Postprint of: Niedźwiecki M., Gańcza A., Karhunen-Loeve-based approach to tracking of rapidly fading wireless communication channels, SIGNAL PROCESSING Vol. 209 (2023), 109043, DOI: 10.1016/j.sigpro.2023.109043

Karhunen-Loève-based approach to tracking of rapidly fading wireless communication channels

Maciej Niedźwiecki[a], Artur Gańcza[a]

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

Abstract

When parameters of wireless communication channels vary at a fast rate, simple estimation algorithms, such as weighted least squares (WLS) or least mean squares (LMS) algorithms, cannot estimate them with accuracy needed to secure reliable operation of the underlying communication systems. In cases like this, the local basis function (LBF) estimation technique can be used instead, significantly increasing the achievable tracking accuracy. We show that when some prior knowledge of statistical properties of parameter changes is available, such as the bandwidth of their variation, both the type and the number of basis functions, used in the LBF approach to approximate time-varying channel parameters, can be optimized using the Karhunen-Loève (KL) decomposition based technique.

Keywords: Time-varying channels, local basis functions, Karhunen-Loève expansion

1. Introduction

Many wireless communication channels (terrestrial, underwater) can be well approximated by a time-varying FIR model of the form [1, 2]

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

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

*Corresponding author

Email address: maciekn@eti.pg.edu.pl (Maciej Niedźwiecki)

where $t = \ldots, -1, 0, 1, \ldots$ denotes discrete (normalized) time, $y(t)$ denotes the received complex-valued signal, $\boldsymbol{\varphi}(t) = [u(t), \ldots, u(t-n+1)]^{\mathrm{T}}$ denotes regression vector made up of past values of the complex-valued transmitted signal, $\boldsymbol{\theta}(t) = [\theta_1(t), \ldots, \theta_n(t)]^{\mathrm{T}}$ denotes the vector of time-varying channel impulse response coefficients, and $\{e(t)\}$ denotes circular white noise independent of $\{u(t)\}$ and $\{\boldsymbol{\theta}(t)\}$. The symbol $*$ denotes complex conjugate and H – complex conjugate transpose (Hermitian transpose). We will further assume that $\{u(t)\}$ is a sequence of zero-mean independent and identically distributed circular random variables with variance σ_u^2.

The application, studied recently, which particularly well fits the technique developed in this paper, is adaptive self-interference (SI) cancellation in full-duplex (FD) underwater acoustic (UWA) communication systems [3]. An interesting feature of this application is that it allows one to work with a decision delay, which means that estimation of channel parameters can be based not only on past signal samples but also on a certain number of "future" (with respect to the moment of interest) ones. Hence, channel identification can be carried out using noncausal estimation algorithms with improved tracking capabilities, such as the one described in this paper.

In the LBF approach each parameter trajectory is locally approximated by known functions of time, called basis functions [4]. When nothing is known about the way system parameters vary with time, some general purpose approximation schemes are applied, resulting in such popular choices of basis functions as powers of time (Taylor series approximation) or sine/cosine functions (Fourier series approximation) [5]. We will show that when channel parameters can be regarded as wide-sense stationary processes with known (or experimentally determined) second order characteristics, the choice of the basis set can be optimized using the Karhunen-Loève decomposition technique. The main contribution of the paper is solution of the estimation bias-variance trade-off problem - derivation of the decision rule that allows to determine the number of basis functions guaranteeing minimization of the mean squared parameter tracking error.

2. Local basis function estimators and their properties

In the LBF approach, proposed and analyzed in [4], system parameters are modeled, in a sliding analysis window $T_k(t) = [t-k, t+k]$ of length $K = 2k+1$, as linear combinations of known, linearly independent functions of time $f_1(j), \ldots, f_m(j), j \in I_k = [-k, k]$, further referred to as basis functions. This leads to the following local description of the parameter trajectory

$$\theta_i(t+j) = \sum_{l=1}^{m} f_l^*(j) a_{il}(t) = \mathbf{f}_m^{\mathrm{H}}(j)\boldsymbol{\alpha}_i(t),\ j \in I_k,\ i = 1, \ldots, n, \tag{2}$$

where $\mathbf{f}_m(j) = [f_1(j), \ldots, f_m(j)]^{\mathrm{T}}$ and $\boldsymbol{\alpha}_i(t) = [a_{i1}(t), \ldots, a_{im}(t)]^{\mathrm{T}}$ denotes the vector of basis expansion coefficients. Since the value of $\boldsymbol{\alpha}_i$, characterizing evolution of the i-th system parameter in the analysis window $T_k(t)$, may change along with the position of the window, it is written down as a function of time.

If no prior knowledge of the type or speed of parameter variation is available the selection of basis functions is usually based on some universal approximation arguments. For example, when the polynomial basis is adopted [4, 5, 6, 7]: $f_l(j) = j^{l-1},\ l = 1, \ldots, m$, the model (2) can be interpreted as a local Taylor series expansion of the true parameter trajectory in the neighborhood of the time instant t. Without any loss of generality we will further assume that basis functions are orthonormal, i.e., $\sum_{j=-k}^{k} \mathbf{f}_m(j)\mathbf{f}_m^{\mathrm{H}}(j) = \mathbf{I}_m$, where $\mathbf{I}_m$ denotes the $m \times m$ identity matrix. In the case of the polynomial basis, orthogonalization, which can be performed using the Gram-Schmidt procedure, results in a set of discrete-time Legendre polynomials.

Combining (1) and (2), one arrives at

$$y(t+j) = \boldsymbol{\beta}_m^{\mathrm{H}}(t)\boldsymbol{\psi}_m(t, j) + e(t+j),\ j \in I_k, \tag{3}$$

where $\boldsymbol{\beta}_m(t) = [\boldsymbol{\alpha}_1^{\mathrm{T}}(t), \ldots, \boldsymbol{\alpha}_m^{\mathrm{T}}(t)]^{\mathrm{T}}$, and $\boldsymbol{\psi}_m(t, j) = \boldsymbol{\varphi}(t+j) \otimes \mathbf{f}_m(j)$ denotes the $mn \times 1$ generalized regression vector. The symbol $\otimes$ stands for the Kronecker product of the

respective matrices/vectors. The LBF estimator of $\boldsymbol{\theta}(t)$ takes the form

$$\widehat{\boldsymbol{\beta}}_m(t) = \arg\min_{\boldsymbol{\beta}_m} \sum_{j=-k}^{k} |y(t+j) - \boldsymbol{\beta}_m^{\mathrm{H}} \boldsymbol{\psi}_m(t,j)|^2 = \mathbf{P}_m^{-1}(t)\mathbf{p}_m(t)$$

$$\widehat{\boldsymbol{\theta}}_m(t) = \mathbf{F}_m \widehat{\boldsymbol{\beta}}_m(t), \tag{4}$$

where $\mathbf{P}_m(t) = \sum_{j=-k}^{k} \boldsymbol{\psi}_m(t,j)\boldsymbol{\psi}_m^{\mathrm{H}}(t,j)$ is the $mn \times mn$ generalized regression matrix, $\mathbf{p}_m(t) = \sum_{j=-k}^{k} y^*(t+j)\boldsymbol{\psi}_m(t,j)$ is the $mn \times 1$ vector, and $\mathbf{F}_m = \mathbf{I}_n \otimes \mathbf{f}_m^{\mathrm{H}}(0)$.

The estimated vector of hyperparameters $\widehat{\boldsymbol{\beta}}_m(t)$ allows one to approximate parameter trajectory in the entire analysis interval $[t-k, t+k]$. However, in agreement with the local estimation policy, for each position t of the sliding analysis window $T_k(t)$, only the "central" estimate $\widehat{\boldsymbol{\theta}}_m(t)$ is computed and further utilized.

The LBF estimators are noncausal as they rely on both past and "future" (with respect to the instant t) input/output measurements. They introduce decision delay (latency) equal to k sampling intervals. The computational burden associated with evaluation of $\widehat{\boldsymbol{\theta}}_m(t)$ is dominated by the cost of computing $\widehat{\boldsymbol{\beta}}_m(t)$. When $\widehat{\boldsymbol{\beta}}_m(t) = \mathbf{P}_m^{-1}(t)\mathbf{p}_m(t)$ is computed in a naïve, straightforward manner, the cost is very high as it involves $\mathcal{O}(m^2n^2K)$ complex MAC (multiply and accumulate) operations per time update, needed to evaluate $\mathbf{P}_m(t)$ and $\mathbf{p}_m(t)$, and $\mathcal{O}(m^3n^3)$ operations required to invert $\mathbf{P}_m(t)$ and multiply it by $\mathbf{p}_m(t)$. As shown recently in [8], both cost components can be significantly reduced. by taking into account the transversal structure of the regression vector $\boldsymbol{\varphi}(t)$ and using the dichotomous coordinate descent iterative technique to solve the matrix equation $\mathbf{P}_m(t)\widehat{\boldsymbol{\beta}}_m(t) = \mathbf{p}_m(t)$.

As shown in [4], under the assumptions made above the mean path of LBF estimates can be expressed in the form

$$\bar{\boldsymbol{\theta}}_m(t) = \mathrm{E}\left[\widehat{\boldsymbol{\theta}}_m(t)\right] \cong \sum_{j=-k}^{k} h_m(j)\boldsymbol{\theta}(t+j), \tag{5}$$

where averaging is carried over $\{\boldsymbol{\varphi}(t)\}$ and $\{e(t)\}$, and

$$h_m(j) = \mathbf{f}_m^{\mathrm{H}}(0)\mathbf{f}_m(j), \ j \in I_k, \tag{6}$$

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

If the estimated parameter trajectories belong to the subspace spanned by the basis functions, the covariance matrix of LBF estimates can be expressed in the form

$$\mathrm{cov}[\widehat{\boldsymbol{\theta}}_m(t)] \cong \sigma_e^2 \boldsymbol{\Phi}^{-1} \sum_{j=-k}^{k} h_m^2(j) = \sigma_e^2 \boldsymbol{\Phi}^{-1} \mathbf{f}_m^{\mathrm{H}}(0) \mathbf{f}_m(0), \tag{7}$$

where $\boldsymbol{\Phi} = \mathrm{cov}[\boldsymbol{\varphi}(t)] = \sigma_u^2 \mathbf{I}_n$. When system parameters cannot be exactly represented as linear combinations of basis functions, the accuracy of this approximation quickly increases with growing m.

3. Optimization of tracking performance

Suppose that some statistical knowledge of parameter time variation is available. In particular, we will assume that $\{\theta_i(t)\}$, $i = 1, \ldots, n$ are mutually uncorrelated random sequences with the same - up to the scaling factors $\gamma_i \geq 0, i = 1, \ldots, n$ which reflect the power decay profile of the channel - autocorrelation function $\mathrm{E}[\theta_i(t+\tau)\theta_i^*(t)] = \gamma_i r_\theta(\tau)$, $i = 1, \ldots, n$. Note that this assumption is fulfilled for typical communication channels. In the sequel we will also assume that the autocorrelation function $r_\theta(\tau)$ and scaling coefficients γ_i are known *a priori* or were determined experimentally. For example, in the terrestrial wireless communication systems, operated under uncorrelated scattering and Rayleigh fading, the Jakes' model is often adopted leading to $r_\theta(\tau) = \sigma_\theta^2 J_0(\omega_d \tau)$, where $J_0(\cdot)$ denotes the zero order Bessel function and ω_d denotes the maximum normalized Doppler frequency, which is upper bounded by the moving speed of the mobile and scatterers, divided by the carrier wavelength [9].

The performance optimization approach described below is based on the idea of choosing as the basis functions the eigenvectors of the $K \times K$ correlation matrix

$$\mathbf{R}_\theta = \begin{bmatrix} r_\theta(0) & \ldots & r_\theta(K-1) \\ \vdots & \ddots & \vdots \\ r_\theta^*(K-1) & \ldots & r_\theta(0) \end{bmatrix}.$$

Consider the eigendecomposition of the correlation matrix $\mathbf{R}_\theta = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{H}}$, where $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_1, \ldots, \lambda_K\}$, $\lambda_i \geq 0, i = 1, \ldots K$, is the diagonal matrix made up of the eigenvalues of $\mathbf{R}_\theta$, arranged in the decreasing order, and the $K \times K$ matrix $\mathbf{Q}$ of the form

$$\mathbf{Q} = [\mathbf{q}_1 | \ldots | \mathbf{q}_K], \qquad \mathbf{q}_i^{\mathrm{H}} \mathbf{q}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

is made up of the corresponding eigenvectors. We postulate to set

$$[f_i(-k), \ldots, f_i(k)]^{\mathrm{T}} = \mathbf{q}_i.$$

It is straightforward to show that, in the case considered, it holds that $\mathbf{J}_K \mathbf{q}_i^* = \mathbf{q}_i$, where $\mathbf{J}_K$ is the $K \times K$ antidiagonal matrix with all antidiagonal elements equal to 1, which flips the vector $\mathbf{q}_i^*$ "upside down". Actually, note that $\mathbf{R}_\theta \mathbf{q}_i = \lambda_i \mathbf{q}_i$ entails $\lambda_i \mathbf{J}_K \mathbf{q}_i^* = \mathbf{J}_K \mathbf{R}_\theta^* \mathbf{q}_i^* = \mathbf{J}_K \mathbf{R}_\theta^* \mathbf{J}_K \mathbf{J}_K \mathbf{q}_i^* = \mathbf{R}_\theta \mathbf{J}_K \mathbf{q}_i^*$, where the last two transitions follow from the facts that $\mathbf{J}_K \mathbf{J}_K = \mathbf{I}_K$ and $\mathbf{J}_K \mathbf{R}_\theta^* \mathbf{J}_K = \mathbf{R}_\theta$. As a consequence of this property, one obtains $f_i(j) = f_i^*(-j), j \in I_k$. Additionally, the quantities $f_i(0), i = 1, \ldots, K$, must be real-valued.

Values of the selected number of the largest eigenvalues and the corresponding eigenvectors of $\mathbf{R}_\theta$ can be efficiently computed (without the need to calculate all K eigenvalues/eigenvectors) using bisection and inverse iteration algorithms, after converting the matrix $\mathbf{R}_\theta$ into a symmetric tridiagonal (Hessenberg) form via orthogonal similarity transformations [10, 11]. This step can be performed off-line prior to identification.

Since, according to the Karhunen-Loève expansion theorem, every zero-mean random process $\{\theta(t)\}$ with correlation matrix $\mathbf{R}_\theta$ can be in the interval $T_k(t)$ expressed (exactly) as a linear combination of eigenvectors of $\mathbf{R}_\theta$, the basis set proposed above is a pretty straightforward choice [12, 13]. However, selection of the most appropriate number of eigenvectors used for parameter approximation is in the case of LBF estimators an unsolved practically important problem. When the number m of basis functions is increased, the bias component $B(m)$ of the mean squared parameter estimation error (MSE) decreases, but its variance component $V(m)$ increases and *vice versa* [4]. Since

MSE is the sum of its bias and variance components, a compromise value of m should be chosen, which constitutes the well-known bias-variance trade-off in system identification. Based on the results established for LBF estimators, we will show how one can find m which minimizes $\mathrm{MSE}(m) = B(m) + V(m)$.

Denote by $\mathbf{Q}_m = [\mathbf{q}_1|\ldots|\mathbf{q}_m]$ the $K \times m$ matrix made up of the first m eigenvectors of $\mathbf{R}_\theta$, and denote by $\mathbf{h}_m = [h_m(-k), \ldots, h_m(k)]^{\mathrm{T}}$ the $K \times 1$ vector of the impulse response coefficients. Finally, let $\mathbf{r}_\theta = [r_\theta(-k), \ldots, r_\theta(k)]^{\mathrm{T}}$. Under the assumptions made above, the bias component of $\mathrm{MSE}(m)$ can be expressed in the form

$$B(m) = \mathrm{E}\left[\| \bar{\boldsymbol{\theta}}_m(t) - \boldsymbol{\theta}(t) \|^2\right] = \eta[r_\theta(0) - \mathbf{r}_\theta^{\mathrm{H}}\mathbf{h}_m^* - \mathbf{r}_\theta^{\mathrm{T}}\mathbf{h}_m + \mathbf{h}_m^{\mathrm{H}}\mathbf{R}_\theta\mathbf{h}_m] \tag{9}$$

where $\eta = \sum_{i=1}^{n}\gamma_i$ and $\|\cdot\|$ denotes the Euclidean norm. Note that under (6) it holds that $\mathbf{h}_m = \mathbf{Q}_m\mathbf{f}_m^*(0)$ and $\mathbf{f}_m^*(0) = \mathbf{Q}_m^{\mathrm{H}}\mathbf{c}$, where $\mathbf{c} = [0, \ldots, 0, 1, 0, \ldots 0]^{\mathrm{T}}$ is the $K \times 1$ vector with only one nonzero element located in its center. This leads to $\mathbf{h}_m = \mathbf{Q}_m\mathbf{Q}_m^{\mathrm{H}}\mathbf{c}$ and $\mathbf{h}_m^{\mathrm{H}}\mathbf{R}_\theta\mathbf{h}_m = \mathbf{c}^{\mathrm{T}}\mathbf{Q}_m\mathbf{Q}_m^{\mathrm{H}}\mathbf{R}_\theta\mathbf{Q}_m\mathbf{Q}_m^{\mathrm{H}}\mathbf{c}$. It holds that

$$\mathbf{Q}_m\mathbf{Q}_m^{\mathrm{H}} = \sum_{i=1}^{m}\mathbf{q}_i\mathbf{q}_i^{\mathrm{H}}, \qquad \mathbf{R}_\theta = \sum_{i=1}^{K}\lambda_i\mathbf{q}_i\mathbf{q}_i^{\mathrm{H}}.$$

Hence, exploiting (8), one arrives at $\mathbf{h}_m^{\mathrm{H}}\mathbf{R}_\theta\mathbf{h}_m = \sum_{i=1}^{m}\lambda_i\mathbf{c}^{\mathrm{T}}\mathbf{q}_i\mathbf{q}_i^{\mathrm{H}}\mathbf{c} = \sum_{i=1}^{m}\lambda_i f_i^2(0)$. Similarly, since $r_\theta = \mathbf{R}_\theta^*\mathbf{c}$ and $\mathbf{R}_\theta^{\mathrm{H}} = \mathbf{R}_\theta$, one obtains $\mathbf{r}_\theta^{\mathrm{T}}\mathbf{h}_m = \mathbf{r}_\theta^{\mathrm{H}}\mathbf{h}_m^* = \mathbf{h}_m^{\mathrm{H}}\mathbf{R}_\theta\mathbf{h}_m$ which finally leads to

$$B(m) = \eta\left[r_\theta(0) - \sum_{i=1}^{m}\lambda_i f_i^2(0)\right]. \tag{10}$$

As to the variance component of MSE, one obtains

$$V(m) = \mathrm{E}\left[\| \widehat{\boldsymbol{\theta}}_m(t) - \bar{\boldsymbol{\theta}}_m(t) \|^2\right] = \mathrm{tr}\{\mathrm{cov}[\widehat{\boldsymbol{\theta}}_m(t)]\} \cong \frac{n\sigma_e^2}{\sigma_u^2}\mathbf{h}_m^{\mathrm{H}}\mathbf{h}_m = \frac{n\sigma_e^2}{\sigma_u^2}\sum_{i=1}^{m} f_i^2(0). \tag{11}$$

Note that, as expected, the bias component $B(m)$ decreases, and the variance component $V(m)$ increases, with growing m.

To guarantee invertibility of the generalized regression matrix $\mathbf{P}_m(t)$, the width of the analysis window K should be at least equal to the number of estimated hyperparameters mn. To avoid numerical problems caused by near singularity of $\mathbf{P}_m(t)$, we will restrict

the search of m_{opt} to the range $[1, M_0]$, where $M_0 = \lfloor (K-10)/n \rfloor$ and $\lfloor \cdot \rfloor$ denotes the floor function. Under this restriction the optimal value of m for a given K can be obtained from

$$\begin{aligned} m_{\text{opt}} &= \underset{1 \le m \le M_0}{\arg\min}\, \text{MSE}(m) = \underset{1 \le m \le M_0}{\arg\min} \sum_{i=1}^{m} \left[\frac{n\sigma_e^2}{\sigma_u^2} - \eta\lambda_i \right] f_i^2(0) \\ &= \underset{1 \le m \le M_0}{\arg\max} \left\{ m \text{ s.t. } \lambda_m > \frac{n\sigma_e^2}{\eta\sigma_u^2} \right\}. \end{aligned} \tag{12}$$

i.e., one should reject all eigenvectors $\mathbf{q}_i$ for which it holds that $\lambda_i \le (n\sigma_e^2)/(\eta\sigma_u^2) = n\sigma_\theta^2/\text{SNR}$, where $\text{SNR} = \text{E}[|\boldsymbol{\theta}^{\text{H}}(t)\boldsymbol{\varphi}(t)|^2]/\sigma_e^2 = \eta\sigma_u^2\sigma_\theta^2/\sigma_e^2$ denotes the signal-to-noise ratio.

The summary of the proposed LBF KL algorithm is presented below.

Summary of the LBF KL algorithm

Given/assumed:

- $\{u(t), y(t), t = \ldots, -1, 0, 1, \ldots\}$ – input/output data
- n – number of estimated parameters
- k – decision delay (latency)
- $K = 2k + 1$ – width of the sliding analysis window $T_k(t) = [t-k, t+k]$
- σ_u^2 – variance of the input signal
- σ_e^2 – variance of the measurement noise (known or estimated)
- $\mathbf{R}_\theta$ – autocorrelation matrix of parameter changes (known or estimated)
- $\gamma_1, \ldots, \gamma_n$ – channel power decay profile (known or estimated)

Off-line calculations (performed once prior to identification):

- Compute $M_0 = \lfloor (K-10)/n \rfloor$ largest eigenvalues of $\mathbf{R}_\theta$ along with the corresponding eigenvectors (using standard numerical procedures available e.g. in MATLAB).
- Choose the optimal number of basis functions (eigenvectors of $\mathbf{R}_\theta$) m using (12).

On-line computations (performed for consecutive locations t of the sliding analysis window):

- Compute the generalized regression matrix $\mathbf{P}_m(t)$ and the vector $\mathbf{p}_m(t)$.
- Evaluate the estimate of the vector of basis expansion coefficients $\widehat{\boldsymbol{\beta}}_m(t)$ by solving the associated set of linear equations $\mathbf{P}_m(t)\widehat{\boldsymbol{\beta}}_m(t) = \mathbf{p}_m(t)$.
- Evaluate the estimate of the vector of channel coefficients $\widehat{\boldsymbol{\theta}}_m(t)$ using (4).

Optimization carried out above was based on the assumption that the autocorrelation function $r_\theta(\tau)$ is known. In practice, however, the autocorrelation function can be usually expressed in the form $r_\theta(\tau, \boldsymbol{\xi})$, where $\boldsymbol{\xi}$ denotes the vector of environmental factors. For example, in the case of Jakes' model the maximum Doppler frequency ω_d depends on the vehicle speed, which is usually time-varying. To obtain robust tracking algorithms, one can use the parallel estimation technique. In this case not one but several parameter tracking algorithms, corresponding to different values of design parameters $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_M$, are run concurrently and compared using the local measure of fit. At each time instant t the "locally the best" algorithm is chosen - see [4] for more details on this approach. Alternatively, if the prior distribution of $\boldsymbol{\xi}$, denoted by $\pi(\boldsymbol{\xi})$, is known, one can replace $r_\theta(\tau)$ with $\bar{r}_\theta(\tau) = \int r_\theta(\tau, \boldsymbol{\xi})\pi(\boldsymbol{\xi})d\boldsymbol{\xi}, \forall\tau$. Note that averaging retains the positive definiteness property of the autocorrelation function.

4. Numerical results

The taps of a complex-valued FIR system of order $n = 20$ governed by (1), modeling a FD UWA channel, were simulated as independent zero-mean unit-variance Gaussian processes with uniform power spectral density $S_\theta(\omega)$ obeying [6]

$$S_\theta(\omega) = \begin{cases} S_0 & \text{for} \quad |\omega| \leq \omega_0 \\ 0 & \text{for} \quad |\omega| > \omega_0 \end{cases}, \tag{13}$$

with $\omega_0 = 2\pi f_0$, $f_0 = 0.001$. The autocorrelation function corresponding to this flat (nonpreferential) model of Doppler spectrum has the form $r_\theta(\tau) = \sigma_\theta^2 \, \frac{\sin \omega_0 \tau}{\omega_0 \tau}$. In this case the eigenvectors of $\mathbf{R}_\theta$ are made up of discrete prolate spheroidal (Slepian) sequences, which can be efficiently calculated (usually a small number of them) using the algorithm described in [14].

The FFT-based procedure described in [15] was used to generate realizations of $\{\theta_i(t), t = 1, \ldots, N\}, i = 1, \ldots, n$. The plots of real parts of selected parameters of the simulated channel are shown in Fig. 1. The measurement noise and the input signal were white (complex) Gaussian. The variance of the input signal was set to $\sigma_u^2 = 1$ and

it was assumed that $\gamma_i = 1$, $\forall i$, i.e., $\eta = n$. Note that in FD UWA applications the input signal, generated by the near-end transmitter, is *known* all the time.

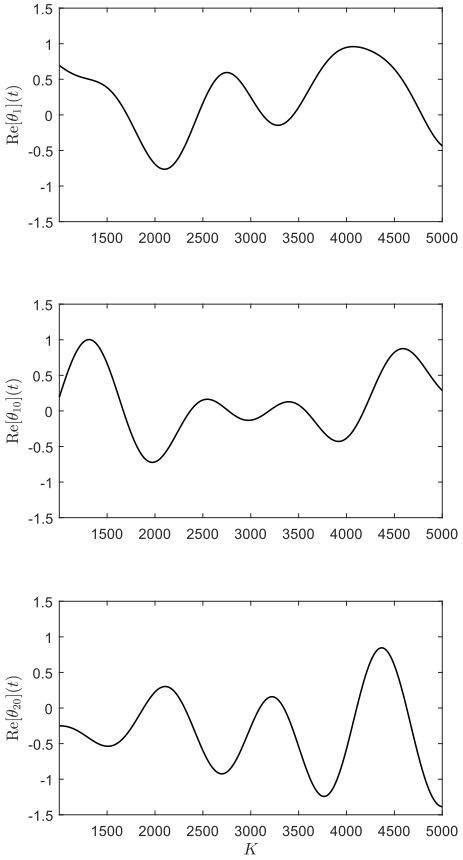

Fig. 1 Real parts of selected parameters of the simulated channel.

Three variants of the signal-to-noise ratio were considered: 10 dB, 30 dB and 50 dB. For comparison, in addition to the optimized LBF algorithm, 3 variants of the Legendre-based LBF algorithms were run, corresponding to $m = 1$, $m = 3$ and $m = 5$, respectively. Fig. 2 shows the dependence of the steady state mean squared parameter estimation errors, averaged over $N = 10^5$ time steps, on the width K of the analysis

interval. Additionally, in each case the plots illustrating the dependence of the optimal number of basis functions on K were provided. Note that the optimized LBF algorithm almost always yields results that are better than or equal to those obtained using the Legendre-based LBF algorithms. When SNR=50 dB the performance gain is significant and close to 10 dB for almost all values of K (note that in the full-duplex scenario, due to the short distance between the transmit antenna and the receive antenna, the typical SNR values are usually high, often in excess of 50 dB).

To obtain good tracking results using the Legendre-based LBF approach, the width of the analysis window must be carefully tuned to the number of basis functions. Too small or too large values of K for a chosen m may result in severe performance degradation. In contrast with this, the optimized LBF algorithm is much easier to handle since the best tracking performance is guaranteed for any preselected value of K, i.e., for any chosen latency k. The lower performance bounds, shown in Fig. 1, were obtained by means of *unconstrained* minimization of MSE regarded as a function of $\mathbf{h} = [h(-k), \ldots, h(k)]^{\mathrm{T}}$:

$$\mathrm{MSE}(\mathbf{h}) = \frac{n\sigma_e^2}{\sigma_u^2}\mathbf{h}^{\mathrm{H}}\mathbf{h} + \eta[r_\theta(0) - \mathbf{r}_\theta^{\mathrm{H}}\mathbf{h}^* - \mathbf{r}_\theta^{\mathrm{T}}\mathbf{h} + \mathbf{h}^{\mathrm{H}}\mathbf{R}_\theta\mathbf{h}]. \tag{14}$$

The optimal value of $\mathbf{h}$ is given by

$$\mathbf{h}_{\mathrm{opt}} = \left(\mathbf{R}_\theta + \frac{n\sigma_e^2}{\eta\sigma_u^2}\mathbf{I}_K\right)^{-1}\mathbf{r}_\theta^*. \tag{15}$$

The plots of $\mathrm{MSE}(\mathbf{h}_{\mathrm{opt}})$, shown in Fig. 2, constitute theoretical bounds on MSE which could have been attained if one had a full control, via selection of basis functions, over the shape of the impulse response $\{h(j)\}$. The limitations of the constrained optimization, governed by (12), are clearly visible for small values of K.

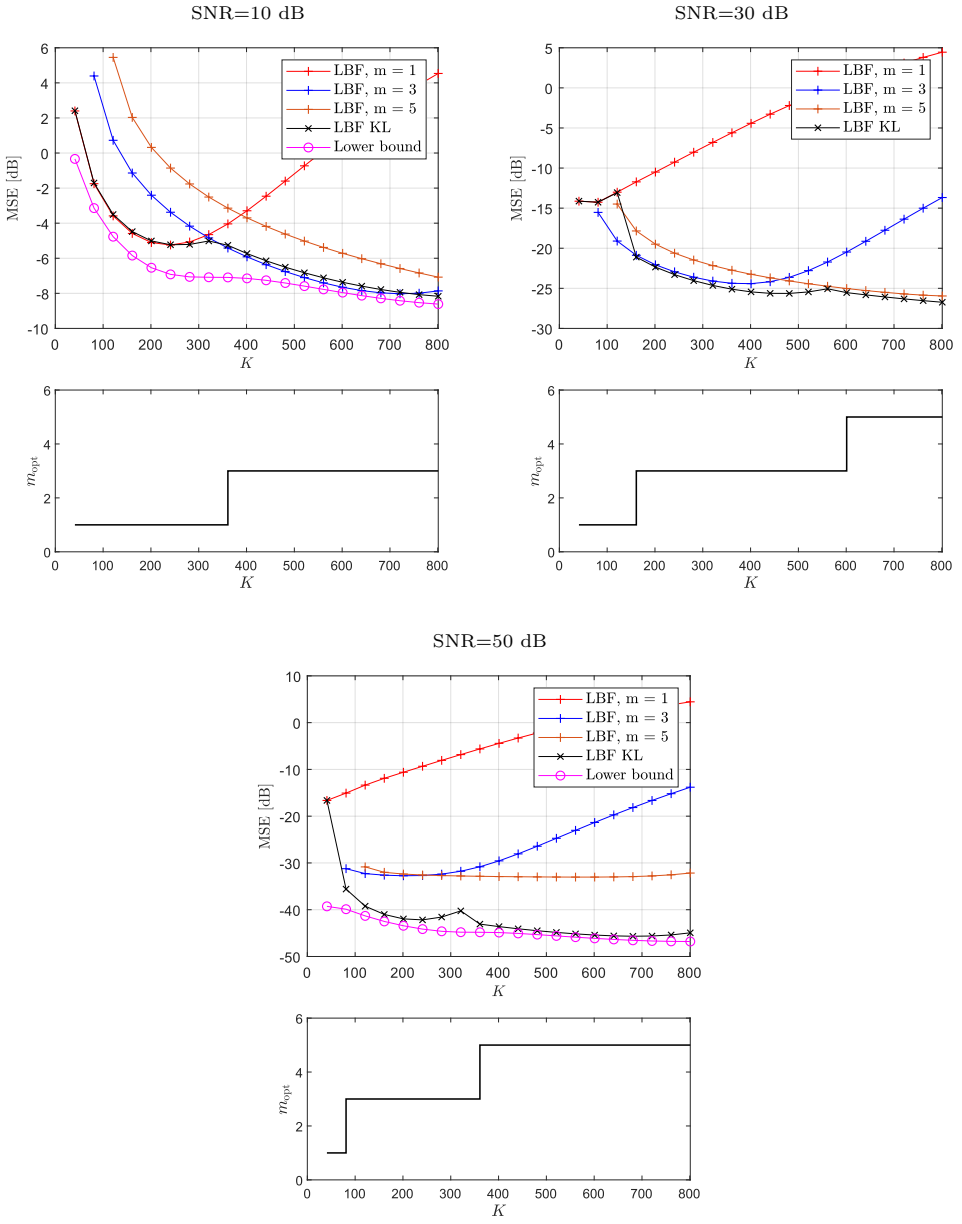

Fig. 2 Mean squared parameter tracking errors obtained for the Legendre-based LBF estimators and for the optimized variant of the LBF scheme based on the Karhunen-Loève expansion (KL). The lower performance bounds correspond to the optimal (unconstrained) choice of the impulse response. K denotes the width of the local analysis interval and m is the number of applied basis functions. The number of estimated impulse response coefficients was equal to $n = 20$. The graphs displayed below MSE plots show the dependence of the

optimal
number of basis functions on K.

Finally, Table 1 shows the best results that can be obtained using the TU-RLS (time-updated recursive least squares) algorithm, considered until recently the state-of-the-art in UWA communication [16, 17].

SNR [dB]	MSE [dB]	λ	μ
10	0.77	0.95	3.5E-03
30	-10.19	0.88	6.00E-07
50	-12.94	0.81	1.25E-02

Tab. 1 The best achievable performance yielded by the TU-RLS algorithm (λ and μ denote tuning parameters specified in [16]).

5. Conclusion

It was shown that prior knowledge of the correlation structure of parameter changes can be taken advantage of when selecting both the number and the type of basis functions used in the local basis function approach to identification of time-varying communication channels. Minimization of the mean squared parameter tracking error was performed using a truncated Karhunen-Loève expansion of the estimated parameter trajectories.

References

[1] M. K. Tsatsanis, G. B. Giannakis, Modelling and equalization of rapidly fading channels, International Journal of Adaptive Control and Signal Processing, 10 (2-3) (1996) 159-176.

[2] M. Stojanovic, J. Preisig, Underwater acoustic communication channels: Propagation models and statistical characterization, IEEE Communications Magazine, 47 (1) (2009) 84-89.

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

[4] M. Niedźwiecki, M. Ciołek, Generalized Savitzky-Golay filters for identification of nonstationary systems, Automatica, 108 (2019) 108477.

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

[6] 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.

[7] D.K. Borah, B.D. Hart, Frequency-selective fading channel estimation with a polynomial time-varying channel model, IEEE Transactions on Communications, 47 (6) (1999) 862-873.

[8] L. Shen, y. Zakharov, L. Shi, B. Henson, BEM adaptive filtering for SI cancellation in full-duplex underwater acoustic systems, Signal Processing, 191 (2022) 108366.

[9] W.C. Jakes, Microwave Mobile Communication, Wiley, 1974.

[10] L.N. Trefethen and D. Bau, III, Numerical Linear Algebra, SIAM, 1997.

[11] G.H. Golub and C.F. Van Loan, Matrix Computations, 4th edition, The Johns Hopkins University Press, 2013.

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

[13] K-W. Yip, T-S. Ng, Karhunen–Loève expansion of the WSSUS channel output and its application to efficient simulation, IEEE Journal on Selected Areas in Communications, 15 (4) (1997) 640-646.

[14] D.M. Gruenbacher and D.R. Hummels, A simple algorithm for generating discrete prolate spheroidal equences, IEEE Transactions on Signal Processing, 42, (1994) 3276-3278.

[15] D.B. Percival, Simulating Gaussian random processes with specified spectra, Computing Science and Statistics, 24 (1992) 534-538.

[16] T.H. Eggen, A.B. Baggeroer and J.C. Preisig, Communication over Doppler spread channels. Part I: Channel and receiver presentation, IEEE Journal of Oceanic Engineering, 25 (2000) 62-71.

[17] M.R. Lewis, Evaluation of Vector Sensors for Adaptive Equalization in Underwater Acoustic Communication, PhD thesis, 2014.