BLIND EQUALIZATION OF NONLINEAR CHANNELS FROM SECOND ORDER STATISTICS USING PRECODING AND CHANNEL DIVERSITY

Roberto López-Valcarce, Xuejie Song and Soura Dasgupta

Department of Electrical and Computer Engineering
University of Iowa, Iowa City, IA 52242, USA

ABSTRACT

This paper considers the blind equalization problem for nonlinear channels of Volterra type excited by real i.i.d. symbols. Previous work has shown that under the right conditions the equalizers can be found from the second order statistics of the channel output as long as the number of subchannels exceeds the number of kernels. In order to alleviate this requirement, we consider the use of a simple precoding device previous to transmission which provides a trade-off between effective data rate and number of subchannels required. Necessary and sufficient conditions for blind equalizability under this scheme are given, and an algorithm for the computation of the equalizers is presented.

1. INTRODUCTION

Blind equalization of single-input multiple-output linear channels has received considerable attention, motivated by the fact that these channels can be perfectly equalized from the second-order statistics (SOS) of the received signal [4] if the subchannels are coprime.

Many systems such as digital satellite and radio links, high-density magnetic and optical storage channels, etc., exhibit non-trivial nonlinearities. Recently blind equalization techniques for *nonlinear channels*, both deterministic [1] and SOS-based [2] have been proposed, which exploit the fact that if several subchannels are available (e.g. if the received signal is oversampled and/or multiple sensors are used), linear FIR equalizers have the potential to completely remove both linear and nonlinear ISI. The discrete-time equivalent single-input, p-output channel model that we consider has the form

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

where the $p \times 1$ vectors $y(\cdot)$, $v(\cdot)$, h_{ij} denote respectively the received signal, additive noise, and channel coefficients. We adopt a truncated Volterra series approximation [3] of the nonlinear channel so that the 'basis functions' $s_i(\cdot)$ are taken as monomials: with $a(\cdot)$ the scalar-valued transmitted symbols, and for some t_i, $0 \le k_{i,1} \le \cdots \le k_{i,t_i}$,

$$s_i(n) = a(n)a(n-k_{i,1}) \cdots a(n-k_{i,t_i}), \tag{2}$$

and $s_j(n) = a(n)$ for some j, i.e. a linear kernel is present.

A first requirement in the schemes of [1] and [2] is that $p > q$, i.e. the number of subchannels must exceed the number of kernels. For a large number of nonlinear kernels this may require too much diversity. In this paper we investigate blind equalization from the output SOS through a simple data precoding scheme that reduces the number of nonlinear kernels at the expense of the effective data rate. We assume real i.i.d. symmetrically distributed input symbols. Whereas [2] only gives sufficient conditions on the input under which blind equalization from SOS can be effected, here we give conditions that are both necessary and sufficient.

In our notation, $(\cdot)^T$ denotes the transpose; I_m, J_m denote respectively the $m \times m$ identity matrix and the shift matrix with ones in the first subdiagonal and zeros elsewhere; e_k denotes the k-th unit vector, and $\oplus$ stands for block diagonal concatenation, e.g. $A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$.

2. THE PRECODING SCHEME

Observe from (2) that the $s_i(\cdot)$ are products of $a(n)$ and several of its delays. Now suppose that the sequence transmitted through the channel, $a(\cdot)$, is obtained from the actual symbol stream $x(\cdot)$ by inserting $N-1$ zeros between any consecutive symbols:

$$a(n) = \begin{cases} x(n/N) & \text{if } n \bmod N = 0, \\ 0 & \text{else.} \end{cases} \tag{3}$$

If for the i-th kernel we define $j(i) = \min\{j \mid k_{i,j} > 0\}$, and if N satisfies $N > \max\{k_{i,j(i)},\ 1 \le i \le q\}$, then due to (3) all the terms $s_i(n)$ become zero for all n except those of the form

$$s_i(n) = [a(n)]^{m_i}. \tag{4}$$

Hence the effective number of kernels has been reduced by the precoding operation by the removal of all nonlinear terms except pure powers of $a(n)$. In what follows it is assumed that (3) is implemented so that all terms satisfy (4). We shall denote the total number of surviving kernels by q' (thus $q' \le q$).

We shall consider the use of linear equalizers for the precoded nonlinear channel. Due to the upsampling operation (3) at the transmitter, the equalizer output must be downsampled by a factor of N. With $x(n)$, the data to be transmitted and $a(n)$ its precoded mutation, the overall configuration is depicted in Figure 1, where ZMNL is a single-input, q'-output 'Zero-Memory Nolinearity', $\mathbf{H}(z)$ is a $p \times q'$ matrix transfer function whose i-th column is $\sum_{j=0}^{l_i} h_{ij} z^{-j}$, and $\mathbf{G}(z)$ is the $1 \times p$ equalizer transfer function. In view of (4), ZMNL only consists pure powers of its input, i.e. has the form $[1, (.)_1^m, (.)_2^m, \cdots]^T$. By using

$$\mathbf{H}(z) = \sum_{l=0}^{N-1} z^{-l} \mathbf{E}_l(z^N), \qquad \mathbf{G}(z) = \sum_{l=0}^{N-1} z^{-l} \mathbf{R}_l(z^N),$$

Supported by NSF grants ECS-9970105 and CCR-9973133.

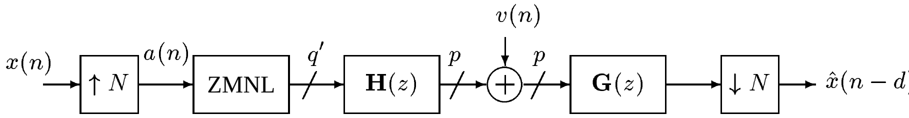

Figure 1: Overall precoder-channel-equalizer configuration.

(polyphase representations), and noting that the upsampler and the ZMNL commute, one obtains the equivalent block diagram of Figure 2, where $v_l(n) = v(nN - l)$, $0 \le l \le N-1$.

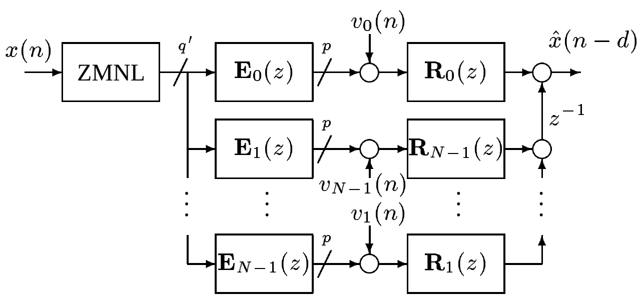

Figure 2: Equivalent configuration.

3. PROBLEM STATEMENT

Observe that the $p \times 1$ vector process $y(\cdot)$ is cyclostationary. We can block it to form the $Np \times 1$ vector process

$$\mathbf{y}(n) \triangleq [\; y(nN)^T \quad y(nN-1)^T \quad \cdots \quad y(nN-N+1)^T \;]^T,$$

which now is wide-sense stationary. This process satisfies

$$\mathbf{y}(n) = \sum_{i=1}^{q'} \sum_{j=0}^{l'_i} \mathbf{h}_{ij} [x(n-j)]^{m_i} + \mathbf{v}(n), \qquad (5)$$

with $l'_i = \lfloor l_i / N \rfloor$ and the $Np \times 1$ vectors $\mathbf{h}_{ij}$, $\mathbf{v}(n)$ given by

$$\mathbf{h}_{ij} \triangleq \begin{bmatrix} h_{i,jN} \\ h_{i,jN+1} \\ \vdots \\ h_{i,jN+N-1} \end{bmatrix}, \quad \mathbf{v}(n) \triangleq \begin{bmatrix} v_0(n) \\ v_1(n) \\ \vdots \\ v_{N-1}(n) \end{bmatrix}.$$

With this, we can stack M consecutive samples of the processes $\mathbf{y}(\cdot)$, $\mathbf{v}(\cdot)$ to define the $MNp \times 1$ vectors

$$\begin{aligned} Y(n) &= [\; \mathbf{y}(n)^T \quad \mathbf{y}(n-1)^T \quad \cdots \quad \mathbf{y}(n-M+1)^T \;]^T, \\ V(n) &= [\; \mathbf{v}(n)^T \quad \mathbf{v}(n-1)^T \quad \cdots \quad \mathbf{v}(n-M+1)^T \;]^T. \end{aligned}$$

These processes satisfy the following relation:

$$Y(n) = \mathcal{F} S(n) + V(n), \qquad (6)$$

where the signal regressor $S(n)$ is given by

$$\begin{aligned} S(n) &= [\; S_1(n)^T \quad \cdots \quad S_{q'}(n)^T \;]^T, \\ S_i(n) &= [\; x(n)^{m_i} \quad x(n-1)^{m_i} \quad \cdots \quad x(n-d_i+1)^{m_i} \;]^T, \end{aligned}$$

with $d_i = M + l'_i$, and the channel matrix $\mathcal{F}$ is given by $\mathcal{F} = [\; \mathcal{F}_1 \quad \mathcal{F}_2 \quad \cdots \quad \mathcal{F}_{q'} \;]$, with

$$\mathcal{F}_i \triangleq \begin{bmatrix} \mathbf{h}_{i0} & \cdots & \mathbf{h}_{il'_i} & & \\ & \ddots & & \ddots & \\ & & \mathbf{h}_{i0} & \cdots & \mathbf{h}_{il'_i} \end{bmatrix}, \qquad NMp \times (M + l'_i).$$

Therefore the covariance sequence of $Y(\cdot)$ can be written as

$$C_y(k) \triangleq \mathrm{cov}[Y(n), Y(n-k)] = \mathcal{F} C_s(k) \mathcal{F}^T + C_v(k), \quad (7)$$

with the signal and noise covariance matrices

$$C_s(k) \triangleq \mathrm{cov}[S(n), S(n-k)], \quad C_v(k) \triangleq \mathrm{cov}[V(n), V(n-k)].$$

We adopt the following assumptions:
A1: The channel matrix $\mathcal{F}$ has full column rank.
A2: $v(\cdot)$ is zero-mean, white, with covariance $\sigma_v^2 I_p$.
A3: $x(\cdot)$ is real, zero-mean, i.i.d., symmetrically distributed about the origin.
A4: $C_s(0)$ is positive definite.

Observe that for **A1** to hold $\mathcal{F}$ must be tall, i.e.

$$NMp > Mq + (l'_1 + \cdots + l'_{q'})$$

must hold. This in turn implies $p > q'/N$, in contrast with the requirement $p > q$ when no precoding is used (also note $q' \le q$ since the 'cross-terms' kernels have been eliminated by the precoding scheme). **A1** ensures the existence of vectors g_d with $g_d^T \mathcal{F}$ a unit vector such that in the noiseless case, for $0 \le d \le d_j - 1$ one has $g_d^T Y(n) = x(n-d)$; i.e. g_d provides a FIR zero-forcing (ZF) linear equalizer with associated delay d.

Under **A2**, the process $\mathbf{v}(\cdot)$ is also white with $\mathrm{cov}[\mathbf{v}(n), \mathbf{v}(n)] = \sigma_v^2 I_{Np}$, so that

$$C_v(k) = \sigma_v^2 \left(J_{NMp}^{Np} \right)^k.$$

Then σ_v^2 can be estimated as the smallest eigenvalue of $C_y(0)$ and thus the effect of the noise can be removed from $C_y(k)$. Henceforth we shall assume that this has been done so that $C_y(k) = \mathcal{F} C_x(k) \mathcal{F}^T$. Assumption **A4** is a 'persistent excitation' requirement. The problem is stated as follows.
Blind Equalizability Problem: *Let $\tilde{\mathcal{F}}$ be a matrix of the same size as $\mathcal{F}$ such that*

$$\tilde{\mathcal{F}} C_s(k) \tilde{\mathcal{F}}^T = \mathcal{F} C_s(k) \mathcal{F}^T, \qquad k = 0, 1, \ldots \bar{k}. \qquad (8)$$

We say that $\tilde{\mathcal{F}}$ is compatible *with the SOS of $Y(\cdot)$, up to lag $\bar{k}$. Determine conditions under which a ZF equalizer g_d for any compatible $\tilde{\mathcal{F}}$ is also a ZF equalizer for $\mathcal{F}$. That is, if the j-th kernel is the linear one, then with $0 \le d \le d_j - 1$ and $c \ne 0$,*

$$\begin{aligned} & g_d^T \tilde{\mathcal{F}} = [\; 0^T_{d_1 + \cdots + d_{j-1}} \quad e^T_{d+1} \quad 0^T_{d_{j+1} + \cdots + d_{q'}} \;] \\ \Rightarrow \quad & g_d^T \mathcal{F} = [\; 0^T_{d_1 + \cdots + d_{j-1}} \quad c \cdot e^T_{d+1} \quad 0^T_{d_{j+1} + \cdots + d_{q'}} \;]. \quad (9) \end{aligned}$$

Observe that if (9) is satisfied, then the matrices $C_y(0), \ldots, C_y(\bar{k})$ effectively contain enough information for the determination of the equalizers. In that case we say that the channel is blindly equalizable from the SOS of $Y(\cdot)$.

4. EQUALIZABILITY CONDITIONS

Assume that $l'_1 \geq l'_2 \geq \cdots \geq l'_{q'}$. Define

$$\begin{aligned} L_{ij} &\triangleq [\ I_{d_i} \quad 0_{d_i \times (d_j - d_i)}\ ] \qquad (d_i \times d_j), \\ \alpha_{ij} &\triangleq \operatorname{cov}[x^{m_i}(n), x^{m_j}(n)], \qquad 1 \leq i, j \leq q', \end{aligned}$$

and let A be the matrix $A = (\alpha_{ij})_{1 \leq i,j \leq q'}$. Observe that $A > 0$ because of **A4**. Let $A = BB^T$ be the Cholesky factorization of A with $B = (\beta_{ij})_{1 \leq i,j \leq q'}$ (thus $\beta_{ij} = 0$ if $j > i$). We have the following result:

Lemma 1 *The source covariance matrices $C_s(k)$ admit the factorization $C_s(k) = \Omega \mathcal{J}^k \Omega^T$, where $\mathcal{J} = J_{d_1} \oplus J_{d_2} \oplus \cdots \oplus J_{d_{q'}}$ and Ω is the $(d_1 + \cdots + d_{q'})$-square matrix*

$$\Omega \triangleq \begin{bmatrix} \beta_{11} L_{11} & & & \\ \beta_{21} L_{21} & \beta_{22} L_{22} & & \\ \vdots & \vdots & \ddots & \\ \beta_{q'1} L_{q'1} & \beta_{q'2} L_{q'2} & \cdots & \beta_{q'q'} L_{q'q'} \end{bmatrix}. \qquad (10)$$

Then in view of Lemma 1 and (7), one has

$$C_y(k) = \mathcal{F} \Omega \mathcal{J}^k \Omega^T \mathcal{F}^T. \qquad (11)$$

Now let $\tilde{\mathcal{F}}$ be a compatible matrix. For $k = 0$, (8) reads as $(\tilde{\mathcal{F}}\Omega)(\tilde{\mathcal{F}}\Omega)^T = (\mathcal{F}\Omega)(\mathcal{F}\Omega)^T$. Since by **A1** $\mathcal{F}$ has full column rank and Ω is nonsingular, it follows that $(\tilde{\mathcal{F}}\Omega) = (\mathcal{F}\Omega)P$ for some orthogonal matrix P, or equivalently $\tilde{\mathcal{F}} = \mathcal{F}\Omega P \Omega^{-1}$. In addition, for $k = 1, \ldots, \bar{k}$, (8) gives

$$P \mathcal{J}^k = \mathcal{J}^k P. \qquad (12)$$

Observe that if (12) is satisfied for $k = 1$, then it is satisfied for all $k > 0$. Therefore it suffices to consider blind equalizability from $C_y(0)$ and $C_y(1)$. The next result gives necessary and sufficient conditions.

Lemma 2 *The channel (5) is blindly equalizable from the output SOS iff (i) there is no nonlinear kernel with the same length as the linear one, and (ii) there is no odd-order kernel longer than the linear one. That is, if the j-th kernel is the linear one, then $l'_i \neq l'_j$ for $i \neq j$; and if $l'_i > l'_j$ then m_i must be even.*

Note that condition (ii) in Lemma 2 amounts to having $\alpha_{j1} = \cdots = \alpha_{j,j-1} = 0$ or equivalently

$$\beta_{j1} = \cdots = \beta_{j,j-1} = 0. \qquad (13)$$

5. OBTAINING THE EQUALIZERS

Assuming that the channel satisfies the conditions of Lemma 2, the equalizers can be found as follows. First, it will be convenient to reorder the kernels so as to have the linear one first, then all those longer than the linear one, and then all those shorter. Then (11) still holds, but now Ω is obtained from (10) by suitably permuting rows and columns. (Note that the resulting Ω still is lower triangular because of (13)). As in [5], perform an SVD of $C_y(0)$:

$$C_y(0) = [\ U_1 \quad U_2\ ] \begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix},$$

where $\Sigma > 0$ is $r \times r$ diagonal and U_1 has r columns, r being the number of columns of $\mathcal{F}$. Due to **A1** and (11), one has $\mathcal{F} = U_1 \Sigma V \Omega^{-1}$ for some $r \times r$ orthogonal V. Let V_1 comprise the first d_1 columns of V. It can be readily checked that with $\sigma_x^2 \triangleq E[x^2(n)]$, the ZF equalizers are given by

$$\mathcal{G}_{\mathrm{ZF}} \triangleq [\ g_0 \quad \cdots \quad g_{d_1-1}\ ] = \sigma_x U_1 \Sigma^{-1} V_1. \qquad (14)$$

In order to obtain V_1 we shall make use of the matrix

$$R \triangleq \Sigma^{-1} U_1^T C_y(1) U_1 \Sigma^{-1} = V \mathcal{J} V^T. \qquad (15)$$

From (15), it follows that

$$R V_1 = V_1 J_{d_1}, \qquad R^T V_1 = V_1 J_{d_1}^T, \qquad (16)$$

so that V_1 can be recovered from its first or last column.

Due to our reordering of the kernels, $\mathcal{J} = J_{d_1} \oplus Y \oplus \hat{Y}$ where $\hat{Y}$ comprises the J blocks of size less than d_1 and Y those of size greater than d_1. Partitioning $V = [\ V_1 \quad V_2 \quad V_3\ ]$ where V_2, V_3 have the same number of columns as Y, $\hat{Y}$ respectively, one has

$$R = \underbrace{V_1 J_{d_1} V_1^T}_{\bar{R}} + \underbrace{V_2 Y V_2^T}_{\Gamma} + \underbrace{V_3 \hat{Y} V_3^T}_{\hat{\Gamma}}. \qquad (17)$$

Further $R^k = \bar{R}^k + \Gamma^k + \hat{\Gamma}^k$ for all k. Let $\mathcal{J}_{ij} \triangleq J_j \oplus \overset{(i \text{ times})}{\cdots} \oplus J_j$; then for some u, $r_1, \ldots, r_u$, and $t_1 > t_2 > \cdots > t_u > d_1$, one can write without loss of generality

$$Y = \mathcal{J}_{r_1 t_1} \oplus \cdots \oplus \mathcal{J}_{r_u t_u}. \qquad (18)$$

Partitioning accordingly $V_2 = [\ V_{21} \quad \cdots \quad V_{2u}\ ]$ one has $\Gamma = V_2 Y V_2^T = \sum_{i=1}^u \Gamma_i$ where $\Gamma_i \triangleq V_{2i} \mathcal{J}_{r_i t_i} V_{2i}^T$; and

$$\Gamma^k = \sum_{i=1}^u \Gamma_i^k \qquad (19)$$

holds for all k. From (15), one finds that for $i = 1, \ldots, u$,

$$R V_{2i} = V_{2i} \mathcal{J}_{r_i t_i}, \qquad R^T V_{2i} = V_{2i} \mathcal{J}_{r_i t_i}^T. \qquad (20)$$

Using (20), it can be shown that for $i = 1, \ldots, u$,

$$\Gamma_i = \sum_{k=1}^{t_i - 1} (R^T)^{t_i - k - 1} \Gamma_i^{t_i - 1} (R^T)^{k-1}. \qquad (21)$$

Now set $R_1 = R$, and for $i = 1, 2, \ldots, u$, do:

$$\begin{aligned} \Gamma_i &= \sum_{k=1}^{t_i - 1} (R^T)^{t_i - k - 1} R_i^{t_i - 1} (R^T)^{k-1}, \\ R_{i+1} &= R_i - \Gamma_i. \end{aligned}$$

Combining (17)-(21), one finds that at the end of the iteration, $R_{u+1} = R - \Gamma = \bar{R} + \hat{\Gamma}$. Therefore, since $\hat{\Gamma}^{d_1 - 1} = 0$,

$$R_{u+1}^{d_1-1} = V_1 J_{d_1}^{d_1-1} V_1^H = (V_1 e_{d_1})(V_1 e_1)^H. \qquad (22)$$

Thus an SVD of $R_{u+1}^{d_1-1}$ provides the first and last columns of V_1 up to a unitary constant. The remaining columns are found via (16) and then (14) yields the ZF equalizers.

subchannel	$\mathbf{h}_{10}$	$\mathbf{h}_{11}$	$\mathbf{h}_{12}$	$\mathbf{h}_{20}$	$\mathbf{h}_{21}$	$\mathbf{h}_{22}$	$\mathbf{h}_{23}$	$\mathbf{h}_{30}$	$\mathbf{h}_{31}$
1	1.0	0.2	0.5	-0.9	0.2	-0.3	-0.6	-0.1	0.4
2	-0.4	0.5	0.8	0.3	-0.4	-0.2	0.3	0.4	0.3
3	0.5	-0.1	-0.9	1.0	0.2	0.1	-0.3	-0.1	0.3
4	-0.3	0.7	0.3	0.5	-0.1	-0.4	0.3	0.3	0.9

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

6. SIMULATION RESULTS

We considered an equivalent channel as in (5) with $q' = 3$ surviving kernels after precoding, with lengths $l_1' = 2$, $l_2' = 3$, $l_3' = 1$ and exponents $m_1 = 1$, $m_2 = 2$, $m_3 = 3$. The modulation is 4-PAM with equiprobable levels $\pm\frac{1}{3}$ and ± 1, and the number of subchannels available after precoding was $Np = 4$. The coefficients are shown in table 1, giving a 2.2 dB linear-to-nonlinear distortion ratio.

For illustration purposes the phase ambiguity inherent to the algorithm was removed before computing the error rates, which were averaged based on 100 independent runs. The noise is white Gaussian with variance σ_v^2. The SNR is defined as

$$\mathrm{SNR} = 10 \log_{10} \frac{\sigma_x^2}{Np\sigma_v^2} \sum_{j=0}^{l_1} \|\mathbf{h}_{1j}\|^2.$$

We chose an equalizer length $M = 7$. Figure 3 shows the symbol error rate attained with the ZF equalizers as a function of the SNR, when $K = 1000$ symbols were used for estimating the channel output covariance matrices. For simplicity these covariance matrices were not denoised. For this example, the equalizers with intermediate associated delays ($1 \le d \le 7$) performed similarly and about 3 dB better than those of extremal delays ($d = 0$ and $d = 8$). Figure 4 shows the variation of the symbol error rate with the number of symbols K for covariance estimation, for a SNR value of 20 dB.

7. CONCLUSIONS

We have presented a simple precoding scheme that considerably alleviates the need for a large number of subchannels for the blind equalization of nonlinear Volterra channels, by reducing the effective number of nonlinear kernels that need to be equalized. For i.i.d. symmetrically distributed input data, we show that the channel can be equalized from its second order output statistics iff in its effective precoded model no nonlinear kernel has the same length as the linear kernel, and no odd degree nonlinear kernel has larger length than the linear one. An algorithm for the computation of the equalizers is presented.

8. REFERENCES

[1] G.B. Giannakis and E. Serpedin, "Linear multichannel blind equalizers of nonlinear FIR Volterra channels", *IEEE Trans. on 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 ICASSP*, vol. 5, pp. 2769-2772, Istanbul, Turkey.

[3] M. Schetzen, "The Volterra and Wiener theories of nonlinear systems", Wiley, 1980.

[4] L. Tong and S. Perreau, "Multichannel blind equalization: from subspace to maximum likelihood methods", *Proc. of the IEEE*, vol. 86. no. 10, pp. 1951-1968, Oct. 1998.

[5] L. Tong, G. Xu and T. Kailath, "Blind identification and equalization based on second-order statistics: a time-domain approach", *IEEE Trans. on Information Theory*, vol. 40, no. 2, pp. 340-350, March 1994.

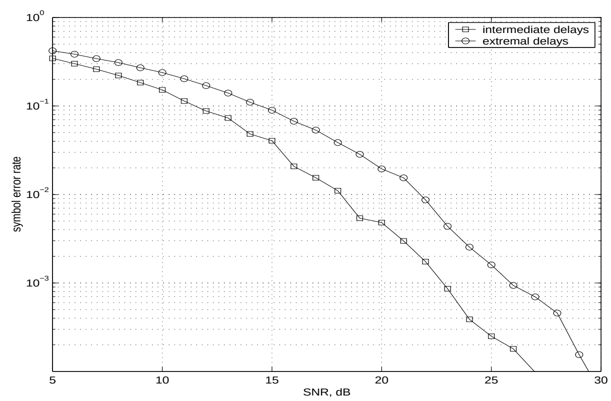

Figure 3: SER vs. SNR. $K = 1000$ symbols for covariance estimation.

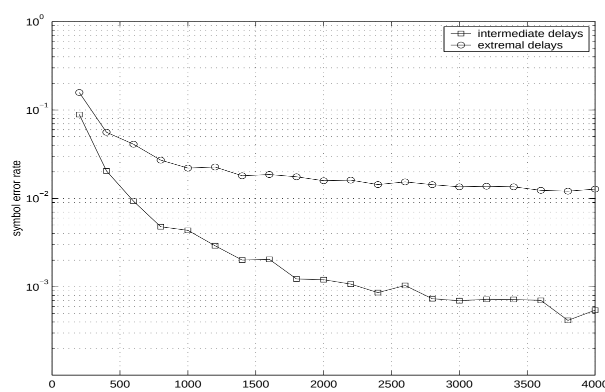

Figure 4: SER vs. sample size K. SNR = 20 dB.