2019 27th European Signal Processing Conference (EUSIPCO)

Fast Basis Function Estimators for Identification of Nonstationary Stochastic Processes

1st Maciej Niedźwiecki
Department of Automatic Control
Gdańsk University of Technology
Gdańsk, Poland
maciekn@eti.pg.edu.pl

2nd Marcin Ciołek
Department of Automatic Control
Gdańsk University of Technology
Gdańsk, Poland
marcin.ciolek@pg.edu.pl

3rd Artur Gańcza
Department of Automatic Control
Gdańsk University of Technology
Gdańsk, Poland
artgancz@student.pg.edu.pl

***Abstract*—The problem of identification of a linear nonstationary stochastic process is considered and solved using the approach based on functional series approximation of time-varying parameter trajectories. The proposed fast basis function estimators are computationally attractive and yield results that are better than those provided by the local least squares algorithms. It is shown that two important design parameters – the number of basis functions and the size of the local analysis interval – can be selected on-line in an adaptive way.**

***Index Terms*—Identification of nonstationary processes, basis function estimators, adaptive estimation**

I. Introduction

The method of basis functions is a well-known and long-standing approach to identification of nonstationary systems described by linear regression equations [1] – [9]. In this framework, which is an extension of the local estimation approach [10], [11], each of n process parameters is approximated by a linear combination of m known functions of time, called basis functions. In this way the problem of tracking n time-varying parameters can be converted into a simpler problem of estimation of nm time-invariant hyperparameters – the coefficients that appear in basis function expansions of parameter trajectories. Estimation of hyperparameters can be easily carried out using the method of least squares. However, since estimation of hyperparameters requires inversion of a $nm \times mn$ generalized regression matrix, the obvious price one has to pay when the conversion described above is used, is in terms of computational burden. Another problem associated with this class of estimators, important from the practical viewpoint, is the choice of design parameters such as the approximation range (to comply with the rate of process nonstationarity) and the number of basis functions (to avoid overparametrization). In this paper we deal with all problems mentioned above. First, based on some large sample approximations, we derive a computationally fast version of the basis function estimator, which requires inversion of a $n \times n$ regression matrix only, i.e., has computational complexity comparable with that of the local least squares estimators [10]. Second, we propose an adaptive scheme for on-line adjustment of the number of basis functions and the size of the local analysis window.

This work was partially supported by the National Science Center under the agreement UMO-2018/29/B/ST7/00325. The authors are with the Faculty of Electronics, Telecommunications and Informatics. Calculations were carried out at the Academic Computer Centre in Gdańsk.

II. Local basis function approach

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

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

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

In the classical basis function framework, one assumes that in a selected time interval T each parameter trajectory $\{\boldsymbol{\theta}^j(t), t \in T\}$ can be modeled as a linear combination of a certain number of known functions of time $\{f_1(t), \ldots, f_m(t), t \in T\}$, further referred to as basis functions (BF). Based on the input-output information $\{y(t), \boldsymbol{\varphi}(t), t \in T\}$, gathered in the interval T, one can estimate time-invariant approximation coefficients and the corresponding parameter trajectories $\{\boldsymbol{\theta}^j(t), t \in T\}, j = 1, \ldots, n$. While the classical BF approach yields interval estimates of time varying parameters $\{\widehat{\boldsymbol{\theta}}^j(t), t \in T\}$, the local basis function (LBF) approach, developed in this paper, provides a sequence of point estimates evaluated independently for each position of the sliding local analysis window $T_k(t) = [t-k, t+k]$ centered at t. It can be shown that such point estimates are more accurate than the interval ones, especially at locations close to both ends of $T_k(t)$ (which is pretty obvious considering the fact that only one-sided information is available at both ends of the analysis interval) [10].

Denote by

$$f_{l|k}(i) = f_l^0\left(\frac{i}{k}\right), \quad l = 1, \ldots, m, \quad i \in I_k = [-k, k] \tag{2}$$

the set of linearly independent discrete-time functions obtained by sampling their continuous-time analogs $f_l^0(s), s \in [-1, 1]$ – the square integrable basis generating functions defined on the interval $[-1, 1]$. Denote by $\mathcal{F}^k$ the space of all square summable sequences defined on I_k and by $\mathcal{F}_{m|k}$ – the subspace of $\mathcal{F}^k$ spanned by the basis functions (2). Finally, denote by

$$\{\widetilde{f}_{1|k}(t), \ldots, \widetilde{f}_{m|k}(t), t \in T_k(t)\} \tag{3}$$

the orthonormal basis set of $\mathcal{F}_{m|k}$, i.e., the set obeying the condition

$$\sum_{t=-k}^{k} \widetilde{\mathbf{f}}_{m|k}(t)\widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(t) = \mathbf{I}_m \tag{4}$$

where $\widetilde{\mathbf{f}}_{m|k}(t) = [\widetilde{f}_{1|k}(t), \ldots, \widetilde{f}_{m|k}(t)]^{\mathrm{T}}$. Orthonormalization can be carried out, for example using the Gram-Schmidt procedure. In agreement with our local approximation strategy, we will assume that in the interval T_k each parameter can be written down as a linear combination of basis functions (3)

$$\theta^j(t+i) = \sum_{l=1}^{m} b^j_{l,m|k}\widetilde{f}_{l|k}(i), \quad i \in I_k, \ j = 1, \ldots, n. \tag{5}$$

Denote by $\boldsymbol{\psi}_{m|k}(t,i) = \boldsymbol{\varphi}(t+i) \otimes \widetilde{\mathbf{f}}_{m|k}(i)$ the generalized regression vector ($\otimes$ is a symbol of a Kronecker product) and by $\boldsymbol{\beta}_{m|k} = [b^1_{1,m|k}, \ldots, b^1_{m,m|k}, \ldots, b^n_{1,m|k}, \ldots, b^n_{m,m|k}]^{\mathrm{T}}$ the vector of all approximation coefficients. Using these shorthands, the local process model can be expressed in the form

$$y(t+i) = \boldsymbol{\psi}^{\mathrm{T}}_{m|k}(t,i)\boldsymbol{\beta}_{m|k} + e(t+i), \quad i \in I_k. \tag{6}$$

Based on (6), the LBF estimate of $\boldsymbol{\theta}(t)$ can be obtained using the method of least squares

$$\widehat{\boldsymbol{\theta}}_{m|k}(t) = \mathbf{F}_{m|k}\widehat{\boldsymbol{\beta}}_{m|k}(t), \quad \mathbf{F}_{m|k} \quad = \mathbf{I}_n \otimes \widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(0). \tag{7}$$

where

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{m|k}(t) &= \arg\min_{\boldsymbol{\beta}_{m|k}} \sum_{i=-k}^{k} [y(t+i) - \boldsymbol{\psi}^{\mathrm{T}}_{m|k}(t,i)\boldsymbol{\beta}_{m|k}]^2 \\ &= \mathbf{R}^{-1}_{m|k}(t)\mathbf{r}_{m|k}(t)\end{aligned} \tag{8}$$

$$\begin{aligned}\mathbf{R}_{m|k}(t) &= \sum_{i=-k}^{k} \boldsymbol{\psi}_{m|k}(t,i)\boldsymbol{\psi}^{\mathrm{T}}_{m|k}(t,i) \\ \mathbf{r}_{m|k}(t) &= \sum_{i=-k}^{k} y(t+i)\boldsymbol{\psi}_{m|k}(t,i) \ .\end{aligned} \tag{9}$$

When operated in the sliding window mode, the LBF estimators are computationally expensive. The computational burden can be reduced when basis functions are recursively computable. Such is the case when the basis is made up of the powers of time, namely

$$\left\{1, \frac{i}{k}, \ldots, \left(\frac{i}{k}\right)^{m-1}, \ i \in I_k\right\} \tag{10}$$

which corresponds to choosing $f^0_l(s) = s^{l-1}, l = 1, \ldots, m$. For such a choice of basis functions the formula (5) can be interpreted as a local Taylor series approximation of the true parameter trajectory.

It is straightforward to show that under (9) the vector $\widetilde{\mathbf{f}}_{m|k}(i)$ is recursively computable

$$\widetilde{\mathbf{f}}_{m|k}(i-1) = \mathbf{A}_{m|k}\widetilde{\mathbf{f}}_{m|k}(i) \tag{11}$$

which allows one to recursively compute the regression matrix $\mathbf{R}_{m|k}(t)$ and the vector $\mathbf{r}_{m|k}(t)$

$$\begin{aligned}\mathbf{R}_{m|k}(t) &= \mathbf{B}_{m|k}\big[\mathbf{R}_{m|k}(t-1) \\ &\quad - \boldsymbol{\psi}_{m|k}(t-1,-k)\boldsymbol{\psi}^{\mathrm{T}}_{m|k}(t-1,-k)\big]\mathbf{B}^{\mathrm{T}}_{m|k} \\ &\quad + \boldsymbol{\psi}_{m|k}(t,k)\boldsymbol{\psi}^{\mathrm{T}}_{m|k}(t,k) \\ \mathbf{r}_{m|k}(t) &= \mathbf{B}_{m|k}\big[\mathbf{r}_{m|k}(t-1) \\ &\quad - y(t-k-1)\boldsymbol{\psi}^{\mathrm{T}}_{m|k}(t-1,-k)\big] \\ &\quad + y(t+k)\boldsymbol{\psi}_{m|k}(t,k)\end{aligned} \tag{12}$$

where $\mathbf{B}_{m|k} = \mathbf{I}_n \otimes \mathbf{A}_{m|k}$.

However, even if the recursive computational scheme (12) is employed, the need to invert the $mn \times mn$ matrix $\mathbf{R}_{m|k}(t)$ at each time instant t makes the computational effort substantial. In the next section we will show that this effort can be reduced significantly if the simplified version of the LBF estimator is used instead of (7) - (9).

III. Properties of the Local Basis Function Estimator

In order to derive the simplified version of the LBF estimator, we will make some technical assumptions about the input and noise sequences

- (A1) $\{u(t)\}$ is a zero-mean wide sense stationary Gaussian sequence, persistently exciting of order at least n, with an exponentially decaying autocorrelation function $r_u(i)$.
- (A2) $\{e(t)\}$, independent of $\{u(t)\}$, is a sequence of zero-mean independent and identically distributed random variables.
- (A3) $\{\boldsymbol{\theta}(t)\}$ is independent of $\{u(t)\}$ and $\{e(t)\}$.

According to [6], under (A1) and (A2) it holds that

$$\lim_{k\to\infty} \mathbf{R}_{m|k}(t) = \boldsymbol{\Phi} \otimes \mathbf{I}_n = \bar{\mathbf{R}}_m \tag{13}$$

where $\boldsymbol{\Phi} = \mathrm{E}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)]$, and convergence takes place in the mean square sense (and hence also in probability). This means that for sufficiently large values of k the LBF estimator can be approximately expressed as

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_{m|k}(t) &\cong \mathbf{F}_{m|k}\bar{\mathbf{R}}^{-1}_m\mathbf{r}_{m|k}(t) \\ &= \Big[\mathbf{I}_n \otimes \widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(0)\Big]\left[\boldsymbol{\Phi}^{-1} \otimes \mathbf{I}_n\right] \\ &\quad \times \sum_{i=-k}^{k} y(t+i)\boldsymbol{\varphi}(t+i) \otimes \widetilde{\mathbf{f}}_{m|k}(i) \\ &= \boldsymbol{\Phi}^{-1}\sum_{i=-k}^{k} \widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(0)\widetilde{\mathbf{f}}_{m|k}(i)y(t+i)\boldsymbol{\varphi}(t+i) \\ &= \bar{\boldsymbol{\theta}}_{m|k}(t)\end{aligned} \tag{14}$$

where the second transition follows from the identity $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC} \otimes \mathbf{BD})$ (provided that the dimensions of the corresponding matrices/vectors match each other).

Note that (14) can be rewritten in the form

$$\bar{\boldsymbol{\theta}}_{m|k}(t) = \sum_{i=-k}^{k} h_{m|k}(i)\boldsymbol{\theta}^*(t+i) \tag{15}$$

where

$$h_{m|k}(i) = \widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(0)\widetilde{\mathbf{f}}_{m|k}(i), \quad i \in I_k \tag{16}$$

is the impulse response associated with the LBF estimator, and the quantity

$$\boldsymbol{\theta}^*(t) = \boldsymbol{\Phi}^{-1} y(t)\boldsymbol{\varphi}(t) \tag{17}$$

will be further referred to as the pre-estimate of $\boldsymbol{\theta}(t)$. Combining (1) and (17), and using assumptions (A1) - (A3), one arrives at

$$\begin{aligned} \mathrm{E}[\boldsymbol{\theta}^*(t)] &= \boldsymbol{\Phi}^{-1}\mathrm{E}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)]\boldsymbol{\theta}(t) \\ &\quad + \boldsymbol{\Phi}^{-1}\mathrm{E}[\boldsymbol{\varphi}(t)e(t)] = \boldsymbol{\theta}(t) \end{aligned} \tag{18}$$

Which means that $\boldsymbol{\theta}^*(t)$ is an unbiased estimator of $\boldsymbol{\theta}(t)$. However, even though unbiased, the pre-estimate has a very large variability which makes it useless in practice.

It can be shown [6] that if $1 \in \mathcal{F}_{m|k}$, i.e., if a constant function can be expressed as a linear combination of basis functions [note that this condition is fulfilled if the basis (10) is adopted], it holds that $\sum_{i=-k}^{k} h_{m|k}(i) = 1$, which means that the associated filter is lowpass. Combining (14), (15) and (18), one obtains

$$\mathrm{E}[\widehat{\boldsymbol{\theta}}_{m|k}(t)] \cong \mathrm{E}[\bar{\boldsymbol{\theta}}_{m|k}(t)] = \sum_{i=-k}^{k} h_{m|k}(i)\boldsymbol{\theta}(t+i) \tag{19}$$

i.e., the mean path of LBF estimates can be regarded as the result of passing the true parameter trajectory through the basis-dependent lowpass filter .

Note that in the "idealistic" case, where the true parameter trajectory can be exactly modeled as a linear combination of basis functions, i.e.,

$$\boldsymbol{\theta}(t+i) = [\mathbf{I}_n \otimes \widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(i)]\boldsymbol{\beta}_{m|k}, \quad i \in I_k$$

it holds that

$$\begin{aligned} &\mathrm{E}[\widehat{\boldsymbol{\theta}}_{m|k}(t)] \cong \mathrm{E}[\bar{\boldsymbol{\theta}}_{m|k}(t)] \\ &= \sum_{i=-k}^{k} \widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(0)\widetilde{\mathbf{f}}_{m|k}(i)[\mathbf{I}_n \otimes \widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(i)]\boldsymbol{\beta}_{m|k} \\ &= \left\{ \mathbf{I}_n \otimes \left[\widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(0) \sum_{i=-k}^{k} \widetilde{\mathbf{f}}_{m|k}(i)\widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(i) \right] \right\} \boldsymbol{\beta}_{m|k} \\ &\qquad = [\mathbf{I}_n \otimes \widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(0)]\boldsymbol{\beta}_{m|k} = \boldsymbol{\theta}(t) \end{aligned}$$

which means that under such conditions the LBF estimator is (approximately) unbiased. For arbitrary parameter changes the mean path of parameter estimates is a sequence of pointwise projections of $\{\boldsymbol{\theta}(t)\}$ on the subspace $\mathcal{F}_{m|k}$.

IV. Fast local basis function estimator

The simplified, fast version of the local basis function estimator (fLBF) can be obtained by replacing in (14) the covariance matrix $\boldsymbol{\Phi}$ with its local estimate. This results in

$$\widetilde{\boldsymbol{\theta}}_{m|k}(t) = \widehat{\boldsymbol{\Phi}}_k^{-1}(t) \sum_{i=-k}^{k} h_{m|k}(i) y(t+i)\boldsymbol{\varphi}(t+i) \tag{20}$$

where

$$\widehat{\boldsymbol{\Phi}}_k(t) = \frac{1}{L_k} \sum_{i=-k}^{k} \boldsymbol{\varphi}(t+i)\boldsymbol{\varphi}^{\mathrm{T}}(t+i) \tag{21}$$

denotes the local estimate of $\boldsymbol{\Phi}$ and $L_k = 2k+1$ is the width of the local analysis window.

The fLBF estimator is computationally cheap. First of all, note that while evaluation of (7) - (8) requires inversion of a $mn \times mn$ matrix $\mathbf{R}_{m|k}(t)$, in the case of the fast estimator the size of the inverted matrix is reduced to (only) $n \times n$. Moreover, since the matrix $\widehat{\boldsymbol{\Phi}}_k(t)$ can be updated recursively

$$\begin{aligned} \widehat{\boldsymbol{\Phi}}_k(t) &= \widehat{\boldsymbol{\Phi}}_k(t-1) + \frac{1}{L_k}\boldsymbol{\varphi}(t+k)\boldsymbol{\varphi}^{\mathrm{T}}(t+k) \\ &\quad - \frac{1}{L_k}\boldsymbol{\varphi}(t-k-1)\boldsymbol{\varphi}^{\mathrm{T}}(t-k-1) \end{aligned} \tag{22}$$

one can also easily derive a recursive algorithm for evaluation of $\widehat{\boldsymbol{\Phi}}_k^{-1}(t)$ – see e.g. Section 3.2 in [10].

Second, if basis functions are recursively computable, the elements of the vector

$$\mathbf{p}_{m|k}(t) = \sum_{i=-k}^{k} h_{m|k}(i) y(t+i)\boldsymbol{\varphi}(t+i) \tag{23}$$

can be also computed recursively. Actually, denote by $p^l_{m|k}(t)$ the l-th component of $\mathbf{p}_{m|k}(t)$ and observe that

$$\begin{aligned} p^l_{m|k}(t) &= \widetilde{\mathbf{f}}^{\mathrm{T}}_{m|k}(0)\mathbf{s}^l_{m|k}(t) \\ \mathbf{s}^l_{m|k}(t) &= \sum_{i=-k}^{k} \widetilde{\mathbf{f}}_{m|k}(i) y(t+i) u(t-l+i). \end{aligned} \tag{24}$$

Finally, note that in the case of the polynomial basis (10), the vectors $\mathbf{s}^1_{m|k}(t), \ldots, \mathbf{s}^m_{m|k}(t)$ can be computed recursively using the formula

$$\begin{aligned} \mathbf{s}^l_{m|k}(t) = \mathbf{A}_{m|k}\big[&\mathbf{s}^l_{m|k}(t-1) \\ &- y(t-k-1)u(t-k-l-1)\widetilde{\mathbf{f}}_{m|k}(i-k)\big] \\ &+ y(t+k)u(t+k-l)\widetilde{\mathbf{f}}_{m|k}(i+k) \end{aligned} \tag{25}$$

Remark 1

Note that when the polynomial basis is adopted, and $m = 1$, the LBF estimator is identical with the sliding window least squares estimator.

Remark 2

Since all eigenvalues of the matrix $\mathbf{A}_{m|k}$ lie on the unit circle in the complex plane, the algorithm (25) is not exponentially stable, but only marginally stable. For this reason, to prevent numerical errors from unbounded growth, it is recommended that the quantities $\mathbf{s}^l_{m|k}(t), l = 1, \ldots, n$, are periodically reset to the values obtained via direct (nonrecursive) computation. The same advice applies to the subtract-add algorithms for computation of $\widehat{\mathbf{\Phi}}_k(t)$ and $\widehat{\mathbf{\Phi}}_k^{-1}(t)$

V. Adaptive tuning of fast basis function estimators

Tracking performance of LBF estimators depends on two design parameters – the size of the local analysis window (k) and the number of adopted basis functions (m). It is known that the bias component of the mean square parameter tracking error (MSE) is inversely proportional to m and proportional to k, and the variance component of MSE is proportional to m and inversely proportional to k. Since MSE is the sum of both components, to guarantee good tracking, some sort of a compromise should be reached.

The solution proposed in this paper is based on parallel estimation. Suppose that MK fLBF estimation algorithms, equipped with different settings $m \in \mathcal{M} = \{m_1, \ldots, m_M\}$ and $k \in \mathcal{K} = \{k_1, \ldots, k_K\}$, are run simultaneously and compared using the local quality measure $J_{m|k}(t)$. At each time instant t we will choose from the set of MK competing estimates the one that is locally the best $\widehat{\boldsymbol{\theta}}_{\widehat{m}(t)|\widehat{k}(t)}(t)$ where

$$\{\widehat{m}(t), \widehat{k}(t)\} = \arg \min_{\substack{m \in \mathcal{M} \\ k \in \mathcal{K}}} J_{m|k}(t) \; . \tag{26}$$

To assess the local tracking capability of the compared estimation algorithms, we will use the method of cross-validation. Our key observation is that the fLFB estimator is a solution of the following quadratic minimization problem

$$\begin{aligned} &\widetilde{\boldsymbol{\theta}}_{m|k}(t) = \\ &= \arg\min_{\boldsymbol{\theta}} \sum_{i=-k}^{k} [L_k h_{m|k}(i) y(t+i) - \boldsymbol{\varphi}^{\mathrm{T}}(t+i)\boldsymbol{\theta}]^2 \\ &= \mathbf{W}_k^{-1}(t)\mathbf{w}_{m|k}(t) \end{aligned} \tag{27}$$

where $\mathbf{W}_k(t) = L_k \widehat{\mathbf{\Phi}}_k(t)$ and $\mathbf{w}_{m|k}(t) = L_k \mathbf{p}_{m|k}(t)$.

As a measure of fit we will use the sum of squared unbiased interpolation errors

$$J_{m|k}(t) = \sum_{l=-L}^{L} [\varepsilon^{\circ}_{m|k}(t+l)]^2$$

where L determines the size of the local decision window,

$$\varepsilon^{\circ}_{m|k}(t) = y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widetilde{\boldsymbol{\theta}}^{\circ}_{m|k}(t) \tag{28}$$

and $\widetilde{\boldsymbol{\theta}}^{\circ}_{m|k}(t)$ denotes parameter estimate obtained after excluding the measurement collected at the instant t from the available data set

$$\begin{aligned} &\widetilde{\boldsymbol{\theta}}^{\circ}_{m|k}(t) = \\ &= \arg\min_{\boldsymbol{\theta}} \sum_{\substack{i=-k \\ i \neq 0}}^{k} [L_k h_{m|k}(i) y(t+i) - \boldsymbol{\varphi}^{\mathrm{T}}(t+i)\boldsymbol{\theta}]^2 \\ &= [\mathbf{W}^{\circ}_k(t)]^{-1}\mathbf{w}^{\circ}_{m|k}(t) \end{aligned} \tag{29}$$

where

$$\begin{aligned} \mathbf{W}^{\circ}_k(t) &= \mathbf{W}_k(t) - \boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t) \\ \mathbf{w}^{\circ}_{m|k}(t) &= \mathbf{w}_{m|k}(t) - L_k h_{m|k}(0) y(t)\boldsymbol{\varphi}(t). \end{aligned} \tag{30}$$

Incorporating (30) and using the matrix inversion lemma [10], one obtains

$$\begin{aligned} &\boldsymbol{\varphi}^{\mathrm{T}}(t)\widetilde{\boldsymbol{\theta}}^{\circ}_{m|k}(t) \\ &= \boldsymbol{\varphi}^{\mathrm{T}}(t)\left[\mathbf{W}_k^{-1}(t) + \frac{\mathbf{W}_k^{-1}(t)\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\mathbf{W}_k^{-1}(t)}{1 - \boldsymbol{\varphi}^{\mathrm{T}}(t)\mathbf{W}_k^{-1}(t)\boldsymbol{\varphi}(t)}\right] \\ &\times \left[\mathbf{W}_k(t)\widetilde{\boldsymbol{\theta}}_{m|k}(t) - L_k h_{m|k}(0) y(t)\boldsymbol{\varphi}(t)\right] = \boldsymbol{\varphi}^{\mathrm{T}}(t)\widetilde{\boldsymbol{\theta}}_{m|k}(t) \\ &- \frac{q_k(t)}{1 - q_k(t)}\left[L_k h_{m|k}(0) y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widetilde{\boldsymbol{\theta}}_{m|k}(t)\right] \end{aligned}$$

where

$$q_k(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\mathbf{W}_k^{-1}(t)\boldsymbol{\varphi}(t).$$

Combining the last result with (28), one finally arrives at

$$\begin{aligned} \varepsilon^{\circ}_{m|k}(t) &= \varepsilon_{m|k}(t) \\ &+ \frac{q_k(t)}{1 - q_k(t)}\left[L_k h_{m|k}(0) y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widetilde{\boldsymbol{\theta}}_{m|k}(t)\right] \end{aligned} \tag{31}$$

where

$$\varepsilon_{m|k}(t) = y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widetilde{\boldsymbol{\theta}}_{m|k}(t) \; .$$

According to (31), evaluation of unbiased interpolation errors does not require evaluation of the modified estimates (29).

Remark 3

For the basis (10) and $m = 1$ it holds that $L_k h_{1|k}(0) = 1$, leading to

$$\varepsilon^{\circ}_{1|k}(t) = \frac{\varepsilon_{1|k}(t)}{1 - q_k(t)} \; .$$

VI. Simulation results

In our simulation experiment parameter estimation was carried out for a stationary two-tap FIR system governed by

$$y(t) = \theta_1(t)u(t-1) + \theta_2(t)u(t-2) + e(t).$$

The input signal was autoregressive Gaussian ($r_u(i) = (0.8)^{|i|}$). System parameters were modeled as triangular chirps with two different linearly increasing beat frequencies – see Fig. 1. The variance of the measurement noise was set to $\sigma_e^2 = 0.005$ which corresponds to SNR=15 dB.

Table I shows comparison of the mean squared parameter estimation errors obtained for 9 LBF/fLBF estimators (polynomial basis) corresponding to different choices of design parameters: local analysis window size k (50, 100, 200) and the number of basis functions m (1, 3, 5), and for the adaptive parallel estimation schemes based on cross-validation ($L = 30$). The dynamics of the model selection process is illustrated by decision histograms shown in Fig. 2.

Note that the adaptive schemes yield results that are better or equal to those provided by the best LBF/fLBF estimators with fixed settings. As expected, the price paid for computational simplicity of fLBF estimators is paid in terms of estimation

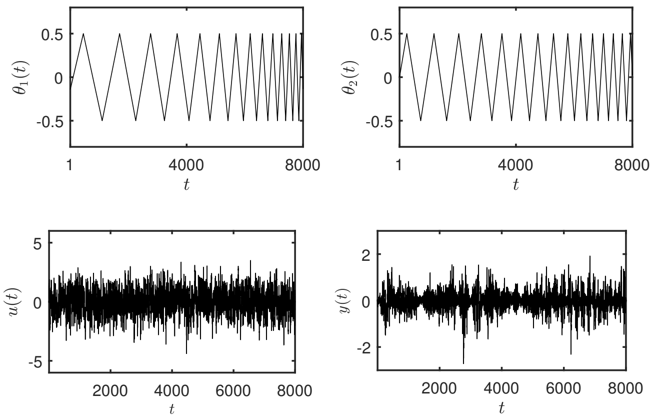

Figure 1: Evolution of system parameters (two upper plots) and typical realizations of the input/output signals (two lower plots).

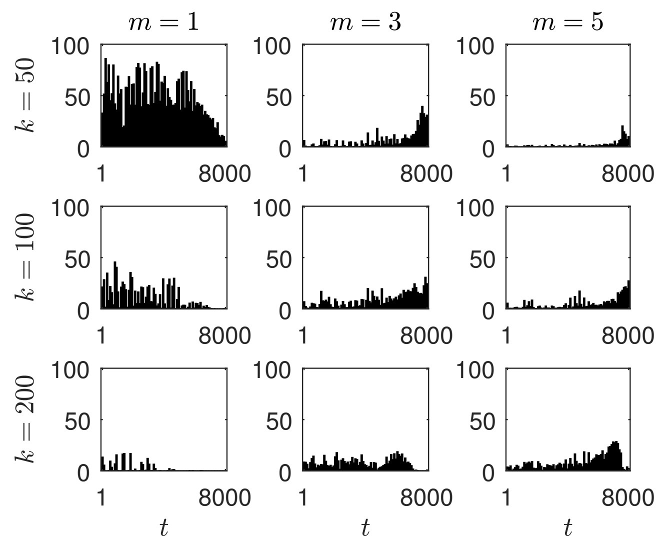

Figure 2: Histograms of decisions on the window size (k) and the number of basis functions (m) obtained for 100 realizations of the fLBF estimator. Each time bin corresponds to 100 consecutive samples.

accuracy – the MSE yielded by the LBF-based scheme was was (for the same process realizations) more than 3 times lower than the analogous score yielded by the fLBF-based scheme (see also Fig. 3).

VII. Conclusion

Identification of nonstationary stochastic systems can be carried out using the local basis function (LBF) approach. While providing good parameter tracking results, the LBF estimators are computationally demanding, especially for high dimensions of the adopted functional basis. We have shown that this computational load can be significantly reduced (at the cost of some estimation accuracy deterioration) if the simplified, fast version of the LBF algorithm (fLBF) is used instead of the original one. It was also shown that two important parameters of fLBF estimators – the number of basis functions and the size of the local analysis window – can be chosen in an adaptive way using the cross-validation technique. The proposed parallel estimation scheme yields good tracking results at a moderate computational cost.

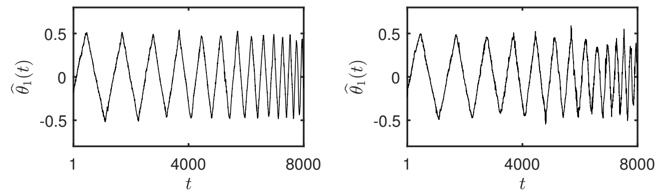

Figure 3: Comparison of parameter estimates yielded by the original parallel estimation scheme (LBF - left plot) and its fast (simplified) version (fLBF - right plot).

Table I: Mean squared parameter estimation errors obtained for 9 LBF/fLBF estimators corresponding to different choices of design parameters k (50, 100, 200) and m (1, 3, 5), and for the proposed adaptive estimation schemes (A). All averages were computed for 100 process realizations.

	LBF			fLBF		
$k \backslash m$	1	3	5	1	3	5
50	0.0038	0.0008	0.0011	0.0038	0.0122	0.0236
100	0.0193	0.0017	0.0009	0.0193	0.0077	0.0128
200	0.0772	0.0162	0.0035	0.0772	0.0192	0.0098
A	0.0008			0.0027		

References

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

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

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

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

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

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

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

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

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

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

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