Adaptive Identification of Wiener Type Nonlinear Systems with Memory

Christian A. Schmidt
Universidad Nacional del Sur - CONICET

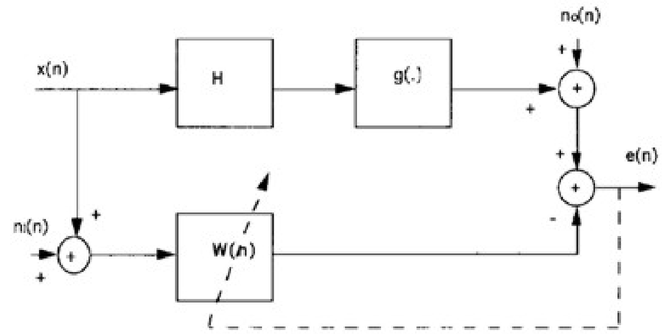

Fig. 1. First phase of the identification process.

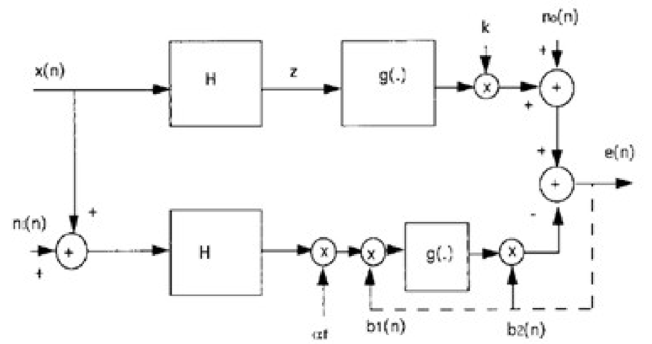

Fig. 2. Second phase of the identification process.

***Abstract*— In this paper, the behavior of an adaptive gradient search algorithm for the identification of a Wiener type nonlinear system is analyzed. Such system is formed by a linear discrete time system H followed by a static nonlinearity g(.). It is assumed that the shape of the nonlinearity is known a priori, in this case described by an erf function, which is commonly used to model saturation effects.**

The method analyzed is a sequential adaptive schema, which is divided in two parts. First, an LMS algorithm is used to estimate the linear part of the system, except for a scale error. Then, the parameters of the static nonlinearity are estimated using a gradient search algorithm.

I. INTRODUCTION

The field of nonlinear system identification has been studied for many years, and is still an active research area. Many nonlinear systems can be modeled by wiener type systems. In general, its parameters can be estimated by any of these methods:

* Deterministic approaches, such as orthogonal least squares expansion.

* Stochastic techniques, based on recursive adaptive algorithms.

* Nonadaptive block processing.

In this work, adaptive schemes are used to recursively estimate the Wiener model parameters. The nonlinearity is assumed to be of saturation type, but its parameters unknown.

Both input and output signals are corrupted by independent noise sources, and all inputs to the system are modeled by Gaussian variables. Figure 1 shows a bloc diagram of the system and the first stage of the identification process.

The input to the adaptive system is comprised of the sum of two zero-mean independent Gaussian white sequences $x(n)$ and $n_I(n)$ with variances σ_x^2 and σ_I^2, respectively. The output of the system is affected by a third zero-mean independent Gaussian white sequence $n_0(n)$,with variance σ_0^2. The input and output noises are measurement noises.

In the first phase of the identification process, an LMS adaptive algorithm is used to estimate the linear part of the system. Because of the inherent linear structure of the LMS algorithm, it will not be able to model the nonlinear part of the system, but it can correctly estimate its linear part within a scale factor. As it will be shown later, this scale factor depends on the systems nonlinearity and the number of training samples. After the first adaptation phase is completed, the estimated filter coefficients are frozen.

In the second part of the identification process, the scale factor from the first stage is estimated, along with the gain of the saturation type nonlinearity. In figure 2, a bloc diagram of this part of the algorithm is shown.

II. LINEAR ADAPTATION ANALYSIS: LMS ALGORITHM

The recursive expression for the calculation of the filter coefficient vector in an LMS algorithm is given by [4]:

$$W(n+1) = W(n) + \mu.e(n).Y(n) \tag{1}$$

Where $Y(n) = X(n) + N_I(n)$, $X^T(n) = [x(n), x(n-1), ..., x(n-N+1)]$, N is the number of coefficients in the adaptive filter, $N_I^T(n) = [n_I(n), n_I(n-1), ..., n_i(n-N+1)]$, and the instantaneous error equation is:

$$e(n) = g(H^T.X(n)) + n_0(n) - W^T(n).Y(n) \tag{2}$$

Insterting (2) in (1) yields:

$$W(n+1) = [I - \mu.Y(n).Y(n)^T].W(n) + \mu.[g(H^T.X(n)) + n_0(n)].Y(n) \quad (3)$$

If we assume μ sufficiently small, averaging both sides of (3), it can be shown that [1]:

$$\lim_{n\to\infty} E[W(n)] = \frac{\sigma_x^2}{\sigma_I^2}.E[g^{'}(z(n))].H \quad (4)$$

Where $z(n) = H^T.X(n)$; σ_x^2, σ_I^2 and $\sigma_y^2 = \sigma_x^2 + \sigma_I^2$ are the variances of $X(n)$, $N_I(n)$, and $Y(n)$ respectively.

The Wiener optimum filter for this problem satisfies the orthogonality condition $E[e(n).Y(n)] = 0$ [4], which can be expressed as:

$$E[(g(z(n)) + n_0(n) - W_0^T.Y(n)).Y(n)] = 0 \quad (5)$$

The solution to (5) is [1], [4]:

$$W_0 = \frac{\sigma_x^2}{\sigma_y^2} E[g^{'}(z(n))].H \quad (6)$$

Therefore, under the assumptions of n sufficiently large and μ small, the LMS algorithm converges on average to the optimum Wiener filter.

It can be shown [1] that for an arbitrary number of iterations n, $W(n) \cong \alpha(n).H$, where:

$$\alpha(n) = \frac{\sigma_x^2}{\sigma_y^2}.E[g^{'}(z(n))].(1 - [1 - \mu.\sigma_y^2]^n) \quad (7)$$

Which is the expression for the misadjustment scale factor from the first phase of the system identification process.

III. Nonlinear Adaptation Analysis: Gradient Search Algorithm

We now consider the scheme shown in figure 2 for the identification of the nonlinear part of the system $g(.)$. α_f is the scale error factor in the identification of the linear part of the system, once the adaptive filter coefficients have been frozen.That is, α_f is the value for $\alpha(n)$ after the last iteration of the LMS algorithm. We assumed that the shape of the nonlinear function $g(z)$ is known, we now have to estimate its gain k. The powers at the input and output of $g(z)$ are assumed to be equal, and are described by $E[z^2(n)] = \sigma_x^2.H^T.H$ and $k^2.E[g^2(z(n))] = \sigma_x^2.H^T.H$, respectively.

Solving for k yields:

$$k^2 = \frac{\sigma_x^2.H^T.H}{E[g^2(z(n))]} \quad (8)$$

Then, the proposed estimation problem consists in adapting coefficients b_1 and b_2 to the unknown system parameters α_f and k. This coefficient adaptation will be performed by a gradient search algorithm. In such algorithm, given the equations that describe the system, the instantaneous error is [1]:

$$e(n) = k.g(z(n)) + n_0(n) - b_2(n).g(b_1(n).W^T.Y(n)) \quad (9)$$

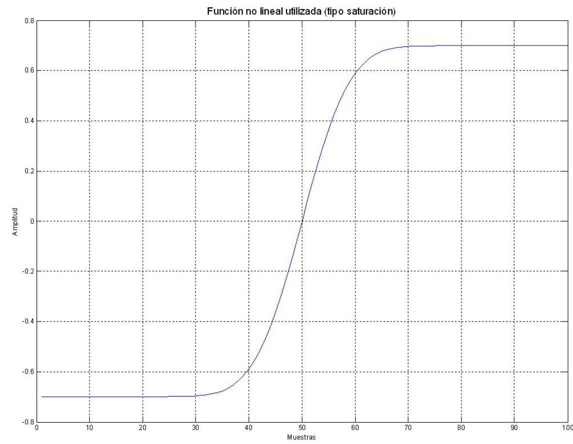

Fig. 3. Nonlinear function used in the simulation.

The optimum result would obviously be $b_1 = 1/\alpha_f$ and $b_2 = k$. However, this result can not be reached exactly because of the input measurement noise to the system $n_I(n)$ [1]. fact, it can be shown [1] that the error committed is small if and only if $\sigma_x^2 >> \sigma_I^2$.

The recursive stochastic gradient search algorithm for $b_1(n)$ and $b_2(n)$ is [1]:

$$b_1(n+1) = b_1(n) + \mu.e(n).b_2(n).W^T.Y(n).g^{'}(b_1(n).W^T.Y(n))$$

$$b_2(n+1) = b_2(n) + \mu.e(n).g(b_1(n).W^T.Y(n))$$

IV. Simulation Results

The algorithm was implemented in MATLAB. An analysis of Mean Square Error (MSE), as well as the quality of the estimation of the linear filter coefficients obtained in the first phase of the identification process is carried out. The parameters calculated in the second phase of the algorithm and the overall response of the obtained model are then compared with the optimum values.

The simulation of a linear system with an impulse response consisting of a shifted cosine pulse of unitary amplitude and a length of 40 samples was performed. The simulation includes a saturation type nonlinear function with gain equal to 0.7 at the output of this system, which is shown in Figure 3.

The dimensionality of the estimated filter was taken into account for the linear part of the system identification problem. It was observed that if the number of coefficients in the estimation filter is less than the length of the impulse response to be estimated, the obtained filter is not capable of modeling the system. On the other hand, if the number of coefficients in the filter is greater than length of the impulse response of the system, the filter can correctly estimate the system's response and the extra coefficients converge to zero. This is shown in Figure 4.

Therefore, it can be assumed without any loss of generality that the lengths of the estimation filter and the impulse response of the linear part of the system coincide.

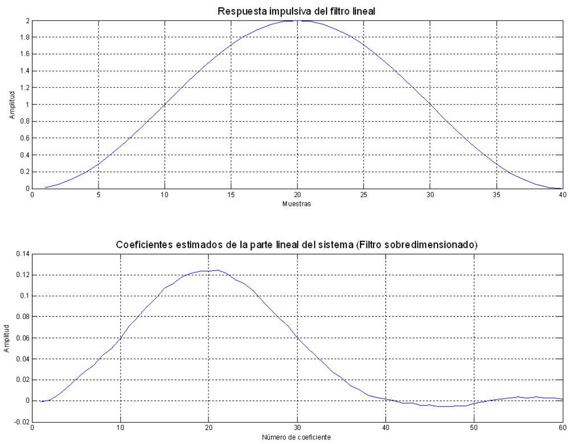

Fig. 4. Overdimension of the estimation filter.

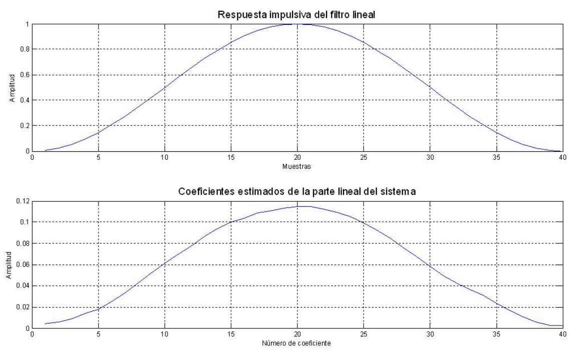

Fig. 5. H (top) and Estimated filter coefficients (lower).

First, the LMS algorithm for the identification of H was implemented, H being the impulse response of the linear part of the system. The parameters used in the simulation are the following:

* Number of samples of the input data sequence: 8x10^6
* $\sigma_x^2 = 1.2$; variance of the input data sequence.
* $\sigma_I^2 = 0.01$; variance of the input measurement noise.
* $\sigma_0^2 = 0.05$; variance of the output measurement noise.
* Length of the impulse response of H: 40.
* Number of coefficients in the estimation filter: 40.
* $k = 0.7$; gain of the nonlinear function.
* $\mu = 0.0002$, learning step for the LMS algorithm.
* $\mu_2 = 0.02$, learning step for the gradient search algorithm.

Fifty percent of the data samples were used in each estimation phase, and the initial values for the filter coefficients were chosen to be zero.

As expected, the obtained filter coefficients do estimate H correctly, except for a scale factor. This result can be seen in Figure 5.

Then, the second part of the estimation process was performed, taking one as the initial value for b_1 and b_2. The obtained values for the coefficients were the following: $b_1 = 7.2576$ and $b_2 = 0.7009$. In Figure 6, H and $b_1.W$ are shown together.

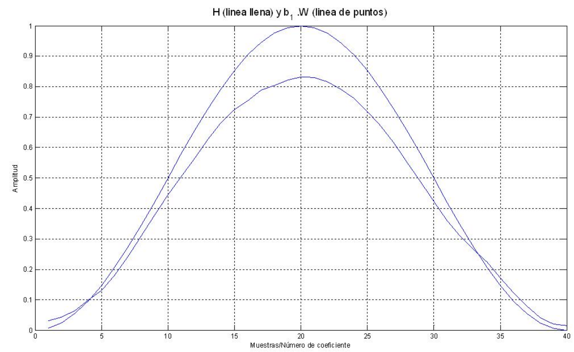

Fig. 6. H vs. $b_1.W$.

In the figure it can be seen that there is a misadjustment of about 15 percent in b_1, while the error in b_2 is much lower. This result was forseen in the theoretic analysis [1], and it is due to the input measurement noise.

Then, a performance analysis of the joint estimation was carried out by means of the MSE calculation. For this purpose, a new input data sequence of 100,000 samples of length was used, and the outputs of the actual system and the previously estimated model were compared for validation.

The expression for the MSE is [4]:

$$MSE = \frac{1}{N}.\sum_{n=1}^{N} e^2(n) \tag{10}$$

Where the instantaneous error $e(n) = y(n) - \hat{y}(n)$ [4] is given by [1]:

$$\begin{aligned} e(n) &= k.g(x(n) \otimes H + n_0(n)) \\ &- b_2.g([x(n) + n_I(n)] \otimes b_1.W) \end{aligned} \tag{11}$$

The MSE calculated for these conditions was MSE=0.0048.

Finally, the whole process was simulated again for more adverse noise conditions, taking $\sigma_I^2 = 0.05$ and $\sigma_0^2 = 0.1$. In this case, the MSE obtained was MSE=0.015.

V. Conclusions and future work

This work is based on the results presented in [1], particularly the first algorithm described in that paper. The theory used to implement the simulations was acquired during the course "Adaptive Signal Processing", its basic reference being [4].

In [2], the same algorithm is analyzed for combined Wiener-Hammerstein systems. This type of systems are composed of a linear system H_1 at the input, followed by a static nonlinearity $g(.)$, and a second linear system H_2 at the output. The difference in this case consists in that the LMS algorithm correctly estimates the convolution between systems H_1 and H_2, except for a scale factor. Again, this scale factor is estimated during the second phase of the identification process.

For future work, we propose an analysis for the case of a non Gaussian information source, taking into account arbitrary nonlinear functions. In that sense, a model for the nonlinearity

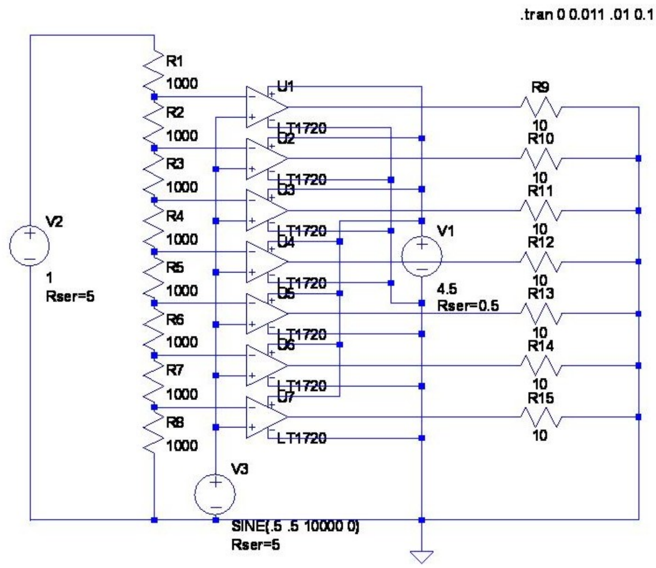

Fig. 7. Simplified flash ADC schematic.

based on hermitic polynomial expansions is proposed in [3]. This would allow for the use of this method in a more general type of systems, in particular for modeling the nonlinear behavior of AD converters.

In [5], a Volterra model and its approximation by a parallel Hammerstein system are studied for this application.

Another alternative would be the use of models based on neural networks to characterize the behavior of the nonlinear part of the system.

Currently, a 4 bit flash ADC architecture is being simulated in Spice and MATLAB. The theory for these simulations is based on [6]. The elements used for the Spice part of the simulation are discrete commercial components, such as resistors, comparators, and operational amplifiers. The obtained output signals are then exported to MATLAB in order to carry out spectral analysis and to test different compensation schemes. A simplified Spice schematic of such a converter is shown in figure 7.

This procedure is aimed to collect realistic data to study the nonlinear effects found in ADC converters and analyze possible compensation techniques.

A flash architecture was chosen for simulation because of its fast response and one clock cycle convertion time. In this type of converter, timing errors are one of the most important sources of distortion [6]. Therefore, it is of great interest to obtain precise data to model the behavior of the converter under fast timing conditions.

In the near future, it is forseen to increment the resolution of the simulated converter to 6 or 8 bits, using interpolation techniques as discussed in [6]. Also, a VHDL design of the final converter architecture form the previous simulation process would be implemented in a chip, to validate the results obtained by means of actual measurements.

VI. References

[1] "Stochastic Analysis of Gradient Adaptive Identification of Nonlinear Systems with Memory for Gaussian Data and Noisy Input and Output Measurements", N. J. Bershad, P. Celka, and J. M. Vesin, IEEE Transactions On Signal Processing, VOL. 47, NO. 3, MARCH 1999.

[2] "Stochastic Analysis of Adaptive Gradient Identification of Wiener-Hammerstein Systems for Gaussian Inputs", N. J. Bershad, S. Bouchired, and F. Castanie, IEEE Transactions On Signal Processing, VOL. 48, NO. 2, FEBRUARY 2000

[3] "Analysis of Stochastic Gradient Identification of Wiener-Hammerstein Systems for Nonlinearities with Hermite Polynomial Expansions", N. J. Bershad, P. Celka, and S. McLaughlin, IEEE Transactions On Signal Processing, VOL. 49, NO. 5, MAY 2001.

[4] "Adaptive Filter Theory", S. Haykin, Prentice Hall.

[5] "Modeling Analog to Digital Converters at Radio Frequency", N. Björsell, Doctoral Thesis in Telecommunications, Stockholm, Sweden 2007

[6] "CMOS Integrated Analog-to-Digital and Digital-to-Analog Converters", R. Van de Plassche, Kluwer Academic Publishers, 2003.