

ON BLIND EQUALIZATION OF RANK DEFICIENT NONLINEAR CHANNELS

Roberto López-Valcarce

Soura Dasgupta

Dept. Tecnologías de las Comunicaciones
Universidad de Vigo, 36200 Vigo, Spain
valcarce@gts.tsc.uvigo.es

Dept. Electrical & Computer Engineering
University of Iowa, 52242 IA, USA
dasgupta@engineering.uiowa.edu

ABSTRACT

We consider the problem of blind equalization of nonlinear channels from the second-order statistics of the channel output. The channel model is linear in the parameters, with additive terms that are nonlinear functions of the transmitted symbols. All previous approaches assume that the corresponding channel matrix has full column rank, which ensures the existence of linear FIR zero forcing equalizers. We show that this assumption is not necessary, and that under certain circumstances linear FIR equalizers can be found despite the violation of this assumption. An important consequence of this fact is that equalization can be effected with a smaller level of diversity. In this paper necessary and sufficient conditions on the channel matrix are given. An algorithm for the computation of the equalizers is also given for those channels satisfying these conditions, assuming an i.i.d. symbol sequence and memory dominance of the linear part.

1. INTRODUCTION

Recently blind equalization of single-input multiple-output (SIMO) channels has received considerable attention, due to the fact that these channels can be perfectly equalized if the equalizer is long enough and the subchannels are coprime. This equalizer can be obtained from the second-order statistics (SOS) of the received signal [7].

With a few exceptions [1, 6, 9], almost all the available literature on blind equalization is devoted to the linear channel case. However, many real world communication systems, such as radio links with high power amplifiers, high-density magnetic and optical storage channels, etc., exhibit a considerable degree of nonlinearity. Thus it is of interest to consider blind equalization of *nonlinear channels*. Our 1-input, p -output channel model has the form

$$y(k) = \sum_{i=1}^q \sum_{j=0}^{l_i} h_{ij} s_i(k-j) + n(k), \quad (1)$$

where $s_1(k) = a(k)$ is the scalar, stationary input, the terms $s_i(k) = f_i(a(k), a(k-1), \dots)$ for $i = 2, \dots, q$ are scalar nonlinear causal functions of $a(\cdot)$, h_{ij} are $p \times 1$ coefficient vectors, and $n(k)$, $y(k)$ are $p \times 1$ signal vectors representing an additive disturbance and the observed signal, respectively. $n(\cdot)$ and $a(\cdot)$ are assumed independent. This model accommodates polynomial approximations of nonlinear channels (Volterra models), but the ‘basis functions’ $s_i(\cdot)$ need not be monomials in principle.

We are interested in equalizer design for the class of channels (1) using only the SOS of $y(\cdot)$. As shown in [1], under certain conditions linear finite impulse response (FIR) filters can perfectly equalize nonlinear SIMO channels of the type (1). For those cases, [1] presented a blind, deterministic approach for equalizer design. However it has been shown in [2] that the conditions in [1] are in fact conservative. More general sufficient conditions on the channel and the input signal statistics for SOS-based blind equalizability were presented in [3].

Observe that (1) could be seen as a linear multiple-input multiple-output (MIMO) system if we regard the nonlinear terms $s_i(\cdot)$ as additional inputs. However, standard SOS-based equalization techniques for MIMO systems usually assume that the different inputs are uncorrelated (which is no longer true in our setting), and they only resolve the inputs to within a mixing matrix [7]. In addition, in our case only the term $s_1(\cdot)$ is of interest.

All previous approaches [1, 2, 3] assume that the so-called *channel matrix* constructed from the channel coefficients has full column rank. In that case linear FIR equalizers always exist. However, a consequence is that the number of subchannels required by these schemes must exceed the number of distinct kernels in (1). This level of diversity may at times be unacceptably high. In an earlier paper, [4], we had shown that in a linear multi-user multichannel setting, this full column rank condition can be relaxed, and a lower level of diversity can be tolerated. In particular suppose in (1) the $s_i(k)$ are independent users, and the goal is only to equalize $s_1(k)$. Then [4] gives a necessary and sufficient condition for equalization of that $s_1(k)$. Clearly this same condition will also ensure the existence of a linear FIR equalizer for the nonlinear setting of this paper. Should l_1 exceed all other l_i , and the $s_i(k)$ are white and mutually uncorrelated then [4] also provides an algorithm that permits the construction of the equalizer from the output SOS alone. However, in the nonlinear setting one cannot assume that the $s_i(k)$ are mutually uncorrelated even if $a(k)$ is white, as $s_i(k)$ are nonlinear functions of $a(k)$. Thus the algorithm of [4] cannot be applied to nonlinear channels. The key contribution of this paper is to formulate an algorithm that provides the required equalizer from the output SOS, provided the equalizability condition of [4] is met, and even if the $s_i(k)$ are statistically dependent. This algorithm assumes that the memory of the nonlinear part is strictly less than that of the linear part. Simulation results are given as evidence of the feasibility of this procedure.

In our notation, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^\#$ denote transpose, conjugate transpose and pseudoinverse respectively; I_n , J_n denote respectively the $n \times n$ identity matrix and the shift matrix with ones in the first subdiagonal and zeros elsewhere, and e_n denotes the n -th unit vector.

2. CONDITIONS FOR THE EXISTENCE OF LINEAR FIR ZF EQUALIZERS

By stacking m consecutive samples of $y(\cdot)$ into

$$Y(k)^T = [y(k)^T \quad y(k-1)^T \quad \cdots \quad y(k-m+1)^T],$$

one gets

$$Y(k) = \mathcal{H}S(k) + N(k), \quad (2)$$

with $N(k)^T = [n(k)^T \quad n(k-1)^T \quad \cdots \quad n(k-m+1)^T]$, $S(k)^T = [S_1^T(k) \quad S_2^T(k)]$ the noise and signal vectors,

$$S_1^T(k) = [a(k) \quad \cdots \quad a(k-l_1-m+1)], \quad (3)$$

$$S_2^T(k) = [s_2(k) \quad \cdots \quad s_s(k-l_2-m+1) \quad | \quad \cdots \\ \cdots \quad | \quad s_q(k) \quad \cdots \quad s_q(k-l_q-m+1)], \quad (4)$$

and the channel matrix $\mathcal{H} = [\mathcal{H}_1 \quad \mathcal{H}_2 \quad \cdots \quad \mathcal{H}_q]$, with every \mathcal{H}_i block Toeplitz:

$$\mathcal{H}_i = \begin{bmatrix} h_{i0} & \cdots & h_{il_i} & & \\ & \ddots & & \ddots & \\ & & h_{i0} & \cdots & h_{il_i} \end{bmatrix} \quad pm \times (m+l_i).$$

For convenience, let $d_1 = m + l_1$, which is the size of $S_1(k)$, the linear part of the regressor; and $d_2 = l_2 + \cdots + l_q + (q-1)m$, which is the size of $S_2(k)$ (thus $S(k)$ is $(d_1 + d_2) \times 1$).

Observe that if the channel matrix \mathcal{H} has full column rank, then its pseudoinverse $\mathcal{H}^\#$ satisfies $\mathcal{H}^\# \mathcal{H} = I_{d_1+d_2}$. In that case, in the noiseless case ($N(k) = 0$) one obtains from (2) $\mathcal{H}^\# Y(k) = S(k)$. Thus the first d_1 rows of $\mathcal{H}^\#$ provide zeroforcing (ZF) equalizers with associated delays 0 through $d_1 - 1$. However, this also shows the existence of vectors (the last d_2 rows of $\mathcal{H}^\#$) that recover all the nonlinear terms $s_i(k)$ and their delays, which is clearly not necessary since these terms are of no interest to the receiver. This leads us to ask for necessary and sufficient conditions on \mathcal{H} for the ZF equalizers to exist. First, let us introduce the following partition of the channel matrix:

$$\mathcal{H} = [\mathcal{H}_1 \quad \mathcal{H}_{nl}] \quad \text{with} \quad \mathcal{H}_{nl} = [\mathcal{H}_2 \quad \cdots \quad \mathcal{H}_q]. \quad (5)$$

That is, \mathcal{H}_{nl} comprises the ‘nonlinear part’ of the channel matrix. Recall that \mathcal{H}_1 and \mathcal{H}_{nl} have sizes $pm \times d_1$ and $pm \times d_2$ respectively. We shall make the following assumption:

A1: \mathcal{H}_1 has full column rank, and with $r_1 = \text{rank}(\mathcal{H}_1)$, $r_2 = \text{rank}(\mathcal{H}_{nl})$, \mathcal{H} satisfies $\text{rank}(\mathcal{H}) = r_1 + r_2 \leq pm$.

Observe that if \mathcal{H} has full column rank, then Assumption **A1** is satisfied but not conversely. The significance of this condition is reflected in the following result from [4]:

Theorem 1 *There exists a $pm \times d_1$ matrix \mathcal{G} such that*

$$\mathcal{G}^H \mathcal{H} = [I_{d_1} \quad 0_{d_1 \times d_2}] \quad (6)$$

*if and only if Assumption **A1** holds.*

The columns of \mathcal{G} constitute the desired ZF equalizers. The geometrical interpretation of Theorem 1 is as follows.

Lemma 1 *The condition $\text{rank}([\mathcal{H}_1 \quad \mathcal{H}_{nl}]) = \text{rank}(\mathcal{H}_1) + \text{rank}(\mathcal{H}_{nl})$ is equivalent to $\text{range}(\mathcal{H}_1) \cap \text{range}(\mathcal{H}_{nl}) = \{0\}$, with $\text{range}(A)$ the subspace spanned by the columns of A .*

Thus linear FIR ZF equalizers exist iff \mathcal{H}_1 has full column rank and no nonzero vector lies in the range space of both \mathcal{H}_1 and \mathcal{H}_{nl} .

3. SOS-BASED EQUALIZER DESIGN

We turn our attention now to the problem of extracting the equalizers from the SOS of the received signal, assuming that \mathcal{H} satisfies the relaxed rank condition **A1**. From (2), the covariance of the received vector $Y(\cdot)$ is given by

$$C_y(l) = \text{cov}[Y(k), Y(k-l)] = \mathcal{H}C_s(l)\mathcal{H}^H + C_n(l), \quad (7)$$

with $C_s(l) = \text{cov}[S(k), S(k-l)]$, $C_n(l) = \text{cov}[N(k), N(k-l)]$ the signal and noise covariance matrices. In addition to **A1**, we adopt the following standard assumptions:

A2: $n(\cdot)$ is zero-mean, white, with covariance $\sigma_n^2 I_p$.

A3: The covariance matrix $C_s(0)$ is positive definite.

Observe that [4] assumes that $C_s(0)$ is diagonal. This assumption is not needed here. Under **A1** and **A2**, σ_n^2 can be estimated as the smallest eigenvalue of $C_y(0)$. Thus the effect of the noise can be removed from $C_y(l)$; henceforth we shall assume that $C_y(l) = \mathcal{H}C_s(l)\mathcal{H}^H$. **A3** is a ‘persistent excitation’ condition on $a(\cdot)$, which allows us to write

$$C_s(0) = QQ^H \quad \text{with } Q \text{ invertible.} \quad (8)$$

Now let Q be a square root of $C_s(0)$ as in (8), and define the *normalized* channel and source covariance matrices respectively as

$$H = \mathcal{H}Q, \quad \bar{C}_s(l) = Q^{-1}C_s(l)Q^{-H}. \quad (9)$$

Using (9), the matrices $C_y(l)$ become

$$C_y(l) = H\bar{C}_s(l)H^H, \quad \text{with } \bar{C}_s(0) = I_{d_1+d_2}. \quad (10)$$

The following result relates the ZF equalizers to the normalized channel matrix H .

Lemma 2 *Under **A1-A3**, let the square root of $C_s(0)$, Q , be block lower triangular:*

$$Q = \begin{bmatrix} Q_{11} & 0 \\ Q_{21} & Q_{22} \end{bmatrix}, \quad \text{with } Q_{ij} \text{ of size } d_i \times d_j. \quad (11)$$

Then the matrix \mathcal{G} satisfying (6) (ZF equalizers) is given by

$$\mathcal{G}^H = Q_{11} [I_{d_1} \quad 0_{d_1 \times d_2}] H^\#. \quad (12)$$

Thus if $H = U_1 \Sigma V$ is an SVD of H , with $U_1: pm \times (r_1+r_2)$, $\Sigma: (r_1+r_2) \times (r_1+r_2)$, $V: (r_1+r_2) \times (d_1+d_2)$, and partitioning V as

$$V = [V_1 \quad V_2], \quad V_i \text{ of size } (r_1+r_2) \times d_i, \quad (13)$$

then the equalizers are given by

$$\mathcal{G}^H = Q_{11} V_1^H \Sigma^{-1} U_1^H. \quad (14)$$

Observe that Q_{11} is known to us from the source statistics, and that Σ, U_1 can be obtained from an SVD of $C_y(0)$ since

$$C_y(0) = \mathcal{H}C_s(0)\mathcal{H}^H = HH^H = U_1 \Sigma^2 U_1^H. \quad (15)$$

Therefore if V_1 could be somehow estimated, the ZF equalizers could be computed. Note that $VV^H = V_1 V_1^H + V_2 V_2^H = I_{r_1+r_2}$. An additional property is shown by the next result.

Lemma 3 Under **A1-A3**, let Q be block lower triangular as in (11). Let $H = \mathcal{H}Q$ have a singular value decomposition $H = U_1 \Sigma V$, and partition V as in (13). Then

$$V_1^H V_1 = I_{d_1}, \quad V_1^H V_2 = 0_{d_1 \times d_2}. \quad (16)$$

This property is obvious in the full column rank case (for which V is square), but it is somewhat surprising that it still holds even under the relaxed rank condition **A1**. Now consider the matrix

$$R(1) = \Sigma^{-1} U_1^H C_y(1) U_1 \Sigma^{-1}, \quad (17)$$

which satisfies $R(1) = V \bar{C}_s(1) V^H$. From (16), this gives

$$R(1) V_1 = V \bar{C}_s(1) \begin{bmatrix} I_{d_1} \\ 0 \end{bmatrix}, \quad R^H(1) V_1 = V \bar{C}_s^H(1) \begin{bmatrix} I_{d_1} \\ 0 \end{bmatrix}. \quad (18)$$

These relations will allow us to estimate V_1 under the following additional assumptions:

A4: The symbol sequence $a(\cdot)$ is a zero-mean i.i.d. process with $\text{cov}[a(k), a(k)] = \sigma_a^2$.

A5: $S_2(k)$ satisfies $S_2(k) = f(a(k), a(k-1), \dots, a(k-d_1+2))$, with $f(\cdot, \dots, \cdot)$ a memoryless mapping.

Basically **A5** amounts to saying that the memory of the nonlinear part of the channel is strictly shorter than that of the linear part. One has the following result:

Lemma 4 Under **A1-A5**, a lower block triangular square root Q as in (11) exists such that $Q_{11} = \sigma_a^2 I_{d_1}$ and

$$\bar{C}_s(1) = \begin{bmatrix} J_{d_1} & 0 \\ 0 & C \end{bmatrix} \text{ for some } d_2 \times d_2 \text{ } C, \quad (19)$$

$$\bar{C}_s(d_1 - 1) = e_{d_1} e_1^H. \quad (20)$$

Substituting (19) in (18) one obtains the Jordan chains

$$R(1) V_1 = V_1 J_{d_1}, \quad R^H(1) V_1 = V_1 J_{d_1}^H, \quad (21)$$

which show how V_1 can be estimated once its first or last column is available. Partition $V_1 = [v_1 \ v_2 \ \dots \ v_{d_1}]$ columnwise, and consider the matrix

$$R(d_1 - 1) = \Sigma^{-1} U_1^H C_y(d_1 - 1) U_1 \Sigma^{-1}, \quad (22)$$

which satisfies $R(d_1 - 1) = V \bar{C}_s(d_1 - 1) V^H$. Using (20),

$$R(d_1 - 1) = V e_{d_1} e_1^H V^H = v_{d_1} v_1^H. \quad (23)$$

Thus $R(d_1 - 1)$ is a rank one matrix and its only nonzero singular value equals 1. The vectors v_1, v_{d_1-1} can be obtained up to a constant of the form $e^{j\theta}$ from an SVD of $R(d_1 - 1)$, or alternatively they can be estimated as

$$\hat{v}_{d_1} = \frac{R(d_1 - 1) e_{i_{\max}}}{\|R(d_1 - 1) e_{i_{\max}}\|}, \quad \hat{v}_1^H = \frac{e_{j_{\max}}^H R(d_1 - 1)}{\|e_{j_{\max}}^H R(d_1 - 1)\|}, \quad (24)$$

where

$$i_{\max} = \arg \max\{\|R(d_1 - 1) e_i\|, 1 \leq i \leq d_1 + d_2\},$$

$$j_{\max} = \arg \max\{\|e_j^H R(d_1 - 1)\|, 1 \leq j \leq d_1 + d_2\}.$$

In this way these estimates \hat{v}_{d_1}, \hat{v}_1 are related to the true quantities v_{d_1}, v_1 by some complex constants with unit modulus. From these, the remaining columns of V_1 can be estimated via either

of the Jordan chains (21), thus obtaining an estimate \hat{V}_1 satisfying $\hat{V}_1 = e^{j\theta} V_1$ for some real θ . Therefore the matrix $\hat{\mathcal{G}}_{\text{ZF}} = \sigma_a U_1 \Sigma^{-1} \hat{V}_1$ satisfies $\hat{\mathcal{G}}_{\text{ZF}}^H \mathcal{H} = e^{j\theta} I_{d_1}$, providing equalization up to an unknown phase rotation. This is acceptable since the need for a phase reference can be sidestepped by differentially encoding the data. Finally, it is possible to obtain the Minimum Mean-Squared Error (MMSE) equalizers in the spirit of [5]:

Lemma 5 Under Assumptions **A1-A3**, the MMSE equalizers $\mathcal{G}_{\text{MMSE}}$ minimizing trace $E[|\mathcal{G}^H Y(k) - S_1(k)|^2]$ are related to the ZF equalizers by

$$\mathcal{G}_{\text{MMSE}} = [I - \sigma_n^2 C_y^{-1}(0)] \mathcal{G}_{\text{ZF}} \quad (25)$$

where now $C_y(0) = \mathcal{H} C_s(0) \mathcal{H}^H + \sigma_n^2 I_{pm}$ represents the undenoised channel output covariance matrix.

The resulting algorithm is summarized next.

Blind equalization algorithm

1. Compute estimates $\hat{C}_y(0), \hat{C}_y(1), \hat{C}_y(d_1 - 1)$.
2. Estimate $\hat{\sigma}_n^2$ as the smallest eigenvalue of $\hat{C}_y(0)$ and subtract the noise effect from $\hat{C}_y(\cdot)$.
3. Perform an SVD of $\hat{C}_y(0)$ as in (15) to obtain U_1, Σ .
4. Compute $R(1), R(d_1 - 1)$ as in (17), (22) respectively.
5. Form the estimates \hat{v}_{d_1}, \hat{v}_1 via (24).
6. For $i = 2, 3, \dots, d_1$, let $\hat{v}_i = R(1) \hat{v}_{i-1}$. Alternatively, for $j = d_1, d_1 - 1, \dots, 2$, let $\hat{v}_{j-1} = R^H(1) \hat{v}_j$.
7. ZF equalizers: $\hat{\mathcal{G}}_{\text{ZF}} = \sigma_a U_1 \Sigma^{-1} [\hat{v}_1 \ \dots \ \hat{v}_{d_1}]$.
8. Compute the MMSE equalizers via (25).

4. SIMULATION RESULTS

We present now a numerical example of the results obtained by the algorithm. For illustration purposes, the phase ambiguity inherent to the method was removed before computing the error rates. Averages were computed based on 100 independent runs.

The channel we consider is real with $q = 3, l_1 = 4, l_2 = l_3 = 1$ and i.i.d. symbols taking the values ± 1 with equal probabilities. The number of subchannels is $p = 4$; the coefficients are given in table 1. The nonlinear terms are $s_2(k) = a(k)a(k-1), s_3(k) = a(k)a(k-2)$. The resulting linear-to-nonlinear distortion ratio for this channel is 8 dB. The equalizer length that we consider is $m = 6$. The corresponding channel matrix \mathcal{H} (of size 24×24) is not full column rank but it satisfies the relaxed rank condition **A1**:

$$23 = \text{rank}(\mathcal{H}) = \text{rank}(\mathcal{H}_1) + \text{rank}([\mathcal{H}_2 \ \mathcal{H}_3]) = 10 + 13.$$

Note that this channel satisfies **A5**, and that σ_n^2 can still be estimated as the smallest eigenvalue of $C_y(0)$ even though the channel matrix \mathcal{H} is square, since \mathcal{H} is rank deficient. Figure 1(a) shows the symbol error rate (SER) vs. SNR using $K = 2000$ samples for covariance estimation, while figure 1(b) shows the variation of the SER with K for a fixed value of SNR = 24 dB, for the equalization delays 0, 3, 8 and 9. In this case the equalizer with maximal delay ($d = 9$) provides the poorest performance of all. The best results are obtained with the equalizer of delay $d = 3$.

channel	h_{10}	h_{11}	h_{12}	h_{13}	h_{14}	h_{20}	h_{21}	h_{30}	h_{31}
1	1.0	0.5	0.4	0.2	-0.2	0.2	0.5	0.1	-0.1
2	0.1	0.6	1.0	-0.4	0.2	0.1	0.25	0.2	-0.2
3	-0.2	0.6	0.6	0.1	-0.3	0.2	0.5	0.2	-0.2
4	0.3	1.0	0.7	-0.5	0.2	0.1	0.25	0.1	-0.1

Table 1: Coefficients of the Volterra channel used in the simulations

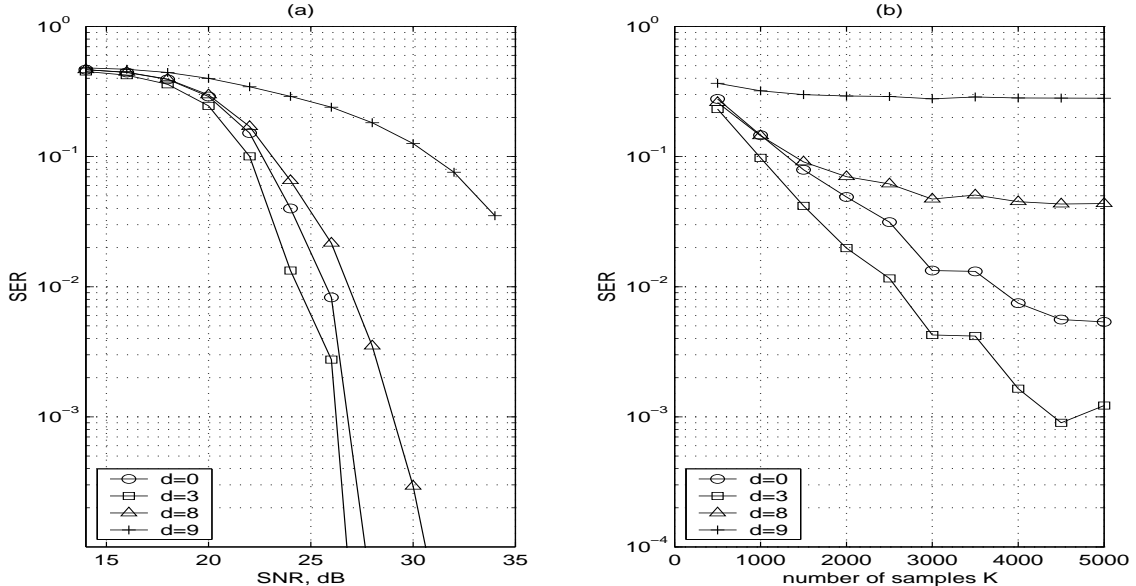


Figure 1: MMSE equalizer performance, $m = 6$. (a) SER vs. SNR, $K = 2000$ symbols. (b) SER vs. sample size K , SNR = 24 dB.

5. CONCLUSIONS

In contrast with the linear channel case, for equalizability of nonlinear channels with linear FIR filters it is not necessary that the channel matrix have full column rank. We have given necessary and sufficient conditions on the channel matrix for this property to hold. If in addition the input symbol sequence is i.i.d. and the memory of the nonlinear part of the channel is strictly shorter than that of the linear part, a blind algorithm based on the second-order statistics of the channel output provides the equalizers.

6. REFERENCES

- [1] G. B. Giannakis and E. Serpedin, "Linear multichannel blind equalizers of nonlinear FIR Volterra channels", *IEEE Trans. Signal Processing*, vol. 45 no. 1, pp. 67-81, Jan. 1997.
- [2] R. López-Valcarce and S. Dasgupta, "Blind identifiability/equalizability of single input multiple output nonlinear channels from second order statistics", *Proc. 2000 IEEE ICASSP*, Istanbul, Turkey.
- [3] R. López-Valcarce and S. Dasgupta, "The role of second-order statistics in blind equalization of nonlinear systems", *Proc. 2000 IEEE SSAP Workshop*, pp. 211-215, Pocono Manor, PA.
- [4] Z. Ding, S. Dasgupta and R. López-Valcarce, "Interference cancellation and blind equalization for linear multi-user systems", *Proc. 2000 IEEE ICASSP*, Istanbul, Turkey.
- [5] C. Papadias and D. T. M. Slock, "Fractionally spaced equalization of linear polyphase channels and related blind techniques based on multichannel linear prediction", *IEEE Trans. Signal Processing*, vol 47 no. 3, pp. 641-654, March 1999.
- [6] G. M. Raz and B. D. Van Veen, "Blind equalization and identification of nonlinear and IIR systems—A Least Squares approach", *IEEE Trans. Signal Processing*, vol. 48 no. 1, pp. 192-200, January 2000.
- [7] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods", *Proc. IEEE*, vol. 86 no. 10, pp. 1951-1968, Oct. 1998.
- [8] L. Tong, G. Xu and T. Kailath, "Blind identification and equalization based on second-order statistics: a time-domain approach", *IEEE Trans. Information Theory*, vol. 40, no. 2, pp. 340-350, March 1994.
- [9] M. Tsatsanis and H. Cirpan, "Blind identification of nonlinear channels excited by discrete alphabet inputs", *Proc. 1996 IEEE SSAP Workshop*, vol. 1, pp. 176-179, Corfu, Greece.