A DECADE OF PROGRESS IN ITERATIVE PROCESS CONTROL DESIGN : FROM THEORY TO PRACTICE

Michel Gevers *,

* *CESAME, Université Catholique de Louvain*
Bâtiment EULER, 4 av. Georges Lemaitre, B-1348 Louvain-la-Neuve, Belgium *E-mail:* Gevers@csam.ucl.ac.be

Abstract:
One of the most active areas of research in the nineties has been the study of the interplay between system identification and robust control design. It has led to the development of "control-oriented identification design", the paradigm being that, since the model is only a tool for the design of a controller, its accuracy (or its error distribution) must be tuned towards the control design objective. This observation has led to the concept of "iterative identification and control design" and, subsequently, to model-free iterative controller design, in which the controller parameters are iteratively tuned on the basis of successive experiments performed on the real plant, leading to better and better closed loop behaviour. These iterative methods have found immediate applications in industry; they have also been applied to the optimal tuning of PID controllers. This paper presents the progress that has been accomplished in iterative process control design over the last decade. It is illustrated with applications in the chemical industry.

Keywords: identification for control, validation, iterative controller tuning

Acknowledgements

The work reported herein is the result of collaborations with a large number of co-authors over the last ten years. I would like to acknowledge B.D.O. Anderson, P. Ansay, R.R. Bitmead, X. Bombois, B. Codrons, F. De Bruyne, G.C. Goodwin, S. Gunnarsson, H. Hjalmarsson, C. Kulcsar, J. Leblond, O. Lequin, L. Ljung, B. Ninness, A.G. Partanen, M.A. Poubelle, G. Scorletti, L. Triest, P.M.J. Van den Hof, V. Wertz, and Z. Zang.

[1] The author acknowledges the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with the author.

1. INTRODUCTION

Iterative process control is a control design methodology that has emerged in the nineties as a result of intense research efforts aimed at bridging the gap between system identification and robust control analysis and design. In order to give the reader an idea about how wide this gap was, we quote from a keynote lecture delivered at the 1991 IFAC Symposium on Identification (Gevers, 1991) [1]

The last ten years have seen the emergence of robust control theory as a major research activity. During the same period, research in system identification has dwindled, and it might be tempting to believe that most of the theoretical questions in identification theory have been resolved for some

[1] Our extensive quote should of course not be construed as approval of the ideas expressed in that paper.

time. The surprising fact is that much of robust control theory is based on prior descriptions of model uncertainty or model errors which classical identification theory has been incapable of delivering. Conversely, until recently identification theorists have not spent much effort in trying to produce the accurate uncertainty bounds around their estimated models that their robust control design colleagues were taking for granted. It is as if, until a few years ago, the control design community and the identification community had not been talking much to each other. The gap between the surrealistic premises on which much of robust control design theory is built and the failure of identification theory to deliver accurate uncertainty bounds in the face of unmodelled dynamics has brought to light major deficiencies in both theories, and a sudden awareness from around 1988 of the need to understand better the interactions between both theories.

Surely, a natural place to search for an understanding of the interactions between identification and robust control design is in the adaptive control community. Indeed, adaptive control combines the design of an on-line identifier with that of a control law. ...... However, in their pursuit of an adaptive controller possessing some degree of robustness to unmodelled dynamics, the adaptive control community split into two camps. There were those who focused upon the specification of more and more sophisticated control law design procedures in the hope that a super control law would be able to overcome the vagaries of any identification method. And there were those who, by cajoling and forcing the parameter estimates to conform to certain boundedness conditions whatever the control law*, were aiming at developing a super identifier able to leap over any reasonable controller in a single bound. An essential feature of adaptive control, however, is that the identification is performed in closed loop and that the controller therefore impacts on the estimated model and on its quality (i.e. its error with respect to the true system). It is therefore to be expected that the separate designs of the identifier and of the controller without regard for the effect of the control law on the identified model, or of the identified model on the robustness of the control law, may not lead to a maximization of the global robustness of the identifier/controller schema.*

This last sentence was to become the program for much of the research activity in the nineties: going from separate designs to a synergistic design. In 1990, any observer of the scene was aware of the many different technical inconsistencies between the newly emerged robust control theory and the more classical prediction error identification theory. For example, prediction error identification theory had very little to offer in terms of quantification of the model error, and whatever tools that were available were totally incompatible with the frequency domain uncertainty descriptions required for robust control analysis and design.

However, the most crucial manifestation of the "identification/control gap" was not so much these technical incompatibilities, but rather the total absence of synergy between the two parts of the design: identification design and model-based control design. The prevailing philosophy at the time was: "First identify a model with a method that also allows the estimation of error bounds on this model; then design a controller based on this model and its error bounds." The problem is that an identification method whose sole merit is to deliver error bounds on a restricted complexity model may well produce a nominal model and an uncertainty set that are ill-suited for robust control design.

Due to a lack of understanding of the interplay between identification and robust control, most of the earlier work focused on producing suitable estimates of model quality (or uncertainty), and on bridging the gap between identification and robust control. The most obvious manifestation of this gap, and the one that triggered most of the research activity in the early nineties, was the realization that robust control theory requires a priori *hard bounds* on the model error, whereas classical identification theory delivers at best *soft bounds*, i.e. confidence ellipsoids in a probabilistic sense. This led to the development of new identification theories that were called "control-oriented" only because they delivered model uncertainty descriptions that were compatible with those required by robust control theory. The question of whether the identified models and their uncertainty descriptions were likely to deliver high performance controllers was not addressed, at least initially.

It later became clear that the great 'hard-versus-soft' debate was not the real issue. To quote from another plenary lecture (Gevers, 1993): *An identification and control design method that leads to a closed loop system that is stable with probability* 99% *is of course just as acceptable as an H_∞-based design that leads to a 'guaranteed stable' closed loop, but that is based on prior error bounds that cannot be verified.* However, even though the 'hard-versus-soft' question proved to be a non-issue, numerous other technical stumbling blocks had to be conquered before robust control analysis and design could be applied to models identified from data, rather than just to models and model uncertainty sets obtained from prior assumptions.

To summarize, the intense research effort of the nineties on identification for control has pursued two major objectives:

- Obtaining better estimation procedures for the quantification of the model uncertainty for identified models; in particular, produce uncertainty descriptions that are compatible with robust control theory.
- Understanding the interaction between identification and model-based control in order to produce control-oriented identification design guidelines.

In this paper, we shall mainly focus on the progress accomplished in *identification for control design*, i.e. the second issue. This line of research has led to such important new concepts as iterative model-based controller redesign, cautious model and controller updates, and eventually iterative model-free controller redesign. But before we venture in this direction, let us first briefly elaborate on the question of model uncertainty, if only to clearly distinguish it from the question of identification for control.

The quantification of the model error is of course a very important objective, whatever the goal of the identification step that has produced this model. A reputable engineer should never deliver a product to his client without some statement about the quality of that product, whether it be a machine tool, a measurement device, or a dynamical model. When the product is a model, and when the client is a robust control designer, then this client expects a model quality statement that is compatible with his/her robust control design tools. There is no sense telling the robust control engineer that the bias error on the delivered model is characterized by some complicated minimization formula, and that the noise-induced error is characterized by ellipsoidal confidence regions on a meaningless parameter vector, if all the control engineer can handle for his robust control design is a frequency domain error bound. When that happens - and this is exactly what did happen ten years ago - then the robust control engineer leaves the room in disgust and starts developing a new identification theory which he calls 'control-oriented', only because it can deliver model error bounds that are compatible with existing robust control theory.

All through the nineties, we have witnessed a tremendous activity, on the part of both communities, in the area of model quality estimation and model uncertainty description, with a view of bridging the technical incompatibilities between the two theories. We cannot possibly hope to reference the hundreds of relevant papers. One of the better surveys of this line of research, up to the middle of the decade, can be found in (Ninness and Goodwin, 1995).

The results on model quality are necessary for the construction of a synergistic design of identification and robust control, but they do not constitute this synergy. They are the technical building blocks. Indeed, identification **for** control is a "design" problem, as its name indicates. We now explain how to give meaning to this problem, and why the solution leads to iterative controller design.

2. FROM IDENTIFICATION FOR CONTROL TO ITERATIVE IDENTIFICATION AND CONTROL

2.1 *The setup*

All through this paper we consider the situation where there is a "*true system*" which, for the sake of analysis, we assume to be linear time-invariant. For the sake of simplicity, we also consider a single-input single-output system only in this paper. Thus, the true system is represented by

$$\mathcal{S} : y_t = G_0(z)u_t + v_t, \qquad (1)$$

where $G_0(z)$ is a linear time-invariant causal operator, y is the measured output, u is the control input, and v is noise, assumed to be quasistationary.

We now consider the situation where we can perform experiments on this system with the purpose of designing a feedback controller. We also consider that, most often, the system is already under feedback control, and that the task is to replace the present controller by one that achieves better performance. This situation is representative of very many practical industrial situations.

It is also typical of many industrial applications that the system to be controlled is very complex, and that it would therefore require a very high order dynamical model to represent the system with high fidelity. Any model-based control design procedure would therefore lead to a very high order controller, since the complexity of a model-based controller is of the same order as that of the system. The practical situation, considered in this paper, is where we want the to-be-designed controller to have low complexity.

2.2 *In search of a low complexity controller*

There are many ways of obtaining a low-complexity controller for a high order system. These include identification, model reduction, or controller reduction, in open or in closed loop, etc. A comparison between identification methods and reduction methods, on an industrial example, can be found in (Bendotti *et al.*, 1998).

Here we consider the rather obvious idea which consists of identifying a low order model from data collected on the real system, from which a model-based controller is then computed. Given that the low order model cannot possibly represent the true system over the whole frequency range, it will have a systematic error called the *bias error*, in addition to the inevitable noise-induced error called the *variance error*. This bias error - and indeed the total error - must be taken into account in the control design; hence the importance of producing methods for the estimation of model errors. But what is even much more important than estimating the model error a posteriori is to design the identification in such a way that the bias error does not harm the performance of the controller that will be designed on the basis of this approximate model. This is based on the observation that one can design a high performance controller with a model that has large error with respect to the real system (i.e. a very wrong model), as long as this model represents with high accuracy the dynamical features of the true system that are essential for control design. For example, it is essential that the model be very accurate around the crossover frequency of the to-be-designed closed loop system, but the error in its steady state gain can be huge.

The idea of tuning the bias error for control design is at the core of "identification for control". It is an *identification design* problem, whose objective is to produce, within a specified class of restricted complexity models, a nominal model whose bias error distribution is tuned towards the control design objective. As we shall see later in this paper, this can only be achieved through a succession of model and controller iterations; hence the iterative schemes that have emerged in identification for control.

The tuning of the bias error has led to iterative schemes for the estimation of a 'control-oriented' nominal model. However, to fully take advantage of robust control theory, one must develop an 'identification for control' theory not just for the nominal model, but also for the uncertainty regions around this nominal model. Indeed, robust control is a model-based design methodology in which the controller is designed on the basis of a nominal model together with an uncertainty region around the nominal model: see e.g. (Zhou *et al.*, 1995). It is the task of model validation to construct an uncertainty region around a nominal model. When the model and its uncertainty region are to be used for robust control design, then this validation step must also be tuned towards the control objective. This is a much harder problem for which few results are presently available. In the prediction error framework, recent results have been obtained in (Gevers *et al.*, 1999) and (Gevers *et al.*, 2000). We shall not elaborate on them in this paper, where we focus on iterative designs.

2.3 *Matching identification and control objectives*

We have described the context in which we operate, and we have introduced the motivation for identification for control. We now show that the matching of identification and control objectives leads to iterative identification and control design. To illustrate the need for iterative design, we take the simplest possible control design objective: model reference control. Thus, consider the true system (1) and suppose we have identified a model $\hat{G}(z) \stackrel{\triangle}{=} G(z,\hat{\theta})$ of G_0, from some parametrized set of low order models $\{G(z,\theta) \mid \theta \in D_\theta\}$. Consider a control law

$$u_t = C(z)[r_t - y_t], \tag{2}$$

and assume that our control design objective is to design $C(z)$ such that the closed loop transfer function from v_t to y_t is some prespecified $S(z)$. Then, given a model $\hat{G}(z)$, the controller $C(z)$ is computed from

$$\frac{1}{1+\hat{G}(z)C(z)} = S(z), \tag{3}$$

We compare the real closed loop system of Figure 1 with the designed closed loop system of Figure 2, with both loops driven by the same reference signal r_t.

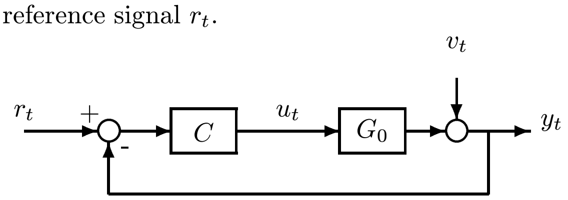

Fig. 1. Actual closed loop system

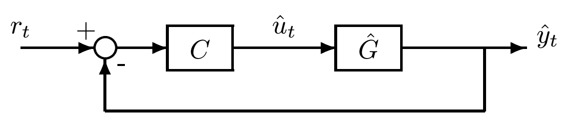

Fig. 2. Designed (or nominal) closed loop system

Now, staring at Figures 1 and 2, and remembering that there are no disturbances, one observes that:

$$\begin{aligned} y_t &= \frac{G_0C}{1+G_0C}r_t + \frac{1}{1+G_0C}v_t \\ u_t &= \frac{C}{1+G_0C}r_t - \frac{C}{1+G_0C}v_t, \\ \hat{y}_t &= \frac{\hat{G}C}{1+\hat{G}C}r_t. \end{aligned} \tag{4}$$

The 'control performance error'[2], defined as the error between the actual and the designed outputs, is given by:

$$y_t - \hat{y}_t = \left[\frac{G_0 C}{1+G_0 C} - \frac{\hat{G}C}{1+\hat{G}C}\right] r_t + \frac{1}{1+G_0 C}\, v_t \quad (5)$$

After some straightforward manipulations, this can be rewritten as

$$y_t - \hat{y}_t = S(z)[y_t - G(z,\hat{\theta})u_t]. \quad (6)$$

Equation (6) can be seen as an equality between a control performance error on the left hand side (LHS) and a filtered identification error on the right hand side (RHS). Indeed, the RHS is a filtered (by $S(z)$) version of the output error $y_t - G(z,\hat{\theta})u_t$, where u_t and y_t are collected on the actual closed loop system of Figure 1. Thus, it appears that if θ is obtained by minimizing the Mean Square of the RHS of (6), i.e. by closed loop identification with a filter $S(z)$, then this will minimize the Mean Square control performance error. In other words, apparently (6) shows that we get a perfect match between control error and identification error. However, life is more subtle and complicated. Indeed, the controller $C(z)$ is also a function of the model parameter vector θ via (3). Since the data collected on the real closed loop system of Figure 1 are a function of $C(z)$, they are also dependent on θ. Hence, a more suggestive and correct way to write (6) is as follows:

$$y_t - \hat{y}_t = S(z)[y_t(\theta) - G(z,\theta)u_t(\theta)]. \quad (7)$$

Even though the RHS of (7) looks like a closed loop prediction error, it cannot be minimized by standard identification techniques, because θ appears everywhere and not just in $G(z,\theta)$. We have illustrated the fact that with the simplest possible control design mechanism, namely Model Reference Control, one can apparently equate a 'control performance error' to an 'identification error', but that this identification error cannot be minimized by standard identification techniques because the parameter vector appears in more than just the model. In other words, we know that, to make the control error small, we should minimize the RHS of (7) with respect to θ, but we don't know how to do this.

For optimization-based control design criteria, the control performance criterion also defines an identification criterion that one would want to minimize with respect to the model parameters: see (Zang *et al.*, 1995), (Van den Hof and Schrama, 1995). This identification error is typically a norm of the following error:

$$y_t - \hat{y}_t = S(z,\theta)[y_t(\theta) - G(z,\theta)u_t(\theta)], \quad (8)$$

where the data filter $S(z,\theta)$ is proportional to the sensitivity function of the design loop (compare with (3)) and is now also θ-dependent.

As a consequence, the approach suggested in all known 'identification for control' schemes is to perform identification and control design steps in an iterative way, whereby the i-th identification step is performed on filtered closed loop data collected on the actual closed loop system with the $(i-1)$-th controller operating in the loop. This corresponds to an i-th identification step in which the following filtered prediction error is minimized with respect to θ, for fixed $\hat{\theta}_{i-1}$:

$$y_t - \hat{y}_t = S(z,\hat{\theta}_{i-1}) \times [y_t(\hat{\theta}_{i-1}) - G(z,\theta)u_t(\hat{\theta}_{i-1})]. \quad (9)$$

We refer the reader to (Gevers, 1993), (Bitmead, 1993) and (Van den Hof and Schrama, 1995) for details and for a survey on such iterative schemes.

An interesting question is whether these iterative identification and control schemes converge to the minimum of the achieved cost over the set $\mathcal{C} \triangleq \{C(G(z,\theta))\ \forall\theta \in D_\theta\}$ of all certainty equivalence controllers. This corresponds to asking whether by successively minimizing over θ the mean square of the prediction errors defined by (9) one will converge to the minimum of

$$J(\theta) \triangleq E\{S(z,\theta)[y_t(\theta) - G(z,\theta)u_t(\theta)]\}^2. \quad (10)$$

This question was analyzed in (Hjalmarsson *et al.*, 1995*b*) for model reference control; it was shown there that the iterative identification and control schemes do not generically converge to the minimum of the achieved control cost.

This does not mean that iterative identification and control schemes have failed. In fact, the idea of using available data, collected on the actual closed loop system, to obtain a model that is better suited for the design of a new controller, has found immediate and widespread applications because it is easy and intuitively reasonable. In typical applications large numbers of closed loop data are flowing into the control computer, and it makes a lot of sense to use these data to replace the existing controller by one that achieves better performance. In addition, the theoretical work on iterative model and controller adjustments has shown that, in order to compute a new controller with better performance, the optimal experiment is to perform closed loop identification. Thus, no

[2] It was called that way in (Åström, 1993), (Åström and Nilsson, 1994) because, if the closed loop transfer function of the actual system was equal to the reference model $S(z)$, this error would contain only the noise contribution.

special experiments are required, and there is no need to "open the loop" in order to design the new model and the corresponding new controller.

Thus, this is one area where the transfer of technology from theoretical research to applications has been extremely fast. The first applications of control-oriented identification and iterative model-based controller tuning were reported within months after the theoretical results had been produced. Representative examples can be found in (Partanen and Bitmead, 1995), (Schrama and Bosgra, 1993), (de Callafon *et al.*, 1993), (de Callafon, 1998). The practical impact of iterative closed loop identification and controller redesign has been assessed in (Landau, 1999*a*), where some interesting observations are made on the distinction between this batch-like mode of operation and the more classical theory and methods of adaptive control.

The guidelines that emerged during the nineties for the application of iterative identification and control schemes were supported by intense research that brought to light two essential features.

- The benefits of *closed loop identification* when the model is identified with a view of designing a new controller: see e.g. (Liu and Skelton, 1990), (Schrama, 1992*a*), (Hakvoort *et al.*, 1994), (Lee *et al.*, 1995), (Hjalmarsson *et al.*, 1996), (Landau, 1999*b*). This produced a revival of interest in the design of closed loop identification methods: see e.g. (Hansen *et al.*, 1989), (Van den Hof and Schrama, 1993), (Van den Hof *et al.*, 1995).
- The need for *cautious adjustments* between successive model and controller updates, in order to guarantee closed loop stability or performance improvement with the new controller: see e.g. (Schrama, 1992*b*), (Bitmead *et al.*, 1997), (Anderson *et al.*, 1998) and (Anderson and Gevers, 1998).

Despite its practical successes, the failure of all attempts to establish convergence of iterative identification and control schemes was a major worry, more from a theoretical than from a practical point of view. Indeed, in practice it was observed that major improvements in performance of the closed loop systems were obtained within the first few iterations, after which the improvements were very minor. Divergence typically occurred only if one continued to iterate beyond these initial steps.

It is the analysis of (Hjalmarsson *et al.*, 1995*b*) that revealed the reason for the possible divergence. This analysis led the authors to reformulate the iterative identification and control design scheme as a parameter optimization problem, in which the optimization is carried directly on the controller parameters, thereby abandoning the identification step altogether. This idea led to a gradient-based algorithm for the iterative optimization of the parameters of any restricted complexity controller (Hjalmarsson *et al.*, 1994), which was later called IFT, for Iterative Feedback Tuning. In the next section we describe the IFT algorithm and some of its more recent developments.

3. IFT: A MODEL-FREE ITERATIVE CONTROLLER RETUNING METHOD

3.1 *Introduction to the IFT algorithm*

The key feature of the IFT algorithm is that an unbiased estimate of the gradient of a control performance criterion is computed from signals obtained from closed loop experiments with the present controller operating on the actual system. For a controller of given (typically low-order) structure, the minimization of the criterion is then performed iteratively by a Gauss-Newton based scheme. For a two-degree-of-freedom controller, three batch experiments are performed at each step of the iterative design. The first and third simply consist of collecting data under normal operating conditions; the only real experiment is the second batch which requires feeding back, at the reference input, the output measured during normal operation. Hence the acronym Iterative Feedback Tuning (IFT) given to this scheme. For a one-degree-of-freedom controller, only the first and third experiments are required. No identification procedure is involved.

The optimal IFT scheme was initially derived in (Hjalmarsson *et al.*, 1994). Given its simplicity, it became clear that this new scheme had wide-ranging potential, from the optimal tuning of simple PID controllers to the systematic design of controllers of increasing complexity that have to meet some prespecified specifications. In particular, the IFT method is appealing to process control engineers because the controller parameters can be successively improved without ever opening the loop. In addition, in many process control applications the objective of the controller design is to achieve disturbance rejection. With the IFT scheme the tuning of the controller parameters for disturbance rejection is driven by the disturbances themselves; there is no need for the injection of a persistently exciting reference signal.

Since 1994, much experience has been gained with the IFT scheme:

- It has been shown to compare favourably with identification-based schemes in simulation examples: see (Hjalmarsson *et al.*, 1994), and its accuracy has been analyzed in (Hjalmarsson and Gevers, 1997).

- It has been successfully applied to the flexible transmission benchmark problem posed by I.D. Landau for ECC95, where it achieved the performance specifications with the simplest controller structure (Hjalmarsson *et al.*, 1995*a*).
- It has been tested on the flexible arm of the Laboratoire d'Automatique de Grenoble (Ceysens and Codrons, 1995), on a ball-on-beam system (De Bruyne and Carrette, 1997), for the temperature control of a water tube and for the control of a suspended plate (Molenaar, 1995).
- It has been adapted to linear time-invariant MIMO systems (Hjalmarsson and Birkeland, 1998) and to time varying, and in particular periodically time-varying, systems (Hjalmarsson, 1995).
- It has been applied by the chemical multinational Solvay S.A. to the tuning of PID controllers for a number of critical control loops for which opening the loop or creating limit cycles for PID tuning was not allowed.

Here we present the fundamentals of the IFT algorithm, and we then review the performance achieved by the scheme at S.A. Solvay.

3.2 *The basic control design criterion*

We present here a basic version of the IFT algorithm; we refer to (Hjalmarsson *et al.*, 1998) for a more complete derivation and discussion. We consider the unknown true system (1), to be controlled by a two degrees of freedom controller:

$$u_t = C_r(\rho) r_t - C_y(\rho) y_t \tag{11}$$

where $C_r(\rho)$ and $C_y(\rho)$ are linear time-invariant transfer functions parametrized by some parameter vector $\rho \in \mathbf{R}^{n_\rho}$, and $\{r_t\}$ is an external reference signal, independent of $\{v_t\}$: see Figure 3.

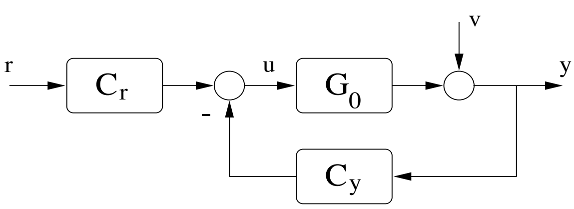

Fig. 3. Block diagram of the closed loop system

To ease the notation we will from now on omit the time argument of the signals; however, we shall use the notation $y(\rho)$, $u(\rho)$ for the output and the control input of the system (1) in feedback with the controller (11), in order to make explicit the dependence of these signals on the controller parameter vector ρ.

Let y^d be a desired output response to a reference signal r for the closed loop system. This response may be defined as the output of a reference model T_d, i.e. $y^d = T_d r$, but for the IFT method knowledge of the signal y^d is sufficient. The error between the achieved and the desired response is

$$\begin{aligned}\tilde{y}(\rho) &\stackrel{\triangle}{=} y(\rho) - y^d \\ &= \left(\frac{C_r(\rho)G_0}{1 + C_y(\rho)G_0} r - y^d \right) + \frac{1}{1 + C_y(\rho)G_0} v.\end{aligned} \tag{12}$$

If a reference model is used this error can also be written as

$$\tilde{y}(\rho) = \left(\frac{C_r(\rho)G_0}{1 + C_y(\rho)G_0} - T_d \right) r + \frac{1}{1 + C_y(\rho)G_0} v. \tag{13}$$

This error consists of a contribution due to incorrect tracking of the reference signal r and an error due to the disturbance. With IFT the following quadratic control performance criterion is used:

$$J(\rho) = \frac{1}{2N} \mathrm{E} \left[\sum_{t=1}^{N} (\tilde{y}_t(\rho))^2 + \lambda \sum_{t=1}^{N} (u_t(\rho))^2 \right] \tag{14}$$

but any other differentiable signal-based criterion can be used. In (14) $\mathrm{E}[\cdot]$ denotes expectation w.r.t. the weakly stationary disturbance v. The optimal controller parameter ρ is defined by

$$\rho^* = \arg\min_{\rho} J(\rho), \tag{15}$$

The errors $\tilde{y}_t(\rho)$ and $u_t(\rho)$ appearing in the criterion can be filtered by frequency weighting filters L_y and L_u to give added flexibility to the design: see (Hjalmarsson *et al.*, 1998).

Let $T_0(\rho)$ and $S_0(\rho)$ denote the achieved closed loop response and sensitivity function with the controller $\{C_r(\rho), C_y(\rho)\}$, i.e.

$$T_0(\rho) = \frac{C_r(\rho)G_0}{1 + C_y(\rho)G_0}, \qquad S_0(\rho) \frac{1}{1 + C_y(\rho)G_0}. \tag{16}$$

Given the independence of r and v, $J(\rho)$ can be written as

$$\begin{aligned}J(\rho) = &\frac{1}{2N} \sum_{t=1}^{N} \left\{ (y_t^d - T_0(\rho) r_t) \right\}^2 \\ &+ \frac{1}{2} \mathrm{E} \left[\left\{ S_0(\rho) v \right\}^2 \right] + \lambda \frac{1}{2N} \mathrm{E} \left[\sum_{t=1}^{N} (u_t(\rho))^2 \right].\end{aligned} \tag{17}$$

The first term is the tracking error, the second term is the disturbance contribution, and the last term is the penalty on the control effort.

3.3 *Criterion minimization*

We now address the minimization of $J(\rho)$ given by (14) with respect to the controller parameter vector ρ for a controller of specified structure. It is evident from (12) that $J(\rho)$ depends in a fairly complicated way on ρ, on the unknown system G_0 and on the unknown spectrum of $\{v\}$. To obtain the minimum of $J(\rho)$ we would like to find a solution for ρ to the equation

$$\begin{aligned} 0 &= \frac{\partial J}{\partial \rho}(\rho) \qquad (18) \\ &= \frac{1}{N}\mathrm{E}\left[\sum_{t=1}^{N} \tilde{y}_t(\rho)\frac{\partial \tilde{y}_t}{\partial \rho}(\rho) + \lambda \sum_{t=1}^{N} u_t(\rho)\frac{\partial u_t}{\partial \rho}(\rho)\right]. \end{aligned}$$

If the gradient $\frac{\partial J}{\partial \rho}$ could be computed, then the solution of (19) would be obtained by the following iterative algorithm:

$$\rho_{i+1} = \rho_i - \gamma_i R_i^{-1} \frac{\partial J}{\partial \rho}(\rho_i). \qquad (19)$$

Here R_i is some appropriate positive definite matrix, typically a Gauss-Newton approximation of the Hessian of J, while the sequence γ_i must obey some constraints for the algorithm to converge to a local minimum of the cost function $J(\rho)$: see (Hjalmarsson *et al.*, 1994).

Such problem can be solved by using a stochastic approximation algorithm of the form (19) such as suggested in (Robbins and Monro, 1951), provided the gradient $\frac{\partial J}{\partial \rho}(\rho_i)$ evaluated at the current controller can be replaced by an unbiased estimate. In order to solve this problem, one thus needs to generate the following quantities:

(1) the signals $\tilde{y}(\rho_i)$ and $u(\rho_i)$;
(2) the gradients $\frac{\partial \tilde{y}}{\partial \rho}(\rho_i)$ and $\frac{\partial u}{\partial \rho}(\rho_i)$;
(3) unbiased estimates of the products $\tilde{y}(\rho_i)\frac{\partial \tilde{y}}{\partial \rho}(\rho_i)$ and $u(\rho_i)\frac{\partial u}{\partial \rho}(\rho_i)$.

The computation of the last two quantities has always been the key stumbling block in solving this direct optimal controller parameter tuning problem. The main contribution of (Hjalmarsson *et al.*, 1994) was to show that these quantities can indeed be obtained by performing experiments on the closed loop system formed by the actual system in feedback with the controller $\{C_r(\rho_i), C_y(\rho_i)\}$. This is done as follows.

3.4 *The IFT algorithm*

At iteration i of the controller tuning algorithm, the controller $C(\rho_i) \triangleq \{C_r(\rho_i), C_y(\rho_i)\}$ operates on the actual plant. We then perform three experiments, each of which consists of collecting a sequence of N data. Two of these experiments (the first and third) just consist of collecting data under normal operating conditions; the second is a real (i.e. *special*) experiment. We denote N-length reference signals by $\{r_i^j\}$, $j = 1, 2, 3$, and the corresponding output signals by $\{y^j(\rho_i)\}$, $j = 1, 2, 3$. Thus we have

$$\begin{aligned} r_i^1 &= r, \quad \text{yielding} \qquad (20) \\ & y^1(\rho_i) = T_0(\rho_i)r + S_0(\rho_i)v_i^1; \\ r_i^2 &= r - y^1(\rho_i), \quad \text{yielding} \qquad (21) \\ & y^2(\rho_i) = T_0(\rho_i)(r - y^1(\rho_i)) + S_0(\rho_i)v_i^2; \\ r_i^3 &= r, \quad \text{yielding} \qquad (22) \\ & y^3(\rho_i) = T_0(\rho_i)r + S_0(\rho_i)v_i^3. \end{aligned}$$

Here v_i^j denotes the disturbance acting on the system during experiment j at iteration i. These experiments yield an exact realization of $\tilde{y}(\rho_i)$:

$$\tilde{y}(\rho_i) = y^1(\rho_i) - y^d, \qquad (23)$$

while it is shown in (Hjalmarsson *et al.*, 1998) that

$$\begin{aligned} \widehat{\frac{\partial y}{\partial \rho}}(\rho_i) \triangleq \frac{1}{C_r(\rho_i)} &\left[\left(\frac{\partial C_r}{\partial \rho}(\rho_i) - \frac{\partial C_y}{\partial \rho}(\rho_i\right) y^3(\rho_i)\right. \\ &\left. + \frac{\partial C_y}{\partial \rho}(\rho_i) y^2(\rho_i)\right] \qquad (24) \end{aligned}$$

is an unbiased estimate of $\frac{\partial y}{\partial \rho}(\rho_i)$.

The three experiments described above generate corresponding control signals:

$$\begin{aligned} u^1(\rho_i) &= S_0(\rho_i)\left[C_r(\rho_i)r - C_y(\rho_i)v_i^1\right], \\ u^2(\rho_i) &= S_0(\rho_i)\left[C_r(\rho_i)(r - y^1(\rho_i)) - C_y(\rho_i)v_i^2\right], \\ u^3(\rho_i) &= S_0(\rho_i)\left[C_r(\rho_i)r - C_y(\rho_i)v_i^3\right]. \end{aligned}$$

These signals can similarly be used to generate the estimates of the input related signals required for the estimation of the gradient (19). Indeed, $u^1(\rho_i)$ is a perfect realization of $u(\rho_i)$,

$$u(\rho_i) = u^1(\rho_i), \qquad (25)$$

while

$$\begin{aligned} \widehat{\frac{\partial u}{\partial \rho}}(\rho_i) \triangleq \frac{1}{C_r(\rho_i)} &\left[\left(\frac{\partial C_r}{\partial \rho}(\rho_i) - \frac{\partial C_y}{\partial \rho}(\rho_i)\right) u^3(\rho_i)\right. \\ &\left. + \frac{\partial C_y}{\partial \rho}(\rho_i) u^2(\rho_i)\right] \qquad (26) \end{aligned}$$

is an unbiased estimate of $\frac{\partial u}{\partial \rho}(\rho_i)$.

An experimentally based estimate of the gradient of J can be formed by taking

$$\begin{aligned} &\widehat{\frac{\partial J}{\partial \rho}}(\rho_i) = \qquad (27) \\ &\frac{1}{N}\sum_{t=1}^{N}\left(\tilde{y}_t(\rho_i)\widehat{\frac{\partial y_t}{\partial \rho}}(\rho_i) + \lambda u_t(\rho_i)\widehat{\frac{\partial u_t}{\partial \rho}}(\rho_i)\right). \end{aligned}$$

The next controller parameters are then computed by replacing, in the iteration (19), the gradient of the cost criterion by this estimate:

$$\rho_{i+1} = \rho_i - \gamma_i R_i^{-1} \widehat{\frac{\partial J}{\partial \rho}}(\rho_i) \tag{28}$$

where $\{\gamma_i\}$ is a sequence of positive real numbers that determines the step size, and where $\{R_i\}$ is a sequence of positive definite matrices. The key feature of our construction of $\widehat{\frac{\partial J}{\partial \rho}}(\rho_i)$, and also the motivation for the third experiment, is that this estimate is unbiased:

$$\mathrm{E}\left[\widehat{\frac{\partial J}{\partial \rho}}(\rho_i)\right] = \frac{\partial J}{\partial \rho}(\rho_i), \tag{29}$$

As a result, the controller parameters converge under reasonable conditions to a stationary point of the performance criterion, provided the sequence of controllers along the way are all stabilizing: see (Hjalmarsson *et al.*, 1998).

A number of implementation issues as well as design choices are addressed in detail in (Hjalmarsson *et al.*, 1998). They concern the choice of the step size γ_i and of the matrix R_i in (28), the choice of frequency weighting filters, the elimination of possibly unstable controllers in the filtering operations (24) and (26), the enforcement of integral action, the attenuation of the effect of disturbances, as well as the simplification that occurs in the case of a one-degree-of-freedom controller, where the third experiment is not necessary. One interesting design parameter is the step size, which determines how much the controller changes from one iteration to the next one. Before implementing a new controller one can compare its Bode plot with that of the previous one, and possibly reduce the step size if one feels that the change is too large. Similarly, one can predict the effect of a new controller on the closed loop response and on the achieved cost using a Taylor series expansion: see (Hjalmarsson *et al.*, 1998).

4. APPLICATIONS OF IFT IN CHEMICAL PROCESS CONTROL

The IFT scheme has been applied by the chemical multinational Solvay S.A. for the optimal tuning of industrial PID controllers operating on a range of different control loops. In each of these loops, PID controllers were already operating. Important performance improvements were achieved using the IFT method, both in tracking and in regulation applications. The reductions in variance achieved after a few (typically 2 to 6) iterations of the algorithm ranged from 25 % in a flow regulation problem in an evaporator, to 87 % in a temperature control problem for the tray of a distillation column. Here we present the results obtained with a temperature regulation problem for a tray of a distillation column. An application to a flow control problem in an evaporator is presented in (Hjalmarsson *et al.*, 1998).

The controller used was an industrial 2-degree-of-freedom PID controller where the derivative action is applied to y only, and where a first order filter is applied to y in order to limit the gain of the controller at high frequencies when the derivative action is used. The time constant of this filter is chosen as $Td/8$, Td being the derivative time constant. The PID regulator parameters were iteratively tuned using the IFT scheme, with the following design choices: Gauss-Newton direction, step-size $\gamma_i = 1\ \forall i$, control weighting $\lambda = 0$, sampling period of 15 seconds, $r_d = y_d = 0$ during 5 hours. The deadtime and the time constants of the process were unknown.

Figure 4 presents temperature deviations with respect to setpoint in a tray of a distillation column, over a 24-hour period, first with the original PID parameters, then with the PID controller obtained after 6 iterations of the new scheme. Figure 5 shows the corresponding histograms of these deviations over 2-week periods. The control error has been reduced by 70 %.

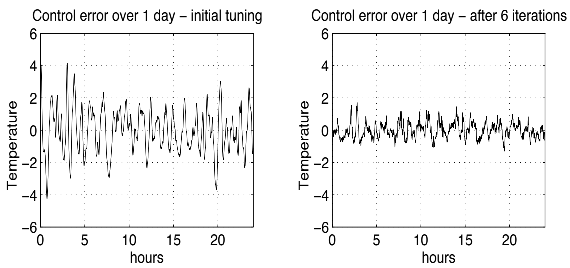

Fig. 4. Control error over a 24-hour period before tuning and after 6 iterations of IFT

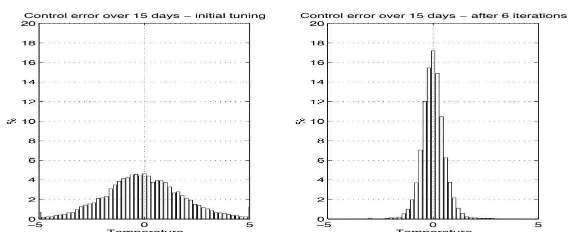

Fig. 5. Histogram of control error over 2-week period before tuning and after 6 iterations of IFT

Figure 6 shows the Bode plots of the two-degree of freedom controller (C_r, C_y) before optimal tuning (full line), after 3 iterations of the IFT algorithm (dashed line) and after 6 iterations (dotted line). The gain was too low and the derivative action underused.

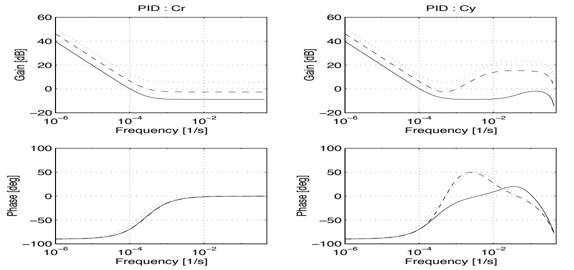

Fig. 6. Bode diagram of the two-degree-of-freedom controller before tuning (full), after 3 iterations (dashed) and after 6 iterations of the algorithm (dotted).

Table 1 shows the measured cost J with the 6 successive controllers, as well as the predicted value of the cost, calculated at each iteration with the new controller parameters, as explained above. The prediction was good except for the 2nd iteration which was perturbed by an abnormal disturbance.

Iteration	Cost (measured)	Next cost (predicted)
1	0.80	0.36
2	1.00	0.59
3	0.57	0.35
4	0.37	0.18
5	0.22	0.15
6	0.14	0.11

Table 1 : Calculated and predicted cost

5. MINIMIZING THE SETTLING TIME WITH IFT

The criterion (14) is well suited when the objective is to follow a specific reference trajectory, but is not so appropriate if the objective is to change the output from one setpoint to another one. Indeed, in such case the goal is typically to reach the new setpoint with a minimum settling time, and one does not care about the transient trajectory, provided it does not produce too much overshoot. By constraining the output to follow some particular reference trajectory y^d during the transient, one puts too much emphasis on the transient phase of the response at the expense of the settling time at the new setpoint value.

To cope with this situation Lequin observed in (Lequin, 1997) that one can add nonnegative weighting factors to each element of $\tilde{y}_t$ and u_t in the criterion (14). A simple way to obtain a satisfactory closed loop response to a desired setpoint change is then to set the weighting factors on $\tilde{y}_t$ to zero during the transient period and to one afterwards, while the weights on the control are put to one everywhere:

$$J_m(\rho) = \frac{1}{2N}\mathrm{E}\left[\sum_{t=t_0}^{N}(\tilde{y}_t(\rho))^2 + \lambda \sum_{t=1}^{N}(u_t(\rho))^2\right].$$

We say in such case that a *mask* of length t_0 is put on the transient response of the tracking error. Often it is not known a priori how much time is required to achieve a setpoint change without overshoot. In such case, one can perform the IFT iterations by initially applying a long mask, and then gradually reducing the length of this length of this mask until oscillations start occuring.

We illustrate this idea with an example presented in (Lequin *et al.*, 1999). Consider the plant

$$G(s) = \frac{1}{s^2 + 0.1s + 1}$$

One wishes to tune a PID controller in order to achieve a settling time of 20 seconds for the closed loop system. The initial PID parameter values were taken as $K = 0.025$, $T_i = 2$ and $T_d = 1$, yielding the very sluggish response shown in Figure 7.

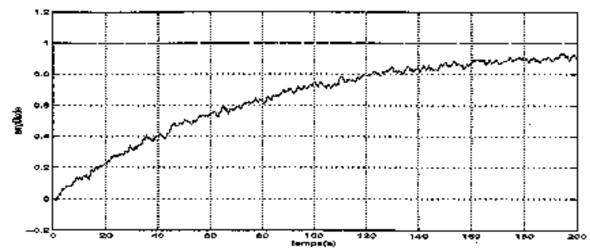

Fig. 7. Closed loop step response with initial PID parameters

The classical IFT criterion was then applied with a desired response shown in dotted line in Figure 8, with the achieved response shown in full line on that same figure. This response is very unsatisfactory, in large part due to an unfortunate choice of initial parameters.

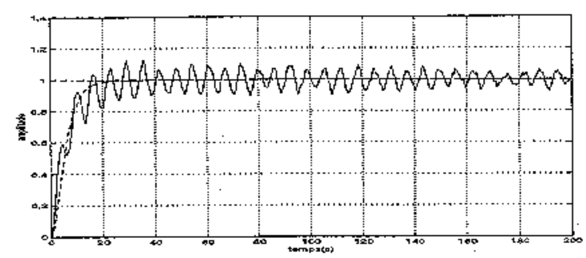

Fig. 8. Optimal closed loop step response (full) obtained with the classical IFT criterion and using the desired response (dashed)

The IFT criterion was then applied with a mask of decreasing length, with an initial length of 80 seconds, and with the same initial parameters. At every iteration of the IFT scheme, the length of the mask was decreased by 20 seconds, until a mask of length 20 was reached. This led to the closed loop response shown in Figure 9.

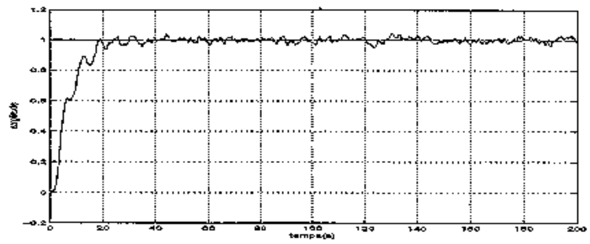

Fig. 9. Optimal closed loop step response obtained with the IFT criterion using masks of decreasing length

Observe the dramatic improvement of the closed loop response.

6. CONCLUSIONS

Iterative redesign of controllers using data collected on the operating closed loop system has emerged as a new, powerful and successful control design methodology, as a result of significant progress accomplished in the nineties on the understanding of the interplay between identification and control design. Most of the schemes are based on model and controller updates; they require safeguards such as cautious changes between successive controllers. The study of these 'identification for control' schemes has somewhat surprisingly also led to iterative schemes that are entirely model-free.

In this paper we have focused on the design of the nominal model and/or controller via these iterative schemes, since these have given rise to the more practical design methods, well suited for process control applications. We have barely touched upon the vast amount of progress accomplished on model uncertainty estimation, and have completely left aside our very recent theoretical work on model and robust controller validation.

7. REFERENCES

Anderson, B.D.O. and M. Gevers (1998). Fundamental problems in adaptive control. *Perspectives in Control, a Tribute to Ioan Doré Landau.*

Anderson, B.D.O., M. Gevers and C. Kulcsár (1998). Caution in iterative modeling and control design. In: *IFAC Workshop on Adaptive Control and Signal Processing.* Glasgow.

Åström, K.J. (1993). Matching criteria for control and identification. In: *Proc. 2nd European Control Conference.* Groningen, The Netherlands. pp. 248–251.

Åström, K.J. and J. Nilsson (1994). Analysis of a scheme for iterated identification and control. In: *Proc. IFAC Symp. on Identification.* Copenhagen, Denmark. pp. 171–176.

Bendotti, P., B. Codrons, C. M. Falinower and M. Gevers (1998). Control oriented low order modelling of a complex pwr plant: a comparison between open loop and closed loop methods. In: *Proc. 37th IEEE Conference on Decision and Control.* Tampa,Florida. pp. 3390–3395.

Bitmead, R.R. (1993). Iterative control design approaches. In: *Proc. 12th IFAC World Congress.* Vol. 9. Sydney, Australia. pp. 381–384.

Bitmead, R.R., M. Gevers and A.G. Partanen (1997). Introducing caution in iterative controller design. In: *11th IFAC Symp. on System Identification.* Fukuoka, Japan. pp. 1701–1706.

Ceysens, B. and B. Codrons (1995). Synthèse itérative de controleurs sans identification. Technical report. Masters Thesis. Université Catholique de Louvain.

De Bruyne, F. and P. Carrette (1997). Synthetic generation of the gradient for an iterative controller optimization method. *CD-ROM of 4th European Control Conference, paper THA-F2.*

de Callafon, R.A. (1998). Feedback Oriented identification for enhanced and robust control: a fractional approach applied to a wafer stage. PhD thesis. Delft University of Technology.

de Callafon, R.A., P.M.J. Van den Hof and M. Steinbuch (1993). Control-relevant identification of a compact disc pick-up mechanism. In: *Proc. 1993 IEEE Conf. on Decision and Control.* San Antonio, Texas. pp. 2050–2055.

Gevers, M. (1991). Connecting identification and robust control: A new challenge. In: *Proc. IFAC/IFORS Symposium on Identification and System Parameter Estimation.* Budapest, Hungary. pp. 1–10.

Gevers, M. (1993). Towards a joint design of identification and control?. In: *Essays on Control: Perspectives in the Theory and its Applications, Birkhauser, New York* (H.L. Trentelman and J.C. Willems, Eds.). Birkhauser. New York. pp. 111–151.

Gevers, M., X. Bombois, B. Codrons, F. De Bruyne and G. Scorletti (1999). The role of experimental conditions in model validation for control. In: *Robustness in Identification and Control - Proc. of Siena Workshop, July 1998* (A. Garulli, A. Tesi and A. Vicino, Eds.). Vol. 245 of *Lecture Notes in Control and Information Sciences.* Springer-Verlag. pp. 72–86.

Gevers, M., X. Bombois, B. Codrons, G. Scorletti and B.D.O. Anderson (2000). Model validation for control and controller validation: a prediction error identification approach. *submitted for publication to Automatica*.

Hakvoort, R.G., R.J.P. Schrama and P.M.J. Van den Hof (1994). Approximate identification with closed-loop performance criterion and application to LQG feedback design. *Automatica* **30**, 679–690.

Hansen, F., G. Franklin and R. Kosut (1989). Closed-loop identification via the fractional representation: Experiment design. *Proc. ACC*.

Hjalmarsson, H. (1995). Model-free tuning of controllers: Experience with time-varying linear systems. *Proc. 3rd European Control Conference* pp. 2869–2874.

Hjalmarsson, H. and M. Gevers (1997). Frequency domain expressions of the accuracy of a model-free control design scheme. In: *Proc. 11th IFAC Symp. on System Identification*. Vol. 1. Fukuoka, Japan. pp. 135–140.

Hjalmarsson, H. and T. Birkeland (1998). Iterative Feedback Tuning of linear time-invariant MIMO systems. *submitted to CDC 98*.

Hjalmarsson, H., M. Gevers and F. De Bruyne (1996). For model-based control design, closed-loop identification gives better performance. *Automatica* **32**, 1659–1673.

Hjalmarsson, H., M. Gevers, S. Gunnarsson and O. Lequin (1998). Iterative Feedback Tuning: theory and applications. *IEEE Control Systems Magazine* **18**, 26–41.

Hjalmarsson, H., S. Gunnarsson and M. Gevers (1994). A convergent iterative restricted complexity control design scheme. In: *Proc. 33rd IEEE Conf. on Decision and Control*. Orlando, Florida. pp. 1735–1740.

Hjalmarsson, H., S. Gunnarsson and M. Gevers (1995*a*). Model-free tuning of a robust regulator for a flexible transmission system. *European Journal of Control* **1**(2), 148–156.

Hjalmarsson, H., S. Gunnarsson and M. Gevers (1995*b*). Optimality and sub-optimality of iterative identification and control design schemes. In: *Proc. American Control Conference 95*. Vol. 4. pp. 2559–2563.

Landau, I.D. (1999*a*). From robust control to adaptive control. *Control Engineering Practice* **7**(9), 1113–1124.

Landau, I.D. (1999*b*). Identification in closed-loop: a powerful design tool (better design models, simple controllers). In: *Proceedings of the workshop "Future directions in Systems and Control Theory"(Org: M. Athans and J. Senteiro)*. Lisboa/Cascais, Portugal.

Lee, W.S., B.D.O. Anderson, I.M.Y Mareels and R.L. Kosut (1995). On some key issues in the windsurfer approach to adaptive robust control. *Automatica* **31**(11), 1619–1636.

Lequin, O. (1997). Optimal closed loop PID tuning in the process industry with the Iterative Feedback Tuning scheme. In: *CD-ROM of European Control Conference, paper TH-A-H6*. Brussels, Belgium.

Lequin, O., M. Gevers and L. Triest (1999). Optimizing the settling time with Iterative Feedback Tuning. In: *Proc. 14th IFAC World Congress, paper I-3b-08-3*. Beijing, P.R. China. pp. 433–437.

Liu, K. and R.E. Skelton (1990). Closed loop identification and iterative controller design. *29th IEEE Conf on Decision and Control* pp. 482–487.

Molenaar, D. (1995). Model-free data-driven optimal tuning of controller parameters: A practical guide with application examples. Technical report. Report LiTH-ISY-R-1722, Department of Electrical Engineering. Linköping University, Sweden.

Ninness, B. and G.C. Goodwin (1995). Estimation of model quality. *Automatica* **31**(12), 32–74.

Partanen, A.G. and R.R. Bitmead (1995). The application of an iterative identification and controller design to a sugar cane crushing mill. *Automatica* **31**, 1547–1563.

Robbins, H. and S. Monro (1951). A stochastic approximation method. *Ann. Math. Stat.* **22**, 400–407.

Schrama, R.J.P. (1992*a*). Accurate identification for control: The necessity of an iterative scheme. *IEEE Trans. on Automatic Control* **37**, 991–994.

Schrama, R.J.P. (1992*b*). Approximate identification and control design. *PhD thesis*.

Schrama, R.J.P. and O.H. Bosgra (1993). Adaptive performance enhancement by itreative identification adn control design. *Int. Journal of Adaptive Control and Signal Processing* **7**(5), 475–487.

Van den Hof, P.M.J. and R.J.P. Schrama (1993). An indirect method for transfer function estimation from closed loop data. *Automatica, 29* **6**, 1523–1527.

Van den Hof, P.M.J. and R.J.P. Schrama (1995). Identification and control - closed-loop issues. *Automatica* **31**, 1751–1770.

Van den Hof, P.M.J., R.J.P. Schrama, R.A. de Callafon and O.H. Bosgra (1995). Identification of normalized coprime plant factors from closed-loop experimental data. *European Journal of Control* **1**, 62–74.

Zang, Z., R.R. Bitmead and M. Gevers (1995). Iterative weighted least-squares identification and weighted LQG control design. *Automatica* **31**(11), 1577–1594.

Zhou, K., J.C. Doyle and K. Glover (1995). *Robust and Optimal Control*. Prentice Hall, New Jersey.