ADAPTIVE NEURAL NETWORKS FOR ROBUST ESTIMATION OF PARAMETERS OF NOISY HARMONIC SIGNALS

A. Cichocki
FRP Riken - ABS Laboratory, Institute of Physical and Chemical Research, Japan
Tel: +81 48 465 2645; fax: +81 48 462 4633
e-mail: cia@kamo.riken.go.jp

P. Kostyla, T. Lobos, Z. Waclawek
Technical University of Wroclaw
pl. Grunwaldzki 13, 50-370 Wroclaw, Poland
Tel: +48 71 203448; fax: +48 71 229725
e-mail: lobos@elektryk.ie.pwr.wroc.pl

ABSTRACT

In many applications, very fast methods are required for estimating and measurement of parameters of harmonic signals distorted by noise. This follows from the fact that signals have often time varying amplitudes. Most of the known digital algorithms are not fully parallel, so that the speed of processing is quite limited. In this paper we propose new parallel algorithms, which can be implemented by analogue adaptive circuits employing some neural network principles. The problem of estimation is formulated as an optimization problem and solved by using the gradient descent method. Algorithms based on the least-squares (LS), the total least-squares (TLS) and the robust TLS criteria are developed and compared. The networks process samples of observed noisy signals and give as a solution the desired parameters of signal components. Extensive computer simulations confirm the validity and performance of the proposed algorithm.

1 INTRODUCTION

Estimation of parameters (amplitudes) of harmonic signals is important in electric power systems and power electronics due to increasing use of nonlinear dynamic loads. The harmonics produced have usually varying amplitudes due to dynamic nature of nonlinear loads. Fast estimation of the parameters is essential for the control and protection of electric power systems. It is also useful in modelling, measurements and compensation of higher harmonics. Various digital and analogue (neural networks) algorithms have been proposed for the estimation of parameters of harmonic signals. They range from simple least-squares (LS) methods [1, 3], least absolute value (LAV) technique [2, 3], Lp-norm criteria [4], minimax (Chebyshev norm) technique [2, 3, 4, 5], methods based on the singular value decomposition (SVD) [6], DFT [7, 8], to the Kalman filtering approach [7].

The purpose of this paper is to present novel on-line techniques for estimation of parameters of harmonics based on the least-squares (LS), total least-squares (TLS) and the robust TLS criteria. The corresponding architectures of analogue neuron-like adaptive processors are also shown. The developed methods are more robust against random noise in comparison with other known algorithms.

2 MATHEMATICAL FORMULATION OF THE PROBLEM

We analyse and solve the following standard problem.

Let y(t) denote measured noisy signal

$$y(t) = \sum_{i=1}^{n} a_i \sin(i\omega t) + \sum_{i=1}^{n} b_i \cos(i\omega t) + r(t) \tag{1}$$

where

$\omega = 2\pi f$ is a known or rather approximately estimated angular frequency,
$r(t)$ is unknown noise or error (residual),
a_i, b_i are unknown amplitudes of harmonic signals.

On basis of values y(t) it is necessary to find or estimate in real time the amplitudes a_i, b_i. Assuming that a continuous-time signal y(t) is sampled and hold with a sampling interval T, the problem can be mathematically reformulated as the problem of solving a large overdetermined system of linear equations

$$\mathbf{D}\mathbf{x} = \mathbf{y} \tag{2}$$

where

$\mathbf{x} = [a_1, b_1, a_2, b_2, \ldots, a_n, b_n]^T \in \mathbf{R}^{2n}$ and
$\mathbf{y} = [y(T), y(2T), \ldots, y(mT)] \in \mathbf{R}^m$

$$\mathbf{D} = \begin{bmatrix} \sin(\omega T) & \cos(\omega T) & \cdots & \sin(n\omega T) & \cos(n\omega T) \\ \sin(2\omega T) & \cos(2\omega T) & \cdots & \sin(2n\omega T) & \cos(2n\omega T) \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \sin(m\omega T) & \cos(m\omega T) & \cdots & \sin(mn\omega T) & \cos(mn\omega T) \end{bmatrix} =$$

$$= \left[\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4, \ldots, \mathbf{d}_{2n-1}, \mathbf{d}_{2n}\right].$$

3 LEAST SQUARES APPROACH

The above formulated problem can be solved using standard LS approach. According to this approach we minimize the energy (objective) function defined as

$$E_2(\mathbf{x}) = \tfrac{1}{2}\|\mathbf{e}\|_2^2 = \tfrac{1}{2}\mathbf{e}^T\mathbf{e} \tag{3}$$

where

$$\mathbf{e} = \mathbf{D}\mathbf{x} - \mathbf{y} = \left[\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_{2n}\right]\mathbf{x} - \mathbf{y}$$

In this case the estimated vector **x** will be given by

$$\mathbf{x} = \left[\mathbf{D}^T\mathbf{D}\right]^{-1}\mathbf{D}^T\mathbf{y} \tag{4}$$

The formula (4) requires to compute inverse matrix and it is rather very time consuming. In order to find the LS estimates in real time we can employ the Hopfield type neural networks [9, 10]. According to this approach we need to solve (simulate) the system of differential equations

$$\frac{d\mathbf{x}}{dt} = -\mu \mathbf{D}^T(\mathbf{D}\mathbf{x} - \mathbf{y}) = -\mu \mathbf{D}^T \mathbf{e} \tag{5}$$

where

$\mu > 0$ is the appropriate learning rate.

The matrix differential equation (5) can be written in a scalar form as

$$\frac{da_i}{dt} = -\mu \sum_{k=1}^{m} e_k \sin(ik\omega T) \tag{6}$$

$$\frac{db_i}{dt} = -\mu \sum_{k=1}^{m} e_k \cos(ik\omega T) \tag{7}$$

where

$$e_k = \sum_{i=1}^{n} [a_i \sin(ik\omega t) + b_i \cos(ik\omega t)] - y(kT)$$

$i=1,2,...,n$; $k=1,2,...,m$

The LS technique is relatively simple. However the approach is optimal only if matrix **D** is exactly known and the vector **y** is perturbated by a Gaussian noise. It provides rather poor estimation if noise has impulsive character, i.e. if the data have large errors called outliers.

4 TOTAL LEAST SQUARES APPROACH

The standard LS method assumes that the matrix **D** is exactly determined and only the vector **y** is contaminated by noise. In practice the matrix **D** is also perturbated by error. In fact, the frequency ω is not exactly known. Moreover, it can be slightly fluctuated during the measurement, and these fluctuations are unknown. Furthermore, the sampling period is sometimes not fixed but also fluctuates (i.e. the sampling of the signal is not ideally regular). For these reasons, in order to obtain more reliable and robust solution we apply the total least squares (TLS) approach, known in the statistics literature as *orthogonal regression* and *errors-in-variable regression*. The TLS criterion assumes errors both in the matrix **D** and in the vector **y**. Therefore, whereas LS minimizes the prediction error, TLS minimizes the error normal to the graph of the linear predictor. Applying the TLS criterion we obtain the following linear matrix equation

$$(\hat{\mathbf{D}} + \mathbf{R})\mathbf{x}_{TLS} = \hat{\mathbf{y}} + \mathbf{r} \tag{8}$$

where

$\mathbf{D} = \hat{\mathbf{D}} + \mathbf{R}$ and $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{r}$

$\hat{\mathbf{D}} \in R^{m\times n}, \hat{\mathbf{y}} \in R^m$ are exact but unknown matrices,

$\mathbf{R} \in R^{m\times n}, \mathbf{r} \in R^m$ are corresponding errors.

In other words, the TLS problem can be formulated as the optimization problem: to find the vector $\mathbf{x}^*_{TLS}$ that minimizes

$$\|\mathbf{R}\|_F^2 + \|\mathbf{r}\|_F^2 \tag{9}$$

subject to the equality constraints (8), where $\|\mathbf{R}\|_F$ denotes the Frobenius norm of **R**. The main numerical tool for solving the TLS problem is the singular value decomposition (SVD) of the extended matrix [11]

$$\tilde{\mathbf{D}} = [\mathbf{D}, \mathbf{y}] = \mathbf{U}\Sigma\mathbf{V}^T \tag{10}$$

The TLS solution is computed as

$$\mathbf{x}^*_{TLS} = -\frac{1}{\upsilon_{2n+1,2n+1}} \cdot \left[\upsilon_{1,2n+1}, \upsilon_{2,2n+1}, \dots, \upsilon_{2n,2n+1}\right]^T \tag{11}$$

where

$\upsilon_{2n+1} = \left[\upsilon_{1,2n+1}, \upsilon_{2,2n+1}, \dots, \upsilon_{2n,2n+1}\right]^T$ is the right singular vector associated to the smallest singular value σ_{n+1} of the extended matrix [**D**, **y**]. As the singular value σ_{n+1} goes to zero, the LS and TLS approaches to each other. It is important to point out that the LS solution is based on the minimization of the sum of the squared errors

$$\mathbf{e} = \mathbf{D}\mathbf{x} - \mathbf{y} \tag{12}$$

while the TLS solution is based on the minimization the sum of weighted squared errors [11]. In other words, the TLS problem can be formulated as the minimization of the energy function

$$E_{TLS}(\mathbf{x}) = \frac{\|\mathbf{D}\mathbf{x} - \mathbf{y}\|_2^2}{1 + \mathbf{x}^T\mathbf{x}} = \sum_{k=1}^{m} \frac{\left(\sum_{j=1}^{2n} d_{kj} x_j - y_k\right)^2}{1 + \mathbf{x}^T\mathbf{x}} \tag{13}$$

In comparison with the standard LS technique obtaining the solution of the TLS problem is generally quite burdensome and very time consuming. This probably because the TLS approach has not been as widely used as the usual LS approach, although the TLS approach has been investigated in robust statistics long ago.

In order to simplify complexity of the algorithm we introduce an instantaneous error defined as

$$e(t) = \mathbf{S}^T(t)(\mathbf{D}\mathbf{x} - \mathbf{y}) = \mathbf{S}^T(t)\mathbf{e} = \sum_{k=1}^{m}\left(\sum_{j=1}^{2n} d_{kj} x_j - y_k\right) S_k(t) = \sum_{j=1}^{2n} \tilde{d}_j x_j(t) - \tilde{y}(t) \tag{14}$$

where

$\mathbf{S}(t) = \left[S_1(t), S_2(t), \dots, S_m(t),\right]^T$ is the vector of zero-mean independent identically distributed (i.i.d.) externally excitation signals (e.g. zero-mean noise sources),

$$\tilde{d}_j(t) \stackrel{df}{=} \sum_{k=1}^{m} d_{kj} S_k(t) \tag{15}$$

$$\tilde{y}(t) \overset{df}{=} \sum_{k=1}^{m} y_k S_k(t) \qquad (16)$$

The TLS problem can be now reformulated as minimization of the following instantaneous energy function

$$E_{TLS}[\mathbf{x}(t)] = \frac{1}{2}\frac{e^2(t)}{\mathbf{x}^T\mathbf{x}+1} \qquad (17)$$

Applying the gradient descent approach we obtain the system of differential equations

$$\begin{aligned}\frac{dx_j(t)}{dt} &= -\mu(t)\frac{\partial E_{TLS}[\mathbf{x}(t)]}{\partial x_j} = \\ &= -\mu(t)e(t)\frac{\tilde{d}_j(t)(\mathbf{x}^T\mathbf{x}+1) - e(t)x_j(t)}{(\mathbf{x}^T\mathbf{x}+1)}\end{aligned} \qquad (18)$$

where

$\mu(t) > 0$.

The above set of differential equations can be further simplified, after linearization, as

$$\frac{dx_j}{dt} = -\mu(t)e(t)[\tilde{d}_j(t) + \tilde{y}(t)x_j(t)] \qquad (19)$$

Functional block diagram illustrating implementation of the algorithm (19) is shown in the Fig. 1. The block diagram can be considered as a single neuron with synapses x_j, learned (adjusted) according to eqn. (19).

5 ROBUST TLS ALGORITHM (RTLS)

In practice, the errors are more or less isotropically - contrary to the assumptions. The TLS algorithm is rather very sensitive to noise and kind of distributed errors, especially in the presence of outliers.

This extreme sensitivity implies that we need to modify or generalize TLS algorithm in order to eliminate, as far as possible, outlying points or large spiky noise. This fact was main motivation for development and investigation of a new generalized algorithm called *Robust Total Least Squares (RTLS)* algorithm.

The learning algorithm (19) can be extended as follows:

$$\frac{dx_j(t)}{dt} = -\mu(t)\Psi[e(t)][\tilde{d}_j(t) + \alpha\tilde{y}(t)x_j(t)] \qquad (20)$$

where

$\Psi(e)$ is nonlinear activation function enabling supress or neglecting large error, e.g. $\Psi(e) = \tanh(\gamma e)$ or

$$\Psi(e) = \begin{cases} e & \text{for } |e| \le \beta \\ 0 & \text{otherwise} \end{cases}$$

and $\alpha \ge 0$ is nonnegative coefficient.

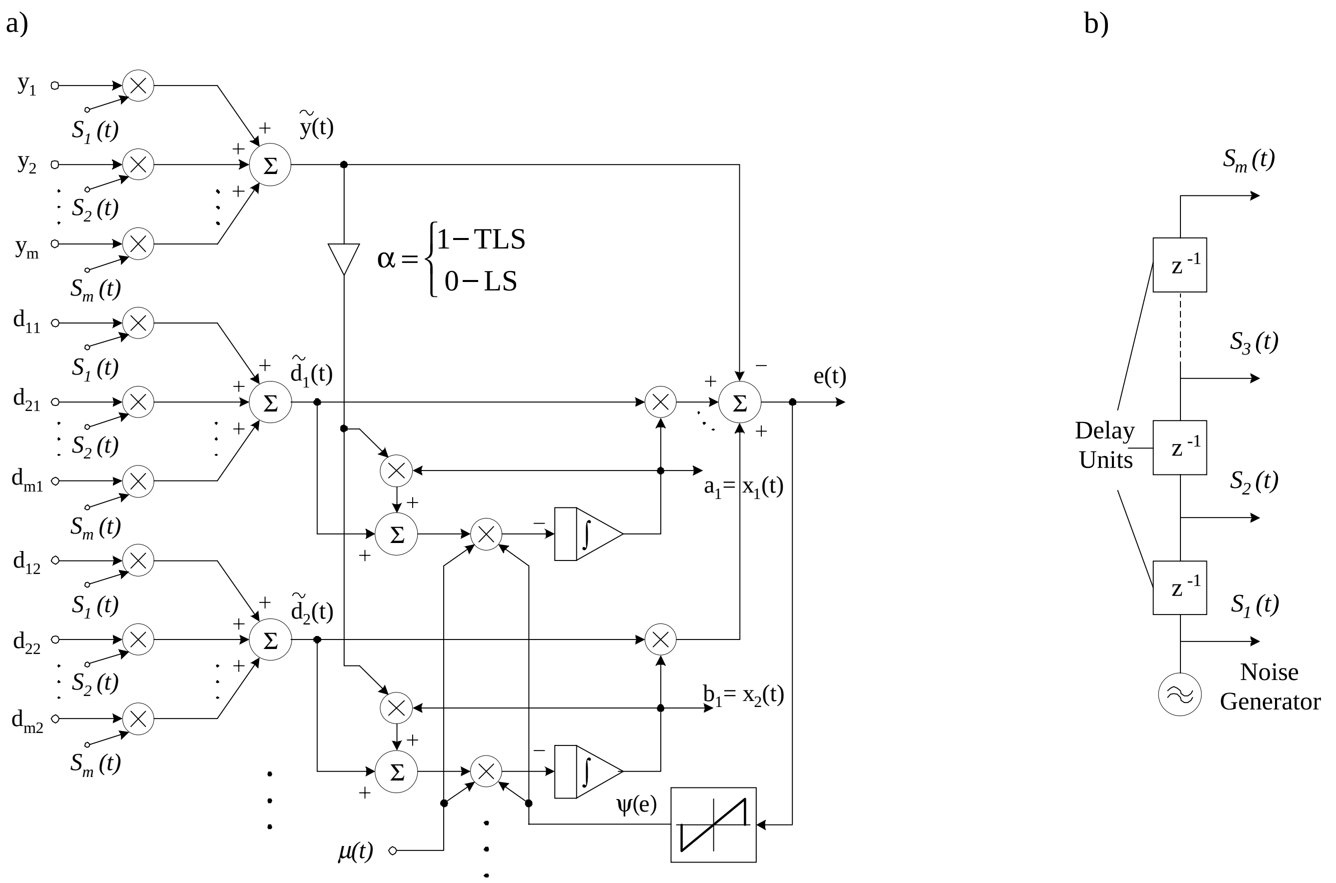

Fig. 1. a) Analog neural network for solving the estimation problem (see Eq.(20)); b) Exemplary method for generation signals S

In the special case $\Psi(e) \equiv e$ and $\alpha = 0$ we obtain standard LS algorithm and for $\Psi(e) \equiv e$ and $\alpha = 1$ obtain standard TLS algorithm. It should be noted, that by changing the value of the parameter α more or less emphasis can be given to errors of the matrix **D** with respect to errors of the vector **y** . If $\alpha = 0$, we assume that error is only in the vector **y**. On the other hand, for large α (say 100) it can be assumed that the vector **y** is almost free of error and the all error lies in the data matrix **D**. Such extreme case is refferred to as so called DLS (data least squares) problem (since error occurs only in **D** but not in **y**). The RTLS algorithm (20) could be transformed to time-discrete form as

$$x_j(k+1) = x_j(k) - \eta(k)\Psi[e(k)][\tilde{d}_j(k) + \alpha\tilde{y}(k)x_j(k)] \qquad (21)$$

6 SIMULATION EXPERIMENTS

Extensive computer simulation experiments have confirmed the validity and performance of the proposed algorithms. The associated network has been simulated on computer. Owing to limited space, we shall present only some illustrative results. First, we have simulated a signal

$$x(t) = 100\sin(\omega t) + 100\cos(\omega t) + r(t) \qquad (22)$$

The algorithm has been investigated for the frequency of the basic component equal to 50 Hz. The investigations have been done, when the frequency of the simulated signal changes. The sampling window was T = 0.02 and T = 0.04s. The number of samples N = 20 - 100. When the frequency of the simulated signal 49,5 Hz the estimation errors were less 1%. Slightly better results we have obtained using the TLS and the RTLS method. We have applied the learning rate $\mu = 1000$. We have also investigated the influence of the random noise in the measured (simulated) signal on the estimation error. The LS, the TLS and especially the RTLS methods show a great immunity against the noise. For the noise level up to 10% the estimation errors were less than 1%.

In all the investigations the level of the auxiliary noise $S_n(t)$ was about 1% of the amplitudes of the simulated signal.

When using the LS algorithm the trajectories of the estimated parameter converge in less than 5 ms, and for TLS or RTLS algorithm in less than 2 ms.

CONCLUSIONS

Adaptive analogue neural networks represent a very promising approach for high-speed estimation of parameters of signals. In this paper new algorithms and architectures of neuron-like adaptive circuits have been developed, according to the LS, TLS and RTLS optimization criteria, applying the gradient descent approach. They are more robust against noise in the measured signal than other known neural network algorithms. The network based on the TLS and RTLS criteria optimize the estimation under the assumption, that the signal model can be also perturbated (frequency or sampling interval fluctuation and so forth). The robust TLS estimates are usually better and more reliable than the corresponding LS estimates. While the TLS and RTLS estimates do have a bias this seems to be dependent relatively weakly on the noise level.

The important conclusion we draw here is that the RTLS estimation may be useful when large amounts of errors and noise occur. The RTLS algorithm is generalization of well known LMS rule and could be in some applications superior to family of LMS algorithms.

Extensive computer simulation experiments confirmed the validity and performance of the proposed algorithms.

REFERENCES

[1] **S. Osowski (1992)**: Neural Networks for Estimation of Harmonic Components in a Power System, *IEE Proc.- C*, 1992, vol. 139, No. 2, pp. 129-135.

[2] **A. Van Den Bos (1988)**: Nonlinear Least-Absolute Values and Minimax Model Fitting, *Automatica*, 1988, vol. 24, No. 6, pp. 803-808.

[3] **A. Cichocki, T. Lobos (1994)**: Artificial Neural Networks for Real-Time Estimation of Basic Waveforms of Voltages and Currents, *IEEE Trans. on Power Systems*,1994, vol. 9, No. 2, pp. 612-618.

[4] **A. Cichocki, T. Lobos (1993)**: Neural Networks for Real-Time Estimation of the Basic Waveforms of Voltages and Currents Encountered in Power Systems, *Int. J. Electronics*, 1993, vol. 74, No. 2, pp. 307-318.

[5] **A. Cichocki, P. Kostyla, T. Lobos, Z. Waclawek (1994)**: Neural Networks for Real-Time Estimation of Signals Encountered in Power Systems, presented at the *IMACS Int. Symp. on Signal Processing, Robotics and Neural Networks*, Lille (France) 1994, Conf. Proceedings, pp. 317-320.

[6] **S. Osowski (1994)**: SVD Technique for Estimation of Harmonic Components in a Power System: a Statistical Approach, *IEE Proc.- C*, 1994, vol. 141, No. 5, pp. 473-479.

[7] **T. Lobos (1989)**: Nonrecursive Methods for Real-Time Determination of Basic Waveforms of Voltages and Currents, *IEE Proc.- C*, 1989, vol. 136, No. 6, pp. 347-351.

[8] **K.F. Eichhorn, T. Lobos, P. Ruczewski (1993)**: Constrained Frequency Domain Algorithms for Determination of Parameters of Fundamental Sinewave of Signals, *IEE Proc.- C,* 1993, vol. 140, No. 6, pp. 477-480.

[9] **S. - I. Amari (1990)**: Mathematical Foundations of Neurocomputing, *Proc. IEEE*, 1990, vol. 78, No. 9, pp. 1443-1463.

[10] **A. Cichocki, R. Unbehauen (1993)**: Neural Network for Optimization, Computing and Signal Processing, Stuttgart, *Teubner -Wiley*, 1993, pp. 70-84

[11] **G.H. Golub, C.F. Van Loan (1980)**: An Analysis of the Total Least Squares Problem. *SIAM J. Numer Anal*, 1980, vol. 17, pp. 883-893.