A New Method of Noncausal Identification of Time-varying Systems

Maciej Niedźwiecki*, Marcin Ciołek†, Artur Gańcza‡ and Piotr Kaczmarek§
Faculty of Electronics, Telecommunications and Informatics, Department of Automatic Control
Gdańsk University of Technology, ul. Narutowicza 11/12, Gdańsk, Poland
*maciekn@eti.pg.edu.pl, †marcin.ciolek@pg.edu.pl, ‡artgancz@student.pg.edu.pl, piotr.kaczmarek@eti.pg.edu.pl§

***Abstract*—The paper shows that the problem of noncausal identification of a time-varying FIR (finite impulse response) system can be reformulated, and solved, as a problem of smoothing of the preestimated parameter trajectories. Characteristics of the smoothing filter should be chosen so as to provide the best trade-off between the bias and variance of the resulting estimates. It is shown that optimization of the smoothing operation can be performed adaptively using the parallel estimation technique.**

I. Introduction

The paper deals with the problem of noncausal identification of a nonstationary FIR system. Noncausality means that the estimates of the time-varying system parameters are functions of both past and "future" (prerecorded) input/output measurements. Models obtained in this way cannot be incorporated in real-time applications (such as prediction or control) but may prove useful in almost real-time ones (such as channel equalization [1]), where the model-based decisions can be made with some delay, and in off-line applications (such as channel simulation).

When system parameters vary sufficiently slowly, they can be tracked using the time-localized versions of the classical estimation algorithms, such as least squares or maximum likelihood [2]– [4]. The more advanced solutions, capable of tracking fast parameter changes, are based on explicit models (hypermodels) of parameter time-variation, either deterministic or stochastic. In the first case system parameters are modeled as linear combinations of known functions of time, called basis functions (BF), and their estimation can be carried out using standard methods such as weighted least squares [5]–[20]. In the second case, the adopted model is stochastic, e.g. parameter changes are modeled as a first-order or higher-order (integrated) random walk process. The problem of parameter estimation can be then stated as a problem of filtering/smoothing in an appropriately defined state space, and its solution can be obtained using the algorithms known as Kalman filters/smoothers [21]– [27].

The approach proposed in this paper is different from all methods mentioned above. It is based on the concept of preestimation. Preestimates are rough estimates of system parameters – unbiased but very noisy. They can be obtained by inverse filtering of exponentially weighted least squares estimates. To get reliable identification results, the preestimated parameter trajectories must be postfiltered – this allows one to trade-off the bias and variance components of the mean squared parameter tracking error. It will be shown that characteristics of such postfilters can be adaptively adjusted to the unknown, and possibly time-varying, functional form and speed of parameter variation. The proposed solution is based on parallel estimation and cross-validation.

This work was partially supported by the National Science Center under the agreement UMO-2018/29/B/ST7/00325. Calculations were carried out at the Academic Computer Centre in Gdańsk.

II. System

Consider the problem of identification, based on the available input-output data, of a nonstationary stochastic process governed by

$$y(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}(t) + e(t) \tag{1}$$

where $t = \ldots, -1, 0, 1, \ldots$ denotes discrete (normalized) time, $\boldsymbol{\varphi}(t) = [u(t-1), \ldots, u(t-n)]^{\mathrm{T}}$ denotes the regression vector made up of the previous samples of the observable input signal $u(t)$, $\boldsymbol{\theta}(t) = [\theta_1(t), \ldots, \theta_n(t)]^{\mathrm{T}}$ is the vector of unknown time-varying process parameters, and $e(t)$ denotes white measurement noise.

Linear time-varying FIR models are used, among others, to describe rapidly fading mobile communication channels. The FIR structure describes well the so-called multi-path effect: due to scattering the transmitted signal reaches the receiver along different paths, i.e., with different time delays; the values of FIR coefficients depend on the strength of "natural reflectors" and their time variation is caused by the receiver motion [1]. Noncausal identification of time-varying parameters of (1), i.e., their estimation based on the prerecorded data set $\Omega(N) = \{y(1), \boldsymbol{\varphi}(1), \ldots, y(N), \boldsymbol{\varphi}(N)\}$ can be used eg. for channel simulation purposes.

We will assume that

(A1) $\{u(t)\}$ is a zero-mean wide sense stationary Gaussian sequence, persistently exciting of order at least n, with an exponentially decaying autocorrelation function.

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

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

III. Preliminaries

The preestimation technique proposed in [14] was based on inverse filtering of causal (forward-time) short memory exponentially weighted least squares (EWLS) estimates. We will show that the estimation results can be further improved if the preestimated parameter trajectories are obtained by combining the forward-time preestimates, based on the estimates yielded by the causal EWLS algorithm, with the analogous backward-time preestimates, obtained by means of inverse filtering of the estimates provided by the anticausal EWLS algorithm. The forward-time EWLS estimator is given by

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_-(t) &= \arg\min_{\boldsymbol{\theta}} \sum_{i=0}^{t-1} \lambda^i [y(t-i) - \boldsymbol{\varphi}^{\mathrm{T}}(t-i)\boldsymbol{\theta}]^2 \\ &= \mathbf{R}_-^{-1}(t)\mathbf{r}_-(t)\end{aligned} \tag{2}$$

where λ, $0 < \lambda < 1$, denotes the so-called forgetting constant, $\mathbf{R}_-(t) = \sum_{i=0}^{t-1} \lambda^i \boldsymbol{\varphi}(t-i)\boldsymbol{\varphi}^{\mathrm{T}}(t-i)$ and $\mathbf{r}_-(t) = \sum_{i=0}^{t-1} \lambda^i y(t-i)\boldsymbol{\varphi}(t-i)$. The effective memory of this estimator can be obtained from

$$L_-(t) = \sum_{i=0}^{t-1} \lambda^i = \lambda L_-(t-1) + 1 \tag{3}$$

The analogous expressions for the backward-time (anticausal) estimator are

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_+(t) &= \arg\min_{\boldsymbol{\theta}} \sum_{i=0}^{N-t} \lambda^i [y(t+i) - \boldsymbol{\varphi}^{\mathrm{T}}(t+i)\boldsymbol{\theta}]^2 \\ &= \mathbf{R}_+^{-1}(t)\mathbf{r}_+(t)\end{aligned} \tag{4}$$

where $\mathbf{R}_+(t) = \sum_{i=0}^{N-t} \lambda^i \boldsymbol{\varphi}(t+i)\boldsymbol{\varphi}^{\mathrm{T}}(t+i)$, $\mathbf{r}_+(t) = \sum_{i=0}^{N-t} \lambda^i y(t+i)\boldsymbol{\varphi}(t+i)$ and

$$L_+(t) = \sum_{i=0}^{N-t} \lambda^i = \lambda L_+(t+1) + 1. \tag{5}$$

Both estimates can be computed recursively using the formula

$$\widehat{\boldsymbol{\theta}}_\pm(t) = \widehat{\boldsymbol{\theta}}_\pm(t \pm 1) + \mathbf{R}_\pm^{-1}(t)\boldsymbol{\varphi}(t)\varepsilon_\pm(t) \tag{6}$$

where $\varepsilon_\pm(t) = y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widehat{\boldsymbol{\theta}}_\pm(t \pm 1)$ denotes the forward/backward output prediction error. Finally, we note that inverses of both regression matrices $\mathbf{R}_-(t)$ and $\mathbf{R}_+(t)$ can be computed recursively [3].

IV. Unidirectional preestimates and their properties

The forward/backward preestimates will be defined in the following form

$$\begin{aligned}\boldsymbol{\theta}^*_\pm(t) &= L_\pm(t)\widehat{\boldsymbol{\theta}}_\pm(t) - \lambda L_\pm(t \pm 1)\widehat{\boldsymbol{\theta}}_\pm(t \pm 1) \\ &= [L_\pm(t) - \lambda L_\pm(t \pm 1)]\widehat{\boldsymbol{\theta}}_\pm(t \pm 1) + \widehat{\boldsymbol{\Phi}}_\pm^{-1}(t)\boldsymbol{\varphi}(t)\varepsilon_\pm(t)\end{aligned} \tag{7}$$

where the quantity

$$\widehat{\boldsymbol{\Phi}}_\pm(t) = \frac{\mathbf{R}_\pm(t)}{L_\pm(t)} \tag{8}$$

can be recognized as local exponentially weighted estimate of the covariance matrix $\boldsymbol{\Phi} = \mathrm{E}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)] > 0$.

It can be shown that, under assumption (A1), it holds that $\lim_{\lambda \to 1} \widehat{\boldsymbol{\Phi}}_\pm(t) = \boldsymbol{\Phi}$ and $\lim_{\lambda \to 1} \widehat{\boldsymbol{\Phi}}_\pm^{-1}(t) = \boldsymbol{\Phi}^{-1}$, where convergence takes place in the mean squared sense [2]. This justifies the following approximation

$$\widehat{\boldsymbol{\Phi}}_\pm^{-1}(t) \cong \boldsymbol{\Phi}^{-1} \tag{9}$$

which will be further used to study parameter tracking properties of preestimates. Applying (9) and noting that $L_\pm(t) - \lambda L_\pm(t \pm 1) = 1$, one can rewrite (7) in the form

$$\begin{aligned}&\boldsymbol{\theta}^*_\pm(t) \cong \boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}(t) \\ &+[\mathbf{I}_n - \boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)]\widehat{\boldsymbol{\theta}}_\pm(t \pm 1) + \boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}(t)e(t).\end{aligned} \tag{10}$$

Since $\mathrm{E}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)] = \boldsymbol{\Phi}$, $\mathrm{E}[\boldsymbol{\varphi}(t)e(t)] = \mathbf{0}$ and $\boldsymbol{\varphi}(t)$ is asymptotically independent of $\widehat{\boldsymbol{\theta}}_\pm(t \pm 1)$, for λ sufficiently close to 1 and under (A1)-(A3), one obtains

$$\mathrm{E}[\boldsymbol{\theta}^*_\pm(t)] \cong \boldsymbol{\theta}(t) \tag{11}$$

which means that the forward/backward preestimates are approximately unbiased. Furthermore, based on (10), one arrives at the following error equation

$$\begin{aligned}&\Delta\boldsymbol{\theta}^*_\pm(t) = \boldsymbol{\theta}^*_\pm(t) - \boldsymbol{\theta}(t) \\ &\cong [\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t) - \mathbf{I}_n][\boldsymbol{\theta}(t) - \widehat{\boldsymbol{\theta}}_\pm(t \pm 1)] + \boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}(t)e(t)\end{aligned} \tag{12}$$

which will be used for comparative purposes.

V. Bidirectional preestimates

According to (12), the variability of forward/backward preestimates depends on estimation errors $\boldsymbol{\theta}(t) - \widehat{\boldsymbol{\theta}}_\pm(t \pm 1)$ yielded by the corresponding EWLS algorithms. Assesment of the local estimation capabilities of both algorithms can be based on comparison of their predictive abilities. As a local estimate of the variance of the forward/backward prediction error, one can use the quantity

$$\mathcal{P}_\pm(t) = \frac{\mathcal{E}_\pm(t)}{L^0_\pm(t)} \tag{13}$$

$$\mathcal{E}_-(t) = \sum_{i=0}^{t-1} \lambda_0^i \varepsilon_-^2(t-i) = \lambda_0 \mathcal{E}_-(t-1) + \varepsilon_-^2(t)$$

$$\mathcal{E}_+(t) = \sum_{i=0}^{N-t} \lambda_0^i \varepsilon_+^2(t+i) = \lambda_0 \mathcal{E}_+(t+1) + \varepsilon_+^2(t)$$

where $\lambda_0 \le \lambda$ is another forgetting constant and $L^0_\pm(t)$ is defined in the analogous way as $L_\pm(t)$: $L^0_\pm(t) = L_\pm(t)|_{\lambda=\lambda_0}$.

Denote by $d(t)$ the local decision statistic

$$d(t) = \begin{cases} 1 & \text{if} \quad \mathcal{P}_-(t) \le \mathcal{P}_+(t) \\ 0 & \text{if} \quad \mathcal{P}_-(t) > \mathcal{P}_+(t) \end{cases}. \tag{14}$$

Bidirectional preestimate $\boldsymbol{\theta}^*(t)$ can be defined as follows

$$\boldsymbol{\theta}^*(t) = d(t)\boldsymbol{\theta}^*_-(t) + [1 - d(t)]\boldsymbol{\theta}^*_+(t). \tag{15}$$

Similar to unidirectional preestimates, the bidirectional preestimate is approximately unbiased: $\mathrm{E}[\boldsymbol{\theta}^*(t)] \cong \boldsymbol{\theta}(t)$. The associated error equation has the form

$$\begin{aligned}\Delta\boldsymbol{\theta}^*(t) &= \boldsymbol{\theta}^*(t) - \boldsymbol{\theta}(t)\\ &\cong d(t)[\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t) - \mathbf{I}_n][\boldsymbol{\theta}(t) - \widehat{\boldsymbol{\theta}}_-(t-1)]\\ &+[1-d(t)][\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t) - \mathbf{I}_n][\boldsymbol{\theta}(t) - \widehat{\boldsymbol{\theta}}_+(t+1)]\\ &+ \boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}(t)e(t).\end{aligned} \tag{16}$$

Fig. 1 shows the preestimated parameter trajectories obtained for a nonstationary two-tap FIR system governed by

$$y(t) = \theta_1(t)u(t-1) + \theta_2(t)u(t-2) + e(t) \tag{17}$$

excited by a zero-mean stationary autoregressive Gaussian process with autocorrelation function $\mathrm{E}[u(t)u(t-i)] = (0.8)^i$, and corrupted by white Gaussian noise with variance $\sigma_e^2 = 0.0025$ (SNR=25 dB). Parameter $\theta_1(t)$ was changing in a chirp-like way, and parameter $\theta_2(t)$ was piecewise constant – see Fig. 1. The forgetting constants were set to $\lambda = \lambda_0 = 0.9$.

As expected, the bidirectional preestimates have smaller variability than the one-sided ones. The improvement is particularly evident in the vicinity of parameter jumps. Note that the forward parameter estimates $\widehat{\boldsymbol{\theta}}_-(t)$ are more accurate than backward estimates $\widehat{\boldsymbol{\theta}}_+(t)$ just before the jump, which results in $\mathcal{P}_-(t) < \mathcal{P}_+(t)$, and the converse is true just after the jump. As a result, the combined estimate (15) is free of the jump-related artifacts.

VI. Competitive smoothing

Even though the preestimates $\boldsymbol{\theta}^*(t)$ are approximately unbiased, they have a very large variability and hence they should be smoothed to become practically useful. In this paper we will use for this purpose a parallel filtering scheme combining FIR smoothers of the form

$$\widehat{\boldsymbol{\theta}}_{\alpha|k}(t) = \sum_{i=-k}^{k} h_{\alpha|k}(i)\boldsymbol{\theta}^*(t+i) \tag{18}$$

where $h_{\alpha|k}(i)$ denotes impulse response of a lowpass filter obeying $\sum_{i=-k}^{k} h_{\alpha|k}(i) = 1$, and α denotes parameter, or a set of parameters, used to shape characteristics of the filter. Consider a number of such algorithms running simultaneously, each one equipped with different filter settings: $\alpha \in \mathcal{A} = \{\alpha_1, \ldots, \alpha_M\}$ and $k \in \mathcal{K} = \{k_1, \ldots, k_K\}$. At each time instant the estimated parameter vector will have the form $\widehat{\boldsymbol{\theta}}_{\widehat{\alpha}(t)|\widehat{k}(t)}(t)$ where

$$\{\widehat{\alpha}(t), \widehat{k}(t)\} = \arg\min_{\substack{\alpha\in\mathcal{A}\\ k\in\mathcal{K}}} J_{\alpha|k}(t) \tag{19}$$

and $J_{\alpha|k}(t)$ denotes the local decision statistic.

The proposed selection criterion is based on the cross-validation approach. Denote by $\widehat{\boldsymbol{\theta}}^{\circ}_{\alpha|k}(t)$ the holey estimator of $\boldsymbol{\theta}(t)$, i.e., the one that eliminates from the estimation process the central measurement $y(t)$, and by

$$\varepsilon^{\circ}_{\alpha|k}(t) = y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widehat{\boldsymbol{\theta}}^{\circ}_{\alpha|k}(t) \tag{20}$$

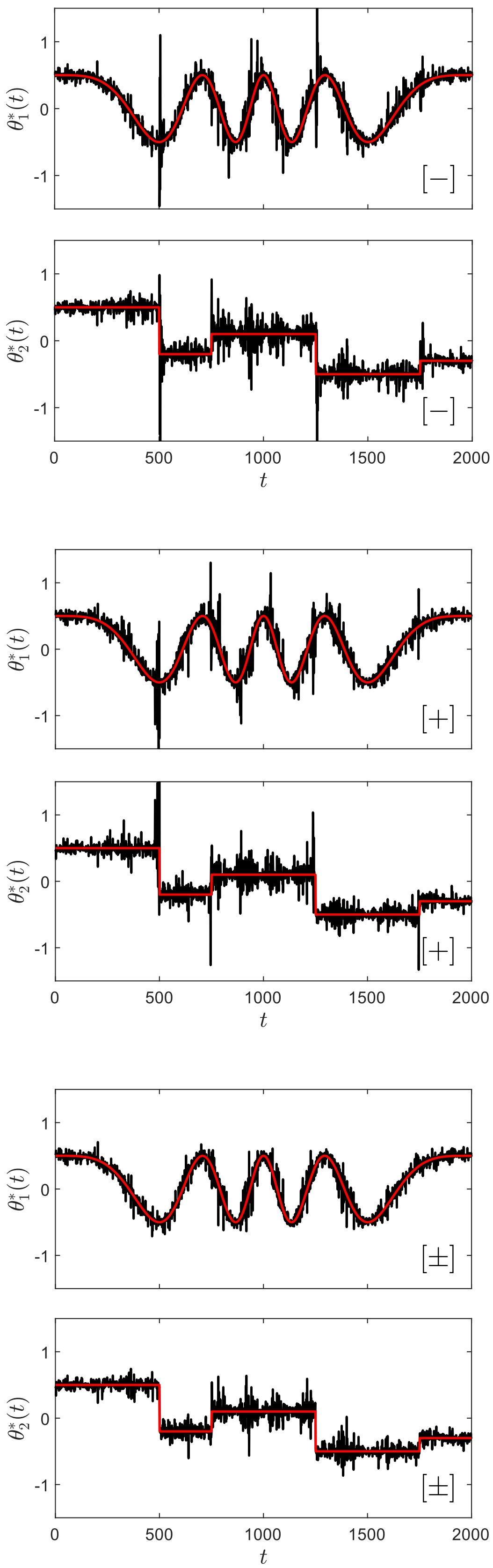

Figure 1: Preestimated parameter trajectories of a nonstationary two-tap FIR system: forward-time preestimates [−], backward-time preestimates [+], and bidirectional preestimates [±]. Preestimates (black lines) are superimposed on true parameter trajectories (red lines).

the corresponding leave-one-out interpolation error (deleted residual). For (18) the holey estimator can be built as follows

$$\widehat{\boldsymbol{\theta}}^{\circ}_{\alpha|k}(t) = \frac{\sum_{\substack{i=-k \\ i\neq 0}}^{k} h_{\alpha|k}(i)\boldsymbol{\theta}^{*}(t+i)}{1 - h_{\alpha|k}(0)} \tag{21}$$

where the normalizing factor $1 - h_{\alpha|k}(0) = \sum_{\substack{i=-k \\ i\neq 0}}^{k} h_{\alpha|k}(i)$ is applied to retain the unity DC gain of the filter. Simple calculations lead to

$$\widehat{\boldsymbol{\theta}}^{\circ}_{\alpha|k}(t) = \frac{\widehat{\boldsymbol{\theta}}_{\alpha|k}(t) - h_{\alpha|k}(0)\boldsymbol{\theta}^{*}(t)}{1 - h_{\alpha|k}(0)} \,. \tag{22}$$

When several algorithms are run in parallel, the best-local values of α and k can be chosen using the decision rule (19) with

$$J_{\alpha|k}(t) = \sum_{i=-D}^{D} [\varepsilon^{\circ}_{\alpha|k}(t+i)]^2 \tag{23}$$

where D determines the size of the local decision window. Note that while denoising of preestimates is carried out using classical *signal processing* tools, selection of the best smoothing variant relies on comparison of local system modeling errors (20), i.e., it is based on *system identification* inference.

Note also that the selection criterion described above can be easily extended to the case where different preestimates are denoised using different smoothers. Such a decentralized strategy may prove useful when system parameters vary at different rates.

VII. Computer simulations

Simulations were carried out for a two-tap FIR system (17). Discrete-time parameter trajectories were generated by sampling the continuous-time prototypes. To check the compared algorithms under different operating conditions, in addition to the medium speed parameter variation scenario, depicted in Fig. 1, the two times faster and two times slower variations were considered. In the first case the sampling rate was decreased, and in the second case – increased by a factor of 2. As a result, the length of the simulation interval T_s was equal to 1000, 2000 and 4000 for fast, medium speed and slow changes, respectively. For each speed of parameter variation, data generation was started 1000 instants prior to $t = 1$ and was continued for 1000 instants after $t = T_s$, so that, no matter what k and D, the estimation process and evaluation of its results could be, for all algorithms, started at the instant 1 and ended at the instant T_s. For $t < 1$ and $t > T_s$ system parameters were constant and equal to $\boldsymbol{\theta}(1)$ and $\boldsymbol{\theta}(T_s)$, respectively. To check behavior of the compared algorithms under different noise conditions, two average signal-to-noise ratios were considered: 25 dB ($\sigma_e^2 = 0.0025$) and 15 dB ($\sigma_e^2 = 0.025$).

The competitive FIR smoother was made up of 3 lowpass filters designed, for $k = 100$ and 3 different cutoff frequencies ω_c, using the well-known window method [28]

$$h_{\omega_c|k}(i) = \begin{cases} \mu_{\omega_c|k} v_k(i) h^{\mathrm{id}}_{\omega_c}(i) & \text{for} \quad |i| \le k \\ 0 & \text{for} \quad |i| > k \end{cases} ,$$

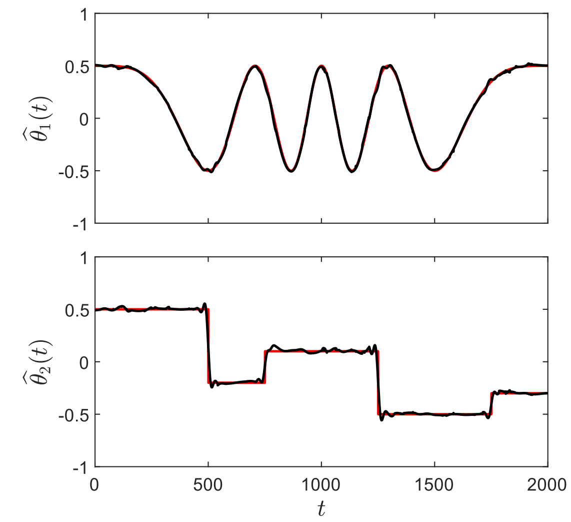

Figure 2: Estimated parameter trajectories of a two-tap FIR system. Smoothed bidirectional preestimates (black lines) are superimposed on true parameter trajectories (red lines).

where $v_k(i)$ denotes the Blackman window, $\mu_{\omega_c|k} = [\sum_{i=-k}^{k} v_k(i) h^{\mathrm{id}}_{\omega_c}(i)]^{-1}$ is the scaling coefficient, and $h^{\mathrm{id}}_{\omega_c}(i)$ denotes impulse response of the ideal lowpass filter:

$$h^{\mathrm{id}}_{\omega_c}(i) = \begin{cases} \frac{\omega_c}{\pi} & \text{for} \quad i = 0 \\ \frac{\sin \omega_c i}{\pi i} & \text{for} \quad i \neq 0 \,. \end{cases}$$

The cutoff frequencies were set to $\omega_c^1 = 0.05$, $\omega_c^2 = 0.1$ and $\omega_c^3 = 0.2$, respectively.

Table I compares results – separately for $\theta_1(t)$ and $\theta_2(t)$ – obtained using adaptive FIR smoothing of forward preestimates (S_-), backward preestimates (S_+) and bidirectional preestimates ($\mathrm{S}_\pm$), respectively, with the analogous results yielded by the state-of-the-art local basis function (LBF) approach described in [20] and the multi-resolution wavelet (MW) approach described in [13], [17]. The competitive LBF smoother was made up of 9 LBF estimators corresponding to different choices of design parameters k (25, 50, 100) and m (1, 3, 5), where m denotes the number of basis functions (powers of time). The multi-resolution wavelet scheme was used in the configuration recommended in [17] (cardinal B-splines, resolution level 3, analysis interval of length 501, overlap-add synthesis with Hann window and 50% overlap). All scores were obtained by means of combined time averaging (over the simulation interval) and ensemble averaging (over 100 independent realizations of $\{e(t)\}$ and $\{u(t)\}$). The width of the decision window was set to $2D + 1 = 61$ ($D = 30$).

According to the simulation evidence summarized in Table I, the proposed approach yields results that are better or comparable with those provided by the computationally much more involved LBF and MW approaches (note that FIR smoothing can be efficiently realized in the frequency domain using the FFT-based routine).

Table I: Comparison of tracking results obtained – for different values of SNR and different speeds of parameter variation (SoV) – using adaptive FIR smoothing of forward preestimates (S_-), backward preestimates (S_+) and bidirectional preestimates ($S_\pm$), with the analogous results yielded by the local basis function (LBF) approach and multi-resolution wavelet (MW) approach. In each case the best result is shown in boldface.

$\theta_1(t)$

SNR	Method	Fast	Medium	Slow
15 dB	S_-	1.66E-03	1.07E-03	7.31E-04
	S_+	1.63E-03	1.04E-03	7.30E-04
	$S_\pm$	1.52E-03	9.82E-04	**7.25E-04**
	LBF	1.54E-03	1.08E-03	7.71E-04
	MW	**1.29E-03**	**8.01-04**	7.35E-04
25 dB	S_-	8.61E-04	2.73E-04	1.79E-04
	S_+	8.48E-04	2.77E-04	1.91E-04
	$S_\pm$	5.88E-04	1.88E-04	**1.03E-04**
	LBF	4.09E-04	2.05E-04	1.58E-04
	MW	**3.39E-04**	**1.79E-04**	1.36E-04

$\theta_2(t)$

SNR	Method	Fast	Medium	Slow
15 dB	S_-	3.21E-03	1.89E-03	1.19E-03
	S_+	3.35E-03	1.91E-03	1.18E-03
	$S_\pm$	**3.01E-03**	**1.70E-03**	**1.09E-03**
	LBF	3.06E-03	1.86E-03	1.19E-03
	MW	3.74E-03	2.28E-03	1.47E-03
25 dB	S_-	2.50E-03	1.11E-03	6.13E-04
	S_+	2.54E-03	1.15E-03	6.07E-04
	$S_\pm$	2.08E-03	9.31E-04	**4.63E-04**
	LBF	**1.91E-03**	**9.14E-04**	5.14E-04
	MW	1.94E-03	1.14E-03	5.99E-04

VIII. Conclusion

The paper describes a new method of identification of time-varying FIR systems subject to (locally) stationary excitation. The proposed approach is based on smoothing of preestimated parameter trajectories obtained by inverse filtering of short-memory EWLS parameter estimates, and yields results that are better, or at least comparable with those provided by the more sophisticated and computationally much more demanding methods based on local basis function approximation and multi-resolution wavelet approximation. It was shown that optimization of the smoothing operation can be carried out adaptively using the parallel estimation technique and cross-validation based variant selection strategy.

References

[1] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *Int. J. Adaptive Contr. Signal Process.* vol. 10, pp. 159–176, 1996.

[2] M. Niedźwiecki, *Identification of Time-varying Processes*. New York: Wiley, 2000.

[3] T. Söderström and P. Stoica, *System Identification*. Englewoods Cliffs NJ: Prentice-Hall, 1988.

[4] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, 2003.

[5] T. Subba Rao, "The fitting of nonstationary time-series models with time-dependent parameters," *J. R. Statist. Soc. B*, vol. 32, pp. 312–322, 1970.

[6] J. M. Mendel, *Discrete Techniques of Parameter Estimation: The Equation Error Formulation*, New York: Marcel Dekker, 1973.

[7] J. M. Liporace, "Linear estimation of nonstationary signals," *J. Acoust. Soc. Amer.*, vol. 58, pp. 1288–1295, 1975.

[8] M. Hall, A. V. Oppenheim, and A. Willsky, "Time-varying parametric modeling of speech," *Signal Processing*, vol. 5, pp. 267–285, 1983.

[9] R. Charbonnier, M. Barlaud, G. Alengrin, and J. Menez, "Results on AR-modelling of non-stationary signals," *Signal Processing*, vol. 12, pp. 143–151, 1987.

[10] M. Niedźwiecki, "Functional series modeling approach to identification of nonstationary stochastic systems," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 955–961, 1988.

[11] M. Niedźwiecki, "Recursive functional series modeling estimators for identification of time-varying plants – more bad news than good?," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 610–616, 1990.

[12] R. B. Mrad, S. D. Fassois, and J. A. Levitt, "A polynomial-algebraic method for non-stationary TARMA signal analysis - Part I: The method," *Signal Processing*, vol. 65, pp. 1–19, 1998.

[13] H.L. Wei, J. Liu and S.A. Billings, "Identification of time-varying systems using multi-resolution wavelet models," *International Journal of Systems Science*, vol. 33, pp. 1217–1228, 2002.

[14] M. Niedźwiecki and T. Kłaput, "Fast recursive basis function estimators for identification of nonstationary systems," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1925–1934, 2002.

[15] R. B. Zou, H. Wang H., and K. H. Chon, "A robust time-varying identification algorithm using basis functions," *Annals of Biomedical Engineering*, vol. 31, pp. 840–853, 2003.

[16] A. G. Poulimenos and S. D. Fassois, "Parametric time-domain methods for non-stationary random vibration modelling and analysis – a critical survey and comparison," *Mechanical Systems and Signal Processing*, vol. 20, pp. 763–816, 2006.

[17] Y. Li, H.L., Wei, S.A. Billings, and P. Sarrigiannis, "Time-varying model identification for time-frequency feature extraction from EEG data," *Journal of Neuroscience Methods*, vol. 196, pp. 151-8, 2011.

[18] S.C. Chan and Z.G. Zhang, "Local Polynomial Modeling and Variable Bandwidth Selection for Time-Varying Linear Systems", *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 3, pp. 1102-1117, 2011.

[19] M. Niedźwiecki and S. Gackowski, "New approach to noncausal identification of nonstationary FIR systems subject to both smooth and abrupt parameter changes," *IEEE Trans. Automat. Contr.*, vol. 58, pp. 1847–1853, 2013.

[20] M. Niedźwiecki and M. Ciołek, "Generalized Savitzky-Golay filters for identification of nonstationary systems," *Automatica*, vol. 108, no. 108522, 2019.

[21] D.Q. Mayne, "Optimal nonstationary estimation of the parameters of a linear system with Gaussian inputs," Journal of Electronics and Control, vol. 14, pp. 101–112, 1963.

[22] J. P. Norton, "Optimal smoothing in the identification of linear time-varying systems," *Proc. IEE*, vol. 122, pp. 663–668, 1975.

[23] T. Bohlin, "Four cases of identification of changing systems," In P.K. Mehra and D.G. Lainiotis (eds.), System Identification: Advances and Case Studies, Academic Press, New York, pp. 441–518, 1976.

[24] H. Akaike, "Likelihood and the Bayes procedure," in *Bayesian Statistics*, J.M. Bernado, M.H. DeGroot, D.V. Lindley & A.F.M. Smith, Eds. Valencia, Spain, University Press, pp. 141–166, 1980.

[25] G. Kitagawa and W. Gersch, "A smoothness priors time-varying AR coefficient modeling of nonstationary covariance time series," *IEEE Trans. Autom. Control*, vol. 30, pp. 48–56, 1985.

[26] M. Arnold, W.H.R. Miltner, H. Witte, R. Bauer, and C. Braun, "Adaptive AR modeling of nonstationary time series by means of Kalman filtering," *IEEE Transactions on Biomedical Engineering*, vol. 45, pp. 553–562 , 1998

[27] M. Niedźwiecki, "Locally adaptive cooperative Kalman smoothing and its application to identification of nonstationary stochastic systems," *IEEE Trans. Signal Process.*, vol. 60, pp. 48–59, 2012.

[28] S. J. Orfanidis, *Introduction to Signal Processing*, New York: Prentice-Hall, 1996.