Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Robustly ℓ_∞-Stable Implementation of the Adaptive Inverse Control Scheme for Noise Cancelation

Néstor O. Pérez Arancibia and Tsu-Chin Tsao

Abstract— **This paper addresses the issues relating to the enforcement of robust stability when implementing the Adaptive Inverse Control (AIC) scheme. In this scheme, an adaptive FIR filter is added to a closed-loop system in order to reduce the output error caused by external disturbances, enhancing the performance achieved by linear time-invariant controllers alone. A Small-Gain-Theorem-based sufficient stability condition, which accounts for the feedback interaction between the time-varying adaptive filter and the unmodeled dynamics existing in the closed-loop plant, is derived. This condition suggests that system stability can be imposed by reducing the feasible region where the FIR filter gains can lie. Thus, a constrained convex optimization problem is formulated and solved for the FIR filter form. A relation between this optimization problem and the recursive least-squares (RLS) algorithm is established. Also, a suboptimal solution, which is implementable recursively, is proposed. Finally, to demonstrate their effectiveness, these algorithms are implemented on a laser beam steering experiment.**

I. Introduction

In this work, we address the stability issues involved in the implementation of the AIC scheme used for noise cancelation [1], [2]. The AIC scheme has been demonstrated to be very useful in a wide range of applications. In particular, it has been seen that using AIC, it is possible to enhance the performance achieved by linear time-invariant (LTI) controllers employed in laser beam jitter suppression applications [4], [5]. The block diagram of the AIC system is shown in Fig. 1.

In general, the stability problem arises when we have a system that consists of smaller subsystems interconnected in feedback configurations. Even though each of these subsystems is internally stable, the bigger system could be unstable. In the particular case of AIC, ensuring that the adaptive filter employed is stable does not guarantee stability of the scheme as a whole. Arguments for stability of the AIC scheme have been discussed in [1]. However, those are based on conditions that are impossible to impose on a real case experiment. Such arguments do not take into account the inherent uncertainty in any identified model, due to the inability of an LTI model to capture the real dynamics of a physical system, which in general, is time-varying and nonlinear.

When implementing the AIC scheme, the adaptation process can be performed using a wide range of stochastic methods, such as the least-mean-squares (LMS), and deterministic methods, such as the recursive least-squares (RLS) method. Convergence and stability of adaptive filtering implementation, using these algorithms, are studied in [6]–[10]. However, as mentioned earlier, stability of the subsystems does not guarantee stability of the whole scheme. Assuming that in the block diagram of Fig. 1, all the input signals are bounded and all the subsystems are internally stable, a condition for ℓ_∞-stability is determined. Imposing this condition leads to the formulation of a constrained optimization problem. Specifically, it can be shown that one way to ensure stability under uncertainty is to enforce certain constraints on the adaptive filter gains. This problem has not been addressed in the adaptive filtering literature, though, it is mentioned briefly in [9]. We refer to this as the constrained adaptive filtering problem.

In particular, we focus on the recursive solution to the least-squares (LS) problem, subject to the constraints on the adaptive filter coefficients that impose robust stability under uncertainty. The consequence of constraining the filter coefficients is the reduction of the feasible region, where the set of optimal coefficients can lie. Fortunately, this constrained problem is convex [11], allowing us to formulate two classes of algorithms. The first type exploits the relation between the constrained LS problem and the RLS algorithm. The second type is based on a suboptimal solution obtained by computing projections of the unconstrained solution onto the convex feasible region. The applicability of these algorithms is demonstrated on the optical experiment described in [4] and [5].

It is important to emphasize that the approach, chosen in this paper, allows us to address a practical issue in the core of many adaptive schemes. This type of controllers require strong and rich excitation of the output signals, which is usually against the control objectives. Therefore, it is common to encounter cases, where high performance diminishes the excitation of the measured outputs, causing the adjustable parameters to drift to unstable regions [12].

The paper is organized as follows. In Section II, we introduce some notation and formulate the problem in consideration. In Section III, a connection between the constrained LS problem and the RLS algorithm is established. After that, following a different approach, a suboptimal solution is found. Finalizing this section, we propose two recursive algorithms. In Section IV, the experimental implementation of these algorithms is demonstrated, and finally in Section V we draw some conclusions and discuss some directions for future work.

This work was supported, in part, by the U.S. Air Force Office of Scientific Research under Grant F49620-02-01-0319.

The authors are with the Mechanical and Aerospace Engineering Department, University of California, Los Angeles, CA 90095-1597, USA. nestor@seas.ucla.edu, ttsao@seas.ucla.edu.

0-7803-9568-9/05/$20.00 ©2005 IEEE

II. Problem Formulation

A. The Adaptive Inverse Control Scheme

In principle, the AIC scheme intends to find a transfer function $F(z)z^{-1}$, which is the inverse of the plant G in a certain frequency range. This frequency range depends on the bandwidth of the disturbance sequence $n(i)$. In reality, the finding for $F(z)$ can be seen as a predictive process based on $\hat{G}(z)$, which is the plant model of the system G. Usually, $\hat{G}(z)$ is computed off line.

In every step, an estimate $\hat{n}(i)$ for the disturbance $n(i)$ is computed using $\hat{G}(z)$. For analysis purposes, we consider that the reference signal $r(i)$ is set to 0 for all i. Thus, once $\hat{n}(i)$ is computed, it is filtered through $-F(z)z^{-1}$ and, ideally, the output of $-F(z)z^{-1}$ is the required input for G, such that, the output of G is $v(i) \approx -\hat{n}(i)$. Therefore, if $\hat{n}(i) \approx n(i)$, the noise $n(i)$ is counteracted, making the control error smaller.

In this work, $F(z)$ is chosen to be a finite impulse response (FIR) filter, whose coefficients are found solving an LS problem. The data used for solving this LS problem are generated using $\hat{G}(z)$ and $\hat{n}(i)$. This can be done off-line, as in [3], or in an adaptive-recursive manner [4], [5]. In the following sections, we show that stability of the whole interconnected system is not guaranteed for any of these two estimation methods when $\hat{G} \neq G$, even if, $F(z)$ is stable. We propose a method to handle this problem.

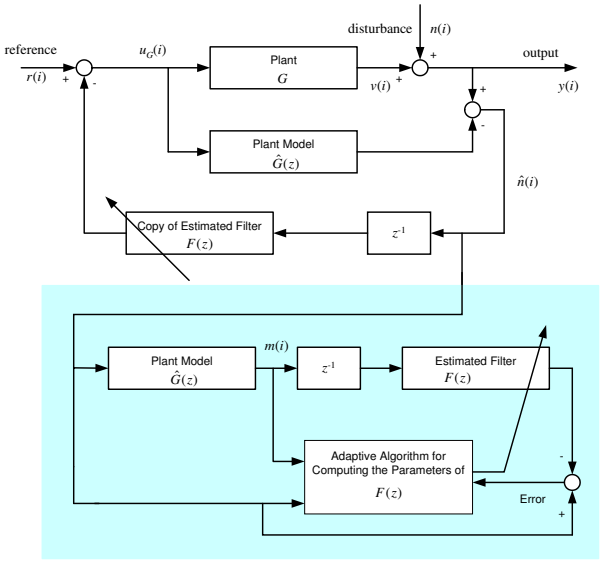

Fig. 1. Adaptive Inverse Control Scheme.

B. Notation and Mathematical Preliminaries

The space of all bounded scalar-valued sequences is denoted by $\ell_\infty(Z)$. Thus, if $x = \{\ldots, x(-1), x(0), x(1), \ldots\}$, with $x(k) \in \Re$, is a sequence in $\ell_\infty(Z)$, then

$$\|x\|_{\ell_\infty} = \sup_k |x(k)| < \infty. \tag{1}$$

From a mathematical point of view, a system is an operator that maps sequences between two signal spaces. In this case, the operators of interest are the operators that map signals from ℓ_∞ to ℓ_∞ (operators on ℓ_∞), that are linear and causal, but not necessarily time-invariant. The space of all linear, causal operators from ℓ_∞ to ℓ_∞ can be represented by an infinite dimensional block triangular matrix

$$R = \begin{bmatrix} R(0,0) & 0 & 0 & \ldots \\ R(1,0) & R(1,1) & 0 & \ldots \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix}, \tag{2}$$

where, $R(i,j) \in \Re$ [13]. Consequently, if the sequence $\{u(k)\}_{k=0}^{\infty}$ in ℓ_∞ is the input to the system associated with the matrix R and the sequence $\{y(k)\}_{k=0}^{\infty}$ in ℓ_∞ is the output from the system associated with R, then $y = Ru$.

An operator F from ℓ_∞ to ℓ_∞ is called bounded if its induced norm defined as

$$\|F\|_{\ell_\infty \to \ell_\infty} = \sup_{x \in \ell_\infty, x \neq 0} \frac{\|Fx\|_{\ell_\infty}}{\|x\|_{\ell_\infty}} \tag{3}$$

is finite. Also, an operator $F : \ell_\infty \to \ell_\infty$ is said to be stable with respect to a signal space ℓ_∞ if it is bounded on ℓ_∞ [13]–[16].

In particular, we would like to find the induced norm from ℓ_∞ to ℓ_∞ for the linear time-varying (LTV) system $F(k,z) = w_0(k) + w_1(k)z^{-1} + \ldots + w_N(k)z^{-N}$, with $w_i(k) \in \Re$, and where z^{-1} denotes the unit delay operator. Pursuing that, consider an input sequence $u = \{u(k)\}_{k=0}^{\infty}$. Then, the relationship $y = Ru$, becomes

$$\begin{bmatrix} y(0) & y(1) & \ldots & y(N) & \ldots & \ldots \end{bmatrix}^T =$$

$$\begin{bmatrix} w_0(0) & 0 & 0 & \ldots \\ w_1(1) & w_0(1) & 0 & \ldots \\ \vdots & \vdots & \ddots & \vdots \\ w_N(N) & w_{N-1}(N) & \ldots & \ldots \\ 0 & w_N(N+1) & \ldots & \ldots \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(N) \\ u(N+1) \\ \vdots \\ \vdots \end{bmatrix}.$$

Now, define $w(k)$ as the vector that contains all the elements that are different than 0 in the k^{th} row, with $k = \{0, 1, 2, \ldots\}$. For example, $w(0) = w_0(0)$, $w(1) = [w_1(1)\ w_0(1)]^T$, $w(N) = [w_N(N)\ \ldots\ w_0(N)]^T$ and $w(N+1) = [w_N(N+1)\ \ldots\ w_0(N+1)]^T$. This allows us to write the following.

Lemma 1: The induced norm from ℓ_∞ to ℓ_∞ of the LTV system F is given by

$$\|F\|_{\ell_\infty \to \ell_\infty} = \sup_k \|w(k)\|_{\ell_1}. \tag{4}$$

Proof: It is straightforward that

$$\|F\|_{\ell_\infty \to \ell_\infty} = \sup_{\|x\|_{\ell_\infty} \leq 1} \|Fx\|_{\ell_\infty} = \sup_k w(k)^T x_s(k), \tag{5}$$

with $x_s(k) = sign([w_N(k)\ w_{N-1}(k)\ \dots\ w_0(k)]^T)$. Then, it is clear that (4) holds. ■

It is worth to mention that when F is LTI, with w being a constant vector of gains, the relation $\|F\|_{\ell_\infty \to \ell_\infty} = \|w\|_{\ell_1}$ holds. This result is consistent with the well known theorem that states that the induced norm from ℓ_∞ to ℓ_∞ of an LTI bounded system H is given by $\|h\|_{\ell_1}$, where h denotes the impulse response of the system H [13].

Next, consider the block diagram in Fig. 2. It is clear that the relations

$$\begin{aligned} y_1 &= P_2 y_2 + u_1 \\ y_2 &= P_1 y_1 + u_2 \end{aligned} \tag{6}$$

hold. The feedback connection in Fig. 2 is called well posed if (6) gives a unique output $\{y_1, y_2\}$ for any input $\{u_1, u_2\}$ in ℓ_∞ [13]–[15]. A special case of wellposedness is given when the operator P_1P_2 is strictly causal.

Lemma 2: If the operator P_1P_2 is strictly causal, then the feedback connection of Fig. 2 is well posed.

Proof: It is clear that $y_2 = P_1P_2y_2 + P_1u_1 + u_2$, therefore, for any $\{u_1, u_2\} \in \ell_\infty$, a unique solution, y_2, can be computed recursively, since P_1P_2 is strictly causal. Similarly, y_1 can be computed in the same way ■

The last mathematical tool that we need is a particular version of the *Small Gain Theorem*.

Theorem 1: Let $P_1 : \ell_\infty \to \ell_\infty$ and $P_2 : \ell_\infty \to \ell_\infty$ be two stable operators and assume that the closed-loop system, in Fig. 2, is well posed. Then, the closed-loop system is ℓ_∞-stable if $\|P_1\|_{\ell_\infty \to \ell_\infty}\|P_2\|_{\ell_\infty \to \ell_\infty} < 1$.

Proof: See [13]. ■

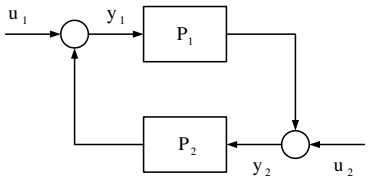

Fig. 2. Typical Feedback Connection.

C. Model Uncertainty and Robust Stability in the AIC scheme

Consider Fig. 1, and let Q be a system operator, such that, $y = Qr$. It is immediate that $Q = G$ if the condition $\hat{G} = G$ holds. Consequently, it is possible to conclude that Q is ℓ_∞-stable for any $F(z)$ if the system plant G is stable. Unfortunately, this stability condition is not useful in a real case scenario, since it is theoretically impossible to have a model $\hat{G}(z)$ such that $\hat{G} = G$.

We could think that a way to be confident about the stability of our system is to have a model very close to the real physical system. However, no matter how *good* our model is, one can always find gains for $F(z)$ so that the system is unstable. On the other hand, empirical evidence shows that, in many cases, the control scheme works in a stable way, even when, $\hat{G} \neq G$ [4], [5]. This indicates that there should be a less conservative and more realistic sufficient condition that ensures the stability of the adaptive scheme as a whole if G belongs to a known set of system operators.

Let the signals r and n in Fig. 1 be in ℓ_∞ and let $\hat{G}$, G and F be operators on ℓ_∞. In general, G is a bounded, nonlinear, time-varying and causal operator, and therefore it is natural to consider that

$$G = \hat{G} + \Delta, \tag{7}$$

where, in general Δ is a bounded, nonlinear, time-varying and causal operator, representing the differences between modeling and reality. Another interpretation of Δ is as an uncertainty. In other words, Δ represents what we do not know about the physical system G.

On the other hand, let us define the system operator F_a as

$$F_a(z) = -F(z)z^{-1}, \tag{8}$$

and notice that F_a is strictly causal and linear. Also, notice that considering (7) and (8), the system in Fig. 1 and the system in Fig. 3 are equivalent. Hence, one can write

$$\begin{aligned} u_G &= F_a\hat{n} + r \\ \hat{n} &= \Delta u_G + n. \end{aligned} \tag{9}$$

The development, done thus far, allows us to conclude some facts about our adaptive system.

Lemma 3: The system in Fig. 3 is well posed.

Proof: F_a is a strictly causal operator and Δ is a causal operator, therefore the operator ΔF_a is strictly causal. Then, by lemma 2, the system in Fig. 3 is well posed. ■

Theorem 2: Let $w(i) = [w_N(i)\ \dots\ w_0(i)]^T$ be the adaptively computed vector of gains for the prediction problem in Fig. 1 at time i. Furthermore, let $\|\Delta\|_{\ell_\infty \to \ell_\infty} < \frac{1}{\gamma}$, with $\gamma \in \Re^{++}$. Then the system in Fig. 3 is ℓ_∞-stable if

$$\|w(i)\|_{\ell_1} \leq \gamma,\ \forall\ i = 0, 1, 2, \dots \tag{10}$$

Proof: It is immediate that $\|F_a\|_{\ell_\infty \to \ell_\infty} = \|F\|_{\ell_\infty \to \ell_\infty}$. Then, by Lemmas 1–3 and Theorem 1, the system in Fig. 3 is ℓ_∞-stable if $\sup_i \|w(i)\|_{\ell_1} \leq \gamma$, which is equivalent to (10). ■

Since the systems in Fig. 1 and Fig. 3 are equivalent, Theorem 2 gives us a sufficient condition for enforcing ℓ_∞-stability on the AIC scheme in Fig. 1. Experimentally, it is possible to include this condition by imposing bounds on the value of the ℓ_1-norm of the adaptively computed vector of gains. Thus, (10) leads to the formulation of a new problem, namely, the constrained adaptive filtering problem.

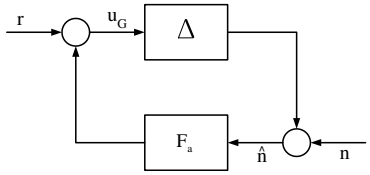

Fig. 3. Δ and F_a in Typical Feedback Connection.

D. The Constrained Adaptive Filtering Problem

As mentioned earlier, we need to formulate a least-squares problem that allows us to estimate coefficients for $F(z) = w_0 + z^{-1}w_1 + \ldots + z^{-M+1}w_{M-1}$ in Fig. 4. We define $d(i)$ as the desired output, $u(i)$ as the input and $\hat{d}(i)$ as the actual output from $F(z)$ with coefficients estimated at time $i-1$.

It is common to define the so-called regressor row vector as $U(i) = [u(i)\ u(i-1)\ \ldots\ u(i-M+1)]$. Then, one can write $\hat{d}(i) = U(i)w(i-1)$, where $w(i-1) = [w_0(i-1)\ w_1(i-1)\ \ldots\ w_{M-1}(i-1+M)]^T$. The relations between the variables in Fig. 1 and Fig. 4 are: $u(i) = m(i-1)$ and $d(i) = \hat{n}(i)$.

Thus, we formulate the optimization problem

$$\min_{w} \|Aw - b\|_{\ell_2}^2, \tag{11}$$

where the matrices for $N-1$ samples are

$$b = \begin{bmatrix} d(0) \\ d(1) \\ \vdots \\ d(N-1) \end{bmatrix} \ and\ A = \begin{bmatrix} U(0) \\ U(1) \\ \vdots \\ U(N-1) \end{bmatrix}. \tag{12}$$

Notice that this optimization problem is unconstrained, which implies that the feasible region is $\Re^M$, where $M-1$ is the order of the filter to be found. The recursive solution to the regularized version of this problem is the well known RLS algorithm. In this case, in order to impose the condition given by Theorem 2, we formulate the constrained least-squares problem

$$\min_{w} \|Aw - b\|_{\ell_2}^2 \quad s.t. \quad \|w\|_{\ell_1} \leq \gamma. \tag{13}$$

This optimization problem is convex, which means that an optimal solution can be found, and that this solution is unique if $rank(A) = M$. However, in this case we have the additional difficulty of finding this optimal point in a recursive manner. Thus, it could be useful to replace the constraint given by the ℓ_1-norm by the constraint $\|w\|_{\ell_2} \leq \alpha$, where, $\alpha = \frac{\gamma}{\sqrt{M}}$. This is possible, because for a vector $x \in \Re^n$, the relationship $\|x\|_{\ell_1} \leq \sqrt{n}\|x\|_{\ell_2}$ holds.

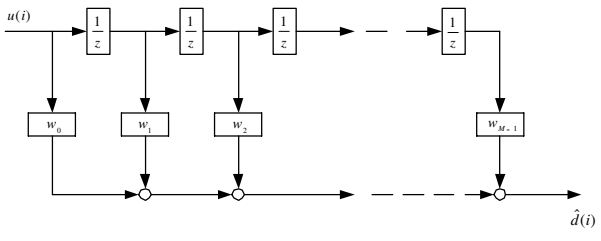

Fig. 4. FIR Adaptive Filter.

III. Some Approaches to the Solution of the Constrained Problem

A. Optimal Solution to the ℓ_2-norm Constrained Least-Squares Problem Using SVD Decomposition

Our first goal is to have a solution to the constrained least-squares problem

$$\min_{w} \|Aw - b\|_{\ell_2}^2 \quad s.t. \quad \|w\|_{\ell_2}^2 \leq \alpha^2 = \frac{\gamma^2}{M}, \tag{14}$$

which allows us to have a better understanding of what is needed to be done in order to ensure the conditions of stability.

To begin with, consider the SVD of matrix A. $A = U\Sigma_A V^T$, where U and V are unitary matrices and where Σ_A is a diagonal matrix with r elements greater than zero and $r = rank(A)$ [17]. Multiplication of a vector by unitary matrices preserve the *length* of the vector. Hence, $\|Aw - b\|_{\ell_2}^2 = \|U^T(Aw - b)\|_{\ell_2}^2$ and $\|w\|_{\ell_2} = \|V^T w\|_{\ell_2}$. Furthermore, if we define $\tilde{b} = U^T b$ and $y = V^T w$, it is possible to formulate the equivalent minimization problem

$$\min_{y} \|\Sigma_A y - \tilde{b}\|_{\ell_2}^2 \quad s.t. \quad \|y\|_{\ell_2}^2 \leq \alpha^2. \tag{15}$$

In this case, we assume that $r = n$, i.e., the rank of matrix A is equal to the number of columns of A. The case $r \neq n$ is a trivial generalization of this solution.

Strong duality allows us to solve (15) using Lagrange multipliers and the Karush-Kuhn-Tucker (KKT) conditions [11]. Thus, we write the corresponding Lagrangian function

$$h(y, \lambda) = \|\Sigma_A y - \tilde{b}\|_{\ell_2}^2 + \lambda(\|y\|_{\ell_2}^2 - \alpha^2) =$$

$$\sum_{i=1}^{n}(\sigma_i - y_i\tilde{b}_i)^2 + \sum_{i=n+1}^{m} \tilde{b_i}^2 + \lambda\left(\sum_{i=1}^{n} y_i^2 - \alpha^2\right) \tag{16}$$

and the corresponding KKT conditions

$$\lambda \geq 0 \tag{17}$$

$$\frac{\partial h(y,\lambda)}{\partial y_i} = (\sigma_i y_i - \tilde{b}_i)\sigma_i + y_i\lambda = 0 \tag{18}$$

$$\frac{\partial h(y,\lambda)}{\partial \lambda} = \sum_{i=1}^{n} y_i^2 - \alpha^2 = 0. \tag{19}$$

In this way, the optimal w^o can be found as follows.

1) Find the optimal solution to the unconstrained problem. If this solution lies outside the feasible region go to step 2, otherwise stop.
2) Find the optimal $\lambda^o \geq 0$, that satisfies
$$\sum_{i=1}^{n} \frac{\tilde{b}_i\sigma_i}{\sigma_i^2 + \lambda^o} = \alpha^2. \tag{20}$$
3) Compute the vector y^o as
$$y_i^o = \frac{\tilde{b}_i\sigma_i}{\sigma_i^2 + \lambda^o}. \tag{21}$$

4) Compute the optimal w^o for the original problem (14) as

$$w^o = V y^o. \tag{22}$$

The previous algorithm allows us to solve the problem, nevertheless, this solution is not implementable in a recursive manner. Despite this fact, it sheds light to understanding the nature of the original problem (14). Clearly, (20)–(22) indicate that there exists a set $\Lambda = \{\lambda : \lambda > \lambda^o\}$, such that, any vector $w(\lambda)$, with $\lambda \in \Lambda$, is inside the feasible region and therefore, such a point $w(\lambda)$ is a vector of gains robustly stable in the sense of (10).

B. The Constrained ℓ_2-norm LS problem as a Regularized LS problem

In the previous section we found an explicit solution for the constrained least-squares problem and we noticed that the optimal solution to the constrained problem (14) can be found by scaling the i^{th} entry of the unconstrained optimal vector by $\frac{\sigma_i^2}{\sigma_i^2+\lambda^o}$. Therefore $w(\lambda)$, with $\lambda \geq \lambda^o$, is inside the feasible region. These facts indicate that if we are able to find, from the data for example, an estimate for the optimal λ^o, we could find a recursive method for computing $w(\lambda^o)$.

To begin with, we found optimality conditions for the original problem (14). In this case, the Lagrangian is

$$L(w,\mu) = \|Aw-b\|_{\ell_2}^2 + \mu(\|w\|_{\ell_2}^2 - \alpha^2) =$$

$$w^T A^T A w - 2w^T A^T b + b^T b + \mu\left(\sum_{i=1}^{n} w_i^2 - \alpha^2\right) \tag{23}$$

and the KKT conditions are given by

$$\mu \geq 0 \tag{24}$$

$$\nabla_w L(w,\mu) = 2A^T A w - 2A^T b + 2\mu w = 0 \tag{25}$$

$$\frac{\partial L(w,\mu)}{\partial \mu} = \|w\|^2 - \alpha^2 = 0. \tag{26}$$

From (25) we obtain

$$(A^T A + \mu I) w = A^T b, \tag{27}$$

which is the normal equation for the regularized problem, stated as

$$\min_w \left\| \begin{pmatrix} A \\ \sqrt{\mu} I \end{pmatrix} w - \begin{pmatrix} b \\ 0 \end{pmatrix} \right\|_{\ell_2}^2 = \min_w \|Aw-b\|_{\ell_2}^2 + \mu \|w\|_{\ell_2}^2. \tag{28}$$

Now, we establish a strong relation between (16) and (27), that will be used later.

Lemma 4: Let (y^o, λ^o) be the solution to (17)–(19) and let (w^o, μ^o) be the solution to (24)–(26). Then, $\mu^o = \lambda^o$ and $h(y^o, \lambda^o) = L(w^o, \mu^o)$.

Proof: From (16), it is immediate that $(\Sigma_A^T \Sigma_A + \lambda^o I) y^o = \Sigma_A^T \tilde{b}$. Thus, $(V^T A^T U U^T A V + \lambda^o I) V^T w^o = V^T A^T U U^T b$. Then, $V^T (A^T A + \lambda^o I) w^o = V^T A^T b$, which implies $(A^T A + \lambda^o I) w^o = A^T b$, since V^T is always invertible. By (27), μ^o must satisfy the same relationship. Therefore, $\mu^o = \lambda^o$ and $h(y^o, \lambda^o) = L(w^o, \mu^o)$. ■

Thus, in principle, (14) can be solved in two steps. First, an estimate $\widehat{\lambda^o}$ for λ^o is found, and then, the problem (28) is solved. $\widehat{\lambda^o}$ must satisfy $\widehat{\lambda^o} \geq \lambda^o$, in order to have a solution point inside the feasible region.

It is essential to note that the problem (28), for a fixed estimate $\widehat{\lambda^o}$, is nothing but the regularized least-squares problem that can be solved using the recursions

$$w(i) = w(i-1) +$$

$$\frac{P(i-1)U^T(i)}{1+U(i)P(i-1)U^T(i)}(d(i) - U(i)w(i-1)) \tag{29}$$

$$P(i) = P(i-1) - \frac{P(i-1)U^T(i)U(i)P(i-1)}{1+U(i)P(i-1)U^T(i)}, \tag{30}$$

with $w(-1) = 0$ and $P(-1) = \widehat{\lambda^o}^{-1} I$ [6]. In other words, when using the RLS algorithm, one is solving a ℓ_2-norm constrained problem with an unknown α. It is important to mention that, for numerical reasons, it is preferable to implement (29)–(30) using some equivalent method. For example, the inverse QR-RLS algorithm [6], [7].

Thus, one can solve (14), recursively, employing the RLS algorithm. However, if our estimate of λ^o is too conservative or not big enough, we would like to modify the value of $P(-1)$. Since RLS is recursive, this is not possible. Nevertheless, one can reinitialize the filter with a new $P(-1)$. Considering these ideas, we define Algorithm 1. For reasons that will become obvious, from this point onwards, we distinguish between the regularized gains $w(i)$ computed using (29)-(30) and the constrained gains $w^c(i)$ that meet the stability condition. Consequently, the control scheme is always implemented using $w^c(i)$.

Algorithm 1 (Variable P-RLS):

1) Compute $w(i)$ and $P(i)$ using (29)–(30) or a numerically more reliable equivalent method (for example the inverse QR-RLS algorithm).
2) If $\|w(i)\|_{\ell_2} < \alpha$ set $w^c(i) = w(i)$, return to step 1 and compute $w(i+1)$ and $P(i+1)$. If $\|w(i)\|_{\ell_2} \geq \alpha$ go to step 3.
3) Set $w^c(i+k) = w(i-1)$ for $k = 0, \ldots, M_o - 1$. Set $w(i) = 0$ and $P(i) = \widehat{\lambda^o}^{-1} I$, with $\widehat{\lambda^o}^{-1} := \max\{\widehat{\lambda^o}^{-1} + K_p(\alpha - \|w(i)\|_{\ell_2}),\ \epsilon > 0\}$, where K_p and ϵ are in $\Re^{++}$. Use (29)–(30) to update w and P. When $k = M_o$, set $w^c(i+M_o) = w(i+M_o)$, return to step 1 and compute $w(i+M_o+1)$ and $P(i+M_o+1)$.

C. Suboptimal Solution to the Constrained LS problem

In this section, we derive suboptimal solutions for the constrained LS problem. This is done by assuming that we know the optimal point, w^{uo}, of the unconstrained or slightly regularized LS problem. In general, we will have two possible cases. The case when w^{uo} is inside the feasible region, and the case when w^{uo} is outside. Since the norm constraints define convex feasible regions, if w^{uo} is outside,

the optimal point of the constrained problem must be in the boundary of the feasible region. Considering the previous facts, we propose to compute a suboptimal point, w^{co}, by projecting w^{uo}, orthogonally, onto the feasible region. This idea is illustrated by Fig. 5.

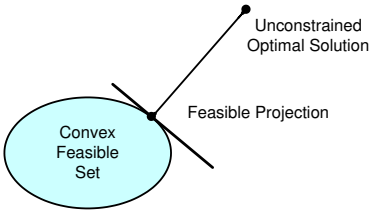

Fig. 5. Feasible Set.

1) Suboptimal Solution to the ℓ_1-norm Constrained LS problem: The projection problem for the ℓ_1-norm constraint case can be stated as

$$\min_{w^c} \|w^c - w^{uo}\|_{\ell_2}^2 \quad s.t. \quad \|w^c\|_{\ell_1} \leq \gamma. \tag{31}$$

Noting that the feasible region of (31) is determined by the ℓ_1-norm of the points w^cs, the optimal point, w^{co}, of (31) must have the same direction that the unconstrained optimal vector w^{uo}. Thus, we can write the equivalent optimization problem

$$\min_{w^c} \|w^c - w^{uo}\|_{\ell_2}^2 \quad s.t. \quad c^T w^c \leq \gamma, \tag{32}$$

where c is the vector that contains the signs of the entries of w^{uo}. This equivalent problem remains convex, because the new constraint is a halfplane. The Lagrangian is given by

$$h(w^c, \lambda) = \sum_{i=1}^{n} (w_i^c - w_i^{uo})^2 + \lambda \left(\sum_{i=1}^{n} c_i w_i^c - \gamma \right). \tag{33}$$

Thus, deriving the corresponding KKT conditions, the optimal point becomes

$$w_i^{co} = w_i^{uo} - \frac{\lambda^o c_i}{2}, \tag{34}$$

where the optimal λ^o is given by

$$\lambda^o = \frac{2}{\sum_{i=1}^{n} c_i^2} \left(\sum_{i=1}^{n} c_i w_i^{uo} - \gamma \right) = \frac{2}{n} \left(\sum_{i=1}^{n} c_i w_i^{uo} - \gamma \right). \tag{35}$$

2) Suboptimal Solution to the ℓ_2-norm Constrained LS problem: The projection problem for the ℓ_2-norm constraint case can be stated as

$$\min_{w^c} \|w^c - w^{uo}\|_{\ell_2}^2 \quad s.t. \quad \|w^c\|_{\ell_2}^2 \leq \alpha^2 \ . \tag{36}$$

In this case, the Lagrangian function is

$$h(w^c, \lambda) = \sum_{i=1}^{n} (w_i^c - w_i^{uo})^2 + \lambda \left(\sum_{i=1}^{n} (w_i^c)^2 - \alpha^2 \right). \tag{37}$$

Thus, deriving the corresponding KKT conditions, the optimal point becomes

$$w_i^{co} = \frac{w_i^{uo}}{1 + \lambda^o}, \tag{38}$$

where λ^o solves

$$\frac{\|w^{uo}\|_{\ell_2}^2}{(1+\lambda)^2} = \alpha^2. \tag{39}$$

3) The Simple Projection-RLS Algorithm: Based on the previous results, it is possible to formulate another algorithm that enforces the stability condition (10).

Algorithm 2 (The Simple Projection-RLS Algorithm):

1) Compute $w(i)$ and $P(i)$ using (29)-(30) or a numerically more reliable equivalent method.
2) If $\|w(i)\|_{\ell_2\ (\ell_1)} < \alpha\ (\gamma)$ set $w^c(i) = w(i)$, return to step 1 and compute $w(i+1)$ and $P(i+1)$. If $\|w(i)\|_{\ell_2\ (\ell_1)} \geq \alpha\ (\gamma)$ go to step 3.
3) Compute $w^c(i)$ according to (34) in the ℓ_1-norm case or according to (38) in the ℓ_2-norm case. Set $w(i) = w^c(i)$ and go back to step 1.

IV. Experimental Implementation of the Algorithms

A. Experiment Description

The algorithms developed in the previous sections were implemented, in real time, on an optical experiment. For this purpose, we used a Texas Instruments TMS320C6701 digital signal processor (DSP). The sampling rate was set to 2KHz.

From a physical viewpoint, this experiment is a free-space path through which a laser beam travels. Disturbance is added during the trajectory, which produces jittering in the laser beam spot to be measured at the end of the optical path. The control objective is to adaptively suppress this jitter using the scheme shown in Fig. 1. The experimental setup is shown in Fig. 6 and the diagram of this optical system is shown in Fig. 7. The main optical components in this system are the laser source, two MEMS beam steering mirrors (BSM), and a position sensing device (sensor). The diagram in Fig. 7 shows the path of the laser beam from the source to the sensor. After leaving the laser source, the beam reflects of the mirror BSM 1, which serves as the control actuator, then reflects of the mirror BSM 2, which adds disturbance to the beam direction, and finally goes to the sensor. Each mirror rotates about horizontal and vertical axes and the outputs of the sensor are the horizontal and vertical displacements of the centroid of the laser spot on the sensor plane.

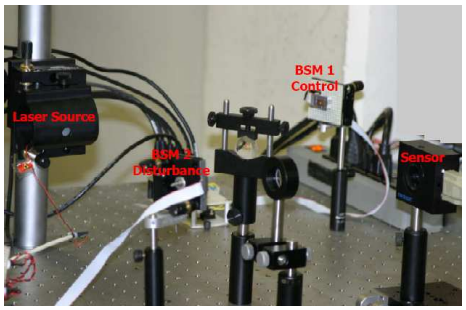

Fig. 6. Laser Beam Steering Experiment.

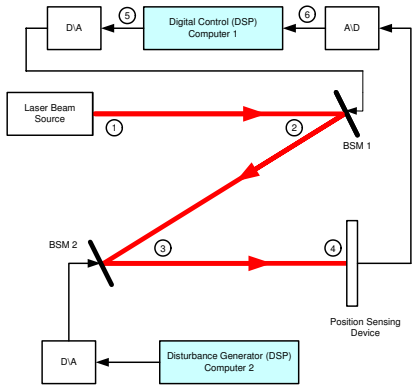

Fig. 7. Diagram of Experiment.

TABLE I
RMS VALUES OF THE OUTPUT ERROR y (8000 SAMPLES).

Method	RMS values of y
Open Loop	0.5944
LTI Feedback Controller	0.5834
Adaptive Scheme with Algorithm 1	0.0536
Adaptive Scheme with Algorithm 2	0.0573

For practical reasons, in this subsection, we consider that all the physical systems, involved in this experiment, are LTI. Thus, the discrete-time open-loop transfer function, $P(z)$, is defined as everything between points 5 and 6, in Fig. 7. In principle, this transfer function is a MIMO system, because the MEMS mirror has two axes. However, these two axes are decoupled, reason by which, in this paper we consider the horizontal jitter only. The closed-loop transfer function, $G(z)$, is determined by the block diagram in Fig. 8, where $C(z)$ is an LTI controller designed to reject the disturbance n_0. Notice that n_0 and n of Fig. 1 are related by $n = (1 - PC)^{-1} n_o$. The Bode plot of the computed sensitivity function, $S(z) = (1 - \hat{P}(z)C(z))^{-1}$, is shown in Fig. 9. For further information about the experimental implementation, system identification, LTI control design, and other aspects relating to this system, refer to [4] and [5].

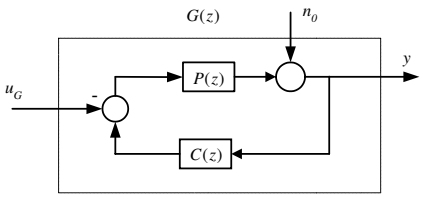

Fig. 8. Block Diagram of LTI Feedback Control System. $P(z)$: Open-Loop Plant; $C(z)$: LTI Feedback Controller; $G(z) = Y(z)/U_G(z)$.

B. Experimental Results and Analysis

Let d_o be the sequence generated inside the DSP 2 to be applied, as a disturbance, to the laser beam through the actuator mirror BSM 2. Then, the open-loop output measurement is $n_o = P_2 d_o$, where P_2 is the BSM 2 plant. Similarly, when the optical system is under the control of the LTI feedback controller C, the disturbance to be predicted and canceled adaptively is $n = (1 - PC)^{-1} P_2 d_o$. In this case, d_o was generated as the sum of various sequences with different bandwidths. These bandwidths are: 0-130Hz, 225-235Hz, 340-350Hz and 455-465Hz.

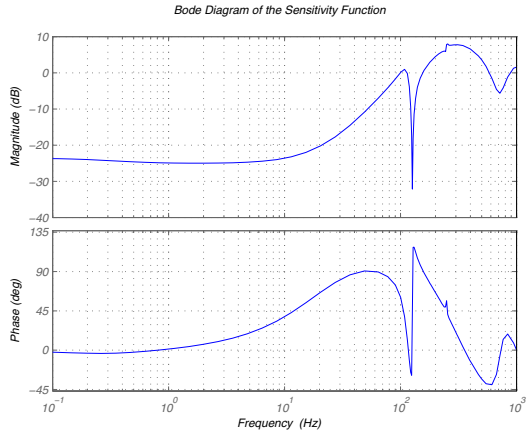

Fig. 9. Bode Plot for the Computed Model, $S(z)$, of the Sensitivity Function.

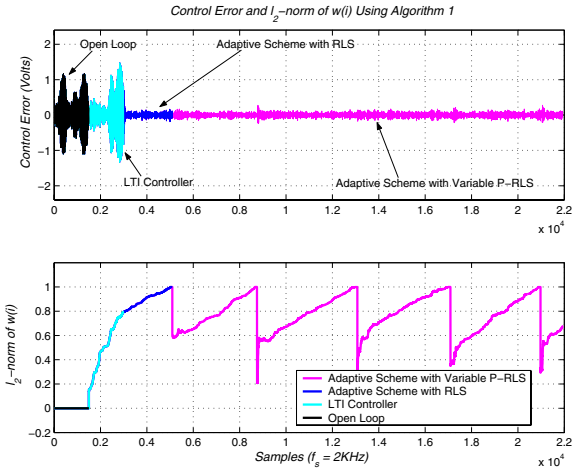

Fig. 10. Algorithm 1, with $\alpha = 1.0$, $K_P = 10$, and $M_o = 100$. The Filter Order is 16. *Top Plot*: Time Series. *Bottom Plot*: ℓ_2-norm of $w^c(i)$ Over Time.

The effectiveness of the constrained adaptive algorithms is demonstrated in Fig. 10 and Fig. 11. In Fig. 10, the top plot shows the time series of the measured output y, using Algorithm 1, with $\alpha = 1.0$, $K_p = 10$, $M_o = 100$ and initial $\widehat{\lambda^o}^{-1} = 10^7$, for a filter of order 16. The bottom plot shows the evolution on time of $\|w^c(i)\|_{\ell_2}$. In order to have many transition points, K_p was chosen small. Similarly, in Fig. 11, the top plot shows the time series of the measured output y, using Algorithm 2, with $\alpha = 1.0$ and $\widehat{\lambda^o}^{-1} = 10^7$, for a filter of order 16. The bottom plot shows the evolution on time of $\|w^c(i)\|_{\ell_2}$. The plots are segmented in four sections. The first section shows the system in open loop,

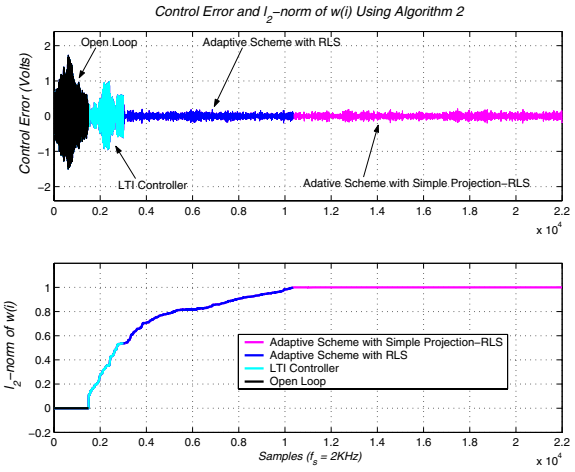

Fig. 11. Algorithm 2, with $\alpha = 1.0$ and $\widehat{\lambda^o}^{-1} = 10^7$. The Filter Order is 16. *Top Plot*: Time Series. *Bottom Plot*: ℓ_2-norm of $w^c(i)$ Over Time.

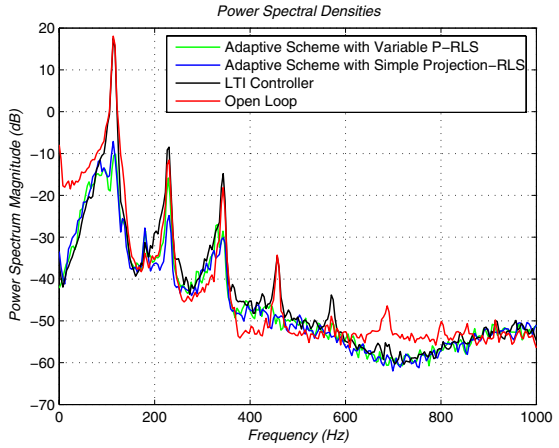

Fig. 12. PSDs of the Outputs (8000 Samples).

the second section shows the system operating under the LTI controller, designed in [4], and the last two sections show the system operating under the adaptive scheme, with the unconstrained RLS first, and then with the constrained RLS algorithms. Both constrained algorithms are able to bound the filter coefficients, as desired, and improve significatively the performance achieved by the LTI feedback controller alone. This can be seen in Table I, Fig. 10, and Fig. 11. Fig. 12 shows that the noise with bandwidth below to 80Hz is suppressed by the LTI controller. In higher frequencies the LTI controller amplifies more than rejects, which explains the performance of the LTI controller in Table I, Fig. 10 and Fig. 11. The former is predicted by Fig. 9. Thus, the noise cancelation above 80Hz is mostly done by the AIC scheme.

Finally, it is essential to mention that under the constrained adaptive scheme, the optical system is able to endure strong disturbances, which cannot be tolerated by the unconstrained scheme, such as, blocking of the laser beam path and high amplitude mechanical vibrations, returning to the steady-state performance shown in Table I. Based on empirical evidence, in this system, ℓ_∞-stability is enforced if $\|w(i)\|_{\ell_2} \leq 10$, which means that in this experiment we used a scheme 10 times more conservative than the critical limit.

V. CONCLUSIONS

In this paper we presented some initial results on the implementation of a robustly stable AIC scheme. We worked on the $\ell_\infty(Z)$ space. Assuming additive uncertainty, a sufficient condition for enforcing ℓ_∞-stability was found. We showed that this condition can be imposed by solving a constrained convex optimization problem. A connection between this constrained convex problem and the RLS algorithm was established, leading us to the formulation of the Variable P-RLS algorithm. Also, we proposed a projection algorithm which gives us a suboptimal solution implementable recursively (Algorithm 2). Finally, we demonstrate these algorithms on a laser beam jitter suppression application.

Empirical evidence shows that the control error converges, satisfactorily, to a small level, when either of the two algorithms is implemented on this experiment. However, the question of the convergence properties of the AIC scheme, in general, and the algorithms proposed here, in particular, still remains. A beginning would be to enforce the convergence properties of the RLS algorithm [10], when bounding the norms of $w(i)$.

REFERENCES

[1] B. Widrow and E. Walach, *Adaptive Inverse Control.* Upper Saddle River, NJ: Prentice-Hall, 1996.

[2] B. Widrow and S. D. Stearns, *Adaptive Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1985.

[3] B.-S. Kim, S. Gibson and T.-C. Tsao, "Adaptive control of a tilt mirror for laser beam steering," in *Proc. American Control Conference*, Boston, MA, Jun. 2004, pp. 3417–3421.

[4] N. O. Pérez Arancibia, N. Chen, S. Gibson and T.-C. Tsao, "Adaptive control of a MEMS steering mirror for suppression of laser beam jitter," in *Proc. American Control Conference*, Portland, OR, Jun. 2005, pp. 3586–3591.

[5] N. O. Pérez Arancibia, S. Gibson and T.-C. Tsao, "Adaptive control of MEMS mirrors for beam steering," in *Proc. ASME International Mechanical Engineering Congress*, Anaheim, CA, Nov. 2004, IMECE2004–60256.

[6] A.H. Sayed, *Fundamentals of Adaptive Filtering.* New York, NY: John Wiley & Sons, 2003.

[7] S. Haykin, *Adaptive Filter Theory.* Upper Saddle River, NJ: Prentice-Hall, 1996.

[8] S.T. Alexander, *Adaptive Signal Processing.* New York, NY: Springer-Verlag, 1986.

[9] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1984.

[10] G. C. Goodwin, P. J. Ramadge and P. A. Caines, "Discrete-Time Multivariable Adaptive Control," *IEEE Transactions on Automatic Control*, vol 25, NO. 3, pp. 449-455, Jun. 1980.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambrige, UK: Cambrige University Press, 2004.

[12] T.-T. Tay, I. Mareels and J. B. Moore, *High Performance Control.* Boston, MA: Birkäuser, 1998.

[13] M. A. Dahle and I. J. Diaz-Bobillo, *Control of Uncertain Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[14] J. Doyle, B. Francis and A. Tannenbaum, *Feedback Control Theory.* Basingstoke Hampshire, UK: Macmillan, 1992.

[15] G. E. Dullerud and F. Paganini, *A Course in Robust Control Theory.* New York, NY: Springer-Verlag, 2000.

[16] G. E. Dullerud, *Control of Uncertain Sampled-Data Systems.* Boston, MA: Birkäuser, 1996.

[17] G.H. Golub and C.F. Van Loan, *Matrix Computations.* Baltimore, MD: The Johns Hopkins University Press, 1989.