Practical Issues in the Implementation of Self-Tuning Control

Wittenmark, Björn; Åström, Karl Johan

1984

Document Version:
Publisher's PDF, also known as Version of record

CODEN:LUTFD2/(TFRT-7268)/1-030/(1984)

PRACTICAL ISSUES IN THE IMPLEMENTATION OF SELF-TUNING CONTROL

BJÖRN WITTENMARK
KARL JOHAN ÅSTRÖM

DEPARTMENT OF AUTOMATIC CONTROL
LUND INSTITUTE OF TECHNOLOGY
MARCH 1984

LUND INSTITUTE OF TECHNOLOGY
DEPARTMENT OF AUTOMATIC CONTROL
Box 725
S 220 07 Lund 7 Sweden

Document name
Report

Date of issue
March 1984

Document number
CODEN:LUTFD2/(TFRT-7268)/1-030/(1984)

Supervisor

Sponsoring organization

Author(s)

Björn Wittenmark and Karl Johan Åström

Title and subtitle

PRACTICAL ISSUES IN THE IMPLEMENTATION OF SELF-TUNING CONTROL.

Abstract

Implementation aspects of self-tuning regulators are discussed in the paper. There is a large discrepancy between simulation or academic algorithms and practical algorithms. In the idealized environment of simulation it is easy to get different types of adaptive algorithms to perform well. In practice the situation is quite opposite. The adaptive or self-tuning controller must be able to handle nonlinearities, unmodeled dynamics, and unmodeled disturbances over a wide range of operating conditions. Some aspects of how to implement self-tuning controllers are discussed in the paper. This includes robustness, signal conditioning, parameter tracking, estimator windup, reset action and start-up. Different ways to use the prior knowledge about the process are also discussed.

Key words

Adaptive control, Control applications, Process control, Robustness, Self-tuning regulators.

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title

ISBN

Language
English

Number of pages
30 pages

Recipient's notes

Security classification

Distribution: The report may be ordered from the Department of Automatic Control or borrowed through the University Library 2, Box 1010, S-221 03 Lund, Sweden, Telex: 33248 lubbis lund.

PRACTICAL ISSUES IN THE IMPLEMENTATION OF SELF-TUNING CONTROL

Björn Wittenmark and Karl Johan Åström

Department of Automatic Control
Lund Institute of Technology
Box 725, S-220 07 Lund, Sweden

Abstract: Implementation aspects of self-tuning regulators are discussed in the paper. There is a large discrepancy between simulation or academic algorithms and practical algorithms. In the idealized environment of simulations it is easy to get different types of adaptive algorithms to perform well. In practice the situation is quite opposite. The adaptive or self-tuning controller must be able to handle nonlinearities, unmodeled dynamics, and unmodeled disturbances over a wide range of operating conditions. Some aspects of how to implement self-tuning controllers are discussed in the paper. This includes robustness, signal conditioning, parameter tracking, estimator windup, reset action, and start-up. Different ways to use the prior knowledge about the process are also discussed.

Key words: Adaptive control, Control applications, Process control, Robustness, Self-tuning regulators.

1. INTRODUCTION

A non-specialist who tries to get an understanding of adaptive control is confronted with contradicting information like:

- Adaptive control works like a beauty in this particular practical application.
- Adaptive control is ridiculous. Look how the algorithm behaves in this simulation.
- You can not possibly use adaptive control because there is no proper theory that guarantee stability and convergence of the algorithm.

Use of adaptive control is very tempting in many situations, especially if little is known about the process to be controlled. The sad fact of life is, however, that there are seldom any shortcuts that really pay off. Nothing, not even an adaptive controller, can replace good engineering and physical insight. A newcomer to the field of adaptive control must find the situation somewhat confusing. There is a large number of proposed algorithms. Some are quite different while other only differ in minor details. Are the differences important or do they just reflect personal preferences of the designer?

Many adaptive schemes can be represented with the block diagram in Fig 1. The adaptive controller may be regarded as composed of two loops. An inner loop, which is the basic control loop for the system, and an outer loop which adjusts the parameters of the regulator in the inner loop by parameter estimation and control design. When considering an adaptive control scheme one must be sure that both loops work satisfactory, both individually and together. The design method used must work well if the process is totally known. It should also be robust and insensitive to the underlying assumptions, for instance to the distribution of the noise, the character of the reference signal, and to unmodeled high frequency process dynamics. This means that all physical knowledge about the process should be used to check if the design method is appropriate. Further the a priori knowledge should be used to check the design specifications. For instance that the desired bandwidth is not too high compared to allowed control signals. The choices of presampling filters and sampling interval are also influenced by the knowledge about the process.

The estimation routine in the outer loop must be such that it can give good estimates in the intended application. Issues to be considered are for instance identifiability, the richness of the input signal, identification in closed loop, and the possibility to follow changes in process parameters.

In summary the two loops must be based on good engineering practice. Each loop must work individually. The interaction between the loops is also important. This is, however, a very difficult problem to investigate. Interactions can normally be reduced by ensuring that the outer loop is much slower than the inner loop. This will unfortunately limit the adaptation rate.

Great progress has been made in the theory of adaptive control during the last years. Many theoretical problems such as stability and convergence have been solved for idealized cases. Shortcomings of adaptive algorithms have, however, also been pointed out. The limitations are often due to violation of the assumptions for the algorithms. These limitations must be circumvented in order to obtain robust algorithms that can be used in practice. An overview of theory and applications of adaptive control is given in Åström (1983a).

Adaptive control laws may be used in many different ways. One possibility is to make a control algorithm with automatic tuning. One example of such a controller is described in Åström and Hägglund (1984). The tuning may then be switched off when a satisfactory performance is obtained. The tuning can in this case be supervised by an operator and it is not necessary to have too much logic built into the algorithm. Another way to use an adaptive controller is to have the tuning switched on all the time in order to follow possible changes in the dynamic of the process. This is a much more difficult situation which requires a more robust algorithm.

Many feasibility studies and commercial installations indicate that adaptive control can be used successfully to control industrial processes, see for instance Åström (1983a). However, in all applications different types of 'safety-nets' and special tricks have to be used. The purpose of this paper, which is based on Wittenmark and Åström (1982), is to give some reflections on the state of art and experiences from practical work on adaptive control.

The paper is organized in the following way. Section 2 discusses theory and practice. For instance what can be said theoretically and what are the practical limitations. Robustness issues are treated in Section 3. The idea of self-tuning algorithms is briefly reviewed in Section 4. Implementation aspects such as estimator windup, signal conditioning, and reset action is covered in Section 5. Section 6 gives some conclusions and references are given in Section 8.

2. THEORY AND PRACTICE

It is important to consider both theoretical and practical aspects when discussing adaptive control algorithms. The theory deals with idealized situations where all the conditions are under control. The theory thus gives the ultimate limit of what can be achieved and expected under idealized conditions. The practical situation is, however, such that there are all kinds of violations of the conditions of the theory.

Some important theoretical problems related to self-tuning control have been solved during the last years. Stability and convergence proofs for simple algorithms under idealized conditions are available. See for instance Egardt (1979, 1980), Goodwin et al (1980), Morse (1980) and Narendra et al (1980). There are, however, no results available for more realistic assumptions. One main criticism against adaptive control is the fact that the adaptive controllers, like most other controllers, are designed on models that are simpler than the real processes. The effect of unmodeled high frequency dynamics on some adaptive schemes is discussed for instance in Rohrs et al (1981, 1982) and Gawthrop and Lim (1982). Other problems are due to process nonlinearities and actuator saturation. To avoid these kind of problems it is necessary to provide the practical algorithms with a safety-net or a supervisory level which can take care of unusual and undesired situations in a safe way. Much of the practical work is thus not done on a firm theoretical basis but consists of ad hoc solutions that often will depend on the considered application. It is often verified by extensive experimentation and simulation.

Some of the main issues in adaptive control are

- how to use prior information about the process
- how to determine realistic specifications of the closed loop system
- how to make robust estimation
- unmodeled high frequency dynamics
- signal conditioning.
- numerical problems
- startup and bumpless transfer
- process and actuator nonlinearities

Several of these points are not specific for adaptive control but are valid also for design of digital controllers in general.

3. ROBUSTNESS

Before going into the details of the implementation aspects of adaptive control it is relevant to discuss robustness issues in general. Many adaptive controllers are based on the certainty equivalence hypothesis. The regulators can then be regarded as a combination of a design procedure for known systems and a recursive estimation scheme. See Fig 1.

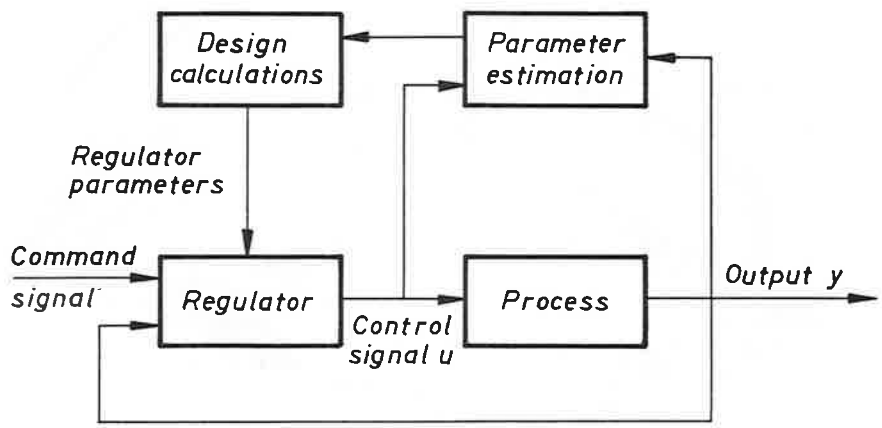

Fig. 1 Block diagram of a self-tuning regulator.

The true parameters of the process are replaced by the estimated parameters. All control design must be based on sound principles. For an adaptive controller it is necessary to have a robust design method and a robust estimation scheme in order to get a robust adaptive controller.

Robust Control Design

Robustness properties of the design of a controller can be discussed in terms of the loop gain of the system in cascade with the controller. A Bodeplot of a typical loop gain is shown in Fig 2.

It is common practice to make the design such that the loop gain is high below the cross-over frequency, ω_c. Further the loop gain should fall off rapidly above the cross-over frequency, see Horowitz (1963) and Doyle and Stein (1981). The high loop gain at low frequencies is obtained for instance by introducing integral action. The high loop gain at low frequencies will make it possible to eliminate low frequency disturbances. It will also make the system insensitive to the low frequency characteristics of the process model. The high frequency roll-off is necessary to eliminate the influence of high frequency disturbances or unmodeled high frequency dynamics. The high frequency roll-off is obtained by filtering the signals. For sampled data systems the sampling procedure will restrict the high frequency content in the sampled signal. To avoid aliasing it is necessary to provide the controller with an effective anti-aliasing filter. The choice of the sampling interval will then naturally be related to the closed loop behavior of the system. Aliasing and choice of sampling period are discussed for instance in

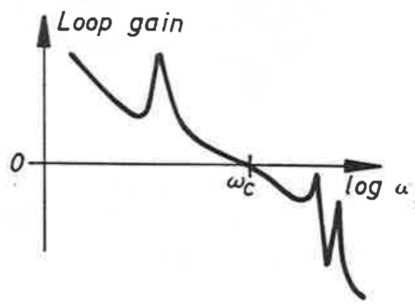

Fig. 2 Bodeplot of a typical loop gain.

Åström and Wittenmark (1984).

The discussion of the loop gain results in the conclusion that it is necessary to have a good process model for the frequencies around the cross-over frequency. It is possible to make a more quantitative statement for design procedures based on model following. Consider a system with the transfer function G_0. Assume that a design is based on the process model G. Select the feedforward and the feedback transfer functions such that the closed loop system from u_c to y is G_m. See Fig. 3.

The closed loop system is stable if

$$|G_0 - G| < \left|\frac{G}{G_m}\right| \left|\frac{T}{S}\right| \qquad (3.1)$$

on the imaginary axis and at infinity (or on the unit circle for a discrete time system). The statement is proven for discrete time systems in Åström (1980b). It also follows from the result of Doyle and Stein (1981). Other theorems of similar nature are given in Mannerfelt (1981) and Åström and Wittenmark (1984).

The left hand side of (3.1) is the error in the transfer function of the model. The advantage of (3.1) is that the right hand side depends only on known quantities on the used model, the desired model, and the resulting controller. Using the inequality it is possible to investigate how the desired closed loop characteristics will influence the accuracy that is needed for the model. From (3.1) it also follows that it is necessary to have a good model for frequencies

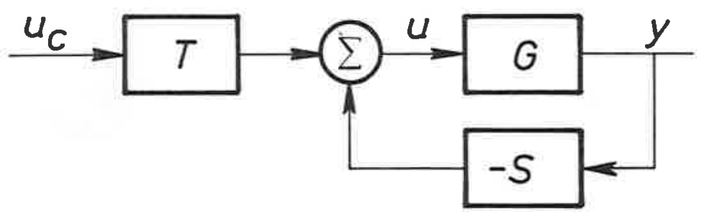

Fig. 3 Block diagram of process and regulator.

around the cross-over frequency. It also follows that requirements on model precision increases with increasing closed loop bandwidth.

The conclusion of this part is that a reliable design method should be used and that the design must be based on a model that is accurate at least for frequencies where the loop gain is around unity.

Robust Estimation

The accuracy requirements for the design procedure lead to the question of how to obtain a good model. The key issues are good data and an appropriate model structure. The models used will invariably be simplified e.g. linear and of low order. It is known from the theory of system identification that the estimates obtained in such a case will depend crucially on the properties of the input signal.

Above it has been demonstrated that robust control requires a model which is accurate around the cross-over frequency. To estimate an accurate reduced order model with this property it is essential that the input signal has sufficient energy content around the cross-over frequency and that it is so rich in frequency that it is persistently exciting. See Åström and Bohlin (1966). The conditions on persistent excitation are related to the complexity of the estimated model. This implies that the requirements on the input signal become more severe if the model order is increased. Since the input signal in an adaptive system is generated by feedback there is no guarantee that it will be persistently exciting. To guarantee a good model it is thus necessary to monitor the excitation and the energy of the input signal in the relevant frequency bands.

When the input signal generated by the feedback lopp is not persistently exciting or when the signal energy is too low the estimated parameters will be poor. There are two ways to avoid this, by injecting extra perturbation signals or by switching off the adaptation when excitation is poor. The results by Egardt (1979) and Peterson and Narendra (1982) indicate that it is reasonable to estimate only when the absolute value of the useful input energy is above a certain level. Ways to give the signals a proper frequency content is to filter the signals or to introduce external perturbation signals.

There are other safe-guards of similar nature to make sure that the estimation only is done when the data is reasonable.

The difficulties with adaptive control reported by Rohrs and others (1981, 1982) are due to high frequency reference signals and measurement noise in combination with unmodeled dynamics. The reported problems are analyzed in Åström (1983b) and the difficulties can be eliminated if the precautions above are taken.

<u>Example 3.1 - Rohrs' example</u>

A counterexample to adaptive control is presented in Rohrs et al (1982). The algorithm is a continuous time model reference controller. The process is of third order. The dominating dynamics is a pole at -1. The other modes of the system represent dynamics with a natural frequency that is high compared to the dominating part and to the desired response of the system. The adaptive controller estimates a first order model, i.e. two parameters. The problems with the controller show up when the reference signal is constant and a low amplitude high frequency disturbance is added.

If the only excitation is a high frequency sinusoid the adaptive system will attempt to match the model to the system at this frequency. This will result in a low order model that will represent the dominating dynamics very poorly with expected disastrous results. (See Fig 4).

The remedy is either to switch off the adaptation when there is no excitation in the appropriate frequency band or to inject a relevant perturbation signal. One way to improve the algorithm is thus to filter the output of the process and to introduce an excitation signal of low frequency. Signal energy at low frequencies will make it possible to better estimate the dominating dynamics of the process. Fig 4 also shows the parameter estimates when the output is filtered with a low pass filter and when the excitation signal has the same amplitude as the high frequency disturbance. These precautions will prevent the parameter estimates from diverging. The estimates will now stay bounded and the closed loop system will be stable.

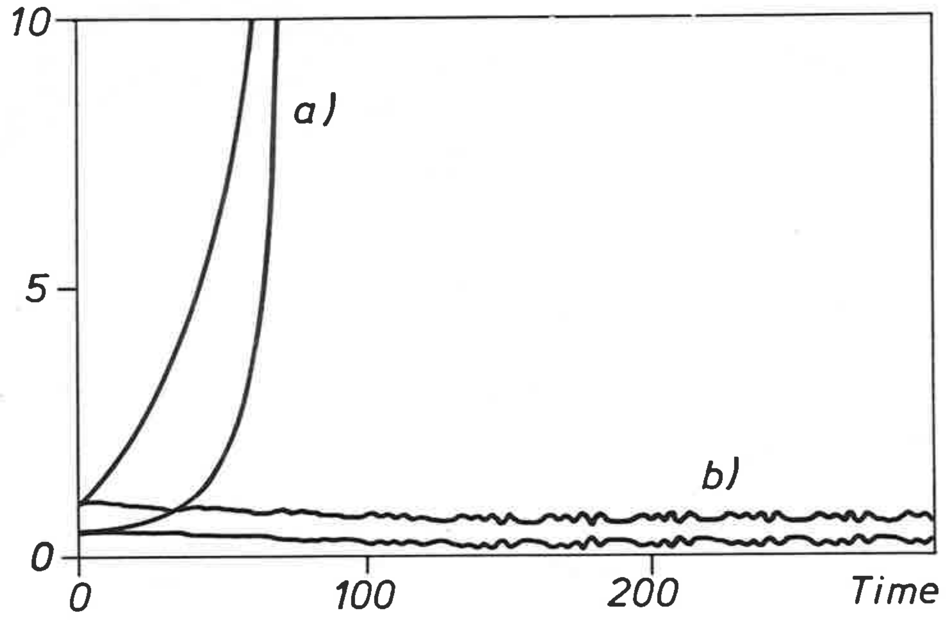

Fig. 4 Parameter estimates for the process in Example 3.1.
a) Rohrs' algorithm
b) The modified algorithm.

□

Robust Adaptive Control

To obtain a robust adaptive control algorithm it is necessary to use both robust control and robust estimation. In the adaptive context there is also some new trade-offs to be made. Consider for instance the robustness properties obtained by having a high open loop gain at low frequencies. This may be obtained by having integral action in the control loop. It can also be obtained via adaptation. An adaptive controller with enough parameters will automatically introduce a high gain at those frequencies where there are low frequency disturbances. This will be discussed further in Section 5.

4. SELF-TUNING ALGORITHMS

Before discussing implementation aspects of self-tuning controllers it is necessary to be little more specific about the algorithms. A summary of self-tuning algorithms is given for instance in Clarke (1982) and in Åström (1983a). Assume that the process to be controlled can be described by the discrete time system

$$A^*(q^{-1})y(k) = B^*(q^{-1})u(k-d_o) + C^*(q^{-1})e(k) \qquad (4.1)$$

where y(k) is the output, u(k) is the input and e(k) is a white Gaussian noise disturbance. The time scale is normalized such that one sampling interval is one time unit. A^*, B^*, and C^* are polynomials in the delay operator q^{-1}. The polynomials are defined as

$$A^*(q^{-1}) = 1 + a_1q^{-1} + \ldots + a_nq^{-n}$$

etc. A general linear controller can be written on the form

$$R^*(q^{-1})u(k) = -S^*(q^{-1})y(k) + T^*(q^{-1})u_c(k) \qquad (4.2)$$

where u_c is the command signal or reference value.

Self-tuning controllers can be divided into two categories, explicit and implicit algorithms, see Åström and Wittenmark (1980).

Explicit or Indirect Self-Tuning Algorithms

When using an explicit algorithm an explicit process model is estimated, i.e. the coefficients of the polynomials A^*, B^*, and C^* in (4.1). An explicit algorithm can then be described by two steps. The first step is to estimate the polynomials A^*, B^*, and C^* of the process model (4.1). In the second step a design method is used to determine the polynomials in the regulator (4.2) using the estimated parameters from the first step. The two steps are repeated at each sampling interval. The design procedure in the second step can be any good design method that is suitable for the problem on hand. A typical application of an explicit algorithm is for instance a pole placement algorithm described in Åström and Wittenmark (1980).

Implicit or Direct Self-Tuning Algorithms

In an implicit algorithm the parameters of the regulator are estimated directly. This can be made possible by a reparameterization of the process model. A typical example of an implicit algorithm is the minimum variance self-tuner in Åström and Wittenmark (1973).

One advantage with the implicit algorithms over the explicit algorithms is that the design computations are eliminated, since the controller parameters are estimated directly. The implicit algorithms usually have more parameters to estimate than the explicit algorithms, especially if there are long time delays in the process. Simulations and practical experiments indicate, however, that the implicit algorithms are more robust. The implicit algorithms usually have the disadvantage that all process zeros are cancelled. This implies that the implicit methods are intended only for processes with a stable inverse or minimum phase systems. Sampling of a continuous time system often gives a discrete time system with zeros on the negative real axis, inside or outside the unit circle. It is not good practice to cancel these zeros even if they are inside the unit cirlce. Cancellation of these zeros will give rise to 'ringing' in the control signal. Many implicit algorithms can, however, be used also if the system is nonminimum phase through a proper choice of parameters. An example is given in Example 5.5 below.

Feedforward

Feedforward control is a very effective way to reduce the influence of measurable disturbances. It requires, however, knowledge about the process dynamics. In adaptive controllers it is easy to include feedforward from different signals, see Åström and Wittenmark (1973). The estimation algorithm will automatically give the required dynamics. To be effective the models must also be reasonably accurate. Adaptation is thus almost a prerequisite for practical use of feedforward. Adaptive feedforward has been used in several of the applications of adaptive controllers discussed in Åström (1983b).

Real time estimation

Both explicit and implicit algorithms need a recursive estimation scheme. The most common is the method of least squares or its modifications. The least squares estimator is described by the equations

$$\theta(k+1) = \theta(k) + P(k+1)\varphi(k+1)\varepsilon(k+1) \tag{4.3}$$

$$P(k+1) = [P(k) - P(k)\varphi(k)R(k)\varphi^T(k)P(k)]/\lambda \tag{4.4}$$

$$R(k) = [\lambda + \varphi^T(k)P(k)\varphi(k)]^{-1} \tag{4.5}$$

where θ is a vector consisting of the parameters to be estimated, φ is a vector of delayed inputs and outputs (possibly filtered), and ε is the prediction error. P can be interpreted as the covariance matrix of the estimation error. Finally λ is an exponential forgetting factor. The forgetting factor is used to make it possible to follow time varying parameters. Further details of recursive estimation schemes are found in Ljung and Söderström (1983).

5. IMPLEMENTATION ASPECTS

The empirical knowledge that the authors have in the field of adaptive control is mainly from using self-tuning regulators of different types, see Åström et al (1977) and Källström et al (1979). The estimation has mostly been done using the method of least squares. Different design methods have been used for instance minimum variance control, linear quadratic Gaussian control, and pole placement design based on complex and simplified models. Some of the experiences are discussed in the following.

Signal Conditioning

In all digital control applications it is important to have a proper conditioning of the signals. Due to the aliasing problem connected with the sampling procedure it is necessary to eliminate all frequencies above the Nyquist frequency before sampling the signals. High frequencies may otherwise be interpreted as low frequencies and may introduce disturbances in the closed loop system. The filtering of the signals also has the effect that the system will be excited by frequencies where it is important to have good process models. Compare the example in Section 3.

Example 5.1 - The effect of prefilter

A laboratory plant for concentration control has been used in Åström and Zhao-Ying (1982) for experiments with a linear quadratic Gaussian self-tuner. Fig 5 shows the effect of prefiltering of the measurement. In Fig 5a a typical sample

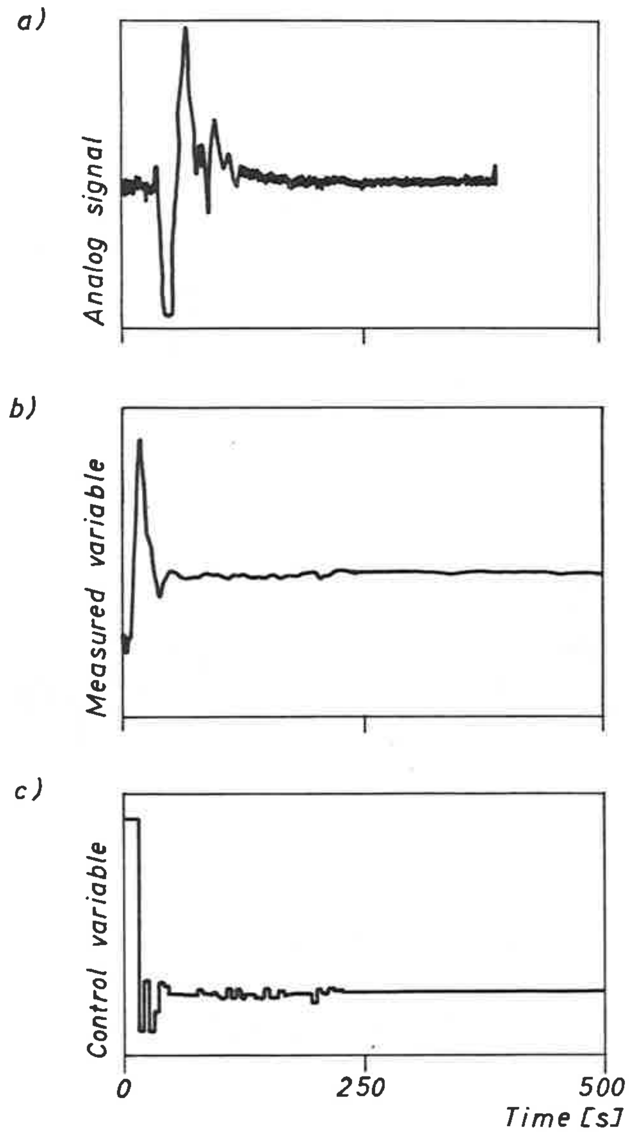

Fig. 5 The effect of a prefilter when controlling concentration in a laboratory process using a LQG self-tuner. a) Analog process output b) Sampled process output c) Control signal (Adapted from Åström and Zhao-Ying (1982)) Notice that curves a) and b) are not from the same experiment.

of the time continuous process output is shown. In Fig 5b and c the sampled process output and the control signals are shown. (Curves a) and b) are not from the same experiment). The sampling period is 15s. At t = 600 an analog antialiasing filter is connected to the process output. The filter is a first order filter with a time constant of 75s. The high frequency disturbance in the concentration measurement introduces a low frequency disturbance in the sampled process output and consequently also in the control signal. This disturbance is effectively eliminated by using a prefilter.

□

Parameter Tracking and Estimator Windup

The key property of an adaptive controller is the ability to track variations in the process dynamics. To do so it is necessary to discount old data. This will involve compromises. Too fast discounting will make the estimates uncertain if the parameters are constant. Too slow discounting will make it impossible to track rapid parameter variations.

Exponential forgetting is one way to discard old data. The algorithm described by (4.3) - (4.5) minimizes the loss function

$$\sum_{k=0}^{N} \lambda^{N-k} \varepsilon(k)^2$$

where ε is the prediction error. With $\lambda = 1$ all data have the same weight. With $\lambda < 1$ more recent data are weighted more than old data. It is possible to generalize the method with exponential forgetting and have different forgetting factors for different parameters. Exponential forgetting works well only if the process is properly excited all the time. There are several problems with exponential forgetting when the excitation of the process changes. A typical situation is when the main source of excitation is changes in the set point. There may then be long periods with no excitation. The estimator will then forget the proper values of the parameters and the uncertainties will grow. This may be called <u>estimator windup</u> (compare with integrator windup in conventional integral control). The problem can be understood from (4.4). The negative term on the right hand side represents the reduction in uncertainty due to the last measurement. If there is no information in the last measurement then $P(k)\varphi(k)$ will be zero and (4.4) reduces to

$$P(k+1) = P(k)/\lambda$$

$P(k)$ will thus grow exponentially until φ changes direction if $\lambda < 1$. If there is no excitation for a long period of time then $P(k)$ may be very large. Since $P(k)$ also is the gain in (4.3) then there may be large changes in the estimated parameters when new information is coming into the system, for instance when the reference value changes. The estimator windup may then cause a <u>burst</u> in the output of the process.

There are several ways to avoid estimator windup. The main idea is to ensure that P stays bounded. This can be done for instance by ensuring that the trace of P is constant in each iteration, see Irving (1979). Another possibility is to adjust the forgetting factor automatically. Ways to do this is given in Forstescue et al (1978) and Wellstead and Sanoff (1981). The automatic adjustment of λ in these references does, however, not guarantee that P stays bounded.

It has also been proposed to stop the updating of the parameters and the covariance matrix when $P\varphi$ or ε are sufficiently small, see Egardt (1979). Hägglund (1984) has proposed superior algorithms which only discount data in the directions where there are new information.

<u>Example 5.2 - Estimator windup</u>

The problem with estimator windup is illustrated by a simple simulated example. Let the process to be controlled be described by

$$y(k) - 0.9y(k-1) = u(k-1)$$

It is desired that the pulse transfer operator from the reference signal to the output has a pole in 0.7 and that the gain is unity. This is achieved with the controller

$$u(k) = 0.3y(k) - 0.35y(k-1) + u_c(k) - 0.5u_c(k-1)$$

The process is controlled using an implicit pole placement algorithm where the parameters in the controller are estimated. Fig 6 shows the diagonal elements of the P-matrix (4.4) when different estimation schemes are used. The reference signal is a square wave with unit magnitude and period 100 up to time 300. After that the reference signal is constant.

Parameter estimates when controlling the process in Example 5.5 with the basic self-tuning algorithm with $d = 2$ and when there are two parameters in the controller. The dashed lines corresponds to the parameters of the suboptimal minimum variance controller.

In Fig 6a the estimation algorithm described by (4.3)-(4.5) is used with $\lambda = 0.99$. When the reference signal is constant and the output has settled there is no information in the measurements. The variance will then start to increase. In Fig 6b a new estimation routine described in Hägglund (1984) is used. The variances of the estimates now settle on constant

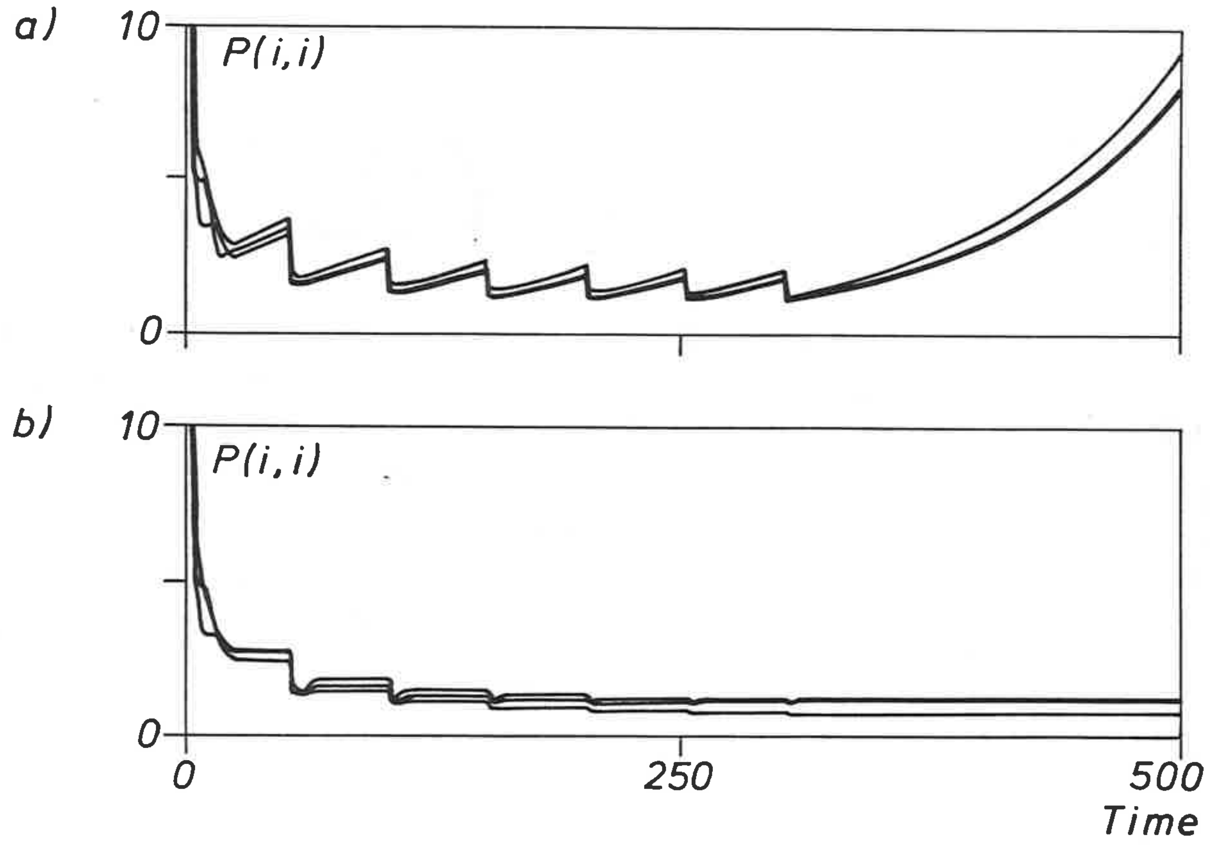

Fig. 6 The diagonal elements of the P-matrix when controlling the process in Example 5.2.
a) Constant exponential forgetting factor $\lambda = 0.99$.
b) Forgetting according to Hägglund (1984).

values and there is no estimator windup.

□

There may be numerical problems in the updating of the parameters and the covariance matrix. This especially true if the excitation of the process is poor. The square root method or the U-D factorization method, see Bierman (1977), are then good ways to organize the computations.

<u>Start-Up Procedures</u>

There are several ways in which a self-tuning algorithm can be initialized depending on the a priori information about the process.

One case is if the process has been controlled before with a conventional or an adaptive controller. The initial values should then be such that they correspond to the controller used before.

Another situation occur if nothing is known about the process. The initial values of the parameters in the estimator can then be chosen to be zero or such that the initial controller is a proportional or integral controller with low gain. The inputs and outputs of the process should be scaled such that they are of the same magnitude. This will improve the numerical conditions in the estimation and the control parts of the algorithm. The initial value of the covariance matrix can be 1 - 100 times a unit matrix. These values are usually not crucial since the estimator will get reasonable values in a very short period of time. Our experience is that 10 - 50 samples is sufficient to get a very good controller. During the initial phase it can be advantageous to add a perturbation signal to speed up the convergence of the estimator. Auto-tuning discussed in Åström and Hägglund (1984) is a convenient way to initialize the algorithm because it generates a suitable input signal and safe initial values of the parameters.

Sometimes it is important to have as small disturbances as possible due to the start-up of the self-tuning algorithm. There are then two precautions that can be taken. First the estimator can be used for some sampling periods before the self-tuning algorithm is allowed to put out any control actions. During that time a safe simple controller should be used. It is also possible and desirable to limit the control signal. The allowable magnitude can be very small during the first period of time and can then be increased when better parameter estimates are obtained. This kind of soft start-up is for instance used in Asea's Novatune, see Bengtsson (1979). The drawback of having small input signals is that the excitation of the process will be poor and it will take longer time to get good parameter estimates.

Reset Action

It is important that a controller has the ability to eliminate steady state errors when the reference value is constant. The steady state error can be generated by many different mechanisms: calibration errors, nonlinearities, disturbances etc. In a conventional controller, reset action is obtained by introducing an integrator in the controller. When using self-tuning controllers there are several ways to introduce reset action. Since there is no method that is uniformly best different alternatives will be discussed. The introduction of integrators is discussed in Åström (1980a) and also in Allidina and Hughes (1982).

The simplest way to get reset action is to let the self-tuning regulator take care of the problem. Since it estimates a model of the process and the environment it can be expected that the self-tuner tries to model the offset and compensate for it. It is easy to check if a particular self-tuner has this ability by investigating possible stationary solutions. A typical example is given below.

Example 5.3 - Automatic reset action

Consider the simple implicit self-tuning controller in Åström and Wittenmark (1973) which is based on least squares estimation and minimum variance control. The estimation is based on the model

$$y(k+d) = R^*(q^{-1})\ u(k) + S^*(q^{-1})\ y(k) \tag{5.1}$$

where d is an estimate of d_o in (4.1) and the regulator is

$$u(k) = -\frac{S^*}{R^*}\ y(k) \tag{5.2}$$

The conditions for a stationary solution are that

$$\hat{r}_y(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} y(k+\tau)y(k) = 0 \quad \tau = d,\ldots,d+\deg S^* \quad (5.3)$$

$$\hat{r}_{yu}(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} y(k+\tau)u(k) = 0 \quad \tau = d,\ldots,d+\deg R^* \quad (5.4)$$

These conditions are not satisfied unless the mean value of y is zero. When there is an offset the parameter estimates will get values such that $R^*(1) = 0$, i.e. there is an integrator in the controller. The convergence to the integrator may, however, be slow. It can be shown that other both explicit and implicit self-tuning regulators also can give reset automatically.

□

A second way to introduce reset action is to estimate the bias in the process. A simple way to do this is to include a bias term, b, in the model (4.1) The model is then

$$A^*y(k) = B^*u(k-d_0) + C^*e(k) + b$$

With this model it is easy to estimate b and to compensate for it. Such a scheme was proposed in Clarke and Gawthrop (1979). One drawback is, that an extra parameter has to be estimated. Further it is necessary to have different forgetting factors on the bias estimate and the other estimates. Otherwise the convergence to a new level will be very slow. Finally if the bias is estimated in this way it is not possible to use the self-tuner as a tuner since there will be no reset when the estimation is switched off.

A third way to get reset action is to force an integrator into the regulator. That means that the controller has the form

$$R^*\nabla u(k) = -S^*y(k) + T^*u_c(k)$$

where

$$\nabla u(k) = (1 - q^{-1})u(k) = u(k) - u(k-1)$$

This form has several advantages. First, there will always be an integrator even if the regulator parameters are not optimally tuned. Second, the high gain at low frequencies will increase the robustness of the system due to uncertainties in the process dynamics at low frequencies. This implies that the estimation can be concentrated at frequencies around the cross-over frequency. One drawback with this method is that the self-tuner will try to eliminate the integral action when it is not needed. This implies that the regulator will try to cancel a pole at the stability boundary.

Actuator Nonlinearities

A controller, adaptive or not, usually contains several nonlinearities. For instance the magnitude of the control signal is limited. When using a self-tuning regulator it is especially important that the estimator is fed with the control signal that is sent out to the process. The estimator will otherwise get incorrect estimates for instance of the gain of the process.

Antireset windup

An integrator is an unstable system and it may happen that the integral can assume very large values if the control signal saturates when there is an error. This is called reset windup or integrator windup. Special precautions must be taken in order to avoid this. Ways to do this are discussed in Åström and Wittenmark (1984) for different controller structures.

Example 5.4 - Antireset windup controller

Consider a regulator described by (4.2) where the regulator may contain unstable modes. One way to solve the reset windup problem is to rewrite (4.2) by adding $A_o^*(q^{-1})u(k)$ on both sides. This gives

$$A_o^* u(k) = T^* u_c - S^* y + (A_o^* - R^*)u$$

A regulator with antireset windup compensation is then given by

$$\begin{cases} A_o^* v(k) = T^* u_c - S^* y + (A_o^* - R^*)u \\ u(k) = \mathrm{sat}[v(k)] \end{cases} \tag{5.5}$$

where sat[·] is the saturation function. This regulator is equivalent to (4.2) when the control signal does not saturate. The polynomial A_o^* should be stable. It can be interpreted as the dynamics of the observer associated with the controller. A block diagram of (5.5) is shown in

Fig. 7. A particular simple case is when $A_o^*=1$, which corresponds to a dead beat observer. The controller is then

$$u(k) = \text{sat}\left[T^* u_c(k) - S^* y(k) + (1 - R^*) u(k)\right]$$

□

Tuning parameters

When adaptive control is mentioned the vision of the ideal black box easily appears, i.e. a system without any tuning knobs that would give a good closed loop performance no matter what it is connected to. Today we are far away from such a solution and it can also be questioned if such a solution is desired. It is at least necessary to tell the controller what we expect it to do. Usually it is also necessary to provide much more a priori information. Only for specialized applications it may be possible and desirable to have a regulator without any parameters to tune. The important issue is instead what kind of parameters that are tuned. One possibility is to introduce performance related knobs, i.e. knobs which relate to the performance of the closed loop system. Examples of performance related knobs are the bandwidth and the damping of the closed loop system or the gain and phase margins. Such parameters are easy to tune since they are directly related to the behaviour of the closed loop system.

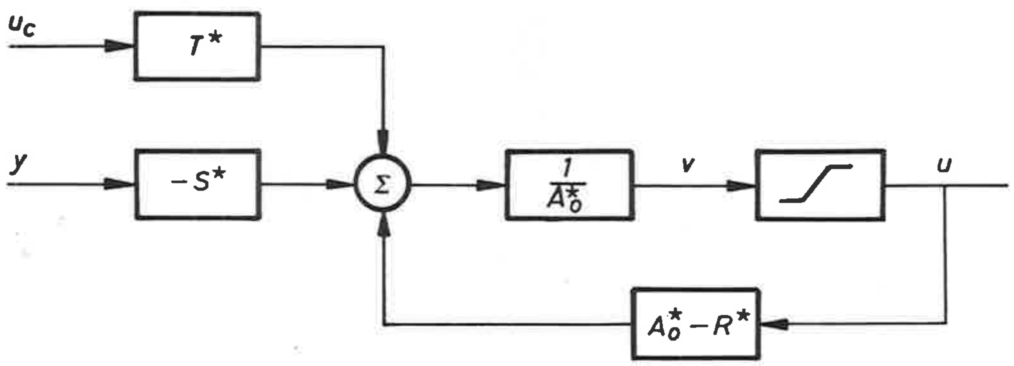

Fig. 7 Block diagram of (5.5) which avoids windup.

There are two categories of parameters to be chosen in adaptive controllers, integers and reals. For self-tuning regulators the integer parameters are typically the order of the process model and possibly the time-delay in the process to be controlled. The real parameters are the performance related parameters and initial values for the estimation routine. The integer parameters are usually quite easy to determine. The performance can also be made insensitive to the initial values of the estimator. This was further discussed above in connection with start-up procedures for adaptive algorithms. A discussion of the choice of the parameters in self-tuning regulators is found for instance in Wellstead and Zanker (1982).

One parameter that can influence the behavior of the algorithm very much is the sampling period. The choice of the sampling interval is not specific for adaptive controllers but is an important design parameter for all sampled data design methods. In general it is important that the sampling period is related to the desired performance of the closed loop system. There are several rules of thumb that can be given. One is to relate the sampling period, h, to the desired rise time, T_r, and define

$$N_r = T_r/h$$

To get a good servo performance of the closed loop system N_r should be chosen in the range 2 - 4. A similar rule of thumb is to relate the sampling period to the natural frequency, ω, of the control poles of the desired closed loop system. The sampling period can then be chosen such that $\omega \cdot h$ is in the range 0.25 - 1. The choice of the sampling period is also influenced by the character of the disturbances acting on the system. More details about the choice of the sampling period for different sampled data design methods are found in Åström and Wittenmark (1984). The choice of sampling period for minimum variance controllers is also discussed in MacGregor (1976) and Söderström and Lennartson (1981).

Example 5.5 - Choice of parameters

In this example the choice of parameters are discussed for the self-tuning regulator in Åström and Wittenmark (1973). The controller is based on implicit estimation and minimum variance control. The controller parameters are estimated from the model (5.1) and the controller is given by (5.2). The tuning parameters are deg S^*, deg R^*,

d, λ, $P(0)$, $\theta(0)$ and the sampling period, h.

The estimator parameters λ, $P(0)$, and $\theta(0)$ are not crucial and can often be given standard values.

The parameter d is the prediction horizon of the controller. For the optimal minimum variance controller is d the delay of the system in number of sampling periods, i.e. d is given by

$$d = d_o = \text{int}[\tau/h] + 1$$

where τ is the time delay of the process. The self-tuning controller can be used with a longer time horizon. This will decrease the variance of the control signal at the expense of a larger variance of the output. It is important that dh is not shorter than τ while it is safe to have it longer than τ.

For the minimum variance controller the order is given by

$$\deg S^* = n - 1$$
$$\deg R^* = n + d_o$$

where n and d_o are defined in (4.1). In practice it is often sufficient to choose $n = 1 - 3$. It is also possible to test if $\deg S^*$ and $\deg R^*$ are sufficiently large by monitoring $\hat{r}_y(\tau)$ and $\hat{r}_{yu}(\tau)$, see (5.3) and (5.4). If the system is controlled by an optimal minimum variance controller then $\hat{r}_y(\tau)$ and $\hat{r}_{yu}(\tau)$ are zero if $\tau \geq d - 1$.

The algorithm discussed here gives the optimal minimum variance controller only if the system to be controlled has a stable inverse. The algorithm will diverge if the process has zeros outside the unit circle. It is, however, possible to stabilize such systems by increasing the prediction horizon d. If $d = \max[\deg A^*, \deg B^* + d_o]$ then it can be shown that there exists a locally stable point for the parameter estimates. These parameter values correspond to a regulator such that the output is a moving average of order $n - 1$ and such that no process zeros are cancelled by the regulator. This regulator will thus not have any ringing in the control signal due to cancellation of process zeros. The algorithm will in this case converge to the suboptimal minimum variance controller for nonminimum phase systems given in Åström (1970). This new result is illustrated with a simulated example. Let the process be

$$y(k) + ay(k-1) = u(k-1) + bu(k-2) + e(k) + ce(k-1)$$

with a = -1, b = 2 and c = -0.5. Fig. 8 shows the parameters estimates when the basic algorithm is used with d = 2, deg R^* = 1 and deg S^* = 0.

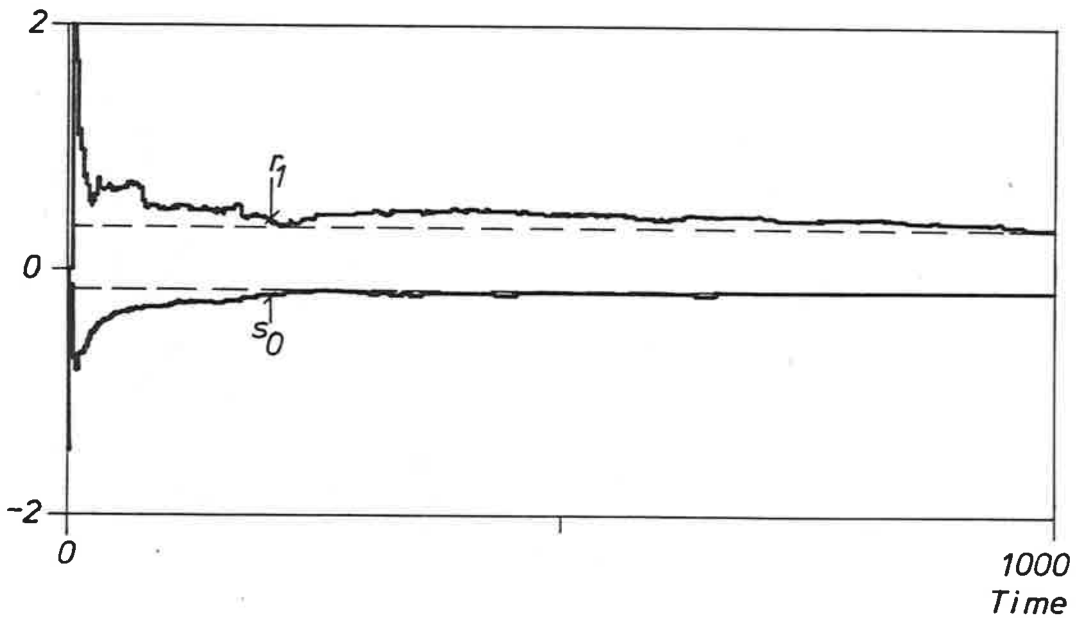

Fig. 8 Parameter estimates when controlling the process in Example 5.5 with the basic self-tuning algorithm with d = 2 and when there are two parameters in the controller. The dashed lines corresponds to the parameters of the suboptimal minimum variance controller.

The dashed lines are the parameters to which the regulator should converge. The suboptimal minimum variance controller gives in this case an average loss of 1.11 per step while the stable optimal minimum variance controller gives a loss of 1.08 per step. The basic self-tuning regulator with increased prediction horizon will perform as well as the suboptimal minimum variance controller after a few steps of time.

□

6. CONCLUSIONS

The experience from many feasibilities studies and applications of self-tuning regulators indicate that they can be successfully used in many situations. The conclusions from the applications indicate that it is not quite straightforward to use self-tuning regulators. There are many precautious that must be taken, as when using conventional controllers. Some of these practical aspects of self-tuning regulators have been discussed.

To summarize it is necessary to stress that it is important to have as much a priori knowledge about the process as possible. This knowledge is primarily used to choose design method and specifications. The parameters of the controller are then estimated and tuned by the controller. The advantages compared to conventional control are that the controller can follow process variations, that more complex controllers can be used, and that dead-time and feedforward compensations are easily handled. There are, however, much to be done both theoretically and practically before adaptive and self-tuning controllers can be applied routinely by unexperienced users.

7. ACKNOWLEDGEMENT

The research in adaptive control at the Department of Automatic Control has for many years been supported by the Swedish Board of Technical Development (STU). This report was prepared under contract 82-3430. This support is gratefully acknowledged. We also want to thank colleagues and students who have shared their knowledge and experience. A special thank to Tore Hägglund for the simulations given in Example 5.2.

8. REFERENCES

Allidina, A.Y. and Hughes, F.M. (1982): Self-Tuning Controller with Integral Action. Optimal Control Applications & Methods, 3, 355-362.

Åström, K.J. (1970): Introduction to Stochastic Control Theory, Academic Press, New York.

Åström, K.J. (1980a): Design Principles for Self-Tuning Regulators. In Unbehauen, H. (ed): Methods and Applications in Adaptive Control, Springer-Verlag, Berlin.

Åström, K.J. (1980b): Robustness of a Design Method Based on Assignment of Poles and Zeros, IEEE Trans on Automatic Control, AC-25, 588-591.

Åström, K.J. (1983a): Theory and Applications of Adaptive Control - A Survey, Automatica, 19, 471-486.

Åström, K.J. (1983b): Analysis of Rohrs Counterexamples to Adaptive Control. Proceedings of the 22nd Conference on Decision & Control, San Antonio, TX, USA, 982-987.

Åström, K.J. and Bohlin, T. (1966): Numerical Identification of Linear Dynamic Systems from Normal Operating Records. In P.N Hammond (Ed), Theory of Self-Adaptive Control Systems, Plenum Press, New York.

Åström, K.J., Borisson, U., Ljung, L. and Wittenmark, B. (1977): Theory and Applications of Self-Tuning Regulators, Automatica, 13, 457-476.

Åström, K.J. and Hägglund, T. (1984): Automatic Tuning of Simple Regulators with Specifications on Phase and Amplitude Margins, Automatica, 20, this issue.

Åström, K.J. and Wittenmark, B. (1973): On Self-Tuning Regulators, Automatica, 9, 185-199.

Åström, K.J. and Wittenmark, B. (1980): Self-Tuning Controllers Based on Pole-Zero Placement, IEE Proc, 127, PtD, 120-130.

Åström, K.J. and Wittenmark, B. (1984): Computer Controlled Systems - Theory and Design, Prentice-Hall, Englewood Cliffs, N.J.

Åström, K.J. and Zhao-Ying, Z. (1982): A Linear Quadratic Gaussian Self-Tuner, Proc. Workshop on Adaptive Control, Florence, Italy.

Bengtsson, G. (1979): Industrial Experience with a Microcomputer Based Self-Tuning Regulator. Proceedings of the 18th IEEE Conference on Decision & Control, Fort Lauderdale, Florida.

Bierman, G.J. (1977): Factorization Methods for Discrete Sequential Estimation, Academic Press, New York.

Clarke, D.W. (1982): Model Following and Pole-Placement Self-Tuners, Optimal Control Applications & Methods, 3, 323-335.

Clarke, D.W. and Gawthrop, P.J. (1979): Implementation and Application of Microprocessor Based Self-Tuners, Preprints 5th IFAC Symposium on Identification and System Parameter Estimation, Darmstadt, September 1979, 197-208.

Doyle, J.C. and Stein, G. (1981): Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis, IEEE Trans on Automatic Control, AC-26, 4-16.

Egardt, B. (1979): Stability of Adaptive Controllers, Lecture Notes in Control and Information Sciences, 20, Springer-Verlag, Berlin.

Egardt, B. (1980): Stability Analysis of Continuous-Time Adaptive Control Systems, SIAM J. of Control and Optimization, 18, 540-557.

Fortescue, T.R., Kershenbaum, L.S. and Ydstie, B.E. (1981): Implementation of Self-Tuning Regulators with Variable Forgetting Factors, Automatica, 17, 831-835.

Gawthrop, P.J. and Lim, K.W. (1982): Robustness of Self-Tuning Controllers, IEE Proceeding, 129, PtD, 21-29.

Goodwin, G.C., Ramadge, P.J. and Caines, P.E. (1980): Discrete Multivariable Adaptive Control, IEEE Trans on Automatic Control, AC-25, 449-456.

Horowitz, I.M. (1963): Synthesis of Feedback Systems. Academic Press, New York.

Hägglund, T. (1984): New Estimation Techniques for Adaptive Control, PhD Thesis Report TFRT-1025, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Irving, E. (1979): Improving Power Network Stability and Unit Stress with Adaptive Generator Control, Automatica, 15, 31-46.

Källström, C.G., Åström, K.J., Thorell, N.E., Eriksson, J. and Sten, L. (1979): Adaptive Autopilots for Tankers, Automatica, 15, 241-254.

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

MacGregor, J.F. (1976): Optimal Choice of the Sampling Interval for Discrete Process Control, Technometrics, 18, 151-160.

Mannerfelt, C.F. (1981): Robust Control Design with Simplified Models, PhD Thesis TFRT-1021, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Morse A.S. (1980): Global Stability of Parameter Adaptive Control Systems, IEEE Trans on Automatic Control, AC-25, 433-439.

Narendra, K.S., Lin, Y.H. and Valavani, L.S.(1980): Stable Adaptive Control Design, IEEE Trans on Automatic Control, AC-25, 440-448.

Peterson, B.B. and Narendra, K.S. (1982): Bounded Error Adaptive Control, IEEE Trans on Automatic Control, AC-27, 1161-1168.

Rohrs, C.E. Valavani, L., Athans, M. and Stein, G. (1981): Analytical Verification of Undesirable Properties of Direct Model Reference Adaptive Control Algorithms, Proceedings of the 20th IEEE Conference on Decision & Control, San Diego, 1271-1284.

Rohrs, C.E., Valavani, L, Athans, M. and Stein, G. (1982): Robustness of Adaptive Control Algorithms in the Presence of Unmodeled Dynamics, Proceedings of the 21st IEEE Conference on Decision & Control, Orlando Florida, 3-11.

Söderström, T. and Lennartson, B. (1981): On Linear Optimal Control with Infrequent Output Sampling, In Marshall J.E. et al (eds): Third IMA Conference on Control Theory, Academic Press, New York.

Wellstead, P.E. and Sanoff, S.P. (1981): Extended Self-Tuning Algorithm, Int. J. Control, 34, 433-455.

Wellstead, P.E. and Zanker, P. (1982): Techniques of Self-Tuning, Optimal Control Applications and Methods, 3, 305-322.

Wittenmark, B, and Aström, K.J. (1982): Implementation Aspects of Adaptive Controllers and Their Influence on Robustness, Proceedings of the 21st IEEE Conference on Decision & Control, Orlando, Florida.