IEEE Workshop on Stat. Signal and Array Processing, pp. 181–185, Pocono Manor, Usa, Aug. 14-16, 2000.

BLIND IDENTIFICATION OF LINEAR-QUADRATIC CHANNELS WITH USUAL COMMUNICATION INPUTS

Nicolas PETROCHILOS [1,2]

(1) CAS, Dept EE, Delft Univ. of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands
petro@ieee.org

Pierre COMON [2]

(2) I3S, Algorithmes-Euclide-B
2000 route des Lucioles, BP 121
F-06903 Sophia-Antipolis cedex, France
comon@unice.fr

ABSTRACT

This article presents a method to blindly identify linear quadratic channels (LQC). The method is designed for the single-input/single-output (SISO) case with white inputs with specific distributions (as those usually found in digital communications). Using High-Order Statistics (HOS) of the input, the method is able to match the third-order moments with the LQC model, yielding an original simple relation. Several simulations are performed and show a fair accuracy given sufficiently long observation records.

1. INTRODUCTION

Nonlinear systems provide a better approximation to real life channels, and many examples of nonlinearities can be found in nonlinear control systems [5], hydrodynamics [4], satellite communication systems [1], or underwater acoustics, among others. Blind methods are attractive when the input is unknown, and to avoid the reduction of the information rate caused by the insertion of training sequences.

Blind identification of Volterra systems has been already widely studied in the past. For instance, in [7], the authors derive the cumulant-matching equations, allowing to blindly identify a pure real quadratic system, with i.i.d. inputs of unknown distribution. Next in [2], P.Bondon goes much further, and derives identifiability conditions, when two input sequences are observed, one Gaussian and one non Gaussian.

In this paper, we focus our attention on linear-quadratic systems, with specific discrete inputs, encountered in $n-$PSK and QAM digital modulations. So this contribution differs from the previous ones in two respects: the system is not purely quadratic, and the inputs are imposed to be discrete and of known distribution. The scope is thus less general.

This work was partly supported by ENS Lyon, ENSEA, TU-Delft, and the RNRT project "Paestum". The first author thanks A. Trindade, A. Heldring, S. Halford, and A. Elmilady for their moral support, E. Serpedin for useful discussions, and G. Giannakis for having attracted his attention on the non-linear blind identification problem.

2. MODEL FORMULATION

The problem is modeled here by the parameterization of the channel and by the statistics of the inputs.

2.1. Volterra kernel model

The model is described by the noisy output of a nonlinear system moving average Volterra model (which can be of any order). Sampling at a rate T_S and restricting to the Linear-Quadratic case, the channel can be modeled as:

$$\begin{aligned} y(n) &= \sum_{l=0}^{L_1} h_1(l)\, x(n-l) + v(n) \\ &+ \sum_{i,j=0}^{L_2} h_2(i,j)\, x(n-i)\, x(n-j) \end{aligned} \tag{1}$$

where $x(n)$ is the input signal, $v(n)$ denotes the additive noise, and h_n is called the nth-order Volterra non-linear operator (here, we only have the linear and the quadratic term: $h_1(l_1)$ and $h_2(l_1, l_2)$). Without loss of generality, we consider that h_n is symmetric in its arguments [6, pp.80-81].

2.2. Usual communication inputs

For the sake of convenience, denote:

$$\epsilon_{ab} \stackrel{\text{def}}{=} E\left[x^a x^{b*}\right].$$

In this article, we consider inputs commonly used in digital communications, sharing the high-order properties:

$$\epsilon_{21} = \epsilon_{31} = \epsilon_{32} = \epsilon_{41} = \epsilon_{42} = 0 \tag{2}$$

Among these inputs, two groups have been identified (see [9] and [3]):

- Distributions that are symmetric about both axes in the complex plane: $p(z) = f(\Re\{z\}) \cdot g(\Im\{z\})$. Corresponding random variables can be rewritten as $z = \varepsilon + \jmath\varepsilon'$, where ε and ε' are real, independent, and symmetrically distributed, and $\jmath^2 \stackrel{\text{def}}{=} -1$. QAM constellations, in digital communications, belong to this class.
- Discrete distributions that are invariant by a rotation of an angle of the form $\frac{2\pi}{K}$, $(K \in N)$. QPSK, double QPSK, and any n-PSK are included in this class as soon that $n \geq 4$.

3. CHANNEL IDENTIFICATION

First, the basis of the identification process is presented, then the algorithms are derived, and a proof of uniqueness is eventually given.

3.1. Moment-matching relations

Consider the following assumptions:

(AS1) The channel is Linear-Quadratic of finite known length.

(AS2) The input is stationary independent identically distributed (i.i.d.), and must comply with the properties (2); $\sigma_x^2 = \epsilon_{11}$ and $\mu_{4x} = \epsilon_{22}$ are also assumed to be known.

(AS3) The noise is signal-independent white Gaussian.

Let us now define the complex bicorrelation as:

$$C_{12y}(l,k) \stackrel{\text{def}}{=} E\left\{y^*(n)y(n+l)y(n+k)\right\} \qquad (3)$$

Under assumptions (**AS1-AS3**), the bicorrelation of the output (3) and the channel model (1) should match, which gives the following relations:

$$C_{12y}(l,k) = \sum_{i,j} h_1(i+l)h_1(j+k)\hat{h}_2^*(i,j) \qquad (4)$$

with $(l,k) \in [-L_2, L_1] \times [-L_2, L_1]$, and where $\hat{h}_2^*(i,j) \stackrel{\text{def}}{=} \left[2\epsilon_{11}^2 + \delta(i-j)(\varepsilon_{22} - 2\epsilon_{11}^2)\right] \cdot h_2^*(i,j)$. The Z-transform of $C_{12}(l,k)$ gives in the (Z_1, Z_2) domain:

$$S_{12y}(Z_1, Z_2) = H_1(Z_1)\, H_1(Z_2)\, \hat{H}_2^*\left(\frac{1}{Z_1^*}, \frac{1}{Z_2^*}\right) \qquad (5)$$

Equations (4) or (5) form the core of the algorithms subsequently proposed. By stacking the elements of $C_{12y}(l,k)$ in a matrix **C12Y** as:

$$\mathbf{C12Y} \stackrel{\text{def}}{=} \begin{bmatrix} C_{12y}(-L_2,-L_2) & \cdots & C_{12y}(-L_2,L_1) \\ \vdots & & \vdots \\ C_{12y}(L_1,-L_2) & \cdots & C_{12y}(L_1,L_1) \end{bmatrix},$$

we get the matrix formulation:

$$\mathbf{C12Y} = \mathbf{A}^\top \cdot \mathbf{B} \cdot \mathbf{A} \qquad (6)$$

where $\mathbf{A}$ is the $(L_2+1) \times (L_1+L_2+1)$ Upper Triangular Band (UTB) Töplitz matrix containing $\mathbf{h}_1 = [h(0) \ldots h(L_1)]$ in the first row and zeros elsewhere:

$$\mathbf{A} \stackrel{\text{def}}{=} \begin{bmatrix} \boxed{\quad\mathbf{h}_1\quad} & & \mathbf{0} \\ & \ddots & \ddots \\ \mathbf{0} & & \boxed{\quad\mathbf{h}_1\quad} \end{bmatrix}$$

while $\mathbf{B}$ is symmetric complex and contains the values of the kernel $\hat{h}_2^*$:

$$\mathbf{B} \stackrel{\text{def}}{=} \begin{bmatrix} \hat{h}_2^*(L_2,L_2) & \cdots & \hat{h}_2^*(0,L_2) \\ \vdots & & \vdots \\ \hat{h}_2^*(L_2,0) & \cdots & \hat{h}_2^*(0,0) \end{bmatrix}$$

We propose to identify the channel coefficients by using either relation (5), or (6) with the estimate $\widehat{C_{12y}}(l,k) \stackrel{\text{def}}{=} \frac{1}{N}\sum_{n=1}^{N} y^*(n)\, y(n+l)\, y(n+k)$.

One can notice that **C12Y** is a (L_1+L_2+1) square matrix of rank (L_2+1). This observation allows to detect the length of the channels $(\mathbf{h}_1, \mathbf{h}_2)$ from an estimate $\widehat{\mathbf{C12Y}}$ of **C12Y**.

3.2. Proposed algorithms

We propose several algorithms: (i) a Root-Finding method (RF), (ii) a Sub-Space Intersection method (SSI), (iii) a method that forces the row span to have certain triangular properties (UTB), and (iv) an iterative Multidimensional Search method (MS).

(i) One can give several values to Z_2 in (5), and get several functions of Z_1: $F_{z_2}(Z_1)$. These functions $F_{z_2}(Z_1)$ share the roots of $H_1(Z_1)$: r_i, which are detected by clustering. The channel $h_1(n)/h_1(0)$ is the inverse Z-transform of $\prod_{i=1}^{L_1}(Z - r_i)$, and one can build $\mathbf{A}$. Denoting $\mathbf{A}^-$ the Moore-Penrose pseudo-inverse of $\mathbf{A}$, $\hat{h}_2$ is recovered via the "deconvolution":

$$\mathbf{B} = \mathbf{A}^{\top -} \cdot \mathbf{C12Y} \cdot \mathbf{A}^-. \qquad (7)$$

(ii) Alternatively, one can factorize the matrix **C12Y** in order to recover the vector $\mathbf{h}_1$ in a similar fashion as in [8]. In the noiseless case, given that $\mathbf{B}$ has no null eigenvalue, the matrix model (6) implies clearly that:

$$\begin{cases} \text{row}(\mathbf{A}) = \text{row}(\mathbf{C12Y}) \\ \text{col}(\mathbf{A}^\top) = \text{col}(\mathbf{C12Y}) \end{cases} \qquad (8)$$

Considering the singular value decomposition (SVD) of the symmetric complex matrix $\mathbf{C12Y} = \mathbf{V}^\top . \mathbf{S} . \mathbf{V}$, we define $\hat{\mathbf{V}}$ as the L_2+1 first rows of $\mathbf{V}$, associated with the L_2+1 dominant singular values. Let $\hat{\mathbf{V}}^{(i)}$ be the $L_2+1 \times L_1+1$ submatrix extracted from $\hat{\mathbf{V}}$ that gathers the columns i to L_1+i. Then the conditions (8) are restated as: $\mathbf{h}_1 \in \hat{\mathbf{V}}^{(i)}$, $\forall i \in [1,..,L_2+1]$. Thus $\mathbf{h}_1{}^{\mathrm{H}}$ can be obtained by computing the dominant right singular vector of the matrix $\mathcal{V}$ containing all $\hat{\mathbf{V}}^{(i)}$ stacked one above the other:

$$\mathcal{V} = \begin{bmatrix} \hat{\mathbf{V}}^{(1)} \\ \vdots \\ \hat{\mathbf{V}}^{(L_2+1)} \end{bmatrix}.$$

Then the matrix $\mathbf{B}$ can be estimated afterwards by the "deconvolution" procedure (7).

(iii) Another technique consists of forcing the UTB structure of $\mathbf{A}$ beforehand by combining the rows of matrix $\hat{\mathbf{V}}$; this is possible because of Lemma 1. Then, one extracts the L_2+1 dimensional row vectors $\mathbf{v}^{(i)}$ contained in the UTB matrix $\mathbf{T}\hat{\mathbf{V}}$, and stacks them in a matrix $\mathcal{V}$. The rest of the procedure is identical to the previous approach (ii).

(iv) Lastly, one can perform an iterative search in the $(L_1+L_2(L_2+1)/2)$ dimensional space of the matrix product of (6) in order to find the parameters $\theta\,(h_1,h_2)$ that minimize the error in the sense of the Frobenius norm:

$$\theta(h_1,h_2) = \arg\min_{\theta} \left\| \mathbf{C12Y} - \left[\mathbf{A}^{\top} \cdot \mathbf{B} \cdot \mathbf{A}\right](\theta) \right\|_F^2$$

3.3. Uniqueness

Lemma 1 *Let N and P be two positive integers. Under certain regularity conditions, any $N \times (N+P)$ rectangular matrix $\mathbf{M}$ can be put in UTB form by premultiplication by a square invertible matrix $\mathbf{T}$. The matrix $\mathbf{T}$ is unique up to an invertible diagonal multiplicative matrix.*

Proof: The constructive algorithm is very similar to Gaussian elimination. Assume there are two matrices $\mathbf{T}_1$ and $\mathbf{T}_2$ such that $\mathbf{M} = \mathbf{T}_1\,\mathbf{U}_1$ and $\mathbf{M} = \mathbf{T}_2\,\mathbf{U}_2$, where $\mathbf{U}_1$ and $\mathbf{U}_2$ are UTB. Then, considering the N first columns of both sides shows that the matrix $\mathbf{T}_1\mathbf{T}_2^{-1}$ relates two Lower Triangular (LT) matrices, and is thus LT itself. Similarly, considering the N last columns shows that $\mathbf{T}_1\mathbf{T}_2^{-1}$ is Upper Triangular (UT). Thus, it is diagonal, which eventually shows that $\mathbf{T}_1$ and $\mathbf{T}_2$ are related by a diagonal multiplicative matrix. □

Lemma 2 *Any symmetric complex matrix $\mathbf{C}$ can be factorized as $\mathbf{C} = \mathbf{L}\,\mathbf{L}^{\top}$, where $\mathbf{L}$ is lower triangular. Matrix $\mathbf{L}$ is unique up to the post-multiplication of a diagonal matrix Δ formed of signs $\{\pm 1\}$.*

Proposition 3 *If $\mathbf{B}$ is square full rank, and $\mathbf{A}$ is UTB, then the decomposition of a complex symmetric matrix $\mathbf{C} = \mathbf{A}^{\top}\,\mathbf{B}\,\mathbf{A}$ is unique up to a multiplicative diagonal matrix.*

Proof: The proposition is a direct consequence of lemmas 1 and 2. It is easily seen that if $(\mathbf{A},\mathbf{B})$ is solution, then so is $(\Delta\mathbf{A}, \Delta^{-1}\mathbf{B}\Delta^{-1})$, where Δ is any diagonal regular matrix. □

Corollary 4 *Let $\mathbf{B}$ be full rank symmetric complex and $\mathbf{A}$ Töplitz UTB. When the decomposition of a symmetric matrix as $\mathbf{C12Y} = \mathbf{A}^{\top}\,.\,\mathbf{B}\,.\,\mathbf{A}$ exists, then it is unique up to a scalar multiplicative factor.*

Proof: From proposition 3, if $\mathbf{A}$ is solution, then so is $\Delta\mathbf{A}$, with Δ diagonal. But because $\mathbf{A}$ is Töplitz, $\Delta\mathbf{A}$ can be Töplitz only if Δ is proportional to the Identity matrix. □

4. SIMULATIONS

In order to illustrate the Root-Finding (RF) method step by step, we first present a typical example with only RF and MS methods. Later we show a a more exhaustive study with all the methods. Because we are mainly interested in direct methods, the MS is given only as a reference.

In all simulations, the input x was 4-PSK. We used the real channel given by [9] ($h_1 = [1, 0.5, -0.8, 1.6, 0.4]$ and $h_2 = [1, 0.6; 0.6, -0.3]$).

Typical example: The input is QPSK; the number of samples is 16284 points; and the SNR is 10 dB. Figure *I.a.* illustrates the clustering method. It shows all the roots calculated for different Z_2, the true roots, and the ones estimated by the method, the estimated roots (stars) are fairly accurate and match the real ones (square). Figure *I.b.* shows the spectra of the real and estimated linear channels. Both estimated spectra are fairly accurate.

Computer comparisons: A first study showed that the estimation noise of $\hat{C}_{12y}$ is rapidly predominant over the additive noise contribution. As expected, the Gaussian noise does not interact in the third-order moment as soon as the length of integration is long enough. So we mainly tried to estimate the influence of the number of samples. For each number of samples we took 1000 independent realizations, and the SNR is 10 dB. For each realization, we estimated $\hat{C}_{12y}$, on which we applied all the algorithms. Since in our case $\mathbf{C12Y}$ is 6×6, the most computational intensive step is its estimation for the direct methods. Due to its iterative nature, up to several thousand of samples, the most intensive step for the MS method is the multidimensional search.

Figure *II* presents the influence of the integration length on the mean and variance of both estimates.

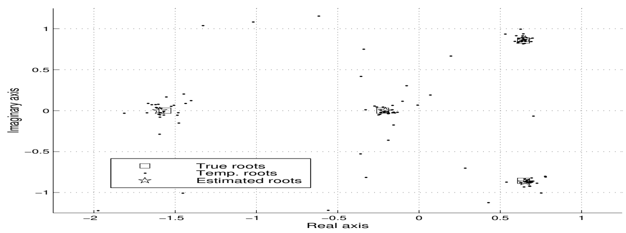
Figures $II.a.$ and $II.b.$ show that all methods converge to the true channel, the bias behaves well from 4096 points. The RF is the slowest method to converge to the expected value, while the MS is the fastest to converge. The SSI and the UTB follow similar patterns.

Figures $II.c.$ and $II.d.$ present the variances of both methods. The variances follow approximately a linear slope. It is difficult to decide which method behaves the best. One can notice that the MS has stationary performance after 64000 samples, this is because this method was implemented in a too rustic way, and it happens that a few times the MS algorithm is stuck in local minima, thus degrading the quality of the standard deviation. While not visible on the figure, the best method varies for each element of $\mathbf{h}_1$, and generally around 4096 samples the best method changes. Nevertheless, above 4096 samples clearly the best method is the SSI.

The variance shows well the usual problem with High Order Statistics: in order to have consistent high-order moment estimate, the integration length must be long enough: a minimum of 8192 seems to be required here.

5. CONCLUDING REMARKS

Several methods have been proposed to blindly identify a linear-quadratic channel for communication applications. The idea is to use the specificities of the distribution of the inputs. The methods have shown to converge with a good accuracy, with a rather large number of samples.



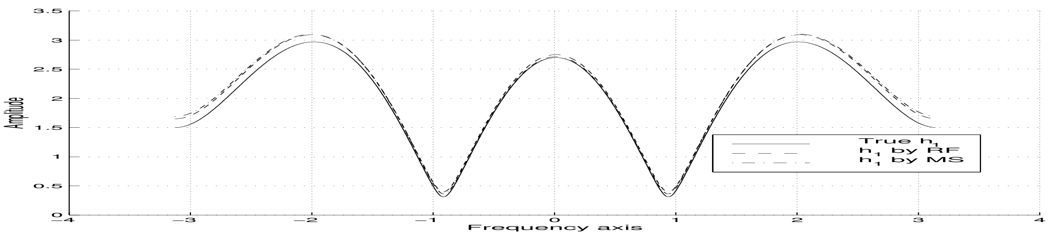


Figure I: *Example of Identification: 10dB, 16284 points.*

REFERENCES

[1] S. BENEDETTO, E. BIGLIERI, V. CASTELLANI, *Digital Transmission Theory*, Prentice-Hall Inc., New Jersey, 1987.

[2] P. BONDON, M. KROB, "Blind identifiability of quadratic stochastic system", *IEEE trans. on Information Theory*, vol. 41, no. 1, pp. 245–254, Jan. 1998.

[3] N. PETROCHILOS, "Elements for blind identification of non-linear channels", supervision by G. Giannakis, Master's thesis, ENSEA / ENS Lyon, Sept. 1996, In archive of ENSEA, France.

[4] E. J. POWERS, S. IM et al., "Applications of hos to nonlinear hydrodynamics", in *IEEE-ATHOS Workshop on Higher-Order Statistics*, Begur, Spain, 12–14 June 1995, pp. 414–418.

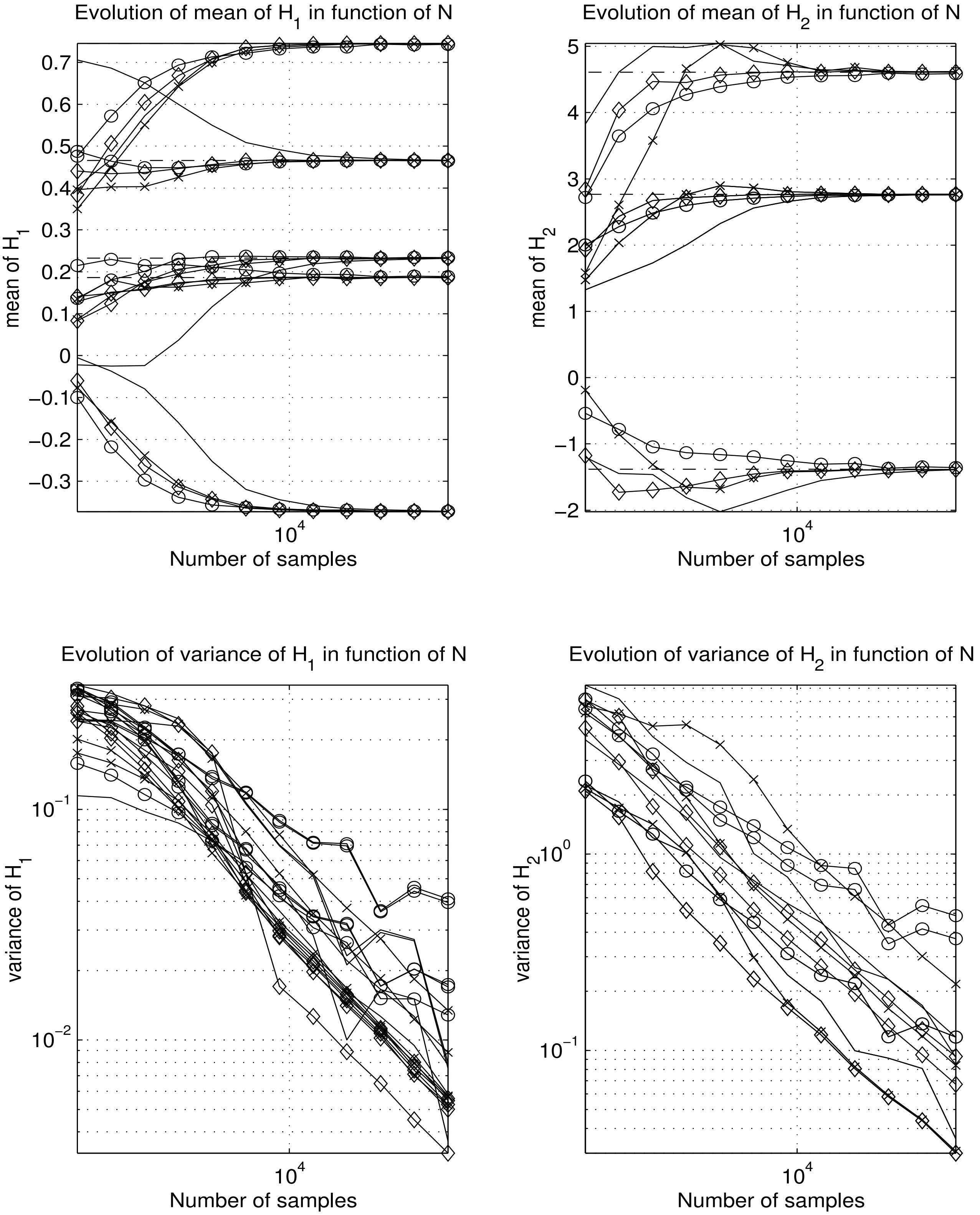
[5] W. J. RUGH, *Nonlinear System Theory*, Johns Hopkins Univ. Press, Baltimore, MD, 1981.

[6] M. SCHETZEN, *The Volterra and Wiener Theories of Nonlinear Systems*, Wiley, New York, 1980.

[7] H-Z. TAN, Z-Y. MAO, "Blind identifiability of quadratic non-linear systems in higher-order statistics domain", *Int. Jour. Adapt. Control Signal Processing*, vol. 12, pp. 567–577, 1998.

[8] A. J. van der VEEN, S. TALWAR, A. PAULRAJ, "A subspace approach to blind space-time signal processing for wireless communication systems", *IEEE trans. on Signal Processing*, vol. 45, no. 1, pp. 173–190, Jan. 1997.

[9] G. T. ZHOU, G. B. GIANNAKIS, "Nonlinear channel identification and performance analysis with PSK inputs", in *Proc. of 1st IEEE Signal Processing Workshop on Wireless Communications*, Paris, France, 16-18 April 1997, pp. 337–340.



,

Figure II: *Means and standard deviations for all methods with* 1000 *independent realizations. Simple line: RF, x-: UTB, d-: SSI, o-:MS.*