Accepted manuscript of: M. Niedźwiecki and A. Gańcza, "Fast Algorithms for Identification of Time-Varying Systems with Both Smooth and Discontinuous Parameter Changes," 2023 62nd IEEE Conference on Decision and Control (CDC), Singapore, Singapore, 2023, pp. 3519-3526, doi: 10.1109/CDC49753.2023.10384258.

Fast algorithms for identification of time-varying systems with both smooth and discontinuous parameter changes

Maciej Niedźwiecki and Artur Gańcza

***Abstract*— The problem of noncausal identification of a time-varying linear system subject to both smooth and occasional jump-type changes is considered and solved using the preestimation technique combined with the basis function approach to modeling the variability of system parameters. The proposed estimation algorithms yield very good parameter tracking results and are computationally attractive.**

I. Introduction

The problem of noncausal identification of a time-varying finite impulse response (FIR) system, subject to both smooth parameter changes and occasional parameter jumps, will be considered and solved using a new identification paradigm based on the concept of preestimation.

An example of application enabling the use of the proposed approach is adaptive self-interference (SI) cancellation in full-duplex (FD) underwater acoustic (UWA) communication systems [1] – [4]. In this case the transmit and receive antennas operate simultaneously in the same frequency bandwidth which allows to increase the channel throughput. As a consequence, the far-end signal is strongly contaminated by the self-interference introduced by the near-end transmitter - the effect caused by multiple reflections of the emitted signal from the sea surface, the bottom and surrounding scattering objects. Channel coefficients (coefficients of its impulse response) change smoothly over time, due to the Doppler effect caused by the transmitter/receiver motion, but may be also subject to occasional jumps caused by a sudden appearence or disappearence of scatterers (fish, vessel etc.) or by a sharp change in weather conditions. Channel identification is needed to secure reliable communication as it allows one to eliminate, or at least significantly reduce self-interference. Moreover, when the entire data packets are transmitted/received/decoded the noncausal estimation techniques developed in this paper, which operate on the prerecorded input/output data, are admissible and allow one to achieve better tracking results compared to conventional causal algorithms.

Most of the statistical literature on identification of dynamic systems with jump-type changes is devoted to linear Markovian switching systems [6] – [11]. In this case system parameters are assumed to switch among a finite set of unknown but constant values. The switchings are modeled by a finite state ergodic Markov chain, and parameter estimation is carried out using the maximum likelihood approach or Bayesian reasoning. Apart from the fact that the resulting estimation algorithms are computationally very intense (especially when smoothing is involved), such a problem formulation does not meet our current needs, as we are interested in estimation/tracking of parameters that are subject to both smooth continuous and jump-type variation.

As shown in [12], [13], when parameter jumps occur infrequently, the solution to the identification problem mentioned above can be obtained by combining in an appropriate way the results yielded by the causal (forward-time), anticausal (backward-time) and noncausal (bidirectional) parameter tracking algorithms based on stochastic [12] or deterministic [13] models (hypermodels) of system parameter changes. The current contribution is another step in this direction. Unlike [12] and [13], where system parameters are estimated "directly", the approach developed in this paper is based on a two-step identification procedure described in [14]. In the first step, system parameters are preestimated. Since preestimates are "raw" parameter estimates, approximately unbiased but very "noisy" (with a large variability), they must be further processed (postfiltered) - this constitutes the second step of the identification procedure.

The contribution of the paper is twofold. First, we propose a new (improved) preestimation scheme, capable of coping adequately with parameter jumps. Second, we design a new postfiltering algorithm capable, at a very low computational cost (linearly proportional to the number of estimated parameters), of accurately reproducing both smooth and jump-type parameter changes.

II. Problem statement

Many nonstationary systems, including telecommunication channels [15], [16], can be well approximated by a time-varying finite impulse response model of the form

$$\begin{aligned} y(t) &= \sum_{i=1}^{n} \theta_i(t)u(t-i+1) + e(t) \\ &= \boldsymbol{\theta}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t) + e(t) \end{aligned} \tag{1}$$

where $t = \ldots, -1, 0, 1, \ldots$ denotes discrete (normalized) time, $y(t)$ denotes the output signal, $\boldsymbol{\varphi}(t) = [u(t), \ldots, u(t-n+1)]^{\mathrm{T}}$ denotes regression vector made up of past samples of the input signal $u(t)$, $\boldsymbol{\theta}(t) = [\theta_1(t), \ldots, \theta_n(t)]^{\mathrm{T}}$ is the vector of time-varying system coefficients, and $\{e(t)\}$ denotes noise. Note that, unlike the conventional communication systems working in the simplex mode, in the full-duplex case the input sequence $\{u(t)\}$, which is emitted

*This work was partially supported by the National Science Center under the agreement UMO-2018/29/B/ST7/00325. Both authors are with the Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics, Department of Automatic Control, ul. Narutowicza 11/12, Gdańsk, Poland: artgancz@student.pg.edu.pl, maciekn@eti.pg.edu.pl

by the near-end transmitter, is *known*. In this case $\{e(t)\}$ is a mixture of the far-end signal and the channel noise, and the underlying goal of identification is extraction of the signal $\{e(t)\}$ from $\{y(t)\}$. This can be easily done provided that channel parameters are known. The sequence $\{\theta_i(t), i = 1, \ldots, n\}$ can be interpreted as a time-varying impulse response of the system (1).

We will assume that

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

These assumptions are met in typical communication systems. Furthermore, we will assume that the entire segment of the input/output data $\Omega(N) = \{u(t), y(t), t = 1, \ldots, N\}$, of length N, is available and can be used to estimate the time-varying trajectory of system parameters $\{\boldsymbol{\theta}(t), t = 1, \ldots, N\}$. In the UWA FD case, under typical operating conditions, the length N of the transmitted data packet is within from several hundred to several thousand samples/symbols.

III. Preestimation technique

A. Forward/backward preestimates

1) Forward preestimates: Forward-time preestimates can be obtained by "inverse filtering" the estimates yielded by the short-memory forward-time exponentially weighted least squares (EWLS) algorithm

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_-(t) &= \arg\min_{\boldsymbol{\theta}} \sum_{j=1}^{t} \lambda^{t-j}[y(j) - \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(j)]^2 \\ &= \left[\sum_{j=1}^{t} \lambda^{t-j}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\right]^{-1} \left[\sum_{j=1}^{t} \lambda^{t-j}\boldsymbol{\varphi}(j)y(j)\right].\end{aligned} \tag{2}$$

The effective width $M_-(t)$ of the exponential window is given by $M_-(t) = \sum_{j=1}^{t} \lambda^{t-j}$.

The inverse filtering formula, derived and analyzed in [14], which can be used to obtain forward-time preestimates, has the form

$$\boldsymbol{\theta}^*_-(t) = M_-(t)\widehat{\boldsymbol{\theta}}_-(t) - \lambda M_-(t-1)\widehat{\boldsymbol{\theta}}_-(t-1). \tag{3}$$

The preestimates $\boldsymbol{\theta}^*_-(t)$ are approximately unbiased (no matter how true parameters change). The term "preestimator" is used because the estimates (3) have a very large variance. Hence, to obtain reliable identification results, preestimates must be further processed ("denoised") by means of postfiltering.

The forgetting constant λ should be "as small as possible" to guarantee that fast parameter changes will be tracked successfully. On the other hand, λ shouldn't be "too small" to guarantee that the number of system parameters is not greater than the steady-state equivalent number of observations $N_\infty = (1+\lambda)/(1-\lambda) \cong 2/(1-\lambda)$ (different from the effective number of observations [17]) used for their estimation - otherwise the estimation results would be questionable from the statistical viewpoint. This leads to the following recommendation

$$\lambda = \max\left\{0.9, 1 - \frac{2}{n}\right\}. \tag{4}$$

2) Backward preestimates: When causality of the estimation scheme is not required, i.e., one has access to both "past" and "future" (with respect to the current time instant t) input/output data, parameter preestimates can be equally well obtained by processing the estimates yielded by the backward-time EWLS algorithm

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_+(t) &= \arg\min_{\boldsymbol{\theta}} \sum_{j=t}^{N} \lambda^{j-t}[y(j) - \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(j)]^2 \\ &= \left[\sum_{j=t}^{N} \lambda^{j-t}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\right]^{-1} \left[\sum_{j=t}^{N} \lambda^{j-t}\boldsymbol{\varphi}(j)y(j)\right].\end{aligned} \tag{5}$$

The effective width $M_+(t)$ of the corresponding exponential window is given by $M_+(t) = \sum_{j=t}^{N} \lambda^{j-t}$.

The backward-time preestimates can be defined in an analogous way to (3)

$$\boldsymbol{\theta}^*_+(t) = M_+(t)\widehat{\boldsymbol{\theta}}_+(t) - \lambda M_+(t+1)\widehat{\boldsymbol{\theta}}_+(t+1). \tag{6}$$

B. Bidirectional preestimates

When local parameter variation is smooth, the forward and backward EWLS estimates can be combined yielding the following estimation formula

$$\widehat{\boldsymbol{\theta}}_\pm(t) = \frac{M_-(t)\widehat{\boldsymbol{\theta}}_-(t) + M_+(t)\widehat{\boldsymbol{\theta}}_+(t)}{M_-(t) + M_+(t)}. \tag{7}$$

The corresponding bidirectional preestimates can be defined in the form

$$\boldsymbol{\theta}^*_\pm(t) = \frac{M_-(t)\boldsymbol{\theta}^*_-(t) + M_+(t)\boldsymbol{\theta}^*_+(t)}{M_-(t) + M_+(t)}. \tag{8}$$

The combined estimates $\widehat{\boldsymbol{\theta}}_\pm(t)$, after a slight modification, will be further used to locally evaluate the "quality" of bidirectional preestimates.

IV. Preestimation revisited

In this section a new preestimation scheme will be proposed, which combines unidirectional and bidirectional preestimates in a way that allows one to benefit from the advantages of both approaches while avoiding their weaknesses.

When system parameters change in a discontinuous way, both unidirectional and bidirectional preestimation schemes, which are in fact based on highpass filtering of EWLS estimates, are prone to generate impulsive disturbances around the points where the jumps occur: just after the jump in the case of forward preestimates, just before the jump in the case of backward preestimates, and both before and after the

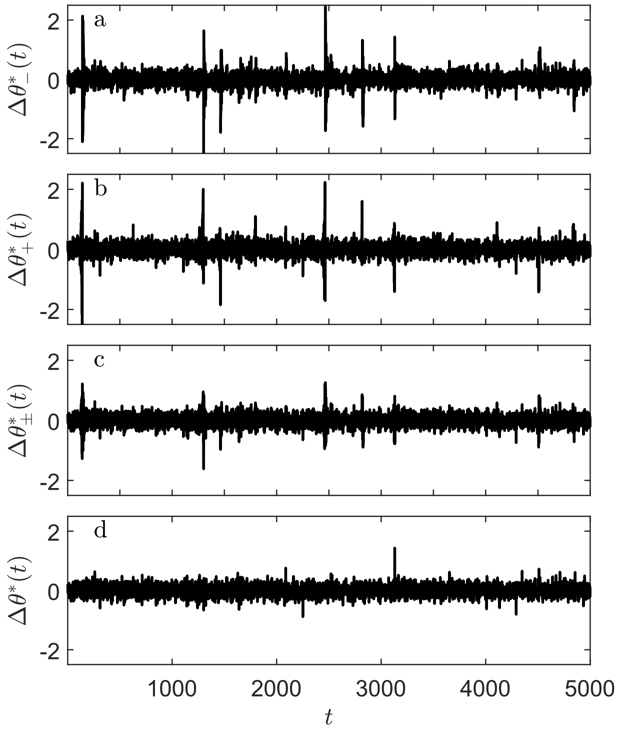

Fig. 1. Preestimation errors for different types of preestimates: forward (a), backward (b), bidirectional (c) and competitive (d). All errors were evaluated for $\theta_1(t)$.

jump in the case of bidirectional preestimates. The effects mentioned above are illustrated in Figs. 1a-1c, which show preestimation errors obtained for the second-order ($n = 2$) FIR system governed by

$$y(t) = \theta_1(t)u(t) + \theta_2(t)u(t-1) + e(t) \qquad (9)$$

where $\{\theta_1(t)\}$ is "piecewise-lowpass" and $\{\theta_2(t)\}$ is piecewise-constant – both parameter trajectories, depicted in Fig. 2, have the same ℓ_2 norm. The applied input signal was pseudo-random binary ($u(t) = \pm 1$) and the variance of the white Gaussian noise $\{e(t)\}$ was set to $\sigma_e^2 = 0.028$, which corresponds to the average signal-to-noise ratio SNR=20 dB. The jump-related artifacts are easy to spot.

A. Competitive preestimates

The competitive preestimation will be based on checking, at each time instant t, which estimation algorithm provides locally the best description of the identified system. For forward/backward EWLS algorithms, as a local performance measures one can adopt the estimates of the variance of the corresponding one-step-ahead output prediction errors

$$\mathcal{E}_-(t) = \frac{1}{K+1}\sum_{i=0}^{K} \varepsilon_-^2(t-i)$$
$$\mathcal{E}_+(t) = \frac{1}{K+1}\sum_{i=0}^{K} \varepsilon_+^2(t+i) \qquad (10)$$

where $\varepsilon_-(t) = y(t) - \widehat{\boldsymbol{\theta}}_-^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)$, $\varepsilon_+(t) = y(t) - \widehat{\boldsymbol{\theta}}_+^{\mathrm{T}}(t+1)\boldsymbol{\varphi}(t)$ and $K+1 = 2k+1$ denotes the width of the local decision window.

In the case of the combined EWLS algorithm, the one-step ahead prediction errors can be replaced with leave-one-out output interpolation errors, leading to

$$\mathcal{E}_\pm(t) = \frac{1}{K+1}\sum_{i=-k}^{k} \varepsilon_\pm^2(t+i) \qquad (11)$$

where $\varepsilon_\pm(t) = y(t) - [\widehat{\boldsymbol{\theta}}_\pm^\circ(t)]^{\mathrm{T}}\boldsymbol{\varphi}(t)$ and $\widehat{\boldsymbol{\theta}}_\pm^\circ(t)$ denotes the leave-one-out version of $\widehat{\boldsymbol{\theta}}_\pm(t)$, obtained by eliminating from the estimation process the central sample $y(t)$

$$\widehat{\boldsymbol{\theta}}_\pm^\circ(t) = \frac{M_-(t-1)\widehat{\boldsymbol{\theta}}_-(t-1) + M_+(t+1)\widehat{\boldsymbol{\theta}}_+(t+1)}{M_-(t-1) + M_+(t+1)}\,. \qquad (12)$$

Let

$$\mathcal{E}_{\min}(t) = \min\{\mathcal{E}_-(t), \mathcal{E}_\pm(t), \mathcal{E}_+(t)\}.$$

The competitive (winner-takes-all) preestimates can be defined as follows[1]

$$\boldsymbol{\theta}^*(t) = \begin{cases} \boldsymbol{\theta}_-^*(t) & \text{if} \quad \mathcal{E}_-(t) = \mathcal{E}_{\min}(t) \\ \boldsymbol{\theta}_\pm^*(t) & \text{if} \quad \mathcal{E}_\pm(t) = \mathcal{E}_{\min}(t) \\ \boldsymbol{\theta}_+^*(t) & \text{if} \quad \mathcal{E}_+(t) = \mathcal{E}_{\min}(t) \end{cases}. \qquad (13)$$

The competitive preestimates obtained for the system (9) are shown in Fig.1d. Note that these preestimates are almost free of jump-related artifacts typical of their forward, backward and bidirectional counterparts.

Remark 1

The value of the forgetting constant λ sets the lower bound on the distance between subsequent parameter jumps $T_{\min}$ guaranteeing sharp reproduction of jump changes. It is known that the EWLS algorithm needs approximately $N_\infty \cong 2M_\infty$ time steps to "forget" completely about the parameter step change, i.e., to reduce to (almost) zero the step-invoked transient bias error [18]. Hence, one can set $T_{\min} = N_\infty$. If this condition is met, the forward/backward EWLS algorithms manage to fully recover from the parameter step change before the next one occurs.

B. Collaborative preestimates

According to [18], instead of the "competitive" estimation formula (13), one can use the following Bayesian "collaborative" rule

$$\boldsymbol{\theta}^*(t) = \mu_-(t)\boldsymbol{\theta}_-^*(t) + \mu_\pm(t)\boldsymbol{\theta}_\pm^*(t) + \mu_+(t)\boldsymbol{\theta}_+^*(t) \qquad (14)$$

where $\mu_-(t)$, $\mu_\pm(t)$ and $\mu_+(t)$, obeying $\mu_-(t) + \mu_\pm(t) + \mu_+(t) = 1$, denote the so-called model credibility coefficients (related to posterior probabilities of different parameter "patterns" [19]), which can be obtained from

$$\mu_\star(t) \propto \left[\frac{\mathcal{E}_{\min}(t)}{\mathcal{E}_\star(t)}\right]^{\frac{K+1}{2}}, \quad \star \in \{-, \pm, +\} \qquad (15)$$

where $\propto$ denotes proportionality.

[1] If the score $\mathcal{E}_{\min}(t)$ is attained by more than one algorithm (which is extremely unlikely), any of them can be chosen.

V. Postfiltering

Denote by $\{\theta_j^*(t), t = 1, \dots, N\}$ the preestimated trajectory of the j-th system parameter $\theta_j(t)$. Since the preestimated trajectory can be regarded as a true trajectory contaminated with a zero-mean noise of large variance, to obtain statistically meaningful estimation results, preestimates have to be further processed (denoised). We will show that if the identified FIR system is subject to both smooth and abrupt parameter changes, excellent results can be obtained using a new variant of the local basis function (LBF) approach developed in [14].

Denote by $\{f_1(i), \dots, f_m(i)\}$ the set of m discrete-time basis functions (BF), linearly independent on $[-\infty, \infty]$, and by λ_0, $0 < \lambda_0 < 1$, the forgetting constant (different from λ) which will be used for estimation localization purposes. For purely computational reasons, in the sequel we will use powers of time as basis functions, namely

$$f_l(i) = i^{l-1}, \quad l = 1, \dots, m \tag{16}$$

Note that the basis functions (16) can be computed recursively using

$$\mathbf{f}(i+1) = \mathbf{A}\mathbf{f}(i) \tag{17}$$

where $\mathbf{f}(i) = [f_1(i), \dots, f_m(i)]^{\mathrm{T}}$ and $[\mathbf{A}]_{ij} = \{\binom{i-1}{j-1}$ if $i \geq j, 0$ otherwise$\}$. The LBF approach is based on the assumption that parameter changes can be locally approximated by a linear combination of basis functions. To be able to sharply reproduce parameter jumps, whenever they occur, we will use a solution similar to that applied in the case of preestimation, namely, we will combine results yielded by causal, anticausal and noncausal LBF algorithms.

A. Causal LBF estimation

To derive the causal (forward-time) exponentially weighted basis function (EWBF) estimate of $\theta_j(t)$, we will adopt the following backward-time model of parameter evolution

$$\theta_j(t-i) = \sum_{l=1}^{m} a_{jl} f_l(i) = \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j, i = 0, \dots, t-1 \tag{18}$$

where $\boldsymbol{\alpha}_j = [a_{j1}, \dots, a_{jm}]^{\mathrm{T}}$.

In agreement with the LBF paradigm, we will regard this model as trustworthy only, or predominatly, in the "recent past" (relative to t). The simplest way of making estimation results dependent mainly on the recently observed data is by means of exponential forgetting. This leads to the following EWBF estimate of $\theta_j(t)$

$$\begin{aligned}
\widetilde{\boldsymbol{\alpha}}_j^-(t) &= \arg\min_{\boldsymbol{\alpha}} \sum_{i=0}^{t-1} \lambda_0^i [\theta_j^*(t-i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}]^2 \\
&= [\mathbf{V}^-(t)]^{-1}\mathbf{v}_j^-(t) \\
\widetilde{\theta}_j^-(t) &= \mathbf{f}^{\mathrm{T}}(0)\widetilde{\boldsymbol{\alpha}}_j^-(t)
\end{aligned} \tag{19}$$

where

$$\begin{aligned}
\mathbf{V}^-(t) &= \sum_{i=0}^{t-1} \lambda_0^i \mathbf{f}(i)\mathbf{f}^{\mathrm{T}}(i) \\
&= \mathbf{V}^-(t-1) + \lambda_0^{t-1}\mathbf{f}(t-1)\mathbf{f}^{\mathrm{T}}(t-1) \\
\mathbf{v}_j^-(t) &= \sum_{i=0}^{t-1} \lambda_0^i \theta_j^*(t-i)\mathbf{f}(i) \\
&= \lambda_0 \mathbf{A}\mathbf{v}_j^-(t-1) + \theta_j^*(t)\mathbf{f}(0) \\
t &= 1, \dots, N\ .
\end{aligned} \tag{20}$$

Denote by $\mathbf{W}^-(t)$ the inverse of the matrix $\mathbf{V}^-(t)$. It is easy to derive the formula for recursive computation of $\mathbf{W}^-(t)$. Actually, note that $\lambda_0^{t-1}\mathbf{f}(t-1)\mathbf{f}^{\mathrm{T}}(t-1) = \widetilde{\mathbf{f}}(t-1)\widetilde{\mathbf{f}}^{\mathrm{T}}(t-1)$, where $\widetilde{\mathbf{f}}(t)$ can be computed recursively using $\widetilde{\mathbf{f}}(t) = \widetilde{\mathbf{A}}\widetilde{\mathbf{f}}(t-1)$ with $\widetilde{\mathbf{A}} = \sqrt{\lambda_0}\mathbf{A}$ and $\widetilde{\mathbf{f}}(0) = \mathbf{f}(0)$. Then, using the matrix inversion lemma [20], one arrives at

$$\begin{aligned}
&\mathbf{W}^-(t) = [\mathbf{V}^-(t-1) + \widetilde{\mathbf{f}}(t-1)\widetilde{\mathbf{f}}^{\mathrm{T}}(t-1)]^{-1} \\
&= \mathbf{W}^-(t-1) - \frac{\mathbf{W}^-(t-1)\widetilde{\mathbf{f}}(t-1)\widetilde{\mathbf{f}}^{\mathrm{T}}(t-1)\mathbf{W}^-(t-1)}{1 + \widetilde{\mathbf{f}}^{\mathrm{T}}(t-1)\mathbf{W}^-(t-1)\widetilde{\mathbf{f}}(t-1)}\ .
\end{aligned} \tag{21}$$

It is easy to show that, since $\lim_{t\longrightarrow\infty} \widetilde{\mathbf{f}}(t) = 0$, it holds that the matrices $\mathbf{W}^-(t)$ and $\mathbf{V}^-(t)$ converge to their constant steady state values $\mathbf{W}^-(\infty)$ and $\mathbf{V}^-(\infty)$, respectively. Since the matrix $\mathbf{W}^-(\infty)$ can be precomputed, this leads to the following asymptotic (valid for sufficiently large values of t) matrix-inversion-free variant of (19)

$$\widetilde{\boldsymbol{\alpha}}_j^-(t) = \mathbf{W}^-(\infty)\mathbf{v}_j^-(t). \tag{22}$$

If the backward-time model of parameter trajectory (18) is replaced with a more straightforward forward-time description

$$\theta_j(i) = \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}_j,\ i = 1, \dots, t \tag{23}$$

the corresponding regression matrix, unlike $\mathbf{V}^-(t)$, indefinitely grows with time. This means that the EWBF algorithm based on (23) would need periodic resetting. It can be easily shown that the forward-time and backward-time models are equivalent in the sense that they yield the same estimates of $\theta_j(t)$ for all values of t.

B. Anticausal LBF estimation

The anticausal (backward-time) EWBF algorithm is a simple modification of the causal one

$$\begin{aligned}
\widetilde{\boldsymbol{\alpha}}_j^+(t) &= \arg\min_{\boldsymbol{\alpha}} \sum_{i=0}^{N-t} \lambda_0^i [\theta_j^*(t+i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}]^2 \\
&= [\mathbf{V}^+(t)]^{-1}\mathbf{v}_j^+(t) \\
\widetilde{\theta}_j^+(t) &= \mathbf{f}^{\mathrm{T}}(0)\widetilde{\boldsymbol{\alpha}}_j^+(t)
\end{aligned} \tag{24}$$

where

$$\begin{aligned}\mathbf{V}^{+}(t) &= \sum_{i=0}^{N-t} \lambda_0^i \mathbf{f}(i)\mathbf{f}^{\mathrm{T}}(i) \\ \mathbf{v}_j^{+}(t) &= \sum_{i=0}^{N-t} \lambda_0^i \theta_j^*(t+i)\mathbf{f}(i) \\ &= \lambda_0 \mathbf{A}\mathbf{v}_j^{+}(t+1) + \theta_j^*(t)\mathbf{f}(0) \\ t &= N, \dots, 1.\end{aligned} \tag{25}$$

Note that $\mathbf{V}^{+}(N-t+1) = \mathbf{V}^{-}(t)$, $t = N, \dots, 1$.

C. Noncausal LBF estimation

As a noncausal (bidirectional) LBF solution we will use the fast LBF (fLBF) estimator described in [14]

$$\begin{aligned}\widetilde{\boldsymbol{\alpha}}_j^{\pm}(t) &= \arg\min_{\boldsymbol{\alpha}} \sum_{i=-L}^{L} [\theta_j^*(t+i) - \mathbf{f}^{\mathrm{T}}(i)\boldsymbol{\alpha}]^2 \\ \widetilde{\theta}_j^{\pm}(t) &= \mathbf{f}^{\mathrm{T}}(0)\widetilde{\boldsymbol{\alpha}}_j^{\pm}(t)\end{aligned} \tag{26}$$

where $L = \mathrm{int}[1/(1-\lambda_0)]$. It is straightforward to show that

$$\widetilde{\theta}_j^{\pm}(t) = \sum_{i=-L}^{L} h(i)\theta_j^*(t+i) \tag{27}$$

where

$$h(i) = \mathbf{f}^{\mathrm{T}}(0) \left[\sum_{i=-L}^{L} \mathbf{f}(i)\mathbf{f}^{\mathrm{T}}(i) \right]^{-1} \mathbf{f}(i). \tag{28}$$

Since $\sum_{i=-L}^{L} h(i) = 1$, the fLBF estimator can be regarded as a result of passing the preestimates through a linear lowpass FIR filter with an impulse response $\{h(i), i \in [-L, L]\}$. Since for the basis (16) the impulse response (28) is recursively computable, the convolution (27) can be computed in a recursive way. Alternatively, for any set of basis functions, the off-line computation of (27) can be efficiently carried out using the FFT-based procedure.

D. Competitive LBF estimation

The competitive LBF estimation scheme will be designed in an analogous way as the competitive preestimation scheme. Let $\widetilde{\boldsymbol{\theta}}_{-}(t) = [\widetilde{\theta}_1^{-}(t), \dots, \widetilde{\theta}_n^{-}(t)]^{\mathrm{T}}$, $\widetilde{\boldsymbol{\theta}}_{+}(t) = [\widetilde{\theta}_1^{+}(t), \dots, \widetilde{\theta}_n^{+}(t)]^{\mathrm{T}}$, $\widetilde{\boldsymbol{\theta}}_{\pm}(t) = [\widetilde{\theta}_1^{\pm}(t), \dots, \widetilde{\theta}_n^{\pm}(t)]^{\mathrm{T}}$ and

$$\begin{aligned}\widetilde{\boldsymbol{\theta}}_{-}(t|t-1) &= [\widetilde{\theta}_1^{-}(t|t-1), \dots, \widetilde{\theta}_n^{-}(t|t-1)]^{\mathrm{T}} \\ \widetilde{\boldsymbol{\theta}}_{+}(t|t+1) &= [\widetilde{\theta}_1^{+}(t|t+1), \dots, \widetilde{\theta}_n^{+}(t|t+1)]^{\mathrm{T}} .\end{aligned}$$

where $\widetilde{\theta}_j^{-}(t|t-1) = \mathbf{f}^{\mathrm{T}}(-1)\widetilde{\boldsymbol{\alpha}}_j^{-}(t)$ and $\widetilde{\theta}_j^{+}(t|t+1) = \mathbf{f}^{\mathrm{T}}(-1)\widetilde{\boldsymbol{\alpha}}_j^{+}(t)$ denote the one-step-ahead predictions of $\theta_j(t)$ based on the information gathered prior to t, or after t, respectively. Finally, denote by $\widetilde{\boldsymbol{\theta}}_{\pm}^{\circ}(t)$ the leave-one-out version of $\widetilde{\boldsymbol{\theta}}_{\pm}(t)$:

$$\widetilde{\boldsymbol{\theta}}_{\pm}^{\circ}(t) = \frac{\sum_{\substack{i=-L \\ i\neq 0}}^{L} h(i)\boldsymbol{\theta}^*(t+i)}{\sum_{\substack{i=-L \\ i\neq 0}}^{L} h(i)} = \frac{\widetilde{\boldsymbol{\theta}}_{\pm}(t) - h(0)\boldsymbol{\theta}^*(t)}{1-h(0)}. \tag{29}$$

and by $K_0 + 1 = 2k_0 + 1$ – the width of the local decision window. The competitive estimate of $\boldsymbol{\theta}(t)$ can be obtained from

$$\widetilde{\boldsymbol{\theta}}(t) = \begin{cases} \widetilde{\boldsymbol{\theta}}_{-}(t) & \text{if} \quad \mathcal{D}_{-}(t) = \mathcal{D}_{\min}(t) \\ \widetilde{\boldsymbol{\theta}}_{\pm}(t) & \text{if} \quad \mathcal{D}_{\pm}(t) = \mathcal{D}_{\min}(t) \\ \widetilde{\boldsymbol{\theta}}_{+}(t) & \text{if} \quad \mathcal{D}_{+}(t) = \mathcal{D}_{\min}(t) \end{cases} . \tag{30}$$

where

$$\begin{aligned}\mathcal{D}_{-}(t) &= \frac{1}{K_0+1}\sum_{i=0}^{K_0} d_{-}^2(t-i) \\ \mathcal{D}_{\pm}(t) &= \frac{1}{K_0+1}\sum_{i=-k_0}^{k_0} d_{\pm}^2(t+i) \\ \mathcal{D}_{+}(t) &= \frac{1}{K_0+1}\sum_{i=0}^{K_0} d_{+}^2(t+i) \\ \mathcal{D}_{\min}(t) &= \min\{\mathcal{D}_{-}(t), \mathcal{D}_{\pm}(t), \mathcal{D}_{+}(t)\}\end{aligned} \tag{31}$$

and

$$\begin{aligned}d_{-}(t) &= y(t) - \widetilde{\boldsymbol{\theta}}_{-}^{\mathrm{T}}(t|t-1)\boldsymbol{\varphi}(t) \\ d_{\pm}(t) &= y(t) - [\widetilde{\boldsymbol{\theta}}_{\pm}^{\circ}(t)]^{\mathrm{T}}\boldsymbol{\varphi}(t) \\ d_{+}(t) &= y(t) - \widetilde{\boldsymbol{\theta}}_{+}^{\mathrm{T}}(t|t+1)\boldsymbol{\varphi}(t)\end{aligned}$$

denote the corresponding prediction/interpolation errors.

E. Collaborative LBF estimation

Similarly as in the preestimation case, the collaborative LBF estimates have the form

$$\widetilde{\boldsymbol{\theta}}(t) = \eta_{-}(t)\widetilde{\boldsymbol{\theta}}_{-}(t) + \eta_{\pm}(t)\widetilde{\boldsymbol{\theta}}_{\pm}(t) + \eta_{+}(t)\widetilde{\boldsymbol{\theta}}_{+}(t) \tag{32}$$

where the credibility coefficients $\eta_{-}(t)$, $\eta_{\pm}(t)$ and $\eta_{+}(t)$, $\eta_{-}(t) + \eta_{\pm}(t) + \eta_{+}(t) = 1$, can be obtained from

$$\eta_{\star}(t) \propto \left[\frac{\mathcal{D}_{\min}(t)}{\mathcal{D}_{\star}(t)}\right]^{\frac{K_0+1}{2}} , \quad \star \in \{-, \pm, +\} . \tag{33}$$

The accuracy of the competitive and collaborative LBF estimates can be further increased by means of postfiltration, namely the estimates (30) and (33) can be smoothed in the analogous way (using the same settings) as the preestimates (13).

F. Adaptive selection of λ_0 and m

So far we have assumed that the forgetting constant λ_0, which determines the effective memory span of EWBF/fLBF algorithms, and the number of basis functions m, which decides upon the flexibility of the basis function model (18), are fixed design parameters, selected prior to identification. It is known that small values of $1/(1-\lambda_0)$ and/or large values of m result in parameter estimates with a small bias but large variance, and that the opposite is true when $1/(1-\lambda_0)$ is large and/or m is small [14]. Since the mean square parameter tracking error is the sum of its bias and variance components, to guarantee good tracking performance of the identification algorithm, one should choose compromise values of λ_0 and m, trading-off estimation bias and estimation variance. The adaptive solution to this problem can be obtained via parallel estimation. In this approach several identification algorithms, equipped with different values of $\lambda_0 \in \Lambda = \{\lambda_1, \ldots, \lambda_k\}$ and $m \in \mathcal{M} = \{m_1, \ldots, m_l\}$, yielding the estimates $\widetilde{\boldsymbol{\theta}}_\star(t|\lambda_0, m)$, are run concurrently and compete/collaborate with each other. The best-local parameter estimate can be obtained using the formula

$$\widetilde{\boldsymbol{\theta}}_{\widehat{\star}(t)}(t|\widehat{\lambda}_0(t), \widehat{m}(t)) \tag{34}$$

where

$$\{\widehat{\star}(t), \widehat{\lambda}_0(t), \widehat{m}(t)\} = \underset{\substack{\lambda_0 \in \Lambda, m \in \mathcal{M} \\ \star \in \{-,\pm,+\}}}{\arg\min}\ \mathcal{D}_\star(t|\lambda_0, m). \tag{35}$$

Alternatively, one can implement the collaborative estimation scheme analogous to (32)

$$\widetilde{\boldsymbol{\theta}}(t) = \sum_{\substack{\lambda_0 \in \Lambda, m \in \mathcal{M} \\ \star \in \{-,\pm,+\}}} \eta_\star(t|\lambda_0, m) \widetilde{\boldsymbol{\theta}}_\star(t|\lambda_0, m) \tag{36}$$

where

$$\eta_\star(t|\lambda_0, m) \propto \left[\frac{\mathcal{D}_{\min}(t)}{\mathcal{D}_\star(t|\lambda_0, m)}\right]^{\frac{K_0+1}{2}}$$
$$\mathcal{D}_{\min}(t) = \min\Big\{\mathcal{D}_\star(t|\lambda_0, m) : \lambda_0 \in \Lambda, m \in \mathcal{M}, \star \in \{-,\pm,+\}\Big\} \tag{37}$$

subject to

$$\sum_{\substack{\lambda_0 \in \Lambda, m \in \mathcal{M} \\ \star \in \{-,\pm,+\}}} \eta_\star(t|\lambda_0, m) = 1.$$

VI. Computational complexity

Forward/backward EWLS estimates, and hence also the parameter preestimates, can be computed at the cost of $O(n)$ multiply-add operations per time step using one of the available fast transversal filter (FTF) algorithms [20], [21] (some numerical safety measures are recommended if FTF algorithms are implemented in finite-precision arithmetic). Since postfiltering is carried out independently for each system parameter, the cost of performing this step is $O(m^2 n)$. Finally, the cost of updating the statistics $\mathcal{E}_-(t)$, $\mathcal{E}_\pm(t)$, $\mathcal{E}_+(t)$ and $\mathcal{D}_-(t)$, $\mathcal{D}_\pm(t)$, $\mathcal{D}_+(t)$ is $O(1)$ and does not depend on K and K_0, respectively. Hence, the overall cost of evaluating parameter estimates is $O(m^2 n)$ per time step, i.e., it linearly depends on the number of estimated parameters. In contrast with this, the computational burden of the algorithms presented in [12], [13] is $O(m^3 n^3)$ due to the need to multiply and/or invert $mn \times mn$-dimensional matrices.

VII. Computer simulations

To check performance of the proposed approach, a parallel estimation scheme was implemented for the system (9), combining causal/anticausal/noncausal indirect LBF algorithms, described in Section V, designed for 3 different values of m (1, 2, 3) and 3 different equivalent estimation memory spans L_∞ (10, 30, 90) of the forward/backward EWBF algorithms. For $m = 1$ it holds that $L_\infty = (1+\lambda_0)/(1-\lambda_0)$; for $m > 1$ the analytical formulas allowing one to compute L_∞ are given in [13]. The values of λ_0 corresponding to different choices of m and L_∞ are shown in Table I. The forgetting constant used at the preestimation stage was set to $\lambda = 0.9$, and the widths of the decision windows - to $K = K_0 = 30$.

TABLE I

The values of the forgetting constant λ_0 corresponding to different choices of the number of basis functions m and the equivalent memory of EWBF trackers L_∞.

$m \setminus L_\infty$	10	30	90
1	0.818	0.920	0.975
2	0.936	0.973	0.984
3	0.978	0.991	0.995

Table IIa summarizes results – averaged mean squared parameter estimation errors – obtained for all 27 algorithms (A_-, A_+, $A_\pm$) under 3 different SNR levels (10 dB, 20 dB, 30 dB), and the results yielded by the corresponding competitive and collaborative algorithms. Ensemble averaging was performed over 100 realizations of the measurement noise. Table IIb shows results obtained after the second round of smoothing. Typical identification results, obtained for SNR= 20 dB, are shown in Fig. 2. Note that additional smoothing noticeably improves the estimation results (further smoothing does not). Finally, Table III shows the reference results obtained using the state-of-the-art direct LBF approach described in [13] (not based on preestimation).

In the majority of cases, especially for higher values of SNR, the adaptive algorithms yield better results than any of the component algorithms. Note also that while the one-shot indirect LBF algorithm gives slightly worse results than its direct counterpart (in spite of the fact that the component algorithms seem to work better), the additional round of smoothing makes the indirect and direct approaches fully comparable in terms of estimation accuracy. As expected, for the basic indirect LBF approach and the direct LBF approach, the collaborative estimates are more accurate than the competitive ones. This observation does not extend to the smoothed indirect approach, most likely because in this case the input estimation noise is not white any more (unlike the preestimation noise).

The next example is more realistic and involves a simulated underwater acoustic FD channel with 20 taps [3]. A single data packet contained 5000 samples, which corresponds to 5 s under the assumed sampling rate 1 kHz. The time-varying taps were modeled as realizations of independent random Gaussian processes (white noise bandlimited to 3 Hz,

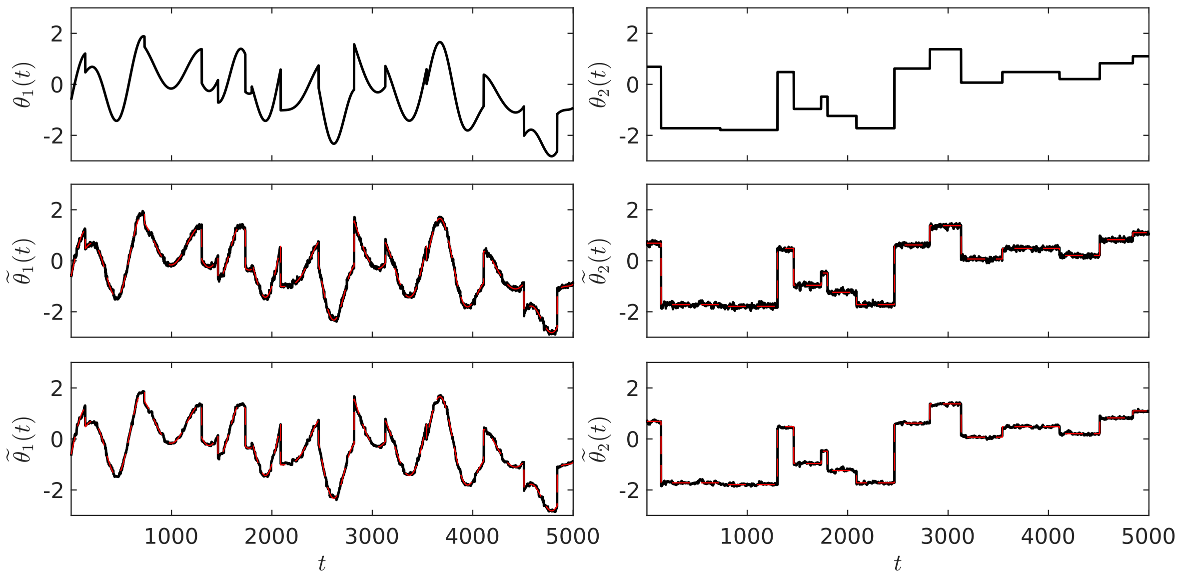

Fig. 2. Parameter trajectories of the simulated nonstationary system (two upper figures) and their estimates: competitive indirect LBF estimates (two middle figures) and their smoothed versions (two lower figures).

TABLE II

MSE SCORES OBTAINED FOR INDIRECT LBF ALGORITHMS AND THEIR COMPETITIVE AND COLLABORATIVE VARIANTS. ALL RESULTS WERE AVEREGED OVER 100 INDEPENDENT REALIZATIONS OF MEASUREMENT NOISE.

(a) Basic estimates

		$L_\infty = 10$			$L_\infty = 30$			$L_\infty = 90$			Competitive	Collaborative
SNR	m	A_-	A_+	$A_\pm$	A_-	A_+	$A_\pm$	A_-	A_+	$A_\pm$		
10 dB	1	7.80E-02	7.37E-02	6.03E-02	8.25E-02	7.66E-02	**4.04E-02**	2.57E-01	2.46E-01	**6.44E-02**	6.51E-02	5.88E-02
	2	7.73E-02	7.21E-02	**4.05E-02**	1.11E-01	1.04E-01	6.04E-02	2.09E-01	1.99E-01	9.87E-02		
	3	1.04E-01	9.70E-02	4.47E-02	3.19E-01	3.01E-01	8.24E-02	6.33E-01	5.86E-01	1.55E-01		
20 dB	1	2.47E-02	2.36E-02	**1.29E-02**	5.84E-02	5.67E-02	**2.07E-02**	2.48E-01	2.41E-01	**5.77E-02**	8.09E-03	6.97E-03
	2	3.34E-02	3.20E-02	2.44E-02	9.06E-02	8.86E-02	5.33E-02	1.97E-01	1.91E-01	9.43E-02		
	3	7.81E-02	7.56E-02	3.14E-02	3.07E-01	2.94E-01	7.69E-02	6.25E-01	5.83E-01	1.52E-01		
30 dB	1	1.95E-02	1.85E-02	**8.20E-03**	5.62E-02	5.46E-02	**1.87E-02**	2.47E-01	2.40E-01	**5.71E-02**	1.08E-03	9.26E-04
	2	2.91E-02	2.79E-02	2.28E-02	8.87E-02	8.68E-02	5.26E-02	1.95E-01	1.90E-01	9.39E-02		
	3	7.55E-02	7.33E-02	3.01E-02	3.06E-01	2.92E-01	7.64E-02	6.25E-01	5.82E-01	1.52E-01		

(a) Smoothed estimates

		$L_\infty = 10$			$L_\infty = 30$			$L_\infty = 90$			Competitive	Collaborative
SNR	m	A_-	A_+	$A_\pm$	A_-	A_+	$A_\pm$	A_-	A_+	$A_\pm$		
10 dB	1	5.20E-02	4.80E-02	4.13E-02	7.70E-02	7.19E-02	**3.86E-02**	2.56E-01	2.46E-01	**6.55E-02**	4.18E-02	5.88E-02
	2	6.26E-02	5.78E-02	**4.00E-02**	1.08E-01	1.02E-01	6.15E-02	2.07E-01	1.99E-01	1.00E-01		
	3	9.93E-02	9.28E-02	4.54E-02	3.16E-01	3.00E-01	8.35E-02	6.32E-01	5.87E-01	1.56E-01		
20 dB	1	2.27E-02	2.18E-02	**1.18E-02**	5.83E-02	5.70E-02	**2.12E-02**	2.47E-01	2.42E-01	**5.82E-02**	5.69E-03	6.97E-03
	2	3.27E-02	3.14E-02	2.49E-02	9.09E-02	8.91E-02	5.38E-02	1.97E-01	1.92E-01	9.45E-02		
	3	7.83E-02	7.57E-02	3.21E-02	3.08E-01	2.96E-01	7.73E-02	6.27E-01	5.85E-01	1.52E-01		
30 dB	1	1.90E-02	1.84E-02	**8.08E-03**	5.58E-02	5.51E-02	**1.88E-02**	2.46E-01	2.41E-01	**5.69E-02**	7.24E-04	9.25E-04
	2	2.92E-02	2.82E-02	2.29E-02	8.94E-02	8.88E-02	5.25E-02	1.97E-01	1.93E-01	9.32E-02		
	3	7.64E-02	7.48E-02	3.03E-02	3.09E-01	2.98E-01	7.64E-02	6.27E-01	5.87E-01	1.51E-01		

TABLE III

MSE SCORES OBTAINED FOR DIRECT LBF ALGORITHMS AND THEIR COMPETITIVE AND COLLABORATIVE VARIANTS. RESULTS WERE AVEREGED OVER 100 INDEPENDENT REALIZATIONS OF MEASUREMENT NOISE.

		$L_\infty = 10$			$L_\infty = 30$			$L_\infty = 90$			Competitive	Collaborative
SNR	m	A_-	A_+	$A_\pm$	A_-	A_+	$A_\pm$	A_-	A_+	$A_\pm$		
10 dB	1	8.24E-02	7.51E-02	4.23E-02	8.55E-02	7.79E-02	**3.95E-02**	2.61E-01	2.48E-01	**1.09E-01**	3.98E-02	3.04E-02
	2	8.19E-02	7.37E-02	**4.29E-02**	1.14E-01	1.04E-01	9.80E-02	2.14E-01	2.01E-01	2.02E-01		
	3	1.08E-01	9.73E-02	8.45E-02	3.26E-01	3.02E-01	3.18E-01	6.46E-01	5.95E-01	6.32E-01		
20 dB	1	3.09E-02	2.38E-02	**1.62E-02**	6.40E-02	5.66E-02	**2.88E-02**	2.55E-01	2.42E-01	**1.05E-01**	5.43E-03	4.18E-03
	2	3.95E-02	3.15E-02	3.44E-02	9.61E-02	8.71E-02	9.45E-02	2.04E-01	1.90E-01	2.00E-01		
	3	8.40E-02	7.41E-02	8.07E-02	3.17E-01	2.93E-01	3.16E-01	6.40E-01	5.90E-01	6.31E-01		
30 dB	1	2.57E-02	1.87E-02	**1.36E-02**	6.18E-02	5.45E-02	**2.77E-02**	2.54E-01	2.41E-01	**1.05E-01**	7.87E-04	6.50E-04
	2	3.53E-02	2.73E-02	3.36E-02	9.43E-02	8.54E-02	9.41E-02	2.02E-01	1.89E-01	2.00E-01		
	3	8.16E-02	7.18E-02	8.03E-02	3.16E-01	2.92E-01	3.16E-01	6.40E-01	5.90E-01	6.31E-01		

TABLE IV
MSE SCORES OBTAINED FOR A SIMULATED UWA CHANNEL FOR DIRECT AND INDIRECT LBF ALGORITHMS AND THEIR COMPETITIVE AND COLLABORATIVE VARIANTS. ALL RESULTS WERE AVEREGED OVER 100 INDEPENDENT REALIZATIONS OF MEASUREMENT NOISE.

		$L_\infty = 10$			$L_\infty = 30$			$L_\infty = 90$			Competitive	Collaborative
Method	m	A_-	A_+	$A_\pm$	A_-	A_+	$A_\pm$	A_-	A_+	$A_\pm$		
Indirect LBF	1	4.71E-02	4.37E-02	3.32E-02	6.38E-02	5.69E-02	**1.63E-02**	3.67E-01	3.52E-01	**2.06E-02**	1.26E-02	1.17E-02
	2	3.76E-02	3.57E-02	1.50E-02	9.04E-02	8.37E-02	1.89E-02	2.79E-01	2.64E-01	5.24E-02		
	3	6.52E-02	6.34E-02	**1.48E-02**	4.70E-01	4.58E-01	4.16E-02	9.88E-01	9.73E-01	2.28E-01		
Direct LBF	1	8.93E-02	8.84E-02	2.12E-02	1.30E-01	1.30E-01	**2.41E-02**	4.80E-01	4.94E-01	**1.44E-01**	6.09E-03	5.12E-03
	2	4.51E-02	4.82E-02	**1.64E-02**	1.59E-01	1.56E-01	9.76E-02	4.11E-01	4.03E-01	3.16E-01		
	3	1.22E-01	1.23E-01	5.77E-02	6.77E-01	6.65E-01	5.68E-01	1.23E+00	1.24E+00	1.12E+00		

which corresponds to fast changes in the UWA case), and tap variances $\mathrm{var}[\theta_j(t)]$ declined exponentially for increasing j to reflect the decaying power delay profile caused by the spreading and absorption loss. The time-varying impulse response generated in this way was subject to two jump changes at instants $t = 1500$ and $t = 3500$, triggered by the sudden appearance and disappearance of an extra scatterer, respectively. The signal (self-interference) to noise ratio was set to 50 dB which is typical of FD communication where SNR is usually large (often in excess of 50 dB). All the remaining technical details (the form of the input signal, the choices of m, λ, λ_0, K and K_0) were exactly the same as in the previous example.

The corresponding parameter tracking results obtained, similarly as in the previous example, by means of combined time and ensemble averaging, are shown in Table IV. Time averaging was restricted to the interval $[101, 4900]$ (the results corresponding to the first and last 100 samples were excluded since both border regions require special treatment). Note that in the UWA channel case the adaptive selection mechanisms still work satisfactorily. Unlike the previous example, there is some performance gap between the direct and indirect algorithms (in this case additional smoothing does not improve tracking results noticeably). This performance deterioration is a price that has to be paid for a reduction of the computational load offered by the indirect approach.

VIII. Conclusion

A new, computationally simple approach to identification of time-varying FIR systems was proposed and compared with the state-of-the-art solution. The new method is based on the preestimation paradigm which allows one to to convert the problem of identification of a time-varying system to the problem of smoothing properly generated preestimates of system parameters. The proposed two-stage algorithm can be used to identify systems with both continuous-smooth and occasional jump-type parameter changes, typical of some telecommunication applications.

References

[1] C. Liu, Y. V. Zakharov, T. Chen, "Doubly selective underwater acoustic channel model for a moving transmitter/receiver," IEEE Transactions on Vehicular Technology 61 (3) (2012) 938-950.

[2] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, pp. 16371652, 2014.

[3] L. Shen, B. Henson, Y. Zakharov, P. Mitchell, "Digital self-interference cancellation for full-duplex underwater acoustic systems," IEEE Transactions on Circuits and Systems II: Express Briefs 67 (1) (2019) 192-196.

[4] 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, vol. 200, 108664, 2022.

[5] J. G. Proakis, *Digital Communications*. McGraw-Hill, 1983.

[6] J. Tugnait, "Adaptive estimation and identification for discrete systems with Markov jump parameters," *IEEE Transactions on Automatic Control*, vol. 27, 1982.

[7] R. E. Helmick, W. D. Blair, and S. A. Hoffman, "Fixed-interval smoothing for Markovian switching systems," *IEEE Transactions on Information Theory*, vol. 41, pp. 18451855, 1995.

[8] A. Logothetis and V. Krishnamurthy, "Expectation maximization algorithms for MAP estimation of jump Markov linear systems," *IEEE Transactions on Signal Processing*, vol. 47, pp. 21392156, 1999.

[9] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: a survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, pp. 103123, 1998.

[10] B. Chen and J.K. Tugnait, "Interacting multiple model fixed-lag smoothing algorithm for Markovian switching systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, pp. 243-250, 2000.

[11] C. Andrieu, A. Doucet, and R. Holenstein, "Particle Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 72, pp. 269342, 2010.

[12] M. Niedźwiecki, "Identification of time-varying systems with abrupt parameter changes," *Automatica*, vol. 30, pp. 447-459, 1994.

[13] M. Niedźwiecki and S. Gackowski, "New approach to noncausal identification of nonstationary FIR systems subject to both smooth and abrupt parameter changes," *IEEE Transactions on Automatic Control*, vol. 58, pp. 1847-1853, 2013.

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

[15] 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.

[16] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Communications Magazine*, vol. 47, pp. 84-89, 2009.

[17] M. Niedźwiecki, "On time and frequency characteristics of weighted least squares estimators applied to nonstationary system identification," 24th IEEE Conference on Decision and Control, Fort Lauderdale, USA, pp. 225-230, 1985.

[18] M. Niedźwiecki, *Identification of Time-varying Processes*, Wiley, 2000.

[19] M. Niedźwiecki, "Locally-adaptive Kalman smoothing and its application to identification of nonstationary stochastic systems," *IEEE Transactions on Signal Processing*, vol. 60, pp. 48-59, 2012.

[20] S. Haykin, *Adaptive Filter Theory*. Prentice-Hall, 1996.

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