IEEE Copyright Notice

- ©20xx IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

- This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author’s copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder.
New Approach to Noncausal Identification of Nonstationary Stochastic FIR Systems Subject to Both Smooth and Abrupt Parameter Changes

Maciej Niedźwiecki and Szymon Gackowski

Abstract—In this technical note, we consider the problem of finite-interval parameter smoothing for a class of nonstationary linear stochastic systems subject to both smooth and abrupt parameter changes. The proposed parallel estimation scheme combines the estimates yielded by several exponentially weighted basis function algorithms. The resulting smoother automatically adjusts its smoothing bandwidth to the type and rate of nonstationarity of the identified system. It also allows one to account for the distribution of the measurement noise.

Index Terms—Identification of nonstationary systems, parameter smoothing.

I. INTRODUCTION

In this technical note, we consider the problem of noncausal identification, i.e., identification based on pre-recorded data, of a nonstationary linear stochastic system subject to a mixed-mode parameter variation—a combination of smooth persistent parameter changes and occasional parameter jumps. In the statistical literature, such an approach to system identification is often referred to as fixed-interval parameter smoothing.

An interesting application which admits such problem formulation is identification of rapidly fading mobile radio channels [1]. Slow time variation of the channel coefficients is due to receiver movement in space, and rapid variation is caused by abrupt changes of the spatial system configuration (e.g., due to switching between base stations). Fixed-interval parameter smoothing can be used to reconstruct trajectories of channel coefficients. The reconstructed trajectories can be later used to perform realistic simulations of a mobile communication system—they serve as “ground truth” in demonstrations and tests.

Most of the existing work on parameter smoothing is based on the stochastic model of parameter variation known as the integrated random walk (IRW) model—the idea goes back to Shiller [2]. Adopting the IRW model of order $m$, one assumes that the vector of system coefficients $\hat{\theta}(t) = [\theta_1(t), \ldots, \theta_n(t)]^T$ obeys the following equation:

$$ \nabla^m \theta(t) = (1 - q^{-1})^m \theta(t) = w(t) $$

(1)

where $\nabla^m (\cdot)$ denotes the $m$-th order difference ($q^{-1}$ stands for the backward shift operator), and $w(t)$ is a white noise process ($I_n$ denotes the $n \times n$ identity matrix).

After state space embedding the problem of estimation of parameters governed by (1) can be solved using the algorithms known as Kalman filters/smoothers [3], [4]. The estimation properties of these algorithms depend on the adopted order $m$ of the IRW model and on the variance quotient (smoothness tradeoff parameter) $\xi = \sigma^2_w / \sigma^2_\epsilon$—see e.g., [5].

To obtain satisfactory estimation results, both quantities should be locally “matched” to the type and speed of parameter variation, as well as to the signal-to-noise ratio. The matching problem can be solved by means of combining results yielded by several, simultaneously operated, Kalman filters/smoothers, equipped with different smoothness constraints. For systems with abrupt parameter changes such robust parallel estimation scheme, called competitive smoother, was described in [6]. It combines, in a statistically meaningful way, the results provided by several forward-time and backward-time Kalman filters. The analogous solution for systems with smooth parameter changes, called cooperative smoother, was proposed in [7]—it combines results yielded by several Kalman smoothing algorithms. Finally, one should mention the interacting multiple model (IMM) approach, originally proposed as a solution to the problem of tracking maneuvering targets [8]. The suitably modified IMM algorithm, which can be used for the purpose of parameter tracking (based on IRW modeling), was described in [9]. In principle, the IMM smoother can be obtained by combining the results presented in [9] and [10]. However, to the best of our knowledge, the problem of IMM-based parameter smoothing has not been explored yet.

The contribution of the technical note is twofold. First, we extend the results presented in [6] and [7] to a new class of estimation algorithms—exponentially weighted basis function (EWBF) trackers/smoothers. The basis function approach to system identification is based on a deterministic model of parameter variation—it is assumed that system parameters can be modeled as linear combinations of deterministic functions of time, the so-called basis functions. The EWBF algorithms are obtained by combining the basis function based system description with the estimation technique known as exponential weighting (or exponential forgetting). Similar to the Kalman-filter-based algorithms, the EWBF algorithms belong to the class of finite memory adaptive filters. However, unlike the Kalman-filter-based algorithms, the estimation memory of EWBF trackers/smoothers can be easily quantified in terms of their design parameters, such as the number of basis functions and the forgetting constant. This makes the EWBF algorithms much easier to handle in practice.

Second, we show how one can merge results yielded by the competitive smoother with those provided by the cooperative smoother. The resulting combined smoother preserves advantages of the component algorithms—similar to the competitive smoother it sharply resolves parameter jumps, and, similar to the cooperative smoother, it accurately follows smooth parameter changes.

II. BASIS FUNCTION APPROACH TO PARAMETER TRACKING AND SMOOTHING

Consider the problem of identification of a discrete-time stochastic system governed by

$$ y(t) = \varphi^T(t) \hat{\theta}(t) + \epsilon(t) $$

(2)

where $\varphi(t) = [\varphi_1(t), \ldots, \varphi_n(t)]^T$ denotes the vector of input (regression) variables and $\epsilon(t)$ denotes white measurement noise.

Suppose that identification of (2) is carried out based on the pre-recorded data set $\Omega(N) = \{y(i), \varphi(i), i \in T\}$, where $T = \{1, \ldots, N\}$ denotes the observation interval of length $N$. Denote by

$$ f_j(t) = t^{j-1}, j = 1, \ldots, m, t \in T $$

(3)

the set of the so-called basis functions. Adopting the basis function approach, one assumes that each time-varying system coefficient can be represented by a linear combination of basis functions (for more...
details on the BF approach, and its comparison with other methods of identification of nonstationary systems, see [5] and [11])

\[
g_i(t) = \sum_{j=1}^{m} \alpha_{ij}f_j(t), \quad i = 1, \ldots, n, \quad t \in T. \tag{4}
\]

Powers of time are the most frequently used general-purpose basis functions (although, in principle, any set linearly independent square summable sequences defined on \(T\) can be used). Such a choice, which stems from the Taylor series expansion theorem and guarantees recursive computability of parameter estimates, can be traced back to Subba Rao [12]. All results presented in this technical note will be derived for the basis (3).

After combining (2) with (4), system equation can be written down in the form

\[
y(t) = \psi^T(t)\alpha + \varepsilon(t) \tag{5}
\]

where \(\alpha = [\alpha_1, \ldots, \alpha_m]^T\) is the \(nm \times 1\) vector of coefficients used to describe the parameter time variation, and \(\psi(t)\) denotes the \(nm \times 1\) generalized regression vector: \(\psi(t) = \psi(t) \otimes f(t)\). The symbol \(\otimes\) is used to denote a Kronecker product of two matrices/vectors.

According to (5), the time-varying process of order \(n\) can be represented by a linear time-invariant model of order \(nm\). The values of the time-varying system parameters \(\theta(t)\) can be easily recovered from \(\alpha\) using the following compact version of (4)

\[
\theta(t) = Z(t)\alpha, \quad Z(t) = I_n \otimes \Gamma^T(t). \tag{6}
\]

A. Parameter Tracking

It would be naive to assume that the expansion coefficients \(\alpha_{ij}\) in (4) are constant in the entire time domain. This assumption would mean, in fact, that the analyzed system could be regarded as reducible, i.e., varying in a perfectly predictable manner. To cope with possible fluctuations in \(\alpha_{ij}\)'s, the method of basis functions can be combined with exponential data weighting (exponential forgetting). Weighting forces the estimation to be more focused on the most recent data samples and therefore allows one to track slow variations in expansion coefficients.

The forward-time and backward-time exponentially weighted basis function (EWBF) parameter trackers take the form

\[
\hat{\theta}_{+}(t) = Z(t)\hat{\alpha}_{+}(t)
\]

\[
\hat{\alpha}_{+}(t) = \arg \min_{\alpha} \sum_{i \in T_{+}(t)} \lambda^{t-i} \left[ y(i) - \psi^T(i)\alpha \right]^2
\]

\[
= G_{+}^{-1}(t)g_{+}(t)
\]

where \(\lambda, 0 < \lambda < 1\), denotes the forgetting constant, \(T_{+}(t) = [1, \ldots, t], T_{-}(t) = [t, \ldots, N]\) and the quantities \(G_{+}(t), g_{+}(t)\) are recursively computable

\[
G_{+}(t) = \lambda G_{+}(t+1) + \psi(t)\psi^T(t)
\]

\[
g_{+}(t) = \lambda g_{+}(t+1) + \psi(t)g_{+}(t).
\]

Using the well-known matrix inversion lemma [5], one can easily set up an algorithm for recursive computation of \([G_{+}(t)]^{-1}\).

B. Parameter Smoothing

The EWBF parameter smoother is a natural complement of the causal and anticausal estimation algorithms presented in the previous subsection. The algorithm has the following form:

\[
\hat{\theta}_{-}(t) = Z(t)\hat{\alpha}_{-}(t)
\]

\[
\hat{\alpha}_{-}(t) = \arg \min_{\alpha} \sum_{i \in T_{-}(t)} \lambda^{t-i} \left[ y(i) - \psi^T(i)\alpha \right]^2
\]

\[
= G_{-}^{-1}(t)g_{-}(t)
\]

where

\[
G_{+}(t) = \sum_{i \in T_{+}(t)} \lambda^{t-i} \psi(i)\psi^T(i), \quad g_{+}(t) = \sum_{i \in T_{+}(t)} \lambda^{t-i} \psi(i)y(i).
\]

It is straightforward to check that the quantities \(G_{+}(t)\) and \(g_{+}(t)\) can be evaluated by means of backward time filtering of the quantities \(G_{-}(t)\) and \(g_{-}(t)\), respectively

\[
G_{+}(t) = \lambda G_{+}(t+1) + (1 - \lambda^2)G_{-}(t)
\]

\[
g_{+}(t) = \lambda g_{+}(t+1) + (1 - \lambda^2)g_{-}(t).
\]

For the basis (3) with only one component \(f_1(t) = 1\), the EWBF smoother (8) is identical with the algorithm proposed in [13].

III. STATISTICAL PROPERTIES OF THE EXPONENTIALLY WEIGHTED BASIS FUNCTION SMOOTHER

To simplify our analysis, we will assume that an infinite observation history is available, i.e., \(G_{+}(t) = \sum_{t=0}^{\infty} \lambda^{t-i} \psi(i)\psi^T(t), g_{+}(t) = \sum_{t=0}^{\infty} \lambda^{t-i} \psi(i)y(i)\). The results of such a steady state analysis will remain valid for the algorithm (8) as long as \(1 < \lambda < \infty\), i.e., everywhere except the boundary regions.

Denote by \(\hat{\theta}_{s}(t) = E[\hat{\theta}_{s}(t)]\) the mean path of the estimates yielded by the EWBF smoother. The mean-squared parameter estimation error can be written down in the form

\[
E \left[ \left( \hat{\theta}_{s}(t) - \theta(t) \right)^2 \right] = \left[ \hat{\theta}_{s}(t) - \theta(t) \right]^2 + \text{tr} \left\{ \text{cov} \left[ \hat{\theta}_{s}(t) \right] \right\}
\]

i.e., it can be decomposed into the bias component and the variance component, respectively. Both components will be examined in some detail below. To arrive at analytical results, we will make the following assumptions:

(A1) The sequence of regression vectors \(\{\psi(t)\}\) is a wide sense stationary and ergodic process with positive definite correlation matrix \(E[\psi(t)\psi^T(t)] = \Phi_0 > 0\).

(A2) \(\{\varepsilon(t)\}\), independent of \(\{\psi(t)\}\), is a sequence of zero-mean, independent and identically distributed random variables with variance \(\sigma_v^2\).

Both assumptions hold true for rapidly fading mobile radio channels.

A. Estimation Bias

Under the assumption (A2), one arrives at

\[
\hat{\theta}_{s}(t) = E \left[ \hat{\theta}_{s}(t) \right] = E \left[ Z(t)G_{+}^{-1}(t)\varepsilon(t) \right],
\]

where \(\varepsilon(t) = \sum_{i=-\infty}^{\infty} \lambda^{t-i} \psi(i)\psi^T(i)\theta(i)\).

To facilitate further analysis, we will express all quantities in terms of orthonormal basis vectors \(f_i(t) = Cf_i(t)\) which obey

\[
\sum_{i=-\infty}^{\infty} \lambda^{t-i} f_i(t) f_i^T(i) = I_m.
\]


The matrix $C$ can be established using the classical Gram-Schmidt procedure. Additionally, we will use the following properties of the basis $(3)$: $g(\cdot)$ is a lower triangular matrix such that $g_{ij} = (\cdot)^+_{ij}$ for $i = 1, \ldots, m$, $j \leq i$, and $B = [b_{ij}]$ is a diagonal matrix such that $b_i = (-1)^{i-1}$ for $i = 1, \ldots, m$. Let $H(t) = I_n \otimes D(t)$ where $D(t) = CA^T$. Observe that $D(t)f(i) = f_i(t - i)$. Note that

$$Z(t)G^{-1}(t)r_i(t) = Z_iG_A^{-1}(t)r_i(t)$$

where $Z_i = Z_iH(t), G_i = H(t)G_iH(t)$ and $r_i(t) = H(t)r_i(t)$. Using the identity $(P \otimes Q)(R \otimes S) = PR \otimes QS$, which holds for Kronecker products, one obtains

$$Z_i = I_n \otimes f_i(0)$$

$$G_i(t) = \sum_{i=-\infty}^{\infty} \lambda^{i-1}\left[\varphi(i)\varphi^T(i)\right] \otimes \left[f_i(t-i) + f_i(t-i)\right]$$

$$r_i(t) = \sum_{i=-\infty}^{\infty} \lambda^{i-1}\left[\varphi(i) \otimes f_i(t-i)\right]$$

Under (A1) and (A2), it holds that $E[G_i(t)] = \Phi_i \otimes I_m$. Furthermore, using the generalized law of large numbers for weighted sums of random variables given by Pruitt [14], one can show that $\lim_{n \to \infty} G_i(t) = E[G_i(t)]$ w.p.1. Hence, when the forgetting constant $\lambda$ is sufficiently close to 1, one can apply the following approximation:

$$G_i^{-1}(t) \approx \left\{E[G_i(t)]\right\}^{-1} = \Phi_i^{-1} \otimes I_m.$$  

Using this approximation, one obtains

$$Z_iG_i^{-1}(t)r_i(t) \approx \sum_{i=-\infty}^{\infty} \lambda^{i-1}\left[\varphi_0^{-1}\varphi(i)\varphi^T(i)\right] \otimes \left[f_i(t-i) + f_i(t-i)\right]$$

which leads to

$$\hat{\theta}_i(t) \approx \sum_{i=-\infty}^{\infty} k_i(t-i)\theta(t-i), \quad k_i(t) = \lambda^{i}\lambda_1^{-1}f_i(0)\lambda_i(t).$$  

According to (12), the mean path of EWBF estimates $\{\hat{\theta}_i(t)\}$ can be approximately viewed as a result of passing the process $\{\theta(t)\}$ through a linear time-invariant filter. We will call the sequence $k_i(t)$ the impulse response associated with the EWBF smoother.

The impulse response $k_i(t)$ can be alternatively expressed in the form

$$k_i(t) = \lambda^{i}\lambda_1^{-1}f_i(t) = \sum_{i=-\infty}^{\infty} \lambda^{i}\varphi(j)\varphi^T(j)^{-1}f(i).$$

Using (11), the error (14) can be rewritten in the following equivalent form:

$$\Delta\hat{\theta}_i(t) = Z_iG_i^{-1}(t)r_i(t) \approx \left[\Phi_i^{-1} \otimes f_i(t)\right] \otimes \left[h_i(t)\right]$$

where $h_i(t) = H(t)h_i(t)$ for $t \geq 0$. Using the identity $(P \otimes Q)(R \otimes S) = PR \otimes QS$, which holds for Kronecker products, one obtains

$$E[h_i(t)]h_i(t) = \sum_{i=-\infty}^{\infty} \lambda^{i-1}\left[\varphi(i) \otimes f_i(t-i)\right]$$

Since $E[h_i(t)] = 0$, it holds that $E[\Delta\theta_i(t)] = \Phi_i^{-1} \otimes f_i(t)$, i.e., the smoothed estimate is approximately unbiased. Exploiting assumptions (A1) and (A2), after elementary calculations, one arrives at

$$E[\Delta\theta_i(t)] = \Phi_i^{-1} \otimes f_i(t)$$

$$\text{cov}[\Delta\theta_i(t)] \approx \left[\Phi_i^{-1} \otimes f_i(t)\right]$$

$$\otimes E[h_i(t)]h_i(t) \left[\Phi_i^{-1} \otimes f_i(t)\right]^{-1}$$

$$\approx \frac{\sigma_\lambda^2}{\lambda_i}$$

where the quantity $\lambda_i$ denotes the equivalent estimation memory length of the EWBF smoother.

The results derived above parallel those obtained earlier for the EWBF tracker [15]. In the case of tracking, one obtains

$$\hat{\theta}\pm(t) = E[\theta(t)] \approx \sum_{i=-\infty}^{\infty} k_i(t)\theta(t \pm i)$$

where

$$k_i(t) = \lambda^i\lambda_1^{-1}f_i(t) = \sum_{i=-\infty}^{\infty} \lambda^{i}\varphi(j)\varphi^T(j)^{-1}f(i), \quad i \geq 0.$$


### TABLE I

**ASSOCIATED IMPULSE RESPONSES AND EQUIVALENT MEMORY SPANS OF EWBF SMOOTHERS AND TRACKERS OF DIFFERENT ORDERS**

**EWBF smoothers**

<table>
<thead>
<tr>
<th>( m )</th>
<th>( k_s(i) = \lambda^i [1 + \frac{1 - \lambda^i}{1 - \lambda}] )</th>
<th>( m_s = \frac{1 - \lambda^i}{1 - \lambda} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( k_s(i) = \lambda^i [1 + \frac{1 - \lambda^i}{1 - \lambda}] )</td>
<td>( m_s = \frac{1 - \lambda^i}{1 - \lambda} )</td>
</tr>
<tr>
<td>2</td>
<td>( k_s(i) = \lambda^i [1 + \frac{1 - \lambda^i}{1 - \lambda}] )</td>
<td>( m_s = \frac{1 - \lambda^i}{1 - \lambda} )</td>
</tr>
<tr>
<td>3</td>
<td>( k_s(i) = \lambda^i [1 + \frac{1 - \lambda^i}{1 - \lambda}] )</td>
<td>( m_s = \frac{1 - \lambda^i}{1 - \lambda} )</td>
</tr>
</tbody>
</table>

**EWBF trackers**

<table>
<thead>
<tr>
<th>( m )</th>
<th>( k_{\pm}(i) = \lambda^i (1 - \lambda) )</th>
<th>( m_{\pm} = \frac{1}{2} \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( k_{\pm}(i) = \lambda^i (1 - \lambda) )</td>
<td>( m_{\pm} = \frac{1}{2} \lambda )</td>
</tr>
<tr>
<td>2</td>
<td>( k_{\pm}(i) = \lambda^i (1 - \lambda) )</td>
<td>( m_{\pm} = \frac{1}{2} \lambda )</td>
</tr>
</tbody>
</table>

**Note:** For Kalman algorithms based on (1) the situation is different. When \( m = 1 \), one can show that the estimation memory takes the form \( m_{\pm} = c_{\pm}/(1 - \lambda) \) and \( m_{\pm} = c_{\pm}/(1 - \lambda) \), where \( c_{\pm} \) and \( c_{\pm} \) are easily computable quantities that depend on the number of basis functions and the forgetting constant, and do not depend on the characteristics of the input-output data.

**Remark:** When estimation is carried out using the exponentially weighted basis function algorithms, the forgetting constant should be chosen with caution—the rules of thumb which work for the exponentially weighted least squares tracker cannot be mechanically extended to higher-order EWBF estimators. For example, when \( \lambda = 0.99 \) the equivalent memory spans of the third-order estimators are equal to \( L_\alpha \approx 33 \) samples and \( L_\alpha \approx 48 \) samples, rather than \( L_\alpha \approx 200 \) samples and \( L_\alpha \approx 400 \) samples, respectively, as one might expect based on the experience with exponentially weighted least squares trackers/smoothers.

### IV. COOPERATIVE EWBF SMOOTHER

Estimation properties of EWBF algorithms depend on the choice of two design parameters: the forgetting constant \( \lambda \) and the number of basis functions \( m \). For “small” values of \( \lambda \) and/or for “large” values of \( m \), the estimation memory of the EWBF algorithm increases. The resulting short-memory trackers/smoothers are “fast” (yield small estimation bias) but “inaccurate” (yield large estimation variance). The converse is also true—long-memory algorithms are “accurate” but “slow”. The best results are obtained if the values of \( \lambda \) and \( m \) are selected so as to match the degree and type of nonstationarity of the identified system, trading off the bias and variance error components. Optimization of the design parameters is possible using parallel estimation techniques. Following this idea, consider simultaneously running EWBF smoothers, equipped with different settings \( \{\lambda_k, m_k\} \) and yielding the estimates

\[
\hat{\theta}_k(t) = \mathbf{Z}_k(t) [\mathbf{G}_k(t)]^{-1} \mathbf{g}_k(t),
\]

\[
\mathbf{Z}_k(t) - \mathbf{I} \otimes \mathbf{f}_k(t)
\]

\[
\mathbf{G}_k(t) = \sum_{i=-\infty}^{\infty} \lambda^{i+1} \mathbf{g}_k(i) \mathbf{g}_k(i)^T
\]

\[
\mathbf{g}_k(t) = \sum_{i=-\infty}^{\infty} \lambda^{i+1} \mathbf{g}_k(i) y(i).
\]

We will combine these estimates using the cooperative smoothing strategy proposed in [13].

Cooperative smoothing is a Bayesian extension of the leave-one-out cross-validation approach to model selection. In this approach, credibility of each smoother \( \hat{\theta}_k(t) \) is assessed (locally) based on observation of matching errors \( e_k(t) = y(t) - \mathbf{\varphi}_k(t) \hat{\theta}_k(t) \),—residual errors yielded by the “holey” smoother \( \hat{\theta}_k(t) \) associated with \( \hat{\theta}_k(t) \). Holey smoother associated with the EWBF algorithm, i.e., the one that excludes \( y(t) \) from the set of measurements used for estimation of \( \theta_k(t) \), takes the form

\[
\hat{\theta}_k(t) = \mathbf{Z}_k(t) [\mathbf{G}_k(t)]^{-1} \mathbf{g}_k(t),
\]

\[
\mathbf{G}_k(t) = \mathbf{G}_k(t) - \mathbf{\varphi}_k(t) \mathbf{\varphi}_k(t)^T
\]

\[
\mathbf{g}_k(t) = \mathbf{g}_k(t) - \mathbf{\varphi}_k(t) y(t).
\]

Suppose that the measurement noise \( \{\epsilon(t)\} \) is a sequence of zero-mean independent random variables obeying the generalized normal law \( \epsilon \sim \mathcal{GN}(\alpha, \beta) \):

\[
\epsilon(t; \alpha, \beta) = \frac{\beta}{2\alpha \Gamma(\frac{3}{2})} \exp \left\{ -\frac{1}{\alpha^2} \right\}\right.^{\beta/2}
\]

where \( \alpha > 0 \) is the unknown scale parameter, \( \beta > 1 \) is the known shape parameter, and \( \Gamma(\cdot) \) denotes the Euler’s gamma function. Generalized normal law incorporates such practically important distributions as Gaussian (\( \beta = 2 \)), Laplace (\( \beta = 1 \)), and uniform (\( \beta \rightarrow \infty \))—see [16].

According to [13], the cooperative estimate of the parameter trajectory can be obtained in the form

\[
\hat{\theta}_k(t) = \sum_{k=1}^{K} \mu_k(t) \hat{\theta}_k(t), \quad \mu_k(t) \propto \left[ \sum_{i \in T_k(t)} |r_k(i)|^\beta \right]^{-M/(\beta - 1)}
\]

where \( r_k(i) = 0, k = 1, \ldots, K, \sum_{i \in T_k(t)} r_k(i) = 1 \) denote credibility coefficients (related to posterior probabilities of different parameter “patterns”), evaluated in the local decision window \( T_k(t) = [t-i, t+i] \) of (user-dependent) width \( M = 2i + 1 \).

Similarly as in [13], one can show that matching errors can be expressed in terms of the residual errors \( e_k(t) = y(t) - \mathbf{\varphi}_k(t) \hat{\theta}_k(t) \), namely

\[
e_k(t) = \frac{e_k(t)}{1 - r_k(t)}
\]

\[
e_k(t) = \frac{e_k(t)}{1 - r_k(t)}
\]

1After replacing the regression vector \( \mathbf{\varphi}_k(t) \) with the generalized regression vector \( \mathbf{\varphi}_k(t) \), the proof is identical with that given in [13].
where
\[ r_k(t) = \psi_\theta^T(t) [G^*_\theta(t)]^{-1} \psi_\theta(t). \]

Hence, credibility coefficients can be evaluated without implementing the corresponding holey smoothers.

V. COMPETITIVE EWBF SMOOTHER

The smoothing formula (8) is based on an implicit assumption that parameter changes are continuous, i.e., that the true parameter trajectory obeys the polynomial model at least in some neighborhood of \( t \). Quite obviously, the presence of jumps in the parameter trajectory violates this “homogeneity” assumption. Note, however, that when parameter changes are infrequent, namely when the time between subsequent jumps is greater than the equivalent memory length of the forward/backward EWBF trackers, at each time instant at least one of them generates accurate estimates of the parameter trajectory. In particular, the forward time tracker yields accurate estimates immediately before the jumps, and the backward time tracker—immediately after the jumps. The degree of accurateness, i.e., “credibility” of different estimators, can be easily judged by the magnitude of the locally observed prediction errors. Following [6], the competitive smoother that combines results yielded by \( K \) forward time and \( K \) backward time EWBF trackers can be set up as follows:

\[ \begin{align*}
\hat{\theta}_t(t) &= \sum_{k=1}^{K} \left[ \mu_k^-(t) \hat{\theta}_k^-(t) + \mu_k^+(t) \hat{\theta}_k^+(t) \right] \\
\mu_k^\pm(t) &= \left[ \sum_{l \in T_k^\pm} \left| \varepsilon_{\theta}^l(t) \right|^2 \right]^{-\frac{1}{2}} \quad (21)
\end{align*} \]

where \( T_k^- = |t - M + 1, t| \) and \( T_k^+ = |t, t + M - 1| \) denote the local decision windows of width \( M \), and

\[ e_k^\pm(t) = y(t) - \varphi^T(t) \hat{\theta}_k^\pm(t \pm 1), \quad k = 1, \ldots, K \]

are the corresponding forward/backward prediction errors.

The term “competitive smoothing” refers to the fact that, within this approach, the forward and backward algorithms compete with each other, rather than cooperate (since, at each time instant, only one of them is assumed to be based on the correct system description).

VI. COMBINED EWBF SMOOTHER

While the competitive smoother yields good estimation results in the regions of parameter jumps, for smooth parameter changes it may lead to parameter “jitter”—the effect caused by random switching between the forward-time and backward-time algorithms—see [6]. The converse is true for the cooperative smoother: it yields good results for smooth parameter changes, but “blurs” all step-like features of the estimated parameter trajectory.

To obtain smoothing algorithm that works satisfactorily in both cases mentioned above, one can fuse the cooperative smoother (18) with the competitive smoother (21). Using the cooperative smoothing rule once again, one arrives at the following formula which will be further referred to as a combined smoother:

\[ \hat{\theta}_c(t) = \mu_A(t) \hat{\theta}_A(t) + \mu_B(t) \hat{\theta}_B(t) \quad (22) \]

where

\[ \begin{align*}
\mu_A(t) &= \frac{\eta_A(t)}{\eta_A(t) + \eta_B(t)} \\
\mu_B(t) &= \frac{\eta_B(t)}{\eta_A(t) + \eta_B(t)} \\
\eta_A(t) &= \left[ \sum_{l \in T_k^A(t)} \left| e_{\theta,k}^l(t) \right|^2 \right]^{-\frac{1}{2}} \\
\eta_B(t) &= \left[ \sum_{l \in T_k^B(t)} \left| e_{\theta,k}^l(t) \right|^2 \right]^{-\frac{1}{2}}
\end{align*} \]
TABLE II
COMPARISON OF PARAMETER ESTIMATION ERRORS OBTAINED FOR THREE EWBF SMOOTHERS BASED ON THE POLYNOMIAL MODEL OF ORDER 0 \(\{S_1^0, S_2^0, S_3^0\}\), THREE EWBF SMOOTHERS BASED ON THE POLYNOMIAL MODEL OF ORDER 1 \(\{S_1^1, S_2^1, S_3^1\}\), THREE EWBF SMOOTHERS BASED ON THE POLYNOMIAL MODEL OF ORDER 2 \(\{S_1^2, S_2^2, S_3^2\}\), COOPERATIVE SMOOTHER \(S_A\), COMPETITIVE SMOOTHER \(S_B\) AND COMBINED SMOOTHER \(S_C\). SIMULATIONS WERE PERFORMED FOR TWO VARIANTS OF PARAMETER CHANGES (I, II). THE BEST RESULTS ARE SHOWN IN BOLDFACE

<table>
<thead>
<tr>
<th>T</th>
<th>(\sigma_u)</th>
<th>(S_1^0)</th>
<th>(S_2^0)</th>
<th>(S_3^0)</th>
<th>(S_1^1)</th>
<th>(S_2^1)</th>
<th>(S_3^1)</th>
<th>(S_1^2)</th>
<th>(S_2^2)</th>
<th>(S_3^2)</th>
<th>(S_A)</th>
<th>(S_B)</th>
<th>(S_C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.05</td>
<td>5.48</td>
<td>16.82</td>
<td>54.80</td>
<td>13.75</td>
<td>44.78</td>
<td>125.92</td>
<td>35.99</td>
<td>110.48</td>
<td>228.80</td>
<td>5.45</td>
<td>1.04</td>
<td>1.00</td>
</tr>
<tr>
<td>0.10</td>
<td>7.02</td>
<td>17.31</td>
<td>54.97</td>
<td>14.39</td>
<td>44.98</td>
<td>125.99</td>
<td>36.24</td>
<td>110.56</td>
<td>228.88</td>
<td>6.19</td>
<td>3.81</td>
<td>3.44</td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td>12.13</td>
<td>18.88</td>
<td>55.32</td>
<td>16.42</td>
<td>45.47</td>
<td>126.02</td>
<td>36.94</td>
<td>110.71</td>
<td>228.78</td>
<td>8.47</td>
<td>8.40</td>
<td>6.91</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>0.05</td>
<td>1.13</td>
<td>2.60</td>
<td>7.90</td>
<td>2.13</td>
<td>6.41</td>
<td>28.53</td>
<td>5.04</td>
<td>18.16</td>
<td>108.53</td>
<td>1.10</td>
<td>0.24</td>
<td>0.14</td>
</tr>
<tr>
<td>0.15</td>
<td>2.66</td>
<td>3.09</td>
<td>8.07</td>
<td>2.76</td>
<td>6.61</td>
<td>28.62</td>
<td>5.29</td>
<td>18.28</td>
<td>108.51</td>
<td>1.69</td>
<td>1.59</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td>7.75</td>
<td>4.76</td>
<td>8.60</td>
<td>4.86</td>
<td>7.27</td>
<td>28.84</td>
<td>6.12</td>
<td>18.54</td>
<td>108.66</td>
<td>3.50</td>
<td>5.38</td>
<td>3.46</td>
<td></td>
</tr>
</tbody>
</table>

Laplace noise

<table>
<thead>
<tr>
<th>T</th>
<th>(\sigma_u)</th>
<th>(S_1^0)</th>
<th>(S_2^0)</th>
<th>(S_3^0)</th>
<th>(S_1^1)</th>
<th>(S_2^1)</th>
<th>(S_3^1)</th>
<th>(S_1^2)</th>
<th>(S_2^2)</th>
<th>(S_3^2)</th>
<th>(S_A)</th>
<th>(S_B)</th>
<th>(S_C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.05</td>
<td>5.47</td>
<td>16.81</td>
<td>54.80</td>
<td>13.74</td>
<td>44.77</td>
<td>125.93</td>
<td>35.98</td>
<td>110.49</td>
<td>228.82</td>
<td>5.38</td>
<td>1.38</td>
<td>1.35</td>
</tr>
<tr>
<td>0.15</td>
<td>7.04</td>
<td>17.34</td>
<td>54.93</td>
<td>14.40</td>
<td>44.96</td>
<td>125.90</td>
<td>36.27</td>
<td>110.50</td>
<td>228.74</td>
<td>6.07</td>
<td>4.02</td>
<td>3.69</td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td>12.29</td>
<td>19.10</td>
<td>55.57</td>
<td>16.60</td>
<td>45.72</td>
<td>126.15</td>
<td>37.18</td>
<td>110.82</td>
<td>228.83</td>
<td>8.15</td>
<td>8.45</td>
<td>7.08</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>0.05</td>
<td>1.13</td>
<td>2.60</td>
<td>7.90</td>
<td>2.14</td>
<td>6.41</td>
<td>28.54</td>
<td>5.04</td>
<td>18.16</td>
<td>108.54</td>
<td>1.02</td>
<td>0.21</td>
<td>0.12</td>
</tr>
<tr>
<td>0.15</td>
<td>2.66</td>
<td>3.08</td>
<td>8.05</td>
<td>2.76</td>
<td>6.59</td>
<td>28.64</td>
<td>5.28</td>
<td>18.24</td>
<td>108.67</td>
<td>1.52</td>
<td>1.39</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td>7.83</td>
<td>4.84</td>
<td>8.66</td>
<td>4.95</td>
<td>7.35</td>
<td>28.81</td>
<td>6.23</td>
<td>18.51</td>
<td>108.68</td>
<td>3.03</td>
<td>4.79</td>
<td>3.04</td>
<td></td>
</tr>
</tbody>
</table>

and the corresponding matching errors are given by

\[
e^c_A(t) = \sum_{k=1}^{K} \mu_k^c(t)z_k(t) \quad e^c_B(t) = \sum_{k=1}^{K} \mu_k^c(t)z_k^c(t) + \mu_k^b(t)z_k^b(t)
\]

Note that the computational load of the combined smoother grows linearly with the number of component algorithms \(K\). The analogous count for the IMM smoother proposed in [8] is proportional to \(K^2\).

VII. COMPUTER SIMULATIONS

To check performance of smoothing algorithms, the following two-tap FIR system (inspired by channel equalization applications) was simulated

\[
y(t) = \theta_1(t)u(t - 1) + \theta_2(t)u(t - 2) + v(t)
\]

where \(u(t) = \pm 1\), \(\sigma_u^2 = 1\), denotes the pseudo-random binary signal (PRBS)—the sequence transmitted over a telecommunication channel—and \(v(t)\) denotes a zero-mean white noise.

Two types of parameter changes were considered: piecewise constant (I), and piecewise sinusoidal (II)—see Fig. 1. Due to appropriate scaling, all parameter trajectories have the same \(L_2\) norm.

For each of the compared algorithms the steady state accumulated mean-squared parameter estimation errors \(E_{\text{A}}[\sum_{t=1}^{T} (\theta(t)^{5000}) - \overline{\theta}(t)^2]\) were computed (to eliminate transient effects, the evaluation was restricted to the interval \([101, 4900]\)). Ensemble averaging \(E_{\text{A}}[\cdot]\) was performed over 50 realizations of the measurement noise \(x(t)\).

The procedure was repeated for each of three noise intensities ranging from \(\sigma_u = 0.05\) (SNR = 26 dB) to \(\sigma_u = 0.30\) (SNR = 10.5 dB). The distribution of noise was either Gaussian (\(\beta = 2\)) or Laplacian (\(\beta - 1\)). The width of the decision window was equal to \(M = 21\).

Table II summarizes results obtained for 9 EWBF smoothers \(\{S_k^j, j = 1, \ldots, 3\}\), the cooperative smoother \(S_A\), the competitive smoother \(S_B\) and the combined smoother \(S_C\). The first 3 EWBF algorithms \(\{S_1^1, S_2^1, S_3^1\}\) were based on the polynomial model of order 0 \((m = 1, \lambda_1^1 = 0.818, \lambda_2^1 = 0.920, \lambda_3^1 = 0.975)\). The next 3 EWBF algorithms \(\{S_1^2, S_2^2, S_3^2\}\) were based on the polynomial model of order 1 \((m = 2, \lambda_1^1 = 0.936, \lambda_2^1 = 0.973, \lambda_3^1 = 0.984)\). The last 3 EWBF algorithms \(\{S_1^3, S_2^3, S_3^3\}\) were based on the polynomial model of order 2 \((m = 3, \lambda_1^3 = 0.978, \lambda_2^3 = 0.991, \lambda_3^3 = 0.995)\). The forgetting constants \(\lambda_k^j\) of constituent smoothers were not optimized in any way—the corresponding values were chosen so that within each group of algorithms \(\{S_1^j, S_2^j, S_3^j\}\), \(\{S_1^2, S_2^2, S_3^2\}\) and \(\{S_3^3, S_3^3\}\)
the equivalent memory spans of the corresponding EWBF trackers were the same and equal to 10 samples, 30 samples and 90 samples, respectively. Typical estimation results (Gaussian noise, $\sigma_e = 0.2$) are shown in Fig. 2.

Note that the combined smoother is uniformly better than all other smoothers, including all component smoothers, i.e., it provides the smallest estimation errors in all cases considered. An obvious advantage of the combined smoother is its increased robustness to unknown and possibly time-varying degree of nonstationarity and the mode of variation of the identified system.

VIII. CONCLUSION

We have shown that, by combining the results yielded by several exponentially weighted basis function parameter trackers/smoothers (run simultaneously and equipped with different settings), one obtains a new fixed-interval smoothing algorithm with improved estimation capabilities. The proposed smoother can be applied to noncausal identifications of nonstationary stochastic systems subject to both smooth and abrupt parameter changes. Additionally, it accounts for the distribution of measurement noise.

REFERENCES


State Feedback Stabilization for Boolean Control Networks

Rui Li, Meng Yang, and Tianguang Chu

Abstract—State feedback stabilization for Boolean control networks is investigated in this technical note. Based on the algebraic representation of logical dynamics in terms of the semi-tensor product of matrices, a necessary and sufficient condition is derived for the existence of a globally stabilizing state feedback controller, and a general control design approach is proposed when global stabilization is feasible via state feedback. Instead of designing the logical form of a stabilizing feedback law directly, we first construct its algebraic representation and then convert the algebraic representation back to the logical form. An example is worked out to illustrate the proposed design procedure.

Index Terms—Algebraic representation, Boolean control networks, semi-tensor product, state feedback stabilization, systems biology.

I. INTRODUCTION

Boolean networks (BNs) are discrete-time dynamical systems with binary state variables. Since Kauffman’s seminal work [1], BNs have been used extensively as abstract modeling schemes in many different fields, including cell differentiation [2], immune response [3], biological evolution [4], [5], neural networks [6], and gene regulation [7], [8]. Despite the fact of their conceptual simplicity, BNs provide a nice framework that allows for qualitative and theoretical analysis.

One of the objectives of BN modeling is to use the BN to develop methods for finding control strategies for complex biological systems. When control inputs are added, the concept of BNs extends naturally to that of Boolean control networks (BCNs). To be consistent with the binary nature of BCNs, the control inputs will be allowed to take on one of two values: a 1 value indicating that a particular intervention is being applied at that time point and a 0 value indicating that the application of that intervention has ceased. The development of a control theory for BCNs may provide new insights into the intrinsic control in biological systems and enable us to develop strategies for manipulating complex biological systems using exogenous inputs. There have been many papers in this direction; see, for example, [9]–[19]. For the control of probabilistic Boolean networks and context-sensitive probabilistic Boolean networks, see [20]–[23] and [24]–[28], respectively.

The stabilization problem is a basic issue when studying BCNs. The consideration of such a problem could be naturally motivated in the context of gene therapy applications, where one may want to design therapeutic interventions that steer a particular gene regulatory network to a desirable state, such as the healthy one, and maintain this state afterwards. Also, systematic analysis of the stabilization issue for complex biological systems modeled using BCNs may reveal how the structure and organization of the system contribute to the system stability. In [29], stabilization of BCNs via open-loop control and state feedback was addressed, and some necessary and sufficient conditions were derived. Later, the study was extended to BCNs with impulsive

Manuscript received July 01, 2012; revised October 08, 2012; accepted December 10, 2012. Date of publication January 04, 2013; date of current version June 19, 2013. This work was supported by the National Basic Research Program of China under Grant 2012CB821200 and by the National Natural Science Foundation of China under Grant 61273111 and Grant 60974064. Recommended by Associate Editor L. Schenato.

The authors are with the State Key Laboratory for Turbulence and Complex Systems, College of Engineering, Peking University, Beijing 100871, China (e-mail: rui.li@pku.edu.cn; ym_jeff@hotmail.com; chu@gpku.edu.cn).

Digital Object Identifier 10.1109/TAC.2013.2238092