Multichannel self-optimizing narrowband interference canceller

Michał Meller *, Maciej Niedźwiecki

Department of Automatic Control, Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Computer Science, Narutowicza 11/12, 80-233 Gdańsk, Poland

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. The algorithm, derived and analyzed assuming a single-tone, complex-valued disturbance with known frequency, can be extended to cope with a range of realistic applications, such as real-valued disturbances, multitone signals, and unknown frequency.

© 2013 Elsevier B.V. All rights reserved.

1. 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. Vaguely speaking, the idea is to generate acoustic waveform that, in the area/point of interest, has the same shape as the disturbance waveform, but opposite polarity [1]. 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. Commercial applications of such systems include reduction of propeller-induced interior noise in aircrafts [2,3], active noise [4,5] and vibration control systems for cars [6], active noise-cancelling mufflers [7], and active noise barriers [8], among many others.

Most of the existing multichannel solutions are based on the classical filtered-X least mean squares (FXLMS) approach [1,9] or its modifications obtained by replacing the LMS adaptive filters with faster converging ones, such as recursive least squares (RLS) [10] or affine projection (AP) [11] – for comparison of different variants see e.g. [12]. 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. A special random perturbation technique is often used [13] for this purpose. Unfortunately, auxiliary noise disturbs operation of the ANC system, which results in performance degradation.

Another conceptually different solution to the multivariable narrowband disturbance rejection problem, based on the “phase-locked” loop structure, was presented in [14]. 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 cancelling was recently proposed in [15,16] (for complex-valued disturbances) and in [17] (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 [18–20], 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.

The main contribution of this study is development and analysis of a multivariate version of SONIC controller. Similar to the univariate case, the extended controller can cope with plant modelling errors and locally converges in mean to the optimal one. Several useful extensions, e.g., frequency adaptive version of the controller, are proposed in the paper. Finally, good behavior of the proposed algorithm is confirmed with simulations and practical experiments.

2. Problem statement

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

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

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

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

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-dimensional wideband noise, and $\mathbf{L}_p(q^{-1}) = [L_{pq}(q^{-1})]$, $q = 1, 2, \ldots, n$, $l = 1, 2, \ldots, n$ 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.

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 as minimum-variance (MV) regulators. It will not be assumed that a reference signal, correlated with the disturbance, is available. For this reason the designed system will be a purely feedback canceller, not incorporating any feedforward compensation loop.

3. Control of a known plant

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

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

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.

3.1. Stabilizing controller

Suppose that the controlled plant is time-invariant, and that its transfer function $\mathbf{L}_p(q^{-1})$ is known. Since the disturbance is a narrowband signal, so must be the cancellation signal, with angular frequency $\omega_0$ and complex amplitude chosen so as to enable destructive interference at the plant’s output. In a case like this Eq. (2) can be approximately written down in the form:

$$\mathbf{y}(t) \approx \mathbf{K}_p\mathbf{u}(t-1) + \mathbf{d}(t) + \mathbf{v}(t)$$

where

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

denotes the nonsingular matrix of plant gains at the frequency $\omega_0$.

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

$$\mathbf{u}(t) = -\mathbf{K}_p^{-1}\mathbf{d}(t+1).$$

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:

$$\mathbf{\hat{d}}(t+1)|t = e^{j\omega_0}[\mathbf{\hat{d}}(t)|t-1] + \mathbf{M}\mathbf{v}(t)$$

$$\mathbf{u}(t) = -\mathbf{K}_p^{-1}\mathbf{\hat{d}}(t+1)|t$$

where $\mathbf{M} = [\mu_{ik}], \quad k = 1, 2, \ldots, n, \quad l = 1, 2, \ldots, n$ denotes a matrix of complex-valued adaptation gains, chosen so as to guarantee stability of the closed loop.

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

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

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

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

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

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

$$\mathbf{c}(t) = e^{j\omega_0}(1-\mathbf{M})\mathbf{c}(t-1) - \mathbf{M}\mathbf{\hat{v}}(t-1) + \mathbf{\hat{e}}(t)$$

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

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 $\mathbf{M}$ belongs to the set of
stabilizing gains:
\[ M \in \Omega_s : |\lambda_i(I-M)| < 1, \quad i = 1, \ldots, n \]
where \( \lambda_i(\cdot) \) denotes the ith eigenvalue of the respective matrix.

3.2. 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 amplitude changes \( e(t) \) form a white noise sequence. This will serve two purposes. First, the tracking analysis under RW-type variations is an important estimation and control benchmark \[21,22\]. 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) \( \{v(t)\} \) is a zero-mean circular white sequence with covariance matrix \( V \).

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

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 \( d(t) \) obeying (7), based on noisy measurements:
\[ s(t) = y(t)|_{\nu(t)} = 0 = d(t) + v(t). \]
(9)

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

Regarding (7) as a state equation of a dynamic system, and (9) as its output (measurement) equation, prediction of \( d(t) \) can be viewed as an estimation problem in the state space. The optimal, in the mean-squared sense, one-step-ahead predictor of \( d(t) \) has the well-known form:
\[ d(t+1|t) = E[d(t+1)|s(t)] \]
where \( s(t) = \{s|i, i \leq t\} \) denotes the observation history available at instant \( t \). Under Gaussian assumptions,

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

the conditional mean estimate (10) can be computed using the Kalman filter:
\[
\begin{align*}
\epsilon(t) &= s(t) - \hat{d}(t|t-1) \\
G(t) &= P(t|t-1)P(t|t-1) + V^{-1} \\
\hat{d}(t+1|t) &= e^{j\omega_0}[\hat{d}(t|t-1) + G(t)e(t)] \\
P(t+1|t) &= [I-G(t)]P(t|t-1) + E \\
\end{align*}
\]
(11)
where \( G(t) \) and \( P(t|t-1) \) 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:
\[
\begin{align*}
\epsilon(t) &= s(t) - \hat{d}(t|t-1) \\
\hat{d}(t+1|t) &= e^{j\omega_0}[\hat{d}(t|t-1) + G_{\infty}\epsilon(t)] \quad (12)
\end{align*}
\]
where \( G_{\infty} \in \Omega_s \) denotes the steady-state gain matrix:
\[
G_{\infty} = \lim_{t \to \infty} G(t) = P_{\infty}[P_{\infty} + V]^{-1} \quad (13)
\]
and \( P_{\infty} = \lim_{t \to \infty} P(t) \) is the positive definite solution of the Riccati equation:
\[
P_{\infty}[P_{\infty} + V]^{-1}P_{\infty} - E = 0 \quad (14)
\]
where \( O \) denotes the zero matrix (in the sequel \( A > O \) will mean that the matrix \( A \) 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 \( M \in \Omega_s \), denote by \( C_s(M) = \lim_{t \to \infty} E[c(t)e^0(t)] \) the covariance matrix of the corresponding steady-state cancellation error. It will be shown that the best performance of the closed-loop cancelling system can be obtained by setting \( M = G_{\infty} \).

**Corollary 1** (Proof – see Appendix A). Under assumptions (A1)–(A3) it holds that \( \inf_{M \in \Omega_s} C_s(M) = P_{\infty} \). The minimum is obtained for \( M = G_{\infty} \).

4. Control of an unknown plant

So far it has been assumed that the “true” gain \( K_p \) of the plant is known. Suppose now that the “idealized” control rule (4) is replaced with
\[
u(t) = -K_n^{-1}d(t+1|t) \]
(15)
where \( K_n, \text{det}(K_n) \neq 0 \), is the nominal (assumed) plant gain at the frequency \( \omega_0 \), generally different from the true gain \( K_p \). Denote by \( B = K_pK_n^{-1} \) the matrix of complex-valued modelling errors. Combining (3) with (5) and (15), one arrives at the following generalized version of (8):
\[
y(t) = c(t) + v(t) \]
(16)
\[
c(t) = e^{j\omega_0}(I - BM)c(t-1) - BMv(t-1) + \epsilon(t) \]
(17)
where
\[
c(t) = d(t) - B\hat{d}(t|t-1) \]
denotes cancellation error.

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

4.1. Self-optimizing controller

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

$$R_{yy}(0; M) \equiv E[y(t; M)y^H(t; M)]$$

where $y(t; M)$ denotes a stationary process that “settles down” in the closed-loop system for a constant value of $M$ such that $BM \in \Omega_r$.

Since it holds that [cf. (16)] $R_{yy}(0; M) = C_{\infty}(M) + V$, minimization of $R_{yy}(0; M)$ 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:

$$\dot{\hat{M}}(t) = \hat{M}(t - 1) + \alpha \Delta \hat{M}(t)$$

(18)

where $\alpha$ denotes a small positive constant, and $\Delta \hat{M}(t)$ is chosen so as to guarantee that

$$R_{yy}(0; \hat{M} + \Delta \hat{M}) \leq R_{yy}(0; \hat{M}).$$

(19)

While derivation of the gain tuning algorithm given in [15] was based on Wirtinger calculus, in the present study another analytic technique, known as directional derivatives, will be used [23].

Directional derivatives are related to Gâteaux differentials. Let $X, Y \in C^{p \times n}$ be complex-valued matrices. The first-order directional derivative of a function $H(X) : C^{m \times n} \to C^{k \times l}$ at $X$ in the direction $Y$ is defined as

$$d_Y H(X) = \frac{d}{d\epsilon} \bigg|_{\epsilon = 0} H(X + \epsilon Y) \in C^{k \times l}$$

(assuming that $H$ is differentiable with respect to $\epsilon$).

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, which considerably complicates analysis.

4.1.1. Gradient update

To find the “right” direction $\Delta \hat{M}$, the first-order Taylor series expansion of $R_{yy}(0; M)$ about $M$ will be used:

$$R_{yy}(0; M + \Delta M) \approx R_{yy}(0; M) + E[d_{\Delta M^H} y(t; M)y^H(t; M)].$$

(20)

Note that

$$d_{\Delta M} y(t; M)y^H(t; M) = d_{\Delta M} y(t; M)y^H(t; M) + y(t; M)\left[d_{\Delta M^H} y(t; M)\right]^H.$$  

(21)

Furthermore, since

$$y(t; M) = e^{i\omega t} (1 - BM)y(t - 1; M) + \hat{e}(t) + \nu(t) - e^{i\omega t} v(t - 1)$$

(22)

one obtains

$$d_{\Delta M} y(t; M) = e^{i\omega t} (1 - BM)d_{\Delta M} y(t - 1; M) - B\Delta My(t - 1; M).$$

(23)

Since the directional derivative of $y(t)$ depends on the modeling error $B$, the recursive formula (23) cannot be used without modification. In the univariate case this problem was dealt with by means of applying the substitution $\beta = c_\nu/\mu$, where $\beta$ and $\mu$ denote the one-dimensional counterparts of $B$ and $M$, respectively, and $c_\nu$ is a positive constant [15]. In the multivariate case the same “trick” will be used by postulating that (see remark below) $B = c_\nu M^{-1}$. Using this substitution, one obtains the modified version of (23):

$$d_{\Delta M} y(t; M) = e^{i\omega t} (1 - c_\nu)\left[d_{\Delta M} y(t - 1; M) - c_\nu M^{-1} \Delta My(t - 1; M)\right].$$

(24)

Let $z(t; M)$ be the quantity defined implicitly by

$$d_{\Delta M} y(t; M) = M^{-1} \Delta M z(t; M).$$

(25)

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

$$d_{\Delta M} y(t; M)y^H(t; M) \leq O.$$  

(26)

After combining (21) with (25), one obtains

$$d_{\Delta M} y(t; M)y^H(t; M) = M^{-1} \Delta M z(t; M)y^H(t; M) + y(t; M)\left[d_{\Delta M^H} y(t; M)\right]^H.$$  

(27)

Therefore, in order to fulfill (26), one should set

$$\Delta M = -My(t)z^H(t)P$$

(28)

where $P$ is an arbitrary positive definite matrix.

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

$$\dot{z}(t) = e^{i\omega t} [(1 - c_\nu)\dot{z}(t - 1) - c_\nu y(t - 1)]$$

$$\dot{M}(t) = \hat{M}(t - 1)[I - \alpha y(t)z^H(t)]$$

(29)

where the first recursion was obtained by rewriting (24) in terms of $\dot{z}(t)$.

Remark. Substitution $B = c_\nu M^{-1}$ 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 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 [18–20].

The proposed substitution is hardly intuitive and there-fore its feasibility cannot be judged a priori. But our main point here is that any way of coping with unknown
modelling errors is feasible, provided that it guarantees
mean convergence of gain estimates to their optimal
values. In [15] it was shown that this is the case for the
univariate version of SONIC. In the sequel we will prove
that the same holds true for its multivariate version

4.1.2. Normalized updates

In addition to (29), two normalized versions of the gain
update can be used: the trace algorithm:

\[ \hat{r}(t) = r\hat{r}(t-1) + |\hat{z}(t)|^2 \]
\[ \tilde{M}(t) = \tilde{M}(t-1) \left[ 1 - \frac{y(t)\hat{z}^H(t)}{\hat{r}(t)} \right] \]

and, the computationally more involved, directional algo-
rithm [obtained by taking \( P = R^{-1} \) in (28)]:

\[ \hat{R}(t) = r\hat{R}(t-1) + \hat{z}(t)\hat{z}^H(t) \]
\[ \tilde{M}(t) = \tilde{M}(t-1) \left[ 1 - y(t)\hat{z}^H(t)\hat{R}^{-1}(t) \right]. \]

The adjective “trace” refers to the fact that the normalizing
factor \( \hat{r}(t) \) in (30) is equal to \( \text{tr}(\hat{R}(t)) \), where \( \hat{R}(t) \) denotes
the normalizing matrix in (31). In both cases \( \rho, 0 < \rho < 1, \)
denotes the forgetting constant determining the effective
length of the local averaging window.

Normalization makes the algorithms (30) and (31) scale
invariant. Suppose that the gradient algorithm (29) is run
on a scaled data sequence \( y'(t), y'(t) = \delta y(t), \delta \neq 0 \), Then,
to obtain results identical with the original ones, the
stepsize \( \alpha \) should be replaced with \( \alpha/\delta^2 \). Both normalized
algorithms are free of this ambiguity. Directional normali-
zation provides more careful scaling than trace normali-
zation as it takes into account the relative “strength” of
different measurements.

4.2. Mean convergence analysis

For simplicity, consider gradient updates (29). It will
be shown that under (A1) and (A2) the estimates \( \tilde{M}(t) \)
converge (locally) in mean to the optimal solution \( \tilde{M}_{\text{opt}} = B^{-1}C_\infty \).

It is known that the tracking behavior of constant-gain
(finite-memory) estimation algorithms, such as (33), can
be studied by examining the properties of the associated
ordinary differential equations (ODEs) [24,25]. Denote by
\( \{y(t), \tilde{M}(t)\} \) and \( \{\hat{z}(t), \tilde{M}(t)\} \) the stationary processes observed in
the closed-loop system for a constant value of \( \tilde{M} : BM \in \mathcal{U}. \)
For sufficiently small values of \( \alpha \), the estimates \( \tilde{M}(t) \)
wander around \( \tilde{M}_0 \) – the stable equilibrium point of ODE
associated with (29):

\[ \tilde{M} = F(\tilde{M}) \]

(32)

where

\[ F(\tilde{M}) = -\tilde{M}E[y(t; \tilde{M})\hat{z}^H(t; \tilde{M})]. \]

It can be shown that

\textbf{Theorem 1} (Proof – see Appendix B). Under assumptions
(A1)-(A2) it holds that \( \tilde{M}_0 = \tilde{M}_{\text{opt}} \) is a unique stable
equilibrium point of ODE (32). The minimum is attained for
\( \tilde{M} = G_\infty \).

The same result can be proved for normalized updates
(30) and (31).

4.3. Multivariate SONIC

Combining all earlier results, the trace (recommended)
version of the multivariate SONIC algorithm can be sum-
marized as follows:

\[ \begin{align*}
\dot{z}(t) &= e^{in_0}[(1-c_p)\hat{z}(t-1)-c_py(t-1)] \\
\dot{r}(t) &= r\hat{r}(t-1) + |\hat{z}(t)|^2 \\
\dot{M}(t) &= \dot{M}(t-1) \left[ 1 - y(t)\hat{z}^H(t) \right] \\
\dot{\tilde{d}}(t+1) &= e^{in_0}\dot{\tilde{d}}(t-1) + \dot{M}(t)y(t) \\
\tilde{u}(t) &= -K^{-1}\dot{\tilde{d}}(t+1). 
\end{align*} \]

(33)

SONIC consists of two loops: the inner, cancellation loop [the
last two recursions of (33)], which predicts and cancels
disturbance, and the outer, self-optimization loop [the first
two recursions of (33)], 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 deductible from
the single input–single output solution presented in [15].

4.4. Selection of design parameters

The influence of design parameters \( c_p \) and \( \rho \) on tracking
properties of SONIC was studied in [15] for cisoids with
random amplitude drift. It was shown there (see Section VI-
E in [15]) that the closer that \( 1-\rho \) becomes to 0, the longer it
takes for the algorithm to readjust the adaptation gain \( \mu(t) \)
when the operating conditions change (plant, noise and/or
interference). On the other hand, small values of \( (1-\rho)/c_p \)
are required to make the steady-state fluctuations of \( \mu(t) \) around
\( \mu_{\text{opt}} \) small. Finally, the value of \( c_p, 0 < c_p < 1, \) should be
trimmed to the rate of plant/disturbance variation – the faster
the changes, the larger the recommended values of \( c_p \).
The following rules of thumb, which extend to the multivariable
case, seem to work pretty well in practice: (1) Choose \( c_p \) in
the range \([0.005/f_s, 0.1/f_s]\), where \( f_s \) denotes sampling frequency
expressed in kHz. (2) Choose \( \rho \) such that \( 1-\rho < c_p \). Under
1 kHz sampling our default choice is \( c_p = 0.05 \) and \( \rho = 0.999 \).

5. Comments and extensions

5.1. Real systems

Any real-valued narrowband disturbance \( \tilde{d}(t) \) can be
be treated as a sum of two cisoids with frequencies \( +\omega_0 \)
and \( -\omega_0 \):

\[ \tilde{d}(t) = \tilde{d}_+(t) + \tilde{d}_-(t), \quad \tilde{d}_+(t) = \tilde{d}_-^*(t). \]

(34)

It follows that a parallel structure, consisting of two
separate blocks of the form (33) tracking each signal
component [26], may be employed in the real-valued case. It is straightforward to verify that all quantities computed by the subcontroller tracking the negative frequency are conjugates of their respective counterparts computed for the positive frequency. This leads to the conclusion that a “quick and dirty” extension to real-valued signals may be obtained by replacing the last equation in (33) with

\[
\mathbf{u}(t) = -\text{Re}[\mathbf{K}_n^{-1} \mathbf{d}(t+1|t)]
\]

\[
= -\frac{1}{2} \mathbf{K}_n^{-1} \mathbf{d}(t(t-1) - \frac{1}{2} (\mathbf{K}_n^*)^{-1} \mathbf{d}^*(t(t-1)).
\]

To evaluate how the proposed modification affects the behavior of the closed loop system observe that [c.f. (3), (34), (35)]

\[
\mathbf{y}(t) \approx -\frac{1}{2} \mathbf{K}_p \mathbf{K}_n^{-1} \mathbf{d}(t(t-1) + \mathbf{d}^*(t)
\]

\[
-\frac{1}{2} \mathbf{K}_p^*(\mathbf{K}_n^*)^{-1} \mathbf{d}^*(t(t-1) + \mathbf{d}^*(t) + \mathbf{v}(t).
\]

Since the proposed controller is narrowband, it can be argued that the negative frequency components will not affect the controller, tuned to track the positive frequency component, significantly. On the other hand, the appearance of the factor 1/2 in the term \(\mathbf{K}_p \mathbf{K}_n^{-1}/2\) effectively makes the modelling errors half of their value in the complex valued case. Note however, that these errors will become compensated by the self-optimization loop of the controller.

5.2. Unknown and time-varying frequency

When frequency of the disturbance is unknown and possibly time varying, the scheme presented in the previous section should be equipped with a frequency estimation mechanism. Suppose that a reference signal \(\mathfrak{r}(t)\) is available

\[
\mathfrak{r}(t) = \hat{d}_i(t)+\mathfrak{v}_i(t)
\]

where \(\hat{d}_i(t)\) denotes the narrowband signal measured at a location that is close to the source of disturbance, and \(\mathfrak{v}_i(t)\) denotes measurement noise independent of \(\mathfrak{v}(t)\). Then one can replace constant frequency \(\omega_0\) in (33) with their estimates, evaluated using the reference signal \(\mathfrak{r}(t)\),

\[
\hat{\omega}_i(t)|_{t-1} = \mathcal{R}[\mathfrak{r}(t)] \text{ where } \mathcal{R}(\mathfrak{r}(t)|_{t|t}) \text{ is the instantaneous frequency estimates may be computed recursively using the following adaptive notch filter:}

\[
\mathfrak{e}_i(t) = \mathfrak{r}(t) - \hat{d}_i(t)|_{t-1}
\]

\[
\hat{d}_i(t+1) = e^{\mathfrak{e}_i(t)|_{t-1}}[\hat{d}_i(t)|_{t-1} + \mu_i \mathfrak{e}_i(t)]
\]

\[
g_i(t) = \frac{\text{Im}[\hat{d}_i^*(t)|_{t-1} + \mu_i \mathfrak{e}_i(t)]}{|\hat{d}_i(t)|_{t-1}}^2
\]

\[
\hat{\omega}_i(t+1) = \hat{\omega}_i(t)|_{t-1} + \gamma_i^\alpha g_i(t)
\]

\[
\hat{\omega}_i(t+1) = \hat{\omega}_i(t)|_{t-1} + \gamma_i^\beta g_i(t).
\]

The quantities \(\mu_i, \gamma_i^\alpha \) and \(\gamma_i^\beta \) (all real-valued) denote small adaptation gains controlling the speed of amplitude, frequency, and frequency rate adaptation, respectively.

5.3. 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 \(\mathbf{M}\) to the optimal value \(\mathbf{M}_0 = \mathbf{M}_{\text{opt}}\), the initial estimate \(\mathbf{M}(0)\) must belong to the domain of attraction of \(\mathbf{M}_0\), further denoted by \(\Omega_0\). Vaguely speaking, the domain \(\Omega_0 \subset \Omega\), “shrinks” as the modelling error grows. Hence, to minimize the risk of divergence during initialization of the algorithm, one can set \(\mathbf{K}_n\) to \(\mathbf{K}_p\), 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 \(i\)th column of the matrix \(\mathbf{K}_n\) using classical frequency-domain identification tools, such as the method of empirical transfer function estimation (ETFE) described in [28]. According to (8), when \(\mathbf{K}_n = \mathbf{K}_p\), i.e., \(\mathbf{B} = \mathbf{I}\), stability of the closed-loop system can be guaranteed by setting \(\mathbf{M} = \mathbf{I}\), where \(\mu \in (0, 2)\). Following this observation, we recommend using \(\mathbf{M}(0) = \mu_0 \mathbf{I}\), 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 (\(\mathbf{K}_p\)) 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 \(\mathbf{M}(t)\) – see [15] for more details.

5.4. Comparison with adaptive FX-LMS

All advantages of SONIC compared to adaptive FXLMS (i.e., FXLMS with on-line plant identification), pointed out in [15] and [16], carry on to the multivariate case. For the readers’ convenience they will be 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 \(M\):

\[
L_{i_1i_2}(q^{-1}) = \sum_{m=1}^{M} p_{i_2i_1} q^{-m}, \quad 1 \leq i_1, i_2 \leq n.
\]

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 cancelling 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 \(\omega_0 \gamma_i^h \approx \sum_{m=1}^{M} p_{i_2i_1} e^{-jm\omega_0}\). This means that, in the case considered,
the FIR representation adopted for FXLMS is grossly over-parameterized – it requires estimation of $n^2M$ 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 [13]. Unfortunately, injection of such an auxiliary noise disturbs operation of the cancelling system and causes deterioration of its performance. Note that SONIC is free of this drawback – due to gain fixing, $M(t)$ converges in mean to its optimal value without additional excitation.

3. **Self-optimization**: SONIC automatically adjusts its gain matrix $M$ 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. [27] and the references therein), we have found out that in the presence of sinusoidal disturbances they do not perform satisfactorily [15].

6. Simulation results

6.1. 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

\[ y(t) = L_p(q^{-1})u(t-1) + d(t) + v(t) \]

with impulse response depicted in Fig. 1 (established experimentally using data from a real acoustic encounter), and for its static counterpart

\[ y(t) = L_p(e^{-j\omega_0})u(t-1) + d(t) + v(t). \]

In both cases

\[ V = 0.03 \begin{bmatrix} 1 & 0.3j \\ -0.3j & 1 \end{bmatrix}. \] (39)

2. For four choices of amplitude modulation intensity:

\[ E = \sigma_e^2 \begin{bmatrix} 1 & 0.5 \\ 0.5 & 2 \end{bmatrix} \]

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

The trace version of the algorithm was used with $\rho = 0.9998$ and $c_0 = 0.005$. While the assumed plant model was equal to the true one, $K_n = K_p$, the algorithm was initialized with

\[ \tilde{M}(0) = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix} e^{j\pi/4}. \] (40)

This means that, initially, a significant phase bias was present in $M(t)$.

Due to the presence of a nonstationary disturbance (note that magnitudes of $d(t)$ grow unbounded), all results
were obtained by combining ensemble averaging (50 realizations of \((e(t))\) and \((w(t))\)) and time averaging (100,000 samples, initial 50,000 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 the shows mean values of \(\lambda_1(\mathbf{M})\) and \(\lambda_2(\mathbf{M})\). 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.

### 6.2. Transient behavior

Transient behavior of the proposed scheme for the dynamic plant is depicted in Fig. 2. This time, the trace algorithm was used with \(\rho = 0.999\), \(c_\rho = 0.005\), \(\mathbf{M}(0)\) was given by (40), and \(\sigma_w^2 = 10^{-5}\). Additionally, to avoid rough start, during the initial 2000 sampling periods, the quantity \(\tilde{z}(t)\) was evaluated, but \(\mathbf{M}(t)\) 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 that to magnitude errors. Similar difference in sensitivity to different kinds of errors was observed for the univariate SONIC.

### 6.3. Comparison with FXLMS

Suppose that the transfer function of the plant \(L_p(\omega^{-1})\) does not change with time, and that the true gain matrix \(\mathbf{K}_p = L_p(e^{-j\omega_0})\) 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

\[
\mathbf{R}_r(t) = \mathbf{K}_p\mathbf{r}(t - 1)
\]

\[
\mathbf{w}(t) = \mathbf{w}(t - 1) + \eta \mathbf{R}_r^H(t)\mathbf{y}(t)
\]

\[
\mathbf{u}(t) = -\mathbf{w}(t)\mathbf{r}(t)
\]

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

In Fig. 3, closed-loop estimates (solid line) of the plant’s frequency response \(L_p(e^{-j\omega})\) (broken line). The sharp notch appears at the frequency of the cancelled narrowband disturbance.
The optimal step-size depends on which channel is optimized – with a step equal to 0.00025. The optimal step-size is artificially generated white noise perturbation [13].

Even if the technique usually works satisfactorily, it underperforms in the presence of narrowband disturbances because the estimated plant gains are strongly biased at the frequency \( \omega_0 \) (which is the frequency of our interest). This phenomenon can be explained as follows: when dealing with narrowband disturbances, the FX-LMS controller is equivalent to a linear time invariant controller with the following transfer function [1]:

\[
K_c(q^{-1}) = b q^{-1} \cos(\omega_0 - \phi) - q^{-2} \cos \phi \\
1 - 2q^{-1} \cos \omega_0 q^{-2}
\]

where \( b \) and \( \phi \) are related to the model of secondary path. The transfer function describing how the auxiliary noise affects the output of the closed loop reads as follows:

\[
C(q^{-1}) = L_p(q^{-1}) / (1 + L_p(q^{-1})K_c(q^{-1})).
\]

Note that the controller has poles at the unit circle, \( p_{1,2} = e^{\pm j\omega_0} \). It follows that, while unaffected outside the passband of the controller \( |C(q^{-1})| \approx L_p(q^{-1}) \) when \( |K_c(e^{-j\omega_0})| \ll 1 \), the closed loop has zeros at \( z_{1,2} = p_{1,2} = e^{\pm j\omega_0} \). This causes a selective bias effect – the estimation results are affected by the notch filtering action of the FXLMS controller. Note that this effect does not occur when FX-LMS is used for broadband disturbance rejection or when the estimation is carried in an open-loop setup.

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 \( \sigma_e^2 = 3 \times 10^{-7} \) and \( \eta = 0.0005 \). Even though the variance of the auxiliary noise \( \sigma_e^2 = 0.0025 \) 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 \( \omega_0 \) – see Fig. 3. When the auxiliary noise was switched off and the plant gain in (41) was frozen at its last closed-loop estimate, the mean-squared cancelling errors increased from \( \text{MSE}_1 = 1 \times 10^{-4} \) and \( \text{MSE}_2 = 2.45 \times 10^{-4} \) (see Table 2) to \( \text{MSE}_1 = 8.13 \times 10^{-4} \) and \( \text{MSE}_2 = 1.49 \times 10^{-3} \), 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.

Table 2 shows comparison of the mean-squared cancellation performance of SONIC and the performance of the optimally tuned FXLMS algorithm (41). The simulated plant and all SONIC settings were identical with those described in Section 6.1 above. The optimal values of \( \eta \) were searched numerically in the interval [0.01, 0.000025] with a step equal to 0.000025. The optimal step-size depends on which channel is optimized – the quantity \( \eta_{\text{opt}}^2 \), shown in Table 2, denotes the value that minimizes the mean-squared cancellation error observed at the first system output (MSE_1), and \( \eta_{\text{opt}}^2 \) – 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 \( \eta \) 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 \( K_0 \) slowly varies with time. In a case like this, the constant-known gain \( K_0 \) can be replaced in (41) with its current estimate \( K_0(t) \) 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 [13]. Even though for wideband disturbances such an auxiliary noise technique usually works satisfactorily, it underperforms in the presence of narrowband disturbances because the estimated plant gains are strongly biased at the frequency \( \omega_0 \) (which is the frequency of our interest).
7. Real-world experiments

7.1. Constant frequency case

A real-world experiment was arranged to check the algorithm’s performance in vibration control. The system consisted of a 4-input, 4-output test stand (Fig. 4), DSP unit, signal conditioning unit and power amplifiers (Fig. 5). The disturbance – a sinusoidal signal with frequency equal to 60 Hz applied on the primary shaker – was generated by a separate source.

The employed DSP unit supports hardware sampling rates from 16 kHz up to 192 kHz. Such values of sampling frequency are grossly above the needs of vibration control system. For this reason the software was equipped with digital sampling frequency down/up-conversion feature, which allowed us to reduce sampling frequency to 200 Hz.

The applied control procedure was two-step: during the first 4 s, the open-loop preliminary plant estimation was performed by sequential application of sinusoidal excitation to all four inputs (1 s each). After this period, the noise cancellation algorithm was started with \( M(0) = 0.0035I \). The following values of adaptation parameters were adopted: \( c_v = 0.0025, \rho = 0.9999 \).

Fig. 6 shows typical results behavior of the adaptive system during initial transient phase which, in this case, lasts approximately 1 s in the first 3 channels and about 5 s in the fourth channel. The poorer rate in the latter case can be attributed to errors of the initial model and crosscoupling between different channels of the plant which made control of the fourth output difficult during the initial phase – the adaptive algorithm required some extra time to compensate modelling errors. However, the steady-state attenuation of disturbance was equal to approximately 30 dB in all channels.

7.2. Time-varying frequency case

In the second experiment, the frequency of the disturbance varied in a sawtooth manner between 55 and 61 Hz. Note that, even though the amplitude of the generated sinusoid was constant, the disturbance observed at the outputs of the plant showed a significant amount of amplitude variation (Fig. 7). Such an effect was caused by a dynamic nature of...

Fig. 5. Additional hardware used in the real-world experiment – DSP unit, signal conditioner and power amplifiers.

Fig. 6. Typical behavior of the adaptive system during the first five seconds of the real-world experiment.
the path linking the primary shaker with the accelerometers. Power spectral densities of the uncancelled disturbance, computed over a single period of frequency variation using Welch’s method with Hamming window, are shown in Fig. 8. A signal with SNR equal to 40 dB served as a reference. The gains of the feedforward estimator were set to the following values: \( \mu_r = 0.1, \gamma_s = 0.05, \gamma_s^2 = 8 \times 10^{-4}. \) Such settings allowed the controller to cope with sweep rate up to 0.8 Hz/s. Fig. 9 depicts power spectral densities of the residual noise for sweep rate equal to 0.5 Hz/s and shows a considerable improvement at all frequencies.

8. Conclusions

The problem of cancellation of a narrowband disturbance acting at the output of an unknown multivariate linear stable
Acknowledgement

This work was supported in part by the National Science Centre.

Appendix A. Proof of Corollary 1

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

\[
\mathbf{C}_\infty(\mathbf{M}) = (\mathbf{I} - \mathbf{M})\mathbf{C}_\infty(\mathbf{M})^H(\mathbf{I} - \mathbf{M})^H + \mathbf{MVM}^H + \mathbf{E}.
\]

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:

\[
\mathbf{\Sigma}(\mathbf{M}) = (\mathbf{I} - \mathbf{M})\mathbf{\Sigma}(\mathbf{M})^H(\mathbf{I} - \mathbf{M})^H + \mathbf{Y}
\]

where \( \mathbf{Y} = (\mathbf{G}_\infty - \mathbf{M})\mathbf{S}_\infty(\mathbf{G}_\infty - \mathbf{M})^H \) and \( \mathbf{S}_\infty = \mathbf{P}_\infty + \mathbf{V} > \mathbf{0} \). Note that \( \mathbf{M} \neq \mathbf{G}_\infty \) entails \( \mathbf{Y} > \mathbf{0} \). Hence, using basic properties of a discrete Lyapunov equation [29], one arrives at \( \mathbf{\Sigma}(\mathbf{M}) > \mathbf{0} \), i.e., \( \mathbf{C}_\infty(\mathbf{M}) > \mathbf{P}_\infty \). The conclusion follows from the fact that \( \mathbf{M} \in \Omega_\infty \) and \( \mathbf{Y} > \mathbf{0} \). The minimum value of \( \mathbf{C}_\infty(\mathbf{M}) \), equal to \( \mathbf{P}_\infty \), is attained when \( \mathbf{Y} = \mathbf{0} \), i.e., when \( \mathbf{M} = \mathbf{G}_\infty \).

Appendix B. Proof of Theorem 1

B.1. Equilibrium point

Since the properties of the closed-loop system depend on the value of \( \mathbf{BM}(t) \), rather than on the value of \( \mathbf{M}(t) \) alone, a new variable \( \mathbf{X} = \mathbf{BM} \) will be introduced. Multiplying both sides of (32) with \( \mathbf{B} \), one arrives at

\[
\mathbf{X} = \mathbf{F}(\mathbf{X})
\]

where \( \mathbf{F}(\mathbf{M}) = -\mathbf{XR}_y(0; \mathbf{X}) \) and \( \mathbf{R}_y(0; \mathbf{X}) = E[\mathbf{y}(t; \mathbf{X})\mathbf{y}^H(t; \mathbf{X})] \).

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

\[
\mathbf{F}(\mathbf{X}_0) = \mathbf{0}.
\]

Note that

\[
\mathbf{y}(t; \mathbf{X}) = e^{j\omega_0}(1 - \mathbf{X})\mathbf{y}(t - 1; \mathbf{X}) + \mathbf{\varepsilon}(t) + \mathbf{v}(t) - e^{j\omega_0}\mathbf{v}(t - 1)
\]

\[
\mathbf{\dot{z}}^H(t; \mathbf{X}) = e^{-j\omega_0}[1 - \mathbf{c}_x]\mathbf{\dot{z}}^H(t - 1; \mathbf{X}) - \mathbf{c}_y\mathbf{y}(t - 1; \mathbf{X})].
\]

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

\[
\mathbf{R}_y(0; \mathbf{X}) = (1 - \mathbf{c}_y)(1 - \mathbf{X})\mathbf{R}_y(0; \mathbf{X}) + \mathbf{c}_y[\mathbf{V} - (1 - \mathbf{X})\mathbf{R}_{yy}(0; \mathbf{X})].
\]

The condition (43) is met provided that

\[
(I - \mathbf{X}_0)\mathbf{R}_{yy}(0; \mathbf{X}_0) = \mathbf{V}.
\]

First, it will be shown that \( \mathbf{X}_0 = \mathbf{G}_\infty \) fulfills this requirement. Note that (according to Corollary 1)

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

Therefore

\[
(I - \mathbf{G}_\infty)\mathbf{R}_{yy}(0; \mathbf{G}_\infty) = (I - \mathbf{G}_\infty)(\mathbf{P}_\infty + \mathbf{V}) = \mathbf{P}_\infty + \mathbf{V} - \mathbf{G}_\infty(\mathbf{P}_\infty + \mathbf{V}) = \mathbf{V}
\]

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 disturbance. 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.
where the last transition stems from the fact that [cf. (13)]
\[ G_{\infty}(P_{\infty} + V) = P_{\infty}. \]
To prove that the solution \( X_0 = G_{\infty} \) is unique, observe that [cf. (44)]
\[ E[\eta(t; X)] = O, \quad R_{yy}(1; X) = (I - X)R_{yy}(0; X) - V \]
\[ R_{yy}(t; X) = (I - X)R_{yy}(\tau - 1; X) \quad \forall \tau \geq 2 \]
It follows that for any solution \( X_0 \) of (45) \( R_{yy}(1; X_0) = R_{yy}(2; X_0) = \ldots = O \), i.e., the sequence \( \{\eta(t; X_0)\} \) is zero-mean and white. Under Gaussian assumptions this property implies that the control loop of (33) is optimal. Since there exists exactly one optimal gain matrix \( M_{\text{opt}} = B^{-1}G_{\infty} \) (which corresponds to \( X_{\text{opt}} = BM_{\text{opt}} = G_{\infty} \)), \( X_0 = G_{\infty} \) is the unique equilibrium point of the ODE (42).

Finally, since it holds that \( X_0 = BM_0 \), one gets \( M_0 = B^{-1}G_{\infty} \) as the unique equilibrium point of (32).

\[ \text{B.2. Local stability} \]
To check stability of the equilibrium point \( X_0 = G_{\infty} \), the ODE (42) will be linearized at \( X_0 \)
\[ \Delta X = d_{\Delta X}F(X_0) \]
where \( \Delta X = X - X_0 \).
To show that the system governed by (42) is locally stable at \( X_0 \), the asymptotic stability of the linearized system (46) will be proved. First of all, note that
\[ d_{\Delta X}F(X_0) = -d_{\Delta X}[X_0R_{yy}(0; X_0)] \]
\[ = -\Delta XR_{yy}(0; X_0) - X_0d_{\Delta X}R_{yy}(0; X_0) \]
\[ = -X_0d_{\Delta X}R_{yy}(0; X_0) \]
where the last transition is a consequence of the fact that (as shown in the preceding subsection)
\[ R_{yy}(0; X_0) = O. \]
Carrying out differentiation of \( R_{yy}(0; X_0) \), one obtains
\[ -d_{\Delta X}R_{yy}(0; X_0) = (1 - c_\mu)(I - X_0)d_{\Delta X}R_{yy}(0; X_0) \]
\[ - (1 - c_\mu)\Delta XR_{yy}(0; X_0) \]
\[ + c_\mu \Delta XR_{yy}(0; X_0) \]
\[ + c_\mu \Delta XR_{yy}(0; X_0). \]
\[ \Delta X = [V - \Delta XR_{yy}(0; X_0)(I - X_0)]^{\dagger} \]
\[ + [V - (I - X_0)R_{yy}(0; X_0)]\Delta X^{\dagger}. \]
Since the last two components on the right hand side of (51) are zero [cf. (45)], one arrives at
\[ d_{\Delta X}R_{yy}(0; X_0) = O. \]
Combining (49) with (48) and (52), one obtains
\[ d_{\Delta X}R_{yy}(0; X_0) = -(1 - c_\mu)(I - X_0)d_{\Delta X}R_{yy}(0; X_0) + c_\mu \Delta XR_{yy}(0; X_0) \]
which finally results in
\[ d_{\Delta X}R_{yy}(0; X_0) = c_\mu D_0 \Delta XR_{yy}(0; X_0) \]
\[ \Delta X = -c_\mu X_0D_0 \Delta XR_{yy}(0; X_0). \]
To show asymptotic stability of the system governed by (56), the Lyapunov approach will be used. Consider the following Lyapunov function
\[ V(\Delta X) = \text{tr}(\Delta X^{\dagger} P_0 \Delta X) \]
where \( P_0 \) is a positive definite matrix yet to be defined. Note that \( V(\Delta X) \geq 0 \), and that the equality holds iff \( \Delta X = 0 \).
Using (56), one obtains
\[ V(\Delta X) = \text{tr}(\Delta X^{\dagger} P_0 \Delta X + \Delta X^{\dagger} P_0 \Delta X) = \text{tr}(P_0 A_0 + A_0^{\dagger} P_0 B_0) \]
where
\[ A_0 = -c_\mu X_0 D_0, \quad B_0 = \Delta XR_{yy}(0; X_0) \Delta X^{\dagger}. \]
To proceed further, the following result will be needed.\[ \begin{align*}
\textbf{Corollary 2.} & \quad \text{For any value } c_\mu, \text{ such that } 0 < c_\mu < 1, \text{ the matrix } A_0 \text{ is stable, i.e., it has eigenvalues with strictly negative real parts} \\
& \quad \text{Re}(\lambda_i(A_0)) < 0, \quad i = 1, \ldots, n. \end{align*} \]
\[ \textbf{Proof.} \quad \text{The proof will start from rewriting } X_0 \text{ in the form } X_0 = Q^{-1}A_0 Q, \text{ where } A_0 = \text{diag}(\{\lambda_i(X_0)\}, \ i = 1, \ldots, n) \text{ is a diagonal matrix made up of the eigenvalues of } X_0. \]
Similarly, \( D_0 \) can be expressed in the form \( D_0 = Q^{-1} \Sigma_0 Q \), where
\[ \Sigma_0 = \text{diag} \left\{ \frac{1}{c_\mu + (1 - c_\mu)\lambda_i(X_0)}, \ i = 1, \ldots, n \right\}. \]
Combining both results, one obtains
\[ \lambda_i(X_0 D_0) = \frac{\lambda_i(X_0)}{c_\mu + (1 - c_\mu)\lambda_i(X_0)}, \quad i = 1, \ldots, n. \]
Since \( X_0 \in \Omega, \) the eigenvalues of \( X_0 \) must obey \( \text{Re}(\lambda_i(X_0)) > 0, \ i = 1, \ldots, n. \) Using this property, it is straightforward to show that for any \( 0 < c_\mu < 1 \) it holds that \( \text{Re}(\lambda_i(A_0)) < 0, \ i = 1, \ldots, n. \)
Using the well-known properties of a continuous Lyapunov equation [29], one can conclude that there exists (exactly one) matrix \( P_0 > 0 \) obeying
\[ P_0 A_0 + A_0^{\dagger} P_0 + I = 0. \]
The conclusion follows from the fact that the matrix \( A_0 \) is stable and the matrix \( I \) is positive definite.
Combining (58) and (60), one arrives at
\[ \dot{V}(\Delta X) = -\text{tr}(B_0) \leq 0 \]
which stems from the fact that the matrix \( R_{yy}(0; X_0) \) is positive definite. Moreover, equality in (61) holds iff \( \Delta X = 0 \). This proves the asymptotic stability of the system governed by (46). \( \square \)
References