Improving performance of adaptive feedforward noise attenuators using a dynamic adaptation gain

Ioan Doré Landau, Bernard Vau, Tudor-Bogdan Airimitoaie, Gabriel Buche

Improving performance of adaptive feedforward noise attenuators using a dynamic adaptation gain

Ioan Doré Landau[a,*], Bernard Vau[b], Tudor-Bogdan Airimitoaie[c], Gabriel Buche[a]

[a] *Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France*
[b] *Ixblue, 94380 Bonneuil-sur-Marne, France*
[c] *Univ. Bordeaux, Bordeaux INP, CNRS, IMS-lab, 33405 Talence, France*

Abstract

The paper explores in detail the use of dynamic adaptation gain/learning rate (DAG) for improving the performance of adaptive feedforward attenuation schemes. The DAG is an ARMA (poles-zeros) filter embedded in gradient type adaptation/learning algorithms and generalizes the various improved gradient algorithms available in the literature. After introducing the DAG algorithm in the context of adaptive feedforward attenuation schemes and providing relationships with other algorithms, its design is developed. Strictly Positive Real (SPR) conditions play an important role in the design of the DAG. Then the stability issues for adaptive/learning systems using a DAG are discussed. The potential of the DAG is then illustrated by experimental results obtained on a relevant adaptive active noise control system.

Keywords: active noise control, adaptive feedforward compensation, Youla–Kučera parametrization, adaptation algorithms, learning algorithms

*Corresponding author
Email addresses: ioan-dore.landau@gipsa-lab.grenoble-inp.fr (Ioan Doré Landau), bernard.vau@ixblue.com (Bernard Vau), tudor-bogdan.airimitoaie@u-bordeaux.fr (Tudor-Bogdan Airimitoaie), gabriel.buche@gipsa-lab.grenoble-inp.fr (Gabriel Buche)

List of Acronyms

ANC - Active Noise Control
ARMA - Auto Regressive Moving Average
AVC - Active Vibration Control
DAG - Dynamic adaptation gain/learning rate
f_s - Sampling frequency
PAA - Parameter adaptation algorithm
PR - Positive real transfer function
SPR - Strictly positive real transfer function
PSD - Power Spectral Density

1. Introduction

Many parameter adaptation algorithms (PAA) used in active noise and vibration control (ANVC) are based on the gradient technique [1]. There are a number of algorithms resulting either from the choice of the criterion to be minimized or from the approximation of the gradient. Other adaptation algorithms used in ANVC result from using a stability point of view for their development [2]. Some of these stability-based algorithms can be interpreted as gradient type algorithms.

One of the most common features in applications is the use of a constant adaptation gain/learning rate in order to keep, on one hand, the alertness of the algorithm in the presence of changes in the characteristics of the disturbances (noise or vibration) to be attenuated, and on the other hand, to have a low computational load. Stability issues not addressed in many cases often lead to the use of very low values for the adaptation gains, resulting in very slow adaptation transients and unsatisfactory performance. The choice of the adaptation gain value is a major problem for applications.

Stability issues have been well understood in adaptive feedforward ANVC schemes. Appropriate algorithms which guarantee the stability of the system even for high adaptation gain values have been developed [3, 4]. These algorithms on one hand take into account the inherent internal positive feedback occurring in adaptive feedforward attenuation schemes and on the other hand use the "a posteriori" residual error as adaptation error (leading to the normalization of the measured residual error). The implementation of these algorithms requires identification of the secondary and reverse paths for design of the appropriate filters acting on the measurements.

Once the stability issue is resolved, a significant problem remains to be addressed: How to choose the value of the adaptation gain assuring the best compromise between the speed of adaptation and steady state performance? More specifically, given a value of the adaptation gain assuring good steady state performance, how can one improve the adaptation speed? The answer to this question is one of the main objectives of this paper.

Solutions derived from the recursive least squares algorithm with forgetting factor have been proposed ([2, 5]). However this approach leads to a significant increase in computational load ([5]). One of these solutions is the so-called "constant trace algorithm" [2, 6]

implementing a time varying matrix adaptation gain[1]. Experimental results in ANVC show a certain improvement in adaptation transients but the computational load becomes much more important than in the case of constant adaptation gains. The signal processing community has also proposed algorithms using a time varying matrix adaptation gain as well as different simplifications in order to try to reduce the computational load (variable step size algorithms) [5]. The question is: does another solution for improving the adaptation transients exist that does not lead to a significant increase in computational load?

Recently, the concept of "dynamic adaptation gain" (DAG) has been introduced in [7] and simulation results (based on an academic example) indicate a high potentiality of this approach. The basic idea behind DAG can be summarized as follows: the gradient (or its approximation) is first filtered before it corrects the current estimated parameters through the (stationary) adaptation gain. It results from [7], that a filter with 3 coefficients allows significant improvement of the adaptation transients (the ARIMA2 algorithm in [7]). This *dynamic adaptation gain* can be interpreted also as a PID (with a filtered derivative) filter acting on the gradient[2].

The objectives of this paper are:

- To investigate the potentiality of *dynamic adaptation gain* in adaptive feedforward noise attenuation schemes using constant adaptation gain/learning rate.
- To provide experimental results obtained on a relevant test-bench dedicated to active noise control (ANC).
- To provide a methodology for design of *dynamic adaptation gains*.

As it will be shown in the paper, "passivity" concepts will play a key role in the design of *dynamic adaptation gains*.

The paper is organized as follows: the experimental set-up is described in Section 2. Then a basic adaptive feedforward configuration using a gradient type PAA is presented in Section 3. The *dynamic adaptation gain/learning rate* is introduced in the context of adaptive feedforward compensators in Section 4. The design of the *dynamic adaptation gain* is discussed in Sections 5 and 6. Experimental results obtained on the ANC test-bench introduced in Section 2 are presented in Section 7.

2. Experimental setup

The view of the test-bench used for experiments and its detailed diagram are shown in Fig. 1. The actual dimensions of the test-bench are given in Fig. 2.

The speaker used as the source of disturbances is labelled 1, while the control speaker is marked 2. At the pipe's open end, the microphone that measures the system's output

[1]The trace of the matrix adaptation gain is kept constant and it is equal to the desired adaptation gain per parameter multiplied by the number of parameters.

[2]The correcting term is no more only proportional to the gradient but it depends also upon the integral of the gradient as well as upon the filtered variations of the gradient.

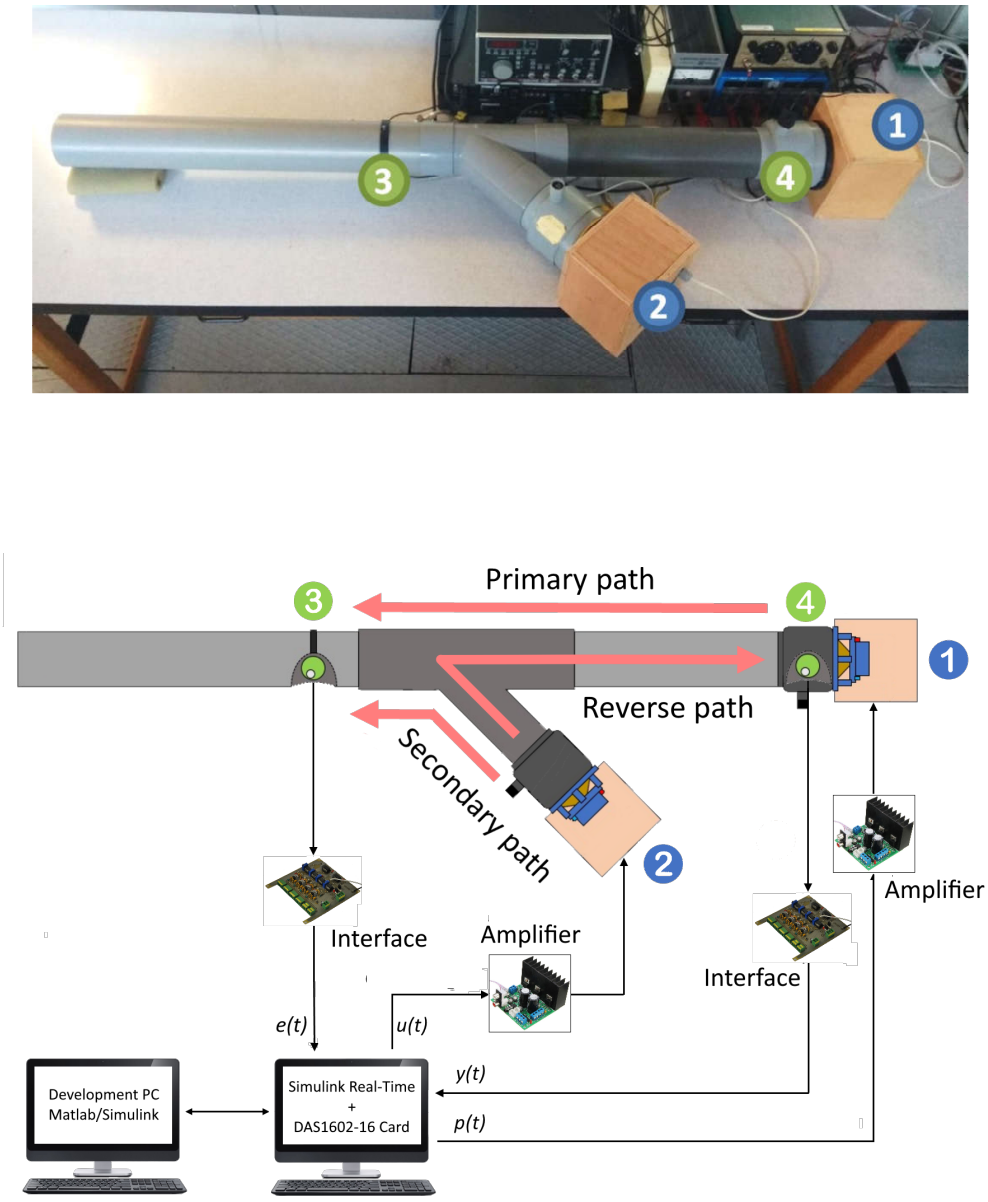

Figure 1: Duct active noise control test-bench photo (top) and block diagram (bottom).

(residual noise $e(t)$) is denoted as 3. $s(t)$ is the disturbance. Inside the pipe, close to the source of the disturbances, the second microphone, labelled 4, measures the disturbance's image, denoted as $y(t)$. $u(t)$ is the control signal. The transfer function between the disturbance speaker and the output microphone (1→3) is called *Global Primary Path*. The

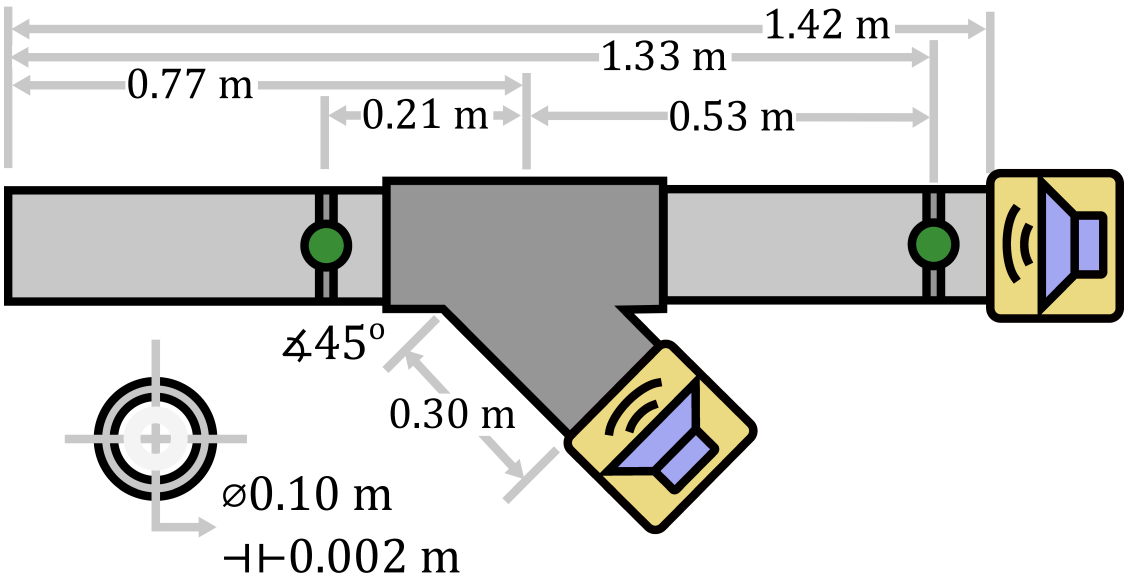

Figure 2: Duct active noise control test-bench dimension.

transfer function between microphones (4→3) is called *Primary Path* (D), while the transfer function between the control speaker and the microphone (2→3) is denoted *Secondary Path* (G). The internal coupling found between (2→4) is denoted *Reverse Path* (M). These marked paths have a double differentiator behavior, since as input we have the voice coil displacement and as output the air acoustical pressure.

In this configuration, speakers are isolated inside wood boxes filled with special foam in order to create anechoic chambers and reduce the produced radiation noise. These boxes have dimensions 0.15 m × 0.15 m × 0.12 m, giving a chamber volume of 2.7 L.

Speakers and microphones are connected to a target computer with Simulink Real-time®. A second computer is used for development and operation with Matlab. The sampling frequency is $f_s = 2500$ Hz.

The various paths are described by models of the form[3]:

$$X(q^{-1}) = q^{-d_x}\frac{B_X(q^{-1})}{A_X(q^{-1})} = q^{-d_x}\frac{b_1^X q^{-1} + ... + b_{n_{B_X}}^X q^{-n_{B_X}}}{1 + a_1^X q^{-1} + ... + a_{n_{A_X}}^X q^{-n_{A_X}}} \tag{1}$$

with $B_X = q^{-1}B_X^*$ for any $X \in \{G, M, D\}$. $\hat{G} = q^{-d_G}\frac{\hat{B}_G}{A_G}$, $\hat{M} = q^{-d_M}\frac{\hat{B}_M}{A_M}$, and $\hat{D} = q^{-d_T}\frac{\hat{B}_T}{A_T}$ denote the identified (estimated) models of G, M, and D. The system's order is defined as (the indexes G, M, and T have been omitted): $n = max(n_A, n_B + d)$.

The frequency characteristics of the identified models for the primary[4], secondary and reverse paths are shown in Fig. 3. These characteristics present multiple resonances (low

[3]The complex variable z^{-1} will be used for characterizing the system's behavior in the frequency domain and the delay operator q^{-1} will be used for describing the system's behavior in the time domain.

[4]The primary path model has been exclusively used for simulation purposes.

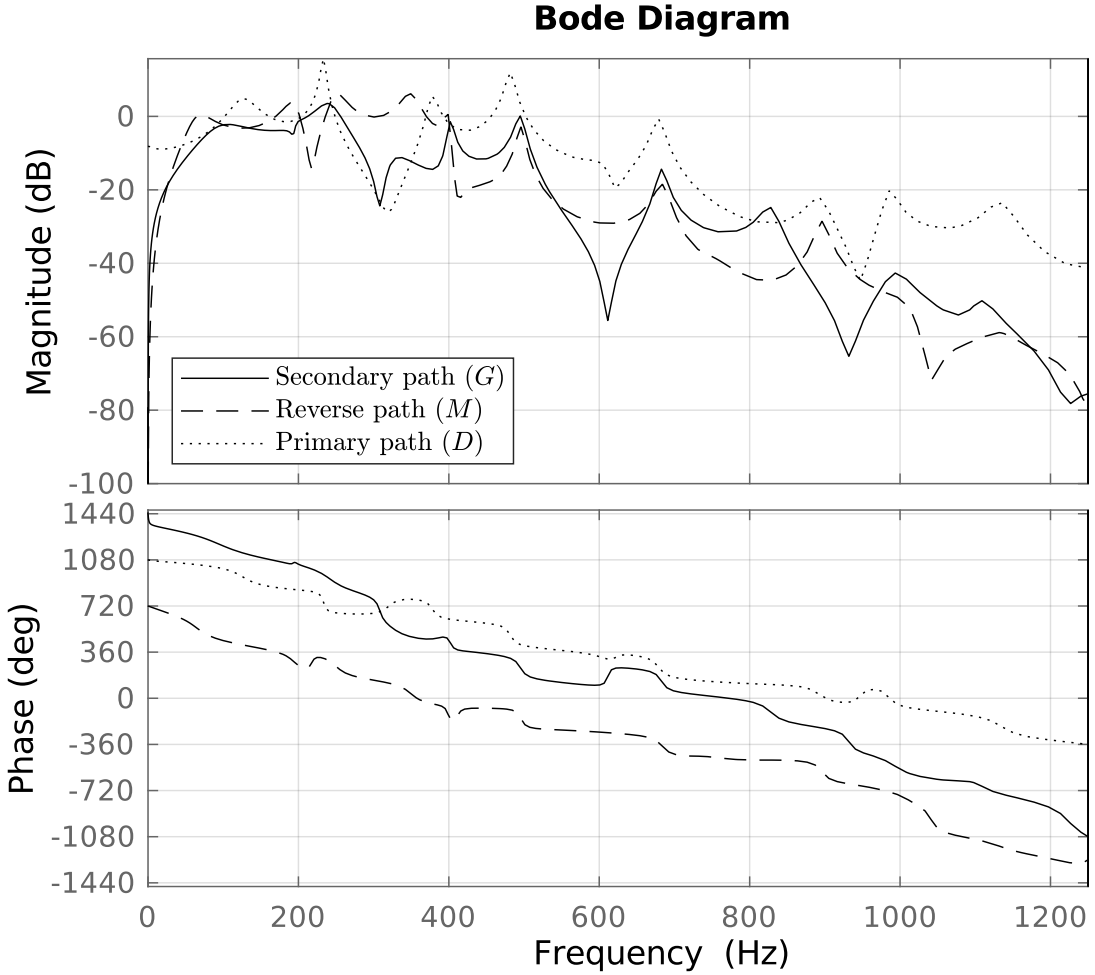

Figure 3: Frequency characteristics of the Primary, Secondary and Reverse paths identified models.

damped complex poles)[5] and anti-resonances (low damped complex zeros). One can see that the secondary path has a high gain between 70 to 250 Hz, which means that disturbances can be efficiently attenuated in this region. It is also clear that the reverse path has a significant gain on a large frequency range (up to 400 Hz) so its effect cannot be neglected. These models have been identified experimentally using the techniques described in [8]. The orders and the pure delays of the various identified models are given in Table 1.

It is important to note that most of the implementations of the adaptive feedforward compensation systems are close to a co-location of the residual noise measurement and of the secondary source used for compensation. See for example [9],[10],[11]. A ratio of 1/6 to 2/6 between the length of the secondary path and the length of the primary path is

[5]The lowest damping is around 0.01.

Model	n_B	n_A	d
Primary (D)	20	19	0
Secondary (G)	22	24	5
Reverse (M)	22	25	5

Table 1: Orders of the identified system paths.

used in these references. Nevertheless, there are new potential applications areas where technological constraints will not allow a configuration to be close to a co-location. For this reason the ratio between the length of secondary path and the length of the primary path has been chosen slightly larger than 4/6.

The objective of the active control system is to attenuate incoming unknown broad-band noise disturbances.

3. The basic adaptive feedforward configuration

For the purpose of this paper (comparison of various adaptation/learning algorithms), the FIR Youla-Kučera configuration for the feedforward compensator has been selected. This configuration has been described previously in the active vibration context [12], [4] and in the active noise context [13, 14]. It offers the great advantage of defining a priori the closed loop poles of the internal positive feedback loop independently of the values of the adaptive FIR filter assuring the attenuation of the incoming disturbance. For a detailed description of the algorithm and of the stability conditions, see for example [13]. In these papers, a stability approach has been considered for the synthesis of the PAAs. In this paper, the PAAs are synthesized using the "gradient rule"[6] and then the stability of the system is analyzed. This approach is necessary for the subsequent introduction of the new algorithms which can be viewed as an improvement of the "gradient" type algorithms.

The corresponding block diagram for the adaptive feedforward noise compensation using the FIR Youla-Kučera (FIR-YK) parametrization of the feedforward compensator is shown in Figure 4. $\hat{y}(t)$ denotes the effective output of the measurement device (microphone in ANC) providing an image of the incoming disturbance and which will serve as input to the adaptive feedforward filter. The output of this filter, denoted by $\hat{u}(t)$, is applied to the actuator through an amplifier. The transfer function G (the secondary path) characterizes the dynamics from the output of the feedforward filter to the residual noise measurement (amplifier + actuator + dynamics of the system). The transfer function D characterizes the primary path. The transfer function M characterizes the positive feedback path.

The adjustable filter $\hat{Q}$ has the structure:

$$\hat{Q}(q^{-1}) = \hat{q}_0 + \hat{q}_1 q^{-1} + ... + \hat{q}_{n_Q} q^{-n_Q} \tag{2}$$

[6]In the gradient rule the correction of the parameters is done in the direction of the gradient of the criterion to be minimized with the minus sign.

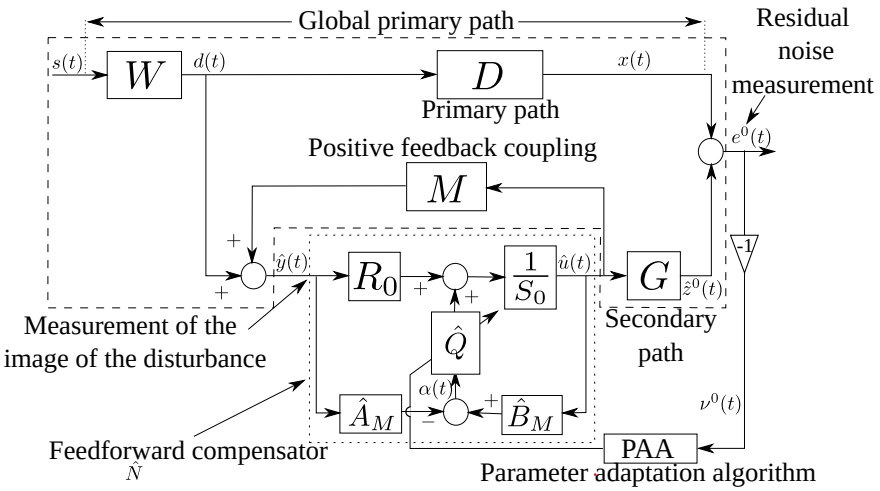

Figure 4: Feedforward AVC with FIR-YK adaptive feedforward compensator.

and the parameters q_i will be adapted in order to minimize the residual error. One defines

$$\mathbf{w}^T = [q_0, q_1, q_2, \ldots, q_{n_Q}] \quad (3)$$

$$\hat{\mathbf{w}}^T = [\hat{q}_0, \hat{q}_1, \hat{q}_2, \ldots, \hat{q}_{n_Q}] \quad (4)$$

$$\mathbf{u}^T(t) = [\alpha(t+1), \alpha(t), \ldots, \alpha(t-n_Q+1)] \quad (5)$$

where:

$$\alpha(t+1) = B_M\hat{u}(t+1) - A_M\hat{y}(t+1) = B_M^*\hat{y}(t) - A_M\hat{u}(t+1) \quad (6)$$

The *a priori* output of the feedforward compensator is given by:

$$\hat{u}^\circ(t+1) = -S_0^*\hat{u}(t) + R_0\hat{y}(t+1) + \hat{\mathbf{w}}^T(t)\mathbf{u}(t) \quad (7)$$

where $\hat{u}(t)$, $\hat{u}(t-1)$, ... are the "a posteriori" outputs of the feedforward filter generated by

$$\hat{u}(t+1) = -S_0^*\hat{u}(t) + R_0\hat{y}(t+1) + \hat{\mathbf{w}}^T(t+1)\mathbf{u}(t) \quad (8)$$

The measured input to the feedforward filter satisfies the following equation (when feedforward compensation is active)

$$\hat{y}(t+1) = d(t+1) + \frac{B_M^*(q^{-1})}{A_M(q^{-1})}\hat{u}(t). \quad (9)$$

The unmeasurable value of the output of the primary path is denoted $x(t)$. The unmeasurable *a priori* output of the secondary path will be denoted $\hat{z}^\circ(t+1)$.

$$\hat{z}^\circ(t+1) = \hat{z}(t+1|\hat{\mathbf{w}}(t)) = \frac{B_G^*(q^{-1})}{A_G(q^{-1})}\hat{y}(t) \quad (10)$$

The *a posteriori* unmeasurable value of the output of the secondary path is denoted:

$$\hat{z}(t+1) = \hat{z}(t+1|\hat{\mathbf{w}}(t+1)) \quad (11)$$

The *a priori* adaptation error is defined as:

$$\nu^{\circ}(t+1) = \nu(t+1|\hat{\mathbf{w}}(t)) = -e^{\circ}(t+1) = -x(t+1) - \hat{z}^{\circ}(t+1) \tag{12}$$

where $e^{\circ}(t+1)$ is the measured residual noise. The *a posteriori* adaptation error (computed) will be given by:

$$\nu(t+1) = \nu(t+1|\hat{\mathbf{w}}(t+1)) = -x(t+1) - \hat{z}(t+1). \tag{13}$$

When using an estimated filter $\hat{N}$ with constant parameters: $\hat{y}^{\circ}(t) = \hat{y}(t)$, $\hat{z}^{\circ}(t) = \hat{z}(t)$ and $\nu^{\circ}(t) = \nu(t)$.

Using the results of [12], the *a posteriori* residual adaptation error can be expressed as:

$$\nu(t+1) = \nu(t+1|\hat{\mathbf{w}}(t+1)) = H(q^{-1})[\mathbf{w} - \hat{\mathbf{w}}(t+1)]^T \mathbf{u}(t) \tag{14}$$

where[7]

$$H(q^{-1}) = \frac{A_M G}{P_0}; \quad P_0 = A_M S_0 - B_M R_0 \tag{15}$$

Assume that we would like to minimize at each step a quadratic criterion in terms of the *a posteriori* residual error

$$\min_{\hat{\mathbf{w}}(t+1)} J(t+1) = [\nu(t+1)]^2 \tag{16}$$

A solution can be provided by the *gradient rule*. The corresponding parameter adaptation/learning algorithm (PALA) will have the form:

$$\hat{\mathbf{w}}(t+1) = \hat{\mathbf{w}}(t) - F \bigtriangledown_{\mathbf{w}} J(t+1) = \hat{\mathbf{w}}(t) - F \frac{\partial J(t+1)}{\partial \hat{\mathbf{w}}(t+1)} \tag{17}$$

where F is the matrix adaptation gain/learning rate and $\bigtriangledown_{\mathbf{w}} J(t+1)$ is the partial gradient of the criterion given in (16) with respect to $\hat{\mathbf{w}}(t+1)$ which has the expression $\frac{\partial J(t+1)}{\partial \hat{\mathbf{w}}(t+1)}$. There are two possible choices for the matrix adaptation gain/learning rate: (i) $F = \gamma I$; $\gamma > 0$; (ii) $F > 0$ (positive definite matrix). For the remainder of the paper we will use the option $F = \gamma I$. The term (constant) *adaptation gain* or *learning rate* is used for γ. Neglecting the non-commutativity of the transfer operator (valid for low values of the adaptation gain), (14) can be re-written as:

$$\nu(t+1) = \nu(t+1|\hat{\mathbf{w}}(t+1)) = [\mathbf{w} - \hat{\mathbf{w}}(t+1)]^T H(q^{-1}) \mathbf{u}(t) \tag{18}$$

Using this equation, the expression of the partial gradient becomes:

$$\bigtriangledown_{\mathbf{w}} J(t+1) = -H(q^{-1}) \mathbf{u}(t) \nu(t+1) \tag{19}$$

[7] P_0 defines the poles of the internal positive closed loop formed by the feedforward compensator and the reverse path.

Unfortunately, $H(q^{-1})$ is not known exactly and should be replaced by its estimation $\hat{H}(q^{-1})$ based on the identified models of the secondary and reverse paths. Therefore the gradient of the criterion will be computed using:

$$\hat{\bigtriangledown}_{\mathbf{w}} J(t+1) = L(q^{-1})\mathbf{u}(t)\nu(t+1); \quad L(q^{-1}) = \hat{H}(q^{-1}) \tag{20}$$

One defines also the regressor vector (a filtered observation vector) as:

$$\mathbf{r}(t) = L(q^{-1})\mathbf{u}(t) = [\alpha_f(t+1), \alpha_f(t), \ldots, \alpha_f(t-nQ+1)] \tag{21}$$

where:

$$\alpha_f(t+1) = L(q^{-1})\alpha(t+1) \tag{22}$$

and the basic gradient type PALA will take the form:

$$\hat{\mathbf{w}}(t+1) = \hat{\mathbf{w}}(t) + \gamma\mathbf{r}(t)\nu(t+1), \tag{23}$$

Using $R_0 = 0$ and $S_0 = 1$, the poles of the internal positive closed loop will be defined by A_M (see Eq. (15)) and they will remain unchanged (see [13] for details). In this context, the filter used in (22) will be $L = \hat{G}$.

Defining the parameter error:

$$\tilde{\mathbf{w}}(t) = \hat{\mathbf{w}}(t) - \mathbf{w} \tag{24}$$

one obtains from (14) and (23):

$$\nu(t+1) = -\frac{H(q^{-1})}{\hat{H}(q^{-1})}\tilde{\mathbf{w}}(t+1)]^T\mathbf{r}(t) \tag{25}$$

$$\tilde{\mathbf{w}}(t+1) = \tilde{\mathbf{w}}(t) + \gamma\mathbf{r}(t)\nu(t+1) \tag{26}$$

$$\tilde{\mathbf{w}}(t+1)^T\mathbf{r}(t) = \tilde{\mathbf{w}}(t)^T\mathbf{r} + \gamma\mathbf{r}^T(t)\mathbf{r}(t)\nu(t+1) \tag{27}$$

Equations (25) and (27) define an equivalent feedback system shown in Fig. 5 and the associated linear transfer operator appearing in the equivalent feedforward path is:

$$H'(q^{-1}) = \frac{H(q^{-1})}{L(q^{-1})} \tag{28}$$

Using the results of [6, Theorem 3.2] one concludes that asymptotic stability will be obtained for any values of the adaptation gain γ and any initial conditions if $H'(z^{-1})$ is a *strictly positive real* (SPR) transfer function. This is a very mild condition as far as a good experimental identification of the models is carried out.

To implement the algorithm, we need an expression of $\nu(t+1)$ as a function of $\nu^\circ(t+1)$. From (18), one gets:

$$\nu(t+1) = \left[\mathbf{w} - \hat{\mathbf{w}}(t+1)\right]^T\mathbf{r}(t) - \hat{H}^*(q^{-1})\nu(t) + H^*(q^{-1})\left[\mathbf{w} - \hat{\mathbf{w}}(t)\right]^T\mathbf{r}(t-1) \tag{29}$$

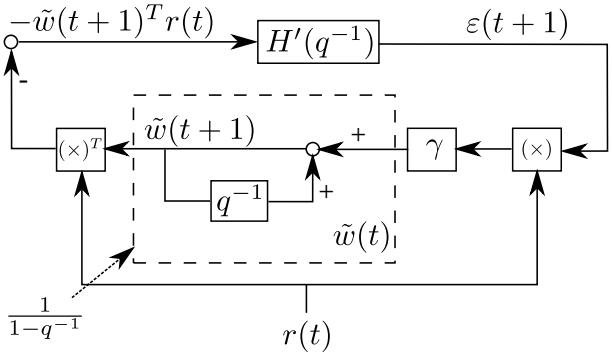

Figure 5: Equivalent feedback system for the adaptive feedforward attenuation scheme.

where

$$H(q^{-1}) = 1 + q^{-1}H^*(q^{-1}); \quad H(\hat{q^{-1}}) = (1 + q^{-1}\hat{H}^*(q^{-1}) \tag{30}$$

The *a priori* adaptation error obeys the equation

$$\nu^\circ(t+1) = [\mathbf{w} - \hat{\mathbf{w}}(t)]^T \mathbf{r}(t) - H^*(q^{-1}) [\mathbf{w} - \hat{\mathbf{w}}(t)]^T \mathbf{r}(t-1) - \hat{H}^*(q^{-1})\nu(t) \tag{31}$$

The *a posteriori* adaptation/learning error can then be written as:

$$\begin{aligned} \nu(t+1) &= \nu^\circ(t+1) - [\hat{\mathbf{w}}(t+1) - \hat{\mathbf{w}}(t)]^T \mathbf{r}(t) \\ &= \nu^\circ(t+1) - \gamma \mathbf{r}(t)^T \mathbf{r}(t)\nu(t+1) \end{aligned} \tag{32}$$

which leads to:

$$\nu(t+1) = \frac{\nu^\circ(t+1)}{1 + \gamma \mathbf{r}^T(t)\mathbf{r}(t)} \tag{33}$$

For the purpose of this paper, it is convenient to express the adaptation algorithm as:

$$\mathbf{r}(t) = \mathbf{u}_f(t) = L(q^{-1})\mathbf{u}(t) \tag{34}$$

$$\nu(t+1) = \frac{\nu^\circ(t+1)}{1 + \gamma \mathbf{r}^T(t)\mathbf{r}(t)} \tag{35}$$

$$\hat{\mathbf{w}}(t+1) = \hat{\mathbf{w}}(t) + \gamma \mathbf{r}(t)\nu(t+1) \tag{36}$$

$$\tag{37}$$

where $\nu^\circ(t+1)$ is given in (12). The asymptotic stability condition for any values of the adaptation gain γ and any initial conditions $\hat{\mathbf{w}}(0)$, $\nu^\circ(0)$ is that $H'(z^{-1})$ given in (28) should be strictly positive real (SPR).

4. An adaptation algorithm using a dynamic adaptation gain/learning rate

The new adaptation algorithm which will be discussed in this section will have the form[8]

$$\hat{\mathbf{w}}(t+1) = \hat{\mathbf{w}}(t) + \gamma \frac{C(q^{-1})}{D'(q^{-1})} [\mathbf{r}(t)\nu(t+1)] \tag{38}$$

[8]It can be interpreted as a dynamic gradient rule.

where $\frac{C(q^{-1})}{D'(q^{-1})}$ is termed the "dynamic adaptation gain/learning rate" (DAG). It is in fact a MIMO diagonal transfer operator having identical terms. All the diagonal terms are of the form:

$$H^{ii}_{DAG}(q^{-1}) = \frac{C(q^{-1})}{D'(q^{-1})} = \frac{1 + c_1 q^{-1} + c_2 q^{-2} + .. + c_{n_C} q^{-n_C}}{1 - d'_1 q^{-1} - d'_2 q^{-2} - ..d'_{n_{D'}} q^{-n'_D}} \tag{39}$$

The effective implementation of the algorithm given in (38) leads to:

$$\begin{aligned}\hat{\mathbf{w}}(t+1) = d_1\hat{\mathbf{w}}(t) + d_2\hat{\mathbf{w}}(t-1) + \ldots + d_{n_D}\hat{\mathbf{w}}(t-n_D) \\ + \gamma[\mathbf{r}(t)\nu(t+1) + c_1\mathbf{r}(t-1)\nu(t) + c_2\mathbf{r}(t-2)\nu(t-1) \\ + \ldots + c_{n_C}\mathbf{r}(t-n_C)\nu(t-n_C+1)]\end{aligned} \tag{40}$$

where:

$$d_i = (d'_i - d'_{i-1}) \;\; ; i = 1, ...n_D; d'_0 = -1, \; d'_{n_D} = 0 \tag{41}$$

To implement the algorithm one needs a computable expression for $\nu(t+1)$. Using a similar development as for the basic gradient adaptation algorithm one finds:

$$\nu(t+1) = \frac{\nu^{\circ}(t+1)}{1 + \gamma\mathbf{r}^T(t)\mathbf{r}(t)} \tag{42}$$

Relations with other algorithms

Many algorithms have been proposed for improving the "gradient rule" in order to accelerate the convergence of the adaptation/learning algorithm for a given adaption gain/learning rate. The algorithm of (38) is termed ARIMA (Autoregressive with Integrator Moving Average) algorithm [7]. The various algorithms described in the literature are of MAI form or ARI form. The MAI form includes "Integral + Proportional" algorithm [6, 15] ($c_1 \neq 0, c_i = 0, \forall\ i > 1,\ d'_i = 0, \forall\ i \geq 0$), "Integral + Proportional + Derivative" algorithm [16, 7] ($c_1, c_2,\ \ c_i = 0\ \ for\ i > 2, d'_i = 0, i \geq 0$), "Averaged gradient" ($c_i,\ for\ i = 1, 2, ..., d'_i = 0, \forall\ i \geq 0$) [17, 18]. The ARI form includes "Conjugate gradient" and "Nesterov" algorithms [19, 20] ($c_i = 0, for\ \ i = 1, 2, ..,\ d'_1 \neq 0,\ d'_i = 0,\ for\ i > 1$) as well as the "Momentum back propagation" algorithm [21] which corresponds to the conjugate gradient plus a normalization of γ by $(1 - d'_1)$[9]. For more details see [7]. Application of the I+P algorithm in active vibration control is discussed in [15].

A particular form of the ARIMA algorithms termed "ARIMA2" ($c_1, c_2,\ c_i = 0, \forall\ i > 2,\ d'_1 \neq 0, d'_i = 0, \forall\ i > 1$) will be studied subsequently and evaluated experimentally.[10]

5. Design of the dynamic adaptation gain - performance issues

The dynamic adaptation gain/learning rate will introduce a phase distortion on the gradient depending on the frequency.

[9] There are very few indications on how to choose the various weights in the above mentioned algorithms.

[10] The algorithms mentioned above can be viewed as particular cases of the ARIMA2 algorithm.

- Assume that the algorithms should operate for all frequencies in the range: 0 to $0.5f_s$.
- Assume that the gradient of the criterion to be minimized contains a single frequency.

In order to minimize the criterion, the phase distortion introduced by the dynamic adaptation gain/learning rate should be less than 90° at all the frequencies. In other terms, the transfer function $H_{DAG} = \frac{C(z^{-1})}{D'(z^{-1})}$ should be SPR. But a transfer function which is SPR has its poles and zeros inside the unit circle (asymptotically stable poles and zeros). For transfer functions having their poles and zeros inside the unit circle one has the following property:

DAG Property 1: *Assume that the polynomials $C(z^{-1})$ and $D'(z^{-1})$ have all their zeros inside the unit circle, then:*

$$I = \int_0^{\pi} \log\left(\left|\frac{C(e^{-i\omega})}{D'(e^{-i\omega})}\right|\right) \mathrm{d}\omega = 0 \tag{43}$$

The proof of this property is given in Appendix A.

This result allows to conclude that the average gain (in dB) over the frequency range 0 to $0.5f_s$ is 0 dB (=1), i.e. on average this filter will not modify the adaptation gain/learning rate. It is just a frequency weighting of the adaptation gain/learning rate. To be more specific, Figure 6 shows the frequency characteristics of two DAGs which will be subsequently used in the experimental section[11]. It can be observed first that the phase is within the range ±90°, i.e. they are SPR. Then one can observe that effectively the average gain over the frequency range 0 to $0.5f_s$ (f_s = 2500 Hz) is 0 dB. Now examining the magnitude, one observes that both are low pass filters amplifying low frequencies. This means that if the frequency content of the gradient is in the low frequency range, the effective adaptation gain/learning rate will be larger than γ, which should have a positive effect upon the adaptation/learning transient. In particular the DAG which has a larger gain in low frequencies (ARIMA2) should provide better performance than the (I+P) DAG (which is indeed the case—see Section 7). Furthermore, since the average gain is 0 dB, the filters shown in Figure 6 will introduce an attenuation in the high frequency range, reducing the effective adaptation gain in this region. This will reduce the influence of the measurement noise located usually in the high frequency region and steady state performance is improved (in general) with respect to the gradient (see Section 7).

Since we need to have a DAG which is SPR, we will provide subsequently the tools for designing a SPR DAG. We will consider the case of the ARIMA2 algorithm introduced in [7]. The DAG in this case will have the form:

$$H_{DAG} = \frac{C(q^{-1})}{D'(q^{-1})} = \frac{1 + c_1 q^{-1} + c_2 q^{-2}}{1 - d_1' q^{-1}} \tag{44}$$

A criterion for the selection of c_1, c_2 and d_1' in order that the DAG be SPR is given below.

SPR criterion for DAG: *The conditions assuring that $H_{DAG}(z) = \frac{1+c_1 z^{-1}+c_2 z^{-2}}{1-d_1' z^{-1}}$ is strictly positive real (SPR) are:*

[11] ARIMA2 filter with $c_1 = 0.99, c_2 = 0, d_1' = 0.9$ and I+P filter with $c_1 = 0.99, c_2 = 0, d_1' = 0$.

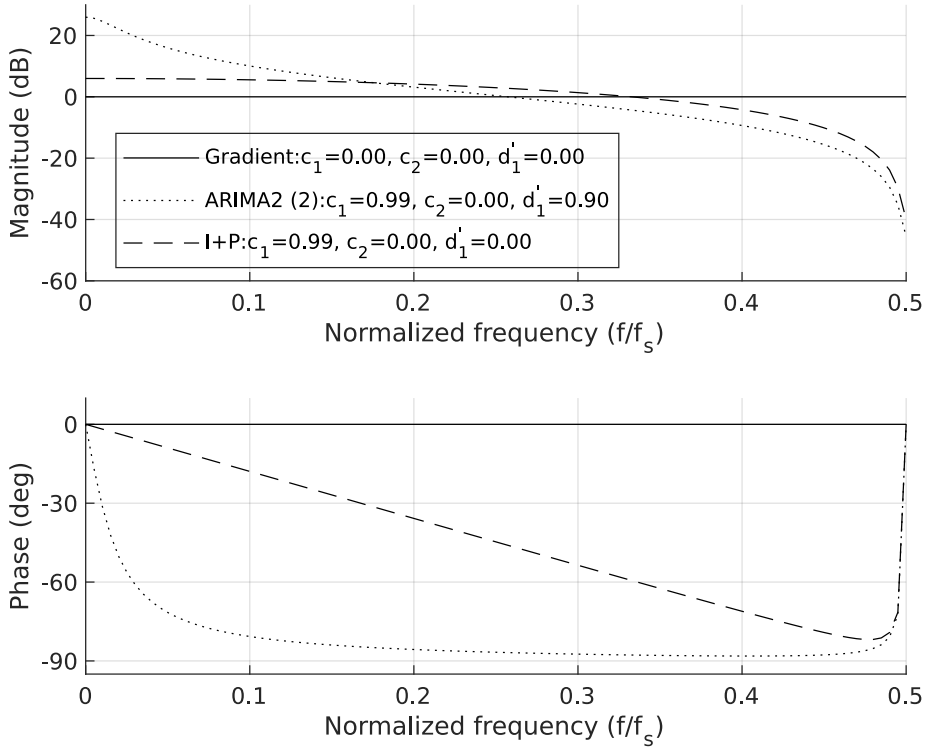

Figure 6: Frequency characteristics of two DAGs (used in the experiments).

- *for* $c_2 \leq 0$, c_1 *must be such that*

$$-1 - c_2 < c_1 < 1 + c_2$$

- *for* $c_2 \geq 0$
  - *if the following condition is satisfied*

$$2(d_1^{'} - c_2) < \sqrt{2(c_2 - c_2^2)(1 - d_1^{'2})} < 2(d_1^{'} + c_2)$$

the maximum bound on c_1 *is given by*

$$c_1 < d_1^{'} - 3d_1^{'}c_2 + 2\sqrt{2(c_2 - c_2^2)(1 - d_1^{'2})}$$

otherwise the maximum bound on c_1 *is given by*

$$c_1 < 1 + c_2$$

– *if the following condition is satisfied*

$$2(d_1^{'} - c_2) < -\sqrt{2(c_2 - c_2^2)(1 - d_1^{'2})} < 2(d_1^{'} + c_2)$$

the minimum bound on c_1 *is given by*

$$c_1 > d_1^{'} - 3d_1^{'}c_2 - \sqrt{2(c_2 - c_2^2)(1 - d_1^{'2})}$$

otherwise the minimum bound on c_1 *is given by*

$$c_1 > -1 - c_2$$

The proof of this result is given in Appendix B.

From these conditions, closed contours in the plane $c_2 - c_1$ can be defined for the different values of $d_1^{'}$ allowing to select c_1 and c_2 for a given value of $d_1^{'}$ so that the DAG is SPR[12]. Such a diagram is presented in Figure 7.

6. Design of the dynamic adaptation gain—stability issues

Eq. (38) can be expressed also as

$$\hat{\mathbf{w}}(t+1) = H_{PAA}(q^{-1})\gamma \mathbf{r}(t)\nu(t+1) \tag{45}$$

where H_{PAA} is a MIMO diagonal transfer operator having identical terms. All the diagonal terms are described by:

$$\begin{aligned} H_{PAA}^{ii}(q^{-1}) &= \frac{1 + c_1 q^{-1} + c_2 q^{-2} + .. + c_{n_C} q^{-n_C}}{(1 - q^{-1})(1 - d_1^{'} q^{-1} - d_2^{'} q^{-2} - ..d_{n_{D'}}^{'} q^{-n_D^{'}})} \\ &= \frac{C(q^{-1})}{(1 - q^{-1})D'(q^{-1})} = \frac{C(q^{-1})}{D(q^{-1})} \end{aligned} \tag{46}$$

The relation between the coefficients of polynomials D and D' is given in (41).

One has the following result for the stability of the adaptive feedforward attenuation scheme using a dynamic adaptation gain/learning rate:

Stability Conditions with DAG: *For the system described by Eqs* (2) *through* (14) *using the PALA of* (40), (41) *and* (42), *one has* $\lim_{t\to\infty} \nu(t+1) = 0$ *for any positive adaptation gain* γ *and initial conditions* $\hat{\mathbf{w}}(0), \nu(0)$, *if:*

- *The transfer function associated with the transfer operator* $H_{PAA}^{ii}(q^{-1})$ *given in* (46) *is a positive real (PR) transfer function with a pole at* $z = 1$.
- *The transfer function* $H'(z^{-1})$ *given in* (28) *is a strictly positive real (SPR) transfer function*

[12] A Matlab routine is available for drawing these contours.

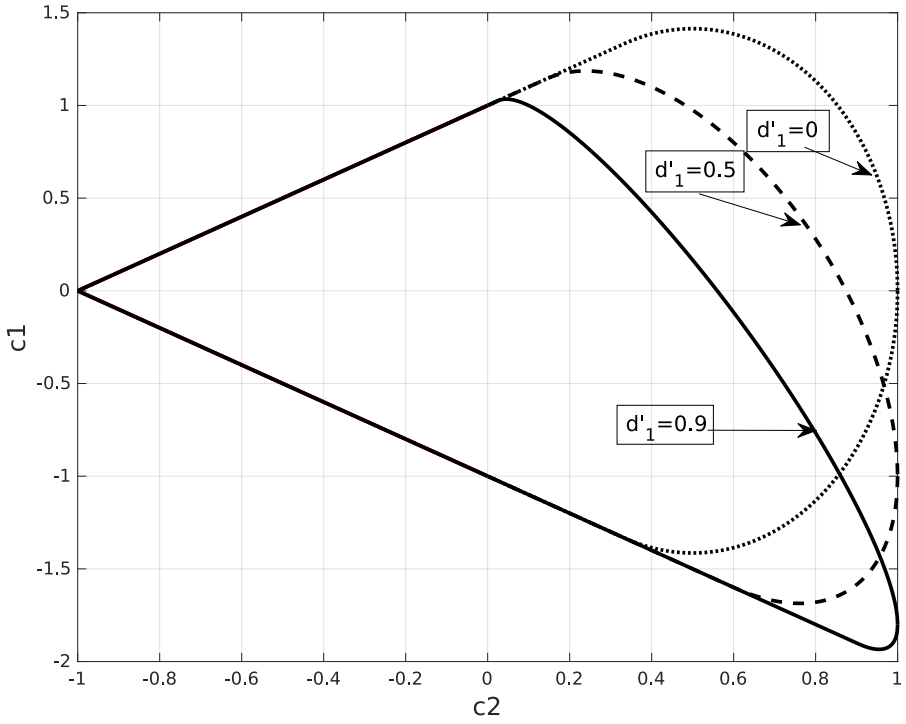

Figure 7: SPR contours in the plane $c_2 - c_1$ for various values of $d_1'(0, 0.5, 0.9)$ for the DAG given in Eq. (44).

The proof of this result is given in Appendix C.

For the particular case of the ARIMA2 algorithm, in order to operate with high adaptation/learning rate, the coefficients c_1, c_2 and d_1' should be chosen such that the DAG is SPR and the H_{PAA}^{ii} is PR, i.e.

$$H_{PAA}^{ii} = \frac{1 + c_1 q^{-1} + c_2 q^{-2}}{1 - d_1 q^{-1} - d_2 q^{-2}} = \frac{1 + c_1 q^{-1} + c_2 q^{-2}}{(1 - q^{-1})(1 - d_1' q^{-1})} \tag{47}$$

should be PR. It is important for applications to give explicit bounds for the selection of the coefficients c_1, c_2, d_1' in order to guarantee the positive realness of the embedded ARIMA filter. One has the following result:

PR conditions for H_{PAA}: *In order that the transfer operator H_{PAA}^{ii} given in* (47) *be characterized by a PR transfer function, the necessary and sufficient conditions are:*

$$-1 < d_1^{'} < 1 \tag{48}$$

$$0 \leq \delta \leq 2 \tag{49}$$

$$-1 \leq d_1^{'} - \frac{\gamma}{1 - \delta/2} \leq 1 \tag{50}$$

where:

$$\delta = \frac{1 + c_1 + c_2}{1 - d_1'}; \quad \gamma = \frac{d_1' c_1 + d_1'^2 + c_2}{d_1' - 1} \tag{51}$$

The proof of this criterion is given in Appendix D.

From these conditions, closed contours in the plane $c_2 - c_1$ can be defined for the different values of d_1' allowing to select c_1 and c_2 for a given value of d_1' so that the H_{PAA} is PR[13]. Such a diagram is presented in Figure 8.

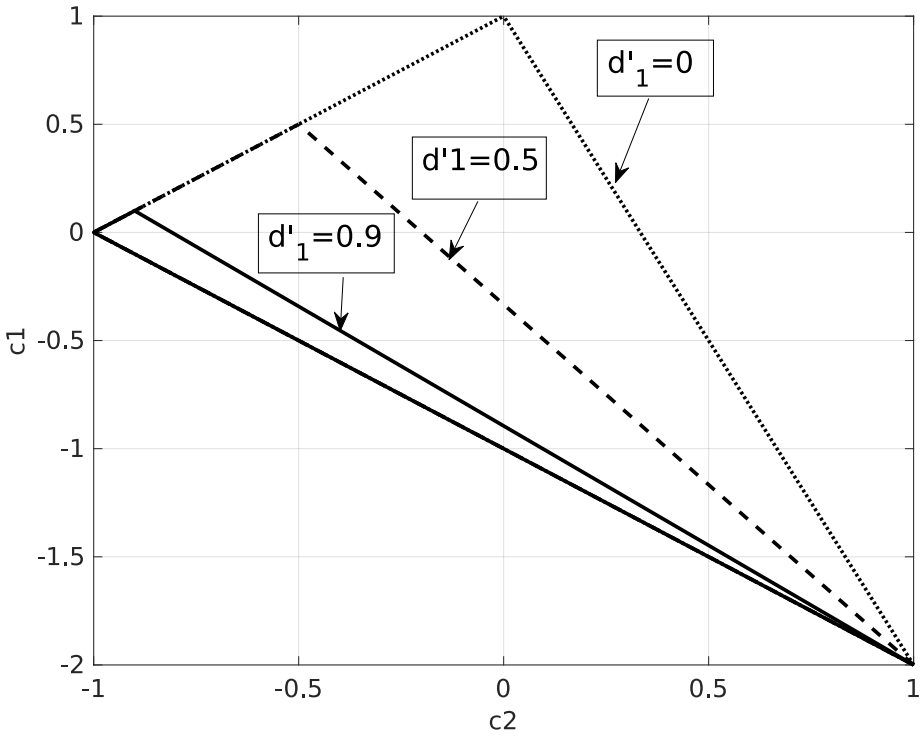

Figure 8: PR contours in the plane $c_2 - c_1$ for various values of $d_1'(0, 0.5, 0.9)$ for the H_{PAA} of Eq. (47)

It is also interesting to see the intersections of the contours assuring the SPR of the H_{GAD}^{ii} with the contours assuring that H_{PAA}^{ii} is PR. Such an intersection is shown in Fig.9. From this figure one can conclude that not all the SPR H_{GAD} will lead to an H_{PAA} PR. In such cases the performances are improved for low adaptation gain, but one cannot guarantee asymptotic stability for large values of the adaptation gain γ (see [7] for a simulated example). Fig. 9 shows also that there is a region where despite H_{PAA} being PR, H_{DAG} is not SPR. For such configurations, large adaptation gains can be used but the adaptation transient is slower than for the basic gradient algorithm.

[13]A Matlab routine is available for drawing these contours.

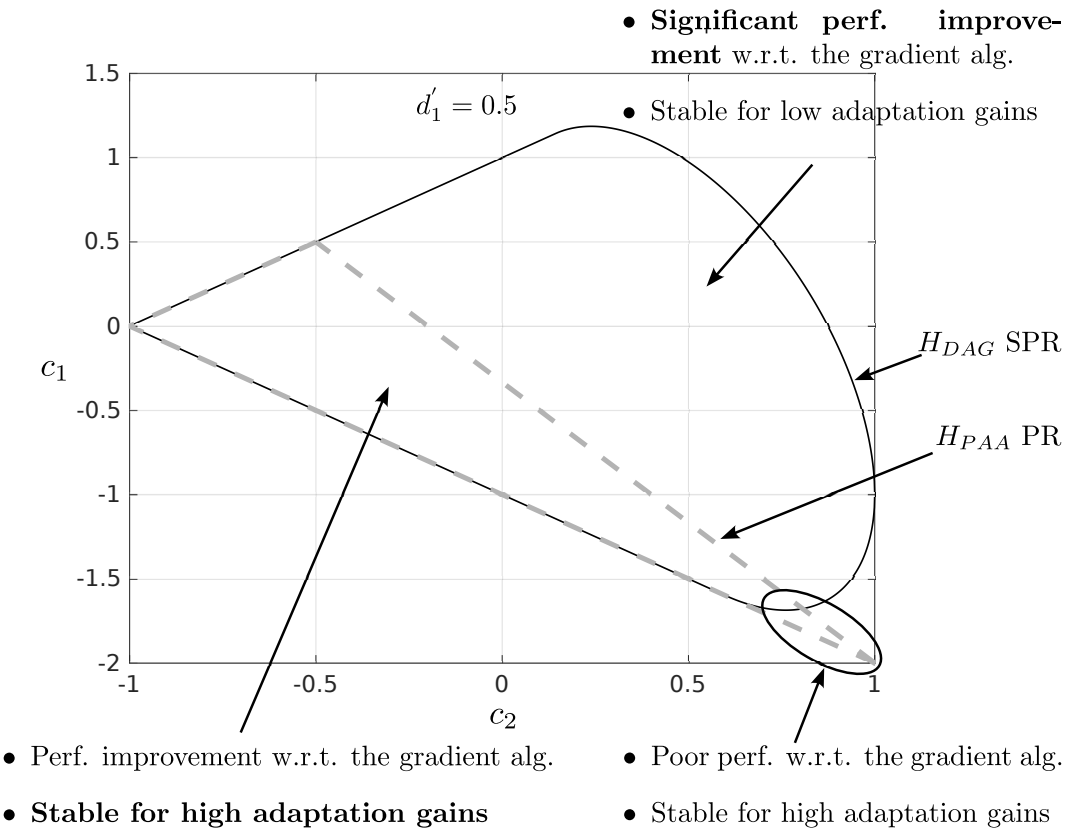

Figure 9: Intersection in the plane $c_1 - c_2$ of the contour $H_{PAA} = PR$ with the contour $H_{GAD} = SPR$ for $d_1' = 0.5$

For small values of the adaptation gains/learning rates the passivity/stability condition can be relaxed using *averaging* [22]. Using the results of [3], under the hypothesis of an input signal spanning all the frequencies up to half of the sampling frequency, passivity in the average will be assured if the frequency interval where H_{PAA}^{ii} is not positive real is smaller than the frequency interval where H_{PAA}^{ii} is positive real. In fact what is important is that the H_{PAA}^{ii} is PR in the frequency region of operation (mainly defined by the spectrum of the input signals to the systems).

7. Experimental results

Two types of signals used as unknown disturbances acting on the system have been considered:

- A broad-band disturbance 70-170 Hz, to which two tonal disturbances located at 100 Hz and 140 Hz have been added
- A broad-band disturbance 70-220 Hz

The steady state and transient attenuation[14] will be evaluated for the various values of the parameters c_1, c_2 and d_1' given in Tables 2 and 3. The system will operate in open loop during the first 15 s. For all the experiments the constant adaptation gain γ has the value 0.2 and the adjustable filter $\hat{Q}$ has 60 parameters.

Table 2: ARIMA algorithms with SPR H_{DAG}

Algorithm	$H_{PAA}PR$	$H_{DAG}SPR$	c_1	c_2	d_1'
Gradient	Y	Y	0	0	0
Conj.Gr/Nest.	N	Y	0	0	0.9
I+P+D ($\gamma_P = -2\gamma_D$)	N	Y	1	0.8	0
I+P	Y	Y	0.99	0	0
ARIMA2 (1)	N	Y	-0.55	0.4	0.75
ARIMA2 (2)	N	Y	0.99	0	0.9

Table 3: ARIMA algorithms with SPR H_{DAG} and H_{PAA} PR.

Algorithm	$H_{PAA}PR$	$H_{DAG}SPR$	c_1	c_2	d_1'
Gradient	Y	Y	0	0	0
Conj.Gr/Nest.	Y	Y	0	0	0.333
I+P+D ($\gamma_P = -2\gamma_D$)	Y	Y	0.1	0.333	0
I+P	Y	Y	0.99	0	0
ARIMA 2	Y	Y	0.1	0.08	0.22

7.1. Broad-band disturbance 70-170 Hz plus two tonal disturbances located at 100 Hz and 140 Hz

Figure 10 shows the time response of the system as well as the evolution of the global attenuation when using the Gradient algorithm (top) and the ARIMA2 (2) algorithm (bottom) (last row of Table 2). One can observe that the ARIMA2 (2) algorithm converges faster than the Gradient algorithm and provides also better attenuation in steady state. Figure 11 shows for the same experiments a comparison of the PSD of the residual noise in the absence of the control and the PSD of the residual noise at 600 s under the control effect. One can observe that the ARIMA2 algorithm provides better attenuation in the low frequencies as well as in the high frequency range when compared with the Gradient algorithm. This is coherent with the frequency characteristics of the ARIMA 2 dynamic adaptation gain, shown in Figure 12. It has a larger gain in low frequencies and introduces

[14]The attenuation is defined as the ratio between the variance of the residual noise in the absence of the control and the variance of the residual noise in the presence of adaptive feedforward compensation. The variance is evaluated over a horizon of 3 s.

attenuation in the high frequencies (in other terms: the adaptation gain is higher in low frequencies and lower in the high frequencies).

The top of Fig. 13 shows the time evolution of the global attenuation for the algorithms considered in Table 2. As it can be seen, there is a clear improvement in the adaptation transient using ARIMA2 (2) (last row of Table 2) with respect to the Gradient algorithm (first row of Table 2). The Conjugate Gradient/Nesterov algorithm also provides a significant improvement of the transient. The ARIMA2 (2) algorithm reduces the adaptation/learning transient by a factor of three and a half. One also sees an improvement in steady state attenuation with respect to gradient adaptation. The other algorithms (from Table 2) also provide an improvement with respect to the gradient algorithms. Note that except for the I+P algorithm, H_{PAA} is not PR, therefore there is a risk of instability for large values of the adaptation gain γ.

The bottom of Fig. 13 shows the time evolution of global attenuation for the algorithms of Table 3 (PR constraint on H_{PAA}). The I+P adaptation algorithm gives the best results. The transient improvement provided by the various algorithms with respect to the gradient algorithm is slightly less significant than for the algorithms of Table 2. However these algorithms will tolerate high values of the adaption/learning rate.

7.2. Broad band disturbance 70-220 Hz

The top of Fig. 14 shows the time evolution of global attenuation for the algorithms given in Table 2. One can draw the same conclusions as for the previous type of disturbance. A significant reduction of the adaptation transient with respect to the gradient is obtained with the various algorithms proposed. The ARIMA2 (2) algorithm (last row of Table 2) provides the best result followed closely by the Conjugate Gradient/Nesterov algorithm.
The bottom of Fig. 14 shows the time evolution of global attenuation for the algorithms of Table 3 (PR constraint on H_{PAA}). Similar results to those obtained for the previous type of disturbance are obtained. The I+P adaptation algorithm gives the best results in terms of adaptation transients.

8. Conclusions

The paper has tried to give a global picture of the use of dynamic adaptation gain/learning rate for improving the convergence speed of adaptive feedforward noise attenuation schemes using the gradient algorithm with constant adaptation gain/learning rate. The dynamic adaptation gain introduces a frequency weighting which allows to tune it for getting high gain in low frequencies and attenuation in high frequencies. It is this behaviour which explains the benefit of using the dynamic adaptation gain for improving both adaptation transients and steady state performance. While the experimental verifications have been carried out in the context of active noise control, similar results can certainly be obtained in the context of active vibration control.

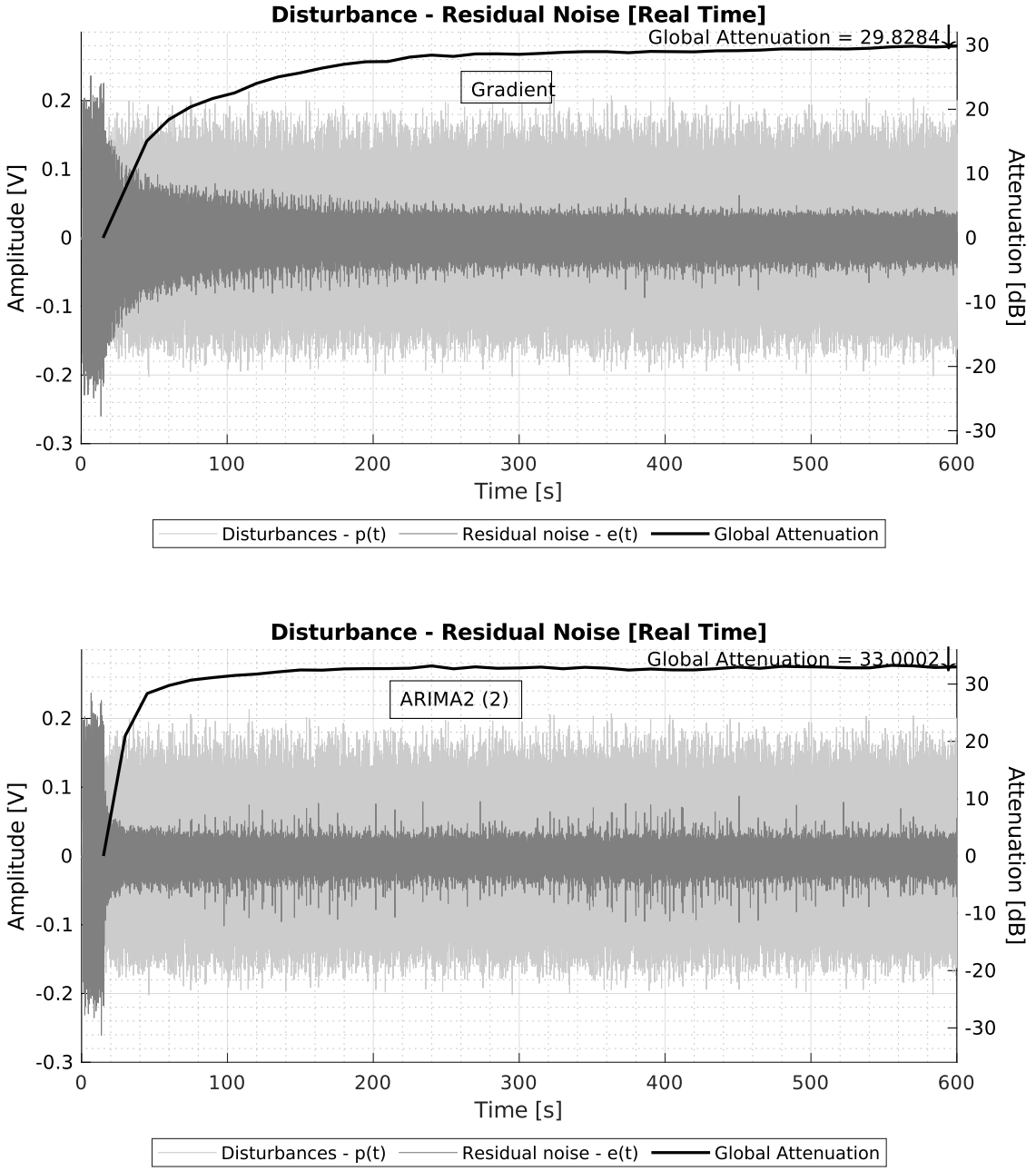

Figure 10: Time evolution of the residual noise using the Gradient algorithm (top) and using the ARIMA2 algorithm (bottom).

Appendix A. Average gain property for strictly positive real DAG

Lets us consider the function $\log\left(\left|\frac{C(z^{-1})}{D'(z^{-1})}\right|\right)$, where $C(z^{-1}) = 1 + \sum_{k=1}^{nC} c_k z^{-k}$, $D'(z^{-1}) = 1 - \sum_{k=1}^{nD'} d'_k z^{-k}$. One seeks to evaluate (43).

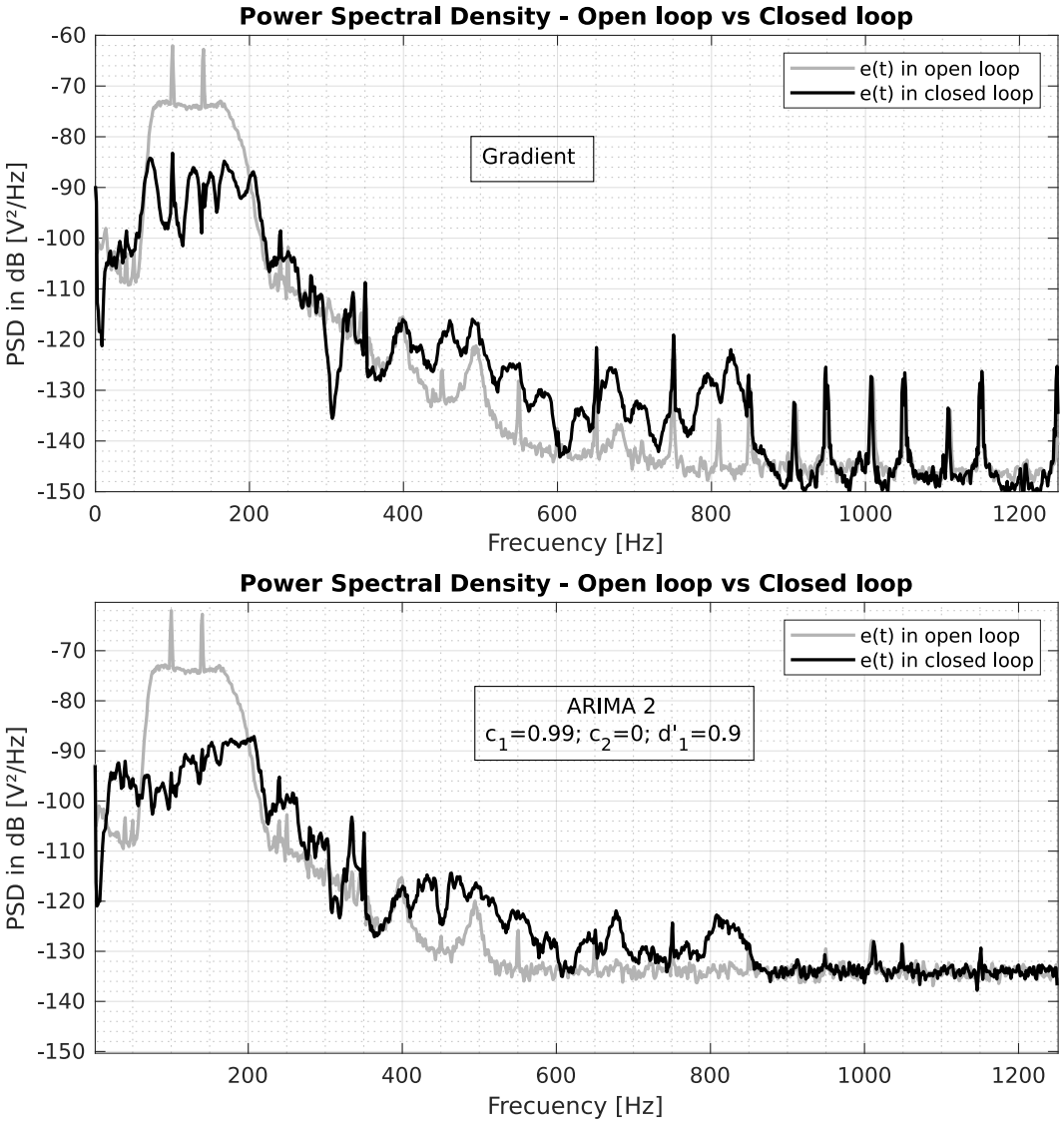

Figure 11: Comparison of the residual noise PSD in open loop and in closed loop (t=600 s) using the Gradient algorithm (top) and using the ARIMA2 (2) algorithm. Disturbance 70 - 170 Hz + 2 tonal disturbances. (bottom).

For $|z| = 1$, one has $z = e^{j\omega}$, $\mathrm{d}z = je^{j\omega}\mathrm{d}\omega$, and $\frac{\mathrm{d}z}{z} = j\mathrm{d}\omega$. Thus one can write

$$I = \frac{1}{j}\left(\oint_{\mathbb{T}} \log\left|1 + \sum_{k=1}^{nC} c_k z^{-k}\right| \frac{\mathrm{d}z}{z} - \oint_{\mathbb{T}} \log\left|1 - \sum_{k=1}^{nD'} d'_k z^{-k}\right| \frac{\mathrm{d}z}{z}\right) \quad \text{(A.1)}$$

where $\mathbb{T}$ is the unit circle.

On the other hand $\left|1 + \sum_{k=1}^{nC} c_k z^{-k}\right| = \left|1 + \sum_{k=1}^{nC} c_k z^{k}\right|$ for $|z| = 1$. The poles of $f(z) = \log\left(\left|1 + \sum_{k=1}^{nC} c_k z^{k}\right|\right)$ are the zeros of $\left|1 + \sum_{k=1}^{nC} c_k z^{k}\right|$, and they all lie outside the unit

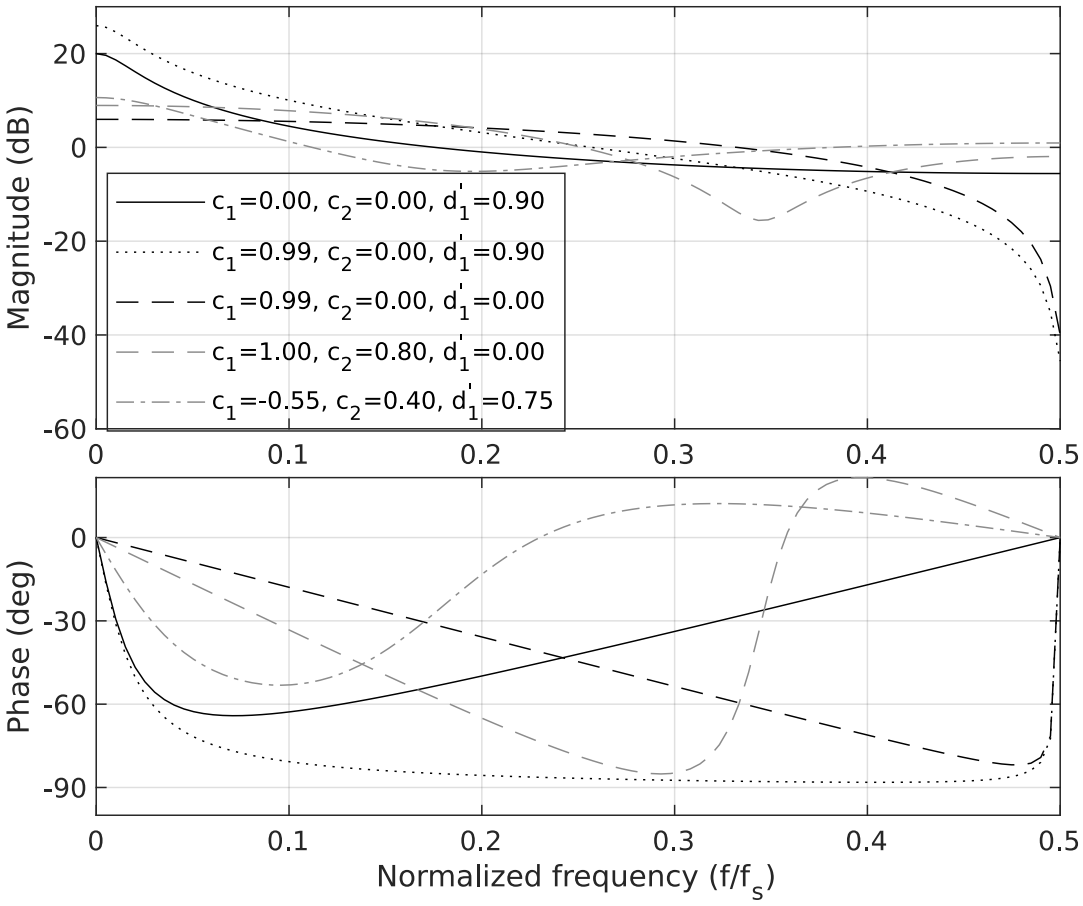

Figure 12: Frequency figcharacteristics of H_{DAG} for the various algorithms given in table 3

circle (by assumption, otherwise H_{PAA} cannot be PR). Therefore the function $f : z \mapsto f(z)$ is holomorphic in the open unit circle, and one can apply the Cauchy Integral's Formula (see [23, pg. 411]). This formula yields

$$\frac{1}{j}\oint_{\mathbb{T}} f(z)\frac{\mathrm{d}z}{z} = Ind(\mathbb{T},0)f(0)$$

where $Ind(\mathbb{T},0)$ is the index of the unit circle with respect to $z=0$. One has $Ind(\mathbb{T},0)=1$, and $f(0)=\log(1)=0$. Therefore one gets

$$\frac{1}{j}\oint_{\mathbb{T}} \log\left(\left|1+\sum_{k=1}^{nC} c_k z^{-k}\right|\right)\frac{\mathrm{d}z}{z} = \frac{1}{j}\oint_{\mathbb{T}} f(z)\frac{\mathrm{d}z}{z} = 0$$

The same machinery can be applied *mutatis mutandis* for $f(z) = \log\left|1-\sum_{k=1}^{nD'} d'_k z^k\right|$

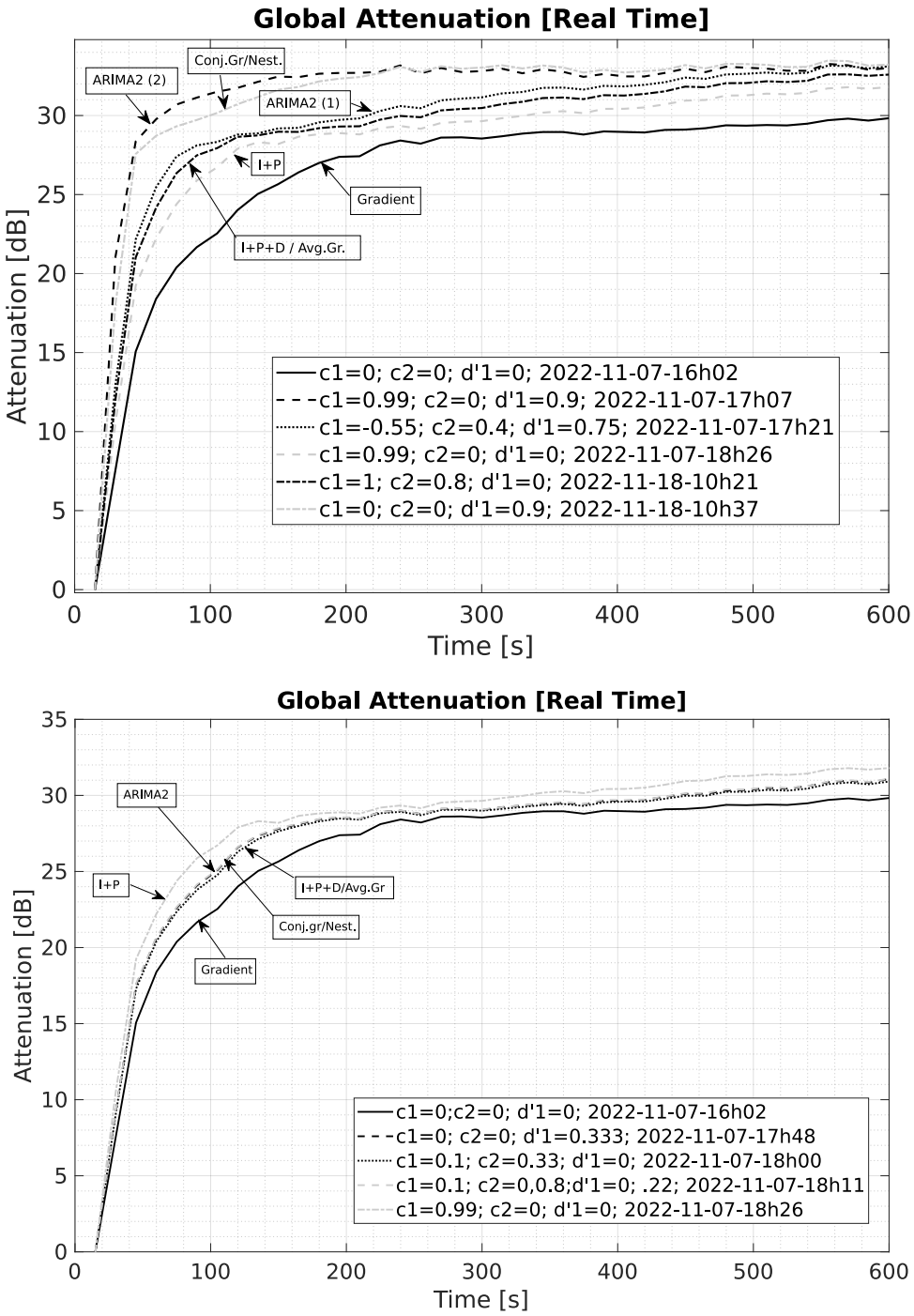

Figure 13: Time evolution of the global attenuation for the algorithms of Table 2 (top) and of Table 3 (bottom). Disturbance 70 - 170 Hz + 2 tonal disturbances.

and one finally obtains

$$\int_{-\pi}^{+\pi} log\left(\left|\frac{C(e^{-i\omega})}{D'(e^{-i\omega})}\right|\right) d\omega = 0$$

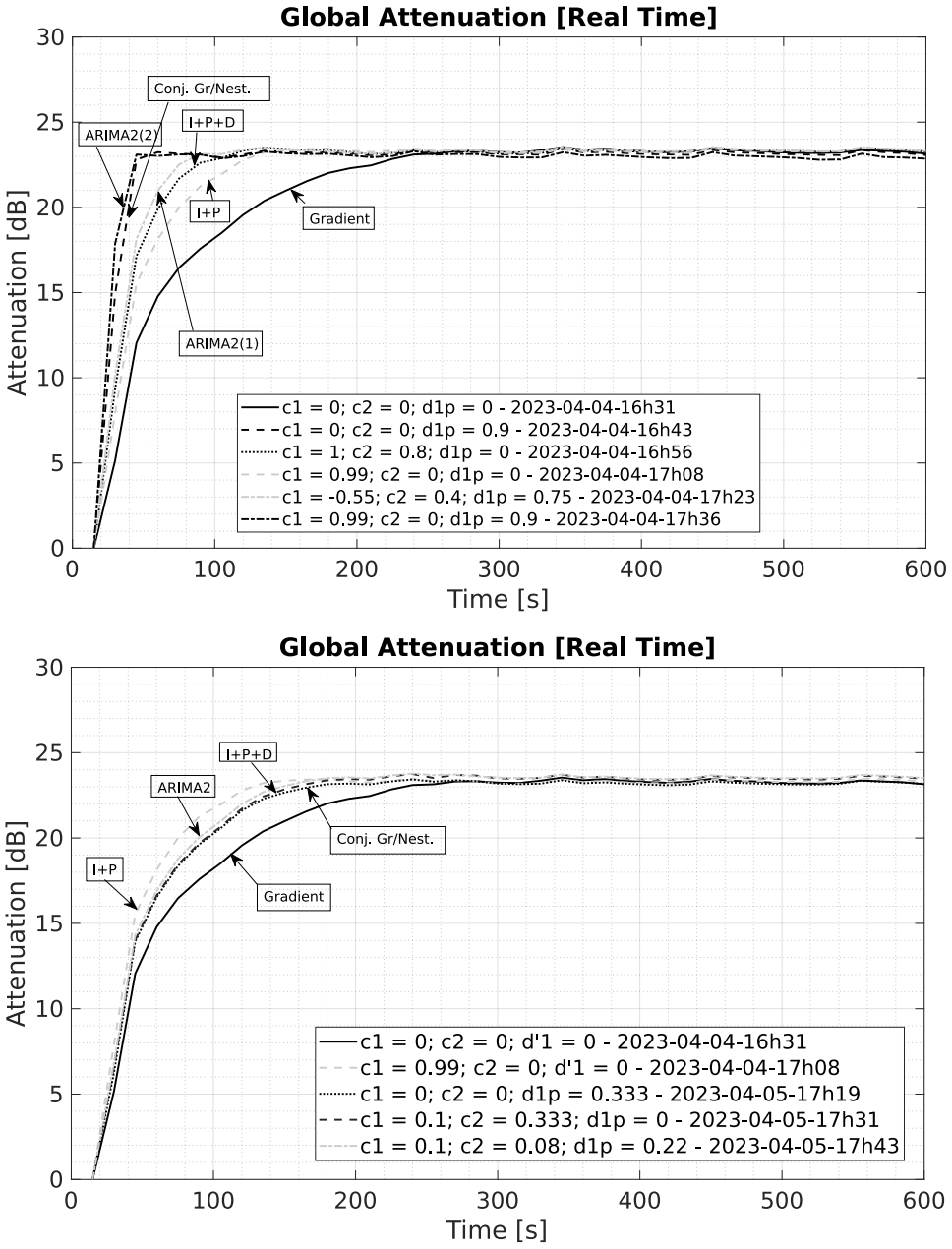

Figure 14: Time evolution of the global attenuation for the algorithms of Table 2 (top) and of Table 3 (bottom). Disturbance 70 -220 Hz.

and since the the function

$$\left| \frac{C(e^{-i\omega})}{D'(e^{-i\omega})} \right|$$

is even one gets the claimed result.

Appendix B. Proof of SPR conditions for H_{DAG}

In order to assess the strict real positivity of $H_{DAG}(z)$ on must check the condition

$$\mathrm{Re}\left((1-d_1^{'}z)(1+c_1z^{-1}+c_2z^{-2})\right)>0 \tag{B.1}$$

Set $z=e^{j\omega}=\cos(\omega)+j\sin(\omega)$, and the condition (B.1) becomes

$$(1-c_2-d_1'c_1)+(c_1-d_1'c_2-d_1')\cos(\omega)+2c_2\cos^2(\omega)>0 \tag{B.2}$$

Set $X=\cos(\omega)$, $x\in[-1,1]$ and $f(X)=2c_2X^2+(c_1-d_1'c_2-d_1')X+(1-c_2-d_1'c_1)$.

- case $c_2\leq 0$
  f has a finite maximum, and it is located at $X_{max}=\frac{-c_1+d_1'c_2+d_1'}{4c_2}$.
  If $X_{max}>1$ one must verify $f(-1)>0$, moreover one has $f(1)>f(-1)$.
  If $X_{max}<-1$ one must verify $f(1)>0$, moreover one has $f(-1)>f(1)$.
  If $-1<X_{max}<1$ one must verify at the same time $f(-1)>0$ and $f(1)>0$.
  In any case one must check that $min(f(-1),f(1))>0$. But $f(1)>0$ implies that $c_1>-c_2-1$, and $f(-1)>0$ implies that $c_1<c_2+1$. Thus for $c_2<0$ the passivity condition is equivalent to $-1-c_2<c_1<1+c_2$.

- case $c_2=0$
  In this case f is represented by a line, and one must again verify that $f(-1)>0$ and $f(1)>0$ that leads to the passivity condition $-1<c_1<1$

- case $c_2>0$
  In this case f has a finite minimum at $X_{min}=\frac{-c_1+d_1^{'}c_2+d_1^{'}}{4c_2}$. A sufficient condition for $f(X)\geq 0\ \forall X$ is that $f(X)=0$ has a unique solution. In such a situation the discriminant of f denoted Δ is given by $\Delta=(c_1-d_1^{'}c_2-d_1^{'})^2-8c_2(1-c_2-d_1^{'}c_1)$, and one must have $\Delta=0$, which is equivalent to

$$c_1^2+c_1(-2d_1^{'}+6d_1^{'}c_2)+d_1^{'2}(c_2+1)^2+8c_2(c_2-1)=0 \tag{B.3}$$

  Thus, one looks for the solutions of (B.3). The discriminant Δ' of (B.3) is $\Delta'=32(c_2-c_2^2)(1-d_1^{'2})$, and the two solutions of (B.3) are

$$c_{1+}^*=d_1^{'}-3d_1^{'}c_2+2\sqrt{2(c_2-c_2^2)(1-d_1^{'2})}$$

$$c_{1-}^*=d_1^{'}-3d_1^{'}c_2-2\sqrt{2(c_2-c_2^2)(1-d_1^{'2})}$$

  On the other hand if $-1\leq X_{min}\leq 1$ one must have (owing to the expression of X_{min})

$$-4c_2+d_1^{'}c_2+d_1^{'}<c_1<4c_2+d_1^{'}c_2+d_1^{'} \tag{B.4}$$

  Now if c_{1+}^* meets (B.4), the upper bound on c_1 is $d_1^{'}-3d_1^{'}c_2+2\sqrt{2(c_2-c_2^2)(1-d_1^{'2})}$, otherwise this upper bound is given by $c_1<1+c_2$, and similarly if c_{1-}^* meets (B.4) the lower bound on c_1 is $d_1^{'}-3d_1^{'}c_2-2\sqrt{2(c_2-c_2^2)(1-d_1^{'2})}$, otherwise this lower bound is given by $c_1>-c_2-1$. This ends the proof.

Appendix C. Stability of the system using the dynamic adaptation gain

Consider Eq. (45). Using the $[A, B, C, D]$ state space representation associated to $H_{PAA}(z)$ one gets:

$$x(t+1) = Ax(t) + B\mathbf{r}(t)\epsilon(t+1) \tag{C.1}$$

$$\tilde{\mathbf{w}}(t+1) = Cx(t) + D\mathbf{r}(t)\nu(t+1) \tag{C.2}$$

and respectively:

$$\mathbf{r}^T(t)\tilde{\mathbf{w}}(t+1) = \mathbf{r}^T(t)Cx(t) + \mathbf{r}^T(t)D\mathbf{r}(t)\nu(t+1) \tag{C.3}$$

Eqs. (C.1), and (C.3) define an equivalent feedback system, the equivalent feedback path being defined by (C.1) and (C.3). In order to apply passivity theorem the equivalent feedback path should be passive. One can use Theorem 3.3.1 from [6]:

Theorem 1. *For a PAA having the form of Eqs* (C.1) *and* (C.2)*, the equivalent feedback path described by Eqs.* (C.1)*, and* (C.3) *is passive, i.e.,*

$$\eta(0, t_1) = \sum_{t=0}^{t_1} \nu(t+1)\mathbf{r}^T(t)\tilde{\mathbf{w}}(t+1) \geq -\gamma^2 \; ;$$

$$\gamma^2 < \infty \; \textit{for all} \; t \geq 0 \tag{C.4}$$

if the associated linear system $[A, B, C, D]$ *described by Eqs.*(C.1) *and* (C.2) *is passive, or equivalently, if* $H_{PAA}(z)$ *given in Eq.* (45) *is a PR transfer matrix.*

Since the equivalent path is passive, Using Theorem C2 from [6] one conclude that if the equivalent feedforward path is characterized by a SPR transfer function, one has that $\lim_{t\to\infty} \nu(t+1) = 0$ for any initial conditions.

Appendix D. Proof of PR conditions for H_{PAA}

A first necessary condition is given by condition (48) assuring that the poles of the transfer function H_{PAA} are inside or on the unit circle. By performing a partial fraction expansion of H_{PAA}, one has:
$H_{PAA}(q^{-1}) = 1 + \frac{\delta q^{-1}}{1-q^{-1}} + \frac{\gamma q^{-1}}{1-d_1' q^{-1}}$. Set β such that $\beta \in]0, 1[$, one can write $H_{PAA}(q^{-1}) = H_1(q^{-1}) + H_2(q^{-1})$ with $H_1(q^{-1}) = \beta + \frac{\delta q^{-1}}{1-q^{-1}} = \beta \frac{1-\frac{\beta-\delta}{\beta} q^{-1}}{1-q^{-1}}$ and $H_2(q^{-1}) = (1-\beta) + \frac{\gamma q^{-1}}{1-d_1' q^{-1}} = (1-\beta)\frac{1-\left(d_1' - \frac{\gamma}{1-\beta}\right)q^{-1}}{1-d_1' q^{-1}}$. Since H_1, H_2 are first order transfer function operators, and since $\beta > 0$, $1-\beta > 0$, a sufficient condition for H_1 and H_2 to be both PR is that their zeros be inside or on the unit circle. This is assured if the two following conditions are met simultaneously

$$-1 \leq \frac{\beta - \delta}{\beta} \leq 1 \qquad (a)$$

$$-1 \leq d_1^{'} - \frac{\gamma}{1-\beta} \leq 1 \qquad (b)$$

If (a) and (b) are met at the same time, there exists at least one $\beta \in]0,1[$ such that the two conditions (a) and (b) are met at the same time, and this value is β_0 the smallest value of β such that $\mathrm{Re}(H_1(e^{i\omega})) \geq 0 \quad \forall\omega$. Condition (a) will be satisfied if condition (49) is met. One has therefore $0 \leq \delta \leq 2$ and $\beta \in [\frac{\delta}{2}, 1[$. Moreover, the function $f(\beta) = d_1^{'} - \frac{\gamma}{1-\beta}$ is monotone for $\beta \in [\frac{\delta}{2}, 1[$. If there exists only one value of β such that condition b) is satisfied, this value is necessarily $\delta/2$ since $-1 < d_1^{'} < 1$. Hence condition (50).

These conditions are also necessary: Let us assume that the condition $\frac{\beta-\delta}{\beta} \leq 1$ is violated, since $\beta > 0$ one has $\delta < 0$. By definition $\delta = \frac{1+c_1+c_2}{1-d_1^{'}}$ and $1 - d_1^{'} > 0$. That leads to $1 + c_1 + c_2 < 0$. But from the Jury criterion [24], a necessary and sufficient condition for $1+c_1q^{-1}+c_2q^{-2}$ be a stable polynomial is that at the same time the three following conditions $|c_2| < 1$, $1 + c_1 + c_2 > 0$, $1 - c_1 + c_2 > 0$ are verified. Therefore if $\frac{\beta-\delta}{\beta} > 1$, one has $1 + c_1 + c_2 < 0$ and $1 + c_1q^{-1} + c_2q^{-2}$ cannot be stable, and H_{PAA} cannot be positive real. Similarly if $\frac{\beta-\delta}{\beta} \leq -1$, one has $\delta \geq 2$ and since $1 - d_1^{'} > 0$ there exists some values of $d_1^{'}$ such that $1 + c_1 + c_2 > 2$: for $c_1 = 0$ this implies $c_2 > 1$ which is not compatible with the first condition of the Jury criterion, thus in this case $1 + c_1q^{-1} + c_2q^{-2}$ cannot be stable and H_{PAA} cannot be positive real.
This ends the proof.

References

[1] S. Elliott, Signal processing for active control, Academic Press, San Diego, California, 2001.

[2] I. D. Landau, T.-B. Airimitoaie, A. Castellanos Silva, A. Constantinescu, Adaptive and Robust Active Vibration Control—Methodology and Tests, Advances in Industrial Control, Springer Verlag, 2017.

[3] I. D. Landau, M. Alma, T.-B. Airimitoaie, Adaptive feedforward compensation algorithms for active vibration control with mechanical coupling, Automatica 47 (10) (2011) 2185 – 2196. doi:10.1016/j.automatica.2011.08.015.

[4] I. D. Landau, T.-B. Airimitoaie, M. Alma, IIR Youla–Kučera parameterized adaptive feedforward compensators for active vibration control with mechanical coupling, IEEE Transactions on Control System Technology 21 (3) (2013) 765–779.

[5] B. Huang, Y. Xiao, J. Sun, G. Wei, A variable step-size fxlms algorithm for narrowband active noise control., IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING 21 (2) (2013) 301–312.

[6] I. D. Landau, R. Lozano, M. M'Saad, A. Karimi, Adaptive control, 2nd Edition, Springer, London, 2011.

[7] I. Landau, T. Airimitoaie, Does a general structure exist for adaptation/learning algorithms?, in: Proceedings of the IEEE CDC Conference,Dec.2022,Cancun, Mexico, 2022, pp. 2650–2655.

[8] R. Melendez, I. Landau, L. Dugard, G. Buche, Data driven design of tonal noise feedback cancellers, in: Proceedings of the 20th IFAC World Congress, Toulouse, France, 2017, pp. 916–921.

[9] J. Zeng, R. de Callafon, Recursive filter estimation for feedforward noise cancellation with acoustic coupling, Journal of Sound and Vibration 291 (3-5) (2006) 1061 – 1079. doi:10.1016/j.jsv.2005.07.016.

[10] R. Fraanje, Robust and fast schemes in broadband active noise and vibration control, Ph.D. thesis, University of Twente, Twente, The Netherlands (2004).

[11] J. Hu, J.-F. Lin, Feedforward active noise controller design in ducts without independent noise source measurements, IEEE Trans. on Control System Technology 8 (3) (2000) 443–455.
[12] I. D. Landau, T.-B. Airimitoaie, M. Alma, A Youla–Kučera parametrized adaptive feedforward compensator for active vibration control with mechanical coupling, Automatica 48 (9) (2012) 2152 – 2158. doi:http://dx.doi.org/10.1016/j.automatica.2012.05.066.
[13] I. D. Landau, R. Melendez, T. Airimitoaie, L. Dugard, Robust and adaptive feedback noise attenuation in ducts, Journal of Sound and Vibration 455 (2019) 339–358.
[14] T.-B. Airimitoaie, I. D. Landau, R. Melendez, L. Dugard, Algorithms for Adaptive Feedforward Noise Attenuation–A Unified Approach and Experimental Evaluation, IEEE Transactions on Control Systems Technology 29 (5) (2021) 1850–1862. doi:10.1109/TCST.2020.3023276.
URL `https://hal.archives-ouvertes.fr/hal-02947816`
[15] T.-B. Airimitoaie, I. D. Landau, Improving adaptive feedforward vibration compensation by using integral+proportional adaptation, Automatica 49 (5) (2013) 1501–1505. doi:10.1016/j.automatica.2013.01.025.
[16] I. Landau, Adaptive control : the model reference approach, Marcel Dekker, New York, 1979.
[17] M.Schmidt, Stochastic average gradient, University of British Columbia. (Winter 2018).
[18] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. P. Reyes, M.-L. Shyu, S.-C. Chen, S. S. Iyengar, A survey on deep learning: Algorithms, techniques, and applications, ACM Comput. Surv. 51 (5) (sep 2018). doi:10.1145/3234150.
[19] I. Livieris, R. Pintelas, A survey on algorithms for training artificial neural networks, University of Patras Report 455 (TR08-01) (2014).
[20] J. E. Gaudio, T. E. Gibson, A. M. Annaswamy, M. A. B. E. Lavretsky, Connections between adaptive control and optimization in machine learning, in: Proceedings of the IEEE CDC Conference,Dec.2019, Nice, France, 2019, pp. 4563–4568.
[21] R. Jacobs, Increased rates of convergence through learning rate adaptation., Neural Networks 27 (1) (1988) 295–307.
[22] B. Anderson, R. Bitmead, C. Johnson, P. Kokotovic, R. Kosut, I. Mareels, L. Praly, B. Riedle, Stability of adaptive systems, The M.I.T Press, Cambridge Massachusetts , London, England, 1986.
[23] H. Bourlès, Linear Systems, Wiley, 2010.
[24] E. Jury, J. Blanchard, A stability test for linear discrete time systems in table form, Proceedings IRE . 49 (1961) 1947–1948.