SEMI-BLIND PARAMETERS AND SYMBOLS JOINT ESTIMATION FOR A HAMMERSTEIN TYPE CHANNEL USING PRECODING

Alain Y. Kibangou and Gérard Favier

Laboratoire I3S/CNRS/ UNSA
2000 Route des Lucioles, B.P. 121, 06903 Sophia Antipolis cedex, France
e-mail: {kibangou, favier}@i3s.unice.fr

ABSTRACT

Some communication systems can be represented with a Hammerstein model; that is the case of the radio over fiber downlink channel. In order to recover the input signal from the received signal, we propose a semi-blind approach based on an input precoding inducing redundancy. The suggested approach makes it possible to jointly estimate the channel and the input sequence. The channel is estimated using a pilot sequence then the symbols are estimated by solving a triangular system of polynomial equations or by calculating the inverse of a given polynomial.

1. INTRODUCTION

The continuing increase in customer demand for broadband applications coupled with mobile cellular and personal communications has become a global phenomenon. Optical fiber based wireless access is a key technology because of its potential to increase system capacity, to enable wideband access and to cover special areas such as tunnels, supermarkets, airports, etc. Such a scheme is especially useful for indoor applications with micro and pico cellular architectures [1]. This technology combines two media: radio and optical. Typically, the optical part is used to interconnect a central radio processing facility with a remote radio antenna, the latter providing coverage to wireless broadband users. When the fiber is short and the radio frequency is only a few GHz, effects of fiber dispersion are negligible [2]. However, nonlinear distorsion of the electrical to optical conversion process becomes the major limitation. This distorsion can be modeled with a static nonlinearity.

Typically, the fiber-wireless access is constituted by two parts: the wireless channel and the nonlinear link. The wireless channel is modeled with a tapped delay line filter. The fiber-wireless downlink, i.e. from the base station to the portable unit, is then modeled by a static nonlinearity (the nonlinear link) followed by a linear filter (the wireless channel), i.e. a Hammerstein model. A discrete-time P-th order Hammerstein model of memory M is described by the following input-output relation:

$$y(n) = \sum_{p=1}^{P}\sum_{i=0}^{M} h_p(i)u^p(n-i) \tag{1}$$

where u and y are respectively the input and the output signals and h_p is the p-th order kernel of the Hammerstein model.

Identification of the Hammerstein model has been mainly studied using input/output data. There are few works devoted to the blind identification of this model. These works only consider second order Hammerstein models and use higher order statistics when the input is Gaussian [3, 4, 5] ; these methods exhibit a high computational cost. This paper considers a semi-blind approach for SISO Hammerstein models of any order under the assumption that the input belongs to a finite symbol set. New methods are proposed for both channel estimation and symbol detection using a precoding.

The organization of the paper is as follows. In the next section, the precoding scheme is presented. In Section 3, the channel estimation method is explained whereas symbols detection procedures are derived in Section 4. The proposed methods are then illustrated by means of simulation results in Section 5 before concluding the paper in Section 6.

2. THE PRECODING SCHEME

We consider a block transmission system where each $L \times 1$ block $U(n) = (s(n)\, u(nL-1)\, \cdots\, u(nL-L+1))^T$ is constituted by one pilot symbol $s(n)$, known to the receiver, followed by $(L-1)$ information-bearing symbols $u(nL-i)$, $i = 1, \cdots, L-1$. In order to decouple channel estimation and unknown symbols estimation, each block $U(n)$ is precoded by using a $(K \times L)$ precoding matrix $\mathbf{C}$. The $(K \times 1)$ precoded block is given by:

$$\begin{aligned} X(n) &= (x(nK)\, x(nK-1) \cdots x(nK-K+1))^T \\ &= \mathbf{C}U(n). \end{aligned}$$

The information rate decreases with the ratio L/K. By using the precoding matrix proposed in [6]:

$$\mathbf{C} = \begin{pmatrix} C_M & 0_{(2M+1)\times(L-1)} \\ 0_{(L-1)\times 1} & \mathbf{I}_{(L-1)\times(L-1)} \end{pmatrix}, \qquad (2)$$

where C_M is the $(2M+1)\times 1$ vector defined as:

$$C_M = \left(\underbrace{0\cdots 0}_{M}\,1\,\underbrace{0\cdots 0}_{M}\right)^T, \qquad (3)$$

the components $x(nK-i), i = 0, \cdots, K-1$, of $X(n)$ are given by:

$$x(nK-i) = \begin{cases} s(n) & if \quad i = M \\ u(nL-i+2M) & if \quad i = 2M+1, \cdots \\ & \quad\quad \cdots, K-1 \\ 0 & if \quad i = 0, \cdots, M-1, \\ & \quad\quad M+1, \cdots, 2M \end{cases} \qquad (4)$$

where $K = L + 2M$. Due to the structure of $\mathbf{C}$, no matrix multiplication is involved to produce the precoded data block, only $2M$ zero-inserting operations are necessary.

Let us consider the ($K \times 1$) received block of data $Y(n)$ such as:
$Y(n) = (y(nK)\; y(nK-1)\; \cdots\; y(nK-K+1))^T$, with

$$\begin{aligned} y(nK-i) &= \sum_{p=1}^{P}\sum_{j=0}^{M} h_p(j)x^p(nK-i-j) + w(nK-i) \\ &= \sum_{p=1}^{P}\sum_{k=i}^{M+i} h_p(k-i)x^p(nK-k) + w(nK-i), \end{aligned} \qquad (5)$$

$i = 0, \cdots, K-1$, where w is the additive noise.
From (4), $x(nK-k)$ is different from zero if and only if $k = M$ or $2M+1 \le k \le K-1$. Consequently the relation (5) can be decomposed into two parts:

- For $i = 0, \cdots, M$

$$y(nK-i) = \sum_{p=1}^{P} h_p(M-i)x^p(nK-M) + w(nK-i); \qquad (6)$$

- For $i = M+1, \cdots, K-1$

$$y(nK-i) = \sum_{p=1}^{P}\sum_{k=i}^{M+i} h_p(k-i)x^p(nK-k) + w(nK-i). \qquad (7)$$

When $M+1 \le i \le 2M$, using (4), equation (7) can be rewritten as:

$$y(nK-i) = \sum_{p=1}^{P}\sum_{k=2M+1}^{M+i} h_p(k-i)x^p(nK-k) + w(nK-i). \qquad (8)$$

When $2M+1 \le i \le K-1$, by taking the causality of the channel into account, equation (7) can also be rewritten as (8). Therefore, this equation is valid for $i = M+1, \cdots, K-1$. Note that for guaranteing that all the inputs $x(nK-k)$, $k = 2M+1, \cdots, M+i$, belong to the n-th block, the maximum value of i is chosen such that $M+i \le K-1$, or $i \le K-M-1$. Then using (4), equations (6) and (8) become respectively:

$$y(nK-i) = \sum_{p=1}^{P} h_p(M-i)s^p(n) + w(nK-i), i = 0, \cdots, M, \qquad (9)$$

and

$$\begin{aligned} y(nK-i) = &\sum_{p=1}^{P}\sum_{k=2M+1}^{M+i} h_p(k-i)u^p(nL-k+2M) \\ &+ w(nK-i),\; i = M+1, \cdots, K-M-1. \end{aligned} \qquad (10)$$

As the pilot sequence $s(n)$ is known by the receiver, the channel can be estimated from (9), and then the symbols can be reconstructed from (10).

3. CHANNEL ESTIMATION

The relation (9) can be viewed as the input/ouput equation of a static polynomial channel indexed by i, with the parameters $\Theta_i = \begin{pmatrix} h_1(M-i) & \cdots & h_P(M-i) \end{pmatrix}^T$, $i = 0, \cdots, M$. The parameters of the $M+1$ subchannels can be estimated in the MMSE (Minimum Mean Square Error) sense. Then, they are given by:

$$\hat{\Theta}_i = \mathbb{E}[S(n)S^T(n)]^{-1}\mathbb{E}[y(nK-i)S(n)],$$

where $S(n) = \left(s(n)\; \cdots\; s^P(n)\right)^T$, and $\mathbb{E}$ denotes the mathematical expectation. In the sequel we present an efficient way of computing $\hat{\Theta}_i$ without any matrix inversion. The following assumptions are adopted:
A1: $\mu_j = \mathbb{E}[s^j(n)] = 0$ for odd powers j.
A2: $w(.)$ is zero-mean, white and independent of $s(.)$.

Let us consider the Hilbert space of P-th degree polynomials associated with the inner product:

$$< A(s(n)), B(s(n)) >= \mathbb{E}\left[A(s(n))B(s(n))\right],$$

where $A(.)$ and $B(.)$ are P-th degree polynomials. A basis of this space is given by the following set of monomials $\left\{s^j(n)\right\}_{j=0}^{P}$. By applying the well-known Gram-Schmidt orthonormalization procedure we get an orthonormal basis $\{\mathcal{P}_j(s(n))\}_{j=0}^{P}$ such as:

$$\begin{aligned} &\mathcal{P}_0(s(n)) = 1, \\ &\tilde{\mathcal{P}}_j(s(n)) = s^j(n) - \sum_{k=0}^{j-1}\mathbb{E}[s^j(n)\mathcal{P}_k(s(n))]\mathcal{P}_k(s(n)), \\ &\mathcal{P}_j(s(n)) = \frac{\tilde{\mathcal{P}}_j(s(n))}{\|\tilde{\mathcal{P}}_j(s(n))\|}. \end{aligned} \qquad (11)$$

For example, for $P = 3$, we get the following set of polynomials:

$$\begin{aligned}
\mathcal{P}_0(s(n)) &= 1, \\
\mathcal{P}_1(s(n)) &= \frac{s(n)}{\sqrt{\mu_2}}, \\
\mathcal{P}_2(s(n)) &= \frac{1}{\sqrt{\mu_4 - \mu_2^2}}\left(s^2(n) - \mu_2\right), \\
\mathcal{P}_3(s(n)) &= \frac{1}{\sqrt{\mu_2^2\mu_6 - \mu_4^2\mu_2}}\left(\mu_2 s^3(n) - \mu_4 s(n)\right).
\end{aligned}$$

The obtained polynomials exhibit the following properties, δ denoting the Kronecker symbol:

$$\begin{aligned}
\mathbb{E}\left[\mathcal{P}_j(s(n))\mathcal{P}_k(s(n))\right] &= \delta_{j,k}, && (12)\\
\mathbb{E}\left[\mathcal{P}_j(s(n))\right] &= 0, \quad j > 0, && (13)\\
\mathbb{E}[s^k(n)\mathcal{P}_j(s(n))] &= 0, \quad if \quad k < j. && (14)
\end{aligned}$$

Properties (12) and (13) result respectively from the basis orthonormality and from the orthogonality of the polynomials $\mathcal{P}_j(s(n))$, $j > 0$, with $\mathcal{P}_0(s(n))$. To prove the property (14), one can note that the expansion of the monomial $s^k(n)$ is such as $s^k(n) = \sum_{l=0}^{k} b_l \mathcal{P}_l(s(n))$, where b_l is an expansion coefficient. Then:

$$\begin{aligned}
\mathbb{E}[s^k(n)\mathcal{P}_j(s(n))] &= \sum_{l=0}^{k} b_l \mathbb{E}[\mathcal{P}_l(s(n))\mathcal{P}_j(s(n))] \\
&= \sum_{l=0}^{k} b_l \delta_{l,j} \\
&= \begin{cases} 0, & if \quad k < j \\ b_j, & else \end{cases}.
\end{aligned}$$

The expansion of the polynomial

$$V_i(s(n)) = \sum_{p=1}^{P} h_p(M-i)s^p(n)$$

on the orthonormal basis $\{\mathcal{P}_j(s(n))\}_{j=0}^{P}$ yields:

$$V_i(s(n)) = \sum_{p=0}^{P} f_{j,i}\mathcal{P}_j(s(n)) = \breve{\Theta}_i^T \breve{S}(n), \qquad (15)$$

with

$$f_{j,i} = \mathbb{E}\left[V_i\left(s\left(n\right)\right)\mathcal{P}_j\left(s\left(n\right)\right)\right],$$

$$\breve{\Theta}_i = \left(f_{0,i} \ \cdots \ f_{P,i}\right)^T,$$

and

$$\breve{S}(n) = \left(\mathcal{P}_0(s(n)) \ \cdots \ \mathcal{P}_P(s(n))\right)^T.$$

Consequently the expansion coefficients, optimal in the MMSE sense, are given by:

$$\hat{f}_{j,i} = \mathbb{E}[y(nK-i)\mathcal{P}_j(s(n))], j = 0, \cdots, P, i = 0, \cdots, M. \qquad (16)$$

Then, using the property (14), the Hammerstein model parameters are obtained by solving the following triangular system of equations:

$$\hat{f}_{j,i} = \sum_{p=j}^{P} h_p(M-i)\mathbb{E}[s^p(n)P_j(s(n))]. \qquad (17)$$

4. SYMBOLS ESTIMATION

Once the channel parameters estimated, the information-bearing symbols are to be reconstructed. Two methods are proposed. The first one, called TPS-root (*Triangular polynomial System root*), needs the solution of a triangular system of polynomial equations, and the second one is based on the determination of a polynomial inverse; it is called PI method.

One can note that by making the following change of coordinates $l = k - 2M$, $j = i - M$, equation (8) can be rewritten as:

$$\begin{aligned}
y(nK - j - M) = \sum_{p=1}^{P}\sum_{l=1}^{j} h_p(l - j + M)u^p(nL - l) \\
+ w(nK - j - M), \; j = 1, \cdots, K - M - 1.
\end{aligned} \qquad (18)$$

4.1. The TPS-root method

Due to the causality of the kernels h_p it is obvious that (18) has a triangular structure and can be solved for $u(nL - i)$, $i = 1, \cdots, L - 1$, using back substitution. In the noiseless case, the corresponding system of equations can be written as:

$$\begin{aligned}
P_1\left(u(nL-1)\right) &= 0 \\
P_2\left(u(nL-1), u(nL-2)\right) &= 0 \\
\vdots &= \vdots \\
P_{L-1}\left(u(nL-1), \cdots, u(nL-L+1)\right) &= 0
\end{aligned}$$

where each $P_j(.)$, $j = 1, \cdots, L - 1$ is the polynomial defined by:

$$\begin{aligned}
&P_j(u(nL-1), \cdots, u(nL-j)) = \\
&\quad \sum_{p=1}^{P}\sum_{l=1}^{j} h_p(l - j + M)u^p(nL - l) - y(nK - j - M).
\end{aligned}$$

Let us assume that the first $(j - 1)$ equations have been already solved and $\hat{u}(nL - l)$, $l = 1, \cdots, j - 1$, are their respective solutions. By substituting these values in the expression of $P_j(u(nL-1),\cdots,u(nL-j))$ we get a polynomial in the single variable $u(nL-j)$:

$$\tilde{P}_j(u(nL-j)) = P_j(u(nL-1),\cdots,u(nL-j)) - \sum_{p=1}^{P}\sum_{l=1}^{j-1} h_p(l-j+M)\hat{u}^p(nL-l). \tag{19}$$

Thus the symbol $u(nL-j)$ can be computed as a root of $\tilde{P}_j(u(nL-j))$. As $\tilde{P}_j(u(nL-j))$ is a P-th order polynomial we can determine the set of its roots $\mathcal{R}_{j,n} = \{r_{p,n}\}_{p=1}^{P}$. Knowing the alphabet $\{\alpha_k\}_{k=1}^{q}$ of the transmitted signal, the estimated symbol $\hat{u}(nL-j)$ is then given by:

$$\hat{u}(nL-j) = \arg\min_{r\in\mathcal{R}_{j,n}} \prod_{k=1}^{q} |r-\alpha_k|^2 . \tag{20}$$

Note that the principle of symbol estimation by solving polynomial equations was first suggested by Redfern and Tong Zhou [7] for Volterra systems. However in their method, the channel coefficients are assumed to be known. In the best of our knowledge the method proposed in this paper is the first one which allows to jointly estimate both the channel and the symbols.

4.2. The PI method

As for the TPS-root method, let us assume that the $(j-1)$ first symbols $u(nL-l)$, $l=1,\cdots,j-1$, are known. Then (18) can be rewritten as:

$$\tilde{y}(nK-j-M) = \sum_{p=1}^{P} h_p(M)u^p(nL-j) = H_M(u(nL-j)), \tag{21}$$

where

$$\tilde{y}(nK-j-M) = y(nK-j-M) - \sum_{p=1}^{P}\sum_{l=1}^{j-1} h_p(l-j+M)\hat{u}^p(nL-l),$$

and $H_M(u(nL-j))$ is a polynomial the coefficients of which are $h_p(M)$. Assuming that this polynomial is invertible, it exists a polynomial J_M such as:

$$J_M(\tilde{y}(nK-j-M)) = J_M\left(H_M\left(u(nL-j)\right)\right) = u(nL-j) \tag{22}$$

The polynomial H_M characterizes a mapping from a finite set (the alphabet of the source) to another finite set. Reciprocally J_M is the inverse mapping from a finite set of data to another. As the alphabet of the source and the polynomial H_M are known we can use the Lagrange interpolation formula:

Definition: Let $\mathcal{Q}$ be a polynomial in the variable x. Given the data (β_j,α_j), $j=1,\cdots,q$, such as $\mathcal{Q}(\beta_j)=\alpha_j$, we have:

$$\mathcal{Q}(x) = \sum_{j=1}^{q} \alpha_j \prod_{k\neq j}\left(\frac{x-\beta_k}{\beta_j-\beta_k}\right). \tag{23}$$

In order to apply this formula to our problem, we must first determine the images of the elements of the source alphabet by the mapping characterized by H_M, i.e.

$$\beta_j = H_M(\alpha_j)$$

where α_j is an element of the source alphabet. Knowing the data (β_j,α_j), $j=1,\cdots,q$, we can apply the Lagrange formula to find the polynomial J_M. Thus the estimated symbols are given by:

$$\hat{u}(nL-j) = J_M(\tilde{y}(nK-j-M)). \tag{24}$$

The PI method only requires the calculation of the J_M polynomial while the TPS-root method needs to solve several polynomial equations. From a computational cost point of view, the PI method needs less operations than the TPS-root method. Note again that the PI-method differs from the pth-order inverse method of Volterra system equalization introduced by Schetzen [8]. The pth-order inverse exists only if the first order kernel is minimum phase. Here, the existence of the inverse polynomial is always guaranteed as explained above.

The TPS-root and the PI methods can be applied for systems having a minimum phase first order kernel or not. However a possible drawback is due to their DFE (Decision Feedback Equalizer) nature. An estimation error can then be propagated on the other symbols. This propagation is hopefully limited to the current block of data.

5. SIMULATION RESULTS

We consider the following Hammerstein channel:

$$\begin{aligned} y(n) \;=\; & u(n) + 0.45u(n-1) + 0.35u(n-2) \\ & +0.3u^2(n) - 0.6u^2(n-2) \\ & +0.1u^3(n) - 0.01u^3(n-1) + w(n). \end{aligned}$$

According to the precoding scheme presented in the previous sections, the subchannels of the SIMO representation of the Hammerstein channel are given by

$$\begin{aligned} y(nK) &= s(n) + 0.3s^2(n) + 0.1s^3(n) + w(nK), \\ y(nK-1) &= 0.45s(n) - 0.01s^3(n) + w(nK-1), \\ y(nK-2) &= 0.35s(n) - 0.6s^2(n) + w(nK-2), \end{aligned}$$

where s is the pilot sequence.

The source belongs to the 4-PAM alphabet. The transmitted block is of size $L = 5$ and $K = 9$. The pilot sequence is also a 4-PAM sequence independent of the informative signal. A white Gaussian noise was added to the received signal. For different levels of Signal-to-Noise-Ratio (SNR), 100 independent experiments were carried out. The channel estimation process was started after the transmission of 200 data blocks.

The results of the channel estimation are given in Table 1.

Parameters	Actual	Estimated (mean $\pm$ standard deviation)
$h_1(0)$	1.00	1.001 ± 0.016
$h_1(1)$	0.45	0.450 ± 0.003
$h_1(2)$	0.35	0.350 ± 0.009
$h_2(0)$	0.30	0.300 ± 0.003
$h_2(2)$	-0.60	-0.600 ± 0.002
$h_3(0)$	0.10	0.100 ± 0.002
$h_3(1)$	-0.01	$-0.010 \pm 4 \times 10^{-4}$

Table 1. Estimated parameters (SNR= 30 dB)

One can note that the estimated mean values of the channel parameters are very close to the actual values.

The results of the symbol estimation procedure are illustrated by Figure 1. The proposed methods are compared with the p-th order inverse method suggested by Schetzen. The proposed methods give a good SER (Symbol Error Rate) for low levels of noise which isn't the case with the p-th order inverse method. The two proposed algorithms provide similar performances in terms of SER, but from a computational complexity point of view the PI method is less time consuming than the TPS-root method.

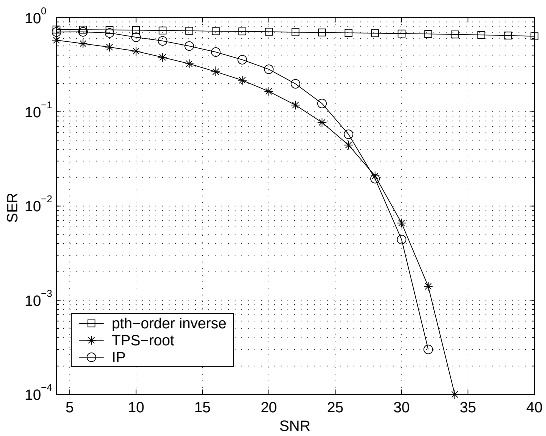

Fig. 1. SER versus SNR

6. CONCLUSION

In this paper, the problem of joint channel and symbols estimation has been solved for Hammerstein type channels with a semi-blind approach. By means of a specific input precoding that induces a redundancy in the received signal, the channel estimation and the symbols detection are carried out separately. The proposed methods are simple to implement and give good results with the cost of a transmission rate reduction.

7. REFERENCES

[1] W.I. Way. Optical fiber based microcellular systems. An overview. *IEICE Trans. Commun.*, E76-B(9):1091–1102, September 1993.

[2] X. N. Fernando and A.B. Sesay. Adaptive asymmetric linearization of radio over fiber links for wireless access. *IEEE Trans. on Vehicular Technology*, 51(6):1576–1586, November 2002.

[3] P. Koukoulas and N. Kalouptsidis. Blind identification of second order Hammerstein series. *Signal Processing*, 83:213–234, 2003.

[4] J.M. Le Caillec and R. Garello. Time series nonlinearity modeling: A Giannakis formula type approach. *Signal Processing*, 83:1759–1788, 2003.

[5] G. Picard and O. Cappé. Blind identification of Hammerstein nonlinear distorsion models. In *Proc. of the IEEE workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz,NY, USA, October, 19-22 2003.

[6] J. Choi. Sufficient conditions of the redundancy-transform for the blind identification. *Signal Processing*, 79:151–160, 1999.

[7] A.J. Redfern and G. Tong Zhou. A root method for Volterra system equalization. *IEEE Signal Processing Letters*, 5(5):285–288, 1998.

[8] M. Schetzen. *The Volterra and Wiener theories of nonlinear systems*. John Wiley and Sons, Inc., New-York, 1980.