15th Triennial World Congress, Barcelona, Spain

USING LOCAL AND GLOBAL INFORMATION IN IDENTIFICATION FOR CONTROL

Henrik Jansson and Håkan Hjalmarsson *

* *Division of Automatic Control, Dept. of Signals, Sensors and Systems, Royal Institute of Technology (KTH), Stockholm, Sweden,* henrik.jansson@s3.kth.se

Abstract: The problem of identifying a restricted complexity model for control purposes is considered in this contribution. A new iterative identification method is proposed in which local and global information about the control design criterion is blended. The global information, which is used for faster convergence, comes from the usual extrapolating property of a model and the local information, used for higher accuracy, is the sensitivity of the closed loop system to the model parameters. It is shown that the method has the same stationary points as the control design criterion and it is also shown that if the approximated Hessian is sufficiently accurate, the method converge locally.

Keywords: Closed-loop identification, Iterative methods, Model-based control

1. PROBLEM STATEMENT

The true system is described by

$$y(t) = G_0(q)u(t), \tag{1}$$

where the dynamics between the output $y(t)$ and the input $u(t)$ is governed by the unknown linear time-invariant discrete transfer operator $G_0(q)$. The system is for simplicity assumed to be single input/single output and to keep the ideas transparent and avoid technicalities we will also assume that it is noise-free. The system is controlled by a one-degree of freedom linear time-invariant controller $C(q)$, i.e.

$$u(t) = C(q)(r(t) - y(t)), \tag{2}$$

where $r(t)$ is the reference signal. The time argument will be omitted frequently to ease the notation.

In this presentation we will consider model based control design, where as an intermediate step a model of G_0 is used to design the controller. Let G_0 be modeled by the transfer function $G(\theta)$ parameterized by the real valued vector θ. Since the design is based on the model, the controller can be expressed as a function of the model, i.e. $C = C(\theta)$. Then the designed closed loop control system can be expressed as $y(\theta) = T(\theta)r$ where $T(\theta) = \frac{G(\theta)C(\theta)}{1+G(\theta)C(\theta)}$. The achieved closed loop response is denoted by $T_0(\theta) = \frac{G_0C(\theta)}{1+G_0C(\theta)}$. The objective of the control design should be to minimize the difference between the designed and the achieved performance. The control performance can be measured by some norm, e.g.

$$J(\theta) = \frac{1}{2N}\sum_{t=1}^{N}[(T_0(\theta) - T(\theta))r \tag{3}$$

The task is then to design a controller that minimizes (3) w.r.t. θ. Introducing the designed and the achieved sensitivity function, $S(\theta) = \frac{1}{1+G(\theta)C(\theta)} = 1 - T(\theta)$ and $S_0(\theta) = \frac{1}{1+G_0(\theta)C(\theta)} = 1 - T_0(\theta)$, respectively, the criterion (3) can be rewritten as

$$\begin{aligned} J(\theta) &= \frac{1}{2N}\sum_{t=1}^{N}\varepsilon_{tc}(\theta)^2 \\ \varepsilon_{tc}(\theta) &= S(\theta)(y(\theta) - G(\theta)u(\theta)) \end{aligned} \tag{4}$$

Here, the criterion has been arranged in a way such that it now looks more like an identification problem. However, since both u and y depends on θ, i.e. they are the closed loop input/output signals which corresponds to $C(\theta)$ in the loop, this is somewhat misleading. The expression is still useful from an identification perspective in that it can be used to express how the model bias should be distributed over various frequencies so that (4)

is minimized. Hence, this expression has been a source of inspiration for developing identification methods that attempts to make (4) small. In the next section we will briefly review some of these approaches.

A new method is presented in Sec. 3. The method uses the idea of identification wherein a model $G(\theta)$ of G_0 is estimated. The model $G(\theta)$ is indirect also a model of the design criterion (4). This is useful since the extrapolating property of the model will give a hint of where the global optimum of (4) is. The proposed method also takes local variations of y and u, w.r.t. θ, into account. Ordinary identification methods cannot not do this, which may make them suboptimal to the problem of minimizing (4).

The local convergence for the proposed method is studied in Sec. 4 and numerical examples are given in Sec. 5. Finally there are some concluding remarks in Sec. 6.

2. PREVIOUS METHODS

2.1 Iterative identification and control schemes

In the early 1990's the problem of turning the control criterion into a combined identification and control problem was attacked using different so called iterative identification and control schemes, e.g. (Zang *et al.*, 1995; Schrama and Bosgra, 1993; Åström, 1993). The basic idea is to approximate the criterion so that it can be interpreted as identification in closed loop followed by control design. An example of approximation made for (4) is, in each iteration, to fix the current model estimate at all places in (4) except in $G(\theta)$. This means that at the ith iteration the current estimate is θ_i and the identification turns into minimizing

$$J_{ID} = \frac{1}{2N}\sum_{t=1}^{N}(S(\theta_i)(y(\theta_i) - G(\theta)u(\theta_i)))^2. \quad (5)$$

The algorithm is at the ith iteration:

(1) Collect $y(\theta_i)$ and $u(\theta_i)$ in closed loop using the controller based on the current model estimate θ_i.
(2) Update the model to obtain θ_{i+1} based on minimizing the criterion (5).
(3) Goto to 1.

This algorithm is quite intuitive. However, it has turned out to be very difficult to prove any type of convergence result. It has also been pointed out in (Hjalmarsson *et al.*, 1995) that possible convergence points of this type of scheme are not necessarily minima of (4). A similar situation arises in adaptive control, see (Ljung and Söderström, 1983).

2.2 Iterative Feedback Tuning

As a continuation of this work a method called Iterative Feedback Tuning(IFT) was proposed, c.f. (Hjalmarsson *et al.*, 1998). This method is focused on finding a local optimum to (3) or equivalently finding a solution to $J'(\theta) = 0$. This is done by taking repeated steps in a descent direction of the criterion, i.e.

$$\theta_{i+1} = \theta_i - \gamma_i R_i^{-1} J'(\theta_i), \quad (6)$$

where γ_i is the step length and R_i is typically a Gauss-Newton approximation of the Hessian of $J(\theta)$. The key contribution in IFT is that an unbiased estimate of the gradient $J'(\theta)$ can be computed using data from closed loop experiments. It is also is shown in (Hjalmarsson *et al.*, 1998) that the method is guaranteed to converge to a local optimum of the design criterion under the assumption of bounded signals and restricted step length. This means that the global optimum of the design criterion also is a stationary point for the IFT method. The drawback is that many experiments might be necessary to perform on the true plant due to the fact that only small steps are taken along the design criterion in the iterative search method and for each step new experiments have to be performed.

2.3 Virtual Reference Feedback Tuning

Recently a new model free tuning method has been proposed in (Campi *et al.*, 2000). Virtual Reference Feedback Tuning(VRFT) is a method of direct tuning of the controller C with a model reference criterion, $\|T_0 - T\|$. Minimizing $T_0 - T = S_0T(C(T^{-1}G_0 - G_0) - 1)$ corresponds to minimizing the sum of squares of $\epsilon = \Phi_{uu}^{-1/2} S_0 T(C(r_T - y) - u)$ where $r_T = T^{-1}y$ is called the virtual reference. In (Campi *et al.*, 2000) S_0 is approximated by $S = 1 - T$. Then the method only need one set of data $\{y, u\}$ to estimate the controller.

It is pointed out in (Hjalmarsson and Lindqvist, 2001) that the method can be interpreted as model based where the model is directly parameterized in terms of the controller parameters. The design criterion can then be rewritten as a frequency weighted identification problem using a certain prefilter.

The main advantage of the method is that only one experiment on the true plant to collect the data $\{y, u\}$ is required. The approximation of S_0 will however make the controller suboptimal to the true design criterion.

2.4 Discussion

In IFT information about the shape of the cost function in a neighborhood of the current model parameter is extracted in each iteration. This is at the same time the strength and the weakness of the method. The advantage is that precise information may be obtained even if the system is complex. It has been shown, see (Hjalmarsson, 1998) and (Sjöberg and Bruyne, 1999), that IFT can be applied to non-linear systems. However, the disadvantage is that the information is local so that only gradual changes may be possible.

Model based methods, on the other hand, can be viewed as methods where the model is used as an intermediate to model the *complete cost function*. Hence, as evidenced by the discussion above, this approach can be sensitive to model imperfections, but on the the other hand may be useful to point out the right region in the parameter space where the optimum is. In the next section we will try to combine the advantages of these two approaches.

3. COMBINING LOCAL AND GLOBAL INFORMATION

In this section two new methods are presented. They are both based on the same idea, which is to use the extrapolating property of a model, see e.g. the identification and control scheme described in Sec. 2, together with local modelling based on gradient information about the design criterion, as in IFT.

The derivations will be based on the design criterion (4). To simplify the presentation direct model parameterization of the controller via model reference design is considered, i.e. $C(\theta) = \frac{1}{G(\theta)}\frac{T_d}{1-T_d}$, where T_d is the reference model. This means that $S(\theta) = S_d = 1 - T_d$ in (4).

3.1 Linearizing y and u

The main difficulty with the criterion (4) from an identification point of view is that $y(\theta)$ and $u(\theta)$ depends on the model. The idea is to linearize $y(\theta)$ and $u(\theta)$ around the current model estimate θ_i, which results in the following approximation of (4),

$$\begin{aligned}\hat{J}_1(\theta,\theta_i) &= \frac{1}{2N}\sum_{t=1}^{N}\varepsilon_{c1}(\theta,\theta_i)^2 \\ \varepsilon_{c1}(\theta,\theta_i) &= S_d(y(\theta_i)+y'(\theta_i)(\theta-\theta_i) \\ &\quad - G(\theta)(u(\theta_i)+u'(\theta_i)(\theta-\theta_i))).\end{aligned} \tag{7}$$

Compare (7) with the identification criterion (5). It looks similar but (7) has two added terms $y'(\theta_i)(\theta-\theta_i)$ and $u'(\theta_i)(\theta-\theta_i)$ which are local approximations of how y and u are varying with respect to changes in the model parameter θ. Remember, that all the signals $y(\theta_i)$, $y'(\theta_i)$, $u(\theta_i)$ and $u'(\theta_i)$ are fixed and the free variable is the θ without subindex. The gradient signals $y'(\theta_i)$ and $u'(\theta_i)$ are obtained performing closed loop experiments, since $y(\theta) = T_0(\theta)r$ and $u(\theta) = C(\theta)S_0(\theta)r$ then

$$\begin{aligned}y'(\theta_i) &= \frac{C'(\theta_i)}{C(\theta_i}T_0(\theta_i)S_0(\theta_i)r \\ u'(\theta_i) &= C'(\theta_i)S_0^2(\theta_i)r.\end{aligned} \tag{8}$$

The approximated criterion (7) could also be compared with the Gauss-Newton search in IFT. Introduce $\tilde{y}(\theta) = y(\theta) - y_d = (T_0(\theta) - T_d)r$, then the following result, which is proven in (Hjalmarsson *et al.*, 1994), holds.

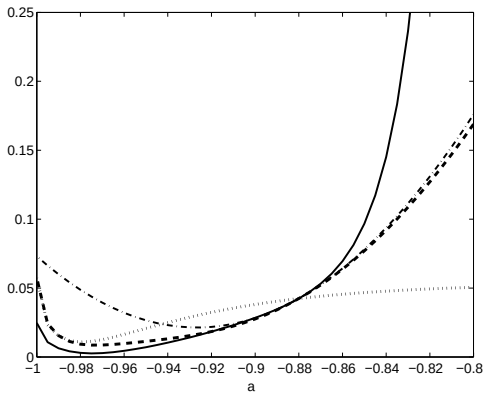

Fig. 1. Comparison of cost functions, solid line - true cost (4), dotted line - identification criterion (5), dashed line - blended criterion (7), dash-dotted line - quadratic approximation in IFT (9). Current model estimate, $a_i = -0.88$.

Lemma 1. Let $\tilde{z}_i(\theta)$ be a first order Taylor expansion of $\tilde{y}(\theta)$ around θ_i,

$$\tilde{z}_i(\theta) = \tilde{y}(\theta_i) + \tilde{y}'(\theta_i)(\theta - \theta_i).$$

Then the Gauss-Newton update θ_{i+1}, with step length one, is the solution of the minimization problem

$$\begin{aligned}\theta_{i+1} &= \arg\min_{\theta} J_{GN}(\theta) \\ J_{GN}(\theta) &= \frac{1}{2N}\sum_{t=1}^{N}\tilde{z}_i^2(\theta).\end{aligned} \tag{9}$$

The difference in (7) from the criterion used in the Gauss-Newton iterations in IFT is that $G(\theta)$ is not linearized, which makes the approximated criterion (7) more flexible w.r.t. the parameter θ. This is illustrated in Fig. 1. The model has one parameter a that is free and the current estimate is $a_i = -0.88$. Based on this estimate the three approximated cost functions, (7), (5) and (9) are compared with the true criterion (4). Here, the plot of $\hat{J}_1(\theta,\theta_i)$ clearly shows the trade off between local and global fitting to the true criterion.

3.2 Linearizing S_0

Reconsider (4)

$$\begin{aligned}J(\theta) &= \frac{1}{2N}\sum_{t=1}^{N}\left[S(\theta)(y(\theta) - G(\theta)u(\theta))\right]^2 \\ &= \frac{1}{2N}\sum_{t=1}^{N}[S(\theta)C(\theta)S_0(\theta)(G_0 - G(\theta))r]^2.\end{aligned} \tag{10}$$

The linearization of the signals $y(\theta)$ and $u(\theta)$ correspond to linearization of $C(\theta)S_0(\theta)$. Notice that the controller is completely known but not $S_0(\theta)$ since it contains the true system G_0. Hence, an alternative to (7) should be to only linearize $S_0(\theta)$. If $S_0(\theta)$ is linearized in (10) the following criterion is obtained:

$$\begin{aligned}\hat{J}_2(\theta,\theta_i) &= \frac{1}{2N}\sum_{t=1}^{N}\varepsilon_{c2}(\theta,\theta_i)^2\\ \varepsilon_{c2}(\theta,\theta_i) &= S_d C(\theta)(S_0(\theta_i)\\ &+ S_0'(\theta_i)(\theta-\theta_i))(G_0 - G(\theta))r\\ &= S_d\frac{C(\theta)}{C(\theta_i)}(y(\theta_i)+y_z(\theta,\theta_i)\\ &- G(\theta)(u(\theta_i)+u_z(\theta,\theta_i))).\end{aligned} \tag{11}$$

The criterion (11) has also the structure of an identification criterion with the added terms

$$\begin{aligned}y_z(\theta,\theta_i) &= -\frac{C'(\theta)}{C(\theta_i)}(\theta-\theta_i)T_0(\theta_i)^2 r\\ u_z(\theta,\theta_i) &= C'(\theta)(\theta-\theta_i)S_0(\theta_i)T_0(\theta_i)r.\end{aligned} \tag{12}$$

3.3 The algorithm

Let

$$\hat{J}_m(\theta,\theta_i) = \frac{1}{2N}\sum_{t=1}^{N}\varepsilon_m(\theta,\theta_i)^2 \tag{13}$$

denote either (7) or (11), then the following algorithm is proposed which at the ith iteration becomes

Algorithm 1.

(1) Collect $y(\theta_i)$ and $u(\theta_i)$ in closed loop using the controller based on the current model estimate θ_i.
(2) Do a second closed loop experiment to compute (8) or (12).
(3) Update the model
$$\theta_{i+1} = \arg\min_{\theta}\hat{J}_m(\theta,\theta_i).$$
(4) Goto to 1.

4. LOCAL CONVERGENCE ANALYSIS

The local convergence of the suggested algorithm Alg. 1 is studied in this section. The analysis will be made locally around the true minimum. The following is a result about the possible convergence points of the proposed algorithm.

Lemma 2. The possible convergence points of Alg. 1 are stationary points to the cost function (4), and the value of the approximated cost function (13) are equal to the true objective function (4) at those points.

Proof: Assume that the algorithm converge to $\theta = \bar{\theta}$ then $\frac{\partial\hat{J}_m(\bar{\theta},\bar{\theta})}{\partial\theta} = 0$ and

$$\begin{aligned}\frac{\partial\hat{J}_m(\bar{\theta},\bar{\theta})}{\partial\theta} &= \frac{1}{N}\sum_{t=1}^{N}\varepsilon_m(\bar{\theta},\bar{\theta})\frac{\partial\varepsilon_m(\bar{\theta},\bar{\theta})}{\partial\theta}\\ &= \frac{1}{N}\sum_{t=1}^{N}[S_d(\frac{\partial y(\bar{\theta})}{\partial\theta} - \frac{\partial P(\bar{\theta})}{\partial\theta}u(\bar{\theta}) - P(\bar{\theta})\frac{\partial u(\bar{\theta})}{\partial\theta})]\\ &\times[S_d(y(\bar{\theta}) - P(\bar{\theta})u(\bar{\theta}))].\end{aligned}$$

From (4) it follows that $\frac{\partial\hat{J}_m(\bar{\theta},\bar{\theta})}{\partial\theta} = \frac{\partial J(\bar{\theta})}{\partial\theta}$ and it is trivial to verify that $\hat{J}_m(\bar{\theta},\bar{\theta}) = J(\bar{\theta})$. □

This means that if the method converge, then the stationary point of $\hat{J}_m(\theta,\theta_i)$ also is a stationary point to the true design criterion $J(\theta)$. This should be compared with the iterative identification procedure described in Sec. 2 where, as pointed out, this is not always the case.

Introduce the following assumptions:

A1. Assume that the number of data N tends to infinity.
A2. Let θ^* be a stable minimum of (4).

Using assumption A1, (13) and (4) can be rewritten as $\hat{J}_m(\theta,\theta_i) = \frac{1}{2}E\varepsilon_m(\theta,\theta_i)^2$ and $J(\theta) = \frac{1}{2}E\varepsilon_{tc}(\theta)^2$.

The suggested algorithm, Alg. 1 update the model, in each iteration, according to $\theta_{i+1} = \arg\min_\theta\hat{J}_m(\theta,\theta_i)$, which, under convexity assumptions, is equivalent of finding the solution to $\frac{\partial\hat{J}_m(\theta_{i+1},\theta_i)}{\partial\theta} = 0$. Introduce $F(\theta_{i+1},\theta_i) = \frac{\partial\hat{J}(\theta_{i+1},\theta_i)}{\partial\theta}$ then the question of convergence for the algorithm turns into a stability problem of the implicit nonlinear autonomous function $F(\theta_{i+1},\theta_i)$. To prove local convergence the following classical result about stability of nonlinear systems is needed, see e.g. (Khalil, 1996).

Theorem 3. (stability in the first approximation) The stationary point θ^* is an asymptotically stable equilibrium point of $\theta_{i+1} = f(\theta_i)$ if all the eigenvalues of $\frac{f(\theta^*)}{\partial\theta_i}$ have magnitude less than one.

From the definition of $F(\theta_{i+1},\theta_i)$ we have that

$$\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_{i+1}} = \frac{\partial^2\hat{J}_m(\theta^*),\theta^*}{\partial\theta^2}. \tag{14}$$

It can be shown that (14) can be written as

$$\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_{i+1}} = \frac{\partial^2 J(\theta^*)}{\partial\theta^2} - E\varepsilon_m(\theta^*,\theta^*)^T g(\theta^*) \tag{15}$$

where $g(\theta^*)$ depends on the method. In the method corresponding to the criterion (7), g is given as

$$g(\theta^*) = S_d(\frac{\partial^2 y(\theta^*)}{\partial\theta^2} - P(\theta^*)\frac{\partial^2 u(\theta^*)}{\partial\theta^2}). \tag{16}$$

It can also be shown that

$$\begin{aligned}\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_i} &= E\varepsilon_m(\theta^*,\theta^*)^T g(\theta^*)\\ &= \frac{\partial^2\hat{J}_m(\theta^*,\theta^*)}{\partial\theta^2} - \frac{\partial^2 J(\theta^*)}{\partial\theta^2}.\end{aligned} \tag{17}$$

Let $\bar{\sigma}(\cdot)$ and $\underline{\sigma}(\cdot)$ denote the largest and the smallest singular value, respectively.

Theorem 4. If $\frac{\partial^2\hat{J}_m(\theta^*,\theta^*)}{\partial\theta^2}$ is nonsingular then there exist a neighborhood around θ^* in which the algorithm Alg. 1 converge to θ^* if

$$\frac{\bar{\sigma}(\frac{\partial^2 \hat{J}(\theta^*,\theta^*)}{\partial\theta^2} - \frac{\partial^2 J(\theta^*)}{\partial\theta^2})}{\underline{\sigma}(\frac{\partial^2 \hat{J}(\theta^*,\theta^*)}{\partial\theta^2})} < 1$$

Proof: Assumption A2 and Lemma 2 give $F(\theta^*,\theta^*) = 0$. The implicit function theorem (c.f. (Courant, 1954)) states that if $F(\theta^*,\theta^*) = 0$ and if $\partial F(\theta^*,\theta^*)/\partial\theta_{i+1}$ is nonsingular, then there exist a continuously differentiable function $f(\cdot)$ such that $\theta_{i+1} = f(\theta_i)$ in a neighborhood of θ^* and the differential is described by

$$\Delta\theta_{i+1} = -\left[\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_{i+1}}\right]^{-1}\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_i}\Delta\theta_i, \quad (18)$$

where $\Delta\theta_{i+1} = \theta_{i+1} - \theta^*$ and $\Delta\theta_i = \theta_i - \theta^*$. Furthermore from Theorem 3, θ^* is a locally asymptotically stable point to $\theta_{i+1} = f(\theta_i)$ if all the eigenvalues of $\frac{f(\theta^*)}{\partial\theta_i} = -\left[\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_{i+1}}\right]^{-1}\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_i}$ have magnitude less than one. An upper bound of the maximum eigenvalue is then, see (Horn and Johnson, 1991)

$$\max_i |\lambda_i(\left[\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_{i+1}}\right]^{-1}\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_i})|$$
$$\leq \frac{\bar{\sigma}(\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_i})}{\underline{\sigma}(\frac{\partial F(\theta^*,\theta^*)}{\partial\theta_{i+1}})} = \frac{\bar{\sigma}(\frac{\partial^2 \hat{J}(\theta^*,\theta^*)}{\partial\theta^2} - \frac{\partial^2 J(\theta^*)}{\partial\theta^2})}{\underline{\sigma}(\frac{\partial^2 \hat{J}(\theta^*,\theta^*)}{\partial\theta^2})}.$$

□

If θ^* is a strict optimum to $\hat{J}(\theta^*,\theta^*)$, then obviously $\frac{\partial^2 \hat{J}(\theta^*,\theta^*)}{\partial\theta^2}$ is nonsingular. The size of $\frac{\partial^2 \hat{J}(\theta^*,\theta^*)}{\partial\theta^2} - \frac{\partial^2 J(\theta^*)}{\partial\theta^2}$ is apparently important in the context of convergence. Reconsider (16)

$$g(\theta^*) = S_d\left(\frac{\partial^2 y(\theta^*)}{\partial\theta^2} - P(\theta^*)\frac{\partial^2 u(\theta^*)}{\partial\theta^2}\right)$$
$$= \underbrace{\frac{S_0(\theta^*)}{C(\theta^*)}\left[\frac{\partial^2 C(\theta^*)}{\partial\theta^2} - 2\frac{\partial C(\theta^*)}{\partial\theta}\frac{\partial C(\theta^*)}{\partial\theta}^T\frac{T_0(\theta^*)}{C(\theta^*)}\right]}_{M_A(\theta^*)}$$
$$\times S_d C(\theta^*) S_0(\theta^*)(P_0 - P(\theta^*))r$$
$$= M_A(\theta^*)\varepsilon_m(\theta^*,\theta^*).$$

Then using the definition (17) and Parseval's formula,

$$\frac{\partial^2 \hat{J}(\theta^*,\theta^*)}{\partial\theta^2} - \frac{\partial^2 J(\theta^*)}{\partial\theta^2} = E\varepsilon_m(\theta^*,\theta^*)^T g(\theta^*)$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} M_A(e^{j\omega},\theta^*)\Phi_{\varepsilon_m}(\omega)d\omega, \quad (19)$$

where $\Phi_{\varepsilon_m}(\omega)$ is the spectrum of the control error $\varepsilon_m(\theta^*,\theta^*)$. Consequently, $\frac{\partial^2 \hat{J}(\theta^*,\theta^*)}{\partial\theta^2} - \frac{\partial^2 J(\theta^*)}{\partial\theta^2}$ is small if the achievable control performance, represented by $\varepsilon_m(\theta^*,\theta^*)$ is small. The same conclusion is also valid for the method where only $S_0(\theta)$ was linearized. The only difference is the matrix M_A which is specific for each method.

5. NUMERICAL EXAMPLES

The aim of this paper has been to try to develop a method that combines both local and global information about the cost function in an efficient way. It is thus of interest to examine how the algorithms behaves when the system to be controlled is nonlinear but such that it can be controlled by a linear controller. In such a situation the use of local information should be crucial for good performance at the same time as an overall model would be useful for fast convergence.

The proposed methods in Sec 3 will be illustrated using two simulation examples. In the control design direct parameterization of the controller via the model is considered, i.e. $C(\theta) = \frac{1}{G(\theta)}\frac{T_d}{1-T_d}$. The system to be controlled is the Hammerstein system $P = G_0(z)f(u)$, where $f(u) = \sqrt{(|u|)}sign(u)$ and $G_0(z) = \frac{0.03792z+0.03595}{z^2-1.845z+0.8521}$. The model for control will be of order one, i.e. $G(z,\theta) = \frac{b}{z-a}$ where a and b are the free parameters. The difference between the two examples is the reference model. The first reference model is $T_{d1} = \frac{0.07688}{z-0.9231}$ and the second is $T_{d2} = \frac{0.1479}{z-0.8521}$, i.e. the bandwidth of the second is twice the bandwidth of the first. The proposed methods will be compared with the iterative identification and control scheme in Sec 2.1 and IFT in Sec 2.2. The reference signal is two consecutive steps of amplitude ± 3, each of duration 250 time instants.

The contour plot of the cost (3) is given as a function of the model parameters a and b in Fig. 2. Three iterations are shown for the methods IFT (ift), iterative identification and control (itidc), Alg. 1 using (7) (linuy) and Alg. 1 using (11) (linSo). The optimum is denoted by optim in the plots. The step response for Alg. 1 in Fig. 3 shows the initial design (dashed line) and after three iterations (solid line) compared with the desired response (dash-dotted line). The experiment is repeated with $T_d = T_{d2}$ in Fig. 4-5.

Since the linear version of IFT is used, it will not converge to the true optimum, but it still works well, see (Hjalmarsson, 1998). The iterative identification and control scheme also works quite well when the performance requirements are low, but when the bandwidth increases the algorithm starts to diverge as is seen in Fig. 4-5. This is a fundamental problem with these methods. In these examples Alg. 1 and Alg. 2 presents a similar behavior as IFT, but with a slightly quicker decrease of the cost, especially in the first iteration.

6. CONCLUSIONS

In this contribution we have considered the problem of how to identify a restricted complexity model that is suitable for control design. We have proposed an iterative identification algorithm which blends information about the closed loop sensitivity to the model parameters, i.e. local information, together with the usual extrapolating property of a model, global system information. We have shown that the algorithm has the same stationary points as the control design

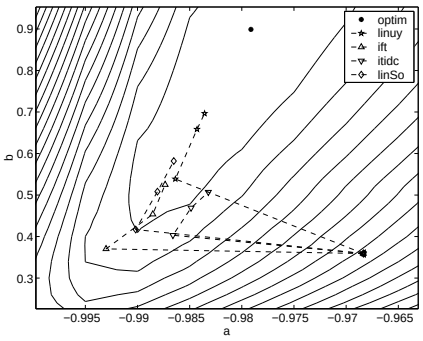

Fig. 2. Contour plot, $T_d = T_{d1}$.

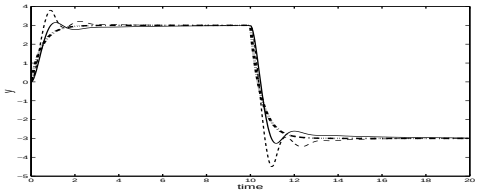

Fig. 3. The step response.

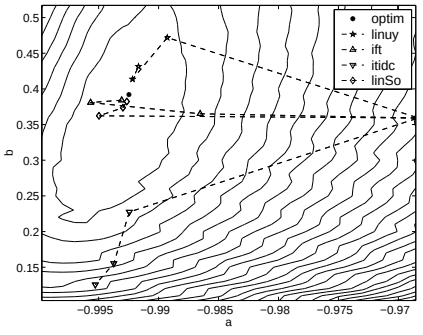

Fig. 4. Contour plot, $T_d = T_{d2}$.

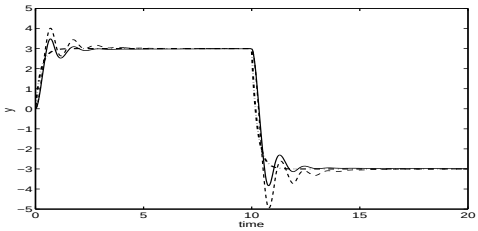

Fig. 5. The step response, $T_d = T_{d2}$.

criterion. Provided the approximated Hessian of the cost function is sufficiently accurate, the algorithm has also been shown to converge locally. We believe that these results, though preliminary, are of great interest given the sparse hard results for iterative identification for control schemes based on restricted complexity models.

REFERENCES

Åström, K. J. (1993). "Matching criteria for control and identification and identification". In: *Proceedings of European Control Conference*. Groningen, The Netherlands. pp. 248–251.

Campi, M.C., A. Lecchini and S.M Savaresi (2000). Virtual reference feedback tuning (vrft): a new direct approach to the design of feedback controllers. In: *Proc. 39th IEEE Conf. on Decision and Control*. Sydney, Australia.

Courant, R (1954). *Differential and Integral Calculus*. Blackie & Son Limited. London and Glasgow.

Hjalmarsson, H. (1998). "Control of nonlinear systems using Iterative Feedback Tuning". In: *Proc. 1998 American Control Conference*. Philadelphia. pp. 2083–2087.

Hjalmarsson, H. and K. Lindqvist (2001). "Identification for Control: L_2 and L_∞ Methods". In: *Proc. 40th IEEE Conf. on Decision and Control*. Orlando, FL.

Hjalmarsson, H., M. Gevers and O. Lequin (1998). "Iterative Feedback Tuning: Theory and Applications". *IEEE Control Systems Magazine* **18**(4), 26–41.

Hjalmarsson, H., S. Gunnarsson and M. Gevers (1994). "Model free data driven optimal tuning of controller parameters". In: *Technical report, Report LiTH-ISY-R-1680, Department of Electrical Engineering,*. Linköping University, Sweden.

Hjalmarsson, H., S. Gunnarsson and M. Gevers (1995). "Optimality and sub-optimality of iterative identification and control schemes". In: *Proc. of American Control Conference*. Vol. 4. pp. 2559–2563.

Horn, R. A. and C. R. Johnson (1991). *Topics in Matrix Analysis*. Cambridge University Press.

Jansson, H. and H. Hjalmarsson (2001). On the use of local and global information in identification for control. Internal Report IR–S3–REG–0104. Division of Automatic Control, Royal Institute of Technology (www.s3.kth.se).

Khalil, H. K. (1996). *Nonlinear systems*. Prentice Hall. Upper Saddle River, NJ.

Ljung, L. and T. Söderström (1983). *Theory and Practice of Recursive Identification*. MIT press. Cambridge, Mass.

Schrama, R. J. P. and O. H. Bosgra (1993). "Adaptive performance enhancement by iterative identification and control design". *Int. J. Adapt. Contr. & Signal Proc.* **7**, 475–487.

Sjöberg, J. and F. De Bruyne (1999). On a nonlinear controller tuning strategy. In: *14th IFAC World Congress*. Vol. I. Beijing, P.R. China. pp. 343–348.

Zang, Z., R.R. Bitmead and M. Gevers (1995). Iterative weighted least-squares identification and weighted LQG control design. *Automatica* **31**, 1577–1594.