Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Identification of Hammerstein and Wiener Models Using Spectral Magnitude Matching

Emad Abd-Elrady and Li Gan

Christian Doppler Laboratory for Nonlinear Signal Processing, Institute of Signal Processing and Speech Communication, Graz University of Technology. Inffeldgasse 16b, A-8010 Graz, Austria. (e-mail: emad.abd-elrady@TUGraz.at, li.gan@TUGraz.at).

Abstract: The problem of the identification of Hammerstein and Wiener models is considered in this paper. The suggested approach in this paper utilizes the spectral magnitude matching method that minimizes the sum squared error between the spectral magnitudes - evaluated for a number of short-time frames - of the measured output signal of the nonlinear system and the output signal of the nonlinear model. The coefficients of Hammerstein and Wiener models are estimated using the generalized Newton iterative algorithm. Simulation results show that the suggested approach gives very good results especially for moderate and high signal to noise ratios.

Keywords: Nonlinear models; Nonlinear systems; Parameter estimation; Recursive algorithms; System identification.

1. INTRODUCTION

Block-structured models are used to model nonlinear systems that can be represented by interconnections of linear dynamics and static nonlinear elements. There are four commonly used block-structured models in the literature. Namely: Hammerstein model (N-L), Wiener model (L-N), Hammerstein-Wiener cascade model (N-L-N) and Wiener-Hammerstein cascade model (L-N-L). For these nonlinear models, it is assumed that only the input and the output signals of the model are measurable. See Abd-Elrady (2005) for more details.

Hammerstein model consists of a static nonlinearity followed by a linear dynamic system, as shown in Fig. 1. It is considered as the easiest nonlinear model to use for identification purposes compared to other nonlinear model structures. This is due to the fact that it is possible to transfer the SISO Hammerstein model to MISO model which is linear in parameters. Hence, linear techniques can be used for estimation purposes.

One can find a quite substantial literature dealing with the identification of Hammerstein models. Generally speaking, existing identification methods for Hammerstein models can be divided into iterative methods Stoica (1981), Vörös (1997), Vörös (1999), over-parameterization methods Boutayeb et al. (1996), stochastic methods, Billings et al. (1978), Greblicki (1996), separable least squares methods Bai (2002), Westwick et al. (2001), blind identification methods Bai et al. (2002) and frequency domain methods Bai (2003).

On the other hand, Wiener model consists of a linear dynamic system followed by a static nonlinearity, as shown in Fig. 2. Wiener models arise in practice whenever a

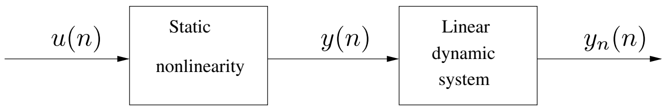

Fig. 1. Hammerstein model (N-L) structure.

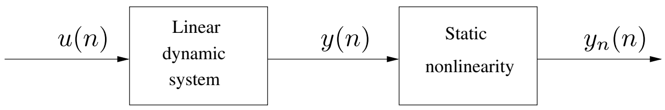

Fig. 2. Wiener model (L-N) structure.

measurement device has a nonlinear characteristic, see Billings et al. (1978), Greblicki (1992), Kalafatis et al. (2001), Koeppl et al. (2002), Nordsjö (1998), Wigren (1990).

On the contrary to Hammerstein model, it is not possible for Wiener model to identify the linear dynamics independently of the static nonlinearity. Independent parameterization of the two blocks requires maintaining the static gain to be constant in one of the blocks, see Abd-Elrady (2002), Wigren (1990). Also, it is important in what way disturbances enter to the system, *i.e.* before or after the static nonlinearity.

Using a parametric description for the linear dynamic block and the static nonlinearity, a prediction error criterion can be used to estimate the parameters of Wiener model as done in Wigren (1993), Wigren (1994). Also, nonparametric approaches have been used to identify the Wiener model in Greblicki (1992), Bai (2003).

In Quatieri et al. (2000), a method was introduced for estimating telephone handset nonlinearity by matching the spectral magnitude of the distorted signal to the output

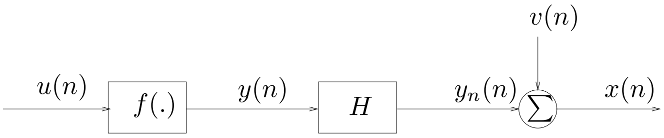

Fig. 3. Hammerstein system with measurement noise.

of a nonlinear channel model. The nonlinear model was chosen as a Wiener-Hammerstein cascade system (L-N-L) with a static nonlinearity described by a finite-order polynomial. The nonlinear model coefficients were estimated using the generalized Newton iterative algorithm Luenberger (1973), Widrow et al. (1985) that minimize a cost function of the sum squared error between the spectral magnitudes - evaluated for a number of short-time frames - of the measured distorted signal and the output signal of the nonlinear model. In this paper, the same approach of Quatieri et al. (2000) is used for the identification of Hammerstein and Wiener models.

The paper is organized as follows. Hammerstein and Wiener systems are discussed in Sec. 2 and Sec. 3, respectively. The spectral magnitude matching (SMM) method is presented in Sec. 4. In Sec. 5, some simulation examples are given. Conclusions are presented in Sec. 6.

2. HAMMERSTEIN SYSTEM

Assume that the nonlinear block of Hammerstein system of Fig. 3 is characterized as

$$\begin{aligned} y(n) &= f(u(n)) \\ &= c_1 u(n) + c_2 u^2(n) + \cdots + c_{n_c} u^{n_c}(n). \end{aligned} \tag{1}$$

Assume also that Hammerstein system has the following output

$$x(n) = y_n(n) + v(n) \tag{2}$$

where $v(n)$ is zero-mean additive white Gaussian noise (AWGN) and $y_n(n)$ is characterized by the following input-output relation

$$\begin{aligned} y_n(n) &= H(z^{-1})y(n) \\ &= \frac{B(z^{-1})}{A(z^{-1})} y(n). \end{aligned} \tag{3}$$

Here $A(z^{-1})$ and $B(z^{-1})$ are polynomials in the shift operator z^{-1} ($z^{-1}y(n) = y(n-1)$) with

$$\begin{aligned} A(z^{-1}) &= 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a} \\ B(z^{-1}) &= b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b}. \end{aligned} \tag{4}$$

In order to have a unique parameterization of the Hammerstein model structure, the first coefficient of the nonlinear function $f(.)$, *i.e.* c_1 is set equal to 1, see Ding et al. (2005), Bai (2002).

Let us define a parameter vector $\boldsymbol{\theta}$ as follows

$$\boldsymbol{\theta} = (\, a_1 \; a_2 \; \cdots \; a_{n_a} \; b_1 \; b_2 \; \cdots \; b_{n_b} \; c_2 \; \cdots \; c_{n_c} \,)^T. \tag{5}$$

The aim of this paper is the estimation of the parameter vector $\boldsymbol{\theta}$ by minimizing the difference between the spectral magnitudes of the true measured output data of the nonlinear system described by Eqs. (1)-(4) and the output of the Hammerstein model structure.

3. WIENER SYSTEM

As it was mentioned in Sec. 1, it is important in what way disturbances enter to Wiener system. Two different

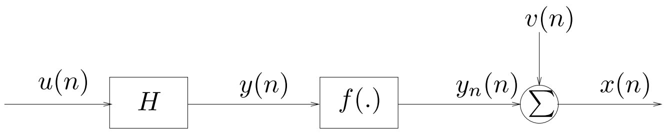

Fig. 4. Wiener system with measurement noise.

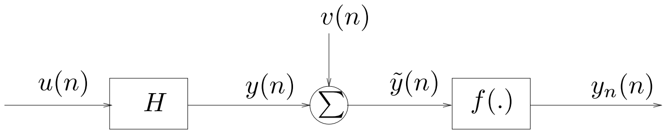

Fig. 5. Wiener system with disturbance noise.

cases are given in Figs. 4-5. In Fig. 4, $v(n)$ is considered as a measurement noise but in Fig. 5 it is considered as a disturbance and $v(n)$ will be nonlinearly mapped at the output of the system. Only the case of Fig. 4 is considered in this section. The two cases will be considered in the simulation study in Sec. 5.

Assume that the linear block of Wiener system of Fig. 4 is characterized as

$$\begin{aligned} y(n) &= H(z^{-1})u(n) \\ &= \frac{B(z^{-1})}{A(z^{-1})} u(n) \end{aligned} \tag{6}$$

where $A(z^{-1})$ and $B(z^{-1})$ are given by (4). Assume also that Wiener system has the following output

$$x(n) = y_n(n) + v(n). \tag{7}$$

Here the output of the nonlinear block $y_n(n)$ is given by

$$\begin{aligned} y_n(n) &= f(y(n)) \\ &= c_1 y(n) + c_2 y^2(n) + \cdots + c_{n_c} y^{n_c}(n). \end{aligned} \tag{8}$$

Similarly as in Sec. 2, the parameter vector $\boldsymbol{\theta}$ which defined in Eq. (5) is estimated by minimizing the difference between the spectral magnitudes of the true measured output data of the nonlinear system described by Eqs. (6)-(8) and the output of the Wiener model structure. This is the topic of the next section.

4. THE SPECTRAL MAGNITUDE MATCHING APPROACH

The suggested SMM method shown in Fig. 6 minimizes the error between spectral magnitudes of the measured signal $x(n)$ and the output signal $z(n)$ through the following cost function, see Quatieri et al. (2000):

$$V_{\boldsymbol{\theta}} = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} \left[|X(\omega_l; k)| - |Z(\omega_l; k; \boldsymbol{\theta})| \right]^2 \tag{9}$$

where $X(\omega_l; k)$ and $Z(\omega_l; k; \boldsymbol{\theta})$ are the short-time DFT of the nonlinear system measured output signal and the nonlinear model output signal, respectively. Here, K is the number of uniformly-spaced short-time frames and L is the DFT length.

The cost function in Eq. (9) can be written as

$$V_{\boldsymbol{\theta}} = \boldsymbol{\Gamma}_{\boldsymbol{\theta}}^T \boldsymbol{\Gamma}_{\boldsymbol{\theta}} \tag{10}$$

where

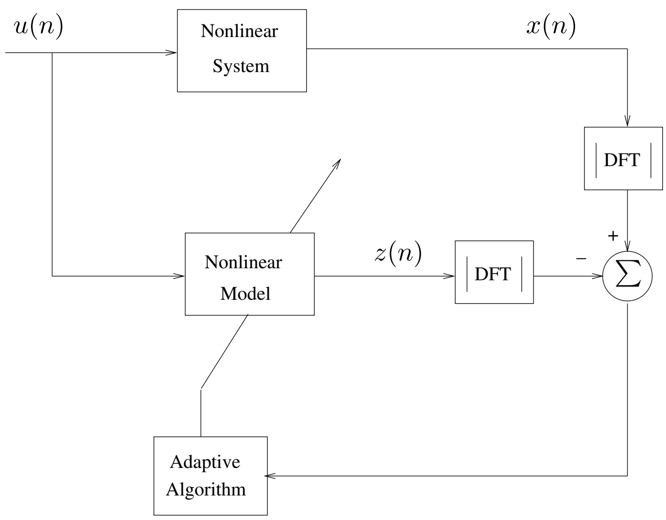

Fig. 6. Identification of Hammerstein and Wiener models using the SMM method.

$$\boldsymbol{\Gamma}_{\boldsymbol{\theta}} = \begin{pmatrix} \boldsymbol{\gamma}^0(\boldsymbol{\theta}) \\ \boldsymbol{\gamma}^1(\boldsymbol{\theta}) \\ \vdots \\ \boldsymbol{\gamma}^{K-1}(\boldsymbol{\theta}) \end{pmatrix} \tag{11}$$

and

$$\boldsymbol{\gamma}^k(\boldsymbol{\theta}) = \begin{pmatrix} |X(\omega_0;k)| - |Z(\omega_0;k;\boldsymbol{\theta})| \\ \vdots \\ |X(\omega_{L-1};k)| - |Z(\omega_{L-1};k;\boldsymbol{\theta})| \end{pmatrix}, k = 0,..,K-1. \tag{12}$$

The parameter vector $\boldsymbol{\theta}$ that minimizes the cost function $V_{\boldsymbol{\theta}}$ can be estimated similarly as in Quatieri et al. (2000) using the generalized Newton iteration algorithm, see Luenberger (1973), Söderström et al. (1989), Widrow et al. (1985). Hence, the estimate of the parameter vector follows as

$$\widehat{\boldsymbol{\theta}}(m+1) = \widehat{\boldsymbol{\theta}}(m) + \mu\ \boldsymbol{\Delta}(m) \tag{13}$$

where m is the iteration index, μ is the adaptation gain, and the gradient $\boldsymbol{\Delta}(m)$ is given by

$$\begin{aligned} \boldsymbol{\Delta}(m) &= -\left[\frac{d^2 V_{\boldsymbol{\theta}}}{d\boldsymbol{\theta}^2}\right]^{-1} \left[\frac{dV_{\boldsymbol{\theta}}}{d\boldsymbol{\theta}}\right] \\ &= -\Big(\boldsymbol{J}^T(m)\boldsymbol{J}(m)\Big)^{-1} \boldsymbol{J}^T(m)\boldsymbol{\Gamma}_{\boldsymbol{\theta}} \,|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}(m)} . \end{aligned} \tag{14}$$

Here $\boldsymbol{J}(m)$ is the Jacobian matrix of first derivative of $\boldsymbol{\Gamma}_{\boldsymbol{\theta}}$ with respect to $\boldsymbol{\theta}$ evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}(m)$, *i.e.*

$$\begin{aligned} \boldsymbol{J}(m) &= \frac{\mathrm{d}\boldsymbol{\Gamma}_{\boldsymbol{\theta}}}{\mathrm{d}\boldsymbol{\theta}} \,|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}(m)} \\ &= \begin{pmatrix} \boldsymbol{J}^0(m) \\ \boldsymbol{J}^1(m) \\ \vdots \\ \boldsymbol{J}^{K-1}(m) \end{pmatrix} \end{aligned} \tag{15}$$

where

$$\begin{aligned} \boldsymbol{J}^k(m) &= \frac{\mathrm{d}\boldsymbol{\gamma}^k(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}} \,|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}(m)} \\ &= -\begin{pmatrix} \dfrac{d|Z(\omega_0;k;\boldsymbol{\theta})|}{d\boldsymbol{\theta}} \\ \vdots \\ \dfrac{d|Z(\omega_{L-1};k;\boldsymbol{\theta})|}{d\boldsymbol{\theta}} \end{pmatrix}_{|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}(m)}} , k = 0, \cdots, K-1. \end{aligned} \tag{16}$$

Due to the fact that there is no close form expression for the gradient $\boldsymbol{\Delta}(m)$, an approximate gradient was evaluated in Quatieri et al. (2000) by finite element approximation. The same approach is considered here. The approximation follows the following lines:

1. Initiate with a parameter vector $\widehat{\boldsymbol{\theta}}(0)$ and compute the DFT magnitude $|X(\omega_l;k)|$.
2. Compute the DFT magnitude $|Z(\omega_l;k;\widehat{\boldsymbol{\theta}})|$ based on the current value of the parameter vector $\widehat{\boldsymbol{\theta}}(m)$ and form $\boldsymbol{\Gamma}_{\boldsymbol{\theta}}$.
3. Recalculate $z(n;\widehat{\boldsymbol{\theta}})$ for each perturbed component of $\widehat{\boldsymbol{\theta}}(m)$ and then compute its DFT magnitude. The (i,j) element of the matrix element $\boldsymbol{J}^k(m)$ denoted as $J^k_{i,j}(m)$ is evaluated using a first backward difference for each element of $\widehat{\boldsymbol{\theta}}(m)$ as
$$\begin{aligned} J^k_{i,j}(m) &= \frac{\partial \gamma^k_i(\widehat{\boldsymbol{\theta}};m)}{\partial \widehat{\theta}_j(m)} \\ &\approx -\frac{1}{\varepsilon_m}\Big(|Z(\omega_i;k;\widehat{\theta}_1(m),\cdots,\widehat{\theta}_j(m)+\varepsilon_m,\cdots)| \\ &\qquad -|Z(\omega_i;k;\widehat{\theta}_1(m),\cdots,\widehat{\theta}_j(m),\cdots)|\Big) \end{aligned} \tag{17}$$
where $\gamma^k_i(\widehat{\boldsymbol{\theta}};m)$ is the ith element of $\boldsymbol{\gamma}^k(\boldsymbol{\theta})$, $\widehat{\theta}_j(m)$ is the jth element of the parameter vector $\widehat{\boldsymbol{\theta}}(m)$ and ε_m is a small adaptive perturbation evaluated as
$$\varepsilon_m = \frac{V_{\boldsymbol{\theta}}(m)}{V_{\boldsymbol{\theta}}(0)}\varepsilon_0 \tag{18}$$
where ε_0 is the initial perturbation, $V_{\boldsymbol{\theta}}(0)$ is the initial value of $V_{\boldsymbol{\theta}}$, and $V_{\boldsymbol{\theta}}(m)$ is the value of $V_{\boldsymbol{\theta}}$ at iteration m. This means that the perturbation decreases proportionally with the error.
4. Finally, evaluate the correction term $\boldsymbol{\Delta}(m)$ from Eq. (14) and update the parameter vector $\widehat{\boldsymbol{\theta}}$ using Eq. (13).

In the SMM method, the spectral magnitudes of the measured output signal $x(n)$ and the output signal of the nonlinear model $z(n)$ are calculated using DFT for a number of short-time frames. In order to have accurate estimates, long data length is needed. Since the SMM method is off-line identification process, in each iteration the output signal corresponding to the input data length has to be evaluated and the DFT for the time frames is calculated. Also, in the gradient calculations, the previous steps are repeated for each perturbed parameter. Meanwhile, a matrix inversion is needed in Eq. (14). Therefore, the computation complexity of the SMM method is quite high.

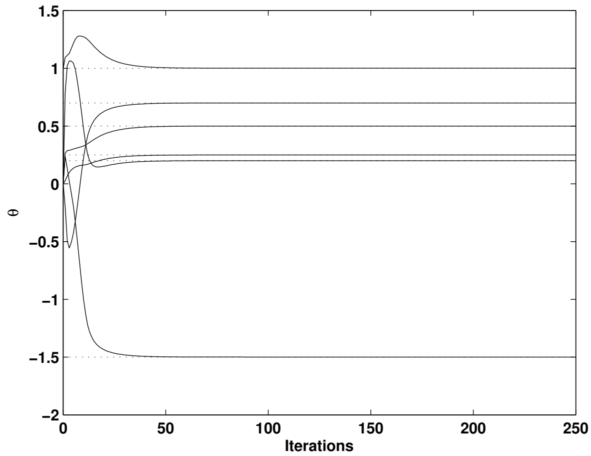

Fig. 7. Parameters convergence of the Hammerstein model structure. True values (dotted) and estimated (solid).

In order to have a feeling of the computation complexity of the SMM method, let us assume the following. The linear block is an IIR filter with $n_a = n_b = 4$, the static nonlinear block is described as a polynomial with $n_c = 4$, the length of the input data is 10^4 which is divided into $K = 10$ short frames and the DFT length is $L = 250$. Straightforward calculations show that 1.73×10^6 additions and 2.77×10^6 multiplications are needed in every iteration. This is without counting the DFT calculations and the matrix inversion. In case the size of the parameter vector and/or the number of the short-time frames increase, the computation complexity will be even higher. Future research will consider the possibility of reducing the computation complexity of the SMM method.

5. SIMULATION EXAMPLES

In order to investigate the performance of the suggested approach of Sec. 4, the following simulation examples were performed.

Example 1: Identification of Hammerstein model structure. In this simulation study the following Hammerstein system of Fig. 3 was considered:

$$\begin{aligned} &A(z^{-1})y_n(n) = B(z^{-1})y(n) \\ &A(z^{-1}) = 1 - 1.5z^{-1} + 0.7z^{-2} \\ &B(z^{-1}) = z^{-1} + 0.2z^{-2} \\ &y(n) = f(u(n)) = u(n) + 0.5u^2(n) + 0.25u^3(n) \\ &\boldsymbol{\theta} = (\, a_1 \;\, a_2 \;\, b_1 \;\, b_2 \;\, c_2 \;\, c_3 \,)^T . \end{aligned} \tag{19}$$

The data were generated from the system described by (19) using PRBS of length 8×10^3 samples. In the estimation process, 8 time frames were used each with 10^3 samples and 250 DFT frequency components, *i.e.* $K = 8$ and $L = 250$. Also the adaptation gain and the initial perturbation were chosen as $\mu = 0.1$ and $\varepsilon_0 = 0.5$, respectively. The initial parameter vector was $\boldsymbol{\theta}(0) = (\,0\;0\;1\;0\;0\;0\,)^T$. The AWGN $v(n)$ was chosen such that a signal to noise ratio (SNR) of 40 dB was achieved.

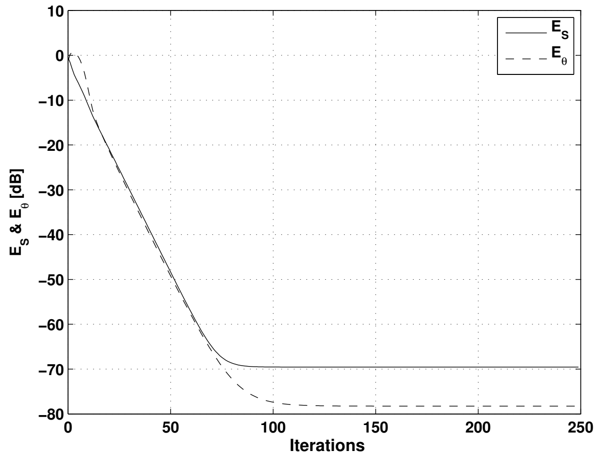

Fig. 8. E_S and $E_{\boldsymbol{\theta}}$ for the Hammerstein model structure.

As a performance measure, the normalized spectral error which is defined as

$$E_s(m) = \frac{\sum_{l=0}^{L-1} (|\bar{X}(\omega_l)| - |\bar{Z}(\omega_l; \widehat{\boldsymbol{\theta}}(m))|)^2}{\sum_{l=0}^{L-1} |\bar{X}(\omega_l)|^2} \tag{20}$$

has been used. Here $\bar{X}(\omega_l)$ and $\bar{Z}(\omega_l; \widehat{\boldsymbol{\theta}}(m))$ are the mean-values of the DFT of $x(n)$ and $z(n)$ over the time frames, and defined as

$$\bar{X}(\omega_l) = \frac{1}{K} \sum_{k=0}^{K-1} |X(\omega_l; k)|, \quad l = 0, \cdots, L-1. \tag{21}$$

$$\bar{Z}(\omega_l; \widehat{\boldsymbol{\theta}}(m)) = \frac{1}{K} \sum_{k=0}^{K-1} |Z(\omega_l; k; \widehat{\boldsymbol{\theta}}(m))|, \quad l = 0, \cdots, L-1. \tag{22}$$

Moreover, the normalized parameter error vector which is defined as

$$E_{\boldsymbol{\theta}}(m) = \frac{\|\widehat{\boldsymbol{\theta}}(m) - \boldsymbol{\theta}^o\|_2}{\|\boldsymbol{\theta}^o\|_2} \tag{23}$$

has been evaluated, where $\boldsymbol{\theta}^o$ is the true parameter vector. The simulation results are given in Figs. 7-8. The achieved values were $E_S = -69.55$ dB and $E_{\boldsymbol{\theta}} = -78.27$ dB.

Also, in order to study the performance of the suggested approach with SNR, the previous simulations were repeated for different SNRs. The mean values of $E_{\boldsymbol{\theta}}$ evaluated at the end of the 250 iterations over 100 different realization experiments are given in Fig. 9. The results of Fig. 9 show that the suggested approach gives very good results for different SNRs.

Example 2: Identification of Wiener model structure. In this simulation study the following Wiener system of Fig. 4 was considered:

$$\begin{aligned} &A(z^{-1})y(n) = B(z^{-1})u(n) \\ &A(z^{-1}) = 1 + 0.2z^{-1} - 0.35z^{-2} \\ &B(z^{-1}) = z^{-1} + 0.5z^{-2} \\ &y_n(n) = f(y(n)) = y(n) + 0.5y^2(n) + 0.25y^3(n) \\ &\boldsymbol{\theta} = (\, a_1 \;\, a_2 \;\, b_1 \;\, b_2 \;\, c_2 \;\, c_3 \,)^T . \end{aligned} \tag{24}$$

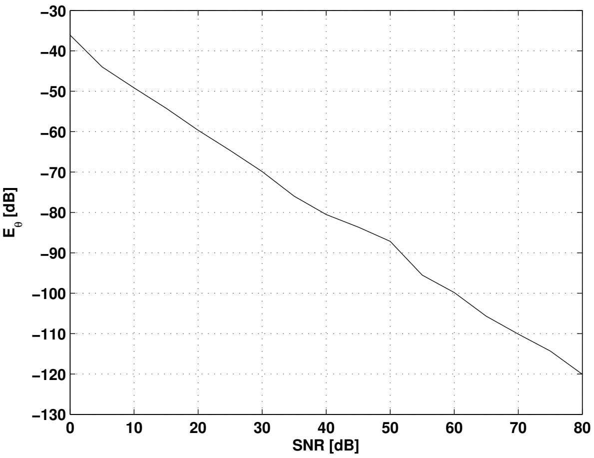

Fig. 9. $E_{\boldsymbol{\theta}}$ vs. SNR for the Hammerstein model structure.

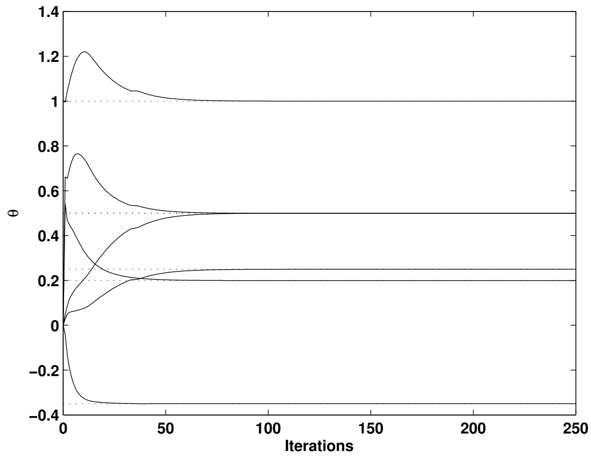

Fig. 10. Parameters convergence of the Wiener model structure. True values (dotted) and estimated (solid).

The data were generated and the suggested approach was initialized as done in Example 1. The simulation results at SNR = 40 dB are given in Figs. 10-11. The achieved values were $E_S = -61.5$ dB and $E_{\boldsymbol{\theta}} = -63.1$ dB.

Also, similarly to Example 1, a simulation study with SNR was performed for the Wiener system with measurement noise (*cf.* Fig. 4) and with disturbance noise (*cf.* Fig. 5). The results are given in Fig. 12. The results of Fig. 12 show that the suggested approach gives good results in both cases especially for moderate and high SNRs. Also, the results in the measurement noise case are more accurate as expected.

6. CONCLUSION

In this paper a suggested approach for the identification of Hammerstein and Wiener models has been presented. The approach utilizes the spectral magnitude matching method and the generalized Newton iterative algorithm to estimate the parameter vector. This estimation is done by minimizing the difference between the spectral magnitudes of the measured data and the output of the nonlinear model over a number of uniform short-time DFT frames.

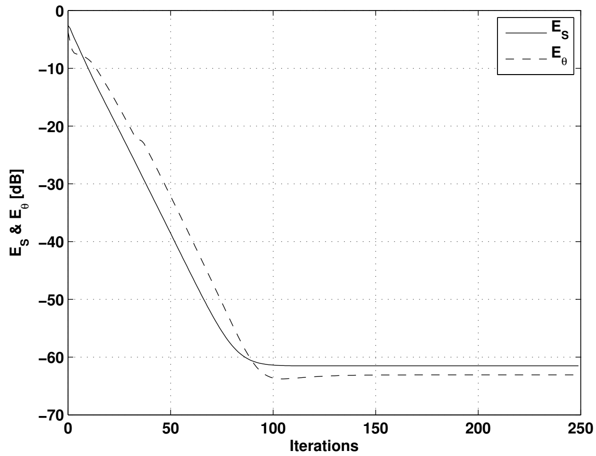

Fig. 11. E_S and $E_{\boldsymbol{\theta}}$ for the Wiener model structure.

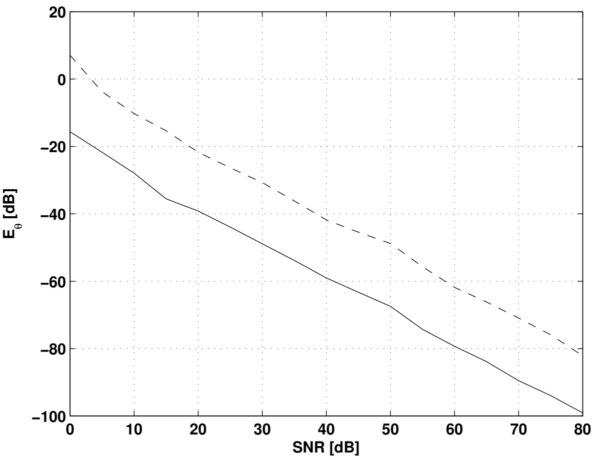

Fig. 12. $E_{\boldsymbol{\theta}}$ vs. SNR. Wiener model with measurement noise (solid) and with disturbance noise (dashed).

The suggested method is simple and straightforward to be applied for identification purposes. On the other hand, the computation complexity is high due to the fact that the gradient is evaluated approximately using DFT and finite element approximation at each iteration. Simulation results show that the suggested approach gives very accurate estimates.

ACKNOWLEDGEMENTS

This work was supported by Infineon Technologies Austria AG. The authors would also like to thank Professor Gernot Kubin, Dr. David Schwingshackl, and Dr. Gerhard Paoli for valuable discussions during this work.

REFERENCES

E. Abd-Elrady. A nonlinear approach to harmonic signal modeling. *Signal Processing*, 84:163–195, 2002.

E. Abd-Elrady. *Nonlinear approaches to periodic signal modeling.* Ph.d. thesis, Department of Information Technology, Uppsala University, Uppsala, Sweden, 2005.

E. Bai. A blind approach to the Hammerstein-Wiener model identification. *Automatica*, 38:967–979, 2002.

E. Bai. Identification of linear systems with hard input nonlinearities of known structure. *Automatica*, 38:853–860, 2002.

E. Bai. Frequency domain identification of Hammerstein models. *IEEE Transaction on Automatic Control*, 48: 530–542, 2003.

E. Bai. Frequency domain identification of Wiener models. *Automatica*, 39:1521–1530, 2003.

E. Bai and M. Fu. A blind approach to Hammerstein model identification. *IEEE Transactions on Signal Processing*, 50:1610–1619, 2002.

S. A. Billings and S. Y. Fakhouri. Identification of a class of nonlinear systems using correlation analysis. *IEE Proceedings*, 125:691–697, 1978.

S. A. Billings and S. Y. Fakhouri. Identification of nonlinear systems using the Wiener model. *Electronics Letters*, 13:502–504, 1978.

M. Boutayeb, H. Rafaralahy and M. Darouach. A robust and recursive identification method for Hammerstein model. *Proc. of IFAC World Congress, San Francisco, CA, USA*, 447–452, 1996.

F. Ding and T. Chen. Identification of Hammerstein nonlinear ARMAX system. *Automatica*, 41:1479–1489, 2005.

W. Greblicki. Nonparametric identification of Wiener systems. *IEEE Transaction on Information Theory*, 38: 1487–1493, 1992.

W. Greblicki. Nonlinearity estimation in Hammerstein systems based on ordered observations. *IEEE Transaction on Signal Processing*, 44:1224–1233, 1996.

A. Kalafatis, N. Arifin, L. Wang and W. R. Cluett. A new approach to the identification of pH processes based on the Wiener model. *Chemical Engineering Science*, 50: 3693–3701, 1995.

H. Koeppl and G. Paoli. Non-linear modeling of a broadband SLIC for ADSL-Lite-over-POTS using harmonic analysis. *Proc. of IEEE International Symposium on Circuits and Systems, Arizona, USA*, 133–136, 2002.

D. G. Luenberger. *Introduction to Linear and Nonlinear Programming.* Addison-Wesley, Reading, Massachusetts, 1973.

A. Nordsjö. *Studies on Identification of Time-Invariant and Time-Varying Nonlinear Systems.* Ph.d. thesis, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, 1998.

T. F. Quatieri and D. A. Reynolds and G. C. O'Leary. Estimation of handset nonlinearity with application to speaker recognition. *IEEE Trans. on Speech and Audio Processing*, 8:567–584, 2000.

T. Söderström and P. Stoica. *System Identification.* Prentice-Hall International, Hemel Hempstead, United Kingdom, 1989.

P. Stoica. On the convergence of an iterative algorithm used for Hammerstein system identification. *IEEE Transaction on Automatic Control*, 26:967–969, 1981.

J. Vörös. Parameter identification of discontinuous Hammerstein systems. *Automatica*, 33:1141–1146, 1997.

J. Vörös. Iterative algorithm for parameter identification of Hammerstein systems with two-segment nonlinearities. *IEEE Transaction on Automatic Control*, 44:2145–2149, 1999.

D. Westwick and R. Kearney. Seperable least squares identification of nonlinear Hammerstein models: application to stretch reflex dynamics. *Annals of Biomedical Engineering*, 29:707–718, 2001.

B. Widrow and S. D. Stearns. *Adaptive Signal Processing.* Prentice-Hall, Upper Saddle River, NJ, USA, 1985.

T. Wigren. *Recursive Identification Based on the Nonlinear Wiener Model.* Ph.d. thesis, Department of Technology, Uppsala University, Uppsala, Sweden, 1990.

T. Wigren. Convergence analysis of recursive identification algorithms based on the nonlinear Wiener model. *IEEE Transaction on Automatic Control*, 39:2191–2206, 1994.

T. Wigren. Recursive prediction error identification using the nonlinear Wiener model. *Automatica*, 29:1011–1025, 1993.