A KERNEL CANONICAL CORRELATION ANALYSIS ALGORITHM FOR BLIND EQUALIZATION OF OVERSAMPLED WIENER SYSTEMS

Steven Van Vaerenbergh, Javier Vía and Ignacio Santamaría

{steven,jvia,nacho}@gtas.dicom.unican.es
Department of Communications Engineering
University of Cantabria, Spain

ABSTRACT

In this paper we present an algorithm for blind equalization of single-input multiple-output (SIMO) nonlinear systems, in which each nonlinear channel is a Wiener system. The proposed method combines ideas from blind linear SIMO identification with kernel canonical correlation analysis (kernel CCA) to identify the nonlinearities. It is shown in the paper that the blind equalization problem can be solved in an iterative manner, alternating between a CCA problem (to estimate the linear filters) and a kernel CCA problem (to estimate the memoryless nonlinearities). The resulting algorithm can be applied to the general case of nonlinear SIMO systems with P outputs. Simulations are included to demonstrate its effectiveness.

1. INTRODUCTION

In the last decade there has been a great interest in blind identification and equalization methods. In digital communications, blind methods permit channel identification or equalization without the need to send known training signals, thus saving bandwidth. In particular, the problem of blind identification of single-input multiple-output (SIMO) linear channels has received considerable attention [1, 2]. In this case, blind identification can be accomplished by resorting only to the second-order statistics (SOS) of the channel output.

While a lot of attention has gone to the analysis of linear SIMO systems, many real-life systems exhibit nonlinear characteristics. Recently, a growing amount of research has been conducted on nonlinear system identification [3]. Nonlinear dynamical system models generally have a high number of parameters, although many problems can be sufficiently well approximated by simplified block-based models. The model consisting of a cascade of a linear dynamic system and a memoryless nonlinearity is known as the Wiener system, as illustrated in Fig. 1. Wiener systems are frequently used in contexts such as digital satellite communications [4], optical fibre communications [5] and digital magnetic recording.

This work was supported by MEC (Ministerio de Educación y Ciencia) under grant TEC2007-68020-C04-02 TCM (MultiMIMO) and FPU grant AP2005-5366.

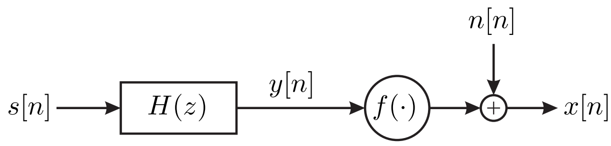


Fig. 1. A Wiener system with additive noise $n[n]$.

A number of supervised approaches have been proposed to identify or equalize these systems, ranging from black-box approaches using different types of structures and training criteria [6, 7], to approaches that explicitly exploit the system structure [8, 9, 10]. However, very little work has been done on blind identification methods. Blind methods generally assume some knowledge on the input signal statistics and/or the channel model. A few blind methods to identify Wiener systems can be found in [11, 12]. Blind identification methods for other nonlinear systems such as Volterra models have also been presented in [13, 14].

In this paper we present a blind method to identify and equalize nonlinear SIMO systems that consist of various Wiener systems, as illustrated in Fig. 2. These systems could represent a sensor array in which every sensor exhibits a nonlinear behavior, or they could be obtained by oversampling the output of a nonlinear communications channel [1]. The presented method combines ideas from the blind linear SIMO identification method in [1] and from the supervised nonlinear equalization technique discussed in [9]. It performs at the same time a kernel-based regression to learn the nonlinearities and a least-squares (LS) method to retrieve the linear channels. The assumptions made by our method are that every nonlinearity in the SIMO Wiener system is invertible, that the linear channels share no common zeros and, finally, that the maximum channel order is known.

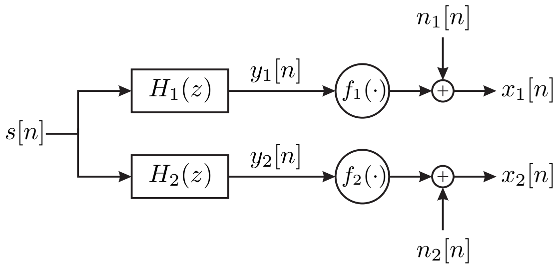


Fig. 2. A SIMO system consisting of 2 Wiener systems.

2. PROBLEM SETTING

Consider a nonlinear SIMO system, in which each channel is modeled as a Wiener system. An example with only two outputs is shown in Fig. 2. In a general case with P outputs the system can be modeled as

$$y_i[n] = \sum_{j=0}^{L-1} h_i[j] s[n-j] \quad (1)$$

$$x_i[n] = f_i(y_i[n]) + n_i[n], \quad (2)$$

where $s[n]$ represents the input symbol sent at time instant n, $h_i[j]$ is the j-th coefficient of the i-th linear FIR channel $H_i(z)$, $f_i(\cdot)$ is the nonlinearity of channel i and $n_i[n]$ represents additive Gaussian noise, for $i = 1, \ldots, P$ and $n = 0, \ldots, N-1$. Without loss of generality, L represents the maximum channel order (which we assume to be known).

The problem considered in this paper is to recover the transmitted signal $s[n]$ when only the output signals $x_i[n]$ are observed.

3. BLIND SIMO WIENER SYSTEM EQUALIZATION

The solution we propose to the equalization problem is mainly based on the linear identification method presented in [1]. In the following this linear method is explained briefly for a 1×2 linear SIMO system, although the generalization to more output channels is straightforward.

3.1. Blind identification of a linear SIMO system

Taking a block of N observations, define the matrix

$$\mathbf{X}_i = \begin{bmatrix} x_i[n+L-1] & \cdots & x_i[n] \\ \vdots & \ddots & \vdots \\ x_i[n+N-1] & \cdots & x_i[n+N-L] \end{bmatrix}, \quad (3)$$

for $i = 1, 2$. Denoting the estimate of the channel impulse response vectors as

$$\hat{\mathbf{h}}_i = \left[\hat{h}_i[0], \ldots, \hat{h}_i[L-1]\right]^T, \quad (4)$$

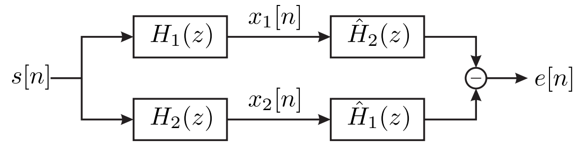


Fig. 3. A blind identification scheme for a linear SIMO model without noise.

it can be easily verified that in a noiseless case the solution should satisfy

$$\mathbf{X}_1 \hat{\mathbf{h}}_2 = \mathbf{X}_2 \hat{\mathbf{h}}_1, \quad (5)$$

as illustrated in Fig. 3.

In order to avoid the trivial solution $\hat{\mathbf{h}}_i = \mathbf{0}$, a restriction must be applied to the solution. Typical restrictions in communications are either to fix the norm of the filters $\hat{\mathbf{h}}_i$, or to fix the norm of the output signal $\mathbf{X}_i \hat{\mathbf{h}}_j$.

A restriction on the filter norm was used in [1] to develop the LS method, also referred to as cross-relation. This problem consists of minimizing the cost function

$$J_{LS} = \frac{1}{2} \left\| \mathbf{X}_1 \hat{\mathbf{h}}_2 - \mathbf{X}_2 \hat{\mathbf{h}}_1 \right\|^2 \quad \text{s.t. } \|\hat{\mathbf{h}}_1\|^2 + \|\hat{\mathbf{h}}_2\|^2 = 1, \quad (6)$$

which is equivalent to the following eigenvalue problem

$$\begin{bmatrix} \mathbf{X}_1^T \mathbf{X}_1 & -\mathbf{X}_1^T \mathbf{X}_2 \\ -\mathbf{X}_2^T \mathbf{X}_1 & \mathbf{X}_2^T \mathbf{X}_2 \end{bmatrix} \hat{\mathbf{h}} = \beta \hat{\mathbf{h}}. \quad (7)$$

The solution $\hat{\mathbf{h}} = [\hat{\mathbf{h}}_2^T, \hat{\mathbf{h}}_1^T]^T$ is found as the eigenvector corresponding to the smallest eigenvalue.

If, instead, the constraint is applied to the norm of output signals as in [2], the cost function to minimize turns out to be

$$J_{CCA} = \frac{1}{2} \left\| \mathbf{X}_1 \hat{\mathbf{h}}_2 - \mathbf{X}_2 \hat{\mathbf{h}}_1 \right\|^2 \quad \text{s.t. } \|\mathbf{X}_1 \hat{\mathbf{h}}_2\|^2 + \|\mathbf{X}_2 \hat{\mathbf{h}}_1\|^2 = 1. \quad (8)$$

This is a canonical correlation analysis (CCA) problem, and its solution is given by the principal eigenvector of the following generalized eigenvalue problem (GEV)

$$\begin{bmatrix} \mathbf{X}_1^T \mathbf{X}_1 & \mathbf{X}_1^T \mathbf{X}_2 \\ \mathbf{X}_2^T \mathbf{X}_1 & \mathbf{X}_2^T \mathbf{X}_2 \end{bmatrix} \hat{\mathbf{h}} = \beta \begin{bmatrix} \mathbf{X}_1^T \mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2^T \mathbf{X}_2 \end{bmatrix} \hat{\mathbf{h}}. \quad (9)$$

Once the channels $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ have been estimated, they can be used to obtain an equalizer by applying the zero-forcing (ZF) or the minimum mean square error (MMSE) approach. Note that both the LS algorithm and the CCA-based algorithm require knowledge of the maximum channel order L, and they assume the linear channels share no common zeroes.

When we consider that each channel of the system is replaced by a Wiener system, the scheme of Fig. 4 can be used for blind identification, where the influence of the nonlinearities $f_i(\cdot)$ is removed first, by estimating the inverse nonlinearities $g_i(\cdot) = \hat{f}_i^{-1}(\cdot)$.

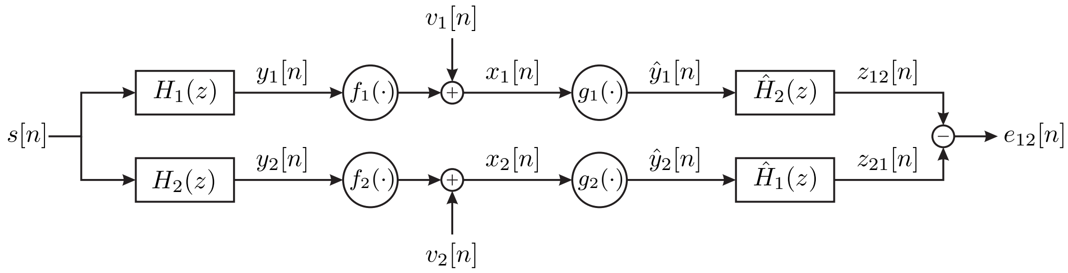


Fig. 4. The identification diagram for a SIMO system consisting of 2 Wiener subsystems, in which $g_i(\cdot) = \hat{f}_i^{-1}(\cdot)$.

To estimate the inverse nonlinearities $g_i(\cdot)$, we will apply a nonparametric identification approach based on kernel methods. Nonparametric approaches do not assume that the nonlinearity corresponds to a given model, but rather let the training data decide which characteristic fits them best.

3.2. Nonlinear regression with kernel methods

Kernel methods [15] are based on a nonlinear transformation Φ of the data from the input space to a high-dimensional *feature space* $\mathcal{H}$, where it is more likely that a problem can be solved in a linear manner,

$$\mathbf{\Phi} : \mathbb{R}^m \rightarrow \mathcal{H}$$

$$\Phi(\mathbf{x}) = \tilde{\mathbf{x}}.$$

Scalar products in feature space can be calculated without the explicit knowledge of the nonlinear transformation Φ, by applying the corresponding *kernel function* $\kappa(\cdot, \cdot)$ on pairs of data points in the input space,

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) := \langle \tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_j \rangle = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle. \quad (10)$$

This property, which is known as the "kernel trick", allows to perform any scalar product-based algorithm in the feature space by solely replacing the scalar products with the kernel function in the input space.

Most kernel algorithms use a *kernel matrix* $\mathbf{K}_i$, which is constructed by applying the kernel function on pairs of points: $k_i(m, n) = \kappa(x_i[m], x_i[n])$, with $m, n = 1, \ldots, N$. An often used kernel function is the Gaussian kernel with width σ

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right),$$

which implies an infinite dimensional feature space [15].

Nonlinear regression with kernels is possible by representing the nonlinearity as a *kernel expansion*

$$\hat{y}_i[n] = g_i(x_i[n]) = \sum_{m=1}^{M} \hat{\alpha}_i[m]\kappa(x_i[n], x_i^s[m]), \quad (11)$$

where $x_i^s[m]$ are called the *support vectors* of the nonlinear representation. In the following we will use the variable $k_i^s(n, m) = \kappa(x_i[n], x_i^s[m])$ to simplify the notation. In a first approach, all available points $x_i[n]$ will be used as support vectors, i.e., $M = N$.

At this point it should be clear that once the inverse nonlinearities $g_i(\cdot)$ have been estimated, retrieval of the linear FIR channels $\hat{\mathbf{h}}_i$ is straightforward through a linear SIMO identification technique such as the CCA- or LS-based algorithms. Given only the outputs $x_i[n]$ of the system, direct estimation of these nonlinearities is difficult, however, since no information on the input signal $s[n]$ is available.

Therefore, since separate estimation of the linear and nonlinear parts of this system is difficult, we will design an algorithm that allows us to obtain both the linear filters and the nonlinearities simultaneously, through a single cost function.

3.3. Proposed cost function

First, we will treat the case where the observed system has only two outputs. Given the representation of the nonlinearity $g_i(\cdot)$ as in (11), the output of the proposed identification scheme can be written as

$$z_{12}[n] = \sum_{i=0}^{L-1} \sum_{m=1}^{M} \hat{h}_2[i] k_1^s(n-i, m)\hat{\alpha}_1[m]. \quad (12)$$

In matrix notation, this becomes

$$z_{12}[n] = \hat{\mathbf{h}}_2^T \mathbf{K}_1[n] \hat{\boldsymbol{\alpha}}_1, \quad (13)$$

where the l-th row of $K_1[n]$ contains the elements from $k_1^s(n + l - 1, 1)$ till $k_1^s(n + l - 1, M)$. The expression for $z_{21}[n]$ is found in the same manner. Combining N output samples of each channel into the vectors $\mathbf{z}_{12}$ and $\mathbf{z}_{21}$, we obtain the following cost function to minimize:

$$J_2 = \|\mathbf{z}_{12} - \mathbf{z}_{21}\|^2 \quad \text{s.t. } \|\mathbf{z}_{12}\|^2 + \|\mathbf{z}_{21}\|^2 = 1. \quad (14)$$

3.4. Proposed iterative solution

The minimization problem (14) has no direct analytical solution. However, if $\hat{\boldsymbol{\alpha}}_1$ and $\hat{\boldsymbol{\alpha}}_2$ were available, it would be possible to obtain the corresponding optimal filters $\hat{\mathbf{h}}_2$ and $\hat{\mathbf{h}}_1$

by applying linear CCA. Moreover, since we are representing the nonlinearities $g_1(\cdot)$ and $g_2(\cdot)$ as linear combinations of support vectors, a similar operation can be carried out to estimate these: if $\hat{\mathbf{h}}_2$ and $\hat{\mathbf{h}}_1$ are given, (14) can be solved to find the optimal coefficients of the kernel expansions $\hat{\boldsymbol{\alpha}}_{\mathbf{1}}$ and $\hat{\boldsymbol{\alpha}}_{\mathbf{2}}$. This suggests an iterative scheme that alternates between updating the linear channels $\hat{\mathbf{h}}_i$ and the memoryless nonlinearities $\hat{\boldsymbol{\alpha}}_{\boldsymbol{i}}$. Convergence is guaranteed because each update may either decrease or maintain the cost.

3.4.1. Iteration 1: given $\hat{\boldsymbol{\alpha}}_i$, obtain $\hat{\mathbf{h}}_i$

If estimates of $\hat{\boldsymbol{\alpha}}_1$ and $\hat{\boldsymbol{\alpha}}_2$ are given, Eq. (12) shows that the output $z_{12}[n]$ of the identification scheme can be obtained as

$$z_{12}[n] = \sum_{i=0}^{L-1} \hat{h}_2[i]\hat{y}_1[n-i], \tag{15}$$

where $\hat{y}_1[n-i]$ is calculated with (11). In matrix form this can be written as $\mathbf{z}_{12} = \hat{\mathbf{Y}}_1\hat{\mathbf{h}}_2$, where n-th row of the matrix $\hat{\mathbf{Y}}_1$ contains the elements from $\hat{y}_1[n]$ until $\hat{y}_1[n+L-1]$. The minimization problem (14) can be rewritten as minimizing

$$J_{\mathbf{h}} = \|\hat{\mathbf{Y}}_1\hat{\mathbf{h}}_2 - \hat{\mathbf{Y}}_2\hat{\mathbf{h}}_1\|^2 \quad \text{s.t. } \|\hat{\mathbf{Y}}_1\hat{\mathbf{h}}_2\|^2 + \|\hat{\mathbf{Y}}_2\hat{\mathbf{h}}_1\|^2 = 1, \tag{16}$$

which can be solved by standard linear CCA.

3.4.2. Iteration 2: given $\hat{\mathbf{h}}_i$, obtain $\hat{\boldsymbol{\alpha}}_i$

If estimates of $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ are given, Eq. (12) shows that the output $z_{12}[n]$ of the identification scheme can be obtained as

$$z_{12}[n] = \sum_{m=1}^{M} w_1[n,m]\hat{\alpha}_1[m], \tag{17}$$

where the variable $w_1[n,m] = \sum_{i=0}^{L-1} \hat{h}_2[i]k_1(n-i,m)$ is introduced. In matrix form this can be written as $\mathbf{z}_{12} = \mathbf{W}_1\hat{\boldsymbol{\alpha}}_1$, where the n-th row of the matrix $\mathbf{W}_1$ contains the elements $w_1[n,1]$ until $w_1[n,M]$. The minimization problem (14) can be rewritten as minimizing

$$J_{\hat{\boldsymbol{\alpha}}} = \|\mathbf{W}_1\hat{\boldsymbol{\alpha}}_1 - \mathbf{W}_2\hat{\boldsymbol{\alpha}}_2\|^2 \quad \text{s.t. } \|\mathbf{W}_1\hat{\boldsymbol{\alpha}}_1\|^2 + \|\mathbf{W}_2\hat{\boldsymbol{\alpha}}_2\|^2 = 1. \tag{18}$$

If all data points $x_i[n]$ are used as support vectors in the kernel expansion (11), i.e., if $M = N$ (which implies $\hat{\boldsymbol{\alpha}}_{\boldsymbol{i}} \in \mathbb{R}^N$), the dimensionality of this problem is significantly higher than its linear counterpart (16). This leads to various difficulties.

First of all, problem (18) will suffer from overfitting when sufficiently “rich” kernel functions are used, i.e., kernels that correspond to feature spaces whose dimension m' is much higher than the number of available data points N. This occurs for instance for the Gaussian kernel, whose feature space is infinite dimensional. Second, the GEV corresponding to this problem requires the retrieval of eigenvectors of $2N \times 2N$ matrices, which in this case implies a high computational cost.

Algorithm 1 Equalization algorithm for nonlinear SIMO channels.

Initialization: obtain $\hat{\mathbf{h}}_i$ by solving the LS problem (7).
Construct the kernel matrices $\mathbf{K}_i$ from $x_i[n]$.
Perform kernel PCA to obtain the reduced matrices $\mathbf{W}_i$.
repeat
  CCA1: With given $\hat{\mathbf{h}}_i$, update $\hat{\boldsymbol{\alpha}}_i$ by solving (18).
  CCA2: With given $\hat{\boldsymbol{\alpha}}_i$, update $\hat{\mathbf{h}}_i$ by solving (16).
until Convergence
Obtain $s[n]$ from $\hat{y}_i[n]$ and $\hat{\mathbf{h}}_i$ by applying linear ZF or MMSE equalizers.

Overfitting is a common issue in kernel CCA that can be solved in different manners [16]. Common workarounds include adding regularization to the problem or reducing the dimensionality of the problem by applying kernel PCA [17]. In this case a dimensionality reduction is desired since at the same time it will avoid overfitting and reduce the computational load. Specifically, kernel PCA reduces the kernel matrix $\mathbf{K}_i \in \mathbb{R}^{N\times N}$ to

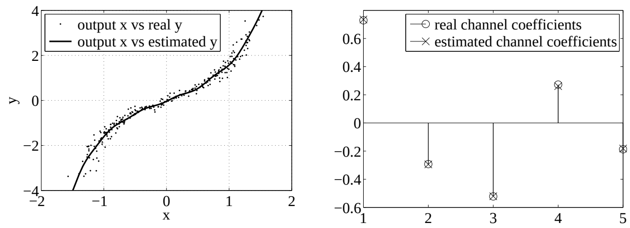
$$\mathbf{V}_i\boldsymbol{\Sigma}_i\mathbf{V}_i^T \approx \mathbf{K}_i, \tag{19}$$

where $\boldsymbol{\Sigma}_i \in \mathbb{R}^{M\times M}$ is a diagonal matrix containing the M largest eigenvalues of $\mathbf{K}_i$ and $\mathbf{V}_i \in \mathbb{R}^{N\times M}$ contains the M corresponding eigenvectors. This allows us to redefine the variable $w_1[n,m]$ as $w_1[n,m] = \sum_{i=0}^{L-1} h_2[i]v_1(n-i,m)$, where $v_1(n,m)$ is the n-th element of the m-th eigenvector in $\mathbf{V}_1$. Thanks to this reduction, the dimensions of the matrices $\mathbf{W}_i$ in (18) are reduced to $N \times M$, with $M \ll N$, and the solutions $\hat{\boldsymbol{\alpha}}_{\boldsymbol{i}}$ can be found by applying CCA.

3.5. Extensions and Further Comments

Analogously to many other iterative techniques, the performance of the proposed approach can depend on the initialization of the linear channels and nonlinearities. Here, we propose to obtain an initial estimate of the linear channels $\hat{\mathbf{h}}_i$ by first applying the LS algorithm from [1] to the outputs $x_i[n]$, i.e, in the first iteration, the estimated nonlinearities are $g_i(x_i[n]) = x_i[n]$. Furthermore, we must note that the final target of the proposed algorithm consists in recovering the source signal $s[n]$. Thus, after obtaining the outputs $\hat{y}_i[n]$ and the linear channels $\hat{\mathbf{h}}_i$, the input can be easily recovered by means of a linear ZF or MMSE equalizer.

In the general case of a system with P sensors, the cost function needs to take into account the difference between each pair of outputs. Note that the output signal $\mathbf{z}_{ij}$ represents the signal $\mathbf{x}_i$ after being transformed by $g_i(\cdot)$ and filtered by



(a) Nonlinearity estimate. (b) Linear filter coefficients.

Fig. 5. Identification results on the 1×3 Wiener SIMO system. (a) shows the noisy output $x_3[n]$ vs. the real internal signal $y_3[n]$, and $x_3[n]$ vs. the estimated $\hat{y}_3[n]$. (b) shows the estimated filter coefficients of $\mathbf{h}_3$ vs. the real coefficients.

$\mathbf{h}_j$. The cost function to minimize now becomes

$$J_P = \sum_{\substack{i,j=1 \\ i \neq j}}^{M} \|\mathbf{z}_{ij} - \mathbf{z}_{ji}\|^2 \quad \text{s.t.} \sum_{\substack{i,j=1 \\ i \neq j}}^{M} \|\mathbf{z}_{ij}\|^2 = 1, \qquad (20)$$

and the resulting algorithm is analogous to that of the two-channel case. The final iterative technique for P output channels is summarized in Algorithm 1.

Finally, we must note that when the SIMO system is obtained by oversampling, the P nonlinearities will be the same. Obviously, this can be exploited to obtain a more accurate estimate. The corresponding GEV can be found easily, but it is omitted here due to space restrictions.

4. EXPERIMENTS

We experimentally tested the proposed algorithm with some numerical examples. All tests were conducted on data sets of $N = 256$ data symbols. The fraction of the signal energy discarded by the kernel PCA procedure in the initialization phase was fixed as 10^{-14}. The resulting number of kept eigenvectors was between $M = 11$ and $M = 15$. In all experiments convergence was obtained in less than 20 iterations.

The first system used is a 1×3 Wiener SIMO system with linear filters $\mathbf{h}_1 = [0.6172, 0.6247, 0.3373, -0.0349, -3.2957]^T$, $\mathbf{h}_2 = [-0.8601, 0.1532, -0.1888, -0.6264, 0.9985]^T$ and $\mathbf{h}_3 = [1.3271, -0.1472, -0.4786, 0.6682, 0.0045]^T$, respectively. The nonlinearity was the same for all the channels, namely $f_i(x) = \tanh(0.8x) + 0.1x$.

A first test was conducted on this system with a zero mean and unit variance Gaussian source. The power of the white Gaussian noise after the nonlinearities was fixed to obtain a 20dB SNR. Fig. 5 shows the true and estimated linear filter and nonlinearity for one of the branches of the Wiener SIMO system, after 15 iterations of the algorithm.

We then compared the proposed algorithm to the linear CCA-based equalizer from [2]. Averages were taken over 50 independent Monte-Carlo simulations, and the MSE was calculated between the true and the estimated input signal. The results are shown in Fig. 6. The curve with solid black dots was obtained by applying the linear CCA-based equalizer on the system that only contained the linear channels $\mathbf{h}_1$, $\mathbf{h}_2$ and $\mathbf{h}_3$. The same algorithm was tested on the nonlinear 1×3 SIMO Wiener system, resulting in the curve with white squares. The curve marked with circles was obtained by applying the proposed blind method on the 1×3 SIMO Wiener system, and in the last curve (dashed line) we show the results when the supervised method of [9] was applied to this system. The proposed method obtains results that are very close to those obtained by the supervised method.

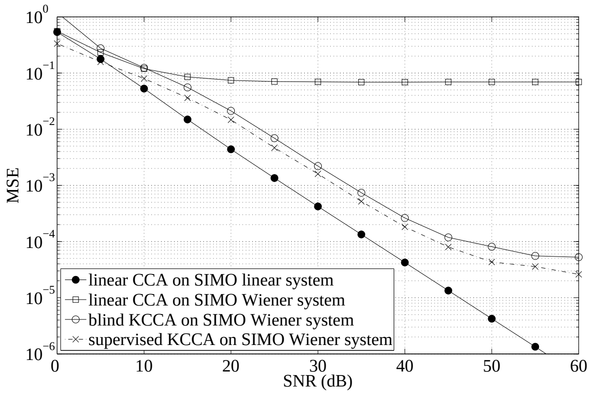


Fig. 6. MSE comparison for different algorithms.

For the second test we compared three SIMO Wiener systems with different numbers of outputs. System 1 was a 1×2 SIMO Wiener system with $\mathbf{h}_1$ and $\mathbf{h}_2$ as defined in the first experiment. System 2 was the discussed 1×3 system. System 3 was a 1×4 SIMO Wiener system that included all three previous Wiener systems and a new linear channel $\mathbf{h}_4 = [-0.1155, -0.9170, 0.5605, 0.4862, -0.8004]^T$ in its fourth branch. The nonlinearity was maintained, and we exploited the fact that it was the same for each channel. The results are shown in Fig. 7. The same test was repeated for a system with a binary input $s[n] \in \{-1, 1\}$, but now we did not exploit the information that the nonlinearity was the same for each channel. The results are shown in Fig. 8.

5. CONCLUSIONS

We proposed a blind equalization algorithm for nonlinear SIMO systems in which every channel is a Wiener system. Basically, the method iterates between a CCA algorithm for estimating the linear channel and a KCCA algorithm for estimating the memoryless nonlinearities. First results show that this iterative algorithm converges fast and achieves performance that is very close to a related supervised method. Future research lines include a comparison to other blind nonlinear equalization methods such as [13, 14].

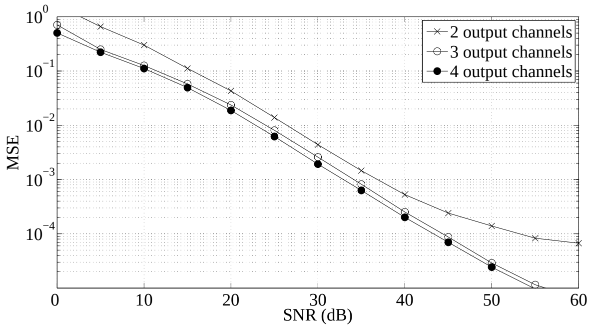


Fig. 7. Blind identification results for different SIMO Wiener systems with a Gaussian input, where the algorithm took into account that each channel had the same nonlinearity.

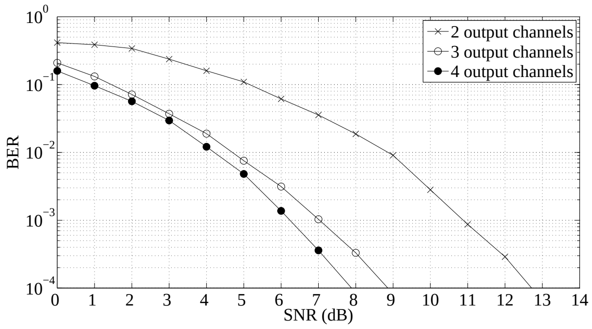


Fig. 8. Results for SIMO Wiener systems with a binary input.

6. REFERENCES

[1] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. on Signal Processing*, vol. 43, no. 12, pp. 2982–2993, Dec. 1995.

[2] J. Vía, I. Santamaría, and J. Pérez, "Effective channel order estimation based on combined identification / equalization," *IEEE Trans. on Signal Processing*, vol. 54, no. 9, pp. 3518–3526, Sept. 2006.

[3] O. Nelles, *Nonlinear System Identification*, Springer-Verlag, Berlin, Germany, 2000.

[4] G. Kechriotis, E. Zarvas, and E. S. Manolakos, "Using recurrent neural networks for adaptive communication channel equalization," *IEEE Trans. on Neural Networks*, vol. 5, pp. 267–278, Mar. 1994.

[5] R. Kawakami Harrop Galvão, S. Hadjiloucas, A. Izhac, V.M. Becerra, and J.W. Bowen, "Wiener-system subspace identification for mobile wireless mm-wave networks," *IEEE Trans. on Vehicular Technology*, vol. 56, no. 4, pp. 1935–1948, July 2007.

[6] I. Santamaría, D. Erdogmus, and J. C. Principe, "Entropy minimization for supervised digital communications channel equalization," *IEEE Trans. on Signal Processing*, vol. 50, no. 2, pp. 1184–1192, May 2002.

[7] T. Adali and X. Liu, "Canonical piecewise linear network for nonlinear filtering and its application to blind equalization," *Signal Processing*, vol. 61, no. 2, pp. 145–155, Sept. 1997.

[8] M. Pawlak, Z. Hasiewicz, and P. Wachel, "On nonparametric identification of Wiener systems," *IEEE Trans. on Sig. Proc.*, vol. 55, no. 2, pp. 482–492, Feb. 2007.

[9] S. Van Vaerenbergh, J. Vía, and I. Santamaría, "Adaptive kernel canonical correlation analysis algorithms for nonparametric identification of Wiener and Hammerstein systems," *EURASIP Journal on Advances in Signal Processing*, vol. 1, pp. 10–22, Apr. 2008.

[10] J.E. Cousseau, J.L. Figueroa, S. Werner, and T.I. Laakso, "Efficient nonlinear Wiener model identification using a complex-valued simplicial canonical piecewise linear filter," *IEEE Trans. on Signal Processing*, vol. 55, pp. 1780–1792, May 2007.

[11] A. Taleb, J. Sole, and C. Jutten, "Quasi-nonparametric blind inversion of Wiener systems," *IEEE Trans. on Signal Processing*, vol. 49, no. 5, pp. 917–924, May 2001.

[12] J.C. Gómez and E. Baeyens, "Subspace-based blind identification of IIR Wiener systems," in *Proceedings of 15th European Signal Processing Conference (EUSIPCO)*, Poznań, Sept. 2007.

[13] G. 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.

[14] R. López-Valcarce and S. Dasgupta, "Blind equalization of nonlinear channels from second-order statistics," *IEEE Trans. on Signal Processing*, vol. 49, no. 12, pp. 3084–3097, Dec. 2001.

[15] B. Schölkopf and A. J. Smola, *Learning with Kernels*, The MIT Press, Cambridge, MA, 2002.

[16] D. R. Hardoon, S. Szedmak, and J. Shawe-Taylor, "Canonical correlation analysis; an overview with application to learning methods," *Neural Computation*, vol. 16, no. 12, pp. 2639–2664, Dec. 2004.

[17] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.