

## **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.

# Generalized Adaptive Notch and Comb Filters for Identification of Quasi-Periodically Varying Systems

Maciej Niedźwiecki and Piotr Kaczmarek

**Abstract**—The problem of identification/tracking of quasi-periodically varying real-valued systems is considered. This problem is a generalization, to the system case, of a classical signal processing task of either elimination or extraction of nonstationary sinusoidal signals buried in noise. The solution is based on the exponentially weighted basis function (EWBF) approach. The proposed algorithms are capable of tracking slow changes in system frequencies, which means that not only the expansion coefficients in the basis function description of the analyzed system but also the basis functions themselves are adjusted in an adaptive manner. First, assuming that the system frequencies are known and constant, the running basis and fixed basis variants of the EWBF algorithm are derived, and their relationship to the classical notch filter with constrained poles and zeros is established. Next, the frequency-adaptive versions of both algorithms are obtained using the gradient search and recursive prediction error principles, respectively. Finally, the interrelated frequencies case is analyzed and two additional parameter tracking algorithms (generalized adaptive comb filters) are derived.

**Index Terms**—Basis function approach, frequency estimation, system identification, time-varying processes.

## I. INTRODUCTION

CONSIDER the problem of identification/tracking of coefficients of a time-varying system governed by

$$y(t) = \sum_{l=1}^n \theta_l(t) \varphi_l(t) + v(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\theta}(t) + v(t) \quad (1)$$

where  $t = 1, 2, \dots$ , denotes the normalized discrete time,  $y(t)$  denotes the system output,  $\boldsymbol{\varphi}(t) = [\varphi_1(t), \dots, \varphi_n(t)]^T$  is the regression vector,  $v(t)$  is an additive white noise, uncorrelated with  $\boldsymbol{\varphi}(t)$ , and  $\boldsymbol{\theta}(t) = [\theta_1(t), \dots, \theta_n(t)]^T$  denotes the vector of time-varying system coefficients, modeled as weighted sums of sinusoidal and cosinusoidal functions

$$\theta_l(t) = c_{l,0}(t) + \sum_{i=1}^k [c_{l,2i-1}(t) \sin \phi_i(t) + c_{l,2i}(t) \cos \phi_i(t)] + c_{l,2k+1}(t) (-1)^t \quad l = 1, \dots, n \quad (2)$$

where

$$\phi_i(t) = \sum_{s=1}^t \omega_i(s). \quad (3)$$

Manuscript received March 29, 2004. This work was supported by KBN under Grant 4 T11A 01225. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ioan Tabus.

The authors are with the Faculty of Electronics, Telecommunications and Computer Science, Department of Automatic Control, Gdańsk University of Technology, Gdańsk, Poland (e-mail: maciekn@eti.pg.gda.pl; piokacz@proterians.net.pl).

Digital Object Identifier 10.1109/TSP.2005.859223

Inclusion of the dc component (with characteristic frequency  $\omega_0 = 0$ ) and ac component (with characteristic frequency  $\omega_{k+1} = \pi$ ) is optional—both terms were added merely for the sake of completeness. Since the amplitudes  $c_{l,i}(t)$  and frequencies  $\omega_i(t)$  in (2) and (3) are time-varying, system parameters change over time in a periodic-like but not exactly periodic manner. We will assume that for every  $i$ , the quantities  $c_{l,i}(t)$ ,  $l = 1, \dots, n$  and  $\omega_i(t)$  are slowly time-varying. The system governed by (1) and (2), which obeys the above-mentioned limitation, will be further referred to as *quasi-periodically* time-varying.

As an example of a process obeying (1), consider the nonstationary autoregressive with exogenous input (ARX) system governed by

$$y(t) = \sum_{l=1}^r a_l(t) y(t-l) + \sum_{l=1}^p b_l(t) u(t-l) + v(t) \quad (4)$$

where  $u(t)$  denotes the measurable input signal. In this case the regression vector  $\boldsymbol{\varphi}(t) = [y(t-1), \dots, y(t-r), u(t-1), \dots, u(t-p)]^T$  is made up of past output and input samples, the parameter vector  $\boldsymbol{\theta}(t) = [a_1(t), \dots, a_r(t), b_1(t), \dots, b_p(t)]^T$  consists of autoregressive coefficients  $a_l(t)$  and direct path coefficients  $b_l(t)$ , and  $n = r + p$ .

One of the challenging applications, based on the ARX system description, is adaptive equalization of rapidly fading communication channels. For mobile radio channels, the sinusoidal model of channel coefficients variation has a long history, which goes back to Aiken [1]. Quite recently, a number of papers explored the possibility of using it for equalization purposes—see, e.g., [2]–[4] and [5]. Even though the references given above deal with complex quadrature amplitude modulation channels, the proposed equalization technique remains valid for fast-varying real (e.g., binary phase-shift keying) channels with a baseband model given by (1) and (2).

Note that in the special case where  $n = 1$ , and  $\varphi(t) = 1, \forall t$ , the model (1) and (2) becomes a description of a nonstationary multifrequency *signal*  $\theta(t)$  buried in noise

$$y(t) = \theta(t) + v(t) \quad (5)$$

$$\theta(t) = c_0(t) + \sum_{i=1}^k [c_{2i-1}(t) \sin \phi_i(t) + c_{2i}(t) \cos \phi_i(t)] + c_{2k+1}(t) (-1)^t. \quad (6)$$

The problem of either elimination or extraction of nonstationary sinusoidal signals contaminated with noise has attracted a great deal of attention of researchers in the field of signal processing. When signal frequencies are unrelated, the solutions proposed in the literature are based on adaptive notch filtering—see

[6]–[9] and the references therein. When the frequencies  $\omega_1(t), \dots, \omega_k(t)$  are interrelated, namely, they are multiples of the same fundamental frequency  $\omega(t)$ , the problem can be solved by means of adaptive comb filtering—see, e.g., [10] and [11]. Since the problem analyzed in our paper is a generalization, to the system case, of the signal elimination/tracking problem mentioned above, the identification algorithms derived for systems with unrelated and interrelated frequency modes can be considered generalized adaptive notch and generalized adaptive comb filters, respectively.

To our best knowledge, this is the first paper dealing with the problem of identification of quasi-periodically varying *real-valued* systems. Identification of complex-valued systems was considered in [3], [5], and [12]. It should be stressed, however, that the results presented below are neither a special case, nor can they be obtained as a trivial extension of known results derived for complex-valued systems.

This paper is organized as follows. In Section II, the identification problem is solved for a system with known constant frequencies of parameter variation. In Section III, the frequency-adaptive versions of two algorithms, described in Section II, are derived for systems with unrelated frequencies. The analogous algorithms for systems with interrelated frequencies are presented in Section IV. Section V provides simulation results, and Section VI concludes.

## II. KNOWN FREQUENCIES

Suppose, for the time being, that both the amplitudes and angular frequencies in (3) are constant, i.e., that system parameters are almost periodic functions of time [13], governed by

$$\theta_l(t) = c_{l,0} + \sum_{i=1}^k [c_{l,2i-1} \sin \omega_i t + c_{l,2i} \cos \omega_i t] + c_{l,2k+1} (-1)^t \quad l = 1, \dots, n. \quad (7)$$

According to (7) each system parameter is modeled as a sum of the dc term, the ac term and  $k$  “regular” terms, which correspond to periodic modes of variation, characterized by frequencies  $0 < \omega_1, \dots, \omega_k < \pi$ .

Denote by  $\mathbf{f}_i(t)$  the basis functions associated with the  $i$ th mode:  $\mathbf{f}_0(t) = 1$ ,  $\mathbf{f}_{k+1}(t) = (-1)^t$ , and

$$\mathbf{f}_i(t) = \begin{bmatrix} \sin \omega_i t \\ \cos \omega_i t \end{bmatrix}, \quad i = 1, \dots, k.$$

Denote by  $\boldsymbol{\alpha}_i$  the corresponding vectors of coefficients  $\boldsymbol{\alpha}_0 = [c_{1,0}, \dots, c_{n,0}]^T$ ,  $\boldsymbol{\alpha}_{k+1} = [c_{1,2k+1}, \dots, c_{n,2k+1}]^T$ ,  $\boldsymbol{\alpha}_i = [c_{1,2i-1}, c_{1,2i}, \dots, c_{n,2i-1}, c_{n,2i}]^T$ ,  $i = 1, \dots, k$ . Finally, let  $\boldsymbol{\psi}_i(t) = \boldsymbol{\varphi}(t) \otimes \mathbf{f}_i(t)$ , where  $\otimes$  denotes the Kronecker product of two matrices/vectors. Note that  $\boldsymbol{\alpha}_i$  is the vector of coefficients associated with a particular frequency  $\omega_i$  and *not* with a particular impulse response parameter  $\theta_i(t)$ . Similarly,  $\boldsymbol{\psi}_i(t)$  is the generalized regression vector associated with the  $i$ th frequency component.

Using the short-hand notation, introduced above, (1) can be rewritten in the form

$$y(t) = \sum_{i=0}^{k+1} \boldsymbol{\psi}_i^T(t) \boldsymbol{\alpha}_i + v(t) = \boldsymbol{\psi}^T(t) \boldsymbol{\alpha} + v(t) \quad (8)$$

where  $\boldsymbol{\psi}(t) = [\boldsymbol{\psi}_0^T(t), \dots, \boldsymbol{\psi}_{k+1}^T(t)]^T$  and  $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_0^T, \dots, \boldsymbol{\alpha}_{k+1}^T]^T$ .

Suppose now that the vector  $\boldsymbol{\alpha}$  is slowly varying with time. It is known that, in the case considered, one can track  $\boldsymbol{\alpha}(t)$  using the method of exponentially weighted least squares (EWLS). The EWLS estimate of  $\boldsymbol{\alpha}(t)$  can be obtained from

$$\begin{aligned} \hat{\boldsymbol{\alpha}}(t) &= [\hat{\boldsymbol{\alpha}}_0^T(t), \dots, \hat{\boldsymbol{\alpha}}_{k+1}^T(t)]^T \\ &= \arg \min_{\boldsymbol{\alpha}} \sum_{s=0}^{t-1} \lambda^s \left( y(t-s) - \boldsymbol{\psi}^T(t-s) \boldsymbol{\alpha} \right)^2 \\ &= \mathbf{R}_{\boldsymbol{\alpha}}^{-1}(t) \mathbf{r}_{\boldsymbol{\alpha}}(t) \end{aligned} \quad (9)$$

where

$$\begin{aligned} \mathbf{R}_{\boldsymbol{\alpha}}(t) &= \sum_{s=0}^{t-1} \lambda^s \boldsymbol{\psi}(t-s) \boldsymbol{\psi}^T(t-s) = \lambda \mathbf{R}_{\boldsymbol{\alpha}}(t-1) + \boldsymbol{\psi}(t) \boldsymbol{\psi}^T(t) \\ \mathbf{r}_{\boldsymbol{\alpha}}(t) &= \sum_{s=0}^{t-1} \lambda^s y(t-s) \boldsymbol{\psi}(t-s) = \lambda \mathbf{r}_{\boldsymbol{\alpha}}(t-1) + y(t) \boldsymbol{\psi}(t) \end{aligned}$$

and  $\lambda$  ( $0 < \lambda < 1$ ,  $1 - \lambda \ll 1$ ) denotes the so-called forgetting constant—the design parameter that controls the memory of the estimator, and hence allows one to trade off between its tracking speed and tracking accuracy.

The recursive algorithm for evaluation of  $\hat{\boldsymbol{\theta}}(t)$  is given by

$$\begin{aligned} \mathbf{f}_i(t) &= \mathbf{G}_i \mathbf{f}_i(t-1) \\ \boldsymbol{\psi}_i(t) &= \boldsymbol{\varphi}(t) \otimes \mathbf{f}_i(t) \quad i = 0, \dots, k+1 \\ \varepsilon(t) &= y(t) - \boldsymbol{\psi}^T(t) \hat{\boldsymbol{\alpha}}(t-1) \\ \mathbf{k}_{\boldsymbol{\alpha}}(t) &= \frac{\mathbf{P}_{\boldsymbol{\alpha}}(t-1) \boldsymbol{\psi}(t)}{\lambda + \boldsymbol{\psi}^T(t) \mathbf{P}_{\boldsymbol{\alpha}}(t-1) \boldsymbol{\psi}(t)} \\ \mathbf{P}_{\boldsymbol{\alpha}}(t) &= \frac{1}{\lambda} [\mathbf{P}_{\boldsymbol{\alpha}}(t-1) - \mathbf{k}_{\boldsymbol{\alpha}}(t) \boldsymbol{\psi}^T(t) \mathbf{P}_{\boldsymbol{\alpha}}(t-1)] \\ \hat{\boldsymbol{\alpha}}(t) &= \hat{\boldsymbol{\alpha}}(t-1) + \mathbf{k}_{\boldsymbol{\alpha}}(t) \varepsilon(t) \\ \hat{\boldsymbol{\theta}}(t) &= \sum_{i=0}^{k+1} \mathbf{D}_i(t) \hat{\boldsymbol{\alpha}}_i(t) \end{aligned} \quad (10)$$

where  $\mathbf{P}_{\boldsymbol{\alpha}}(t) = \mathbf{R}_{\boldsymbol{\alpha}}^{-1}(t) \mathbf{D}_i(t) = \mathbf{I}_n \otimes \mathbf{f}_i^T(t) = \text{diag}\{\mathbf{f}_i^T(t), \dots, \mathbf{f}_i^T(t)\}$ ,  $\mathbf{I}_n$  denotes the  $n \times n$  identity matrix,  $\mathbf{G}_0^n = g_0 = 1$ ,  $\mathbf{G}_{k+1} = g_{k+1} = -1$ , and

$$\mathbf{G}_i = \begin{bmatrix} \cos \omega_i & \sin \omega_i \\ -\sin \omega_i & \cos \omega_i \end{bmatrix} \quad i = 1, \dots, k$$

denote orthogonal rotation matrices ( $\mathbf{G}_i^{-1} = \mathbf{G}_i^T$ ).

Since (10) combines the basis function parameterization with exponentially weighted least squares estimation, it will be further referred to as the exponentially weighted basis function (EWBF) algorithm [14], [15].

Another, equivalent form of the EWBF estimator can be obtained by rewriting (9) and (10) in a different system of coordinates. Using the linear time-varying transformation

$$\begin{aligned} \hat{\boldsymbol{\beta}}(t) &= \mathbf{A}^{-(t+1)} \hat{\boldsymbol{\alpha}}(t) \\ \mathbf{A} &= \text{diag}\{\mathbf{A}_0, \dots, \mathbf{A}_{k+1}\}, \quad \mathbf{A}^{-1} = \mathbf{A}^T \end{aligned} \quad (11)$$

where  $\mathbf{A}_0 = \mathbf{I}_n$ ,  $\mathbf{A}_{k+1} = -\mathbf{I}_n$  and  $\mathbf{A}_i = \mathbf{I}_n \otimes \mathbf{G}_i$ ,  $\mathbf{A}_i^{-1} = \mathbf{A}_i^T$ ,  $i = 1, \dots, k$ , one can convert (9) into

$$\hat{\boldsymbol{\beta}}(t) = \mathbf{R}_{\boldsymbol{\beta}}^{-1}(t) \mathbf{r}_{\boldsymbol{\beta}}(t) \quad (12)$$

where

$$\begin{aligned} \mathbf{R}_\beta(t) &= \mathbf{A}^{-(t+1)} \mathbf{R}_\alpha(t) \mathbf{A}^{t+1} = \mathbf{A}^T [\lambda \mathbf{R}_\beta(t-1) + \boldsymbol{\zeta}(t) \boldsymbol{\zeta}^T(t)] \mathbf{A} \\ \mathbf{r}_\beta(t) &= \mathbf{A}^{-(t+1)} \mathbf{r}_\alpha(t) = \mathbf{A}^T [\lambda \mathbf{r}_\beta(t-1) + y(t) \boldsymbol{\zeta}(t)] \\ \boldsymbol{\zeta}(t) &= \mathbf{A}^{-t} \boldsymbol{\psi}(t) = [\boldsymbol{\zeta}_0^T(t), \dots, \boldsymbol{\zeta}_{k+1}^T(t)]^T \\ \boldsymbol{\zeta}_i(t) &= \mathbf{A}_i^{-t} \boldsymbol{\psi}_i(t) = \boldsymbol{\varphi}(t) \otimes \mathbf{f}_i(0), \quad i = 0, \dots, k+1. \end{aligned}$$

Similarly, using the substitutions  $\mathbf{k}_\beta(t) = \mathbf{A}^{-t} \mathbf{k}_\alpha(t)$  and  $\mathbf{P}_\beta(t) = \mathbf{R}_\beta^{-1}(t) = \mathbf{A}^{-(t+1)} \mathbf{P}_\alpha(t) \mathbf{A}^{t+1}$ , one can transform (10) into

$$\begin{aligned} \varepsilon(t) &= y(t) - \boldsymbol{\zeta}^T(t) \hat{\boldsymbol{\beta}}(t-1) \\ \mathbf{k}_\beta(t) &= \frac{\mathbf{P}_\beta(t-1) \boldsymbol{\zeta}(t)}{\lambda + \boldsymbol{\zeta}^T(t) \mathbf{P}_\beta(t-1) \boldsymbol{\zeta}(t)} \\ \mathbf{P}_\beta(t) &= \frac{1}{\lambda} \mathbf{A}^T [\mathbf{P}_\beta(t-1) - \mathbf{k}_\beta(t) \boldsymbol{\zeta}^T(t) \mathbf{P}_\beta(t-1)] \mathbf{A} \\ \hat{\boldsymbol{\beta}}(t) &= \mathbf{A}^T [\hat{\boldsymbol{\beta}}(t-1) + \mathbf{k}_\beta(t) \varepsilon(t)] \\ \hat{\boldsymbol{\theta}}(t) &= \sum_{i=0}^{k+1} \mathbf{E}_i \hat{\boldsymbol{\beta}}_i(t) \end{aligned} \quad (13)$$

where  $\mathbf{E}_0 = \mathbf{I}_n$ ,  $\mathbf{E}_{k+1} = -\mathbf{I}_n$  and  $\mathbf{E}_i = \mathbf{D}_i(-1) = \mathbf{I}_n \otimes \mathbf{f}_i^T(-1) = \text{diag}\{\mathbf{e}_i^T, \dots, \mathbf{e}_i^T\}$ ,  $\mathbf{e}_i^T = [-\sin \omega_i, \cos \omega_i]$ ,  $i = 1, \dots, k$ .

Following [14] and [15], the algorithms (10) and (13) will be referred to as the running basis (RB) and fixed basis (FB) EWBF algorithms, respectively. It should be stressed that both algorithms are strictly input–output equivalent, that is, they yield identical parameter estimates  $\hat{\boldsymbol{\theta}}(t)$  for identical data sets  $u(t)$ ,  $y(t)$  and conforming initial conditions ( $\hat{\boldsymbol{\beta}}(0) = \mathbf{A} \hat{\boldsymbol{\alpha}}(0)$ ,  $\mathbf{P}_\beta(0) = \mathbf{A}^{-1} \mathbf{P}_\alpha(0) \mathbf{A}$ ). The RB algorithm will be the starting point for derivation of the gradient-based frequency-adaptive version of the EWBF filter, and the FB algorithm—for derivation of its recursive prediction error based version.

To shed more light on the algorithms described above, consider the problem of extraction of a multifrequency signal (6), characterized by known frequencies  $\omega_1, \dots, \omega_k$ , from noisy measurements (5).

Denote by  $N(q^{-1})$  the transfer function of the cancellation filter

$$\varepsilon(t) = N(q^{-1})y(t).$$

We will prove the following corollary, which is a nontrivial extension (to the multiple frequencies case) of the result given in [16].

*Corollary 1:* All poles of  $N(q^{-1})$  lie on the circle of radius  $\lambda$  in the  $z$  plane and all zeros of  $N(q^{-1})$  lie on the unit circle (see Fig. 1).

*Proof:* See Appendix I.

According to Corollary 1, in the signal case the steady-state version of the EWBF algorithm is nothing but a classical notch filter with constrained poles and zeros, governed by

$$N(q^{-1}) \propto \frac{W(q^{-1})}{\bar{W}(\lambda q^{-1})}$$

where  $W(q^{-1})$  is the polynomial with all zeros located on the unit circle and  $0 < \lambda < 1$ . The name stems from the fact that

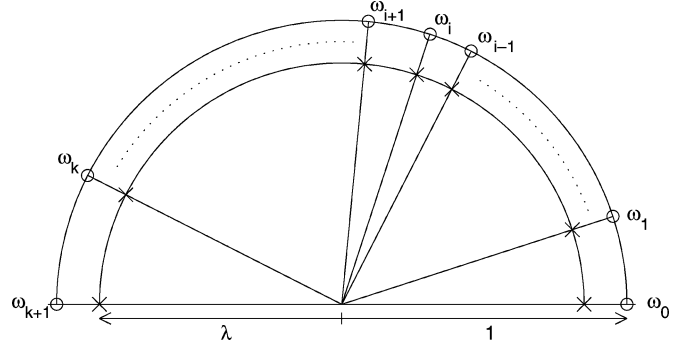


Fig. 1. Location of the poles and zeros of the cancellation filter  $N(q^{-1})$ .

the poles  $p_i$  and zeros  $z_i$  of  $N(q^{-1})$  are constrained to lie on the same radial line with the poles slightly displaced toward the origin ( $p_i = \lambda z_i$ ) by the factor  $\lambda$ , often called the pole contraction factor.

Since the problem solved in this paper is a generalization, to the system case, of the signal elimination/tracking problem mentioned above, the identification algorithms (10) and (13) can be considered generalized notch filters.

### III. UNKNOWN FREQUENCIES

When the functions  $f_i(t)$ ,  $i = 1, \dots, k$  are fixed, the method of basis functions (which can be traced down to Rao [17], Mendel [18], and Liporace [19]) is a straightforward extension of the least squares approach. However, when the basis functions are to be adjusted along with the expansion coefficients  $c_{ij}(t)$ , the estimation problem becomes considerably more difficult. In particular, note that when the frequencies  $\omega_i$ ,  $i = 1, \dots, k$ , in (3) are not known a priori, the generalized regression vector  $\boldsymbol{\psi}(t)$  is also (partially) unknown. As a result, the model (8) becomes nonlinear in the estimated parameters, not to mention that it describes a system which, beyond any doubt, can be termed as rapidly time varying (see [15] for more details on classification of time-varying systems).

In this section we will derive two frequency-adaptive EWBF algorithms, capable of tracking the time-varying frequencies  $\omega_i(t)$ ,  $i = 1, \dots, k$ . To simplify derivations we will neglect the dc and ac terms in (2).

#### A. Gradient Approach

Consider the running basis algorithm (10) and denote by

$$J(t, \boldsymbol{\omega}) = \frac{1}{2} \varepsilon^2(t, \boldsymbol{\omega}) \quad (14)$$

where  $\boldsymbol{\omega} = [\omega_1, \dots, \omega_k]^T$ , the instantaneous measure of fit.

The simplest gradient algorithm that can be used for frequency tracking has the form

$$\hat{\boldsymbol{\omega}}(t+1) = \hat{\boldsymbol{\omega}}(t) - \mu J'(\hat{\boldsymbol{\omega}}(t)) \quad (15)$$

where  $J'(\hat{\boldsymbol{\omega}}(t))$  denotes derivative of  $J(t, \boldsymbol{\omega})$  with respect to  $\boldsymbol{\omega}$ , evaluated at the point  $\hat{\boldsymbol{\omega}}(t)$ , and  $\mu > 0$  is a small stepsize parameter.

Note that

$$\varepsilon(t) = y(t) - \boldsymbol{\psi}^T(t) \hat{\boldsymbol{\alpha}}(t-1) = y(t) - \sum_{i=1}^k \boldsymbol{\psi}_i^T(t) \hat{\boldsymbol{\alpha}}_i(t-1).$$

According to (3), it holds that

$$\mathbf{f}_i(t) = \begin{bmatrix} \sin\left(\sum_{s=1}^t \omega_i(s)\right) \\ \cos\left(\sum_{s=1}^t \omega_i(s)\right) \end{bmatrix}$$

and

$$\mathbf{h}_i(t) = \frac{\partial \mathbf{f}_i(t)}{\partial \omega_i(t)} = \begin{bmatrix} \cos\left(\sum_{s=1}^t \omega_i(s)\right) \\ -\sin\left(\sum_{s=1}^t \omega_i(s)\right) \end{bmatrix} = \mathbf{J}_o \mathbf{f}_i(t)$$

where

$$\mathbf{J}_o = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Therefore

$$-\frac{\partial J(t, \boldsymbol{\omega})}{\partial \omega_i(t)} = \frac{\partial \boldsymbol{\psi}_i^T(t)}{\partial \omega_i(t)} \hat{\boldsymbol{\alpha}}_i(t-1) \varepsilon(t) = \boldsymbol{\chi}_i^T(t) \hat{\boldsymbol{\alpha}}_i(t-1) \varepsilon(t)$$

where

$$\boldsymbol{\chi}_i(t) = \boldsymbol{\varphi}(t) \otimes \mathbf{h}_i(t).$$

Let  $\hat{\mathbf{f}}_i(t) = [\sin(\sum_{s=1}^t \hat{\omega}_i(s)), \cos(\sum_{s=1}^t \hat{\omega}_i(s))]^T$ ,  $\hat{\mathbf{h}}_i(t) = [\cos(\sum_{s=1}^t \hat{\omega}_i(s)), -\sin(\sum_{s=1}^t \hat{\omega}_i(s))]^T$  and  $g_i(t) = -\partial J(t, \boldsymbol{\omega}) / \partial \omega_i(t) |_{\boldsymbol{\omega}(s) = \hat{\boldsymbol{\omega}}(s), s=1, \dots, t}$ .

Using the notation introduced above, the gradient-type frequency-adaptive version of the running basis algorithm can be summarized as follows:

$$\begin{aligned} \hat{\mathbf{f}}_i(t) &= \hat{\mathbf{G}}_i(t) \hat{\mathbf{f}}_i(t-1) \\ \hat{\mathbf{h}}_i(t) &= \hat{\mathbf{G}}_i(t) \hat{\mathbf{h}}_i(t-1) = \mathbf{J}_o \hat{\mathbf{f}}_i(t) \\ \hat{\boldsymbol{\psi}}_i(t) &= \boldsymbol{\varphi}(t) \otimes \hat{\mathbf{f}}_i(t) \\ \hat{\boldsymbol{\chi}}_i(t) &= \boldsymbol{\varphi}(t) \otimes \hat{\mathbf{h}}_i(t), \\ & \quad i=1, \dots, k \\ \varepsilon(t) &= y(t) - \hat{\boldsymbol{\psi}}^T(t) \hat{\boldsymbol{\alpha}}(t-1) \\ g_i(t) &= \hat{\boldsymbol{\chi}}_i^T(t) \hat{\boldsymbol{\alpha}}_i(t-1) \varepsilon(t) \\ \hat{\omega}_i(t+1) &= \hat{\omega}_i(t) + \mu g_i(t), \\ & \quad i=1, \dots, k \\ \mathbf{k}_\alpha(t) &= \frac{\mathbf{P}_\alpha(t-1) \hat{\boldsymbol{\psi}}(t)}{\lambda + \hat{\boldsymbol{\psi}}^T(t) \mathbf{P}_\alpha(t-1) \hat{\boldsymbol{\psi}}(t)} \\ \mathbf{P}_\alpha(t) &= \frac{1}{\lambda} [\mathbf{P}_\alpha(t-1) - \mathbf{k}_\alpha(t) \hat{\boldsymbol{\psi}}^T(t) \mathbf{P}_\alpha(t-1)] \\ \hat{\boldsymbol{\alpha}}(t) &= \hat{\boldsymbol{\alpha}}(t-1) + \mathbf{k}_\alpha(t) \varepsilon(t) \\ \hat{\boldsymbol{\theta}}(t) &= \sum_{i=1}^k \hat{\mathbf{D}}_i(t) \hat{\boldsymbol{\alpha}}_i(t) \end{aligned} \quad (16)$$

where

$$\begin{aligned} \hat{\mathbf{G}}_i(t) &= \begin{bmatrix} \cos \hat{\omega}_i(t) & \sin \hat{\omega}_i(t) \\ -\sin \hat{\omega}_i(t) & \cos \hat{\omega}_i(t) \end{bmatrix} \\ \hat{\mathbf{D}}_i(t) &= \mathbf{I}_n \otimes \hat{\mathbf{f}}_i^T(t) = \text{diag}\{\hat{\mathbf{f}}_i^T(t), \dots, \hat{\mathbf{f}}_i^T(t)\} \end{aligned}$$

and the initial conditions are set to  $\hat{\mathbf{f}}_i(0) = [0, 1]^T$  and  $\hat{\mathbf{h}}_i(0) = [1, 0]^T$ ,  $i = 1, \dots, k$ .

As all gradient-type estimation procedures, the algorithm (16) may have problems with initial convergence unless the starting values  $\hat{\omega}_i(1)$ ,  $i = 1, \dots, k$  are chosen in a judicious way. To guarantee that the initial frequency estimates are close to the true frequencies, one can preestimate  $\boldsymbol{\omega}$  using the nonparametric (Fourier analysis based) approach. In the signal case, both the number of frequencies and their initial values can be estimated using the peak picking technique on periodogram of the short initial fragment of the analyzed signal. In the system case, the initial frequency estimates can be obtained by analyzing the generalized (system) periodogram [20] (see Remark 3 in the next section), by exploiting cyclostationarity of the analyzed system [3], [5], or using the method based on spectral analysis of the short-memory sliding window least squares estimates of system coefficients [21].

### B. Recursive Prediction Error Approach

Consider the following exponentially weighted measure of fit:

$$V(t, \boldsymbol{\omega}) = \frac{1}{2} \sum_{s=0}^{t-1} \gamma^s \varepsilon^2(t-s, \boldsymbol{\omega})$$

where  $\gamma$ ,  $0 < \gamma < 1$ , is the forgetting constant, which will be used to control the speed of the frequency adaptation. This time, to evaluate the estimate

$$\hat{\boldsymbol{\omega}}(t) = \arg \min_{\boldsymbol{\omega}} V(t, \boldsymbol{\omega})$$

we will use the recursive prediction error (RPE) approach. According to Söderström and Stoica [22], the RPE algorithm can be expressed in the form

$$\hat{\boldsymbol{\omega}}(t) = \hat{\boldsymbol{\omega}}(t-1) - [V''(t, \hat{\boldsymbol{\omega}}(t-1))]^{-1} V'(t, \hat{\boldsymbol{\omega}}(t-1))$$

where

$$\begin{aligned} V'(t, \hat{\boldsymbol{\omega}}(t-1)) &\cong \varepsilon(t, \hat{\boldsymbol{\omega}}(t-1)) \frac{\partial \varepsilon(t, \hat{\boldsymbol{\omega}}(t-1))}{\partial \boldsymbol{\omega}} \\ V''(t, \hat{\boldsymbol{\omega}}(t-1)) &\cong \gamma V''(t-1, \hat{\boldsymbol{\omega}}(t-2)) \\ & \quad + \frac{\partial \varepsilon(t, \hat{\boldsymbol{\omega}}(t-1))}{\partial \boldsymbol{\omega}} \frac{\partial \varepsilon(t, \hat{\boldsymbol{\omega}}(t-1))}{\partial \boldsymbol{\omega}^T} \end{aligned}$$

and all derivatives are taken with respect to  $\boldsymbol{\omega}$ .

Note that the regression vector  $\boldsymbol{\zeta}(t)$ , which appears in the fixed basis algorithm (13), does not depend on the frequencies  $\omega_1, \dots, \omega_k$ , while the components of the generalized regression vector  $\boldsymbol{\psi}(t)$ , appearing in the running basis algorithm (10), are frequency dependent. For this reason, the fixed basis algorithm

is a more convenient starting point for derivation of the adaptive RPE version of the EWBF algorithm.

The recursive frequency-adaptive EWBF algorithm, based on the RPE minimization principle, can be written in the form (see Appendix II)

$$\begin{aligned}
\varepsilon(t) &= y(t) - \zeta^T(t)\hat{\beta}(t-1) \\
\eta_i(t) &= -\zeta^T(t)\xi_i(t-1), \quad i = 1, \dots, k \\
\rho(t) &= \gamma + \boldsymbol{\eta}^T(t)\mathbf{F}(t-1)\boldsymbol{\eta}(t) \\
\mathbf{l}(t) &= \frac{\mathbf{F}(t-1)\boldsymbol{\eta}(t)}{\rho(t)} \\
\mathbf{F}(t) &= \frac{1}{\gamma} [\mathbf{F}(t-1) - \mathbf{l}(t)\boldsymbol{\eta}^T(t)\mathbf{F}(t-1)] \\
\hat{\boldsymbol{\omega}}(t) &= \hat{\boldsymbol{\omega}}(t-1) - \mathbf{F}(t)\boldsymbol{\eta}(t)\varepsilon(t) \\
\delta(t) &= \lambda + \zeta^T(t)\mathbf{P}_\beta(t-1)\zeta(t) \\
\mathbf{k}_\beta(t) &= \frac{\mathbf{P}_\beta(t-1)\zeta(t)}{\delta(t)} \\
\mathbf{P}_\beta(t) &= \frac{1}{\lambda} \hat{\mathbf{A}}^T(t)[\mathbf{P}_\beta(t-1) - \mathbf{k}_\beta(t)\zeta^T(t)\mathbf{P}_\beta(t-1)]\hat{\mathbf{A}}(t) \\
\hat{\boldsymbol{\beta}}(t) &= \hat{\mathbf{A}}^T(t)[\hat{\boldsymbol{\beta}}(t-1) + \mathbf{k}_\beta(t)\varepsilon(t)] \\
\tilde{\mathbf{k}}_i(t) &= \frac{1}{\delta(t)} [\tilde{\mathbf{P}}_i(t-1)\zeta(t) - \mathbf{k}_\beta(t)\zeta^T(t)\tilde{\mathbf{P}}_i(t-1)\zeta(t)] \\
\tilde{\mathbf{P}}_i(t) &= \boldsymbol{\Lambda}_i^T\mathbf{P}_\beta(t) + \mathbf{P}_\beta(t)\boldsymbol{\Lambda}_i + \frac{1}{\lambda} \hat{\mathbf{A}}^T(t)[\tilde{\mathbf{P}}_i(t-1) \\
&\quad - \tilde{\mathbf{k}}_i(t)\zeta^T(t)\mathbf{P}_\beta(t-1) - \mathbf{k}_\beta(t)\zeta^T(t)\tilde{\mathbf{P}}_i(t-1)]\hat{\mathbf{A}}(t) \\
\boldsymbol{\xi}_i(t) &= \boldsymbol{\Lambda}_i^T\hat{\boldsymbol{\beta}}(t) + \hat{\mathbf{A}}^T(t)[\boldsymbol{\xi}_i(t-1) + \mathbf{k}_\beta(t)\eta_i(t) + \tilde{\mathbf{k}}_i(t)\varepsilon(t)] \\
&\quad i = 1, \dots, k \\
\hat{\boldsymbol{\theta}}(t) &= \sum_{i=1}^k \hat{\mathbf{E}}_i(t)\hat{\boldsymbol{\beta}}_i(t)
\end{aligned} \tag{17}$$

where

$$\begin{aligned}
\zeta(t) &= \underbrace{[\boldsymbol{\varphi}_0^T(t), \dots, \boldsymbol{\varphi}_0^T(t)]^T}_k \\
\boldsymbol{\varphi}_0(t) &= [0, \varphi_1(t), \dots, 0, \varphi_n(t)]^T \\
\hat{\mathbf{A}}(t) &= \text{diag}\{\hat{\mathbf{A}}_1(t), \dots, \hat{\mathbf{A}}_k(t)\} \\
\hat{\mathbf{A}}_i(t) &= \text{diag}\{\hat{\mathbf{G}}_i(t), \dots, \hat{\mathbf{G}}_i(t)\}_n \\
\hat{\mathbf{E}}_i(t) &= \text{diag}\{\hat{\mathbf{e}}_i^T(t), \dots, \hat{\mathbf{e}}_i^T(t)\}_n \\
\hat{\mathbf{e}}_i^T(t) &= [-\sin \hat{\omega}_i(t), \cos \hat{\omega}_i(t)] \\
\boldsymbol{\Lambda}_i &= \text{diag}\{\underbrace{0, \dots, 0}_{i-1}, \underbrace{1, 0, \dots, 0}_{k-i}\} \otimes \mathbf{J}, \\
\mathbf{J} &= \mathbf{I}_n \otimes \mathbf{J}_o = \text{diag}\{\mathbf{J}_o, \dots, \mathbf{J}_o\}_n.
\end{aligned}$$

The recommended initial conditions are  $\hat{\boldsymbol{\beta}}_i(0) = \mathbf{0}$  and  $\mathbf{P}_\beta(0) = \delta \mathbf{I}_{2kn}$ , where  $\delta$  denotes a large positive constant (this is a standard initialization procedure for all recursive least square type recursive estimation algorithms [22]). As to the remaining quantities, good results were obtained for  $\tilde{\mathbf{P}}_i(0) = \mathbf{0}$ ,  $i = 1, \dots, k$ ,  $\mathbf{F}(0) = \epsilon \mathbf{I}_k$ , where  $\epsilon$  denotes a small positive constant, and for randomly generated values of  $\boldsymbol{\xi}_i(0)$ .

To speed up initial convergence, the rough starting frequency estimates can be obtained by means of nonparametric system identification. However, unlike the gradient approach, in the RPE case frequency preestimation is not prerequisite for convergence. If no prior knowledge about the estimated frequencies is available, one can set  $\hat{\boldsymbol{\omega}}(0) = \mathbf{0}$ .

*Remark 1:* A significant reduction of computational complexity of the algorithm (17) can be achieved by setting  $\tilde{\mathbf{P}}_i(t) = \mathbf{0}$ ,  $\tilde{\mathbf{k}}_i(t) = \mathbf{0}$ ,  $i = 1, \dots, k$ . Simulations show that the results yielded by such simplified version of the RPE algorithm are practically identical with those given by the original algorithm.

*Remark 2:* Using the technique presented in [12] and [23], the “global” EWBF algorithms (16) and (17) can be decomposed into  $k$  interconnected “local” algorithms. The decoupled parallel-form and cascade-form generalized adaptive notch filters, obtained in this way, are made up of identical (from the computational viewpoint) blocks arranged in the parallel and cascade structures, respectively. Each block is designed to track a particular frequency component of the parameter vector  $\boldsymbol{\theta}(t)$ .

*Remark 3:* In this paper, we have assumed that the number of frequency modes  $k$  is known *a priori*. As shown in [20], both the number of frequencies and the initial frequency values, allowing one to efficiently start (or restart) generalized adaptive notch filters, can be inferred from the generalized (system) periodogram, evaluated for a short initial fragment of the input/output data. Such a combined nonparametric/parametric approach allows one to preserve advantages of both frameworks, leading to estimation procedures which guarantee global frequency search, high-frequency resolution, fast initial convergence, and good steady-state tracking capabilities.

### C. Selection of Design Variables

The tracking properties of the RPE algorithm (17) depend on two design variables: the forgetting constant  $\lambda$ , which controls the bandwidth of parameter tracking, and the forgetting constant  $\gamma$ , which controls the bandwidth of frequency tracking. Generally speaking, both forgetting factors should be chosen so as to trade off the tracking speed of a generalized adaptive notch filter (which decreases with growing  $\lambda$  and  $\gamma$ ) and its noise rejection capability (which increases with growing  $\lambda$  and  $\gamma$ ). We have found out experimentally that good results are usually obtained when the frequency tracking bandwidth is larger (but not significantly larger) than the parameter tracking bandwidth, e.g.,  $1 - \gamma = 2(1 - \lambda)$ . The tuning problem is then reduced to selection of a single design variable  $\lambda$ .

The stepsize parameter  $\mu$ , used in the gradient approach, plays an analogous role as the frequency adaptation gain  $1 - \gamma$  in the RPE approach. For a complex-valued version of the gradient algorithm (16), some analytical results are now available, showing how performance of a generalized adaptive notch filter depends on selection of its design variables—see Niedźwiecki and Kaczmarek [24]. Even though restricted to a simple case (single frequency mode, random walk frequency variation), the analysis presented in [24] provides interesting insights into the tracking capabilities and tracking limitations of generalized adaptive notch filters, including the associated speed/accuracy tradeoffs, performance optimization issues, and achievable tracking bounds.

TABLE I  
COMPARISON OF COMPUTATIONAL COMPLEXITY OF THE GRADIENT (G)  
AND RECURSIVE PREDICTION ERROR (RPE) VARIANTS OF THE  
FREQUENCY-ADAPTIVE EXPONENTIALLY WEIGHTED BASIS  
FUNCTION (EWBF) ALGORITHM WITH FIXED BASIS  
AND RUNNING BASIS, RESPECTIVELY;  $n$  DENOTES  
THE NUMBER OF SYSTEM COEFFICIENTS AND  
 $k$  IS THE NUMBER OF BASIS FUNCTIONS

Algorithm	Total complexity
EWBF running basis	$8(nk)^2 + 14nk + 4k$
EWBF fixed basis	$16(nk)^2 + 15nk$
EWBF-G	$8(nk)^2 + 16nk + 6k$
EWBF-RPE	$16(nk)^2 + 15nk$ $+ [22(nk)^2 + 16nk + 3k + 6]k$
simplified EWBF-RPE	$16(nk)^2 + 15nk$ $+ [6nk + 3k + 6]k$

#### D. Computational Complexity

Table I presents a comparison of computational complexity (the number of multiply/add operations per time update) for the gradient-search algorithm (16), the RPE algorithm (17), and the simplified version of the RPE algorithm mentioned in Remark 1 above. To show the computational overhead due to the frequency adaptation, the analogous data for the regular fixed basis EWBF algorithm (13) and regular running basis EWBF algorithm (10) (for a system without the dc and ac terms) were also included in Table I. All evaluations take into consideration special features of some of the involved matrices and vectors, such as symmetry ( $\mathbf{P}_\alpha(t)$ ,  $\mathbf{P}_\beta(t)$ ,  $\mathbf{F}(t)$ ,  $\tilde{\mathbf{P}}_i(t)$ ,  $i = 1, \dots, k$ ), block diagonal structure with repeatable blocks ( $\hat{\mathbf{A}}(t)$ ), or zero elements ( $\zeta(t)$ ). According to Table I, the full version of the RPE algorithm is computationally much more demanding than the gradient algorithm. The simplified version of the RPE algorithm is approximately two times more demanding than the gradient algorithm. Since the steady-state tracking results yielded by the gradient algorithm are only slightly inferior to the analogous results obtained for the RPE-type algorithms (see Section V), in practical applications the EWBF-G approach may be an attractive alternative to the EWBF-RPE approach.

#### IV. INTERRELATED FREQUENCIES

So far we have assumed that the characteristic frequencies  $\omega_1, \dots, \omega_k$ , associated with different modes of parameter variation, are not interrelated. Such conditions are met, for example, for mobile radio channels, where the frequencies correspond to Doppler shifts along different propagation paths and their time-variation is caused by changes of the vehicle speed in the directions of signal arrival.

Suppose now that the estimated frequencies are mutually coupled, namely, that they are multiples of the same fundamental frequency  $\omega(t)$

$$\omega_i(t) = i\omega(t), \quad i = 1, \dots, k. \quad (18)$$

Such multiple frequencies, called harmonics, appear in the Fourier series expansions of periodic signals ( $\omega(t) = \omega = \text{const}$ ). The use of the “time-varying harmonics” (18) can be regarded as a natural extension of the Fourier analysis concepts to quasi-periodically varying processes—see, e.g., Parker and Anderson [11].

The harmonic model (with the dc and ac terms neglected) has the form

$$\theta_l(t) = \sum_{i=1}^k [c_{l, 2i-1}(t) \sin(i\phi(t)) + c_{l, 2i}(t) \cos(i\phi(t))] \quad (19)$$

where  $\phi(t) = \sum_{s=1}^t \omega(s)$ . It can be easily incorporated in the identification schemes, developed in Section II, by setting

$$\mathbf{f}_i(t) = \begin{bmatrix} \sin i\omega t \\ \cos i\omega t \end{bmatrix}, \quad \mathbf{G}_i = \begin{bmatrix} \cos i\omega & \sin i\omega \\ -\sin i\omega & \cos i\omega \end{bmatrix}, \quad i = 1, \dots, k.$$

In the signal case, the corresponding algorithms reduce down to a comb filter, used to cancel periodic interferences (such as the power frequency pickup and its harmonics) or to enhance periodic signals in noise.

When the fundamental frequency is time-varying, it holds that

$$\mathbf{f}_i(t) = \begin{bmatrix} \sin \left( i \sum_{s=1}^t \omega(s) \right) \\ \cos \left( i \sum_{s=1}^t \omega(s) \right) \end{bmatrix}$$

leading to

$$-\frac{\partial J(t, \omega)}{\partial \omega(t)} = \frac{\partial \boldsymbol{\psi}^T(t)}{\partial \omega(t)} \hat{\boldsymbol{\alpha}}(t-1) \boldsymbol{\varepsilon}(t) = \sum_{i=1}^k i \mathbf{x}_i^T(t) \hat{\boldsymbol{\alpha}}_i(t-1) \boldsymbol{\varepsilon}(t).$$

Only minor modifications are needed to convert the gradient-type generalized adaptive notch filter (16) into the generalized adaptive comb filter: all one needs to do is set

$$\hat{\mathbf{G}}_i(t) = \begin{bmatrix} \cos i\hat{\omega}(t) & \sin i\hat{\omega}(t) \\ -\sin i\hat{\omega}(t) & \cos i\hat{\omega}(t) \end{bmatrix}$$

and replace the frequency update section of (16) with

$$g(t) = \sum_{i=1}^k i \hat{\mathbf{x}}_i^T(t) \hat{\boldsymbol{\alpha}}_i(t-1) \boldsymbol{\varepsilon}(t)$$

$$\hat{\omega}(t+1) = \hat{\omega}(t) + \mu g(t).$$

Since the analogous changes to the RPE algorithm (17) are more substantial, the entire modified algorithm is summarized below

$$\boldsymbol{\varepsilon}(t) = y(t) - \boldsymbol{\zeta}^T(t) \hat{\boldsymbol{\beta}}(t-1)$$

$$\boldsymbol{\eta}(t) = -\boldsymbol{\zeta}^T(t) \boldsymbol{\xi}(t-1)$$

$$\nu(t) = \gamma \nu(t-1) + \eta^2(t)$$

$$\hat{\omega}(t) = \hat{\omega}(t-1) - \frac{\eta(t) \boldsymbol{\varepsilon}(t)}{\nu(t)}$$

$$\delta(t) = \lambda + \boldsymbol{\zeta}^T(t) \mathbf{P}_\beta(t-1) \boldsymbol{\zeta}(t)$$

$$\mathbf{k}_\beta(t) = \frac{\mathbf{P}_\beta(t-1) \boldsymbol{\zeta}(t)}{\delta(t)}$$

$$\mathbf{P}_\beta(t) = \frac{1}{\lambda} \hat{\mathbf{A}}^T(t) [\mathbf{P}_\beta(t-1) - \mathbf{k}_\beta(t) \boldsymbol{\zeta}^T(t) \mathbf{P}_\beta(t-1)] \hat{\mathbf{A}}(t)$$

$$\hat{\boldsymbol{\beta}}(t) = \hat{\mathbf{A}}^T(t) [\hat{\boldsymbol{\beta}}(t-1) + \mathbf{k}_\beta(t) \boldsymbol{\varepsilon}(t)]$$

$$\tilde{\mathbf{k}}_\beta(t) = \frac{1}{\delta(t)} [\tilde{\mathbf{P}}_\beta(t-1) \boldsymbol{\zeta}(t) - \mathbf{k}_\beta(t) \boldsymbol{\zeta}^T(t) \tilde{\mathbf{P}}_\beta(t-1) \boldsymbol{\zeta}(t)]$$

$$\tilde{\mathbf{P}}_\beta(t) = \boldsymbol{\Lambda}^T \mathbf{P}_\beta(t) + \mathbf{P}_\beta(t) \boldsymbol{\Lambda} + \frac{1}{\lambda} \hat{\mathbf{A}}^T(t) [\tilde{\mathbf{P}}_\beta(t-1)$$

$$- \tilde{\mathbf{k}}_\beta(t) \boldsymbol{\zeta}^T(t) \mathbf{P}_\beta(t-1) - \mathbf{k}_\beta(t) \boldsymbol{\zeta}^T(t) \tilde{\mathbf{P}}_\beta(t-1)] \hat{\mathbf{A}}(t)$$

$$\boldsymbol{\xi}(t) = \boldsymbol{\Lambda}^T \hat{\boldsymbol{\beta}}(t) + \hat{\mathbf{A}}^T(t) [\boldsymbol{\xi}(t-1) + \mathbf{k}_\beta(t) \eta(t) + \tilde{\mathbf{k}}_\beta(t) \boldsymbol{\varepsilon}(t)]$$

$$\hat{\boldsymbol{\theta}}(t) = \sum_{i=1}^k \hat{\mathbf{E}}_i(t) \hat{\boldsymbol{\beta}}_i(t) \quad (20)$$

where

$$\mathbf{\Lambda} = \text{diag}\{1, \dots, k\} \otimes \mathbf{J}$$

and the quantities  $\hat{\mathbf{E}}_i(t)$  are defined analogously as in Section III, except that  $\hat{\mathbf{e}}_i^T(t) = [-\sin i\hat{\omega}(t), \cos i\hat{\omega}(t)]$ .

*Remark:* Incorporation of the constraint (18), whenever applicable, can significantly increase accuracy of the frequency estimates. The analysis, carried out for stationary harmonic signals by Nehorai and Porat [10], shows that significant improvements in the Cramér–Rao bounds can be achieved if the harmonic structure, imposed by (18), is taken into account in the estimation process.

On the negative side, the frequency estimates based on the harmonic model (19) may occasionally suffer from the frequency mismatch effect— $\hat{\omega}(t)$  may lock onto a fraction or multiple of the true frequency, e.g., onto  $(1/2)\omega(t)$  or  $2\omega(t)$  [11]. Since this effect results in attempts to track the whole bunch of nonexistent harmonics (in the case of the fractional-type convergence) or it results in the failure to track a certain number of the existing ones (in the multiple-type convergence case), it is potentially more damaging than the analogous effect that may occur in the unconstrained algorithm, which follows each frequency separately. Whether it pays off or not to use comb filtering in the interrelated frequencies case depends on the power of the corresponding frequency components. In the case of a strong imbalance [e.g., when only the first and fifth harmonics are significant for a given signal-to-noise ratio (SNR)], one may be better off using notch filter instead of a comb filter.

## V. COMPUTER SIMULATIONS

### A. Generalized Notch Filters

The first set of simulation results demonstrates the frequency and parameter tracking properties of the generalized notch filters: the gradient algorithm (EWBF-G) and the recursive prediction error algorithm (EWBF-RPE).

The identified system (a simulated two-tap quasi-periodically varying channel) had the form

$$y(t) = \theta_1(t)u(t) + \theta_2(t)u(t-1) + v(t)$$

where

$$\begin{aligned} \theta_1(t) &= 0.5 + 2\sin \phi_1(t) - 0.5\sin \phi_2(t) + 0.7\sin \phi_3(t) \\ \theta_2(t) &= -1 - 0.5\cos \phi_1(t) + 1.5\cos \phi_2(t) - 0.7\cos \phi_3(t). \end{aligned}$$

The uncorrelated pseudorandom bit sequence (PRBS) sequence ( $u(t) = \pm 1$ ) was used as the input signal and the noise was white and Gaussian with variance  $\sigma_v^2(t) = 0.05$  (average SNR  $\cong 20$  dB). The forgetting factors  $\lambda = 0.99$  and  $\gamma = 0.98$  for EWBF-RPE algorithm and  $\lambda = 0.99$  and  $\mu = 1.7 \cdot 10^{-4}$  for EWBF-G algorithm were chosen so as to guarantee approximately the same prediction error variance and frequency estimation variance for both algorithms in the case where all basis frequencies are constant.

1) *Initial Convergence Tests:* In all initial convergence tests, the initial frequency estimates  $\hat{\omega}_i(0)$ ,  $i = 1, \dots, k$  were set to zero, the quantities  $\xi_i(0)$ ,  $i = 1, \dots, k$ , were drawn

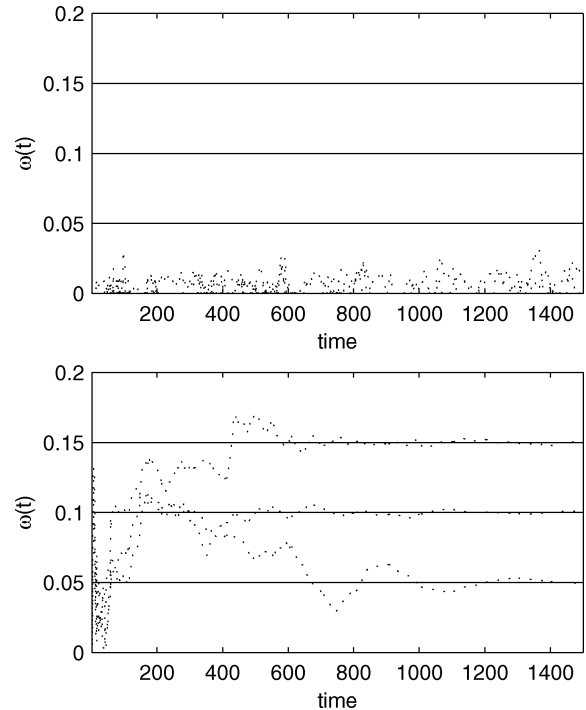


Fig. 2. True basis frequencies (solid lines) and their estimates (dotted lines) obtained in the initial convergence period using the (upper plot) WBF-G algorithm and (lower plot) WBF-RPE algorithm.

as random independent samples from the normal distribution  $\mathcal{N}(\mathbf{0}, \mathbf{I}_{kn})$ , and the initial values of the matrix  $\mathbf{F}(t)$  were set to  $\mathbf{F}(0) = 0.05\mathbf{I}_k$ . It was observed that selection of the starting value  $\mathbf{F}(0)$  has major influence on the initial convergence of the EWBF-RPE filter. In particular, too large initial values may block the adaptation mechanism in the startup phase of frequency estimation.

The results of a typical initial convergence test are shown in Fig. 2. In a majority of test runs (about 80%) the EWBF-RPE algorithm properly matched the true basis frequencies. In contrast with this, the EWBF-G algorithm was unable to find system frequencies in all tests.

It should be stressed that the initial convergence time depends on the constellation of system frequencies, on the average SNR and on the initialization technique used. In the case of our example (illustrated by Fig. 2), the settling time is equal to (roughly) 1000 sampling intervals. However, when the preestimation techniques—proposed in [20]—are used, the average settling time can be brought down to 64 samples. This does not mean that the length of the training sequence must be comparable with the average settling time. There are schemes, proposed recently (see, e.g., [25]), which combine data from many consecutive frames whenever the interframe consistency test is positive. This means that the decoding results for the first frame can be (conditionally) regarded as a training sequence for the second frame, etc., until the change of propagation conditions forces one to start channel identification “from scratch.” Multi-frame schemes allow one to reduce the bit error rate for a given transmission rate—or conversely, to increase transmission rate for a given bit error rate level. We think our algorithms can be useful in such a multiframe context.

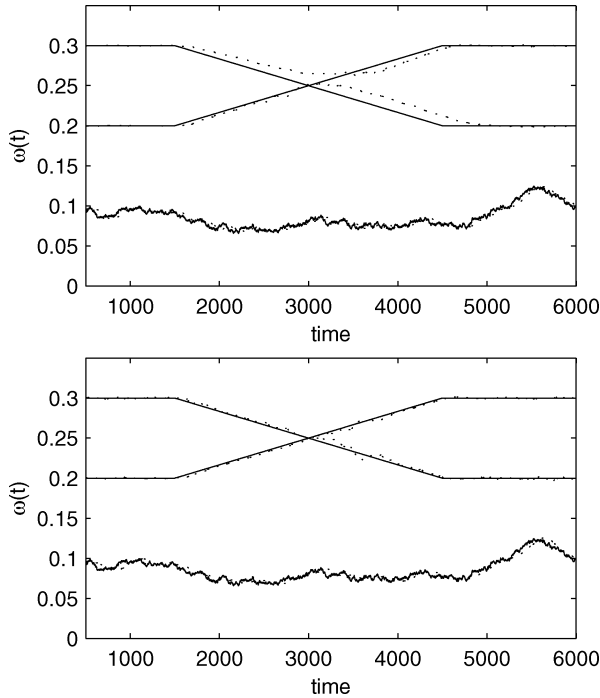


Fig. 3. Instantaneous Doppler frequencies of a simulated channel (solid lines) and their estimates (dotted lines) obtained using the (upper plot) WBF-G algorithm and (lower plot) WBF-RPE algorithm.

2) *Steady-State Tracking Tests:* In all steady-state tracking tests, it was assumed that the initial values of the frequencies are approximately known (the training part of the transmitted data can be used for the initial frequency estimation purposes—see, e.g., [20]). Furthermore, the frequency estimates were frozen at their initial values during the first 300 steps of the algorithm. Some of the simulated frequency changes (see Fig. 3) were difficult from the tracking viewpoint—note that two frequency trajectories intersect in the middle of the analysis interval.

Fig. 3 shows evolution of the frequency estimates and one-step-ahead prediction errors yielded by the EWBF-G and EWBF-RPE algorithms. Figs. 4 and 5 show the estimated parameter trajectories obtained for a typical simulation run as well as the mean trajectories of parameter estimates, obtained by averaging results of 50 simulation runs.

The main conclusion that can be drawn after examining the plots shown in Figs. 3–5, as well as many other plots (not shown here) obtained under the same experimental conditions for different realizations of the measurement noise, is that the EWBF-RPE algorithm performs (slightly) better than the EWBF-G algorithm, in terms of both parameter estimation and frequency tracking.

### B. Generalized Comb Filters

Another two-coefficient finite impulse response system, excited by the PRBS input signal and white noise disturbance ( $\sigma_v^2(t) = 0.05$ ), was simulated to test the steady-state behavior of the EWBF-RPE generalized adaptive comb filtering algorithm. The time-varying system parameters were modeled as linear combinations of three harmonics

$$\begin{aligned}\theta_1(t) &= \sin \phi(t) + 0.5 \sin 2\phi(t) + 0.7 \sin 3\phi(t) \\ \theta_2(t) &= 0.6 \cos \phi(t) - 1.5 \cos 2\phi(t) + 1.4 \cos 3\phi(t)\end{aligned}$$

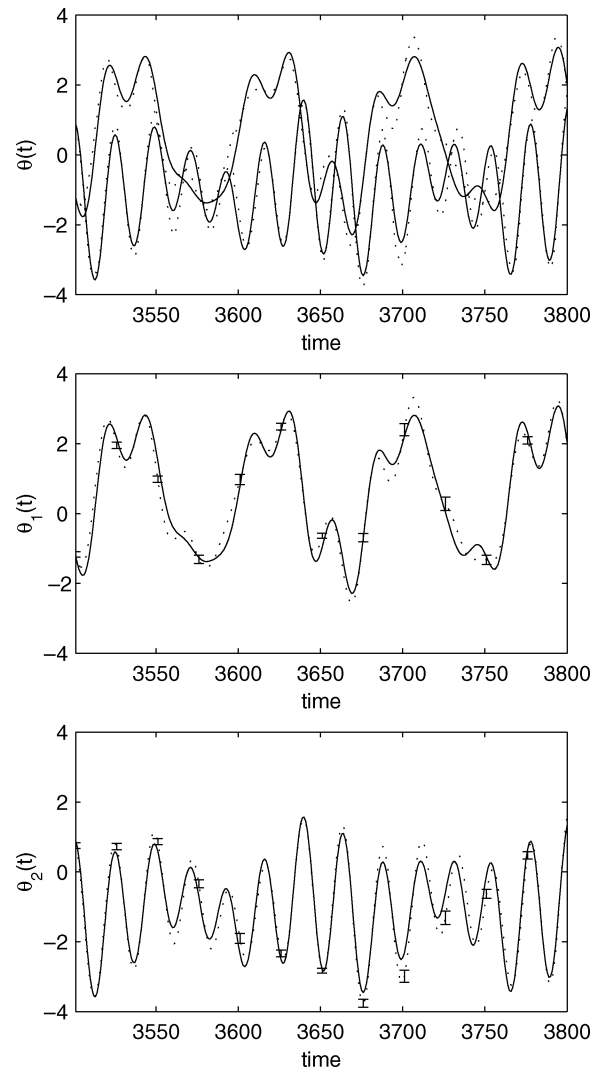


Fig. 4. Evolution of true system parameters  $\theta_1(t)$  and  $\theta_2(t)$  (solid lines) along with their estimates (dotted lines) obtained using the EWBF-G algorithm (upper plot). The two lower plots show mean trajectories of parameter estimates, obtained by averaging results of 50 simulation runs; the corresponding standard deviations ( $\pm$ ) are marked with short vertical lines.

and the fundamental frequency was changing according to the random walk model

$$\omega(t) = \omega(t-1) + \varrho(t) \quad (21)$$

where  $\varrho(t)$  denotes white Gaussian noise with zero mean and variance  $\sigma_\varrho^2 = 4 \cdot 10^{-4}$ .

Fig. 6 shows simulation results obtained for the EWBF-RPE generalized comb algorithm with forgetting factors  $\lambda = 0.99$  and  $\gamma = 0.98$ . Again, the results were satisfactory, in terms of both parameter estimation and frequency tracking.

## VI. CONCLUSION

The problem of identification/tracking of quasi-periodically varying real systems was considered and solved using the frequency-adaptive basis function approach. The proposed algorithms are capable of tracking slow changes in system frequencies, which means that not only the expansion coefficients in the basis function description of the analyzed system but also the

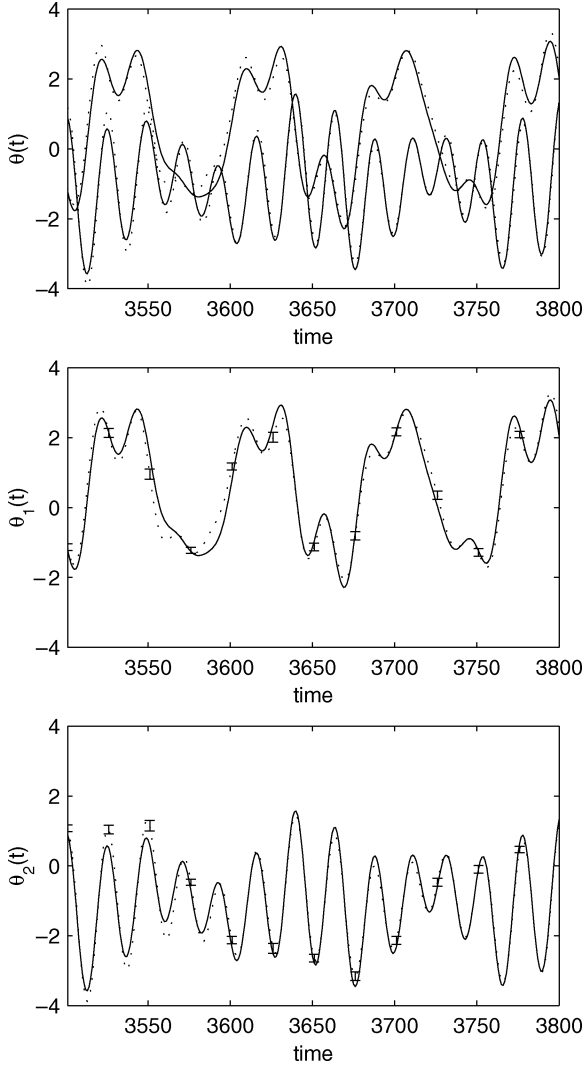


Fig. 5. Evolution of true system parameters  $\theta_1(t)$  and  $\theta_2(t)$  (solid lines) along with their estimates (dotted lines) obtained using the EWBF-RPE algorithm (upper plot). The two lower plots show mean trajectories of parameter estimates, obtained by averaging results of 50 simulation runs; the corresponding standard deviations ( $\pm$ ) are marked with short vertical lines.

basis functions themselves are adjusted in an adaptive manner. Two versions of frequency-adaptive algorithms, gradient search based and recursive prediction error based, were derived both in the unrelated frequencies case (generalized notch filters) and in the interrelated frequencies case (generalized comb filters). Simulation experiments show that the RPE versions of the proposed algorithms have better initial convergence properties and better steady-state tracking capabilities than their gradient counterparts. On the other hand, gradient algorithms are simpler and hence computationally less demanding.

#### APPENDIX I PROOF OF COROLLARY 1

In order to prove the corollary, consider the FB realization of the EWBF filter (due to the equivalence of the RB and FB algorithms, all results derived below remain valid in the RB case). Note that in the case considered,  $n = 1$ ,  $\varphi(t) = 1$ ,  $\forall t$ , and hence  $\zeta(t) = \mathbf{f}(0)$ , where  $\mathbf{f}(t) = [f_0(t), \mathbf{f}_1^T(t), \dots, \mathbf{f}_k^T(t), f_{k+1}(t)]^T$ .

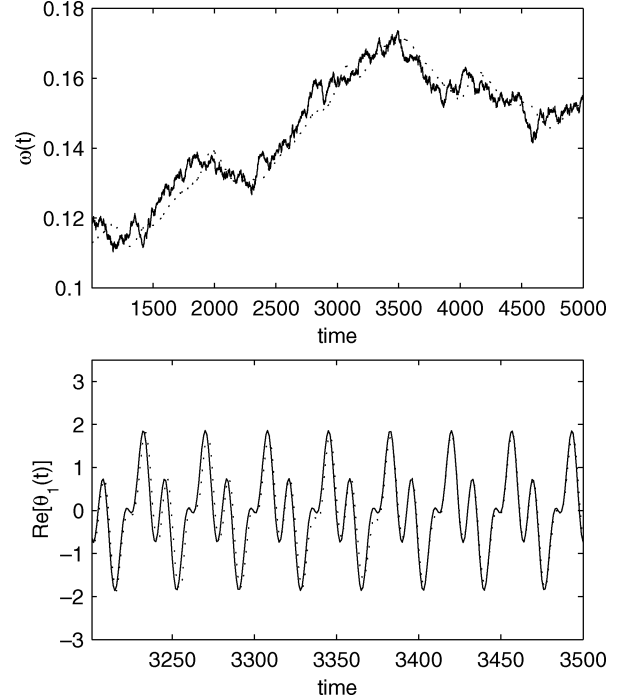


Fig. 6. Instantaneous fundamental basis frequency (solid line) and its estimate (dotted line) obtained using the generalized adaptive comb filter (EWBF-RPE). The lower plot shows the parameter  $\theta_1(t)$  (solid line) and its estimate  $\hat{\theta}_1(t)$  (dotted line).

Note also that the matrix  $\mathbf{P}_\beta(t)$  stabilizes at the constant value  $\mathbf{P}_\beta(\infty)$  as time goes to infinity

$$\lim_{t \rightarrow \infty} \mathbf{P}_\beta(t) = \mathbf{P}_\beta(\infty) = \mathbf{R}_\beta^{-1}(\infty)$$

where

$$\mathbf{R}_\beta(\infty) = \sum_{s=0}^{\infty} \lambda^s \mathbf{f}(-s-1) \mathbf{f}^T(-s-1).$$

Therefore, in the steady state, the estimation algorithm (13) can be written in the form

$$\begin{aligned} \varepsilon(t) &= y(t) - \mathbf{f}^T(0) \hat{\boldsymbol{\beta}}(t-1) \\ \hat{\boldsymbol{\beta}}(t) &= \mathbf{G}^T [\hat{\boldsymbol{\beta}}(t-1) + \mathbf{k}_\beta(\infty) \varepsilon(t)] \end{aligned} \quad (22)$$

where  $\mathbf{G} = \text{diag}\{1, \mathbf{G}_1, \dots, \mathbf{G}_k, -1\}$ , or equivalently [see (12)]

$$\hat{\boldsymbol{\beta}}(t) = \mathbf{R}_\beta^{-1}(\infty) \sum_{s=0}^{\infty} \lambda^s y(t-s) \mathbf{f}(-s-1). \quad (23)$$

Denote by  $\mathbf{T}(q^{-1})$  the (vector) transfer function of the extraction filter

$$\hat{\boldsymbol{\beta}}(t) = \mathbf{T}(q^{-1}) y(t).$$

Note that (22) can be rewritten in the form

$$\hat{\boldsymbol{\beta}}(t) = \mathbf{B} \hat{\boldsymbol{\beta}}(t-1) + \mathbf{b} y(t)$$

where  $\mathbf{B} = \mathbf{G}^T - \mathbf{G}^T \mathbf{k}_\beta(\infty) \mathbf{f}^T(0)$  and  $\mathbf{b} = \mathbf{G}^T \mathbf{k}_\beta(\infty)$ . Therefore, the poles of  $\mathbf{T}(q^{-1})$  coincide with the eigenvalues of the state transition matrix  $\mathbf{B}$ , further denoted as  $\sigma[\mathbf{B}]$ .

Multiplying by  $\lambda \mathbf{G}^T \mathbf{P}_\beta^{-1}(\infty)$  both sides of the steady-state equation for  $\mathbf{P}_\beta(\infty)$

$$\mathbf{P}_\beta(\infty) = \frac{1}{\lambda} \mathbf{G}^T [\mathbf{P}_\beta(\infty) - \mathbf{k}_\beta(\infty) \mathbf{f}^T(0) \mathbf{P}_\beta(\infty)] \mathbf{G}$$

one obtains

$$\mathbf{B} = \lambda \mathbf{P}_\beta(\infty) \mathbf{G}^T \mathbf{P}_\beta^{-1}(\infty).$$

Hence

$$\sigma[\mathbf{B}] = \sigma[\lambda \mathbf{P}_\beta(\infty) \mathbf{G}^T \mathbf{P}_\beta^{-1}(\infty)] = \lambda \sigma[\mathbf{G}^T].$$

Since all eigenvalues of the matrix  $\mathbf{G}^T$  lie on the unit circle in the  $z$ -plane ( $\sigma[\mathbf{G}_i^T] = \sigma[\mathbf{G}_i] = e^{\pm j\omega_i}$ ,  $i = 0, \dots, k+1$ ), all poles of  $\mathbf{T}(q^{-1})$  must lie on the circle of radius  $\lambda$ . Finally, since it holds

$$N(q^{-1}) = 1 - \mathbf{f}^T(0)q^{-1}\mathbf{T}(q^{-1})$$

the poles of  $N(q^{-1})$  must have the same locations as the poles of  $\mathbf{T}(q^{-1})$ , which proves the first assertion of Corollary 1.

The proof of the second assertion is based on the following identities:

$$\begin{aligned} q^{-1}\mathbf{T}(q^{-1})|_{q=e^{j0}} &= \mathbf{1}_0 = [1, 0, \dots, 0]^T \\ q^{-1}\mathbf{T}(q^{-1})|_{q=e^{j\omega_i}} &= \mathbf{1}_i = [0, \dots, 0, \underset{2i-1}{j}, \underset{2i}{1}, 0, \dots, 0]^T \\ &\quad i = 1, \dots, k \\ q^{-1}\mathbf{T}(q^{-1})|_{q=e^{j\pi}} &= \mathbf{1}_{k+1} = [0, \dots, 0, 1]^T \end{aligned}$$

which can be easily checked by combining [see (23)]

$$\mathbf{T}(q^{-1}) = \mathbf{R}_\beta^{-1}(\infty) \sum_{s=0}^{\infty} \lambda^s \mathbf{f}(-s-1)q^{-s}$$

with

$$\mathbf{R}_\beta(\infty) = [\mathbf{r}_0 | \dots | \mathbf{r}_{2i-1} | \mathbf{r}_{2i} | \dots | \mathbf{r}_{2k+1}]$$

where

$$\begin{aligned} \mathbf{r}_0 &= -\mathbf{r}_{2k+1} = \sum_{s=0}^{\infty} \lambda^s \mathbf{f}(-s-1) \\ \mathbf{r}_{2i-1} &= -\sum_{s=0}^{\infty} \lambda^s \mathbf{f}(-s-1) \sin \omega_i(s+1) \\ \mathbf{r}_{2i} &= \sum_{s=0}^{\infty} \lambda^s \mathbf{f}(-s-1) \cos \omega_i(s+1) \quad 1 \leq i \leq k. \end{aligned}$$

Actually, observe that

$$\begin{aligned} \mathbf{R}_\beta(\infty) \mathbf{1}_i &= \mathbf{r}_{2i} + j \mathbf{r}_{2i-1} = \sum_{s=0}^{\infty} \lambda^s \mathbf{f}(-s-1) e^{-j\omega_i(s+1)} \\ &= q^{-1} \sum_{s=0}^{\infty} \lambda^s \mathbf{f}(-s-1) q^{-s} |_{q=e^{j\omega_i}} \\ &= \mathbf{R}_\beta(\infty) q^{-1} \mathbf{T}(q^{-1}) |_{q=e^{j\omega_i}}, \quad i = 1, \dots, k \end{aligned}$$

and hence

$$q^{-1} \mathbf{T}(q^{-1}) |_{q=e^{j\omega_i}} = \mathbf{1}_i, \quad i = 1, \dots, k.$$

The remaining two identities (for  $\omega_0 = 0$  and  $\omega_{k+1} = \pi$ ) can be verified in the analogous way.

Finally, one obtains

$$\begin{aligned} N(e^{-j\omega_i}) &= 1 - \mathbf{f}^T(0) e^{-j\omega_i} \mathbf{T}(e^{-j\omega_i}) \\ &= 1 - \mathbf{f}^T(0) \mathbf{1}_i = 0, \quad i = 0, \dots, k+1. \end{aligned}$$

## APPENDIX II

### DERIVATION OF THE RPE ALGORITHM

Consider the fixed basis EWBF algorithm (13). Observe that

$$\begin{aligned} \frac{\partial \varepsilon(t, \boldsymbol{\omega})}{\partial \boldsymbol{\omega}} &= -\frac{\partial \hat{\boldsymbol{\beta}}^T(t-1, \boldsymbol{\omega})}{\partial \boldsymbol{\omega}} \boldsymbol{\zeta}(t) \\ \frac{\partial \hat{\boldsymbol{\beta}}^T(t, \boldsymbol{\omega})}{\partial \boldsymbol{\omega}} &= \left[ \frac{\partial \hat{\boldsymbol{\beta}}(t, \boldsymbol{\omega})}{\partial \omega_1} \dots \frac{\partial \hat{\boldsymbol{\beta}}(t, \boldsymbol{\omega})}{\partial \omega_k} \right]^T \\ \frac{\partial \mathbf{k}_\beta^T(t, \boldsymbol{\omega})}{\partial \boldsymbol{\omega}} &= \left[ \frac{\partial \mathbf{k}_\beta(t, \boldsymbol{\omega})}{\partial \omega_1} \dots \frac{\partial \mathbf{k}_\beta(t, \boldsymbol{\omega})}{\partial \omega_k} \right]^T. \end{aligned}$$

Furthermore, for  $1 \leq i \leq k$ , it holds

$$\begin{aligned} \frac{\partial \hat{\boldsymbol{\beta}}(t, \boldsymbol{\omega})}{\partial \omega_i} &= \frac{\partial \mathbf{A}^T(\boldsymbol{\omega})}{\partial \omega_i} \left[ \hat{\boldsymbol{\beta}}(t-1, \boldsymbol{\omega}) + \mathbf{k}_\beta(t, \boldsymbol{\omega}) \varepsilon(t, \boldsymbol{\omega}) \right] \\ &\quad + \mathbf{A}^T(\boldsymbol{\omega}) \left[ \frac{\partial \hat{\boldsymbol{\beta}}(t-1, \boldsymbol{\omega})}{\partial \omega_i} + \frac{\partial \mathbf{k}_\beta(t, \boldsymbol{\omega})}{\partial \omega_i} \varepsilon(t, \boldsymbol{\omega}) \right. \\ &\quad \left. + \mathbf{k}_\beta(t, \boldsymbol{\omega}) \frac{\partial \varepsilon(t, \boldsymbol{\omega})}{\partial \omega_i} \right] \\ \frac{\partial \mathbf{k}_\beta(t, \boldsymbol{\omega})}{\partial \omega_i} &= \frac{\frac{\partial \mathbf{P}_\beta(t-1, \boldsymbol{\omega})}{\partial \omega_i} \boldsymbol{\zeta}(t)}{\lambda + \boldsymbol{\zeta}^T(t) \mathbf{P}_\beta(t-1, \boldsymbol{\omega}) \boldsymbol{\zeta}(t)} \\ &\quad - \frac{\mathbf{P}_\beta(t-1, \boldsymbol{\omega}) \boldsymbol{\zeta}(t) \boldsymbol{\zeta}^T(t) \frac{\partial \mathbf{P}_\beta(t-1, \boldsymbol{\omega})}{\partial \omega_i} \boldsymbol{\zeta}(t)}{\left( \lambda + \boldsymbol{\zeta}^T(t) \mathbf{P}_\beta(t-1, \boldsymbol{\omega}) \boldsymbol{\zeta}(t) \right)^2} \\ \frac{\partial \mathbf{P}_\beta(t, \boldsymbol{\omega})}{\partial \omega_i} &= \frac{1}{\lambda} \frac{\partial \mathbf{A}^T(\boldsymbol{\omega})}{\partial \omega_i} [\mathbf{P}_\beta(t-1, \boldsymbol{\omega}) \\ &\quad - \mathbf{k}_\beta(t, \boldsymbol{\omega}) \boldsymbol{\zeta}^T(t) \mathbf{P}_\beta(t-1, \boldsymbol{\omega})] \mathbf{A}(\boldsymbol{\omega}) \\ &\quad + \frac{1}{\lambda} \mathbf{A}^T(\boldsymbol{\omega}) \left[ \frac{\partial \mathbf{P}_\beta(t-1, \boldsymbol{\omega})}{\partial \omega_i} \right. \\ &\quad \left. - \frac{\partial \mathbf{k}_\beta(t, \boldsymbol{\omega})}{\partial \omega_i} \boldsymbol{\zeta}^T(t) \mathbf{P}_\beta(t-1, \boldsymbol{\omega}) \right. \\ &\quad \left. - \mathbf{k}_\beta(t, \boldsymbol{\omega}) \boldsymbol{\zeta}^T(t) \frac{\partial \mathbf{P}_\beta(t-1, \boldsymbol{\omega})}{\partial \omega_i} \right] \mathbf{A}(\boldsymbol{\omega}) \\ &\quad + \frac{1}{\lambda} \mathbf{A}^T(\boldsymbol{\omega}) [\mathbf{P}_\beta(t-1, \boldsymbol{\omega}) \\ &\quad - \mathbf{k}_\beta(t, \boldsymbol{\omega}) \boldsymbol{\zeta}^T(t) \mathbf{P}_\beta(t-1, \boldsymbol{\omega})] \frac{\partial \mathbf{A}(\boldsymbol{\omega})}{\partial \omega_i}. \end{aligned}$$

Note that

$$\begin{aligned} \frac{\partial \mathbf{G}_i(\omega_i)}{\partial \omega_i} &= \mathbf{G}_i(\omega_i) \mathbf{J}_o, \quad \mathbf{J}_o = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\ \frac{\partial \mathbf{A}_i(\omega_i)}{\partial \omega_i} &= \mathbf{A}_i(\omega_i) \mathbf{J}, \quad \mathbf{J} = \mathbf{I}_n \otimes \mathbf{J}_o = \text{diag}\{\mathbf{J}_o, \dots, \mathbf{J}_o\}. \end{aligned}$$

Let  $\varepsilon(t) = \varepsilon(t, \hat{\boldsymbol{\omega}}(t))$ ,  $\hat{\boldsymbol{\beta}}(t) = \hat{\boldsymbol{\beta}}(t, \hat{\boldsymbol{\omega}}(t))$ ,  $\hat{\mathbf{A}}(t) = \mathbf{A}(\hat{\boldsymbol{\omega}}(t))$ ,  $\mathbf{k}_\beta(t) = \mathbf{k}_\beta(t, \hat{\boldsymbol{\omega}}(t))$ ,  $\mathbf{P}_\beta(t) = \mathbf{P}_\beta(t, \hat{\boldsymbol{\omega}}(t))$ ,  $\eta_{\hat{\boldsymbol{\beta}}}(t) = \partial \varepsilon(t, \hat{\boldsymbol{\omega}}(t-1)) / \partial \omega_i$ ,  $\xi_{\hat{\boldsymbol{\beta}}}(t) = \partial \hat{\boldsymbol{\beta}}(t, \hat{\boldsymbol{\omega}}(t)) / \partial \omega_i$ ,  $\mathbf{k}_i(t) = \partial \mathbf{k}_\beta(t, \hat{\boldsymbol{\omega}}(t-1)) / \partial \omega_i$ ,  $\mathbf{P}_i(t) = \partial \mathbf{P}_\beta(t, \hat{\boldsymbol{\omega}}(t)) / \partial \omega_i$ , and  $\mathbf{F}(t) = [\mathbf{V}''(t, \hat{\boldsymbol{\omega}}(t-1))]^{-1}$ .

Using the shorthands introduced above one arrives at (17).

## REFERENCES

- [1] R. T. Aiken, "Communication over the discrete-path fading channel," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 346–347, 1967.
- [2] L. Lindbom, M. Sternad, and A. Ahlén, "A Viterbi detector based on sinusoid modeling of fading mobile radio channels: An illustration of the utility of deterministic models of time-variation in adaptive systems," in *STU Workshop Digital Communication*, Gothenburg, Sweden, 1991.
- [3] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *Int. J. Adaptive Contr. Signal Process.*, vol. 10, pp. 159–176, 1996.
- [4] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, 1998.
- [5] J. Bakkoury, D. Roviras, M. Ghogho, and F. Castanie, "Adaptive MLSE receiver over rapidly fading channels," *Signal Process.*, vol. 80, pp. 1347–1360, 2000.
- [6] B. D. V. Rao and S. -Y. Kung, "Adaptive notch filtering for the retrieval of sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 791–802, 1984.
- [7] A. Nehorai, "A minimal parameter adaptive notch filter with constrained poles and zeros," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 158–161, 1985.
- [8] T. S. Ng, "Some aspects of an adaptive digital notch filter with constrained poles and zeros," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 158–161, 1987.
- [9] P. Händel and A. Nehorai, "Tracking analysis of an adaptive notch filter with constrained poles and zeros," *IEEE Trans. Signal Process.*, vol. 42, pp. 281–291, 1994.
- [10] A. Nehorai and B. Porat, "Adaptive comb filtering for harmonic signal enhancement," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1124–1138, 1986.
- [11] P. J. Parker and B. D. O. Anderson, "Frequency tracking of nonsinusoidal periodic signals in noise," *Signal Process.*, vol. 20, pp. 127–152, 1990.
- [12] M. Niedźwiecki and P. Kaczmarek, "Estimation and tracking of quasiperiodically varying processes," in *Proc. 13th IFAC Symp. System Identification*, Rotterdam, The Netherlands, 2003, pp. 1102–1107.
- [13] N. Wiener, "Generalized harmonic analysis," *Acta Math.*, pp. 117–258, 1930.
- [14] M. Niedźwiecki, "Recursive functional series modeling approach to identification of time-varying plants – More bad news than good?," *IEEE Trans. Autom. Control*, vol. AC-35, pp. 610–616, 1990.
- [15] —, *Identification of Time-Varying Processes*. New York: Wiley, 2000.
- [16] B. Carlsson and P. Händel, "A notch filter based on recursive least-squares modeling," *Signal Process.*, vol. 35, pp. 231–239, 1994.
- [17] S. T. Rao, "The fitting of nonstationary time-series models with time-dependent parameters," *J. Roy. Statist. Soc. B*, vol. 32, pp. 312–322, 1970.
- [18] J. M. Mendel, *Discrete Techniques of Parameter Estimation: The Equation Error Formulation*. New York: Dekker, 1973.
- [19] J. M. Liporace, "Linear estimation of nonstationary signals," *J. Acoust. Soc. Amer.*, vol. 58, pp. 1288–1295, 1975.
- [20] M. Niedźwiecki and P. Kaczmarek, "Identification of quasi-periodically varying systems using the combined nonparametric/parametric approach," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4588–4598, Dec. 2005.
- [21] T. Klaput and M. Niedźwiecki, "A novel approach to estimation of Doppler frequencies of a time-varying communication channel," in *Proc. 2002 Amer. Control Conf.*, Anchorage, AK, 2002, pp. 3213–3218.
- [22] T. Söderström and P. Stoica, *System Identification*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
- [23] M. Niedźwiecki and P. Kaczmarek, "Generalized adaptive notch filters for identification of real quasiperiodically varying systems," in *XIIIth Eur. Signal Processing Conference (EUSIPCO 2004)*, Vienna, Austria, 2004, pp. 1915–1918.
- [24] —, "Tracking analysis of a generalized adaptive notch filter," *IEEE Trans. Signal Process.*, 2005, to be published.
- [25] F. Bui and D. Hatzinakos, "Time-varying channel modeling and variable-size burst for spatio-temporal interference suppression in DS-CDMA systems," in *Proc. 2004 IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Montreal, PQ, Canada, 2004, pp. IV-981–IV-984.



**Maciej Niedźwiecki** was born in Poznań, Poland, in 1953. He received the M.Sc. and Ph.D. degrees from Gdańsk University of Technology, Gdańsk, Poland, and the Dr.Hab. (D.Sc.) degree from the Technical University of Warsaw, Warsaw, Poland, in 1977, 1981, and 1991, respectively.

He spent three years as a Research Fellow with the Department of Systems Engineering, Australian National University (1986–1989). From 1990 to 1993, he was a Vice Chairman of the Technical Committee on Theory, International Federation of Automatic Control (IFAC). He is the author of *Identification of Time-varying Processes* (New York: Wiley, 2000). He is a Professor and Head of the Department of Automatic Control, Faculty of Electronics, Telecommunications and Computer Science, Gdańsk University of Technology. His main areas of research interests include system identification, signal processing, and adaptive systems.



**Piotr Kaczmarek** received the M.Sc. degree in automatic control from Gdańsk University of Technology, Gdańsk, Poland, in 2000, where he is currently pursuing the Ph.D. degree. He graduated from the European Master Degree Course in Control and Management of Lean Manufacturing in Network Systems conducted in cooperation between Gdańsk University of Technology, Catholic University of Louvain, Belgium, and University of Karlsruhe, Germany.

Since 1998 he has been involved in many projects for the biggest Polish air conditioning systems manufacturers. His interests include system identification and adaptive filtering as well as optimization of production techniques.