# Factorized All-pass Based IIR Adaptive Notch Filters

Juan E. Cousseau,<sup>\*,1</sup> Senior Member, IEEE, Stefan Werner,<sup>2</sup> Member, IEEE and Pedro D. Donate<sup>1</sup>

### Abstract—

This paper introduces a new family of IIR adaptive notch filters that forms multiple notches using a second-order factorization of an all-pass transfer function. The new orthogonal realization is amenable for adaptive filtering to obtain the unknown frequencies of interest. Two new adaptive filtering algorithms are presented that can achieve fast convergence at low computational cost. Local convergence analysis for the new algorithms is performed, and a detailed discussion of their properties is provided. The new all-pass based notch realization introduces a different compromise between bias and signal-to-noise ratio (SNR) when compared with realizations previously reported in the literature. Specifically, it achieves lower bias than other approaches at low SNR. This property is particularly attractive for the estimation and tracking of multiple sinusoids. Furthermore, the bias can be made arbitrarily small or can be accurately estimated and compensated for. Extensive computer simulations are provided to illustrate performance of the proposed adaptive notch filters in terms of bias, speed of convergence, and tracking capability.

# I. INTRODUCTION

The classical problem of low complexity multiple sinusoid frequency estimation can be traced back to the *Adaptive Line Enhancer* [1] where mean-square error (MSE) minimization was achieved using an FIR prediction filter structure. FIR solutions have proved to be inefficient to recover sinusoids in noise, mainly because a high-order filter is required to model a deep notch filter. A more practical alternative is IIR based adaptive notch filters (ANF) or narrow bandpass filters with a very selective frequency characteristic. Although an exact solution is not available, efficient approximations can be obtained using IIR ANF realizations of adequate order.

IIR ANF are today a low-complexity alternative for frequency estimation in many application areas. Examples of communications systems applications are narrowband interference cancellation in direct sequence spread spectrum signals [2] and fast fading channel estimation [3][4].

A popular approach for IIR ANF, proposed in [5], utilizes a canonical <sup>1</sup> direct-form realization of order 2M, M being the number of unknown frequencies. The zeros of the ANF are located on the unit circle, and the modulus of the poles (with the same angles as the zeros but inside the unit circle) is a user defined parameter. The properties and the accuracy of the ANF in [5] have been extensively studied in the literature, see, e.g., [6], [7], [8], [9], [10]. Although classical estimation properties can be related to this model [11], the estimated frequencies are not explicitly available and need to be recursively computed from the roots of a high-order (2M) polynomial. Alternative ANF realizations based on the approach in [5] have been studied in the past, see [12], [13], [14]. These works proposed to use a cascade of second-order ANF to enable simple calculation of the frequencies of interest. Two different realizations for the second-order ANF with constant quality factor or notch bandwidth can be considered. As concluded in [12], the frequency estimates obtained with this kind of ANF are biased due to input measurement noise.

In [15], a cascade of second-order all-pass based ANF sections on a serial - sinusoid canceler configuration was considered. The individual second-order sections are based on constant bandwidth notch realizations. An ad-hoc updating algorithm using local errors was proposed in [16]. The serial-sinusoid canceling strategy for multiple sinusoid estimation has low computational complexity. However, the use of local errors in the ad-hoc updating algorithm may lead to an unstable behavior, particularly in tracking applications at low signal-to-noise ratio (SNR).

A different approach for low-complexity frequency estimation follows the phase differences concept [17], [18]. A representative algorithm, based on a high SNR assumption, is the Multiple Frequency Tracker [9] that uses phase differences to estimate the instantaneous frequencies of multiple sinusoids. The generalized adaptive notch filter [19] is a recently proposed family of algorithms that extend the phase differences technique from signal to system modeling problems using the classical basis function approach.

Local convergence analysis of both ANF approaches (ARMA modeling and phase differences) have been performed for the case of one or two sinusoids, e.g., [8], [20], [10], [21]. The general requirement of minimum separation of frequencies and high SNR seems to be necessary for all realizations.

This paper proposes an ANF model for multiple sinusoid estimation based on a high order all-pass filter factorized using second-order all-pass sections. The model can be seen as a generalization of the model discussed in [15] for frequency estimation of real sinusoids, and for the single-sinusoid case the model in [15] coincide with ours. The proposed factorization have several advantages when compared with previous approaches. In particular, the estimated notch frequencies are not influenced in the mean by the measurement noise like

<sup>\*</sup> Corresponding author. This work was partially supported by Universidad Nacional del Sur, Project PGI # 24-K023, Agencia Nacional de Promocion Cientifica y Tecnologica, Project PICT 21723, and the Academy of Finland, Smart and Novel Radios (SMARAD) Center of Excellence. <sup>1</sup>CONICET-Department of Electrical and Computer Eng. Universidad Nacional del Sur, Argentina. <sup>2</sup>Helsinki University of Technology, Signal Processing Laboratory, Finland.

 $<sup>^{1}</sup>$ Minimum number of parameters corresponding to each of M unknown frequencies to estimate

previously proposed multiple notch IIR adaptive filters [12] [13]. Also, the proposed ANF introduces a different source of bias in the estimated frequencies. This bias is directly related to the all-pass based realization and, if required, it can be accurately estimated and removed. The proposed realization guarantees unit gain maxima between notches. As a result, good behavior at low SNR can be expected, which is not the case for conventional cascaded notch realizations. We also present two efficient adaptation algorithms that are based on different minimization criteria. The algorithms use an orthogonal basis (with respect to the measurement noise) that leads to fast convergence. Computer simulations show that the new model presents low bias and fast convergence even at low SNR. Preliminary results related to the proposed realization and one of the related algorithms were presented in [22].

The paper is organized as follows. Section II introduces the new all-pass based model for IIR multiple notch filters, and discusses its implementation using a set of orthogonal basis functions. In Section III, two novel ANF algorithms are derived; one using a recursive prediction error approach, and the other using an iterative Steiglitz-McBride based approach. In this section we also provide local convergence analysis of the algorithms and detail their most important characteristics. Section IV compares the proposed ANF with other ANF available in the literature via computer simulations. Finally, conclusions are presented in Section V.

#### II. ALL-PASS BASED NOTCH FILTERS

This section first describes the notch filtering problem and briefly reviews the approach of using second-order notch filter sections as building blocks. Thereafter, a new all-pass based factorization is presented. Finally, orthogonal realizations are discussed that have good properties in presence of measurement noise, and render fast converging adaptive implementations (see Section III.)

#### A. Problem description

Consider a signal u(n) formed by M sinusoids that is buried in additive broadband (white) noise  $\nu(n)$  with variance  $\sigma_{\nu}^2$ :

$$u(n) = \sum_{i=1}^{M} p_i \sin(w_{oi}n + \eta_i) + \nu(n)$$

where  $p_i$  is the amplitude,  $w_{oi}$  is the frequency and  $\eta_i$  is the phase of sinusoid *i*, respectively. This signal is processed by a linear filter H(z) to obtain an output signal y(n) that is ideally free of noise, i.e., composed only by the sinusoids. Therefore, H(z) has to be a multiple notch or multiple narrowband adaptive filter with capability to detect and track the input signal. The notch or passband frequencies are related to the roots of the numerator and the denominator polynomials of H(z). Finding these roots means solving high order polynomials (2M), which is a costly operation considering the dynamic nature of H(z). The process of root finding is greatly simplified if H(z) is built from a cascade of second order filters. A new all-pass based form is also possible. These two realizations are detailed in the following.

# B. Cascaded second order notch filter

The overall transfer function has the form

$$H_C(z) = \prod_{i=1}^M H_i(z) \tag{1}$$

where  $H_i(z)$  is the transfer function of a suitably parameterized second-order notch filter [16]. An efficient and popular realization for the second order sections is the lattice form,

$$H_i(z) = \frac{1}{2}[1 + V_i(z)]$$

with  $V_i(z) = \frac{\overline{D}_i(z)}{D_i(z)} = \frac{s_{2i}+s_{1i}(1+s_{2i})z^{-1}+z^{-2}}{1+s_{1i}(1+s_{2i})z^{-1}+s_{2i}z^{-2}}$ . The parameters  $s_{1i}$  and  $s_{2i}$  are related with the two lattice

The parameters  $s_{1i}$  and  $s_{2i}$  are related with the two lattice parameters  $\theta_1^i$  and  $\theta_2^i$  of each section through the equalities  $s_{1i} = \sin \theta_1^i$  and  $s_{2i} = \sin \theta_2^i$ ,  $1 \le i \le M$ . Furthermore,  $s_{1i}$  and  $s_{2i}$  relate to the notch frequency and the 3 dB notch bandwidth  $B_i$  as

$$s_{1i} = -\cos w_{oi}$$
  $s_{2i} = \frac{1 - \tan(B_i/2)}{1 + \tan(B_i/2)}$  (2)

In order to obtain a fast initial convergence, a predefined exponential profile can be associated with  $s_{2i}$  to tune each section [5]. It is common to use the following exponential profile:

$$s_{2i}(n+1) = \rho_i s_{2i}(n) + (1-\rho_i) s_{2i}^{\infty}$$
(3)

where  $\rho_i$  is related to the exponential decay time constant and  $s_{2i}^{\infty}$  is the asymptotic value of  $s_{2i}$ .

Let  $\phi_i(w)$  denote the phase of each section when evaluated on the unit circle. Then  $V_i(e^{jw}) = e^{j\phi_i(w)}$  is given by

$$V_i(e^{jw}) = \left\{ \cos^2[\phi_i(w)/2] - \sin^2[\phi_i(w)/2] ... + j2\sin[\phi_i(w)/2]\cos[\phi_i(w)/2] \right\}^2$$

where the phase relates to the coefficients of the section as

$$\cos\left[\frac{\phi_i(w)}{2}\right] = (1+s_{2i})\frac{(s_{1i}+\cos w)}{|D_i(w)|}$$

$$\sin\left[\frac{\phi_i(w)}{2}\right] = -(1-s_{2i})\frac{\sin w}{|D_i(w)|}$$
(4)

with  $|D_i(w)|^2 = (1 + s_{2i})^2 (s_{1i} + \cos w)^2 + (1 - s_{2i})^2 \sin^2 w$ . The all-pass section  $V_i(z)$  can also be related to  $H_i(z)$  and  $G_i(z)$ , i.e., the transfer functions of the second-order notch and narrowband filters, respectively. On the unit circle, these functions can be described using the lattice parameters of Eq. (4)

$$H_{i}(e^{jw}) = \frac{1}{2} \left[ 1 + V_{i}(e^{jw}) \right] = \cos \left[ \phi_{i}(w)/2 \right] e^{j\phi_{i}(w)/2}$$
(5)  
$$G_{i}(e^{jw}) = \frac{1}{2} \left[ 1 - V_{i}(e^{jw}) \right] = -j \sin \left[ \phi_{i}(w)/2 \right] e^{j\phi_{i}(w)/2}$$

# C. Factorized all-pass based notch filter

In this paper we propose to use a *factorized all-pass based realization* (Factorized Adaptive Notch Filter, FANF) for the multiple notch filter. This realization is described by

$$H_A(z) = \frac{1}{2} \left[ 1 + V(z) \right] = \frac{1}{2} \left[ 1 + \prod_{i=1}^M V_i(z) \right]$$
(6)

Evaluating Eq. (6) on the unit circle gives

$$H_A(e^{jw}) = e^{j\phi_i(w)/2} e^{j\phi^i(w)/2} \cos\left\{\frac{1}{2} \left[\phi_i(w) + \phi^i(w)\right]\right\}$$
(7)

where  $\phi^i(w) = \sum_{k=1, k\neq i}^M \phi_k(w)$  is the sum of all the contributions to the phase of V(z) except the one corresponding to the *i*th section. From this equation it is possible to verify that

$$||H_A(e^{jw})||^2 = \cos^2\left\{\frac{1}{2}\left[\phi_i(w) + \phi^i(w)\right]\right\}$$
(8)

Below we will discuss important properties of the FANF model in Eqs. (6)–(8). In particular, we provide details on how the notch frequencies relate to the lattice parameters and properties of the frequency response.

1) Model parameters: The notch frequencies  $\overline{w}_i$ ,  $1 \le i \le M$  of the factorized all-pass IIR notch filter can be obtained by solving the *M*th-order polynomial in  $\cos w$  in (8). Considering an adaptive implementation, it is more efficient to exploit the relationship that exists between the notch frequencies  $\overline{w}_i$  and 3 dB bandwidth  $\overline{B}_i$  of the signal model and the parameters of the FANF realization (i.e.,  $s_{1i}$  and  $s_{2i}$ ).

Using (5) and (6) it is easy to show that at a notch frequency  $\overline{w}_i$  the following must hold

$$\cos \overline{w}_i = -s_{1i} + \left(\frac{1 - s_{2i}}{1 + s_{2i}}\right) \sin \overline{w}_i \tan \phi^i(\overline{w}_i)/2$$

Since  $\phi^i(\overline{w}_i) \cong 2\pi k$  when  $s_{2i} \to 1$ , we have  $\tan \phi^i(\overline{w}_i)/2 \cong \sum_{j=1, j \neq i}^M \tan \phi_j(\overline{w}_i)/2$ . Therefore, the previous equation can be well approximated by

$$\cos \overline{w}_i \cong -s_{1i} + \left(\frac{1-s_{2i}}{1+s_{2i}}\right)^2 \sin^2 \overline{w}_i \sum_{j=1, j \neq i}^M \left(\frac{1}{s_{1j} + \cos \overline{w}_i}\right)$$
(9)

For the case when  $s_{2i} \rightarrow 1$ , the second term of Eq. (9) becomes negligible and we get  $\cos \overline{w}_i \cong -s_{1i}$ . When  $s_{2i} < 1$ , the frequencies  $w_{oi}$  are different than  $\overline{w}_i$ . However, Eq. (9) can be used to obtain an improved estimate (bias removal) if necessary (more details in Section IV).

Finally, the  $s_{2i}$  parameters can be related to the 3 dB bandwidth parameters  $B_i$  using Eq. (2). In fact, the true  $\overline{B}_i$  related to the FANF realization are upper bounded by Eq. (2) [23].

2) Frequency response: The phase of a stable all-pass filter  $\phi(w)$  decreases monotonically from 0 to  $-2M\pi$  when  $0 \le w \le \pi$  [24]. Based on the all-pass characteristic of the FANF model, we conclude that there exist M frequencies  $\overline{w}_1 < \overline{w}_2 < \cdots < \overline{w}_M$  with  $\phi(\overline{w}_i) = -(2i-1)\pi$  (for  $i = 1, \dots, M$ ) that correspond to the minima of (7), i.e.,

$$H_A(e^{j\overline{w}_i}) = \cos\left\{\phi(\overline{w}_i)/2\right\} = 0 \tag{10}$$

There are also M + 1 frequencies  $0 \le \overline{w}_1 < \cdots < \overline{w}_M$ with  $\phi(\overline{w}_i) = -2i\pi$  (for  $i = 1, \cdots, M$ ) that correspond to the maxima of Eq. (7), i.e.,  $H(e^{j\overline{w}_i}) = 1$ . Those frequencies can be obtained by

$$\tan\left\{\phi(\overline{w}_i)/2\right\} = 0\tag{11}$$

As a consequence of the parameterization, the frequency response of Eq. (6) is different from that of the cascaded notch filter described by Eq. (1). The FANF realization *always provides maxima of unit magnitude*, while the maxima of the cascaded notch realization will be smaller. Therefore, FANF provides a better isolation of notch frequencies which is an important property when considering a low SNR.

#### D. Orthogonal FANF

One advantage of the FANF model is the possibility to implement it using an orthogonal set of functions,  $F_i(z)$  that can be generated from the all-pass sections. Thus, enabling adaptive implementations that provide fast convergence at low computational cost.

Since the basic construction blocks for the FANF are second-order sections, there are many ways to build the complete filter. A possible solution is to use the set of orthogonal functions proposed in [25] (see also [26])

$$F_{i}(z) = \frac{z^{-1}}{D_{i}(z)} \prod_{k=1, : k \neq i}^{M} \frac{\overline{D}_{k}(z)}{D_{k}(z)}$$
(12)

or the set considered in [14]

$$F'_{i}(z) = \frac{z^{-1}}{D_{i}(z)} \prod_{k=1}^{i-1} \frac{\overline{D}_{k}(z)}{D_{k}(z)}$$
(13)

where, as before,  $D_i(z)$  is the corresponding denominator of the second-order all-pass  $V_i(z)$ . Note that each second-order section is related to a different sinusoid. Therefore, both basis functions  $F_i(z)$  and  $F'_i(z)$  are equivalent.

The orthogonality of these realizations introduces useful properties when measurement noise is present, which is the typical case in notch filtering applications. Consider the minimization of the output signal variance, defined as

$$\mathbf{E}[y^2(n)] = \langle H_A(z), H_A(z) \rangle_{S_u} + \langle H_A(z), S_\nu(z) H_A(z) \rangle$$
(14)

where the first term describes the inner product induced by the signal components, i.e.,

$$\langle H_A(z), H_A(z) \rangle_{S_u} = \sum_{i=1}^M p_i^2 |H_A(e^{jw_i})|^2$$
 (15)

and the second term is the standard inner product in  $L_2$ ,

$$\langle H_A(z), S_\nu(z)H_A(z)\rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_\nu(e^{jw}) |H_A(e^{jw})|^2 dw$$
 (16)

with  $S_{\nu}(z)$  denoting the power spectral density of the white noise sequence  $\nu(n)$ . Using Eqs. (15) and (16), Eq. (14) can be written as

$$E[y^{2}(n)] = \sum_{i=1}^{M} p_{i}^{2} |H_{A}(e^{jw_{i}})|^{2} + \sigma_{\nu}^{2} ||H(z)||_{2}^{2}$$
(17)

where the noise-induced term (assuming  $s_2 = s_{2i}$ ,  $1 \le i \le M$ for clarity)

$$\sigma_{\nu}^{2} \|H(z)\|_{2}^{2} = \sigma_{\nu}^{2} \|1 + \prod_{i=1}^{M} V_{i}(z)\|_{2}^{2} = \frac{\sigma_{\nu}^{2}}{2} \left[1 + s_{2}^{M}\right]$$
(18)

is independent of the lattice parameter  $\theta_1^i$ , i.e., independent of  $s_{1i} = \sin \theta_1^i$ . Therefore, the measurement noise only introduces a change in the residual noise when an orthogonal structure is used. This is different to the case of using a cascade realization, where white measurement noise introduces a noise-induced component in the parameter descriptions [12].

Next section considers an efficient adaptive notch filtering algorithms based on the factorized all-pass IIR notch model that uses the orthogonal realization  $F_i(z)$  in Eq. (12).

# III. ADAPTIVE ALGORITHMS FOR FANF

A family of all-passed based ANF can be developed as in [27]. Two types of algorithms are proposed here: the first algorithm is designed using a *Recursive Prediction Error* (RPE) approach (similar to [5]), and the second minimizes an iterative criterion that can be related to the Steiglitz-McBride error (SME) [28]. As a consequence of the orthogonal realization used here, the considered normalized stochastic gradient implementations will enjoy fast convergence at low computational cost.

In the end of each section, an analysis of the stationary points and local convergence is provided. The convergence study is performed using the associated ordinary differential equation (ODE) method [11]. The general ODE approach to convergence analysis comes from the field of stochastic approximation theory, and changes the convergence study of a stochastic nonlinear equation by a stability study of the solutions to a deterministic differential equation. As a consequence, the convergence properties of the discrete parameter adaptation algorithm are strongly related to the stability of the solutions to the differential equation. Two different kind of algorithms can be studied in this form: 1) vanishing gain algorithms (i.e., with the stepsize  $\mu \to 0$ ), mostly oriented to estimation in a stationary environment, and; 2) constant gain algorithms, where the stepsize  $\mu$  is kept constant to enable tracking studies. Our interest here is to study the latter case of constant gain for which the ODE method guarantees that the adaptation algorithm converges in probability (not with probability one as with vanishing gain algorithms).

The ODE association introduces certain regularity conditions that needs to be verified for the proposed algorithms. These conditions are related to the differentiability and boundedness of the criterion and the average direction used. They are easily verified in general, as discussed in [12] for the cascaded 4

realization and in [8] for the direct-form realization, since the only difference with these approaches is related to the notch filter realization. Due to space limitations, we refer the reader to [8] and [12] for the technical details of the ODE association.

#### A. RPE factorized all-pass IIR ANF (RFANF)

Similarly to the ANF in [5], we can derive a FANF whose parameters minimize the recursive prediction error (RPE) criterion. The updating equations of the resulting normalized stochastic gradient based algorithm, referred to as the *Recursive Prediction Error all-pass IIR ANF* (RFANF), are given by

$$\theta_1^i(n+1) = \theta_1^i(n) - \frac{\mu_{rpe}}{r_{rpe}^i(n)} \Psi_{\theta_1^i}(n) y(n)$$
  
$$r_{rpe}^i(n+1) = (1-\lambda) r_{rpe}^i(n) + \mu_{rpe} |\Psi_{\theta_1^i}(n)|^2$$

where  $\Psi_{\theta_1^i}(n)$  is the regressor

$$\Psi_{\theta_i^i}(n) = -(1+s_{2i})c_{1i}G_i(q)F_i(q)u(n-1)$$

where  $c_{1i} = \cos \theta_1^i$ ,  $\lambda$  is the forgetting factor ( $0 < \lambda < 1$ ) and  $\mu_{rpe} \cong 1 - \lambda$  is the step size. The RFANF algorithm, without the terms  $G_i(z)$  (see discussion in the following section), is summarized in Table I.

The computational complexity of this realization is similar to (slightly lower than) that of [12] and [13]. Due to the interchangeability of the sections, the RFANF algorithm has a multimodal mean square error surface (M! equivalent minima). Thus, the order in which each section converge depends on the initial conditions. This is a common characteristic of adaptive algorithms related to cascaded realizations [12], [13].

Characterization and properties: As previously discussed, the noise-induced term that contributes to the signal variance (17) is independent of the parameters  $\theta_1^i$ . Therefore, we study only the signal-induced part of  $E[y^2(n)]$ . From Eq. (8) we get

$$\frac{\partial \theta_1^i(t)}{\partial t} = -2 \left\langle \frac{\partial H_A(z)}{\partial \theta_1^i}, H_A(z) \right\rangle_{S_u} \\ = -\sum_{k=1}^M p_k^2 \cos\left\{ \left[ \phi_i(w_{ok}) + \phi^i(w_{ok}) \right] / 2 \right\} \times \\ \left\{ \left[ \frac{\partial}{\partial \theta_1^i} \cos \phi_i(w_{ok}) / 2 \right] \cos \phi^i(w_{ok}) / 2 \\ - \left[ \frac{\partial}{\partial \theta_1^i} \sin \phi_i(w_{ok}) / 2 \right] \sin \phi^i(w_{ok}) / 2 \right\}$$

Assuming that  $|\theta_1^i| < \pi/2$ , i.e.,  $c_{1i} > 0$  gives

$$\frac{\partial \theta_{1}^{i}(t)}{\partial t} = -\sum_{k=1}^{M} p_{k}^{2} \frac{(1+s_{2i})c_{1i}}{|D_{i}(e^{jw_{ok}})|} \sin \phi_{i}(w_{ok})/2 \times \cos\left\{ \left[\phi_{i}(w_{ok}) + \phi^{i}(w_{ok})\right]/2 \right\} \times \sin\left\{ \left[\phi_{i}(w_{ok}) + \phi^{i}(w_{ok})\right]/2 \right\}$$
(19)

As expected, the stationary points of this algorithm, i.e., the solutions when the previous equation is equal to zero, coincide with the minima and the maxima of the FANF realization, Eqs. (10) and (11), respectively.

In order to compare the RFANF with the algorithms in [12], [13], we need the ODE associated to the RPE algorithm employing a cascade realization (1). This ODE is given by

$$\frac{\partial \theta_1^i(t)}{\partial t} = -\sum_{k=1}^M p_k^2 \frac{(1+s_{2i})c_{1i}}{|D_i(e^{jw_{ok}})|} \sin \phi_i(w_{ok})/2 \times \left(\prod_{n=1,n\neq i}^M \cos^2 \phi_n(w_{ok})/2\right) \cos \phi_i(w_{ok})/2$$
(20)

Using Eq. (1), we realize that the term  $\left(\prod_{n=1,n\neq i}^{M} \cos^2 \phi_n(w)/2\right)$  in the cascaded algorithm can be written as  $\left(\prod_{n=1,n\neq i}^{M} |H_n(e^{jw})|^2\right)$ . The counterpart of this term in Eq. (19) has the form  $\cos \phi^i(w)/2$ . Both terms have a similar behavior for  $s_{2i} \rightarrow 1$ .

The term  $\sin \phi_i(w_{ok})/2$  in (19) and (20) can be associated to the narrow passband transfer function  $G_i(e^{jw})$  and can be eliminated in a practical simplified implementation [16]. Since this term reduces the magnitude of the gradient component, faster convergence can be expected if it is left out.

Furthermore, if the term  $\sin \phi_i(w_{ok})/2$  is left out, we can perform an approximate local convergence analysis. In a neighborhood of a stationary point, we can assume that  $\cos \{ [\phi_k(w) + \phi^k(w)]/2 \} \cong 0$  for all  $k \neq i$  (and  $\sin \{ [\phi_k(w) + \phi^k(w)]/2 \} \cong 1$ ). As a consequence, Eq. (19) can be rewritten as

$$\frac{\partial \theta_1^i(t)}{\partial t} \cong -p_i^2 \frac{(1+s_{2i})c_{1i}}{|D_i(e^{jw_{oi}})|} \cos\left\{ \left[\phi_i(w) + \phi^i(w)\right]/2 \right\} \\ = -p_i^2 \frac{(1+s_{2i})c_{1i}}{|D_i(e^{jw_{oi}})|} \left[\cos\phi^i(w_{oi})/2\cos\phi_i(w_{oi})/2 - \sin\phi^i(w_{oi})/2\sin\phi_i(w_{oi})/2\right] (21)$$

Note that the second term in Eq. (21) tends to zero faster than the first term when  $s_{2i} \rightarrow 1$ . A similar behavior is expected for the cascade realization when  $s_{2i} \rightarrow 1$ . Invoking this assumption, Eq. (21) will in a neighborhood of a stationary point reduce to

$$\frac{\partial \theta_1^i(t)}{\partial t} \cong -p_i^2 \frac{(1+s_{2i})c_{1i}}{|D_i(e^{jw_{oi}})|} \cos \phi^i(w_{oi})/2 \cos \phi_i(w_{oi})/2$$
(22)

Equation (22) corresponds (when  $s_{2i} \rightarrow 1$ ) to the ODE obtained with the RPE methodology for a second-order section. We can, therefore, conclude that it converges to  $s_{1i} \cong$  $-\cos w_{oi}$  [16]. Since only  $s_{1i}$  is involved in the computation of the gradient component *i*, the gradient components at a stationary point are approximately orthogonal. Therefore, the Hessian matrix related to a general recursive prediction error algorithms is diagonal. This justifies the normalization factor  $r_{rpe}^{i}(n), 1 \le i \le M$  included in the RFANF algorithm.

Finally, we note that the RFANF estimates are obtained by means of a stochastic gradient variant of a general recursive prediction error algorithm [11]. Therefore, we can expect low variance of the MSE for a parameter close to the minimum. On the other hand, recursive prediction error algorithms like the RFANF can be sensitive to initial conditions. An alternative approach that overcomes this problem is developed in the following section.

#### B. SME factorized all-pass IIR ANF (SFANF)

Using the factorized all-pass model an *a posteriori* off-line error, linear in the parameters, can be used to minimize the output signal variance. The minimization of the following error is considered

$$e(n+1) = \frac{1}{2} \left[ \prod_{i=1}^{M} \frac{\overline{D}_{i}^{n+1}(q)}{D_{i}^{n}(q)} + \prod_{i=1}^{M} \frac{D_{i}^{n+1}(q)}{D_{i}^{n}(q)} \right] u(n) \quad (23)$$

In order to obtain an on-line algorithm that is suitable for adaptive notch filtering, the *a priori* error obtained from Eq. (23) can be written as

$$e(n) = \frac{1}{2} \left[ 1 + \prod_{i=1}^{M} \frac{\overline{D}_{i}^{n}(q)}{D_{i}^{n}(q)} \right] u(n)$$

Following the instantaneous gradient of the on-line MSE, the regressor is given by

$$\Psi_{\theta_1^i}(n) = -\left[\frac{q^{-1}}{D_i(q)} + F_i(q)\right]u(n)$$
(24)

where the set of orthogonal functions  $F_i(z)$  previously introduced are used. Finally, the update equations for the *Steiglitz-McBride Factorized all-pass based IIR ANF* (SFANF) are

$$\theta_{1}^{i}(n+1) = \theta_{1}^{i}(n) - \frac{\mu_{sm}}{r_{sm}^{i}(n)}e(n)\Psi_{\theta_{1}^{i}}(n)$$
  
$$r_{sm}^{i}(n+1) = (1-\lambda)r_{sm}^{i}(n) + \mu_{sm}|\Psi_{\theta_{1}^{i}}(n)|^{2} (25)$$

where  $0 < \lambda < 1$  is the forgetting factor and  $\mu_{sm} \cong 1 - \lambda$  is the step size. A summary of SFANF algorithm is provided in Table II.

A justification for the normalization factor  $r_{sm}^i(n)$  will be given in the analysis provided below. The SFANF algorithm can be seen as an efficient implementation of the approach in [28]. The good behavior for low SNR scenarios reported in [28] is also inherited by the SFANF algorithm.

*Characterization and properties:* To study the SFANF stationary points and convergence properties we consider the ODE associated to (25), without normalization. Employing the inner product notation gives us the following ODE

$$\frac{\partial \theta_1^i(t)}{\partial t} = -\left\langle \left[ \frac{z^{-1}}{D_i(z)} + F_i(z) \right], [1 + V(z)] \right\rangle_{S_u} -\sigma_\nu^2 \left\langle \left[ \frac{z^{-1}}{D_i(z)} + F_i(z) \right], [1 + V(z)] \right\rangle (26)$$

which is formed by a signal-induced part and a noise-induced part, respectively. Using the fact that  $F_i(z)$  is an orthogonal basis related to V(z) [16][26], we get for the noise-induced part

$$\left\langle \left[ \frac{z^{-1}}{D_i(z)} + F_i(z) \right], [1+V(z)] \right\rangle = \left\langle \frac{z^{-1}}{D_i(z)}, 1 \right\rangle + \left\langle \frac{z^{-1}}{D_i(z)}, \prod_{k=1}^M V_k(z) \right\rangle + \left\langle \frac{z^{-1}}{D_i(z)} \prod_{k=1, k \neq i}^M V_k(z), 1 \right\rangle + \left\langle \frac{z^{-1}}{D_i(z)} \prod_{k=1, k \neq i}^M V_k(z), \prod_{k=1}^M V_k(z) \right\rangle = \left\langle \frac{z^{-1}}{D_i(z)}, 1 \right\rangle + 2 \left\langle 1, z^{-1} F_i(z) \right\rangle + \left\langle \frac{z^{-1}}{D_i(z)}, V_i(z) \right\rangle = 0$$

The first two terms are zero because they are inner products that involve projections of strictly causal functions on a constant. The last term is zero due to the orthogonality of the basis function with respect to the corresponding all-pass section [16]. This indicates that the noise-induced term does not have any influence on the stationary points. As mentioned before, this is not the case with the multiple notch IIR ANF previously reported in literature.

Using the second-order section all-pass phase  $\phi_i(w)$ , the signal-induced part can after some algebraic manipulations be written as (assuming  $|\theta_1^i| < \pi/2$ , i.e.,  $c_{1i} > 0$ ),

$$\frac{\partial \theta_{1}^{i}(t)}{\partial t} = -\sum_{k=1}^{M} p_{k}^{2} \frac{(1+s_{2i})c_{1i}}{|D_{i}(w_{ok})|} \cos\left[\phi^{i}(w_{ok})/2\right] \times (27)$$

$$\cos\left[\phi_{i}(w_{ok})/2 + \phi^{i}(w_{ok})/2\right] \quad (28)$$

where, as before,  $\phi^i(w) = \sum_{m=1, m \neq i}^M \phi_i(w)$ .

From (28) we conclude that the stationary points of the SFANF, i.e,  $\cos \left[\phi_i(w)/2 + \phi^i(w)/2\right] = 0$ , coincide with the notch frequencies of the factorized all-pass based IIR notch realization.

On the other hand, the term  $\cos \left[\phi^i(w)/2\right]$  in Eq. (28) is zero if  $\phi^i(w) = (2n-1)\pi$ , for some integer  $n, n = 1, \dots, M$ . Due to the phase constraint of the FANF realization this implies that  $\cos \left[\phi_i(w)/2\right] = 0$ . However, from Eq. (11) we see that the particular solution for  $\theta_1^i$  that fulfill this condition is not a zero related to the FANF realization. Therefore, the term  $\cos \left[\phi^i(w)/2\right]$  does not introduce any stationary points in Eq. (28).

The main difficulty to perform a conventional convergence analysis that follows the Liapunov second method [16] is the existence of multiple equivalent stationary points in the SFANF algorithm. In order to overcome this problem we follow a local convergence study showing that for  $s_{2i} \rightarrow$ 1,  $s_{1i} \cong \cos w_{ok}$ , for  $i, k = 1, \dots, M$  are locally stable stationary points of Eq. (28). After that, it is shown that the only attractive stationary points correspond to those of the FANF realization.

Using Eq. (4), Eq. (28) can be written as

$$\frac{\partial \theta_1^i(t)}{\partial t} = B_{1i} + B_{2i}$$

where

$$B_{1i} = -\sum_{k=1}^{M} p_k^2 \frac{(1+s_{2i})^2 c_{1i}}{|D_i(w_{ok})|^2} \cos^2 \left[\frac{\phi^i(w_{ok})}{2}\right] (s_{1i} + \cos w_{ok})$$
$$B_{2i} = \sum_{k=1}^{M} p_k^2 \frac{(1-s_{2i}^2) c_{1i}}{|D_i(w_{ok})|^2} \cos \left[\frac{\phi^i(w_{ok})}{2}\right] \sin \left[\frac{\phi^i(w_{ok})}{2}\right] \sin w_{ok}$$

The term  $B_{2i}$  represents a deterministic bias (for a constant  $w_{oi}$ ), that can be asymptotically eliminated using  $s_{2i} \rightarrow 1$ . This is directly related to the factorized all-pass based model as stated previously.

Let us assume that  $\cos^2 \phi^i(w)/2 \neq 0$  and  $|D_i(w_{ok})|^2$  nonzero at  $w_{oi}$ , i.e., all sections converge except the *i*th section as  $s_{2i} \rightarrow 1$ . Then, the dominant term in  $B_{1i}$  is

$$\frac{\partial \theta_1^i(t)}{\partial t} \cong -p_i^2 \frac{(1+s_{2i})^2 c_{1i}}{|D_i(w_{oi})|^2} (s_{1i} + \cos w_{oi})$$
(29)

which has as stable solution  $s_{1i} \cong -\cos w_{oi}$  [16]. There exist M stationary points as determined by the factorized all-pass based realization and, at least locally, all are attractive.

Since only  $s_{1i}$  appears in Eq. (29) (i.e., the gradients are approximately orthogonal at the stationary points), then a normalized stochastic gradient algorithm will have similar convergence behavior to a complete Gauss-Newton algorithm [13]. This justifies the choice of the normalization factor  $r_{sm}^i(n)$ . To illustrate this behavior, the solutions of Eq. (28) for  $s_{11}$  and increasing values of  $s_{2i}$  are depicted in Figure 1. As can be observed, there are four possible solutions for the ODE of the FANF realization. We see that only one solution has appreciable magnitude. This is different to what happens with the partial cascaded realization [15], where similar gradient magnitude is obtained for each solution. An exponential profile for  $s_{2i}$ , as the one in (3), can be used to overcome the gradient regions with low magnitude (right part of Figure 1). Also, we can use different parameters for each profile to reduce problems related with multiple equivalent minima.

It can be verified that the simplified version of the RFANF algorithm (without the narrow passband filters at each regressor component) has the same stationary point equation as obtained in Eq. (28) for SFANF. However, the simplified RFANF is an approximate prediction error algorithm, whereas SFANF is not. Therefore, local convergence properties may be similar but the general convergence properties are in general different.

If computational complexity is a concern, certain simplifications can be done to reduce the complexity of the SFANF algorithm. For example, the regressor computation can be reduced to the essential filtering verified at the stationary point, i.e., by replacing Eq. (24) by

$$\Psi_{\theta_1^i}(n) = -\left[\frac{q^{-1}}{D_i(q)}\right] u(n)$$

that for the single sinusoid case coincides with the algorithm proposed in [17]. However, such simplifications (and other ones too) will lead to poor performance at low SNR or in a tracking scenario with several sinusoids. Therefore, simplifications of the SFANF algorithm will not be considered here.

By replacing  $\cos \phi^i(w)/2$  by  $\prod_{n=1,n\neq i}^M \cos \phi_i(w)/2$  in (26), we can conclude that the notch frequencies can be accurately estimated as  $s_2 \rightarrow 1$ . For the case when  $s_2 < 1$ , an efficient correction to obtain  $w_{oi}$  can be found from Eq. (9), as discussed in the next section.

#### IV. EVALUATION AND COMPARISONS

In this section, a complete performance study is presented for the new FANF model and its related adaptive filtering algorithms. In particular, we study how the bias is affected by the SNR and filter parameters, and how to accurately estimate it if needed. We also illustrate the good isolation of notches obtained by the new FANF model, which is a useful property at low SNR. Finally, convergence and tracking properties (chirps and frequency jumps) of the new algorithms are illustrated. For comparison purposes, the algorithms of [12] and [20] were implemented.

# A. Frequency response and bias properties of FANF

This example illustrate important properties of the proposed FANF realization. Firstly, the phase and magnitude responses are studied. Thereafter, we show how the bias relates to parameter  $s_{2,i}$  and the filter order. Finally, we show how to successfully estimate the bias.

Figure 2 shows magnitude and phase responses of the proposed FANF realization and the cascade realization of [12], both having an order of eight (i.e., M = 4 notches). To better visualize the differences, parameter  $s_{2i}$  is kept fixed and chosen to 0.6. We see that the FANF realization confirms the results in Section II and have minima (notches) and maxima fixed on the unit circle. The maxima between notch frequencies is lower for the cascade-realization. This property becomes important in low SNR contexts. For these cases the good isolation between notches provided by the FANF realization can reduce the risk that sections exchange equivalent estimated parameters. Note that the bias of the FANF is large due to the choice of the user-defined parameter  $s_{2i}$ . Next we illustrate that the bias approaches zero as  $s_{2i}$  takes on more practical values.

Figure 3 shows the *normalized bias* as a function of  $s_{2i}$  for three realizations of order 4, 6 and 8 (i.e, with M = 2, M = 3 and M = 4 frequency notches). The bias is calculated by finding the roots of (8). The frequencies were chosen uniformly distributed in  $(0, \pi)$ .

The normalized bias is defined as

$$\mathcal{B}_{i} = \sqrt{\frac{\sum_{k=1}^{M} (\overline{w}_{ok}^{i} - w_{ok}^{i})^{2}}{\sum_{k=1}^{M} w_{ok}^{i}}^{2}}$$

where  $\overline{w}_{ok}^i$  is the *k*th notch frequency that corresponds to the *i*th-order FANF realization, and  $w_{ok}^i$  is the true *k*th notch

frequency of an *i*th-order cascade realization. We see that the bias is mostly independent of the order of the realization. This indicates that only proximity between notch frequencies will affect the FANF frequency notches. This is common for the available algorithms for frequency estimation [10]. The bias related to a FANF realization that uses practical values of  $s_{2i}$  (0.95 and higher) is usually very low.

If required, Eq. (9) can be used for bias correction. The bias estimate for each notch frequency can be obtained as

$$\operatorname{Bias}_{\overline{w}_{i}} \cong K \left(\frac{1 - s_{2i}}{1 + s_{2i}}\right)^{2} c_{1i}^{2} \sum_{j=1, j \neq i}^{M} \left(\frac{1}{s_{1j} - s_{1i}}\right)$$
(30)

by substituting the unknown frequencies  $\cos \overline{w}_i$  with the available parameters  $s_{1i}$  in (9), i.e,  $\cos \hat{w}_i \cong -s_{1i}$ . The bias estimate will be more accurate when the frequency separation increases. To account for this fact we introduced a constant K in Eq. (30). Computer simulations verify that, for high SNR,  $K \cong 1$ . To illustrate the performance of the bias estimation, Figure 4 shows the actual bias of an eighth order FANF with uniformly chosen frequencies and the bias estimated using (30). We see that Eq. (30) can accurately estimate the bias.

# B. Bias versus SNR

The bias in the SFANF algorithm is reduced as  $s_{2i} \rightarrow 1$ , and its mean value is not related to the input SNR (see Section III.B). This is not the case with the algorithms in [12], [13]. To illustrate the performance in terms of bias, the solutions of the ODE associated to the SFANF algorithm and to that of the algorithm in [12] were evaluated, and the bias was calculated for different SNRs. We consider a fourth order realization and an eighth order realization.

As in the previous example, the bias of the SFANF algorithm can be obtained from the stationary points of (28), since they coincide with those of (8). The bias related to the algorithm of [12] can calculated from

$$Bias = \frac{1}{2} \left[ 1 + s_{2i} \right] \left[ 2s_{1i}r_i(0) + r_i(1) \right]$$

where

$$r_{i}(m) = \left\langle z^{m-1} \left| \frac{\prod_{k=1, k \neq i}^{M} H_{k}(z)}{D_{i}(z)} \right|^{2}, S_{\nu}(z) \right\rangle, \quad i = 1, \cdots, M$$

and  $S_{\nu}(z)$  is white noise with unit variance ( $\sigma_{\nu}^2 = 1$ ). Figures 5 and 6 show the results for the fourth order and eighth order cases, respectively. We see that the solutions obtained for the algorithm of [12] depend on the SNR. The SFANF algorithm has a lower bias at low SNRs. These results show, at least from the point of view of the consistency of the estimates, that in low SNR contexts the new algorithm should have a competitive (if not better) performance when compared with that of the cascade realization in [12].

# C. Adaptive implementations: convergence and tracking behavior

In this example, the proposed SFANF and RFANF algorithms are used for estimating four sinusoids buried in white measurement noise. We consider the cases of high and low SNRs, and stationary and non-stationary environments. For the high SNR case, convergence speed and bias properties of the algorithms are evaluated for a stationary environment. The low SNR case consider the tracking of time-varying frequencies. The results are compared with those obtained with the stochastic gradient version of the algorithm in [20]. In our simulations, both algorithms employ the exponential profile in Equation (3) for the  $s_{2i}$  parameters (for all sections). The parameters used with Equation (3) and additional algorithm constants are given in Table III. The parameters were chosen to optimize convergence speed. The notch frequencies were initialized to zero, which corresponds to the worst situation for both the cascade and the FANF realizations.

For the stationary case, the input signal is modeled by four sinusoids with normalized notch frequencies  $w_{o1} = 0.1$ ,  $w_{o2} = 0.2, w_{o3} = 0.3$  and  $w_{o4} = 0.4$ . The SNR was set to 80 dB. The learning curves obtained by averaging 100 independent runs are shown in Figure 7. The figure clearly shows the improved convergence speed obtained by the FANF algorithms. As previously discussed, the increased convergence speed is achieved at the expense of a large bias in the estimated frequencies. Figure 8 plots the normalized bias in the estimated frequencies (expressed in dB) versus the number of iterations for different values of  $s_{2i}^{\infty}$  parameters, i.e., 0.95, 0.98 and 0.99. The corresponding MSE learning curves are plotted in Figure 9. A similar behavior in terms of MSE learning curves is verified for the algorithm of [20], except that for each case convergence is slower. Despite that lower bias is obtained using larger  $s_{2i}^{\infty}$ , as can be observed in this figure, the values are for most cases acceptable. After convergence has taken place, the bias can be reduced by 15 dB using (30).

A more difficult scenario for frequency estimation is to consider low SNR. Two time-varying frequency cases with the example of four sinusoids are now considered: 1) instantaneous jumps and; 2) linear variations in the sinusoid frequencies (chirps). The SNR was set to 0 dB.

In the first case, instantaneous jumps are introduced at frequencies  $w_{o2}$  and  $w_{04}$ . The objective is to study the recovering capabilities of the algorithms when subject to these abrupt changes. The step size  $\mu$  was chosen equal to 0.02. The remaining constants are chosen as before. Figures 10– 12 show the results obtained with the SFANF algorithm, the RFANF algorithm, and the algorithm of [20], respectively. The fast convergence of the SFANF and RFANF algorithms is maintained even when operating in this difficult environment. In addition, the SFANF and RFANF are seen to track abrupt frequency changes better than the algorithm of [20].

Finally, to illustrate the capability to follow slow changes in the estimated parameters we consider the case when each sinusoid frequency vary linearly with time.

The constants for each algorithm and the SNR scenario are the same as in previous example. Figures 13 and 14 show the results obtained with the SFANF algorithm and the algorithm of [20], respectively. We observe that both algorithms are able to follow the frequency variation having a small tracking error. Also, the recovering capability for the SFANF algorithm is better than for that of the algorithm of [20]. The performance of the RFANF algorithm is for this example almost identical (plots not shown here).

#### V. CONCLUSIONS

This paper presented a new family of IIR adaptive notch filters (ANF). A cascade of second-order all-pass sections is used to form multiple notches. A characterization of the main properties of the new realization is detailed. One of the advantages of the new ANF approach is that it guarantees unit maxima between notches. Therefore, we obtain a better isolation of notch frequencies when compared with other approaches available in literature, which is an important property at low SNR. Furthermore, the new realization can be implemented using orthogonal functions that have good properties in the presence of measurement noise and enable efficient adaptive implementations. Two low-complexity adaptive algorithms were addressed. A study of their stationary points and a local convergence analysis was also provided. The new ANF realization introduces a small bias that can be made arbitrarily small or be accurately estimated and removed after algorithm convergence. Simulations carried out for both stationary and nonstationary scenarios indicate that the proposed ANF can provide favorable results in terms of convergence speed, tracking capability, and notch separation. Further research is being performed to analyze the tracking performance of the adaptive implementations.

#### REFERENCES

- B. Widrow *et. al*, "Adaptive noise canceling: principles and applications, *Proc. IEEE*, vol. 63, pp. 1692-1716, Dec. 1975.
- [2] J. W. Choi, N. I. Cho, "Suppression of narrow-band interference in DS-spread spectrum systems using adaptive IIR notch filter," *Signal Process.*, vol. 82, pp. 2003-2013, 2002.
- [3] J. Bakkoury, D. Roviras, M. Ghogho, F. Castanie, "Adaptive MLSE receiver over rapidly fading channels," *Signal Process.*, vol. 80, pp. 1347-1360, 2000.
- [4] M. K. Tsatsanis, G. B. Giannakis, "Modeling and equalization of fast fading channels," *Int. J. Adaptive Contr. Signal Proc.*, vol.10, pp. 159-176, 1996.
- [5] A. Nehorai, "A minimal parameter adaptive notch filter with constrained poles and zeros," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 43, pp. 983-996, Aug. 1985.
- [6] 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, Feb. 1994.
- [7] T. S. Ng, "Some aspects of an adaptive digital notch filter with constrained poles and zeros," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 158-161, Feb. 1987.
- [8] P. Stoica and A. Nehorai, "Performance analysis of an adaptive notch filter with constrained poles and zeros," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, pp. 911-919, June 1988.
- [9] P. Tichavsky and P. Händel, "Two algorithms for adaptive retrieval of slowly time-varying multiple cisoids in noise," *IEEE Trans. Signal Process.*, vol. 43, pp. 1116-1127, May 1995.
- [10] P. Tichavsky and A. Nehorai, "Comparative study of four adaptive frequency trackers," *IEEE Trans. Signal Process.*, vol. 45, pp. 1473-1484, June 1997.
- [11] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA., MIT Press, 1983.

#### TABLE I

# RECURSIVE PREDICTION ERROR BASED FACTORIZED ALL-PASS ADAPTIVE NOTCH FILTERING ALGORITHM (RFANF).

# Definitions:

| u(n), y(n), ANF input and output signals   |  |  |  |  |
|--|--|--|--|--|
| $H_A(q)$ , following Eq. (7). $F_i(q)$ , $1 \le i \le M$ , following Eq. (13)                          |  |  |  |  |
| Parameters:  |  |  |  |  |
| M, number of frequencies to estimate   |  |  |  |  |
| $\lambda$ , forgetting factor (0 < $\lambda$ < 1)  |  |  |  |  |
| $\mu_{rpe}$ , step size ( $\cong 1 - \lambda$ )  |  |  |  |  |
| $s_{2i}^{\infty}$ , asymptotic value of $s_{2i}(n)$ (typically 0.98)                                   |  |  |  |  |
| $\rho_i$ , exponential decay time constant (typically 0.9)   |  |  |  |  |
| Initialization (for $1 \le i \le M$ )  |  |  |  |  |
| $\theta_1^i(0) = 0, \ \Psi_{\theta_1^i}(0) = 0, \ r_{rpe}^i(0) = 0, \ s_{2i}(0) = 0$                   |  |  |  |  |
| For each $n = 1, 2, \cdots$  |  |  |  |  |
| $y(n) = H_A(q)u(n)$  |  |  |  |  |
| For $1 \leq i \leq M$ ,  |  |  |  |  |
| $	heta_{1}^{i}(n+1) = 	heta_{1}^{i}(n) - rac{\mu_{rpe}}{r_{rpe}^{i}(n)} \Psi_{	heta_{1}^{i}}(n) y(n)$ |  |  |  |  |
| $\Psi_{\theta_1^i}(n) = -(1+s_{2i})c_{1i}F_i(q)u(n-1)$   |  |  |  |  |
| $r_{rpe}^{i}(n+1) = (1-\lambda)r_{rpe}^{i}(n) + \mu_{rpe} \Psi_{\theta_{1}^{i}}(n) ^{2}$               |  |  |  |  |
| $s_{2i}(n+1) = \rho_i s_{2i}(n) + (1-\rho_i) s_{2i}^{\infty}$  |  |  |  |  |

- [12] N. I. Cho and S. U. Lee, "On the adaptive lattice notch filter for the detection of sinusoids," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 40, pp. 405-416, July 1993.
- [13] T. Kwan and K. Martin, "Adaptive detection and enhancement of multiple sinusoids using a cascade IIR filter," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 937-947, July 1989.
- [14] Y. Liu, P. S. R. Diniz and T. I. Laakso, "Adaptive Steiglitz-McBride notch filter design for radio interference suppression in VDSL systems," *Proc. IEEE Globecom*, San Antonio, USA, 2001.
- [15] P. A. Regalia, "An improved lattice-based adaptive IIR notch filter," *IEEE Trans. Signal Process.*, vol. 39, pp. 2124-2128, Sept. 1991.
- [16] P. A. Regalia, Adaptive IIR filtering in signal processing and control, Marcel Dekker, NY, 1995.
- [17] S.M. Kay, "A fast and accurate single frequency estimator," *IEEE Trans.* Acoust., Speech, Signal Process., vol.37, pp. 1987-1990, Dec. 1989.
- [18] M. Padmanahhan and K. Martin, "A second-order hyperstable adaptive filter for frequency estimation," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 40, pp. 398-403, June 1993.
- [19] M. Niedzwiecki and P. Kaczmarek, "Generalized adaptive notch and comb filters for identification of quasi-periodically varying systems," *IEEE Trans. Signal Process.*, vol. 53, Dec. 2005.
- [20] N. I. Cho and S. U. Lee, "Tracking analysis of an adaptive lattice notch filter," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 42, pp.186-195, March 1995.
- [21] M. Niedzwiecki and P. Kaczmarek, "Tracking analysis of a generalized adaptive notch filter," *IEEE Trans. Signal Process.*, vol. 54, pp. 304-314, Jan. 2006.
- [22] J. E. Cousseau, P. D. Donate and Y. Liu, "Factorized allpass based adaptive notch filters," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, pp. 661-664, May 2004.
- [23] Y. V. Joshi and S. C. Dutta Roy, "Design of IIR multiple notch filters based on all-pass filters," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process*, vol. 46, pp. 134-138, Feb. 1999.
- [24] P. A. Regalia, S. K. Mitra and P. P. Vaidyanathan, "The digital all pass filter: a versatile signal processing building block," *Proc. IEEE*, vol. 76, pp. 19-37, Jan. 1988.
- [25] A. Bultheel, P. Gonzales-Vera, E. Hendriksen and O. Njastad, "A Szego theory for rational functions," *Report TW 131*, Katholieke Universiteit Leuven, May, 1990.
- [26] J. E. Cousseau, P. S. R. Diniz, G. Sentoni and O. Agamennoni, "On orthonormal realizations for adaptive IIR filters," *Int. J. Circ. Theory and Applicat.*, Wiley, vol. 28, pp. 481-500, 2000.
  [27] J. P. Scoppa, "New algorithms for IIR adaptive notch filtering," *Final*
- [27] J. P. Scoppa, "New algorithms for IIR adaptive notch filtering," *Final Project Dissertation* (in Spanish), Dept. of Electrical and Computer Engineering, Universidad Nacional del Sur, Argentina, 2002.
- [28] S.M. Kay, "Accurate frequency estimation at low signal-to-noise ratio," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, pp. 540-547, June 1984.

#### TABLE II

# STEIGLITZ-MCBRIDE BASED FACTORIZED ALL-PASS ADAPTIVE NOTCH FILTERING ALGORITHM (SFANF).

| <b>D</b> (1)   |  |  |  |  |
|--|--|--|--|--|
| Definitions:   |  |  |  |  |
| u(n), y(n), ANF input and output signals   |  |  |  |  |
| $H_A(q)$ , following Eq. (7). $F_i(q)$ , $1 \le i \le M$ , following Eq. (13)                        |  |  |  |  |
| Parameters:  |  |  |  |  |
| M, number of frequencies to estimate   |  |  |  |  |
| $\lambda$ , forgetting factor (0 < $\lambda$ < 1)  |  |  |  |  |
| $\mu_{sm}$ , step size ( $\cong 1 - \lambda$ )   |  |  |  |  |
| $s_{2i}^{\infty}$ , asymptotic value of $s_{2i}(n)$ (typically 0.98)                                 |  |  |  |  |
| $\rho_i$ , exponential decay time constant (typically 0.9)   |  |  |  |  |
| Initialization (for $1 \le i \le M$ )  |  |  |  |  |
| $\theta_1^i(0) = 0,  \Psi_{\theta_1^i}(0) = 0,  r_{rpe}^i(0) = 0,  s_{2i}(0) = 0$                    |  |  |  |  |
| For each $n = 1, 2, \cdots$  |  |  |  |  |
| $y(n) = H_A(q)u(n)$  |  |  |  |  |
| For $1 \leq i \leq M$ ,  |  |  |  |  |
| $	heta_{1}^{i}(n+1) = 	heta_{1}^{i}(n) - rac{\mu_{sm}}{r_{sm}^{i}(n)} \Psi_{	heta_{1}^{i}}(n) y(n)$ |  |  |  |  |
| $\Psi_{\theta_1^i}(n) = -\left[\frac{q^{-1}}{D_i(q)} + F_i(q)\right] u(n)$                           |  |  |  |  |
| $r_{sm}^{i}(n+1) = (1-\lambda)r_{sm}^{i}(n) + \mu_{sm} \Psi_{\theta_{1}^{i}}(n) ^{2}$                |  |  |  |  |
| $s_{2i}(n+1) = \rho_i s_{2i}(n) + (1-\rho_i) s_{2i}^{\infty}$  |  |  |  |  |

TABLE III Constants used with each algorithm.

|                   | [12]                   | RFANF                  | SFANF                  |
|-------------------|------------------------|------------------------|------------------------|
| $s_{2i}^{\infty}$ | 0.92, 0.93, 0.94, 0.95 | 0.92, 0.93, 0.94, 0.95 | 0.92, 0.93, 0.94, 0.95 |
| $\bar{\rho_i}$    | 0.99                   | 0.99                   | 0.99                   |
| $\lambda$         | 0.9                    | 0.9                    | 0.9                    |
| $\mu$             | 0.1                    | 0.1                    | 0.1                    |

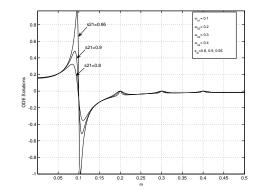


Fig. 1. ODE solutions associated to SFANF for an eighth order example.

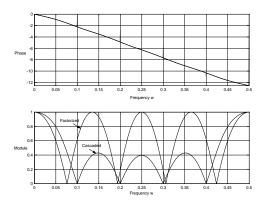


Fig. 2. Magnitude and phase of FANF and cascaded ANF realizations for an 8-order example.

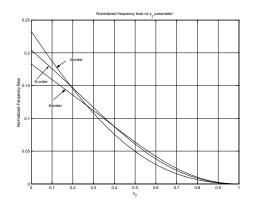


Fig. 3. Normalized bias for different FANF realization orders.

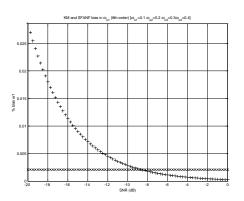


Fig. 6. Bias in  $w_{o1}$  using SFANF (x) and the algorithm of [12] (+), for a eighth order example, for different input SNR.

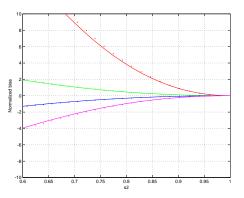


Fig. 4. True (solid) and estimated (dots) bias for an eight order FANF realization.

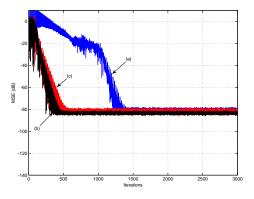


Fig. 7. MSE learning curves for SFANF algorithm (c), RFANF algorithm (b), and [12] (a) (eighth order example).

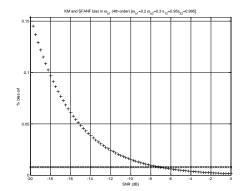


Fig. 5. Bias in  $w_{o1}$  using SFANF (*x*) and the algorithm of [12] (+), for a fouth order example, for different input SNR.

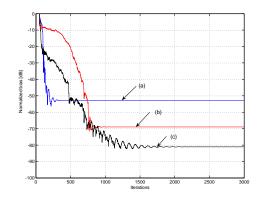


Fig. 8. Bias curves for SFANF algorithm using (a)  $s_{2i}^\infty=0.95,$  (b)  $s_{2i}^\infty=0.98$  and (c)  $s_{2i}^\infty=0.99.$ 

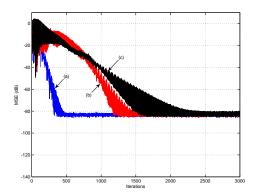


Fig. 9. MSE learning curves for SFANF algorithm using (a)  $s_{2i}^{\infty} = 0.95$ , (b)  $s_{2i}^{\infty} = 0.98$  and (c)  $s_{2i}^{\infty} = 0.99$ .

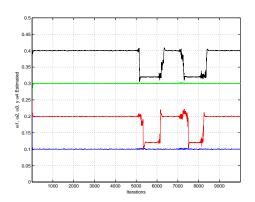


Fig. 12. Frequency estimation using the algorithm of [20] (eighth order example).

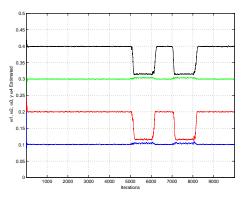


Fig. 10. Frequency estimation using the SFANF algorithm (eighth order example).

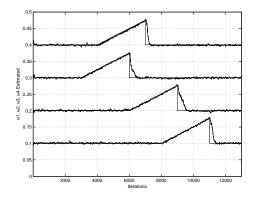


Fig. 13. Frequency estimation using SFANF algorithm in a tracking application (eighth order example).

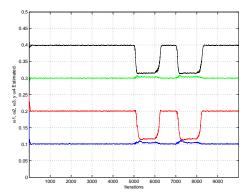


Fig. 11. Frequency estimation using the RFANF algorithm (eighth order example).

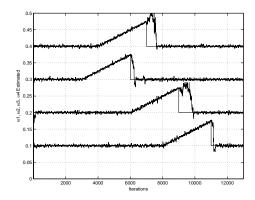


Fig. 14. Frequency estimation using the algorithm of [20] in a tracking application (eighth order example).