Postprint of: Meller M., Niedźwiecki M., Multichannel self-optimizing narrowband interference canceller, Signal Processing, Vol. 98 (2014), pp. 396-409, DOI: 10.1016/j.sigpro.2013.12.001

Multichannel Self-optimizing Narrowband Interference Canceller

Michał Meller and Maciej Niedźwiecki, *Member, IEEE*

***Abstract*—The problem of cancellation of a nonstationary sinusoidal interference, acting at the output of an unknown multivariable linear stable plant, is considered. No reference signal is assumed to be available. The proposed feedback controller is a nontrivial extension of the SONIC (self-optimizing narrowband interference canceller) algorithm, developed earlier for single-input, single-output plants. The algorithm consists of two loops: the inner, control loop, which predicts and cancels disturbance, and the outer, self-optimization loop, which automatically adjusts the gain matrix so as to optimize the overall system performance. The proposed scheme is capable of adapting to slow changes in disturbance characteristics, measurement noise characteristics, and plant characteristics. It is shown that in the important benchmark case – for disturbances with random-walk-type amplitude changes – the designed closed-loop control system converges locally in mean to the optimal one.**

***Index Terms*—active noise and vibration control, time-varying processes, adaptive signal processing**

I. Introduction

NARROWBAND interferences (acoustic noise and/or vibration) usually originate from rotation of an engine, compressor, fan, or propeller. In the range of small frequencies (below 1 kHz) such interferences are difficult to eliminate using passive methods, but can be efficiently removed using active noise control (ANC) techniques, i.e., by means of destructive interference. Multichannel ANC systems are becoming increasingly popular as they allow one to create larger (and spatially diversified) quiet zones compared to single-channel systems, albeit at the expense of higher equipment cost and increased computational requirements. Typical commercial applications of such systems include reduction of propeller-induced interior noise in aircrafts [1], [2], [3], or suppression of engine-induced vibrations in automotive vehicles [4], [5], [6], among many others.

Most of the existing multichannel solutions are based on the classical filtered-X least mean squares (FXLMS) approach [7], [8] or its modifications obtained by replacing the LMS adaptive filters with faster converging ones, such as recursive least squares (RLS) [9] or affine projection (AP) [10] – for comparison of different variants see e.g. [11]. In all cases mentioned above impulse response coefficients of all secondary paths, linking actuators with sensors, are supposed to be constant and known. In practice this means that the controlled acoustic/vibration field should be identified prior to starting the ANC algorithm, and that it should be re-estimated each time the spatial configuration of the system (positions of actuators and sensors) changes. To exert full control over the system in the presence of nonstationarities, on-line plant identification is required. To avoid closed-loop identifiability problems, a special random perturbation technique is often used [12]. Unfortunately, since auxiliary noise disturbs operation of the ANC system, identifiability is restored at the expense of unnecessary performance degradation.

Another, conceptually different, solution to the multivariable narrowband disturbance rejection problem, based on the "phase-locked" loop structure, was presented in [13]. However, in order to use this technique characteristics of the controlled plant (complex gains at given frequencies) need to be known *a priori*.

An entirely new approach to narrowband disturbance canceling was recently proposed in [14], [15] (for complex-valued disturbances) and in [16] (for real-valued disturbances). The developed scheme, called SONIC (self-optimizing narrowband interference canceller), combines the coefficient fixing technique, used to "robustify" self-tuning minimum-variance regulators [17], [18], [19], with automatic gain tuning. It can be used to control nonstationary plants subject to nonstationary narrowband disturbances and compares favorably, both in terms of cancellation quality and computational complexity, with the FXLMS scheme. In this study a multivariate version of SONIC is proposed and analyzed. Additionally, an extended, frequency-adaptive version of the algorithm is briefly sketched, allowing one to cope with quasi-periodic disturbances.

II. Problem statement

Consider the problem of cancellation of an n-dimensional complex-valued narrowband disturbance

$$\mathbf{d}(t) = \boldsymbol{\alpha}(t) e^{j\omega_0 t} \tag{1}$$

where $t = \ldots, -1, 0, 1, \ldots$ is a discrete, normalized time, $\omega_0 \in [-\pi, \pi)$ is a known angular frequency, and $\boldsymbol{\alpha}(t) = [\alpha_1(t), \ldots, \alpha_n(t)]^{\mathrm{T}}$ denotes the unknown time-varying vector of complex-valued "amplitudes", acting at the output of a multidimensional stable plant governed by

$$\mathbf{y}(t) = \mathbf{L}_{\mathrm{p}}(q^{-1})\mathbf{u}(t-1) + \mathbf{d}(t) + \mathbf{v}(t) \tag{2}$$

where $\mathbf{y}(t)$ is the n-dimensional output signal, $\mathbf{u}(t)$ is the n-dimensional input (cancellation) signal, $\mathbf{v}(t)$ is an n-

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, Tel: +48 58 3472519; fax: +48 58 3415821 (e-mails: michal.meller@eti.pg.gda.pl; maciekn@eti.pg.gda.pl).

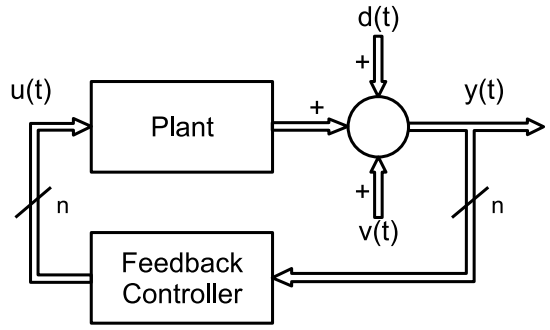

Fig. 1. Block diagram of the disturbance rejection system.

dimensional wideband noise, and

$$\mathbf{L}_{\mathrm{p}}(q^{-1}) = \begin{bmatrix} L_{11}(q^{-1}) & \dots & L_{1n}(q^{-1}) \\ \vdots & \ddots & \vdots \\ L_{n1}(q^{-1}) & \dots & L_{nn}(q^{-1}) \end{bmatrix}$$

denotes the $n \times n$-dimensional transfer function (q^{-1} denotes the backward shift operator) which will be further assumed unknown and possibly time-varying.

In acoustic active noise control systems, $L_{ij}(q^{-1}), 1 \leq i, j \leq n$, are transfer functions of the so-called secondary paths, linking each of the n actuators (canceling loudspeakers) with each of n sensors (error microphones). Note that the number of actuators is assumed to be equal to the number of sensors, which is usually referred to as the fully determined case.

We will look for a feedback controller allowing for cancellation, or near cancellation, of the sinusoidal disturbance, i.e., controller generating the signal $\mathbf{u}(t)$ that minimizes the system output in the mean-squared sense – in the control literature such devices are usually termed minimum-variance (MV) regulators. It will be not assumed that a reference signal, correlated with the disturbance, is available. For this reason the designed system, depicted in Fig. 1, will be a purely feedback canceller, not incorporating any feedforward compensation loop.

III. Control of a known plant

We will look for a steady-state MV regulator, i.e., for a control rule which guarantees that

$$\lim_{t\to\infty} \mathrm{E}[\mathbf{y}(t)\mathbf{y}^{\mathrm{H}}(t)] \longrightarrow \min .$$

We will start from a very simple controller that requires full prior knowledge of the plant and disturbance. Then we will gradually step back from restrictive assumptions to make our design work under more realistic conditions.

A. Stabilizing Controller

Suppose that the controlled plant is time-invariant, and that its transfer function $\mathbf{L}_{\mathrm{p}}(q^{-1})$ is known. Since the disturbance is a narrowband signal, the cancellation signal must also be narrowband, with angular frequency ω_0 and complex amplitude chosen so as to enable destructive interference at the plant's output. In a case like this equation (2) can be approximately written down in the form

$$\mathbf{y}(t) \cong \mathbf{K}_{\mathrm{p}}\mathbf{u}(t-1) + \mathbf{d}(t) + \mathbf{v}(t) \tag{3}$$

where

$$\mathbf{K}_{\mathrm{p}} = \mathbf{L}_{\mathrm{p}}(e^{-j\omega_0}), \quad \det(\mathbf{K}_{\mathrm{p}}) \neq 0$$

denotes the nonsingular matrix of plant gains at the frequency ω_0.

Should the disturbances be measurable and known ahead of time, the MV controller could be expressed in the form

$$\mathbf{u}(t) = -\mathbf{K}_{\mathrm{p}}^{-1}\mathbf{d}(t+1) \ . \tag{4}$$

When $\mathbf{d}(t+1)$ is unknown, it can be replaced in (4) with the one-step-ahead prediction, evaluated recursively by a simple gradient algorithm. This leads to the following control rule

$$\widehat{\mathbf{d}}(t+1|t) = e^{j\omega_0}[\widehat{\mathbf{d}}(t|t-1) + \mathbf{M}\mathbf{y}(t)] \tag{5}$$

$$\mathbf{u}(t) = -\mathbf{K}_{\mathrm{p}}^{-1}\widehat{\mathbf{d}}(t+1|t) \tag{6}$$

where

$$\mathbf{M} = \begin{bmatrix} \mu_{11} & \dots & \mu_{1n} \\ \vdots & \ddots & \vdots \\ \mu_{n1} & \dots & \mu_{nn} \end{bmatrix}$$

denotes a matrix of complex-valued adaptation gains μ_{ij}, chosen so as to guarantee stability of the closed loop system.

To arrive at stability conditions, the time-varying amplitude in (1) will be rewritten in the form

$$\boldsymbol{\alpha}(t) = \boldsymbol{\alpha}(t-1) + \mathbf{e}(t)$$

where $\mathbf{e}(t)$ denotes the one-step amplitude change. Using this notation, one can rewrite $\mathbf{d}(t)$ in the form

$$\mathbf{d}(t) = e^{j\omega_0}\mathbf{d}(t-1) + \tilde{\mathbf{e}}(t) \tag{7}$$

where $\tilde{\mathbf{e}}(t) = e^{j\omega_0 t}\mathbf{e}(t)$. Denote by $\mathbf{c}(t) = \mathbf{d}(t) - \widehat{\mathbf{d}}(t|t-1)$ the cancellation error. After combining (3) with (6) and (7), one arrives at

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{c}(t) + \mathbf{v}(t) \\ \mathbf{c}(t) &= e^{j\omega_0}(\mathbf{I} - \mathbf{M})\mathbf{c}(t-1) - \mathbf{M}\tilde{\mathbf{v}}(t-1) + \tilde{\mathbf{e}}(t) \end{aligned} \tag{8}$$

where $\tilde{\mathbf{v}}(t) = e^{j\omega_0}\mathbf{v}(t)$.

It is clear from (8) that when the processes $\{\mathbf{v}(t)\}$ and $\{\mathbf{e}(t)\}$ are bounded in the mean-squared sense, so is the output signal, provided that

$$\mathbf{M} \in \boldsymbol{\Omega}_{\mathrm{s}}: \quad |\lambda_i(\mathbf{I} - \mathbf{M})| < 1, \ i = 1, \dots, n$$

where $\lambda_i(\cdot)$ denotes the ith eigenvalue of the respective matrix.

B. Optimal Controller

In this section it will be examined how the canceller (6) performs in the presence of random-walk (RW) amplitude drift, arising when the sequence of amplitude changes $\{\mathbf{e}(t)\}$ forms white noise. This will serve two purposes. First, the tracking analysis under RW-type variations is an important estimation and control benchmark [20], [21]. Even though the RW model is not very realistic, it is quite demanding, as the resulting amplitude trajectories are not bounded. Second,

the RW case is usually analytically tractable, which means that one can derive closed-form expressions characterizing the system's performance, and compare them with the analogous expressions obtained for the optimal controller. This allows one to examine the statistical efficiency of the proposed solution, i.e., to evaluate it in absolute, rather than in relative, terms.

To arrive at analytical results, the following assumptions will be made:

(A1) $\{\mathbf{v}(t)\}$ is a zero-mean circular white sequence with covariance matrix $\mathbf{V}$.

(A2) $\{\mathbf{e}(t)\}$, independent of $\{\mathbf{v}(t)\}$, is a zero-mean circular white sequence with covariance matrix $\mathbf{E}$.

First, the Cramér-Rao-type lower tracking bound, which limits from below the cancellation efficiency, will be derived. In order to do this, consider the open-loop problem of finding the optimal, in the mean-squared sense, one-step-ahead predictor of $\mathbf{d}(t)$ obeying (7), based on noisy measurements

$$\mathbf{s}(t) = \mathbf{y}(t)|_{\mathbf{u}(t)\equiv\mathbf{0}} = \mathbf{d}(t) + \mathbf{v}(t). \tag{9}$$

Note that under (A2) the sequence $\{\tilde{\mathbf{e}}(t)\}$, appearing in (7), is circular white with covariance matrix $\mathbf{E}$.

Regarding (7) as a state equation of a dynamic system, and (9) as its output (measurement) equation, prediction of $\mathbf{d}(t)$ can be viewed as an estimation problem in the state space. The optimal, in the mean-squared sense, one-step-ahead predictor of $\mathbf{d}(t)$ has the well-known form

$$\widehat{\mathbf{d}}(t+1|t) = \mathrm{E}[\mathbf{d}(t+1)|\mathcal{S}(t)] \tag{10}$$

where $\mathcal{S}(t) = \{s(i), i \le t\}$ denotes the observation history available at instant t. Under Gaussian assumptions

(A3) The sequences $\{\mathbf{v}(t)\}$ and $\{\mathbf{e}(t)\}$ are normally distributed: $\mathbf{v}(t) \sim \mathcal{CN}(0, \mathbf{V})$, $\mathbf{e}(t) \sim \mathcal{CN}(0, \mathbf{E})$

the conditional mean estimate (10) can be computed recursively using the celebrated Kalman filter

$$\begin{aligned}
\boldsymbol{\varepsilon}(t) &= \mathbf{s}(t) - \widehat{\mathbf{d}}(t|t-1) \\
\mathbf{G}(t) &= \mathbf{P}(t|t-1)[\mathbf{P}(t|t-1) + \mathbf{V}]^{-1} \\
\widehat{\mathbf{d}}(t+1|t) &= e^{j\omega_0}[\widehat{\mathbf{d}}(t|t-1) + \mathbf{G}(t)\boldsymbol{\varepsilon}(t)] \\
\mathbf{P}(t+1|t) &= [\mathbf{I} - \mathbf{G}(t)]\mathbf{P}(t|t-1) + \mathbf{E}
\end{aligned} \tag{11}$$

where $\mathbf{G}(t)$ and $\mathbf{P}(t|t-1)$ denote the $n \times n$ Kalman gain and posterior covariance matrices, respectively.

The steady-state version of this algorithm can be written down in the form

$$\begin{aligned}
\boldsymbol{\varepsilon}(t) &= \mathbf{s}(t) - \widehat{\mathbf{d}}(t|t-1) \\
\widehat{\mathbf{d}}(t+1|t) &= e^{j\omega_0}[\widehat{\mathbf{d}}(t|t-1) + \mathbf{G}_\infty\boldsymbol{\varepsilon}(t)]
\end{aligned} \tag{12}$$

where $\mathbf{G}_\infty \in \boldsymbol{\Omega}_\mathrm{s}$ denotes the steady-state gain matrix

$$\mathbf{G}_\infty = \lim_{t\to\infty} \mathbf{G}(t) = \mathbf{P}_\infty[\mathbf{P}_\infty + \mathbf{V}]^{-1} \tag{13}$$

and $\mathbf{P}_\infty = \lim_{t\to\infty} \mathbf{P}(t)$ is the positive definite solution of the Riccati equation

$$\mathbf{P}_\infty[\mathbf{P}_\infty + \mathbf{V}]^{-1}\mathbf{P}_\infty - \mathbf{E} = \mathbf{O} \tag{14}$$

where $\mathbf{O}$ denotes the zero matrix (in the sequel $\mathbf{A} > \mathbf{O}$ will mean that the matrix $\mathbf{A}$ is positive definite).

Note that the disturbance estimation part of the control rule (6) resembles that of a steady-state Kalman filter (12). For any stabilizing gain $\mathbf{M} \in \boldsymbol{\Omega}_\mathrm{s}$, denote by $\mathbf{C}_\infty(\mathbf{M}) = \lim_{t\to\infty} \mathrm{E}[\mathbf{c}(t)\mathbf{c}^\mathrm{H}(t)]$ the covariance matrix of the corresponding steady-state cancellation error. It will be shown that the best performance of the closed-loop interference cancelling system can be obtained by setting $\mathbf{M} = \mathbf{G}_\infty$.

Corollary 1

Under assumptions (A1)–(A3) it holds that

$$\inf_{\mathbf{M}\in\boldsymbol{\Omega}_\mathrm{s}} \mathbf{C}_\infty(\mathbf{M}) = \mathbf{P}_\infty \ .$$

The minimum is obtained for $\mathbf{M} = \mathbf{G}_\infty$.

Proof - see Appendix I.

IV. Control of an unknown plant

So far it has been assumed that the "true" gain $\mathbf{K}_\mathrm{p}$ of the plant is known. Suppose now that the "idealized" control rule (4) is replaced with

$$\mathbf{u}(t) = -\mathbf{K}_\mathrm{n}^{-1}\widehat{\mathbf{d}}(t+1|t) \tag{15}$$

where $\mathbf{K}_\mathrm{n}$, $\det(\mathbf{K}_\mathrm{n}) \neq 0$, is the nominal (assumed) plant gain at the frequency ω_0, generally different from the true gain $\mathbf{K}_\mathrm{p}$. Denote by

$$\mathbf{B} = \mathbf{K}_\mathrm{p}\mathbf{K}_\mathrm{n}^{-1} = \begin{bmatrix} \beta_{11} & \dots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \dots & \beta_{nn} \end{bmatrix}$$

the matrix of complex-valued modeling errors. Combining (3) with (5) and (15), one arrives at the following generalized version of (8)

$$\mathbf{y}(t) = \mathbf{c}(t) + \mathbf{v}(t) \tag{16}$$

$$\mathbf{c}(t) = e^{j\omega_0}(\mathbf{I} - \mathbf{BM})\mathbf{c}(t-1) - \mathbf{BM}\tilde{\mathbf{v}}(t-1) + \tilde{\mathbf{e}}(t) \tag{17}$$

where

$$\mathbf{c}(t) = \mathbf{d}(t) - \mathbf{B}\widehat{\mathbf{d}}(t|t-1)$$

denotes cancellation error.

Note that equation (17) is almost identical with equation (8), the only difference being that the matrix $\mathbf{M}$ is now replaced with $\mathbf{BM}$. This has important practical implications as it means that by making the proper choice of the adaptation gain $\mathbf{M}$, one can not only "undo" modeling errors, but also optimize the overall system performance. In the RW case this can be achieved by setting $\mathbf{M} = \mathbf{M}_\mathrm{opt} = \mathbf{B}^{-1}\mathbf{G}_\infty$. Under such a choice the covariance matrix of the steady-state mean-squared cancellation error will reach its smallest possible value $\mathbf{P}_\infty$ – in spite of adopting an incorrect plant gain in (15)! Since the matrix $\mathbf{B}$ is unknown, an automatic gain adjustment procedure will be designed. As will be shown later, this procedure yields the estimates $\hat{\mathbf{M}}(t)$ that converge locally in mean to $\mathbf{M}_\mathrm{opt}$, exactly as desired.

A. Self-optimizing Controller

The quantity $\hat{\mathbf{M}}(t)$ will be adjusted recursively by minimizing the following measure of fit

$$\mathbf{R}_{\mathbf{yy}}(0;\mathbf{M}) = \mathrm{E}[\mathbf{y}(t;\mathbf{M})\mathbf{y}^{\mathrm{H}}(t;\mathbf{M})]$$

where $\{\mathbf{y}(t;\mathbf{M})\}$ denotes a stationary process that "settles down" in the closed-loop system for a constant value of $\mathbf{M}$ such that $\mathbf{BM} \in \mathbf{\Omega}_{\mathrm{s}}$.

Since it holds that [cf. (16)] $\mathbf{R}_{\mathbf{yy}}(0;\mathbf{M}) = \mathbf{C}_{\infty}(\mathbf{M}) + \mathbf{V}$, minimization of $\mathbf{R}_{\mathbf{yy}}(0;\mathbf{M})$ is equivalent to minimization of the covariance matrix of the steady-state mean-squared cancellation error. A simple stochastic gradient algorithm will be designed, of the form

$$\hat{\mathbf{M}}(t) = \hat{\mathbf{M}}(t-1) + \alpha\Delta\mathbf{M}(t) \tag{18}$$

where α denotes a small positive constant, and $\Delta\mathbf{M}(t)$ is chosen so as to guarantee that[1]

$$\mathbf{R}_{\mathbf{yy}}(0;\hat{\mathbf{M}} + \Delta\mathbf{M}) \leq \mathbf{R}_{\mathbf{yy}}(0;\hat{\mathbf{M}}). \tag{19}$$

While derivation of the gain tuning algorithm given in [14] was based on Wirtinger calculus, in the present study another analytic technique, known as directional derivatives, will be used [22].

Directional derivatives are related to Gâteaux differentials. Let $\mathbf{X}, \mathbf{Y} \in \mathcal{C}^{m\times n}$ be complex-valued matrices. The first-order directional derivatives of functions $f(\mathbf{X}) : \mathcal{C}^{m\times n} \to \mathcal{C}$, $\mathbf{g}(\mathbf{X}) : \mathcal{C}^{m\times n} \to \mathcal{C}^{k}$ and $\mathbf{H}(\mathbf{X}) : \mathcal{C}^{m\times n} \to \mathcal{C}^{k\times l}$ at $\mathbf{X}$ in the direction $\mathbf{Y}$ are defined as

$$\begin{aligned} d_{\overrightarrow{\mathbf{Y}}} f(\mathbf{X}) &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} f(\mathbf{X}+\epsilon\mathbf{Y}) \in \mathcal{C} \\ d_{\overrightarrow{\mathbf{Y}}} \mathbf{g}(\mathbf{X}) &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathbf{g}(\mathbf{X}+\epsilon\mathbf{Y}) \in \mathcal{C}^{k} \\ d_{\overrightarrow{\mathbf{Y}}} \mathbf{H}(\mathbf{X}) &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathbf{H}(\mathbf{X}+\epsilon\mathbf{Y}) \in \mathcal{C}^{k\times l} \end{aligned}$$

(assuming that all functions are differentiable with respect to ϵ). Unlike much more popular gradients, directional derivatives do not expand dimensions – directional derivatives of scalar, vector-valued and matrix-valued functions on matrix domain are scalar, vector-valued and matrix-valued, respectively. The analogous gradient operators have two-dimensional (matrix), three-dimensional (cubix) and four-dimensional (quatrix) representations, respectively, which considerably complicates analysis.

1) Gradient Update: To find the "right" direction $\Delta\mathbf{M}$, the first-order Taylor series expansion of $\mathbf{R}_{\mathbf{yy}}(0;\mathbf{M})$ about $\mathbf{M}$ will be used:

$$\begin{aligned} \mathbf{R}_{\mathbf{yy}}(0;\mathbf{M}+\Delta\mathbf{M}) \cong\ & \mathbf{R}_{\mathbf{yy}}(0;\mathbf{M}) \\ &+ \mathrm{E}\left\{ d_{\overrightarrow{\Delta\mathbf{M}}} \left[\mathbf{y}(t;\mathbf{M})\mathbf{y}^{\mathrm{H}}(t;\mathbf{M})\right] \right\}. \end{aligned} \tag{20}$$

[1]From here on, Δ will always denote an increment of the accompanying matrix ($\Delta\mathbf{M}$, $\Delta\mathbf{X}$, etc.), i.e., it will not be used as a stand-alone quantity.

Note that

$$\begin{aligned} d_{\overrightarrow{\Delta\mathbf{M}}} \left[\mathbf{y}(t;\mathbf{M})\mathbf{y}^{\mathrm{H}}(t;\mathbf{M})\right] =\ & d_{\overrightarrow{\Delta\mathbf{M}}} \mathbf{y}(t;\mathbf{M})\mathbf{y}^{\mathrm{H}}(t;\mathbf{M}) \\ &+ \mathbf{y}(t;\mathbf{M}) \left[d_{\overrightarrow{\Delta\mathbf{M}}} \mathbf{y}(t;\mathbf{M}) \right]^{\mathrm{H}}. \end{aligned} \tag{21}$$

Furthermore, since

$$\begin{aligned} \mathbf{y}(t;\mathbf{M}) =\ & e^{j\omega_0}(\mathbf{I}-\mathbf{BM})\mathbf{y}(t-1;\mathbf{M}) + \tilde{\mathbf{e}}(t) \\ &+ \mathbf{v}(t) - e^{j\omega_0}\mathbf{v}(t-1) \end{aligned} \tag{22}$$

one obtains

$$\begin{aligned} d_{\overrightarrow{\Delta\mathbf{M}}} \mathbf{y}(t;\mathbf{M}) = e^{j\omega_0} \Big[& (\mathbf{I}-\mathbf{BM}) d_{\overrightarrow{\Delta\mathbf{M}}} \mathbf{y}(t-1;\mathbf{M}) \\ &- \mathbf{B}\Delta\mathbf{M}\mathbf{y}(t-1;\mathbf{M}) \Big]. \end{aligned} \tag{23}$$

Since the directional derivative of $\mathbf{y}(t)$ depends on the modeling error $\mathbf{B}$, the recursive formula (23) cannot be used without modification. In the univariate case this problem was dealt with by means of applying the substitution $\beta = c_\mu/\mu$, where β and μ denote the one-dimensional counterparts of $\mathbf{B}$ and $\mathbf{M}$, respectively, and c_μ is a positive constant – for more comments on this choice see [14]. In the multivariate case the same "trick" will be used by postulating that (see remark below)

$$\mathbf{B} = c_\mu \mathbf{M}^{-1}. \tag{24}$$

Using this substitution, one obtains the modified version of (23)

$$\begin{aligned} d_{\overrightarrow{\Delta\mathbf{M}}} \mathbf{y}(t;\mathbf{M}) = e^{j\omega_0} \Big[& (1-c_\mu) d_{\overrightarrow{\Delta\mathbf{M}}} \mathbf{y}(t-1;\mathbf{M}) \\ &- c_\mu \mathbf{M}^{-1}\Delta\mathbf{M}\mathbf{y}(t-1;\mathbf{M}) \Big]. \end{aligned} \tag{25}$$

Let $\tilde{\mathbf{z}}(t;\mathbf{M})$ be the quantity defined implicitly by

$$d_{\overrightarrow{\Delta\mathbf{M}}} \mathbf{y}(t;\mathbf{M}) = \mathbf{M}^{-1}\Delta\mathbf{M}\tilde{\mathbf{z}}(t;\mathbf{M}) \ . \tag{26}$$

According to (20), one should choose $\Delta\mathbf{M}$ in such a way that

$$d_{\overrightarrow{\Delta\mathbf{M}}} \left[\mathbf{y}(t;\mathbf{M})\mathbf{y}^{\mathrm{H}}(t;\mathbf{M})\right] \leq \mathbf{O}. \tag{27}$$

After combining (21) with (26), one obtains

$$\begin{aligned} d_{\overrightarrow{\Delta\mathbf{M}}} \left[\mathbf{y}(t;\mathbf{M})\mathbf{y}^{\mathrm{H}}(t;\mathbf{M})\right] =\ & \mathbf{M}^{-1}\Delta\mathbf{M}\tilde{\mathbf{z}}(t;\mathbf{M})\mathbf{y}^{\mathrm{H}}(t;\mathbf{M}) \\ &+ \mathbf{y}(t;\mathbf{M})\tilde{\mathbf{z}}^{\mathrm{H}}(t;\mathbf{M})\Delta\mathbf{M}^{\mathrm{H}}\mathbf{M}^{-\mathrm{H}}. \end{aligned} \tag{28}$$

Therefore, in order to fulfill (27), one should set

$$\Delta\mathbf{M} = -\mathbf{M}\mathbf{y}(t)\tilde{\mathbf{z}}^{\mathrm{H}}(t)\mathbf{P} \tag{29}$$

where $\mathbf{P}$ is an arbitrary positive definite matrix.

Summarizing all steps of our derivation, and setting $\mathbf{P} = \mathbf{I}$ for simplicity, one arrives at the following gradient algorithm for updating the weight matrix

$$\begin{aligned} \tilde{\mathbf{z}}(t) &= e^{j\omega_0}\left[(1-c_\mu)\tilde{\mathbf{z}}(t-1) - c_\mu\mathbf{y}(t-1)\right] \\ \hat{\mathbf{M}}(t) &= \hat{\mathbf{M}}(t-1)\left[\mathbf{I} - \alpha\mathbf{y}(t)\tilde{\mathbf{z}}^{\mathrm{H}}(t)\right] \end{aligned} \tag{30}$$

where the first recursion was obtained by rewriting (25) in terms of $\tilde{\mathbf{z}}(t)$.

Remark: Substitution (24) is the "controversial" part of our derivation. Does one have the right to introduce such modifications? To answer this question, one should realize that the gain update algorithm is operated in a closed loop. It is known that feedback, under certain circumstances, can correct design errors. A good example of exploiting such self-correction capabilities of adaptive feedback controllers is the coefficient fixing technique, used to robustify self-tuning minimum-variance regulators. Coefficient fixing means that, for controller design purposes, one of the system coefficients is deliberately set to an (almost) arbitrary value and not estimated. In spite of this obviously erroneous assumption (the adopted value of the fixed coefficient generally differs from the true value), the closed-loop control system can be shown to converge locally in mean to the optimal one, which means that the design error is compensated by feedback [17], [18], [19].

The substitution (24) is hardly intuitive and therefore its feasibility cannot be judged *a priori*. But our main point here is that *any* way of coping with unknown modeling errors is feasible, provided that it guarantees mean convergence of gain estimates to their optimal values. In [14] it was shown that this is the case for the univariate version of SONIC. In the next section we will prove that the same holds true for its multivariate version

2) Normalized Updates: In addition to (30), two normalized versions of the gain update can be used: the trace algorithm

$$\begin{aligned}\tilde{r}(t) &= \rho\tilde{r}(t-1) + |\tilde{\mathbf{z}}(t)|^2 \\ \hat{\mathbf{M}}(t) &= \hat{\mathbf{M}}(t-1)\left[\mathbf{I} - \frac{\mathbf{y}(t)\tilde{\mathbf{z}}^{\mathrm{H}}(t)}{\tilde{r}(t)}\right]\end{aligned} \tag{31}$$

and, the computationally more involved, directional algorithm [obtained by taking $\mathbf{P} = \tilde{\mathbf{R}}^{-1}$ in (29)]

$$\begin{aligned}\tilde{\mathbf{R}}(t) &= \rho\tilde{\mathbf{R}}(t-1) + \tilde{\mathbf{z}}(t)\tilde{\mathbf{z}}^{\mathrm{H}}(t) \\ \hat{\mathbf{M}}(t) &= \hat{\mathbf{M}}(t-1)\left[\mathbf{I} - \mathbf{y}(t)\tilde{\mathbf{z}}^{\mathrm{H}}(t)\tilde{\mathbf{R}}^{-1}(t)\right].\end{aligned} \tag{32}$$

The adjective "trace" refers to the fact that the normalizing factor $\tilde{r}(t)$ in (31) is equal to $\mathrm{tr}\{\tilde{\mathbf{R}}(t)\}$, where $\tilde{\mathbf{R}}(t)$ denotes the normalizing matrix in (32). In both cases ρ, $0 < \rho < 1$, denotes the forgetting constant determining the effective length of the local averaging window. The recommended values of ρ are those from the interval [0.999,0.9999].

Normalization makes the algorithms (31) and (32) scale invariant. Suppose that the gradient algorithm (30) is run on a scaled data sequence $\{\mathbf{y}'(t)\}$, $\mathbf{y}'(t) = \delta\mathbf{y}(t)$, $\delta \neq 0$. Then, to obtain results identical with the original ones, the stepsize α should be replaced with α/δ^2. Both normalized algorithms are free of this ambiguity. Directional normalization provides more careful scaling than trace normalization as it takes into account the relative "strength" of different measurements.

Note that the matrix $\tilde{\mathbf{P}}(t) = \tilde{\mathbf{R}}^{-1}(t)$ can be computed recursively

$$\tilde{\mathbf{P}}(t) = \frac{1}{\rho}\left[\tilde{\mathbf{P}}(t-1) - \frac{\tilde{\mathbf{P}}(t-1)\tilde{\mathbf{z}}(t)\tilde{\mathbf{z}}^{\mathrm{H}}(t)\tilde{\mathbf{P}}(t-1)}{\rho + \tilde{\mathbf{z}}^{\mathrm{H}}(t)\tilde{\mathbf{P}}(t-1)\tilde{\mathbf{z}}(t)}\right]. \tag{33}$$

It will be shown that in the univariate case ($n = 1$) both normalized updates are approximately equivalent to those incorporated in the SONIC controller:

$$\begin{aligned}z(t) &= e^{j\omega_0}\left[(1-c_\mu)z(t-1) - \frac{c_\mu}{\widehat{\mu}(t-1)}y(t-1)\right] \\ r(t) &= \rho r(t-1) + |z(t)|^2 \\ \widehat{\mu}(t) &= \widehat{\mu}(t-1) - \frac{y(t)z^*(t)}{r(t)}\end{aligned} \tag{34}$$

where μ denotes a complex-valued scalar gain.

To show equivalence, note that the scalar counterpart of both normalized updates can be written down in the form

$$\begin{aligned}\tilde{z}(t) &= e^{j\omega_0}\left[(1-c_\mu)\tilde{z}(t-1) - c_\mu y(t-1)\right] \\ \tilde{r}(t) &= \rho\tilde{r}(t-1) + |\tilde{z}(t)|^2 \\ \widehat{\mu}(t) &= \widehat{\mu}(t-1)\left[1 - \frac{y(t)\tilde{z}^*(t)}{\tilde{r}(t)}\right].\end{aligned} \tag{35}$$

When $\widehat{\mu}(t)$ and $\tilde{r}(t)$ change slowly compared to $\tilde{z}(t)$, i.e., $\widehat{\mu}(t) \cong \widehat{\mu}(t-1) \cong \mu$ (which takes place for ρ close to 1), one can apply the following substitutions

$$\tilde{z}(t) = \mu z(t), \quad \tilde{r}(t) = |\mu|^2 r(t)$$

which convert (35) into (34).

B. Mean Convergence Analysis

For simplicity, consider gradient updates (30). It will be shown that under (A1)–(A2) the estimates $\hat{\mathbf{M}}(t)$ converge (locally) in mean to the optimal solution $\mathbf{M}_{\mathrm{opt}} = \mathbf{B}^{-1}\mathbf{G}_\infty$.

It is known that the tracking behavior of constant-gain (finite-memory) estimation algorithms, such as (37), can be studied by examining the properties of the associated ordinary differential equations (ODEs) [23], [24]. Denote by $\{\mathbf{y}(t;\mathbf{M})\}$ and $\{\tilde{\mathbf{z}}(t;\mathbf{M})\}$ the stationary processes observed in the closed-loop system for a constant value of $\mathbf{M} : \mathbf{BM} \in \mathbf{\Omega}_{\mathrm{s}}$. For sufficiently small values of α, the estimates $\hat{\mathbf{M}}(t)$ wander around $\mathbf{M}_0$ – the stable equilibrium point of ODE associated with (30)

$$\dot{\mathbf{M}} = \mathbf{F}(\mathbf{M}) \tag{36}$$

where

$$\mathbf{F}(\mathbf{M}) = -\mathbf{M}\,\mathrm{E}\left[\mathbf{y}(t;\mathbf{M})\tilde{\mathbf{z}}^{\mathrm{H}}(t;\mathbf{M})\right].$$

It can be shown that

Theorem 1

Under assumptions (A1)–(A2) it holds that $\mathbf{M}_0 = \mathbf{M}_{\mathrm{opt}}$ is a unique stable equilibrium point of ODE (36). The minimum is attained for $\mathbf{M} = \mathbf{G}_\infty$.

Proof - see Appendix II.

The same result can be proved for normalized updates (31) and (32), respectively.

C. Multivariate SONIC

Combining all earlier results, the trace (recommended) version of the multivariate SONIC algorithm can be summarized

as follows

$$\begin{aligned}\tilde{\mathbf{z}}(t) &= e^{j\omega_0}\left[(1-c_\mu)\tilde{\mathbf{z}}(t-1) - c_\mu \mathbf{y}(t-1)\right]\\ \tilde{r}(t) &= \rho\tilde{r}(t-1) + |\tilde{\mathbf{z}}(t)|^2\\ \hat{\mathbf{M}}(t) &= \hat{\mathbf{M}}(t-1)\left[\mathbf{I} - \frac{\mathbf{y}(t)\tilde{\mathbf{z}}^{\mathrm{H}}(t)}{\tilde{r}(t)}\right]\\ \widehat{\mathbf{d}}(t+1|t) &= e^{j\omega_0}[\widehat{\mathbf{d}}(t|t-1) + \hat{\mathbf{M}}(t)\mathbf{y}(t)]\\ \mathbf{u}(t) &= -\mathbf{K}_{\mathrm{n}}^{-1}\widehat{\mathbf{d}}(t+1|t). \end{aligned} \quad (37)$$

SONIC consists of two loops: the inner, cancellation loop [the last two recursions of (37)], which predicts and cancels disturbance, and the outer, self-optimization loop [the first two recursions of (37)], which automatically adjusts the gain matrix so as to optimize the overall system performance. Even though multivariate SONIC resembles its univariate counterpart, some of its details (such as the multiplicative, rather than additive, gain updates) are hardly deducible from the single input - single output solution presented in [14].

D. Selection of Design Parameters

The influence of design parameters c_μ and ρ on tracking properties of SONIC were studied in [14] for cisoids with random amplitude drift. It was shown there (see Section VI-E in [14]) that the closer that $1-\rho$ becomes to 0, the longer it takes for the algorithm to readjust the adaptation gain $\widehat{\mu}(t)$ when the operating conditions change (plant, noise and/or interference). On the other hand, small values of $(1-\rho)/c_\mu$ are required to make the steady-state fluctuations of $\widehat{\mu}(t)$ around μ_{opt} small. Finally, the value of c_μ, $0 < c_\mu < 1$, should be trimmed to the rate of plant/disturbance variation – the faster the changes, the larger the recommended values of c_μ. The following rules of thumb, which extend to the multivariable case, seem to work pretty well in practice:

- Choose c_μ in the range $[0.005/f_{\mathrm{s}}, 0.1/f_{\mathrm{s}}]$, where f_{s} denotes sampling frequency expressed in kHz.
- Choose ρ such that $1-\rho \ll c_\mu$.

Under 1 kHz sampling our default choice is: $c_\mu = 0.01$ and $\rho = 0.999$. As long as the conditions specified above are fulfilled, cancellation performance of SONIC does not critically depend on the selection of c_μ and ρ. The algorithm's mean transient response to large abrupt plant and/or disturbance changes last for several (usually less than 10) periods $T_0 = 2\pi/\omega_0$ of the sinusoidal interference. When the changes are less significant, the length of the transient phase is usually much shorter, often taking values smaller than T_0.

Although the operating range for SONIC is usually limited to the bandwidth [0, 500 Hz] (where passive methods fail to work), the algorithm can be safely used in the extended (kHz) frequency range, provided that the sampling frequency f_{s} is chosen appropriately. Note that the recommended parameter settings depend on f_{s}.

E. Comparison with the Adaptive FXLMS Algorithm

All advantages of SONIC compared to adaptive FXLMS (i.e., FXLMS with on-line plant identification), pointed out in [14] and [15], carry on to the multivariate case. For the readers' convenience they will be briefly summarized below.

1) Parsimony. When the FXLMS approach is taken, each secondary path is usually modelled as a finite impulse response (FIR) system of order M

$$L_{i_1 i_2}(q^{-1}) = \sum_{m=1}^{M} l_m^{i_1 i_2} q^{-m}, \quad 1 \le i_1, i_2 \le n.$$

To maintain stability and satisfactory performance of the closed loop system, M must be sufficiently large, typically in excess of 100 under 1 kHz sampling (for higher sampling rates, the value of M must be proportionally increased). Large values of M are unavoidable when canceling wideband disturbances (in a feedforward compensation configuration), simply because the frequency response of the plant must be modelled over the entire frequency band $(-\pi, \pi]$. For sinusoidal disturbances the situation is different – all that is needed for steady-state cancellation purposes, is an estimate of the plant's gain at the frequency ω_0: $\widehat{k}_{\mathrm{p}}^{i_1 i_2} = \sum_{m=1}^{M} \widehat{l}_m^{i_1 i_2} e^{-jm\omega_0}$. This means that, in the case considered, the FIR representation adopted for FXLMS is grossly overparameterized – it requires estimation of $n^2 M$ parameters instead of $2n^2$ parameters (SONIC). This lack of parsimony results in a slower response of FXLMS to plant and/or disturbance changes, as well as in some degradation of its steady state performance.

2) Identifiability. When identification/tracking of secondary paths is carried out on-line, i.e., in a closed loop, linear feedback may cause identifiability problems: parameter estimates may not converge to their true values, if they converge at all. To restore identifiability, a low-intensity random perturbation (dither) can be added to the input signal [12]. Unfortunately, injection of such an auxiliary noise disturbs operation of the canceling system and causes deterioration of its performance. Note that SONIC is free of this drawback – due to gain fixing, $\widehat{\mathbf{M}}(t)$ converges in mean to its optimal value without additional excitation.

3) Computational complexity. Computational burden associated with the adaptive FXLMS algorithm is equal to $n^2(6M+12)$ real multiply/add operations per time update for regular LMS, and increases to $n^2(8M+13)$ operations when the normalized LMS (NLMS) recursions are used – see [14].

The computational complexity of the gradient SONIC algorithm (37) is equal to $16n^2+14n$ real multiply/add operations per time update. When the trace algorithm (31) is used, instead of the gradient one, the computational burden increases slightly to $16n^2+16n+2$ real multiply/add operations and 2 real division operations per time update. Finally, in the computationally most expensive version, incorporating directional normalization (32), SONIC requires $22n^2+16n+2$ real multiply/add operations and $2n$ real division operations per time update. Therefore, for a single sinusoidal disturbance and for realistic values of M ($M \ge 100$), SONIC, compared to FXLMS, offers huge computational savings (it is at least 30 times less demanding).

When more than one harmonic component is to be eliminated, the computational complexity of SONIC grows linearly

with the number of harmonics (see Section V-B), while the complexity of FXLMS grows at a slower rate (because the order M of the identified plant model does not need to be increased). In spite of this, when the number of attenuated harmonics is smaller than 30, SONIC remains computationally more attractive than FXLMS.

4) Self-optimization. SONIC automatically adjusts its gain matrix $\mathbf{M}$ so as to minimize the mean-squared cancellation error. Even though several extended FXLMS schemes, equipped with mechanisms for on-line adjustment of LMS step-sizes, were proposed in the literature (see e.g. [25] and the referenses therein), we have found out that in the presence of sinusoidal disturbances they do not perform satisfactorily [14].

A more detailed comparison of SONIC and FXLMS, including the results of simulation experiments, can be found in [14] and [15].

V. Comments and extensions

A. Frequency-adaptive Algorithm

Extending SONIC to the case of unknown, and possibly time-varying, frequency is straightforward. A frequency estimation block can be added and, using the certainty equivalence principle, one can replace the constant-known frequency ω_0 with its estimated value. Suppose that an extra (scalar) reference signal $s(t)$ is available from the detection sensor, localized close to the source of sound/vibration, and unaffected by the control signal $\mathbf{u}(t)$. Then the frequency-adaptive extended SONIC (xSONIC) algorithm can be designed as follows [26]:

$$\begin{aligned}
\tilde{\mathbf{z}}(t) &= e^{j\widehat{\omega}(t|t-1)}\left[(1-c_\mu)\tilde{\mathbf{z}}(t-1) - c_\mu \mathbf{y}(t-1)\right] \\
\tilde{r}(t) &= \rho\tilde{r}(t-1) + |\tilde{\mathbf{z}}(t)|^2 \\
\hat{\mathbf{M}}(t) &= \hat{\mathbf{M}}(t-1)\left[\mathbf{I} - \frac{\mathbf{y}(t)\tilde{\mathbf{z}}^{\mathrm{H}}(t)}{\tilde{r}(t)}\right] \\
\widehat{\mathbf{d}}(t+1|t) &= e^{j\widehat{\omega}(t|t-1)}[\widehat{\mathbf{d}}(t|t-1) + \hat{\mathbf{M}}(t)\mathbf{y}(t)] \\
\widehat{\omega}(t+1|t) &= g[\mathcal{S}(t)] \\
\mathbf{u}(t) &= -\mathbf{K}_{\mathrm{n}}^{-1}\widehat{\mathbf{d}}(t+1|t)
\end{aligned} \tag{38}$$

where $\widehat{\omega}(t+1|t)$ denotes the one-step-ahead instantaneous frequency predictor based on the history of the reference signal: $\mathcal{S}(t) = \{s(i), i \le t\}$.

The frequency estimation routine $g[\cdot]$ can be either parametric, i.e., model-based (different variants of adaptive notch filters), or nonparametric, i.e., based on decomposition of the signal space (FFT, MUSIC, ESPRIT etc.) When the reference signal is not available, frequency estimation can be based on examination of a selected error signal, preferably the one with the highest SNR – see [26] for more details on the frequency-adaptive case.

B. Multiple Frequencies

When disturbance consists of many sinusoidal components, one can use the parallelized version of the proposed algorithm. A separate block of the form (37) or (38) is required for each signal component - for more details see [27]. When the number of frequency components is unknown, one can estimate it using the method proposed in [28].

C. Initialization and Safety Measures

All results presented in the previous section characterize *local* properties of the cancellation algorithm, which means that the gain estimates may, but not necessarily have to, converge to the desired values. To guarantee mean convergence of $\hat{\mathbf{M}}$ to the optimal value $\mathbf{M}_0 = \mathbf{M}_{\mathrm{opt}}$, the initial estimate $\hat{\mathbf{M}}(0)$ must belong to the domain of attraction of $\mathbf{M}_0$, further denoted by $\mathbf{\Omega}_0$. Vaguely speaking, the domain $\mathbf{\Omega}_0 \in \mathbf{\Omega}_{\mathrm{s}}$ "shrinks" as the modeling error grows. Hence, to minimize the risk of divergence during initialization of the algorithm, one can set $\mathbf{K}_{\mathrm{n}}$ to $\widehat{\mathbf{K}}_{\mathrm{p}}$ – the estimated gain matrix, obtained as a result of an open-loop identification experiment. During such experiment, actuators are sequentially activated (one at a time) to generate a probing signal of the form $u_i(t) = \sin\omega_0 t$. Then, based on steady-state responses collected by all sensors, one can estimate the ith column of the matrix $\mathbf{K}_{\mathrm{p}}$ using classical frequency-domain identification tools, such as the method of empirical transfer function estimation (ETFE) described in [29]. According to (8), when $\mathbf{K}_{\mathrm{n}} = \mathbf{K}_{\mathrm{p}}$, i.e., $\mathbf{B} = \mathbf{I}$, stabitity of the closed-loop system can be guaranteed by setting $\mathbf{M} = \mu\mathbf{I}$, where $\mu \in (0, 2)$. Following this observation we recommend using $\hat{\mathbf{M}}(0) = \mu_0\mathbf{I}$, where $\mu_0 \ll 1$ denotes a small positive constant.

All fixes described above are needed *only* in the initial control phase, to smoothly start operation of the closed-loop system. After successful initialization, SONIC is capable of adapting on-line to slow changes in disturbance ($\mathbf{E}$), measurement noise ($\mathbf{V}$), and system ($\mathbf{K}_{\mathrm{p}}$) characteristics.

Similar to the univariate case, to avoid erratic behavior of the canceller in the startup phase, or in the presence of rapid disturbance and/or plant changes, some safety measures are advisable, such as limitation of the "magnitude" and one-step rate of change of $\hat{\mathbf{M}}(t)$ – see [14] for more details.

VI. Simulation results

A. Mean Convergence

To check the asymptotic mean convergence properties of the proposed rejection scheme, several simulations were performed under different conditions:

1) For two choices of the plant: for a dynamic system

$$\mathbf{y}(t) = \mathbf{L}_{\mathrm{p}}(q^{-1})\mathbf{u}(t-1) + \mathbf{d}(t) + \mathbf{v}(t)$$

with impulse response depicted in Fig. 2 (established experimentally using data from a real acoustic encounter), and for its static counterpart

$$\mathbf{y}(t) = \mathbf{L}_{\mathrm{p}}(e^{-j\omega_0})\mathbf{u}(t-1) + \mathbf{d}(t) + \mathbf{v}(t).$$

In both cases

$$\mathbf{V} = 0.03\begin{bmatrix} 1 & 0.3j \\ -0.3j & 1 \end{bmatrix}. \tag{39}$$

2) For four choices of amplitude modulation intensity:

$$\mathbf{E} = \sigma_e^2\begin{bmatrix} 1 & 0.5 \\ 0.5 & 2 \end{bmatrix}$$

where $\sigma_e^2 \in \{3\cdot 10^{-7}, 10^{-6}, 3\cdot 10^{-6}, 10^{-5}\}$.

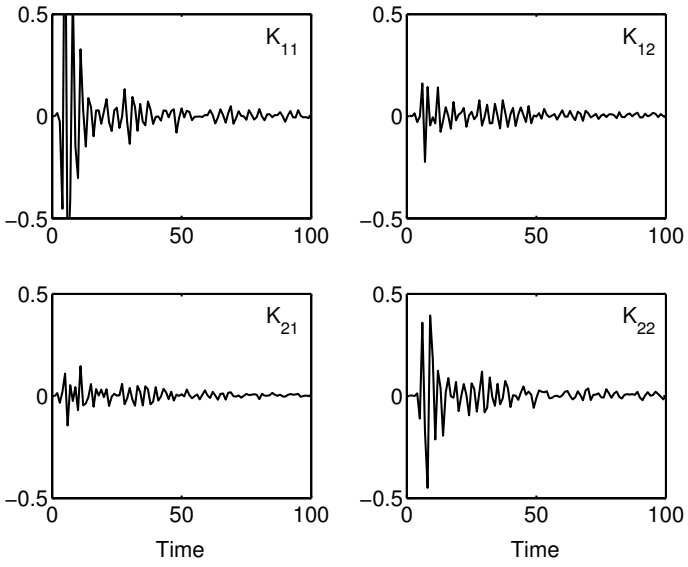

Fig. 2. Impulse response of the plant used in all simulations.

The trace version of the algorithm was used with $\rho = 0.9998$ and $c_\mu = 0.005$. While the assumed plant model was equal to the true one, $\mathbf{K}_\mathrm{n} = \mathbf{K}_\mathrm{p}$, the algorithm was initialized with

$$\hat{\mathbf{M}}(0) = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix} e^{j\pi/4} . \qquad (40)$$

This means that, initially, a significant phase bias was present in $\hat{\mathbf{M}}(t)$.

Due to the presence of a nonstationary disturbance (note that magnitudes of $\mathbf{d}(t)$ grow unbounded), all results were obtained by combining ensemble averaging (50 realizations of $\{\mathbf{e}(t)\}$ and $\{\mathbf{v}(t)\}$) and time averaging (100000 samples, initial 50000 samples were discarded to ensure that steady state was reached). The frequency of the disturbance was in all cases equal to $\omega_0 = \pi/2$.

The results are gathered in Table 1. Since it is more meaningful to check the eigenvalues of the gain matrix $\mathbf{M}$, rather than the elements of $\mathbf{M}$, Table 1 shows mean values of $\lambda_1(\mathbf{M})$ and $\lambda_2(\mathbf{M})$, respectively.

In case of a static plant, almost perfect agreement of theory and simulations can be observed. Although some discrepancies occur for the dynamic plant, they are within acceptable limits. Most importantly, the phase is estimated with only small biases which guarantees stability of the closed loop system.

B. Transient Behavior

Transient behavior of the proposed scheme for the dynamic plant is depicted in Fig. 3. This time, the trace algorithm was used with $\rho = 0.999$, $c_\mu = 0.005$, $\hat{\mathbf{M}}(0)$ was given by (40), and $\sigma_e^2 = 10^{-5}$. Additionally, to avoid rough start, during the initial 2000 sampling periods, the quantity $\widetilde{\mathbf{z}}(t)$ was evaluated, but $\hat{\mathbf{M}}(t)$ was kept at its initial value and not estimated. The adaptation lock was released at $t = 2001$.

Observe that the response to phase errors is quicker than to magnitude errors. Similar difference in sensitivity to different kinds of errors was observed for the univariate SONIC.

σ_e^2	λ_1^{opt}	λ_2^{opt}	$\lvert\lambda_1\rvert$	$\lvert\lambda_2\rvert$	Arg λ_1	Arg λ_2
$3 \cdot 10^{-7}$	0.0027	0.0050	0.0028	0.0052	1.25°	0.17°
$1 \cdot 10^{-6}$	0.0050	0.0092	0.0052	0.0093	-0.48°	0.25°
$3 \cdot 10^{-6}$	0.0087	0.0158	0.0088	0.0161	-0.78°	-0.04°
$1 \cdot 10^{-5}$	0.0158	0.0287	0.0159	0.0289	0.22°	0.13°

σ_e^2	λ_1^{opt}	λ_2^{opt}	$\lvert\lambda_1\rvert$	$\lvert\lambda_2\rvert$	Arg λ_1	Arg λ_2
$3 \cdot 10^{-7}$	0.0027	0.0050	0.0032	0.0059	5.23°	2.54°
$1 \cdot 10^{-6}$	0.0050	0.0092	0.0056	0.0108	4.28°	2.74°
$3 \cdot 10^{-6}$	0.0087	0.0158	0.0099	0.0188	3.19°	2.60°
$1 \cdot 10^{-5}$	0.0158	0.0287	0.0184	0.0345	4.30°	3.89°

TABLE I
MEAN CONVERGENCE RESULTS FOR THE STATIC PLANT (UPPER TABLE) AND DYNAMIC PLANT (LOWER TABLE).

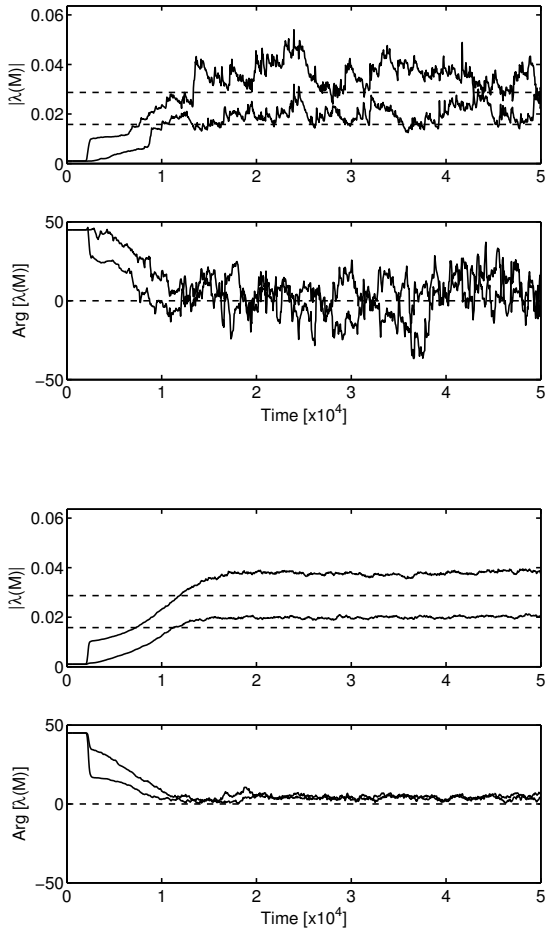

Fig. 3. Transient behavior of the proposed scheme. Two upper figures – results of a typical simulation. Two lower figures – average of 50 simulations.

C. Comparison with FXLMS

Suppose that the transfer function of the plant $\mathbf{L}_\mathrm{p}(q^{-1})$ does not change with time, and that the true gain matrix $\mathbf{K}_\mathrm{p} = \mathbf{L}_\mathrm{p}(e^{-j\omega_0})$ is known, e.g. it was established based on the results of an open-loop identification experiment. In this simple case cancellation of the narrowband disturbance can be

achieved using the following FXLMS algorithm

$$\begin{aligned}\mathbf{R}_{\mathrm{f}}(t) &= \mathbf{K}_{\mathrm{p}} r(t-1)\\ \mathbf{w}(t) &= \mathbf{w}(t-1) + \eta \mathbf{R}_{\mathrm{f}}^{\mathrm{H}}(t)\mathbf{y}(t)\\ \mathbf{u}(t) &= -\mathbf{w}(t) r(t)\end{aligned} \tag{41}$$

where $\mathbf{w}(t)$ denotes the vector of control gains, $r(t) = e^{j\omega_0 t}$ denotes the artificially generated reference signal, $\mathbf{R}_{\mathrm{f}}(t) \cong \mathbf{L}_{\mathrm{p}}(q^{-1})r(t)$ is the filtered reference signal matrix, and $\eta > 0$ denotes the adaptation step-size of the LMS algorithm.

Table II shows comparison of the mean-squared cancellation performance of SONIC and the performance of the optimally tuned FXLMS algorithm (41). The simulated plant as well as all SONIC settings were identical with those described in Section VI A above. The optimal values of η were searched numerically in the interval $[0.01, 0.00025]$ with a step equal to 0.00025. The optimal step-size depends on which channel is optimized – the quantity η_1^{opt}, shown in Table II, denotes the value that minimizes the mean-squared cancellation error observed at the first system output (MSE_1), and η_2^{opt} – the one that minimizes cancellation error at the second output (MSE_2).

Note that, in the channels that are optimized, the well-tuned FXLMS controller yields similar results as SONIC (in order to simultaneously optimize performance in both channels, the scalar gain η should be replaced with a matrix gain). However, unlike SONIC – which has self-optimization and self-correction capabilities – FXLMS may work rather poorly under nonstationary operating conditions. Suppose, for example, that the gain matrix $\mathbf{K}_{\mathrm{p}}$ slowly varies with time. In a case like this, the constant-known gain $\mathbf{K}_{\mathrm{p}}$ can be replaced in (41) with its current estimate $\widehat{\mathbf{K}}_{\mathrm{p}}(t)$ obtained by means of on-line plant identification. To avoid nonidentifiability problems, which occur when identification is carried out in a closed loop, it is usually recommended to add to the input signal an artificially generated white noise perturbation [12]. Even though for wideband disturbances such an auxiliary noise technique usually works satisfactorily, it fails in the presence of narrowband disturbances because the estimated plant gains are strongly biased at the frequency ω_0 (which is the frequency of our interest). This frequency selective bias effect is caused by the fact that identification is carried in a closed loop and hence the estimation results are affected by the notch filtering action of the FXLMS controller.

To demonstrate this effect, the frequency response of the (time-invariant) simulated plant was estimated by means of fitting 4 FIR models of order $M = 512$. Identification was performed in a closed loop using the auxiliary noise technique, in the case where[2] $\sigma_{\mathrm{e}}^2 = 3 \cdot 10^{-7}$ and $\eta = 0.0005$. Even though the variance of the auxiliary noise $\sigma_{\mathrm{a}}^2 = 0.0025$ was pretty high, compared to the measurement noise variance, the FIR-based estimates of the plant's frequency response were strongly biased at the frequency ω_0 – see Fig. 4. When the auxiliary noise was switched off and the plant gain in (41) was frozen at its last closed-loop estimate, the mean-squared canceling errors increased from $\mathrm{MSE}_1 = 1{\cdot}10^{-4}$ and $\mathrm{MSE}_2 = 2.45{\cdot}10^{-4}$ (see Table II) to $\mathrm{MSE}_1 = 8.13{\cdot}10^{-4}$ and $\mathrm{MSE}_2 = 1.49{\cdot}10^{-3}$, respectively, i.e., they grew approximately 8 times. In the presence of auxiliary noise, i.e., in the on-line tuning mode, system performance deteriorated even further.

[2]Similar effect was observed for the other settings shown in Table II.

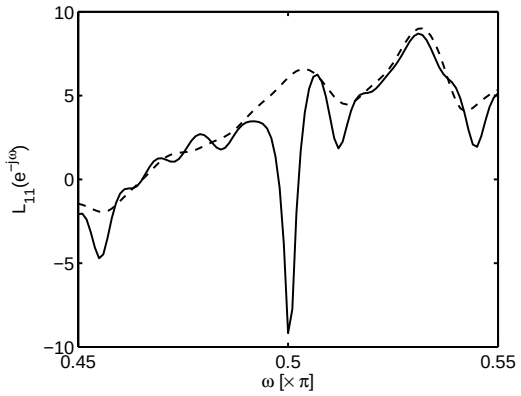

Fig. 4. Closed-loop estimates (solid line) of the plant's frequency response $L_{11}(e^{-j\omega})$ (broken line). The sharp notch appears at the frequency of the cancelled narrowband disturbance.

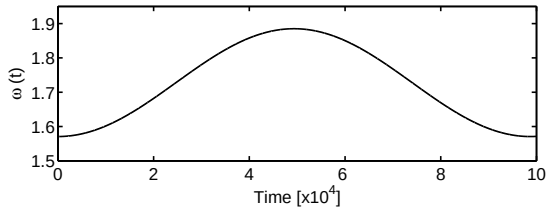

Fig. 5. Frequency trajectory for $\gamma = 10^{-5}$.

D. Cancellation of Disturbances with Time-varying Frequency

In this experiment, the instantaneous frequency of the disturbance was governed by

$$\omega(t) = \frac{\omega_1 + \omega_2}{2} + \frac{\omega_2 - \omega_1}{2}\cos[2\gamma t/(\omega_2 - \omega_1)]$$

where $\omega_1 = \pi/2$ and $\omega_2 = 6\pi/10$. A fragment of the frequency trajectory obtained for $\gamma = 10^{-5}$ is depicted in Fig. 5. Assuming 1 kHz sampling, for such a value of γ the frequency varies between 250 Hz and 300 Hz with the largest rate of variation approximately equal to 1.6 Hz/s.

The frequency of the disturbance was estimated from a reference signal $s(t)$ using the following, improved version of the frequency estimation algorithm presented in [30], which combines frequency tracking with frequency rate tracking

$$\begin{aligned}\epsilon(t) &= s(t) - \widehat{s}(t|t-1)\\ \widehat{s}(t+1|t) &= e^{j\widehat{\omega}(t|t-1)}[\widehat{s}(t|t-1) + 0.05\epsilon(t)]\\ \widehat{\alpha}(t) &= \widehat{\alpha}(t-1) + 0.0002 g(t)\\ \widehat{\omega}(t+1|t) &= \widehat{\omega}(t|t-1) + \widehat{\alpha}(t) + 0.02 g(t)\\ g(t) &= \mathrm{Arg}\left[\frac{\widehat{s}(t+1|t)}{\widehat{s}(t|t-1)e^{j\widehat{\omega}(t|t-1)}}\right].\end{aligned}$$

In the algorithm shown above $\widehat{\alpha}(t)$ is an estimate of the rate of change $\alpha(t) = \omega(t) - \omega(t-1)$.

The reference was contaminated by zero-mean circular white Gaussian noise $v_s(t)$. The signal-to-noise ratio was equal to $10\log_{10}\{|s(t)|^2/\mathrm{E}[|v_s(t)|^2]\} = 20$ dB.

σ_e^2	FXLMS				SONIC
	η_1^{opt}	MSE_1 MSE_2	η_2^{opt}	MSE_1 MSE_2	MSE_1 MSE_2
$3 \cdot 10^{-7}$	$5.00 \cdot 10^{-4}$	$1.00 \cdot 10^{-4}$ $2.45 \cdot 10^{-4}$	$1.75 \cdot 10^{-3}$	$2.07 \cdot 10^{-4}$ $1.43 \cdot 10^{-4}$	$1.05 \cdot 10^{-4}$ $1.42 \cdot 10^{-4}$
$1 \cdot 10^{-6}$	$1.00 \cdot 10^{-3}$	$1.89 \cdot 10^{-4}$ $4.21 \cdot 10^{-4}$	$2.50 \cdot 10^{-3}$	$3.05 \cdot 10^{-4}$ $2.76 \cdot 10^{-4}$	$1.86 \cdot 10^{-4}$ $2.72 \cdot 10^{-4}$
$3 \cdot 10^{-6}$	$1.50 \cdot 10^{-3}$	$3.29 \cdot 10^{-4}$ $9.01 \cdot 10^{-4}$	$6.00 \cdot 10^{-3}$	$7.34 \cdot 10^{-4}$ $5.14 \cdot 10^{-4}$	$3.34 \cdot 10^{-4}$ $4.97 \cdot 10^{-4}$
$1 \cdot 10^{-5}$	$3.00 \cdot 10^{-3}$	$6.54 \cdot 10^{-4}$ $1.62 \cdot 10^{-3}$	$1.00 \cdot 10^{-2}$	$1.52 \cdot 10^{-3}$ $1.02 \cdot 10^{-3}$	$6.45 \cdot 10^{-4}$ $9.78 \cdot 10^{-4}$

TABLE II
COMPARISON OF THE MEAN-SQUARED CANCELLATION PERFORMANCE OF THE OPTIMALLY TUNED FXLMS ALGORITHM WITH THE PERFORMANCE OF SONIC.

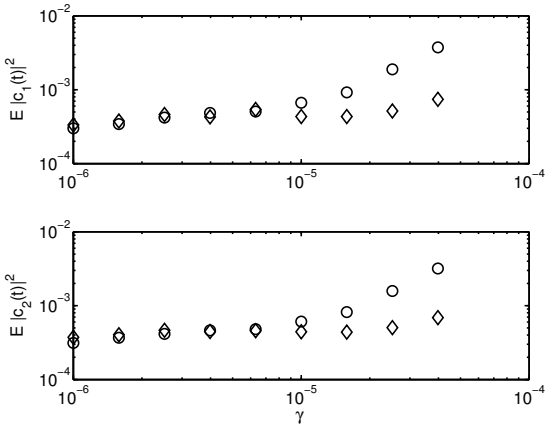

Fig. 6. Comparison of mean-squared cancellation errors for time-varying frequency and two choices of nominal plant model. Diamonds - full knowledge of the plant. Circles - knowledge of the plant's gain at a single frequency.

The amplitudes of disturbance components were constant and equal to 1 (i.e., $\mathbf{E} = \mathbf{O}$), and the covariance matrix of the wideband system noise $\mathbf{v}(t)$ was given by (39).

Fig. 6 shows comparison of steady-state mean-squared cancellation errors for $\rho = 0.999$, $c_\mu = 0.005$ and for two choices of the nominal model: $\mathbf{K}_{\mathrm{n}} = \mathbf{L}_{\mathrm{p}}(e^{-j\widehat{\omega}(t+1|t)})$ (which results in $\mathbf{B} \approx \mathbf{I}$) and $\mathbf{K}_{\mathrm{n}} = \mathbf{L}_{\mathrm{p}}(e^{-j\pi/2})$ (which results in a time varying modeling error). For $\gamma < 10^{-5}$ the loss caused by poor knowledge of the plant is marginal. For $\gamma = 4 \cdot 10^{-5}$ the difference reaches 10 dB, but the cancellation performance is still satisfactory (more than 20 dB). However, for even larger values of γ occasional bursts were observed.

The error-compensating behavior of the proposed algorithm for $\gamma = 10^{-5}$ and the second choice of nominal model is illustrated in Fig. 7. Again, the algorithm was initialized using (40). The evolution of the eigenvalues of the matrix $\mathbf{BM}$ is caused by rapid changes in modeling errors that occur when a local resonance of the plant (typical of acoustic systems) is crossed over – this effect is *not* caused by the proposed algorithm.

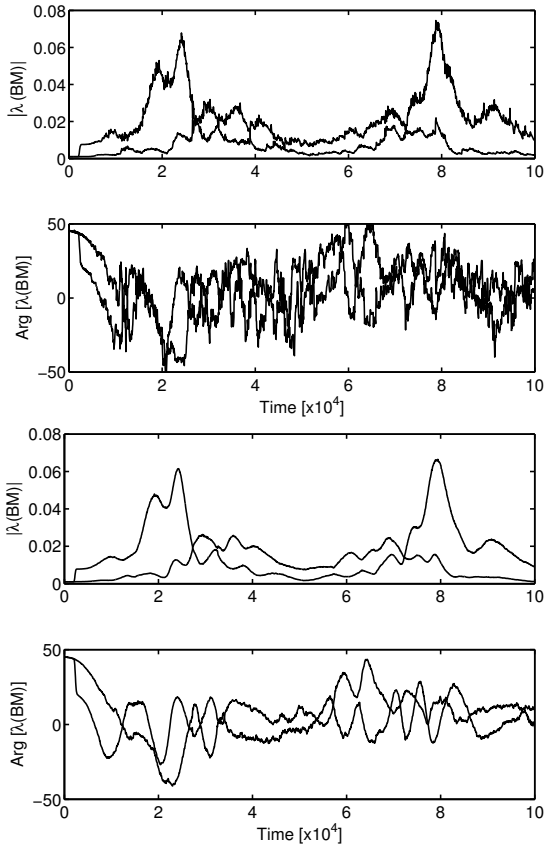

Fig. 7. Evolution of the eigenvalues of the matrix $\mathbf{BM}$, caused by rapid changes in modeling errors. Two upper figures – results of a typical simulation. Two lower figures – average of 50 simulations.

VII. REAL-WORLD EXPERIMENT

A simple real-world experiment was arranged to check the algorithm's performance in a typical indoor acoustic environment. The active noise control system consisted of a signal processing unit (operating with a sampling frequency of 8 kHz), analog amplifiers, two loudspeakers and two microphones placed next to the loudspeakers in an asymmetric configuration (2 m and 1 m away, respectively). The disturbance – a sinusoidal signal with frequency equal to 120 Hz – was generated by a third, PC-controlled loudspeaker.

The applied control procedure was two-step: during the first 0.5 s (4000 samples), the open-loop frequency estimation was carried out using the algorithm described in [30]. After this period, the noise cancellation algorithm was started with $\mathbf{K}_{\mathrm{n}} = \mathbf{I}$ and $\mathbf{M}(0) = 0.05\mathbf{I}$. The following values of adaptation parameters were adopted: $c_\mu = 0.001$, $\rho = 0.9999$.

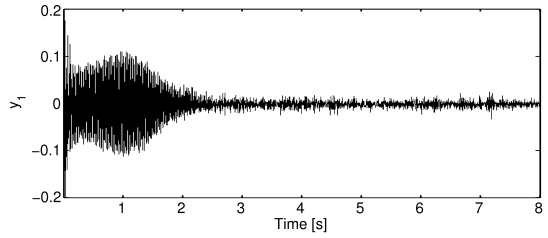

Fig. 8. Transient behavior of the proposed algorithm in a real-world experiment.

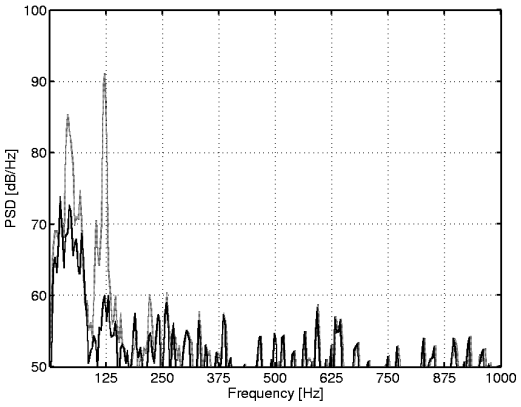

Fig. 9. Power spectral density of the signal recorded by one of the error microphones with (dotted line) and without (solid line) noise cancellation.

Fig. 8 shows the transient behavior of the error signal recorded by the first microphone (the second signal looked very similar). Note that initially the closed-loop system was unstable, but the adaptation loop quickly coped with this problem. Fig. 9 shows the power spectral densities of the signal recorded by the first error microphone before and after cancellation, respectively. The peak at 120 Hz was reduced by 30 dB, almost to the level of the noise floor. Once the initial convergence period was over and SONIC started to operate in the tracking mode, reactions to system changes were pretty quick – changes in the location of error microphones did not cause noticeable changes in the system performance.

VIII. Conclusions

The problem of cancellation of a narrowband disturbance acting at the output of an unknown multivariate linear stable plant was considered. The proposed scheme is an extension of the recently introduced SONIC canceller and consists of two loops. The inner, cancellation loop, predicts and cancels the distrubance. The second, outer loop, adjusts gains of the control loop so as to optimize the canceller's performance. Both theoretical analysis and simulations confirm that, under the Gaussian random-walk-type assumption, the proposed controller converges in mean to the optimal one.

Appendix I
[proof of Corollary 1]

After "squaring" both sides of (8), evaluating the steady-state expectations, and noting that $\mathrm{E}[\tilde{\mathbf{v}}(t)\tilde{\mathbf{v}}^{\mathrm{H}}(t)] = \mathrm{E}[\mathbf{v}(t)\mathbf{v}^{\mathrm{H}}(t)] = \mathbf{V}$, one obtains

$$\mathbf{C}_\infty(\mathbf{M}) = (\mathbf{I}-\mathbf{M})\mathbf{C}_\infty(\mathbf{M})(\mathbf{I}-\mathbf{M})^{\mathrm{H}} + \mathbf{M}\mathbf{V}\mathbf{M}^{\mathrm{H}} + \mathbf{E}\ .$$

Let $\mathbf{\Sigma}(\mathbf{M}) = \mathbf{C}_\infty(\mathbf{M}) - \mathbf{P}_\infty$. It is easy to check that $\mathbf{\Sigma}(\mathbf{M})$ obeys the following discrete Lyapunov equation

$$\mathbf{\Sigma}(\mathbf{M}) = (\mathbf{I}-\mathbf{M})\mathbf{\Sigma}(\mathbf{M})(\mathbf{I}-\mathbf{M})^{\mathrm{H}} + \mathbf{Y}$$

where $\mathbf{Y} = (\mathbf{G}_\infty - \mathbf{M})\mathbf{S}_\infty(\mathbf{G}_\infty - \mathbf{M})^{\mathrm{H}}$ and $\mathbf{S}_\infty = \mathbf{P}_\infty + \mathbf{V} > \mathbf{O}$. Note that $\mathbf{M} \neq \mathbf{G}_\infty$ entails $\mathbf{Y} > \mathbf{O}$. According to Theorem 4.2.1 in [31], one arrives at $\mathbf{\Sigma}(\mathbf{M}) > \mathbf{O}$, i.e., $\mathbf{C}_\infty(\mathbf{M}) > \mathbf{P}_\infty$. The conclusion follows from the fact that $\mathbf{M} \in \mathbf{\Omega}_{\mathrm{s}}$ and $\mathbf{Y} > \mathbf{O}$. The minimum value of $\mathbf{C}_\infty(\mathbf{M})$, equal to $\mathbf{P}_\infty$, is attained when $\mathbf{Y} = \mathbf{O}$, i.e., when $\mathbf{M} = \mathbf{G}_\infty$.

Appendix II
[proof of Theorem 1]

A. Equilibrium Point

Since the properties of the closed-loop system depend on the value of $\mathbf{B}\hat{\mathbf{M}}(t)$, rather than on the value of $\hat{\mathbf{M}}(t)$ alone, a new variable $\mathbf{X} = \mathbf{B}\mathbf{M}$ will be introduced. Multiplying both sides of (36) with $\mathbf{B}$, one arrives at

$$\dot{\mathbf{X}} = \mathbf{F}(\mathbf{X}) \tag{42}$$

where

$$\mathbf{F}(\mathbf{M}) = -\mathbf{X}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0;\mathbf{X})$$

and

$$\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0;\mathbf{X}) = \mathrm{E}\left[\mathbf{y}(t;\mathbf{X})\tilde{\mathbf{z}}^{\mathrm{H}}(t;\mathbf{X})\right].$$

It will be shown that $\mathbf{X}_0 = \mathbf{G}_\infty$ is the unique equilibrium point of the ODE (42), i.e., it obeys

$$\mathbf{F}(\mathbf{X}_0) = \mathbf{O}. \tag{43}$$

Note that

$$\begin{aligned}\mathbf{y}(t;\mathbf{X}) &= e^{j\omega_0}(\mathbf{I}-\mathbf{X})\mathbf{y}(t-1;\mathbf{X}) + \tilde{\mathbf{e}}(t)\\ &\quad + \mathbf{v}(t) - e^{j\omega_0}\mathbf{v}(t-1)\end{aligned} \tag{44}$$

$$\tilde{\mathbf{z}}^{\mathrm{H}}(t;\mathbf{X}) = e^{-j\omega_0}\left[(1-c_\mu)\tilde{\mathbf{z}}^{\mathrm{H}}(t-1;\mathbf{X}) - c_\mu\mathbf{y}^{\mathrm{H}}(t-1;\mathbf{X})\right].$$

Multiplying these equations sidewise, taking the steady-state expectations, and noting that $\mathbf{R}_{\mathbf{yv}}(0;\mathbf{X}) = \mathbf{R}_{\mathbf{vy}}(0;\mathbf{X}) = \mathbf{V}$, one arrives at

$$\begin{aligned}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0;\mathbf{X}) &= (1-c_\mu)(\mathbf{I}-\mathbf{X})\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0;\mathbf{X})\\ &\quad + c_\mu\left[\mathbf{V} - (\mathbf{I}-\mathbf{X})\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X})\right].\end{aligned}$$

The condition (43) is met provided that

$$(\mathbf{I}-\mathbf{X}_0)\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X}_0) = \mathbf{V}. \tag{45}$$

First, it will be shown that $\mathbf{X}_0 = \mathbf{G}_\infty$ fulfills this requirement.

Note that (according to Corollary 1)

$$\mathbf{R}_{\mathbf{yy}}(0;\mathbf{G}_\infty) = \mathbf{C}_\infty(\mathbf{G}_\infty) + \mathbf{V} = \mathbf{P}_\infty + \mathbf{V}.$$

Therefore

$$\begin{aligned}(\mathbf{I}-\mathbf{G}_\infty)\mathbf{R}_{\mathbf{yy}}(0;\mathbf{G}_\infty) &= (\mathbf{I}-\mathbf{G}_\infty)(\mathbf{P}_\infty+\mathbf{V})\\ &= \mathbf{P}_\infty + \mathbf{V} - \mathbf{G}_\infty(\mathbf{P}_\infty+\mathbf{V}) = \mathbf{V}\end{aligned}$$

where the last transition stems from the fact that [cf. (13)] $\mathbf{G}_\infty(\mathbf{P}_\infty + \mathbf{V}) = \mathbf{P}_\infty$.

To prove that the solution $\mathbf{X}_0 = \mathbf{G}_\infty$ is unique, observe that [c.f. (44)]

$$\begin{aligned}
&\mathrm{E}\left[\mathbf{y}(t;\mathbf{X})\right] = \mathbf{O}\\
&\mathbf{R}_{\mathbf{yy}}(1;\mathbf{X}) = (\mathbf{I}-\mathbf{X})\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X}) - \mathbf{V}\\
&\mathbf{R}_{\mathbf{yy}}(\tau;\mathbf{X}) = (\mathbf{I}-\mathbf{X})\mathbf{R}_{\mathbf{yy}}(\tau-1;\mathbf{X}) \quad \forall \tau \geq 2
\end{aligned}$$

It follows that for any solution $\mathbf{X}_0$ of (45)

$$\mathbf{R}_{\mathbf{yy}}(1;\mathbf{X}_0) = \mathbf{R}_{\mathbf{yy}}(2;\mathbf{X}_0) = \cdots = \mathbf{O}\ ,$$

i.e., the sequence $\{\mathbf{y}(t;\mathbf{X}_0)\}$ is zero-mean and white. According to theorem 5.6.1 from [31], under Gaussian assumptions this property implies that the control loop of (37) is optimal. Since there exists exactly one optimal gain matrix $\mathbf{M}_{\mathrm{opt}} = \mathbf{B}^{-1}\mathbf{G}_\infty$ (which corresponds to $\mathbf{X}_{\mathrm{opt}} = \mathbf{B}\mathbf{M}_{\mathrm{opt}} = \mathbf{G}_\infty$), $\mathbf{X}_0 = \mathbf{G}_\infty$ is the unique equilibrium point of the ODE (42).

Finally, since it holds that $\mathbf{X}_0 = \mathbf{B}\mathbf{M}_0$, one gets $\mathbf{M}_0 = \mathbf{B}^{-1}\mathbf{G}_\infty = \mathbf{M}_{\mathrm{opt}}$ as the unique equilibrium point of (36).

B. Local Stability

To check stability of the equilibrium point $\mathbf{X}_0 = \mathbf{G}_\infty$, the ODE (42) will be linearized at $\mathbf{X}_0$

$$\Delta\dot{\mathbf{X}} = d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{F}(\mathbf{X}_0) \tag{46}$$

where $\Delta\mathbf{X} = \mathbf{X} - \mathbf{X}_0$.

To show that the system governed by (42) is locally stable at $\mathbf{X}_0$, the asymptotic stability of the linearized system (46) will be proved. First of all, note that

$$\begin{aligned}
d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{F}(\mathbf{X}_0) &= -d_{\underset{\Delta\mathbf{X}}{\rightarrow}}[\mathbf{X}_0\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0)]\\
&= -\Delta\mathbf{X}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0) - \mathbf{X}_0 d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0)\\
&= -\mathbf{X}_0 d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0)
\end{aligned} \tag{47}$$

where the last transition is a consequence of the fact that (as shown in the preceding subsection)

$$\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0) = \mathbf{O}. \tag{48}$$

Carrying out differentiation of $\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0)$, one obtains

$$\begin{aligned}
d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0) &= (1-c_\mu)(\mathbf{I}-\mathbf{X}_0)d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0)\\
&- (1-c_\mu)\Delta\mathbf{X}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0)\\
&+ c_\mu\Delta\mathbf{X}\mathbf{R}_{\mathbf{yy}}(0,\mathbf{X}_0)\\
&+ c_\mu\mathbf{X}_0 d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{yy}}(0,\mathbf{X}_0).
\end{aligned} \tag{49}$$

After squaring both sides of (44) and evaluating the steady-state expectations, one obtains the following implicit formula for $\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X})$

$$\begin{aligned}
\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X}) &= (\mathbf{I}-\mathbf{X})\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X})(\mathbf{I}-\mathbf{X})^{\mathrm{H}} + \mathbf{E}\\
&+ \mathbf{X}\mathbf{V} + \mathbf{V}\mathbf{X}^{\mathrm{H}}.
\end{aligned} \tag{50}$$

Therefore

$$\begin{aligned}
d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X}_0) &= (\mathbf{I}-\mathbf{X})d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X}_0)(\mathbf{I}-\mathbf{X})^{\mathrm{H}}\\
&+ \Delta\mathbf{X}\left[\mathbf{V} - \mathbf{R}_{\mathbf{yy}}(0;\mathbf{X}_0)(\mathbf{I}-\mathbf{X}_0)^{\mathrm{H}}\right]\\
&+ \left[\mathbf{V} - (\mathbf{I}-\mathbf{X}_0)\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X}_0)\right]\Delta\mathbf{X}^{\mathrm{H}}.
\end{aligned} \tag{51}$$

Since the last two components on the right hand side of (51) are zero [cf. (45)], one arrives at

$$d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X}_0) = \mathbf{O}. \tag{52}$$

Combining (49) with (48) and (52), one obtains

$$\begin{aligned}
d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0) &= (1-c_\mu)(\mathbf{I}-\mathbf{X}_0)d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0)\\
&+ c_\mu\Delta\mathbf{X}\mathbf{R}_{\mathbf{yy}}(0,\mathbf{X}_0)
\end{aligned} \tag{53}$$

which finally results in

$$d_{\underset{\Delta\mathbf{X}}{\rightarrow}}\mathbf{R}_{\mathbf{y}\tilde{\mathbf{z}}}(0,\mathbf{X}_0) = c_\mu\mathbf{D}_0\Delta\mathbf{X}\mathbf{R}_{\mathbf{yy}}(0,\mathbf{X}_0) \tag{54}$$

where

$$\mathbf{D}_0 = \left[\mathbf{I} - (1-c_\mu)(\mathbf{I}-\mathbf{X}_0)\right]^{-1}. \tag{55}$$

After combining (46), (47) and (54), the linearized ODE can be written down in the form

$$\Delta\dot{\mathbf{X}} = -c_\mu\mathbf{X}_0\mathbf{D}_0\Delta\mathbf{X}\mathbf{R}_{\mathbf{yy}}(0,\mathbf{X}_0). \tag{56}$$

To show asymptotic stability of the system governed by (56), the Lyapunov approach will be used. Consider the following Lyapunov function

$$V(\Delta\mathbf{X}) = \mathrm{tr}\left\{\Delta\mathbf{X}^{\mathrm{H}}\mathbf{P}_0\Delta\mathbf{X}\right\} \tag{57}$$

where $\mathbf{P}_0$ is a positive definite matrix yet to be defined. Note that $V(\Delta\mathbf{X}) \geq 0$, and that the equality holds iff $\Delta\mathbf{X} = \mathbf{O}$.

Using (56), one obtains

$$\begin{aligned}
\dot{V}(\Delta\mathbf{X}) &= \mathrm{tr}\left\{\Delta\dot{\mathbf{X}}^{\mathrm{H}}\ \mathbf{P}_0\Delta\mathbf{X} + \Delta\mathbf{X}^{\mathrm{H}}\mathbf{P}_0\ \Delta\dot{\mathbf{X}}\right\}\\
&= \mathrm{tr}\left\{(\mathbf{P}_0\mathbf{A}_0 + \mathbf{A}_0^{\mathrm{H}}\mathbf{P}_0)\mathbf{B}_0\right\}
\end{aligned} \tag{58}$$

where

$$\mathbf{A}_0 = -c_\mu\mathbf{X}_0\mathbf{D}_0, \quad \mathbf{B}_0 = \Delta\mathbf{X}\mathbf{R}_{\mathbf{yy}}(0;\mathbf{X}_0)\Delta\mathbf{X}^{\mathrm{H}}.$$

To proceed further, the following result will be needed

Corollary 2

For any value c_μ, such that $0 < c_\mu < 1$, the matrix $\mathbf{A}_0$ is stable, i.e., it has eigenvalues with strictly negative real parts

$$\mathrm{Re}\left\{\lambda_i(\mathbf{A}_0)\right\} < 0, \quad i = 1,\ldots,n. \tag{59}$$

Proof

The proof will start from rewriting $\mathbf{X}_0$ in the form $\mathbf{X}_0 = \mathbf{Q}^{-1}\mathbf{\Lambda}_0\mathbf{Q}$, where

$$\mathbf{\Lambda}_0 = \mathrm{diag}\left\{\lambda_i(\mathbf{X}_0), i = 1,\ldots,n\right\}$$

is a diagonal matrix made up of the eigenvalues of $\mathbf{X}_0$.

Similarly, $\mathbf{D}_0$ can be expressed in the form $\mathbf{D}_0 = \mathbf{Q}^{-1}\mathbf{\Sigma}_0\mathbf{Q}$, where

$$\mathbf{\Sigma}_0 = \mathrm{diag}\left\{\frac{1}{c_\mu + (1-c_\mu)\lambda_i(\mathbf{X}_0)}\ , i = 1,\ldots,n\right\}.$$

Combining both results, one obtains

$$\lambda_i(\mathbf{X}_0\mathbf{D}_0) = \frac{\lambda_i(\mathbf{X}_0)}{c_\mu + (1-c_\mu)\lambda_i(\mathbf{X}_0)}\ , \quad i = 1,\ldots,n.$$

Since $\mathbf{X}_0 \in \mathbf{\Omega}_{\mathrm{s}}$, the eigenvalues of $\mathbf{X}_0$ must obey $\mathrm{Re}\left\{\lambda_i(\mathbf{X}_0)\right\} > 0, i = 1,\ldots,n$. Using this property, it is

straightforward to show that for any $0 < c_\mu < 1$ it holds that $\mathrm{Re}\,\{\lambda_i(\mathbf{A}_0)\} < 0, i = 1, \ldots, n$.

■

Using the well-known properties of a continuous Lyapunov equation [32], one can conclude that there exists (exactly one) matrix $\mathbf{P}_0 > \mathbf{O}$ obeying

$$\mathbf{P}_0\mathbf{A}_0 + \mathbf{A}_0^{\mathrm{H}}\mathbf{P}_0 + \mathbf{I} = \mathbf{O}. \tag{60}$$

The conclusion follows from the fact that the matrix $\mathbf{A}_0$ is stable and the matrix $\mathbf{I}$ is positive definite.

Combining (58) and (60), one arrives at

$$\dot{V}(\Delta\mathbf{X}) = -\mathrm{tr}\{\mathbf{B}_0\} \leq 0 \tag{61}$$

which stems from the fact that the matrix $\mathbf{R}_{\mathbf{yy}}(0; \mathbf{X}_0)$ is positive definite. Moreover, equality in (61) holds iff $\Delta\mathbf{X} = \mathbf{O}$. This proves the asymptotic stability of the system governed by (46).

REFERENCES

[1] S. Elliott, P. Nelson, I. Stothers, and C. Boucher, "In-flight experiments on the active control of propeller-induced cabin noise," *J. Sound and Vibration*, vol. 140, pp. 219–238, 1990.

[2] S. Johansson, T. Lago, S. Nordebo, and I. Claesson, "Control approaches for active noise control of propeller-induced cabin noise evaluated from data from a dornier 328 aircraft," in *Proc. 6th Int. Congress on Sound and Vibration*, Denmark, 1999, pp. 1611–1618.

[3] O. Pabst, T. Kletschkowski, and D. Sachau, "Active noise control in light jet aircraft," in *Proc. Acoustics'08*, Paris, 2008, pp. 735–5740.

[4] S. Elliott, I. Stothers, P. Nelson, A. McDonald, and D. Quinn, "The active control of engine noise inside cars," in *Proc. Inter-noise*, 1988, pp. 987–990.

[5] C. Bohn, A. Cortabarria, V. Härtel, and K. Kowalczyk, "Active control of engine-induced vibrations in automotive vehicles using disturbance observer gain scheduling," *Control Eng. Practice*, vol. 12, pp. 1029–1039, 2004.

[6] R. S. T. Knurek, "Automotive applications of a hybrid active noise and vibration control," *IEEE Control Systems Magazine*, vol. 16, pp. 72–78, 1996.

[7] S. Elliot, *Signal Processing for Active Control*. Academic Press, 2001.

[8] S. Kuo and D. Morgan, *Active Noise Control Systems*. New York: Wiley, 1996.

[9] M. Bouchard and S. Quednau, "Multichannel recursive least-squares algorithms for active noise control and sound reproduction systems," *IEEE Trans. Speech Audio Processing*, vol. 8, pp. 606–618, 2000.

[10] M. Bouchard, "Multichannel affine projection algorithms for active noise control and acoustic equalization systems," *IEEE Trans. Speech Audio Processing*, vol. 11, pp. 54–60, 2003.

[11] M. Diego, A. Gonzalez, M. Ferrer, and G. Piñero, "An adaptive algorithms comparison for real multichannel active noise control," in *Proc. 12th European Signal Processing Conference*, Vienna, Austria, pp. 925–928, 2004.

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

[13] B. Wu and M. Bodson, "Direct adaptive cancellation of periodic disturbances for multivariable plants," *IEEE Trans. Speech Audio Processing*, vol. 11, pp. 538–548, 2003.

[14] M. Niedźwiecki and M. Meller, "A new approach to active noise and vibration control – Part I: the known frequency case," *IEEE Trans. Signal Processing*, vol. 57, pp. 3373–3386, 2009.

[15] ——, "A new approach to active noise and vibration control – Part II: the unknown frequency case," *IEEE Trans. Signal Processing*, vol. 57, pp. 3387–3398, 2009.

[16] ——, "Self-optimizing adaptive vibration controller," *IEEE Trans. Automat. Contr.*, vol. 54, pp. 2087–2099, 2009.

[17] K. J. Aström, U. Borisson, L. Ljung and B. Wittenmark, "Theory and application of self-tuning regulators," *Automatica*, vol. 13, pp. 457-476, 1977.

[18] D. W. Clarke, "Implementation of self-tuning controllers," In C. J. Harris and S. A. Billings (Eds.), "Self-tuning and Adaptive Control: Theory and Applications," IEE Control Engng Series 15, 1981.

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

[20] L. Guo and L. Ljung, "Performance analysis of general tracking algorithms," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1388–1402, 1995.

[21] M. Niedźwiecki, *Identification of Time-varying Processes*. New York: Wiley, 2001.

[22] J. Dattorro, *Convex Optimization & Euclidean Distance Geometry*. $\mathcal{M}\epsilon\beta oo$ Publishing, 2005.

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

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

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

[26] M. Niedźwiecki and M. Meller, "SONIC – self-optimizing narrowband interference canceler: comparison of two frequency tracking strategies," in *Proc. 8th IEEE Int. Conf. on Control and Automation*, Xiamen, China, 2010, pp. 1892–1896.

[27] ——, "Multifrequency self-optimizing narrowband interference canceller," in *Proc. 18th International Workshop on Acoustic Echo and Noise Control*, Tel Aviv, Israel, pp. 1–4, 2010.

[28] M. Niedźwiecki and P. Kaczmarek, "Identification of quasi-periodically varying systems using the combined nonparametric/parametric approach," *IEEE Trans. Signal Process.*, vol. 53, pp. 4588–4598, 2005.

[29] L. Ljung and T. Glad, *Modeling of Dynamic Systems*. Englewood Cliffs NJ: Prentice Hall, 1993.

[30] M. Meller and M. Niedźwiecki, "An improved frequency estimator for an adaptive active noise control scheme," in *Proc. 18th European Signal Processing Conference*, Aalborg, Denmark, pp. 353-357, 2010.

[31] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs NJ: Prentice Hall, 1979.

[32] G. Golub and C. V. Loan, *Matrix Computations*. Baltimore MD: Johns Hopkins Univ. Press, 1996.