ROBUSTNESS ANALYSIS OF A GRADIENT IDENTIFICATION METHOD FOR A NONLINEAR WIENER SYSTEM

Ernst Aschbacher, Markus Rupp

Institute of Communications and Radio-Frequency Engineering
Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria
email: {ernst.aschbacher, markus.rupp}@tuwien.ac.at
web: www.nt.tuwien.ac.at/rapid_prototyping

ABSTRACT

The gradient identification of the linear filter part of a nonlinear Wiener system with unknown output non-linearity is investigated with respect to robustness in a deterministic sense. In order to estimate accurately the linear filter coefficients, an adaptive nonlinear filter is placed at the output of the Wiener system which compensates the output nonlinearity. Therefore, two adaptive algorithms work simultaneously. Local and global passivity relations are derived from which information on the robustness of the algorithm can be extracted.

1. INTRODUCTION

Nonlinear systems occur frequently in applications, e.g., in wireless and satellite communications the power amplifier is driven near saturation for efficiency reasons. A simple and low-complex model for such nonlinear power amplifiers with memory is a Wiener system [1]. A Wiener system is a series connection of a linear filter and a static (memoryless) non-linearity in that order. Such a model can be used to model the AM/AM conversion [2] of power amplifiers, describing the (nonlinear) dependence of the amplitude of the output signal from the amplitude of the input signal including memory effects.

The problem is that most often in real applications neither the linear part nor the nonlinear part is known and must be identified, e.g., with a low-complex gradient algorithm. A statistical analysis of the adaptation algorithm is hardly possible, often only information about convergence-in-the mean can be extracted [3], assuming a specific output nonlinearity. A further problem is that the Wiener system is nonlinear-in-parameters, thus, without a good initial estimate of the parameters the gradient algorithm locates possibly only a local minimum, resulting in a very inaccurate parameter estimation.

In this paper a method for the identification of the linear filter parameters with a gradient procedure is presented which avoids the problem of local minima. The convergence properties of the algorithm are analysed in a deterministic sense [4].

Research reported here was performed in the context of the network TARGET (Top Amplifier Research Groups in a European Team) and supported by the Information Society Technologies Programme of the EU under contract IST-1-507893-NOE, www.target-org.net.
The work was also supported by the Christian Doppler Pilot Laboratory for Design Methodology of Signal Processing Algorithms, Gusshausstr. 25/389, 1040 Vienna, Austria.

2. IDENTIFICATION METHOD

A graphical representation of the Wiener system is shown in Fig. 1.

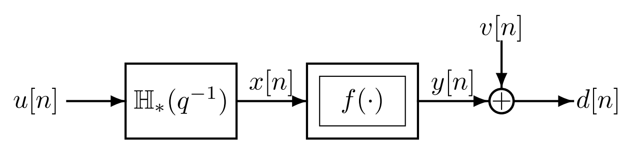

Fig. 1. A nonlinear Wiener system

Both, the linear filter and the static output non-linearity are assumed to be unknown. The signals $x[n]$ and $y[n]$ are not observable, which complicates the system identification. The noise $v[n]$ is the measurement noise. Since $f(\cdot)$ is a nonlinear function, the input/output relation is nonlinear in the parameters $h_{*,m}$ of the linear filter $\mathbb{H}_*(q^{-1}) = \sum_{m=0}^{M-1} h_{*,m} q^{-m}$,

$$d[n] = f\big(\mathbb{H}_*(u[n])\big) + v[n] \; . \tag{1}$$

The minimisation of a quadratic cost-function for the estimation of the parameters of the linear filter

$$\mathbf{h}^o = \arg\min_{\mathbf{h}} \left(\mathbb{E}(d[n] - \mathbf{h}\mathbf{u}_n)^2\right) \tag{2}$$

with $\mathbf{h} = [h_0, \ldots, h_{M-1}]$, $\mathbf{u}_n = \big[u[n], \ldots, u[n-M+1]\big]^T$ and $\mathbb{E}(\cdot)$ denoting the expectation operator, can lead to poor estimation results, depending on the output non-linearity, see also Section 3 further ahead.

Following an idea in [5] the adaptive scheme in Fig. 2 is proposed. Here, the non-linearity $g(\cdot)$ tries to compen-

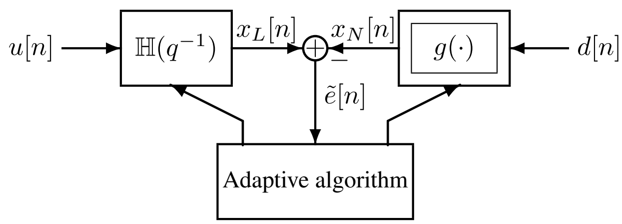

Fig. 2. Proposed adaptive identification scheme

sate the negative effects of the output non-linearity $f(\cdot)$ on the estimation of the linear filter parameters. The objective function which is minimised by the adaptive algorithm is

$$J[n] = \mathbb{E}\big(x_L[n] - x_N[n]\big)^2 \ . \tag{3}$$

It is assumed that $f(\cdot)$ is invertible. The inverse of the non-linear map $f^{-1}(\cdot)$ is approximated by $g(\cdot)$, which is represented using a series

$$f^{-1}(\cdot) \approx g(\cdot) = \sum_{k=1}^{K} w_k \psi_k(\cdot) \ , \tag{4}$$

$\{\psi_k\}_{k=1}^K$ being a set of basis-functions. The objective function is therefore

$$J[n] = \mathbb{E}\left(\mathbf{h}\mathbf{u}_n - \mathbf{w}\boldsymbol{\psi}_n\right)^2 \ , \tag{5}$$

with the signal-vector $\boldsymbol{\psi}_n = \big[\psi_1\big(d[n]\big), \ldots, \psi_K\big(d[n]\big)\big]^T$ and the parameter-vector $\mathbf{w} = [w_1, \ldots, w_K]$. The trivial solution $\mathbf{h} = \mathbf{w} = \mathbf{0}$ has to be excluded. Here, the coefficient h_0 is fixed to one, $h_0 = 1$. This is no restriction since Wiener systems are invariant to scaling [6], meaning that the linear filter can be assumed w.l.o.g. to be monic. Now, a reduced parameter vector has to be estimated, $\mathbf{h}^{'} = [h_1, \ldots, h_{M-1}]$, giving the error

$$\tilde{e}[n] = u[n] + \mathbf{h}^{'}\mathbf{u}_n^{'} - \mathbf{w}\boldsymbol{\psi}_n \ , \tag{6}$$

with $\mathbf{u}_n^{'} = \big[u[n-1], \ldots, u[n-M+1]\big]^T$.

2.1. Gradient updates

Derivation of the objective function with respect to the parameter vectors $\mathbf{h}^{'}$ and $\mathbf{w}$ and simplification of the expectation operator leads to the update equations

$$\mathbf{h}_n^{'} = \mathbf{h}_{n-1}^{'} - \mu_h[n]\tilde{e}_a[n]\mathbf{u}_n^{'T} \ , \ n \geq 0, \ \mathbf{h}_0^{'} \text{ given} \tag{7}$$

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \mu_w[n]\tilde{e}_a[n]\boldsymbol{\psi}_n^T \ , \ n \geq 0, \ \mathbf{w}_0 \text{ given} \ . \tag{8}$$

Here,

$$\tilde{e}_a[n] = u[n] + \mathbf{h}_{n-1}^{'}\mathbf{u}_n^{'} - \mathbf{w}_{n-1}\boldsymbol{\psi}_n \tag{9}$$

is the disturbed a-priori error. In the error-vector form the update equations read

$$\tilde{\mathbf{h}}_n^{'} = \tilde{\mathbf{h}}_{n-1}^{'} + \mu_h[n]\tilde{e}_a[n]\mathbf{u}_n^{'T} \tag{10}$$

$$\tilde{\mathbf{w}}_n = \tilde{\mathbf{w}}_{n-1} - \mu_w[n]\tilde{e}_a[n]\boldsymbol{\psi}_n^T \ , \tag{11}$$

with the parameter error-vectors $\tilde{\mathbf{h}}_n^{'} = \mathbf{h}_* - \mathbf{h}_n^{'}$, $\tilde{\mathbf{w}}_n = \mathbf{w}_* - \mathbf{w}_n$, where the $*$ denotes the optimal values, $\mathbf{h}_n^{'}$, $\mathbf{w}_n$ are the estimated parameter-vectors at time n. The disturbance $v[n]$ is indirectly included in the above definition of the a-priori error, namely via the signal $\boldsymbol{\psi}_n$.

The task is to analyse the stability of the algorithm in (10) and (11) and possibly device bounds for the step-sizes $\mu_h[n]$ and $\mu_w[n]$ which guarantee stable operation of the algorithm. Since a stochastic analysis without a-priori knowledge of the output non-linearity is not feasible, the analysis is carried out deterministically.

2.2. Local passivity relations

Decomposition of the disturbed a-priori error into

$$\tilde{e}_a[n] = e_{a,w}[n] - e_{a,h}[n] + v_e[n] \tag{12}$$

with $e_{a,w}[n] = \tilde{\mathbf{w}}_{n-1}\boldsymbol{\psi}_n$, $e_{a,h}[n] = \tilde{\mathbf{h}}_{n-1}^{'}\mathbf{u}_n^{'}$, and $v_e[n] = v_g[n] + v_v[n]$, whereby

$$v_g[n] = f^{-1}\big(d[n]\big) - \mathbf{w}_*\boldsymbol{\psi}_n \tag{13}$$

$$\begin{aligned} v_v[n] &= f^{-1}\big(y[n]\big) - f^{-1}\big(d[n]\big) \\ &= -\partial_y f^{-1}\big(y[n]\big)v[n] + \mathcal{O}\big(v[n]\big) \end{aligned} \tag{14}$$

are reflecting the errors due to the approximation of $f^{-1}(\cdot)$ using $g(\cdot)$, see (13), and the influence of the measurement noise, see (14), which primarily depends on the first derivative of the function $f^{-1}(\cdot)$.

The error $\tilde{e}_a[n]$ can be decomposed into two different expressions,

$$\tilde{e}_a[n] = e_{a,w}[n] + v_w[n] = -e_{a,h}[w] - v_h[n] \tag{15}$$

with the noise terms

$$v_w[n] = -e_{a,h}[n] + v_e[n] \tag{16}$$

$$v_h[n] = -e_{a,w}[n] - v_e[n] \ . \tag{17}$$

Assuming no noise $v[n]$ the noise $v_v[n]$, see (14), vanishes. The remaining disturbance is $v_g[n]$, due to the approximation of the inverse nonlinear function with a truncated series, cf. (4). At the other hand, if $v[n]$ does not vanish but $f^{-1}(\cdot) = g(\cdot)$, the noise $v_g[n] = 0$ but not $v_v[n]$, resulting in $v_e[n] = v_v[n]$. Further, additional disturbance, i.e., the undisturbed a-priori errors $e_{a,h}[n]$ and $e_{a,w}[n]$ appear and couple the two adaptive systems.

The update equations in the error-vector form read now

$$\tilde{\mathbf{h}}'_n = \tilde{\mathbf{h}}'_{n-1} - \mu_h[n](e_{a,h}[n] + v_h[n])\mathbf{u}_n'^T \quad (18)$$

$$\tilde{\mathbf{w}}_n = \tilde{\mathbf{w}}_{n-1} - \mu_w[n](e_{a,w}[n] + v_w[n])\boldsymbol{\psi}_n^T \ . \quad (19)$$

The equations are coupled via the noise terms $v_h[n], v_w[n]$ which depend on the undisturbed a-priori errors of the respectively other system, see (17) and (16).

The local passivity relations can now be stated as follows [4]: If

$$\mu_h[n] < \frac{1}{\|\mathbf{u}'_n\|_2^2} = \bar{\mu}_h[n] \quad (20)$$

$$\mu_w[n] < \frac{1}{\|\boldsymbol{\psi}_n\|_2^2} = \bar{\mu}_w[n] \quad (21)$$

then

$$\frac{\left\|\tilde{\mathbf{h}}'_n\right\|_2^2 + \mu_h[n]\big(e_{a,h}[n]\big)^2}{\left\|\tilde{\mathbf{h}}'_{n-1}\right\|_2^2 + \mu_h[n]\big(v_h[n]\big)^2} < 1 \quad (22)$$

$$\frac{\left\|\tilde{\mathbf{w}}_n\right\|_2^2 + \mu_w[n]\big(e_{a,w}[n]\big)^2}{\left\|\tilde{\mathbf{w}}_{n-1}\right\|_2^2 + \mu_w[n]\big(v_w[n]\big)^2} < 1 \ . \quad (23)$$

The two relations a coupled via the noise terms – the adaptation processes are not independent. The relations (22) and (23) with the conditions (20) and (21) fulfilled, define two contractive forward maps: as long as $\mu_h[n] < \bar{\mu}_h[n]$ and $\mu_w[n] < \bar{\mu}_h[n]$ local stability of the individual parts is guaranteed.

For one iteration step the error energy, the ℓ_2-norm of the parameter error-vector and the squared magnitude of the undisturbed a-priori error, are guaranteed to remain smaller than the disturbance energy, the ℓ_2-norm of the parameter error-vector at the previous iteration step with the squared magnitude of the noise terms $v_h[n]$ and $v_w[n]$. From (16) and (17) can be seen that these noise terms contain the undisturbed a-priori errors of the respectively other system. Thus, no matter how large these errors are, the error energies (nominators in the passivity relations) are smaller than the energy of the disturbance.

2.2.1. *Feed-back structure*

Straightforward manipulation leads to the update equations

$$\tilde{\mathbf{h}}'_n = \tilde{\mathbf{h}}'_{n-1} - \bar{\mu}_h[n]\big(e_{a,h}[n] + \bar{v}_h[n]\big)\mathbf{u}_n'^T \quad (24)$$

$$\tilde{\mathbf{w}}_n = \tilde{\mathbf{w}}_{n-1} - \bar{\mu}_w[n]\big(e_{a,w}[n] + \bar{v}_w[n]\big)\boldsymbol{\psi}_n^T \quad (25)$$

with (time-index n omitted)

$$\bar{v}_h = -\underbrace{\left(1 - \frac{\mu_h}{\bar{\mu}_h}\right) e_{a,h}}_{\text{feedback}} - \underbrace{\frac{\mu_h}{\bar{\mu}_h}\left(e_{a,w} + v_e\right)}_{\text{disturbance}} \quad (26)$$

$$\bar{v}_w = -\underbrace{\left(1 - \frac{\mu_w}{\bar{\mu}_w}\right) e_{a,w}}_{\text{feedback}} - \underbrace{\frac{\mu_w}{\bar{\mu}_w}\left(e_{a,h} - v_e\right)}_{\text{disturbance}} \ . \quad (27)$$

Since in (24) and (25) the step-sizes are equal to $\bar{\mu}_h[n]$ and $\bar{\mu}_w[n]$, the energy relations are now

$$\frac{\left\|\tilde{\mathbf{h}}'_n\right\|_2^2 + \bar{\mu}_h[n]\big(e_{a,h}[n]\big)^2}{\left\|\tilde{\mathbf{h}}'_{n-1}\right\|_2^2 + \bar{\mu}_h[n]\big(\bar{v}_h[n]\big)^2} = 1 \quad (28)$$

$$\frac{\left\|\tilde{\mathbf{w}}_n\right\|_2^2 + \bar{\mu}_w[n]\big(e_{a,w}[n]\big)^2}{\left\|\tilde{\mathbf{w}}_{n-1}\right\|_2^2 + \bar{\mu}_w[n]\big(\bar{v}_w[n]\big)^2} = 1 \ , \quad (29)$$

and define two lossless forward maps, $\bar{\mathbb{T}}_h$ and $\bar{\mathbb{T}}_w$. The feedback is defined via (26) and (27).

Using the small-gain theorem [7] the conditions for stability of the system (24) and (25) can be devised. Since the forward path is lossless, the stability depends entirely on the gain of the feedback paths. Therefore, if

$$\left|1 - \frac{\mu_h[n]}{\bar{\mu}_h[n]}\right| < 1 \quad (30)$$

$$\left|1 - \frac{\mu_w[n]}{\bar{\mu}_w[n]}\right| < 1 \ , \quad (31)$$

the system is locally stable. This yields the following limits for the step-sizes:

$$0 < \mu_h[n] < 2\bar{\mu}_h[n] \quad (32)$$

$$0 < \mu_w[n] < 2\bar{\mu}_w[n] \ . \quad (33)$$

The upper bounds are now twice as large as the bounds obtained in (20) and (21).

In Fig. 3 the feedback structure corresponding to the update equation for the linear-filter parameters is shown. The a-priori error $e_{a,w}[n]$ constitutes an additional disturbance, which adds to the disturbance $v_e[n]$. The condition for applying the small-gain theorem is that the energy of the complete disturbance-signal (time n omitted)

$$\bar{\bar{v}}_h = -\frac{\mu_h}{\bar{\mu}_h}\sqrt{\bar{\mu}_h} v_e - \frac{\mu_h}{\sqrt{\bar{\mu}_h \bar{\mu}_w}}\sqrt{\bar{\mu}_w} e_{a,w} \quad (34)$$

is bounded. This can be guaranteed if the energy of both components, $e_{a,w}[n]$ and $v_e[n]$ are also bounded. The energy of the a-priori error $e_{a,w}[n]$ depends on the stability of the second update-equation, which can be assured if the step-size is in the derived limit, cf. (33). The energy of the noise $v_e[n]$ depends on the approximation of the inverse $f^{-1}(\cdot)$, cf. (13), and on the noise $v[n]$, cf. (14), and is assumed to be bounded.

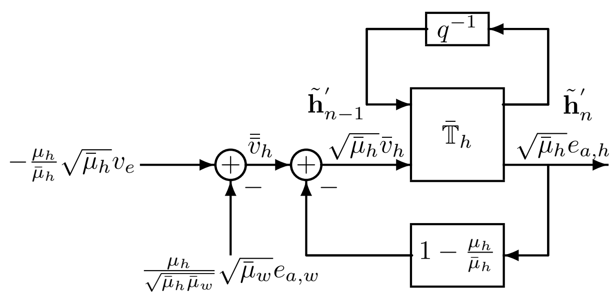

Fig. 3. Feedback structure – only the branch corresponding to the update-equation for the linear system is shown. The other branch, corresponding to the update-equation for the parameters of the nonlinear filter, is similar

2.3. Global passivity relations

Rewriting the local relations (28) and (29)

$$\left\|\tilde{\mathbf{h}}'_n\right\|_2^2 + \bar{\mu}_h[n]\big(e_{a,h}[n]\big)^2 = \left\|\tilde{\mathbf{h}}'_{n-1}\right\|_2^2 + \bar{\mu}_h[n]\big(\bar{v}_h[n]\big)^2 \tag{35}$$

$$\left\|\tilde{\mathbf{w}}_n\right\|_2^2 + \bar{\mu}_w[n]\big(e_{a,w}[n]\big)^2 = \left\|\tilde{\mathbf{w}}_{n-1}\right\|_2^2 + \bar{\mu}_w[n]\big(\bar{v}_w[n]\big)^2 \tag{36}$$

and summation of both sides over the the finite time-horizon $n = 1, \dots, N$ gives

$$\sum_{n=1}^{N} \bar{\mu}_h[n]\big(e_{a,h}[n]\big)^2 \le \left\|\tilde{\mathbf{h}}'_{-1}\right\|_2^2 + \sum_{n=1}^{N} \bar{\mu}_h[n]\big(\bar{v}_h[n]\big)^2 \tag{37}$$

$$\sum_{n=1}^{N} \bar{\mu}_w[n]\big(e_{a,w}[n]\big)^2 \le \left\|\tilde{\mathbf{w}}_{-1}\right\|_2^2 + \sum_{n=1}^{N} \bar{\mu}_w[n]\big(\bar{v}_w[n]\big)^2 , \tag{38}$$

which, after insertion of the noise terms (26) and (27) and simple manipulations gives (omitting the time-index n)

$$\sum_{n=1}^{N} \bar{\mu}_h {e_{a,h}}^2 \le \frac{1}{1-\gamma_h^2}\Big(\left\|\tilde{\mathbf{h}}'_{-1}\right\|_2^2 + \delta_h^2 \sum_{n=1}^{N} \bar{\mu}_h (e_{a,w}+v_e)^2\Big) \tag{39}$$

$$\sum_{n=1}^{N} \bar{\mu}_w {e_{a,w}}^2 \le \frac{1}{1-\gamma_w^2}\Big(\left\|\tilde{\mathbf{w}}_{-1}\right\|_2^2 + \delta_w^2 \sum_{n=1}^{N} \bar{\mu}_w (e_{a,h}+v_e)^2\Big) , \tag{40}$$

whereby

$$\gamma_h = \max_{n=1,\dots,N} \left|1 - \frac{\mu_h[n]}{\bar{\mu}_h[n]}\right| \tag{41}$$

$$\gamma_w = \max_{n=1,\dots,N} \left|1 - \frac{\mu_w[n]}{\bar{\mu}_w[n]}\right| \tag{42}$$

$$\delta_h = \max_{n=1,\dots,N} \frac{\mu_h[n]}{\bar{\mu}_h[n]} \tag{43}$$

$$\delta_w = \max_{n=1,\dots,N} \frac{\mu_w[n]}{\bar{\mu}_w[n]} . \tag{44}$$

Therefore, for global stability the following conditions must hold:

$$\gamma_h < 1 \Rightarrow 0 < \mu_h[n] < 2\bar{\mu}_h[n] , \, n = 1, \dots, N \tag{45}$$

$$\gamma_w < 1 \Rightarrow 0 < \mu_w[n] < 2\bar{\mu}_w[n] , \, n = 1, \dots, N , \tag{46}$$

which are the same conditions as for local stability, except that the conditions must hold for the whole time-horizon.

The error-energies on the left-hand side of (39) and (40) depend on the size of the feedback-gains γ_h and γ_w, as well as on the size of δ_h and δ_w. E.g., choosing $\mu_h[n] = \bar{\mu}_h[n]$ minimises the gain $1/(1-\gamma_h^2)$, since $\gamma_h = 0$, but $\delta_h = 1$, which can result in a relatively high error-energy, since the disturbance-energy is not attenuated.

3. SIMULATION RESULTS

For illustration purposes an example is analysed and simulations are carried out. A reference Wiener system with a linear FIR filter with 17 taps defining a bandpass filter and a static non-linearity given by $f(x) = \tanh(1.5x)$ is feeded with a white gaussian sequence with variance equal to one. For this example no measurement noise is added. The filter $g(\cdot)$ is a Taylor series with only uneven terms up to the seventh order, the basis-functions being $\{\psi(d)\}_{k=1}^{7} = \{d, d^3, d^5, d^7\}$. The algorithm is run 50 times and averages are taken to approximate the relative misadjustment

$$m_h[n] = \mathbb{E}\left(\frac{\|\tilde{\mathbf{h}}'_n\|_2^2}{\|\tilde{\mathbf{h}}'_0\|_2^2}\right) . \tag{47}$$

Four different cases are investigated, in each case the initial values for the filter parameter are $w_0 = [1, 0, 0, 0]$ and $h_0 = [1, 0, \dots, 0]$. The learning curves are shown in Fig. 4:

Case 1 Here, the step-sizes are $\mu_h[n] = \bar{\mu}_h[n], \mu_w[n] = 0$, thus only the linear part is identified. A relative misadjustment of minimally 12.5 dB is achieved. The system is stable. Since no noise at the output is added, the remaining misadjustment is entirely due to the influence of the nonlinear output filter in the Wiener system. By a reduction of the amplitude of the input signal, the error can be reduced, since the output non-linearity looses influence, i.e., the Wiener system behaves more linear.

Case 2 In this case, the step-sizes are $\mu_h[n] = \bar{\mu}_h[n]$ and $\mu_w[n] = 0.01$. Both step-sizes are smaller than the respective limits, as can be assured by observing $\bar{\mu}_w$. A significant reduction of the relative misadjustment of approx. 5 dB compared to case 1 can be achieved if the output of the Wiener system is nonlinearly filtered by $g(\cdot)$.

Case 3 If $\mu_h[n] = \mu_w[n] = \bar{\mu}[n]$ with $\bar{\mu}[n] = 1/\left(\|\mathbf{u}_n^{'}\|_2^2 + \|\boldsymbol{\psi}_n\|_2^2\right)$, both step-sizes are within their limits. The system is guaranteed to be stable. The remaining error is in this case relatively large.

Case 4 A further increase of the step-sizes to $\mu_w[n] = 0.55$ and $\mu_h[n] = 1.5\bar{\mu}_h[n]$ brings the system to instability in the time-horizon simulated. The bound for the step-size $\mu_w[n]$, namely $2\bar{\mu}_w[n]$, is exceeded approx. in only 0.2 % of the time (averaged over the realisations). From this it can be seen that the obtained bounds are relatively tight.

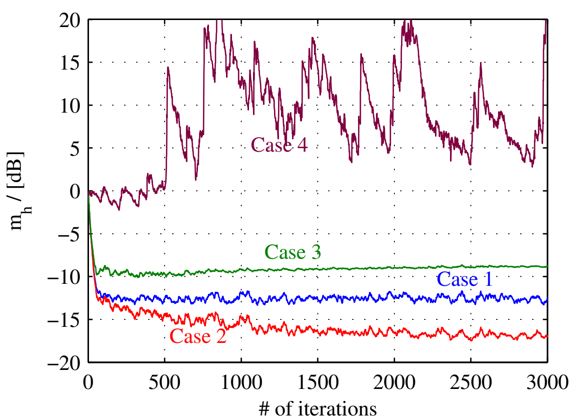

Fig. 4. Learning curves for the first simulation example. No noise is added at the output, $v[n] = 0$.

In the second example, see Fig. 5, white, zero-mean gaussian noise $v[n]$ with a standard-deviation of $\sigma_v = 0.05$ is added at the output. The resulting SNR, defined as

$$\text{SNR [dB]} = 10 \log \left(\frac{\|y[n]\|_2^2}{\|v[n]\|_2^2} \right) \qquad (48)$$

is approx. 21 dB. Case 1 and Case 2 are simulated again. Although significant noise is added only a slightly larger misadjustment compared to the first example of approximately 2 dB results, emphasising that the principal cause of the misadjustment is due to the influence of the nonlinear function at the output of the linear filter of the Wiener system.

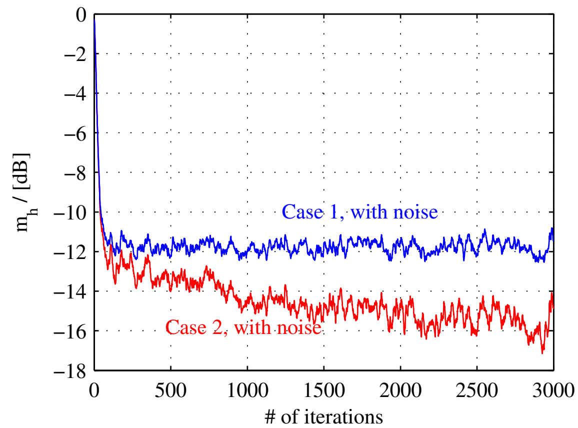

Fig. 5. Learning curves for the second simulation example. Noise is added at the output, resulting in an SNR $\approx$ 21 dB.

4. CONCLUSIONS

A method for the identification of the linear filter of a Wiener system has been analysed with respect to robustness in the ℓ_2-sense. The influence of the unknown output nonlinearity is reduced by an adaptive static nonlinear filter in order to improve the estimation quality for the linear-filter parameters. Therefore, two adaptive algorithms run simultaneously, which complicates the stability analysis. Bounds for the step-sizes of both update-equations for a robust behaviour could be derived and simulation results confirmed the theoretical predictions.

5. REFERENCES

[1] E. Aschbacher and M. Rupp, "Modelling and identification of a nonlinear power-amplifier with memory for nonlinear digital adaptive pre-distortion," in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC'03)*, pp. 555–559, June 2003.

[2] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. COM-29, no. 11, pp. 1715–1720, Nov. 1981.

[3] N. J. Bershad, S. Bouchired, and F. Castanie, "Stochastic analysis of adaptive gradient identification of Wiener-Hammerstein systems for gaussian inputs," *IEEE Trans. Signal Processing*, vol. 48, no. 2, pp. 557–560, Feb. 2000.

[4] A. H. Sayed and M. Rupp, "Error-energy bounds for adaptive gradient algorithms," *IEEE Trans. Signal Processing*, vol. 44, no. 8, pp. 1982–1989, Aug. 1996.

[5] A. D. Kalafatis, L. Wang, and W.R. Cluett, "Identification of Wiener-type nonlinear systems in a noisy environment," *International Journal of Control*, vol. 66, no. 6, pp. 923–941, April 1997.

[6] S. Boyd and L. O. Chua, "Uniquiness of a basic nonlinear structure," *IEEE Trans. Circuits Syst.*, vol. CAS-30, no. 9, pp. 648–651, Sep. 1983.

[7] H. K. Khalil, *Nonlinear Systems*, Macmillan Pub., 1992.