

# Constrained Optimization Methods for Direct Blind Equalization

Michail K. Tsatsanis, *Member, IEEE*, and Zhengyuan (Daniel) Xu, *Student Member, IEEE*

**Abstract**—Constrained optimization techniques are studied in this paper for direct design of linear multichannel equalizers. Novel blind algorithms are derived by minimizing the equalizer's output variance subject to appropriate constraints. The constraints are chosen to guarantee no desired signal cancellation, and their parameters are jointly optimized to maximize the signal component at the output. The resulting blind algorithm was observed to have near optimal performance at high signal-to-noise ratio, i.e., close to the performance of the trained minimum mean-square-error receiver. Also, the proposed method is not sensitive to the color of the transmitted sequence. Analytical expressions are derived to quantify the algorithm's performance.

**Index Terms**—Blind equalization, constrained optimization, multichannel equalization.

## I. INTRODUCTION

MULTIPATH effects and intersymbol interference have always presented major obstacles in high-speed transmission through a wireless medium. The high data rates of current cellular time division multiple access (TDMA) systems (e.g., [1]) further complicate the problem, requiring sophisticated channel equalization and signal processing techniques at the receiver [1]. Diversity combining is also typically employed, where possible, to guard against channel fading [19].

While both channel equalization and diversity combining are hardly new ideas in digital communications, their synergies and consequent implications had not been fully explored until recently. The pioneering work of Tong *et al.* [23], initiated a renewed interest in single-input-multiple-output (SIMO) problems by showing that when multiple channels are available, they can be blindly estimated using only second-order statistics. The possible introduction of diversity through fractional sampling at the receiver has also received considerable attention [22], [23], [29].

One might argue that the SIMO setup is inadequate to describe the cochannel interference which is present in many wireless applications. However, the SIMO framework is directly applicable to situations where cochannel interference is weak and can be neglected, e.g., in cellular systems with power control and large cluster size. Furthermore, the SIMO

scenario has great theoretical importance. It is usually the first step in developing algorithms which may be later extended to the multiple-input-multiple-output (MIMO) case. It is therefore not surprising that the SIMO equalization problem has attracted significant attention.

Most of the work in this area has focused on channel estimation [11], [18], [23] for the case where no training data are available. For many applications, the feasibility and/or desirability of a completely self-recovering system has yet to be determined. However, the techniques developed in this research may be combined with classical trained equalization methods to yield improved semiblind systems [3], [6], [9].

An added disadvantage of some blind channel equalization techniques is that the equalizer parameters are not provided directly as in the case of trained or constant modulus algorithm (CMA) based equalization [4], [24]. They rather rely on a two-step procedure of estimating the channel first and then constructing the equalizer based on the channel estimates. Since the channel estimation errors are not taken into account in the second step, the resulting equalizer is not optimal. For this reason, recent efforts have focused on direct design of the equalizer [7], [8].

The goal of this paper is to derive optimal solutions for the direct equalizer design problem with performance close to that of the trained minimum mean-square-error (MMSE) equalizer. The proposed method is based on constrained optimization ideas widely used in array processing and beamforming problems [30].

There have been many instances of fruitful exchange of ideas between the array processing and system/channel identification areas. In particular, the multichannel deconvolution problem has strong analogies with the beamforming and direction of arrival (DOA) problems [19]. In fact, one of the most successful blind estimation techniques, the subspace method proposed by Moulines *et al.* [18], borrowed ideas from corresponding subspace DOA algorithms (e.g., [21]). However, this analogy has not been fully explored, and a host of techniques developed in the context of adaptive beamforming have not been exploited.

In this paper, we apply constrained optimization ideas on the multichannel equalization problem. These ideas were originally developed in the design of minimum variance distortionless response (MVDR) and Capon beamformers (e.g., [14]). We study the applicability of these methods to the current problem and derive novel, blind equalization algorithms. We optimize their performance and establish appropriate identifiability conditions. Based on these conditions, we study the

Manuscript received August 26, 1997; revised January 11, 1999. This work was supported in part by the National Science Foundation Grant NSF-NCR 9706658, NSF-CCR CAREER 9733048, and the Army Research Office Grant DAAG55-98-1-0224.

The authors are with the Electrical and Computer Engineering Department, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: mtsatsan@stevens-tech.edu).

Publisher Item Identifier S 0733-8716(99)02828-0.

performance of the blind method at high signal-to-noise ratio (SNR) and compare it to that of the trained MMSE equalizer.

Similar to subspace methods (e.g., [18]), our approach requires no common zeros among the multiple channels. In addition, it requires knowledge of the channel order  $q$ . Papers that estimate the channel order can be found in [10] and [17].

Constrained optimization methods have been used successfully (although in a different context) for interference suppression in multi-user code division multiple access (CDMA) systems [13], [25], [28]. However, for the current narrowband multichannel system, the case of multiple users is outside the scope of this paper.

The rest of the paper is organized as follows. In Section II, the system model and problem statement are developed. Constrained optimization methods are proposed in Section III, while their applicability and performance are studied in Section IV. The connections with existing methods are explored in Section V, while some simulation examples are provided in Section VI. Finally, conclusions are drawn in Section VII.

## II. PROBLEM STATEMENT

Let us consider a system with linear modulation, where the user transmits symbols  $w(k)$  taking values from a finite set (e.g., quadrature amplitude modulation). Then the transmitted baseband signal is<sup>1</sup>

$$s_c(t) = \sum_{k=-\infty}^{\infty} w(k)g_{c,tr}(t - kT_s) \quad (1)$$

where  $g_{c,tr}(t)$  is the transmitter's spectral shaping pulse and  $T_s$  is the symbol period.

We assume that the signal is received through  $P$  receiving antennas or other diversity channels. Oversampling at the receiver may also create diversity channels provided that excess bandwidth is available [23], [26], [29]. Let the  $i$ th antenna receive the signal through  $L$  paths with different angles  $\theta_l$  and/or delays  $\tau_l$ ,  $l = 1, \dots, L$ . Then the received baseband signal is

$$y_{c,i}(t) = \sum_{l=1}^L a_i(\theta_l)c_l \sum_{k=-\infty}^{\infty} w(k)g_c(t - \tau_l - kT_s) + v_{c,i}(t) \quad (2)$$

where  $c_l$  is the  $l$ th path's complex gain,  $a_i(\theta_l)$  is the gain of the  $i$ th antenna for direction  $\theta_l$ , and  $g_c(t)$  is the impulse response of the combined transmitter and receiver filters. Finally  $v_{c,i}(t)$  is Gaussian additive noise. All  $c_l$ 's are assumed constant, which corresponds to a system with channel coherence time larger than the packet period.

If we collect the signals from all antennas in a vector  $\mathbf{y}_{P,c}(t) = [y_{c,1}(t), \dots, y_{c,P}(t)]^T$ , then from (2) we obtain

$$\mathbf{y}_{P,c}(t) = \sum_{k=-\infty}^{\infty} w(k)\mathbf{h}_{P,c}(t - kT_s) + \mathbf{v}_{P,c}(t) \quad (3)$$

<sup>1</sup>We use subscript  $c$  to denote continuous-time signals.

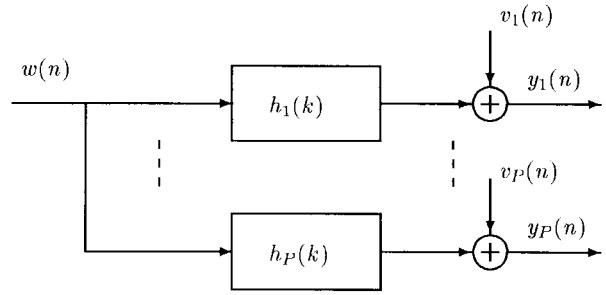


Fig. 1. Discrete-time multichannel model.

where  $\mathbf{v}_{P,c}(t) = [v_{c,1}(t), \dots, v_{c,P}(t)]^T$  and

$$\begin{aligned} \mathbf{h}_{P,c}(t) &= \sum_{l=1}^L \mathbf{a}(\theta_l)g_c(t - \tau_l), \\ \mathbf{a}(\theta_l) &= c_l[a_{1l}(\theta_l), \dots, a_{Pl}(\theta_l)]^T. \end{aligned} \quad (4)$$

Finally, if the received signal is sampled at the symbol rate,  $\mathbf{y}_P(n) = \mathbf{y}_{P,c}(t)|_{t=nT_s}$  then we arrive at<sup>2</sup>

$$\mathbf{y}_P(n) = \sum_{k=-\infty}^{\infty} w(k)\mathbf{h}_P(n - k) + \mathbf{v}_P(n) \quad (5)$$

where  $\mathbf{h}_P(n) = \mathbf{h}_{P,c}(t)|_{t=nT_s}$ ,  $\mathbf{v}_P(n) = \mathbf{v}_{P,c}(t)|_{t=nT_s}$ . The discrete-time impulse response model of (5) corresponds to a SIMO system as shown in Fig. 1. The following assumptions will be imposed on the model of (5) for the rest of the paper:

- AS1)  $w(k)$  is an independently identically distributed (i.i.d.) zero-mean sequence taking finite complex values;
- AS2)  $\mathbf{v}_P(n)$  is white Gaussian, zero-mean noise with covariance matrix  $\sigma_v^2\mathbf{I}$ ;
- AS3)  $\mathbf{h}_P(k)$  is a multichannel finite impulse response (FIR) with order  $q$ ;
- AS4) the  $z$ -transforms  $H_i(z)$  of the impulse responses  $h_i(k)$ ,  $i = 1, \dots, P$  have no common zeros.

AS1)–AS3) are common assumptions in communications problems, although in some cases they may not be satisfied. AS4) is also important in multichannel equalization and has been the subject of past studies [29].

The goal of this paper is to design a multichannel equalizer based on the model of (5). If we focus on FIR linear equalizers, the problem is equivalent to determining a vector  $\mathbf{f}$  which, when operating on the received data, provides an estimate of the transmitted signal  $\hat{w}(n - d)$  with a possible delay of  $d$  samples. Let  $\mathbf{f}$  have length  $MP$  and let us collect  $M$  received vectors in  $\mathbf{y}(n) = [\mathbf{y}_P^T(n), \mathbf{y}_P^T(n-1), \dots, \mathbf{y}_P^T(n-M+1)]^T$ ; then we consider the equalizer

$$\mathbf{f}^H \mathbf{y}(n) = \hat{w}(n - d). \quad (6)$$

If training samples are available, the design of  $\mathbf{f}$  can be formulated as a standard Wiener estimation problem and solved using well-known batch or adaptive methods (e.g.,

<sup>2</sup>Oversampling of the received signal can easily be incorporated in this framework (e.g., [19]). Details are omitted here for the sake of clarity of presentation.

[12]). In the absence of training samples, however, the problem becomes more challenging.

In this paper, we explore constrained optimization methods to derive optimal blind equalizers. We integrate ideas from beamforming and array processing problems, which share common features with the current formulation of (5) and (6). In order to further reveal these analogies, let us use (5) to write  $\mathbf{y}(n)$  as

$$\mathbf{y}(n) = \mathcal{T}(\mathbf{h})\mathbf{w}(n) + \mathbf{v}(n) \quad (7)$$

where

$$\mathcal{T}(\mathbf{h}) = \begin{bmatrix} \mathbf{h}_P(0) & \cdots & \mathbf{h}_P(q) & \cdots & \mathbf{0} \\ \vdots & \ddots & & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{h}_P(0) & \cdots & \mathbf{h}_P(q) \end{bmatrix} \quad (8)$$

is a  $(PM) \times (M+q)$  block Toeplitz matrix,  $\mathbf{w}(n) = [w(n), w(n-1), \dots, w(n-M-q+1)]^T$ , and  $\mathbf{v}(n) = [\mathbf{v}_P^T(n), \mathbf{v}_P^T(n-1), \dots, \mathbf{v}_P^T(n-M+1)]^T$ . The description of (6) and (7) emphasizes the analogies with beamforming applications where we wish to suppress interference from other ‘‘sources.’’ Recall that the standard beamforming problem is defined by

$$\mathbf{y}_P(n) = \mathbf{A}\mathbf{w}(n) + \mathbf{v}_P(n) \quad (9)$$

where the vector of  $\mathbf{w}(n)$  represents all sources and the columns of  $\mathbf{A}$  correspond to the signatures of those sources. By comparing (7) with (9), it is clear that suppression of intersymbol interference in (7) is analogous to suppression of interfering sources in (9). This description will be exploited in the following to derive blind equalizers.

### III. CONSTRAINED OPTIMIZATION METHODS

In certain array processing applications the direction of arrival and signature of the user of interest is known or estimated [i.e.,  $\mathbf{A}$  is known in (9)]. In that case, a popular beamforming approach is to minimize the array output variance while constraining the response of the array to the user of interest to a constant (e.g., [14]). In this way, the MVDR beamformer suppresses interfering users without impeding the signal of interest.

Given the similarities of (7) and (9), one might wonder whether such constrained optimization ideas are applicable to the current framework. Here, the signature of each contributing symbol (or ‘‘source’’) is the corresponding column of  $\mathcal{T}(\mathbf{h})$ . Let us denote by  $\mathbf{h}_i$  the  $i$ th column of  $\mathcal{T}(\mathbf{h})$ , i.e.,

$$\mathcal{T}(\mathbf{h}) = [\mathbf{h}_1, \dots, \mathbf{h}_{M+q}]. \quad (10)$$

Then, the MVDR principle would translate into the following optimization problem

$$\begin{aligned} \min_{\mathbf{f}} E\{|\hat{w}(n-d)|^2\} &= \min_{\mathbf{f}} \mathbf{f}^H \mathbf{R}_y \mathbf{f} \\ \text{subject to } \mathbf{f}^H \mathbf{h}_{d+1} &= 1 \end{aligned} \quad (11)$$

where  $\mathbf{R}_y = E\{\mathbf{y}(n)\mathbf{y}^H(n)\}$ . The solution can be derived in closed form using Lagrange multipliers and is given by (e.g., [30])

$$\mathbf{f}_{\text{MV}} = \left( \mathbf{h}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{h}_{d+1} \right)^{-1} \mathbf{R}_y^{-1} \mathbf{h}_{d+1} \quad (12)$$

while the minimum output variance achieved is

$$\text{MV}(\mathbf{h}_{d+1}) = \left( \mathbf{h}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{h}_{d+1} \right)^{-1}. \quad (13)$$

Notice that the solution (12) differs from the optimal MMSE equalizer by a scalar multiple only (e.g., [12]), since  $\mathbf{h}_{d+1}$  is a multiple of the cross-correlation vector  $E\{w^*(n-d)\mathbf{y}(n)\}$ . Hence, the MVDR and MMSE equalizers are expected to exhibit identical performance with respect to output signal-to-interference-plus-noise ratio (SINR).

The optimal performance of MVDR beamformers explains their popularity. It should be stressed however, that they do not represent blind solutions since they require knowledge of the signature  $\mathbf{h}_{d+1}$ . Extensions of MVDR ideas to the case of unknown signatures were investigated by Capon providing blind solutions [2]. His approach was to design an MVDR solution for every possible (hypothesized) angle of arrival; then select the one with the maximum output variance. The rationale for this max/min approach is that one should seek to maximize the signal component at the output (after the interference has been suppressed).

In the current setup the signature  $\mathbf{h}_{d+1}$  is given by [cf. (8)]

$$\mathbf{h}_{d+1} = [0, \dots, 0, \mathbf{h}_P^T(q), \dots, \mathbf{h}_P^T(0), 0, \dots, 0]^T \quad (14)$$

for  $M \geq d+1 \geq q+1$ . Notice that  $\mathbf{h}_{d+1}$  is linearly parameterized by the channel coefficients

$$\mathbf{h} = [\mathbf{h}_P^T(q), \dots, \mathbf{h}_P^T(0)]^T \quad (15)$$

i.e., it can be written as

$$\mathbf{h}_{d+1} = \mathbf{C}_{d+1} \mathbf{h}, \quad \mathbf{C}_{d+1} = \begin{bmatrix} \mathbf{0}_{P(d-q) \times P(q+1)} \\ \mathbf{I}_{P(q+1) \times P(q+1)} \\ \mathbf{0}_{P(M-d-1) \times P(q+1)} \end{bmatrix} \quad (16)$$

where  $\mathbf{I}$  denotes the identity matrix. Hence, a Capon equalizer in the current setup would be the solution to the following optimization problem:

$$\begin{aligned} \max_{\substack{\hat{\mathbf{h}}_{d+1} = \mathbf{C}_{d+1} \hat{\mathbf{h}} \\ \|\hat{\mathbf{h}}_{d+1}\| = 1}} \left[ \min_{\mathbf{f}} \mathbf{f}^H \mathbf{R}_y \mathbf{f} \right] \quad \text{subject to } \mathbf{f}^H \hat{\mathbf{h}}_{d+1} &= 1 \end{aligned} \quad (17)$$

where we use ‘‘hat’’ to denote the dummy variables of optimization. Substituting the optimal solution (13) in (17) we obtain the equivalent problem

$$\max_{\hat{\mathbf{h}}_{d+1} = \mathbf{C}_{d+1} \hat{\mathbf{h}}} \frac{\hat{\mathbf{h}}_{d+1}^H \hat{\mathbf{h}}_{d+1}}{\hat{\mathbf{h}}_{d+1}^H \mathbf{R}_y^{-1} \hat{\mathbf{h}}_{d+1}} \quad (18)$$

where we eliminated the constraint  $\|\hat{\mathbf{h}}_{d+1}\| = 1$  by multiplying the cost function with  $\hat{\mathbf{h}}_{d+1}^H \hat{\mathbf{h}}_{d+1}$  and making it insensitive to the length of  $\hat{\mathbf{h}}_{d+1}$ . By further manipulating (18) we obtain

$$\begin{aligned} \hat{\mathbf{h}}_{\text{capon}} &= \arg \max_{\hat{\mathbf{h}}} \frac{\hat{\mathbf{h}}^H \mathbf{C}_{d+1}^H \mathbf{C}_{d+1} \hat{\mathbf{h}}}{\hat{\mathbf{h}}^H \mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1} \hat{\mathbf{h}}} \\ &= \arg \min_{\hat{\mathbf{h}}} \frac{\hat{\mathbf{h}}^H \mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1} \hat{\mathbf{h}}}{\hat{\mathbf{h}}^H \hat{\mathbf{h}}} \end{aligned} \quad (19)$$

where the last equality is due to the fact that  $\mathbf{C}_{d+1}^H \mathbf{C}_{d+1} = \mathbf{I}$  [cf. (16)]. The cost function in (19) is a Rayleigh quotient,

TABLE I  
PROPOSED ALGORITHM

Step 1:	Estimate $\mathbf{R}_y$ and compute $\mathbf{A} = \mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1}$ .
Step 2:	Compute $\hat{\mathbf{h}}$ as the eigenvector of $\mathbf{A}$ corresponding to the minimum eigenvalue
Step 3:	Compute the equalizer using $\mathbf{f} = \mathbf{R}_y^{-1} \mathbf{C}_{d+1} \mathbf{A}^{-1} \hat{\mathbf{h}}$ .

and hence the solution is the eigenvector corresponding to the minimum eigenvalue of  $\mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1}$ . Finally, once  $\hat{\mathbf{h}}_{\text{capon}}$  has been determined, the Capon equalizer is given by (12).

A related algorithm can be derived from a different viewpoint through an optimization problem with multiple constraints. Consider the optimization problem

$$\min_{\mathbf{f}} \mathbf{f}^H \mathbf{R}_y \mathbf{f} \quad \text{subject to } \mathbf{C}_{d+1}^H \mathbf{f} = \hat{\mathbf{h}} \quad (20)$$

where  $\hat{\mathbf{h}}$  is a parameter vector to be determined. The constraints of (20) also guarantee a constant response for the signal of interest

$$\mathbf{f}^H \mathbf{h}_{d+1} = \mathbf{f}^H \mathbf{C}_{d+1} \mathbf{h} = \hat{\mathbf{h}}^H \mathbf{h} = \text{constant}. \quad (21)$$

Similar techniques in array processing are referred as ‘‘generalized sidelobe cancellers.’’ The solution to (20) is (e.g., [14])

$$\mathbf{f}_{\text{GSC}} = \mathbf{R}_y^{-1} \mathbf{C}_{d+1} (\mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1})^{-1} \hat{\mathbf{h}} \quad (22)$$

while the minimum achieved variance is

$$\text{MV}_{\text{GSC}}(\hat{\mathbf{h}}) = \hat{\mathbf{h}}^H (\mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1})^{-1} \hat{\mathbf{h}}. \quad (23)$$

Notice that in this case too, the minimum output variance depends on  $\hat{\mathbf{h}}$ , hence the same max/min ideas are applicable. Therefore,  $\hat{\mathbf{h}}_{\text{GSC}}$  may be determined as the vector maximizing the minimum variance of (23)

$$\hat{\mathbf{h}}_{\text{GSC}} = \arg \max_{\hat{\mathbf{h}}} \frac{\hat{\mathbf{h}}^H [\mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1}]^{-1} \hat{\mathbf{h}}}{\hat{\mathbf{h}}^H \hat{\mathbf{h}}}. \quad (24)$$

The cost function in (24) is also a Rayleigh quotient and in fact the solution is identical to that of (19). The steps of the proposed algorithm are summarized in Table I. This algorithm presents a complete solution for the design of batch linear equalizers. Adaptive versions are certainly possible, and they are the focus of our current research. However, details on adaptive implementation are outside the scope of the present paper.

Capon and generalized sidelobe canceller (GSC) beamformers are known to provide good performance in array processing applications. The applicability of such approaches in the current framework, however, is not clear from the discussion up to this point. In Section IV, we investigate the conditions under which this approach is viable and we study its performance. It turns out that under standard assumptions AS1)–AS4) the proposed method can successfully equalize the received signal with a performance which is asymptotically close to that of the MMSE receiver.

#### IV. PERFORMANCE ANALYSIS

We consider the output SINR as a performance measure for our analysis. Here the interference comes from other interfering symbols rather than from other users. In our simulation studies we observed an SINR performance close to that of the MMSE equalizer for a wide range of SNR (see Section VI). We have managed to back up these observations with analytical results for the high SNR region. In particular, in the following we show that as  $\text{SNR} \rightarrow \infty$ :

- the parameter vector  $\hat{\mathbf{h}}_{\text{capon}}$  obtained from (19) converges to the channel vector  $\mathbf{h}$ ;
- the ratio  $\text{SINR}_{\text{mse}}/(\text{SINR}(\hat{\mathbf{h}}_{\text{capon}}))$  converges to a constant  $1 + \delta$ .

In this section, we derive an explicit expression for the excess penalty  $\delta$ . Let us first investigate the behavior of  $\hat{\mathbf{h}}_{\text{capon}}$  as  $\text{SNR} \rightarrow \infty$ . We start by establishing the limit of the matrix  $\mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1}$  involved in (19) as  $\text{SNR} \rightarrow \infty$ .

The major difficulty in this task comes from the fact that  $\mathbf{R}_y^{-1}$  is not well defined if  $\sigma_v^2 = 0$  ( $\mathbf{R}_y$  loses rank). We therefore involve the eigendecomposition of  $\mathbf{R}_y$

$$\mathbf{R}_y = [\mathbf{V}_s \quad \mathbf{V}_n] \begin{bmatrix} \boldsymbol{\Lambda}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_s^H \\ \mathbf{V}_n^H \end{bmatrix} + \sigma_v^2 \mathbf{I} \quad (25)$$

where  $\boldsymbol{\Lambda}_s = \text{diag}\{\lambda_1, \dots, \lambda_{M+q}\}$  and  $\mathbf{V}_s, \mathbf{V}_n$  represents the signal and noise subspaces, respectively. Using (25),  $\mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1}$  can be expressed as an explicit function of  $\sigma_v^2$  as follows.

*Lemma 1:* It holds that

$$\sigma_v^2 \mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1} = \mathbf{C}_{d+1}^H [\mathbf{I} - \mathbf{V}_s \mathbf{D} \mathbf{V}_s^H] \mathbf{C}_{d+1} \quad (26)$$

$$= \mathbf{A}_0 + \sigma_v^2 \mathbf{A}_1 + \sigma_v^4 \mathbf{A}_2 + \mathbf{O}(\sigma_v^6) \quad (27)$$

where

$$\mathbf{D} = \text{diag} \left\{ \frac{\lambda_1}{\sigma_v^2 + \lambda_1}, \dots, \frac{\lambda_{M+q}}{\sigma_v^2 + \lambda_{M+q}} \right\}$$

$$\mathbf{A}_0 = \mathbf{C}_{d+1}^H \mathbf{V}_n \mathbf{V}_n^H \mathbf{C}_{d+1} \quad (28)$$

$$\mathbf{A}_1 = \mathbf{C}_{d+1}^H \mathbf{V}_s \boldsymbol{\Lambda}_s^{-1} \mathbf{V}_s^H \mathbf{C}_{d+1}$$

$$\mathbf{A}_2 = -\mathbf{C}_{d+1}^H \mathbf{V}_s \boldsymbol{\Lambda}_s^{-2} \mathbf{V}_s^H \mathbf{C}_{d+1}. \quad (29)$$

□

*Proof:* The proof relies on applying the matrix inversion lemma (e.g., [12]) to (25) to obtain

$$\begin{aligned} \mathbf{R}_y^{-1} &= \frac{1}{\sigma_v^2} \mathbf{I} - \frac{1}{\sigma_v^2} \mathbf{V}_s (\sigma_v^2 \boldsymbol{\Lambda}_s^{-1} + \mathbf{V}_s^H \mathbf{V}_s)^{-1} \mathbf{V}_s^H \\ &= \frac{1}{\sigma_v^2} [\mathbf{I} - \mathbf{V}_s \mathbf{D} \mathbf{V}_s^H]. \end{aligned} \quad (30)$$

By pre- and postmultiplying (30) by  $\mathbf{C}_{d+1}^H$  and  $\mathbf{C}_{d+1}$ , we arrive at (26). By further using  $\mathbf{I} = \mathbf{V}_s \mathbf{V}_s^H + \mathbf{V}_n \mathbf{V}_n^H$  and the

fact that

$$\frac{1}{\lambda_i + \sigma_v^2} = \frac{1}{\lambda_i} - \frac{1}{\lambda_i^2} \sigma_v^2 + O(\sigma_v^4) \quad (31)$$

in (26) obtain (27).  $\square$

We can see from (27) that

$$\sigma_v^2 \mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1} \rightarrow \mathbf{C}_{d+1}^H \mathbf{V}_n \mathbf{V}_n^H \mathbf{C}_{d+1} \quad (32)$$

as  $\sigma_v^2 \rightarrow 0$ . Hence, the eigenvector minimizing (19) converges to the null eigenvector of  $\mathbf{C}_{d+1}^H \mathbf{V}_n \mathbf{V}_n^H \mathbf{C}_{d+1}$ .

Notice that the solution  $\hat{\mathbf{h}} = \mathbf{h}/\|\mathbf{h}\|$  qualifies for a null eigenvector since  $\mathbf{C}_{d+1} \hat{\mathbf{h}} = \mathbf{h}_{d+1} \in \mathbf{V}_s$  and is orthogonal to the noise subspace  $\mathbf{V}_n^H \mathbf{C}_{d+1} \hat{\mathbf{h}} = 0$ . The question that remains to be answered is whether the solution  $\hat{\mathbf{h}} = \mathbf{h}/\|\mathbf{h}\|$  is unique, i.e., whether  $\mathbf{C}_{d+1}^H \mathbf{V}_n \mathbf{V}_n^H \mathbf{C}_{d+1}$  has a unique null eigenvector, or its null subspace has dimension greater than one.

The following proposition clarifies the conditions which guarantee unique identifiability.

*Proposition 1:* Under AS1)–AS4) and the assumption

AS5a)  $M \geq 3(q+1)$

AS5b)  $M - q - 2 \geq d \geq 2q + 1$

there does not exist  $\mathbf{h}'$  independent from  $\mathbf{h}$  such that  $\mathbf{C}_{d+1} \mathbf{h}' \in \mathbf{V}_s$ .  $\square$

*Proof:* See Appendix A.

AS5b) imposes some restrictions on the choice of the lag  $d$ . According to AS5b),  $d$  is not allowed to take any of the first or last  $2q$  allowable lags, and is in this way restricted to the central area of the allowable window. However, it has been observed in simulations [7] that the best lag  $d$  is usually around the center of the window. Therefore, AS5b) does not present a severe limitation and is not expected to have adverse effects on the system's performance.

By combining the identifiability result of Proposition 1 with our convergence discussion, we conclude the following result.

*Proposition 2:* Under AS1)–AS5) and if  $\hat{\mathbf{h}}_{\text{capon}}$  is the minimizer of (19), then

$$\hat{\mathbf{h}}_{\text{capon}} \rightarrow \frac{\mathbf{h}}{\|\mathbf{h}\|} \quad \text{as } \sigma_v^2 \rightarrow 0. \quad \square$$

Proposition 2 shows that the method can be used as a channel estimator at high SNR. In the current setup however, we are more interested in a complete equalizer design by combining (19) with (12) or (22). We next investigate the SINR performance of an equalizer derived through this constrained optimization approach.

#### A. SINR Analysis

For a general linear receiver  $\mathbf{f}$ , the output SINR is defined as

$$\text{SINR} = \frac{\mathbf{f}^H \mathbf{R}_s \mathbf{f}}{\mathbf{f}^H \mathbf{R}_i \mathbf{f}} \quad (33)$$

where  $\mathbf{R}_s = \sigma_w^2 \mathbf{h}_{d+1} \mathbf{h}_{d+1}^H$  is the correlation matrix due to the signal of interest, while  $\mathbf{R}_i = \mathbf{R}_y - \mathbf{R}_s$  is the correlation matrix of the interference and noise.

Substituting the MMSE solution  $\mathbf{f}_{\text{mse}} = \sigma_w^2 \mathbf{R}_y^{-1} \mathbf{h}_{d+1}$  into (33) we obtain

$$\text{SINR}_{\text{mse}} = \frac{\sigma_w^2 \mathbf{h}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{h}_{d+1}}{1 - \sigma_w^2 \mathbf{h}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{h}_{d+1}}. \quad (34)$$

It is well known that  $\text{SINR}_{\text{mse}} \rightarrow \infty$  as  $\sigma_v^2 \rightarrow 0$  (as the MMSE solution converges to the zero forcing one). Hence, from (34) we must have  $\sigma_w^2 \mathbf{h}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{h}_{d+1} \rightarrow 1$  as  $\sigma_v^2 \rightarrow 0$ . Moreover,  $\text{SINR}(\hat{\mathbf{h}}_{\text{capon}}) \rightarrow \infty$  also as  $\sigma_v^2 \rightarrow 0$ , making the evaluation of the limit  $\lim_{\sigma_v^2 \rightarrow 0} (\text{SINR}_{\text{mse}} / \text{SINR}(\hat{\mathbf{h}}_{\text{capon}}))$  nontrivial. For this reason we use Lemma 1 to express  $\text{SINR}_{\text{mse}}$  in terms of  $\sigma_v^2$ . By pre- and postmultiplying (27) by  $\mathbf{h}^H$  and  $\mathbf{h}$ , and recalling that  $\mathbf{h}_{d+1} = \mathbf{C}_{d+1} \mathbf{h}$ , and that  $\mathbf{V}_n^H \mathbf{h}_{d+1} = 0$ , we obtain

$$\sigma_w^2 \mathbf{h}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{h}_{d+1} = \sigma_w^2 \mathbf{h}^H \mathbf{A}_1 \mathbf{h} + \sigma_v^2 \sigma_w^2 \mathbf{h}^H \mathbf{A}_2 \mathbf{h} + O(\sigma_v^4). \quad (35)$$

According to our discussion on the convergence of the left-hand side of (35) we expect  $\sigma_w^2 \mathbf{h}^H \mathbf{A}_1 \mathbf{h} = 1$ . Substituting (35) into (34) we obtain

$$\text{SINR}_{\text{mse}} = \frac{1 + O(\sigma_v^2)}{\sigma_v^2 \sigma_w^2 \mathbf{h}^H \mathbf{A}_2 \mathbf{h} + O(\sigma_v^4)}. \quad (36)$$

We now turn our attention to  $\text{SINR}(\hat{\mathbf{h}}_{\text{capon}})$ . If we substitute  $\mathbf{f}$  from (22) in (33) we obtain (37) shown at the bottom of the page. After some manipulation and using the fact that  $\mathbf{h}_{d+1} = \mathbf{C}_{d+1} \mathbf{h}$  and

$$\hat{\mathbf{h}}_{\text{capon}}^H \mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1} \hat{\mathbf{h}}_{\text{capon}} = \gamma_{\min}$$

i.e.,  $\gamma_{\min}, \hat{\mathbf{h}}_{\text{capon}}$  are the minimum eigenvalue/eigenvector of  $\mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1}$  we obtain

$$\text{SINR}(\hat{\mathbf{h}}_{\text{capon}}) = \frac{\sigma_w^2 \gamma_{\min} \|\hat{\mathbf{h}}_{\text{capon}} \mathbf{h}\|^2}{1 - \sigma_w^2 \gamma_{\min} \|\hat{\mathbf{h}}_{\text{capon}} \mathbf{h}\|^2}. \quad (38)$$

In order to express  $\text{SINR}(\hat{\mathbf{h}}_{\text{capon}})$  as a function of  $\sigma_v^2$ , we first establish the following result using perturbation theory.

*Lemma 2:* If  $\sigma_v^2 \mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1}$  is expressed as a function of  $\sigma_v^2$  as in (27), then its minimum eigenvalue/eigenvector  $\gamma_{\min}, \hat{\mathbf{h}}_{\text{capon}}$  are given by

$$\hat{\mathbf{h}}_{\text{capon}} = \frac{\mathbf{h}}{\|\mathbf{h}\|} - \sigma_v^2 \mathbf{A}_0^\dagger \mathbf{A}_1 \frac{\mathbf{h}}{\|\mathbf{h}\|} + O(\sigma_v^4) \quad (39)$$

$$\begin{aligned} \gamma_{\min} &= \frac{1}{\sigma_w^2 \|\mathbf{h}\|^2} + \sigma_v^2 \frac{\mathbf{h}^H}{\|\mathbf{h}\|} \mathbf{A}_2 \frac{\mathbf{h}}{\|\mathbf{h}\|} \\ &\quad - \sigma_v^2 \frac{\mathbf{h}^H}{\|\mathbf{h}\|} \mathbf{A}_1 \mathbf{A}_0^\dagger \mathbf{A}_1 \frac{\mathbf{h}}{\|\mathbf{h}\|} + O(\sigma_v^4) \end{aligned} \quad (40)$$

where  $\dagger$  denotes the pseudoinverse.  $\square$

$$\text{SINR}(\hat{\mathbf{h}}_{\text{capon}}) = \frac{\sigma_w^2 \|\hat{\mathbf{h}}_{\text{capon}}^H (\mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1})^{-1} \mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{h}_{d+1}\|^2}{\hat{\mathbf{h}}_{\text{capon}}^H (\mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1})^{-1} \hat{\mathbf{h}}_{\text{capon}} - \sigma_w^2 \|\hat{\mathbf{h}}_{\text{capon}}^H (\mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1})^{-1} \mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{h}_{d+1}\|^2} \quad (37)$$

*Proof:* See Appendix B.

Substituting (39) in (38) and noting that  $\hat{\mathbf{h}}_{\text{capon}}^H \mathbf{A}_0^\dagger = \mathbf{0}$  since  $\hat{\mathbf{h}}_{\text{capon}}$  is a null eigenvector of  $\mathbf{A}_0$  and  $\mathbf{A}_0^\dagger$ , we may write

$$\|\hat{\mathbf{h}}_{\text{capon}} \mathbf{h}\|^2 = \|\mathbf{h}\|^2 + O(\sigma_v^4). \quad (41)$$

Finally, substituting (41) and (40) in (38) we arrive at

$$\begin{aligned} & \text{SINR}(\hat{\mathbf{h}}_{\text{capon}}) \\ &= \frac{1 + O(\sigma_v^2)}{\sigma_v^2 \sigma_w^2 \mathbf{h}^H \mathbf{A}_2 \mathbf{h} - \sigma_v^2 \sigma_w^2 \mathbf{h}^H \mathbf{A}_1 \mathbf{A}_0^\dagger \mathbf{A}_1 \mathbf{h} + O(\sigma_v^4)}. \end{aligned} \quad (42)$$

Having managed to express the SINR as a function of  $\sigma_v^2$ , we may now combine (42) with (36) to compute the limit

$$\begin{aligned} & \lim_{\sigma_v^2 \rightarrow 0} \frac{\text{SINR}_{\text{mse}}}{\text{SINR}(\hat{\mathbf{h}}_{\text{capon}})} \\ &= \lim_{\sigma_v^2 \rightarrow 0} \left\{ \left( \frac{1 + O(\sigma_v^2)}{1 + O(\sigma_v^2)} \right) \cdot \left( \frac{\sigma_v^2 \sigma_w^2 \mathbf{h}^H \mathbf{A}_2 \mathbf{h} - \sigma_v^2 \sigma_w^2 \mathbf{h}^H \mathbf{A}_1 \mathbf{A}_0^\dagger \mathbf{A}_1 \mathbf{h} + O(\sigma_v^4)}{\sigma_v^2 \sigma_w^2 \mathbf{h}^H \mathbf{A}_2 \mathbf{h} + O(\sigma_v^4)} \right) \right\} \\ &= 1 - \frac{\mathbf{h}^H \mathbf{A}_1 \mathbf{A}_0^\dagger \mathbf{A}_1 \mathbf{h}}{\mathbf{h}^H \mathbf{A}_2 \mathbf{h}}. \end{aligned} \quad (43)$$

Equations (29) and (43) provide the proof for the following proposition.

*Proposition 3:* Under AS1)–AS5) and if  $\hat{\mathbf{h}}_{\text{capon}}$ ,  $\mathbf{f}$  are obtained from (19) and (22), respectively, then

$$\frac{\text{SINR}_{\text{mse}}}{\text{SINR}(\hat{\mathbf{h}}_{\text{capon}})} \rightarrow 1 + \delta \quad \text{as } \sigma_v^2 \rightarrow 0$$

where

$$\delta = \frac{\mathbf{h}^H \mathbf{A}_1 \mathbf{A}_0^\dagger \mathbf{A}_1 \mathbf{h}}{\mathbf{h}^H \mathbf{C}_{d+1}^H \mathbf{V}_s \Lambda_s^{-2} \mathbf{V}_s^H \mathbf{C}_{d+1} \mathbf{h}}. \quad \square$$

Our simulation results confirm the conclusion of Proposition 3. While  $\delta$  generally depends on the channel parameters, it was observed in several occasions to be a small penalty (less than 1 dB). However, the SINR was not always observed to be monotonically decreasing with  $\sigma_v^2$ . Hence, there might be room for further performance analysis results for the low SNR region. This effort however, is beyond the scope of the current paper.

## V. DISCUSSION

In this section, we would like to put our method into the context of existing subspace methods [18] by briefly discussing their differences and similarities. The method of [18] is based on the orthogonality of the desired signal with the noise subspace  $\mathbf{V}_n$ . Our method also implicitly exploits the same principle [see discussion after (32)]. Although our method does not have to explicitly estimate the signal and noise subspaces [see (19)], it still requires knowledge of the system order similar to [18]. It can be shown that the identifiability result in Proposition 1 does not hold if  $q$  is overestimated. Therefore,

the method is not applicable to cases where the channel order is unknown.

Another common feature of our method and that of [18] is that they are both insensitive to the color of the input. Indeed (30) and (32), as well as Propositions 1 and 2, hold regardless of the color of  $w(n)$ . We only use the whiteness of  $w(n)$  in Section IV to simplify the SINR analysis.

## VI. SIMULATIONS

In our simulations, a communication system with BPSK modulation and a raised cosine pulse shaping filter with  $\alpha = 0.25$  was used. The transmitted i.i.d. signal, taking values  $\{-1, +1\}$ , was passed through a three-ray multipath channels to arrive at a receiver with several antennas. Results for four up to ten antennas spaced at half wavelength ( $d = \lambda/2$ ) are presented. The channel effects were simulated according to (2). The gain for each path was set to one ( $c_l = 1$ ). The symbol period was  $T_s = 5 \mu\text{s}$ ,  $L = 3$ , and the two delayed signals arrived with delays of  $\tau_2 = 5.5 \mu\text{s}$ ,  $\tau_3 = 8 \mu\text{s}$ , respectively, resulting in a channel of length equal to 3 bits. The first arrived signal and other delayed copies reached each antenna at angles of  $\theta_1 = -30^\circ$ ,  $\theta_2 = 15^\circ$ ,  $\theta_3 = -4^\circ$ . An ideally calibrated narrowband sensor array was assumed with  $i$ th sensor response for direction  $\theta_l$  as  $a_i(\theta_l) = \cos((2\pi/\lambda)id \sin \theta_l)$ , as is common practice in array signal processing [14]. According to the identifiability conditions, we collected  $M = 9$  samples at each antenna and chose  $d = 5$ .

We implemented the proposed receiver according to Table I and compared it with the MMSE receiver in terms of output SINR for a wide range of input noise levels (from  $-10$  dB to  $30$  dB). Fig. 2 compares the performance of the proposed method with that of the MMSE receiver for 4, 6, 8, 10 antennas [Fig. 2(a)–(d), respectively] using ideal correlation estimates for both equalizers. It can be seen that the difference between  $\text{SINR}_{\text{mse}}$  and  $\text{SINR}(\hat{\mathbf{h}}_{\text{capon}})$  approaches a small constant at high SNR. Notice that the performance of both the MMSE and the proposed method improves as the number of antennas increases. However, their difference does not necessarily become smaller. For example, in Fig. 2 the proposed method agrees most closely with the MMSE when there are six antennas.

The ratio of  $\text{SINR}(\hat{\mathbf{h}}_{\text{capon}})/\text{SINR}_{\text{mse}}$  is plotted in Fig. 3 versus SNR for the case of  $P = 6$  antennas. Notice that the SINR ratio tends to the limit 0.96 which is the same as that computed according to Proposition 3 ( $\delta$  turns out to be 0.0408 in this case). The dashed line in Fig. 3 represents the theoretically predicted limit.

Next we compared the performance of our proposed batch algorithm with that of the subspace method in [18] and the direct blind equalizer in [8] with delay  $d = 0$ . The method of [8] deals with  $P$ -fractionally sampled data while our method uses samples from  $P$  antennas. In order to implement the method of [8] in the current setup, we exploit the well-known equivalence between fractional sampling and multiple antennas [23]. In particular, the method of [8] uses a data window of  $[y(nP), \dots, y(nP - P + 1)]$  of the oversampled data (see [8, (3)]). It can be shown that in our framework, this

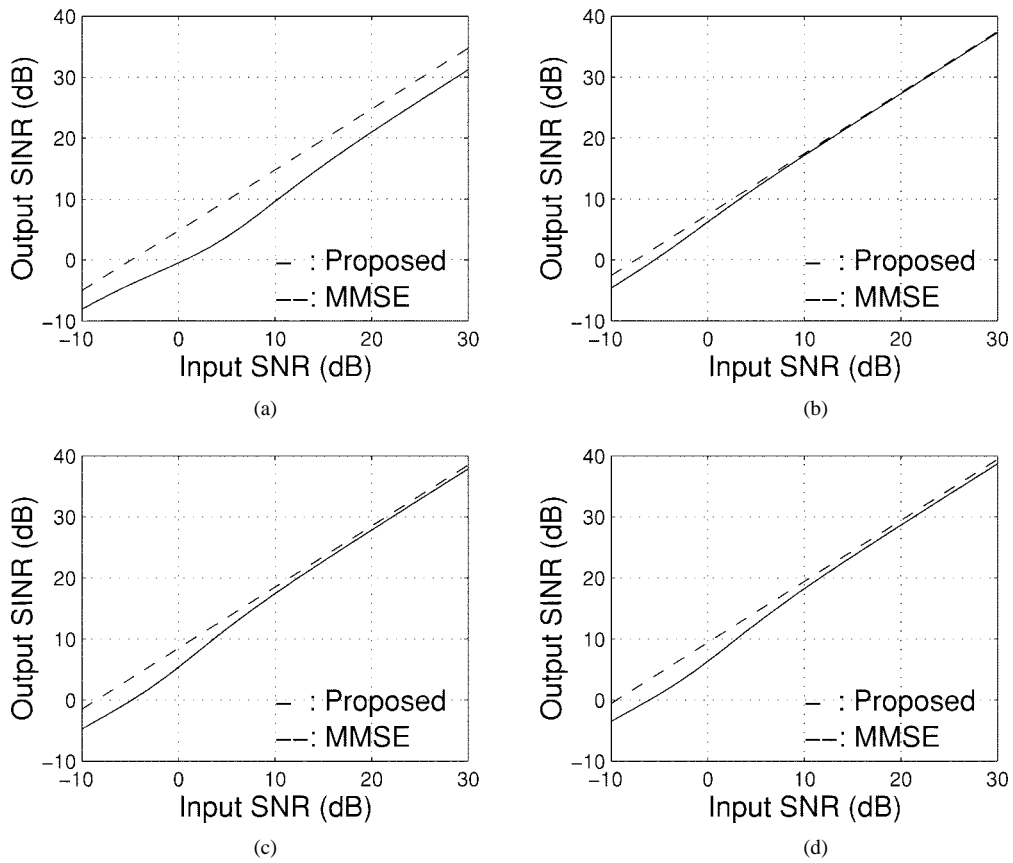


Fig. 2. SINR comparison with that of MMSE with: (a) four antennas; (b) six antennas; (c) eight antennas; and (d) ten antennas.

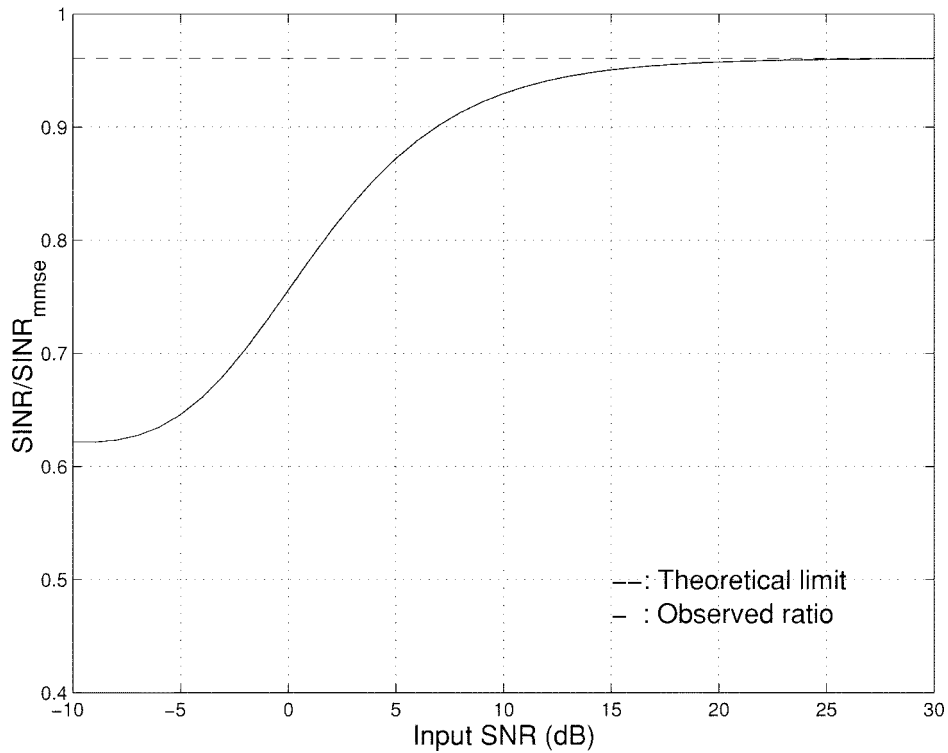


Fig. 3. Performance of ratio  $SINR/SINR_{mmse}$  with six antennas.

maps to a data window of  $[y_0(n), y_{P-1}(n-1), y_{P-2}(n-1), \dots, y_1(n-1)]$ .

The proposed equalizer was implemented according to (19) and (22). An estimate of the correlation matrix was used in

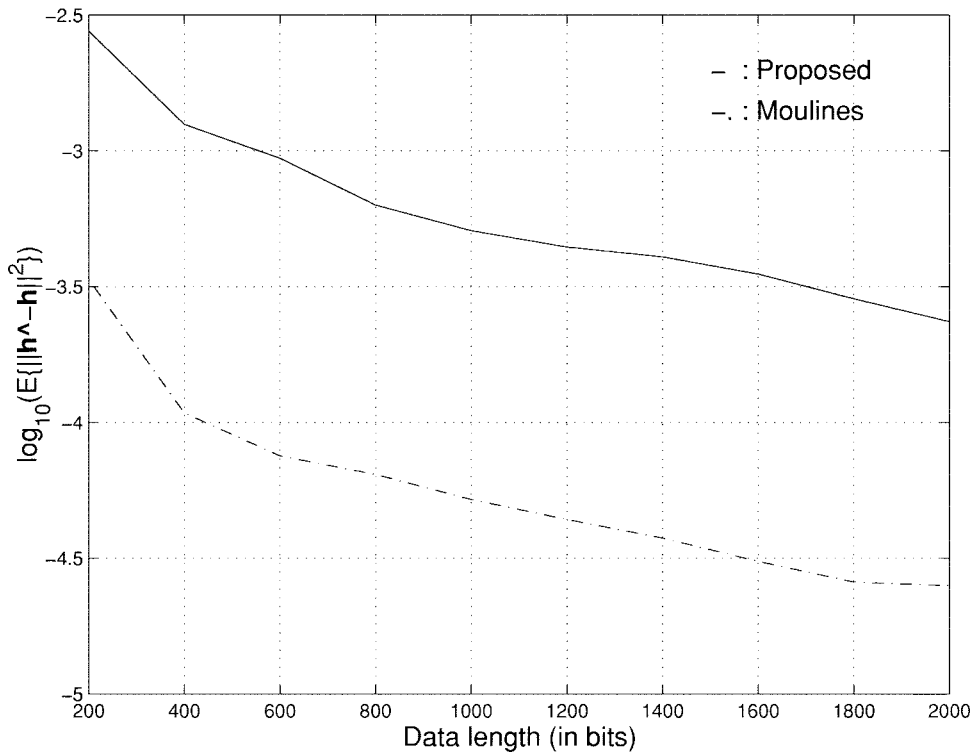


Fig. 4. Estimation error of channel vector with six antennas.

those equations obtained through sample averaging

$$\hat{\mathbf{R}}_y = \frac{1}{N} \sum_{n=1}^N \mathbf{y}(n) \mathbf{y}^H(n)$$

where  $N$  is the number of collected snapshots. In our simulation, the noise level was set to 15 dB and the number of antennas to  $P = 6$ . Figs. 4 and 5 show the average results of 200 Monte Carlo runs. Fig. 4 shows that the channel estimates obtained from [18] (dashed line) are more accurate than those from the proposed method (solid line). The mean square estimation error (log scale) is plotted in that figure versus the number of collected snapshots. Despite the better channel estimates, however, the proposed equalizer exhibits better performance than an MMSE equalizer derived from the estimated channel according to [18]. This can be verified in Fig. 5, where the output SINR of the proposed method (solid line) and that of [18] (dashed line) are plotted versus the number of collected snapshots. The method of [18] achieves worse performance, illustrating the deficiencies of a two-step approach compared with a direct equalizer design one.

Finally, we should point out that we only implemented the method of [8] for the  $d = 0$  case in Fig. 5. Extension of [8] to the  $d \neq 0$  case is not straightforward, as it requires estimates of the noiseless data correlation matrix (see [8, (33)]). In our method, the  $d \neq 0$  case does not need this estimate or any further computations.

## VII. CONCLUSIONS

The integration of array processing ideas into the equalization context has resulted in significant progress in the past [18]. It appears, though, that there is room for further

fruitful interaction between those two areas. In this paper, we exploit constrained optimization methods, widely used in array processing and beamforming problems, for developing novel blind equalizers. The resulting algorithms have near optimal performance and are not sensitive to the color of the input. Further research is needed to evaluate the performance of these methods in low SNR and develop adaptive implementations.

## APPENDIX A

### PROOF OF PROPOSITION 1

Let us decompose matrix  $\mathcal{T}(\mathbf{h})$  [see (8)] into the submatrices

$$\mathcal{T}(\mathbf{h}) = \left[ \underbrace{\mathbf{H}_1}_{d \text{ cols}} \mid \underbrace{\mathbf{h}_{d+1}}_{1 \text{ col}} \mid \underbrace{\mathbf{H}_2}_{M+q-d-1 \text{ cols}} \right] \quad (44)$$

and suppose that both  $\mathbf{C}_{d+1}\mathbf{h} \in \mathbf{V}_s$  and  $\mathbf{C}_{d+1}\mathbf{h}' \in \mathbf{V}_s$  for  $\mathbf{h} \neq \mathbf{h}'$ . Then, since  $\text{span}\{\mathbf{V}_s\} = \text{span}\{[\mathbf{H}_1 \ \mathbf{h}_{d+1} \ \mathbf{H}_2]\}$ , there exist  $\boldsymbol{\theta}'_1, \boldsymbol{\theta}'_2, \theta$  such that

$$\mathbf{C}_{d+1}\mathbf{h}' = \mathbf{H}_1\boldsymbol{\theta}'_1 + \mathbf{H}_2\boldsymbol{\theta}'_2 + \theta\mathbf{h}_{d+1}. \quad (45)$$

But since  $\mathbf{h}_{d+1} = \mathbf{C}_{d+1}\mathbf{h}$ , (45) implies

$$\mathbf{C}_{d+1}(\mathbf{h}' - \theta\mathbf{h}) = \mathbf{H}_1\boldsymbol{\theta}'_1 + \mathbf{H}_2\boldsymbol{\theta}'_2. \quad (46)$$

If  $\mathbf{h}'$  and  $\mathbf{h}$  are not collinear, then  $(\mathbf{h}' - \theta\mathbf{h}) \neq \mathbf{0}$  and (46) can only hold if the matrix  $[\mathbf{C}_{d+1} \ \mathbf{H}_1 \ \mathbf{H}_2]$  loses rank. Hence, we may complete the proof by contradiction if we establish the following lemma.

*Lemma 3:* Under AS1)–AS5),  $[\mathbf{C}_{d+1} \ \mathbf{H}_1 \ \mathbf{H}_2]$  has full rank.  $\square$

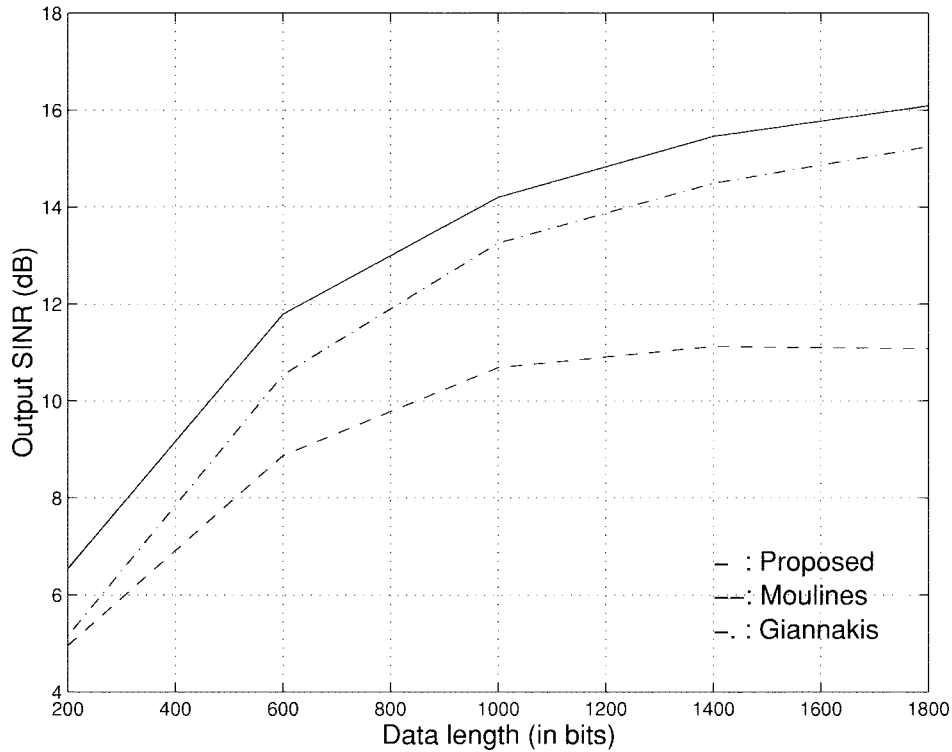


Fig. 5. Output SINR comparison for different data lengths with six antennas.

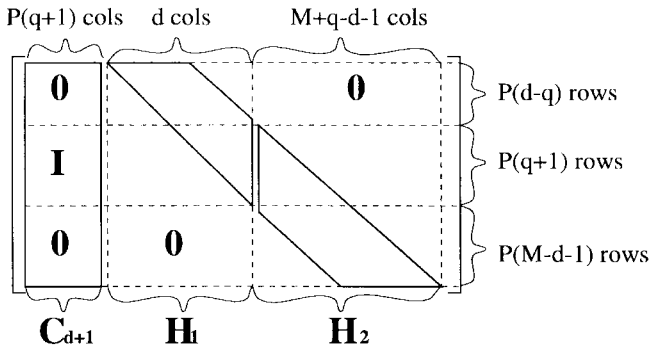


Fig. 6. Structure of matrix  $[C_{d+1} \ H_1 \ H_2]$ .

*Proof:* The proof is based on the structure of this matrix as depicted in Fig. 6, and proceeds through the following three steps.

- 1) The first  $P(d-q)$  rows of  $\mathbf{H}_1$  form a block Sylvester matrix with  $d-q$  block rows (see Fig. 6). From AS5b),  $d-q \geq q+1$ , and if AS4) is satisfied, the Sylvester matrix has full rank (e.g., [15]). Therefore  $\mathbf{H}_1$  has full rank.
- 2) Through similar reasoning the last  $P(M-d-q)$  rows of  $\mathbf{H}_2$  form a full block Sylvester matrix [notice from AS5b) that  $M-d-1 \geq q+1$ ]. Hence  $\mathbf{H}_2$  has full rank.
- 3) Matrix  $\mathbf{C}_{d+1}$  contains an identity submatrix with  $P(q+1)$  rows (see Fig. 6). Hence, the corresponding rows of  $\mathbf{H}_1, \mathbf{H}_2$  can be zeroed using column operation. This will not affect the rank of  $\mathbf{H}_1, \mathbf{H}_2$  according to 1) and 2). Hence  $[C_{d+1} \ H_1 \ H_2]$  has full rank due to its construction from nonoverlapping blocks of full rank submatrices (see Fig. 6).  $\square$

APPENDIX B

PROOF OF LEMMA 2

According to perturbation theory [20], if a matrix  $\mathbf{A}$  can be expressed as

$$\mathbf{A} = \mathbf{A}_0 + \sigma_v^2 \mathbf{A}_1 + \sigma_v^4 \mathbf{A}_2 + \dots \quad (47)$$

and  $(\gamma_0, \mathbf{v}_0), (\gamma, \mathbf{v})$  are eigenvalue/eigenvector pairs of  $\mathbf{A}_0$  and  $\mathbf{A}$ , respectively, then there exist  $\gamma_1, \gamma_2, \dots$  and  $\mathbf{v}_1, \mathbf{v}_2, \dots$  such that

$$\gamma = \gamma_0 + \sigma_v^2 \gamma_1 + \sigma_v^4 \gamma_2 + \dots \quad (48)$$

$$\mathbf{v} = \mathbf{v}_0 + \sigma_v^2 \mathbf{v}_1 + \sigma_v^4 \mathbf{v}_2 + \dots \quad (49)$$

$\forall \sigma_v^2$  in some neighborhood of  $\sigma_v^2 = 0$ . By substituting (48) and (49) in the eigenvalue problem  $\mathbf{A}\mathbf{v} = \gamma\mathbf{v}$  and equating equal powers of  $\sigma_v^2$  we obtain the following equations (cf. [5]):

$$\mathbf{A}_0 \mathbf{v}_0 = \gamma_0 \mathbf{v}_0 \quad (50)$$

$$\mathbf{A}_0 \mathbf{v}_1 + \mathbf{A}_1 \mathbf{v}_0 = \gamma_0 \mathbf{v}_1 + \gamma_1 \mathbf{v}_0 \quad (51)$$

$$\mathbf{A}_0 \mathbf{v}_2 + \mathbf{A}_1 \mathbf{v}_1 + \mathbf{A}_2 \mathbf{v}_0 = \gamma_0 \mathbf{v}_2 + \gamma_1 \mathbf{v}_1 + \gamma_2 \mathbf{v}_0. \quad (52)$$

Equation (50) offers no new information. The other two however, may be premultiplied by  $\mathbf{v}_0$  to yield

$$\gamma_1 = \mathbf{v}_0^H \mathbf{A}_1 \mathbf{v}_0 \quad (53)$$

$$\gamma_2 = \mathbf{v}_0^H \mathbf{A}_1 \mathbf{v}_1 + \mathbf{v}_0^H \mathbf{A}_2 \mathbf{v}_0 - \gamma_1 \mathbf{v}_0^H \mathbf{v}_1. \quad (54)$$

In our case, consider  $\mathbf{A} = \sigma_v^2 \mathbf{C}_{d+1}^H \mathbf{R}_y^{-1} \mathbf{C}_{d+1}$ ; then  $\mathbf{A}_0, \mathbf{A}_1,$  and  $\mathbf{A}_2$  are given by (29) and we are interested in the perturbation of the null eigenvector  $\mathbf{v}_0 = \mathbf{h}/\|\mathbf{h}\|$  corresponding to the eigenvalue  $\gamma_0 = 0$ . According to the discussion after (35), (53) yields  $\gamma_1 = 1/\sigma_w^2 \|\mathbf{h}\|^2$ . Further, it can be shown [16] that

$$\mathbf{v}_1 = -\mathbf{A}_0^\dagger \mathbf{A}_1 \mathbf{v}_0. \quad (55)$$

Equation (55) shows that  $\mathbf{v}_1^H \mathbf{v}_0 = 0$  since  $\mathbf{v}_0^H \mathbf{A}_0 = \mathbf{0}$ , and hence  $\mathbf{v}_0^H \mathbf{A}_0^\dagger = \mathbf{0}$ . Thus the last term of (54) can be ignored

$$\gamma_2 = -\mathbf{v}_0^H \mathbf{A}_1 \mathbf{A}_0^\dagger \mathbf{A}_1 \mathbf{v}_0 + \mathbf{v}_0^H \mathbf{A}_2 \mathbf{v}_0. \quad (56)$$

Substituting (56) and (55) into (48) and (49), and dividing by  $\sigma_v^2 [(1/\sigma_v^2)\mathbf{A}]$  has eigenvalue  $\gamma_{\min} = \gamma/\sigma_v^2$ , we obtain the desired result.  $\square$

## REFERENCES

- [1] G. D'Aria, F. Muratore, and V. Palestini, "Simulation and performance of the pan-European land mobile radio system," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 177–189, May 1992.
- [2] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 2408–2418, Aug. 1969.
- [3] E. De Carvalho and D. T. M. Slock, "Cramer–Rao bounds for semi-blind, blind, and training sequence based channel estimation," in *Proc. Signal Processing Advances in Wireless Communications (SPAWC'97)*, Paris, France, Apr. 16–18, 1997, pp. 129–132.
- [4] R. A. Casas, and C. R. Johnson, "On the blind adaptation of an FSE+DFE combination," in *Proc. Signal Processing Advances in Wireless Communications (SPAWC'97)*, Paris, France, Apr. 16–18, 1997, pp. 113–116.
- [5] B. Champagne, "Adaptive eigendecomposition of data covariance matrices based on first-order perturbations," *IEEE Trans. Signal Processing*, vol. 42, pp. 2758–2770, Oct. 1994.
- [6] H. A. Cirpan and M. K. Tsatsanis, "Stochastic maximum likelihood methods for semi-blind channel equalization," *IEEE Signal Processing Lett.*, vol. 5, pp. 21–24, Jan. 1998.
- [7] D. Gesbert, P. Duhamel, and S. Mayrargue, "Blind multichannel adaptive MMSE equalization with controlled delay," in *Proc. 8th IEEE Signal Processing Workshop Statistical Signal and Array Processing*, Corfu, Greece, June 24–26, 1996, pp. 172–175.
- [8] G. B. Giannakis and S. D. Halford, "Blind fractionally spaced equalization of noisy FIR channels: Direct and adaptive solutions," *IEEE Trans. Signal Processing*, vol. 45, pp. 2277–2292, Sept. 1997.
- [9] A. Gorokhov and P. Loubaton, "Semi-blind second order identification of convolutive channels," in *Proc. Int. Conf. ASSP(ICASSP'97)*, vol. 5, Munich, Germany, Apr. 21–24, 1997, pp. 3905–3908.
- [10] S. D. Halford and G. B. Giannakis, "Channel order determination based on sample cyclic correlations," in *Proc. 28th Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 1994, pp. 425–429.
- [11] ———, "Optimal blind equalization and symbol error analysis of fractionally-sampled channels," in *Proc. 29th Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 29–Nov. 1, 1995, pp. 1332–1336.
- [12] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [13] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.
- [14] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [15] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.
- [16] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: Unification, simplification and observations," *IEEE Trans. Aerosp., Electron. Syst.*, vol. 29, Oct. 1993.
- [17] A. P. Liavas, P. A. Regalia, and J. P. Delmas, "Blind channel approximation: Effective channel order determination," in *Proc. 32nd Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1–4, 1998.
- [18] 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, pp. 516–525, Feb. 1995.
- [19] C. B. Papadias and A. Paulraj, "Space-time signal processing for wireless communications: A survey," in *1st Signal Processing Workshop Signal Processing Advances in Wireless Communications (SPAWC'97)*, Paris, France, Apr. 16–18, 1997, pp. 285–288.
- [20] F. Rellich, *Perturbation Theory of Eigenvalue Problems*. New York: Gordon and Breach, 1969.
- [21] R. O. Schmidt, "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. dissertation, Stanford University, Stanford, CA, 1981.
- [22] D. T. M. Slock and C. B. Papadias, "Blind fractionally-spaced equalization based on cyclostationarity," in *Proc. Vehicular Technology Conf.*, Stockholm, Sweden, June, 1994.
- [23] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time-domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.
- [24] A. Touzni and I. Fijalkow, "Channel robust blind fractionally-spaced equalization," in *Proc. Signal Processing Advances in Wireless Communications (SPAWC'97)*, Paris, France, Apr. 16–18, 1997, pp. 33–36.
- [25] M. K. Tsatsanis, "Inverse filtering criteria for CDMA systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 102–112, Jan. 1997.
- [26] M. K. Tsatsanis and G. B. Giannakis, "Blind estimation of direct sequence spread spectrum signals in multipath," *IEEE Trans. Signal Processing*, vol. 45, pp. 1241–1252, May 1997.
- [27] ———, "Optimal decorrelating receivers for DS-SS systems: A signal processing framework," *IEEE Trans. Signal Processing*, vol. 44, pp. 3044–3055, Dec. 1996.
- [28] M. K. Tsatsanis and Z. Xu, "Performance analysis of minimum variance CDMA receivers," in *Proc. 13th Int. Conf. Digital Signal Processing (DSP'97)*, vol. 1, Santorini-Hellas, Greece, July 2–4, 1997, pp. 379–382.
- [29] J. Tugnait, "On blind identifiability of multipath channels using fractional sampling and second-order cyclostationary statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 308–311, Jan. 1995.
- [30] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoust., Speech, Signal Processing Mag.*, vol. 5, pp. 4–24, Apr. 1988.



**Michail K. Tsatsanis** (S'88–M'92) was born in Patras, Greece, in 1964. He received the diploma degree in electrical engineering from the National Technical University of Athens, Greece, in 1987 and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Virginia, Petersburg, in 1990 and 1993, respectively.

From 1986 until 1988, he was with Binary Logic Applications, Athens, Greece, where he worked on the design and development of digital systems for industrial control. From 1994 to 1995, he worked as a Research Associate at the Department of Electrical Engineering, University of Virginia. In 1995, he joined the Electrical Engineering and Computer Science Department, Stevens Institute of Technology, Hoboken, NJ, as an Assistant Professor. His general research interests lie in the areas of statistical signal and array processing, system identification, pattern recognition, higher order statistics, and wavelet theory. His current interests focus on signal processing techniques for wireless communications including blind equalization, multiuser detection, fading channel estimation and tracking, and signal processing methods for networking problems.

Dr. Tsatsanis is a member of the IEEE Technical Committee on SPCOM and is an Associate Editor for IEEE COMMUNICATION LETTERS and IEEE TRANSACTIONS ON SIGNAL PROCESSING. He has served as a member of the organizing committee for the 1996 IEEE Signal Processing Workshop on SSAP and is the Cochair of the organizing committee for the 1999 IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC'99). He received the NSF Career Award in 1998 and the IEEE Signal Processing Society 1998 Young Author Best Paper Award.



**Zhengyuan (Daniel) Xu** (S'97) received the B.E. and M.E. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1989 and 1991, respectively. He is currently pursuing the Ph.D. degree in electrical and computer engineering at Stevens Institute of Technology, Hoboken, NJ.

From 1991 to 1996, he worked as an Electrical Engineer and Department Manager in Tsinghua Unisplendour Group, Tsinghua University. His work focused on scanner manufacturing, office automation, and hardware and software design of digital control systems. Since August 1996, he has been a Graduate Research Assistant at Stevens Institute of Technology. His current research interests include adaptive and array signal processing, blind channel estimation and equalization for wireless communications, and single and multiuser detection for CDMA systems.

Mr. Xu is the recipient of an Outstanding Student Fellowship and Guanghua Scholarship from Tsinghua University.