IEEE Copyright Notice

Generalized Adaptive Notch Smoothing Revisited

Maciej Niedźwiecki, *Member, IEEE*

***Abstract*—The problem of identification of quasi-periodically varying dynamic systems is considered. This problem can be solved using generalized adaptive notch filtering (GANF) algorithms. It is shown that the accuracy of parameter estimates can be significantly increased if the results obtained from GANF are further processed using a cascade of appropriately designed filters. The resulting generalized adaptive notch smoothing (GANS) algorithm can be employed in many off-line or nearly real-time on-line applications, such as elimination of sinusoidal interference from a prerecorded signal or identification of a rapidly varying telecommunication channel.**

***Index Terms*—Adaptive signal processing, noncausal estimation.**

I. Introduction

GENERALIZED Adaptive notch filters (GANFs) [1] were designed for the purpose of identification/tracking of quasi-periodically varying complex-valued systens, i.e., systems governed by

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

where $t = 1, 2, \ldots$ denotes the normalized discrete time, $y(t)$ denotes the system output, $\boldsymbol{\varphi}(t) = [\varphi_1(t), \ldots, \varphi_n(t)]^{\mathrm{T}}$ denotes regression vector, $v(t)$ denotes measurement noise, and $\boldsymbol{\theta}(t) = [\theta_1(t), \ldots, \theta_n(t)]^{\mathrm{T}}$ is the vector of time-varying system coefficients, modeled as weighted sums of complex exponentials

$$\boldsymbol{\theta}(t) = \sum_{i=1}^{k} \boldsymbol{\alpha}_i(t) e^{j\sum_{l=1}^{t}\omega_i(l)}. \tag{2}$$

All quantities in (1)–(2), except angular frequencies $\omega_1(t), \ldots, \omega_k(t)$, are complex-valued. Since the complex "amplitudes" $\boldsymbol{\alpha}_i(t) = [a_{i1}, \ldots, a_{in}]^{\mathrm{T}}$ incorporate both magnitude and phase information, there is no explicit phase component in (2).

An interesting application, which under certain conditions admits the formulation presented above, is adaptive equalization of rapidly fading multipath telecommunication channels—see, e.g., [2]–[4]. In this particular case, $y(t)$ is the sampled baseband signal, received by the mobile radio system, the regression vector is made up of past input (transmitted) symbols, $\boldsymbol{\theta}(t)$ is the vector of time-varying impulse response coefficients of the channel, and the angular frequencies $\omega_1, \ldots, \omega_k$ correspond to Doppler shifts along different paths of signal arrival (when the speed of the vehicle changes over time, Doppler shifts are also time-varying).

In the special case where $n = 1$ and $\varphi(t) \equiv 1$, (1)–(2) describe a mixture of complex-valued sinusoidal signals (called cisoids) buried in noise $(s(t) = \theta(t))$

$$\begin{aligned} y(t) &= s(t) + v(t) \\ s(t) &= \sum_{i=1}^{k} a_i(t) e^{j\sum_{l=1}^{t}\omega_i(l)} \end{aligned} \tag{3}$$

and GANF filters become "ordinary" adaptive notch filters (ANFs)—devices used for a variety of purposes, such as line enhancement [5], mitigation of narrowband interferences in communication channels [6], active noise and vibration control [7], [8], biomedical signal processing [9]–[11], or elimination of narrowband disturbances (generated by power lines and/or electronic circuitry) from audio signals [12]. For an overview of different ANF algorithms, see e.g., [13], [14], and references therein.

GANFs/ANFs are causal estimation algorithms, which means that at any instant t they yield estimates $\hat{\boldsymbol{\theta}}(t)$ and $\hat{s}(t)$ that are functions of current and past measurements $\mathcal{Y}_{-}(t) = \{y(1), \ldots, y(t)\}$. While causality is an obvious requirement in real-time applications, such as active noise or vibration control, many other applications exist that allow one to either partially or entirely remove this constraint. Generally speaking, such applications fall into two categories.

1) Near Real-Time Applications: Consider the problem of on-line elimination of a narrowband interference from a speech signal transmitted over a telecommunication channel. Since the received signal is already a delayed version of the transmitted one (due to channel and processing delays), one can usually afford an additional decision delay of τ sampling intervals to improve estimation accuracy, namely to evaluate smoothed (noncausal) estimates $\tilde{s}(t-\tau) = \tilde{s}[\mathcal{Y}_{-}(t)]$ that are functions of the observation history $\mathcal{Y}_{-}(t-\tau)$ available at instant $t-\tau$, and of τ "future" samples: $y(t-\tau+1), \ldots, y(t)$. In a statistical literature such solution is known as fixed-lag smoothing. When appropriately designed, smoothed estimators yield smaller estimation errors than their causal counterparts. Another group of on-line applications, allowing one to break causality limitations, are those exploiting the fact that in modern telecommunication systems, signals are transmitted in a block-by-block, rather than sample-by-sample, fashion. Once such a block (or "frame") of data $\mathcal{Y} = \{y(1), \ldots, y(N)\}$, of length N, is received, local parameter and signal estimates $\tilde{\boldsymbol{\theta}}(t) = \tilde{\boldsymbol{\theta}}[\mathcal{Y}]$, $\tilde{s}(t) = \tilde{s}[\mathcal{Y}]$, $t = 1, \ldots N$, required, e.g., for channel equalization or interference mitigation, can be obtained using procedures known as fixed-interval smoothing.

Manuscript received May 12, 2009; accepted October 20, 2009. First published November 20, 2009; current version published February 10, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Hideaki Sakai.

The author is 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).

Digital Object Identifier 10.1109/TSP.2009.2037071

2) Off-Line Applications: Elimination of narrowband disturbances from archived signals (audio, biomedical), or reconstruction of trajectories of time-varying channel parameters based on prerecorded input/output sequences (e.g., for simulation purposes), are examples of tasks that can be performed off-line, using the entire data record. Again, this can be done by employing fixed-interval smoothing.

In spite of clear advantages of smoothing, noncausal estimation techniques are surprisingly absent from the literature on adaptive notch filtering, except for a handful papers devoted to frequency smoothing [16], [17]. The concept of adaptive notch smoothing (ANS) and its system identification extension (GANS) was originally developed in our earlier work [18], and later refined in [19]. The fixed-lag smoothers proposed there were based on compensation of estimation delays that arise in the frequency tracking and amplitude tracking loops of GANF/ANF algorithms.

The contribution of the current paper is twofold. First, using the postfiltering technique (backward-time filtering of frequency and amplitude estimates), we derive fixed-interval GANS/ANS algorithms and study their estimation properties. Second, we propose new versions of fixed-lag GANS/ANS algorithms, with increased estimation capabilities. Moving from compensation of estimation delay to postfiltering has important implications. Since compensation techniques reduce only the bias component of the mean-squared estimation error (MSE), without changing its variance component, the GANS/ANS algorithms derived in [18] and [19] show advantage over their GANF/ANF counterparts mainly in the range of small values of adaptation gains (because for small gains, MSE is dominated by bias errors). In contrast with this, postfiltering allows one to reduce both bias *and* variance error components and, hence, to improve results yielded by GANF/ANF algorithms for all values of adaptation gains.

II. Problem Statement

To simplify further considerations, we will assume that the analyzed system has a single frequency mode ($k = 1$), i.e., that it is governed by

$$\begin{aligned} y(t) &= \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}(t) + v(t) \\ \boldsymbol{\theta}(t) &= \boldsymbol{\alpha}(t)e^{j\sum_{l=1}^{t}\omega(l)}. \end{aligned} \tag{4}$$

Later, we will comment on the possibility of extending the obtained results to the multiple frequencies case.

We will assume that the complex-valued vector of "amplitudes" $\boldsymbol{\alpha}(t) = [a_1(t), \ldots, a_n(t)]^{\mathrm{T}}$ and the real-valued instantaneous frequency $\omega(t) \in (-\pi, \pi]$ are slowly varying quantities (this statement will be made more precise later on), but we will not request that $\omega(t)$ should be small, which means that system parameters $\boldsymbol{\theta}(t)$ may arbitrarily fast vary with time. Additionally, we will assume that

(A1) The measurement noise $\{v(t)\}$ is a zero-mean circular white sequence with variance σ_v^2.

(A2) The sequence of regression vectors $\{\boldsymbol{\varphi}(t)\}$, independent of $\{v(t)\}$, is zero-mean, wide-sense stationary and ergodic with known covariance matrix $\boldsymbol{\Phi} = \mathrm{E}[\boldsymbol{\varphi}^*(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)]$.

The algorithm that will serve as a basis for our further considerations, referred to as a pilot GANF, is a normalized version of the GANF algorithm proposed and analyzed in [15]

$$\begin{aligned} \hat{f}(t) &= e^{j\hat{\omega}(t)}\hat{f}(t-1) \\ \varepsilon(t) &= y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{f}(t)\hat{\boldsymbol{\alpha}}(t-1) \\ \hat{\boldsymbol{\alpha}}(t) &= \hat{\boldsymbol{\alpha}}(t-1) + \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}^*(t)\hat{f}^*(t)\varepsilon(t) \\ \hat{\omega}(t+1) &= \hat{\omega}(t) - \gamma\,\mathrm{Im}\left[\frac{\varepsilon^*(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{f}(t)\hat{\boldsymbol{\alpha}}(t-1)}{\hat{\boldsymbol{\alpha}}^{\mathrm{H}}(t-1)\boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}(t-1)}\right] \\ \hat{\boldsymbol{\theta}}(t) &= \hat{\boldsymbol{\alpha}}(t)\hat{f}(t) \end{aligned} \tag{5}$$

where $*$ denotes complex conjugation and H denotes conjugate transpose. Tracking properties of this algorithm are determined by two user-dependent tuning coefficients: the adaptation gain $0 < \mu \ll 1$, which determines the rate of amplitude adaptation, and another adaptation gain $0 < \gamma \ll \mu$, which determines the rate of frequency adaptation.

Pilot GANF is a causal estimation algorithm, i.e., it yields estimates $\hat{\boldsymbol{\theta}}(t)$ that are functions of current and past measurements $\mathcal{Y}_-(t)$. Based on analysis of tracking properties of the pilot algorithm, we will design a cascade of postprocessing filters increasing accuracy of amplitude and frequency estimation. We will show that, using such multistage estimation scheme, one can significantly improve identification results.

III. Frequency Smoothing

We will show that, similar to the signal case tackled in [19], statistically efficient smoothing can be achieved by means of backward-time filtering of frequency estimates yielded by the pilot GANF algorithm.

We will start from considering a general postfiltering scheme, incorporating any linear noncausal filter. Then we will show that the best results can be obtained when the smoothing filter is anticausal and "matched" to the frequency tracking characteristic of optimally tuned GANF. We will prove, and confirm this using simulation results, that under Gaussian assumptions and for random-walk (RW) frequency variations, the resulting estimation scheme is statistically efficient, i.e., the mean-squared frequency estimation error achieves the Cramér-Rao-type lower smoothing bound. Finally, we will explain why the proposed scheme should work satisfactorily (although not optimally) for *any* slow frequency variations, not necessarily of the RW type, and for *any* adaptation gains, not necessarily optimally tuned.

A. Frequency Tracking Properties of Pilot GANF

To analyze frequency tracking properties of the pilot algorithm (5), we will assume that the signal amplitude is unknown but constant $\boldsymbol{\alpha}(t) = \boldsymbol{\alpha}$, $\forall t$, and that the instantaneous frequency $\omega(t)$ changes according to the RW model

$$\omega(t) = \omega(t-1) + w(t) \tag{6}$$

where

(A3) The sequence of one-step frequency changes, independent of $\{\boldsymbol{\varphi}(t)\}$ and $\{v(t)\}$, is a zero-mean white sequence of real-valued random variables with variance σ_w^2.

The RW model of frequency variation is often used in tracking studies as it leads to analytical results. It will allow us to reveal important features of the frequency tracking loop.

Denote by $\Delta\hat{\omega}(t) = \hat{\omega}(t) - \omega(t)$ the frequency estimation error. Using the approximating linear filter (ALF) technique,[1] one can show that [15]

$$\Delta\hat{\omega}(t) \cong \frac{F(q^{-1})-1}{1-q^{-1}}\, w(t) + (1-q^{-1})F(q^{-1})e(t) \tag{7}$$

where $e(t) = \mathrm{Im}[\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\varphi}^{*}(t)f^{*}(t)v(t)/a^2]$, $a^2 = \boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{\alpha}$, q^{-1} denotes the backward shift operator

$$F(q^{-1}) = \frac{\gamma q^{-1}}{1-(\lambda+\delta)q^{-1}+\lambda q^{-2}}$$

and $\delta = 1-\gamma$, $\lambda = 1-\mu$. Note that, similarly to $\{v(t)\}$, $\{e(t)\}$ is a zero-mean circular white sequence with variance $\sigma_e^2 = \sigma_v^2/(2a^2)$.

Due to mutual orthogonality of $\{w(t)\}$ and $\{e(t)\}$, the mean-squared frequency estimation error can be expressed in the form

$$\mathrm{E}\{|\Delta\hat{\omega}(t)|^2\} \cong \frac{1}{2\pi}\int_{-\pi}^{\pi} h\left[F(e^{-j\xi})\right] d\xi \tag{8}$$

where ξ denotes standard Fourier-domain normalized angular frequency variable

$$h[F] = (F-1)(F^{*}-1)\Delta\Delta^{*}\sigma_w^2 + \frac{FF^{*}}{\Delta\Delta^{*}}\sigma_e^2$$

and $\Delta(q^{-1}) = 1/(1-q^{-1})$.

B. Postfiltering

To obtain a smoothed estimate of $\omega(t)$, further denoted by $\tilde{\omega}(t)$, we will pass the estimates $\hat{\omega}(t)$ through a noncausal filter $G(q^{-1}) = \ldots + g_{-1}q^{-1} + g_0 + g_1 q^{1} + \ldots$

$$\tilde{\omega}(t) = G(q^{-1})\hat{\omega}(t). \tag{9}$$

The filter $G(q^{-1})$ will be designed so as to minimize the mean-squared frequency estimation error $\mathrm{E}\{[\Delta\tilde{\omega}(t)]^2\}$ where $\Delta\tilde{\omega}(t) = \tilde{\omega}(t) - \omega(t)$. Combining (7) with (9), one arrives at

$$\Delta\tilde{\omega}(t) = \frac{X(q^{-1})-1}{1-q^{-1}}\, w(t) + (1-q^{-1})X(q^{-1})e(t) \tag{10}$$

$$\mathrm{E}\{[\Delta\tilde{\omega}(t)]^2\} \cong \frac{1}{2\pi}\int_{-\pi}^{\pi} h[X(e^{-j\xi})]d\xi \tag{11}$$

where

$$X(q^{-1}) = F(q^{-1})G(q^{-1}). \tag{12}$$

[1]The ALF technique is a linearization approach introduced by Tichavský and Händel [20] for the purpose of analysis of classical adaptive notch filters. When carrying out the ALF analysis one examines dependence of estimation errors $\Delta\hat{\omega}(t)$ on $v(t)$ and $w(t)$, neglecting higher than first-order terms of all quantities listed above, including all cross-terms.

Minimization of (11) is pretty straightforward—the problem can be solved by minimizing $h[X(e^{-j\xi})]$ for every value of $\xi \in [-\pi, \pi]$. Requiring that $\partial h/\partial X^{*}|_{X=X_0} = 0$, where $X_0(q^{-1})$ is the optimal transfer function and

$$\frac{\partial}{\partial z} = \frac{1}{2}\left[\frac{\partial}{\partial \mathrm{Re}[z]} - j\frac{\partial}{\partial \mathrm{Im}[z]}\right]$$
$$\frac{\partial}{\partial z^{*}} = \frac{1}{2}\left[\frac{\partial}{\partial \mathrm{Re}[z]} + j\frac{\partial}{\partial \mathrm{Im}[z]}\right]$$

denote the so-called Wirtinger derivatives—symbolic derivatives with respect to a complex variable z, applicable to nonanalytic functions, such as $h(\,\cdot\,)$. Using Wirtinger calculus [21], one obtains

$$X_0(q^{-1}) = \frac{2\kappa_\omega}{2\kappa_\omega + (1-q^{-1})^2(1-q)^2} \tag{13}$$

where

$$\kappa_\omega = \frac{a^2\sigma_w^2}{\sigma_v^2} \tag{14}$$

is a scalar coefficient that can be regarded as a measure of system nonstationarity [15].

One can check that

$$X_0(q^{-1}) = F_0(q^{-1})F_0(q) \tag{15}$$

where $F_0(q^{-1}) = F(q^{-1}|\mu = \mu_\omega, \gamma = \gamma_\omega)$, and $\mu_\omega, \gamma_\omega$ denote optimal values of μ, γ that can be obtained by solving

$$\frac{\gamma_\omega^2}{1-\mu_\omega} = 2\kappa_\omega, \quad \frac{\mu_\omega^2}{2-\mu_\omega} = \gamma_\omega. \tag{16}$$

After combining both conditions in (16), one obtains

$$\mu_\omega^4 - 2(1-\mu_\omega)(2-\mu_\omega)^2\kappa_\omega = 0. \tag{17}$$

Note that the substitution

$$u = \frac{\mu_\omega^2}{1-\mu_\omega} \tag{18}$$

turns the fourth-order (17) into the second-order equation $u^2 - 2\kappa_\omega u - 8\kappa_\omega = 0$ which, for $u > 0$, leads to

$$u = \kappa_\omega + \sqrt{\kappa_\omega^2 + 4\kappa_\omega}. \tag{19}$$

Solving (18) for μ_ω, one obtains

$$\mu_\omega = \frac{-u + \sqrt{u^2+4u}}{2}, \quad \gamma_\omega = \frac{{\mu_\omega}^2}{2-\mu_\omega}.$$

To evaluate the steady-state mean-squared frequency estimation error (11), one can use the method of residue calculus [22]. Let

$$I[W(z)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} W(e^{-j\xi})W(e^{j\xi})d\xi = \frac{1}{2\pi j}\oint W(z^{-1})W(z)\frac{dz}{z}$$

where $W(z^{-1})$ denotes any stable proper rational transfer function and the second integral is evaluated along the unit circle in the z-plane.

For $\mu = \mu_\omega$ and $\gamma = \gamma_\omega$, (11) can be rewritten in the form

$$\mathrm{E}\{[\Delta\tilde{\omega}(t)]^2 \,|\, \mu = \mu_\omega, \gamma = \gamma_\omega\} \cong I[A^+(z^{-1})]\sigma_w^2 + I[B^+(z^{-1})]\sigma_e^2 \quad (20)$$

where A^+ and B^+ denote stable factors of $A = (X_0 - 1)(X_0^* - 1)\Delta\Delta^*$ and $B = X_0 X_0^*/\Delta\Delta^*$, respectively

$$A^+(q^{-1}) = \frac{1}{4\kappa_\omega^2}(1 - q^{-1})^3 F_0^2(q^{-1})$$
$$B^+(q^{-1}) = (1 - q^{-1})F_0^2(q^{-1}). \quad (21)$$

Based on (21), the mean-squared estimation error can be evaluated numerically using formulas given in [22].

C. Comparison With the Lower Smoothing Bound

Suppose that, in addition to (A1)–(A3), the identified signal/system obeys the following condition.

(A4) The processes $\{v(t)\}$ and $\{w(t)\}$ are Gaussian.

1) Signal Case: The signal case ($n = 1, \varphi(t) \equiv 1$) was studied in [17] and [23]. The analytical expression for the Cramér-Rao-type lower smoothing bound (LSB), derived by Tichavský and Händel [17], has the form

$$\mathrm{E}\{[\Delta\tilde{\omega}(t)]^2\} \geq \frac{4(4 - \vartheta)(2 - \vartheta)}{\vartheta(32 - 16\vartheta + \vartheta^2)}\sigma_w^2 = \mathrm{LSB}_\omega \quad (22)$$

where

$$\vartheta = -u + \sqrt{u^2 + 4u}$$

and u is given by (19). Note that for sufficiently small values of κ_ω it holds that $u \cong \sqrt{8\kappa_\omega}, \vartheta \cong 2\sqrt{u} = 2\sqrt[4]{8\kappa_\omega}$ and $\mathrm{LSB}_\omega \cong \sigma_w^2/\vartheta$.

Fig. 1 shows the plot of the relative difference between the mean-squared frequency estimation error $\mathrm{E}\{[\Delta\tilde{\omega}(t)]^2 | \mu = \mu_\omega, \gamma = \gamma_\omega\}$, evaluated according to (20), and the lower smoothing bound obtained from (22)

$$\epsilon(\kappa_\omega) = \frac{|\mathrm{E}\{[\Delta\tilde{\omega}(t)]^2 | \mu = \mu_\omega, \gamma = \gamma_\omega\} - \mathrm{LSB}_\omega|}{\mathrm{LSB}_\omega}.$$

Although it is difficult to prove this analytically, very small relative differences (smaller than $5 \cdot 10^{-6}$ for $\kappa_\omega \in [10^{-6}, 10^{-2}]$) suggest that (20) and (22) may well be mathematically equivalent.

It is clear from the analysis carried up above that for random-walk frequency variations, and under Gaussian assumptions, the optimally tuned two-stage frequency smoothing algorithm described earlier is—in spite of its simplicity—a statistically efficient noncausal frequency estimation procedure.

2) System Case: For quasi-periodically varying systems obeying (A1)–(A4), the lower smoothing bound is identical with (22), and can be obtained by combining results presented in [15] (derivation of the posterior information matrix), [23] (derivation of the implicit LSB formula), and [17] (derivation of the explicit LSB formula). Hence, the final conclusion is identical with that reached in the signal case.

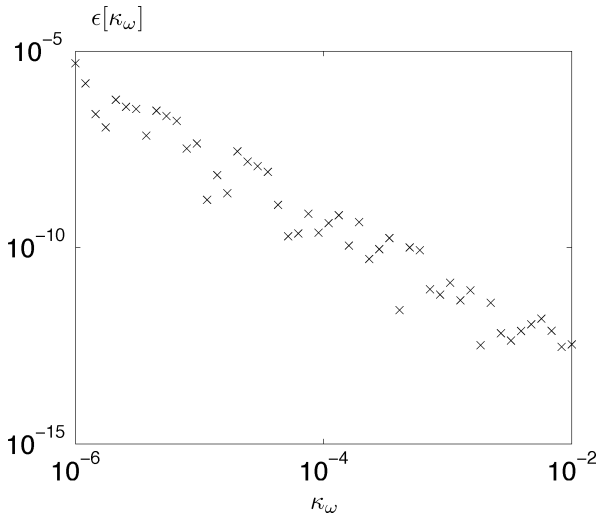

Fig. 1. Dependence of the relative approximation error on κ_ω.

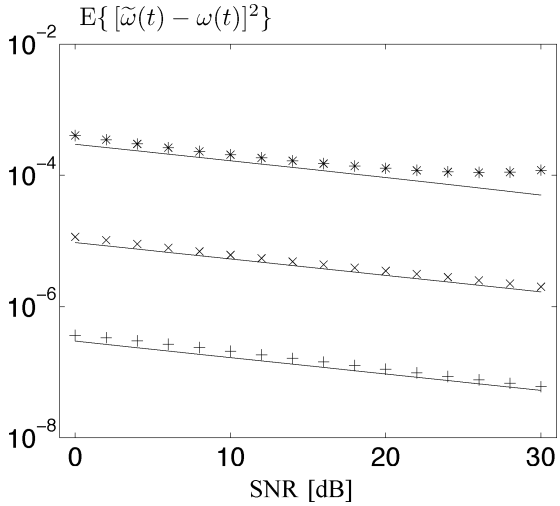

Fig. 2. Comparison of the theoretical values of the lower smoothing bound (solid lines) with experimental results obtained for the signal with randomly drifting frequency for three different speeds of frequency variation: $\sigma_w = 0.01$ $(*)$, $\sigma_w = 0.001(\times)$, $\sigma_w = 0.0001(+)$, and 16 different SNR values.

To check validity of theoretical results, the following two-tap FIR system (inspired by channel equalization applications) was simulated

$$y(t) = \theta_1(t)u(t) + \theta_2(t)u(t - 1) + v(t) \quad (23)$$

where $u(t)$ denotes a white 4-QAM input sequence ($u(t) = \pm 1 \pm j, \sigma_u^2 = 2$) and $v(t)$ denotes a complex-valued Gaussian measurement noise. The impulse response coefficients of this system were modeled as nonstationary cisoids $\theta_i(t) = a_i e^{j\sum_{\tau=1}^{t}\omega(\tau)}, i = 1, 2$, with time-invariant complex "amplitudes" $\boldsymbol{\alpha} = [a_1, a_2]^{\mathrm{T}} = [2 - j, 1 + 2j]^{\mathrm{T}}$. Note that in this case $\boldsymbol{\varphi}(t) = [u(t), u(t - 1)]^{\mathrm{T}}$ and $\boldsymbol{\Phi} = \mathbf{I}_2\sigma_u^2$.

Fig. 2 shows a comparison of the theoretical values of the lower smoothing bound with experimental results obtained for the analyzed system, for three different speeds of frequency variation σ_w (0.01, 0.001, and 0.0001) and 16 different SNR values, ranging from 0 to 30 dB. The mean-squared frequency

estimation errors were evaluated (for the optimally tuned GANS algorithm) by means of joint time and ensemble averaging. First, for each realization of the measurement noise sequence and each realization of the frequency trajectory, the mean-squared errors were computed from 200 iterations of the GANS filter (after the algorithm has reached its steady state). The obtained results were next averaged over 50 realizations of $\{w(t)\}$, 50 realizations of $\{u(t)\}$ and 50 realizations of $\{v(t)\}(50 \times 50 \times 50)$.

Note the good agreement between the theoretical curves and the results of computer simulations. The worst fit can be observed for the fastest frequency changes ($\sigma_w = 0.01$), which is understandable since in this case the associated degrees of signal nonstationarity κ_ω range from 10^{-4} (for SNR $= 0$ dB) to 10^{-1} (for SNR $= 30$ dB), i.e., they violate the "satisfactory tracking" condition $\kappa_\omega \leq 10^{-5}$ [15].

D. Frequency Smoothing Procedure

Transfer function of the optimal smoothing filter is given by $G_0(q^{-1}) = X_0(q^{-1})/F_0(q^{-1}) = F_0(q)$. This suggests the following form of frequency smoothing

$$\tilde{\omega}(t) = F(q)\hat{\omega}(t). \tag{24}$$

Since the filter $F(q)$ is anticausal, the smoothed estimate $\tilde{\omega}(t)$ can be obtained by means of backward-time filtering of the estimates yielded by the pilot algorithm:

$$\tilde{\omega}(t) = (\lambda+\delta)\tilde{\omega}(t+1) - \lambda\tilde{\omega}(t+2) + \gamma\hat{\omega}(t+1), \quad t = N-1, \ldots, 1 \tag{25}$$

with initial conditions set to: $\tilde{\omega}(N+1) = \hat{\omega}(N+1), \tilde{\omega}(N) = \hat{\omega}(N)$.

The proposed smoothing formula was derived under idealized assumptions. Therefore some robustness analysis is needed to confirm its usefulness under more realistic conditions, e.g., for frequency changes that are not governed by the RW model and/or in the presence of amplitude variations.

First, it should be noticed that the relationship (7), which is the cornerstone of the smoothing procedure, remains valid even if the sequence of one-step frequency changes $\omega(t)-\omega(t-1) = w(t)$ is not a white noise process, i.e., it holds for *arbitrary* slow frequency variations.

Second, and equally importantly, careful examination of the derivation presented in [15] shows that the relationship (7) remains approximately valid even if the vector of amplitudes $\boldsymbol{\alpha}(t)$ is not constant, but slowly varies with time—the only thing that should be changed in this more general case is the definition of $e(t)$: $e(t) = \mathrm{Im}[\boldsymbol{\alpha}^{\mathrm{H}}(t)\boldsymbol{\varphi}^*(t)f^*(t)v(t)/a^2(t)]$, $a^2(t) = \boldsymbol{\alpha}^{\mathrm{H}}(t)\boldsymbol{\Phi}\boldsymbol{\alpha}(t)$. This observation is consistent with the known fact that the results of frequency estimation of narrowband signals are usually pretty insensitive to the results of their amplitude estimation (but not *vice versa*!) [24].

For zero-mean measurement noise it holds that $\mathrm{E}[e(t)] = 0$, hence, (7) entails

$$\mathrm{E}[\hat{\omega}(t) \,|\, \omega(s), s \leq t] \cong F(q^{-1})\omega(t).$$

Since $F(q^{-1})$ is a lowpass filter with unity static gain $F(1) = 1$, when the instantaneous frequency varies slowly with time, the mean path of frequency estimates is roughly the time-delayed version of the true trajectory

$$\mathrm{E}[\hat{\omega}(t) \,|\, \omega(s), s \leq t] \cong \omega(t - t_\omega)$$

where $t_\omega = \mathrm{int}[\mu/\gamma]$ denotes nominal (low-frequency) delay introduced by the filter $F(q^{-1})$ [18]. The lag error results in estimation bias which, especially for small adaptation gains, may severely degrade cancellation/extraction efficiency of the pilot algorithm.

The situation is different when smoothing is applied. Since the nominal delay of the filter $F(q^{-1})F(q)$ is equal to zero, one arrives at the following approximate relationship:

$$\mathrm{E}[\tilde{\omega}(t)|\omega(s), -\infty < s < \infty] \cong F(q^{-1})F(q)\omega(t) \cong \omega(t)$$

which shows that smoothing, in the proposed form, reduces estimation bias *irrespective* of the shape of the estimated frequency trajectory. Additionally, by the very nature of smoothing, the variance component of the mean-squared estimation error is also reduced.

Remark: The idea of two-directional processing, which is a cornerstone of fixed-interval Kalman smoothing, inspired the authors of [16] to design a simple *ad hoc* frequency smoothing algorithm: it was proposed to process the data independently forward and backward using a tracking (ANF) algorithm and combine both estimates—for example, estimate the instantaneous frequency as an arithmetic average of the frequency estimates obtained in the forward and backward runs, respectively. We note that this simple scheme, which can be easily extended to the system identification case, has two obvious drawbacks compared to the solution based on postfiltering. First, as later remarked by the same authors [17], it does not utilize all available information efficiently—for example, under RW frequency variation, averaging the forward and backward frequency estimates reduces the variance of the frequency smoothing error, compared to the variance of the frequency tracking error, by the factor of two, whereas in the statistically efficient scheme, such as the one based on posfiltering, the variance is reduced by the factor of (approximately) four. Second, due to a very low computational cost of postfiltering (equal to 3 real multiplications per time update) the postfiltering scheme is computationally about two times less demanding than the forward-backward filtering scheme.

IV. Amplitude Smoothing

Smoothed frequency estimates can be used for the purpose of re-estimation of amplitude coefficients. The resulting frequency-guided GANF algorithm will serve as a basis for amplitude smoothing. We will show that, similar to frequency smoothing, the best results can be achieved when the amplitude estimates, yielded by frequency-guided GANF, are processed by a "matched" filter operating backward in time. Finally, we will point to certain robustness features of the proposed amplitude smoothing scheme, which make it work satisfactorily under much more realistic conditions than those initially assumed.

A. Frequency-Guided GANF and Its Amplitude Tracking Properties

The smoothed frequency estimates $\tilde{\omega}(t)$ can be used to obtain more accurate amplitude estimates. A simple way of doing this, suggested in [18], is run—in addition to (5)—the following frequency-guided GANF

$$\begin{aligned}\tilde{f}(t) &= e^{j\tilde{\omega}(t)}\tilde{f}(t-1)\\ \bar{\varepsilon}(t) &= y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\tilde{f}(t)\bar{\boldsymbol{\alpha}}(t-1)\\ \bar{\boldsymbol{\alpha}}(t) &= \bar{\boldsymbol{\alpha}}(t-1) + \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}^{*}(t)\tilde{f}^{*}(t)\bar{\varepsilon}(t)\\ \bar{\boldsymbol{\theta}}(t) &= \bar{\boldsymbol{\alpha}}(t)\tilde{f}(t).\end{aligned} \tag{26}$$

We will analyze amplitude tracking properties of this algorithm under the assumption that the time-varying instantaneous frequencies are known exactly, i.e., $\hat{\omega}(t) = \omega(t)$, $\forall t$. Even though obviously violated for the pilot algorithm (5), this assumption is approximately fulfilled [for small values of μ and γ in (5), and for sufficiently slow frequency variations] by the frequency-guided algorithm (26). Additionally, we will assume that the vector of "amplitudes" $\boldsymbol{\alpha}(t)$ evolves according to the random-walk model

$$\boldsymbol{\alpha}(t) = \boldsymbol{\alpha}(t-1) + \mathbf{n}(t) \tag{27}$$

where

(A5) The sequence of one-step amplitude changes $\{\mathbf{n}(t)\}$, independent of $\{\boldsymbol{\varphi}(t)\}$ and $\{v(t)\}$, is zero-mean circular white, and has covariance matrix $\boldsymbol{\Sigma}_n = \sigma_n^2\mathbf{I}$.

After setting $\tilde{\omega}(t) \equiv \omega(t)$ in (26), one arrives at

$$\begin{aligned}f(t) &= e^{j\omega(t)}f(t-1)\\ \bar{\varepsilon}(t) &= y(t) - \boldsymbol{\psi}^{\mathrm{T}}(t)\bar{\boldsymbol{\alpha}}(t-1)\\ \bar{\boldsymbol{\alpha}}(t) &= \bar{\boldsymbol{\alpha}}(t-1) + \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\psi}^{*}(t)\bar{\varepsilon}(t)\\ \bar{\boldsymbol{\theta}}(t) &= \bar{\boldsymbol{\alpha}}(t)f(t)\end{aligned} \tag{28}$$

where $\boldsymbol{\psi}(t) = f(t)\boldsymbol{\varphi}(t)$.

Denote by $\Delta\bar{\boldsymbol{\alpha}}(t) = \bar{\boldsymbol{\alpha}}(t) - \boldsymbol{\alpha}(t)$ and $\Delta\bar{\boldsymbol{\theta}}(t) = \bar{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}(t)$ the amplitude tracking and parameter tracking errors, respectively. Observe that $\|\Delta\bar{\boldsymbol{\theta}}(t)\| = \|\Delta\bar{\boldsymbol{\alpha}}(t)\|$. Using

$$y(t) = \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\alpha}(t) + v(t)$$

one arrives at

$$\begin{aligned}\Delta\bar{\boldsymbol{\alpha}}(t) = {}& [\mathbf{I} - \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\psi}^{*}(t)\boldsymbol{\psi}^{\mathrm{T}}(t)]\Delta\bar{\boldsymbol{\alpha}}(t-1)\\ & -[\mathbf{I} - \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\psi}^{*}(t)\boldsymbol{\psi}^{\mathrm{T}}(t)]\mathbf{n}(t) + \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\psi}^{*}(t)v(t).\end{aligned} \tag{29}$$

Since, for small values of μ and for slow amplitude changes, the estimation error $\Delta\bar{\boldsymbol{\alpha}}(t)$ varies slowly compared to the modified regression vector $\boldsymbol{\psi}(t)$, recursion (29) can be analyzed using the stochastic averaging technique. This results in the following approximation:

$$\Delta\bar{\boldsymbol{\alpha}}(t) \cong \lambda\Delta\bar{\boldsymbol{\alpha}}(t-1) - \lambda\mathbf{n}(t) + \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\psi}^{*}(t)v(t)$$

or equivalently

$$\Delta\bar{\boldsymbol{\alpha}}(t) \cong \frac{H(q^{-1}) - 1}{1 - q^{-1}}\,\mathbf{n}(t) + H(q^{-1})\boldsymbol{\zeta}(t) \tag{30}$$

where $\boldsymbol{\zeta}(t) = \boldsymbol{\Phi}^{-1}\boldsymbol{\psi}^{*}(t)v(t)$, $\boldsymbol{\Sigma}_{\zeta} = \mathrm{cov}[\boldsymbol{\zeta}(t)] = \boldsymbol{\Phi}^{-1}\sigma_v^2$ and

$$H(q^{-1}) = \mu/(1 - \lambda q^{-1}).$$

Due to mutual orthogonality of $\{\mathbf{n}(t)\}$ and $\{\boldsymbol{\zeta}(t)\}$, the mean-squared parameter estimation error can be expressed in the form

$$\mathrm{E}\{\|\Delta\bar{\boldsymbol{\alpha}}(t)\|^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} g[H(e^{-j\xi})]\,d\xi \tag{31}$$

where

$$g[H] = (H-1)(H^{*}-1)\Delta\Delta^{*}\,\mathrm{tr}[\boldsymbol{\Sigma}_n] + HH^{*}\,\mathrm{tr}[\boldsymbol{\Sigma}_{\zeta}]. \tag{32}$$

B. Postfiltering

Consider the smoothed estimate of $\boldsymbol{\alpha}(t)$

$$\tilde{\boldsymbol{\alpha}}(t) = E(q^{-1})\bar{\boldsymbol{\alpha}}(t) \tag{33}$$

where $E(q^{-1})$ is a transfer function of a linear noncausal filter. The filter $E(q^{-1})$ will be designed so as to minimize the mean-squared estimation error $\mathrm{E}\{\|\Delta\tilde{\boldsymbol{\theta}}(t)\|^2\} = \mathrm{E}\{\|\Delta\tilde{\boldsymbol{\alpha}}(t)\|^2\}$, where $\Delta\tilde{\boldsymbol{\theta}}(t) = \tilde{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}(t), \tilde{\boldsymbol{\theta}}(t) = \tilde{\boldsymbol{\alpha}}(t)f(t)$, and $\Delta\tilde{\boldsymbol{\alpha}}(t) = \tilde{\boldsymbol{\alpha}}(t) - \boldsymbol{\alpha}(t)$. Combining (30) with (33), one obtains

$$\Delta\tilde{\boldsymbol{\alpha}}(t) = \frac{Y(q^{-1}) - 1}{1 - q^{-1}}\mathbf{n}(t) + Y(q^{-1})\boldsymbol{\zeta}(t) \tag{34}$$

$$\mathrm{E}\{\|\Delta\tilde{\boldsymbol{\alpha}}(t)\|^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} g[Y(e^{-j\xi})]d\xi \tag{35}$$

where

$$Y(q^{-1}) = H(q^{-1})E(q^{-1}). \tag{36}$$

Minimization of (35) can be carried out in an analogous way as minimization of (11). Requiring that $\partial g/\partial Y^{*}|_{Y=Y_0} = 0$, where $Y_0(q^{-1})$ is the optimal transfer function, one obtains

$$Y_0(q^{-1}) = \frac{\kappa_{\alpha}}{\kappa_{\alpha} + (1 - q^{-1})(1 - q)} \tag{37}$$

where

$$\kappa_{\alpha} = \frac{\mathrm{tr}[\boldsymbol{\Sigma}_n]}{\mathrm{tr}[\boldsymbol{\Sigma}_{\zeta}]} = \frac{n\sigma_n^2}{\sigma_v^2\mathrm{tr}[\boldsymbol{\Phi}^{-1}]}. \tag{38}$$

It can be checked that

$$Y_0(q^{-1}) = H_0(q^{-1})H_0(q) \tag{39}$$

where $H_0(q^{-1}) = H(q^{-1}|\mu = \mu_{\alpha})$ and μ_{α}—the optimal value of μ—can be obtained by solving the following equation:

$$\frac{\mu_{\alpha}^2}{1 - \mu_{\alpha}} = \kappa_{\alpha}.$$

This leads to

$$\mu_\alpha = \frac{-\kappa_\alpha + \sqrt{\kappa_\alpha^2 + 4\kappa_\alpha}}{2}. \tag{40}$$

After elementary but tedious calculations, one can show that

$$\mathrm{E}\{\|\Delta\tilde{\boldsymbol{\alpha}}(t)\|^2 | \mu = \mu_\alpha\} = \frac{\mu_\alpha}{2-\mu_\alpha}\sigma_v^2 \mathrm{tr}[\boldsymbol{\Phi}^{-1}]. \tag{41}$$

C. Comparison With the Lower Smoothing Bound

Suppose that the identified signal/system obeys assumptions (A1)–(A3) and (A5).

1) Signal Case: Under Gaussian assumptions the efficient noncausal estimate of signal amplitude (given that its instantaneous frequency is known exactly) is provided by the Kalman-type Rauch-Tung-Streibel smoothing algorithm [25], which can be expressed in the form

$$\begin{aligned}
&\text{forward pass:}\\
&\hat{a}(t|t-1) = \hat{a}(t-1|t-1)\\
&p(t|t-1) = p(t-1|t-1) + \sigma_n^2\\
&k(t) = \frac{p(t|t-1)f(t)}{\sigma_v^2 + p(t|t-1)}\\
&\varepsilon(t) = y(t) - \hat{a}(t|t-1)f(t)\\
&\hat{a}(t|t) = \hat{a}(t-1|t-1) + k(t)\varepsilon(t)\\
&p(t|t) = p(t|t-1) - \frac{p^2(t|t-1)}{\sigma_v^2 + p(t|t-1)}\\
&\qquad\qquad t = 1,\ldots,N
\end{aligned} \tag{42}$$

$$\begin{aligned}
&\text{backward pass:}\\
&\hat{a}(t|N) = \hat{a}(t|t) + \frac{p(t|t)}{p(t+1|t)}\\
&\qquad\times[\hat{a}(t+1|N) - \hat{a}(t+1|t)]\\
&p(t|N) = p(t|t) + \frac{p^2(t|t)}{p^2(t+1|t)}\\
&\qquad\times[p(t+1|N) - p(t+1|t)]\\
&\qquad\qquad t = N-1,\ldots,1
\end{aligned} \tag{43}$$

where $\hat{a}(t|t-1) = \mathrm{E}[a(t)|\mathcal{Y}_-(t-1)]$, $\hat{a}(t|t) = \mathrm{E}[a(t)|\mathcal{Y}_-(t)]$, and $\hat{a}(t|N) = \mathrm{E}[a(t)|\mathcal{Y}]$ denote predictive, filtered and smoothed amplitude estimates, respectively, and $p(t|t-1), p(t|t), p(t|N)$ denote the corresponding estimation error variances.

Denote the steady-state variances by

$$\begin{aligned}
\bar{p}_\infty &= \lim_{t\to\infty} p(t|t-1),\\
p_\infty &= \lim_{t\to\infty} p(t|t)\\
\tilde{p}_\infty &= \lim_{t\to\infty} p(t|2t)
\end{aligned}$$

Then, according to (42) and (43), it holds that

$$\bar{p}_\infty = p_\infty + \sigma_n^2$$

$$\begin{aligned}
p_\infty &= \bar{p}_\infty - \frac{\bar{p}_\infty^2}{\sigma_v^2 + \bar{p}_\infty}\\
\tilde{p}_\infty &= p_\infty + \frac{p_\infty^2}{\bar{p}_\infty^2}[\tilde{p}_\infty - \bar{p}_\infty].
\end{aligned} \tag{44}$$

Solving (44) for $\tilde{p}_\infty$, one finally arrives at the following expression for the lower smoothing (posterior Cramér-Rao) bound

$$\mathrm{LSB}_a = \tilde{p}_\infty = \frac{\kappa_a}{\sqrt{\kappa_a^2 + 4\kappa_a}}\sigma_v^2. \tag{45}$$

where

$$\kappa_a = \frac{\sigma_n^2}{\sigma_v^2}.$$

Since it can be easily checked that $\mu_a/(2-\mu_a) = \kappa_a/\sqrt{\kappa_a^2 + 4\kappa_a}$, one obtains

$$\mathrm{E}\{|\Delta\tilde{a}(t)|^2 | \mu = \mu_a\} = \mathrm{LSB}_a$$

i.e., for random-walk amplitude variations, and under Gaussian assumptions, the optimally tuned two-step amplitude smoothing procedure described above is statistically efficient.

2) System Case: According to [26], when the rate of system nonstationarity κ_α is sufficiently small, the lower smoothing bound for the system governed by (27) can be expressed in the form

$$\mathrm{E}\{\|\Delta\tilde{\boldsymbol{\alpha}}(t)\|^2\} \geq \frac{1}{2}\sigma_n\sigma_v \mathrm{tr}[\boldsymbol{\Phi}^{-1/2}] = \mathrm{LSB}_\alpha. \tag{46}$$

Note that for small κ_α it holds that $\mu_\alpha \cong \sqrt{\kappa_\alpha}$, $2-\mu_\alpha \cong 2$, leading to [cf. (40)]

$$\begin{aligned}
&\mathrm{E}\{\|\Delta\tilde{\boldsymbol{\alpha}}(t)\|^2 | \mu = \mu_\alpha\}\\
&\quad\geq \frac{\sigma_v^2}{2}\sqrt{\kappa_\alpha} = \sigma_n\sigma_v\sqrt{n\mathrm{tr}[\boldsymbol{\Phi}^{-1}]}\\
&\quad= \mathrm{LSB}_\alpha.
\end{aligned}$$

Using the Cauchy-Shwartz inequality, one obtains

$$\mathrm{tr}[\boldsymbol{\Phi}^{-1/2}] \leq \sqrt{n\mathrm{tr}[\boldsymbol{\Phi}^{-1}]}$$

where equality holds iff $\boldsymbol{\Phi}$ is similar to the identity matrix: $\boldsymbol{\Phi} = \sigma_\varphi^2\mathbf{I}$. Hence, apart from the special case mentioned above, the two-step amplitude smoothing algorithm is not statistically efficient. To guarantee efficiency one would have to run a much more complicated Kalman smoothing algorithm—a system counterpart of (42)–(43).

D. Amplitude Smoothing Procedure

According to (39), when the tracking algorithm (28) is optimally tuned, i.e., $H(q^{-1}) = H_0(q^{-1})$, transfer function of the optimized smoothing filter is given by $E_0(q^{-1}) = H_0(q)$. This suggests the following form of amplitude smoothing

$$\tilde{\boldsymbol{\alpha}}(t) = H(q)\bar{\boldsymbol{\alpha}}(t). \tag{47}$$

Since the filter $H(q)$, similarly as $F(q)$, is anticausal, the smoothed estimate $\tilde{\boldsymbol{\alpha}}(t)$ can be obtained by means of backward-time filtering of the estimates yielded by the frequency-guided algorithm

$$\tilde{\boldsymbol{\alpha}}(t) = \lambda\tilde{\boldsymbol{\alpha}}(t+1) + \mu\bar{\boldsymbol{\alpha}}(t), \quad t = N-1,\ldots,1 \tag{48}$$

with $\tilde{\boldsymbol{\alpha}}(N)$ set equal to $\bar{\boldsymbol{\alpha}}(N)$.

The results of robustness analysis of the amplitude smoothing procedure resemble those obtained for frequency smoothing. Since (30) remains valid for arbitrary amplitude changes (the sequence of one-step amplitude changes $\boldsymbol{\alpha}(t)-\boldsymbol{\alpha}(t-1) = \mathbf{n}(t)$ must not be white for (30) to hold), for zero-mean measurement noise and slow amplitude variations one arrives at

$$\mathrm{E}[\bar{\boldsymbol{\alpha}}(t)|\boldsymbol{\alpha}(s), s \le t] = H(q^{-1})\boldsymbol{\alpha}(t) \cong \boldsymbol{\alpha}(t-t_\alpha)$$

where $t_\alpha = \mathrm{int}[\lambda/\mu]$ is the nominal delay introduced by the lowpass filter $H(q^{-1})$.

When smoothing is applied one obtains

$$\mathrm{E}[\tilde{\boldsymbol{\alpha}}(t)|\boldsymbol{\alpha}(s), -\infty < s < \infty] = H(q^{-1})H(q)\boldsymbol{\alpha}(t) \cong \boldsymbol{\alpha}(t)$$

which stems from the fact that the nominal delay of the filter $H(q^{-1})H(q)$ is zero and $H(1) = 1$. Therefore, whether optimal or not, for small adaptation gains the proposed smoothing procedure can significantly improve accuracy of amplitude estimation.

V. Fixed-Interval Generalized Adaptive Notch Smoother

After combining the results of frequency smoothing and amplitude smoothing, the smoothed estimate of the parameter vector $\boldsymbol{\theta}(t)$ can be obtained in the form

$$\begin{aligned}\tilde{f}(t) &= e^{j\tilde{\omega}(t)}\tilde{f}(t-1)\\ \tilde{\boldsymbol{\theta}}(t) &= \tilde{\boldsymbol{\alpha}}(t)\tilde{f}(t).\end{aligned} \tag{49}$$

The proposed adaptive smoothing algorithm is a five-step procedure, combining results yielded by the pilot GANF algorithm (5), frequency smoother (25), frequency-guided GANF algorithm (26), amplitude smoother (48), and output filter (49). For the reader's convenience, all steps are summarized in Table I.

Computational complexity of this algorithm is equal to $12n^2 + 28n + 29$ real multiplications and 1 real division per time update (1 complex multiplication is counted as an equivalent of 4 real ones).

When the covariance matrix $\boldsymbol{\Phi}$ is not known and/or it is time-varying, one can replace it with the following exponentially weighted estimate:

$$\hat{\boldsymbol{\Phi}}(t) = \lambda_o\hat{\boldsymbol{\Phi}}(t-1) + (1-\lambda_o)\boldsymbol{\varphi}^*(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)$$

where $0 < \lambda_o < 1$ denotes the forgetting constant. We note that the inverse of $\hat{\boldsymbol{\Phi}}(t)$ can also be evaluated recursively by exploiting the well-known matrix inversion lemma [21].

In the signal case, the algorithm listed above can be rewritten in an equivalent form that alleviates the need to compute the quantities $\hat{f}(t)$, $\hat{a}(t)$, $\bar{f}(t)$, $\bar{a}(t)$ and $\tilde{a}(t)$. This reduces the computational burden by approximately 40% from 44 real multiplications and 1 real division, to 28 real multiplications and 1 real division, per time update. The resulting cost-optimized ANS algorithm is summarized in Table II.

TABLE I
Fixed-Interval Generalized Adaptive Notch Smoother

$$\begin{aligned}&pilot\ filter:\\ \widehat{f}(t) &= e^{j\widehat{\omega}(t)}\widehat{f}(t-1)\\ \varepsilon(t) &= y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widehat{f}(t)\widehat{\boldsymbol{\alpha}}(t-1)\\ \widehat{\boldsymbol{\alpha}}(t) &= \widehat{\boldsymbol{\alpha}}(t-1) + \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}^*(t)\widehat{f}^*(t)\varepsilon(t)\\ \widehat{\omega}(t+1) &= \widehat{\omega}(t) - \gamma\,\mathrm{Im}\left[\frac{\varepsilon^*(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\widehat{f}(t)\widehat{\boldsymbol{\alpha}}(t-1)}{\widehat{\boldsymbol{\alpha}}^{\mathrm{H}}(t-1)\boldsymbol{\Phi}\widehat{\boldsymbol{\alpha}}(t-1)}\right]\\ t &= 1,\ldots,N\\ &frequency\ smoother:\\ \widetilde{\omega}(N+1) &= \widehat{\omega}(N+1)\\ \widetilde{\omega}(N) &= \widehat{\omega}(N)\\ \widetilde{\omega}(t) &= (\lambda+\delta)\widetilde{\omega}(t+1) - \lambda\widetilde{\omega}(t+2) + \gamma\widehat{\omega}(t+1)\\ t &= N-1,\ldots,1\\ &frequency{-}guided\ filter:\\ \widetilde{f}(t) &= e^{j\widetilde{\omega}(t)}\widetilde{f}(t-1)\\ \bar{\varepsilon}(t) &= y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widetilde{f}(t)\bar{\boldsymbol{\alpha}}(t-1)\\ \bar{\boldsymbol{\alpha}}(t) &= \bar{\boldsymbol{\alpha}}(t-1) + \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}^*(t)\widetilde{f}^*(t)\bar{\varepsilon}(t)\\ t &= 1,\ldots,N\\ &amplitude\ smoother:\\ \widetilde{\boldsymbol{\alpha}}(N) &= \bar{\boldsymbol{\alpha}}(N)\\ \widetilde{\boldsymbol{\alpha}}(t) &= \lambda\widetilde{\boldsymbol{\alpha}}(t+1) + \mu\bar{\boldsymbol{\alpha}}(t)\\ t &= N-1,\ldots,1\\ &output\ filter:\\ \widetilde{\boldsymbol{\theta}}(t) &= \widetilde{\boldsymbol{\alpha}}(t)\widetilde{f}(t)\\ t &= 1,\ldots,N\end{aligned}$$

Based on the results presented in [18], the proposed GANS/ANS algorithms can be easily extended to real-valued systems/signals, and to the multiple frequencies system and signal cases, governed by (1)–(2) and (3), respectively.

VI. Fixed-Lag Generalized Adaptive Notch Smoother

The fixed-interval GANS/ANS algorithms are suitable for off-line or block-oriented on-line applications. When the smoothed estimates must be continuously updated to maintain a fixed decision delay of τ sampling intervals, one needs a fixed-lag smoother. The fixed-lag GANS/ANS can be obtained by restricting postprocessing of GANF/ANF estimates to the recent τ time-steps only. The resulting "sawtooth" smoothing algorithms (inspired, to some extent, by the work of Johnston and Krishnamurthy on sawtooth extended Kalman

TABLE II
FIXED-INTERVAL ADAPTIVE NOTCH SMOOTHER

$$
\begin{aligned}
&pilot\ filter:\\
\varepsilon(t) &= y(t) - e^{j\widehat{\omega}(t)}\widehat{s}(t-1)\\
\widehat{s}(t) &= e^{j\widehat{\omega}(t)}\widehat{s}(t-1) + \mu\varepsilon(t)\\
\widehat{\omega}(t+1) &= \widehat{\omega}(t) - \gamma\,\mathrm{Im}\left[\frac{\varepsilon^*(t)e^{j\widehat{\omega}(t)}}{\widehat{s}^*(t-1)}\right]\\
t &= 1,\ldots,N\\
&frequency\ smoother:\\
\widetilde{\omega}(N+1) &= \widehat{\omega}(N+1)\\
\widetilde{\omega}(N) &= \widehat{\omega}(N)\\
\widetilde{\omega}(t) &= (\lambda+\delta)\widetilde{\omega}(t+1) - \lambda\widetilde{\omega}(t+2) + \gamma\widehat{\omega}(t+1)\\
t &= N-1,\ldots,1\\
&frequency{-}guided\ filter:\\
\bar{\varepsilon}(t) &= y(t) - e^{j\widetilde{\omega}(t)}\bar{s}(t-1)\\
\bar{s}(t) &= e^{j\widetilde{\omega}(t)}\bar{s}(t-1) + \mu\bar{\varepsilon}(t)\\
t &= 1,\ldots,N\\
&output\ filter:\\
\widetilde{s}(N) &= \bar{s}(N)\\
\widetilde{s}(t) &= \lambda e^{-j\widetilde{\omega}(t+1)}\widetilde{s}(t+1) + \mu\bar{s}(t)\\
t &= N-1,\ldots,1
\end{aligned}
$$

TABLE III
FIXED-LAG GENERALIZED ADAPTIVE NOTCH SMOOTHER

$$
\begin{aligned}
&pilot\ filter:\\
\widehat{f}(t) &= e^{j\widehat{\omega}(t)}\widehat{f}(t-1)\\
\varepsilon(t) &= y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\widehat{f}(t)\widehat{\boldsymbol{\alpha}}(t-1)\\
\widehat{\boldsymbol{\alpha}}(t) &= \widehat{\boldsymbol{\alpha}}(t-1) + \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}^*(t)\widehat{f}^*(t)\varepsilon(t)\\
\widehat{\omega}(t+1) &= \widehat{\omega}(t) - \gamma\,\mathrm{Im}\left[\frac{\varepsilon^*(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\widehat{f}(t)\widehat{\boldsymbol{\alpha}}(t-1)}{\widehat{\boldsymbol{\alpha}}^{\mathrm{H}}(t-1)\boldsymbol{\Phi}\widehat{\boldsymbol{\alpha}}(t-1)}\right]\\
&frequency\ smoother:\\
\widetilde{\omega}_t(t+1) &= \widehat{\omega}(t+1)\\
\widetilde{\omega}_t(t) &= \widehat{\omega}(t)\\
\widetilde{\omega}_t(i) &= (\lambda+\delta)\widetilde{\omega}_t(i+1) - \lambda\widetilde{\omega}_t(i+2) + \gamma\widehat{\omega}(i+1)\\
i &= t-1,\ldots,t-\tau\\
&frequency{-}guided\ filter:\\
\widetilde{f}_t(t-\tau-1) &= \widetilde{f}_{t-1}(t-\tau-1)\\
\widetilde{\boldsymbol{\alpha}}_t(t-\tau-1) &= \widetilde{\boldsymbol{\alpha}}_{t-1}(t-\tau-1)\\
\widetilde{f}_t(i) &= e^{j\widetilde{\omega}_t(i)}\widetilde{f}_t(i-1)\\
\bar{\varepsilon}_t(i) &= y(i) - \boldsymbol{\varphi}^{\mathrm{T}}(i)\widetilde{f}_t(i)\bar{\boldsymbol{\alpha}}_t(i-1)\\
\bar{\boldsymbol{\alpha}}_t(i) &= \bar{\boldsymbol{\alpha}}_t(i-1) + \mu\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}_t^*(i)\widetilde{f}_t^*(i)\bar{\varepsilon}_t(i)\\
i &= t-\tau,\ldots,t\\
&amplitude\ smoother:\\
\widetilde{\boldsymbol{\alpha}}_t(t) &= \bar{\boldsymbol{\alpha}}_t(t)\\
\widetilde{\boldsymbol{\alpha}}_t(i) &= \lambda\widetilde{\boldsymbol{\alpha}}_t(i+1) + \mu\bar{\boldsymbol{\alpha}}_t(i)\\
i &= t-1,\ldots,t-\tau\\
&output\ filter:\\
\widetilde{\boldsymbol{\theta}}_t(t-\tau) &= \widetilde{\boldsymbol{\alpha}}_t(t-\tau)\widetilde{f}_t(t-\tau)
\end{aligned}
$$

filters/smoothers [27]), are summarized in Tables III and IV. The smoothed estimates of $\boldsymbol{\theta}(t-\tau)$ and $s(t-\tau)$, evaluated at instant t, are denoted by $\widetilde{\boldsymbol{\theta}}_t(t-\tau)$ and $\tilde{s}_t(t-\tau)$, respectively. To avoid confusion, all other quantities updated during the post-filtering steps (frequency smoothing, amplitude re-estimation, and amplitude smoothing) were also indexed by t.

Additional computational cost of carrying out postprocessing steps, i.e., the computational overhead of smoothing, grows linearly with the lag τ, and is equal to $(4n^2+16n+9)\tau+4$ real multiplications per time update for the GANS algorithm, and 17τ real multiplications per time update for the ANS algorithm.

Remark: The estimation accuracy improvements, offered by smoothing, gradually saturate with growing τ. In the case considered, only marginal improvements can be expected when τ is increased beyond the "principal" delay $\tau_0 = \max\{\tau_\omega, \tau_\alpha\}$.

VII. COMPUTER SIMULATIONS

Our optimization study, presented in Sections III and IV, is of more theoretical than practical importance. From a practical viewpoint the most important question is whether the derived algorithms can cope favorably with "arbitrary" frequency and amplitude variations, including deterministic changes. The simulation study presented below aims to show that, even though our analysis was carried out under assumption that either the instantaneous frequency or complex amplitudes drift according to random-walk model, the proposed GANS/ANS algorithms are robust to frequency/amplitude model misspecification, and that they yield better estimation results, compared to GANF/ANF, *irrespectively* of the choice of adaptation gains. Even though it is difficult to prove this theoretically, we have rich simulation evidence suggesting that, no matter how system/signal parameters vary with time, and no matter how the coefficients μ and γ are chosen, one always gains, in terms of estimation accuracy, by switching from GANF/ANF to GANS/ANS. Additionally, and rather paradoxically, the attainable rates of improvement are usually much higher for systems/signals subject to smooth deterministic changes than for systems/signals with stochastic changes—even though the latter better fit the assumed model of time-variation.

A. ANS Algorithms

To check performance of the fixed-interval ANS algorithm, a noisy quasi-periodically varying signal (3) was generated with fast sinusoidal amplitude and frequency changes

$$a(t) = \cos(2\pi t/2000), \quad \omega(t) = \sin(2\pi t/2000). \qquad (50)$$

TABLE IV
FIXED-LAG ADAPTIVE NOTCH SMOOTHER

$$
\begin{aligned}
&\textit{pilot filter}:\\
\varepsilon(t) &= y(t) - e^{j\widehat{\omega}(t)}\widehat{s}(t-1)\\
\widehat{s}(t) &= e^{j\widehat{\omega}(t)}\widehat{s}(t-1) + \mu\varepsilon(t)\\
\widehat{\omega}(t+1) &= \widehat{\omega}(t) - \gamma\,\mathrm{Im}\left[\frac{\varepsilon^*(t)e^{j\widehat{\omega}(t)}}{\widehat{s}^*(t-1)}\right]\\
&\textit{frequency smoother}:\\
\widetilde{\omega}_t(t+1) &= \widehat{\omega}(t+1)\\
\widetilde{\omega}_t(t) &= \widehat{\omega}(t)\\
\widetilde{\omega}_t(i) &= (\lambda+\delta)\widetilde{\omega}_t(i+1) - \lambda\widetilde{\omega}_t(i+2) + \gamma\widehat{\omega}(i+1)\\
i &= t-1,\ldots,t-\tau\\
&\textit{frequency-guided filter}:\\
\bar{s}_t(t-\tau-1) &= \bar{s}_{t-1}(t-\tau-1)\\
\bar{\varepsilon}_t(i) &= y(i) - e^{j\widetilde{\omega}_t(i)}\bar{s}_t(i-1)\\
\bar{s}_t(i) &= e^{j\widetilde{\omega}_t(i)}\bar{s}_t(i-1) + \mu\bar{\varepsilon}_t(i)\\
i &= t-\tau,\ldots,t\\
&\textit{output filter}:\\
\widetilde{s}_t(t) &= \bar{s}_t(t)\\
\widetilde{s}_t(i) &= \lambda e^{-j\widetilde{\omega}_t(i+1)}\widetilde{s}_t(i+1) + \mu\bar{s}_t(i)\\
i &= t-1,\ldots,t-\tau
\end{aligned}
$$

Fig. 3 shows comparison of the steady-state mean-squared signal estimation errors, yielded by the pilot ANF algorithm and by the proposed ANS algorithm, for different values of the adaptation gain μ and for two noise intensities: $\sigma_v = 0.56$ ($\mathrm{SNR} = 5$ dB) and $\sigma_v = 0.01$ ($\mathrm{SNR} = 20$ dB). To reduce the number of design degrees of freedom the adaptation gain γ was set to $\mu^2/2$—see [15] for further explanations. All MSE values were obtained by means of joint time averaging (the evaluation interval [2001,4000] was placed inside a wider analysis interval [1,6000]) and ensemble averaging (100 realizations of measurement noise were used). As expected, the ANS algorithm yielded uniformly better results than the ANF algorithm. The peak-to-peak (or, more adequately, bottom-to-bottom) variance reduction is approximately equal to 10 dB.

Fig. 4 shows MSE plots, analogous to those presented in Fig. 3, obtained for a signal with very fast amplitude and frequency variations, governed by

$$a(t) = \cos(2\pi t/200), \quad \omega(t) = \sin(2\pi t/200). \tag{51}$$

Even though the corresponding changes—ten times faster than those considered in the previous example—can be hardly regarded as "slow," qualitative comparison of ANS versus ANF leads to identical conclusions: smoothing can considerably improve estimation results in the entire range of adaptation gains.

Fig. 5 shows true signal and its estimates obtained for a typical realization of measurement noise ($\mathrm{SNR} = 10$ dB) in the case where $\mu = 0.08$. Close-up views of these plots are shown in Fig. 6.

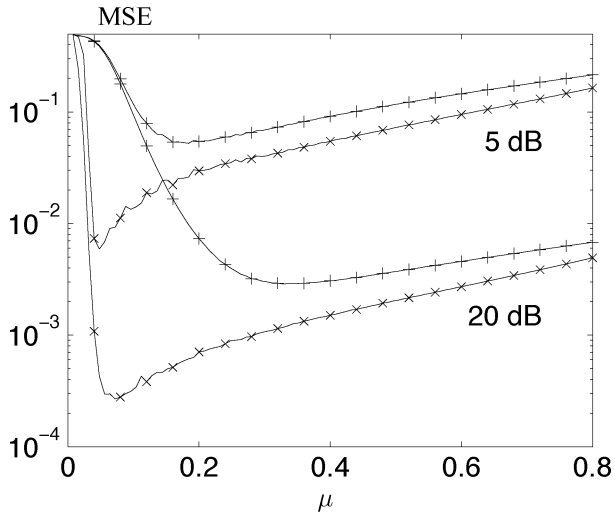

Fig. 3. Dependence of the mean-squared cancellation error on adaptation gain $\mu(\gamma = \mu^2/2)$ for the pilot estimate $\hat{s}(t)(+)$ and smoothed estimate $\bar{s}(t)(\times)$, for two signal-to-noise ratios (SNRs): 5 dB (two upper plots) and 20 dB (two lower plots). All plots (solid lines) were evaluated on a grid of 100 equidistant values of μ. The analyzed signal was subject to fast amplitude and frequency changes.

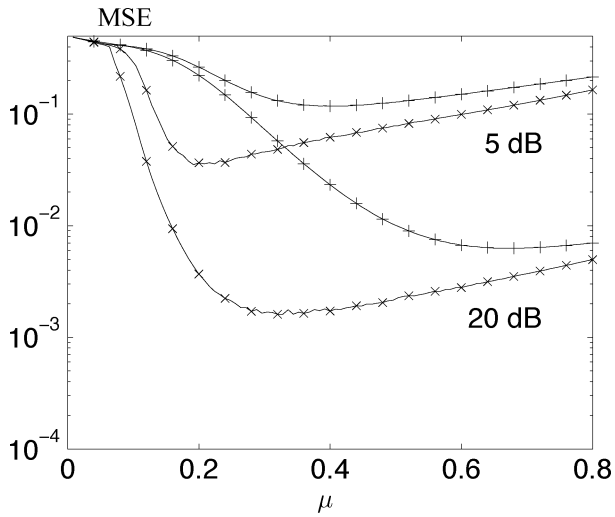

Fig. 4. Dependence of the mean-squared cancellation error on adaptation gain $\mu(\gamma = \mu^2/2)$ for the pilot estimate $\hat{s}(t)(+)$ and smoothed estimate $\bar{s}(t)(\times)$, for two SNRs: 5 dB (two upper plots) and 20 dB (two lower plots). All plots (solid lines) were evaluated on a grid of 100 equidistant values of μ. The analyzed signal was subject to very fast amplitude and frequency changes.

Finally, Fig. 7 shows how efficacy of a fixed-delay ANS, compared to ANF, depends on decision delay τ. The presented results were obtained by combined time and ensemble averaging (in the same way as aforedescribed), for the signal governed by (50), $\mathrm{SNR} = 10$ dB, and $\mu = 0.08$. Reduction of MSE attained for $\tau = 5, 10, 20, 40$, and 100, is equal to 2.8, 6.7, 12.7, 18.7, and 25, respectively. Note that when τ is increased beyond $\tau_0 = 40$, only a small further reduction of MSE can be achieved.

B. GANS Algorithm

The simulated plant was governed by (23). The instantaneous frequency evolved according to the random-walk model

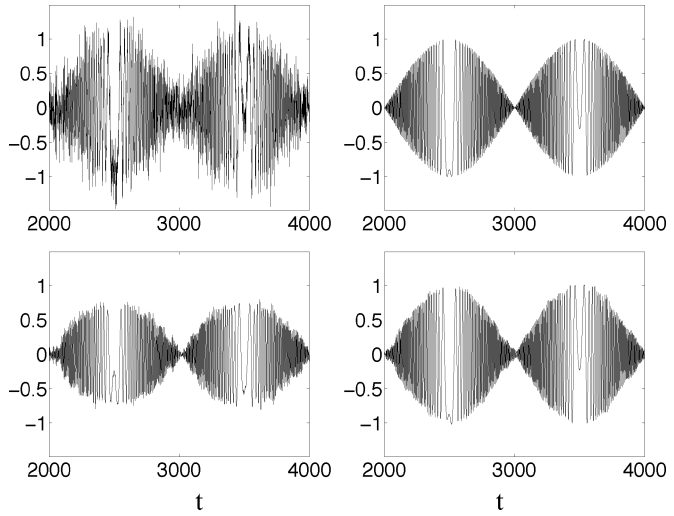

Fig. 5. Real parts of the noisy signal $y(t)$ (top left figure), noiseless signal $s(t)$ (top right figure), pilot estimate $\hat{s}(t)$ (bottom left figure), and smoothed estimate $\bar{s}(t)$ (bottom right figure).

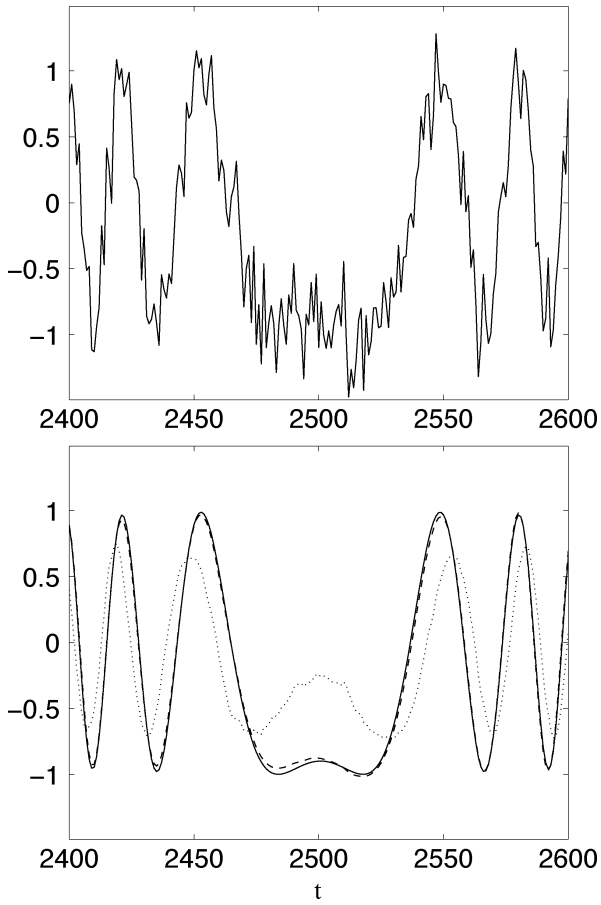

Fig. 6. Real parts of a selected fragment of the noisy signal $y(t)$ (top figure), noiseless signal $s(t)$ (bottom figure—solid line), pilot estimate $\hat{s}(t)$ (bottom figure—dotted line), and smoothed estimate $\bar{s}(t)$ (bottom figure—broken line).

$(\sigma_w = 0.001)$, and complex amplitudes were subject to sinusoidal changes

$$a_1(t) = (2 - j)\sin(2\pi t/2000)$$
$$a_2(t) = (1 + 2j)\cos(2\pi t/2000).$$

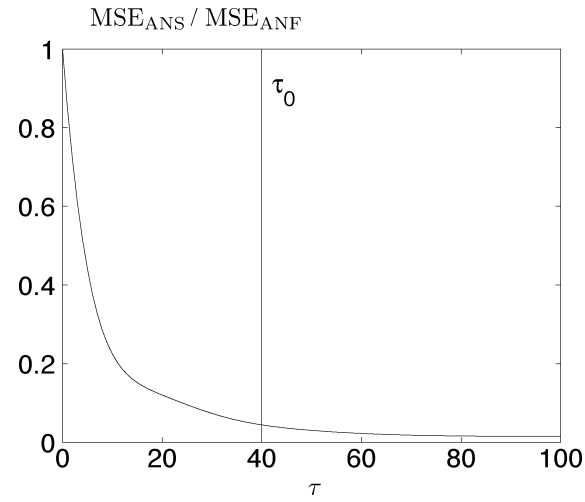

Fig. 7. ANS/ANF estimation error variance reduction ratio as a function of decision delay τ; the principal delay τ_0 is marked with a vertical line.

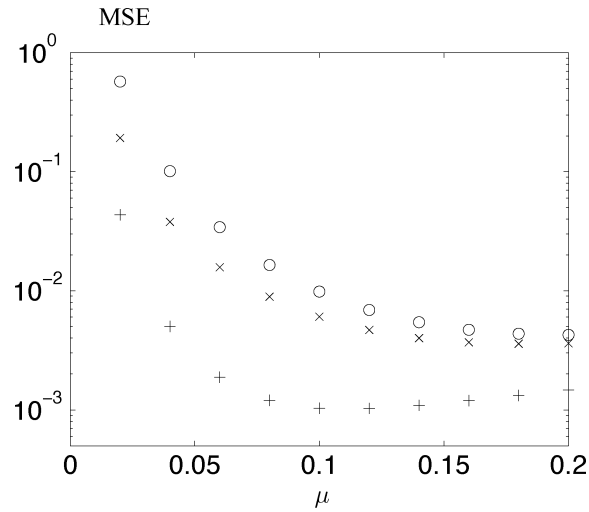

Fig. 8. Comparison of steady-state mean-squared parameter estimation errors yielded by the pilot GANF algorithm ($\bigcirc$), frequency-guided GANF algorithm ($\times$), and fixed-interval GANS algorithm ($+$), for the plant with sinusoidally varying amplitudes and randomly drifting frequency.

White 4-QAM sequence was used as an input signal $(u(t) = \pm 1 \pm j)$. Finally, the measurement noise variance σ_v^2 was set equal 0.4.

Denote by

$$\Sigma_{\hat{\theta}} = \frac{1}{2000}\sum_{t=2001}^{4000} ||\hat{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}(t)||^2$$

the time-averaged parameter estimation error computed for the pilot estimator $\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\alpha}}(t)\hat{f}(t)$, and denote by $\Sigma_{\bar{\theta}}$ and $\Sigma_{\tilde{\theta}}$ the analogous errors evaluated for the frequency-guided estimator $\bar{\boldsymbol{\theta}}(t) = \bar{\boldsymbol{\alpha}}(t)\bar{f}(t)$ and fixed-interval GANS estimator $\tilde{\boldsymbol{\theta}}(t) = \tilde{\boldsymbol{\alpha}}(t)\tilde{f}(t)$, respectively. Similarly to the signal case, to arrive at steady-state results the evaluation interval [2001,4000] was placed inside a wider analysis interval [1, 6000].

Fig. 8 shows dependence on μ of the average values of $\Sigma_{\hat{\theta}}, \Sigma_{\bar{\theta}}$ and $\Sigma_{\tilde{\theta}}$. All ensemble averages were computed for 50 realizations of $\{u(t)\}$, 50 realizations of $\{w(t)\}$, and 50

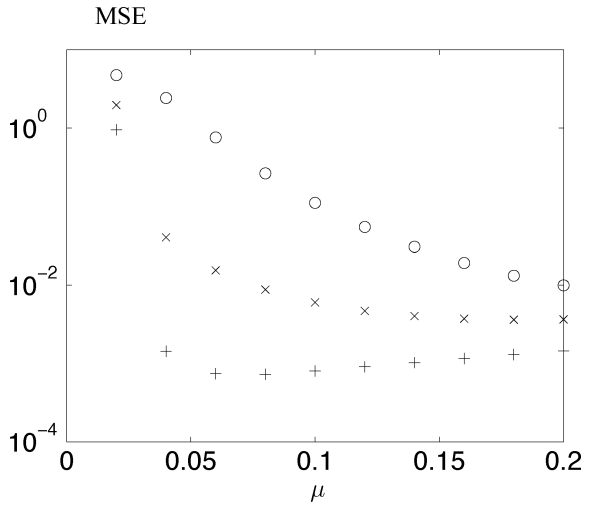

Fig. 9. Comparison of steady-state mean-squared parameter estimation errors yielded by the pilot GANF algorithm ($\bigcirc$), frequency-guided GANF algorithm ($\times$), and fixed-interval GANS algorithm ($+$), for the plant with sinusoidally varying amplitudes and linearly varying frequency.

realizations of $\{v(t)\}$ $(50 \times 50 \times 50)$. Similarly as before, γ was set to $\mu^2/2$. Note that while the frequency-guided GANF algorithm is capable of improving estimation results in the range of small values of μ only, the proposed GANS algorithm yields uniformly better results for all values of μ.

Fig. 9 shows the analogous results obtained for the system with linear frequency changes

$$\omega(t) = 0.2 + 0.008t$$

(all other details, including the model of amplitude variations, remained unchanged). Note that although there are no qualitative differences compared to the previous case, the attainable margin of improvement is, under such a fully deterministic scenario, considerably larger.

VIII. Conclusion

Identification of quasi-periodically varying dynamic systems can be carried out using GANFs. Based on analysis of parameter tracking properties of GANF algorithms, we have designed a cascade of postprocessing filters increasing accuracy of frequency and amplitude estimation. We have shown that the resulting noncausal GANS algorithms yield significantly improved estimation results compared to their causal (GANF) counterparts.

References

[1] M. Niedźwiecki and P. Kaczmarek, "Generalized adaptive notch filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Montreal, Canada, 2004, pp. II-657–II-660.

[2] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *Int. J. Adapt. Contr. Signal Process.*, vol. 10, pp. 159–176, 1996.

[3] G. B. Giannakis and C. Tepedelenlioğlu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, 1998.

[4] J. Bakkoury, D. Roviras, M. Ghogho, and F. Castanie, "Adaptive MLSE receiver over rapidly fading channels," *Signal Process.*, vol. 80, pp. 1347–1360, 2000.

[5] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[6] M. Soderstrand, T. G. Johnson, R. H. Strandberg, H. H. Loomis, Jr., and K. V. Rangarao, "Suppression of multiple narrow-band interferences using real-time adaptive notch filters," *IEEE Trans. Circuits Syst.—II*, vol. 44, pp. 217–225, 1997.

[7] C. R. Fuller and A. H. von Flotow, "Active control of sound and vibration," *IEEE Syst. Control Mag.*, vol. 15, pp. 9–19, 1995.

[8] Y. Chen, V. Wickramasinghe, and D. Zimcik, "Smart spring impedance control algorithm for helicopter blade harmonic vibration suppression," *J. Vibr. Contr.*, vol. 11, pp. 543–560, 2005.

[9] N. V. Thakor and Y. S. Zhou, "Application of adaptive filtering to ECG analysis: Noise cancellation and arrhytmia detection," *IEEE Trans. Biomed. Eng.*, vol. 38, pp. 785–794, 1991.

[10] M. Ferdjallah and R. E. Barr, "Adaptive digital notch filter design on the unit circle for the removal of powerline noise from biomedical signals," *IEEE Trans. Biomed. Eng.*, vol. 41, pp. 529–536, 1994.

[11] W. K. Ma, Y. T. Zhang, and F. S. Yang, "A fast recursive-least-squares adaptive notch filter and its applications to biomedical signals," *Med. Biol. Eng. Comput.*, vol. 37, pp. 99–103, 1999.

[12] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*. New York: Wiley, 2001.

[13] P. A. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*. New York: Marcel Dekker, 1995.

[14] P. Tichavský and A. Nehorai, "Comparative study of four adaptive frequency trackers," *IEEE Trans. Signal Process.*, vol. 45, pp. 1473–1484, 1997.

[15] M. Niedźwiecki and P. Kaczmarek, "Tracking analysis of a generalized adaptive notch filter," *IEEE Trans. Signal Process.*, vol. 54, pp. 304–314, 2006.

[16] P. Tichavský and P. Händel, "Multicomponent polynomial phase signal analysis using a tracking algorithm," *IEEE Trans. Signal Process.*, vol. 47, pp. 1390–1395, 1999.

[17] P. Tichavský and P. Händel, "Estimation and smoothing of instantaneous frequency of noisy narrowband signals," in *Proc. X Eur. Signal Process. Conf., EUSIPCO 2000*, Tampere, Finland, 2000, pp. 1803–1806.

[18] M. Niedźwiecki and A. Sobociński, "A simple way of increasing estimation accuracy of generalized adaptive notch filters," *IEEE Signal Process. Lett.*, vol. 14, pp. 217–220, 2007.

[19] M. Niedźwiecki and A. Sobociński, "Generalized adaptive notch smoothers for real-valued signals and systems," *IEEE Trans. Signal Process.*, vol. 56, pp. 125–133, 2008.

[20] P. Tichavský 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, 1995.

[21] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[22] M. Jury, *Theory and Application of the $\mathcal{Z}$-Transform Method*. New York: Wiley, 1964.

[23] M. Niedźwiecki, "Statistically efficient smoothing algorithm for time-varying frequency estimation," *IEEE Trans. Signal Process.*, vol. 56, pp. 3846–3854, Aug. 2008.

[24] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[25] J. S. Meditch, "A survey of data smoothing for linear and nonlinear dynamic systems," *Automatica*, vol. 9, pp. 151–162, 1973.

[26] M. Niedźwiecki, "On the lower smoothing bound in identification of time-varying systems," *Automatica*, vol. 44, pp. 459–464, 2008.

[27] L. A. Johnston and V. Krishnamurthy, "Derivation of a sawtooth iterated extended Kalman smoother via the AECM algorithm," *IEEE Trans. Signal Process.*, vol. 49, pp. 1899–1909, 2001.

Maciej Niedźwiecki (M'08) was born in Poznań, Poland, in 1953. He received the M.Sc. and Ph.D. degrees from the 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 (1986–1989) as a Research Fellow with the Department of Systems Engineering, Australian National University. During 1990–1993, he served as a Vice Chairman of Technical Committee on Theory of the International Federation of Automatic Control (IFAC). He is the author of the book *Identification of Time-varying Processes* (New York: Wiley, 2000). Currently, 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.

Dr. Niedźwiecki is an Associate Editor for the IEEE Transactions on Signal Processing.