-TEC and P11904-TEC.

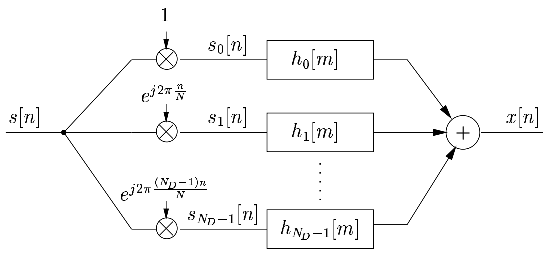

Fig. 1: Multichannel LTI representation of an LTV channel.

Using the *(delay-Doppler) spreading function* [3]

$$S_h[m,l] \triangleq \sum_{n=0}^{N-1} h[n,m]\, e^{-j2\pi\frac{ln}{N}}\,,$$

we can rewrite (1) as

$$x[n] = \sum_{l=0}^{N_D-1} \sum_{m=0}^{L-1} h_l[m]\, s_l[n-m]\,, \tag{2}$$

with $h_l[m] \triangleq S_h[m,l]\, e^{j2\pi\frac{lm}{N}}$, $s_l[n] \triangleq s[n]\, e^{j2\pi\frac{ln}{N}}$, and N_D the maximum Doppler shift. This corresponds to a *multichannel LTI representation* of the LTV channel where each subchannel consists of a modulator (Doppler shift) and an LTI filter (see Fig. 1). Thus, practically arbitrary channels—including channels with diffuse scattering—are characterized by a finite set of LTI filters associated with *uniformly spaced*, discrete Doppler shifts. Equivalently, $h[n,m]$ is expanded (cf. [4]) into the basis functions $e^{j2\pi\frac{ln}{N}}$ that naturally arise through discretization of a continuous-time channel. Estimation of basis functions or Doppler shifts is not required.

In what follows, we assume that K of the N_D subchannels in Fig. 1, corresponding to specific Doppler shifts $l_k \in [0, N_D-1]$ with $k = 1, 2, \cdots, K$, are active. That is, only the subchannel impulse responses $h_{l_k}[m]$ are nonzero. (This is no restriction since we allow $K = N_D$.) Thus, (2) becomes

$$x[n] = \sum_{k=1}^{K} \sum_{m=0}^{L-1} h_{l_k}[m]\, s_{l_k}[n-m]\,. \tag{3}$$

The *model parameters* that will be relevant to our method are the "active" Doppler shift indices l_k and the lengths of the corresponding channel impulse responses $h_{l_k}[m]$. For WSSUS channels, these model parameters can be deduced from the channel's scattering function [3, 5]. Because the scattering function of a WSSUS channel does not change with time, it is much easier to estimate than the channel itself [6, 7].

As is often done in the context of blind estimation (e.g., [8]), we assume that we observe M output signals of the type (3)

$$x_i[n] = \sum_{k=1}^{K} \sum_{m=0}^{L-1} h_{l_k}^{(i)}[m]\, s_{l_k}[n-m]\,, \qquad i = 1, 2, \cdots, M\,. \tag{4}$$

These are obtained from an array of M sensors/antennas or by oversampling the output of a single sensor/antenna by the factor M. The corresponding M channels are assumed to possess the same scattering function and thus the same channel model parameters [2]. For now, we assume that the LTI impulse responses $h_{l_k}^{(i)}[m]$ have equal length L; the case of unequal lengths will be addressed later.

3. BLIND EQUALIZATION

The LTV channel representation in Fig. 1 is closely related to multiuser communication over LTI channels (with $s_{l_k}[n]$ associated to the l_kth user). This correspondence allows us to adapt the multiuser method in [2] to our problem. In contrast to [9], this permits direct calculation of the input signal $s[n]$ without first calculating the channel.

We start by reformulating (4) as follows [2]. Let

$$\mathbf{H}[m] \triangleq \begin{bmatrix} h_{l_1}^{(1)}[m] & \cdots & h_{l_K}^{(1)}[m] \\ \vdots & & \vdots \\ h_{l_1}^{(M)}[m] & \cdots & h_{l_K}^{(M)}[m] \end{bmatrix}$$

and $\mathbf{H} \triangleq \left[\mathbf{H}[0] \cdots \mathbf{H}[L-1]\right]$, and define the following channel matrix of size $Mu \times K(L+u-1)$, in which $\mathbf{H}$ is stacked u times with shifts to the left by K positions each (the stacking parameter u is sometimes called *smoothing factor* [2]),

$$\boldsymbol{\mathcal{H}} \triangleq \begin{bmatrix} \mathbf{0} & & \boxed{\mathbf{H}} \\ & \boxed{\mathbf{H}} & \\ & \iddots & \\ \boxed{\mathbf{H}} & & \mathbf{0} \end{bmatrix}.$$

Next, we define a vector of modulated input samples, $\mathbf{s}[n] \triangleq \left[s_{l_1}[n] \cdots s_{l_K}[n]\right]^T$, and we form the following block-Toeplitz input matrix of size $K(L+u-1) \times (N-u+1)$,

$$\boldsymbol{\mathcal{S}} \triangleq \begin{bmatrix} \mathbf{s}[u-1] & \mathbf{s}[u] & \cdots & \mathbf{s}[N-1] \\ \mathbf{s}[u-2] & \mathbf{s}[u-1] & \cdots & \mathbf{s}[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{s}[-L+1] & \mathbf{s}[-L+2] & \cdots & \mathbf{s}[N-L-u+1] \end{bmatrix}.$$

Finally, we define the output vector $\mathbf{x}[n] \triangleq \left[x_1[n] \cdots x_M[n]\right]^T$ and form the following block-Hankel output matrix of size $Mu \times (N-u+1)$,

$$\boldsymbol{\mathcal{X}} \triangleq \begin{bmatrix} \mathbf{x}[0] & \mathbf{x}[1] & \cdots & \mathbf{x}[N-u-1] & \mathbf{x}[N-u] \\ \mathbf{x}[1] & \mathbf{x}[2] & \cdots & \mathbf{x}[N-u] & \mathbf{x}[N-u+1] \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ \mathbf{x}[u-1] & \mathbf{x}[u] & \cdots & \mathbf{x}[N-2] & \mathbf{x}[N-1] \end{bmatrix}.$$

Now (4) can be written as the matrix relation (cf. [9])

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{S}}\,. \tag{5}$$

Blind equalization corresponds to calculation of the input/symbol matrix $\boldsymbol{\mathcal{S}}$ from the known (observed) matrix $\boldsymbol{\mathcal{X}}$. For this to be possible, $\boldsymbol{\mathcal{S}}$ must be a wide matrix and the row span of $\boldsymbol{\mathcal{X}}$ must be equal to the row span of $\boldsymbol{\mathcal{S}}$ [9, 2]. This, in turn, requires that $\boldsymbol{\mathcal{H}}$ is a square or tall matrix and has full rank. These requirements lead to the necessary conditions [2] $u \geq \frac{K(L-1)}{M-K}$, with $M > K$, and

$$N \,>\, KL + (K+1)(u-1)\,. \tag{6}$$

For $u = 1$, condition (6) becomes $KL < N$. If all subchannels are active (i.e., $K = N_D$), this equals the discrete underspread condition [6] that guarantees identifiability of the channel (cf. [10]). For $u > 1$ and $K = N_D$, condition (6) becomes $N_D L < N - (N_D+1)(u-1)$, which is a more restrictive underspread condition.

The input matrix $\boldsymbol{\mathcal{S}}$ is a block-Toeplitz matrix that is generated by the element vectors $\mathbf{s}[-L+1], \mathbf{s}[-L+2], \cdots, \mathbf{s}[N-1]$ or, equivalently, by the "generating matrix"

$$\mathbf{S} \triangleq \left[\mathbf{s}[-L+1]\;\, \mathbf{s}[-L+2]\; \cdots\; \mathbf{s}[N-1]\right]$$

of size $K \times (N+L-1)$. Thus, our equalization problem can be equivalently phrased as calculation of $\mathbf{S}$ from $\boldsymbol{\mathcal{X}}$. This can be done as follows:

Step 1: Using a singular value decomposition (SVD), the row span of $\boldsymbol{\mathcal{S}}$ is calculated from $\boldsymbol{\mathcal{X}}$ [2].

Step 2: Another SVD is used to construct a $K(L+u-1) \times (N-u+1)$ block-Toeplitz matrix $\boldsymbol{\mathcal{S}}_A$ whose row span equals that of $\boldsymbol{\mathcal{S}}$ [9, 2]. It can be shown that the $K \times (N+L-1)$ generating matrix of $\boldsymbol{\mathcal{S}}_A$ can be written as $\mathbf{S}_A = \mathbf{A}\mathbf{S}$, where $\mathbf{A}$ is an unknown invertible matrix of size $K \times K$ [2]. Due to the SVD construction, the rows of $\mathbf{S}_A$ are orthonormal.

At this point, $\mathbf{S}$ is determined up to an ambiguity corresponding to the unknown matrix $\mathbf{A}$.

Step 3: This ambiguity is resolved whereby $\mathbf{S}$ and, in turn, the scalar input sequence $s[-L+1], s[-L+2], \cdots, s[N-1]$ (which is the first row of $\mathbf{S}$) are obtained. A closed-form solution to the matrix ambiguity problem has been presented in [9]. However, we next present a POCS method that is computationally more efficient and allows to incorporate *a-priori* knowledge such as the symbol alphabet.

4. RESOLVING THE MATRIX AMBIGUITY

Resolution of the matrix ambiguity (Step 3) relies on the relations between the $s_{l_k}[n]$ (recall that $s_{l_k}[n] = s[n]\, e^{j2\pi \frac{l_k n}{N}}$).

Uniqueness theorem. The structure of the $K \times (N+L-1)$ generating matrix $\mathbf{S}$ can be captured by factoring $\mathbf{S}$ as

$$\mathbf{S} = \mathbf{M}(n_0)\mathbf{D}\,, \tag{7}$$

with the $K \times (N+L-1)$ "modulation matrix"

$$\mathbf{M}(n_0) \triangleq \begin{bmatrix} W_N^{-n_0 l_1} & W_N^{-(n_0+1)l_1} & \cdots & W_N^{-(n_0+N+L-1)l_1} \\ \vdots & \vdots & & \vdots \\ W_N^{-n_0 l_K} & W_N^{-(n_0+1)l_K} & \cdots & W_N^{-(n_0+N+L-1)l_K} \end{bmatrix},$$

where $W_N^k \triangleq e^{-j2\pi\frac{k}{N}}$, and with the $(N+L-1) \times (N+L-1)$ diagonal input matrix $\mathbf{D} \triangleq \mathrm{diag}\{s[-L+1], s[-L+2], \cdots, s[N-1]\}$. Note that by the definition of $\mathbf{M}(n_0)$, we allow for a time offset n_0 that will generally occur in practice. The next theorem, proved in the Appendix, states the following result: If $\mathbf{S}_A = \mathbf{A}\mathbf{S} = \mathbf{A}\mathbf{M}(n_0)\mathbf{D}$ is given and $\mathbf{M}(n_0)$ is known except for the time offset n_0 (i.e., the active Doppler shifts l_k are known) but $\mathbf{A}$ is unknown, then $\mathbf{D}$ is uniquely defined up to a scalar factor.

Theorem 1. *Let $\mathbf{M}(n_0)$ be defined as above, and assume $0 \leq l_k \leq \lceil N/2 \rceil - 1$. Then for diagonal matrices $\mathbf{D}$ and $\widetilde{\mathbf{D}}$, with $\mathbf{D}$ nonsingular, the equation*

$$\mathbf{A}\mathbf{M}(n_0)\mathbf{D} = \mathbf{M}(\tilde{n}_0)\widetilde{\mathbf{D}} \tag{8}$$

implies $\widetilde{\mathbf{D}} = c\mathbf{D}$ with $c \in \mathbb{C}$ and $\mathbf{A} = c\,\mathrm{diag}\{W_N^{(n_0-\tilde{n}_0)l_1}, W_N^{(n_0-\tilde{n}_0)l_2}, \cdots, W_N^{(n_0-\tilde{n}_0)l_K}\}$.

Thus, given $\mathbf{S}_A = \mathbf{AS}$ and the active Doppler shifts l_k, the input samples $s[-L+1], s[-L+2], \cdots, s[N-1]$ are uniquely defined up to a common factor c; this is true even if the time offset n_0 is unknown. The Doppler shift constraint $0 \leq l_k \leq \lceil N/2 \rceil - 1$ is always satisfied in communications applications. The condition that $\mathbf{D}$ be nonsingular means that the input samples $s[-L+1], \cdots, s[N-1]$ are nonzero.

Theorem 1 can be adapted to the channel model used in [1]. It can even be extended to the case of unknown active Doppler shifts l_k; in that case, $s[n]$ is uniquely defined up to a factor $c\,e^{j2\pi \frac{ln}{N}}$ with some $l \in [0, N-1]$. (This ambiguity can still be resolved if $s[n]$ is a symbol sequence with known symbol alphabet.)

POCS algorithm. We can reformulate our uniqueness result as follows. Given a matrix $\mathbf{S}_A$ and a modulation matrix $\mathbf{M}(n_0)$ (with n_0 arbitrary but fixed), the desired generating matrix $\mathbf{S}$ is uniquely determined (up to a scalar factor) by the following two properties:

1. $\mathbf{S} = \mathbf{M}(n_0)\mathbf{D}$ with $\mathbf{D}$ diagonal;
2. the row span of $\mathbf{S}$ lies in the row span of $\mathbf{S}_A$.

Equivalently, $\mathbf{S} \in \mathcal{A} \cap \mathcal{B}$ where $\mathcal{A}$ denotes the linear subspace of all matrices $\mathbf{M}(n_0)\mathbf{D}$ with $\mathbf{M}(n_0)$ given and $\mathbf{D}$ diagonal, and $\mathcal{B}$ denotes the linear subspace of all matrices whose row span lies in the row span of $\mathbf{S}_A$, i.e., of all matrices of the form $\mathbf{B}\mathbf{S}_A$ with some $K \times K$ matrix $\mathbf{B}$.

Since both $\mathcal{A}$ and $\mathcal{B}$ are linear subspaces and thus convex, the formulation $\mathbf{S} \in \mathcal{A} \cap \mathcal{B}$ suggests a *POCS algorithm* [11] for matrix ambiguity resolution. This algorithm is iterative and consists in alternately projecting the iterated version of the matrix $\mathbf{S}$ onto $\mathcal{A}$ and $\mathcal{B}$.

Projection onto $\mathcal{A}$: The projection onto $\mathcal{A}$ amounts to forming $\mathbf{S}^{(i)} = \mathbf{M}(n_0)\mathbf{D}^{(i)}$, where the nonzero (diagonal) elements of $\mathbf{D}^{(i)}$ can be shown to be given by

$$\left(\mathbf{D}^{(i)}\right)_{n,n} = \frac{1}{K}\sum_{k=1}^{K} \left(\mathbf{S}^{(i-1)}\right)_{k,n} \left(\mathbf{M}(n_0)\right)^*_{k,n}.$$

Here, $\mathbf{S}^{(i-1)}$ is the result of the previous iteration (i.e., the projection onto $\mathcal{B}$).

Projection onto $\mathcal{B}$: The projection onto $\mathcal{B}$ amounts to forming $\mathbf{S}^{(i)} = \mathbf{B}^{(i)}\mathbf{S}_A$, where it can be shown that

$$\mathbf{B}^{(i)} = \mathbf{S}^{(i-1)}\mathbf{S}_A^{\#}.$$

Here, $\mathbf{S}^{(i-1)}$ is the result of the previous iteration (i.e., the projection onto $\mathcal{A}$) and $\mathbf{S}_A^{\#}$ is the pseudo-inverse of $\mathbf{S}_A$. Since $\mathbf{S}_A$ is a wide matrix with orthonormal rows, there is simply $\mathbf{S}_A^{\#} = \mathbf{S}_A^H$.

The POCS algorithm is guaranteed to converge to an intersection point, i.e., $\mathbf{S}^{(\infty)} \in \mathcal{A} \cap \mathcal{B}$ [11]. Thus, $\mathbf{S}^{(\infty)} = c\,\mathbf{S}$ where $\mathbf{S}$ is the desired generating matrix (which is unique according to Theorem 1) and $c \in \mathbb{C}$. The convergence speed depends on the initialization, $\mathbf{S}^{(0)}$. In the semiblind case, we could use some known symbols to calculate a good initialization. However, our simulations showed that initialization is not very critical since the convergence of the POCS algorithm always was much faster than the two SVDs required to obtain $\mathbf{S}_A$ from $\boldsymbol{\mathcal{X}}$. The convergence speed can be increased by *relaxation* [11] and/or by using knowledge of the symbol alphabet [12]. The latter approach, however, introduces a non-convex set and thus convergence to the desired solution is no longer guaranteed. We found that, typically, the POCS method is much more efficient than direct calculation of the closed-form solution provided in [9].

Unequal channel lengths. We now allow the K active subchannels to have unequal lengths. (However, physical considerations motivate the assumption that corresponding active subchannels of the M different channels in (4)—i.e., $h_{l_k}^{(i)}[m]$ with different i values but the same k—have equal length.) Specifically, we assume that there are P sets of active subchannels, where the pth set consists of K_p subchannels with the same length L_p and with Doppler shifts $l_k^{(p)}$ $(k = 1, 2, \cdots, K_p)$. (Note that $\sum_{i=1}^{P} K_p = K$.) Here, a problem is that the $Mu \times K(L+u-1)$ matrix $\boldsymbol{\mathcal{H}}$ (with $L = \max_p L_p$) will not have full rank. However, we can replace (5) by $\boldsymbol{\mathcal{X}} = \sum_{p=1}^{P} \boldsymbol{\mathcal{H}}_p \boldsymbol{\mathcal{S}}_p$ where the $Mu \times K_p(L_p + u - 1)$ matrices $\boldsymbol{\mathcal{H}}_p$ have full rank. The generating matrix $\mathbf{S}_p$ of $\boldsymbol{\mathcal{S}}_p$ can be written as (cf. (7))

$$\mathbf{S}_p = \mathbf{M}_p(n_0)\mathbf{D}_p\,,$$

with the $K_p \times (N + L_p - 1)$ modulation matrix

$$\mathbf{M}_p(n_0) \triangleq \begin{bmatrix} W_N^{-n_0 l_1^{(p)}} & W_N^{-(n_0+1)l_1^{(p)}} & \cdots & W_N^{-(n_0+N+L_p-1)l_1^{(p)}} \\ \vdots & \vdots & & \vdots \\ W_N^{-n_0 l_{K_p}^{(p)}} & W_N^{-(n_0+1)l_{K_p}^{(p)}} & \cdots & W_N^{-(n_0+N+L_p-1)l_{K_p}^{(p)}} \end{bmatrix}$$

and the $(N + L_p - 1) \times (N + L_p - 1)$ diagonal input matrix $\mathbf{D}_p \triangleq \mathrm{diag}\{s[-L_p+1], s[-L_p+2], \cdots, s[N-1]\}$. The algorithms in [13] and [2] can be used to retrieve $\mathbf{S}_{A,p} = \mathbf{A}_p\mathbf{S}_p$ (with unknown $\mathbf{A}_p$) from $\boldsymbol{\mathcal{X}}$. According to Theorem 1, $\mathbf{S}_p$ is uniquely defined (up to a scalar factor) by $\mathbf{S}_{A,p}$ even if the time offset n_0 is unknown. Our POCS algorithm can be used to calculate $\mathbf{S}_p$ and thus the input samples $s[-L_p+1], \cdots, s[N-1]$. Whereas theoretically one such calculation is sufficient, with noise present it will generally be advisable to use several or all p and average the corresponding results for $s[0], \cdots, s[N-1]$.

5. POCS EQUALIZATION ALGORITHM

A computationally intensive part of the three-step method of Section 3 is the SVD in Step 2 that is used to construct the generating matrix $\mathbf{S}_A$. This SVD can be avoided by the following approach. As mentioned in Section 3, because of its block Toeplitz structure $\mathbf{S}$ can be reconstructed from the row span of $\boldsymbol{\mathcal{X}}$ up to a matrix ambiguity. According to Section 4, this matrix ambiguity can be resolved based on the modulation structure $\mathbf{S} = \mathbf{M}(n_0)\mathbf{D}$ (with $\mathbf{D}$ diagonal). Consequently, $\boldsymbol{\mathcal{S}}$ is uniquely determined (up to a scalar factor) by the following two properties:

1. $\boldsymbol{\mathcal{S}}$ is block Toeplitz and its generating matrix has modulation structure, i.e., $\mathbf{S} = \mathbf{M}(n_0)\mathbf{D}$ with $\mathbf{D}$ diagonal;
2. the row span of $\boldsymbol{\mathcal{S}}$ lies in the row span of $\boldsymbol{\mathcal{X}}$.

Equivalently, $\boldsymbol{\mathcal{S}} \in \mathcal{A} \cap \mathcal{B}$ where $\mathcal{A}$ denotes the linear subspace of all block Toeplitz matrices with generating matrix $\mathbf{M}(n_0)\mathbf{D}$, where $\mathbf{M}(n_0)$ is given with n_0 arbitrary but fixed and $\mathbf{D}$ is diagonal, and $\mathcal{B}$ denotes the linear subspace of all matrices whose row span lies in the row span of $\boldsymbol{\mathcal{X}}$, i.e., of all matrices of the form $\mathbf{B}\boldsymbol{\mathcal{X}}$ with some $K(L+u-1) \times Mu$ matrix $\mathbf{B}$. This formulation again suggests a POCS algorithm for calculating $\boldsymbol{\mathcal{S}}$ that consists in alternately projecting the iterated version of the matrix $\boldsymbol{\mathcal{S}}$ onto $\mathcal{A}$ and $\mathcal{B}$.

Projection onto $\mathcal{A}$: $\boldsymbol{\mathcal{S}}$ being a linear structured matrix [14], it can be shown that the projection onto $\mathcal{A}$ can be performed by the following two steps:

Step 1: Enforce block Toeplitz property. Let $\boldsymbol{\mathcal{S}}^{(i-1)}$ be the result of the previous iteration (projection onto $\mathcal{B}$). From $\boldsymbol{\mathcal{S}}^{(i-1)}$, which is not block Toeplitz, we calculate a $K \times (N+L-1)$ "pseudo generating matrix" $\tilde{\mathbf{S}}^{(i-1)}$ as follows. The first one of the K rows of $\tilde{\mathbf{S}}^{(i-1)}$ is obtained by averaging properly aligned and zero-padded versions of the first, $(K+1)$st, $(2K+1)$st, etc. rows of $\boldsymbol{\mathcal{S}}^{(i-1)}$. More precisely, the first row of $\boldsymbol{\mathcal{S}}^{(i-1)}$ is shifted to the right by one position and added to the $(K+1)$st row of $\boldsymbol{\mathcal{S}}^{(i-1)}$, with zeros appended where necessary. The result is again shifted to the right by one position and added to the $(2K+1)$st row of $\boldsymbol{\mathcal{S}}^{(i-1)}$, etc. Finally the jth element of the resulting row vector of length $N+L-1$ is divided by the jth element of $(1, 2, \cdots, K, K, \cdots, K, K-1, \cdots, 1)$ to yield the first row of $\tilde{\mathbf{S}}^{(i-1)}$. The second row of $\tilde{\mathbf{S}}^{(i-1)}$ is obtained similarly by averaging properly aligned and zero-padded versions of the second, $(K+2)$nd, $(2K+2)$nd, etc. rows of $\boldsymbol{\mathcal{S}}^{(i-1)}$. In this way, all K rows of $\tilde{\mathbf{S}}^{(i-1)}$ are obtained.

Step 2: Enforce modulation structure. Next, we form $\mathbf{S}^{(i)} = \mathbf{M}(n_0)\mathbf{D}^{(i)}$, where the nonzero (diagonal) elements of $\mathbf{D}^{(i)}$ can be shown to be given by

$$\left(\mathbf{D}^{(i)}\right)_{n,n} = \frac{1}{K} \sum_{k=1}^{K} \left(\tilde{\mathbf{S}}^{(i-1)}\right)_{k,n} \left(\mathbf{M}(n_0)\right)^*_{k,n} .$$

We then form the block Toeplitz matrix $\boldsymbol{\mathcal{S}}^{(i)}$ generated by $\mathbf{S}^{(i)}$.

Projection onto $\mathcal{B}$: The projection onto $\mathcal{B}$ amounts to forming $\boldsymbol{\mathcal{S}}^{(i)} = \mathbf{B}^{(i)}\boldsymbol{\mathcal{X}}$, where it can be shown that

$$\mathbf{B}^{(i)} = \boldsymbol{\mathcal{S}}^{(i-1)}\boldsymbol{\mathcal{X}}^{\#} .$$

Here, $\boldsymbol{\mathcal{S}}^{(i-1)}$ is the result of the previous iteration (projection onto $\mathcal{A}$). Note that the pseudo-inverse $\boldsymbol{\mathcal{X}}^{\#}$ need only be calculated once at the beginning of the iterative procedure.

The POCS algorithm is guaranteed to converge to an intersection point, i.e., $\boldsymbol{\mathcal{S}}^{(\infty)} \in \mathcal{A} \cap \mathcal{B}$ [11]. Thus, $\boldsymbol{\mathcal{S}}^{(\infty)} = c\boldsymbol{\mathcal{S}}$ with $c \in \mathbb{C}$. Convergence speed depends on the initialization, $\boldsymbol{\mathcal{S}}^{(0)}$. In the semiblind case, some known consecutive input samples (symbols) $s[n_1], \cdots, s[n_2]$ can be used to construct a matrix $\boldsymbol{\mathcal{S}}^{(0)}$ with generating matrix $\mathbf{S}^{(0)} = \mathbf{M}(n_0)\mathbf{D}^{(0)}$ where $\mathbf{D}^{(0)} = \text{diag}\{0, \cdots, 0, s[n_1], \cdots, s[n_2], 0, \cdots, 0\}$ [15]. Convergence can again be accelerated by relaxation and/or by using knowledge of the symbol alphabet.

The POCS algorithm can be modified to accommodate the case of (known) unequal subchannel lengths, and it can be shown to be robust to unknown unequal subchannel lengths.

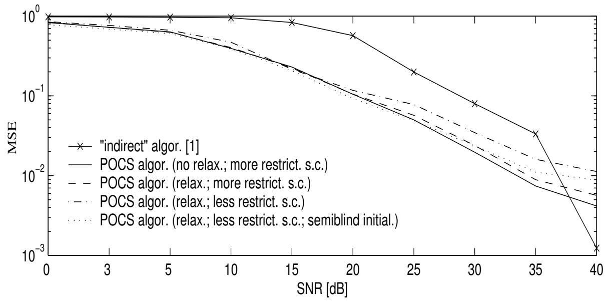

Fig. 2: MSE vs. SNR for the "indirect" algorithm presented in [1] and the POCS equalization algorithm proposed in Section 5. Note that relax. = relaxation, s. c. = stopping criterion, and initial. = initialization.

Algorithm	Megaflops
Indirect	≈370
POCS (no relax.; more restrict. s. c.)	≈160 ··· ≈350
POCS (relax.; more restrict. s. c.)	≈115 ··· ≈170
POCS (relax.; less restrict. s. c.)	≈75
POCS (relax.; less restrict. s. c.; semiblind initialization)	≈50

Table 1: Computational complexity of the "indirect" algorithm presented in [1] and the POCS equalization algorithm proposed in Section 5. Note that relax. = relaxation and s. c. = stopping criterion. Where ranges of megaflop values are given, small (large) values correspond to maximum (minimum) SNR.

6. SIMULATIONS RESULTS

In order to assess the performance of the POCS equalization algorithm proposed in Section 5, we used $M = 5$ channel output signals $x_i[n]$ (see (4)) that were corrupted by additive white Gaussian noise with variance σ^2 and observed over an interval of length $N = 115$. Each of the 5 channels comprised $K = 3$ subchannels with active Doppler shifts $l_k = 0, 1, 2$ and equal filter lengths $L = 5$. For each simulation run, the impulse responses of the subchannel filters were generated randomly.

Fig. 2 compares the mean-square error (MSE) obtained with our POCS algorithm with that obtained with the "indirect" algorithm presented in [1], as a function of the SNR.[1] Table 1 compares the computational complexity of the POCS algorithm with that of the indirect algorithm. The smoothing factor was $u = 6$ for the POCS algorithm and $u = 7$ for the indirect algorithm; in both cases, this was the minimum possible value of u. (We did not consider the "direct" algorithm of [1] since it would require computation of the eigenvector corresponding to the minimal eigenvalue of a matrix of size $J \times J$ with $J = \left[K(L+u-1)\right]^2 = 900$.) The POCS algorithm was implemented with and without relaxation [11] and with a more or less restrictive stopping criterion. We also considered semiblind initialization using 13 known samples. The number of simulation runs was chosen between 25 and 250, depending on the SNR.

[1]The MSE is defined as $\|\mathbf{s} - \hat{c}\hat{\mathbf{s}}\|^2 / \|\mathbf{s}\|^2$ averaged over all simulation runs, where $\mathbf{s} = \left[s[-L+1]\ s[-L+2]\ \cdots\ s[N-1]\right]^T$, $\hat{\mathbf{s}}$ is the estimate of $\mathbf{s}$ obtained with the respective method, and $\hat{c}$ is the least-squares fit for the unknown factor c. The SNR is defined as $\left\|\left[\mathbf{x}[0] \cdots \mathbf{x}[N-1]\right]\right\|_F^2 / (NM\sigma^2)$, with $\|\cdot\|_F$ the Frobenius norm. The same SNR was used for each simulation run.

Our simulation results show that the proposed POCS equalization algorithm performed significantly better (in terms of both MSE and computational complexity) than the indirect algorithm presented in [1]. Computational complexity is reduced mainly because calculation of the closed-form solution to the matrix ambiguity problem (the most expensive part of the indirect algorithm) is avoided. The reduction of MSE appears to be due to the fact that we exploit the structure of $\boldsymbol{\mathcal{S}}$ whereas the indirect algorithm exploits the structure of $\boldsymbol{\mathcal{H}}$ which seems to be weaker. We also see that relaxation reduces computational cost. Using a less restrictive stopping criterion also reduces computational cost, at the expense of some increase of MSE that is significant only at higher SNR.

APPENDIX: PROOF OF THEOREM 1

Let us write $\mathbf{M}(n_0) = \mathbf{M}$ and $\mathbf{M}(\tilde{n}_0) = \widetilde{\mathbf{M}}$ for notational simplicity. Assuming $s[n] \neq 0$, $\mathbf{D}$ is invertible and we can write (8) equivalently as $\widetilde{\mathbf{M}}\mathbf{D}' = \mathbf{A}\mathbf{M}$ with $\mathbf{D}' \triangleq \widetilde{\mathbf{D}}\mathbf{D}^{-1}$. Two consecutive rows of this equation read $\widetilde{M}_{i,k} D'_{k,k} = \sum_{j=1}^{K} A_{i,j} M_{j,k}$ and $\widetilde{M}_{i-1,k} D'_{k,k} = \sum_{j=1}^{K} A_{i-1,j} M_{j,k}$, with $i = 2, \cdots, K$ and $k = 0, \cdots, N+L-2$. Solving both equations for $D'_{k,k}$ and equating the resulting two expressions gives

$$\sum_{j=1}^{K} A_{i,j} \frac{M_{j,k}}{\widetilde{M}_{i,k}} = \sum_{j'=1}^{K} A_{i-1,j'} \frac{M_{j',k}}{\widetilde{M}_{i-1,k}} .$$

Inserting for $\mathbf{M}$ and $\widetilde{\mathbf{M}}$ yields

$$\sum_{j=1}^{K} \tilde{A}_{i,j} W_N^{(l_i - l_j)k} = \sum_{j'=1}^{K} \tilde{A}_{i-1,j'} W_N^{(l_{i-1} - l_{j'})k} \tag{9}$$

with $\tilde{A}_{i,j} \triangleq A_{i,j} W_N^{\tilde{n}_0 l_i - n_0 l_j}$. This must be valid for $i = 2, \cdots, K$ and $k = 0, \cdots, N+L-2$. Let us now consider a specific index tuple i, j and, thus, a specific $\tilde{A}_{i,j}$. Noting that the functions $f_l[k] = W_N^{lk}$ with $l \in [0, N-1]$ are linearly independent, we see that two cases may occur:

Case 1: If we can find j' such that $W_N^{(l_i - l_j)k} = W_N^{(l_{i-1} - l_{j'})k}$, then (9) implies $\tilde{A}_{i,j} = \tilde{A}_{i-1,j'}$ and hence

$$A_{i,j} \;=\; A_{i-1,j'} W_N^{n_0(l_j - l_{j'}) - \tilde{n}_0(l_i - l_{i-1})}. \tag{10}$$

But $W_N^{(l_i - l_j)k} = W_N^{(l_{i-1} - l_{j'})k}$ implies $l_i - l_j = l_{i-1} - l_{j'} + mN$ with $m \in \mathbb{Z}$ or, equivalently, $l_i - l_{i-1} = l_j - l_{j'} + mN$. Assuming that $0 \leq l_k \leq \lceil N/2 \rceil - 1$ for all k, only $m = 0$ is possible and thus $l_i - l_{i-1} = l_j - l_{j'}$. Without loss of generality, we assume that the active Doppler shifts are ordered such that $l_k > l_{k'}$ for $k > k'$. Then $l_i - l_{i-1} > 0$ and thus $l_j - l_{j'} > 0$, which finally gives $j > j'$.

Case 2: If there is no j' such that $W_N^{(l_i - l_j)k} = W_N^{(l_{i-1} - l_{j'})k}$, (9) implies $A_{i,j} = 0$.

We conclude that for $i = 2, \cdots, K$, $A_{i,j}$ can be nonzero only if there exists a j' such that $j > j'$. Thus, we have $A_{i,1} = 0$ for $i = 2, \cdots, K$, i.e., apart from $A_{1,1}$ the first column of $\mathbf{A}$ equals zero. Next, $A_{3,2}$ is either zero *a priori* (Case 2) or $A_{3,2}$ equals $A_{2,1} = 0$ up to a phase factor (Case 1; cf. (10)), so $A_{3,2} = 0$. Continuing this line of reasoning, it is seen that all of the lower diagonals are zero. Similarly, we have $A_{i-1,K} = 0$ for $i = 2, \cdots, K$, i.e., apart from $A_{K,K}$ the last column of $\mathbf{A}$ equals zero. Arguing similarly as above, this implies that all of the upper diagonals are zero. Thus, $\mathbf{A}$ is diagonal. On the diagonal, (10) becomes

$$A_{i,i} = A_{i-1,i-1} W_N^{n_0(l_i - l_{i-1}) - \tilde{n}_0(l_i - l_{i-1})} \tag{11}$$

(note that the condition $l_i - l_{i-1} = l_i - l_{j'}$ is satisfied with $j' = i-1$). Writing $A_{1,1} = c\, W_N^{(n_0 - \tilde{n}_0)l_1}$ with some $c \in \mathbb{C}$ (which is always possible), (11) yields $\mathbf{A} = c\,\mathrm{diag}\{W_N^{(n_0 - \tilde{n}_0)l_1}, W_N^{(n_0 - \tilde{n}_0)l_2}, \cdots, W_N^{(n_0 - \tilde{n}_0)l_K}\}$. With this $\mathbf{A}$, we have $\mathbf{A}\mathbf{M}(n_0) = c\mathbf{M}(\tilde{n}_0)$. Using this in (8) leads to $c\mathbf{M}(\tilde{n}_0)\mathbf{D} = \mathbf{M}(\tilde{n}_0)\widetilde{\mathbf{D}}$ and thus to $\widetilde{\mathbf{D}} = c\mathbf{D}$.

REFERENCES

[1] H. Liu and G. B. Giannakis, "Deterministic approaches for blind equalization of time-varying channels with antenna arrays," *IEEE Trans. Signal Processing*, vol. 46, pp. 3003–3013, Nov. 1998.

[2] A. J. van der Veen, S. Talwar, and A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 173–190, Jan. 1997.

[3] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Comm. Syst.*, vol. 11, pp. 360–393, 1963.

[4] G. B. Giannakis and C. Tepedelenlioğlu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, Oct. 1998.

[5] R. S. Kennedy, *Fading Dispersive Communication Channels.* New York: Wiley, 1969.

[6] H. Artés, G. Matz, and F. Hlawatsch, "Unbiased scattering function estimation during data transmission," in *Proc. IEEE VTC-99 Fall*, (Amsterdam, The Netherlands), pp. 1535–1539, Sept. 1999.

[7] C. Tepedelenlioğlu and G. B. Giannakis, "Precoding for scattering function estimation of mobile channels using output correlations only," in *Proc. 37th Allerton Conf. on Commun., Contr., Comput.*, Sept. 1999.

[8] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 516–525, 1995.

[9] H. Liu and G. Xu, "Closed form blind symbol estimation in digital communications," *IEEE Trans. Signal Processing*, vol. 43, pp. 2714–2723, Nov. 1995.

[10] T. Kailath, "Measurements on time-variant communication channels," *IEEE Trans. Inf. Theory*, vol. 8, pp. 229–236, Sept. 1962.

[11] P. L. Combettes, "The foundations of set theoretic estimation," *Proc. IEEE*, vol. 81, pp. 181–208, Feb. 1993.

[12] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array — Part I: Algorithms," *IEEE Trans. Signal Processing*, vol. 44, pp. 1184–1197, May 1996.

[13] H. Liu and G. Xu, "Multiuser blind channel estimation and spatial channel pre-equalization," in *Proc. IEEE ICASSP-95*, (Detroit (MI)), pp. 1756–1759, May 1995.

[14] J. A. Cadzow, "Signal enhancement–A composite property mapping algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 49–62, Jan. 1988.

[15] J. Laurila, K. Kopsa, R. Schürhuber, and E. Bonek, "Semi-blind separation and detection of co-channel signals," in *Proc. IEEE Int. Conf. Commun. (ICC)*, (Vancouver B.C., Canada), pp. 17–22, June 1999.