Adaptive Recursive Deconvolution and Adaptive Noise Cancellation

Pandolfi, L.[*]
Politecnico di Torino
Dipartimento di Matematica
Corso Duca degli Abruzzi, 24
10129 Torino — Italy
Tel. +39-11-5647516
E–Mail LUCIPAN@POLITO.IT

Abstract

In this paper we apply a recursive deconvolution method to Active Noise Cancellation (ANC) in a linear system: the observation of the output of a linear system of relative degree one, red at discrete time instants, is fed to a deconvolution algorithm which identify the disturbance (with the delay of one step). This information is used in order to reduce the effect of the disturbance itself. Deconvolution being an ill posed problem, a regularization parameter is to be introduced. The choice of the value of the parameter is a delicate issue. We show that, when studying ANC, the discrepancy principle (applied recursively) is a feasible method for the choice of the parameter.

Key words: Active noise cancellation, linear systems, recursive deconvolution.

1 Introduction

Deconvolution problems are ubiquitous in the applications of mathematics and have many different aspects. Here we distinguish between "off line" and "on line" i.e. "recursive" deconvolution algorithms. Off line deconvolution algorithms accumulate all the available pieces of information which are then elaborated at the end of the process. Recursive deconvolution instead evaluates an estimates $\hat{u}$ of the input function u in real time (or with a delay

[*]This paper fits the research program of GNAMPA.

which in principle can be made as small as wanted). Only algorithms of this second class can be used for control or regulation, see [2].

The difficulty of the deconvolution problem stems from the fact that it is "ill posed", i.e. the solution is not a continuous function of the data. Hence, "regularization" methods have to be used, which provide approximate solutions continuously depending on the available data. These regularization methods do depend on the introduction of a new "regularization parameter" which we shall denote α. Moreover, in practice on line computations are performed at discrete time instants so that a time step τ has to be introduced as a second parameter of the algorithm which then depends on the two parameters α and τ. The parameter τ can be seen as a second regularization parameter, see the formulas (8) and (10). A third parameter which enters the algorithm is the tolerance h of the errors of the measures.

Recursive deconvolution algorithms have been proposed for example in [7, 8, 11] and these algorithms can be used in systems theory.

An interesting approach to "on line" deconvolution is described in [7]. A modification of this algorithm, which is a recursive application of Tikonov regularization, is studied in details in [3] and it is described below. This algorithm depends on the three parameters mentioned above, the regularization parameter α, the error tolerance h and the time step τ and constructs a candidate approximants $\hat{u}$ of the unknown input u which converges to u provided that α, h and τ converges to zero while respecting certain "compatibility conditions" (which depends on the type of convergence which is studied) and which require in particular that α fades away slower then h and τ.

In general, the parameters α and τ should be "small" in order to guarantee fast convergence but in this case both the noise and the round off errors are amplified so that if α and τ are too small with respect to h then wild oscillations appear which may prevent every reasonable use of the algorithm.

In the case of "off line" reconstruction algorithms, several methods have been proposed for an acceptable choice of the regularization parameters, see [6]. The most commonly used is the "discrepancy principle" introduced by Morozov, which essentially requires that the signal produced by $\hat{u}$ should approximate the observed signal (produced by u) so that the error be of the order of (a multiple of) h. *So, the parameters α and τ of the algorithm are deemed acceptable when the output produced by $\hat{u}$ is "close" to the measures taken on the real output. This is precisely the criteria to be used for noise cancellation, the problem we are going to study.*

We note that quality of the reconstruction of the output is not the sole parameter to be considered in practical problems. For example, in practice fast oscillations should be avoided. This goal could be achieved by penalizing also the derivatives. However, in this paper we are not going to consider additional goals like this.

1.1 Deconvolution and Active Noise Control

A recent approach proposed to perform noise reduction has been described as follows: the plant is subject to a known input f_0, corrupted by the noise ω, and to a control u. The observed output y is filtered so to reconstruct an approximation $\hat{\omega}$ of the disturbance (this is the "deconvolution" problem) which is then fed back to the system (with a "small" delay) so to approximately cancel the effect of the noise. This approach has been successfully used, often on a purely intuitive basis, in ANC (Active Noise Cancellation), see [4, 13] with interesting applications to mobile communications.

Most of the algorithms used in ANC resembles those studied in [2, 3] where and the application of the method has been completely justified. In this context, the regularization parameter α has the role of a high gain: if it is too small then disturbances and round off errors in computations are amplified.

The choice of the regularization parameters is most delicate since it must take into account conflicting goals. Often it is made on the basis of off-line experience, on the basis of the "a priori" information on the disturbances to be canceled. In this case α is chosen before the process starts and it is kept constant in time. Instead, an on-line version of the discrepancy principle for the adaptive choice of α has been proposed in [9, 10]. In this paper we are going to show that the idea of an adaptive choice of the regularization parameter α can be used to solve ANC for systems of relative degree one (the reason of this restriction will be discussed in section 2.4.1.) The time step τ is still to be "a priori" chosen.

We note that when the regularization parameter α is chosen in an adaptive way then it will not be constant in time: the "regularization parameter" is now a function.

For future reference we explicitly note:

Remark 1 The reason for an on-line determination of the regularization parameter α, based on the discrepancy principle, is now explained. If sufficient information on the class of possible errors and disturbances is available, off-line experience may show that very small values of α and τ can be used so to achieve a very accurate reconstruction of the disturbance. However, unexpected errors or disturbances can destabilize the system. *The goal of the adaptive determination of α is a more robust and stable algorithm, at the expenses of the fidelity of the reconstruction.* ∎

The plan of the paper is as follows: we shall study linear finite dimensional systems, possibly time varying. In section 2 (and subsections) we study the deconvolution problem in details. The obtained results are then applied to the ANC in Section 3. Simulations of a realistic case, from [2, 14], are given in Section 4.

The structure of Section 2 is as follows: the algorithm is described first in the (nonrealistic) case of the full state observation and the technical proofs are in subsection 2.2. Subsection 2.4 shows that the algorithm can be extended to systems of relative degree one and discuss the reasons of this limitation.

2 The algorithm

As we said, the algorithm we use has been precisely studied (with α fixed once and for all the times) in [3].

We shall assume that the unknown input signal is bounded but knowledge of the value of the bound of the input is not required. It is only used in the proofs that this bound is finite.

The system which elaborates the signal is the input-output control system

$$\dot{x} = Ax + Bu\,, \qquad 0 \le t \le T \tag{1}$$

where x and u are respectively n and m–vectors while A and B denotes matrices of suitable dimensions, which are continuous function of time t (the matrices are often constant).

The system is initialized at zero, $x(0) = 0$ so that

$$x(t) = \int_0^t X(t,s)B(s)u(s)\,\mathrm{d}s \tag{2}$$

where $X(t,s)$ is the evolution matrix generated by A. If A is constant, $X(t,s) = e^{A(t-s)} = X(t-s)$ and the input-output relation (2) is of convolution type.

We assume that the observation is red at discrete times and contaminated by errors of known tolerance h. Hence available data are

$$\xi_k = Cx(\tau_k) + \theta_k\,, \qquad |\theta_k| < h\,.$$

It is not restrictive to assume $h < 1$.

It will be $C = I$ in subsections 2.1 and 2.2.

The time instants τ_k are equispaced for simplicity, $\tau_k = k\tau$.

We put

$$I_k = [\tau_k, \tau_{k+1})\,, \qquad \tau_k = k\tau\,.$$

Our goal is the construction of a signal v such that:

- the vector $v(t)$ depends on the available data, the fixed step τ of the observations and a penalization parameter α which may change at every step;

- the vector $v(t)$ at time t is constructed only using the data ξ_k available at previous time instants;

- we want that when $h \to 0$, $\tau \to 0$ then automatically $\alpha \to 0$ and in such a way that

$$||v - v_*||_{L^2(0,T)} \to 0\,, \qquad \text{where } v_* \text{ satisfies } Bv_* = Bu.$$

We are going to describe an algorithm for the online determination of α.

2.1 Full state observation

In this subsection we assume $C = I$.

In practical applications, in particular applications to control theory, the full state x is rarely measured. In general, only an output $y(t) = Cx(t)$ is measured. The study of the case $C = I$ is instrumental for understanding the general case.

Clearly, two different inputs $u_1(t)$ and $u_2(t)$ might produce the same evolution $x(t)$ of the state. This happens when their difference belongs to $\ker B(t)$ a.e. $t \in (0, T)$. So we can replace Eq. (2) with the new input-output relation

$$\dot{x} = Ax + b\,, \qquad x(0) = 0\,, \qquad \xi_k = x(\tau_k) + \theta_k \tag{3}$$

and we can approximate the function $b(t) = B(t)u(t)$ with a function $v(t)$. After that, the input $\hat{u}(t) = B^\dagger v(t)$ can be singled out, even if in general $v(t) \notin \mathrm{im}B(t)$.

Remark 2 When $n - m$ is large, as it usually is, the practical computation of $B^\dagger v$ will require the use of a penalization method. This is a standard procedure in the computation of the pseudoinverse of a matrix (see [5]) and it is not considered here. ∎

As in [3, 7], we associate a "model" to our system

$$\dot{w} = Aw + v\,, \qquad 0 \le t \le T\,. \tag{4}$$

Now we initialize the reconstruction algorithm as follows: we fix the sampling time τ and we fix two exponents γ and ϵ, $0 < \epsilon < \gamma < 1$. Moreover, we fix a coefficient $\mu \ge 1$, whose role is discussed in Remark 3.

Now we describe the operation to be performed at every time instant $\tau_k = k\tau$ in order to construct the function $v(t)$. The notation $v_{(k)}$ is used for the restriction of v to the interval I_k,

$$v_{(k)}(t) = v(t)\,, \qquad t \in I_k = [\tau_k, \tau_{k+1})\,.$$

The operations are recursively described as follows:

i) at time $\tau_0 = 0$ we initialize Eq. (4) with $w(0) = 0$.

We impose the condition

$$\tilde{v}_{(0)} = \arg\min\left\{ \left|\left| \int_0^\tau X(\tau, s)v(s)\, \mathrm{d}s - \xi_1 \right|\right|^2 + \alpha \int_0^\tau ||v(s)||^2\, \mathrm{d}s \right\}. \quad (5)$$

This function depends on the still unspecified value α.

We compute

$$\tilde{w}(t) = \int_0^t X(t,s)\tilde{v}_{(0)}(s)\, \mathrm{d}s\,, \qquad t \in I_0\,. \quad (6)$$

Also this function depends on the still unspecified value α.

We compare $\tilde{w}(\tau)$ and ξ_1 and we apply the **algorithm** described below so to identify the value α_0 of α, to be used on I_0. The function $\tilde{v}_{(0)}$ which corresponds to this value of α is denoted $v_{(0)}$ while $w(t)$ is the corresponding function, given by (6), when $\tilde{v}$ is replaced by $v_{(0)}$.

ii) Once we have reached time $\tau_k = k\tau$, the functions $v(t)$ and $w(t)$ have been computed on $[0, \tau_k)$ and we proceed as follows in order to extend $v(t)$ to the next time interval I_k. First we compute

$$\begin{aligned}\tilde{v}_{(k)} \quad = \quad & \arg\min\left\{ \left|\left| X(\tau_{k+1}, \tau_k)w(\tau_k) + \int_{\tau_k}^{\tau_{k+1}} X(\tau_{k+1}, s)v(s)\, \mathrm{d}s - \xi_{k+1} \right|\right|^2 \right. \\ & \left. +\alpha \int_{\tau_k}^{\tau_{k+1}} ||v(s)||^2\, \mathrm{d}s \right\}. \qquad (7)\end{aligned}$$

The value to be used on I_k for the penalization parameter α has not yet been specified.

We compute now $\tilde{w}(\tau_{k+1})$ using this function $\tilde{v}_{(k)}$, which we compare with ξ_{k+1} and we apply the **algorithm** (described below) so to identify the value α_k of α, to be used on I_k. The function $v_{(k)}$ which corresponds to this value of α is denoted $v_{(k)}$ and we extend w to I_k:

$$w(t) = X(t, \tau_k)w(\tau_k) + \int_{\tau_k}^t X(t,s)v_{(k)}(s)\, \mathrm{d}s\,, \qquad t \in I_k\,.$$

We note that ξ_{k+1} is available at the time τ_{k+1} so that only at this time it will be possible to define $v_{(k)}$: the function $v(t)$ which approximate $b(t)$ is reconstructed with one step delay.

Now the **algorithm** for the choice of the value of α at each step, inspired by the discrepancy principle, see [5, 12], is as follows:

Algorithm: at the time step τ_{k+1} we compare $w(\tau_{k+1})$ and ξ_{k+1}. If it turns out that *for every* $\alpha > 0$ we have

$$||w(\tau_{k+1}) - \xi_{k+1}|| \le \mu h^{\gamma}$$

then the value of α on I_k is

$$\alpha = \mu h^{\gamma - \epsilon} \,.$$

Otherwise, we shall see the existence of a unique value of α such that

$$||w(\tau_{k+1}) - \xi_{k+1}|| = \mu h^{\gamma} \,.$$

This is the value α_k of α we shall use on I_k. ∎

Remark 3 First of all, in order to clarify the **algorithm,** we anticipate (see [5]) that

$$\alpha \longrightarrow ||\alpha[\alpha I + R_k]^{-1}[\Gamma_k w_k - \xi_{k+1}]||$$

is increasing with limits equal to zero for $\alpha \to 0+$ and $||[\Gamma_k w_k - \xi_{k+1}]||$ for $\alpha \to +\infty$. Hence we have

$$||w_{k+1} - \xi_{k+1}|| \le \mu h^{\gamma} \text{ for all } \alpha \text{ iff } ||[\Gamma_k w_k - \xi_{k+1}]|| \le \mu h^{\gamma} \,.$$

In this case $\alpha_k = \mu h^{\gamma-\epsilon}$. Otherwise there exists only one value of α for which

$$||w_{k+1} - \xi_{k+1}|| = \mu h^{\gamma} \,.$$

The role of μ is now clear: it represent the additional error that it is admissible in the reconstruction process.

A last observation is that the algorithm we presented penalizes the L^2 norm of the input v to the model system. We could add a penalization of the incremental quotient $[v(\tau_{k+1}) - v(\tau_k)]/h$ in order to reduce the oscillations in the reconstructed signal, hence also in the regulated output. As we said already we are not going to consider this problem here. ∎

We shall prove the following consistency result:

Theorem 4 *Let u be bounded and $T < +\infty$. We relate τ to h so that $\tau \to 0$ when $h \to 0$, in such a way that*

$$\frac{h}{\tau} \to 0\,, \qquad \frac{\tau}{h^\gamma} \to 0\,.$$

Under these conditions, we have that v converges to b in $L^2(0,T)$.

As a consequence, $\hat{u} = B^\dagger v$ converges to the input of minimal L^2 norm which produce the evolution $x(t)$ of the state.

This is a consistency result, whose conditions are satisfied if we choose, for example, $\tau = h^\sigma \tau_0$, τ_0 fixed and $\gamma < \sigma < 1$ but, of course, in practice neither $h \to 0$ nor $\tau \to 0$ and we must content ourselves with suitably small values of them.

Remark 5 The case $T = +\infty$ is of utmost importance for the applications. Theorem 4 can be reformulated also when $T = +\infty$ under additional (stability) conditions. See section 2.3 for this case. In order to be able to treat the case $T = +\infty$, we keep track of all the constants in the proof of Theorem 4, even if this leads to a bit involved formulas. ∎

2.2 The proof of the consistency theorem

It is convenient to introduce the following notations. We denote w_k, x_k the vectors $w(\tau_k)$, $x(\tau_k)$ and we introduce the following operators: the operator Λ_k acts from $L^2(I_k)$ to $\mathbb{R}^n$,

$$\Lambda_k u = \int_{\tau_k}^{\tau_{k+1}} X(\tau_{k+1}, s) u(s)\, \mathrm{d}s$$

so that

$$(\Lambda_k{}^* x)(s) = X^*(\tau_{k+1}, s)x\,, \qquad \Lambda_k \Lambda_k{}^* = \int_{\tau_k}^{\tau_{k+1}} X(\tau_{k+1}, s) X^*(\tau_{k+1}, s)\, \mathrm{d}s = R_k\,.$$

The operator R_k does not depend on k if A is constant.

The operator Γ_k: $\mathbb{R}^n \to \mathbb{R}^n$ is given by

$$\Gamma_k w = X(\tau_{k+1}, \tau_k) w\,.$$

In this section $k \le N = E[T/\tau] + 1$ where $E[\cdot]$ denotes the integer part. Hence there exist constants m, M which do not depend on k and such that

$$\begin{cases} ||\Gamma_k w|| \le (1 + M\tau)||w|| \\ \tau(1 - m\tau)I \le R_k \le \tau(1 + M\tau)I \end{cases} \tag{8}$$

where I denotes the identity matrix.

The upper bound of the unknown input u has no role in the algorithm we described in the previous section and it needs not be known for the implementation of the algorithm. However, it will be used in the proofs that this upper bound is finite. More precisely we introduce the *numbers:*

$$\mathcal{B}_T = \sup_{t\in[0,T]} ||b(t)|| \,, \qquad \mathcal{B}_k = \sup_{t\in I_k} ||b(t)|| \,. \tag{9}$$

The following formulas are easily derived:

$$\left\{\begin{array}{lcl} \tilde{v}_{(k)} & = & -[\alpha I + \Lambda_k{}^*\Lambda_k]^{-1}\Lambda_k{}^*[\Gamma_k w_k - \xi_{k+1}] \,, \\ & = & -\Lambda_k{}^*[\alpha I + R_k]^{-1}(\Gamma_k w_k - \xi_{k+1}) \,, \\ \tilde{w}_{k+1} & = & \alpha[\alpha I + R_k]^{-1}\Gamma_k w_k + [\alpha I + R_k]^{-1} R_k \xi_{k+1} \,, \\ \tilde{w}_{k+1} - \xi_{k+1} & = & \alpha[\alpha I + R_k]^{-1}[\Gamma_k w_k - \xi_{k+1}] \,. \end{array}\right. \tag{10}$$

We note that in these formulas the use of ˜ is consistent with the online determination of α: at the time τ_k we already know the values of α, v and w for $t \le \tau_k$; hence, ˜ does not appear on the right hand side of the formulas. Instead, α is to be determined on I_k so that the ˜ notation is used on the left.

Remark 6 It is seen from (8) that $||(\alpha I + R_k)^{-1}||$ is of the order of $1/(\alpha + m\tau)$ so that the last addendum of w_{k+1} is of the order of $h/(\alpha + m\tau)$. In order to avoid excessive amplification of the noise (and round off errors) both α and τ should not be too small. ∎

A preliminary computation is as follows. Here $b_{(k)}$ is the restriction of b to I_k so that

$$||\Lambda_k b_{(k)}|| \le \tau(1 + M\tau)\mathcal{B}_k \,.$$

$$\begin{array}{l} ||(\alpha I + R_k)^{-1}\{\Gamma_k w_k - \xi_{k+1}\}|| \le \\ ||(\alpha I + R_k)^{-1}\Gamma_k(w_k - \xi_k)|| + ||(\alpha I + R_k)^{-1}\{\Gamma_k\theta_k - \theta_{k+1}\}|| + \\ ||(\alpha I + R_k)^{-1}\Lambda_k b_{(k)}|| \le \\ \dfrac{(1 + M\tau)||w_k - \xi_k|| + [(2 + M\tau)h + \tau(1 + M\tau)\mathcal{B}_k]}{\alpha + \tau(1 - m\tau)} \,. \end{array} \tag{11}$$

We first prove that our **algorithm** gives a lower bound for α:

Lemma 7 *Let α_k be the value of α produced by the* **algorithm** *at the instant τ_k. We have either $\alpha = \mu h^{\gamma-\epsilon}$ or*

$$\alpha_k \ge \frac{\mu h^\gamma}{\mathcal{B}_k + \tilde{M}(h^\gamma + h/\tau)}(1 - m\tau) \,. \tag{12}$$

The number $\tilde{M}$ depends neither on τ nor on k. See Remark 8 for the dependence of $\tilde{M}$ on T.

Proof. In the first interval,

$$||w_1 - \xi_1|| = ||\alpha[\alpha I + R_1]^{-1}\xi_1|| < h < \mu h^\gamma$$

so that $\alpha_0 = \mu h^{\gamma-\epsilon}$.

Let the assertion be true at $k-1$ and let us see what's going on at the next step. We noted that

$$\alpha \longrightarrow ||\alpha[\alpha I + R_k]^{-1}[\Gamma_k w_k - \xi_{k+1}]|| = ||\tilde{w}_{k+1} - \xi_{k+1}||$$

is increasing with limits equal to zero for $\alpha \to 0+$ and $||[\Gamma_k w_k - \xi_{k+1}]||$ for $\alpha \to +\infty$. Hence either

$$||w_{k+1} - \xi_{k+1}|| \le \mu h^\gamma \qquad \forall\alpha\,,$$

so that $\alpha_k = \mu h^{\gamma-\epsilon}$, or there exists only one value of α for which

$$||w_{k+1} - \xi_{k+1}|| = \mu h^\gamma\,.$$

In both the cases, we have $||w_k - \xi_k|| \le \mu h^\gamma$ and, from (11) the chosen value of α gives

$$\begin{aligned} &||w_{k+1} - \xi_{k+1}|| = \mu h^\gamma = ||\alpha(\alpha I + R_k)^{-1}[\Gamma_k w_k - \xi_{k+1}]|| \\ &\le \frac{\alpha}{\alpha + \tau(1 - m\tau)}\left\{(1 + M\tau)||w_k - \xi_k|| + (2 + M\tau)h + \tau(1 + M\tau)\mathcal{B}_k\right\}\,. \end{aligned}$$

We noted that $||w_k - \xi_k|| \le \mu h^\gamma$. Hence we have

$$\alpha \ge \frac{\mu h^\gamma(1 - m\tau)}{M\mu h^\gamma + (2 + M\tau)(h/\tau) + (1 + M\tau)\mathcal{B}_k}$$

as wanted. ∎

Note that if τ is too small (with respect to h) then this lower bound for α is small too.

Remark 8 The number $\tilde{M}$ in (12) is explicitly given by

$$\tilde{M} = \max\{2,\ M\mu + M\mathcal{B}_k\frac{\tau}{h^\gamma} + Mh^{1-\gamma}\}\,.$$

Our assumption is that $\tau/h^\gamma \to 0$, $h \to 0$ so that when $T < +\infty$ we can choose (for h and τ/h^γ small enough)

$$\tilde{M} = \min\{2, M(1 + \mu + \mathcal{B}_k)\}\,.$$

If instead $T = +\infty$ it might be impossible to chose a finite value for $\tilde{M}$. We find that $\tilde{M} < +\infty$ if $T = +\infty$ when u is bounded and $X(t,s)$ is bounded in the angle $0 \le s \le t$. This is a stability condition. ∎

We observe that the estimate from below on α implies that

$$\frac{1}{\alpha+\tau(1-m\tau)} \le \frac{1}{\mu h^{\gamma-\epsilon}+\tau(1-m\tau)} \tag{13}$$

when $\alpha = \mu h^{\gamma-\epsilon}$. Otherwise we have

$$\frac{1}{\alpha+\tau(1-m\tau)} \le \frac{\mathcal{B}_k+\tilde{M}(h^\gamma+h/\tau)}{\mu h^\gamma(1-m\tau)}\,. \tag{14}$$

These estimates are now used to study the convergence of v in $L^2(0,T)$.

We prove first boundedness of the set of the functions v constructed by the algorithm. We use

$$v_{(k)}(s) = X^*(\tau_{k+1},s)(\alpha I+R_k)^{-1}[\Gamma_k w_k-\xi_{k+1}]\,,$$

the decomposition (11) and the estimate (12) so to obtain

$$\begin{aligned}&||v_{(k)}(t)|| \le \frac{1+M\tau}{\alpha+\tau(1-m\tau)}\left\{||\Gamma_k(w_k-\xi_k)||+||\Gamma_k\theta_k-\theta_{k+1}||+||\Lambda_k b_{(k)}||\right\}\\ &\le \frac{1+M\tau}{\alpha+\tau(1-m\tau)}\left\{(1+M\tau)\mu h^\gamma+(2+M\tau)h+\tau(1+M\tau)\mathcal{B}_k\right\}\,.\end{aligned}$$

When (14) holds we have

$$||v_{(k)}(t)|| \le \frac{\mathcal{B}_k}{1-m\tau}+\mathrm{o}(1) \tag{15}$$

where $\mathrm{o}(1)$ denotes

$$\begin{aligned}&\frac{M\tau\mathcal{B}_k}{1-m\tau}+\frac{\tilde{M}(h^\gamma+(h/\tau))(1+M\tau)}{1-m\tau}\\ &+\frac{\mathcal{B}_k+\tilde{M}(h^\gamma+h/\tau)}{\mu h^\gamma(1-m\tau)}\left[(2+M\tau)h+\tau(1+M\tau)\mathcal{B}_k\right]\,.\end{aligned}$$

Otherwise, when inequality (13) holds, we have

$$||v_k(t)|| \le \mathrm{o}(1) \tag{16}$$

and

$$\begin{aligned}\mathrm{o}(1) = {}&\frac{1+M\tau}{\mu+(\tau/h^\gamma)h^\epsilon(1-m\tau)}\,\cdot\\ &\cdot\left\{(1+M\tau)h^\epsilon+(2+M\tau)h^{1-\gamma+\epsilon}+(\tau/h^\gamma)h^\epsilon(1+M\tau)\mathcal{B}_k\right\}\,.\end{aligned}$$

So, boundedness of the function b implies boundedness of the set of the functions v constructed by the algorithm, both uniformly and in $L^2(0,T)$.

We note that v depends on τ and the errors θ, $|\theta| < h$. We choose arbitrary sequences $\{h_n\}$, $\{\tau_n\}$ which converges to zero while satisfying the conditions in Theorem 4 and any sequence of errors θ, of tolerance h. For each value of n the algorithm gives a function v on $[0, T]$, which we denote v^n. Let $\{w^n\}$ be the corresponding sequence produced by (4).

We already noted that $\{v^n\}$ is bounded. We are going to prove that $\{v^n\}$ converges to b in $L^2(0, T)$.

Boundedness of $\{v^n\}$ implies equicontinuity of the bounded sequence $\{w^n\}$. From this and the condition

$$||w^n(\tau_k) - \xi_k|| \leq \mu h^\gamma$$

it follows that

$$\lim w^n = x$$

uniformly on $[0, T]$.

Consequently the set $\mathcal{V}$ has a unique weak limit point, which is b.

In order to prove that b is also the limit in the norm topology we use a standard property of the Hilbert spaces: the conditions

$$v^n \rightharpoonup b\,, \qquad ||v^n|| \to ||b||$$

imply that $\{v^n\}$ converges to b in norm. So, it remains to be proved

$$||v^n||_{L^2(0,T)} \to ||b||_{L^2(0,T)}.$$

We note that, by lower weak semicontinuity of the norm,

$$||b||_{L^2(0,T)} \leq \liminf ||v^n||_{L^2(0,T)}$$

while, using the estimates (15) and (16),

$$||v^n||^2_{L^2(0,T)} \leq \tau \sum_{k=0}^{N-1} ||v^n_{(k)}||^2_{L^2(I_k)} \leq \tau \sum_{k=0}^{N-1} \left[\frac{\mathcal{B}_k}{1 - m\tau} + \mathrm{o}(1) \right]^2 \leq \left\{ \mathrm{o}(1) + \tau \sum_{k=0}^{N-1} \mathcal{B}_k^2 \right\}.$$

Hence we have

$$\begin{aligned} &||b||^2_{L^2(0,T)} \leq \liminf ||v^n||^2_{L^2(0,T)} \leq \limsup ||v^n||^2_{L^2(0,T)} \\ &\leq \limsup \left\{ \mathrm{o}(1) + \tau \sum_{k=0}^{N-1} \mathcal{B}_k^2 \right\} = ||b||^2_{L^2(0,T)} \end{aligned}$$

so that we also have

$$\lim v = b \qquad \text{in } L^2(0, T). \quad \blacksquare$$

This proves the consistency of the algorithm. We repeat that a suitable choice for τ which satisfies the assumptions of the theorem is for example $\tau = h^\sigma \tau_0$, τ_0 "a priori" given. The determination of the coefficient τ_0 is a matter of practice.

2.3 Infinite horizon problems

In every practical application T is finite, but often not "a priori" known. In order to include also this case in the theory, it is usual to study the case $T = +\infty$, a case we are going to consider now.

We observe that the condition that u be square integrable on $[0, +\infty)$ is not realistic, in particular when u is a disturbance, which does not fades with time. A realistic assumption is that u is bounded and every bounded (measurable) u belongs to the Hilbert space K whose elements are the (equivalence classes of) functions such that

$$||u||^2 = \lim_{T\to+\infty} \frac{1}{T}\int_0^T ||u(s)||^2 \, \mathrm{d}s\,.$$

This number is the norm of K, which turns out to be a Hilbert space.

It is easily seen that theorem 4 holds also in the Hilbert space K provided that u is bounded on $[0, +\infty)$, (finite) values of the constants M and $\tilde{M}$ can be found on $[0, +\infty)$ and furthermore the estimate from below in (8) holds on $[0, +\infty)$.

We noted in Remark 8 that the constant $\tilde{M}$ can be found also on $[0, +\infty)$ provided that $X(t, s)$ is bounded on $0 \le s \le t$. This is a Liapunov stability condition which always holds in the applications when the system is linear (because when the null solution is not stable then the system will enter a nonlinear regime).

The estimate from below in (8) is more delicate. Once the step τ and T have been fixed, then the best choice for $(1 - m\tau)$ is the supremum of the numbers $\tilde{m}$ such that

$$\tilde{m}I \le \frac{1}{\tau}\int_{\tau_k}^{\tau_{k+1}} X(\tau_{k+1}, s)X^*(\tau_{k+1}, s)\, \mathrm{d}s = \frac{1}{\tau}R_k \tag{17}$$

(in fact we must also take $\tilde{m} < 1$). Inequality (17) must hold for each k such that $k\tau < T$. Hence for each natural number k if $T = +\infty$.

The condition to be checked is that we can find $\tilde{m} > 0$ which satisfies the previous inequality

It is important to note that *if $T = +\infty$, the number $\tilde{m}$ defined as above is strictly positive in the important case that A is a constant matrix* since, as we noted, in this case R_k does not depend on k. If instead the system is time dependent and $T = +\infty$ then $\tilde{m}$ can be zero, as it is seen in the following example: the system is scalar and $A(t) = -t$. In this case

$$X(t, s) = e^{(t^2-s^2)/2}\,, \qquad R_k = \int_{\tau_k}^{\tau_{k+1}} e^{-(\tau_{k+1}^2 - s^2)}\, \mathrm{d}s\,.$$

A standard computation shows that

$$\left[\int_{\tau_k}^{\tau_{k+1}} e^{s^2}\, \mathrm{d}s\right]^2 \le \frac{\pi}{4}\left[e^{2k\tau}e^{\tau} - 1\right]$$

so that

$$R_k \leq \frac{\sqrt{\pi}}{2} e^{-k^2\tau/2} e^{k\tau/2} e^{-\tau/2} \to 0$$

when $k \to +\infty$, with fixed τ. Hence, in this case the conditions used in the proof of Theorem 4 are not satisfied.

2.4 Systems of relative degree one

We consider now systems of relative degree one. This means that the system is time-invariant, i.e. the matrices A, B and C are constant, and that

$$\ker CB = 0\,,$$

see [1]. This case is of great importance in robustness theory. We are going to show that this case can be reduced to the case of full state transformation and, in principle, treated as previously described. Of course, in concrete examples, different procedures might be more convenient, see the comments in Remark 10 and the example in Section 4. As in [3], we consider the matrix

$$\hat{J} = BB^*C^*C\,.$$

It is seen in [3] that this matrix has nonnegative eigenvalues and it is diagonalizable. So, in a suitable coordinate system we have

$$\begin{aligned} \hat{J} &= \begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix}, & \hat{A} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\ \hat{B} &= \begin{bmatrix} B_+ \\ B_- \end{bmatrix}, & \hat{C} &= \begin{bmatrix} C_+ & C_- \end{bmatrix}. \end{aligned} \tag{18}$$

Here J is diagonal with positive diagonal elements. The condition $\ker CB = 0$ implies that $B_- = 0$, see [3, Th. 23]. In fact, we have more:

Lemma 9 *We have* $\ker C_+ = 0$, $\ker B_+ = 0$ *and* $C_+^* C_- = 0$.

<u>Proof</u>. As $B_- = 0$ $\ker CB = 0$ we see that $\ker C_+B_+ = 0$ so that we have also $\ker B_+ = 0$. We compute $\hat{J}$ and we find that it has the form

$$\hat{J} = \begin{bmatrix} B_+B_+^*C_+^*C_+ & B_+B_+^*C_+^*C_- \\ 0 & 0 \end{bmatrix}$$

Hence,

$$J = B_+B_+^*C_+^*C_+\,.$$

We know that J is invertible so that B_+ is surjective so that $\ker B_+B_+^* = 0$. From $0 = B_+B_+^*C_+^*C_-$ we get $C_+^*C_- = 0$. ∎

This lemma has an important consequence: we can multiply the output y by C_+^* and we get a new output, which does not "see" the component y_-. Moreover, the property $\ker C_+ = 0$ shows that $C_+^* C_+$ is invertible: a suitable transformation of coordinates reduces the system to the following form:

$$\left\{ \begin{array}{lcl} \dot{x}_+ & = & A_{11}x_+ + A_{12}x_- + B_+u \\ \dot{x}_- & = & A_{21}x_+ + A_{22}x_- \end{array} \right. \qquad\qquad y = x_+ \,.$$

Noisy observation is now

$$\xi_k = x_+(\tau_k) + \theta$$

where the error θ is of the order of h. Hence, x_+ is known (wit tolerance h) and the initial condition is zero. Consequently, the output y is given by

$$y(\tau_k) = \left[\int_0^{\tau_k} e^{A_{11}(\tau_k - s)} b(s)\, \mathrm{d}s + \zeta_k \right]$$

where $b(s) = B_+u(s)$ as in the previous sections and

$$\zeta_k = \int_0^{\tau_k} e^{A_{11}(\tau_k - s)} [A_{12}x_-(s)]\, \mathrm{d}s$$

can be computed (with an error of the order of h). This can be subtracted from the noisy observation so to obtain a new observation

$$\hat{\xi}_k = \xi_k - \zeta_k$$

which is the full state (noisy) observation for a system of the form

$$\dot{x}_+ = A_{11}x_+ + B_+u \,.$$

The observation error is still of the order of h and we are so reduced to the case we solved in the previous sections.

Remark 10 We observe that even if the matrix A is stable, the submatrix A_{11} might be unstable so that the methods just described might be practically not feasible. This problem can be handled with the device described in [2, Sect. B]. ∎

2.4.1 Why relative degree one

It is well known that any system of the form (1) is equivalent to its Morse quasicanonical form. This observation has been used in [3], where it is proved that the identification problem in the general case $C \neq I$ can be reduced to the solution of a (finite) chain of identification problems as those described in Sect. 2. The case treated in [3] is the case that the value of α is fixed once and for all.

The arguments in [3] (and applied to the disturbance reduction in [2]) can be applied to every linear time invariant system. The restriction that the system be of relative degree one was not used. In order to understand this point, let us consider the simple case

$$x_1' = x_2\,, \quad x_2' = u\,, \qquad y = x_1\,.$$

The goal is the reconstruction of u which is achieved in two steps: we use y and our algorithm in order to obtain an estimate $y_1 = \hat{x}_2$ of the "input" x_2 which is then seen as a new "observation" in order to estimate u. This can be done on-line (with the delay of two steps). Thanks to Morse canonical form, this example contains all the features of the general case.

Of course, this two-steps algorithm requires a very precise reconstruction of x_2 in order to achieve a reasonable reconstruction of u and, as seen in [2, Formula (42)], this require small values of α (i.e. of the error say h_1 in the "observation" of x_2) which cannot be achieved with the adaptive method, due to the estimate in Lemma 7. In fact, we repeat that the goal of this recursive choice of α is not sharp reconstruction of the input signal, but robustness in the case of poor information on the class of possible inputs, see Remark 1.

3 Active noise cancellation

We assume now that the input u to the systems is

$$(f_0 + \omega) - u$$

where f_0 is a known signal and ω is the disturbance. the control u has to be constructed so to reduce the effect of ω. We proceed as described in [2]. On the first interval we put $u = u_{(0)} = 0$ and we feed $Bf_0(t) + u_{(0)} + v(t)$ to the model system. The signal v has to be determined. It is determined from condition (5). The term which is penalized is the norm of v and not that of $Bf_0 + v$ since, due to the linearity, the contribution of Bf_0 in the system and in the model cancel out.

The algorithm provides a function $v_{(0)}$ and $\hat{\omega}(t) = B^\dagger v_{(0)}(t)$. In the next interval $I_1 = [\tau_1, \tau_2)$ we feed

$$u_{(1)}(t) = B^\dagger \hat{\omega}(t - \tau)$$

to the system and

$$Bf_0(t) - Bu_{(1)}(t) + v$$

to its model. The function v is then penalized as in (7) in order to construct $\hat{\omega}(t)$ on I_2. We then choose

$$u_{(2)}(t) = B^\dagger \hat{\omega}(t - \tau)\,, \qquad t \in I_3\,.$$

We proceed in a similar way in each one of the next steps. Alternatively, as in [2] we can consider $u_{(1)}(t)$ as an "approximant" of the noise. We feed it to the model and not to the system. In this way the system and the model are subject to similar noises which is then "estimated" in the next step. We then subtract this estimate of the noise, i.e. $B^{\dagger}\hat{\omega}(t-2\tau)$, both from the system and its model. This has the role of the control which suppress the noise with a two-steps delay both from the system and from its model. The same proof as in [2] shows that asymptotically, for $\tau \to 0$, $\alpha \to 0+$ while respecting suitable relations the output y converges to the nominal output of the system, i.e. the output driven by f_0 without noise. The condition on α is automatically achieved when $h \to 0+$ if the adaptive algorithm is used.

4 Example and simulations

Simulations of realistic problems have been presented in [2]. In that paper however the choice of the regularization parameter α was off-line. We consider now one of the systems already studied in [2], but now we use the adaptive algorithm as studied in this paper for the determination of α. The interest of this system to flight control is described in [14]. Here we confine ourselves to present the equations. the system is scalar, i.e. $\dim u = \dim y = 1$ while the state has dimension 3. It has the form

$$\dot{\mathbf{x}} = A\mathbf{x} + B\{\zeta f_0 - u\}\,, \qquad y = C\mathbf{x}$$

where the nominal value of ζ is 1. The value of ζ may change due to the failure of some component and the effect of u should be to restore the nominal value of ζf_0. This problem fits into our framework once we represent

$$\zeta f_0 - u = f_0 + (\zeta - 1) f_0 - u$$

where now $(\zeta-1)f_0$ is seen as a disturbance to be evaluated and canceled.

The matrices of the system are

$$A = \begin{bmatrix} -0.5162 & 26.96 & 178.9 \\ -0.6896 & -1.225 & -30.38 \\ 0 & 0 & -13 \end{bmatrix}$$
$$B = \begin{bmatrix} -175,6 \\ 0 \\ 14 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 12.43 & 0 \end{bmatrix}.$$

This system is of relative degree one. The matrix A is exponentially stable but the transformation described in Section 2.4.1 is not convenient since the submatrix corresponding to the matrix J (i.e. the matrix that after the transformation will appear in position $(1,1)$, now of dimension 1) is not stable. So, we proceed as in [2] where it is shown that a suitable coordinate transformation can be used to represent the system as

$$\dot{x} = -0.0385x - \{0.7101x_2 - 0.0902x_3\} + \zeta f_0 - u$$
$$\begin{bmatrix} x_2 \\ x_3 \end{bmatrix}' = \begin{bmatrix} 7.3467 & -2.4221 \\ 124 & -29.8436 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1.7325 \\ 20.7556 \end{bmatrix} x\,.$$

In an appropriate scale, the output is $y = x$.

The matrix of the second subsystem is stable, with eigenvalues of the order of -17.8 and -4.6 so that the values of x_1 and x_2 get soon very small (in spite of the fact that the eigenvectors of the matrix are almost colinear). This suggests that we proceed as if the system were described only by its first component, the effect of the second and third components being approximately computed from the output. So, we proceed as follows: we use as a model the system

$$\dot{w} = -0.0385x - \{0.7101w_2 - 0.0902w_3\} + \zeta f_0 - u$$
$$\begin{bmatrix} w_2 \\ w_3 \end{bmatrix}' = \begin{bmatrix} 7.3467 & -2.4221 \\ 124 & -29.8436 \end{bmatrix} \begin{bmatrix} w_2 \\ w_3 \end{bmatrix} + \begin{bmatrix} 1.7325 \\ 20.7556 \end{bmatrix} \xi$$

where ξ is the function of the observations, $\xi(t) = \xi_k$ for $t \in [\tau_k, \tau_{k+1})$.

The output y is read on a time interval of 50 time units, $0 \le t \le 50 = T$. Preliminarily we fix the sample rate. We decide to sample twice each time unit, $\tau = .5$. At each time $\tau_k = k\tau$ we read the output corrupted by noise,

$$\xi_k = y(\tau_k)(1 + h\omega)$$

where ω is a random error uniformly distributed in $[-0.5, 0.5]$. Hence h represents here a relative error.

Instead then choosing γ and ϵ these parameters are determined from the tolerance h on the measures and the tolerance μh admitted in the reconstruction, as

$$\gamma = \log(\mu h)/\log h\,, \qquad \epsilon = 0.9\gamma$$

(the choices are such that h and μh are less then 1 so that $\gamma \in (0,1)$).

The nominal value of ζ is $\zeta = 1$ and we consider the case that, due to the failure of an actuator,

$$\zeta(t) = \begin{cases} 1 & \text{if} \quad t < T/3 \\ -10 & \text{if} \quad T/3 \le t < 2T/3 \\ -20 & \text{if} \quad t > 2T/3\,. \end{cases}$$

The function $f_0(t)$ is $f_0(t) = 1/(t+1)$.

Figure 1: Left: desired output (tick) and uncompensated output. Right: the jump function.

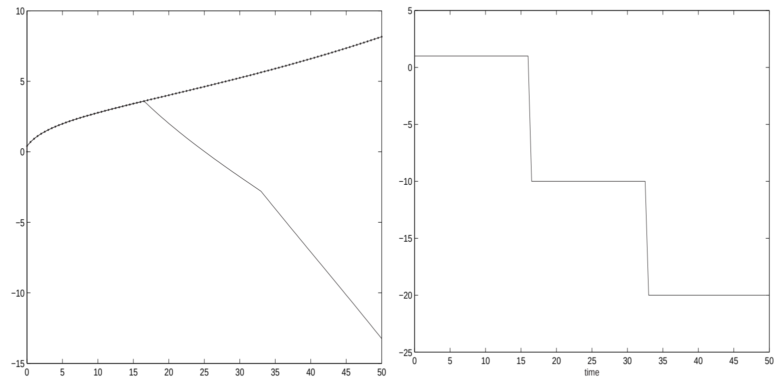
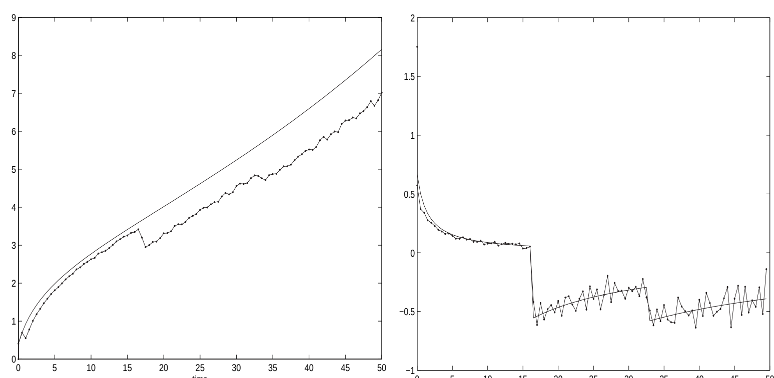


Figure 1 plots the desired and the uncompensated evolution of the output on the left and the jump function $\zeta(t)$ on the right.

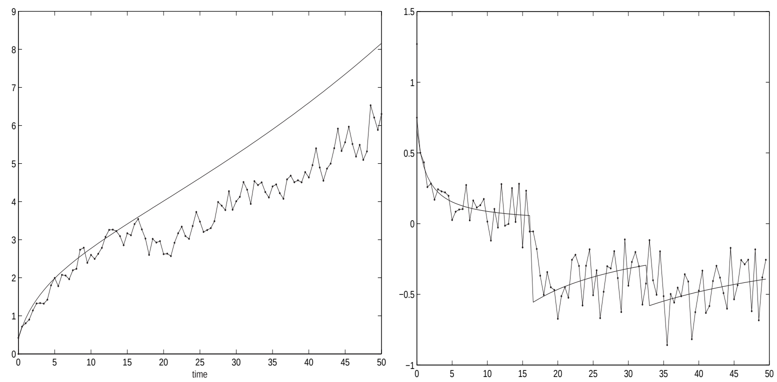
The figures we present next show the effect of the compensator. We consider first $\mu = 3$ and $h = 0.01$ (i.e. 1% of relative error on each individual measure). In Fig. 2 we represent the compensated evolution and the desired evolution of the output on the left, the estimate of $\zeta(t)f_0(t)$ on the right. In order to present a case which seems not acceptable, we represent the

Figure 2: Error tolerance $h = 0.01$, $\mu = 3$.



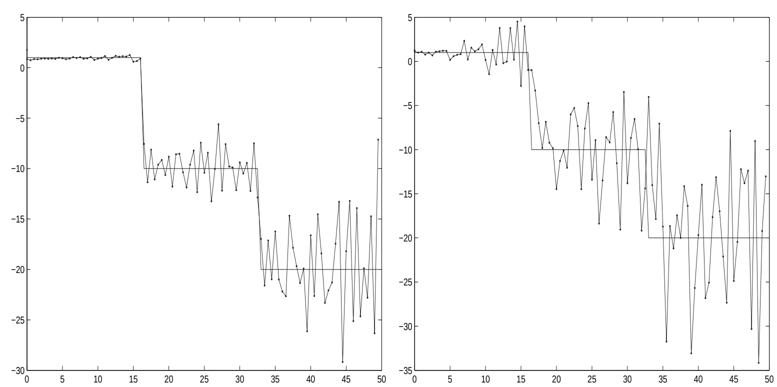
same plots in Fig. 3, but with $h = 0.1$ (i.e. 10% error) and $\mu = 10$. The unacceptable feature of this simulation is not so much high error ($\mu = 10$, this was our choice) but the fact that cancellation has been obtained at the expenses of fast oscillations. In fact, in this paper we did not make any

Figure 3: Error tolerance $h = 0.1$, $\mu = 10$.



attempt to attenuate the derivative of the reconstruction of the signal. Such fast oscillations are due to poor reconstruction of the jump function $\zeta(t)$. The next figure (Fig. 4) shows in fact the functions $\zeta(t)$ and its reconstructions on the left for the case examined in Fig. 2 and on the right for the case examined in Fig. 3.

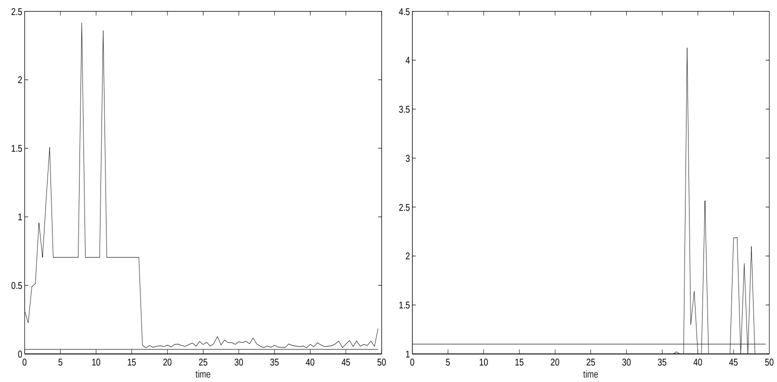
Figure 4: The jump function and its reconstruction.



Fidelity of reconstruction is not a key point of ANC but it is clear that too poor a reconstruction is not compatible with a satisfactory performance.

Finally, Fig. 5 shows the function $\alpha(t)$ on the left for the case examined in Fig. 2 and on the right for the case examined in Fig. 3.

Figure 5: The values of α and μh^γ.



References

[1] V. Dragan, A. Halanay, *Stabilization of linear systems,* Birkhäuser, Boston, 1999.

[2] F. Fagnani, V. Maksimov, L. Pandolfi, A recursive deconvolution approach to disturbance reduction, *IEEE Trans. Automat. Control,* **49** 907-921, 2004.

[3] F. Fagnani, L. Pandolfi, A singular perturbation approach to a recursive deconvolution problem, *SIAM J. Control Optim.* **40** (5), 1384-1405 (2002).

[4] W.S. Gan, S.M. Kuo, An integrated audio and active noise headset, *IEEE Trans. Consumer Electronics,* **48** 242-247, 2002.

[5] C.W. Groetsch, *The theory of Tikhonov regularization for Fredholm equations of the first kind,* Pitman, Boston, 1984.

[6] M. Hanke, P.C. Hensen, Regularization methods for large-scale problems, Surv. Math. Ind., **3** 253–315, 1993

[7] Yu. S. Osipov, A.V. Kryazhimskii *Inverse problems for ordinary differential equations: dynamical solutions,* Gordon and Breach, London, 1995.

[8] P.K. Lamm, Future-sequential regularization methods for ill-posed Volterra equations, *J. Mathematical Analysis Appl.* **195** 465–494, 1995.

[9] P.K. Lamm, Full convergence of sequential local regularization methods for inverse problems, *Inverse Problems,* **21** 785-803, 2005.

[10] P.K. Lamm, Variable smoothing regularization methods for inverse problems, in *Theory and Practice of Control and Systems,* A. Tornambè, G. Conte, A.M. Perdon Ed.s, World Scientific, Singapore, 1999.

[11] V.I. Maksimov, *Dynamical inverse problems of distributed systems* VSP, Utrecht, the Netherlands, 2002.

[12] V.A. Morozov, *Methods for solving incorrectly posed problems,* Springer, New York, 1984

[13] Y. Shiang-Hwua, H. Jwu-Sheng, Controller design for active noise cancellation headphones using experimental row data, *IEEE/ASME Trans. Mechatron.,* **6** 483-490, 2001.

[14] J.-S. Yee, J. L. Wang, B. Jiang, A fault tolerant flight controller design for actuator faults, in *Proc. 3th Int. Conference in Nonlinear Problems in Aviation and Aerospace,* S. Sivasundaram Ed., Cambridge U.K., 2001, pp. 723-730.