Towards robust identification of nonstationary systems

1st Artur Gańcza
Department of Automatic Control
Gdansk University of Technology
Gdansk, Poland
artgancz@pg.edu.pl

2nd Marta Chełkowska
Department of Automatic Control
Gdansk University of Technology
Gdansk, Poland
s175673@student.pg.edu.pl

3rd Natalia Kleinschmidt
Department of Automatic Control
Gdansk University of Technology
Gdansk, Poland
s171704@student.pg.edu.pl

***Abstract*—The article proposes a fast, two-stage method for the identification of nonstationary systems. The method uses iterative reweighting to robustify the identification process against the outliers in the measurement noise and against the numerical errors that may occur at the first stage of identification. We also propose an adaptive algorithm to optimize the values of the hyperparameters that are crucial for this new method.**

***Index Terms*—nonstationary systems, identification, adaptive algorithms, local basis function method, iterative reweighting**

I. Introduction

Recently, a new two-step method for the identification of nonstationary systems has been proposed, called the fast local basis function (fLBF) method [1]. This new method is a close approximation of the local basis function method proposed in [2]. It allows one to obtain accurate estimates of time-varying system parameters while maintaining low computational complexity. The most recent application perfectly suited for this identification method is self-interference cancellation in full-duplex (FD) underwater acoustic (UWA) communications [3], [4]. In full-duplex communication, both communication devices exchange information simultaneously using the same bandwidth. As a result, the recorded signal from the far-end transmitter is contaminated by the signal from the near-end transmitter. This self-interference signal consists of the known transmitted signal convolved with a time-varying impulse response of the self-interference channel [3]. Self-interference cancellation can be performed effectively if one has reliable estimates of the time-varying parameters of the self-interference channel. An important feature of this application is that it allows one to operate with some decision delay, which means that one can use non-causal identification methods such as the fLBF technique.

The fLBF method is a two-stage algorithm. The first stage (preestimation) provides "raw" but approximately unbiased estimates of parameter trajectories, regardless of the type and rate of parameter variation. The preestimates are obtained by "inverse filtering" the exponentially weighted least squares (EWLS) estimates of the parameter trajectories. The resulting equations can be solved using numerical methods, but these methods increase the variance of the preestimation errors and may introduce impulsive disturbances into the preestimates. Therefore, at the second stage (postfiltering) we propose an improved filtering scheme that is more robust to such disturbances. The new scheme, called iterative fLBF (IfLBF), can improve estimation results also when the ambient noise is likely to contain impulsive disturbances.

Computer simulations were carried out at the Academic Computer Centre CI TASK in Gdańsk.

II. Problem formulation

Consider the nonstationary finite impulse response system governed by the following equation

$$y(t) = \sum_{i=1}^{n} \theta_i(t)u(t-i+1) + e(t) = \boldsymbol{\theta}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t) + e(t), \quad (1)$$

where $t = \ldots, -1, 0, 1, \ldots$, denotes discrete (dimensionless) time, $\{y(t)\}$ is the output signal, $\boldsymbol{\varphi}(t) = [u(t), \ldots, u(t-n+1)]^{\mathrm{T}}$, denotes the regression vector, containing time-delayed samples of the input signal $\{u(t)\}$, $\boldsymbol{\theta}(t) = [\theta_1(t), \ldots, \theta_n(t)]^{\mathrm{T}}$ denotes the vector of unknown time-varying parameters, and $\{e(t)\}$ is measurement noise. Such models are very popular in modeling both terrestrial [5] and underwater acoustic telecommunication channels [6], [7]. In the remainder of this paper, we will make the following assumptions, which are typically met in telecommunications applications

(A1) $\{u(t)\}$ is a sequence of zero-mean independent and identically distributed random variables with variance σ_u^2.

(A2) $\{e(t)\}$, independent of $\{u(t)\}$, is a sequence of zero-mean independent and identically distributed random variables with variance σ_e^2.

(A3) $\{\boldsymbol{\theta}(t)\}$, independent of $\{u(t)\}$ and $\{e(t)\}$, is a uniformly bounded sequence.

The main contribution of this paper is application of the iterative reweighting methods in identification of nonstationary systems and derivation of the algortihm for adaptive choice of the design parameters.

III. Preestimation

The two-stage identification method described in this paper consists of preestimation and postfiltering. The first stage yields approximately unbiased but very noisy estimates of the parameter trajectories, which need to be further processed to obtain reliable estimates - hence the name preestimates. Preestimates were introduced in [1] and further analyzed in [8]. They are based on exponentially weighted least squares (EWLS) estimates of system parameters, defined as

$$\widehat{\boldsymbol{\theta}}(t) = \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{t} \lambda^{t-i}[y(i) - \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t)]^2, \tag{2}$$

where $\lambda \in (0,1)$ is the so-called forgetting constant. The EWLS estimates are obtained as a solution to the following system of linear equations

$$\mathbf{R}(t)\widehat{\boldsymbol{\theta}}(t) = \mathbf{r}(t), \tag{3}$$

where

$$\begin{aligned} \mathbf{R}(t) &= \sum_{i=1}^{t} \lambda^{t-i}\boldsymbol{\varphi}(i)\boldsymbol{\varphi}^{\mathrm{T}}(i), \\ \mathbf{r}(t) &= \sum_{i=1}^{t} \lambda^{t-i}\boldsymbol{\varphi}(i)y(i). \end{aligned} \tag{4}$$

Then the preestimates are defined as

$$\widetilde{\boldsymbol{\theta}}(t) = L(t)\widehat{\boldsymbol{\theta}}(t) - \lambda L(t-1)\widehat{\boldsymbol{\theta}}(t-1), \tag{5}$$

where $L(t) = \sum_{i=1}^{t} \lambda^{t-i} = \lambda L(t-1) + 1$ is the effective number of observations of the EWLS algorithm. In the steady state, when $L(t) \cong L_\infty = \frac{1}{1-\lambda}$, the preestimation formula simplifies to

$$\widetilde{\boldsymbol{\theta}}(t) = \frac{1}{1-\lambda}[\widehat{\boldsymbol{\theta}}(t) - \lambda\widehat{\boldsymbol{\theta}}(t-1)]. \tag{6}$$

Note that under the assumptions listed in the previous section, it can be shown [9] that the mean path of the EWLS estimates is a result of lowpass filtering the true parameter trajectories

$$\mathrm{E}[\widehat{\boldsymbol{\theta}}(t)] \cong H(q^{-1})\boldsymbol{\theta}(t), \tag{7}$$

where $H(q^{-1}) = \frac{1-\lambda}{1-\lambda q^{-1}}$ denotes the filter associated with the EWLS method, and q^{-1} is a backward shift operator, namely $q^{-1}\boldsymbol{\theta}(t) = \boldsymbol{\theta}(t-1)$. The EWLS estimates can be written down as

$$\widehat{\boldsymbol{\theta}}(t) \cong H(q^{-1})\boldsymbol{\theta}(t) + \boldsymbol{\eta}(t), \tag{8}$$

where $\boldsymbol{\eta}(t)$ is a zero-mean noise. As a consequence, the steady-state preestimates can be expressed as

$$\widetilde{\boldsymbol{\theta}}(t) \cong \boldsymbol{\theta}(t) + \frac{1}{1-\lambda}[\boldsymbol{\eta}(t) - \lambda\boldsymbol{\eta}(t-1)]. \tag{9}$$

This means that the preestimates are approximately unbiased. Since the preestimates can be seen as an effect of the “inverse” (highpass) filtering of the EWLS estimates, the value of the forgetting constant should be chosen carefully. A too large value of λ will result in a larger variability of the preestimation noise, while using a too small value of λ can make the steady-state equivalent memory of the EWLS algorithm ($N_\infty = \frac{1+\lambda}{1-\lambda} \cong \frac{2}{1-\lambda}$) smaller than the number of system parameters n, which is questionable from a statistical point of view. Hence, the rule for choosing λ that works well in practice turns out to be

$$\lambda = \max\left\{0.9, 1 - \frac{2}{n}\right\}. \tag{10}$$

Remark: It was recently noted in [11] that preestimates tend to inherit some delay introduced by the EWLS algorithm, even though the initial analysis suggests that they are approximately unbiased. Methods for dealing with this delay have been presented in [11].

IV. EWLS WITH THE DICHOTOMOUS COORDINATE DESCENT

Recently, many numerical algorithms for solving the system of linear equations (3) have been proposed to reduce the numerical complexity associated with the EWLS method. One of them is the Dichotomous Coordinate Descent (DCD) algorithm proposed in [10]. The authors suggest to solving numerically the system of auxiliary equations

$$\mathbf{R}(t)\Delta\widehat{\boldsymbol{\theta}}^{\mathrm{DCD}}(t) = \mathbf{r}_0(t), \tag{11}$$

and find the estimates at the current time instant using estimates from the previous time instant

$$\widehat{\boldsymbol{\theta}}^{\mathrm{DCD}}(t+1) = \widehat{\boldsymbol{\theta}}^{\mathrm{DCD}}(t) + \Delta\widehat{\boldsymbol{\theta}}^{\mathrm{DCD}}(t), \tag{12}$$

where

$$\mathbf{r}_0(t) = \Delta\mathbf{r}(t) - \Delta\mathbf{R}(t)\widehat{\boldsymbol{\theta}}^{\mathrm{DCD}}(t-1) + \boldsymbol{\varepsilon}_r(t-1), \tag{13}$$

and $\Delta\mathbf{r}(t) = \mathbf{r}(t) - \mathbf{r}(t-1)$, $\Delta\mathbf{R}(t) = \mathbf{R}(t) - \mathbf{R}(t-1)$,

$$\boldsymbol{\varepsilon}_r(t) = \mathbf{r}(t) - \mathbf{R}(t)\widehat{\boldsymbol{\theta}}^{\mathrm{DCD}}(t). \tag{14}$$

Note, however, that the preestimates obtained using the DCD algorithm will have a larger variability of estimation errors because the combination of (3) and (14) yields

$$\widehat{\boldsymbol{\theta}}^{\mathrm{DCD}}(t) = \widehat{\boldsymbol{\theta}}(t) - \boldsymbol{\delta}(t), \tag{15}$$

where $\boldsymbol{\delta}(t) = \mathbf{R}^{-1}(t)\boldsymbol{\varepsilon}_r(t)$, which leads to

$$\begin{aligned} \widetilde{\boldsymbol{\theta}}^{\mathrm{DCD}}(t) &= \frac{1}{1-\lambda}[\widehat{\boldsymbol{\theta}}^{\mathrm{DCD}}(t) - \lambda\widehat{\boldsymbol{\theta}}^{\mathrm{DCD}}(t-1)] \\ &\cong \boldsymbol{\theta}(t) + \frac{1}{1-\lambda}[\boldsymbol{z}(t) - \lambda\boldsymbol{z}(t-1)], \end{aligned} \tag{16}$$

where $\boldsymbol{z}(t) = \boldsymbol{\eta}(t) - \boldsymbol{\delta}(t)$ is a noise of a larger variance than $\boldsymbol{\eta}(t)$.

Furthermore, because the solution to the system of auxiliary equations (11) provided by the DCD algorithm uses a very limited number of bits, the errors are of a similar nature to quantization errors. This fact, combined with the highpass nature of preestimation algorithms, causes the distribution of preestimation errors to have heavier tails when the DCD algorithm is used. Therefore, later in the paper, we propose an iterative reweighting technique to improve the accuracy of the final estimates obtained after filtering the DCD-based preestimates.

V. Postfiltering

Due to the large variability of the preestimation noise, it is necessary to perform additional filtering in order to obtain statistically meaningful results. The method proposed in [1] is a Savitzky-Golay filtering [12]. Assume that locally, within the analysis window $T_k(t) = [t-k, t+k]$ centered around the current time t, each parameter trajectory can be modeled as a linear combination of m known functions of time, called basis functions

$$\theta_j(t+i) = \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j(t), \quad j=1,\ldots,n,\ i \in I_k = [-k,k], \tag{17}$$

where $\mathbf{f}(i) = [f_1(i),\ldots,f_m(i)]^{\mathrm{T}},\ i \in I_k$ is a vector of basis functions and $\boldsymbol{\alpha}_j(t) = [\alpha_{j1},\ldots,\alpha_{jm}(t)]^{\mathrm{T}},\ j=1,\ldots,n,\ i \in I_k$ is a vector of basis function coefficients.

The final fLBF estimates are defined as follows

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t) &= \arg\min_{\boldsymbol{\alpha}_j} \sum_{i=-k}^{k} w(i)[\widetilde{\theta}_j(t+i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j]^2 \\ \widehat{\theta}_j^{\mathrm{fLBF}}(t) &= \mathbf{f}^{\mathrm{T}}(0)\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t), \quad j=1,\ldots,n, \end{aligned} \tag{18}$$

where $w(i),\ i \in I_k$ is typically a nonnegative, bell-shaped weighting sequence, obeying $w(0)=1$, used to put more emphasis on data closer to the center of the analysis window.

It is easy to check that

$$\widehat{\theta}_j^{\mathrm{fLBF}}(t) = \sum_{i=-k}^{k} h(i)\widetilde{\theta}_j(t+i),\ j=1,\ldots,n, \tag{19}$$

where $h(i),\ i \in I_k$ is the impulse response associated with the basis functions

$$h(i) = \mathbf{f}^{\mathrm{T}}(0)\left[\sum_{i=-k}^{k} w(i)\mathbf{f}(i)\mathbf{f}^{\mathrm{T}}(i)\right]^{-1} w(i)\mathbf{f}(i), \quad i \in I_k. \tag{20}$$

It was shown in [1] that, under the assumptions listed above, the fLBF estimates closely approximate the estimates obtained by the computationally more demanding LBF method [2].

VI. Iterative reweighted fLBF (IfLBF)

The least absolute deviation (LAD) method is a well-known method for robustifying the estimation results when the presence of outliers in the output signal is suspected [13], [14], [15]. It also turns out that LAD is a maximum likelihood estimator when the noise obeys the Laplace distribution [13]. A step towards the application of this type of regression in identification was made in [16] for stationary systems. In the aforementioned article the well-known iterative reweighting technique [17], [18] was used. Applying such a method to the identification of nonstationary systems with the original LBF estimator would be computationally expensive. Here we show how to incorporate iterative reweighting to make the identification process more robust while keeping the computational complexity proportional to the number of system parameters. Note that the ℓ^1 cost function

$$J_j(t) = \sum_{i=-k}^{k} |\widetilde{\theta}_j(t+i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j|, \quad j=1,\ldots,n, \tag{21}$$

can be approximated as

$$\begin{aligned} J_j(t) &= \sum_{i=-k}^{k} \frac{[\widetilde{\theta}_j(t+i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j]^2}{|\widetilde{\theta}_j(t+i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j|} \\ &= \sum_{i=-k}^{k} \frac{[\widetilde{\theta}_j(t+i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j]^2}{|\widetilde{\theta}_j(t+i) - \theta_j(t+i)|} \\ &\cong \sum_{i=-k}^{k} w_j^{\mathrm{I}}(t+i)[\widetilde{\theta}_j(t+i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j]^2, \quad j=1,\ldots,n, \end{aligned} \tag{22}$$

where the weighting sequence is defined as $w_j^{\mathrm{I}}(t) = \frac{1}{|\varepsilon_j(t)|},\ i \in I_k,\ j=1,\ldots,n,$ and

$$|\varepsilon_j(t)| = \max\{|\widetilde{\theta}_j(t) - \widehat{\theta}_j^{\mathrm{fLBF}}(t)|, \varepsilon_0\}, \quad j=1,\ldots,n, \tag{23}$$

where ε_0 is a small positive constant introduced to avoid numerical problems. Minimizing such a cost function yields an iterative reweighted fLBF (IfLBF) estimate

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}_j^{\mathrm{IfLBF}}(t) &= \arg\min_{\boldsymbol{\alpha}_j} \sum_{i=-k}^{k} w_j^{\mathrm{I}}(t+i)[\widetilde{\theta}_j(t+i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j]^2 \\ \widehat{\theta}_j^{\mathrm{IfLBF}}(t) &= \mathbf{f}^{\mathrm{T}}(0)\widehat{\boldsymbol{\alpha}}_j^{\mathrm{IfLBF}}(t), \quad j=1,\ldots,n. \end{aligned} \tag{24}$$

Similar to the fLBF case, the final estimates can be expressed as

$$\widehat{\theta}_j^{\mathrm{IfLBF}}(t) = \sum_{i=-k}^{k} h_j^{\mathrm{I}}(t,i)\widetilde{\theta}_j(t+i),\ j=1,\ldots,n, \tag{25}$$

where the corresponding impulse response is time-dependent

$$\begin{aligned} h_j^{\mathrm{I}}(t,i) = \mathbf{f}^{\mathrm{T}}(0)&\left[\sum_{i=-k}^{k} w_j^{\mathrm{I}}(t+i)\mathbf{f}(i)\mathbf{f}^{\mathrm{T}}(i)\right]^{-1} \times \\ &\times w_j^{\mathrm{I}}(t+i)\mathbf{f}(i), \quad i \in I_k,\ j=1,\ldots,n. \end{aligned} \tag{26}$$

Note that, unlike in the typical reweighting methods, here we use the global estimation errors $\widetilde{\theta}_j(t+i) - \widehat{\theta}_j^{\mathrm{fLBF}}(t+i),\ i \in I_k,\ j=1\ldots,n$, instead of local estimation errors $\widetilde{\theta}_j(t+i) - \mathbf{f}^{\mathrm{T}}(i)\widehat{\boldsymbol{\alpha}}_j^{\mathrm{IfLBF}}(t),\ i \in I_k,\ j=1\ldots,n$, which can lead to better accuracy of the IfLBF estimates.

Note also that the proposed iterative method increases the computational complexity of the final algorithm, which is still proportional to the number of system parameters n.

VII. Hyperparameter optimization

In the case of the IfLBF algorithm, one must decide on the value of the constant ε_0, which can seriously affect the quality of the final estimates. One can develop a technique similar to the leave-one-out cross-validation described in [1]. In this approach, one runs several algorithms in parallel, each with different constants $\varepsilon_0 \in \mathcal{E} = \{\varepsilon_0^1,\ldots,\varepsilon_0^L\}$, and at each point in time one chooses the algorithm that minimizes the local sum of squared leave-one-out interpolation errors

$$J_0(t) = \sum_{i=-M}^{M} [\epsilon_0(t+i|\varepsilon_0)]^2, \tag{27}$$

where $\epsilon_0(t|\varepsilon_0) = y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widehat{\boldsymbol{\theta}}_0^{\mathrm{IfLBF}}(t|\varepsilon_0)$, and

$$\widehat{\boldsymbol{\alpha}}_{j,0}^{\mathrm{IfLBF}}(t|\varepsilon_0) = \arg\min_{\boldsymbol{\alpha}_j} \sum_{\substack{i=-k \\ i\neq 0}}^{k} w_j^{\mathrm{I}}(t+i|\varepsilon_0)[\widetilde{\theta}_j(t+i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j]^2$$
$$\widehat{\theta}_{j,0}^{\mathrm{IfLBF}}(t|\varepsilon_0) = \mathbf{f}^{\mathrm{T}}(0)\widehat{\boldsymbol{\alpha}}_{j,0}^{\mathrm{IfLBF}}(t|\varepsilon_0), \quad j = 1, \ldots, n. \tag{28}$$

Using the well-known Sherman-Morrison formula [19], one can show that

$$\widehat{\theta}_{j,0}^{\mathrm{IfLBF}}(t|\varepsilon_0) = \frac{1}{1-\beta_j(t|\varepsilon_0)}[\widehat{\theta}_j^{\mathrm{fLBF}}(t|\varepsilon_0) - \beta_j(t|\varepsilon_0)\widetilde{\theta}_j(t)], \quad j = 1, \ldots, n, \tag{29}$$

where $\beta_j(t|\varepsilon_0) = w_j^{\mathrm{I}}(t|\varepsilon_0)\mathbf{f}^{\mathrm{T}}(0)\mathbf{F}_j^{-1}(t|\varepsilon_0)\mathbf{f}(0)$, and $\mathbf{F}_j(t|\varepsilon_0) = \sum_{i=-k}^{k} w_j^{\mathrm{I}}(t+i|\varepsilon_0)\mathbf{f}(i)\mathbf{f}^{\mathrm{T}}(i)$. Note that using the results of the averaging theory [20], one obtains

$$\beta_0 = \mathrm{E}[\beta_j(t|\varepsilon_0)] \cong \mathbf{f}^{\mathrm{T}}(0)\left[\sum_{i=-k}^{k}\mathbf{f}(i)\mathbf{f}^{\mathrm{T}}(i)\right]^{-1}\mathbf{f}(0), \tag{30}$$

which stems from the fact that under assumptions listed above, the fLBF estimator is approximately unbiased [1], hence $\mathrm{E}[w_j^{\mathrm{I}}(t|\varepsilon_0)] \cong \frac{1}{\varepsilon_0}$. Therefore, one can use the simplified formula for adaptation

$$\widehat{\theta}_{j,0}^{\mathrm{IfLBF}}(t|\varepsilon_0) \cong \frac{1}{1-\beta_0}[\widehat{\theta}_j^{\mathrm{fLBF}}(t) - \beta_0\widetilde{\theta}_j(t)], \quad j = 1, \ldots, n, \tag{31}$$

Note that the same procedure can be incorporated to choose the length of the analysis interval $K = 2k+1$ and the number of basis functions m.

VIII. Computer simulations

The algorithms described in this paper have been tested on a two-tap FIR system

$$y(t) = \theta_1(t)u(t) + \theta_2(t)u(t-1) + e(t), \tag{32}$$

which parameter trajectories are shown in Fig. 1 The input signal $\{u(t)\}$ was a white Gaussian noise. The simulations were performed for two signal-to-noise (SNR) ratios - 10 dB and 20 dB, corresponding to the variance of the measurement noise $\{e(t)\}$ (which was white Gaussian and independent of the input signal) equal to $\sigma_e^2 = 0.1$ and $\sigma_e^2 = 0.01$, respectively. In our simulations, Legendre polynomials were used as basis functions with $m = 5$. The rectangular weighting sequence ($w(i) \equiv 1, i \in I_k$) was used with the fLBF algorithm.

The table I shows the mean squared estimation errors (MSE) (in decibels) averaged over $T = 5000$ time samples (to avoid boundary problems, data generation started 500 samples before $t = 1$ and ended 500 samples after $t = 5000$) and 100 independent realizations of the measurement noise. The value of the constant ε_0 was chosen from the set of candidates $\mathcal{E} = \{0.5, 0.05, 0.005, 0.0005\}$, and the length of a local decision window $2M+1$ was set to 61. The algorithms with the abbreviation DCD were based on the preestimates obtained with the DCD algorithm. To compute the EWLS estimates based on the DCD algorithm, we used the following settings: $H = 1$, $M_b = 16$, $N_u = 4$ (see [10] for details). In the table, "A" denotes the exact adaptive algorithm using (29), while "A0" denotes the simplified adaptive algorithm using (31).

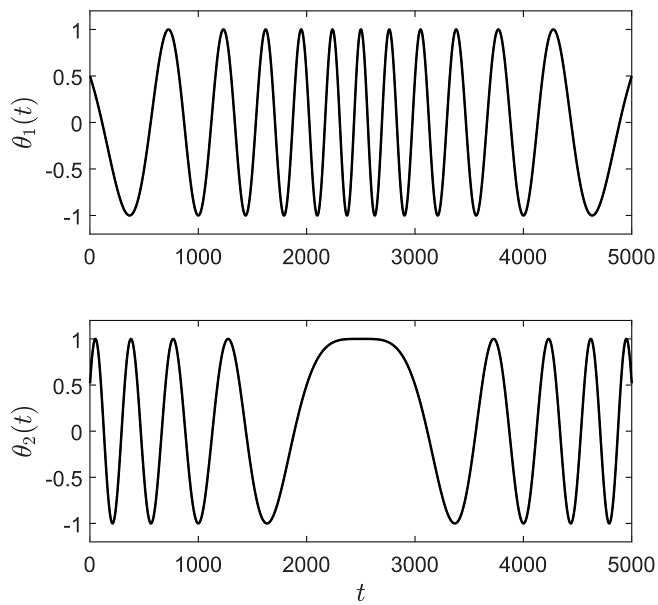

Fig. 1. True parameter trajectories of a simulated time-varying system.

In the second simulation, we used the same setup except for the noise, which this time followed the Student's t-distribution with 3 degrees of freedom. Such a distribution is characterized by the presence of outliers, which can degrade the quality of the estimates. Again, the noise variance was set to $\sigma_e^2 = 0.1$ and $\sigma_e^2 = 0.01$, corresponding to SNRs of 10 and 20 dB, respectively. The resulting MSE values in decibels are shown in table II. The simulation results show that the adaptive algorithms (both the exact (29) and the simplified (31)) provide estimates of comparable quality to the best of the algorithms included in the parallel scheme. Moreover, the proposed IfLBF estimator can significantly improve the estimation accuracy when the numerical algorithms are used to compute the EWLS estimates or in the presence of outliers in the measurement noise. The improvement is particularly noticeable at low SNR values. When the exact values of the EWLS estimates are used and there are no outliers in the measurement noise, the IfLBF algorithm provides estimates of slightly inferior quality.

IX. Conclusion

This paper presents the new iterative reweighted fLBF algorithm for the identification of nonstationary systems. The proposed method can improve the estimation results in the presence of strong noise with outliers, or when the numerical algorithms are used to find the preestimates. We also propose an adaptive algorithm, based on leave-one-out cross-validation,

TABLE I

MSE [dB] averaged over time and 100 independent realizations of Gaussian measurement noise, for a time-varying two-tap FIR system. “A” denotes the exact adaptive algorithm using (29), “A0” denotes the simplified adaptive algorithm, using (31).

SNR [dB]	Algorithm	ε_0	k = 40	k = 80	k = 120	k = 160
10	fLBF		-19.72	**-22.78**	**-24.50**	**-25.13**
	IfLBF	0.5	-19.49	-21.79	-23.02	-23.30
		0.05	-19.76	-21.99	-23.12	-23.04
		0.005	**-20.03**	-22.67	-24.05	-23.84
		0.0005	-20.00	-22.73	-24.19	-23.85
		A	-19.76	-22.18	-23.42	-23.28
		A0	-19.77	-22.18	-23.42	-23.28
	DCD + fLBF		-17.97	-21.05	-22.70	-23.52
	DCD + IfLBF	0.5	-18.80	-21.50	-23.06	-23.65
		0.05	**-19.18**	-22.01	-23.64	-23.85
		0.005	-19.11	**-22.12**	**-23.79**	**-23.89**
		0.0005	-19.04	-22.03	-23.67	-23.62
		A	-19.08	-21.97	-23.58	-23.68
		A0	-19.07	-21.96	-23.57	-23.68
20	fLBF		**-27.35**	**-30.91**	**-32.52**	**-30.70**
	IfLBF	0.5	-26.83	-29.45	-30.47	-29.04
		0.05	-24.89	-26.25	-26.58	-25.59
		0.005	-26.08	-28.24	-28.82	-27.10
		0.0005	-26.19	-28.68	-29.45	-27.17
		A	-25.76	-27.60	-28.04	-26.52
		A0	-25.75	-27.60	-28.03	-26.52
	DCD + fLBF		-23.59	-26.15	-27.40	-27.37
	DCD + IfLBF	0.5	-24.86	-27.49	-29.03	**-28.70**
		0.05	-25.24	-28.03	-29.46	-28.29
		0.005	**-25.32**	**-28.38**	**-29.87**	-28.25
		0.0005	-25.19	-28.15	-29.45	-27.58
		A	-25.27	-28.24	-29.75	-28.22
		A0	-25.27	-28.23	-29.74	-28.22

TABLE II

MSE [dB] averaged over time and 100 independent realizations of a Student’s t-distribution measurement noise with 3 degrees of freedom, for a time-varying two-tap FIR system. “A” denotes the exact adaptive algorithm using (29), “A0” denotes the simplified adaptive algorithm using (31).

SNR [dB]	Algorithm	ε_0	k = 40	k = 80	k = 120	k = 160
10	fLBF		-19.68	-22.74	-24.48	**-25.10**
	IfLBF	0.5	**-20.80**	**-23.22**	-24.49	-24.60
		0.05	-20.42	-22.65	-23.74	-23.56
		0.005	-20.51	-23.15	-24.51	-24.21
		0.0005	-20.45	-23.15	**-24.59**	-24.13
		A	-20.60	-23.01	-24.19	-23.85
		A0	-20.60	-23.01	-24.20	-23.85
	DCD + fLBF		-17.94	-21.03	-22.70	-23.52
	DCD + IfLBF	0.5	**-19.98**	-22.79	-24.44	**-24.94**
		0.05	-19.90	-22.82	**-24.49**	-24.61
		0.005	-19.63	-22.67	-24.37	-24.37
		0.0005	-19.53	-22.52	-24.14	-23.99
		A	-19.91	**-22.85**	-24.49	-24.45
		A0	-19.91	-22.85	-24.49	-24.45
20	fLBF		**-27.32**	**-30.90**	**-32.51**	**-30.69**
	IfLBF	0.5	-27.16	-29.81	-30.82	-29.24
		0.05	-25.14	-26.50	-26.81	-25.79
		0.005	-26.27	-28.46	-28.99	-27.25
		0.0005	-26.37	-28.89	-29.58	-27.37
		A	-26.02	-27.88	-28.27	-26.74
		A0	-26.01	-27.87	-28.27	-26.73
	DCD + fLBF		-23.58	-26.14	-27.40	-27.37
	DCD + IfLBF	0.5	-25.06	-27.67	-29.21	**-28.81**
		0.05	-25.57	-28.39	-29.85	-28.64
		0.005	**-25.56**	**-28.61**	**-30.12**	-28.45
		0.0005	-25.40	-28.33	-29.65	-27.67
		A	-25.55	-28.53	-30.05	-28.42
		A0	-25.55	-28.52	-30.03	-28.41

for choosing the constant ε_0, which yields estimates with accuracy comparable to the accuracy of the best estimator involved in the parallel scheme. Furthermore, we have shown that the simplified version of this adaptive algorithm yields results that are almost indistinguishable from the exact version.

References

[1] M. Niedźwiecki, M. Ciołek and A. Gańcza, “A new look at the statistical identification of nonstationary systems,” *Automatica*, vol. 118, article no. 109037, 2020.

[2] M. Niedźwiecki and M. Ciołek, “Generalized Savitzky-Golay filters for identification of nonstationary systems,” *Automatica*, vol. 108, article no. 108477, 2018.

[3] L. Shen, B. Henson, Y. Zakharov and P. Mitchell, “Digital self-interference cancellation for underwater acoustic systems,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, pp. 192–196, 2020.

[4] G. Qiao, S. Gan, S. Liu, L. Ma and Z. Sun, “Digital self-interference cancellation for asynchronous in-band full-duplex underwater acoustic communication,” *Sensors*, vol. 18, pp. 1700-1716, 2018.

[5] M. K. Tsatsanis and G. B. Giannakis, “Modelling and equalization of rapidly fading channels,” *International Journal of Adaptive Contorl and Signal Processing*, vol. 10, pp. 159–176, 1996.

[6] M. Stojanovic and J. Preisig, “Underwater acoustic communication channels: Propagation models and statistical characterization,” *IEEE Communication Magazine*, vol. 47, pp. 84–89, 2009.

[7] W. Li and J. C. Preisig, “Estimation of rapidly time-varying sparse channels,” *IEEE Journal of Oceanic Engineering*, vol. 32, pp. 927-939, 2007.

[8] M. Niedźwiecki, A. Gańcza and M. Ciołek, “On the preestimation technique and its application to identification of nonstationary systems,” *Proceedings of the 59th IEEE Conference on Decision and Control*, 2020.

[9] M. Niedźwiecki, Identification of Time-varying Processes. Wiley & Sons: New York, 2000.

[10] Y. V. Zakharov, G. P. White, and J. Liu, “Low-complexity RLS algorithms using dichotomous coordinate descent iterations,” *IEEE Transactions on Signal Processing*, vol. 56, pp. 3150-3161, 2008.

[11] M. Niedźwiecki, A. Gańcza, L. Shen and Y. Zakharov, “Adaptive identification of sparse underwater acoustic channels with a mix of static and time-varying parameters,” *Signal Processing*, article no. 108664, 2022 .

[12] R. W. Schafer, “What is a Savitzky-Golay filter?,” *IEEE Signal Processing Magazine*, vol. 28, pp. 111–117, 2011.

[13] P. Bloomfield and W. Steiger, Least Absolute Deviations: Theory, Applications and Algorithms. Birkhäuser: Boston, 1983.

[14] J. F. Claerbout and F. Muir, “Robust modelling with erratic data,” *Geophysics*, vol. 38, pp. 826–844, 1973.

[15] P. Bloomfield and W. Steiger, “Least Absolute Deviations curve-fitting,” *SIAM Journl on scientific and statistical Computing*, vol. 1, pp. 290–301, 1980.

[16] K. B. Janiszowski, “Towards estimation in the sense of the least sum of absolute errors,” *IFAC Proceedings Volumes*, vol. 31, pp. 605–610, 1998.

[17] M. S. Bani and B. L. Chalmers, “Best approximation in L^∞ via iterative Hilbert space procedures,” *Journal of Approximation Theory*, vol. 42, pp. 173–180, 1984.

[18] C. S. Burrus, J. A. Barreto and I. W. Selesnick, “Iterative reweighted least-squares design of FIR filters,” *IEEE Transactions on Signal Processing*, vol. 42, pp. 2926–2936, 1994.

[19] T. Söderstöm and P. Stoica, System Identification. Prentice Hall: New York, 1989.

[20] E. Bai, L. Fu and S. S. Sastry, “Averaging analysis for discrete time and sampled data adaptive systems,” *IEEE Transactions on Circuits and Systems*, vol. 35, pp. 137–147, 1988.