Postprint of: Gańcza A., Niedźwiecki M., Regularized Identification of Time-Varying FIR Systems Based on Generalized Cross-Validation, 29th European Signal Processing Conference (EUSIPCO), 2021, pp. 2035-2039, DOI: 10.23919/EUSIPCO54536.2021.9616058

Regularized Identification of Time-Varying FIR Systems Based on Generalized Cross-Validation

1st Artur Gańcza
Department of Automatic Control
Gdańsk University of Technology
ul. Narutowicza 11/12, Gdańsk, Poland
artgancz@student.pg.edu.pl

2nd Maciej Niedźwiecki
Department of Automatic Control
Gdańsk University of Technology
ul. Narutowicza 11/12, Gdańsk, Poland
maciekn@pg.edu.pl

***Abstract*—A new regularization method is proposed and applied to identification of time-varying finite impulse response systems. We show, that by a careful design of the regularization constraint, one can improve estimation results, especially in the presence of strong measurement noise. We also show that the the most appropriate regularization gain can be found by direct optimization of the generalized cross-validation criterion.**

***Index Terms*—Identification of nonstationary systems, basis functions, generalized cross-validation**

I. Introduction

Identification of nonstationary systems, exploited in many disciplines such as telecommunications [1], geophysics [2] or biomedicine [3], can be particularly demanding in the presence of strong measurement noise or when system parameter changes are fast. Recently, a new approach to this problem was described in [4]. The proposed local basis function (LBF) approach builds upon the assumption that inside a local analysis window, parameter changes can be approximated by a linear combination of some known functions of time. Since the method requires inversion of large-sized matrices and computations are carried out in a sliding window mode, it can be computationally quite demanding. In order to overcome this problem, a novel, simplified algorithm was developed in the follow-up paper [5]. The new, fast local basis function (fLBF) method allows one to convert the original identification problem into a task of smoothing the preestimated parameter trajectories. We will show that the accuracy of fLBF estimates can be further increased by applying the regularization technique. It is well-known that a carefully designed regularization can lower the mean square parameter estimation error (MSE) by improving the bias-variance trade-off [6]. In order to properly tune the regularization constant we will minimize the generalized cross-validation criterion [7].

The main interest of this paper lies in the identification of parameters of the finite impulse response (FIR) model described by the equation

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

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

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

A new challenging application of nonstationary FIR system identification techniques, proposed recently, is tracking and equalization of underwater acoustic (UWA) channels [8], [9]. The time variability of channel impulse response coefficients results from the Doppler effect due to the transmitter/receiver and water motion [10]. It is worth noting that noncausal estimators, like the one presented in this paper, can be used in this application since some decision delay is acceptable. Another problem connected with UWA communication, which requires identification of a time-varying channel impulse response, is self-interference cancellation in full-duplex communication systems [11], [12].

II. Preestimation

Preestimates are raw estimates of parameter trajectories. They are approximately unbiased regardless of the type and speed of parameter changes. However, the price for their unbiasedness is a large variance. Hence, in order to obtain reliable estimates, additional filtration is performed. As shown in [5], preestimates can be obtained via "inverse filtration" of short-memory exponentially weighted least squares (EWLS) estimates,

$$\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t) = \arg\min_{\boldsymbol{\theta}} \sum_{i=0}^{t-1} \lambda^i [y(t-i) - \boldsymbol{\varphi}^{\mathrm{T}}(t-i)\boldsymbol{\theta}]^2, \tag{2}$$

namely

$$\boldsymbol{\theta}^*(t) = L_t \widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t) - \lambda L_{t-1} \widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t-1), \tag{3}$$

where $\lambda \in (0,1)$ denotes the so-called forgetting constant and $L_t = \sum_{i=0}^{t-1} \lambda^i = \lambda L_{t-1} + 1$ is the effective width of the exponential window. The EWLS estimates can be computed recursively using the well-known algorithm [13]

$$\begin{aligned}
\varepsilon(t) &= y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t-1) \\
\mathbf{k}(t) &= \frac{\mathbf{P}(t-1)\boldsymbol{\varphi}(t)}{\lambda + \boldsymbol{\varphi}^{\mathrm{T}}(t)\mathbf{P}(t-1)\boldsymbol{\varphi}(t)} \\
\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t) &= \widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t-1) + \mathbf{k}(t)\varepsilon(t) \\
\mathbf{P}(t) &= \frac{1}{\lambda}\left[\mathbf{P}(t-1) - \frac{\mathbf{P}(t-1)\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\mathbf{P}(t-1)}{\lambda + \boldsymbol{\varphi}^{\mathrm{T}}(t)\mathbf{P}(t-1)\boldsymbol{\varphi}(t)}\right],
\end{aligned} \tag{4}$$

with initial conditions $\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(0) = 0$ and $\mathbf{P}(0) = c\mathbf{I}_n$ where $\mathbf{I}_n$ denotes the $n \times n$ identity matrix and c is a large positive constant.

For large values of t, when the effective window width reaches its steady state value $L_\infty = 1/(1-\lambda)$, the formula (3) can be replaced with

$$\boldsymbol{\theta}^*(t) = \frac{1}{1-\lambda}\left[\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t) - \lambda\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t-1)\right]. \tag{5}$$

The rule of thumb for choosing the value of λ, which works well in practice, is $\lambda = \max\{0.9, 1-2/n\}$. When $\lambda = 1-2/n$ the equivalent width of the exponential window $N_\infty = (1+\lambda)/(1-\lambda) \cong 2/(1-\lambda)$, different from its effective width L_∞ [13], is approximately equal to the number of estimated coefficients n.

It can be shown that if the input signal is (locally) stationary, and the noise $\{e(t)\}$ is white, the preestimates are approximately unbiased

$$\boldsymbol{\theta}^*(t) \cong \boldsymbol{\theta}(t) + \mathbf{z}(t), \tag{6}$$

where $\mathbf{z}(t)$ denotes zero-mean, white noise [5].

III. REGULARIZED FAST LOCAL BASIS FUNCTION ESTIMATORS

A. Fast local basis function approach

In order to reduce the variance of the noise contaminating preestimates, a postfiltering technique should be incorporated. Assume that inside the local analysis window $T_k(t) = [t-k, t+k]$ of width $K = 2k+1$, centered at t, each parameter trajectory can be approximated by a linear combination of known, linearly independent functions of time $f_1(i), \ldots, f_m(i),\ i \in I_k = [-k, k]$, called basis functions, namely

$$\boldsymbol{\Theta}_j(t) = \mathbf{F}\boldsymbol{\alpha}_j, \quad j = 1, \ldots, n, \tag{7}$$

where $\boldsymbol{\Theta}_j(t) = [\theta_j(t-k), \ldots, \theta_j(t+k)]^{\mathrm{T}}$, $\boldsymbol{\alpha}_j = [\alpha_{1,j}, \ldots, \alpha_{m,j}]^{\mathrm{T}}$, $\mathbf{F} = [\mathbf{f}(-k), \ldots, \mathbf{f}(k)]^{\mathrm{T}}$, and $\mathbf{f}(i) = [f_1(i), \ldots, f_m(i)]^{\mathrm{T}}$. Without any loss of generality, we will require the basis to be orthonormal, namely

$$\mathbf{F}^{\mathrm{T}}\mathbf{F} = \mathbf{I}_m. \tag{8}$$

Note that orthogonalization of any set of basis functions can be carried out sequentially using the well-known Gram-Schmidt procedure.

Following [5], the fLBF estimates at the time instant t can be obtained in the form

$$\begin{aligned}
\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t) &= \arg\min_{\boldsymbol{\alpha}_j} ||\boldsymbol{\Theta}_j^*(t) - \mathbf{F}\boldsymbol{\alpha}_j||^2 \\
&= (\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\boldsymbol{\Theta}_j^*(t) = \mathbf{F}^{\mathrm{T}}\boldsymbol{\Theta}_j^*(t) \\
\widehat{\theta}_j^{\mathrm{fLBF}}(t) &= \mathbf{f}_0^{\mathrm{T}}\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t),
\end{aligned} \tag{9}$$

where $\mathbf{f}_0 = \mathbf{f}(0)$ and $||\boldsymbol{x}||$ denotes the L_2 norm. Once the estimates are evaluated, the analysis window is moved to the next position ($t \to t+1$) and the procedure is repeated. For some choices of basis functions, computations can be carried out in a recursive manner [4]. In this study, we will use Legendre polynomials, obtained by orthonormalization of the basis made up of powers of time

$$g_l(i) = \left(\frac{i}{k}\right)^{l-1}, \quad i \in I_k,\ l = 1, \ldots, m.$$

It is worth mentioning that the fLBF procedure described above allows one to obtain estimates that in many cases are almost indistinguishable from those yielded by the, computationally much more demanding LBF procedure [5]. Additionally, the fLBF scheme is more robust to numerical errors.

B. Regularization

Regularization was originally introduced as a way of imposing some smoothness constraints on a solution of the estimation problem [14]. Later on it was reinvented as a remedy for solving some ill-posed inverse problems [15], [16]. Recently, it gained a lot of attention in identification of linear time-invariant (LTI) systems [6], [17] since it can improve the MSE score by reducing the estimation variance at the cost of slightly increasing the bias.

The regularized fLBF (fRLBF) estimates can be written down in the following form

$$\begin{aligned}
\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fRLBF}}(t) &= \arg\min_{\boldsymbol{\alpha}_j} \left[||\boldsymbol{\Theta}_j^*(t) - \mathbf{F}\boldsymbol{\alpha}_j||^2 + \mu_j||\mathbf{f}_0^{\mathrm{T}}\boldsymbol{\alpha}_j||^2\right] \\
&= (\mathbf{F}^{\mathrm{T}}\mathbf{F} + \mu_j\mathbf{f}_0\mathbf{f}_0^{\mathrm{T}})^{-1}\mathbf{F}^{\mathrm{T}}\boldsymbol{\Theta}_j^*(t) \\
\widehat{\theta}_j^{\mathrm{fRLBF}}(t) &= \mathbf{f}_0^{\mathrm{T}}\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fRLBF}}(t),
\end{aligned} \tag{10}$$

where $\mu_j > 0$ denotes a regularization constant. Note that the applied variant of regularization penalizes the norm of the parameter vector $\boldsymbol{\theta}(t) = \mathbf{f}_0^{\mathrm{T}}\boldsymbol{\alpha}_j$, the estimation of which is the purpose of system identification, instead of the norm of the vector of hyperparameters $\boldsymbol{\alpha}_j$. The latter one is penalized in the classical regularization approach.

Using the Sherman-Morrison formula [18], one can compute the fRLBF estimates in terms of the fLBF estimates

$$\begin{aligned}
\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fRLBF}}(t) &= \left[\mathbf{I}_m - \frac{\mu_j\mathbf{f}_0\mathbf{f}_0^{\mathrm{T}}}{1 + \mu_j\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0}\right]\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t) \\
\widehat{\theta}_j^{\mathrm{fRLBF}}(t) &= \frac{\widehat{\theta}_j^{\mathrm{fLBF}}(t)}{1 + \mu_j\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0}.
\end{aligned} \tag{11}$$

IV. Optimization via Generalized Cross Validation

One of the well-known methods for choosing the regularization constant is cross-validation. Cross-validation can be applied in many different forms, the predicted residual sum of squares (PRESS) [19] being probably the most popular one. However, although intuitive, PRESS can be computationally challenging. Moreover, it is well-known, that under some operating conditions [7] better results can be obtained via the so-called generalized cross-validation (GCV), which also asymptotically minimizes the following quality measure (see theorem 1 in [20]) $\mathrm{E}[||\mathbf{F}\boldsymbol{\alpha}_j - \mathbf{F}\widehat{\boldsymbol{\alpha}}_j(t)||^2]$. Since the basis is orthonormal, this is equivalent to minimization of $\mathrm{E}[||\boldsymbol{\alpha}_j - \widehat{\boldsymbol{\alpha}}_j(t)||^2]$. We will show that in the case considered, the GCV criterion can be optimized in a direct way, leading to a closed-form solution.

The GCV quality measure is given by [7]

$$\begin{aligned} \mathrm{GCV}(t,\mu_j) &= \frac{||\boldsymbol{\Theta}_j^*(t) - \widehat{\boldsymbol{\Theta}}_j^{\mathrm{fRLBF}}(t)||^2}{\{\mathrm{tr}[\mathbf{I}_K - \mathbf{A}(\mu_j)]\}^2} \\ &= \frac{||[\mathbf{I}_K - \mathbf{A}(\mu_j)]\boldsymbol{\Theta}_j^*(t)||^2}{\{\mathrm{tr}[\mathbf{I}_K - \mathbf{A}(\mu_j)]\}^2}, \end{aligned} \tag{12}$$

where $\widehat{\boldsymbol{\Theta}}_j^{\mathrm{fRLBF}}(t) = \mathbf{A}(\mu_j)\boldsymbol{\Theta}_j^*(t)$ is the predicted "output" and

$$\begin{aligned} \mathbf{A}(\mu_j) &= \mathbf{F}(\mathbf{I}_m + \mu_j\mathbf{f}_0\mathbf{f}_0^{\mathrm{T}})^{-1}\mathbf{F}^{\mathrm{T}} \\ &= \mathbf{F}\mathbf{F}^{\mathrm{T}} - \frac{\mu_j}{1+\mu_j\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0}\mathbf{F}\mathbf{f}_0\mathbf{f}_0^{\mathrm{T}}\mathbf{F}^{\mathrm{T}}. \end{aligned}$$

Consequently

$$\begin{aligned} \mathrm{tr}[\mathbf{I}_K - \mathbf{A}(\mu_j)] &= K - m + \frac{\mu_j}{1+\mu_j\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0}\mathrm{tr}[\mathbf{B}] \\ &= K - m + \frac{\mu_j\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0}{1+\mu_j\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0}, \end{aligned} \tag{13}$$

where $\mathbf{B} = \mathbf{F}\mathbf{f}_0\mathbf{f}_0^{\mathrm{T}}\mathbf{F}^{\mathrm{T}}$. All transitions in the above formula follow directly from the well-known trace identity $\mathrm{tr}[\mathbf{AB}] = \mathrm{tr}[\mathbf{BA}]$, which holds provided that dimensions of the corresponding matrices match.

Straightforward calculations lead to the following expression

$$\mathrm{GCV}(t,x_j) = \frac{a_j(t)x_j^2 + b_j(t)}{(c+dx_j)^2}, \tag{14}$$

where

$$\begin{aligned} x_j &= \frac{\mu_j}{1+\mu_j\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0} \in \left(0, \frac{1}{\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0}\right) \\ a_j(t) &= \mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0[\boldsymbol{\Theta}_j^*(t)]^{\mathrm{T}}\mathbf{B}\boldsymbol{\Theta}_j^*(t) \\ &= \mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0[\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t)]^{\mathrm{T}}\mathbf{f}_0\mathbf{f}_0^{\mathrm{T}}\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t) = \mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0[\widehat{\theta}_j^{\mathrm{fLBF}}(t)]^2 \\ b_j(t) &= [\boldsymbol{\Theta}_j^*(t)]^{\mathrm{T}}[\mathbf{I}_K - \mathbf{F}\mathbf{F}^{\mathrm{T}}][\mathbf{I}_K - \mathbf{F}\mathbf{F}^{\mathrm{T}}]^{\mathrm{T}}\boldsymbol{\Theta}_j^*(t) \\ &= [\boldsymbol{\Theta}_j^*(t)]^{\mathrm{T}}[\mathbf{I}_K - \mathbf{F}\mathbf{F}^{\mathrm{T}}]\boldsymbol{\Theta}_j^*(t) \\ &= [\boldsymbol{\Theta}_j^*(t)]^{\mathrm{T}}\boldsymbol{\Theta}_j^*(t) - [\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t)]^{\mathrm{T}}\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t) \\ c &= K - m \\ d &= \mathrm{tr}[\mathbf{B}] = \mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0. \end{aligned} \tag{15}$$

Note that both $a_j(t)$ and $b_j(t)$ can be expressed in terms of the fLBF estimates. The lack of the linear term in the numerator of (14) is due to the fact that

$$[\mathbf{I}_K - \mathbf{F}\mathbf{F}^{\mathrm{T}}]\mathbf{F}\mathbf{f}_0\mathbf{f}_0^{\mathrm{T}}\mathbf{F}^{\mathrm{T}} = \mathbf{F}\mathbf{f}_0\mathbf{f}_0^{\mathrm{T}}\mathbf{F}^{\mathrm{T}} - \mathbf{F}\mathbf{f}_0\mathbf{f}_0^{\mathrm{T}}\mathbf{F}^{\mathrm{T}} = 0.$$

The restricted domain of x_j stems from the fact that $\mu_j > 0$. It is straightforward to check that when $a_j(t), b_j(t), c, d > 0$, which is the case, the unrestricted (global) minimum of (14) is attained for

$$x_j^{\min}(t) = \frac{b_j(t)d}{a_j(t)c}. \tag{16}$$

Furthermore, since $x_j = \mu_j/(1+\mu_j\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0)$ and

$$\frac{\partial x_j}{\partial \mu_j} = \frac{1}{(1+\mu_j\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0)^2} > 0,$$

the global minimum of (12) is obtained for

$$\mu_j^{\min}(t) = \frac{x_j^{\min}(t)}{1 - x_j^{\min}(t)\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0} = \frac{b_j(t)d}{a_j(t)c - b_j(t)d^2}. \tag{17}$$

Finally, taking into account the positivity constraint imposed on μ_j, the optimal-local regularization gain can be expressed in the form

$$\mu_j^{\mathrm{opt}}(t) = \begin{cases} \mu_j^{\min}(t) & \text{if} \quad x_j^{\min}(t) < \frac{1}{\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0} \\ \infty & \text{if} \quad x_j^{\min}(t) \geq \frac{1}{\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0}. \end{cases} \tag{18}$$

The practical way of implementing this rule is by the following formula

$$\widehat{\theta}_j^{\mathrm{fRLBF}}(t) = \begin{cases} \frac{\widehat{\theta}_j^{\mathrm{fLBF}}(t)}{1+\mu_j^{\min}(t)\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0} & \text{if} \quad \mu_j^{\min}(t) \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

Remark 1: Note that the unbiased estimate of the variance of noise contaminating the j-th parameter trajectory can be expressed as

$$\widehat{\sigma}_{z_j}^2(t) = \frac{1}{K-m}||\boldsymbol{\Theta}_j^*(t) - \mathbf{F}\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t)||^2 = \frac{b_j(t)}{c}. \tag{20}$$

Hence, combining (15), (17) and (20), one arrives at the following equivalent expression for $\mu_j^{\min}(t)$

$$\begin{aligned} \mu_j^{\min}(t) &= \frac{\widehat{\sigma}_{z_j}^2(t)}{[\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t)]^{\mathrm{T}}\mathbf{f}_0\mathbf{f}_0^{\mathrm{T}}\widehat{\boldsymbol{\alpha}}_j^{\mathrm{fLBF}}(t) - \mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0\widehat{\sigma}_{z_j}^2(t)} \\ &= \frac{\widehat{\sigma}_{z_j}^2(t)}{[\widehat{\theta}_j^{\mathrm{fLBF}}(t)]^2 - \mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0\widehat{\sigma}_{z_j}^2(t)}. \end{aligned} \tag{21}$$

Since typically $\mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0 \ll 1$, for most of the time $[\widehat{\theta}_j^{\mathrm{fLBF}}(t)]^2 > \mathbf{f}_0^{\mathrm{T}}\mathbf{f}_0\widehat{\sigma}_{z_j}^2(t)$ and the value of $\mu_j^{\min}(t)$ is positive. According to (19), when the parameter estimate is greater than some noise-dependent threshold, its value is shrinked. However, whenever its value falls below the threshold, it is disregarded as unreliable and set to zero.

Remark 2: The computational complexity of the EWLS algorithm (4), used to obtain parameter preestimates, is of order $O(n^2)$ per time update; it can be further reduced to as little as $3n$ operations per time update if the iterative dichotomous coordinate descent (DCD) technique, described in [21], is applied. Computation of fLBF estimates is matrix-inversion-free and requires $O(nmK)$ operations per time update. It can be reduced to $O(mn)$ operations if the incorporated basis functions are recursively computable (note that Legendre polynomials have this property). From the quantities needed for evaluation of the regularization constant $\mu_j(t)$, only $a_j(t)$ and $b_j(t)$ are data-dependent and therefore need to be updated at each time step. However, since the first term in $b_j(t)$ can be also computed in a recursive fashion, the number of additional operations per time update is of order $O(mn)$.

V. Computer simulations

To show advantages of the regularization technique, a 45-tap nonstationary FIR system was simulated using the morphing technique. The time-varying impulse response $\boldsymbol{\theta}(t)$ of this system was generated as result of a smooth transition from $\boldsymbol{\theta}_{\mathrm{A}}$, through $\boldsymbol{\theta}_{\mathrm{B}}$ to $\boldsymbol{\theta}_{\mathrm{C}}$, where $\boldsymbol{\theta}_{\mathrm{A}}$, $\boldsymbol{\theta}_{\mathrm{B}}$ and $\boldsymbol{\theta}_{\mathrm{C}}$ denote the truncated impulse responses of time-invariant infinite impulse response (IIR) systems governed by

$$H_{\mathrm{A}}(q^{-1}) = \frac{0.36q^{-1}}{1+0.34q^{-1}+0.7225q^{-2}}$$

$$H_{\mathrm{B}}(q^{-1}) = \frac{0.1606+0.1606q^{-1}+0.3214q^{-2}}{1-1.4038q^{-1}+0.7225q^{-2}}$$

$$H_{\mathrm{C}}(q^{-1}) = \frac{0.015q^{-1}-0.1q^{-2}}{1+2.178q^{-1}+2.2625q^{-2}+1.2347q^{-3}+0.3540q^{-4}}.$$

In the interval $[1, T/2]$ the time-varying impulse response was generated using the formula

$$\boldsymbol{\theta}(t) = [1-w(t)]\boldsymbol{\theta}_{\mathrm{A}} + w(t)\boldsymbol{\theta}_{\mathrm{B}}, \quad t \in [1, T/2]$$

where $\{w(t)\}$ denotes the Hann window

$$w(t) = 0.5\left[1-\cos\left(\frac{2\pi t}{T}\right)\right].$$

The analogous formula was applied to generate the time-varying impulse response in the interval $[T/2+1, T]$:

$$\boldsymbol{\theta}(t) = [1-w(t-T/2)]\boldsymbol{\theta}_{\mathrm{B}} + w(t-T/2)\boldsymbol{\theta}_{\mathrm{C}}, \quad t \in [T/2+1, T].$$

Figure 2 shows snapshots of the true impulse responses of the system at time instants 1, $T/2+1$ and T, respectively, where T denotes the length of the simulation interval. The corresponding fRLBF estimates are shown in the same figure.

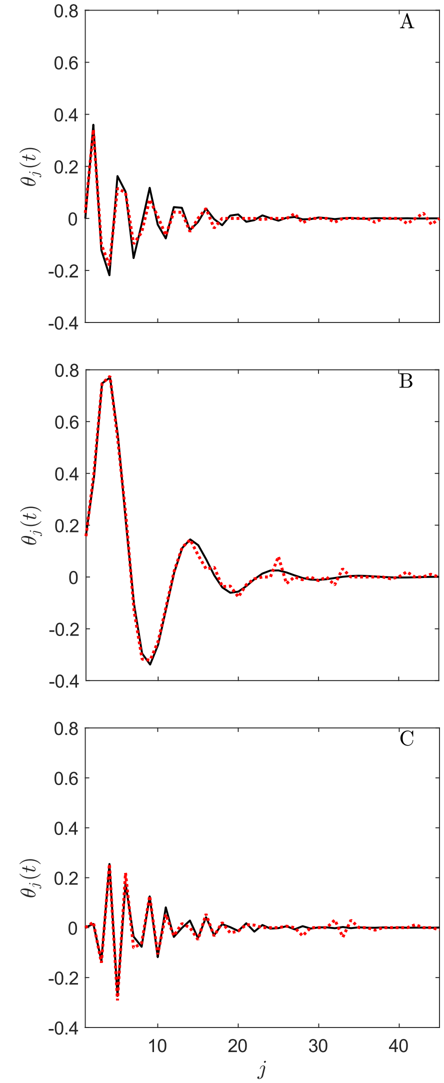

Fig. 1. True (solid lines) and estimated (dotted lines) impulse responses of the identified time-varying FIR system at $t=1$ (A), $t=T/2+1$ (B) and $t=T$ (C); the fRLBF estimates were obtained for $k=200$, $m=3$ and SNR=20 dB.

The simulated system was excited using an autoregressive signal $u(t) = 0.8u(t-1) + v(t)$, where $\{v(t)\}$ denotes white noise with unit variance, independent of $\{e(t)\}$. For the identification purpose, we used $k=200$, $n=45$ and $\lambda = 1-2/n \cong 0.956$. Computations were carried out for three estimators with m equal 1, 3 and 5, respectively, two different speeds of parameter changes, corresponding to $T=1000$ (fast changes) and $T=2000$ (slow changes), and for four different variances of measurement noise, namely 0.4, 0.04, 0.004 and 0.0004, corresponding to SNR equal to 10, 20, 30 and 40 dB, respectively. To avoid transient effects, data generation was started 1000 samples prior to $t=1$ and was continued 1000 samples after $t=T$ for a system with a constant impulse response.

Estimation accuracy was evaluated using the FIT measure

TABLE I

FIT (%) SCORES, AVERAGED OVER TIME AND 100 INDEPENDENT REALIZATIONS OF THE MEASUREMENT NOISE, FOR FAST (UPPER TABLE) AND SLOW (LOWER TABLE) PARAMETER CHANGES.

Fast parameter changes

		m		
SNR	Method	1	3	5
10 dB	fRLBF	58.5	44.9	33.2
	fLBF	50.6	30.7	14.5
20 dB	fRLBF	79.5	80.9	77.3
	fLBF	77.9	76.9	71.8
30 dB	fRLBF	83.2	89.5	88.0
	fLBF	83.4	89.7	88.0
40 dB	fRLBF	83.5	90.6	89.4
	fLBF	84.2	92.3	91.4

Slow parameter changes

		m		
SNR	Method	1	3	5
10 dB	fRLBF	62.3	45.9	34.0
	fLBF	54.8	33.1	17.0
20 dB	fRLBF	86.5	81.8	77.9
	fLBF	84.7	78.5	73.5
30 dB	fRLBF	93.1	93.0	91.8
	fLBF	92.9	92.4	90.8
40 dB	fRLBF	94.1	95.2	94.6
	fLBF	94.5	95.9	95.3

proposed in [6]

$$\mathrm{FIT}(t) = 100 \left(1 - \left[\frac{\sum_{j=1}^{45} |\theta_j(t) - \widehat{\theta}_j(t)|^2}{\sum_{j=1}^{45} |\theta_j(t) - \overline{\theta}_j(t)|^2} \right]^{1/2} \right), \quad (22)$$

where $\overline{\theta}_j(t) = \frac{1}{45}\sum_{j=1}^{45} \theta_j(t)$. The maximum value of this measure, equal to 100, corresponds to the perfect match between the true and estimated impulse responses. The final scores, further referred to as FIT (%), were obtained by combined time (over $[1, T]$) and ensemble (over 100 independent realizations of $\{e(t)\}$) averaging. Numerical results were summarized in Table I. It can be noticed that regularization improves the estimation results most significantly for the SNR equal 10 or 20 dB. For high SNR values, the proposed regularization method can worsen the results slightly, but the fRLBF method still yields high-quality estimates of system parameters.

VI. CONCLUSIONS

It was shown that the identification of a time-varying FIR system can be effectively carried out by means of smoothing the appropriately generated sequence of parameter preestimates. For smoothing purposes, one can use regularized least squares method, locally optimized using the GCV approach. The resulting identification algorithm outperforms the currently available solutions in terms of estimation accuracy, computational complexity, and numerical robustness.

REFERENCES

[1] M. K. Tsatsanis and G. B. Giannakis, "Modelling and equalization of rapidly fading channels," *International Journal of Adaptive Control and Signal Processing*, vol. 310, pp. 515-522, 1996.

[2] P. Lesage, F. Glangeaud, and J. Mars, "Applications of autoregressive models and time-frequency analysis to the study of volcanic tremor and long-period events," *J. Volc. Geotherm. Res.*, vol. 114, pp. 391–417, 2002.

[3] S.G. Fabri, K.P. Camilleri, and T. Cassar, "Parametric modelling of EEG data for the identification of mental tasks," *Biomed. Eng. Trends in Electron., Commun. & Software*, (A. Laskovski (Ed.), pp. 367-386, 2011.

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

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

[6] L. Ljung, T. Chen, "What can regularization offer for estimation of dynamical systems?," *Proc. of the 11th IFAC Workshop on Adaptation and Learning in Control and Signal Processing*, Caen, France, pp. 1-8, 2013.

[7] G. Golub, M. Heath and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, pp. 215-223, 1979.

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

[9] I. Kochańska, Adaptive identification of time-varying impulse response of underwater acoustic communication channel,", *Hydroacoustics*, vol. 18, pp. 87-94, 2015.

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

[11] 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.

[12] L. Shen, B. Henson, Y. Zakharov, N. Morozs and P. Mitchell, "Adaptive filtering for full-duplex UWA systems with time-varying self-interference channel," *Submitted to J. Oceanic Eng.* TechRxiv. https://doi.org/10.36227/techrxiv.11967813.v1, 2019.

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

[14] E. T. Whittaker, "On a new method of graduation," Proc. Edinborough Mathematical Association, vol. 78, pp. 81-89, 1923.

[15] D. L. Phillips, "A technique for the numerical solution of certain integral equations of the first kind," *Journal of the ACM*, vol. 9, pp. 84-97, 1962.

[16] A. Tikhonov and V. Arsenin, *Solutions of Ill-posed Problems*. Winston/Wiley, 1977.

[17] T. Chen, H. Ohlsson and L. Ljung "On the estimation of transfer functons, regularizations and Gaussian processes - Revisited," *Automatica*, vol. 48, pp. 1525-1535, 2012.

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

[19] D. Allen, "The relationship between variable selection and data augmentation and a method for prediction," *technometrics*, vol. 16, pp. 125-127, 1974.

[20] B. Mu, T. Chen and L. Ljung, "Asymptotic properties of generalized cross-validation estimators for regularized system identification," *IFAC-PapersOnLine*, vol. 51, pp. 203-208, 2018.

[21] 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.