Signal Processing 200 (2022) 108664

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

# Adaptive identification of sparse underwater acoustic channels with a mix of static and time-varying parameters☆

Maciej Niedźwiecki[a], Artur Gańcza[a], Lu Shen[b,*], Yuriy Zakharov[b]

[a] Faculty of Electronics, Telecommunications and Computer Science, Department of Automatic Control, Gdańsk University of Technology, Gdańsk 80-233, Poland

[b] Department of Electronic Engineering, University of York, UK

ARTICLE INFO

*Article history:*
Received 2 March 2022
Revised 1 June 2022
Accepted 13 June 2022
Available online 16 June 2022

*Keywords:*
Adaptive filter
FLBF estimator
Preestimate
Regularization
Sparse channel
Time-varying system
Underwater acoustics

ABSTRACT

We consider identification of sparse linear systems with a mix of static and time-varying parameters. Such systems are typical in underwater acoustics (UWA), for instance, in applications requiring identification of the acoustic channel, such as UWA communications, navigation and continuous-wave sonar. The recently proposed fast local basis function (fLBF) algorithm provides high performance when identifying time-varying systems. In this paper, we further improve the performance of the fLBF algorithm by exploiting properties of the system. Specifically, we propose an adaptive time-invariance test to identify whether a particular system tap is static or time-varying and exploit this knowledge for choosing the number of basis functions. We also propose a regularization scheme that exploits the system sparsity and an adaptive technique for estimating the regularization parameter. Finally, a debiasing technique is proposed to reduce an inherent bias of fLBF estimates. The high performance of the fLBF algorithm with the proposed techniques is demonstrated in scenarios of UWA communications, using numerical and real experiments.



## 1. Introduction

There are many applications that require accurate estimation of parameters of time-varying linear systems with only a part of the parameters being time-varying. Such applications are typical in underwater acoustics (UWA) and include UWA communications, navigation, and sonar applications, which deal with estimation of the UWA channel often modelled as a time-varying linear system [1–3]. The UWA channel is characterised by multipath propagation and often is described as a finite impulse response (FIR) filter, whose parameters are varying in time due to the Doppler effect caused by the moving transmitter, receiver and/or the sea surface [4]. The Doppler effect is known to be different for different propagation paths [5].

In an UWA communication system, after correction of the dominant Doppler effect in the received signal, the signal paths not interacting with the sea surface will typically be almost static (direct path and bottom reflections), whereas the sea surface reflections will be fast time-varying [6]. In addition, there will be reflections from objects that may appear/disappear or change their speed in the vicinity of the transducer (fish, vessel, etc.). Therefore, the channel contains a mix of static and time-varying multipaths.

Our special attention will be payed to full-duplex (FD) UWA communications [7,8], where the transceiver simlultaneously transmits and receives signal in the same frequency bandwidth. In FD UWA systems, apart from the weak signal from the far-end user, the near-end reciever will receive a strong self-interference (SI) from the near-end transmitter within the same transceiver. To allow FD operation, the SI signal should be accurately recovered and removed from the received signal. This requires high-precision SI channel estimates. The SI channel may include reflections from the sea surface; even if the SI sea-surface signal components are of a low power (tens of decibels lower than the signal components due to the line-of-sight propagation), they still need to be estimated with a high precision to achieve a level of SI cancellation required for the FD operation (at least 60 dB). Therefore, the channel estimation should deal with a mix of static and time-varying parameters. In this paper, we show how this property of linear systems can be exploited to improve the identification performance.

☆ The work of Y. Zakharov and L. Shen was supported in part by the U.K. EPSRC through Grants EP/V009591/1 and EP/R003297/1. The work of M. Niedźwiecki and A. Gańcza was partially supported by the National Science Center under the agreement UMO-2018/29/B/ST7/00325.

* Corresponding author.
*E-mail address:* lu.shen@york.ac.uk (L. Shen).

https://doi.org/10.1016/j.sigpro.2022.108664

For identifying time-varying systems, adaptive filters are widely used [9,10]. Classical least-mean square and affine projection adaptive filters have slow convergence. A faster convergence is provided by the classical recursive least squares (RLS) based algorithms. Although they are of high complexity when directly implemented, their fast versions are also available [10,11]. Sparse RLS adaptive algorithms [12–14] are widely used for estimating UWA channels, thus exploiting the fact that most of the channel coefficients are close to zero. The Time-Updated RLS (TU-RLS) algorithm is considered as the state-of-the-art algorithm for time-varying UWA channel estimation as it takes into account both the delay and Doppler spread [15], i.e., the fact that propagation paths may have different speed of variation in time. However, the RLS algorithms might still be limited in their tracking performance due to the predictive (causal) nature: for their operation, only current and past signal samples are used. There are multiple applications, such as UWA applications discussed above, which allow the use of interpolating (non-causal) adaptive filters. For example, in communications, it is normally accepted that there exists a delay (latency) in the channel estimation. In such scenarios, non-causal estimators are preferable since they can provide a better tracking performance, which is essential for fast-varying channels.

The local basis function (LBF) principle [16–19] exploits the non-causality and thus can provide a high identification accuracy when estimating (tracking) parameters of nonstationary systems (channels). However, LBF adaptive algorithms are very complicated. The fast LBF (fLBF) approach has recently been proposed to reduce the complexity [17,18]. The fLBF approach divides the estimation procedure in two steps: preestimation and postfiltering. This allows dealing with the regressors and basis functions separately at these two steps, respectively, thus reducing the complexity. The essential feature of the preestimator is that it provides an unbiased estimate, although with a high variance. The preestimator based on the exponentially weighted least squares (EWLS) algorithm with inverse filtering provides approximately unbiased estimates and it allows a high performance of fLBF algorithms [17,18].

As we discussed above, there are scenarios where only some of the system parameters (channel taps) are time-varying, whereas the others are static. We will be focusing on these scenarios. Therefore, we re-formulate the LBF estimation problem taking this model into account. We call the corresponding estimators oracle (e.g., oracle fLBF estimator) since they exploit the perfect knowledge of the time-varying or time-invariant nature of the estimated parameters. In practice, such knowledge is not available, therefore an adaptive algorithm is proposed to identify the static or time-varying nature of parameters based on counting sign changes in the appropriately centered preestimates; this simple for implementation algorithm shows remarkably good performance. In order to further improve the fLBF performance when the system parameters are sparse, we propose a regularization scheme with adaptive estimation of the regularization parameter. A debiasing approach is further proposed to compensate for the bias introduced at the preestimation step of the fLBF algorithms.

The contributions of the paper are as follows.

1. We formulate the problem of adaptive estimation of parameters of a time-varying linear system, in which only a part of the parameters are time-varying, while the other parameters are static.
2. An adaptive technique (statistical test) is proposed for deciding on each parameter if it is static or time-varying.
3. A regularized fLBF (fRLBF) algorithm is proposed, which allows an improved performance when estimating sparse systems.
4. An adaptive estimator of the optimal regularization parameter is proposed.
5. A debiasing technique is proposed, which significantly improves the performance of the fLBF algorithms.
6. The proposed techniques are evaluated in underwater acoustic communication scenarios numerically and using experimental data.

Part of the materials from this paper was presented at the 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). In addition to the techniques proposed in the ICASSP paper (contributions 1, 2 and 5), we also propose a regularization scheme that exploits the system sparsity and an adaptive technique for estimating the optimal regularization parameter, which leads to the regularized fLBF algorithm (contributions 3 and 4). The debiasing procedure is applied to the regulaized fLBF estimates to further improve the identification performance. The proposed techniques are evaluated in underwater acoustic communication scenarios numerically and using experimental data.

The paper is organised as follows. In Section 2, we formulate the estimation problem. Section 3 describes the LBF and fLBF approaches in the context of the perfect knowledge of the time-varying or time-invariant nature of the estimated parameters. Section 4 introduces the adaptive statistical test for deciding on the time variability of parameters. In Section 5, we introduce the regularization scheme and consider practical issues of its implementation. Section 6 introduces the new debiasing technique. Section 8 presents results of numerical simulation and experimental investigation of the proposed algorithms in UWA communication scenarios. Finally, Section 9 presents conclusions.

*Notation:* We use the following notation. The symbol $*$ stands for complex conjugate and H - complex conjugate transpose (Hermitian transpose). We denote by $\mathbf{I}_m$ an $m \times m$ identity matrix; $n_S = \mathrm{card}\{S\}$ denotes the cardinality of a set $S$; $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ are the real and imaginary parts of a complex number, respectively. A block diagonal matrix $\mathbf{F}(j)$ is denoted as $\mathrm{bl\,diag}\{\mathbf{F}_1(j), \ldots, \mathbf{F}_n(j)\}$, where the main-diagonal blocks are vectors $\mathbf{F}_1(j), \ldots, \mathbf{F}_n(j)$ and all off-diagnaol blocks are zeros. The determinant of a matrix $A$ is denoted by $|A|$.

## 2. Problem statement

Many nonstationary systems, such as communication channels (terrestrial, underwater) can be well approximated by a time-varying FIR model of the form Stojanovic and Preisig [4], Tsatsanis and Giannakis [20]

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

where $t = \ldots, -1, 0, 1, \ldots$ denotes discrete (normalized) time, $y(t)$ the complex-valued system output (a received signal in a communication system), $\boldsymbol{\varphi}(t) = [u(t), \ldots, u(t-n+1)]^{\mathrm{T}}$ the regression vector made up of past samples of the complex-valued (transmitted) signal $u(t)$, $\boldsymbol{\theta}(t) = [\theta_1(t), \ldots, \theta_n(t)]^{\mathrm{T}}$ is the vector of system parameters (e.g., channel taps), and $e(t)$ denotes a measurement noise. The sequence $\{\theta_i(t)\}$ can be interpreted as a time-varying impulse response of the channel to be estimated.

We will assume that:

(A1) $\{u(t)\}$ are zero-mean independent and identically distributed circular random variables with variance $\sigma_u^2$;
(A2) $\{e(t)\}$ is a zero-mean circular white noise, independent of $\{u(t)\}$, with variance $\sigma_e^2$;
(A3) $\{\boldsymbol{\theta}(t)\}$ is a sequence independent of $\{u(t)\}$ and $\{e(t)\}$.

The LBF principle is based on the assumption that in a local analysis interval $T_k(t) = [t-k, t+k]$ of length $K = 2k+1$, centered at $t$, system parameters can be expressed as linear combinations of a certain number of linearly independent complex-valued

functions of time $f_1(j), \ldots, f_m(j)$, $j \in I_k = [-k, k]$, further referred to as basis functions. Without any loss of generality we will assume that the basis functions are orthonormal, namely by denoting $\mathbf{f}(j) = [f_1(j), \ldots, f_m(j)]^{\mathrm{T}}$ we have

$$\sum_{j=-k}^{k} \mathbf{f}(j)\mathbf{f}^{\mathrm{H}}(j) = \mathbf{I}_m. \tag{2}$$

In this paper, we adopt the complex exponential basis set of the form (see Tsatsanis and Giannakis [20], Sayeed and Aazhang [21], Zakharov and Kodanev [22] for a physical justification of such a choice)

$$\{f_1(j), \ldots, f_m(j),\ j \in I_k\} = \left\{\frac{1}{\sqrt{K}}e^{\mathrm{i}j\omega_1}, \ldots, \frac{1}{\sqrt{K}}e^{\mathrm{i}j\omega_m},\ j \in I_k\right\}, \tag{3}$$

where $\mathrm{i} = \sqrt{-1}$, $\omega_1 = 0$, $m = 2m_0 + 1$, and

$$\omega_{2l} = -\frac{2\pi l}{K}, \quad \omega_{2l+1} = \frac{2\pi l}{K}, \quad l = 1, \ldots, m_0.$$

It is straightforward to check that the basis (3) is orthonormal. We will denote $f_0 = f_1(j) = 1/\sqrt{K}$, $j \in I_k$.

Unlike [16–19], we will assume that only some of the estimated parameters $\theta_i(t)$ in (1) vary in the local analysis interval $T_k(t)$, while the remaining parameters are constant.

Denote by $S$ the set indicating, within $\Omega = \{1, \ldots, n\}$, positions of time-invariant taps, and by $\bar{S} = \Omega - S$ the set of time-varying taps. Furthermore, let $n_S = \mathrm{card}\{S\}, n_{\bar{S}} = \mathrm{card}\{\bar{S}\}$, then $n_S + n_{\bar{S}} = n$, and denote $\ell = n_S + mn_{\bar{S}}$. In the sequel we will adopt the following *mixed-mode* model of local parameter variation within the interval $T_k(t)$:

$$\theta_i(t+j) = \begin{cases} f_0 a_{i1}(t) & \text{if} \quad i \in S \\ \sum_{l=1}^{m} f_l^*(j) a_{il}(t) & \text{if} \quad i \in \bar{S} \end{cases} \quad j \in I_k, \quad i = 1, \ldots, n. \tag{4}$$

In agreement with the local estimation paradigm, estimation of parameter trajectories, based on the hypermodel (4), will be carried out independently for each location of the analysis interval $T_k(t)$, i.e., it will be performed in the sliding window manner. Therefore, even though system hyperparameters (expansion coefficients) $a_{il}$ are assumed to be constant in the interval $[t-k, t+k]$, their values are allowed to change along with the position of the analysis window. For this reason they are written down as functions of $t$.

The hypermodel (4) can be expressed in a more compact form

$$\boldsymbol{\theta}(t+j) = \mathbf{F}(j)\boldsymbol{\alpha}(t), \quad j \in I_k, \tag{5}$$

where $\boldsymbol{\alpha}(t)$ is an $\ell$-dimensional vector of hyperparameters,

$$\boldsymbol{\alpha}(t) = [\boldsymbol{\alpha}_1^{\mathrm{T}}(t), \ldots, \boldsymbol{\alpha}_n^{\mathrm{T}}(t)]^{\mathrm{T}},$$

$$\boldsymbol{\alpha}_i(t) = \begin{cases} a_{i1}(t) & \text{if} \quad i \in S \\ [a_{i1}(t), \ldots, a_{im}(t)]^{\mathrm{T}} & \text{if} \quad i \in \bar{S} \end{cases}$$

and $\mathbf{F}(j)$ denotes the $n \times \ell$ matrix,

$$\mathbf{F}(j) = \mathrm{bl\ diag}\{\mathbf{F}_1(j), \ldots, \mathbf{F}_n(j)\}$$

$$\mathbf{F}_i(j) = \begin{cases} f_0 & \text{if} \quad i \in S \\ \mathbf{f}^{\mathrm{H}}(j) & \text{if} \quad i \in \bar{S} \end{cases}$$

Using (5), the system model (1) in the local analysis interval $T_k(t)$ can be written in the form

$$y(t+j) = \boldsymbol{\alpha}^{\mathrm{H}}(t)\boldsymbol{\psi}(t, j) + e(t+j), \quad j \in I_k, \tag{6}$$

where $\boldsymbol{\psi}(t, j) = \mathbf{F}^{\mathrm{H}}(j)\boldsymbol{\varphi}(t+j)$ denotes the generalized regression vector.

The LBF approach allows estimating the vector $\boldsymbol{\alpha}(t)$, and, consequently, the vector $\boldsymbol{\theta}(t)$.

## 3. Oracle LBF and fLBF estimators

By oracle estimation algorithms we will mean algorithms exploiting the perfect knowledge of the support sets $S$ and $\bar{S}$. In Section 4, we will propose an adaptive estimator capable of identifying the support sets.

The LBF estimator has the form Niedźwiecki and Ciołek [16]

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}^{\mathrm{LBF}}(t) &= \arg\min_{\boldsymbol{\alpha}} \textstyle\sum_{j=-k}^{k} |y(t+j) - \boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\psi}(t, j)|^2 = \mathbf{R}^{-1}(t)\mathbf{r}(t), \\ \widehat{\boldsymbol{\theta}}^{\mathrm{LBF}}(t) &= \mathbf{F}_0\widehat{\boldsymbol{\alpha}}^{\mathrm{LBF}}(t), \end{aligned} \tag{7}$$

where

$$\begin{aligned} \mathbf{R}(t) &= \sum_{j=-k}^{k} \boldsymbol{\psi}(t, j)\boldsymbol{\psi}^{\mathrm{H}}(t, j), \\ \mathbf{r}(t) &= \sum_{j=-k}^{k} \boldsymbol{\psi}(t, j)y^*(t+j), \end{aligned} \tag{8}$$

and $\mathbf{F}_0 = \mathbf{F}(0)$. Unfortunately, the computation of the matrix $\mathbf{R}(t)$ and its inversion at every instant $t$ requires a high computational load; the direct computation may result in $\mathcal{O}(Kn^2m^2)$ and $\mathcal{O}(n^3m^3)$ arithmetic operations, respectively.

As shown in Niedźwiecki et al. [17], 23], under assumptions (A1)-(A3), the LBF estimates $\widehat{\boldsymbol{\alpha}}^{\mathrm{LBF}}(t)$ and $\widehat{\boldsymbol{\theta}}^{\mathrm{LBF}}(t)$ can be approximated by the fLBF estimates

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}^{\mathrm{fLBF}}(t) &= \arg\min_{\boldsymbol{\alpha}} \sum_{j=-k}^{k} ||\widetilde{\boldsymbol{\theta}}(t+j) - \mathbf{F}(j)\boldsymbol{\alpha}||^2 = \sum_{j=-k}^{k} \mathbf{F}^{\mathrm{H}}(j)\widetilde{\boldsymbol{\theta}}(t+j), \\ \widehat{\boldsymbol{\theta}}^{\mathrm{fLBF}}(t) &= \mathbf{F}_0\widehat{\boldsymbol{\alpha}}^{\mathrm{fLBF}}(t), \end{aligned} \tag{9}$$

where $\{\widetilde{\boldsymbol{\theta}}(t)\}$ denotes a sequence of preestimated system parameters and we denote $||\mathbf{x}||^2 = \mathbf{x}^{\mathrm{H}}\mathbf{x}$. The fLBF estimates can be obtained in a significantly more computationally efficient way than the LBF estimates, thus the word 'fast'. The preestimation step requires $\mathcal{O}(n^2)$ or $\mathcal{O}(n)$ arithmetic operations per time instance if the classical EWLS algorithm or its fast versions are used, respectively, whereas the complexity of the postfiltering step is $\mathcal{O}(mnK)$ in general case. For the exponential basis used in this paper, the fLBF estimates can be computed recursively using a diagonal transition matrix [18] as

$$\mathbf{f}(j) = \boldsymbol{\Lambda}\mathbf{f}(j+1), \quad \boldsymbol{\Lambda} = \mathrm{diag}\left\{1, e^{-\mathrm{i}\omega_2}, \ldots, e^{-\mathrm{i}\omega_m}\right\},$$

which leads to the following recursive formula

$$\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t+1) = \begin{cases} \widehat{\alpha}_i^{\mathrm{fLBF}}(t) - \widetilde{\theta}_i(t-k)f_0 + \widetilde{\theta}_i(t+k+1)f_0 & \text{if} \quad i \in S \\ \boldsymbol{\Lambda}\left[\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t) - \widetilde{\theta}_i(t-k)\mathbf{f}(-k)\right] + \widetilde{\theta}_i(t+k+1)\mathbf{f}(k) & \text{if} \quad i \in \bar{S} \end{cases}$$

Since the matrix $\boldsymbol{\Lambda}$ is diagonal, fLBF estimates can be obtained at a cost of $12mn_{\bar{S}} + 4n_S$ real-valued multiply and accumulate (MAC) operations per time update. Importantly, this cost does not depend on the width of the analysis window $K$. Preestimates are raw estimates of parameter trajectories - approximately unbiased (no matter how system parameters change over time) but with a high variance. Therefore, to obtain statistically meaningful results preestimates must be further processed (postfiltered).

Fast LBF estimators can be rewritten in a decomposed form as follows. For $i \in S$, we have

$$\begin{aligned} \widehat{\alpha}_i^{\mathrm{fLBF}}(t) &= f_0 \sum_{j=-k}^{k} \widetilde{\theta}_i(t+j), \\ \widehat{\theta}_i^{\mathrm{fLBF}}(t) &= f_0\widehat{\alpha}_i^{\mathrm{fLBF}}(t) = \frac{1}{K}\sum_{j=-k}^{k} \widetilde{\theta}_i(t+j), \end{aligned} \tag{10}$$

and, for $i \in \bar{S}$,

$$\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t) = \sum_{j=-k}^{k} \widetilde{\theta}_i(t+j)\mathbf{f}(j),$$

$$\widehat{\theta}_i^{\mathrm{fLBF}}(t) = \mathbf{f}_0^{\mathrm{H}}\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t) = \sum_{j=-k}^{k} h(j)\widetilde{\theta}_i(t+j), \tag{11}$$

where $\mathbf{f}_0 = \mathbf{f}(0)$ and

$$h(j) = \mathbf{f}_0^{\mathrm{H}}\mathbf{f}(j), \quad j \in I_k, \tag{12}$$

denotes the impulse response of an FIR filter associated with the LBF estimator.

The preestimates proposed in Niedźwiecki and Kłaput [24] and further developed in Niedźwiecki et al. [17],23] can be obtained, for every time instant $t$, by 'inverse filtering' of estimates yielded by the EWLS algorithm [9,10]

$$\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t) = \arg\min_{\boldsymbol{\theta}} \sum_{j=1}^{t} \lambda^{t-j}|y(j) - \boldsymbol{\theta}^{\mathrm{H}}\boldsymbol{\varphi}(j)|^2 = \mathbf{G}^{-1}(t)\mathbf{g}(t), \tag{13}$$

where $\lambda$, $0 < \lambda < 1$, is a forgetting factor and

$$\mathbf{G}(\mathbf{t}) = \sum_{j=1}^{t} \lambda^{t-j}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{H}}(j),$$

$$\mathbf{g}(t) = \sum_{j=1}^{t} \lambda^{t-j}\boldsymbol{\varphi}(j)y^*(j). \tag{14}$$

The effective width $M(t)$ of the exponential window can be evaluated using

$$M(t) = \sum_{i=1}^{t} \lambda^{t-i} = \lambda M(t-1) + 1 \tag{15}$$

with the initial condition $M(0) = 0$. The forgetting factor $\lambda$ should be "as small as possible", thus keeping the EWLS estimation memory also small, to maximize the estimation bandwidth, i.e., the frequency range in which system parameters can be tracked successfully. It is recommended that $\lambda \geq 0.9$ since for smaller values of $\lambda$ the EWLS algorithm may behave in an erratic way due to the possible poor conditioning of the exponentially weighted regression matrix $\mathbf{G}(t)$. On the other hand, $\lambda$ should not be "too small"to guarantee that the number of system parameters is not greater than the steady-state equivalent number of observations $N_\infty = (1+\lambda)/(1-\lambda) \cong 2/(1-\lambda)$ used for their estimation (different from the effective number of observations [25]) - otherwise the estimation results would be questionable from the statistical viewpoint. This leads to the following recommendation

$$\lambda = \max\{0.9, 1 - \frac{2}{n}\}. \tag{16}$$

For the numerical simulations in this paper, $\lambda = 1 - \frac{2}{n}$ is used as the number of channel coefficients is higher than 20, which is typical for UWA channels.

The inverse filtering derived and analysed in Niedźwiecki et al. [17] is given by

$$\widetilde{\boldsymbol{\theta}}(t) = M(t)\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t) - \lambda M(t-1)\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t-1). \tag{17}$$

For large values of $t$, when $M(t)$ reaches its steady-state value $M_\infty = 1/(1-\lambda)$, (17) can be replaced with

$$\widetilde{\boldsymbol{\theta}}(t) = \frac{1}{1-\lambda}\left[\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t) - \lambda\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t-1)\right]. \tag{18}$$

As shown in Niedźwiecki et al. [17], when the sequence $\{\boldsymbol{\varphi}(t)\}$ is (locally) stationary with exponentially decaying autocorrelation function, the preestimates are approximately unbiased, i.e.

$$\mathrm{E}[\widetilde{\boldsymbol{\theta}}(t)] \cong \boldsymbol{\theta}(t), \tag{19}$$

where the expectation is over $\{e(t)\}$ and $\{\boldsymbol{\varphi}(t)\}$. Under assumptions (A1)–(A3), the preestimation noise $\mathbf{z}(t) = \widetilde{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}(t)$ is approximately white.

## 4. Adaptive time-invariance test

A clear advantage of the preestimation approach is that it allows the system dynamics to be 'X-rayed' prior to its formal identification. We will use this property to adaptively decide, at each time instant $t$, which parameters can be regarded as time-invariant and which are time-varying.

To assess existence of a trend in the sequence of preestimates (which justifies choosing $m > 1$), one can use the classical approach based on counting the number of sign changes amongst residuals [26]. When the system parameter $\theta_i(t)$ is constant in the analysis interval $T_k(t)$, the residual noise defined as

$$\varepsilon_i(t+j|t) = \widetilde{\theta}_i(t+j) - \bar{\theta}_i(t), \quad j \in I_k, \tag{20}$$

where

$$\bar{\theta}_i(t) = \frac{1}{K}\sum_{j=-k}^{k} \widetilde{\theta}_i(t+j),$$

is approximately equal to the preestimation noise $z_i(t) = \widetilde{\theta}_i(t+j) - \theta_i(t)$, which is zero-mean and white, $\varepsilon_i(\bullet|t)$ denotes elements of the residual noise vector computed at $t$.

Consider the real part of the residual noise:

$$\varepsilon_i^{\mathrm{R}}(t+j|t) = \mathrm{Re}\{\varepsilon_i(t+j|t)\}.$$

Let

$$p_i^{\mathrm{R}}(t+j|t) = \beta[\varepsilon_i^{\mathrm{R}}(t+j-1|t)]\beta[\varepsilon_i^{\mathrm{R}}(t+j|t)],$$

where $j \in [-k+1, k]$, $p_i^{\mathrm{R}}(t+j|t) \in \{-1, 1\}$ and

$$\beta[x] = \begin{cases} 1 & \text{if} \quad x > 0 \\ -1 & \text{if} \quad x \leq 0 \end{cases}.$$

Finally, denote by

$$q_i^{\mathrm{R}}(t) = \mathrm{card}\{\mathcal{Q}_i^{\mathrm{R}}(t)\} \in \{1, \ldots, K-1\},$$

where $\mathcal{Q}_i^{\mathrm{R}}(t) = \{j \in [-k+1, k] : p_i^{\mathrm{R}}(t+j|t) = -1\}$, the number of sign changes of $\varepsilon_i^{\mathrm{R}}(\cdot|t)$ observed in the analysis interval $T_k(t)$. By $q_i^{\mathrm{I}}(t)$ we will denote the analogous count for $\varepsilon_i^{\mathrm{I}}(t+j|t) = \mathrm{Im}\{\varepsilon_i(t+j|t)\}$.

For a real-valued white noise sequence, the sign change can be observed on average every second sample. Hence, when the number of sign changes is smaller than some threshold, one has to assume that the parameter trajectory is not constant inside the analysis window.

Consider the following null hypothesis:

$\mathcal{H}_0^{\mathrm{R}}(t)$: $\{\varepsilon_i^{\mathrm{R}}(t+j|t), j \in I_k\}$ is a sequence of independent random variables obeying the condition

$$P\big(\varepsilon_i^{\mathrm{R}}(t+j|t) > 0\big) = P\big(\varepsilon_i^{\mathrm{R}}(t+j|t) \leq 0\big), \quad \forall j \in I_k,$$

where $P(\cdot)$ is a probability. Note that this hypothesis is true when the sequence $\{\varepsilon_i^{\mathrm{R}}(\cdot|t)\}$ is uncorrelated, zero-mean and Gaussian, but the requirements are in fact much weaker. Note also, that for the null hypothesis $\mathcal{H}_0^{\mathrm{R}}(t)$ to be true there is no need to assume that random variables $\varepsilon_i^{\mathrm{R}}(\cdot|t)$ have the same variance, or even that their variance exists (is well defined).

If the null hypothesis is true, $q_i^{\mathrm{R}}(t)$ is a discrete random variable with an "almost[1] binomial" distribution characterized by the probability of success 0.5:

$$P(q_i^{\mathrm{R}}(t) = q|\mathcal{H}_0^{\mathrm{R}}(t)) = \frac{(K-1)!}{(2^{K-1}-1)q!(K-1-q)!}. \tag{21}$$

[1] As remarked by Geary [26], since the sum of residuals in the interval $T_k(t)$ is - by construction - zero, the value $q_i^{\mathrm{R}}(t) = 0$ is inadmissible.

Furthermore, for any $q_0 \in [1, K-1]$ it holds that

$$P(q_i^{\mathrm{R}}(t) \le q_0 | \mathcal{H}_0^{\mathrm{R}}(t)) = \sum_{q=1}^{q_0} \frac{(K-1)!}{(2^{K-1}-1)q!(K-1-q)!} = \eta_0. \quad (22)$$

The sign statistic $q_i^{\mathrm{R}}(t)$ can be used to verify the null hypothesis for a given probability of Type I error $\eta_0$:

$$\begin{cases} \text{accept } \mathcal{H}_0^{\mathrm{R}}(t) & \text{if } q_i^{\mathrm{R}}(t) > q_0 \\ \text{reject } \mathcal{H}_0^{\mathrm{R}}(t) & \text{if } q_i^{\mathrm{R}}(t) \le q_0 \end{cases} \quad (23)$$

The exemplary thresholds evaluated for the significance level $\eta_0 = 0.05$ are $q_0 = 87, 183$ and $376$ for $K = 2k+1 = 201, 401$ and $801$, respectively.

The same analysis can be carried out for the imaginary components of the residuals $\varepsilon_i^{\mathrm{I}}(\cdot|t)$. Combining both inferences, one arrives at the following decision rule, which can be used to determine whether a given parameter $\theta_i(t)$ should be regarded as constant or time-varying in the analysis interval $T_k(t)$:

$$\begin{cases} i(t) \in S(t) & \text{if} \quad q_i^{\mathrm{R}}(t) > q_0 \text{ and } q_i^{\mathrm{I}}(t) > q_0 \\ i(t) \in \bar{S}(t) & \text{otherwise} \end{cases} \quad (24)$$

**Remark 1.** At each time instant $t$ the proposed time-invariance test is based on the number of sign changes evaluated for the sequence of residuals

$$\mathcal{E}_i(t) = \{\varepsilon_i(t-k|t), \ldots, \varepsilon_i(t+k|t)\}.$$

When the parameter trajectory is already decided to be locally constant and $k \gg 10$, one can use a simplified approach. Assuming that $\bar{\theta}_i(t) \cong \bar{\theta}_i(t-1)$, one arrives at

$$\varepsilon_i(t+j|t) \cong \varepsilon_i(t+j+1|t-1), \quad j = -k, \ldots, k-1.$$

Denote by $\kappa(\cdot)$ the operator which removes the first element from the list of arguments, namely

$$\kappa[\mathcal{E}_i(t)] = \{\varepsilon_i(t-k+1|t), \ldots, \varepsilon_i(t+k|t)\},$$

and define the set of modified (approximate) residuals in the recursive form as follows. If $i(t-1) \in S(t-1)$, then

$$\widetilde{\mathcal{E}}_i(t) = \{\kappa[\widetilde{\mathcal{E}}_i(t-1)], \varepsilon_i(t+k|t)\}.$$

If $i(t-1) \in \bar{S}(t-1)$, then

$$\widetilde{\mathcal{E}}_i(t) = \mathcal{E}_i(t),$$

with the initial condition $\widetilde{\mathcal{E}}_i(1) = \mathcal{E}_i(1)$. Since evaluation of the statistics $q_i^{\mathrm{R}}(t)$ and $q_i^{\mathrm{I}}(t)$ for the modified set $\widetilde{\mathcal{E}}_i(t)$ can be performed recursively using the results obtained for $\widetilde{\mathcal{E}}_i(t-1)$, for large values of $k$ the computational savings can be substantial. To prevent long-term discrepancy between results based on analysis of $\widetilde{\mathcal{E}}_i(t)$ and those based on analysis of $\mathcal{E}_i(t)$, a periodic resetting is recommended in the form $\widetilde{\mathcal{E}}_i(jt_0) = \mathcal{E}_i(jt_0)$, $j = 1, 2, \ldots$ where $t_0$ denotes the resetting period (e.g. $t_0 = k$).

## 5. Regularized fLBF estimators

It is well known that the accuracy of parameter estimates can be improved by means of regularization. The idea is to add to the minimized cost function a term, called regularizer, which reduces a norm of the solution. Regularization allows one to reach a better bias-variance compromise which results in smaller mean-square parameter estimation errors [27]. While regularization is a well established technique in system identification, most of the existing work in this area is focused on identification of time-invariant systems [28], [29]. The recently published papers [18,19] are an exception to this rule. In [18], where estimation is carried out for time-varying systems using the LBF/fLBF approach, the $L_2$ regularization is applied, penalizing excess values of the squared norm of $\boldsymbol{\alpha}(t)$. Since the ultimate goal of identification is estimation of the vector of system parameters $\boldsymbol{\theta}(t)$ (estimation of hyperparameters is only a means to achieve this goal), similar to Niedźwiecki et al. [19], in this paper we will use a different strategy involving penalization of norm of $\boldsymbol{\theta}(t)$. Unlike [16], where the main purpose of regularization was to impose some soft smoothness constraints on the estimated impulse response (which is not the case for channel identification due to the lack of correlation between taps), our main goal here is to reduce the values of estimates of coefficients that can be regarded as statistically insignificant.

### 5.1. Fast regularized LBF (fRLBF) algorithm

In this approach, for each position of the analysis window, we will penalize the weighted norm of $\boldsymbol{\theta}(t)$ using the weights $\gamma_i = 1/\sigma_i^2$, $\sigma_i^2 = \mathrm{var}[\theta_i(t)]$, which reflect the known power delay profile $\{\sigma_i^2\}$, $i = 1, \ldots, n$, of the channel. The corresponding regularizer has the form

$$\mu \sum_{i=1}^{n} \frac{|\theta_i(t)|^2}{\sigma_i^2} = \mu ||\boldsymbol{\theta}(t)||_{\boldsymbol{\Gamma}}^2 = \mu ||\boldsymbol{\alpha}(t)||_{\boldsymbol{\Sigma}}^2, \quad (25)$$

where $\boldsymbol{\Gamma} = \mathrm{diag}\{\gamma_1, \ldots, \gamma_n\}$, $\mu > 0$ denotes the regularization parameter, $\boldsymbol{\Sigma}$ is the $\ell \times \ell$ regularization matrix

$$\boldsymbol{\Sigma} = \mathbf{F}_0^{\mathrm{H}} \boldsymbol{\Gamma} \mathbf{F}_0 = \mathrm{bl\ diag}\{\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_n\}, \quad (26)$$

and

$$\boldsymbol{\Sigma}_i = \begin{cases} \gamma_i f_0^2 & \text{if} \quad i \in S \\ \gamma_i \mathbf{f}_0 \mathbf{f}_0^{\mathrm{H}} & \text{if} \quad i \in \bar{S}. \end{cases}$$

Such regularization will make the identification algorithm focus more on the estimation of "large" parameters and less on estimation of "small", potentially insignificant, ones.

The regularized fLBF estimators can be defined as:

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}^{\mathrm{fRLBF}}(t) &= \arg\min_{\boldsymbol{\alpha}} \left\{ \sum_{j=-k}^{k} ||\widetilde{\boldsymbol{\theta}}(t+j) - \mathbf{F}(j)\boldsymbol{\alpha}||^2 + \mu ||\boldsymbol{\alpha}||_{\boldsymbol{\Sigma}}^2 \right\} \\ &= [\mathbf{I}_\ell + \mu \boldsymbol{\Sigma}]^{-1} \widehat{\boldsymbol{\alpha}}^{\mathrm{fLBF}}(t), \\ \widehat{\boldsymbol{\theta}}^{\mathrm{fRLBF}}(t) &= \mathbf{F}_0 \widehat{\boldsymbol{\alpha}}^{\mathrm{fRLBF}}(t), \end{aligned} \quad (27)$$

leading, for $i \in S$, to

$$\begin{aligned} \widehat{\alpha}_i^{\mathrm{fRLBF}}(t) &= \frac{\widehat{\alpha}_i^{\mathrm{fLBF}}(t)}{1 + \mu \gamma_i f_0^2}, \\ \widehat{\theta}_i^{\mathrm{fRLBF}}(t) &= \frac{\widehat{\theta}_i^{\mathrm{fLBF}}(t)}{1 + \mu \gamma_i f_0^2}, \end{aligned} \quad (28)$$

and, for $i \in \bar{S}$, to

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}_i^{\mathrm{fRLBF}}(t) &= [\mathbf{I}_m + \mu \gamma_i \mathbf{f}_0 \mathbf{f}_0^{\mathrm{H}}]^{-1} \widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t), \\ \widehat{\theta}_i^{\mathrm{fRLBF}}(t) &= \frac{\widehat{\theta}_i^{\mathrm{fLBF}}(t)}{1 + \mu \gamma_i \mathbf{f}_0^{\mathrm{H}} \mathbf{f}_0}. \end{aligned} \quad (29)$$

### 5.2. Optimization of the regularization parameter

Minimization of the quadratic norm in (27) is equivalent to maximization of

$$\exp\left\{ -\tfrac{1}{\sigma_z^2} \textstyle\sum_{j=-k}^{k} ||\widetilde{\boldsymbol{\theta}}(t+j) - \mathbf{F}(j)\boldsymbol{\alpha}||^2 \right\} \times \exp\left\{ -\tfrac{\mu}{\sigma_z^2} ||\boldsymbol{\alpha}||_{\boldsymbol{\Sigma}}^2 \right\}, \quad (30)$$

where $\sigma_z^2$ is the variance of the preestimation noise. The expression (30) can be given a probabilistic interpretation. Such proba-

bilistic embedding of the original deterministic optimization problem is usually referred to as the empirical Bayes approach [30].

Assuming that the preestimation noise is circular white Gaussian and $\mathrm{cov}[\mathbf{z}(t)] = \sigma_z^2 \mathbf{I}_n$, the first term in (30) can be attributed to the conditional data distribution (likelihood)

$$p(\Theta(t)|\boldsymbol{\alpha}, \sigma_z^2) = \frac{1}{(\pi\sigma_z^2)^{Kn}} \times \exp\left\{ -\frac{1}{\sigma_z^2}||\boldsymbol{\alpha}||^2 + \frac{1}{\sigma_z^2}\boldsymbol{\alpha}^{\mathrm{H}}\widehat{\boldsymbol{\alpha}}^{\mathrm{fLBF}}(t) + \frac{1}{\sigma_z^2}[\widehat{\boldsymbol{\alpha}}^{\mathrm{fLBF}}(t)]^{\mathrm{H}}\boldsymbol{\alpha} - c \right\}, \tag{31}$$

where $\Theta(t) = \{\widetilde{\boldsymbol{\theta}}(t+j), j \in I_k\}$ and

$$c = \frac{1}{\sigma_z^2}\sum_{j=-k}^{k} ||\widetilde{\boldsymbol{\theta}}(t+j)||^2.$$

The second term in (30) corresponds to a prior (complex singular Gaussian) distribution of $\boldsymbol{\alpha}$,

$$\pi(\boldsymbol{\alpha}|\boldsymbol{\Sigma}, \sigma_z^2, \mu) = \frac{\mu^\ell |\boldsymbol{\Sigma}|_+}{(\pi\sigma_z^2)^\ell} \exp\left\{ -\frac{\mu}{\sigma_z^2}\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\Sigma}\boldsymbol{\alpha} \right\}, \tag{32}$$

where $|\boldsymbol{\Sigma}|_+$ denotes the pseudodeterminant of the matrix $\boldsymbol{\Sigma}$ (the product of all nonzero eigenvalues of $\boldsymbol{\Sigma}$):

$$|\boldsymbol{\Sigma}|_+ = \prod_{i=1}^{n} |\boldsymbol{\Sigma}_i|_+ = \left(\prod_{i\in S} \gamma_i f_0^2\right)\left(\prod_{i\in\bar{S}} \gamma_i \mathbf{f}_0^{\mathrm{H}}\mathbf{f}_0\right).$$

The likelihood for the unknown parameters $\boldsymbol{\Sigma}$, $\sigma_z^2$ and $\mu$ can be obtained from

$$L(\boldsymbol{\Sigma}, \sigma_z^2, \mu) = \int_{\mathbb{C}^\ell} p(\Theta(t)|\boldsymbol{\alpha}, \sigma_z^2)\pi(\boldsymbol{\alpha}|\boldsymbol{\Sigma}, \sigma_z^2, \mu) d\boldsymbol{\alpha}. \tag{33}$$

By using the relationship (e.g., see Baronkin et al. [31])

$$\int_{\mathbb{C}^\ell} \exp\left\{ \boldsymbol{\alpha}^{\mathrm{H}}\mathbf{b} + \mathbf{b}^{\mathrm{H}}\boldsymbol{\alpha} - \boldsymbol{\alpha}^{\mathrm{H}}\mathbf{A}^{-1}\boldsymbol{\alpha} \right\} d\boldsymbol{\alpha} = (\pi)^\ell |\mathbf{A}| \exp\left\{ \mathbf{b}^{\mathrm{H}}\mathbf{A}\mathbf{b} \right\}, \tag{34}$$

which holds for any positive-definite matrix $\mathbf{A}$ and vector $\mathbf{b}$ of proper sizes, after replacement $\mathbf{A} = \sigma_z^2(\mathbf{I}_\ell + \mu\boldsymbol{\Sigma})^{-1}$ and $\mathbf{b} = \widehat{\boldsymbol{\alpha}}^{\mathrm{fLBF}}(t)$, we arrive at

$$L(\boldsymbol{\Sigma}, \sigma_z^2, \mu) = \frac{(\pi\mu)^\ell |\boldsymbol{\Sigma}|_+}{(\pi\sigma_z^2)^M |\mathbf{I}_\ell + \mu\boldsymbol{\Sigma}|} \exp\left\{ -\frac{\zeta(t; \boldsymbol{\Sigma}, \mu)}{\sigma_z^2} \right\}, \tag{35}$$

where $M = Kn + \ell$ and

$$\zeta(t; \boldsymbol{\Sigma}, \mu) = \sum_{j=-k}^{k} ||\widetilde{\boldsymbol{\theta}}(t+j)||^2 - [\widehat{\boldsymbol{\alpha}}^{\mathrm{fLBF}}(t)]^{\mathrm{H}}[\mathbf{I}_\ell + \mu\boldsymbol{\Sigma}]^{-1}\widehat{\boldsymbol{\alpha}}^{\mathrm{fLBF}}(t) \tag{36}$$

denotes the residual sum of squares. Good [32] referred to the maximization of (35) as a type II maximum likelihood (ML) method, but recently it has been more frequently referred to as empirical Bayes approach [28,30].

Since the ML estimate of the variance $\sigma_z^2$ found from (35) is given by $\widehat{\sigma}_z^2 = \zeta(t; \boldsymbol{\Sigma}, \mu)/M$, the optimal value of the regularization parameter $\mu$ can be obtained by maximizing the concentrated likelihood function $L(\boldsymbol{\Sigma}, \widehat{\sigma}_z^2, \mu)$, or equivalently by minimizing the quantity

$$-\log L(\boldsymbol{\Sigma}, \widehat{\sigma}_z^2, \mu) = \mathrm{const} + M\log\widehat{\sigma}_z^2 - \ell\log\mu + \log|\mathbf{I}_\ell + \mu\boldsymbol{\Sigma}|. \tag{37}$$

Note that

$$|\mathbf{I}_\ell + \mu\boldsymbol{\Sigma}| = \left(\prod_{i\in S}(1 + \mu\gamma_i f_0^2)\right)\left(\prod_{i\in\bar{S}}(1 + \mu\gamma_i \mathbf{f}_0^{\mathrm{H}}\mathbf{f}_0)\right),$$

which leads to the following estimate of the optimal regularization parameter $\mu$:

$$\widehat{\mu}(t) = \arg\min_{\mu} \left\{ M\log\zeta(t; \boldsymbol{\Sigma}, \mu) - \ell\log\mu + \sum_{i\in S}\log(1 + \mu\gamma_i f_0^2) + \sum_{i\in\bar{S}}\log(1 + \mu\gamma_i \mathbf{f}_0^{\mathrm{H}}\mathbf{f}_0) \right\}. \tag{38}$$

**Remark 2.** The regularization described in Section 5.1 is based on the knowledge of the power delay profile $\{\sigma_i^2\}$ for $i = 1, \ldots, n$. In practice, such information is usually unavailable. However, when performing the regularization according to (28) and (29), the fLBF estimates $\widehat{\theta}_i^{\mathrm{fLBF}}(t)$ are available, and we can use $|\widehat{\theta}_i^{\mathrm{fLBF}}(t)|^2$ for approximation of $\sigma_i^2$. Since the regularization matrix $\boldsymbol{\Sigma}$ in (26) is now data-dependent, the decision rule which parallels (38) can only be regarded as "pseudo-Bayesian".

The modified scheme described above bears some resemblance to the LASSO (least absolute shrinkage and selection operator) approach [33]. The well-known property of LASSO estimators, which are based on the $L_1$ regularization, is their ability to discard (shrink to zero) insignificant components of the estimated vector of parameters. Although the proposed fRLBF scheme does not share this property with LASSO, it makes a step in this direction - note that very small values of the estimates $\widehat{\theta}_i^{\mathrm{fLBF}}(t)$ make the weights $\gamma_i(t) = 1/|\widehat{\theta}_i^{\mathrm{fLBF}}(t)|^2$ in $\boldsymbol{\Gamma} = \mathrm{diag}\{\gamma_1, \ldots, \gamma_n\}$ very large, which shrinks the corresponding fRLBF estimates towards zero.

**Remark 3.** Evaluation of $\zeta(t; \boldsymbol{\Sigma}, \mu)$ is not computationally demanding. Observe that

$$\zeta(t; \boldsymbol{\Sigma}, \mu) = \sum_{i=1}^{n} \zeta_i(t; \boldsymbol{\Sigma}, \mu)$$

and, for $i \in \bar{S}$,

$$\zeta_i(t; \boldsymbol{\Sigma}, \mu) = \sum_{j=-k}^{k} |\widetilde{\theta}_i(t+j)|^2 - \frac{|\widehat{\alpha}_i^{\mathrm{fLBF}}(t)|^2}{1 + \mu\gamma_i f_0^2}, \tag{39}$$

whereas, for $i \in \bar{S}$,

$$\zeta_i(t; \boldsymbol{\Sigma}, \mu) = \sum_{j=-k}^{k} |\widetilde{\theta}_i(t+j)|^2 - [\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t)]^{\mathrm{H}}[\mathbf{I}_m + \mu\gamma_i \mathbf{f}_0\mathbf{f}_0^{\mathrm{H}}]^{-1}\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t). \tag{40}$$

Finally note that, using the matrix inversion lemma, the second term on the right hand side of (40) can be rewritten in the form more suitable for implementation:

$$\begin{aligned} &[\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t)]^{\mathrm{H}}[\mathbf{I}_m + \mu\gamma_i \mathbf{f}_0\mathbf{f}_0^{\mathrm{H}}]^{-1}\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t) \\ &= [\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t)]^{\mathrm{H}}\left[\mathbf{I}_m - \frac{\mu\gamma_i \mathbf{f}_0\mathbf{f}_0^{\mathrm{H}}}{1 + \mu\gamma_i \mathbf{f}_0^{\mathrm{H}}\mathbf{f}_0}\right]\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t) \\ &= ||\widehat{\boldsymbol{\alpha}}_i^{\mathrm{fLBF}}(t)||^2 - \frac{\mu\gamma_i}{1 + \mu\gamma_i \mathbf{f}_0^{\mathrm{H}}\mathbf{f}_0}|\widehat{\theta}_i^{\mathrm{fLBF}}(t)|^2. \end{aligned} \tag{41}$$

**Remark 4.** So far we assumed that the same regularization parameter $\mu$ is applied to all system parameters. However, with the preestimation, there is more freedom to choose the way the regularization is carried out. For example, if the regularization is restricted only to parameters that vary with time, i.e., to parameters that belong to the set $\{\theta_i(t), i \in \bar{S}\}$, the optimization rule for the regularization parameter can be easily deduced from the general rule, and it takes the form

$$\widehat{\mu}(t) = \arg\min_{\mu} \left\{ M\log\zeta(t; \boldsymbol{\Sigma}, \mu) + \sum_{i\in\bar{S}}\log\frac{1 + \mu\gamma_i \mathbf{f}_0^{\mathrm{H}}\mathbf{f}_0}{\mu} \right\}, \tag{42}$$

where $M = (K - m + 1)n_{\bar{S}}$, $\zeta(t; \boldsymbol{\Sigma}, \mu) = \sum_{i\in\bar{S}} \zeta_i(t; \boldsymbol{\Sigma}_i, \mu)$ and $\boldsymbol{\Sigma}_i = \gamma_i \mathbf{f}_0\mathbf{f}_0^{\mathrm{H}}$.

Finally, we note that the regularization can also be carried out independently for each system parameter. In this case, one can select $\gamma_1 = \ldots = \gamma_n = 1$ (since relative weighting does not apply), which leads to the following decentralized optimization rule

$$\widehat{\mu}_i(t) = \arg\min_{\mu} J_i(t;\mu), \tag{43}$$

where

$$J_i(t;\mu) = \begin{cases} K\log\zeta_i(t; f_0^2, \mu) + \log\frac{1+\mu f_0^2}{\mu} & \text{if } i \in S \\ (K-m+1)\log\zeta_i(t; \mathbf{f}_0\mathbf{f}_0^{\mathrm{H}}, \mu) & \\ \quad +\log\frac{1+\mu\mathbf{f}_0^{\mathrm{H}}\mathbf{f}_0}{\mu} & \text{if } i \in \bar{S} \end{cases} \tag{44}$$

**Remark 5.** So far we assumed that the number of basis functions $m$ and the analysis window width $K$ were set prior to identification. Since the choice of these parameters may have a strong influence on identification/tracking results, they should be selected with caution. In general, better performance can be expected for larger values of $m$ at the expense of increased decision delay (latency) $k$. For the UWA channels considered in this paper, we use $K \approx 150m + 1$.

The optimal values of $m$ and $K$ depend on the type and speed of parameter variations. For each analysis window location, optimization of $m$ and $K$ can be carried out using the parallel estimation approach described in [15]. In this framework, several identification algorithms, equipped with different settings, are run concurrently yielding competitive fLBF estimates. At each time instant $t$ the best-fitting estimate is chosen using the cross-validation decision rule proposed in [15]. The selected estimate may then be regularized in the way described above.

## 6. Debiasing

It was observed that the estimated parameter trajectory, obtained using the fLBF approach, lags behind the true parameter trajectory, and that the size of this delay depends on the forgetting factor $\lambda$ used in the preestimation stage. This effect, which becomes pronounced as $\lambda$ approaches 1 (this is the case when the number of estimated parameters becomes large) is evidently caused by the estimation delay feature of the EWLS algorithm used to generate preestimates. As shown in Niedźwiecki [25], when the sequence $\{\boldsymbol{\varphi}(t)\}$ is stationary and persistently exciting, the mean path of the steady-state EWLS estimates can be approximately viewed as an output of a linear time-invariant lowpass filter of the form

$$H_{\mathrm{EWLS}}(z^{-1}) = \frac{1-\lambda}{1-\lambda z^{-1}}$$

excited by the process $\{\boldsymbol{\theta}(t)\}$. The consequence of this fact is that when system parameters vary slowly with time, the mean path of EWLS estimates can be considered, to some extent, a delayed version of the true trajectory, i.e.,

$$\mathrm{E}\left[\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t)\right] \cong \boldsymbol{\theta}(t-\Delta),$$

where

$$\Delta = \mathrm{int}\left[\frac{\lambda}{1-\lambda}\right]$$

denotes the nominal (low-frequency) delay of the filter $H_{\mathrm{EWLS}}(z^{-1})$, and the expectation is taken with respect to $\{\boldsymbol{\varphi}(t)\}$ and $\{e(t)\}$; int$[x]$ denotes the integer closest to $x$.

Since the time shift between $\widehat{\boldsymbol{\theta}}^{\mathrm{EWLS}}(t)$ and $\boldsymbol{\theta}(t)$, which is "inherited" by the fLBF scheme, depends on the frequency distribution of $\{\boldsymbol{\theta}(t)\}$ (usually unknown), the nominal delay $\Delta$ is only an approximation of the true delay. Additionally, the true delay may vary with time should the spectral content of $\{\boldsymbol{\theta}(t)\}$ change. For this reason we propose a simple adaptive scheme for minimization of the time shift between $\widehat{\boldsymbol{\theta}}^{\mathrm{fLBF}}(t)$ and $\boldsymbol{\theta}(t)$. The search will be carried out around $\Delta$.

Let

$$\varepsilon_\delta(t) = y(t) - [\widehat{\boldsymbol{\theta}}^{\mathrm{fLBF}}(t+\delta)]^{\mathrm{H}}\boldsymbol{\varphi}(t) \tag{45}$$

and $D = [\Delta - \delta_0, \Delta + \delta_0]$. Define the exponentially weighted sum of squares of $\varepsilon_\delta(t)$ evaluated recursively for every $t$ and every $\delta \in D$

$$J(t,\delta) = \lambda_0 J(t-1,\delta) + |\varepsilon_\delta(t)|^2, \;\; 0 < \lambda_0 < 1.$$

The (approximately) debiased fLBF estimates can be obtained using the formula

$$\widehat{\boldsymbol{\theta}}^{\mathrm{dfLBF}}(t) = \widehat{\boldsymbol{\theta}}^{\mathrm{fLBF}}(t + d(t)), \tag{46}$$

where

$$d(t) = \arg\min_{\delta \in D} J(t,\delta). \tag{47}$$

Note that time-shifting fLBF estimates by a constant amount of time $\delta$ does not affect the variance of the mean-square parameter estimation error $\mathrm{E}[||\widehat{\boldsymbol{\theta}}^{\mathrm{fLBF}}(t+\delta) - \boldsymbol{\theta}(t)||^2]$. This means that the proposed debiasing procedure can reduce the bias of fLBF estimates *without* increasing their variance (which, by the way, is quite unusual in system identification). Hence, application of this procedure can only improve the MSE score of the corrected estimates. This debiasing procedure can also be applied to the regularized fLBF estimates to further improve the identification performance.

## 7. Complexity of proposed algorithms

The complexity of the EWLS and proposed fLBF algorithms at every sample are summarized in Table 1, where $n_d = 2\delta_0$ is the number of delays included in the delay search, $n_\mu$ is the number of μ values considered in the range of search for the regularization parameter, $n_{\bar{S}}$ and $n_S$ are the cardinality of $\bar{S}$ and $S$, respectively. For the numerical simulations in this work, we use the classical recursive EWLS algorithm which has a complexity of $\mathcal{O}(n^2)$ MACs. Note that the EWLS algorithm can also be implemented by fast transversal filtering (FTF) algorithm [9], which offers a lower complexity of $\mathcal{O}(n)$ MACs. If the FTF algorithm is used, the complexity of the fLBF algorithm will be reduced to $\mathcal{O}(mn)$ MACs.

In Section 8, we will perform quantitative analysis and show the complexity in number of real-valued MAC operations for specific UWA communication scenarios.

## 8. Numerical and experimental results

In this section, we consider the following UWA communication experiments:

- Numerical full-duplex (FD) experiment;
- Real FD experiment;
- Two numerical experiments with communication between a static transmitter and a static/moving receiver.

### 8.1. Numerical UWA FD experiment

Here, we investigate the mean square deviation (MSD) performance of the following fLBF algorithms:

- *fLBF*: the original fLBF algorithm, when all taps are assumed to be time-varying;
- *Oracle fLBF*: the fLBF algorithm as described in Section 3, i.e., with perfect knowledge of taps being static or time-varying;

**Table 1**
Complexity of the adaptive filters per sample (analytical expressions).

| Algorithm | MAC | Division | log(·) |
|---|---|---|---|
| EWLS | $7n^2+12n$ | 1 | |
| fLBF | $7n^2+16n+16mn$ | 1 | |
| Adaptive fLBF | $7n^2+16n+16mn_{\bar{S}}+8n_S$ | 1 | |
| Adaptive fLBF with debiasing | $7n^2+(16+4n_d)n+16mn_{\bar{S}}+8n_S$ | 2 | |
| Adaptive fRLBF with debiasing | $7n^2+(18+4n_d)n+(18m+3)n_{\bar{S}}+13n_S+2n_\mu$ | $2n+1$ | $n_\mu n$ |

- *Adaptive fLBF*: the fLBF algorithm with adaptive selection of static/time-varying taps as described in Section 4;
- *Adaptive fRLBF*: regularized version of the adaptive fLBF algorithm with optimal regularization parameter μ in (28)-(29), obtained via numerical search or with μ adaptively estimated according to (38);
- *Adaptive dfLBF* and *Adaptive dfRLBF (recommended)*: fLBF algorithms with debiasing according to (46), adaptive selection of static/time-varying taps according to (24), and (in the case of the dfRLBF algorithm) adaptive tuning of μ using (38).

The algorithms were chosen so as to demonstrate the rate of improvement provided by different components of the final (recommended) solution, namely: adaptive selection of static/time-varying taps, adaptive regularization, and debiasing. Our intention was also to show how adaptive procedures compare with the "ground truth" based ones (known localization of static taps, optimally chosen regularization gain). The adaptive dfRLBF algorithm, which combines all techniques mentioned above, constitutes the main result of this paper.

The MSD performance of the EWLS and TU-RLS algorithms [15] are used as benchmarks for comparison with the performance of the fLBF algorithms, where the TU-RLS algorithm is considered as the state-of-the-art channel estimator in UWA communications [15,34].

The MSD at time instant $t$ is defined as:

$$\mathrm{MSD}(t)=\frac{||\widehat{\boldsymbol{\theta}}(t)-\boldsymbol{\theta}(t)||^2}{||\boldsymbol{\theta}(t)||^2}, \tag{48}$$

where $\widehat{\boldsymbol{\theta}}(t)$ is an estimate of $\boldsymbol{\theta}(t)$. The MSD in (48) is then averaged in time (in this experiment, over an interval of 5000 samples, i.e., 5 s for the assumed sampling rate 1 kHz) and over 50 simulation trials. In every simulation trial, new realizations of the channel impulse response, noise and regressor are used.

Fig. 1 shows the SI channel impulse response measured in the real FD experiment (see Section 8.2 for details). To mimick the channel impulse response in our numerical FD experiment, the SI channel is modelled as follows. The SI channel length is $n=80$. It contains strong taps, which are slowly varying in time; these are reflections from stationary parts of the experimental equipment and the lake bottom. It also contains fast-varying taps associated with reflections from the lake surface. Based on the analysis of the experimental impulse response, we defined the power delay profile of the SI channel as shown in Fig. 2. Most of the channel taps are static, but there are 5 taps, which are time-varying with the highest frequencies in their spectra also shown in Fig. 2. In the simulation trials, the time-varying taps are modelled as realizations of independent random Gaussian processes with uniform power spectral densities within the frequencies shown in Fig. 2. The complex amplitudes of the static taps are random Gaussian variables independent for different taps. The tap variances are defined by the power delay profile in Fig. 2. In this experiment, the transmitted baseband signal is a sequence of independent zero-mean complex-valued Gaussian random numbers of a unit variance. The signal (self-interference) to noise ratio is around 60.4 dB, the same as in the real experiment described in Section 8.2.

**Table 2**
The best MSD performance of the EWLS, TU-RLS and fLBF algorithms in the numerical FD experiment.

| Algorithm | MSD, dB | | | |
|---|---|---|---|---|
| | | m = 3 | m = 5 | m = 7 |
| EWLS | −49.4 | | | |
| TU-RLS | −49.4 | | | |
| fLBF | | −55.2 | −55.4 | −55.5 |
| Adaptive fLBF | | −57.0 | −57.6 | −57.9 |
| Oracle fLBF | | −57.2 | −57.9 | −58.2 |

**Table 3**
The best MSD performance of the adaptive fLBF and adaptive fRLBF algorithms with 'optimal' and adaptive regularization in the numerical FD experiment.

| Algorithm | MSD, dB | | |
|---|---|---|---|
| | m = 3 | m = 5 | m = 7 |
| Adaptive fLBF | −57.0 | −57.6 | −57.9 |
| Adaptive fRLBF ('optimal' μ) | −57.5 | −57.8 | −58.1 |
| Adaptive fRLBF (adaptive μ) | −57.5 | −57.8 | −58.1 |

It was found that for the FD scenario, the best MSD performance provided by the EWLS and TU-RLS algorithms is −49.4 dB, which is achieved at $\lambda=0.96$ and a step size of $\eta=10^{-5}$ (for the TU-RLS algorithm). For preestimation in fLBF algorithms, the forgetting factor is $\lambda=1-2/n=0.975$ in agreement with (16). Fig. 3 shows the performance of the fLBF algorithms against the length $K$ of the analysis window and the number of basis functions $m$. The MSD performance of the original fLBF algorithm (−55.2 dB, −55.4 dB, −55.5 dB for $m=3,5,7$, respectively) is about 6 dB better than the performance of the EWLS and TU-RLS algorithms. Although the performance of the original fLBF algorithm improves when using a higher number of basis functions, the improvement is not significant. Note that a higher $m$ requires a higher $K$ for optimal performance (for a given $m$, $K_{\mathrm{opt}}\approx 150m+1$), and thus the complexity of the implementation increases. It can be seen that taking into account the fact that only a few taps are time-varying, the oracle fLBF algorithm outperforms the original fLBF algorithm by about 2 dB, 2.5 dB and 2.7 dB for $m=3,5,7$, respectively. The adaptive fLBF algorithm with the time-invariance test proposed in Section 4 closely approaches the performance of the oracle fLBF algorithm. For the FD scenario, a high estimation accuracy is required. Therefore, we use a relatively small significance level $\eta_0$ for the time-invariance test to reduce the possibility of identifying a time-varying path as a static path. It is found that $\eta_0=0.005$ works very well in this scenario. Based on this value $\eta_0$ and using (22), the parameter $q_0$ is computed as shown in captions of Fig. 3, as well as Figs. 4 and 5 below. It can be concluded that the adaptive time-invariance test works very well and the adaptive fLBF algorithm significantly improves the MSD performance compared to the EWLS and TU-RLS algorithms (about 8.4 dB for $m=7$).

Fig. 4 shows the performance of the adaptive fLBF algorithm and adaptive fRLBF algorithm with 'optimal' and adaptive regularization. The 'optimal' regularization scheme uses a fixed value of μ (shown in the legend of Fig. 4), which was found to minimise the MSD. It can be seen that the adaptive fLBF algorithms with regular-

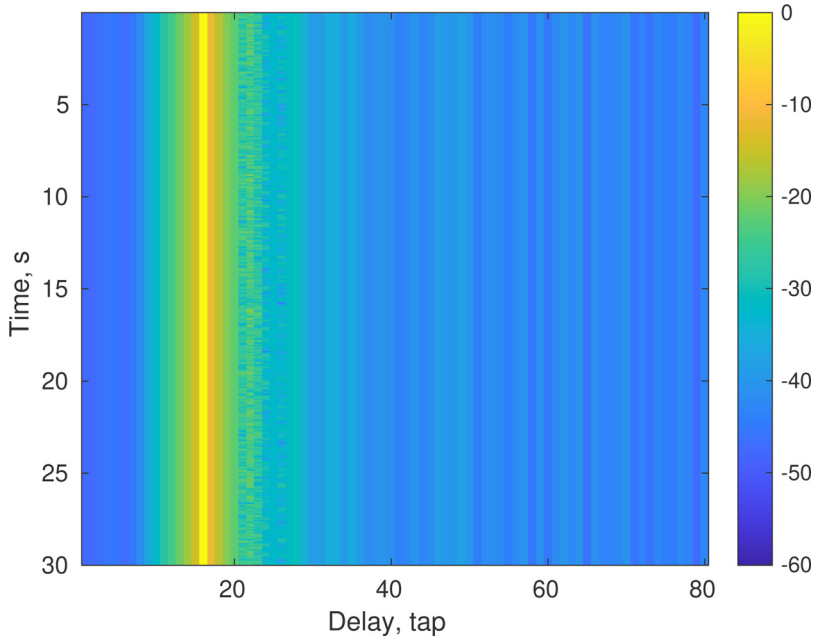


**Fig. 1.** Normalized amplitude (in dB) of the impulse response measured in the real FD experiment.

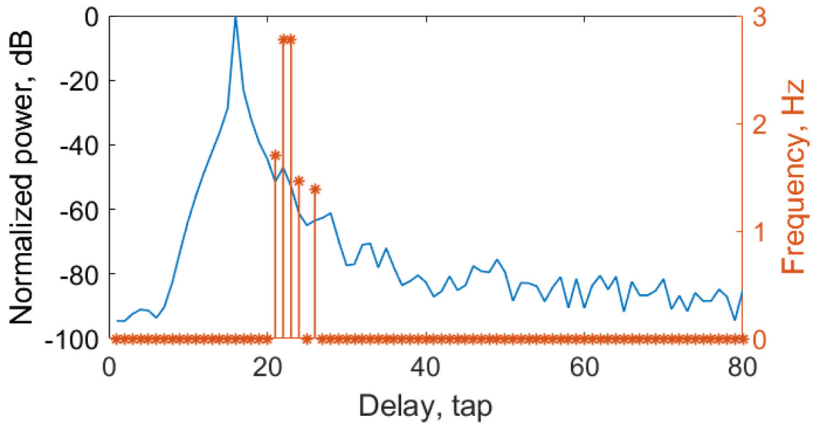


**Fig. 2.** The power delay profile and tap highest frequencies used in the numerical FD experiment.

**Table 4**
The best MSD performance of adaptive fLBF and adaptive fRLBF algorithms with the debiasing in the numerical FD experiment.

| Algorithm | MSD, dB | | |
|---|---|---|---|
| | m = 3 | m = 5 | m = 7 |
| Adaptive dfLBF | −60.6 | −62.4 | −63.4 |
| Adaptive dfRLBF (adaptive μ) | −62.0 | −63.2 | −63.9 |

**Table 5**
Complexity of the adaptive filters per sample .

| Algorithm | MACs |
|---|---|
| EWLS | $4.6 \times 10^4$ |
| fLBF and Adaptive fLBF | $5.5 \times 10^4$ |
| Adaptive fLBF with debiasing | $7.5 \times 10^4$ |
| Adaptive fRLBF with debiasing | $7.6 \times 10^4$ |

ization (either 'optimal' or adaptive) provide further improvement of the MSD performance against the adaptive fLBF algorithm. Another conclusion is that similar MSD performance is achieved with both 'optimal' and adaptive regularization. This demonstrates the effectiveness of the proposed adaptive regularization scheme.

We now show that the debiasing technique proposed in Section 6 allows significant improvement in the performance of the fLBF algorithms. Fig. 5 shows the MSD performance of the debiased adaptive fLBF (adaptive dfLBF) and debiased adaptive fRLBF algorithms (adaptive dfRLBF). By comparing these results with results shown in Fig. 4, it can be seen that the debiasing significantly improves the performance of the adaptive fLBF and adaptive fRLBF algorithms (when $m = 7$) by 5.5 dB and 5.8 dB, respectively.

With the parameters used in this scenario (for $m = 7$), the complexity of the proposed algorithms per sample are summarized in Table 5. For the algorithms with adaptive identification of static and time-varying parameters, we consider the worst-case scenario by assuming that all the parameters in the channel are time-varying. It can be seen that the complexity of the (most complicated) adaptive fRLBF algorithm with debiasing is not much higher than that of the original fLBF algorithm.

### 8.2. Real UWA FD experiment

In the real FD experiment, the true impulse response is unavailable. We therefore investigate the SI cancellation (SIC) perfor-

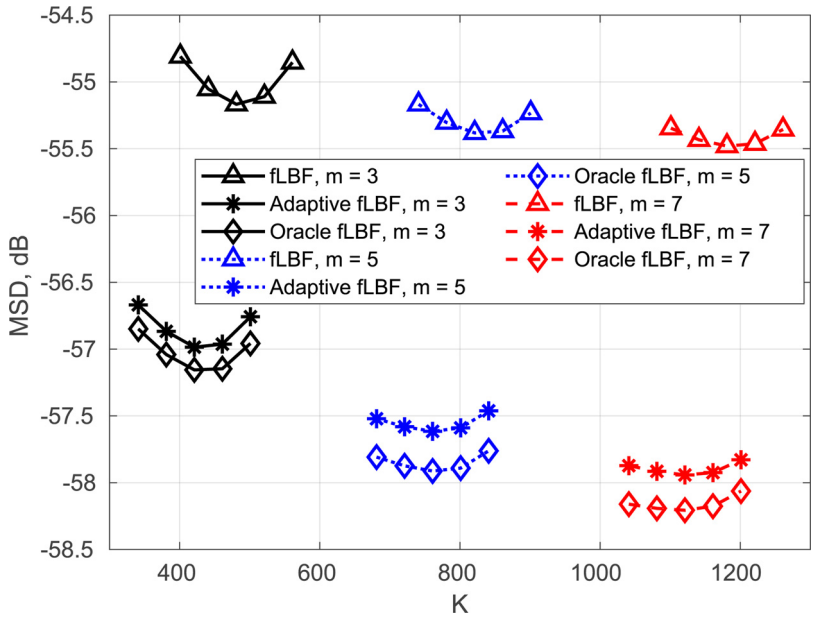


**Fig. 3.** The MSD performance of the oracle fLBF and adaptive fLBF algorithms against the window length $K$ and the number of basis functions $m$ in the numerial FD experiment. Simulated version of channel in the FD experiment; SNR= 60.4 dB; $\lambda = 0.975$. For the adaptive invariance test, we set $q_0 = 0.43$ K, 0.45 K, 0.46 K for $m = 3, 5, 7$, respectively. The best MSD performance achieved by the aforementioned algorithms is summarized in Table 2.

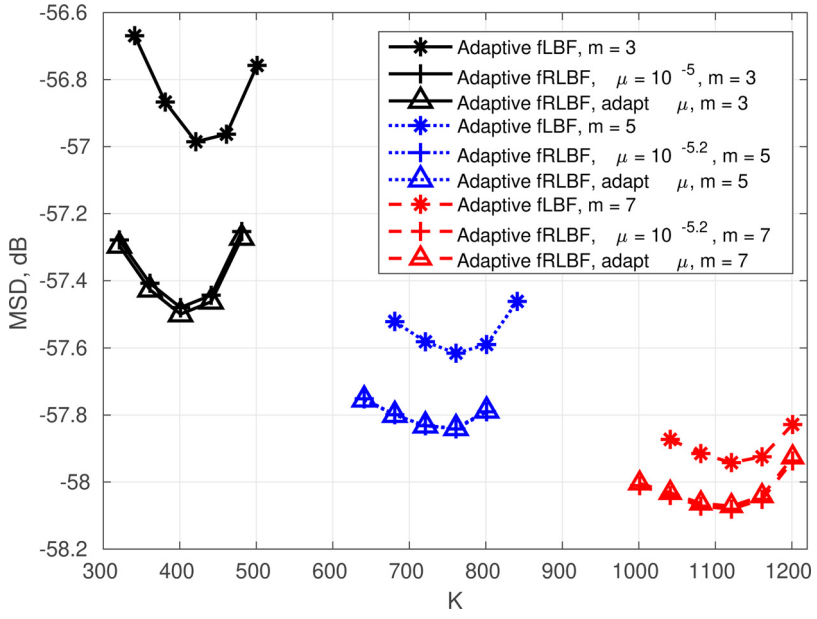


**Fig. 4.** The MSD performance of the adaptive fLBF and adaptive fRLBF algorithms with 'optimal' and adaptive regularization in the numerical FD experiment; SNR= 60.4 dB; $\lambda = 0.975$. For the adaptive invariance test, we set $q_0 = 0.43$ K, 0.45 K, 0.46 K for $m = 3, 5, 7$, respectively. The range of search for the adaptive regularization parameter [ is $\{10^{(-6+0.2a)}, a = 0, 1, \ldots, 10\}$. The best MSD performance achieved by the adaptive fLBF algorithms with 'optimal' and adaptive regularization is summarized in Table 3.

mance measured using the SIC factor, which shows how much the signal-to-interference ratio (SIR) at the output of the SI canceller is reduced compared to the SIR at the input of the canceller; the methodology of measuring the SIC factor is described in Shen et al. [7]. We consider the following fLBF algorithms:

- *fLBF*;
- *Adaptive fLBF*;
- *Adaptive fRLBF*;
- *Adaptive dfLBF* and *Adaptive dfRLBF (recommended)*.

The FD experiment was conducted in a lake of depth 8 m. The distance between the near-end transmitter and receiver, both positioned at a depth of 4 m, is 7 cm. We are interested in the near-end SIC performance, and the far-end transmission is not considered in this experiment. In the experiment, binary-shift keying (BPSK) symbols are transmitted with a rate of 1000 symbols/s at the carrier frequency 32 kHz; a root raised cosine filter with a roll-off factor of 0.2 is used for the pulse shaping [35]. The received signal after analog-to-digital conversion is down shifted in frequency, low-pass filtered and down sampled to the sampling rate 1 kHz. These samples are applied to the adaptive filter as the desired signal. The same operation is performed on the analogue signal applied to the transmit antenna [36]; these samples are used as the regressor in the adaptive filter.

In the experiment, the self-interference to noise ratio is 60.4 dB. The SIC factor is computed over a 10 s interval after the conver-

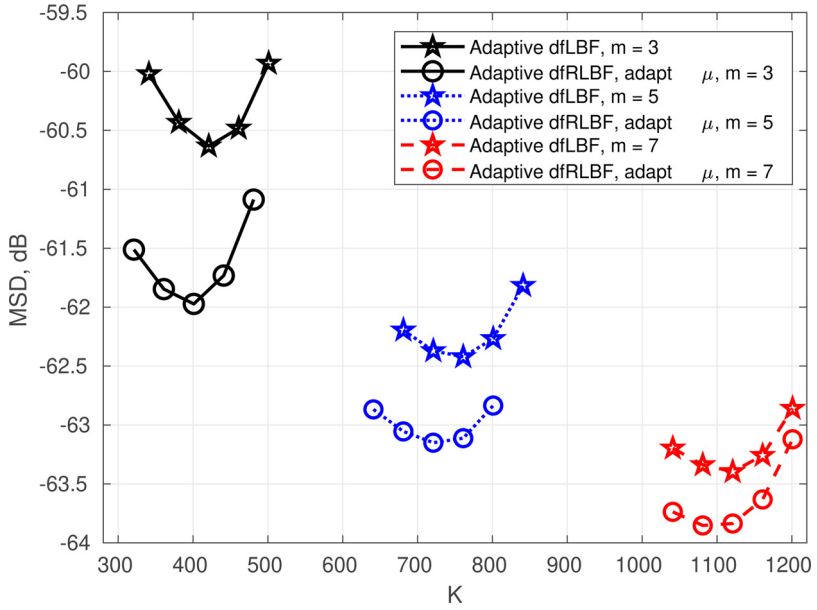


**Fig. 5.** The MSD performance of the adaptive fLBF and adaptive fRLBF algorithms with the debiasing in the numerical FD experiment; SNR = 60.4 dB; $\lambda = 0.975$. For the adaptive invariance test, we set $q_0 = 0.43$ K, 0.45 K, 0.46 K for $m = 3, 5, 7$, respectively. The range of search for the adaptive regularization parameter [ is $\{10^{(-6+0.2a)}, a = 0, 1, \ldots, 10\}$. The corresponding nominal delay is equal to $\Delta = 39$, $\lambda_0$ was set to 0.98 and $\delta_0 = 30$. The best MSD performance achieved by the adaptive fLBF algorithms with debaising is summarized in Table 4.

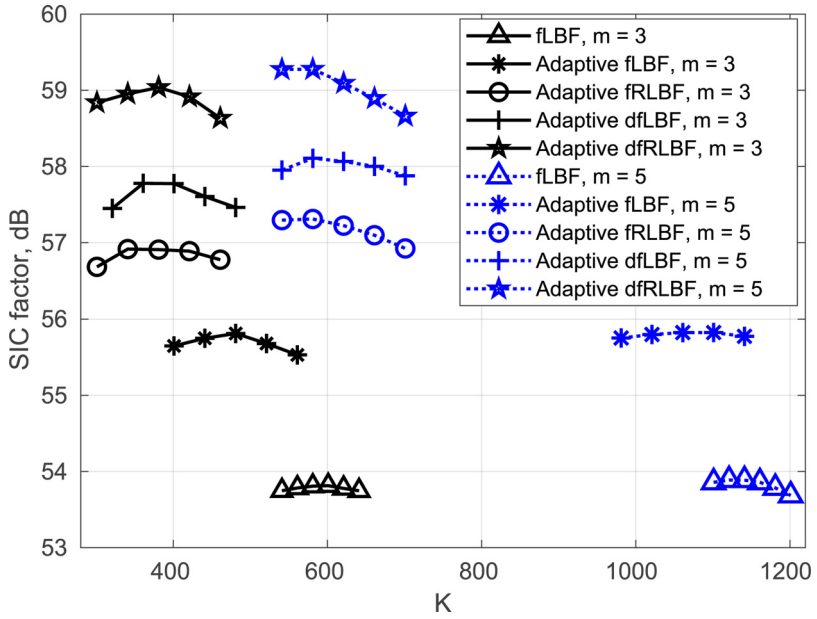


**Fig. 6.** The SIC performance of the proposed fLBF algorithms in cancelling the self-interference in the real UWA FD experiment. For the adaptive invariance test, we set $q_0 = 0.45$ K for $m = 3, 5$. The range of search for the adaptive regularization parameter μ is $\{10^{(-6+0.2a)}, a = 0, 1, \ldots, 10\}$. The corresponding nominal delay is equal to $\Delta = 39$, $\lambda_0$ was set to 0.98 and $\delta_0 = 30$. The best SIC performance achieved by the EWLS, TU-RLS, and proposed fLBF algorithms is summarized in Table 6.

gence of the adaptive filter (see more details on the procedure in Shen et al. [7]). The SIC factor is measured as the improvement in the signal-to-interference (SIR) ratio due to the SI cancellation. Note that for the FD operation, the higher SIC factor the better.

When applying the EWLS algorithm to the experimental data, the highest achievable SIC factor is 50.9 dB. Similar performance is achieved by the TU-RLS algorithm at $\lambda = 0.975$ and a step size of $\eta = 10^{-4}$. Fig. 6 shows the SIC factor achieved by the fLBF adaptive algorithms. The original fLBF algorithm provides a SIC factor of 53.8 dB and 53.9 dB for $m = 3, 5$, respectively. When using the adaptive selection of time-varying taps (adaptive fLBF), the SIC factor increases to 55.8 dB for $m = 3, 5$. The regularization scheme

**Table 6**
The best SIC performance of the EWLS, TU-RLS and fLBF algorithms in the real UWA FD experiment.

| Algorithm | SIC, dB | |
|---|---|---|
| | m = 3 | m = 5 |
| EWLS | 50.9 | |
| TU-RLS | 50.9 | |
| fLBF | 53.8 | 53.9 |
| Adaptive fLBF | 55.8 | 55.8 |
| Adaptive fRLBF | 56.9 | 57.3 |
| Adaptive dfLBF | 57.8 | 58.1 |
| Adaptive dfRLBF | 59.0 | 59.3 |

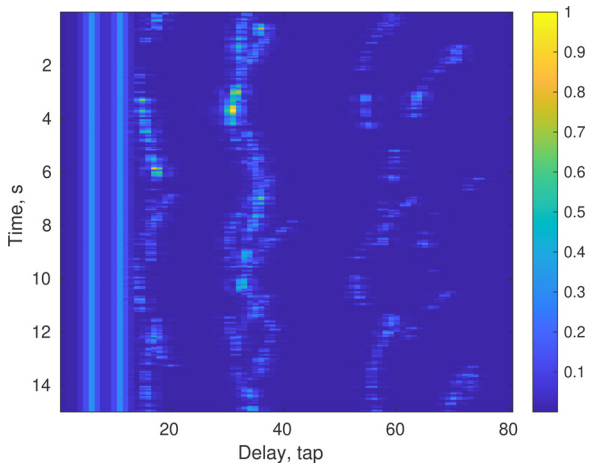


**Fig. 7.** The normalized magnitude of the time-varying impulse response in the UWA communication scenario with static transmitter and receiver.

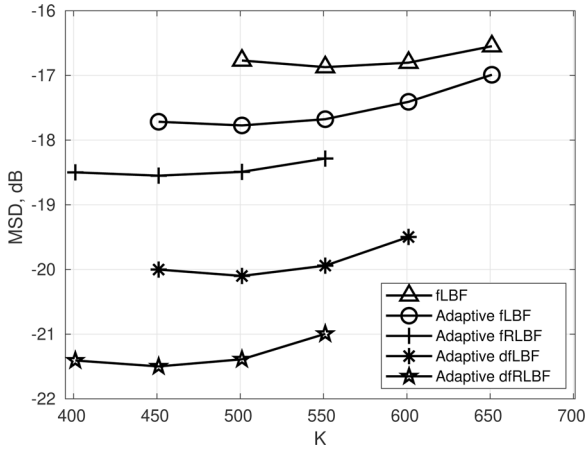


**Fig. 8.** The MSD performance of the fLBF algorithms ($m = 3$) in the UWA communication scenario with static transmitter and receiver in the Korean Sea environment; SNR= 20 dB; $\lambda = 0.975$. For the adaptive invariance test, we set $q_0 = 0.46$ K. The range of search for the adaptive regularization parameter μ is $\{10^{(-2+0.25a)}, a = 0, 1, \ldots, 10\}$. The corresponding nominal delay is equal to $\Delta = 39$, $\lambda_0$ was set to 0.98 and $\delta_0 = 30$. The best MSD performance of the EWLS, TU-RLS and fLBF algorithms in the UWA communication scenario with static transmitter and receiver is summarized in Table 7.

with adaptive computation of the regularization parameter μ further increases the SIC factor to 56.9 dB and 57.3 dB for $m = 3, 5$, respectively. When the debiasing scheme is applied, the SIC factor is increased to 58.1 dB and 59.3 dB (for $m = 5$) for the adaptive fLBF and fRLBF algorithms, respectively. It is interesting that with a higher number of basis functions $m$ the optimal window length $K$ is reduced. This happens because such processing shifts the balance between the bias (approximation error) and variance (noise error) of the estimate towards being dominated by the bias. With the same basis functions (the same $m$), the bias is reduced for smaller $K$ due to better local approximation [8].

Thus, we can conclude that the proposed techniques, specifically the adaptive time-invariance test, adaptive regularization, and debiasing allow significant improvement in the SI cancellation performance in this FD lake experiment.

### 8.3. Numerical UWA communication experiments

The UWA communication scenario is implemented using the Waymark simulator [2,37], which models the acoustic signal transmission between a transmitter and receiver moving underwater.

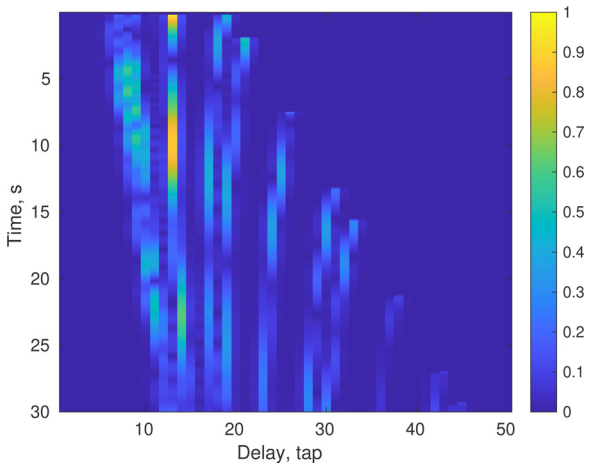


**Fig. 9.** The normalized magnitude of the time-varying impulse response in the UWA communication scenario with a static transmitter and moving receiver.

The propagation channel is represented as a time-varying linear filter, whose impulse response at every time instant is computed using the ray tracing implemented with the acoustic toolbox BELLHOP [38]. As a by-product, the Waymark simulator produces the time-varying baseband impulse response of the propagation channel with the Doppler correction corresponding to the dominant Doppler effect caused by the motion; this Doppler correction is typical at the front-end of UWA modems. We are using this time-varying impulse response for FIR filtering of a communication signal, which is modelled as a sequence of independent zero-mean complex-valued Gaussian random numbers of a unit variance, assumed to be transmitted at a carrier frequency of 12 kHz.

#### 8.3.1. Static transmitter and receiver

For this experiment, we consider an acoustic environment of a shallow sea (the sea depth is 20 m), where both the transmitter and receiver are static and placed at a depth of 10 m. The distance between the transmitter and receiver is 50 m. The generation of the time-varying surface waves is performed by the WAFO software toolbox [39] using the Pierson–Moscowitz spectrum [40] for a wind speed of 10 m/s. For our numerical experiment, we are using the acoustic environment of the test case 'Korean Sea' in Porter [38]. The duration of the transmission is 15 s. The symbol rate of the transmitted signal is 4 kHz and the signal is sampled at the symbol rate.

The time-varying channel impulse response of the communication scenario is shown in Fig. 7. The channel contains 80 taps, which is equivalent to a delay interval of 20 ms. It can be seen that there are two static paths, corresponding to the direct path and the bottom reflection. The rest of the multipath components are time-varying due to the moving sea surface.

The SNR in this scenario is 20 dB. As before, we use the performance of EWLS and TU-RLS algorithms as benchmarks. In this scenario, the best MSD performance of the EWLS algorithm is $-10.9$ dB, which is achieved at $\lambda = 0.965$. The best MSD performance of the TU-RLS algorithm is $-14.8$ dB, and is achieved with $\lambda = 0.945$ and $\eta = 0.007$; thus, the TU-RLS algorithm significantly outperforms the EWLS algorithm in this scenario. Fig. 8 shows the performance of the fLBF algorithms. As the improvement with a higher number of basis functions is not significant, we only show the results for three basis functions ($m = 3$). The original fLBF algorithm provides an MSD of $-16.9$ dB. As the SNR is not high in this scenario, a significance level of $\eta_0 = 0.05$ is used for the time-invariance test. Based on this value $\eta_0$ and using (22), the parameter $q_0$ is computed as shown in captions of Figs. 8 and 11 be-

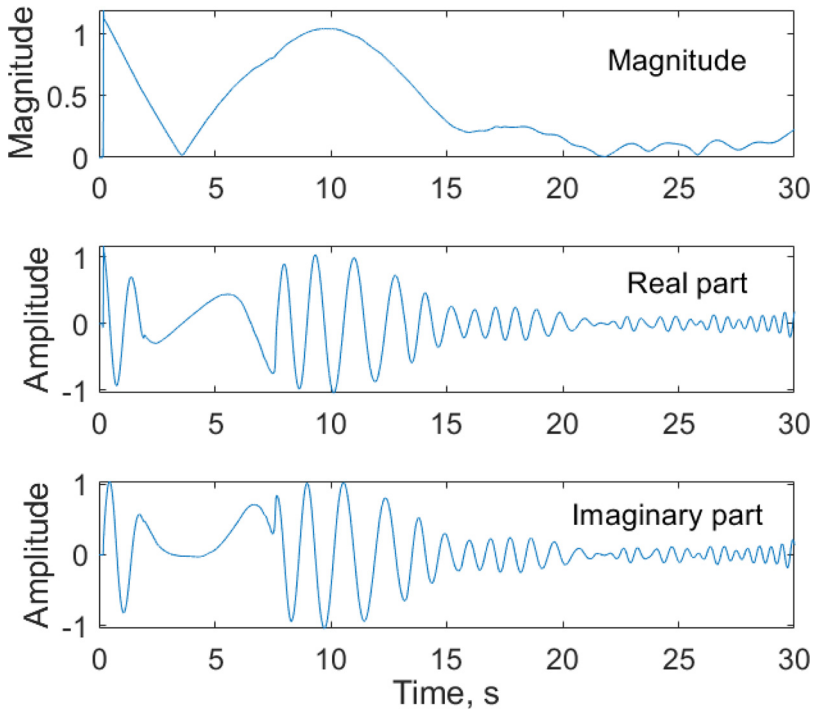


**Fig. 10.** The magnitude, real part and imaginary part of the 13th channel tap in the UWA communication scenario with static transmitter and moving receiver.

**Table 7**
The best MSD performance of the EWLS, TU-RLS and fLBF algorithms in the UWA communication scenario with static transmitter and receiver in the Korean Sea environment.

| Algorithm | MSD, dB |
|---|---|
| EWLS | −10.9 |
| TU-RLS | −14.8 |
| fLBF | −16.9 |
| Adaptive fLBF | −17.8 |
| Adaptive fRLBF | −18.6 |
| Adaptive dfLBF | −20.1 |
| Adaptive dfRLBF | −21.5 |

low. With the adaptive selection of static and time-varying taps, the MSD is reduced to −17.8 dB. It further reduces to −18.5 dB when adaptive regularization is used. Finally, with the debiasing, the MSD of the adaptive dfLBF and adaptive dfRLBF algorithm is further reduced to −20.1 dB and −21.5 dB, respectively. It can be concluded that the identification performance is significantly improved when using the debiased adaptive fRLBF algorithm, with 10.4 dB and 5.7 dB of improvement against the EWLS and TU-RLS algorithms, respectively.

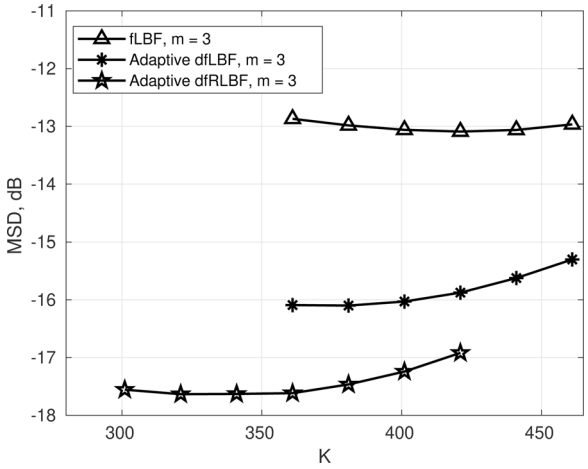


**Fig. 11.** The MSD performance of the fLBF algorithms ($m = 3$) in the UWA communication scenario with static transmitter and moving receiver in the Korean Sea environment; SNR = 15 dB; $\lambda = 0.96$. For the adaptive invariance test, we set $q_0 = 0.46$ K. The range of search for the adaptive regularization parameter μ is $\{10^{(-2+0.25a)}, a = 0, 1, \ldots, 10\}$. The corresponding nominal delay is equal to $\Delta = 24$, $\lambda_0$ was set to 0.98 and $\delta_0 = 20$. The best MSD performance of the EWLS, TU-RLS and fLBF algorithms in this scenario is summarized in Table 8.

*8.3.2. Static transmitter and moving receiver*

For this experiment, we consider a scenario with the receiver moving at a speed of 2.5 m/s away from the transmitter, starting from the distance 50 m. A flat sea surface is considered. The rest of the experimental setup is the same as in the previous experiment with the static transmitter and receiver. The duration of the transmission is 30 s. The symbol rate of the transmitted signal is 1 kHz and the signal is sampled at the symbol rate.

Fig. 9 shows the time-varying impulse response in this UWA communication scenario. It contains 50 taps, which is equivalent to a delay interval of 50 ms. It can be seen that the channel is very dynamic, in particular new multipath components appear as the receiver moves in the channel. It is also seen that the delays (taps) of multipath components vary in time, i.e., a tap moves into a neighbour tap, which is typical for UWA channels with moving transmitter/receiver. However, this plot does not reveal the real speed of the channel fluctuations. Fig. 10 shows the magnitude of the 13th channel tap and its real and imaginary parts. It can be seen that the real speed of the channel variation (in fact, the real and imaginary parts of the tap are estimated) is much faster than can be assumed from Fig. 9. Note that, in this scenario, there are no clear static multipaths.

In this scenario, we set the SNR in the received signal to SNR = 15 dB. It was found that, in this scenario, the best EWLS perfor-

**Table 8**
The best MSD performance of the EWLS, TU-RLS and fLBF algorithms in the UWA communication scenario with static transmitter and moving receiver in the Korean Sea environment.

| Algorithm | MSD, dB |
|---|---|
| EWLS | −6.8 |
| TU-RLS | −16.5 |
| fLBF | −13.1 |
| Adaptive dfLBF | −16.1 |
| Adaptive dfRLBF | −17.6 |

mance is MSD = −6.8 dB, which is achieved at $\lambda = 0.938$. The TU-RLS algorithm achieves an excellent performance of −16.5 dB with $\lambda = 0.96$ and $\eta = 0.002$. Fig. 11 shows the performance of fLBF algorithms with three basis functions ($m = 3$). The original fLBF algorithm provides an improvement of about 6.3 dB compared to the EWLS algorithm, and thus approaches the TU-RLS performance. An extra 3 dB improvement in the performance can be observed when we combine the adaptive selection of static and time-varying taps with the debiasing (adaptive dfLBF algorithm). Finally, we apply the debiasing to the adaptive fLBF algorithm with adaptive regularization, thus arriving at the adaptive dfRLBF algorithm, and obtain an extra improvement of 1.5 dB. Note that the adaptive dfRLBF algorithm outperforms the TU-RLS algorithm by 1.1 dB (Table 8).

## 9. Conclusions

We have considered the problem of identifying linear systems with a mix of time-invariant (or slowly varying) and time-varying parameters. We have further developed adaptive algorithms built on the LBF principle, in particular, its fast version, the fLBF algorithm. We have proposed an fLBF algorithm, which exploits the fact that only a part of the system parameters are time-varying. We have also proposed a simple statistical test to identify whether a particular parameter is static or time-varying. We have further proposed a regularized fLBF algorithm to exploit the sparsity in the parameters and an adaptive technique for estimation of the regularization parameter. Finally, we have proposed a debiasing technique that allows further improvement in the fLBF performance. The performance of the proposed techniques has been investigated in scenarios of UWA communications, using numerical and real experiments.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRediT authorship contribution statement

**Maciej Niedźwiecki:** Conceptualization, Writing – original draft. **Artur Gańcza:** Methodology, Software, Investigation, Writing – review & editing. **Lu Shen:** Software, Investigation, Validation, Writing – review & editing. **Yuriy Zakharov:** Methodology, Supervision, Writing – review & editing.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.sigpro.2022.108664

## References


[1] M. Siderius, M.B. Porter, Modeling broadband ocean acoustic transmissions with time-varying sea surfaces, J. Acoust. Soc. Am. 124 (1) (2008) 137–150.
[2] C. Liu, Y.V. Zakharov, T. Chen, Doubly selective underwater acoustic channelmodel for a moving transmitter/receiver, IEEE Trans. Veh. Technol. 61 (3) (2012) 938–950.
[3] P.A. van Walree, F.-X. Socheleau, R. Otnes, T. Jenserud, The watermarkbenchmark for underwater acoustic modulation schemes, IEEE J. Ocean. Eng. 42 (4) (2017) 1007–1018.
[4] M. Stojanovic, J. Preisig, Underwater acoustic communication channels:propagation models and statistical characterization, IEEE Commun. Mag. 47 (1) (2009) 84–89.
[5] Z. Wang, S. Zhou, J.C. Preisig, K.R. Pattipati, P. Willett, Clustered adaptation for estimation of time-varying underwater acoustic channels, IEEE Trans. Signal Process. 60 (6) (2012) 3079–3091.
[6] Y. Zakharov, B. Henson, R. Diamant, Y. Fei, P.D. Mitchell, N. Morozs, L. Shen, T.C. Tozer, Data packet structure and modem design for dynamic underwater acoustic channels, IEEE J. Ocean. Eng. 44 (4) (2019) 837–849.
[7] L. Shen, Y. Zakharov, B. Henson, N. Morozs, P.D. Mitchell, Adaptive filtering for full-duplex UWA systems with time-varying self-interference channel, IEEE Access 8 (2020) 187590–187604.
[8] L. Shen, Y. Zakharov, L. Shi, B. Henson, Adaptive filtering based on Legendre polynomials, TechRxiv (2020b). 10.36227/techrxiv.13084460
[9] S.S. Haykin, Adaptive Filter Theory, Prentice-Hall, 1996.
[10] A.H. Sayed, Fundamentals of Adaptive Filtering, John Wiley & Sons, 2003.
[11] Y.V. Zakharov, G.P. White, J. Liu, Low-complexity RLS algorithms using dichotomous coordinate descent iterations, IEEE Trans. Signal Process. 56 (7) (2008) 3150–3161.
[12] C. Qi, X. Wang, L. Wu, Underwater acoustic channel estimation based on sparse recovery algorithms, IET Signal Process. 5 (8) (2011) 739–747.
[13] Y.V. Zakharov, J. Li, Sliding-window homotopy adaptive filter for estimation of sparse UWA channels, in: 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), IEEE, 2016, pp. 1–4.
[14] G. Qiao, S. Gan, S. Liu, L. Ma, Z. Sun, Digital self-interference cancellation for asynchronous in-band full-duplex underwater acoustic communication, Sensors 18 (6) (2018) 1700.
[15] T.H. Eggen, A.B. Baggeroer, J.C. Preisig, Communication over Doppler spread channels. Part I: channel and receiver presentation, IEEE J. Ocean. Eng. 25 (1) (2000) 62–71.
[16] M. Niedźwiecki, M. Ciołek, Generalized Savitzky–Golay filters for identification of nonstationary systems, Automatica 108 (2019). Article 108477
[17] M. Niedźwiecki, M. Ciołek, A. Gańcza, A new look at the statistical identification of nonstationary systems, Automatica 118 (2020). Article 109037
[18] A. Gańcza, M. Niedźwiecki, M. Ciołek, Regularized local basis function approach to identification of nonstationary processes, IEEE Trans. Signal Process. 69 (2021) 1665–1680.
[19] M. Niedźwiecki, M. Ciołek, A. Gańcza, P. Kaczmarek, Application of regularized Savitzky–Golay filters to identification of time-varying systems, Automatica 133 (2021). Article 109865
[20] M.K. Tsatsanis, G.B. Giannakis, Modelling and equalization of rapidly fading channels, Int. J. Adapt. Control Signal Process. 10 (2–3) (1996) 159–176.
[21] A.M. Sayeed, B. Aazhang, Joint multipath-Doppler diversity in mobile wireless communications, IEEE Trans. Commun. 47 (1) (1999) 123–132.
[22] Y. Zakharov, V. Kodanev, Multipath-Doppler diversity of OFDM signals in an underwater acoustic channel, in: Proceedings of IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP), vol. 5, 2000, pp. 2941–2944.
[23] M. Niedźwiecki, A. Gańcza, M. Ciołek, On the preestimation technique and its application to identification of nonstationary systems, in: 59th IEEE Conference on Decision and Control (CDC-2020), Jeju Island, Republic of Korea, 2020, pp. 286–293.
[24] M. Niedźwiecki, T. Kłaput, Fast recursive basis function estimators for identification of time-varying processes, IEEE Trans. Signal Process. 50 (8) (2002) 1925–1934.
[25] M. Niedźwiecki, On tracking characteristics of weighted least squares estimators applied to nonstationary system identification, IEEE Trans. Autom. Control 33 (1) (1988) 96–98.
[26] R.C. Geary, Relative efficiency of count of sign changes for assessing residual autoregression in least squares regression, Biometrika 57 (1970) 123–127.
[27] L. Ljung, T. Chen, What can regularization offer for estimation of dynamical systems, in: Proc. of the 11th IFAC Workshop on Adaptation and Learning in Control and Signal Processing, Caen, France, vol 46, 2013, pp. 1–8.
[28] T. Chen, H. Ohlsson, L. Ljung, On the estimation of transfer functions regularizations and gaussian processes - revisited, Automatica 48 (8) (2012) 1525–1535.
[29] G. Pillonetto, F. Dinuzzo, T. Chen, G. De Nicolao, L. Ljung, Kernel methods in system identification, machine learning and function estimation: a survey, Automatica 50 (3) (2014) 657–682.
[30] B.P. Carlin, T.A. Louis, Bayes and Empirical Bayes Methods for Data Analysis, Chapman and Hall/CRC, 2000.
[31] V.M. Baronkin, Y. Zakharov, T.C. Tozer, Frequency estimation in slowly fading multipath channels, IEEE Trans. Commun. 50 (11) (2002) 1848–1859.
[32] I.J. Good, The Estimation of Probabilities, MIT Press, 1965.
[33] R. Tibshirani, Regression shrinkage and selection via the LASSO, J. R. Stat. Soc. 58 (1) (1996) 267–288.

[34] M.R. Lewis, et al., Evaluation of Vector Sensors for Adaptive Equalization in Underwater Acoustic Communication, Massachusetts Institute of Technology, 2014 Ph.D. thesis.
[35] J.G. Proakis, Digital Communications, McGraw-Hill, 1983.
[36] L. Shen, B. Henson, Y. Zakharov, P. Mitchell, Digital self-interference cancellation for full-duplex underwater acoustic systems, IEEE Trans. Circuits Syst. II 67 (1) (2019) 192–196.
[37] B. Henson, J. Li, Y. Zakharov, C. Liu, Waymark baseband underwater acoustic propagation model, in: Proceedindgs of the IEEE Underwater Communications and Networking (UComms), 2014, pp. 1–5.
[38] M.B. Porter, BELLHOP 3D User Guide, Heat, Light, and Sound Research, Inc, Tech. Rep, La Jolla, CA, USA, 2016.
[39] P.A. Brodtkorb, P. Johannesson, G. Lindgren, I. Rychlik, J. Rydén, E. Sjö, WAFO-A matlab toolbox for analysis of random waves and loads, in: 10th International Offshore and Polar Engineering Conference, Seattle, Washington, USA, 2000.
[40] W.J. Pierson Jr., L. Moskowitz, A proposed spectral form for fully developed wind seas based on the similarity theory of S. A. Kitaigorodskii, J. Geophys. Res. 69 (24) (1964) 5181–5190.