ADAPTIVE ACOUSTIC ECHO CANCELLATION IN THE PRESENCE OF MULTIPLE NONLINEARITIES

Kun Shi, Xiaoli Ma, and G. Tong Zhou

School of Electrical and Computer Engineering, Georgia Tech, Atlanta, GA 30332-0250, USA

ABSTRACT

A Hammerstein-Wiener system consists of a linear time invariant subsystem sandwiched between two memoryless nonlinear blocks as is the case of an acoustic system with a nonlinear loudspeaker and a nonlinear microphone. We propose to model the memoryless nonlinear blocks of the Hammerstein-Wiener system using a linear combination of nonlinear basis functions, and concentrate on the task of parameter estimation for the nonlinear blocks. An adaptive algorithm is proposed using a pseudo magnitude squared coherence (PMSC) function-based criterion. The proposed method carries out nonlinearity identification without knowing the linear block in the Hammerstein-Wiener system. This is particularly useful for nonlinear acoustic echo cancellation (NAEC) applications, where dealing with the linear and nonlinear blocks together can be computationally challenging due to the long room impulse response. Numerical examples are provided to illustrate the performance of the proposed method.

***Index Terms*—** Hammerstein-Wiener system, pseudo magnitude squared coherence (PMSC) function, system identification, nonlinearity, acoustic echo cancellation

1. INTRODUCTION

A recent trend in mixed signal applications is to extend the use of analog components beyond their linear regions in order to gain advantages such as power efficiency, but to rely on the power of digital signal processing to “clean up the mess”. Nonlinear modeling and identification techniques are thus of interest. Volterra models are general but are too complex; simpler block based models have been considered including the Hammerstein system, the Wiener system, the Wiener-Hammerstein system, and the Hammerstein-Wiener system, etc [1–8]. In this paper, we are interested in the Hammerstein-Wiener system which consists of a linear time-invariant (LTI) system “sandwiched” between two memoryless nonlinear blocks [3]. Our interest in the Hammerstein-Wiener system stems from our desire to work with an acoustic environment where both the microphone and the power amplifier or the loudspeaker exhibit nonlinear behavior.

Numerous Hammerstein-Wiener system identification algorithms have been proposed in the literature. In [3], an identification scheme for single-input single-output (SISO) Hammerstein-Wiener systems was developed. A very specific model structure was assumed in [3] which limits its practical applicability. Building upon [3], a more general blind identification technique for SISO systems was proposed in [4]. An iterative method was developed in [5], and a linear subspace intersection algorithm was extended in [6] for the identification of Hammerstein-Wiener systems. Multiple-input multiple-output (MIMO) Hammerstein-Wiener systems were investigated in [7] with certain restrictions on the inputs. These existing methods identify the linear block together with the nonlinear blocks requiring heavy computational load when the linear block has a long memory.

In this paper, we develop an adaptive nonlinearity identification method which estimates the nonlinear blocks without any prior knowledge of the linear block. This is desirable in terms of computational complexity especially when the linear block has a long duration. Moreover, our method does not impose restrictive assumptions on the inputs, and does not impose restrictions on the nonlinear models other than that the output nonlinearity has to be invertible.

2. PROBLEM STATEMENT

Consider the Hammerstein-Wiener model shown in Fig. 1, which consists of an input memoryless nonlinearity $f(\cdot)$, a linear time-invariant system $h(n)$, and an output memoryless nonlinearity $g(\cdot)$. Mathematically, because the Hammerstein-Wiener model includes both the Hammerstein model and the Wiener model as special cases, it covers a more general class of nonlinear systems than either of the two models alone. The input $s(n)$ and the output $r(n)$ of the overall system are related by

$$r(n) = g(f(s(n)) * h(n)), \tag{1}$$

where $*$ denotes linear convolution.

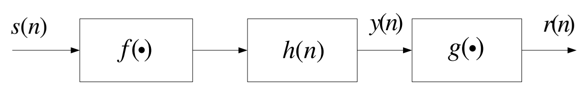

Fig. 1. Hammerstein-Wiener model.

Now the problem is on adaptively identifying the three blocks in Fig. 1 based on measurements of the input $s(n)$ and the output $r(n)$.

3. NONLINEARITY IDENTIFICATION USING THE PMSC FUNCTION

Let us approximate the memoryless nonlinearity $f(\cdot)$ by a linear combination of nonlinear basis functions $f_k(\cdot)$ with corresponding coefficients α_k; possible choices for $f_k(\cdot)$ include the polynomial and the spline bases. The approximated output of $f(\cdot)$ is given as

$$\tilde{f}(s;\boldsymbol{\alpha}) = \sum_{k=1}^{K_f} \alpha_k f_k(s), \;\; \boldsymbol{\alpha} = \left[\alpha_1, \alpha_2, ..., \alpha_{K_f}\right]^T, \tag{2}$$

where $\boldsymbol{\alpha}$ is unknown and needs to be estimated. We use the term approximated to account for modeling error in $f(\cdot)$.

We assume that the output nonlinear function $g(\cdot)$ is memoryless and is invertible. We strive to approximate the inverse of the

This work was supported in part by the Texas Instruments DSP Leadership University Program.

output nonlinearity also by a linear combination of nonlinear basis functions

$$\widetilde{g}^{-1}(r;\boldsymbol{\beta}) = \sum_{k=1}^{K_g} \beta_k g_k(r), \;\; \boldsymbol{\beta} = \left[\beta_1, \beta_2, ..., \beta_{K_g}\right]^T, \qquad (3)$$

where $g_k(r)$ is the k-th basis function whose input is r. Note that the nonlinearity $g(\cdot)$ is not identified directly in our proposed technique; instead, the inverse of $g(\cdot)$ is sought. The motivation is to decouple the estimation of the linear and the nonlinear blocks and also to identify the two nonlinear blocks in an alternating fashion as we shall see later.

We seek to estimate the unknown parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ from the input-output measurement data. Throughout the paper, $f_k(\cdot)$ ($k = 1, 2, ..., K_f$) and $g_k(\cdot)$ ($k = 1, 2, ..., K_g$) are assumed as known nonlinear basis functions and their orders K_f and K_g are assumed to be known as well.

Let $x(n)$ and $y(n)$ be real-valued discrete-time random processes. Define the pseudo magnitude squared coherence (PMSC) function between $x(n)$ and $y(n)$ at normalized frequency f as [9]

$$C_{xy}(f) = \frac{|S_{xy}(f)|^2}{S_{xx}(f)\sigma_y^2}, \; -0.5 \leq f \leq 0.5, \qquad (4)$$

where $S_{xy}(f)$ denotes the cross spectral density between $x(n)$ and $y(n)$ at frequency f; $S_{xx}(f)$ is the power spectral density (PSD) of $x(n)$ at frequency f, and $\sigma_y^2 = E[y^2(n)]$ is the power of $y(n)$. It can be shown that $0 \leq \int_{-0.5}^{0.5} C_{xy}(f)\mathrm{d}f \leq 1$; $\int_{-0.5}^{0.5} C_{xy}(f)\mathrm{d}f = 1$ if and only if $x(n)$ and $y(n)$ are linearly related, i.e., $y(n) = a(n) * x(n) + b(n)$, where $*$ denotes the linear convolution; $a(n)$ and $b(n)$ are deterministic quantities and $a(n) \neq 0$ [9].

Define vectors of basis functions

$$\boldsymbol{f}(n) = \left[f_1(s(n)), f_2(s(n)), ..., f_{K_f}(s(n))\right], \qquad (5)$$

$$\boldsymbol{g}(n) = \left[g_1(r(n)), g_2(r(n)), ..., g_{K_g}(r(n))\right], \qquad (6)$$

and output signals of the nonlinear modules

$$x(n;\boldsymbol{\alpha}) = \tilde{f}(s(n);\boldsymbol{\alpha}) = \boldsymbol{\alpha}^T \boldsymbol{f}(n), \qquad (7)$$

$$y(n;\boldsymbol{\beta}) = \widetilde{g}^{-1}(r(n);\boldsymbol{\beta}) = \boldsymbol{\beta}^T \boldsymbol{g}(n). \qquad (8)$$

If $\tilde{f}(\cdot;\boldsymbol{\alpha})$ is a perfect match to $f(\cdot)$ and $\widetilde{g}^{-1}(\cdot;\boldsymbol{\beta})$ is the inverse of $g(\cdot)$ up to a scalar, then the processes $x(n)$ and $y(n)$ will be perfectly linearly related. Since the metric $\int_{-0.5}^{0.5} C_{xy}(f)\mathrm{d}f$ provides a means for quantifying the linear association between two stationary random processes, we propose to solve for the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in the nonlinear blocks as follows:

$$\left[\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}\right] = \arg\max_{\boldsymbol{\alpha},\boldsymbol{\beta}} J(\boldsymbol{\alpha},\boldsymbol{\beta}), \qquad (9)$$

where

$$J(\boldsymbol{\alpha},\boldsymbol{\beta}) = \int_{-0.5}^{0.5} \hat{C}_{xy}(f;\boldsymbol{\alpha},\boldsymbol{\beta})\mathrm{d}f. \qquad (10)$$

Globally searching for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ will incur high computational complexity. However, given one of the unknown parameters, for instance $\boldsymbol{\beta}$, we can form the signal $y(n)$ according to (8). We also infer from (7) that

$$\sigma_x^2 = \boldsymbol{\alpha}^T E\left[\boldsymbol{f}(n)\boldsymbol{f}^T(n)\right]\boldsymbol{\alpha}, \qquad (11)$$

$$S_{yx}(f;\boldsymbol{\alpha}) = \boldsymbol{\alpha}^T \boldsymbol{s}_{yf}(f), \qquad (12)$$

where $\boldsymbol{s}_{yf}(f)$ is a vector whose k-th element is the cross spectral density between $y(n)$ and $f_k(s(n))$, $1 \leq k \leq K_f$. Substituting (11) and (12) into (10), the objective function given the parameter vector $\boldsymbol{\beta}$ can be reduced to

$$J(\boldsymbol{\alpha}|\boldsymbol{\beta}) = \frac{\boldsymbol{\alpha}^T \mathbf{R}_1 \boldsymbol{\alpha}}{\boldsymbol{\alpha}^T \mathbf{R}_2 \boldsymbol{\alpha}}, \qquad (13)$$

where

$$\mathbf{R}_1 = \int_{-0.5}^{0.5} S_{yy}^{-1}(f)\boldsymbol{s}_{yf}(f)\boldsymbol{s}_{yf}^H(f)\mathrm{d}f, \qquad (14)$$

$$\mathbf{R}_2 = E\left[\boldsymbol{f}(n)\boldsymbol{f}^T(n)\right], \qquad (15)$$

^{H} denotes the Hermitian transpose, and $y(n)$ is formed given the current $\boldsymbol{\beta}$. A similar form holds for $\boldsymbol{\beta}$ given $\boldsymbol{\alpha}$. Therefore, the objective function (10) is a generalized Rayleigh's quotient in $\boldsymbol{\alpha}$ for given $\boldsymbol{\beta}$ and *vice versa*. An alternating parameter estimation procedure is then the following relaxation algorithm [8]

$$\hat{\boldsymbol{\alpha}}(k) = \arg\max_{\boldsymbol{\alpha}} J(\boldsymbol{\alpha}, \hat{\boldsymbol{\beta}}(k-1)), \qquad (16)$$

$$\hat{\boldsymbol{\beta}}(k) = \arg\max_{\boldsymbol{\beta}} J(\hat{\boldsymbol{\alpha}}(k), \boldsymbol{\beta}). \qquad (17)$$

Similar to [9], we solve an eigenvalue decomposition problem

$$\mathbf{R}_1\hat{\boldsymbol{\alpha}} = \lambda_{\max}\mathbf{R}_2\hat{\boldsymbol{\alpha}}, \qquad (18)$$

where $\lambda_{\max}$ is the largest generalized eigenvalue for the pair ($\mathbf{R}_1$, $\mathbf{R}_2$). Similarly, given $\boldsymbol{\alpha}$, the parameters $\boldsymbol{\beta}$ can be solved using the same equation (18), except that $y(n)$ and $\boldsymbol{f}(n)$ are replaced by $x(n)$ and $\boldsymbol{g}(n)$, respectively. An adaptive algorithm was developed in [9] to update the parameter $\boldsymbol{\theta}$ (which can be $\boldsymbol{\alpha}$ or $\boldsymbol{\beta}$):

$$\boldsymbol{\theta}(n) = \frac{\boldsymbol{\theta}^T(n-1)\mathbf{R}_2(n)\boldsymbol{\theta}(n-1)}{\boldsymbol{\theta}^T(n-1)\mathbf{R}_1(n)\boldsymbol{\theta}(n-1)}\mathbf{R}_2^{-1}(n)\mathbf{R}_1(n)\boldsymbol{\theta}(n-1). \qquad (19)$$

Since the estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ depend on each other, we propose an iterative method for identifying the nonlinear parameters as summarized in Table 1, where L denotes the data segment length. Once the two nonlinear blocks have been identified, the linear block can be found via least squares. We point out that the convergence of the proposed iterative method is not guaranteed [8]. However, good initialization usually leads to convergence, which has been demonstrated by simulations. Note that the proposed method decouples the identification of the linear part from the nonlinear part, since the PMSC function is insensitive to the presence of an unknown linear block [9]. This feature is desirable for the NAEC problem, in which case the length of the room impulse response has no effect on the computational complexity. In the following section, we will propose a novel structure for the NAEC problem and apply the proposed Hammerstein-Wiener system identification method.

Table 1. Iterative method to estimate parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in the nonlinear blocks.

```
Initialize α(0) and β(0).
for k = 0, 1, ... do
  All n ∈ [kL, (k+1)L): update y(n) using (8) based on β(k).
  update α(k+1) using (19).
  All n ∈ [kL, (k+1)L): update x(n) using (7) based on α(k+1).
  update β(k+1) using (19).
end for
```

4. APPLICATIONS TO NONLINEAR ACOUSTIC ECHO CANCELLATION

Acoustic echo is a common phenomenon in telecommunication systems, such as in teleconferencing and hands-free telephone systems. Traditionally, adaptive filters have been widely employed to remove the acoustic echo based on the assumption of a completely linear loudspeaker enclosure microphone system (LEMS) (including the amplifier, the loudspeaker, the acoustic echo path and the microphone). A competitive audio consumer market can favor low-cost and small-sized analog components (such as the loudspeaker) which usually exhibit nonlinear characteristics. Research results have shown that linear acoustic echo cancellers fail when nonlinearity is present in the LEMS [10].

Several methods have been proposed for NAEC. By considering the memoryless nonlinearity in the loudspeaker, the LEMS can be well represented by the Hammerstein model [11, 12]. In this paper, we take into account the nonlinearity in both the loudspeaker and the microphone, in which case the LEMS can be described by the Hammerstein-Wiener model. However, to the best of our knowledge, none of the existing Hammerstein-Wiener system identification methods are suitable for the NAEC problem on hand, because: (1) they are nonadaptive and thus can not be readily applied to a real time echo canceller design; and (2) they incur large computational load due to the presence of a long room impulse response.

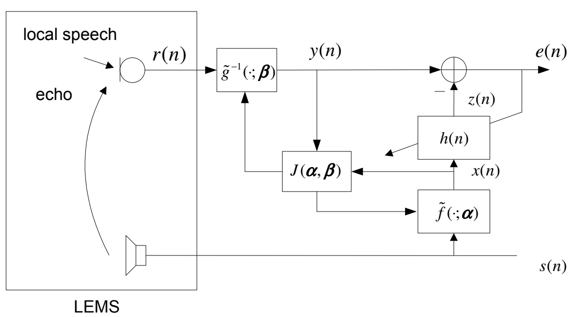

Fig. 2. Proposed structure for NAEC.

We propose a new structure for NAEC design as shown in Fig. 2. The adaptive NAEC consists of three blocks. The nonlinear block $\widetilde{g}^{-1}(\cdot;\boldsymbol{\beta})$ models the inverse of the microphone nonlinearity. Thus, the concatenation of the LEMS system with $\widetilde{g}^{-1}(\cdot;\boldsymbol{\beta})$ yields a Hammerstein system. For echo cancellation, we use a nonlinear block $\tilde{f}(\cdot;\boldsymbol{\alpha})$ and a finite impulse response (FIR) filter $h(n)$ to model the loudspeaker nonlinearity and room impulse response, respectively. As usual, the goal of NAEC is to minimize the power of the residual echo signal

$$e(n) = y(n) - z(n) = \widetilde{g}^{-1}(r(n);\boldsymbol{\beta}) - \tilde{f}(s(n);\boldsymbol{\alpha}) * h(n). \quad (20)$$

The algorithm for estimating the nonlinear parameters is summarized in Table 1. Afterwards, $h(n)$ can be estimated using the normalized least mean square (NLMS) algorithm as in existing approaches. We point out that the insertion of the $\widetilde{g}^{-1}(\cdot)$ block in Fig. 2 also enhances the quality of the local speech by canceling out the microphone nonlinearity.

5. SIMULATION RESULTS

In this section, we demonstrate the performance of the proposed PMSC function-based criterion for (approximately) identifying the unknown nonlinearities in the Hammerstein-Wiener system and its application to the NAEC problem. In the simulations for nonlinearity identification, the source signal $s(n)$ was generated according to an i.i.d. Gaussian distribution. Both nonlinear blocks obey the following input-output relationship:

$$\tanh(s) = (e^{2s} - 1)/(e^{2s} + 1), \quad (21)$$

where $\tanh$ denotes the hyperbolic tangent function. The linear block is an FIR filter of length 256 whose coefficients were generated randomly. We approximate the nonlinear functions $f(\cdot)$ and $g^{-1}(\cdot)$ by eqs. (2) and (3), respectively, both with polynomial bases and orders $K_f = K_g = 7$ in the simulations. We thus encounter certain modeling errors, since $f(\cdot)$ and $g^{-1}(\cdot)$ are not really polynomial functions. The simulations were carried out in a noise free environment. The total number of samples was $N = 32,768$. The block size used in the Welch method for power spectral density (PSD) estimation was $L = 256$, and the overlap between blocks was $P = 64$. $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are initialized such that $s(n) = x(n)$ and $y(n) = r(n)$, respectively.

Figs. 3 (a) and (b) show the performance of nonlinearity identification. Fig. 3 (a) shows the input nonlinearity $f(\cdot)$ and its estimate $\tilde{f}(\cdot)$; it can be seen that the estimate approximates well the nonlinearity $f(\cdot)$ in the system. Fig. 3 (b) shows the output nonlinearity $g(\cdot)$, the estimate of its inverse $\widetilde{g}^{-1}(\cdot)$, as well as the concatenated system consisting of $g(\cdot)$ followed by the nonlinear block $\widetilde{g}^{-1}(\cdot)$, i.e., $\widetilde{g}^{-1}(g(\cdot))$, which displays an approximate linear characteristic.

In Fig. 4, we show the estimate of the objective function in (10) as a function of the number of iterations. It can be seen that $J(\boldsymbol{\alpha},\boldsymbol{\beta})$ approaches one as the number of iterations increases. This implies that the two signals $x(n)$ and $y(n)$ are increasingly linearly related, indicating that $\tilde{f}(\cdot)$ and $\widetilde{g}^{-1}(\cdot)$ approach $f(\cdot)$ and $g^{-1}(\cdot)$ respectively, when a sufficient number of samples are available.

When applying to the NAEC problem, echo return loss enhancement (ERLE) [11] is used to measure the performance of the proposed nonlinear echo canceller

$$\text{ERLE (dB)} = 10\log_{10}\frac{E\left[y^2(n)\right]}{E\left[e^2(n)\right]}, \quad (22)$$

where $y(n)$ and $e(n)$ represent the microphone received echo signal and the residual echo signal, respectively. The microphone received signal $r(n)$ was generated under the single-talk scenario with the signal-to-noise ratio (SNR) set at 30 dB. Figs. 5 (a) and (b) show the ERLE for NAEC with respectively, noise and speech signal as the input; both demonstrate the effectiveness of the nonlinear echo cancellation algorithm.

6. CONCLUSIONS

In this paper, we proposed a pseudo magnitude squared coherence function-based criterion for the identification of the memoryless nonlinear blocks in a Hammerstein-Wiener system. The proposed method decouples the identification of the nonlinear parts from the linear part in the Hammerstein-Wiener system. This technique is particularly suitable for nonlinear acoustic echo cancellation applications, since the long room impulse response imposes heavy computational burden in existing NAEC approaches.

7. REFERENCES

[1] A. E. Nordsjö and L. H. Zetterberg, “Identification of certain time-varying nonlinear Wiener and Hammerstein Systems,” *IEEE Trans. on Signal Processing*, vol. 49, pp. 577–592, Mar. 2001.

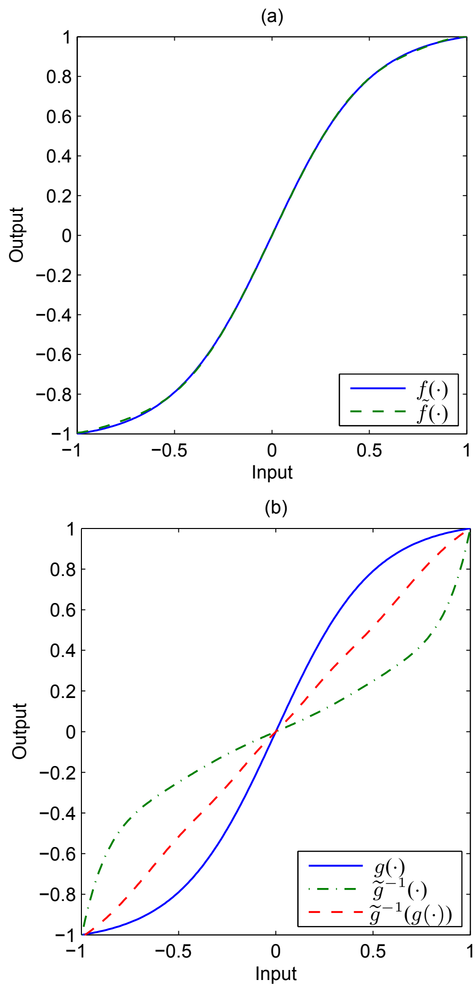

Fig. 3. Nonlinearity identification: (a) input nonlinearity $f(\cdot)$; (b) inverse of output nonlinearity $g^{-1}(\cdot)$.

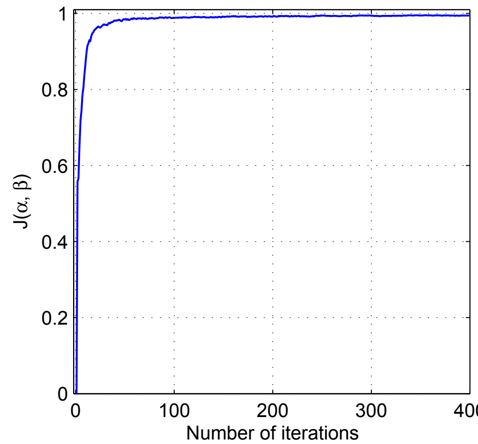

Fig. 4. The objective function J approaches one as the number of iterations increases.

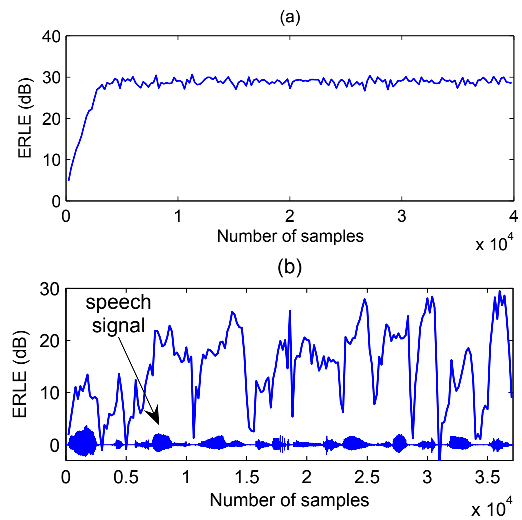

Fig. 5. Performance of nonlinear acoustic echo cancellation: (a) with i.i.d. Gaussian signal as input; (b) with speech signal as input.

[2] N. J. Bershad, S. Bouchired, and F. Castanie, "Stochastic analysis of adaptive gradient identification of Wiener-Hammerstein systems for Gaussian inputs," *IEEE Trans. on Signal Processing*, vol. 48, pp. 557–560, Feb. 2000.

[3] E. W. Bai, "An optimal two–stage identification algorithm for Hammerstein-Wiener nonlinear systems," *Automatica*, vol. 34, pp. 333–338, Mar. 1998.

[4] E. W. Bai, "A blind approach to the Hammerstein–Wiener model identification," *Automatica*, vol. 38, pp. 967-979, 2002.

[5] Y. Zhu, "Estimation of an N-L-N Hammerstein-Wiener model," *Automatica*, vol. 38, pp. 1607-1614, 2002.

[6] I. Goethals, K. Pelckmans, L. Hoegaerts, J. Suykens, and B. D. Moor, "Subspace intersection identification of Hammerstein–Wiener systems," *Proc. IEEE Conference on Decision and Control, and the European Control Conference*, pp. 7108–7113, Seville, Spain, Dec. 2005.

[7] P. Crama and J. Schoukens, "Hammerstein–Wiener system estimator initialization," *Proc. International Conference on Noise and Vibration Engineering*, pp. 1169-1176, Leuven, Belgium, Sept. 2002.

[8] P. Stoica, "On the convergence of an iterative algorithm used for Hammerstein system," *IEEE Trans. on Automatic Control*, vol. 26, pp. 967–969, Aug. 1981.

[9] K. Shi, X. Ma, and G. T. Zhou, "Adaptive nonlinearity identification in a Hammerstein system using a pseudo coherence function," *IEEE Workshop on Statistical Signal Processing*, pp. 745–748, Madison, WI, Aug. 2007.

[10] A. N. Birkett and R. A. Goubran, "Limitations of handsfree acoustic echo cancellers due to nonlinear loudspeaker distortion and enclosure vibration effects," *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 103-106, New Paltz, New York, Oct. 1995.

[11] F. Küch, A. Mitnacht, and W. Kellermann, "Nonlinear acoustic echo cancellation using adaptive orthogonalized power filters," *Proc. IEEE ICASSP*, vol. 3, pp. 105-108, Philadelphia, PA, Mar. 2005.

[12] G. -Y. Jiang and S. -F. Hsieh, "Nonlinear acoustic echo cancellation using orthogonal polynomial," *Proc. IEEE ICASSP*, pp. 273-276, Toulouse, France, May 2006.