A New Approach to Active Noise and Vibration Control – Part II: the Unknown Frequency Case

Maciej Niedźwiecki, *Member, IEEE* and Michał Meller, *Student member, IEEE*

Abstract—**This paper presents a new approach to rejection of complex-valued sinusoidal disturbances acting at the output of a discrete-time stable linear plant with unknown and possibly time-varying dynamics. It is assumed that both the instantaneous frequency of the sinusoidal disturbance and its amplitude may be slowly varying with time and that the output signal is contaminated with wideband measurement noise. It is not assumed that a reference signal, correlated with the disturbance, is available. The proposed disturbance rejection algorithm is an extension of the algorithm derived for the constant-known-frequency case, described in Part I of this paper.**

Index Terms—**Adaptive filtering, system identification, disturbance rejection.**

I. INTRODUCTION

CONSIDER the problem of eliminating a narrowband disturbance acting at the output of a discrete-time system governed by

$$y(t) = K_\mathrm{p}(q^{-1})u(t-1) + d(t) + v(t) \tag{1}$$

where $y(t)$ denotes the corrupted complex-valued system output, $t = \ldots, -1, 0, 1, \ldots$ denotes normalized (dimensionless) time, $K_\mathrm{p}(q^{-1})$ denotes the unknown transfer function of a linear stable single-input single-output plant, q^{-1} is the backward shift operator, $d(t)$ denotes a nonstationary narrowband disturbance, $v(t)$ is a wideband measurement noise, and $u(t)$ denotes the input signal.

We will look for the minimum-variance feedback controller, i.e., for a control rule that minimizes the system output in the mean-squared sense:

$$\mathrm{E}[\,|\,y(t)\,|^2\,] \longmapsto \min .$$

We assume that no reference signal (usually obtained from a transducer placed in the vicinity of the source of vibration) is available. This makes the task of disturbance rejection more difficult, as application of the feedforward compensation technique is, in such a case, not possible.

Practically important, the problem of vibration control has attracted a great deal of attention in recent years [1]–[3], solved by many authors using different approaches, such as filtered-X LMS (FXLMS) compensation, internal model principle, or phase-locked loop – for references see [4]. Methodological

aspects aside, the proposed solutions differ in the assumptions made about the controlled plant and sinusoidal disturbance. The most important design premises are those concerning:

- Known/unknown/time-varying plant dynamics
- Known/unknown/time-varying amplitude of the disturbance
- Known/unknown/time-varying frequency of the disturbance

Only a few algorithms, out of a large number of proposed ones, are capable of dealing with the most general case: a *nonstationary* sinusoidal disturbance, with time-varying amplitude and frequency, corrupting the output of a *nonstationary* linear plant — the adaptive FXLMS algorithm being one of the noticeable exceptions.

To perform satisfactorily, all available methods require tuning of several design parameters. For example, in the case of the adaptive FXLMS algorithm, one should adjust two step-size coefficients (determining the speed of adaptation of the control and system identification loops, respectively), the variance of auxiliary noise (injected into the control loop to facilitate plant identifiability), and at least two adaptation gains of the adaptive line enhancer (used to "extract" the reference signal from the output signal). Since different settings may be required under different operating conditions, tuning those design parameters may be a difficult task, especially in a nonstationary environment. To the best of our knowledge, the algorithm derived in this paper is the first *self-optimizing* narrowband noise canceller capable of eliminating disturbances in their most general form, which incorporates both amplitude and frequency changes. The proposed solution is an extension of the self-optimizing narrowband noise canceling (SONIC) algorithm, described in the first part of the paper [4] and derived under the assumption that frequency of the narrowband disturbance is constant and known. The extended algorithm will be called xSONIC. The new approach compares favorably, both in terms of control quality and computational complexity, with the adaptive FXLMS scheme.

II. ASSUMPTIONS

We assume that the measurement noise in (1) obeys

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

In its most general form, the nostationary narrowband disturbance will be modeled using the following equations

$$\begin{aligned} d(t) &= e^{j\omega(t-1)}d(t-1) + \widetilde{e}(t) \\ \omega(t) &= \omega(t-1) + w(t) \end{aligned} \tag{2}$$

where

(A2) $\{\widetilde{e}(t)\}$ and $\{w(t)\}$, mutually independent and independent of $\{v(t)\}$, are zero-mean white sequences with variances σ_e^2 and σ_w^2, respectively, $\{\widetilde{e}(t)\}$ is complex-valued circular, and $\{w(t)\}$ is real-valued.

When $\sigma_w^2 = 0$, which entails $\omega(t) = \omega(0) = \omega_0$, the disturbance can be expressed in the more explicit form

$$d(t) = a(t)e^{j\omega_0 t} \qquad (3)$$

where the complex-valued "amplitude" $a(t)$ obeys the random-walk model

$$a(t) = a(t-1) + e(t), \quad e(t) = e^{-j\omega_0 t}\widetilde{e}(t) \qquad (4)$$

(note that the sequence $\{e(t)\}$ is circular white with variance σ_e^2). Hence, in this special case, $d(t)$ is a complex-valued sinusoidal signal (cisoid) with constant frequency and randomly drifting amplitude. Disturbances of this form were considered in Part I of this paper [4].

When $\sigma_w^2 \neq 0$, the instantaneous frequency drifts according to the random-walk model. Therefore, the model (2) describes a narrowband signal subject to random variations of both amplitude and phase.

Even though the proposed adaptive control algorithm is capable of eliminating the disturbance (2) in its most general form ($\sigma_e^2 \neq 0, \sigma_w^2 \neq 0$), the majority of our analytical results will be derived for the somewhat simpler case

$$d(t) = e^{j\omega(t-1)}d(t-1)$$
$$\omega(t) = \omega(t-1) + w(t) \qquad (5)$$

which can be obtained from (2) by setting $\sigma_e^2 = 0$. Note that the signal governed by (5) is a constant-modulus cisoid ($|d(t)|$ constant) with randomly drifting frequency.

To arrive at analytical results, we will assume that the plant is unknown, time-invariant, and stable, and that it has nonzero gain over the entire frequency range:

(A3) $K_p(q^{-1}) = \sum_{i=0}^{\infty} k_i q^{-i}, \ \sum_{i=0}^{\infty} |k_i| < \infty,$
$K_p(e^{-j\omega}) \neq 0, \ \forall \omega \in (-\pi, \pi].$

Later on, in Section VII, we will show that the proposed controller also works satisfactorily in the case where the plant dynamics change over time.

III. Basic Control Algorithm and its Properties

A. Control Algorithm

The control algorithm that will serve as a basis for our further considerations is an extended version of the algorithm proposed in [4] for elimination of narrowband disturbances described by (3), i.e., with a constant-known frequency ω_0. The extended algorithm can be summarized as follows:

$$\widehat{d}(t+1\,|\,t) = e^{j\widehat{\omega}(t\,|\,t-1)}[\widehat{d}(t\,|\,t-1) + \mu y(t)]$$

$$\widehat{\omega}(t+1\,|\,t) = (1-\eta)\widehat{\omega}(t\,|\,t-1) + \eta \, \mathrm{Arg}\left[\frac{\widehat{d}(t+1\,|\,t)}{\widehat{d}(t\,|\,t-1)}\right]$$

$$u(t) = -\frac{\widehat{d}(t+1\,|\,t)}{k_n} \qquad (6)$$

and incorporates frequency tracking. In the above algorithm, μ, $0 < \mu \ll 1$, denotes a small gain that controls the rate of amplitude adaptation, and η, $0 < \eta \ll 1$, is another gain that controls the rate of frequency adaptation, k_n denotes the nominal (assumed) gain, and $\mathrm{Arg}[x] \in (-\pi, \pi]$ denotes principal argument of a complex number x. The frequency estimation loop in (6) is identical with that proposed in [5].

Remark: The nominal plant gain, adopted in (6), is a time-invariant quantity. Another possible solution, which incorporates the "instantaneous" gain of the nominal plant model $K_n(e^{-j\omega})$, has the form

$$u(t) = -\frac{\widehat{d}(t+1\,|\,t)}{k_n(t)}, \quad k_n(t) = K_n(e^{-j\widehat{\omega}(t+1\,|\,t)}). \qquad (7)$$

B. Tracking Properties

To derive analytical results, we will assume that the disturbance is governed by (5), i.e., it is a constant-modulus cisoid[1] with unknown magnitude $a = |d(t)|$ and randomly drifting frequency $\omega(t)$. Furthermore, we will assume that the (unknown) plant is time-invariant. Our approach will be based on averaging. Consider a local analysis window $\mathcal{T} = [t_1, t_2]$, covering $T = t_2 - t_1 + 1$ consecutive time instants ($T \gg 2\pi/\omega(t), \forall t \in \mathcal{T}$). If the transfer function of the plant $K_p(e^{-j\omega})$ is a smooth function of ω, and if the instantaneous frequency of the disturbance changes sufficiently slowly with time, the true response of the plant to the narrowband excitation $u(t)$ can be approximated as

$$K_p(q^{-1})u(t-1) \cong k_{\mathcal{T}} u(t-1), \ t \in \mathcal{T}$$

where $k_{\mathcal{T}} = \sum_{t \in \mathcal{T}} K_p(e^{-j\omega(t)})/T$ denotes the average plant gain over the interval T. Using this approximation, one can express plant output in the form

$$y(t) \cong d(t) - \beta\widehat{d}(t\,|\,t-1) + v(t) \, , \ t \in \mathcal{T} \qquad (8)$$

where $\beta = k_{\mathcal{T}}/k_n$ — the ratio of the average plant gain to the nominal (assumed) gain — denotes the local modeling error. In our local analysis, β will be regarded as a time-invariant quantity.

Denote the cancellation error by $\Delta\widehat{d}(t) = d(t) - \beta\widehat{d}(t\,|\,t-1)$, and the one-step-ahead frequency prediction error by $\Delta\widehat{\omega}(t) = \omega(t) - \widehat{\omega}(t\,|\,t-1)$. To establish the dependence of $\Delta\widehat{d}(t)$ and $\Delta\widehat{\omega}(t)$ on $v(t)$ and $w(t)$, we will employ the approximating linear filter (ALF) technique, proposed by Tichavský and Händel [5], for the purpose of analyzing adaptive notch filters. Using this approach, one arrives at the following approximations (see Appendix):

$$\Delta\widehat{x}(t) = (1-\mu\beta)\Delta\widehat{x}(t) + ja^2\Delta\widehat{\omega}(t-1) - \mu\beta z(t-1)$$
$$\Delta\widehat{\omega}(t+1) = \Delta\widehat{\omega}(t) - \frac{\eta}{a^2}\mathrm{Im}[\mu\beta\Delta\widehat{x}(t)] - \frac{\eta}{a^2}\mathrm{Im}[\mu\beta z(t)]$$
$$+ w(t+1) \qquad (9)$$

where $\Delta\widehat{x}(t) = \Delta\widehat{d}(t)d^*(t)$ and $z(t) = v(t)d^*(t)$. Note that

$$\mathrm{E}[|\Delta\widehat{x}(t)|^2] = a^2\mathrm{E}[|\Delta\widehat{d}(t)|^2] \qquad (10)$$

---

[1]From a qualitative viewpoint, this is not a serious limitation, as all information about the frequency of the analyzed signal is contained in its phase – see e.g. [6].

and that $\{z(t)\}$ is a circular white noise with variance $\sigma_z^2 = \mathrm{E}[|z(t)|^2] = a^2\sigma_v^2$.

IV. TRACKING ANALYSIS IN THE ABSENCE OF MODELING ERRORS

Based on (9), we will study the tracking properties of (6) in the absence of modeling errors ($\beta = 1$). Let $\Delta\widehat{x}_{\mathrm{R}}(t) = \mathrm{Re}[\Delta\widehat{x}(t)]$, $\Delta\widehat{x}_{\mathrm{I}}(t) = \mathrm{Im}[\Delta\widehat{x}(t)]$, $z_{\mathrm{R}}(t) = \mathrm{Re}[z(t)]$, and $z_{\mathrm{I}}(t) = \mathrm{Im}[z(t)]$. Since, for the time being, we have assumed that the adaptation gain μ is real-valued, the recursive relationship (9) can be written in the form

$$\Delta\widehat{x}_{\mathrm{R}}(t) = (1-\mu)\Delta\widehat{x}_{\mathrm{R}}(t-1) - \mu z_{\mathrm{R}}(t-1)$$
$$\Delta\widehat{x}_{\mathrm{I}}(t) = (1-\mu)\Delta\widehat{x}_{\mathrm{I}}(t-1) + a^2\Delta\widehat{\omega}(t-1) - \mu z_{\mathrm{I}}(t-1)$$
$$\Delta\widehat{\omega}(t+1) = \Delta\widehat{\omega}(t) - \frac{\eta\mu}{a^2}\Delta\widehat{x}_{\mathrm{I}}(t) - \frac{\eta\mu}{a^2} z_{\mathrm{I}}(t)$$
$$+ w(t+1). \qquad (11)$$

Solving the linear equations (11) with respect to $\Delta\widehat{x}_{\mathrm{R}}(t)$, $\Delta\widehat{x}_{\mathrm{I}}(t)$, and $\Delta\widehat{\omega}(t)$, one obtains

$$\Delta\widehat{x}_{\mathrm{R}}(t) = F(q^{-1})z_{\mathrm{R}}(t)$$
$$\Delta\widehat{x}_{\mathrm{I}}(t) = G_1(q^{-1})z_{\mathrm{I}}(t) + G_2(q^{-1})w(t)$$
$$\Delta\widehat{\omega}(t) = H_1(q^{-1})z_{\mathrm{I}}(t) + H_2(q^{-1})w(t) \qquad (12)$$

where

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

$$G_1(q^{-1}) = -\frac{\mu[1 + (\eta-1)q^{-1}]q^{-1}}{1 - (1+\lambda)q^{-1} + (\lambda + \eta\mu)q^{-2}}$$

$$G_2(q^{-1}) = \frac{a^2}{1 - (1+\lambda)q^{-1} + (\lambda + \eta\mu)q^{-2}}$$

$$H_1(q^{-1}) = -\frac{\eta\mu(1 - q^{-1})q^{-1}}{a^2[1 - (1+\lambda)q^{-1} + (\lambda + \eta\mu)q^{-2}]}$$

$$H_2(q^{-1}) = \frac{1 - \lambda q^{-1}}{1 - (1+\lambda)q^{-1} + (\lambda + \eta\mu)q^{-2}}$$

and $\lambda = 1 - \mu$. All filters are asymptotically stable for any μ and η in the interval (0,1).

A. Frequency Tracking

Since under (A1) and (A2) the process $\{z_{\mathrm{I}}(t)\}$ is orthogonal to $\{w(t)\}$, one arrives at

$$\mathrm{E}\{[\Delta\widehat{\omega}(t)]^2\} = I[H_1(z^{-1})]\mathrm{E}[z_{\mathrm{I}}^2(t)] + I[H_2(z^{-1})]\mathrm{E}[w^2(t)]$$

where

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

is an integral evaluated along the unit circle in the z-plane, and $X(z^{-1})$ denotes any stable proper rational transfer function. By means of residue calculus (see e.g. [7]) one obtains

$$I[H_1(z^{-1})] = \frac{2\eta^2\mu}{a^4(1-\eta)(2 + 2\lambda + \eta\mu)} \cong \frac{\eta^2\mu}{2a^4}$$

$$I[H_2(z^{-1})] = \frac{\mu^2(1+\lambda) + \eta\mu(1+\lambda^2)}{\eta\mu^2(1-\eta)(2 + 2\lambda + \eta\mu)} \cong \frac{1}{2\eta} + \frac{1}{2\mu}$$

where the approximations hold for sufficiently small values of μ and η. Since $\mathrm{E}[z_{\mathrm{I}}^2(t)] = a^2\sigma_v^2/2$, one arrives at

$$\mathrm{E}\{[\Delta\widehat{\omega}(t)]^2\} \cong \frac{\eta^2\mu}{4a^2}\sigma_v^2 + \left[\frac{1}{2\eta} + \frac{1}{2\mu}\right]\sigma_w^2. \qquad (13)$$

Denote by μ_ω and η_ω the values of μ and η that minimize the mean-squared frequency estimation error (13). It is straightforward to check that

$$\mu_\omega = \sqrt[4]{8\xi}, \quad \eta_\omega = \sqrt[4]{\xi/2} \qquad (14)$$

where

$$\xi = \frac{a^2\sigma_w^2}{\sigma_v^2}$$

is a scalar coefficient that can be regarded as a measure of nonstationarity of a signal governed by (5).

Under the Gaussian assumptions imposed on $\{v(t)\}$ and $\{w(t)\}$, the lower frequency tracking bound [called the posterior Cramér-Rao bound (PCRB)] was established in [8]

$$\mathrm{PCRB} \cong \sigma_w^2\sqrt[4]{2\xi^{-1}}. \qquad (15)$$

Note that, after combining (13) with (14), one obtains

$$\mathrm{E}\{[\Delta\widehat{\omega}(t)]^2|\,\mu_\omega, \eta_\omega\} \cong \sigma_w^2\sqrt[4]{2\xi^{-1}} \qquad (16)$$

which is identical to (15). Hence, despite its simplicity, the optimally tuned algorithm (6) is a statistically efficient scheme for tracking randomly-drifting instantaneous frequency. In practical terms, this means that there seems to be no incentive to replace the simple gradient frequency update, incorporated in (6), with a more elaborate frequency estimation mechanism.

B. Disturbance Cancellation/Tracking

Of course, our main interest lies in minimization of the mean-squared cancellation error. In order to evaluate $\mathrm{E}[|\Delta\widehat{d}(t)|^2]$, we will exploit (10). Due to the orthogonality of $z_{\mathrm{R}}(t)$ and $z_{\mathrm{I}}(t)$, it holds that

$$\mathrm{E}[|\Delta\widehat{x}(t)|^2] = \mathrm{E}\{[\Delta\widehat{x}_{\mathrm{R}}(t)]^2\} + \mathrm{E}\{[\Delta\widehat{x}_{\mathrm{I}}(t)]^2\} \qquad (17)$$

where [cf. (12)]

$$\mathrm{E}\{[\Delta\widehat{x}_{\mathrm{R}}(t)]^2\} = I[F(z^{-1})]\mathrm{E}[z_{\mathrm{R}}^2(t)]$$
$$\mathrm{E}\{[\Delta\widehat{x}_{\mathrm{I}}(t)]^2\} = I[G_1(z^{-1})]\mathrm{E}[z_{\mathrm{I}}^2(t)]$$
$$+ I[G_2(z^{-1})]\mathrm{E}[w^2(t)]. \qquad (18)$$

Since

$$I[F(z^{-1})] = \frac{\mu}{1+\lambda} \cong \frac{\mu}{2}$$

$$I[G_1(z^{-1})] = \frac{(2 - 2\eta + \eta^2)\mu + \eta(1+\lambda)}{(1-\eta)(2 + 2\lambda + \eta\mu)} \cong \frac{\mu}{2} + \frac{\eta}{2}$$

$$I[G_2(z^{-1})] = \frac{a^4(1 + \lambda + \eta\mu)}{\eta\mu^2(1-\eta)(2 + 2\lambda + \eta\mu)} \cong \frac{a^4}{2\eta\mu^2}$$

and $\mathrm{E}[z_{\mathrm{R}}^2(t)] = \mathrm{E}[z_{\mathrm{I}}^2(t)] = a^2\sigma_v^2/2$, after combining (10), (17), and (18), one arrives at

$$\mathrm{E}[|\Delta\widehat{d}(t)|^2] \cong \frac{a^2}{4}(\mu + \eta)\sigma_v^2 + \frac{a^4}{2\eta\mu^2}\sigma_w^2. \qquad (19)$$
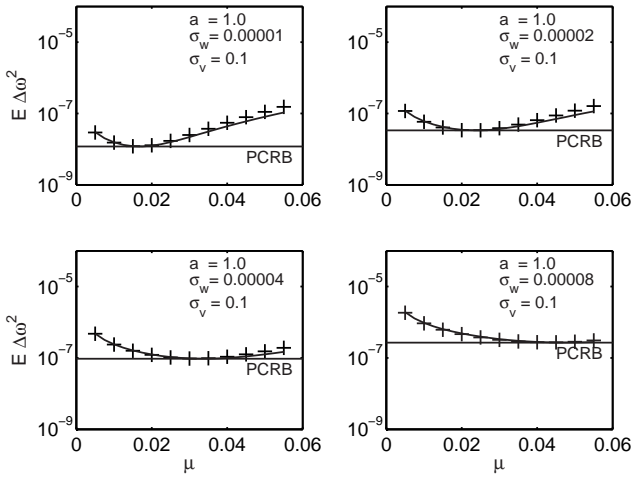
4

Fig. 1. Comparison of theoretical values of the mean-squared frequency tracking error, obtained using the steady-state plant approximation (solid line), with the experimental values (crosses).

Denote by μ_d and η_d the values of μ and η that minimize the mean-squared cancellation error (19). One can easily check that

$$\mu_d = \eta_d = \sqrt[4]{2\xi} \qquad (20)$$

and

$$\mathrm{E}[|\Delta\widehat{d}(t)|^2|\,\mu_d,\eta_d] \cong \sqrt[4]{2\xi}\,\sigma_v^2\ . \qquad (21)$$

Note that the settings that minimize the mean-squared cancellation error differ from those that minimize the mean-squared frequency tracking error established earlier.

C. Numerical Example

Since the ALF approach is heuristic (no strict mathematical analysis of the conditions of its applicability was presented in [5]), "theoretical" results, such as (13) or (19), should be treated with caution. A special simulation experiment was arranged to check how well the formula (13) fits the true error values. The simulated discrete-time plant (Guo-Bodson)

$$K_{\mathrm{p}}(z^{-1}) = \frac{0.0952}{1 - 0.9048z^{-1}} \qquad (22)$$

was adopted from [9], and corresponds to a continuous-time plant with transfer function $K_{\mathrm{p}}(s) = 1/(1 + 0.01s)$ sampled at the rate of 1 kHz.

To (nearly) eliminate modeling errors, the modified control rule (7) was used with $K_{\mathrm{n}}(\cdot) \equiv K_{\mathrm{p}}(\cdot)$, which requires perfect knowledge of the plant's frequency response (note, however, that the nominal gain in (7) is evaluated at the estimated frequency instead of the true frequency).

Simulations were carried on for $\sigma_v = 0.1$ and for four different rates of amplitude variation $\sigma_w \in \{0.00001, 0.00002, 0.00004, 0.00008\}$. To decrease the number of design degrees of freedom from 2 (μ, η) to 1 (η), the value of η was set to $\mu/2$ — such a choice was motivated by the fact that $\eta_\omega = \mu_\omega/2$.

For each (σ_w, μ) pair, the results, summarized in Fig. 1, were obtained by means of combined ensemble averaging

(50 different realizations of $\{v(t)\}$ and $\{w(t)\}$) and time averaging (50000 consecutive time-steps, after discarding the first 10000 samples, to ensure that the steady-state conditions are reached). In all cases, $\omega(0)$ was set to 0.1. Note the good agreement of theoretical expectations with simulation results for small values of μ. Discrepancies that occur for larger values of the adaptation gain can be attributed to the frozen gain approximation (8) and to errors introduced by the ALF analysis (linearization and neglecting higher-order terms). Following [10], one can argue that approximation errors should not compromise our statistical efficiency conclusion as long as tracking is "satisfactory", i.e., as long as the minimum attainable mean-squared cancellation error (often called excess prediction error) is much smaller than the variance of the measurement error. This leads to the following condition, expressed in terms of the rate of system nonstationarity: $\sqrt[4]{2\xi} \ll 1$, or equivalently, $\xi \le 10^{-4}$.

V. SELF-OPTIMIZING CANCELLATION SCHEME

Even though we have been assuming that the adaptation gain μ is a real-valued quantity, the derivation of approximating linear equations is not restricted to this case — equations (9) remain valid also for complex-valued gains $\mu \in \mathbb{C}$. When the true plant characteristics are not known, i.e., when $\beta = k_\mathcal{T}/k_\mathrm{n} \ne 1$, incorporation of a complex-valued gain has an obvious advantage, as it allows one to compensate modeling error. Actually, according to (9), when μ is chosen so that the condition $\mu\beta = \mu_0 > 0$ is met (which can be achieved provided that $\mathrm{Im}[\mu\beta] = 0$, i.e., $\mathrm{Arg}[\mu] = -\mathrm{Arg}[\beta]$), the control algorithm (6) with a complex-valued gain, used in in the presence of modeling errors ($\beta \ne 1$), should yield the same results as the same algorithm equipped with a real-valued adaptation gain μ_0 operated in the absence of modeling errors ($\beta = 1$). In particular, when μ is set to μ_ω/β and η is set to η_ω, the closed-loop system will guarantee statistically efficient frequency tracking. Similarly, when μ is set to μ_d/β and η is set to η_d, the mean-squared cancellation error will achieve its minimum value (21) — even though the assumed plant gain differs from the true value.

Since in practice β and ξ are unknown quantities, we will propose a special mechanism for automatic adjustment of μ and η. Generally, we would like to adjust both adaptation gains so as to minimize the mean-squared value of the output signal

$$\mathrm{E}[\,|\,y(t; \mu, \eta)|^2\,] \longmapsto \min .$$

Since $\mathrm{E}[\,|y(t)|^2\,] = \mathrm{E}[\,|\Delta\widehat{d}(t)|^2\,] + \sigma_v^2$, this is equivalent to minimization of the mean-squared cancellation error. To make the estimation scheme more flexible and to avoid problems with mixed optimization (joint optimization of a complex-valued gain μ and a real-valued gain η), we will design two separate loops for adjustment of μ and η, respectively. Both tuning procedures will be based on the recursive prediction error (RPE) optimization strategy.

A. Adjustment of μ

Consider the following local measure of fit, made up of exponentially weighted system outputs [the output signal $y(t)$

is regarded here as a function of μ]

$$V(t;\mu) = \sum_{\tau=1}^{t} (\rho_\mu)^{t-\tau} |y(\tau;\mu)|^2. \tag{23}$$

The forgetting constant ρ_μ $(0 < \rho_\mu < 1)$ determines the effective averaging range. To evaluate the estimate $\widehat{\mu}(t) = \arg\min_\mu V(t;\mu)$, we will use the RPE approach [11]

$$\widehat{\mu}(t) = \widehat{\mu}(t-1) - \frac{\frac{\partial}{\partial\mu^*} V(t;\widehat{\mu}(t-1))}{\frac{\partial^2}{\partial\mu^*\partial\mu} V(t;\widehat{\mu}(t-1))} \tag{24}$$

where

$$\frac{\partial}{\partial\mu} = \left[\frac{\partial}{\partial\mu_R} - j\frac{\partial}{\partial\mu_I}\right], \quad \frac{\partial}{\partial\mu^*} = \left[\frac{\partial}{\partial\mu_R} + j\frac{\partial}{\partial\mu_I}\right]$$

$$\mu_R = \mathrm{Re}[\mu], \quad \mu_I = \mathrm{Im}[\mu]$$

denote operations of symbolic differentiation used in the so-called Wirtinger calculus, applicable to nonanalytic functions, such as (23) — for references, see [4].

Since

$$\frac{\partial|y(\tau)|^2}{\partial\mu^*} = y(\tau)\left(\frac{\partial y(\tau)}{\partial\mu}\right)^* + \frac{\partial y(\tau)}{\partial\mu^*}y^*(\tau)$$

the gradient and Hessian that appear in (24) can be computed recursively using approximations typical of RPE

$$\frac{\partial V[t;\widehat{\mu}(t-1)]}{\partial\mu^*} = y[t;\widehat{\mu}(t-1)]\left(\frac{\partial y[t;\widehat{\mu}(t-1)]}{\partial\mu}\right)^*$$
$$+ \frac{\partial y[t;\widehat{\mu}(t-1)]}{\partial\mu^*} y^*[t;\widehat{\mu}(t-1)]$$

$$\frac{\partial^2 V[t;\widehat{\mu}(t-1)]}{\partial\mu^*\partial\mu} \cong \rho_\mu \frac{\partial^2 V[t;\widehat{\mu}(t-2)]}{\partial\mu^*\partial\mu}$$
$$+ \left|\frac{\partial y[t;\widehat{\mu}(t-1)]}{\partial\mu}\right|^2 + \left|\frac{\partial y[t;\widehat{\mu}(t-1)]}{\partial\mu^*}\right|^2. \tag{25}$$

Note that [cf. (8)]

$$\frac{\partial y(t)}{\partial\mu} = -\beta\frac{\partial\widehat{d}(t|t-1)}{\partial\mu}, \quad \frac{\partial y(t)}{\partial\mu^*} = -\beta\frac{\partial\widehat{d}(t|t-1)}{\partial\mu^*} \tag{26}$$

and [cf. (6)]

$$\frac{\partial\widehat{d}(t+1|t)}{\partial\mu} = j\frac{\partial\widehat{\omega}(t|t-1)}{\partial\mu}\widehat{d}(t+1|t)$$
$$+ e^{j\widehat{\omega}(t|t-1)}\left[\frac{\partial\widehat{d}(t|t-1)}{\partial\mu} + y(t) + \mu\frac{\partial y(t)}{\partial\mu}\right]$$

$$\frac{\partial\widehat{d}(t+1|t)}{\partial\mu^*} = j\frac{\partial\widehat{\omega}(t|t-1)}{\partial\mu^*}\widehat{d}(t+1|t)$$
$$+ e^{j\widehat{\omega}(t|t-1)}\left[\frac{\partial\widehat{d}(t|t-1)}{\partial\mu^*} + \mu\frac{\partial y(t)}{\partial\mu^*}\right] \tag{27}$$

where the last recursion stems from the fact that, in Wirtinger calculus, $\partial\mu/\partial\mu^* = \partial\mu^*/\partial\mu = 0$.

To evaluate $\partial\widehat{\omega}(t+1|t)/\partial\mu$ and $\partial\widehat{\omega}(t+1|t)/\partial\mu^* = [\partial\widehat{\omega}(t+1|t)/\partial\mu]^*$, we will employ the following relationship

$$\frac{\partial\mathrm{Arg}[x(\mu)]}{\partial\mu} = \frac{\partial\arctan[x_I/x_R]}{\partial\mu}$$
$$= \frac{j}{2|x|^2}\left[x\left(\frac{\partial x}{\partial\mu^*}\right)^* - x^*\frac{\partial x}{\partial\mu}\right] \tag{28}$$

which can be derived using the chain rule. Using (28), one arrives at

$$\frac{\partial\widehat{\omega}(t+1|t)}{\partial\mu} = (1-\eta)\frac{\partial\widehat{\omega}(t|t-1)}{\partial\mu}$$
$$+ \frac{j\eta}{2}\left[\frac{[\partial\widehat{d}(t+1|t)/\partial\mu^*]^*}{\widehat{d}^*(t+1|t)} - \frac{(\partial\widehat{d}[t|t-1]/\partial\mu^*]^*}{\widehat{d}^*(t|t-1)}\right]$$
$$- \frac{j\eta}{2}\left[\frac{\partial\widehat{d}(t+1|t)/\partial\mu}{\widehat{d}(t+1|t)} - \frac{\partial\widehat{d}(t|t-1)/\partial\mu}{\widehat{d}(t|t-1)}\right]. \tag{29}$$

Remark: When the disturbance frequency is constant and known, i.e., $\widehat{\omega}(t|t-1) = \omega_0$ and $\partial\widehat{\omega}(t|t-1)/\partial\mu = \partial\omega_0/\partial\mu = 0$, it holds that (under suitable stability conditions) $\lim_{t\to\infty}\partial\widehat{d}(t|t-1)/\partial\mu^* = \lim_{t\to\infty}\partial y(t)/\partial\mu^* = 0$. In a case like this, the RPE recursions (25)–(27) reduce to those given in [4].

B. Adjustment of η

Consider another exponentially weighted measure of fit

$$W(t;\eta) = \frac{1}{2}\sum_{\tau=1}^{t}(\rho_\eta)^{t-\tau}|y(\tau;\eta)|^2 \tag{30}$$

where ρ_η, $0 < \rho_\eta < 1$, denotes another forgetting constant and $y(t)$ is now regarded a function of η. Using the RPE approach, one arrives at the following recursive scheme for evaluation of $\widehat{\eta}(t) = \arg\min_\eta W(t;\eta)$

$$\widehat{\mu}(t) = \widehat{\mu}(t-1) - \frac{\partial W[t;\widehat{\eta}(t-1)]/\partial\eta}{\partial^2 W[t;\widehat{\eta}(t-1)]/\partial\eta^2}$$

$$\frac{\partial W[t;\widehat{\eta}(t-1)]}{\partial\eta} = \mathrm{Re}\left\{y[t;\widehat{\eta}(t-1)]\frac{\partial y^*[t;\widehat{\eta}(t-1)]}{\partial\eta}\right\}$$

$$\frac{\partial^2 W[t;\widehat{\eta}(t-1)]}{\partial\eta^2} = \rho_\eta\frac{\partial^2 W[t;\widehat{\eta}(t-2)]}{\partial\eta^2}$$
$$+ \left|\frac{\partial y[t;\widehat{\eta}(t-1)]}{\partial\eta}\right|^2. \tag{31}$$

Furthermore

$$\frac{\partial y(t)}{\partial\eta} = -\beta\frac{\partial\widehat{d}(t|t-1)}{\partial\eta} \tag{32}$$

$$\frac{\partial\widehat{d}(t+1|t)}{\partial\eta} = j\frac{\partial\widehat{\omega}(t|t-1)}{\partial\eta}\widehat{d}(t+1|t)$$
$$+ e^{j\widehat{\omega}(t|t-1)}\left[\frac{\partial\widehat{d}(t|t-1)}{\partial\eta} + \mu\frac{\partial y(t)}{\partial\eta}\right]. \tag{33}$$

In order to evaluate $\partial\widehat{\omega}(t|t-1)/\partial\eta$, note that

$$\frac{\partial\mathrm{Arg}[x(\eta)]}{\partial\eta} = \frac{\partial\mathrm{Im}[\log x]}{\partial\eta} = \mathrm{Im}\left[\frac{\partial\log x}{\partial\eta}\right] = \mathrm{Im}\left[\frac{1}{x}\cdot\frac{\partial x}{\partial\eta}\right].$$

6

Using this relationship, one obtains

$$\frac{\partial \widehat{\omega}(t+1\,|\,t)}{\partial \eta} =$$

$$= (1-\eta)\frac{\partial \widehat{\omega}(t\,|\,t-1)}{\partial \eta} - \widehat{\omega}(t\,|\,t-1) + \mathrm{Arg}\left[\frac{\widehat{d}(t+1\,|\,t)}{\widehat{d}(t\,|\,t-1)}\right]$$

$$- \eta\,\mathrm{Im}\left[\frac{\partial \widehat{d}(t+1\,|\,t)/\partial \eta}{\widehat{d}(t+1\,|\,t)} - \frac{\partial \widehat{d}(t\,|\,t-1)/\partial \eta}{\widehat{d}(t\,|\,t-1)}\right]. \qquad (34)$$

C. Coping with Modeling Error

Since the quantities $\partial y(t)/\partial \mu$, $\partial y(t)/\partial \mu^*$, and $\partial y(t)/\partial \eta$ depend explicitly on the modeling error β, which is unknown, recursive formulas for the evaluation of sensitivity derivatives, derived in Sections V-A and V-B, can't be used in their present form. Following [4], we will replace β in (26) and (32) with

$$\beta = \frac{c_\mu}{\mu} \qquad (35)$$

where c_μ denotes a small positive constant. As shown in [4], in the constant-known-frequency case, this particular variant of the gain fixing technique guarantees convergence in mean of the adaptive disturbance rejection scheme to the optimal solution, irrespective of the phase error $\mathrm{Arg}[\beta]$ (which cannot be achieved when β is simply set to 1). Simulation experiments confirm that similar effect can be observed when (35) is used in combination with the self-optimizing control algorithm described in the previous subsections.

D. Summary of the Proposed Control Algorithm

The following shorthand will be used to simplify our notation:

$$\widehat{d}(t) = \widehat{d}(t\,|\,t-1)\,,\ \widehat{\omega}(t) = \widehat{\omega}(t\,|\,t-1)$$

$$y_\mu(t) = \frac{\partial y[t;\widehat{\mu}(t-1)]}{\partial \mu}\,,\ d_\mu(t) = \frac{\partial \widehat{d}[t\,|\,t-1;\widehat{\mu}(t-1)]}{\partial \mu}$$

$$y_{\mu^*}(t) = \frac{\partial y[t;\widehat{\mu}(t-1)]}{\partial \mu^*}\,,\ d_{\mu^*}(t) = \frac{\partial \widehat{d}[t\,|\,t-1;\widehat{\mu}(t-1)]}{\partial \mu^*}$$

$$y_\eta(t) = \frac{\partial y[t;\widehat{\eta}(t-1)]}{\partial \eta}\,,\ d_\eta(t) = \frac{\partial \widehat{d}[t\,|\,t-1;\widehat{\eta}(t-1)]}{\partial \eta}$$

$$r_\mu(t) = \frac{\partial^2 V[t;\widehat{\mu}(t-1)]}{\partial \mu^* \partial \mu}\,,\ r_\eta(t) = \frac{\partial^2 W[t;\widehat{\eta}(t-1)]}{\partial \eta^2}$$

$$\omega_\mu(t) = \frac{\partial \widehat{\omega}[t\,|\,t-1;\widehat{\mu}(t-1)]}{\partial \mu}$$

$$\omega_\eta(t) = \frac{\partial \widehat{\omega}[t\,|\,t-1;\widehat{\mu}(t-1)]}{\partial \eta}\ .$$

Using this symbolism, the proposed xSONIC algorithm can be expressed in the form:

Adjustment of $\widehat{\mu}(t)$

$$d_\mu(t) = j\omega_\mu(t-1)\widehat{d}(t)$$
$$+ e^{j\widehat{\omega}(t-1)}[d_\mu(t-1) + y(t-1) + \widehat{\mu}(t-1)y_\mu(t-1)]$$

$$d_{\mu^*}(t) = j\omega_\mu^*(t-1)\widehat{d}(t)$$
$$+ e^{j\widehat{\omega}(t-1)}[d_{\mu^*}(t-1) + \widehat{\mu}(t-1)y_{\mu^*}(t-1)]$$

$$\omega_\mu(t) = [1-\widehat{\eta}(t-1)]\omega_\mu(t-1) + \frac{j\widehat{\eta}(t-1)}{2}$$

$$\times \left[\frac{(d_{\mu^*}(t))^*}{\widehat{d}^*(t)} - \frac{d_\mu(t)}{\widehat{d}(t)} - \frac{(d_{\mu^*}(t-1))^*}{\widehat{d}^*(t-1)} + \frac{d_\mu(t-1)}{\widehat{d}(t-1)}\right]$$

$$y_\mu(t) = -\frac{c_\mu}{\widehat{\mu}(t-1)}\,d_\mu(t)$$

$$y_{\mu^*}(t) = -\frac{c_\mu}{\widehat{\mu}(t-1)}\,d_{\mu^*}(t)$$

$$r_\mu(t) = \rho_\mu r_\mu(t-1) + |\,y_\mu(t)\,|^2 + |\,y_{\mu^*}(t)\,|^2$$

$$\widehat{\mu}(t) = \widehat{\mu}(t-1) - \frac{y_\mu^*(t)y(t) + y_{\mu^*}(t)y^*(t)}{r_\mu(t)} \qquad (36)$$

Adjustment of $\widehat{\eta}(t)$

$$d_\eta(t) = j\omega_\eta(t-1)\widehat{d}(t) + e^{j\widehat{\omega}(t-1)}[d_\eta(t-1)$$
$$+ \widehat{\mu}(t-1)y_\eta(t-1)]$$

$$\omega_\eta(t) = [1-\widehat{\eta}(t-1)]\omega_\eta(t-1) - \widehat{\omega}(t-1)$$

$$+ \mathrm{Arg}\left[\frac{\widehat{d}(t)}{\widehat{d}(t-1)}\right] + \widehat{\eta}(t-1)\,\mathrm{Im}\left[\frac{d_\eta(t)}{\widehat{d}(t)} - \frac{d_\eta(t-1)}{\widehat{d}(t-1)}\right]$$

$$y_\eta(t) = -\frac{c_\mu}{\widehat{\mu}(t-1)}\,d_\eta(t)$$

$$r_\eta(t) = \rho_\eta r_\eta(t-1) + |\,y_\eta(t)|^2$$

$$\widehat{\eta}(t) = \widehat{\eta}(t-1) - \frac{\mathrm{Re}[y_\eta^*(t)y(t)]}{r_\eta(t)} \qquad (37)$$

Adjustment of $\widehat{d}(t)$ and $\widehat{\omega}(t)$

$$\widehat{d}(t+1) = e^{j\widehat{\omega}(t)}[\widehat{d}(t) + \widehat{\mu}(t)y(t)]$$

$$\widehat{\omega}(t+1) = [1-\widehat{\eta}(t)]\,\widehat{\omega}(t) + \widehat{\eta}(t)\,\mathrm{Arg}\left[\frac{\widehat{d}(t+1)}{\widehat{d}(t)}\right] \qquad (38)$$

Evaluation of control signal

$$u(t) = -\frac{\widehat{d}(t+1)}{k_\mathrm{n}}$$
$$\text{or}$$
$$u(t) = -\frac{\widehat{d}(t+1)}{K_\mathrm{n}(e^{j\widehat{\omega}(t+1)})} \qquad (39)$$

The computational complexity of the xSONIC algorithm is equal to 98 real multiply/add operations, 5 real division operations, and 4 trigonometric operations ($1\times\sin$, $1\times\cos$, $2\times\arctan$) per time update.

E. Simplified Algorithm

Simplification of the xSONIC algorithm can be easily achieved by fixing the gain η. According to the ALF equations (9), the tracking properties of the algorithm (6) can be quantified in terms of $\mu\beta$ and η. Dependence on $\mu\beta$ means that the "true" adaptation gain of the amplitude tracking loop is determined not only by μ, but also by modeling errors β. Since modeling errors are unknown, and they may change with time, choosing a "meaningful" value of μ is an impossible task. This is why automatic adjustment of μ is so important. In the case of the second adaptation gain, controlling the frequency tracking loop, the situation is different. Since modeling errors do not interfere with η, the value of this coefficient can be safely fixed at a constant level, based on our prior knowledge of the expected rate of frequency variation.

To arrive at a simplified version of the algorithm, we will assume that the estimated frequency trajectory coincides with the true trajectory, which leads to: $\omega_\mu(t) = \partial\widehat{\omega}(t)/\partial\mu = \partial\omega(t)/\partial\mu = 0$ and $\omega_\eta(t) = \partial\widehat{\omega}(t)/\partial\eta = \partial\omega(t)/\partial\eta = 0$. Since, under suitable stability constraints, the condition $\omega_\mu(t) = 0$ entails (asymptotically) $d_{\mu^*}(t) = 0$, $y_{\mu^*}(t) = 0$, and the condition $\omega_\eta(t) = 0$ entails $d_\eta(t) = 0$, $y_\eta(t) = 0$, the algorithm (36)-(39) reduces down to

$$d_\mu(t) = e^{j\widehat{\omega}(t-1)}[d_\mu(t-1) + y(t-1)$$
$$+ \widehat{\mu}(t-1)y_\mu(t-1)]$$
$$y_\mu(t) = -\frac{c_\mu}{\widehat{\mu}(t-1)}\, d_\mu(t)$$
$$r_\mu(t) = \rho_\mu r_\mu(t-1) + |\,y_\mu(t)\,|^2$$
$$\widehat{\mu}(t) = \widehat{\mu}(t-1) - \frac{y_\mu^*(t)y(t)}{r_\mu(t)}$$
$$\widehat{d}(t+1) = e^{j\widehat{\omega}(t)}[\widehat{d}(t) + \widehat{\mu}(t)y(t)]$$
$$\widehat{\omega}(t+1) = [1-\eta]\widehat{\omega}(t) + \eta\,\mathrm{Arg}\left[\frac{\widehat{d}(t+1)}{\widehat{d}(t)}\right]$$
$$u(t) = -\frac{\widehat{d}(t+1)}{k_{\mathrm{n}}}\,. \tag{40}$$

The computational burden associated with the simplified algorithm is equal to 37 real multiply/add operations, 2 real division operations, and 4 trigonometric operations ($1\times\sin$, $1\times\cos$, $2\times\arctan$) per time update.

VI. Safety Measures and Extensions

A. Safeguards

Similarly as in [4], to avoid erratic behavior of the algorithm during startup/transient periods, it is advisable to set the maximum allowable values for $|\,\widehat{\mu}(t)\,|$ and $\widehat{\eta}(t)$, further denoted by $\mu_{\max}$ and $\eta_{\max}$, respectively. The modified update rules have the form

$$\widehat{\mu}(t) = \mathrm{sat}\left[\widehat{\mu}(t-1) - \frac{y_\mu^*(t)y(t) + y_{\mu^*}(t)y^*(t)}{r_\mu(t)}\ ,\ \mu_{\max}\right]$$
$$\widehat{\eta}(t) = \min\left[\widehat{\eta}(t-1) - \frac{\mathrm{Re}[y_\eta^*(t)y(t)]}{r_\eta(t)}\ ,\ \eta_{\max}\right] \tag{41}$$

where $\mathrm{sat}(x,a)$, $x \in \mathrm{C}$, $a \in \mathrm{R}_+$ denotes a complex-valued saturation function

$$\mathrm{sat}(x,a) = \left\{ \begin{array}{ccc} x & \text{if} & |\,x\,| \le a \\ a\,\dfrac{x}{|\,x\,|} & \text{if} & |\,x\,| > a \end{array} \right. .$$

B. Extensions

We will describe three extensions of the proposed scheme: to systems with multiharmonic disturbances, systems with an extra transport delay, and real-valued signals.

1) Multiharmonic Disturbances: Suppose that the signal $d(t)$ in (1) has the form

$$d(t) = \sum_{i=1}^m d_i(t)$$

and is made up of m frequency components

$$d_i(t) = e^{j\omega_i(t-1)}d_i(t-1) + \widetilde{e}_i(t)$$
$$\omega_i(t) = \omega_i(t-1) + w_i(t) \tag{42}$$
$$i = 1,\ldots,m$$

evolving independently of one another. Rejection of such a multiharmonic disturbance can be achieved by combining m subalgorithms, each designed to track one component of $d(t)$, into the following parallel estimation scheme

$$\widehat{d}_i(t+1|\,t) = e^{j\widehat{\omega}_i(t|\,t-1)}\left[\,\widehat{d}_i(t|\,t-1) + \widehat{\mu}_i(t)y(t)\,\right]$$
$$\widehat{\omega}_i(t+1|\,t) = [\,1-\widehat{\eta}_i(t)\,]\widehat{\omega}_i(t|\,t-1)$$
$$+ \widehat{\eta}_i(t)\,\mathrm{Arg}\left[\frac{\widehat{d}_i(t+1|\,t)}{\widehat{d}_i(t|\,t-1)}\right]$$
$$i = 1,\ldots,m$$
$$u(t) = -\sum_{i=1}^m \frac{\widehat{d}_i(t+1|\,t)}{k_{\mathrm{n}_i}} \tag{43}$$

where $k_{\mathrm{n}_1},\ldots,k_{\mathrm{n}_m}$ denote the nominal plant gains, and where the adaptation gains $\widehat{\mu}_i(t)$, $\widehat{\eta}_i(t)$ are computed recursively (independently of one another) using the algorithm designed for the single-frequency case.

When the frequencies $\omega_1(t),\ldots,\omega_m(t)$ are known to be mutually related, e.g. when $\omega_i(t) = i\omega_1(t)$, $i = 2,\ldots,m$ (which corresponds to a fundamental and $m-1$ harmonics), more specialized algorithms can be designed to take advantage of the available prior knowledge. Because of the lack of space, such opportunity will be not further explored here.

2) Systems with Delay: Suppose that the plant is governed by

$$y(t) = K_{\mathrm{p}}(q^{-1})u(t - \tau_0) + d(t) + v(t)$$
$$\cong k_{\mathrm{p}}u(t - \tau_0) + d(t) + v(t) \tag{44}$$

where τ_0 denotes transport delay. So far we have considered the one-sample delay case. The modified control rule that will work for $\tau_0 > 1$ has the form

$$\widehat{d}(t + \tau_0 | t) = e^{j\widehat{\omega}(t+\tau_0-1 | t-1)} \widehat{d}(t + \tau_0 - 1 | t - 1)$$
$$+ \widehat{\mu}(t)e^{j\tau_0\widehat{\omega}(t+\tau_0-1 | t-1)}y(t)$$

$$\widehat{\omega}(t + \tau_0 | t) = [1 - \widehat{\eta}(t)]\widehat{\omega}(t + \tau_0 - 1 | t - 1)$$
$$+ \widehat{\eta}(t) \operatorname{Arg}\left[\frac{\widehat{d}(t + \tau_0 | t)}{\widehat{d}(t + \tau_0 - 1 | t - 1)} \right]$$

$$u(t) = -\frac{\widehat{d}(t + \tau_0 | t)}{k_{\mathrm{n}}} . \tag{45}$$

Similar to the algorithm derived in the constant-known-frequency case (see Section VIII-B in Part I), the algorithm (45) is robust to delay misspecification.

Based on (44) and (45), one can easily derive recursions analogous to (26), (27), (29), and (32), (33), (34), needed to evaluate sensitivity derivatives $\partial y(t)/\partial \mu$, $\partial y(t)/\partial \mu^*$, and $\partial y(t)/\partial \eta$.

3) Real-Valued Signals: A simple way of extending the proposed algorithms to real–valued signals can be summarized as follows:

1) Regarding $\{y(t)\}$ as a sequence of complex numbers $[y_{\mathrm{R}}(t) = y(t), y_{\mathrm{I}}(t) = 0]$, compute the complex-valued control signal $u(t) = u_{\mathrm{R}}(t) + ju_{\mathrm{I}}(t)$ using the proposed algorithm.
2) Apply $u_{\mathrm{R}}(t)$ to the input of the controlled plant.

VII. SIMULATION AND EXPERIMENTAL RESULTS

Several experiments were performed to check both the transient and steady-state performance of the proposed disturbance rejection scheme.

A. Mean Convergence Properties

The purpose of this experiment was to examine the steady-state mean convergence properties of the algorithm:

- For two choices of the plant: the dynamic Guo-Bodson plant $K_{\mathrm{p}}(q^{-1})$, given by (22), and its non-inertial counterpart, represented by the time-varying static gain $k_{\mathrm{p}}(t) = K_{\mathrm{p}}(e^{-j\omega(t)})$ — to check to what extent static approximation of dynamic plants, exploited in our analytical study, affects the obtained results.
- For three choices of magnitude errors ($|\beta| = 0.5, 1, 2$) and two choices of phase errors ($\operatorname{Arg}[\beta] = 0\,^\circ, 45\,^\circ$) — to check sensitivity of the cancellation scheme to different kinds of modeling errors.
- For three speeds of frequency variation ($\sigma_w = 4 \cdot 10^{-6}, 10^{-5}, 4 \cdot 10^{-5}$) — to check how the obtained results depend on the rate of frequency variation.

The results, summarized in Tables I and II, were obtained by means of combined ensemble averaging (100 realizations of the processes $\{v(t)\}$ and $\{w(t)\}$) and time averaging ($t \in [20001, 80000]$), after the examined algorithms have reached their steady-state behavior. In all cases the noise variance was constant and equal to $\sigma_v^2 = 0.01$, the initial frequency value was set to $\omega(0) = 0.1$, and the amplitude of the disturbance was equal to $a = 1$. To enforce modeling errors equal to (approximately) β, the nominal model, adopted in (7), had the form $K_{\mathrm{n}}(e^{-j\omega}) = K_{\mathrm{p}}(e^{-j\omega})/\beta$. Finally, the algorithm settings were equal to $\rho_\mu = \rho_\eta = 0.9998$ and $c_\mu = 0.01$.

Several conclusions can be drawn after comparing the average values of the quantities $|\widehat{\mu}(t)\beta|$, $\operatorname{Arg}[\widehat{\mu}(t)\beta]$, and $\widehat{\eta}(t)$, gathered in Tables I and II. In all cases, the proposed adaptive regulator remains close, in the mean sense, to the optimal regulator. Not surprisingly, the smallest estimation biases occur for the slowest frequency changes. There are no substantial differences between the results obtained for the dynamic plant (Table II) and for its static counterpart (Table I). As expected, the static plant experimental results better fit theory.

Figs. 2(a) and 2(b) show typical simulation results obtained in the absence of modeling errors and in the presence of modeling errors, respectively, for the same realizations of $\{v(t)\}$ and $\{w(t)\}$ ($\sigma_v = 0.1$, $\sigma_w = 5 \cdot 10^{-4}$, $a = 1$, $\omega(0) = 0.1$, $c_\mu = 0.01$, $\rho_\mu = \rho_\eta = 0.9992$). In the case depicted in Fig. 2b, modeling error was time-varying – it evolved from $|\beta| = 0.7$, $\operatorname{Arg}[\beta] = -48\,^\circ$ to $|\beta| = 0.4$, $\operatorname{Arg}[\beta] = -60\,^\circ$.

During the first 1000 time-steps, the quantities $y_\mu(t)$, $y_{\mu^*}(t)$, $r_\mu(t)$, $y_\eta(t)$, $r_\eta(t)$ were evaluated, but the adaptation gains $\widehat{\mu}(t)$ and $\widehat{\eta}(t)$ were kept at their starting values and not updated. Then, at the instant $t = 1001$, the adaptation lock was released.

B. Comparison with the Constant-Known-Frequency Solution

Our second experiment aimed at comparing performance of the control algorithm proposed in [4] and the frequency-adaptive algorithms summarized in Sections V-D and V-E. Unlike the previous example, the instantaneous frequency changes were deterministic and governed by

$$\omega(t) = 0.1 + 0.02 \sin(0.001t) \tag{46}$$

i.e., the frequency was subject to small sinusoidal variations around the "nominal" frequency $\omega_0 = 0.1$. Fig. 3 shows the results obtained for the Guo-Bodson plant (22) and a typical (the same in all cases) realization of noise ($\sigma_v = 0.1$). The control algorithm, adopted from [4], was used in Fig. 3(a) with the following settings: $\omega_0 = 0.1$, $k_{\mathrm{n}} = K_{\mathrm{p}}(e^{j0.1})$, $\widehat{d}(0) = 1$, $\widehat{\mu}(0) = 0.05$, $c_\mu = 0.01$, $\rho = 0.9995$, $\mu_{\max} = 0.1$, $r(0) = 100$. The analogous settings for the "full-size" frequency-adaptive algorithm in Fig. 3(b) were: $k_{\mathrm{n}} = K_{\mathrm{p}}(e^{j0.1})$, $\widehat{d}(0) = 1$, $\widehat{\mu}(0) = 0.05$, $\widehat{\eta}(0) = 0.05$, $c_\mu = 0.01$, $\rho_\mu = \rho_\eta = 0.9995$, $\mu_{\max} = 0.1$, $r_\mu(0) = r_\eta(0) = 100$. The adopted upper bounds on $|\mu|$ and η play the role of "safety valves". When, for example, $|\widehat{\mu}(t)|$ is allowed to exceed $\mu_{\max} = 0.1$, the closed-loop system is prone to occasional bursts caused by local

TABLE I
MEAN CONVERGENCE RESULTS FOR DIFFERENT MAGNITUDE AND PHASE ERRORS:
CONTROL ALGORITHM BASED ON THE SUBSTITUTION $\beta = c_\mu/\mu$, STATIC SYSTEM.

σ_w	$\mu_d = \eta_d$	$\beta = 1$			$\beta = 2$			$\beta = 0.5$		
		$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$	$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$	$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$
$4\cdot10^{-6}$	0.0075	0.0079	$0.2\,^\circ$	0.0081	0.0079	$0.2\,^\circ$	0.0080	0.0079	$0.2\,^\circ$	0.0080
$1\cdot10^{-5}$	0.0119	0.0123	$0.2\,^\circ$	0.0112	0.0123	$0.2\,^\circ$	0.0112	0.0123	$0.2\,^\circ$	0.0112
$4\cdot10^{-5}$	0.0241	0.0244	$0.2\,^\circ$	0.0195	0.0244	$0.2\,^\circ$	0.0195	0.0244	$0.2\,^\circ$	0.0195

σ_w	$\mu_d = \eta_d$	$\beta = e^{j\pi/4}$			$\beta = 2e^{j\pi/4}$			$\beta = 0.5e^{j\pi/4}$		
		$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$	$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$	$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$
$4\cdot10^{-6}$	0.0075	0.0079	$0.0\,^\circ$	0.0081	0.0078	$0.0\,^\circ$	0.0082	0.0079	$0.0\,^\circ$	0.0081
$1\cdot10^{-5}$	0.0119	0.0123	$-0.1\,^\circ$	0.0112	0.0123	$-0.1\,^\circ$	0.0112	0.0123	$-0.1\,^\circ$	0.0113
$4\cdot10^{-5}$	0.0241	0.244	$-0.2\,^\circ$	0.0192	0.0244	$-0.2\,^\circ$	0.0192	0.0244	$-0.2\,^\circ$	0.0192

TABLE II
MEAN CONVERGENCE RESULTS FOR DIFFERENT MAGNITUDE AND PHASE ERRORS:
CONTROL ALGORITHM BASED ON THE SUBSTITUTION $\beta = c_\mu/\mu$, DYNAMIC SYSTEM.

σ_w	$\mu_d = \eta_d$	$\beta = 1$			$\beta = 2$			$\beta = 0.5$		
		$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$	$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$	$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$
$4\cdot10^{-6}$	0.0075	0.0082	$-1.4\,^\circ$	0.0073	0.0082	$-1.4\,^\circ$	0.0073	0.0082	$-1.4\,^\circ$	0.0073
$1\cdot10^{-5}$	0.0119	0.0129	$-1.2\,^\circ$	0.0102	0.0129	$-1.2\,^\circ$	0.0102	0.0129	$-1.2\,^\circ$	0.0102
$4\cdot10^{-5}$	0.0241	0.0255	$-1.4\,^\circ$	0.0177	0.0255	$-1.4\,^\circ$	0.0177	0.0255	$-1.4\,^\circ$	0.0177

σ_w	$\mu_d = \eta_d$	$\beta = e^{j\pi/4}$			$\beta = 2e^{j\pi/4}$			$\beta = 0.5e^{j\pi/4}$		
		$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$	$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$	$\overline{\lvert\mu\beta\rvert}$	$\overline{\mathrm{Arg}[\mu\beta]}$	$\overline{\eta}$
$4\cdot10^{-6}$	0.0075	0.0083	$-1.6\,^\circ$	0.0071	0.0083	$-1.6\,^\circ$	0.0072	0.0084	$-1.7\,^\circ$	0.0072
$1\cdot10^{-5}$	0.0119	0.0129	$-1.5\,^\circ$	0.0101	0.0129	$-1.5\,^\circ$	0.0101	0.0129	$-1.4\,^\circ$	0.0101
$4\cdot10^{-5}$	0.0241	0.0255	$-1.5\,^\circ$	0.0177	0.0255	$-1.5\,^\circ$	0.0177	0.0255	$-1.5\,^\circ$	0.0177

instabilities. The simplified frequency-adaptive algorithm in Fig. 3(c) was run with $\eta = 0.05$. For all control algorithms, the gain adjustment mechanisms were switched on at the instant $t = 100$.

Note the relatively poor performance of the simpler controller [Fig. 3(a)], which does not have enough degrees of freedom to compensate all modeling errors. Basically, such a controller works satisfactorily only at moments when the instantaneous frequency is close to the nominal (assumed) frequency. Note also that, after the initial convergence period, the magnitude of the estimated gain $\widehat{\mu}(t)$ stays at its maximum allowable level. In contrast with this, the results obtained for the frequency-adaptive control algorithms [Figs. 3(b) and 3(c)] are satisfactory at all times. The simplified algorithm yields only slightly worse results than the "full-size" algorithm.

C. Comparison with the Adaptive FXLMS Algorithm

We will compare our algorithm with the FXLMS-based solution, described in some detail in Part I of this paper. The FXLMS canceller consists of two loops:

1) The system identification loop, providing estimates of the coefficients of a finite impulse response (FIR) model of the plant: $\widehat{K}_p(q^{-1}) = \sum_{i=1}^{M} \widehat{k}_i q^{-i}$ (to guarantee plant identifiability, a low-intensity random perturbation, e.g., white noise sequence, is added to the input signal).

2) The direct control loop, which forms the compensating signal by means of adjusting the magnitude and phase of the reference signal $r(t)$.

Since, in the case considered, the reference is neither available, nor can it be synthesized (the instantaneous frequency of $d(t)$ is unknown and possibly time-varying), it is replaced with the reconstructed disturbance signal

$$\widehat{d}(t) = \mathrm{ALE}[y(t) - \widehat{K}_p(q^{-1})u(t)]$$

where ALE denotes adaptive line enhancer, an algorithm that extracts sinusoidal waveforms from noisy measurements [12].

Comparison was performed for the same plant that was described in Part I – the first-order inertial system with a "time-varying pole", governed by

$$y(t) = \varrho(t)y(t-1) + 0.0952u(t-1)$$
$$\varrho(t) = 0.7 + 0.25\sin(0.0003t).$$

The amplitude $a(t)$ and the instantaneous frequency $\omega(t)$ of the sinusoidal disturbance

$$d(t) = a(t)e^{j\sum_{\tau=1}^{t}\omega(\tau)}$$

were changing according to

$$a(t) = 1 + 0.2\sin(0.002t)$$
$$\omega(t) = 0.3 + 0.03\sin(0.001t).$$

Fig. 4 shows the results obtained in the absence of measurement noise ($\sigma_v = 0$) for the xSONIC algorithm and the well-tuned adaptive FXLMS algorithm (for a typical realization of
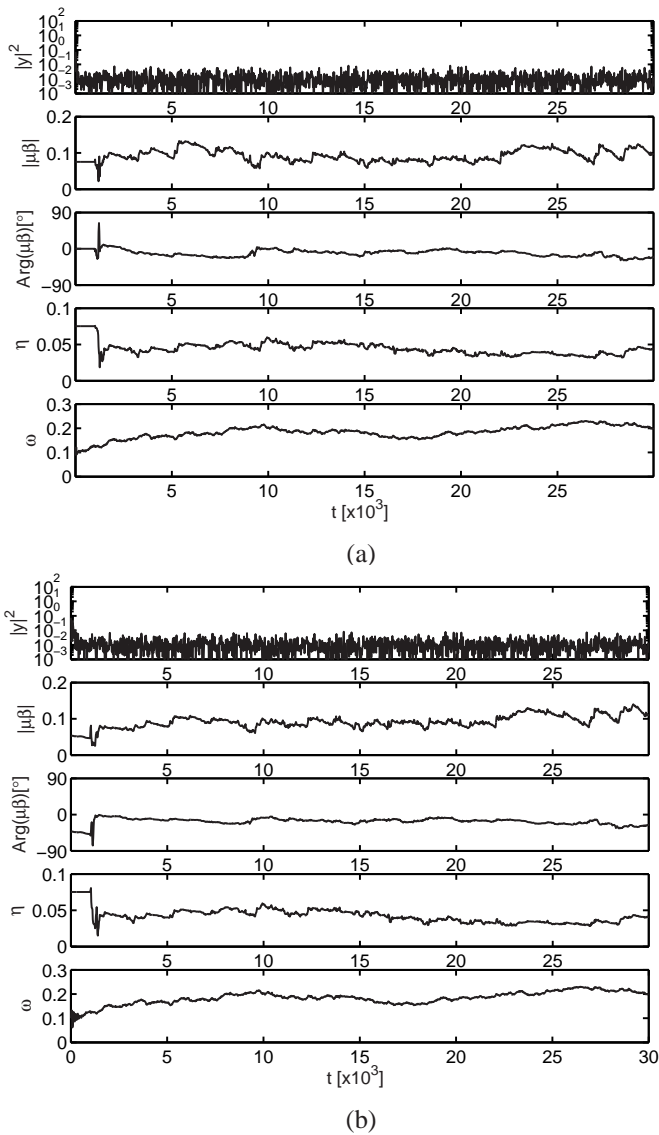
Fig. 2. Behavior of the disturbance rejection algorithm – results of a typical simulation experiment. (a) In the absence of modeling errors. (b) In the presence of modeling errors.
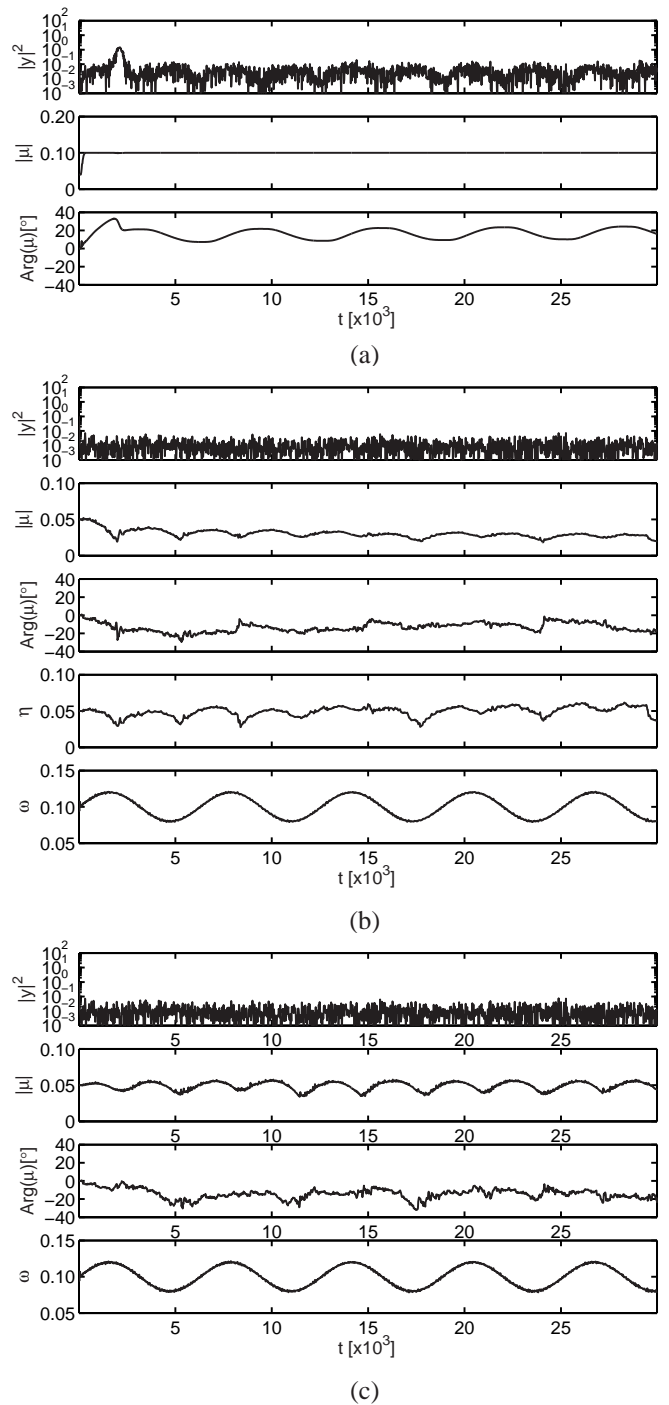

Fig. 3. Performance of adaptive disturbance rejection in the presence of frequency variations. (a) Algorithm based on the constant-known-frequency model. (b) Frequency-adaptive disturbance rejection algorithm. (c) Simplified frequency-adaptive algorithm.

auxiliary noise). The noiseless case is the one most frequently encountered in the literature on acoustic disturbance suppression. xSONIC was used with the following design parameters: $c_\mu = 0.01, \rho_\mu = \rho_\eta = 0.9995, k_n = 1$, none of which had a strong influence on the simulation results. The best settings for FXLMS, determined by means of extensive numerical search, were equal to: $\mu_1 = 0.02$ (adaptation step-size in the system identification loop), $\mu_2 = 0.15$ (adaptation step-size in the control loop), and $\sigma_{u'}^2 = 0.001$ (variance of auxiliary noise). The order of the FIR filter was set to $M = 32$ (for $M < 20$, the closed-loop system was unstable). Finally, the adaptive line enhancer, which is a part of the FXLMS-based canceller, was based on a carefully tuned multiple frequency tracker, described in [5]. It should be stressed, that the performance of FXLMS cannot be improved by further reducing the variance of the auxiliary noise.

According to Fig. 4, xSONIC yields considerably better cancellation results than FXLMS – the corresponding root-mean-squared output error is equal to 0.0034 (virtually the same for the simplified algorithm), compared to 0.0085 for FXLMS (improvement by 8 dB).

In the presence of measurement noise ($\sigma_v = 0.1$), xSONIC offers practically the same cancellation efficiency as the *well-tuned* FXLMS ($\mu_1 = 0.02$, $\mu_2 = 0.1$, $\sigma_{u'}^2 = 0.001$) – the root-mean-squared output errors for xSONIC, simplified xSONIC,
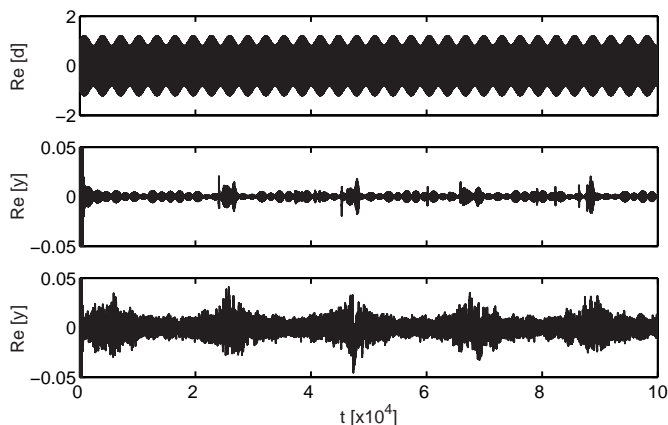
Fig. 4. Real part of the disturbance signal (top) and cancellation results obtained for the xSONIC algorithm (middle) and the adaptive FXLMS algorithm (bottom).

and adaptive FXLMS are equal to 0.1032, 1.041, and 1.032, respectively.

As already remarked in PartI of this paper, since in practice FXLMS seldom is optimally tuned, the comparison presented above is rather unfair for xSONIC.

D. Acoustic Experiment

A real-world acoustic active noise control experiment was conducted using the proposed regulator. The instantaneous frequency of the artificially generated disturbance was changing sinusoidally between 241 and 250 Hz, with a period of 20 s. The error microphone was located approximately 1 m away from the source of disturbance and 15 cm from the noise canceling loudspeaker. The system was operated at a sampling rate of 1 kHz. The nominal filter gain k_n was set to 1. The remaining parameters were chosen as follows: $c_\mu = 0.01$, $\rho_\mu = 0.999$, $\rho_\eta = 0.99$, $\mu_{max} = 0.05$. Fig. 5 shows the results. After an initial convergence phase, which lasted for about 1s, the closed-loop system reached its steady-state behavior. The achieved rate of disturbance attenuation was approximately 20 dB.

In the experiment described above, the rate of frequency variation was limited by a large processing delay, equal to 60 sampling intervals (we did not use specialized hardware). However, our simulation tests show that for smaller delays, xSONIC can satisfactorily track persistent (e.g., linear) frequency changes up to 50 Hz/s under 1 kHz sampling.

VIII. CONCLUSION

The problem of eliminating a sinusoidal disturbance of unknown, slowly time-varying frequency, acting at the output of an unknown (and possibly slowly time-varying) stable linear plant, was considered. The adaptive feedback disturbance rejection scheme, proposed and analyzed in this paper consists of two loops: the inner control loop, which predicts and cancels the disturbance, and the outer, self-optimization loop, which automatically adjusts adaptation gains to the rate of system and/or disturbance nonstationarity. Results of computer simulation and real-world experiments confirm very good rejection/tracking properties of the derived algorithm.
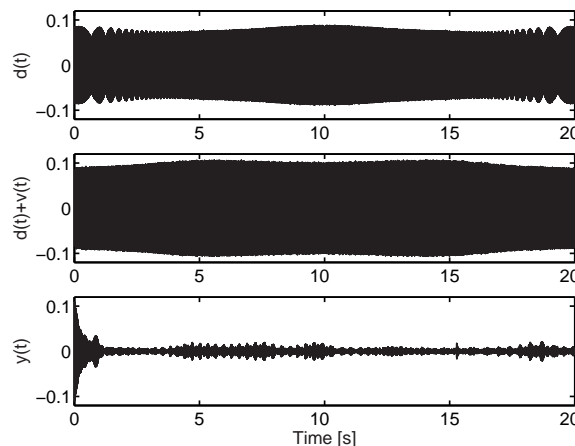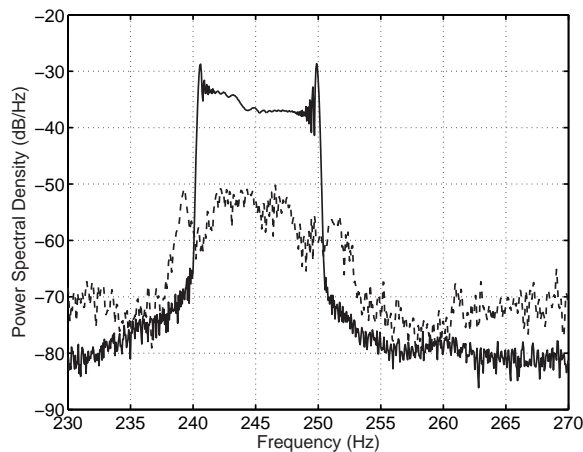

Fig. 5. Power spectral density of the signal before (solid line) and after (dotted line) disturbance cancellation (top figure) and the corresponding time-domain measurements (three lower figures).

REFERENCES

[1] C. R. Fuller, S. J. Elliott, and P. A. Nelson, *Active control of vibration*. New York: Academic Press, 1995.
[2] S. J. Elliott and P. A. Nelson, "Active noise control," *IEEE Signal Process. Mag.*, vol. 10, pp. 12–35, Oct. 1993.
[3] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems: Algorithms and DSP Implementations*. New York: Wiley, 1996.
[4] M. Niedźwiecki and M. Meller, "A new approach to active noise and vibration control – Part I: the known frequency case," submitted to *IEEE Trans. on Signal Process.*
[5] 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.
[6] P.M. Baggenstoss and S.M. Kay, "On estimating the angle parameters of an exponential signal at high SNR," *IEEE Trans. on Acoust., Speech, Signal Process.*, vol. 39, pp. 1203–1205, May 1991.
[7] M. Jury, *Theory and Application of the $\mathcal{Z}$-transform Method*. New York: Wiley 1964.
[8] P. Tichavský, "Posterior Cramér-Rao bound for adaptive harmonic retrieval," *IEEE Trans. on Signal Process.*, vol. 43, pp. 1299–1302, May 1995.
[9] X. Guo and M. Bodson, "Adaptive rejection of multiple sinusoids of unknown frequency," in *Proc. European Control Conference*, Kos, Greece, 2007, pp. 121–128.
[10] M. Niedźwiecki and P. Kaczmarek, "Tracking analysis of a generalized adaptive notch filter," *IEEE Trans. on Signal Process.*, vol. 54, pp. 304–314, Jan. 2006.
[11] T. Söderström and P. Stoica, *System Identification*. Englewood Cliffs NJ: Prentice Hall, 1988.
[12] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs NJ: Prentice Hall, 1996.

APPENDIX

DERIVATION OF (9)

According to [5], when carrying out the ALF analysis of (6), one should examine the dependence of $\Delta\widehat{d}(t)$ and $\Delta\widehat{\omega}(t)$ on $v(t)$ and $w(t)$, neglecting higher than first-order terms of all quantities listed above, including all cross-terms.

To arrive at the recursive formula for $\Delta\widehat{d}(t)d^*(t)$, note that

$$\Delta\widehat{d}(t) = d(t) - \beta\widehat{d}(t\,|\,t-1) = e^{j\omega(t-1)}d(t-1)$$
$$- \beta e^{j\omega(t-1)}e^{-j\Delta\widehat{\omega}(t-1)}[\widehat{d}(t-1\,|\,t-2)$$
$$+ \mu\Delta\widehat{d}(t-1) + \mu v(t-1)]. \tag{47}$$

When tracking is satisfactory, i.e., when $\Delta\widehat{\omega}(t)$ is small, one obtains $e^{-j\Delta\widehat{\omega}(t-1)} \cong 1 - j\Delta\widehat{\omega}(t-1)$ which, after substitution in (47) and neglecting all higher-order terms, leads to the following approximation

$$\Delta\widehat{d}(t) = e^{j\omega(t-1)}d(t-1)$$
$$- \beta e^{j\omega(t-1)}[1 - j\Delta\widehat{\omega}(t-1)]\widehat{d}(t-1\,|\,t-2)$$
$$- \beta\mu e^{j\omega(t-1)}\Delta\widehat{d}(t-1) - \beta\mu e^{j\omega(t-1)}v(t-1)].$$

Furthermore, since $\widehat{d}(t-1\,|\,t-2) = [d(t-1) - \Delta\widehat{d}(t-1)]/\beta$, after substitution, regrouping, and neglecting the term proportional to $\Delta\widehat{\omega}(t-1)\Delta\widehat{d}(t-1)$, one obtains

$$\Delta\widehat{d}(t) = e^{j\omega(t-1)}[1 - \beta\mu]\Delta\widehat{d}(t-1)$$
$$+ je^{j\omega(t-1)}d(t-1)\Delta\widehat{\omega}(t-1) - \beta\mu e^{j\omega(t-1)}v(t-1) . \tag{48}$$

Finally, after multiplying both sides of (48) by $d^*(t) = e^{-j\omega(t-1)}d^*(t-1)$, one arrives at the first recursion of (9).

To derive the second recursion, note that

$$\widehat{\omega}(t+1\,|\,t) = \widehat{\omega}(t\,|\,t-1) + \eta g(t) \tag{49}$$

where

$$g(t) = \text{Arg}\left[\frac{\widehat{d}(t+1\,|\,t)e^{-j\widehat{\omega}(t\,|\,t-1)}}{\widehat{d}(t\,|\,t-1)}\right]$$
$$= \text{Im}\left\{\log\left[1 + \frac{\mu y(t)}{\widehat{d}(t\,|\,t-1)}\right]\right\} \cong \text{Im}\left[\frac{\mu y(t)}{\widehat{d}(t\,|\,t-1)}\right].$$

Assuming that the cancellation error $\Delta\widehat{d}(t)$ is small, which means that $d(t) \cong \beta\widehat{d}(t\,|\,t-1)$, one arrives at

$$g(t) \cong \text{Im}\left[\frac{\mu\beta y(t)}{d(t)}\right] = \frac{1}{a^2}\text{Im}[\mu\beta\Delta\widehat{d}(t)d^*(t)]$$
$$+ \frac{1}{a^2}\text{Im}[\mu\beta v(t)d^*(t)]. \tag{50}$$

Combining (49) with (50), one obtains

$$\widehat{\omega}(t+1\,|\,t) = \widehat{\omega}(t\,|\,t-1) + \frac{\eta}{a^2}\text{Im}[\mu\beta\Delta\widehat{d}(t)d^*(t)]$$
$$+ \frac{\eta}{a^2}\text{Im}[\mu\beta v(t)d^*(t)].$$

Finally, subtracting this equation (sidewise) from $\omega(t+1) = \omega(t) + w(t+1)$, one arrives at the second recursion of (9).

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 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 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, signal processing and adaptive systems.

Michał Meller (S'08) received the M.Sc. degree in automatic control from Gdańsk University of Technology, Gdańsk, Poland, in 2007. Since 2007 he has been working in Telecommuncations Research Institute S.A. Gdańsk Division, Department of Signal and Information Processing. He is also pursuing his Ph.D. in the the Department of Automatic Control, Faculty of Electronics, Telecommunications and Computer Science, Gdańsk University of Technology. His professional interests include signal processing and adaptive systems.