# New Algorithms for Adaptive Notch Smoothing

Maciej Niedźwiecki, *Member, IEEE* and Michał Meller

*Abstract*—The problem of extraction/elimination of a nonstationary complex sinusoidal signal buried in noise is considered. This problem is usually solved using adaptive notch filtering (ANF) algorithms. It is shown that accuracy of signal estimation can be increased if the results obtained from ANF are further processed using a cascade of appropriately designed filters. The resulting adaptive notch smoothing (ANS) algorithms can be employed in many off-line or nearly real-time on-line applications. Whenever signal frequency varies in a sufficiently smooth manner, the proposed fixed-interval and fixed-lag ANS algorithms, based on a new, quasi-linear model of frequency changes, outperform the existing solutions.

*Index Terms*—adaptive notch filtering, adaptive notch smoothing, noncausal estimation

## I. Introduction

IN many applications (e.g. biomedical or telecomunication-oriented) one arises at the problem of either extraction or suppression of a complex sinusoidal signal (cisoid) buried in noise [1], [2], [3]. When frequency of the cisoid is constant and known, this problem can be solved using classical signal processing tools known as notch filters. When frequency of the extracted/eliminated signal is unknown, and possibly time-varying, one can use adaptive versions of notch filtering algorithms. The problem of adaptive notch filtering (ANF) was intensively studied in the past three decades and has resulted in a large number of ANF algorithms, differing in design principles, tracking efficiency and computational complexity. Adaptive notch filters are causal estimation schemes, which means that at each time instant they yield signal estimates that depend on past measurements only. Even though the causality constraint is inevitable when tasks must be performed in real time, quite a number of applications exist that allow one to postpone measurement-based decisions by a certain number of sampling intervals. For example, when removing hum from a signal transmitted over a telecommunication channel, a reasonably long decision delay is acceptable as it will simply add-up to the (unavoidable) transmission delay. In applications such as the one described above, usually referred to as nearly real-time, signal estimation can be based on all past and a certain number of "future" measurements, allowing one to significantly improve estimation results. Adaptive notch algorithms based on this principle are called fixed-lag adaptive notch smoothers (ANS).

There are also quite a number of off-line processing tasks, such as removal of sinusoidal interference from a prerecorded signal, where causality restrictions do not apply at all (since the entire signal is available). Such tasks can be accomplished using fixed-interval ANS algorithms.

Even though adaptive notch smoothers, whenever applicable, yield considerably better results than adaptive notch filters, they are almost absent from the signal processing literature. The problem of adaptive notch smoothing was studied, in a wider context of identification of quasi-periodically varying systems in [4], [5] and [6]. Identification of quasi-periodically varying systems can be carried out using algorithms known as generalized adaptive notch filters/smoothers (GANF/GANS). In the special, signal case, GANF/GANS algorithms reduce themselves to "ordinary" adaptive notch filters/smoothers.

For real-valued systems/signals, simple versions of GANS/ANS algorithms, obtained by means of compensating estimation delays that arise in the frequency tracking and amplitude tracking loops of GANF/ANF algorithms, was proposed in [5] (for complex-valued systems/signals a simplified version of this solution was presented earlier in [4]). The approach described in [6] is more sophisticated. Based on analysis of tracking properties of GANF/ANF designed for complex-valued systems/signals, a cascade of postprocessing filters that increase accuracy of frequency and amplitude estimation was elaborated. It was shown that in the important benchmark case, where the instantaneous frequency of the signal drifts according to the random-walk model, the proposed GANS/ANS algorithms are, under Gaussian assumptions, statistically efficient frequency smoothers, i.e., they reach the corresponding Cramér-Rao-type lower smoothing bounds. They are robust to frequency/amplitude model misspecification, and they yield better estimation results than GANF/ANF algorithms irrespective of the choice of adaptation gains.

In many practical applications narrowband signals (sound, vibration) originating from rotating machinery, such as engine or propeller, have quasi-linearly modulated frequency. Whenever known in advance, this fact can be taken advantage of, as it allows one to design adaptive notch filters with improved tracking capability. In this paper we will propose such an improved adaptive notch tracker, prove its statistical efficiency (under Gaussian assumptions), and show how it can be turned into a statistically efficient adaptive notch smoother. For signals with smooth frequency trajectories the new schemes yield better results than those described in [4]–[6].

2

## II. Problem statement

Consider the problem of extraction or elimination of a nonstationary cisoid $s(t)$ from noisy measurements $y(t)$

$$s(t) = a(t)f(t), \quad f(t) = e^{j \sum_{l=1}^{t} \omega(l)}$$
$$y(t) = s(t) + v(t) \tag{1}$$

where $t = \ldots, -1, 0, 1, \ldots$ denotes the normalized discrete time. We will assume that both the complex-valued amplitude $a(t)$ (incorporating the initial phase shift) and the real-valued instantaneous angular frequency $\omega(t) \in (-\pi, \pi]$ are slowly-varying quantities, and that

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

### A. Quasi-linear Frequency Changes

Consider frequency variations governed by the following model

$$\omega(t) = \omega(t-1) + \alpha(t-1)$$
$$\alpha(t) = \alpha(t-1) + w(t) \tag{2}$$

where $\alpha(t)$ denotes frequency rate and $w(t)$ – the one-step frequency rate change. According to (2) it holds that

$$(1 - q^{-1})^2 \omega(t) = q^{-1} w(t) \tag{3}$$

where $q^{-1}$ is the backward shift operator. Since $(1 - q^{-1})^2 \omega(t) = 0$ implies $\omega(t) = \omega_0 + \delta_\omega t$, where $\omega_0$ and $\delta_\omega$ denote arbitrary constants, (2) can be regarded as a perturbed linear growth/decay model. When

(A2) $\{w(t)\}$, independent of $\{v(t)\}$, is a zero-mean white sequence with variance $\sigma_w^2$

equation (3) defines the so-called second-order random-walk model. The corresponding frequency changes will be further referred to as *quasi-linear*.

### B. Pilot ANF and Its Tracking Properties

Consider the following ANF algorithm, which combines frequency tracking with frequency rate tracking

$$\widehat{f}(t) = e^{j[\widehat{\omega}(t-1) + \widehat{\alpha}(t-1)]} \widehat{f}(t-1)$$
$$\varepsilon(t) = y(t) - \widehat{a}(t-1)\widehat{f}(t)$$
$$\widehat{a}(t) = \widehat{a}(t-1) + \mu \widehat{f}^*(t)\varepsilon(t)$$
$$\widehat{\alpha}(t) = \widehat{\alpha}(t-1) + \gamma_\alpha \delta(t)$$
$$\widehat{\omega}(t) = \widehat{\omega}(t-1) + \widehat{\alpha}(t-1) + \gamma_\omega \delta(t)$$
$$\delta(t) = \mathrm{Im} \left[ \frac{\varepsilon(t)}{\widehat{a}(t-1)\widehat{f}(t)} \right]$$
$$\widehat{s}(t) = \widehat{a}(t)\widehat{f}(t) \tag{4}$$

where $*$ denotes complex conjugation, and $\mu > 0$, $\gamma_\omega > 0$, $\gamma_\alpha > 0$, $\gamma_\alpha \ll \gamma_\omega \ll \mu$, are small adaptation gains determining the rate of amplitude adaptation, frequency adaptation and frequency rate adaptation, respectively. We will show that this algorithm, further referred to as pilot ANF, can favorably cope with quasi-linear frequency changes. Note that for $\gamma_\alpha = 0$ and

under zero initial conditions $\widehat{\alpha}(0) = 0$, (4) reduces down to the algorithm studied in [6] – the equivalence holds for $\gamma = \gamma_\omega$, where $\gamma$ is the adaptation gain used in [6].

Tracking properties of the pilot ANF will be analyzed for the constant-modulus signal $s(t)$:

(A3) $|s(t)| = |a(t)| = a_0$, $s(t) = e^{j\omega(t)} s(t-1)$, $\forall t$.

Using the approximating linear filter (ALF) technique – the stochastic linearization approach proposed in [7] – one can show that (see Appendix I) the frequency and frequency rate estimation errors

$$\Delta \widehat{\omega}(t) = \omega(t) - \widehat{\omega}(t), \quad \Delta \widehat{\alpha}(t) = \alpha(t) - \widehat{\alpha}(t)$$

can be approximately expressed in the form

$$\Delta \widehat{\omega}(t) = H_1(q^{-1})e(t) + H_2(q^{-1})w(t) \tag{5}$$

$$\Delta \widehat{\alpha}(t) = I_1(q^{-1})e(t) + I_2(q^{-1})w(t) \tag{6}$$

where $\{e(t)\}$, $e(t) = -\mathrm{Im}\{v(t)s^*(t)/a_0^2\}$, is a zero-mean white noise, independent of $\{w(t)\}$, with variance $\sigma_e^2 = \sigma_v^2/(2a_0^2)$,

$$H_1(q^{-1}) = (1 - q^{-1})[\gamma_\omega + (\gamma_\alpha - \gamma_\omega)q^{-1}]/D(q^{-1})$$
$$H_2(q^{-1}) = q^{-1}[1 - \gamma_\omega - (1 - \mu)q^{-1}]/D(q^{-1})$$
$$I_1(q^{-1}) = \gamma_\alpha(1 - q^{-1})^2/D(q^{-1})$$
$$I_2(q^{-1}) = [1 + (\mu + \gamma_\omega - 2)q^{-1} + (1 - \mu)q^{-2}]/D(q^{-1})$$

and

$$D(q^{-1}) = 1 + d_1 q^{-1} + d_2 q^{-2} + d_3 q^{-3}$$

$$d_1 = \mu + \gamma_\omega + \gamma_\alpha - 3$$
$$d_2 = 3 - 2\mu - \gamma_\omega$$
$$d_3 = \mu - 1.$$

All filters are asymptotically stable if adaptation gains fulfill the following (sufficient) stability conditions: $0 < \mu < 1$, $0 < \gamma_\omega < 1$, $0 < \gamma_\alpha < 1$ and $\mu(\gamma_\omega + \gamma_\alpha) > \gamma_\alpha$.

It is worth noticing that ALF approximations remain valid for any uniformly small sequences $\{v(t)\}$ and $\{w(t)\}$, i.e., they are not restricted to sequences obeying assumptions (A1) and (A2). Additionally, the functional form of ALF equations does not change when signal amplitude is also slowly varying with time, i.e., when assumption (A3) does not hold true. These facts have important implications when it comes to robustness analysis of ANF/ANS algorithms.

We will show that under Gaussian assumptions:

(A4) The sequences $\{v(t)\}$ and $\{w(t)\}$ are normally distributed

the optimally tuned algorithm (4) is a statistically efficient frequency and frequency rate tracker, i.e., it reaches the corresponding lower tracking bounds. First, note that – due to orthogonality of $e(t)$ and $w(t)$ – the mean-squared tracking errors can be expressed in the form

$$\mathrm{E}\{[\Delta \widehat{\omega}(t)]^2\} = J[H_1(z^{-1})]\mathrm{E}[e^2(t)] + J[H_2(z^{-1})]\mathrm{E}[w^2(t)]$$
$$= \sigma_w^2 F_\omega(\mu, \gamma_\omega, \gamma_\alpha; \kappa) \tag{7}$$

TABLE I
OPTIMAL ANF SETTINGS AND THE CORRESPONDING NORMALIZED
LOWER TRACKING BOUNDS.

| $\kappa$ | $\mu^{\mathrm{opt}}$ | $\gamma_\omega^{\mathrm{opt}}$ | $\gamma_\alpha^{\mathrm{opt}}$ | $\mathrm{LTB}_\omega/\sigma_w^2$ | $\mathrm{LTB}_\alpha/\sigma_w^2$ |
|---|---|---|---|---|---|
| $10^{-10}$ | 0.0472 | 0.00113 | 0.0000138 | $2.05 \cdot 10^5$ | $8.21 \cdot 10^1$ |
| $5 \cdot 10^{-10}$ | 0.0613 | 0.00192 | 0.0000306 | $9.09 \cdot 10^4$ | $6.28 \cdot 10^1$ |
| $10^{-9}$ | 0.0685 | 0.00241 | 0.0000432 | $6.39 \cdot 10^4$ | $5.58 \cdot 10^1$ |
| $5 \cdot 10^{-9}$ | 0.0886 | 0.00407 | 0.0000955 | $2.82 \cdot 10^4$ | $4.26 \cdot 10^1$ |
| $10^{-8}$ | 0.0990 | 0.00509 | 0.000134 | $1.97 \cdot 10^4$ | $3.79 \cdot 10^1$ |
| $5 \cdot 10^{-8}$ | 0.127 | 0.00852 | 0.000295 | $8.66 \cdot 10^3$ | $2.89 \cdot 10^1$ |
| $10^{-7}$ | 0.142 | 0.0106 | 0.000414 | $6.06 \cdot 10^3$ | $2.57 \cdot 10^1$ |
| $5 \cdot 10^{-7}$ | 0.181 | 0.0177 | 0.000905 | $2.63 \cdot 10^3$ | $1.95 \cdot 10^1$ |
| $10^{-6}$ | 0.201 | 0.0219 | 0.00126 | $1.83 \cdot 10^3$ | $1.73 \cdot 10^1$ |
| $5 \cdot 10^{-6}$ | 0.254 | 0.0359 | 0.00273 | $7.81 \cdot 10^2$ | $1.32 \cdot 10^1$ |
| $10^{-5}$ | 0.281 | 0.0443 | 0.00379 | $5.39 \cdot 10^2$ | $1.17 \cdot 10^1$ |
| $5 \cdot 10^{-5}$ | 0.350 | 0.0712 | 0.00806 | $2.25 \cdot 10^2$ | 8.84 |
| $10^{-4}$ | 0.384 | 0.0869 | 0.0111 | $1.54 \cdot 10^2$ | 7.83 |

$$\mathrm{E}\{[\Delta\widehat{\alpha}(t)]^2\} = J[I_1(z^{-1})]\mathrm{E}[e^2(t)] + J[I_2(z^{-1})]\mathrm{E}[w^2(t)]$$
$$= \sigma_w^2 F_\alpha(\mu, \gamma_\omega, \gamma_\alpha; \kappa) \qquad (8)$$

where

$$J[X(z^{-1})] = \frac{1}{2\pi j} \oint X(z)X(z^{-1})\frac{dz}{z}$$

is the integral evaluated along the unit circle [$X(z^{-1})$ is assumed to be a stable proper rational transfer function] and

$$\kappa = \frac{\mathrm{E}[w^2(t)]}{2\mathrm{E}[e^2(t)]} = \frac{a_0^2\sigma_w^2}{\sigma_v^2} = \mathrm{SNR} \cdot \sigma_w^2 \qquad (9)$$

denotes the rate of nonstationarity of the analyzed signal [7] (SNR=$a_0^2/\sigma_v^2$ stands for the signal-to-noise ratio).

Optimal settings $\mu^{\mathrm{opt}}$, $\gamma_\omega^{\mathrm{opt}}$ and $\gamma_\alpha^{\mathrm{opt}}$ should be chosen so as to minimize (7) – to achieve the best frequency tracking, or to minimize (8) – to achieve the best frequency rate tracking. Note that in both cases the optimal values of $\mu$, $\gamma_\omega$ and $\gamma_\alpha$ depend exclusively on the rate of signal nonstationarity $\kappa$. Even though the analytical expressions for $F_\omega(\cdot)$ and $F_\alpha(\cdot)$ can be easily derived using residue calculus (the corresponding formulas can be found e.g. in [8]), they are too lengthy and too complicated to enable optimization in a "symbolic" form. For this reason they are not presented here. For a given value of $\kappa$, the optimal parameter values can be found using numerical search – the results are shown in Table 1, along with the corresponding values of lower tracking bounds $\mathrm{LTB}_\omega$ and $\mathrm{LTB}_\alpha$, derived in Appendix II. Since it was found that the functions $F_\omega(\cdot)$ and $F_\alpha(\cdot)$ attain their minima for the same values of $\mu$, $\gamma_\omega$ and $\gamma_\alpha$, only one set of optimal gains was listed for each value of $\kappa$.

There is a perfect agreement between the lower tracking bounds and the values obtained by minimizing (7) and (8), respectively – in some cases the computed values agreed up to the six decimal place (however, due to space constraints only two decimal places are shown). This means that, at least theoretically, the pilot ANF (4) should be a statistically efficient frequency and frequency rate tracker. A special simulation experiment was arranged to show that this actually holds true. Fig. 1 shows comparison of theoretical values of the lower tracking bounds with experimental results obtained for the signal (2) obeying assumptions (A1)–(A4), 3 different SNR
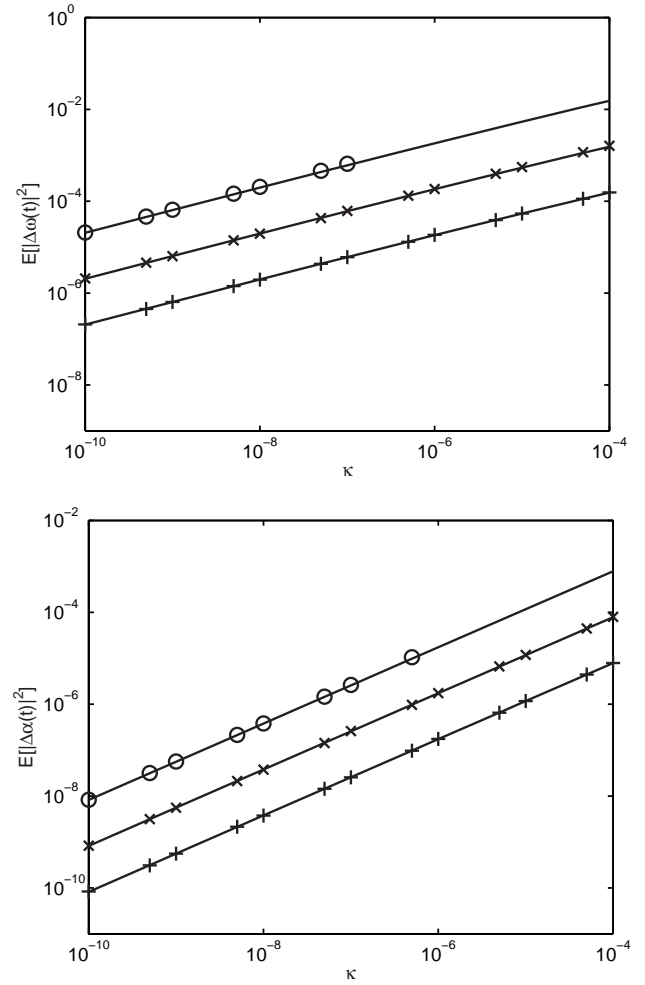




Fig. 1. Comparison of the theoretical values of the lower frequency (upper figure) and frequency rate (lower figure) tracking bounds (solid lines) with experimental results obtained for the signal with quasi-linear frequency changes for 3 different SNR values: SNR=0 dB ($\circ$), SNR=10 dB ($\times$), SNR=20 dB ($+$), and 13 different values of the rate of nonstationarity $\kappa$.

values (0 dB, 10 dB and 20 dB) and 13 different values of the rate of nonstationarity $\kappa$, ranging from $10^{-10}$ to $10^{-4}$. The mean-squared frequency estimation errors were evaluated (for the optimally tuned ANF 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 40000 iterations of the ANF filter (after the algorithm has reached its steady-state). The obtained results were next averaged over 20 realizations of $\{w(t), v(t)\}$. There are no results for SNR=0 dB and $\kappa \geq 10^{-6}$ since the algorithm was unable to track under such extremely difficult conditions.

Note the very good agreement between the theoretical curves and the results of computer simulations.

*Remark:* The algorithm (4) is simpler than the state-of-the-art multiple linear frequency tracker (MFT-L) proposed in [9]. It also performs better than MFT-L. In [9] the authors concluded that MFT-L nearly reaches lower tracking bounds for frequency tracking and frequency rate tracking. The loss in performance, compared to the optimal scheme, ranges from

1% to 7% for frequency tracking and from 9% to 28% for frequency rate tracking. Moreover, the optimal settings for frequency tracking differ from those for frequency rate tracking. We have shown that the algorithm (4) is able to reach both bounds, and that it reaches them simultaneously.

### C. From ANF to ANS

The pilot algorithm (4) and its approximate error model (5)–(6) will serve as a starting point for derivation of an adaptive notch smoother. Basically, we will follow the steps of [6], i.e., we will design a cascade of postprocessing filters increasing accuracy of amplitude, frequency, and frequency rate estimation. We will show that, using such a multistage scheme, one can significantly improve estimation results. Moreover, for quasi-linear frequency changes and under Gaussian assumptions, the ANS algorithm obtained in this way is a statistically efficient frequency and frequency rate smoother, i.e., when optimally tuned it reaches the corresponding lower smoothing bounds.

### III. FREQUENCY RATE SMOOTHING

Smoothed frequency rate estimates can be obtained by means of lowpass filtering of the results yielded by the pilot ANF. Such an approach will be called postfiltering. 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 characteristics of optimally tuned ANF. Finally, we will explain why the proposed scheme should work satisfactorily for any slow frequency rate variations, not necessarily of the random-walk type, and for any adaptation gains, not necessarily optimally tuned.

### A. Optimization

Suppose that the entire measurement record is available up to the instant $N$: $\mathcal{Y}(N) = \{y(1), \ldots, y(N)\}$. To obtain a fixed-interval smoothed estimate of $\alpha(t)$, further denoted by $\widetilde{\alpha}(t)$, we will pass the estimates $\widehat{\alpha}(t)$ through a noncausal filter $R(q^{-1}) = \ldots + r_{-1}q^{-1} + r_0 + r_1 q^1 + \ldots$

$$\widetilde{\alpha}(t) = R(q^{-1})\widehat{\alpha}(t) \, . \tag{10}$$

The filter $R(q^{-1})$ will be designed so as to minimize the mean-squared frequency rate estimation error $\mathrm{E}\{[\Delta\widetilde{\alpha}(t)]^2\}$, where $\Delta\widetilde{\alpha}(t) = \alpha(t) - \widetilde{\alpha}(t)$. After elementary but tedious calculations, one can show that

$$\Delta\widetilde{\alpha}(t) = (1-q^{-1})^2 X(q^{-1})e(t+1) + \frac{1 - X(q^{-1})}{1 - q^{-1}} \, w(t) \tag{11}$$

where

$$
\begin{aligned}
X(q^{-1}) &= R(q^{-1})S(q^{-1}) \\
S(q^{-1}) &= 1 - (1 - q^{-1})I_2(q^{-1}) \\
&= \gamma_\alpha q^{-1}/D(q^{-1}) \, .
\end{aligned} \tag{12}
$$

Let $\Delta(q^{-1}) = 1/(1 - q^{-1})$. Due to orthogonality of $\{e(t)\}$ and $\{w(t)\}$, one obtains

$$\mathrm{E}\{[\Delta\widetilde{\alpha}(t)]^2\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} h\left[X(e^{-j\xi})\right] d\xi \tag{13}$$

where

$$h[X] = \frac{XX^*}{|\Delta|^4} \, \sigma_e^2 + |\Delta|^2(1 - X)(1 - X^*) \, \sigma_w^2 \, . \tag{14}$$

Minimization of (14) can be achieved by minimizing $h[X(e^{-j\xi})]$ for every value of $\xi \in (-\pi, \pi]$.

Denote by

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

the so-called Wirtinger derivatives – symbolic derivatives with respect to a complex variable $z$, applicable to non-analytic functions, such as $h(\cdot)$ [10].

To find the optimal transfer function $X_0(q^{-1})$, i.e., the one that minimizes (13), we will request that

$$\left.\frac{\partial h}{\partial X^*}\right|_{X=X_0} = \frac{X_0}{|\Delta|^4} \, \sigma_e^2 - |\Delta|^2(1 - X_0) \, \sigma_w^2 = 0 \, . \tag{15}$$

Solving (15), one obtains

$$X_0(q^{-1}) = \frac{2\kappa}{2\kappa + (1 - q^{-1})^3(1 - q)^3} \, . \tag{16}$$

One of the key observations of this paper is that the transfer function $X_0(q^{-1})$ can be factorized as follows[1]

$$X_0(q^{-1}) \cong S_0(q^{-1})S_0(q) \tag{17}$$

where $S_0(q^{-1}) = S(q^{-1})|_{(\mu,\gamma_\omega,\gamma_\alpha)^{\mathrm{opt}}}$ and $\mu^{\mathrm{opt}}, \gamma_\omega^{\mathrm{opt}}, \gamma_\alpha^{\mathrm{opt}}$ denote optimal settings for the adaptive notch tracker (4). Since the optimal gain settings for quasi-linear frequency changes can't be established analytically – they are a solution of the set of high-order ($> 4$), and hence algebraically unsolvable, polynomial equations – the verification of (19) *must* be carried out numerically. Our check was based on comparison of frequency characteristics $X_0(e^{-j\xi})$ and $S_0(e^{-j\xi})S_0(e^{j\xi})$. Since an ideal match was observed for all values of $\kappa \in [10^{-10}, 0.1]$, and all values of $\xi \in (-\pi, \pi]$, we have the right to claim that the possible factorization errors, if any, are negligible.

Combining (12) with (17), one arrives at the following transfer function of the optimal postfiltering scheme

$$R_0(q^{-1}) = \frac{X_0(q^{-1})}{S_0(q^{-1})} \cong S_0(q) \, . \tag{18}$$

Since the filter $S_0(q)$ is anticausal, postfiltering can be realized by means of backward-time filtering of the frequency rate estimates $\widehat{\alpha}(t)$ yielded by the pilot algorithm. When postfiltering takes the form (18), the theoretical MSE, evaluated numerically using the formula (13), coincides, with high degree of precision, with the lower smoothing bound $\mathrm{LSB}_\alpha$

---

[1]This relationship can be conjectured from an analogous result proved analytically in [6] for a simpler case of random-walk frequency drift (see equation (15) in [6]).

derived in Appendix II. This completes our theoretical proof of statistical efficiency. Later on, in Section IX, we will confirm, by means of computer simulations, that the performance of the optimized scheme based on postfiltering indeed reaches its theoretical limit, exactly as predicted by the ALF-based analysis.

### B. Robust Smoothing Scheme

Although interesting from the theoretical viewpoint, the facts established so far are of little practical value – the optimality results are restricted to a specific model of frequency rate changes and they were derived under assumption that the optimal settings for the pilot filter are known. We will show that when postfiltering has the form

$$\widetilde{\alpha}(t) = S(q)\widehat{\alpha}(t) \tag{19}$$

improvement in estimation accuracy can be expected for *any* smooth frequency rate trajectory and for *any* choice of adaptation gains. Our robustness analysis will be based on the following relationship

$$\widehat{\alpha}(t) = S(q^{-1})\alpha(t) - I_1(q^{-1})e(t) \tag{20}$$

which is an equivalent of (6). Importantly, as already stressed in Section II-B, ALF approximations (5) and (6) [and hence also the relationship (20)] remain valid for any sequence of small one-step frequency rate changes $\{w(t)\}$, i.e., for any trajectory $\{\alpha(t)\}$ that is sufficiently smooth and hence can be regarded as a lowpass signal.

Note that for zero-mean measurement noise it holds that $\mathrm{E}[e(t)] = 0$, leading to

$$\mathrm{E}[\widehat{\alpha}(t)|\alpha(s), s \leq t] \cong S(q^{-1})\alpha(t) . \tag{21}$$

Note also that $S(q^{-1})$ is a lowpass filter with unity static gain $S(1) = 1$. Hence, when the instantaneous frequency rate varies slowly with time, the mean path of frequency rate estimates is approximately the time-delayed version of the true trajectory

$$\mathrm{E}[\widehat{\alpha}(t)|\alpha(s), s \leq t] \cong \alpha(t - \tau_\alpha) \tag{22}$$

where $\tau_\alpha = \mathrm{int}[t_\alpha]$ denotes the so-called estimation delay, equal to the integer part of the nominal (low-frequency) delay of the filter $S(q^{-1})$. The nominal delay of $S(q^{-1})$ is defined as

$$t_\alpha = -\lim_{\xi \mapsto 0} \frac{d\phi(\xi)}{d\xi} = -\lim_{\xi \mapsto 0} \frac{\phi(\xi)}{\xi}$$

where $\phi(\xi) = \arg\left[S(e^{-i\xi})\right]$ is the phase characteristic of $S(q^{-1})$. One can show that in the case considered

$$t_\alpha = I_2(1) = \frac{\gamma_\omega}{\gamma_\alpha} . \tag{23}$$

Estimation delay is a source of the bias error (sometimes called lag error) which, especially for large values of $\tau_\alpha$, may seriously degrade tracking performance of the pilot ANF.

For the smoothed estimate (19) the situation is different. Combining (19) with (20), one arrives at

$$\widetilde{\alpha}(t) = S(q)S(q^{-1})\alpha(t) - S(q)I_1(q^{-1})e(t) . \tag{24}$$

Since the nominal delay of the filter $S(q)S(q^{-1})$ is equal to zero, and its static gain is equal to 1, the steady-state ($1 \ll t \ll N$) smoothed estimate is approximately unbiased

$$\mathrm{E}[\widetilde{\alpha}(t)|\alpha(s), 1 \leq s \leq N] \cong \alpha(t) \tag{25}$$

which considerably increases estimation accuracy. Importantly, this conclusion holds true irrespective of the shape of the estimated frequency rate trajectory (as long as it is sufficiently smooth), and for any choice of adaptation gains $\mu$, $\gamma_\omega$ and $\gamma_\alpha$ that guarantee stable operation of the pilot ANF.

### IV. FREQUENCY SMOOTHING

Frequency smoothing can be achieved in an analogous way as frequency rate smoothing, i.e., by means of postfiltering the estimated frequency trajectory yielded by the pilot ANF. Similar to frequency rate smoothing, the proposed approach to frequency smoothing is robust to modeling errors.

### A. Optimization

Denote by $\widetilde{\omega}(t)$ the smoothed frequency estimate

$$\widetilde{\omega}(t) = P(q^{-1})\widehat{\omega}(t) \tag{26}$$

where $P(q^{-1}) = \ldots + p_{-1}q^{-1} + p_0 + p_1q^1 + \ldots$ is any noncausal filter. We will proceed similarly as as in Section III-A, i.e., we will find the filter $P(q^{-1})$ that minimizes the mean-squared frequency estimation error $\mathrm{E}\{[\Delta\widetilde{\omega}(t)]^2\}$, where $\Delta\widetilde{\omega}(t) = \omega(t) - \widetilde{\omega}(t)$. One can show that

$$\Delta\widetilde{\omega}(t) = (1 - q^{-1})Y(q^{-1})e(t) + \frac{1 - Y(q^{-1})}{(1 - q^{-1})^2} w(t-1) \tag{27}$$

where

$$\begin{aligned} Y(q^{-1}) &= P(q^{-1})Q(q^{-1}) \\ Q(q^{-1}) &= 1 - q(1 - q^{-1})^2 H_2(q^{-1}) \\ &= [\gamma_\omega + (\gamma_\alpha - \gamma_\omega)q^{-1}]/D(q^{-1}) . \end{aligned} \tag{28}$$

Due to orthogonality of $\{e(t)\}$ and $\{w(t)\}$, (27) leads to

$$\mathrm{E}\{[\Delta\widetilde{\omega}(t)]^2\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} g\left[Y(e^{-j\xi})\right] d\xi \tag{29}$$

where

$$g[Y] = \frac{YY^*}{|\Delta|^2} \sigma_e^2 + |\Delta|^4(1 - Y)(1 - Y^*) \sigma_w^2 . \tag{30}$$

The optimal transfer function $Y_0(q^{-1})$ can be obtained by solving

$$\left.\frac{\partial g}{\partial Y^*}\right|_{Y=Y_0} = \frac{Y_0}{|\Delta|^2} \sigma_e^2 - |\Delta|^4(1 - Y_0) \sigma_w^2 = 0 \tag{31}$$

which leads to

$$Y_0(q^{-1}) = X_0(q^{-1}) . \tag{32}$$

Finally, after combining (32) with (28) and (17), one arrives at

$$P_0(q^{-1}) \cong \frac{S_0(q^{-1})S_0(q)}{Q_0(q^{-1})} \tag{33}$$

where $Q_0(q^{-1}) = Q(q^{-1})|_{(\mu,\gamma_\omega,\gamma_\alpha)^{\mathrm{opt}}}$.

Note that, rather unexpectedly, the optimal filter (33), designed for frequency smoothing, differs from the analogous filter (18) that was earlier designed for frequency rate smoothing.

Similar to frequency rate smoothing, when postfiltering takes the form (33), the theoretical MSE, evaluated numerically using the formula (29), coincides, with high degree of precision, with the lower smoothing bound $\text{LSB}_\omega$ derived in Appendix II. In Section IX we will confirm, by means of computer simulations, that the optimized scheme based on postfiltering is statistically efficient, i.e., that its performance indeed reaches the theoretical limit mentioned above.

### B. Robust Smoothing Scheme

Based on (33) we propose the following postfiltering scheme

$$\widetilde{\omega}(t) = S(q)T(q^{-1})\widehat{\omega}(t) \tag{34}$$

where

$$T(q^{-1}) = \frac{S(q^{-1})}{Q(q^{-1})} = \frac{\gamma_\alpha q^{-1}}{\gamma_\omega + (\gamma_\alpha - \gamma_\omega)q^{-1}} = \frac{b_1 q^{-1}}{1 + c_1 q^{-1}}$$

$$b_1 = \frac{\gamma_\alpha}{\gamma_\omega}, \quad c_1 = \frac{\gamma_\alpha - \gamma_\omega}{\gamma_\omega}.$$

Note that the filter $T(q^{-1})$ is causal and the filter $S(q)$ is anticausal. Therefore, postfiltering can be realized by means of forward-time filtering of the frequency trajectory $\{\widehat{\omega}(t)\}$ using the filter $T(q^{-1})$, followed by backward-time filtering of the results using the filter $S(q)$.

Robustness analysis of (34) is similar to that carried for (19). Using (5), one can show that

$$\widehat{\omega}(t) = Q(q^{-1})\omega(t) - H_1(q^{-1})e(t) \tag{35}$$

leading to (in steady-state)

$$\begin{aligned} \text{E}[\widetilde{\omega}(t)|\omega(s), 1 \leq s \leq N] &= S(q)T(q^{-1})Q(q^{-1})\omega(t) \\ &= S(q)S(q^{-1})\omega(t) \cong \omega(t) \end{aligned} \tag{36}$$

for any frequency trajectory that can be modeled as a lowpass signal and for any selection of stabilizing adaptation gains.

Suppose that postfiltering is restricted to the forward pass

$$\bar{\omega}(t) = T(q^{-1})\widehat{\omega}(t).$$

Then it holds that (in steady-state)

$$\begin{aligned} \text{E}[\bar{\omega}(t)|\omega(s), s \leq t] &= T(q^{-1})Q(q^{-1})\omega(t) \\ &= S(q^{-1})\omega(t) \cong \omega(t - \tau_\omega) \end{aligned}$$

where $\tau_\omega = \tau_\alpha$ is the delay introduced by the filter $S(q^{-1})$ – identical with that established in Section III for frequency rate tracking.

## V. AMPLITUDE SMOOTHING

Having smoothed frequency estimates, one can re-estimate the amplitude coefficients. This can be achieved using the frequency-guided version of the algorithm (4), obtained by replacing the filtering-type (causal) frequency estimates $\widehat{\omega}(t)$ with their smoothed (noncausal) counterparts $\widetilde{\omega}(t)$, derived in Section IV

$$\begin{aligned} \widetilde{f}(t) &= e^{j\widetilde{\omega}(t)}\widetilde{f}(t-1) \\ \bar{\varepsilon}(t) &= y(t) - \bar{a}(t-1)\widetilde{f}(t) \\ \bar{a}(t) &= \bar{a}(t-1) + \mu\widetilde{f}^*(t)\bar{\varepsilon}(t) \\ \bar{s}(t) &= \bar{a}(t)\widetilde{f}(t). \end{aligned} \tag{37}$$

Since the frequency-guided ANF (37) is identical with the analogous algorithm studied in [6], we will simply repeat conclusions reached there: the smoothed estimates of amplitude coefficients can be obtained by means of backward-time filtering of the estimates yielded by the algorithm (37)

$$\widetilde{a}(t) = F(q)\bar{a}(t) \tag{38}$$

where

$$F(q) = \frac{\mu}{1 - (1 - \mu)q}$$

is the unity-gain anticausal lowpass filter "matched" to tracking characteristics of (37). When the amplitude trajectory $\{a(t)\}$ can be modeled as a lowpass process and $1 \ll t \ll N$ (steady-state conditions), it holds that

$$\text{E}[\widetilde{a}(t)|a(s), 1 \leq s \leq N] = F(q)F(q^{-1})a(t) \cong a(t) \tag{39}$$

which means that the smoothed amplitude estimates are approximately unbiased - see [6] for more details.

When smoothing is not incorporated, the corresponding relationship reads

$$\text{E}[\bar{a}(t)|a(s), s \leq t] = F(q^{-1})a(t) \cong a(t - \tau_a) \tag{40}$$

where $\tau_a = \text{int}[t_a]$ denotes estimation delay arising in the amplitude tracking loop of the frequency-guided ANF, equal to the integer part of the nominal delay of the filter $F(q^{-1})$

$$t_a = \frac{1 - \mu}{\mu}.$$

## VI. FIXED-INTERVAL ADAPTIVE NOTCH SMOOTHER

The fixed-interval ANS algorithm can be obtained by combining all steps described in Sections II-V and adding the final signal reconstruction step

$$\widetilde{s}(t) = \widetilde{a}(t)\widetilde{f}(t). \tag{41}$$

To reduce computational burden, we will rewrite all recursions in an equivalent form that alleviates the need to compute the quantities $\widehat{f}(t)$, $\widehat{a}(t)$, $\bar{f}(t)$, $\bar{a}(t)$, $\widetilde{f}(t)$ and $\widetilde{a}(t)$. The resulting cost-optimized ANS is summarized in Table II (the output filter, specified in Table II, incorporates amplitude smoothing). Computational load associated with this algorithm is equal to 37 real multiply/add operations and 1 real division operation per time update for the variant with frequency rate estimation, and 33 real multiply/add operations and 1

real division operation when frequency rate estimation is not needed.

Unless a specific prior knowledge about the analyzed signal is available, we recommend setting $\widehat{s}(1) = y(1)$, $\widehat{\alpha}(1) = 0$, and adopting for $\widehat{\omega}(1)$ the value that stems from spectral analysis (parametric or nonparametric) of a short initial fragment of the available data record. Less careful choice of initial conditions may result in long initialization transients (especially for small adaptation gains) but does not prevent the pilot algorithm from reaching the small-error tracking mode after a sufficiently long initial convergence period.

*Remark:* Denote by $\varphi(t) = \varphi_0 + \sum_{l=1}^{t} \omega(l)$ the instantaneous phase of $s(t)$, which under (2) obeys

$$\varphi(t+1) = \varphi(t) + \omega(t) + \alpha(t) \tag{42}$$

and suppose that the time-varying signal amplitude (real-valued) can be modeled as a random-walk process

$$a(t+1) = a(t) + n(t) \tag{43}$$

where $\{n(t)\}$ is a Gaussian white noise, independent of $\{w(t)\}$ and $\{v(t)\}$. Combining (2) with (42) and (43), one can write down the analyzed signal in the following nonlinear state space form[2]

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\eta(t)$$
$$\mathbf{z}(t) = \mathbf{f}[\mathbf{x}(t)] + \zeta(t) \tag{44}$$

where the state vector is given by $\mathbf{x}(t) = [\varphi(t), \omega(t), \alpha(t), a(t)]^{\mathrm{T}}$, the observation vector $\mathbf{z}(t) = [\mathrm{Re}\{y(t)\}, \mathrm{Im}\{y(t)\}]^{\mathrm{T}}$ is comprised of the real and imaginary parts of $y(t)$, the driving and measurement noise vectors are given by $\eta(t) = [w(t+1), n(t)]^{\mathrm{T}}$ and $\zeta(t) = [\mathrm{Re}\{v(t)\}, \mathrm{Im}\{v(t)\}]^{\mathrm{T}}$, respectively, and

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{f}[\mathbf{x}(t)] = \begin{bmatrix} a(t) \cos \varphi(t) \\ a(t) \sin \varphi(t) \end{bmatrix}.$$

Based on (44), one can seek the solution to our estimation problem by employing the classical extended Kalman filtering/smoothing approach [11]. Not getting into details, we note that extended Kalman smoothing (EKS) can be realized by means of backward-time filtering of the results yielded by the extended Kalman filter (EKF) – the corresponding recursive algorithm is known as the Rauch-Tung-Striebel (RTS) smoother. Note that, from the qualitative viewpoint, the procedure proposed in this paper resembles the classical one. The main difference lies in computational complexity. Even if the special form of the matrices $\mathbf{A}$ and $\mathbf{B}$ is exploited to avoid unnecessary multiplications, the backward-time filtering step of the RTS algorithm requires performing 3 multiplications and 1 inversion of $4 \times 4$ covariance matrices per time update. This computational burden is further increased by the cost of running the extended Kalman filter and saving its results (including all *a priori* and *a posteriori* covariance matrices for

[2]For simplicity we ignore the problem of phase unwrapping.

TABLE II
FIXED-INTERVAL ADAPTIVE NOTCH SMOOTHER.

*pilot filter* :

$$\varepsilon(t) = y(t) - e^{j[\widehat{\omega}(t-1)+\widehat{\alpha}(t-1)]}\widehat{s}(t-1)$$
$$\widehat{s}(t) = e^{j[\widehat{\omega}(t-1)+\widehat{\alpha}(t-1)]}\widehat{s}(t-1) + \mu\varepsilon(t)$$
$$b(t) = |\widehat{s}(t-1)|^2$$
$$\delta(t) = \mathrm{Im}[\varepsilon(t)e^{-j[\widehat{\omega}(t-1)+\widehat{\alpha}(t-1)]}\widehat{s}^*(t-1)]/b(t)$$
$$\widehat{\alpha}(t) = \widehat{\alpha}(t-1) + \gamma_\alpha \delta(t)$$
$$\widehat{\omega}(t) = \widehat{\omega}(t-1) + \widehat{\alpha}(t-1) + \gamma_\omega \delta(t)$$
$$t = 2, \ldots, N$$

*frequency rate smoother [optional]* :

$$\widetilde{\alpha}(N) = \widehat{\alpha}(N)$$
$$\widetilde{\alpha}(N-1) = \widehat{\alpha}(N-1)$$
$$\widetilde{\alpha}(N-2) = \widehat{\alpha}(N-2)$$
$$\widetilde{\alpha}(t) = -d_1\widetilde{\alpha}(t+1) - d_2\widetilde{\alpha}(t+2)$$
$$\qquad -d_3\widetilde{\alpha}(t+3) + \gamma_\alpha\widehat{\alpha}(t+1)$$
$$t = N-3, \ldots, 1$$

*frequency smoother* :

$$\bar{\omega}(1) = \widehat{\omega}(1)$$
$$\bar{\omega}(t) = -c_1\bar{\omega}(t-1) + b_1\widehat{\omega}(t-1)$$
$$t = 2, \ldots, N$$
$$\widetilde{\omega}(N) = \bar{\omega}(N)$$
$$\widetilde{\omega}(N-1) = \bar{\omega}(N-1)$$
$$\widetilde{\omega}(N-2) = \bar{\omega}(N-2)$$
$$\widetilde{\omega}(t) = -d_1\widetilde{\omega}(t+1) - d_2\widetilde{\omega}(t+2)$$
$$\qquad -d_3\widetilde{\omega}(t+3) + \gamma_\alpha\bar{\omega}(t+1)$$
$$t = N-3, \ldots, 1$$

*frequency−guided filter* :

$$\bar{s}(1) = \widehat{s}(1)$$
$$\bar{\varepsilon}(t) = y(t) - e^{j\widetilde{\omega}(t)}\bar{s}(t-1)$$
$$\bar{s}(t) = e^{j\widetilde{\omega}(t)}\bar{s}(t-1) + \mu\bar{\varepsilon}(t)$$
$$t = 2, \ldots, N$$

*output filter* :

$$\widetilde{s}(N) = \bar{s}(N)$$
$$\widetilde{s}(t) = (1-\mu)e^{-j\widetilde{\omega}(t+1)}\widetilde{s}(t+1) + \mu\bar{s}(t)$$
$$t = N-1, \ldots, 1$$

$t = 1, \ldots, N$). In contrast with this, the algorithm summarized in Table II does not involve any matrix calculations and does not require saving any matrix-valued quantities. This results in huge computational savings.

## VII. FIXED-LAG ADAPTIVE NOTCH SMOOTHER

The fixed-lag ANS can be obtained by restricting postprocessing of ANF estimates to the recent $\tau$ time-steps only.

The resulting "sawtooth" smoothing algorithm is summarized in Table III. To avoid confusion most of the quantities were indexed by $t$.

Additional computational cost of carrying out postprocessing steps, i.e., the computational overhead of smoothing, grows linearly with the time lag $\tau$, and is equal to $22\tau$ real multiply/add operations per time update for the algorithm that incorporates frequency rate estimation, and $18\tau + 4$ real multiply/add operations for the variant without frequency rate estimation.

The tracking accuracy improvements offered by smoothing gradually saturate with growing $\tau$. Similarly as in [6], one can argue that only marginal improvements can be expected when $\tau$ is increased beyond the "principal" delay equal to

$$\tau_0 = \max\{\tau_\omega, \tau_a\}.$$

## VIII. Extensions

Using the framework described in [12] (see Section II-C there), the proposed ANS algorithm can be easily extended to the multiple frequencies case, where

$$s(t) = \sum_{i=1}^{k} s_i(t)$$
$$s_i(t) = a_i(t)f_i(t), \quad f_i(t) = e^{j\sum_{l=1}^{t}\omega_i(l)}$$

and $k$ denotes the number of cisoids embedded in noise.

The resulting algorithm is a parallel estimation scheme combining $k$ "local" ANS filters. The component filters are designed to track different frequency components of $s(t)$ and are driven by "global" prediction errors

$$\varepsilon(t) = y(t) - \sum_{i=1}^{k} e^{j[\widehat{\omega}_i(t-1)+\widehat{\alpha}_i(t-1)]}\widehat{s}_i(t-1)$$
$$\bar{\varepsilon}(t) = y(t) - \sum_{i=1}^{k} e^{j\widetilde{\omega}_i(t)}\bar{s}_i(t-1).$$

All algorithms can be applied to real-valued signals $y^{\mathrm{R}}(t)$. In order to process such signals, one may formally regard the input data as complex-valued by setting $y(t) = y^{\mathrm{R}}(t) + j0$, and ignore imaginary parts at the ANF/ANS output: $\widetilde{s}^{\mathrm{R}}(t) = \mathrm{Re}[\widetilde{s}(t)], \widetilde{s}_t^{\mathrm{R}}(t-\tau) = \mathrm{Re}[\widetilde{s}_t(t-\tau)].$

## IX. Simulation and Experimental Results

Our simulation study will focus on two aspects of the proposed adaptive notch smoothing scheme: optimality and robustness.

Although demonstration of the algorithm's optimality, i.e., its ability to reach the Cramér-Rao-type lower smoothing bounds is mainly of theoretical value, it is important as it allows one to specify conditions under which some absolute performance limits can be reached.

From the practical viewpoint, the most important property of the estimation algorithm is its robustness, i.e., insensivity to modeling errors. We will show that, exactly as predicted, the proposed ANS algorithm outperforms its ANF counterpart for a wide range of operating conditions, including different (nonstandard) frequency/amplitude trajectories and different (non-optimal) values of adaptation gains.

TABLE III
Fixed-lag adaptive notch smoother.

*pilot filter* :
$$\varepsilon(t) = y(t) - e^{j[\widehat{\omega}(t-1)+\widehat{\alpha}(t-1)]}\widehat{s}(t-1)$$
$$\widehat{s}(t) = e^{j[\widehat{\omega}(t-1)+\widehat{\alpha}(t-1)]}\widehat{s}(t-1) + \mu\varepsilon(t)$$
$$b(t) = |\widehat{s}(t-1)|^2$$
$$\delta(t) = \mathrm{Im}[\varepsilon(t)e^{-j[\widehat{\omega}(t-1)+\widehat{\alpha}(t-1)]}\widehat{s}^*(t-1)]/b(t)$$
$$\widehat{\alpha}(t) = \widehat{\alpha}(t-1) + \gamma_\alpha\delta(t)$$
$$\widehat{\omega}(t) = \widehat{\omega}(t-1) + \widehat{\alpha}(t-1) + \gamma_\omega\delta(t)$$

*frequency rate smoother [optional]* :
$$\widetilde{\alpha}_t(t) = \widehat{\alpha}(t)$$
$$\widetilde{\alpha}_t(t-1) = \widehat{\alpha}(t-1)$$
$$\widetilde{\alpha}_t(t-2) = \widehat{\alpha}(t-2)$$
$$\widetilde{\alpha}_t(i) = -d_1\widetilde{\alpha}_t(i+1) - d_2\widetilde{\alpha}_t(i+2)$$
$$\qquad -d_3\widetilde{\alpha}_t(i+3) + \gamma_\alpha\widehat{\alpha}(i+1)$$
$$i = t-3,\ldots,t-\tau$$

*frequency smoother* :
$$\bar{\omega}(t) = -c_1\bar{\omega}(t-1) + b_1\widehat{\omega}(t-1)$$
$$\widetilde{\omega}_t(t) = \bar{\omega}(t)$$
$$\widetilde{\omega}_t(t-1) = \bar{\omega}(t-1)$$
$$\widetilde{\omega}_t(t-2) = \bar{\omega}(t-2)$$
$$\widetilde{\omega}_t(i) = -d_1\widetilde{\omega}_t(i+1) - d_2\widetilde{\omega}_t(i+2)$$
$$\qquad -d_3\widetilde{\omega}_t(i+3) + \gamma_\alpha\bar{\omega}(i+1)$$
$$i = t-3,\ldots,t-\tau$$

*frequency−guided filter* :
$$\bar{s}_t(t-\tau-1) = \widehat{s}_t(t-\tau-1)$$
$$\bar{\varepsilon}_t(i) = y(i) - e^{j\widetilde{\omega}(i)}\bar{s}_t(i-1)$$
$$\bar{s}_t(i) = e^{j\widetilde{\omega}(i)}\bar{s}_t(i-1) + \mu\bar{\varepsilon}_t(i)$$
$$i = t-\tau,\ldots,t$$

*output filter* :
$$\widetilde{s}_t(t) = \bar{s}_t(t)$$
$$\widetilde{s}_t(i) = (1-\mu)e^{-j\widetilde{\omega}_t(i+1)}\widetilde{s}_t(i+1) + \mu\bar{s}_t(i)$$
$$i = t-1,\ldots,t-\tau$$

### A. Optimality

Fig. 2 shows comparison of theoretical values of the lower smoothing bounds with experimental results obtained for the signal (2) obeying assumptions (A1)–(A4), 3 different SNR values (0 dB, 10 dB and 20 dB) and 13 different values of the rate of nonstationarity $\kappa$, ranging from $10^{-10}$ to $10^{-4}$. Similarly as in Section II-B, all MSE values were computed for the optimally tuned ANS algorithm by means of joint time and ensemble averaging. Note the excellent agreement between the theoretical curves and the results of computer simulations.

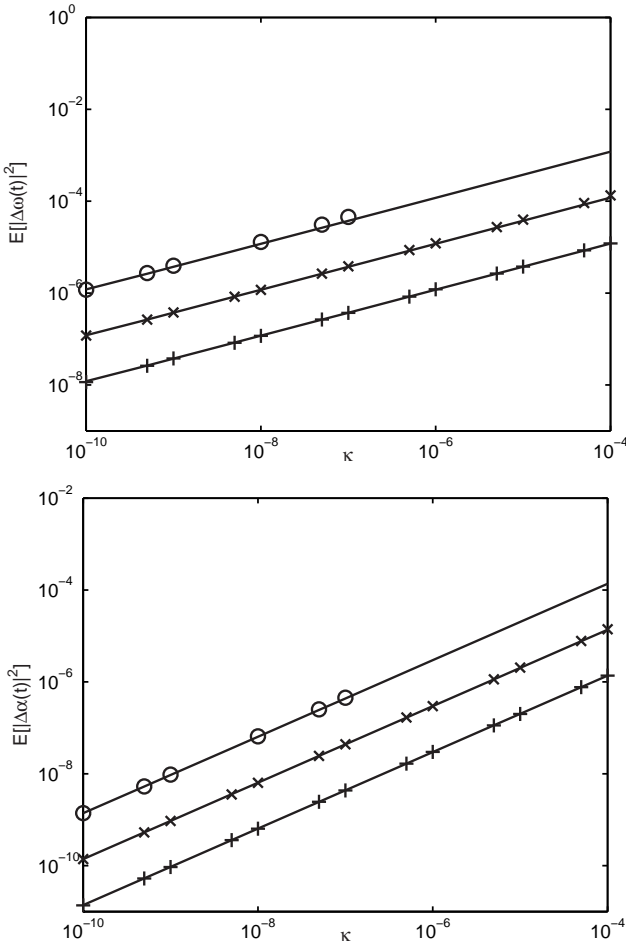The possible margins of improvement $\mathrm{LTB}_\omega/\mathrm{LSB}_\omega$ and

Fig. 2. Comparison of the theoretical values of the lower frequency (upper figure) and frequency rate (lower figure) smoothing bounds (solid lines) with experimental results obtained for the signal with quasi-linear frequency changes, for 3 different SNR values: SNR=0 dB (○), SNR=10 dB (×), SNR=20 dB (+), and 13 different values of the rate of nonstationarity $\kappa$.



Fig. 3. Dependence of the limiting tracking versus smoothing variance ratios on the degree of signal nonstationarity $\kappa$.

$\mathrm{LTB}_\alpha/\mathrm{LSB}_\alpha$, which are functions of $\kappa$, are depicted in Fig. 3.

### B. Robustness

To check performance of the fixed-interval ANS algorithm, a noisy quas-iperiodically varying signal (2) was generated with fast amplitude and frequency changes

$$a(t) = 1 + 0.5\cos(2\pi t/2000) , \quad \omega(t) = \sin(2\pi t/2000) .$$

Figs. 4 and 5 show the comparison of the steady-state mean-squared frequency and signal estimation errors, yielded by the ANF algorithm without frequency rate estimation, the ANS algorithm without frequency rate estimation, the ANF algorithm with frequency rate estimation, and the ANS algorithm with frequency rate estimation. The comparison was made for 40 different values of the adaptation gain $\mu$ and for two noise intensities: $\sigma_v = 0.56$ (SNR=5 dB) and $\sigma_v = 0.1$ (SNR=20 dB). To reduce the number of design degrees of freedom, the two other gains adopted for ANF/ANS algorithms with frequency rate estimation were set to: $\gamma_\omega = \mu^2/2$ and
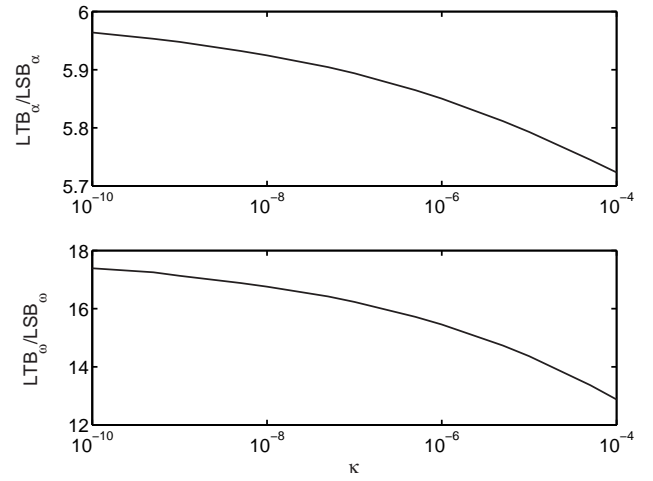
$\gamma_\alpha = \mu\gamma_\omega/4$ – in agreement with the general tendency revealed in Table I. The ANF/ANS algorithms without frequency rate estimation were obtained by setting $\gamma_\alpha = 0$. The results obtained for the MFT-L algorithm [which can be rewritten in a form similar to (4)] were slightly but uniformly inferior (by approximately 3 dB) to those obtained for our pilot tracker. For this reason they are not shown. All MSE values were obtained by means of joint time averaging (the evaluation interval [2001,8000] was placed inside a wider analysis interval [1,10000]), and ensemble averaging (100 realizations of measurement noise were used). As expected, the ANS algorithms yielded uniformly better results than their ANF counterparts. The achievable variance reduction is approximately equal to 20 dB for frequency estimation and 10 dB for signal estimation, respectively. Rather surprisingly, for the particular signal analyzed, a very small signal estimation gain can be achieved by switching from the ANS algorithm without frequency rate estimation, to the ANS algorithm with frequency rate estimation (Fig. 5), even though the frequency estimation benefits are evident (Fig. 4).

Finally, Table IV shows comparison of the mean-squared frequency and signal estimation errors yielded by the well-tuned EKF/EKS algorithms, based on the model (44), with the analogous results provided by the well-tuned ANF/ANS algorithms, based on the (equivalent) model (2). The EKF/EKS algorithms were supported with the true values of the measurement noise intensities ($\sigma_v^2 = 0.31$ or $0.01$). The remaining two variances ($\sigma_a^2$ and $\sigma_w^2$) were also set equal to the true mean-squared rates of change of the corresponding quantities [$a(t)$ and $\alpha(t)$, respectively]: $\sigma_a^2 \cong \omega_0^2/8 = 1.23 \cdot 10^{-6}$, $\sigma_w^2 \cong \omega_0^4/2 = 4.87 \cdot 10^{-11}$, where $\omega_0 = 2\pi/2000$. The numbers shown for the ANF/ANS algorithms correspond to the lowest points on the ○/∗ plots depicted in Figs. 4 and 5. Note that, in spite of their computational simplicity, the proposed ANF/ANS algorithms compare very favorably with the classical EKF/EKS tools.
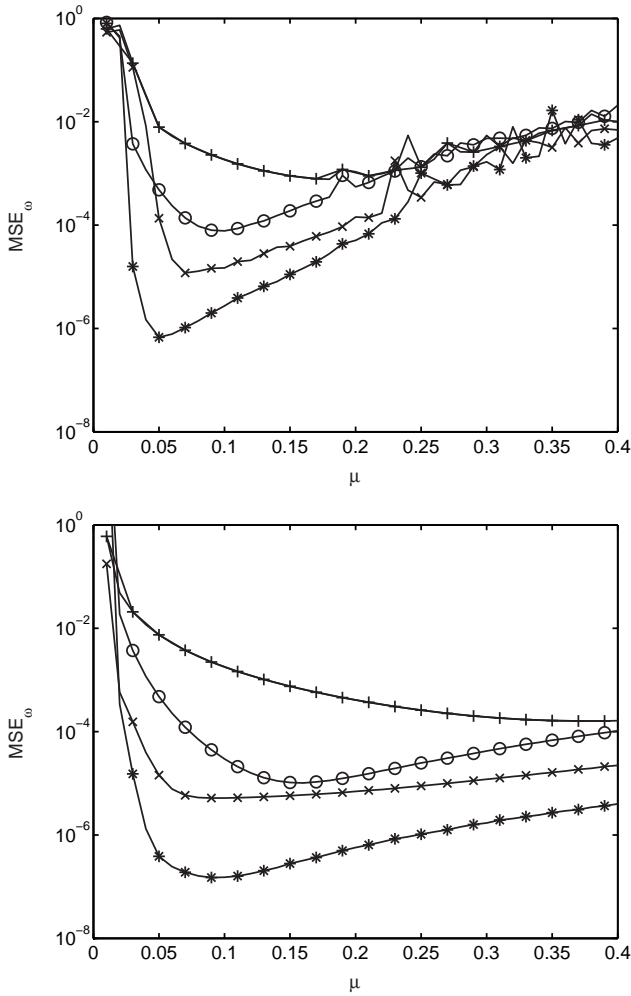
Fig. 4. Dependence of the mean-squared frequency estimation error on adaptation gain $\mu$ for: ANF algorithm without frequency rate estimation (+), ANS algorithm without frequency rate estimation ($\times$), ANF algorithm with frequency rate estimation ($\bigcirc$), and ANS algorithm with frequency rate estimation ($\ast$). Upper figure – SNR=5 dB, lower figure – SNR= 20 dB.
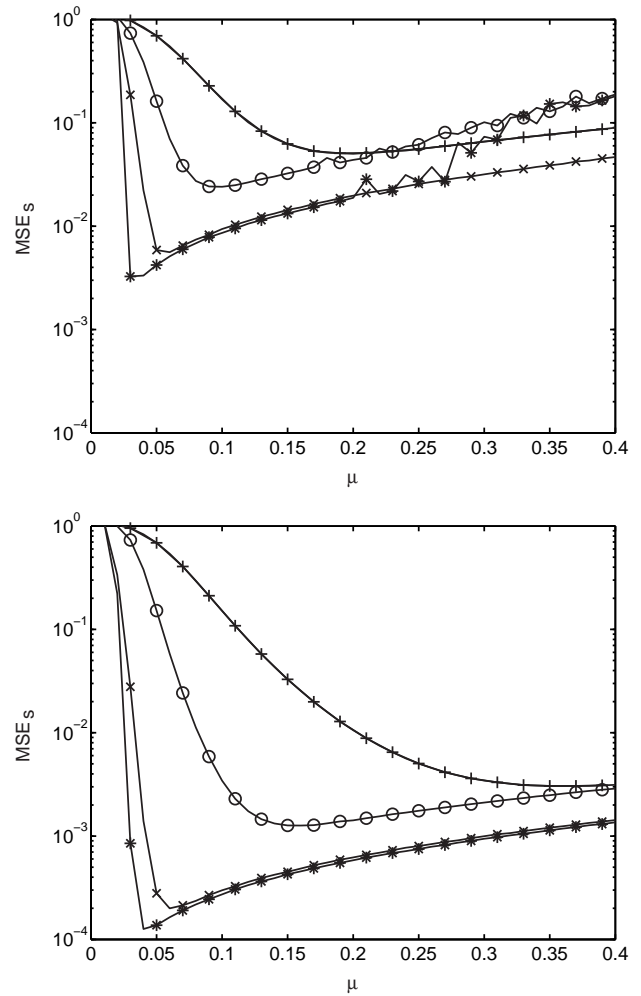


Fig. 5. Dependence of the mean-squared signal estimation error on adaptation gain $\mu$ for: ANF algorithm without frequency rate estimation (+), ANS algorithm without frequency rate estimation ($\times$), ANF algorithm with frequency rate estimation ($\bigcirc$), and ANS algorithm with frequency rate estimation ($\ast$). Upper figure – SNR=5 dB, lower figure – SNR= 20 dB.

TABLE IV
COMPARISON OF THE MEAN-SQUARED FREQUENCY ESTIMATION ERRORS (UPPER TABLE) AND SIGNAL ESTIMATION ERRORS (LOWER TABLE), YIELDED BY THE WELL-TUNED EKF/EKS ALGORITHMS AND WELL-TUNED ANF/ANS ALGORITHMS.

$\mathrm{MSE}_\omega$:

| SNR | EKF | ANF | EKS | ANS |
|-----|-----|-----|-----|-----|
| 5 dB | $5.3 \cdot 10^{-4}$ | $7.7 \cdot 10^{-5}$ | $9.4 \cdot 10^{-7}$ | $6.8 \cdot 10^{-7}$ |
| 20 dB | $4.5 \cdot 10^{-5}$ | $1.0 \cdot 10^{-5}$ | $5.4 \cdot 10^{-8}$ | $1.5 \cdot 10^{-7}$ |

$\mathrm{MSE}_s$:

| SNR | EKF | ANF | EKS | ANS |
|-----|-----|-----|-----|-----|
| 5 dB | $2.4 \cdot 10^{-1}$ | $2.4 \cdot 10^{-2}$ | $7.0 \cdot 10^{-2}$ | $3.3 \cdot 10^{-3}$ |
| 20 dB | $9.3 \cdot 10^{-3}$ | $1.3 \cdot 10^{-3}$ | $3.1 \cdot 10^{-4}$ | $1.2 \cdot 10^{-4}$ |

*C. Application example*

Conventional methods of measuring the rotational speed of a combustion engine rely on dedicated sensors, such as tachometers, photo probes, etc. Deployment of such sensors may be difficult, expensive and/or inconvenient. However, the sound emitted by an operating engine usually provides sufficient information to enable a cheap and accurate remote RPM (revolutions per minute) measurement using acoustic sensing.

The proposed smoothing scheme was used to analyze a 4-second recording of a motorcycle engine noise, sampled at a frequency $f_s = 1.1$ kHz. The analyzed fragment corresponds to subsequent: acceleration, gear change, acceleration and braking, respectively. The complex-valued version of the signal was obtained using the Hilbert transform. The spectrogram of the recording, depicted in Fig. 6, shows that the signal contains $k = 12$ harmonic components with instantaneous frequencies $\omega_i(t)$ that are linearly related to the fundamental

frequency $\omega_1(t)$:

$$\omega_i(t) = i\omega_1(t) , \quad i = 2, \ldots, 12 . \tag{45}$$

To cope with the harmonic signal structure, a slightly modified version of the pilot filter was employed. All harmonic components were tracked using a parallel structure made up of 12 pilot filters

$$\widehat{s}_i(t) = e^{j[\widehat{\omega}_i(t-1)+\widehat{\alpha}_i(t-1)]}\widehat{s}_i(t-1) + \mu\epsilon(t)$$
$$i = 1, \ldots, 12$$

driven by the common prediction error signal

$$\epsilon(t) = y(t) - \sum_{i=1}^{12} e^{j[\widehat{\omega}_i(t-1)+\widehat{\alpha}_i(t-1)]}\widehat{s}_i(t-1) .$$

In order to estimate the instantaneous angular frequencies $\widehat{\omega}_i(t)$ and frequency rates $\widehat{\alpha}_i(t)$, the proposed pilot frequency tracker was applied to the 5-th harmonic signal component (since this component is strong, it guarantees high signal-to-noise ratio):

$$b_5(t) = |\widehat{s}_5(t-1)|^2$$
$$\delta_5(t) = \mathrm{Im}[\epsilon(t)e^{-j[\widehat{\omega}_5(t-1)+\widehat{\alpha}_5(t-1)]}\widehat{s}_5^*(t-1)]/b_5(t)$$
$$\widehat{\alpha}_5(t) = \widehat{\alpha}_5(t) + \gamma_\alpha\delta_5(t)$$
$$\widehat{\omega}_5(t) = \widehat{\omega}_5(t-1) + \widehat{\alpha}_5(t) + \gamma_\omega\delta_5(t) .$$

The estimates of the remaining frequencies and frequency rates were obtained by exploiting (45)

$$\widehat{\alpha}_i(t) = i\widehat{\alpha}_5(t)/5 , \quad \widehat{\omega}_i(t) = i\widehat{\omega}_5(t)/5$$
$$i = 1, \ldots, 12, \ i \neq 5 .$$

The parameters of the pilot filter were set to the following values: $\mu = 0.05$, $\gamma_\omega = \mu^2/2$ and $\gamma_\alpha = \mu\gamma_\omega/4$ – in agreement with the rule of thumb presented in the preceding subsection. The smoothed estimates of $\omega_5(t)$ and $\alpha_5(t)$ were obtained by backward-time filtering of the pilot estimates $\widehat{\omega}_5(t)$ and $\widehat{\alpha}_5(t)$ using the algorithms summarized in Table II.

Fig. 7 shows the comparison of the frequency estimates

$$\widehat{f}_5(t) = \frac{\widehat{\omega}_5(t)f_{\mathrm{s}}}{2\pi}$$

obtained using the pilot filter (top plot) and frequency smoother (bottom plot). The corresponding instantaneous RPM values can be obtained from: $\mathrm{RPM}(t) = 60\widehat{f}_5(t)/5$. Similar comparison of the frequency rate estimates is shown in Fig. 8. In both cases, the results obtained using the noncausal estimation scheme are significantly better than those yielded by the pilot (causal) algorithm.

## X. CONCLUSION

Extraction/elimination of nonstationary sinusoidal signals buried in noise can be carried out using adaptive notch filters (ANF). We have designed a new ANF algorithm based on quasi-linear model of signal frequency changes, and we have shown that, under Gaussian assumptions, it is a statistically efficient frequency and frequency rate tracker. Then, based on analysis of tracking properties of the proposed algorithm, we have designed a cascade of postprocessing filters increasing
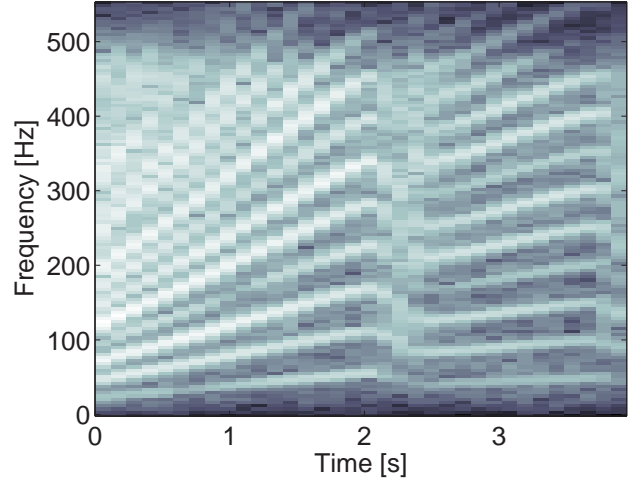


Fig. 6. Spectrogram of the acoustic recording of an accelerating motorcycle. The period between 2 s and 2.5 s corresponds to gear shifting.
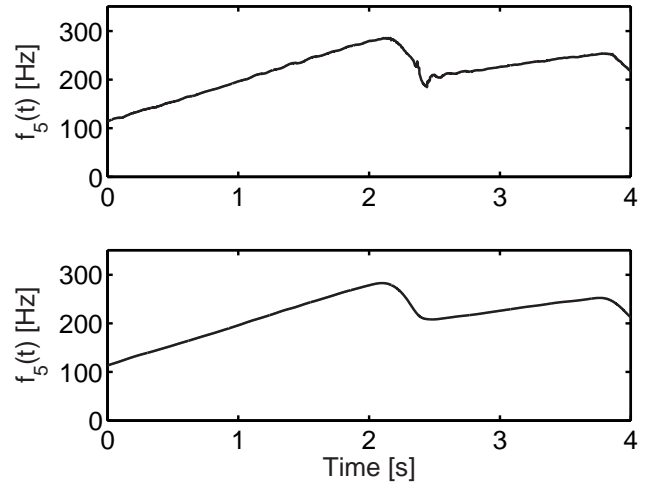


Fig. 7. Estimated instantaneous frequency of the 5-th harmonic signal component. Top plot: pilot estimates. Bottom plot: smoothed estimates.
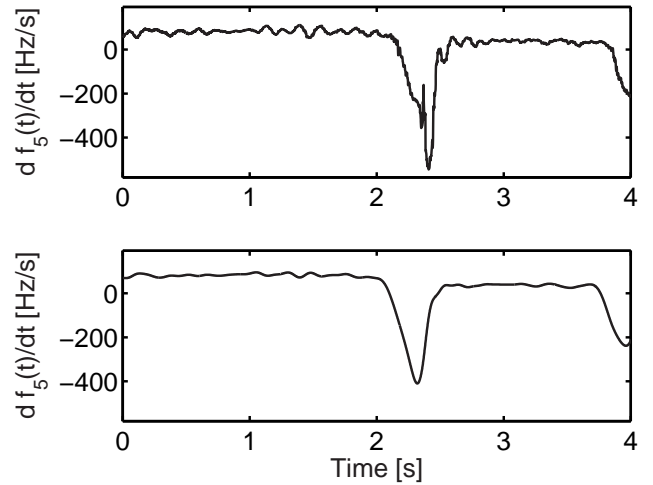


Fig. 8. Estimated instantaneous frequency rate of the 5-th harmonic signal component. Top plot: pilot estimates. Bottom plot: smoothed estimates.

accuracy of signal frequency, frequency rate, and amplitude estimation. We have shown that the resulting adaptive notch smoother (ANS) is statistically efficient (under ideal conditions) and robust to signal model misspecification. It yields improved estimation results compared to its causal (ANF) counterpart. It also works better than the currently available, simpler ANS algorithm.

## APPENDIX I

### Derivation of (5) and (6)

Denote by $\Delta\widehat{s}(t) = s(t) - \widehat{s}(t)$ the signal estimation error and let $\Delta\widehat{x}(t) = \text{Im}[\Delta\widehat{s}(t)s^*(t)/a_0^2]$. According to [7], when carrying ALF analysis one should neglect all terms of order higher than one in $\Delta\widehat{\omega}(t)$, $\Delta\widehat{\alpha}(t)$, $\Delta\widehat{s}(t)$, $w(t)$ and $v(t)$, including all cross-terms.

To derive recursion for $\Delta\widehat{x}(t)$, note that

$$\widehat{s}(t) = \psi(t) + \mu\varepsilon(t)$$
$$\varepsilon(t) = s(t) + v(t) - \psi(t) \qquad (46)$$

where

$$\psi(t) = e^{j[\widehat{\omega}(t-1)+\widehat{\alpha}(t-1)]}\widehat{s}(t-1).$$

Therefore

$$\widehat{s}(t) = \lambda\psi(t) + \mu s(t) + \mu v(t)$$

where $\lambda = 1 - \mu$. This leads to

$$\Delta\widehat{s}(t) = \lambda s(t) - \lambda\psi(t) - \mu v(t).$$

Observe that

$$\psi(t) = e^{j\omega(t)}e^{-j\Delta\widehat{\omega}(t-1)}e^{-j\Delta\widehat{\alpha}(t-1)}[s(t-1) - \Delta\widehat{s}(t-1)].$$

Using the approximations $e^{-j\Delta\widehat{\omega}(t-1)} \cong 1 - j\Delta\widehat{\omega}(t-1)$ and $e^{-j\Delta\widehat{\alpha}(t-1)} \cong 1 - j\Delta\widehat{\alpha}(t-1)$, that hold for small frequency and frequency rate errors, respectively, and applying ALF rules, one arrives at

$$\psi(t) \cong s(t) - e^{j\omega(t)}\Delta\widehat{s}(t-1)$$
$$- js(t)[\Delta\widehat{\omega}(t-1) + \Delta\widehat{\alpha}(t-1)]. \qquad (47)$$

This leads to

$$\Delta\widehat{s}(t) \cong \lambda e^{j\omega(t)}\Delta\widehat{s}(t-1)$$
$$+ j\lambda s(t)[\Delta\widehat{\omega}(t-1) + \Delta\widehat{\alpha}(t-1)] - \mu v(t)$$

or equivalently

$$\Delta\widehat{s}(t)s^*(t) \cong \lambda\Delta\widehat{s}(t-1)s^*(t-1)$$
$$+ j\lambda a_0^2[\Delta\widehat{\omega}(t-1) + \Delta\widehat{\alpha}(t-1)] - \mu v(t)s^*(t).$$

Dividing the last recursion sidewise by $a_0^2$, and taking imaginary parts, one arrives at

$$\Delta\widehat{x}(t) \cong \lambda\Delta\widehat{x}(t-1) + \lambda\Delta\widehat{\omega}(t-1)$$
$$+ \Delta\widehat{\alpha}(t-1)] + \mu e(t). \qquad (48)$$

To derive recursions for $\Delta\widehat{\omega}(t)$ and $\Delta\widehat{\alpha}(t)$, note that in the tracking mode it holds that

$$\delta(t) = \text{Im}\left[\frac{\varepsilon(t)}{\psi(t)}\right] \cong \frac{\text{Im}[\varepsilon(t)\psi^*(t)]}{a_0^2}.$$

Combining (46) with (47), and applying the ALF rules, one arrives at the following approximation

$$\varepsilon(t)\psi^*(t) \cong |s(t)|^2 - s(t-1)\Delta\widehat{s}^*(t-1)$$
$$+ j|s(t)|^2[\Delta\widehat{\omega}(t-1) + \Delta\widehat{\alpha}(t-1)]$$
$$+ v(t)s^*(t) - |\psi(t)|^2$$

leading to

$$\delta(t) \cong \Delta\widehat{x}(t-1) + \Delta\widehat{\omega}(t-1) + \Delta\widehat{\alpha}(t-1) - e(t).$$

Note that

$$\Delta\widehat{\alpha}(t) = \Delta\widehat{\alpha}(t-1) + w(t) - \gamma_\alpha\delta(t)$$
$$\Delta\widehat{\omega}(t) = \Delta\widehat{\omega}(t-1) + \Delta\widehat{\alpha}(t-1) - \gamma_\omega\delta(t).$$

Combining the last three equations, one arrives at

$$\Delta\widehat{\alpha}(t) \cong (1 - \gamma_\alpha)\Delta\widehat{\alpha}(t-1) + w(t) + \gamma_\alpha e(t)$$
$$- \gamma_\alpha\Delta\widehat{\omega}(t-1) - \gamma_\alpha\Delta\widehat{x}(t-1) \qquad (49)$$

$$\Delta\widehat{\omega}(t) \cong (1 - \gamma_\omega)\Delta\widehat{\omega}(t-1) + (1 - \gamma_\omega)\Delta\widehat{\alpha}(t-1)$$
$$+ \gamma_\omega e(t) - \gamma_\omega\Delta\widehat{x}(t-1) \qquad (50)$$

Finally, solving the set of linear equations (48), (49) and (50) for $\Delta\widehat{\omega}(t)$ and $\Delta\widehat{\alpha}(t)$, one arrives at (5) and (6), respectively.

## APPENDIX II

### Computation of Lower Tracking/Smoothing Bounds

In this appendix, we will derive expressions for theoretical upper bounds that limit tracking/smoothing capabilities of *any* causal/noncausal frequency and frequency rate estimation algorithms applied to signals with quasi-linear frequency changes. The corresponding lower tracking bounds (LTB) and lower smoothing bounds (LSB) belong to the class of posterior (or Bayesian) Cramér-Rao bounds, applicable to signals/systems with random parameters[3].

Denote by $\mathbf{y}$ the vector of measurements and let $\widehat{\boldsymbol{\theta}}(\mathbf{y})$ be an estimator of a real-valued random parameter vector $\boldsymbol{\theta}$. Then, under weak regularity conditions, one can show that [13]

$$\text{E}[(\widehat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta})(\widehat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta})^{\text{T}}] \geq \mathbf{J}^{-1}$$

where

$$\mathbf{J} = -\text{E}\left[\frac{\partial^2 \log p(\mathbf{y}, \boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\text{T}}}\right]$$
$$= \text{E}\left[\frac{\partial \log p(\mathbf{y}, \boldsymbol{\theta})}{\partial\boldsymbol{\theta}} \frac{\partial \log p(\mathbf{y}, \boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{\text{T}}}\right]$$

and $p(\mathbf{y}, \boldsymbol{\theta}) = p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})$ is the joint probability density function of the pair $(\mathbf{y}, \boldsymbol{\theta})$.

[3]When the estimated quantities are stochastic variables, rather than unknown deterministic constants, the classical Cramér-Rao inequality does not apply.

In the case considered, let $\boldsymbol{\theta}_t = [\alpha(1), \alpha(2), \ldots, \alpha(t)]^{\mathrm{T}}$ and $\mathbf{y}_t = [y(1), y(2), \ldots, y(t)]^{\mathrm{T}}$. To simplify our analysis (without restricting its generality), we will assume, in addition to (A1)–(A4), that $a_0 = |s(t)|$ is a known deterministic quantity, that $\alpha(0)$ is uniformly distributed over $[\alpha_{\min}, \alpha_{\max}]$, and that $\omega(0) = 0$, $s(0) = a_0$.

Since, under the assumptions listed above, the vector $\boldsymbol{\theta}_t$ fully determines $s(k)$, $k = 1, \ldots, t$

$$s(k) = a_0 e^{j\varphi(k)}$$

$$\varphi(k) = \sum_{n=1}^{k} \omega(n) = \sum_{n=1}^{k} \sum_{m=1}^{n-1} \alpha(m) \qquad (51)$$

one can write

$$\log p(\mathbf{y}_t|\boldsymbol{\theta}_t) = \log p[\mathbf{y}_t|s(1), s(2), \ldots, s(t)]$$

$$= \beta_1 - \frac{1}{\sigma_v^2} \sum_{k=1}^{t} |v(k)|^2$$

$$= \beta_1 - \frac{1}{\sigma_v^2} \sum_{k=1}^{t} |y(k) - s(k)|^2 \qquad (52)$$

where $\beta_1$ is a constant independent of $\boldsymbol{\theta}_t$.

Similarly

$$\log p(\boldsymbol{\theta}_t) = \log p[\alpha(1), \alpha(2) - \alpha(1), \ldots, \alpha(t) - \alpha(t-1)]$$

$$= \beta_\alpha + \beta_2 - \frac{1}{2\sigma_w^2} \sum_{k=2}^{t} w^2(k)$$

$$= \beta_\alpha + \beta_2 - \frac{1}{2\sigma_w^2} \sum_{k=2}^{t} [\alpha(k) - \alpha(k-1)]^2 \qquad (53)$$

where, again, $\beta_\alpha = \log[1/(\alpha_{\max} - \alpha_{\min})]$ and $\beta_2$ are constants independent of $\boldsymbol{\theta}_t$.

Differentiating (52) with respect to $\alpha(i)$, one obtains

$$\frac{\partial \log p(\mathbf{y}_t|\boldsymbol{\theta}_t)}{\partial \alpha(m)} = \frac{2}{\sigma_v^2} \sum_{k=1}^{t} \mathrm{Re} \left\{ [y(k) - s(k)]^* \frac{\partial s(k)}{\partial \alpha(m)} \right\}$$

$$= \frac{2}{\sigma_v^2} \sum_{k=1}^{t} \mathrm{Re} \left\{ v^*(k) \frac{\partial s(k)}{\partial \alpha(m)} \right\}$$

Using (51), one arrives at

$$\frac{\partial s(k)}{\partial \alpha(m)} = j s(k) \sum_{l=1}^{k} \sum_{i=1}^{l-1} \delta_{m,i}$$

where $\delta_{m,i}$ is the Kronecker delta

$$\delta_{m,i} = \begin{cases} 0 & \text{for} \quad m \neq i \\ 1 & \text{for} \quad m = i \end{cases}.$$

This leads to

$$\frac{\partial s(k)}{\partial \alpha(m)} = j s(k) \max(k - m, 0)$$

and

$$\frac{\partial \log p(\mathbf{y}_t|\boldsymbol{\theta}_t)}{\partial \alpha(m)} = -\frac{2}{\sigma_v^2} \sum_{k=1}^{t} \mathrm{Re} \left[ j v^*(k) s(k) \max(k - m, 0) \right]$$

Hence, after averaging, one obtains

$$\mathrm{E} \left[ \frac{\partial \log p(\mathbf{y}_t|\boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}_t} \frac{\partial \log p(\mathbf{y}_t|\boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}_t^{\mathrm{T}}} \right] = \frac{2a_0^2}{\sigma_v^2} \mathbf{A}_t \qquad (54)$$

where

$$[\mathbf{A}_t]_{mn} = \sum_{k=1}^{t} \max(k - m, 0) \cdot \max(k - n, 0)$$

In an analogous way we will compute derivative of the prior density (53)

$$\frac{\log p(\boldsymbol{\theta}_t)}{\alpha(m)} = -\frac{1}{2\sigma_w^2} \sum_{k=2}^{t} \frac{\partial}{\partial \alpha(m)} [\alpha(k) - \alpha(k-1)]^2$$

$$= \frac{1}{\sigma_w^2} \begin{cases} \alpha(2) - \alpha(1) & \text{for} \quad m = 1 \\ [\alpha(m+1) - \alpha(m)] & \\ -[\alpha(m) - \alpha(m-1)] & \text{for} \quad 1 < m < t \\ -[\alpha(t) - \alpha(t-1)] & \text{for} \quad m = t \end{cases}$$

$$= \frac{1}{\sigma_w^2} \begin{cases} w(2) & \text{for} \quad m = 1 \\ w(m+1) - w(m) & \text{for} \quad 1 < m < t \\ -w(t) & \text{for} \quad m = t \end{cases}$$

which leads to

$$\mathrm{E} \left[ \frac{\partial \log p(\boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}_t} \frac{\partial \log p(\boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}_t^{\mathrm{T}}} \right] = \frac{1}{\sigma_w^2} \mathbf{B}_t \qquad (55)$$

where

$$\mathbf{B}_t = \begin{bmatrix} 1 & -1 & 0 & 0 & \ldots & 0 \\ -1 & 2 & -1 & 0 & \ldots & 0 \\ 0 & -1 & 2 & -1 & \ldots & 0 \\ & & & \ddots & & \\ 0 & \ldots & 0 & -1 & 2 & -1 \\ 0 & \ldots & 0 & 0 & -1 & 1 \end{bmatrix}.$$

Combining (54) with (55), one obtains

$$\mathbf{J}_t = \frac{2a_0^2}{\sigma_v^2} \mathbf{A}_t + \frac{1}{\sigma_w^2} \mathbf{B}_t = \frac{1}{\sigma_w^2} [2\kappa \mathbf{A}_t + \mathbf{B}_t].$$

The asymptotic (steady-state) bounds on accuracy of frequency and frequency rate estimates can be obtained from

$$\mathrm{LTB}_\omega = \lim_{t \mapsto \infty} \inf_{\widehat{\omega}(\cdot)} \mathrm{E}\{ [\omega(t) - \widehat{\omega}(t)]^2 \}$$

$$= \lim_{t \mapsto \infty} \mathbf{b}_t^{\mathrm{T}} \mathbf{J}_t^{-1} \mathbf{b}_t$$

$$\mathrm{LTB}_\alpha = \lim_{t \mapsto \infty} \inf_{\widehat{\alpha}(\cdot)} \mathrm{E}\{ [\alpha(t) - \widehat{\alpha}(t)]^2 \}$$

$$= \lim_{t \mapsto \infty} \left[ \mathbf{J}_t^{-1} \right]_{tt}$$

where $\mathbf{b}_t^{\mathrm{T}} = [\mathbf{1}_{t-1}^{\mathrm{T}}, 0]$, and $\mathbf{1}_t$ denotes the vector of ones of length $t$. The analogous expressions for lower smoothing bounds read

$$\mathrm{LSB}_\omega = \lim_{t \mapsto \infty} \inf_{\widetilde{\omega}(\cdot)} \mathrm{E}\{ [\omega(t) - \widetilde{\omega}(t)]^2 \}$$

$$= \lim_{t \mapsto \infty} \mathbf{c}_t^{\mathrm{T}} \mathbf{J}_{2t}^{-1} \mathbf{c}_t$$

$$\mathrm{LSB}_\alpha = \lim_{t \mapsto \infty} \ \inf_{\widetilde{\alpha}(\cdot)} \mathrm{E}\{\,[\alpha(t) - \widetilde{\alpha}(t)]^2\}$$

$$= \lim_{t \mapsto \infty} \left[\mathbf{J}_{2t}^{-1}\right]_{tt}$$

where $\mathbf{c}_t^{\mathrm{T}} = [\mathbf{1}_{t-1}^{\mathrm{T}}, \mathbf{0}_{t+1}^{\mathrm{T}}]$ and $\mathbf{0}_t$ denotes the vector of zeros of length $t$. The values in Table I were computed numerically for $t$ ranging from 100 to 600 (the convergence is slower for smaller values of $\kappa$).

## REFERENCES

[1] N. Thakor and Y. Zhou, "Application of adaptive filtering to ECG analysis: noise cancellation and arrhytmia detection," *IEEE Trans. on Biomed. Eng.*, vol. 38, pp. 785—794, 1991.

[2] Y. Z. W.K. Ma and F. 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.

[3] 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.

[4] 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.

[5] ——, "Generalized adaptive notch smoothers for real-valued signals and systems," *IEEE Trans. Signal Process.*, vol. 56, pp. 125—133, 2008.

[6] M. Niedźwiecki, "Generalized adaptive notch smoothing revisited," *IEEE Trans. Signal Process.*, vol. 58, pp. 1565—1576, 2010.

[7] P. Tichavský and P. Händel, "Two algorithms for adaptive retrieval of slowly time-varying multiple cisoids in noise," *IEEE Trans. on Signal Process.*, vol. 43, pp. 1116—1127, May 1995.

[8] M. Jury, *Theory and Application of the Z-transform Method.* New York: Wiley, 1964.

[9] P. Tichavský and P. Händel, "Recursive estimation of frequencies and frequency rates of multiple cisoids in noise," *Signal Processing*, vol. 58, no. 2, pp. 117–129, April 1997.

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

[11] B.D.O. Anderson and J.B. Moore, *Optimal Filtering.* New York: Prentice-Hall, 1979.

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

[13] H. van Trees, *Detection, Estimation and Modulation Theory.* New York: Wiley, 1968.

**Maciej Niedźwiecki** was born in Poznań, Poland in 1953. He received the M.Sc. and Ph.D. degrees from the Technical University of Gdańsk, Gdańsk, Poland and the Dr.Hab. (D.Sc.) degree from the Technical University of Warsaw, Warsaw, Poland, in 1977, 1981 and 1991, respectively. He spent three years as a Research Fellow with the Department of Systems Engineering, Australian National University, 1986-1989. In 1990 - 1993 he served as a Vice Chairman of Technical Committee on Theory of the International Federation of Automatic Control (IFAC). He is currently Associate Editor for IEEE Transactions on Signal Processing, a member of the IFAC committees on Modeling, Identification and Signal Processing and on Large Scale Complex Systems, and a member of the Automatic Control and Robotics Committee of the Polish Academy of Sciences (PAN). He is the author of the book *Identification of Time-varying Processes* (Wiley, 2000).

He works as 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, statistical signal processing and adaptive systems.

**Michał Meller** received the M.Sc. and Ph.D. degrees in automatic control from the Gdańsk University of Technology, Gdańsk, Poland, in 2007 and 2010, respectively. Since 2007 he has been working in Telecommunications Research Institute S.A., Gdańsk Division, Department of Signal and Information Processing. In 2010 he also joined the Department of Automatic control at the Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Computer Science. His professional interests include signal processing and adaptive systems.