A New Approach to Active Noise and Vibration Control – Part I: the Known 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 dynamics. It is assumed that the frequency of the sinusoidal disturbance is known, and that the output signal is contaminated with wideband measurement noise. The disturbance rejection control rule is first derived and analyzed for a nominal plant model, different from the true model. Then a special adaptation mechanism is added, which is capable of compensating modeling biases (errors in both magnitude and phase) so that, under Gaussian assumptions, the closed-loop system can converge in mean to the optimal solution.

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

I. INTRODUCTION

CONSIDER the problem of reducing of a complex-valued narrowband disturbance at the output of a discrete-time system governed by (see Fig. 1)

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

where $y(t)$ denotes the corrupted complex-valued system output, $t = \ldots, -1, 0, 1, \ldots$ denotes normalized time, $K_{\mathrm{p}}(q^{-1})$ denotes unknown transfer function of a stable linear single-input single-output plant, q^{-1} is the backward shift operator,

$$d(t) = a(t)e^{j\omega_0 t} \qquad (2)$$

is a complex-valued sinusoidal disturbance (cisoid) of known frequency ω_0 and unknown, slowly-varying amplitude $a(t)$, $v(t)$ is a wideband measurement noise, and $u(t)$ denotes the input (control) signal.

We will look for a feedback controller allowing for cancellation, or near-cancellation, of the sinusoidal disturbance, i.e., a controller generating the complex-valued feedback signal $u(t)$ that minimizes the system output in the mean-squared sense: $\mathrm{E}[\,|\,y(t)\,|^2\,] \longmapsto \min.$

The need for vibration control arises in many electromechanical systems, where narrowband disturbances are usually generated by rotating machinery and their suppression is necessary to maintain high quality of the underlying technological process [1]. Similar problems are encountered in acoustic active noise control systems, where the unwanted sound is attenuated by a noise-canceling speaker emitting a sound wave with the same amplitude but opposite polarity

The authors are with the Faculty of Electronics, Telecommunications and Computer Science, Department of Automatic Control, Gdańsk University of Technology, ul. Narutowicza 11/12, Gdańsk, Poland (e-mail: maciekn@eti.pg.gda.pl; michal.meller@eti.pg.gda.pl).
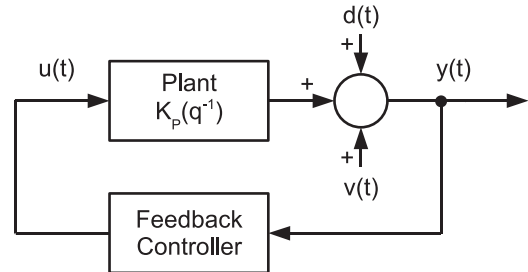
Fig. 1. Block diagram of the disturbance rejection system.

[2], [3]. Both types of applications fall into a more general narrowband disturbance rejection category.

The problem of narrowband disturbance rejection was considered by many authors under different methodologies, such as filtered-X LMS (FXLMS) compensation [3], [4], internal model principle [5], [6], and phase-locked loop control [7], [8].

The proposed self-optimizing narrowband noise canceling (SONIC) algorithm incorporates two adaptation loops: the inner control loop, which predicts and cancels the disturbance, and the outer, self-optimization loop, which automatically adjusts the complex-valued adaptation gain (the only parameter that is estimated) to the unknown, and possibly time-varying, characteristics of the controlled plant, sinusoidal disturbance, and/or measurement noise. We prove that, under certain conditions, the resulting controller converges in mean to the optimal, minimum-variance controller, and we show that it compares favorably, both in terms of canceling efficiency and computational complexity, with the very popular (especially in the signal processing community) FXLMS active noise control scheme.

Unlike most of the existing contributions, our study will focus on complex-valued signals. This is a deliberate choice, motivated by the fact that the analytical results, which can be derived in such a complex case, are simpler and hence easier to interpret than the analogous results in the real case. All qualitative aspects of the disturbance rejection problem, which are of our main interest here, are in both cases the same. Of course, most of the signals and systems encountered in practice are real-valued. Later on we will present a simple way of extending our algorithms to real-valued signals.

II. ASSUMPTIONS

We will assume that the measurement noise in (1) obeys

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

and that the coefficient $a(t)$ in (2) evolves according to the random-walk (RW) model

$$a(t) = a(t-1) + e(t) \qquad (3)$$

where the sequence of one-step changes obeys

(A2) $\{e(t)\}$, independent of $\{v(t)\}$, is a zero-mean circular white sequence with variance σ_e^2.

Even though from a practical viewpoint the RW model can be criticized as rather naïve, it has two obvious advantages. First, it allows one to obtain analytical results. Second, and more importantly, under the random-walk hypothesis, one can establish a Cramér-Rao-type estimation bound that limits the efficiency of *any* disturbance rejection scheme. This will allow us to evaluate the performance of the proposed algorithms in absolute, rather than relative, terms. To derive such a bound, we will need the following additional assumption:

(A3) The sequences $\{v(t)\}$ and $\{e(t)\}$ are normally distributed: $v(t) \sim \mathcal{N}(0, \sigma_v^2)$, $e(t) \sim \mathcal{N}(0, \sigma_e^2)$.

The only assumption that will be made about the unknown plant is that it is stable and has nonzero gain at the frequency ω_0:

(A4) $K_{\mathrm{p}}(q^{-1}) = \sum_{i=0}^{\infty} k_i q^{-i}$, $\sum_{i=0}^{\infty} |k_i| < \infty$, $K_{\mathrm{p}}(e^{-j\omega_0}) \neq 0$.

Although in our primary design we will assume that the transfer function of the plant is fixed and known, later on we will propose a special adaptation mechanism capable of readjusting the control loop so as to make it insensitive to discrepancy between the nominal (assumed) transfer function and the true one. Hence, the nominal gain will be considered only as a convenient starting point for the adaptive regulator.

III. OPEN-LOOP CASE

Consider the problem of one-step-ahead prediction/compensation of a signal governed by

$$s(t) = d(t) + v(t) \qquad (4)$$

where $d(t)$ is a nonstationary cisoid, described in the previous section, and $v(t)$ denotes white measurement noise. Combining (2)–(4), one arrives at the following state-space equations

$$d(t) = e^{j\omega_0} d(t-1) + \widetilde{e}(t)$$
$$s(t) = d(t) + v(t) \qquad (5)$$

where $\widetilde{e}(t) = e^{j\omega_0 t} e(t)$. Under (A2) the sequence $\{\widetilde{e}(t)\}$ is circular white with variance σ_e^2.

Denote by $\mathcal{S}(t) = \{s(1), \ldots, s(t)\}$ the set of measurements available at instant t. The optimal, in the mean-square sense, one-step-ahead predictor of $s(t)$ has the form [9] $\widehat{s}(t|t-1) = \mathrm{E}[s(t)|\mathcal{S}(t-1)] = \widehat{d}(t|t-1)$, where $\widehat{d}(t|t-1) = \mathrm{E}[d(t)|\mathcal{S}(t-1)]$ is a one-step-ahead predictor of $d(t)$. The mean-squared prediction error can be expressed in the form $\mathrm{E}[|s(t) - \widehat{s}(t|t-1)|^2] = \mathrm{E}[|c(t)|^2] + \sigma_v^2$, where $c(t) = d(t) - \widehat{d}(t|t-1)$ will be called the cancellation error.

Under assumptions (A1)–(A3), the optimal estimates of $d(t)$ can be computed recursively using the Kalman filter algorithm

$$\widehat{d}(t|t-1) = e^{j\omega_0} \widehat{d}(t-1|t-1)$$
$$p(t|t-1) = p(t-1|t-1) + \sigma_e^2$$
$$g(t) = \frac{p(t|t-1)}{\sigma_v^2 + p(t|t-1)}$$
$$\varepsilon(t) = s(t) - \widehat{d}(t|t-1)$$
$$\widehat{d}(t|t) = \widehat{d}(t|t-1) + g(t)\varepsilon(t)$$
$$p(t|t) = p(t|t-1) - \frac{p^2(t|t-1)}{\sigma_v^2 + p(t|t-1)} \qquad (6)$$

where $\widehat{d}(t|t) = \mathrm{E}[d(t)|\mathcal{S}(t)]$ denotes the filtered estimate of $d(t)$, and $p(t|t-1) = \mathrm{E}[|d(t) - \widehat{d}(t|t-1)|^2]$ and $p(t|t) = \mathrm{E}[|d(t) - \widehat{d}(t|t)|^2]$ are the *a priori* and *a posteriori* error variances, respectively. The steady-state value of the mean-squared cancellation error $\mathrm{E}[|c(t)|^2] = p(t|t-1)$, equal to $p_\infty = \lim_{t \to \infty} p(t|t-1)$, can be obtained by solving the associated Riccati equation: $p_\infty^2/(\sigma_v^2 + p_\infty) = \sigma_e^2$. This results in

$$p_\infty = \frac{\sigma_e^2 + \sqrt{\sigma_e^4 + 4\sigma_e^2 \sigma_v^2}}{2}. \qquad (7)$$

Note that in steady state, the Kalman filter (6) reduces to

$$\widehat{d}(t+1|t) = e^{j\omega_0}[\,\widehat{d}(t|t-1) + g_\infty \varepsilon(t)\,]$$
$$\varepsilon(t) = s(t) - \widehat{d}(t|t-1) \qquad (8)$$

where $g_\infty = \lim_{t \to \infty} g(t) = p_\infty/(\sigma_v^2 + p_\infty)$, $0 < g_\infty < 1$, is the optimal (real-valued) steady-state gain. Note also that, under assumptions (A1)–(A3), p_∞ is the lower bound on the mean-squared cancellation error, called the Bayesian Cramér-Rao bound [10] (the classical Cramér-Rao bound does not apply to systems/signals with random parameters). This means that no other compensation scheme can do better than the one described above.

IV. CLOSED-LOOP CASE – A PRELIMINARY SOLUTION

Consider a system governed by (1). In order to reject the narrowband disturbance $d(t)$, one should generate another narrowband signal $u(t)$, which, after passing through the plant, will have the same amplitude as $d(t)$ but opposite polarity. Since linear systems basically scale and shift sinusoidal inputs, the narrowband character of $u(t)$ justifies the following steady-state approximation

$$K_{\mathrm{p}}(q^{-1})u(t-1) \cong k_{\mathrm{p}} u(t-1) \qquad (9)$$

where $k_{\mathrm{p}} = K_{\mathrm{p}}(e^{-j\omega_0})$. Some further comments on this approximation will be given in Remark 2 at the end of this section.

To design our feedback controller, we will assume that the transfer function of the plant is equal to some "nominal" transfer function $K_{\mathrm{n}}(q^{-1})$. Since $K_{\mathrm{p}}(q^{-1})$ is usually unknown and/or time-varying, in general it will hold that $K_{\mathrm{n}}(q^{-1}) \neq K_{\mathrm{p}}(q^{-1})$. Denote by $k_{\mathrm{n}} = K_{\mathrm{n}}(e^{-j\omega_0}) \neq 0$ the nominal plant gain at the frequency ω_0. The proposed control rule has the form

$$u(t) = -\frac{\widehat{d}(t+1|t)}{k_{\mathrm{n}}} \qquad (10)$$

where $\widehat{d}(t+1\,|\,t)$ is obtained from

$$\widehat{d}(t+1\,|\,t) = e^{j\omega_0}\big[\,\widehat{d}(t\,|\,t-1) + \mu y(t)\,\big] \qquad (11)$$

and $\mu > 0$ denotes adaptation gain.

To check performance of the closed-loop system, observe that

$$y(t) \cong k_{\mathrm{p}}u(t-1) + d(t) + v(t)$$
$$= d(t) - \beta\widehat{d}(t\,|\,t-1) + v(t) = c(t) + v(t) \qquad (12)$$

where $\beta = k_{\mathrm{p}}/k_{\mathrm{n}}$ is the ratio of the true plant gain to its nominal (assumed) gain and $c(t)$ denotes the cancellation error. Combining (11) and (12) one obtains

$$c(t) = e^{j\omega_0}(1 - \mu\beta)c(t-1) + \widetilde{e}(t) - \mu\beta\widetilde{v}(t-1) \qquad (13)$$

where $\widetilde{v}(t) = e^{j\omega_0}v(t)$. Note that, similarly to $\{v(t)\}$, $\{\widetilde{v}(t)\}$ is a circular white noise signal with variance σ_v^2.

Observe that $\mathrm{E}[c(t)] = e^{j\omega_0}(1 - \mu\beta)\mathrm{E}[c(t-1)]$. Hence, if

$$|\,1 - \mu\beta\,| < 1 \qquad (14)$$

then $\lim_{t\to\infty} \mathrm{E}[c(t)] = \mathrm{E}_\infty[c(t)] = 0$, which means that the steady-state mean value of the cancellation error is zero – even if $\beta \neq 1$, i.e., if the true gain of the plant at frequency ω_0 differs from the assumed gain.

Let $\beta = |\,\beta\,|e^{j\phi}$, where $|\,\beta\,|$ is the magnitude ratio and ϕ denotes the phase modeling error. It is straightforward to check that condition (14) is equivalent to

$$\mu|\,\beta\,| < 2\cos\phi \qquad (15)$$

in agreement with [3, sec. 3.3.5].

According to (15), the disturbance rejection scheme (10)–(11) is unstable if $\cos\phi < 0$, i.e., if the absolute value of the phase error exceeds $\pi/2$. On the contrary, when $|\,\phi\,| < \pi/2$, one can always choose μ so as to satisfy (14) – irrespective of the magnitude error. When the stability condition is fulfilled, one can easily derive and minimize the expression for the steady-state mean-squared cancellation error. Straightforward calculations lead to

$$\mathrm{E}_\infty[|\,c(t)\,|^2] = \frac{\sigma_e^2 + |\,\mu\beta\,|^2\sigma_v^2}{1 - |\,1 - \mu\beta\,|^2} \qquad (16)$$

and

$$\mu_{\mathrm{opt}} = \arg\min_{\mu\in\mathrm{R}_+} \mathrm{E}_\infty[|\,c(t)\,|^2]$$
$$= -\frac{\zeta}{\beta + \beta^*} + \sqrt{\frac{\zeta^2}{(\beta + \beta^*)^2} + \frac{\zeta}{\beta\beta^*}} \qquad (17)$$

where

$$\zeta = \sigma_e^2/\sigma_v^2$$

and $*$ denotes complex conjugation. One can check that when there is no phase error, i.e., $\phi = 0$, it holds that $\mu_{\mathrm{opt}}\beta = \mu_{\mathrm{opt}}|\,\beta\,| = g_\infty$ and $\mathrm{E}_\infty[|\,c(t)\,|^2\,|\,\mu = \mu_{\mathrm{opt}}] = p_\infty$. Therefore, no matter how large the magnitude mismatch, one can always choose an adaptation gain μ that will make the disturbance rejection scheme statistically efficient (in the Gaussian case).

To see what happens in the presence of phase mismatch, we will assume that the ratio ζ is small (slow variation hypothesis), namely $\zeta \ll 4\cos^2\phi$. In this case $\mu_{\mathrm{opt}} \cong \sqrt{\zeta}/|\,\beta\,|$, $p_\infty \cong \sigma_e\sigma_v$ and

$$\mathrm{E}_\infty[|\,c(t)\,|^2\,|\,\mu = \mu_{\mathrm{opt}}] \cong \frac{\sigma_e\sigma_v}{\cos\phi} \cong \frac{p_\infty}{\cos\phi} \geq p_\infty \qquad (18)$$

which means that even if μ is chosen in the optimal way, one may face substantial losses in rejection efficiency for large phase errors.

After the discussion presented above, the suggested disturbance rejection scheme can be summarized as follows

1) Stability can be achieved for arbitrary magnitude errors and limited (smaller than $\pi/2$) phase errors.
2) Statistical efficiency (for arbitrary magnitude errors) is possible only in the absence of phase errors.
3) The "true" adaptation gain of the algorithm (11), used for amplitude tracking, is equal to $\mu\beta$ and hence depends not only on μ but also on modeling errors β, e.g., when $\beta = 0.5$ or $\beta = 2$, the estimation memory of this algorithm is effectively doubled or halved, respectively.
4) The optimal value of μ, guaranteeing the best achievable performance, depends on the value of β, so even if the characteristics of the disturbance (σ_e^2) and noise (σ_v^2) are known, it is not possible to determine μ_{opt}.

In the next section, we will describe an extended disturbance rejection scheme free of the drawbacks mentioned above.

Remark 1: When deriving the feedback control rule (10), we have assumed that $\beta = 1$. This resembles and leads to the same qualitative results as the coefficient fixing technique used to "robustify" adaptive minimum-variance (MV) regulators. The MV control rule is a function of plant parameters. When these parameters are not known *a priori*, or when they vary with time, adaptive versions of MV regulators can be used, where the true (unknown) coefficients are replaced with their current estimates, obtained by means of system identification. However, since identification is carried in a closed loop, the input/output data do not contain enough information to unambiguously estimate all plant parameters – one degree of freedom is "taken away by the loop." To avoid identifiability problems and to prevent parameter estimates from drifting, one of the system coefficients – call it b_0 – is usually fixed at a preselected, nominal value b_{n} and not estimated [11].

It can be shown that for a wide range of ratios b_0/b_{n} the resulting adaptive regulator converges to the optimal (MV) regulator, which means that the modeling bias is automatically compensated by feedback. The coefficient fixing technique has some limitations. First, one should get the sign of b_0/b_{n} right, namely it should hold that $b_0/b_{\mathrm{n}} > 0$ (note that our condition $|\,\phi\,| < \pi/2$ can be interpreted as a sign requirement imposed on the real part of $\beta = k_{\mathrm{p}}/k_{\mathrm{n}}$: $\mathrm{Re}[\beta] > 0$). Second, as shown in [12], [13], when finite-memory identification is combined with coefficient fixing, the "closed-loop memory" of the estimator depends on the ratio b_0/b_{n} in a similar way as was observed for the algorithm (11).

Remark 2: The expression (16), which describes dependence of the mean-squared cancellation error on μ, was based on the assumption that the plant's response to narrowband
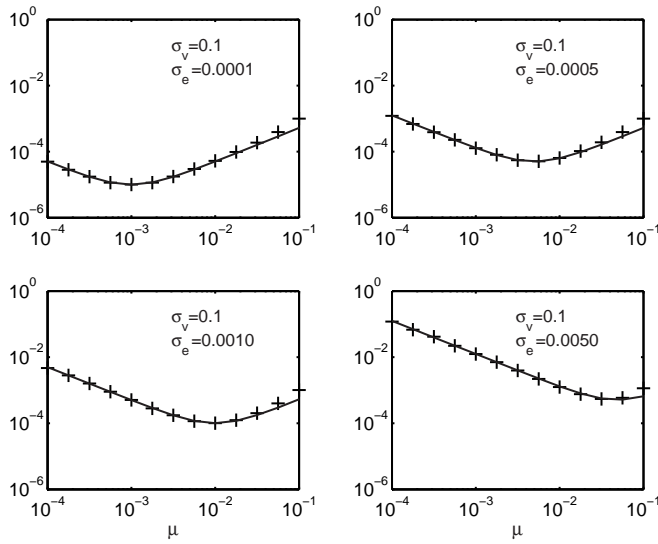
Fig. 2. Comparison of theoretical values of the mean-squared cancellation error, obtained using the steady-state plant approximation (solid line), with the experimental values (crosses).

excitation can be closely approximated by an analogous response of a much simpler "scale and shift" model, based on frequency-domain concepts. A special simulation experiment was arranged to check how well the resulting theoretical formula fits the true error values. The simulated discrete-time plant (Guo–Bodson)

$$K_{\mathrm{p}}(q^{-1}) = \frac{0.0952}{1 - 0.9048q^{-1}} \tag{19}$$

was adopted from [8] 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. Simulations were carried out for $\sigma_v = 0.1$ and for four different rates of amplitude variation $\sigma_e \in \{0.0001, 0.0005, 0.001, 0.005\}$, in the absence of modeling errors ($\beta = 1$). For each (σ_e, μ) pair, the experiment, covering 100000 time-steps, was repeated 500 times for different realizations of $\{v(t)\}$ and $\{e(t)\}$. In all cases, $d(0)$ was set to 1 and $\widehat{d}(0)$ was set to $e^{j\omega_0}$.

The results, summarized in Fig. 2, were obtained by means of combined ensemble and time averaging, after discarding the first 50000 samples (to ensure that the steady-state conditions are reached). Note the good agreement of experimental values with theoretical expectations for the considered (and practically meaningful) range of adaptation gains.

More on the practical side – Fig. 2 clearly shows the benefits that stem from tuning the adaptation gain μ to the degree of system nonstationarity, exactly as predicted by theory.

V. CLOSED-LOOP CASE – THE REFINED SOLUTION

To optimize performance of (10)–(11) and, perhaps even more importantly, to make the proposed disturbance rejection scheme insensitive to modeling errors, we will equip it with an automatic adaptation gain tuning mechanism. It is clear from the discussion in the previous section that in the presence of phase errors ($\phi \neq 0$), no choice of the real-valued adaptation gain μ can prevent performance of the rejection algorithm from

degradation. The situation changes if one allows μ to take complex values. It is straightforward to check that

$$\mu_{\mathrm{opt}} = \arg\min_{\mu \in \mathrm{C}} \mathrm{E}_\infty[|\,c(t)|^2] = \frac{1}{\beta}\left[-\frac{\zeta}{2} + \sqrt{\frac{\zeta^2}{4} + \zeta}\right] = \frac{g_\infty}{\beta} \tag{20}$$

and $\mathrm{E}_\infty[\,|\,c(t)|^2\,|\,\mu = \mu_{\mathrm{opt}}\,] = p_\infty$ for *any* value of ϕ. Having this in mind, we will design an adaptive algorithm for on-line tuning of a complex-valued adaptation gain μ.

We adjust μ recursively by minimizing the following local measure of fit, made up of exponentially weighted squares of system outputs

$$V(t, \mu) = \sum_{\tau=1}^{t} \rho^{t-\tau}\,|\,y(\tau, \mu)|^2 \tag{21}$$

where the forgetting constant ρ ($0 < \rho < 1$) determines the effective averaging range. To evaluate the estimate $\widehat{\mu}(t) = \arg\min_\mu V(t, \mu)$, we will use the recursive prediction error (RPE) approach [14]

$$\widehat{\mu}(t) = \widehat{\mu}(t-1) - \{V''[t, \widehat{\mu}(t-1)]\}^{-1}\,V'[t, \widehat{\mu}(t-1)] \tag{22}$$

where

$$V'[t, \widehat{\mu}(t-1)] \cong \left(\frac{\partial y[t, \widehat{\mu}(t-1)]}{\partial \mu}\right)^* y[t, \widehat{\mu}(t-1)]$$

$$V''[t, \widehat{\mu}(t-1)] \cong \rho V''[t-1, \widehat{\mu}(t-2)] + \left|\frac{\partial y[t, \widehat{\mu}(t-1)]}{\partial \mu}\right|^2$$

and

$$\frac{\partial}{\partial \mu} = \frac{1}{2}\left[\frac{\partial}{\partial \mathrm{Re}[\mu]} - j\frac{\partial}{\partial \mathrm{Im}[\mu]}\right]$$

denotes symbolic differentiation with respect to a complex variable μ. Such an approach, usually referred to as Wirtinger calculus or CR (complex real) calculus, was introduced in [15]. It is based on the concept of real differentiability of complex functions and is applicable to nonanalytic functions, such as (21) – see [16]–[18] for more details.

Applying standard rules of Wirtinger calculus to (11) and (12), one obtains

$$\frac{\partial y(t)}{\partial \mu} = -\beta\frac{\partial \widehat{d}(t\,|\,t-1)}{\partial \mu}$$

$$\frac{\partial \widehat{d}(t+1\,|\,t)}{\partial \mu} = e^{j\omega_0}\frac{\partial \widehat{d}(t\,|\,t-1)}{\partial \mu} + e^{j\omega_0}y(t) + \mu e^{j\omega_0}\frac{\partial y(t)}{\partial \mu}$$

leading to

$$\frac{\partial y(t)}{\partial \mu} = e^{j\omega_0}\left[(1-\mu\beta)\frac{\partial y(t-1)}{\partial \mu} - \beta y(t-1)\right]. \tag{23}$$

Since the value of β is not known, the obtained recursive formula can't be used in its present form. The way out of difficulty is to apply once more the gain fixing technique. Setting $\beta = 1$ in (23), one arrives at

$$\frac{\partial y(t)}{\partial \mu} = e^{j\omega_0}\left[(1-\mu)\frac{\partial y(t-1)}{\partial \mu} - y(t-1)\right]. \tag{24}$$

Note, however, that to guarantee stable operation of (24), one must have $|\,1 - \mu\,| < 1$, which puts some constraints on the phase of μ: $\mathrm{Arg}[\mu] < \pi/2$, where $\mathrm{Arg}[x] \in (-\pi, \pi]$ denotes

principal argument of a complex number x. This means that full correction of phase errors will not be possible using this approach. To overcome the above-mentioned limitation, one can set $\beta = c_\mu/\mu$, where c_μ is a small positive constant. This leads to [1]

$$\frac{\partial y(t)}{\partial \mu} = e^{j\omega_0}\left[(1-c_\mu)\frac{\partial y(t-1)}{\partial \mu} - \frac{c_\mu}{\mu}y(t-1)\right]. \quad (25)$$

Let $r(t) = V''[t, \widehat{\mu}(t-1)]$ and $z(t) = \partial y[t, \widehat{\mu}(t-1)]/\partial \mu$. Then the proposed SONIC algorithm can be summarized as follows

$$z(t) = e^{j\omega_0}\left[(1-c_\mu)z(t-1) - \frac{c_\mu}{\widehat{\mu}(t-1)}y(t-1)\right]$$
$$r(t) = \rho r(t-1) + |z(t)|^2$$
$$\widehat{\mu}(t) = \widehat{\mu}(t-1) - \frac{z^*(t)y(t)}{r(t)}$$
$$\widehat{d}(t+1|t) = e^{j\omega_0}[\,\widehat{d}(t|t-1) + \widehat{\mu}(t)y(t)\,]$$
$$u(t) = -\frac{\widehat{d}(t+1|t)}{k_n}. \quad (26)$$

The recommended values of c_μ and ρ are those from the intervals [0.005,0.05] and [0.999,0.9999], respectively.

The disturbance tracking loop [governed by the fourth equation of (26)] can be easily recognized as an almost standard LMS-type recursion, i.e., a general-purpose estimation algorithm, capable of tracking *any* types of changes. Its only nonstandard (but very important) feature is that the step-size μ is complex-valued and automatically adjusted — this allows one to account for unknown, and possibly time-varying, characteristics of the controlled dynamic process. Note also, that the self-optimizing loop [governed by the first three equation of (26)], by the very nature of the RPE approach, will tend to minimize prediction (output) errors *irrespective* of their origin, e.g., it should respond adequately to abrupt plant changes, or to changes in the level of the measurement noise, even though such forms of nonstationarity were not taken into consideration when deriving the disturbance canceling algorithm.

Hence, although obtained under specific, and rather unrealistic, assumptions (disturbance changing according to the RW model, time-invariant plant, white measurement noise of constant intensity), the algorithm (26) can be safely used under more realistic conditions that are encountered in practical applications – some simulation evidence for this robustness claim will be given in Section X.

VI. TRACKING ANALYSIS

We will show that, under assumptions (A1)–(A4), the proposed adaptive filter converges in mean to the optimal solution

$$\mathrm{E}_\infty[\widehat{\mu}(t)] \cong \mu_{\mathrm{opt}} = \frac{g_\infty}{\beta}. \quad (27)$$

We will also shed light on the role played by design variables ρ and c_μ.

[1]Note that, according to (20), the desired value of $\mu\beta$ is *real*. Hence, setting $\mu\beta = c_\mu$ is to some extent a natural choice. Note also that as long as $c_\mu < 2$ the modified recursion is stable, irrespective of the phase of μ.

Our tracking analysis will be based on studying the properties of a stochastic differential equation (SDE) associated with algorithm (26). Since strict mathematical conditions under which such an SDE-based approach is applicable are not specified (one of the prerequisites is stochastic stability of the analyzed system, which is a difficult problem to resolve on its own), the "theoretical" results derived below must be carefully verified experimentally. This will be done in Section X.

To avoid unnecessary complications, we will initially examine the tracking properties of the simplified (gradient) version of algorithm (26), obtained by replacing the RPE updates [the second and third lines of (26)] with

$$\widehat{\mu}(t) = \widehat{\mu}(t-1) - \alpha y(t)z^*(t) \quad (28)$$

where $\alpha > 0$ denotes a small gain. Later on, we will extend the results of our analysis to the normalized case, where the constant gain is replaced with a recursively computed normalizing factor $1/r(t)$.

It is known that the tracking behavior of constant-gain (finite-memory) estimation algorithms, such as (28), can be studied by examining the properties of the associated difference equations [19], [20]. Denote by $\{y(t;\mu)\}$ and $\{z(t;\mu)\}$ stationary processes that "settle down" in the closed-loop system for a constant value of μ: $\widehat{\mu}(t) \equiv \mu \in \Omega_s$, where Ω_s is the stability region. Furthermore, let μ_0 be the stable "equilibrium" point of (28) obeying

$$f(\mu_0) = 0 \quad (29)$$

$$\mathrm{Re}[f'(\mu_0)] > 0, \quad |f'(\mu_0)|^2 > \left|f^\dagger(\mu_0)\right|^2 \quad (30)$$

where

$$f(\mu) = \mathrm{E}[y(t;\mu)z^*(t;\mu)]$$

and

$$f'(\mu) = \partial f(\mu)/\partial \mu, \quad f^\dagger(\mu) = \partial f(\mu)/\partial \mu^*.$$

According to [19], [20], when the coefficient α is sufficiently small, the evolution of the estimation error $\widehat{\mu}(t) - \mu_0$ can be approximately described by the following linearized stochastic differential equation (provided that tracking is satisfactory, i.e., $\widehat{\mu}(t)$ remains close to μ_0)

$$dX_s = -\alpha f'(\mu_0)X_s ds - \alpha f^\dagger(\mu_0)X_s^* ds$$
$$+ \alpha\sqrt{g(\mu_0)}\,dW_s \quad (31)$$

$$X_s \cong \widehat{\mu}(t) - \mu_0 \quad \text{for} \quad s = t$$

where s denotes continuous time, $\{W_s\}$ is a standard complex-valued Wiener process and

$$g(\mu) = \sum_{\tau=-\infty}^{\infty} \mathrm{E}\left[y(\tau;\mu)z^*(\tau;\mu)y^*(0;\mu)z(0;\mu)\right] \quad (32)$$

(the series being assumed convergent).

A. Equilibrium Point

We will show that $\mu_0 = \mu_{\text{opt}}$. For two jointly stationary processes $\{x(t)\}$ and $\{y(t)\}$, define $R_{xy}(\tau) = \mathrm{E}[x(t)y^*(t-\tau)]$. Note that $f(\mu) = R_{yz}(0; \mu)$. Using the steady-state approximation [cf. (11)–(12)] one arrives at

$$y(t) = (1-\mu\beta)e^{j\omega_0}y(t-1) + \widetilde{e}(t) + v(t) - e^{j\omega_0}v(t-1) \quad (33)$$

(for brevity, the dependence on μ is temporarily dropped). Combining (26) with (33), one obtains

$$\mathrm{E}[y(t)z^*(t)]$$
$$= \mathrm{E}\left\{[(1-\mu\beta)e^{j\omega_0}y(t-1) + \widetilde{e}(t) + v(t) - e^{j\omega_0}v(t-1)]\right.$$
$$\left. \times \left[(1-c_\mu)e^{-j\omega_0}z^*(t-1) - \frac{c_\mu}{\mu^*}e^{-j\omega_0}y^*(t-1)\right]\right\}$$

which, after elimination of cross terms that are zero due to orthogonality, leads to the following recursive relationship

$$R_{yz}(0) = (1-c_\mu)(1-\mu\beta)R_{yz}(0)$$
$$- \frac{c_\mu}{\mu^*}(1-\mu\beta)R_{yy}(0) + \frac{c_\mu}{\mu^*}R_{vy}(0). \quad (34)$$

In an analogous way, the relationship

$$\mathrm{E}[y(t)y^*(t)] = \mathrm{E}\{[(1-\mu\beta)e^{j\omega_0}y(t-1) + \widetilde{e}(t) + v(t)$$
$$- e^{j\omega_0}v(t-1)] \cdot [(1-\mu^*\beta^*)e^{-j\omega_0}y^*(t-1)$$
$$+ \widetilde{e}^*(t) + v^*(t) - e^{-j\omega_0}v^*(t-1)]\}$$

implies

$$R_{yy}(0) = |1-\mu\beta|^2 R_{yy}(0) + \sigma_e^2 + 2\sigma_v^2$$
$$- (1-\mu\beta)R_{yv}(0) - (1-\mu^*\beta^*)R_{vy}(0). \quad (35)$$

Note that $R_{yv}(0) = R_{vy}(0) = \sigma_v^2$. Solving (34) and (35) with respect to $R_{yy}(0)$ and $R_{yz}(0)$, one obtains

$$R_{yy}(0) = \frac{\sigma_e^2 + |\mu\beta|^2\sigma_v^2}{1 - |1-\mu\beta|^2} + \sigma_v^2 \quad (36)$$

$$R_{yz}(0) = -\frac{c_\mu[(1-\mu\beta)R_{yy}(0) - \sigma_v^2]}{[1-(1-c_\mu)(1-\mu\beta)]\mu^*}. \quad (37)$$

The equilibrium point μ_0 can be determined by solving $f(\mu_0) = R_{yz}(0; \mu_0) = 0$ or equivalently

$$(1-\mu_0\beta)R_{yy}(0; \mu_0) = \sigma_v^2. \quad (38)$$

Let $x_0 = 1 - \mu_0\beta$. Equation (38) can be rewritten in the form [cf. (36)]

$$x_0\left(\frac{\zeta + |1-x_0|^2}{1 - |x_0|^2} + 1\right) = 1. \quad (39)$$

Since, according to (39), x_0 must be a real number, one finally obtains

$$x_0 = 1 + \zeta/2 \pm \sqrt{\zeta^2/4 + \zeta}, \quad \mu_0\beta = -\zeta/2 \pm \sqrt{\zeta^2/4 + \zeta}.$$

To guarantee stability of the closed-loop system, one must require that $\mu_0\beta > 0$ which leads to [cf. (20)] $\mu_0\beta = g_\infty$. Since

$$\mu_0 = \arg\min_{\mu \in \mathbb{C}} R_{yy}(0; \mu) \quad (40)$$

the equilibrium point established above corresponds to the optimal (minimum-variance) control strategy.

B. Local Stability

To prove stability of the equilibrium point, we have to verify (30). Note that $R_{yz}(0; \mu) = -N(\mu)/D(\mu)$, where $N(\mu) = c_\mu[(1-\mu\beta)R_{yy}(0) - \sigma_v^2]$ and $D(\mu) = [1-(1-c_\mu)(1-\mu\beta)]\mu^*$. According to (38), it holds that $N(\mu_0) = 0$. Additionally, since [cf. (40)] $R'_{yy}(0; \mu_0) = 0$, one obtains $N'(\mu_0) = -c_\mu\beta R_{yy}(0; \mu_0)$. Combining these results one arrives at

$$f'(\mu_0) = R'_{yz}(0; \mu_0) = -\frac{N'(\mu_0)}{D(\mu_0)}$$
$$= \frac{c_\mu g_\infty R_{yy}(0; \mu_0)}{|\mu_0|^2[1-(1-c_\mu)(1-g_\infty)]} > 0 \quad (41)$$

In an analogous way one can show that $f^\dagger(\mu_0) = -N^\dagger(\mu_0)/D(\mu_0) = 0$ which, when combined with (41), proves that μ_0 is a stable equilibrium point.

C. Results for the Normalized Algorithm

Now consider the normalized algorithm (26) where

$$\widehat{\mu}(t) = \widehat{\mu}(t-1) - \frac{y(t)z^*(t)}{r(t)}$$
$$r(t) = \rho r(t-1) + |z(t)|^2. \quad (42)$$

Let $\gamma = 1 - \rho$. For constant μ and for ρ sufficiently close to 1, it holds that $r(t) \cong h(\mu)/\gamma$ where $h(\mu) = \mathrm{E}[|z(t; \mu)|^2] = R_{zz}(0; \mu)$. Hence, the normalized algorithm can be analyzed analogously to the gradient algorithm (28), provided that the gain α is set to $\gamma R_{zz}^{-1}(0; \mu_0)$. Note that this modification does not qualitatively affect the results reported in the previous subsections – similarly as before, $\mu_0 = \mu_{\text{opt}}$ is the only stable equilibrium point of (42) satisfying (14). Since $f^\dagger(\mu_0) = 0$, the stochastic differential equation associated with (42) has the form

$$dX_s = -\gamma h^{-1}(\mu_0)f'(\mu_0)X_s ds$$
$$+ \gamma h^{-1}(\mu_0)\sqrt{g(\mu_0)}\, dW_s. \quad (43)$$

To evaluate $h(\mu_0) = R_{zz}(0; \mu_0)$, note that

$$\mathrm{E}[z(t)z^*(t)] = \mathrm{E}\left\{[(1-c_\mu)z(t-1) - \frac{c_\mu}{\mu}y(t-1)]\right.$$
$$\left. \times [(1-c_\mu)z^*(t-1) - \frac{c_\mu}{\mu^*}y^*(t-1)]\right\}$$

and hence

$$R_{zz}(0) = (1-c_\mu)^2 R_{zz}(0) - \frac{c_\mu(1-c_\mu)}{\mu^*}R_{zy}(0)$$
$$- \frac{c_\mu(1-c_\mu)}{\mu}R_{yz}(0) + \frac{c_\mu^2}{|\mu|^2}R_{yy}(0).$$

Since $R_{zy}(0; \mu_0) = R_{yz}(0; \mu_0) = 0$, one obtains

$$h(\mu_0) = R_{zz}(0; \mu_0) = R_{zz}^*(0; \mu_0) = \frac{c_\mu R_{yy}(0; \mu_0)}{|\mu_0|^2(2-c_\mu)}. \quad (44)$$

D. Variability

To study stochastic variability of $\widehat{\mu}(t) - \mu_0$ we have to evaluate $g(\mu_0)$ in (31). Note that under (A1)–(A3), the processes $\{y(t)\}$ and $\{z(t)\}$ are zero-mean and Gaussian. Therefore it holds that [18]

$$\eta(\tau; \mu_0) = \mathrm{E}[\, y(\tau; \mu_0) z^*(\tau, \mu_0) y^*(0; \mu_0) z(0; \mu_0)\,]$$
$$= I_1(\tau) + I_2(\tau) + I_3(\tau)$$

where

$$I_1(\tau) = \mathrm{E}[\, y(\tau; \mu_0) z^*(\tau, \mu_0)\,]\, \mathrm{E}[\, y^*(0; \mu_0) z(0; \mu_0)\,]$$
$$= |R_{yz}(0; \mu_0)|^2$$
$$I_2(\tau) = \mathrm{E}[\, y(\tau; \mu_0) y^*(0, \mu_0)\,]\, \mathrm{E}[\, z^*(\tau; \mu_0) z(0; \mu_0)\,]$$
$$= R_{yy}(\tau; \mu_0) R_{zz}^*(\tau; \mu_0)$$
$$I_3(\tau) = \mathrm{E}[\, y(\tau; \mu_0) z(0, \mu_0)\,]\, \mathrm{E}[\, z^*(\tau; \mu_0) y^*(0; \mu_0)\,]$$
$$= R_{yz^*}(\tau; \mu_0) R_{z^*y}(\tau; \mu_0).$$

Since $R_{yz}(0; \mu_0) = 0$, it holds that $I_1(\tau) \equiv 0$. Moreover, since the processes $\{v(t)\}$ and $\{e(t)\}$ were assumed to be circular white, one obtains $R_{vv^*}(\tau) = R_{ww^*}(\tau) = 0, \forall \tau$, which entails $R_{yz^*}(\tau; \mu_0) = R_{z^*y}(\tau; \mu_0) = 0, \forall \tau$ and leads to $I_3(\tau) \equiv 0$. Furthermore, we will show that

$$I_2(\tau) = \begin{cases} R_{yy}(0; \mu_0) R_{zz}(0; \mu_0), & \tau = 0 \\ 0, & \tau \neq 0 \end{cases}. \quad (45)$$

Actually, note that

$$\mathrm{E}[y(t)y^*(t-1)] = \mathrm{E}\{[(1 - \mu\beta)e^{j\omega_0}y(t-1) + e(t)$$
$$+ v(t) - e^{j\omega_0}v(t-1)]\, y^*(t-1)\}$$

which leads to $R_{yy}(1) = e^{j\omega_0}[(1 - \mu\beta)R_{yy}(0) - R_{vy}(0)]$. In an analogous way one can show that $R_{yy}(\tau) = e^{j\omega_0}(1 - \mu\beta)R_{yy}(\tau - 1)$, $\forall \tau > 1$. Since [cf. (38)] $R_{yy}(1; \mu_0) = e^{j\omega_0}[(1 - \mu_0\beta)R_{yy}(0; \mu_0) - \sigma_v^2] = 0$, one obtains $R_{yy}(\tau; \mu_0) = R_{yy}^*(-\tau; \mu_0) = 0, \forall \tau \neq 0$, which leads to (45).

Finally, after combining all results presented above, one gets

$$g(\mu_0) = \sum_{\tau=-\infty}^{\infty} \eta(\tau; \mu_0) = R_{yy}(0; \mu_0) R_{zz}(0; \mu_0). \quad (46)$$

Derivation of (46) in the non-Gaussian case, i.e., under (A1)–(A2) only, is also possible, but much more tedious.

E. Tracking Properties of the Normalized Algorithm

Since the properties of the closed-loop system depend on the value of $\widehat{\mu}(t)\beta$, rather than on the value of $\widehat{\mu}(t)$, we will introduce a new variable $Y_s = \beta X_s$. Multiplying both sides of (43) with β one arrives at the following differential equation

$$dY_s = -\gamma h^{-1}(\mu_0) f'(\mu_0) Y_s ds + \gamma\beta h^{-1}(\mu_0)\sqrt{g(\mu_0)}\, dW_s \quad (47)$$

which can be used to study the evolution of $\widehat{\mu}(t)\beta$ in the neighborhood of the equilibrium point $\mu_0\beta = g_\infty$.

Using (41), (44), and (46) one can rewrite (47) in the form

$$dY_s = -b Y_s ds + c\, dW_s \quad (48)$$

where

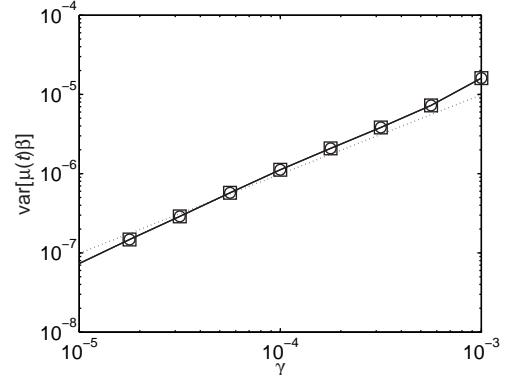$$b = \gamma h^{-1}(\mu_0) f'(\mu_0) = \gamma g_\infty \frac{2 - c_\mu}{1 - (1 - c_\mu)(1 - g_\infty)} \quad (49)$$

Fig. 3. Dependence of the steady-state variance of $\widehat{\mu}(t)\beta$ on γ for the Guo-Bodson's plant $K_1(q^{-1})$. Solid lines – experimental results, dotted lines – theoretical predictions. Circles – $\beta = 1$, squares – $\beta = 4$.

$$c = \gamma\beta h^{-1}(\mu_0)\sqrt{g(\mu_0)} = \gamma|\mu_0|\beta\sqrt{(2 - c_\mu)/c_\mu}. \quad (50)$$

Based on (48)–(50), one can rationalize the choice of γ and c_μ:

1) Selection of γ: Denote by $\mathrm{E}_\infty[|\widehat{\mu}(t)\beta - g_\infty|^2] \cong \mathrm{E}_\infty[|Y_s|^2]$ the variance of fluctuations of $\widehat{\mu}(t)\beta$ around g_∞. Solving the Lyapunov equation associated with (48), one obtains $\mathrm{E}_\infty[|Y_s|^2] = |c|^2/(2b)$, which leads to

$$q_\infty = \mathrm{E}_\infty[|\widehat{\mu}(t)\beta - g_\infty|^2]$$
$$\cong \gamma g_\infty \frac{1 - (1 - c_\mu)(1 - g_\infty)}{2c_\mu}. \quad (51)$$

Quite clearly, to make the steady-state fluctuations of $\widehat{\mu}(t)\beta$ small, one should keep the coefficient γ sufficiently close to 0. On the other hand, as shown in [19], the closer that γ becomes to 0, the longer it takes for the algorithm to readjust the adaptation gain $\widehat{\mu}(t)$ when the plant changes. Hence, selection of γ is a classical variance/bias compromise, typical of identification of nonstationary systems [13]: for "small" values of γ, the estimation algorithm is "slow" (yields large tracking bias) but "accurate" (yields small tracking variance), whereas for "large" values of γ, it is "fast" but "inaccurate".

A special simulation experiment was arranged to check the dependence of the steady-state variance of $\widehat{\mu}(t)\beta$ on $\gamma = 1 - \rho$ for the Guo–Bodson plant (19). The disturbance and noise settings are $\sigma_v = 0.1$, $\sigma_e = 0.001$, $\omega_0 = 0.1$ and $d(0) = 1$. All numerical values are obtained by combined ensemble averaging (50 realizations of $\{v(t)\}$ and $\{e(t)\}$) and time averaging (100000 time-steps). For each realization, the first 25000 samples were discarded to ensure that only the steady-state values are averaged. Fig. 3 shows the results obtained for $\beta = 1$ (no modeling error) and for $\beta = 4$. The constant c_μ was set to 0.01. Note the good correspondence between experimental and theoretical curves in the considered range of RPE gains. The obtained results are practically insensitive to modeling errors.

2) Selection of c_μ: According to [19], the constant c_μ should be chosen so as to minimize the following measure
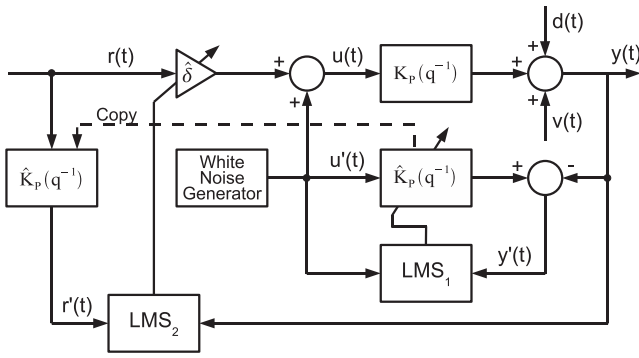
Fig. 4. Block diagram of the adaptive FXLMS algorithm with auxiliary noise plant identification.

of the tracking capability of the algorithm:

$$J(c_\mu) = \frac{b^2}{|c|^2} = \frac{(2 - c_\mu)c_\mu}{[1 - (1 - c_\mu)(1 - g_\infty)]^2}.$$

Straightforward calculations lead to

$$c_\mu^{\mathrm{opt}} = \arg\min_{c_\mu} J(c_\mu) = g_\infty. \qquad (52)$$

Note that in the case where $\sqrt{\zeta} \ll 1$ (slow rate of amplitude variation), it holds that $g_\infty \cong \sqrt{\zeta} = \sigma_e/\sigma_v$ [cf. (5)]. Then, for $c_\mu = g_\infty$, one obtains: $b = \gamma$, $|c| \cong \gamma\sqrt{2g_\infty}$ and $q_\infty \cong \gamma g_\infty$.

Remark: Since neither b nor $|c| = \gamma g_\infty \sqrt{(2 - c_\mu)/c_\mu}$ depend on β, the tracking properties of the RPE algorithm (26) are *independent* of the modeling error.

VII. Comparison with the FXLMS approach

The original FXLMS algorithm was designed for systems operating under stationary conditions, where identification of the controlled plant (e.g., the so-called secondary path, in acoustic applications) can be performed off-line, before the control loop is closed. When the plant characteristics change over time, an adaptive version of the FXLMS algorithm should be used, allowing for on-line plant identification/tracking. Two well-known solutions to this problem are based on the auxiliary noise approach, and the overall system modeling approach, respectively [3]. Both approaches were compared in [21] and, based on evaluation of their steady-state and transient performance, the approach utilizing auxiliary signals was found superior to the approach in which no auxiliary signals are required – for this reason we will focus our attention on the first of them only.

In the auxiliary noise approach, proposed originally by Eriksson and Allie [22], to enable reliable identification/tracking of the plant, a low-variance random perturbation (an artificially generated white noise sequence) is added to the input signal. At each time-step, the updated plant model is used as the reference prefilter for the classical FXLMS algorithm. The resulting scheme, further referred to as adaptive FXLMS, is depicted in Fig. 4. It incorporates two adaptation loops:

1) The system identification loop, which tracks coefficients of a finite impulse response (FIR) model of the plant. It

is served by the first LMS algorithm

$$y'(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\widehat{\mathbf{k}}(t - 1) - y(t)$$
$$\widehat{\mathbf{k}}(t) = \widehat{\mathbf{k}}(t - 1) + \mu_1 \mathrm{Re}\{\boldsymbol{\varphi}(t)y'(t)\} \qquad (53)$$

where $y'(t)$ is a hypothetical (reconstructed) response of the identified system to the auxiliary noise signal $u'(t)$, $\widehat{\mathbf{k}}(t) = [\widehat{k}_1(t), \ldots, \widehat{k}_M(t)]^{\mathrm{T}}$ is a vector of estimated impulse response coefficients, $\boldsymbol{\varphi}(t) = [u'(t - 1), \ldots, u'(t - M)]^{\mathrm{T}}$ is a vector made up of past values of the auxiliary noise signal, and μ_1 is a (real-valued) step-size parameter.

2) Direct control loop, which forms the compensating signal by means of adjusting the magnitude and phase of the measured, or artificially generated, reference signal $r(t) = e^{j\omega_0 t}$. It is served by the second LMS algorithm

$$r'(t) = \boldsymbol{\psi}^{\mathrm{T}}(t)\widehat{\mathbf{k}}(t)$$
$$\widehat{\delta}(t) = \widehat{\delta}(t - 1) + \mu_2 r'(t)y(t) \qquad (54)$$

where $r'(t) = \widehat{K}_{\mathrm{p}}(q^{-1})r(t)$ is a prefiltered reference signal, $\boldsymbol{\psi}(t) = [r(t - 1), \ldots, r(t - M)]^{\mathrm{T}}$ is a vector made up of past values of the reference signal, $\widehat{\delta}(t)$ is a complex-valued weight, and μ_2 is a (real-valued) step-size parameter.

As far as narrowband noise suppression is concerned, the solution described above has some obvious drawbacks, compared to our solution (the comparison presented below does not extend to other variants of FXLMS):

1) Unnecessary Performance Degradation. The method proposed by Eriksson and Allie is an invasive approach – injection of an auxiliary noise disturbs operation of the adaptive noise canceling system, causing deterioration of its performance (especially in the absence of measurement noise). This drawback can be alleviated to some extent by using schemes with dynamic scheduling of noise variance, such as the scheme proposed in [23]. There the auxiliary noise power is large before the adaptive noise control system converges, and becomes small afterwards. Since convergence can only take place under stationary conditions (time-invariant plant), when the system is continuously operated in the tracking mode, only small improvements can be gained by taking this approach.

2) Lack of Parsimony. To maintain satisfactory performance of the closed-loop system, one may be forced to use FIR models of relatively high orders. For example, in our experiments with the Guo-Bodson plant, estimation of at least $M = 20$ impulse response coefficients was necessary to maintain stability of the closed-loop system. Since all that is needed for cancellation purposes, is an estimate of the plant's gain at the frequency ω_0 : $\widehat{k}_{\mathrm{p}} = \sum_{i=1}^{M} \widehat{k}_i e^{-j\omega_0 i}$, the identified model is grossly overparameterized. This lack of parsimony results in a slower response of the system to plant and/or disturbance changes, as well as in an extra degradation of its steady-state performance. Note that in our approach only one complex-valued coefficient is estimated.

3) Excessive Numerical Complexity. Computational burden associated with the adaptive FXLMS algorithm (53)–(54) is

equal to $6M+12$ real multiply/add operations per time update (this count does not include the cost of generating auxiliary noise and reference). When the normalized LMS (NLMS) recursions are used instead of the regular ones, the computational cost further increases to $8M + 13$ real multiply/add operations and 2 real division operations. The analogous count for the algorithm (26) gives 33 real multiply/add operations and 3 real division operations per time update. Note that for $M = 20$ the adaptive FXLMS algorithm is computationally 4 times more demanding than our algorithm. This computational advantage of SONIC diminishes with increasing number of eliminated sinusoidal components m, as it grows linearly with m (see Section VIII-A). For adaptive FXLMS the analogous cost increment is smaller, because the order M of the identified plant model does not need to be increased.

4) Necessity of Tuning. The adaptive FXLMS algorithm (53)–(54) requires selection of three real-valued design parameters: two step-size coefficients and the variance of an auxiliary noise. Since different settings may be required under different operating conditions, no set of fixed values of design parameters can guarantee satisfactory performance of the system in the presence of plant and/or disturbance changes. In contrast with this, the proposed algorithm *automatically* adjusts one complex-valued parameter [note that the constants c_μ, ρ and k_{n}, appearing in (26), have no influence on the steady-state behavior of the system] – and it does it in a way that guarantees minimization of the steady-state mean-squared cancellation error. Of course, one can design extended FXLMS schemes equipped with some additional mechanisms for on-line adjustment of LMS step-sizes μ_1 and μ_2 – see e.g. [24] and the references therein. Generally, we have bad experience with applying variable step-size algorithms to (53)–(54): they are difficult to tune (the method described in [24] requires fixing 3 design parameters for each of the updated step-sizes) and they often cause system instability – see Section X-B. The main problem with such *ad hoc* solutions is that they are designed and optimized for open-loop systems subject to wideband excitation, where they work satisfactorily. However, when incorporated in a frequency selective control system, such as adaptive narrowband noise canceller, they usually misbehave, most likely because of the closed-loop identifiability problems. Note that the adaptation gain update mechanism used in SONIC was derived for a closed-loop system. This explains its good properties.

Remark: To do justice to the adaptive FXLMS approach it should be stressed that while SONIC can be used to cancel narrowband disturbances only, FXLMS is capable of eliminating practically any kind of diswturbance, including wideband noise.

VIII. Extensions

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

A. Multiharmonic Disturbance

Sinusoidal disturbances that occur in vibrating systems often consist of the fundamental (with frequency ω_0) and several harmonics (with frequencies $2\omega_0$, $3\omega_0$, etc.). Suppose that m such components, with slowly-varying amplitudes, are present, i.e.,

$$d(t) = \sum_{i=1}^{m} d_i(t), \quad d_i(t) = a_i(t)e^{ji\omega_0 t}$$
$$a_i(t) = a_i(t-1) + e_i(t) \tag{55}$$

where $\{e_i(t)\}$, $i = 1, \ldots, m$, are mutually independent circular white noise sequences with variances $\sigma_{e_1}^2, \ldots, \sigma_{e_m}^2$, respectively.

Similarly as before, we will base the structure of our adaptive filter on the form of the optimal solution to the open-loop problem. Suppose that one would like to "remove" $d(t)$ from the signal $s(t) = d(t) + v(t)$, where $\{v(t)\}$ denotes circular white measurement noise. Let $\mathbf{d}(t) = [d_1(t), \ldots, d_m(t)]^{\mathrm{T}}$. Note that the signal $s(t)$ admits the following state-space representation

$$\mathbf{d}(t) = \mathbf{A}\mathbf{d}(t-1) + \widetilde{\mathbf{e}}(t)$$
$$s(t) = \mathbf{1}^{\mathrm{T}}\mathbf{d}(t) + v(t) \tag{56}$$

where $\mathbf{A} = \mathrm{diag}\{e^{j\omega_0}, \ldots, e^{jm\omega_0}\}$, $\mathbf{1}^{\mathrm{T}} = [1, \ldots, 1]$, $\widetilde{\mathbf{e}}(t) = [\widetilde{e}_1(t), \ldots, \widetilde{e}_m(t)]^{\mathrm{T}}$ and $\widetilde{e}_i(t) = e_i(t)e^{ji\omega_0 t}$.

In the Gaussian case, the optimal steady-state one-step-ahead predictor of $s(t)$ can be computed using the following Kalman filter recursions

$$\widehat{s}(t|\,t-1) = \mathbf{1}^{\mathrm{T}}\widehat{\mathbf{d}}(t|\,t-1)$$
$$\varepsilon(t) = s(t) - \widehat{s}(t|\,t-1)$$
$$\widehat{\mathbf{d}}(t+1|\,t) = \mathbf{A}\widehat{\mathbf{d}}(t|\,t-1) + \mathbf{A}\mathbf{g}_\infty\varepsilon(t) \tag{57}$$

where $\mathbf{g}_\infty = [g_1, \ldots, g_m]^{\mathrm{T}}$ is the steady-state Kalman gain (note that under the assumptions made, it holds that $\mathbf{\Sigma}_e = \mathrm{cov}[\widetilde{\mathbf{e}}(t)] = \mathrm{diag}\{\sigma_{e_1}^2, \ldots, \sigma_{e_m}^2\}$, i.e., the state-space model of $s(t)$ is time-invariant).

Observe that the last recursion of (57) can be decomposed as

$$\widehat{d}_i(t+1|\,t) = e^{ji\omega_0}[\,\widehat{d}_i(t|\,t-1) + g_i\varepsilon(t)\,], \quad i = 1, \ldots, m \tag{58}$$

i.e., the steady-state Kalman filter can be viewed as a parallel structure made up of m subfilters, designed to track different harmonics and driven by the same global prediction error

$$\varepsilon(t) = s(t) - \sum_{i=1}^{m} \widehat{d}_i(t|\,t-1).$$

Under the slow-variation hypothesis, one can show that $g_i \cong \sqrt{\zeta_i}$ where $\zeta_i = \sigma_{e_i}^2/\sigma_v^2$, $i = 1, \ldots, m$. Based on (58), we propose the following decentralized control rule

$$u(t) = -\sum_{i=1}^{m} \frac{\widehat{d}_i(t+1|\,t)}{k_{\mathrm{n}i}}$$
$$\widehat{d}_i(t+1|\,t) = e^{ji\omega_0}[\,\widehat{d}_i(t|\,t-1) + \widehat{\mu}_i(t)y(t)\,] \tag{59}$$
$$i = 1, \ldots, m$$

where $k_{\mathrm{n}i} = K_{\mathrm{n}}(e^{-ji\omega_0})$ are the nominal plant gains at different frequencies, and the adaptation gains $\widehat{\mu}_i(t)$ are computed (independently of one another) using the algorithm

designed for the single-frequency case. When necessary, a zero-frequency (dc) disturbance component d_0 can be included in an analogous way

$$\widehat{d}_0(t+1\,|\,t) = \widehat{d}_0(t\,|\,t-1) + \widehat{\mu}_0(t)y(t).$$

B. Systems with Delay

Suppose that the plant is governed by

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

where τ_0 denotes transport delay (e.g., secondary path delay). So far, we have considered the unit-delay case. Some modifications are needed to cope with $\tau_0 > 1$.

First, one should replace (10) with

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

Second, since (11) should be replaced with

$$\widehat{d}(t+\tau_0\,|\,t) = e^{j\omega_0}\widehat{d}(t+\tau_0-1\,|\,t-1) + \mu e^{j\omega_0\tau_0}y(t)$$

one arrives at

$$\frac{\partial y(t)}{\partial \mu} = e^{j\omega_0}\frac{\partial y(t-1)}{\partial \mu} - \mu\beta\,e^{j\omega_0\tau_0}\frac{\partial y(t-\tau_0)}{\partial \mu}$$
$$- \beta\,e^{j\omega_0\tau_0}y(t-\tau_0).$$

Finally, setting $\beta = c_\mu/\mu$, one obtains

$$\frac{\partial y(t)}{\partial \mu} = e^{j\omega_0}\frac{\partial y(t-1)}{\partial \mu} - c_\mu\,e^{j\omega_0\tau_0}\frac{\partial y(t-\tau_0)}{\partial \mu}$$
$$- \frac{c_\mu}{\mu}\,e^{j\omega_0\tau_0}y(t-\tau_0) \qquad (60)$$

which can be considered a generalized version of (24).

The corresponding disturbance rejection algorithm can be summarized as follows

$$z(t) = e^{j\omega_0}z(t-1) - c_\mu\,e^{j\omega_0\tau_0}z(t-\tau_0)$$
$$- \frac{c_\mu}{\widehat{\mu}(t-1)}\,e^{j\omega_0\tau_0}\,y(t-\tau_0)$$
$$r(t) = \rho r(t-1) + |\,z(t)\,|^2$$
$$\widehat{\mu}(t) = \widehat{\mu}(t-1) - \frac{z^*(t)y(t)}{r(t)}$$
$$\widehat{d}(t+\tau_0\,|\,t) = e^{j\omega_0}\widehat{d}(t+\tau_0-1\,|\,t-1) + \widehat{\mu}(t)e^{j\omega_0\tau_0}y(t)$$
$$u(t) = -\frac{\widehat{d}(t+\tau_0\,|\,t)}{k_{\mathrm{n}}}. \qquad (61)$$

Remark: The algorithm (26) continues to work satisfactorily for systems with delay – the only price paid for delay underestimation is in longer transient responses and less accurate tuning, compared to (61). Delay overestimation leads to similar effects with one noticeable exeption. Recall that for the purpose of derivation and steady-state analysis of (26) we have adopted a static model of the plant, i.e., we have assumed that $K_{\mathrm{p}}(q^{-1})u(t) \cong k_{\mathrm{p}}u(t)$. Since in all transient phases a more adequate approximation has the form $K_{\mathrm{p}}(q^{-1})u(t) \cong k_{\mathrm{p}}u(t-\tau_\phi)$, where $\tau_\phi = -\arg[K_{\mathrm{p}}(e^{-j\omega_0})]/\omega_0$ denotes the so-called phase delay of the plant, measured at the frequency ω_0, incorporation of an extra delay not exceeding (or only slightly exceeding) τ_ϕ usually *improves* control results! – see Section X-D.

C. Real-Valued Signals

All presented results apply to systems with inputs and outputs described by complex numbers. A "quick and dirty" way of extending our adaptive filter to real-valued signals, which works pretty well in practice, 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 plant.

Remark: A more sophisticated solution to this problem, yielding better results, was described in [25]. One can prove that the control part of the complex-valued algorithm works *identically* to the control part of the real-valued algorithm, derived in [25]. Differences occur in the self-optimizing parts of both algorithms. Note that the complex-valued algorithm tries to minimize the following measure of fit $\mathrm{E}[|y(t)|^2] = \mathrm{E}[y_{\mathrm{R}}^2(t)] + \mathrm{E}[y_{\mathrm{I}}^2(t)]$ while, when applied to real-valued signals, it should minimize $\mathrm{E}[y_{\mathrm{R}}^2(t)]$. For this reason, unlike the algorithm proposed in [25], the modified complex-valued regulator does *not* converge to the optimal regulator designed for a real-valued case – but it continues to work satisfactorily.

IX. Safety jacketing

So far, we have been assuming that the plant is time-invariant. Since the proposed procedure for automatic tuning of the adaptation gain μ is based on minimization of the local measure of fit (21), the algorithm (26) should also cope favorably with slow changes in plant dynamics. To make it robust to abrupt plant changes (note that the initial convergence phase also falls into this category), some further modifications are needed.

First, to avoid erratic behavior of the algorithm during startup/transient periods, it is advisable to set the maximum allowable values for $|\,\widehat{\mu}(t)\,|$, $|\,\widehat{\mu}(t)-\widehat{\mu}(t-1)\,|$, and $r(t)$, further denoted by $\mu_{\max}$, $\Delta\mu_{\max}$, and $r_{\max}$, respectively. These are typical "safety valves" used in adaptive filtering.

Second, instead of a constant forgetting factor ρ, one can use in (26) a time-varying factor dependent on the current value of μ

$$\rho(t) = 1 - c_\rho|\,\widehat{\mu}(t-1)\,|$$

where $0 < c_\rho \ll 1$. This ensures that $1 - \rho$ will be at all times much smaller than $\widehat{\mu}$, which is consistent with the rule saying that the adaptation time constants of a hierarchical multi-layer adaptive system should gradually increase from the shortest (fastest adaptation) to the longest (slowest adaptation). The recommended values of c_ρ are those from the interval [0.01, 0.1].

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

$$\mathrm{sat}(x, a) = \begin{cases} x & ,\quad |\,x\,| \le a \\ a\,\dfrac{x}{|\,x\,|} & ,\quad |\,x\,| > a \end{cases}.$$

Then, the modified disturbance rejection algorithm that combines all "fixes" described above can be summarized as follows

$$z(t) = e^{j\omega_0}\left[(1-c_\mu)z(t-1) - \frac{c_\mu}{\widehat{\mu}(t-1)}\,y(t-1)\right]$$
$$\rho(t) = 1 - c_\rho|\widehat{\mu}(t-1)|$$
$$\widetilde{r}(t) = \rho(t)r(t-1) + |z(t)|^2$$
$$r(t) = \min[\widetilde{r}(t), r_{\max}]$$
$$\Delta\mu(t) = \text{sat}\left[z^*(t)y(t)/r(t)\,,\,\Delta\mu_{\max}\right]$$
$$\widetilde{\mu}(t) = \widehat{\mu}(t-1) - \Delta\mu(t)$$
$$\widehat{\mu}(t) = \text{sat}[\widetilde{\mu}(t),\mu_{\max}]$$
$$\widehat{d}(t+1|t) = e^{j\omega_0}[\,\widehat{d}(t|t-1) + \widehat{\mu}(t)y(t)\,]$$
$$u(t) = -\frac{\widehat{d}(t+1|t)}{k_{\mathrm{n}}}. \tag{62}$$

X. SIMULATION AND EXPERIMENTAL RESULTS

Four computer simulations and one real-world experiment were performed to check both the steady-state and transient performance of the proposed disturbance rejection schemes.

A. Tracking Behavior

The objective of this simulation experiment was to demonstrate the ability of the proposed algorithm to cope with modeling errors, including sudden plant changes. The transfer function of the plant was altered three times during each simulation run – see Table I. The adopted measurement noise and sinusoidal disturbance settings were: $\sigma_v = 0.1$, $\sigma_e = 0.001$, $\omega_0 = 0.1$ rad, $d(0) = 1$.

While the first two changes [from $K_1(q^{-1})$ to $K_2(q^{-1})$ at instant $t = 15000$ and from $K_2(q^{-1})$ to $K_3(q^{-1})$ at instant $t = 30000$] were confined to plant parameters, the last change was more substantial: at instant $t = 45000$, the first-order inertial system $K_3(q^{-1})$ with a single real pole was switched to the second-order nonminimum phase system $K_4(q^{-1})$ with a pair of complex poles. Since the phase shift introduced by $K_4(q^{-1})$ at the frequency ω_0 differs from the analogous shift of $K_3(q^{-1})$ by more than $\pi/2$, the last change causes temporal instability of the closed-loop system, making the task of disturbance rejection even harder.

Fig. 5 (illustrating typical behavior) and Fig. 6 (illustrating mean behavior) show results obtained for the algorihm (62) with the following settings: $c_\rho = 0.05$, $c_\mu = 0.005$, $\mu_{\max} = 0.05$, $\Delta\mu_{\max}(t) = \widehat{\mu}(t-1)/50$, $r_{\max} = 1600$. The nominal plant gain was fixed at the value $k_{\mathrm{n}} = e^{j\omega_0}$ — the corresponding magnitude and phase errors are listed in Table I.

The adaptation process was started from scratch at instant $t = 1$ using the following initial conditions: $\widehat{d}(0) = e^{j\omega_0}$, $r(0) = 100$, $z(0) = 0$, $\widehat{\mu}(0) = 0.02$. The algorithm copes favorably with both the initial convergence problem and with abrupt plant changes. When the experiment is started or when a change to the plant dynamics occurs, the magnitude of the adaptation gain $\widehat{\mu}(t)$ temporarily increases to quickly compensate large initial modeling errors; later on, it gradually decays to settle down around its optimal steady-state value. Note the very quick response to phase errors and usually much

slower response to magnitude errors — the effect caused by diverse sensitivity of system output to two types of modeling errors.

The simulation experiments show that the proposed control scheme has the self-stabilization property (not covered by the SDE-based analysis). When instability occurs at the instant $t = 45000$ [which is unavoidable since, due to the sign mismatch, the stabilizing gain $\widehat{\mu}$ for $K_3(q^{-1})$ does not stabilize $K_4(q^{-1})$], it causes rapid growth of the output signal $y(t)$, which in turn speeds up convergence of $\widehat{\mu}$ to a new stabilizing value. In this way, after a burst observed at the system output, the closed-loop stability is regained.
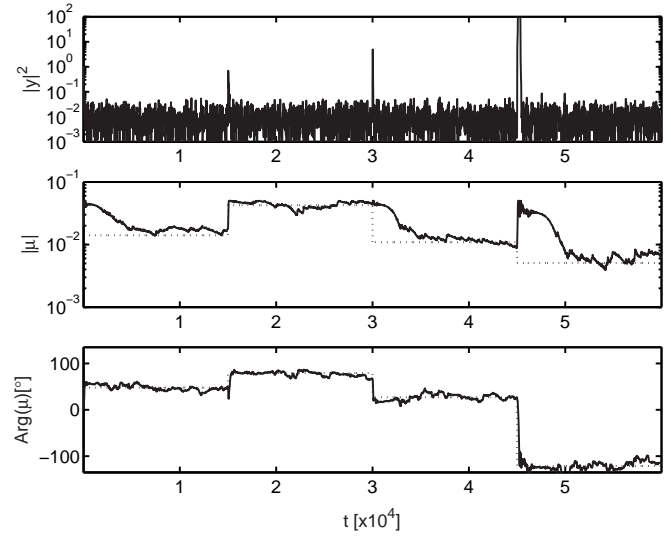

Fig. 5. Transient behavior of the disturbance rejection algorithm (results of a typical simulation experiment). Solid lines – estimated values, dotted lines – optimal steady-state values.

The mean transient responses observed at instants $t = 15000$ and $t = 30000$ are shorter than 500 samples, i.e., they last for less than 8 periods of the disturbance ($T_0 = 2\pi/\omega_0 \cong 63$). For less significant plant changes the length of the transient period is usually much shorter, often taking the values smaller than T_0.

B. Comparison with the Adaptive FXLMS Algorithm

The simulated first-order plant was governed by

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

where the coefficient $\varrho(t) \in [0.45, 0.95]$ determines location of a "time-varying system pole" [the Guo-Bodson plant corresponds to the choice $\varrho(t) = 0.9048, \forall t$]. The frequency of the sinusoidal disturbance was set equal to $\omega_0 = 0.3$, and its amplitude varies according to

$$a(t) = 1 + 0.2\sin(0.002t)$$

Following [21], the order of the FIR filter was set to $M = 32$, and the variance of auxiliary noise – to $\sigma_{u'}^2 = 0.001$.

Table II compares the results yielded by the optimally tuned adaptive FXLMS algorithm with those given by the

TABLE I
PLANT SWITCHING SCHEDULE IN THE TRANSIENT BEHAVIOR EXPERIMENT AND THE CORRESPONDING MODELING ERRORS.

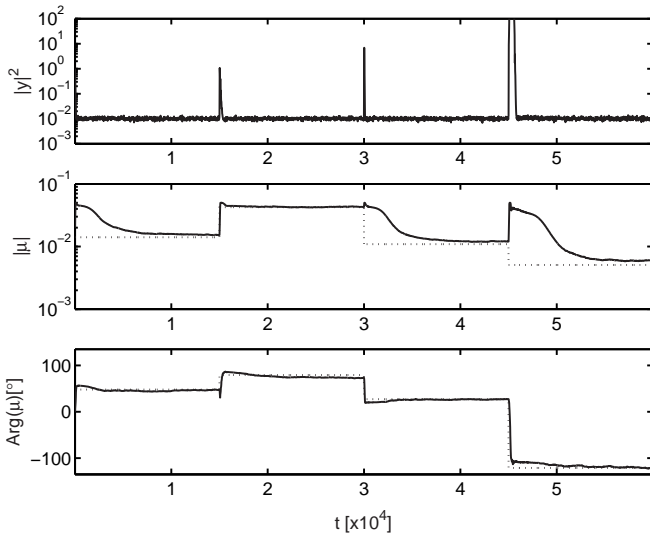| Time interval	Plant	$\lvert\beta\rvert$	$\mathrm{Arg}[\beta]$
$0 < t < 15000$	$K_1(q^{-1}) = \dfrac{0.0952}{1 - 0.9048q^{-1}}$	0.708	-47.9
$15000 \le t < 30000$	$K_2(q^{-1}) = \dfrac{0.0238}{1 - 0.9762q^{-1}}$	0.234	-79.3
$30000 \le t < 45000$	$K_3(q^{-1}) = \dfrac{0.2}{1 - 0.8q^{-1}}$	0.913	-27.1
$45000 \le t \le 60000$	$K_4(q^{-1}) = \dfrac{-0.1 + 0.14q^{-1}}{1 - 1.8391q^{-1} + 0.8649q^{-2}}$	1.960	121.1

Fig. 6. Mean transient behavior of the disturbance rejection algorithm (average of 100 simulation runs). Solid lines – ensemble averages of the estimated values, dotted lines – optimal steady-state values.

TABLE II
COMPARISON OF THE MEAN-SQUARED ESTIMATION ERRORS YIELDED BY THE OPTIMIZED ADAPTIVE FXLMS ALGORITHM AND BY THE SONIC ALGORITHM.

σ_v	MSE	FXLMS	SONIC
0.1	$\overline{\lvert c(t)\rvert^2}$	$3.80 \cdot 10^{-4}$	$3.65 \cdot 10^{-4}$
	$\overline{\lvert y(t)\rvert^2}$	$1.038 \cdot 10^{-2}$	$1.037 \cdot 10^{-2}$
0	$\overline{\lvert y(t)\rvert^2}$	$9.03 \cdot 10^{-5}$	$3.67 \cdot 10^{-7}$

comparison presented above is rather unfair for SONIC.

Our attempts to combine the FXLMS algorithm (53)–(54) with the robust step-size adjustment technique proposed in [24] were a complete failure. Even though we followed the guidelines of [24] for choosing the method's design parameters, the closed-loop system turned out to be unstable in almost every simulation run.

C. Elimination of a Multiharmonic Disturbance

In our third simulation experiment, the output of the Guo–Bodson plant was corrupted with a sinusoidal signal consisting of the first and third harmonics: $d(t) = d_1(t) + d_3(t)$, $d_1(t) = a_1(t)e^{j\omega_0 t}$, $d_3(t) = a_3(t)e^{3j\omega_0 t}$, where $a_1(t)$ and $a_3(t)$ denote complex amplitudes evolving according to the random-walk model: $\sigma_{e_1} = \sigma_{e_3} = 0.001$, $a_1(0) = d_1(0) = 1$, $a_3(0) = d_3(0) = 0.5$. The measurement noise variance $\sigma_v^2 = 0.01$. The decentralized adaptive controller (59) was applied, combining two algorithms of the form (62) with identical settings: $c_{\rho,1} = c_{\rho,3} = 0.05$, $c_{\mu,1} = c_{\mu,3} = 0.005$, $r_{\max,1} = r_{\max,3} = 1600$, $\Delta\mu_{\max,1}(t) = \widehat{\mu}_1(t-1)/50$, $\Delta\mu_{\max,3}(t) = \widehat{\mu}_3(t-1)/50$, $\mu_{\max,1} = \mu_{\max,3} = 0.05$. The initial conditions were set to $\widehat{d}_1(0) = \widehat{d}_3(0) = 0$ and $\widehat{\mu}_1(0) = \widehat{\mu}_3(0) = 0.02$. Finally, the nominal plant gains k_1 and k_3 were set to $e^{j\omega_0}$ and $e^{j3\omega_0}$, respectively, which results in the following modeling errors: $\lvert\beta_1\rvert = 0.708$, $\mathrm{Arg}[\beta_1] = -47.9\,^\circ$, $\lvert\beta_3\rvert = 0.318$, $\mathrm{Arg}[\beta_3] = 80.3\,^\circ$.

Instead of a "cold start", used in the preceding example (which is not recommended in practice, because of the risk of large initialization transients) a two-stage initialization procedure was applied. During the first 300 time-steps, the quantities $[z_1(t),\ r_1(t)]$ and $[z_3(t),\ r_3(t)]$ were evaluated,

SONIC algorithm (26) $[c_\mu = 0.01, \rho = 0.995, k_n = 1]$. The best settings for FXLMS ($\mu_1 = 0.025$, $\mu_2 = 0.7$ for $\sigma_v = 0.1$, and $\mu_1 = 0.025$, $\mu_2 = 1.3$ for $\sigma_v = 0$) were found using a computationally exhaustive trial and error search procedure. All numbers were obtained by means of combined ensemble averaging (10 realizations of $\{v(t)\}$ and $\{u'(t)\}$) and time averaging ($t \in [20001, 70000]$), after the compared algorithms have reached their steady-state behavior. Results for the FXLMS controller *cannot* be improved by decreasing the variance of the auxiliary noise. When the variance $\sigma_{u'}^2$ is further reduced, the mean-squared cancellation/output errors start to increase – this effect is caused by poor identification results due to insufficient plant excitation.

While in the presence of measurement noise ($\sigma_v = 0.1$) SONIC is only marginally better than the optimally tuned FXLMS, in the absence of noise ($\sigma_v = 0$), which is the situation considered in [21], the improvement is quite significant, reaching 24 dB. In both cases, SONIC is computationally almost 6 times less demanding than FXLMS. Note also that, since in practice FXLMS seldom is optimally tuned, the
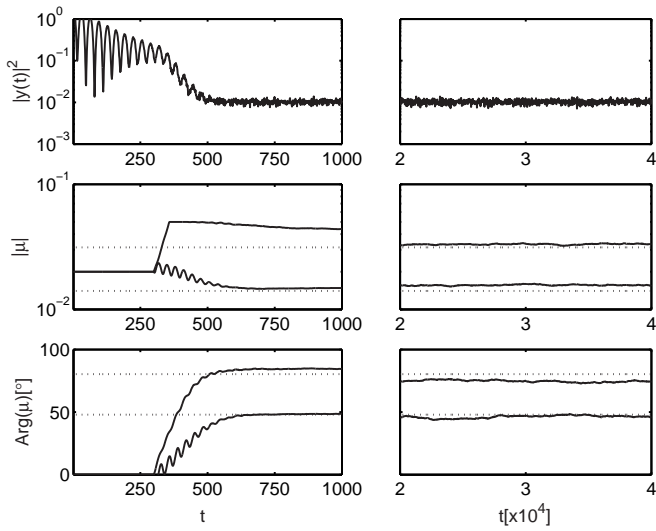
Fig. 7. Rejection of a multiharmonic disturbance (average of 100 simulation runs). Solid lines – ensemble averages of estimated values, dotted lines – optimal steady-state values.
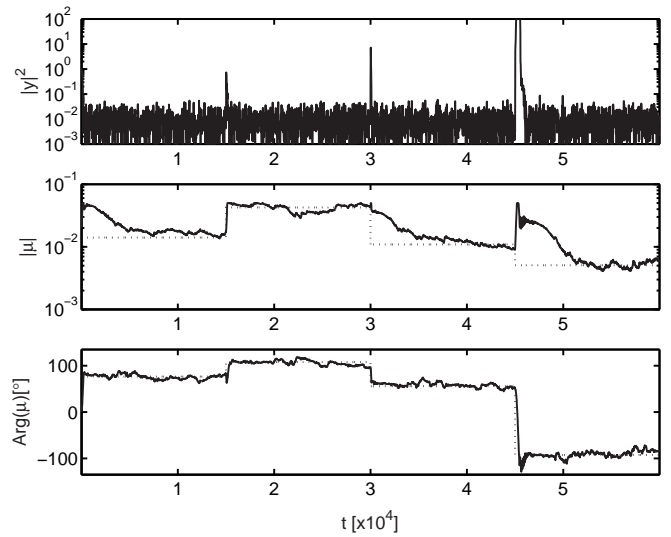

Fig. 8. Transient behavior of the disturbance rejection algorithm in the presence of an extra 5-sample transport delay (results of a typical simulation run). Solid lines – estimated values, dotted lines – optimal steady-state values.

but the adaptation gains $[\widehat{\mu}_1(t), \widehat{\mu}_3(t)]$ were kept at their starting values $[\widehat{\mu}_1(0), \widehat{\mu}_3(0)]$ and not updated. Then, at instant $t = 301$, the adaptation lock was released.

Fig. 7 shows results of a typical simulation run. Similarly, as in the single-frequency case, the phase angles of $\widehat{\mu}_1(t)$ and $\widehat{\mu}_3(t)$ quickly respond to initial phase errors. The reaction of $|\widehat{\mu}_1(t)|$ and $|\widehat{\mu}_3(t)|$ is slower, but eventually both quantities also converge in mean to values that are close to the optimal steady-state values. Small estimation biases can be explained by violation of our small adaptation gain assumption, under which all theoretical results were derived.

D. Systems with Delay

The transient behavior experiment, reported in Section X-A, was repeated with an extra transport delay, equal to 5 sampling intervals ($\tau_0 = 6$), added to all switched plant models. The algorithm (61) was used, equipped with the safety enforcement mechanisms described earlier. The adaptation was started at the instant $t = 1$ in an analogous way as described in Section X-A. The assumed nominal gain k_n was equal to $e^{j\omega_0}$.

Fig. 8 shows the results of a typical simulation run. Note that these results are comparable with those presented in Fig. 5 for a plant with unit delay (in both cases the same realizations of $\{e(t)\}$ and $\{v(t)\}$ were used). The robustness issues are illustrated in Fig. 9.

E. Acoustic Experiment

One real-world active noise control experiment was performed using a standard PC, equipped with a poor-quality sound card. The left loudspeaker served as the noise source, while the right one was used for cancellation purposes. The error microphone was located approximately 1 m away from the left loudspeaker and 30 cm from the right loudspeaker. The system was operated at a sampling rate of 1 kHz. The transport delay was equal to 60 ms, i.e., 60 sampling intervals.
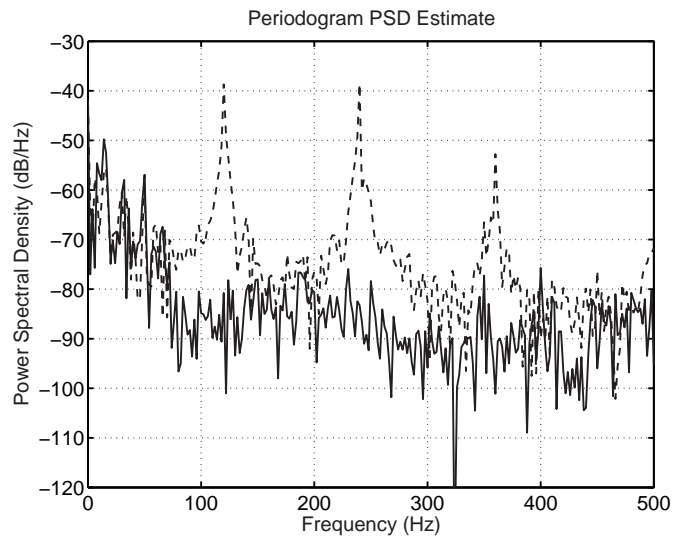

Fig. 10. Power spectral density of the signal before (dashed line) and after (solid line) disturbance cancellation.

The major part of this delay was due to hardware limitations (low processing speed + buffering) and not due to propagation effects.

The generated disturbance consisted of three sinusoidal tones with identical amplitudes and with frequencies 120 Hz, 240 Hz, and 360 Hz, respectively. The nominal filter gains were set to 1 for all frequencies. The remaining parameters were chosen as follows: $\mu_{max} = 0.05$, $c_\mu = 0.005$, $r_{max} = 5$, $\rho = 0.999$ (the same for all three subfilters). After an initial convergence phase, which lasted for a few seconds, the closed-loop system reached its steady-state behavior. Fig. 10 depicts periodogram-based estimates of the power spectral density of the original signal and its "silenced" version, picked up by the error microphone. The peaks at 120 Hz, 240 Hz, and 360 Hz were practically eliminated. Listening tests confirmed that the achieved disturbance reduction was significant.
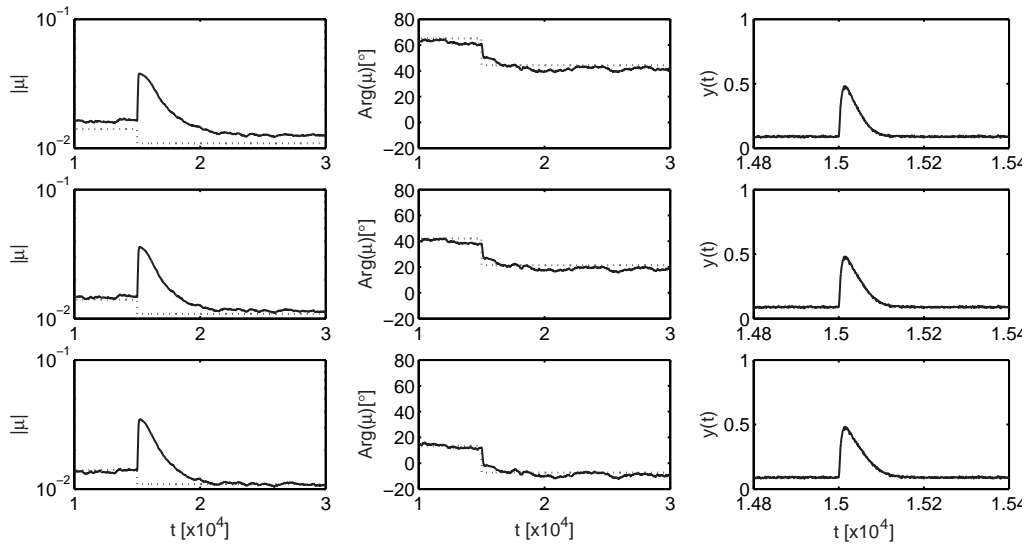
Fig. 9.   Results for the switched plant $[K_1(q^{-1}) \rightarrow K_2(q^{-1}), \ \omega_0 = 0.1, \ \tau_0 = 10]$. Mean transient responses were obtained for regulators designed assuming that the delay τ_0 is equal to 5 (three top figures), 10 (three middle figures), and 15 (three bottom figures), respectively. Phase delay of the plant τ_ϕ evaluated at the frequency ω_0 is equal to 8 before the change, and 15 after the change.

XI. CONCLUSION

The problem of eliminating a complex-valued sinusoidal disturbance of known frequency, acting at the output of an unknown linear stable plant, was considered. The proposed solution is based on the coefficient fixing — a technique originally developed for the purpose of adaptive minimum-variance control — combined with automatic adaptation gain adjustment. It was shown that when the complex-valued amplitude of the disturbance evolves according to the random-walk model, the resulting disturbance rejection scheme converges in mean to the optimal solution. It is also robust to abrupt plant changes. The obtained results were further extended to multiharmonic disturbances, plants with delay, and real-valued signals.

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] D. R. Morgan and C. Sanford, "A control theory approach to the stability and transient analysis of the filtered-X LMS adaptive notch filter," *IEEE Trans. Signal Processing*, vol. 40, pp. 2341–2346, Sep. 1992.

[5] G. Feng and M. Palaniswami, "A stable adaptive implementation of the internal model principle," *IEEE Trans. Automat. Control*, vol. 37, pp. 1220–1225, Aug. 1992.

[6] I. Landau, A. Constantinescu, and D. Rey, "Adaptive narrow band disturbance rejection applied to an active suspension - an internal model principle approach," *Automatica*, vol. 41, pp. 563–574, Apr. 2005.

[7] M. Bodson and S. Douglas, "Adaptive algorithms for the rejection of sinusoidal disturbances with unknown frequency," *Automatica*, vol. 33, pp. 2213–2221, Dec. 1997.

[8] X. Guo and M. Bodson, "Adaptive rejection of multiple sinusoids of unknown frequency," in *Proc. European Control Conference*, Kos, Greece, 2007, pp. 121–128.

[9] F. Lewis, *Optimal Estimation.*   New York: Wiley, 1986.

[10] H. van Trees and E. K. L. Bell, Eds., *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking.*   New York: Wiley, 2007.

[11] K. Åström, U. Borisson, L. Ljung, and B. Wittenmark, "Theory and applications of self-tuning regulators," *Automatica*, vol. 13, pp. 457–476, Sept. 1977.

[12] M. Niedźwiecki, "Steady-state and parameter tracking properties of self-tuning minimum variance regulators," *Automatica*, vol. 25, pp. 597–602, July 1989.

[13] ——, *Identification of Time-varying Processes.*   New York: Wiley, 2000.

[14] T. Söderström and P. Stoica, *System Identification.*   Englewood Cliffs NJ: Prentice Hall, 1988.

[15] W. Wirtinger, "Zur formalen theorie der funktionen von mehr komplexen veränderlichen," *Math. Ann.*, vol. 97, no. 1, pp. 357–375, 1927.

[16] D. H. Brandwood, "A complex gradient operator and its application in array theory," *Proc. Inst. Elect. Eng.*, vol. 130, no. 1, pp. 11–16, 1983.

[17] A. van den Bos, "Complex gradient and Hessian," *IEE Proc. Image Signal Processing*, vol. 141, pp. 380–382, Dec. 1994.

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

[19] A. Benveniste and G. Ruget, "A measure of the tracking capability of recursive stochastic algorithms with constant gains," *IEEE Trans. Autom. Control*, vol. 25, pp. 788–794, June 1982.

[20] A. Benveniste, M. Métivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations.*   Springer-Verlag, 1990.

[21] C. Bao, P. Sas, and H. Van Brussel, "Comparison of two on-line identification algorithms for active noise control," in *Proc. Recent Advances in Active Control of Sound and Vib.*, Blacksburg, VA, USA, 1993, pp. 38–51.

[22] L. J. Eriksson and M. C. Allie, "Use of random noise for on-line transducer modeling in an adaptive active attenuation system," *J. Acoust. Soc. Am.*, vol. 85, pp. 797–802, Feb. 1989.

[23] H. Lan, M. Zhang and W. Ser, "An active noise control system using online secondary path modeling with reduced auxiliary noise," *IEEE Signal Process. Lett.*, vol. 9, pp. 16–18, Jan. 2002.

[24] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: analysis and simulations," *IEEE Trans. Signal Processing*, vol. 45, pp. 631–639, March 1997.

[25] M. Niedźwiecki and M. Meller, "New approach to adaptive vibration control," in *Proc. 47th CDC*, Cancun, Mexico, 2008, pp. 2581–2587.