## Lund University

## A Comparison of Two Suboptimal Dual Controllers on a First-order System

Sternby, Jan; Pernebo, Lars

1977

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

Citation for published version (APA):
Sternby, J., \& Pernebo, L. (1977). A Comparison of Two Suboptimal Dual Controllers on a First-order System. (Technical Reports TFRT-7117). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
2

## General rights

Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

## Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

A COMPARISON OF TWO SUBOPTIMAL DUAL CONTROLLERS ON A FIRST-ORDER SYSTEM

LARS PERNEBO

JAN STERNBY

Lund Institute of Technology Department of Automatic Control April 1977
Dounmontutgavare
Lund Institute of Technology
Handidogere Dept of Automatic Control
Forratare
Lar's Pernebo
Jan Sternby

Dokumentnamn
Dokumentbereckning
REPORT
Utgivningedatum
LUTFD2/(TFRK゙゙ 7117 )/1-30/(1977)
Aprit 1977
Forfattare
Jan Sternby

Dokumentiltel och undertitel
A Comparison of Two Suboptimal Dual Controllers on a First-Order System

## Raferat (ammandrag)

Two suboptimal dual controllers are applied to a first order linear system. The first controller is developed by Sternby and the second by Tse, Bar-Shalom, and Meier. The second controller is slightly modified to serve in a stationary situation. The only unknown parameter of the system is the gain, which varies sinusoidally. It is shown that the effect on the system of the two controllers is much the same even though they are based on completely different calculations

Aaferat okrivat av
Authors
Förslag till ytterligars nyckelord
$\therefore$ : 1 !

Klassitikationssystem och -klass(er)

Indextermer (enge källa)

Omfang
Ovriga bibliografiska uppglfter
30 pages
$1 \because 1$
Sprak
English
Sekretestuppgifter :ISSN

Dokumentet kan erhallas frân
Mottagarens uppgifter
Department of Automatic Control
( $1 .{ }^{\circ}$
Lund Institute of Technology
P 0 Box 725, S-220 07 Lund 7, Sweden
Pris

# A COMPARISON OF TWO SUBOPTIMAL DUAL CONTROLLERS ON A FIRST-ORDER SYSTEM 

Lars Pernebo and Jan Sternby

## Abstract

Two suboptimal dual controllers are applied to a first order linear system. The first controller is developed by Sternby and the second by Tse, Bar-Shalom, and Meier. The second controller is slightly modified to serve in a stationary situation. The only unknown parameter of the system is the gain, which varies sinusoidally. It is shown that the effect on the system of the two controllers is much the same even though they are based on completely different calculations.

CONTENTS

1. INTRODUCTION
2. THE SYSTEM AND THE CRITERION
3. DESCRIPTION OF THE FIRST CONTROLLER
4. DESCRIPTION OF THE SECOND CONTROLLER
4.1 Preliminaries
4.2 The Algorithm
4.3 Some Modifications of the Algorithm
5. SIMULATIONS
6. CONCLUSIONS
7. REFERENCES
8. INTRODUCTION

The concept of dual control was introduced already by Feldbaum (1960). Still very few cases are known, however, where the optimal dual control can actually be calculated. It is therefore necessary to find suitable approximations that will give good suboptimal dual controllers. These will in most cases have to be tested by simulations.

The aim of this report is to compare the performance of two suboptimal dual controllers on a first order single--input single-output system with a time-varying gain.

The first controller is obtained by minimizing an approximation to the expected value of the sum of squares of the output in the next two steps. The second controller is a slight modification of the one described by Tse and Bar-Shalom(1973).

The report is organized as follows. The system is defined in chapter two, where also the optimality criterion is specified. Chapters three and four describe one regulator each. In chapter five the simulation results are shown and discussed. Finally some concluding remarks are given in chapter six.

## 2. THE SYSTEM AND THE CRITERION

Consider a system governed by the difference equation

$$
\begin{equation*}
y(t+1)-y(t)=b(t) u(t)+e(t+1) \tag{2,1}
\end{equation*}
$$

with

$$
\begin{equation*}
\mathrm{b}(\mathrm{t})=1+1.5 \sin 0.04 \cdot t \tag{2.2}
\end{equation*}
$$

The noise e(•) is a sequence of independent, Gaussian random variables with zero mean and variance $Q$. Notice that the sign of the gain is not constant. This makes the system difficult to control.

Now, let the purpose of control be to minimize the criterion

$$
\begin{equation*}
\lim _{\mathrm{N} \rightarrow \infty} E \frac{1}{\mathrm{~N}} \sum_{\mathrm{k}=1}^{\mathrm{N}} \mathrm{y}^{2}(\mathrm{k}) \tag{2.3}
\end{equation*}
$$

i.e. the output should be kept as close to zero as possible. Finally, the admissible control strategies must be specified. For an admissible control strategy the control signal at time $t, u(t)$, may be a function of all the outputs observed up to time $t, i . e . y(t), y(t-1), \ldots$, all previously applied inputs $u(t-1), u(t-2), \ldots$, and the mean and variance of the noise. The coefficient $b(t)$ is considered as completely unknown to the regulator. Thus the eqation (2.1) is supposed to be known, but not eqation (2.2). A good control strategy will therefore include some form of estimation of $b(t)$.

The problem of finding an admissible control strategy for the system (2.1) that minimizes the criterion (2.3) is, however, too difficult to solve. Several suboptimal control strategies for this kind of problem have been proposed in the literature. In this report the performance of the system is examined when using two different suboptimal control strategies.

## 3. DESCRIPTION OF THE FIRST CONTROLLER

One suboptimal control strategy is obtained by minimizing

$$
V_{n}=E\left[\sum_{k=1}^{n}(y(t+k))^{2}\right]
$$

with $\mathrm{n}=2$ at every time t . The example in Sternby (1976) may serve as a motivation for doing so, even if that example is not very practical. There, however, the difference between using $n=2$ or a larger $n$ is small, while the regulator obtained with $\mathrm{n}=\mathrm{l}$ is not as good.

In order to estimate $b(t)$ some kind of model is needed. This is taken as

$$
\begin{equation*}
b(t+1)=b(t)+v(t+1) \tag{3.1}
\end{equation*}
$$

where $v(\cdot)$ is a sequence of independent Gaussian random variables with zero mean and variance $R$. The sequences $e(\cdot)$ and $v(\cdot)$ are assumed to be independent.

Now take $\mathrm{b}(\cdot)$ as the state that obeys equation (3.1) and $y(\cdot)$ as the observation according to equation (2.1). It is then possible to set up a Kalman filter to estimate b(.) as in Aström, Wittenmark (1971). Denote the estimate at time $t$ of $b(t)$ by $\hat{b}(t)$ and its variance by $P_{b}(t)$. Then

$$
\begin{align*}
& \hat{b}(t+1)=\hat{b}(t)+K(t)[y(t+l)-y(t)-\hat{b}(t) u(t)]  \tag{3.2}\\
& P_{b}(t+l)=\frac{P_{b}(t) \cdot Q}{u(t)^{2} P_{b}(t)+Q}+R  \tag{3.3}\\
& K(t)=\frac{P_{b}(t) \cdot u(t)}{u(L)^{2} P_{b}(L)+Q} \tag{3.4}
\end{align*}
$$

A good starting point for the filter is $\hat{b}(0)=E(b(0))$, $P_{b}(0)=\operatorname{Var}(b(0))$. If the distribution of $b(0)$ is not known
$P_{b}(0)$ should be taken large, so that $\hat{b}(t)$ can take large steps in the beginning. Notice that $\hat{b}(0)=0$ is a poor choice if the regulator gives $u=0$ for $\hat{b}=0$, since the result will then be that $\hat{b}(t) \equiv \hat{b}(0)$.

The minimization of $\mathrm{V}_{1}$ and $\mathrm{V}_{2}$ for the system (2.1) assuming model (3.1) for $b(t)$ is discussed in Aström, Wittenmark (1971). Using their results it can be shown that for our case the minimal value $V_{l}^{*}$ of $V_{1}$ is

$$
\begin{equation*}
v_{1}^{*}=\min _{u(t)}\left\{[y(t)+\hat{b}(t) u(t)]^{2}+Q+u(t)^{2} P_{b}(t)\right\} \tag{3.5}
\end{equation*}
$$

The corresponding $u(t)$ is

$$
\begin{equation*}
u_{1}^{*}(t)=-\frac{\hat{b}(t) y(t)}{\hat{b}(t)^{2}+P_{b}(t)} \tag{3.6}
\end{equation*}
$$

This is thus a cautious, but non-dual controller, since it takes into account the present uncertainties, but not the future observation program.

The minimal value $\mathrm{V}_{2}^{*}$ of $\mathrm{V}_{2}$ is given by

$$
\begin{align*}
V_{2}^{*}= & \min _{u(t)}\left\{[Y(t)+\hat{b}(t) u(t)]^{2}+Q+u(t)^{2} P_{b}(t)+\right. \\
& \left.+E\left[\left.Q+\frac{P_{b}(t+1) Y(t+1)^{2}}{\hat{b}(t+1)^{2}+P_{b}(t+1)} \right\rvert\, F_{t}\right]\right\} \tag{3.7}
\end{align*}
$$

The first three terms are the expected loss due to $y(t+l)$ and the last term is the conditional expectation of the minimal loss from $y(t+2)$ given $F_{t}=[y(t), y(t-1), \ldots$ ..., $u(t-1), u(t-2), .$.$] .$

The basic random variable in the expectation is $y(t+1)$, whose conditional distribution given $F_{t}$ is Gaussian. However, $y(t+l)$ occurs in both the numerator and in the denominator (through $\hat{b}(t+1)$ ). It is therefore not possible
to get a simple analytical expression for the expectation. In Sternby (1977) the following approximation was suggested

$$
E\left[\left.\begin{array}{l|}
\frac{P_{b}(t+1) y(t+1)^{2}}{\hat{b}(t+1)^{2}+P_{b}(t+1)} \tag{3.8}
\end{array} \right\rvert\, F_{t}\right] \approx \frac{E\left[P_{b}(t+1) y(t+1)^{2} \mid F_{t}\right]}{E\left[\hat{b}(t+1)^{2}+P_{b}(t+1) \mid F_{t}\right]}
$$

It originates from a series expansion of the denominator.
But using (3.2) - (3.4) gives for the denominator

$$
E\left[\hat{b}(t+1)^{2}+P_{b}(t+1) \mid F_{t}\right]=\hat{b}(t)^{2}+P_{b}(t)+R
$$

Now insert the approximation into (3.7) to get

$$
\begin{align*}
v_{2}^{*} \approx & \min _{u(t)}\left\{\left([y(t)+\hat{b}(t) u(t)]^{2}+Q+u(t)^{2} P(t)\right)\right. \\
& \left.\cdot\left(1+\frac{P_{b}(t) \cdot Q+R\left(u(t)^{2} P_{b}(t)+Q\right)}{\left(\hat{b}(t)^{2}+P_{b}(t)+R\right)\left(u(t)^{2} P_{b}(t)+Q\right)}\right)+Q\right\} \tag{3.9}
\end{align*}
$$

To find the minimum the derivative w.r.t. $u(t)$ is calculated, and its zeroes are determined using a root finding algorithm. Finally the minimum is found by comparing the losses for these values of $u(t)$. For further details about this regulator refer to Sternby (1977).

## 4. DESCRIPTION OF THE SECOND CONTROLLER

### 4.1 Preliminaries

The second controller to be considered is described in Tse, Bar-Shalom (1973). Some of the theoretical background can be found in Tse, Bar-Shalom, Meier (1973).

The system considered in Tse, Bar-Shalom (1973) is described by

$$
\left\{\begin{array}{l}
x(t+1)=A(t, \theta(t)) x(t)+B(t, \theta(t)) u(t)+e(t+1)  \tag{4.1}\\
y(t)=C(t, \theta(t)) x(t)+w(t), \quad t \geqslant 0
\end{array}\right.
$$

where the control $u(t)$ is scalar and $\theta(t) \in R^{\mathbf{S}}$ is a vector of unknown parameters described by the Markov process

$$
\begin{equation*}
\theta(t+1)=D(t) \theta(t)+v(t+l) \quad t \geqslant 0 \tag{4.2}
\end{equation*}
$$

The vectors $\{x(0), \theta(0), e(t), w(t), v(t), t \geqslant 0\}$ are assumed to be mutually independent Gaussian random variables with known statistics

$$
\begin{align*}
& x(0) \sim N\left(\hat{x}(0), P_{x x}(0)\right) \\
& \theta(0) \sim N\left(\hat{\theta}(0), P_{\Theta \Theta}(0)\right) \\
& e(t) \sim N(0, Q(t))  \tag{4.3}\\
& w(t) \sim N(0, G(t)) \\
& v(t) \sim N(0, R(t))
\end{align*}
$$

The unknown parameter $\theta(t)$ is assumed to enter linearly in $A(t, \cdot), B(t, \cdot)$ and $C(t, \cdot)$.

The objective is to find an admissible control sequence $\{u(t)\}_{t=0}^{\mathrm{N}-1}$ such that the cost functional

$$
\begin{align*}
J= & E\left[(x(N)-\rho(N))^{T} W(N)(x(N)-\rho(N))+\right. \\
& \left.+\sum_{k=0}^{N-1}(x(k)-\rho(k))^{T} W(k)(x(k)-\rho(k))+\lambda(k) u^{2}(k)\right] \tag{4.4}
\end{align*}
$$

is minimized. It is assumed that $W(k) \geqslant 0, \lambda(k)>0$ and that the reference trajectory $\{\rho(k)\}_{k=0}^{N}$ is given a priori. Below, the algorithm in Tse, Bar-Shalom (1973) will be described when applied to the system in chapter 2. In this special case most of the formulas of the algorithm take a much simpler form.

The system (2.1) can be described by the scalar equations

$$
\left\{\begin{array}{l}
x(t+1)=x(t)+b(t) u(t)+e(t+1)  \tag{4.5}\\
y(t)=x(t), \quad t \geqslant 0
\end{array}\right.
$$

Since the variation of $b(t)$, given by (2.2), is considered as unknown, $b(t)$ will instead be supposed to be given by

$$
\begin{equation*}
b(t+1)=b(t)+v(t+1), \quad t \geqslant 0 . \tag{4.6}
\end{equation*}
$$

The objective is to minimize the criterion (2.3). However, for this problem to fit in with the formulation in Tse, Bar-Shalom (1973) let us start with considering the cost functional

$$
\begin{equation*}
J=E\left[W x^{2}(N)+\sum_{k=0}^{N-1} W x^{2}(k)+\lambda u^{2}(k)\right] \tag{4.7}
\end{equation*}
$$

The system (4.5) - (4.6) can be written

$$
\left\{\begin{array}{l}
\binom{x(t+l)}{b(t+1)}=\left(\begin{array}{cc}
1 & u(t) \\
0 & 1
\end{array}\right)\binom{x(t)}{b(t)}+\binom{e(t+1)}{v(t+1)} \\
y(t)=\left(\begin{array}{ll}
1 & 0
\end{array}\right)\binom{x(t)}{b(t)} \tag{4.8b}
\end{array}\right.
$$

### 4.2 The Algorithm

Step 1. Denote the present time by $t$ and suppose that the estimates $\binom{\hat{x}(t \mid t-1)}{\hat{b}(t \mid t-1)}$ and $P(t \mid t-1)$ are given. Here $\binom{\hat{x}(t \mid t-l)}{\hat{b}(t \mid t-1)}$ is some estimate of $\binom{x(t)}{b(t)}$ given $y_{t-1}=\{y(t-1), y(t-2), \ldots\}$. Analogously $P(t \mid t-1)$ is some estimate of the covariance of $\binom{x(t)}{b(t)}$ given $y_{t-1}$. Furthermore, suppose that $y(t)$ is available.

Step 2. Compute $\binom{\hat{x}(t \mid t)}{\hat{b}(t \mid t)}$ and $P(t \mid t)$ using a second order filter. See e.g. Jazwinski (1970). Since (4.8b) is linear this will be identical to a Kalmanfilter, which in this case gives

$$
\begin{align*}
& \hat{x}(t \mid t)=y(t)  \tag{4.9}\\
& \hat{b}(t \mid t)=\hat{b}(t \mid t-1)+\frac{P_{x b}(t \mid t-1)}{P_{x x}(t \mid t-1)}(y(t)-\hat{x}(t \mid t-1))  \tag{4.10}\\
& P(t \mid t)=\left(\begin{array}{cc}
0 & 0 \\
0 & P_{b b}(t \mid t-1)-\frac{P_{x b}^{2}(t \mid t-1)}{P_{x x}(t \mid t-1)}
\end{array}\right)
\end{align*}
$$

The matrix $P$ is partitioned as $P=\left(\begin{array}{ll}P_{x x} & P_{x b} \\ P_{x b} & P_{b b}\end{array}\right)$.

Step 3. Solve the optimal control problem for the system

$$
z(j+1)=z(j)+\hat{b}(t \mid t) \bar{u}(j) \quad j=t, \ldots, N-1
$$

with the cost functional

$$
\bar{J}=W z^{2}(N)+\sum_{j=t}^{N-l} W z^{2}(j)+\lambda \bar{u}^{2}(j)
$$

The solution is given by

$$
\bar{u}(j)=-L(j) z(j) \quad j=t, \ldots, N-l
$$

where

$$
\begin{align*}
& L(j)=\frac{S(j+1) \hat{b}(t \mid t)}{\lambda+S(j+1) \hat{b}^{2}(t \mid t)}  \tag{4.12}\\
& S(j)=\left(1-\frac{S(j+1) \hat{b}^{2}(t \mid t)}{\lambda+S(j+1) \hat{b}^{2}(t \mid t)}\right) S(j+1)+W \quad S(N)=W \tag{4.13}
\end{align*}
$$

Step 4. In the sequel the so called remaining dual cost will be calculated. Given $\hat{x}(t \mid t)$ and $\hat{b}(t \mid t)$ it will be a function of $u(t)$ only. The desired value of $u(t)$ is then obtained by minimizing the dual cost numerically with respect to $u(t)$. As a starting value for the numerical minimization algorithm is chosen $u(t)=-L(t) \hat{x}(t \mid t)$.

Step 5. (The calculation of the remaining dual cost as a function of $u(t)$ starts here.)
Calculate $\binom{\hat{X}(t+l \mid t)}{\hat{C}(t+l \mid t)}$ and $P(t+l \mid t)$ using a second order predictor. (See e.g. Jazwinski (1970).) Since, in this case, the right member of (4.8a) is linear (u(t) fixed) we get

$$
\begin{align*}
& \hat{x}(t+l \mid t)=\hat{x}(t \mid t)+\hat{b}(t \mid t) u(t)  \tag{4.14}\\
& \hat{b}(t+l \mid t)=\hat{b}(t \mid t)  \tag{4.15}\\
& P(t+l \mid t)=\left(\begin{array}{cc}
1 & u(t) \\
0 & 1
\end{array}\right) P(t \mid t)\left(\begin{array}{cc}
1 & 0 \\
u(t) & 1
\end{array}\right)+\left(\begin{array}{ll}
Q & 0 \\
0 & R
\end{array}\right) \tag{4.16}
\end{align*}
$$

Step 6. Compute a future nominal trajectory $\left\{x_{0}(j)\right\}_{j=t+1}^{N}$, $\left\{u_{0}(j)\right\}_{j=t+1}^{N-1} \quad$ through

$$
\begin{align*}
& \binom{x_{0}(j+1)}{b_{0}(j+1)}=\left(\begin{array}{cc}
1 & u_{0}(j) \\
0 & 1
\end{array}\right)\left[\begin{array}{l}
x_{0}(j) \\
b_{0}(j)
\end{array}\right)  \tag{4.17}\\
& u_{0}(j)=-L(j) x_{0}(j) \tag{4.18}
\end{align*}
$$

with starting point

$$
\binom{x_{0}(t+l)}{b_{0}(t+l)}=\binom{\hat{x}(t+1 \mid t)}{\hat{b}(t+1 \mid t)}
$$

Here $\{L(j)\}_{j=t+1}^{N-1}$ is given by (4.12).

Step 7. Write

$$
\begin{align*}
& \binom{x(j)}{b(j)}=\binom{x_{0}(j)}{b_{0}(j)}+\binom{\delta x(j)}{\delta b(j)}  \tag{4.19}\\
& u(j)=u_{0}(j)+\delta u(j) \quad j=t+1, \ldots, N
\end{align*}
$$

The deviations from the nominal values are due to the noise and the fact that future measurements will give other estimates of $b(j), j>t+1$, which will give future control laws that differ from $L(j), j>t+l$. It is the fact that the algorithm takes the possibility of future changes of the unknown parameters into account, that makes the control strategy dual.

Introducing (4.19) into (4.8a) gives

$$
\binom{\delta x(j+1)}{\delta b(j+1)}=\left(\begin{array}{cc}
1 & u_{0}(j)  \tag{4.20}\\
0 & 1
\end{array}\right)\binom{\delta x(j)}{\delta b(j)}+\binom{b_{0}(j)}{0} \delta u(j)+\binom{\delta b(j) \delta u(j)}{0}+\binom{e(j+1)}{v(j+1)}
$$

The remaining cost at time $t+l$ can be written

$$
\begin{align*}
& J(t+l)=E\left[W x^{2}(N)+\sum_{j=t+1}^{N-1} W x^{2}(j)+\lambda u^{2}(j)\right]= \\
& =E\left[W x_{0}^{2}(N)+\sum_{j=t+1}^{N-1} W x_{0}^{2}(j)+\lambda u_{0}^{2}(j)\right]+ \\
& +E\left[2 W x_{0}(N) \delta x(N)+W(\delta x(N))^{2}+\right. \\
& \left.+\sum_{j=t+1}^{N-1} 2 W x_{0}(j) \delta x(j)+W(\delta x(j))^{2}+2 \lambda u_{0}(j) \delta u(j)+\lambda(\delta u(j))^{2}\right] \tag{4.21}
\end{align*}
$$

Here the first term does not depend on $\delta u(j)$. Now the second term is minimized with respect to $\{\delta u(j)\}_{j=t+1}^{N-1}$ using (4.20) as a dynamical constraint. This is done using dynamic programming with retention of up to second order terms in the deviations $\delta x(j), \delta b(j)$ and $\delta u(j)$.

The (approximately) optimal value of $J(t+1)$ associated with the given nominal trajectory is

$$
\begin{align*}
J^{*}(t+1)= & S(t+1) \hat{x}^{2}(t+1 \mid t)+ \\
& +\operatorname{tr}\left[\sum_{j=t+1}^{N}(\bar{P}(j \mid j-1)-\bar{P}(j \mid j)) \bar{S}(j)\right], \tag{4.22}
\end{align*}
$$

where $\bar{P}(j \mid j-1)$ and $\bar{P}(j \mid j)$ are the estimation-error covariances given by the Kalman filter for the system (4.20) with $\delta b(j) \delta u(j)$ put equal to zero. The starting value is $\bar{P}(t+l \mid t)=P(t+l \mid t)$.

$$
\bar{S}(j)=\left(\begin{array}{ll}
s(j) & s_{x b}(j)  \tag{4.23}\\
s_{x b}(j) & s_{b b}(j)
\end{array}\right)
$$

where $S(j)$ is given by (4.13) and

$$
\begin{align*}
& S_{x b}(j)=S(j+1) u_{0}(j)+S_{x b}(j+1)- \\
& -\frac{\left(u_{0}(j) S(j+1)+S_{x b}(j+1)\right) b_{0}(j)+S(j+1) x_{0}(j+1)}{\lambda+S(j+1) b_{0}^{2}(j)} b_{0}(j) S(j+1)  \tag{4.24}\\
& S_{b b b}(j)=S(j+1) u_{0}^{2}(j)+2 S_{x b}(j+1) u_{0}(j)+S_{b b}(j+1)- \\
& -\frac{\left[b_{0}(j) S(j+1) u_{0}(j)+b_{0}(j) S_{x b}(j+1)+S(j+1) x_{0}(j+1)\right]^{2}}{\lambda+S(j+1) b_{0}^{2}(j)}  \tag{4.25}\\
& S_{x: b}(N)=S_{b b}(N)=0
\end{align*}
$$

The (approximately) optimal dual loss is now given by

$$
\begin{equation*}
J_{D}(u(t))=\lambda u^{2}(t)+J^{*}(t+l) \tag{4.26}
\end{equation*}
$$

Step 8. The dual loss (4.26) is minimized with respect to $u(t)$ using for instance a quasi-Newton method. Now $\binom{\hat{x}(t+1 \mid t)}{\hat{b}(t+1 \mid t)}$ and $P(t+1 \mid t)$ are calculated from (4.14)-(4.16).

### 4.3 Some Modifications of the Algorithm

The given algorithm demands a fix terminal time N. However the problem, as stated in section 2 , is to find a controller working in a steady state situation. The idea is then to assume that the difference between the terminal time N and the presen $L$ Lime $L$ is, not necessarlly constant, but always large. It is shown below that it is then not necessary to sum from t+l to N in (4.22) when forming the dual loss, but it is sufficient to sum from $t+1$ to $t+n$ for some fixed $n$.

Make the following assumptions.
(i) Define $S=\lim _{j \rightarrow-\infty} S(j)$, where $S(j)$ is given by (4.13). Given $n$, it is supposed that $N$ is always so much larger than $t+n$ so that $S(t+n) \approx s$.
(ii) Let $L$ be given by (4.12) with $S$ substituted for $S(j+1)$. Suppose that $n$ is chosen so large that

$$
\begin{equation*}
(1-\hat{b}(t \mid t) L)^{n-1} \hat{x}(t+1 \mid t) \approx 0 \tag{4.27}
\end{equation*}
$$

(iii) Suppose that $\bar{P}(t+n \mid t+n-1) \approx \lim \bar{P}(j \mid j-1)$ and $\bar{P}(t+n \mid t+n) \approx \lim _{j \rightarrow+\infty} \bar{P}(j \mid j)$.

Since the Riccati equation giving $S(j)$ normally converges in a few steps (for low order systems) then assumption (i) is reasonable for a controller designed to work over long time intervals.

Since L is designed by optimal control methods it will give an asymptotically stable closed loop system. Therefore $n$ can always be chosen such that (4.27) is fulfilled for given $\hat{b}(t \mid t)$ and $\hat{x}(t+l \mid t)$. It is however not possible to choose the same n for all $\hat{\mathrm{b}}(\mathrm{t} \mid \mathrm{t})$ and $\hat{\mathrm{x}}(\mathrm{t}+\mathrm{l} \mid \mathrm{t})$. In assumption (ii) it is supposed that there is an $n$ such that (4.27) is fulfilled for all $\hat{b}(t \mid t)$ and $\hat{x}(t+l \mid t)$ that will occur. Simulations have shown that this assumption is mostly fulfilled.

If assumption (ii) is fulfilled then the nominal trajectory $\left\{x_{0}(j)\right\},\left\{u_{0}(j)\right\}$ tend to zero, which means that the system (4.20) becomes time invariant. This means, in turn, that the limits $\lim \overline{\mathrm{P}}(\mathrm{jlj-l})$ and $\lim \overline{\mathrm{P}}(\mathrm{j} \mid j)$ exist. Therefore $j \rightarrow+\infty \quad j \rightarrow+\infty$
it is always, if necessary, possible to increase $n$ so that assumption (iii) is fulfilled.

Suppose assumptions (i), (ii) and (iii) are fulfilled. It will be shown that summation from $t+1$ to $t+n$ in (4.22) will give the same control $u(t)$ as summation from $t+1$ to $N$. Furthermore it is sufficient to solve (4.24) and (4.25) in the interval $[t+1, t+n]$ with $S_{x b}(t+n)=S_{b b}(t+n)=0$.

Assumptions (i) and (ii) ensure that $u_{0}(j) \approx 0$ and $x_{0}(j) \approx 0$ in the interval $[t+n, N]$. By (4.24) and (4.25) it follows that $S_{x b}(j)$ and $S_{b b}(j)$ is then equal to zero in the same interval. Equation (4.13) shows that $S(j)$ is independent of $u(t)$. It therefore follows that $\bar{S}(j)$ is independent of $u(t)$ in the interval [ $t+n, N]$. Assumption (iii) and (ii) ensure that $\bar{P}(j \mid j-1)$ and $\bar{P}(j \mid j)$ too are independent of $u(t)$ in the interval $[t+n, N]$. Therefore the sum

$$
\sum_{j=t+n+1}^{N}(\bar{P}(j \mid j-1)-\bar{P}(j \mid j)) \bar{S}(j)
$$

in (4.22) is independent of $u(t)$ and can be disregarded when $u(t)$ shall be determined to minimize (4.26).

Motivated by the results above we change the given algorithm so that $t+n$ is substituted for $N$ in (4.22). Furthermore (4.24) and (4.25) will be solved only in the interval $[t+1, t+n]$ with $S_{x b}(t+n)=S_{b b}(t+n)=0$. In step 4 and 6 L is substituted for $L(t)$ and $L(j)$. These modifications eliminate the terminal time N from the algorithm.

One more modification of the given algorithm is made. If $\hat{b}(t \mid t)=0$ then there is no steady state solution to (4.13). Therefore if $|\hat{b}(t \mid t)| \leqslant \varepsilon$ for some $\varepsilon>0$ then $\hat{b}(t \mid t)$ is put equal to $\varepsilon \cdot \operatorname{sign}(\hat{b}(t \mid t))$.

Finally a few words have to be said about the criterion (2.3) that was to be minimized. Since, in this case, $y(t)=x(t)$ Lle modified algorithm will minimize the criterion

$$
J=E\left[W y^{2}(N)+\sum_{t=0}^{N-1} W y^{2}(t)+\lambda u^{2}(t)\right]
$$

for large $N$. The effects occurring at the end of the time interval, when the Riccati equation (4.13) has not reached its steady state solution, are not taken into account. The algorithm is therefore (approximately) minimizing

$$
\lim _{\mathrm{N} \rightarrow+\infty} \frac{1}{\mathrm{~N}} \mathrm{~J} .
$$

## 5. SIMULATIONS

Let the two controllers described in section 3 and 4 be called Reg 1 and Reg 2 respectively. The controllers were applied to the system described by equations (2.1) and (2.2) with noisecovariance $\mathrm{Q}=$ l. Observe that the controllers are not allowed to be based on eq (2.2) but rather on eq (3.1).

The following initial conditions were chosen:

$$
\begin{array}{lll}
\text { The system } & \text { Reg l } & \text { Reg 2 } \\
\mathrm{y}(0)=0 & \hat{\mathrm{~b}}(0)=2 & \hat{\mathrm{x}}(0 \mid-1)=0 \\
& \mathrm{P}_{\mathrm{b}}(0)=1 & \hat{\mathrm{~b}}(0 \mid-1)=2 \\
& & \mathrm{P}(0 \mid-1)=I
\end{array}
$$

Simulations showed that the following parameter values give satisfactory results.

$$
\begin{array}{ll}
\text { Reg } 1 & \text { Reg } 2 \\
R=0.01 & R=0.01 \\
& \lambda / W=0.01 \\
& \varepsilon=0.1
\end{array}
$$

The system was simulated from $t=0$ to $t=300$. The simulation was done 25 times for each controller with 25 independent noiserealizations $\{e(t)\}_{t=0}^{300}$. The same 25 realizations were used for Reg 1 and Reg 2.

Let the value of the costfunctional at the end of simulation $i$ be $V^{(i)}$. Then the mean value of the cost is estimated as

$$
\begin{equation*}
M=\frac{1}{25} \sum_{i=1}^{25} V^{(i)} \tag{5.1}
\end{equation*}
$$

Let $\sigma_{V}$ be the standard deviation of $V$, where $V$ is the value of the costfunctional at the end of a simulation. Then $\sigma_{v}$ is estimated as

The value of $\sigma_{v}$ gives a measure of the magnitude of the expected difference between $M$ and the value of $V$ in $a$ single simulation. Since the simulations are independent the standard deviation of $M$ can be estimated as

$$
\begin{equation*}
\sigma_{\mathrm{M}}=\frac{1}{\sqrt{25}} \sigma_{\mathrm{v}^{\prime}} \tag{5.2a}
\end{equation*}
$$

which can serve as an estimate of the error in $M$.

Reg 1 was simulated for $\mathrm{n}=1$ and 2 while Reg 2 was simulated for $\mathrm{n}=1,2$ and 20 . The results are given below.

|  | Reg 1 |  | Reg 2 |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | $\mathrm{n}=2$ | $\mathrm{n}=1$ | $\mathrm{n}=20$ | $\mathrm{n}=2$ | $\mathrm{n}=1$ |
| $\mathrm{M} \pm \sigma_{M}$ | $475 \pm 14$ | $593 \pm 26$ | $454 \pm 13$ | $462 \pm 12$ | $496 \pm 14$ |
| $\sigma_{v}$ | 68 | 130 | 65 | 61 | 69 |
| CPU-time | 16.5 | 4.1 | 33.4 | 6.0 | 5.6 |

The CPU-time is given in seconds for one simulation on UNIVAC ll08. However, the simulations were performed using a large simulation package, and a part of the CPU-time is due to the administration of this package. Therefore the figures are not quite accurate, but may serve as an indication of the computation time required for each case. For Reg 2 with $n=20$ the condition (4.27) was fulfilled (the left member was less than $10^{-6}$ ) for every $t \in[0,300]$ but for $n=2$ and 1 it was never fulfilled. Despite this the performance for $n=2$ was almost as good as for $\mathrm{n}=20$, but the computation time was much shorter. This may serve as another motivation for taking $\mathrm{n}=2$ in section 3 .

Intuitively, for $\mathrm{n}=1$, Reg 1 and Reg 2 should be very much alike, since both are minimizing a one-step loss. There are, however, certain differences. For Reg 2 with $\mathrm{n}=\mathrm{l}$ the costfunctional to be minimized is given by

$$
\begin{equation*}
J_{D}(u(t))=\lambda u^{2}(t)+S\left[(\hat{x}(t \mid t)+\hat{b}(t \mid t) u(t))^{2}+P_{b b}(t \mid t) u^{2}(t)+R\right] \tag{5.3}
\end{equation*}
$$

$J_{D}(u(t))$ does, in this case, not depend on the future observation program, which means that the control has no dual effect. However, $J_{D}(u(t))$ depends on the present uncertainty of the parameter estimate $P_{b b}(t \mid t)$. This makes the controller a "cautious controller".

For Reg 1 with $n=1$ the cost functional is given by

$$
\begin{equation*}
v_{1}=(y(t)+\hat{b}(t) u(t))^{2}+Q+u^{2}(t) P_{b}(t) \tag{5.4}
\end{equation*}
$$

This also gives a "cautious controller". The main differences between Reg 1 and Reg 2 for $n=1$ are:
(i) $\lambda$ was chosen nonzero in (5.3).
(ii) (5.3) is minimized numerically while (5.4) is minimized analytically.
(iii) $P_{b b}(t \mid t)$ is not the same as $P_{b}(t)$ even though simulations have shown that they are very much alike.
(iv) In Reg $2|\hat{b}(t \mid t)|$ is bounded from below.

It is believed that the difference in performance is mainly due to (iv). Indeed, simulations of Reg 1 have shown that sometimes when $\hat{b}(t)$ is very close to zero it happens that $u(t)$ becomes too small to give a good performance and a turn off phenomenon occurs.

A comparison of the CPU-time for Reg 1 and Reg 2 with $n=2$ shows that Reg 2 requires much less time. Most of the
computation time for Reg 1 is however used by a rootfinding algorithm. This means that the computation time is very insensitive to the number of parameters to be estimated. For Reg 2, on the other hand, the computation time increases rapidly with the number of parameters. Furthermore the rootfinding algorithm of Reg 1 is believed to be not as efficient as it could be. With a better such algorithm the computation time can be reduced.

Another important aspect on the regulators is the number of parameters to be chosen in advance. For Reg l there is only one, R. This has to be chosen for Reg 2 as well. But with the latter regulator there is also $\varepsilon$ to choose and $\lambda / W$. On the other hand $\lambda / W$ adds to the flexibility of Reg 2 , and can be used to improve its performance.

For figures $l$ to 5 a typical simulation is shown in the interval $t \in[60,180]$. This is the most interesting piece, since the gain changes sign twice. The same noise realization is used for the two controllers and Reg l has $n=2$ while Reg 2 has $\mathrm{n}=2$ and $\mathrm{n}=20$.

On all pages corresponding curves for the three cases are shown with Reg 1 on top, Reg 2 with $n=2$ in the middle and Reg 2 with $\mathrm{n}=20$ at the bottom. The curves are all surprisingly similar. Yet Reg 1 and Reg 2 are calculated in completely different ways and use different estimation methods. It is only around the sign changes of the b-parameter that there are some differences.

The estimation method of Reg 1 responds more quickly when the b-parameter changes sign from positive to negative. After a while, however, the estimates are the same again. For b increasiny bolli estimalion methods give step-wise changes in the estimate. This is because the closed-loop system is unstable if $b$ and $\hat{b}$ are constant and $b>2 \hat{b}$. The output will then increase and this will give a large correction of the
estimate $\hat{b}$. This fact also explains why the estimate is much more accurate around the up-crossings of zero than around the down-crossings.

Another difference is that around the zero-crossings of $b$ Reg 1 and Reg 2 with $n=2$ give smaller inputs than Reg 2 with $\mathrm{n}=20$. This is because with a small number of steps to go not so much can be gained by making the input large to get a good estimation.


Figure 1. The parameter b and its estimate $\hat{\mathrm{b}}$.


Figure 2. The estimated variance of $b$.
250.

Figure 3. The lossfunction.


Figure 4. The output Y .


Figure 5. The input u.
6. CONCLUSIONS

The most striking result of the comparison in this report is the similarity in behaviour of the regulators. Even $n=1$ in Reg 1 gives a fairly good performance of the system. This indicates that other qualities than small differences in the resulting loss may be of interest when choosing between suboptimal dual controllers. It is thus of course desirable to have as few parameters as possible to choose before applying the regulator. In that respect Reg 1 is good, because it has only one, while Reg 2 has two. Also, in some applications, the execution time on a computer may be critical. For the simple example of this report Reg 2 with $n=2$ is much faster than Reg 1 , but the execution time for Reg 2 will increase rapidly with the number of parameters and will probably exceed that of Reg 1 already for two or maybe three parameters.

A well-known and much faster alternative is to use Reg 1 with $n=1$, but add an extra probing signal to the input to assure a good estimation. This has been done by several authors. Then again there is an extra parameter to choose, i.e. the amplitude of the extra input.

## 7. REFERENCES

Feldbaum, A A (1960):
Theory of Dual Control, I-IV.
Automat Remote Control 21 (1960) No 9, No 11, and 22 (1961) No 1 and No 2.

Jazwinski, A H (1970):
Stochastic Processes and Filtering Theory.
Academic Press.

Sternby, J (1976):
A Simple Dual Control Problem with an Analytical Solution.
Trans A-C 21, No 6, December 1976.

Sternby, J (1977):
Topics in Dual Control.
Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.
CODEN: LUTFD2/(TFRT-1012)/1-135/(1977)

Tse, E and Bar-Shalom, Y (1973): An Actively Adaptive Control for Linear Systems with Random Parameters via the Dual Control Approach. Trans A-C 18, No 2, April 1973.

Tse, E, Bar-Shalom, Y, and Meier, L (1973): Wide-Sense Adaptive Dual Control for Nonlinear Stochastic Systems. Trans A-C 18, No 2, April 1973.

Åström, K J and Wittenmark, B (1971):
Problems of Identification and Control. J Math Anal \& Appl 34, No 1, April 1971.

