

Almost Optimal Adaptive LQ Control: Observed State Case

Jan Willem Polderman¹, Yvo Boers², Krzysztof Arent³

Abstract—In this paper we propose an almost optimal indirect adaptive controller for input/state dynamical systems. The control part of the adaptive scheme is based on a modified LQ control law: by adding a time varying gain to the certainty equivalent control law we avoid the conflict between identification and control.

Keywords: Adaptive LQ control, Riccati equation, stability radius

Mathematics subject classification: 93C40

I. INTRODUCTION

It has been recognized by many authors that due to lack of excitation, there exists a fundamental conflict between identification and control in adaptive control schemes that are based on an *optimal control* design. See e.g. [2], [9], [10], [12], [14]. Recently it has been shown ([3]) that if the underlying controller design is based on the minimization of a quadratic cost criterion, then the costs incurred may be arbitrarily large. In [1] a first attempt was made to deal with this difficulty by adding a time-varying controller gain to the usual certainty-equivalence controller. In the present paper, this idea will be worked out further. The idea behind the time-varying gain is that it ensures that the true system will be identified and hence also the optimal controller. Of course, the time varying part of the controller should not destroy the stability of the closed-loop system. By exploiting the concept of stability radius, the time-varying gain is chosen such that stability will be preserved. Due to the time variations in the controller, the resulting closed-loop system will not behave optimally. However, an additional scaling factor allows to approximate the optimal behavior arbitrarily well. An interesting by-product of the analysis is that the solution of the algebraic Riccati equation associated with the true system may be obtained as the limit of the solution of the time-varying Riccati difference equation corresponding to the estimated models. The Riccati equation is needed for the determination of the optimal control law, but at the same time it provides a

lower bound for the stability radius.

The paper is organized as follows. In Section II we will introduce the problem precisely and we will provide detailed motivation. It will turn out that we may benefit from a small time varying gain on top of the usual certainty equivalent control law. How to choose this time varying part will be explained in Section III. Section III is divided into three subsections in each of which different aspects of the modified control are discussed. These aspects are: how to identify the true system, how to preserve stability and how to approximate the optimal costs.

The ideas developed in Section III will be used in Section IV to propose an adaptive control algorithm. The main result that we derive there is that the adaptively controlled system asymptotically approaches the optimal behavior. Here asymptotic is with respect to time *and* a design parameter in the time varying part of the feedback.

Finally, in Section V we draw some conclusions and we indicate some possible extensions of the results.

Due to space limitations we do not provide any proofs. The interested reader is referred to [11] where also the non-observed state case is covered.

II. PROBLEM STATEMENT

We consider the class of discrete time input/state dynamical systems

$$x_{k+1} = A_0 x_k + b_0 u_k \quad (1)$$

where $u \in \mathbb{R}$ is the input, $x \in \mathbb{R}^n$ is the state, $(A_0, b_0) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times 1}$ is a controllable, but otherwise unknown pair.

The control objective is the minimization of:

$$J(u, x_0) = \sum_{k=0}^{\infty} x_k^T x_k + r u_k^2 \quad r > 0 \quad (2)$$

The feedback that minimizes (2) is given by (see [8]):

$$u_k = f(A_0, b_0) x_k \quad (3)$$

where

$$f(A_0, b_0) = -(b_0^T P_0 b_0 + r)^{-1} b_0^T P_0 A_0 \quad (4)$$

and P_0 is a unique symmetric positive solution of the Algebraic Riccati Equation:

$$P_0 - A_0^T P_0 A_0 + A_0^T P_0 b_0 (b_0^T P_0 b_0 + r)^{-1} b_0^T P_0 A_0 - I = 0$$

¹Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, e-mail: j.w.polderman@math.utwente.nl (corresponding author)

²Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, e-mail: y.boers@ele.tue.nl

³Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, e-mail: k.arent@math.utwente.nl

The optimal costs are given by:

$$J^*(x_0) = x_0^T P_0 x_0 \quad (5)$$

Because (A_0, b_0) is unknown, we want to generate an input sequence u in an adaptive fashion.

If we proceed in the standard way, we would apply the controller (3) based on estimates of (A_0, b_0) obtained from closed-loop observations. However, in the case of LQ control there exists a conflict between closed-loop identification and control, see [10]. The nature of this conflict can be effectively explained using the notion of *closed-loop unfalsified model*, which we define below, see also [12].

Definition II.1 Consider the dynamical system (1). Let (A, b) be controllable and let $f(A, b)$ be defined as in (3), with (A_0, b_0) replaced by (A, b) . Assume that the sequence x_k is generated as follows

$$x_{k+1} = (A_0 + b_0 f(A, b))x_k \quad x_0 = x_* \quad (6)$$

If for all k

$$(A_0 + b_0 f(A, b))x_k = (A + b f(A, b))x_k \quad (7)$$

then the system

$$z_{k+1} = Az_k + bu_k \quad (8)$$

will be called a closed-loop unfalsified model of (1). The set of closed-loop unfalsified models is denoted by G . ■

Consider a standard adaptive algorithm. If we initialize this algorithm in $(A, b) \in G$ then the one step ahead prediction error will be identically zero (see (II.1)) and therefore the identification procedure used in this algorithm will be frozen. In that case, the true system will be controlled by an input sequence generated as follows:

$$u_k = f(A, b)x_k \quad (9)$$

The question now is what the value of the cost criterion (2) will be if the true system is controlled on the basis of the closed-loop unfalsified model (A, b) . In [10] it was shown that these costs will in general be larger than the optimal costs. More precisely, the only closed-loop unfalsified model on the basis of which the true system could be controlled optimally, is the true system itself. Moreover, it has been proven in [3] for the first order case, that the real costs as a function of the closed-loop unfalsified models, are unbounded.

Remark II.2 The set G can also be seen as the set of possible limit points of the sequence of estimates. Therefore, in the limit, the control law will be based on a closed-loop unfalsified model. ■

It follows from [10] that, if we want to avoid the conflict between identification and control we have to ensure that $G = \{(A_0, b_0)\}$.

Our approach is to use a controller of the form

$$u_k = f(\hat{A}_k, \hat{b}_k)x_k + l_k x_k \quad (10)$$

where l_k is a time varying gain to be designed. The additional time-varying gain should be such that:

- The only closed-loop unfalsified model of (1) controlled by (10) is the true system.
- Stability of the closed-loop system is preserved.
- It is possible to scale l_k such that the value of the cost criterion (2) using (10) can be arbitrary close to the optimal costs.

Seeking l_k that fulfill the above specifications constitutes one of the main goal of this paper. The proposed construction generalizes ideas that were developed for the first order case in [12].

The second goal is to apply these ideas in an adaptive control algorithm.

III. THE TIME VARYING PART OF THE CONTROL LAW

In Section II we suggested to modify the standard LQ control law. The idea is to add a time varying gain l_k to the certainty equivalent LQ feedback. In this section we will construct a time-varying gain that meets the specification described in Section II.

To streamline the discussion we have divided this section into three subsections. In each of these subsections we will discuss one of the three specifications.

A. How to make sure that $G = \{A_0, b_0\}$?

This subsection concerns a time varying gain l_k such that $G = \{(A_0, b_0)\}$.

Consider the equation (7) that defines the set of closed-loop unfalsified models G . The number of unknowns is $n^2 + n$, whereas the number of equations is at most n^2 (this is when $\text{span}\{x_i\} = n$). As a consequence it cannot be expected that this equation will have a *unique* solution. Therefore we will have to increase the number of equations. This is exactly what may be achieved by invoking the time varying feedback. The idea of the time varying feedback is to cyclically switch through an off-line determined set of feedback gains. In particular we will use \tilde{l}_1 for n iterations followed by \tilde{l}_2 for the next n iterations and so on. For convenience of notation we will define the switching mechanism more formally.

Definition III.1 Let $n, r \in \mathbb{N}$, $k \in \mathbb{Z}$. The function $s : \mathbb{Z} \rightarrow \{1, \dots, r\}$ is defined as:

$$s(k) = [(k \bmod nr) \div n] + 1 \quad (11)$$

Example III.2 For a set of r feedbacks $\tilde{l}_i \in \mathbb{R}^{1 \times n}$, we switch through this set as:

$$\underbrace{\tilde{l}_1, \dots, \tilde{l}_1}_{n \text{ times}}, \dots, \underbrace{\tilde{l}_r, \dots, \tilde{l}_r}_{n \text{ times}} \quad (12)$$

For $n = 2$, $r = 3$ this boils down to:

$$\tilde{l}_1, \tilde{l}_1, \tilde{l}_2, \tilde{l}_2, \tilde{l}_3, \tilde{l}_3, \tilde{l}_1, \tilde{l}_1, \dots \quad (13)$$

■

Now we can write l_k as follows

$$l_k = \tilde{l}_{s(k)} \quad (14)$$

Obviously, a minimum requirement on the time varying feedbacks is that the state trajectory spans the state space.

We will now show that the control law (10), where l_k is in the form (14), will indeed yield that (A_0, b_0) is the only closed-loop unfalsified model.

Theorem III.3: Consider the feedback interconnection of (1) and (10), and the set G introduced in Definition II.1. Assume that $x_k \neq 0 \forall k$, l_k in (10) is of the form (14) and $\tilde{l}_1, \dots, \tilde{l}_r$ are chosen such that

$$\bigcap_{i=1}^r \text{Spec}((A_0 + b_0 f(A, b)) + b_0 \tilde{l}_i) = \emptyset.$$

Then

$$G = \{(A_0, b_0)\} \quad (15)$$

■

A natural question concerning Theorem III.3 is whether there exist r and $\tilde{l}_1, \dots, \tilde{l}_r$, independent of the true system, such that the condition on the spectra is satisfied. We now show that the answer is affirmative.

Lemma III.4: Let $(A, b) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times 1}$ be a controllable pair, $\tilde{l}_1, \dots, \tilde{l}_{n+1} \in \mathbb{R}^{1 \times n}$ be such that they do not satisfy the same linear affine relation of the form $la = c$ ($a \in \mathbb{R}^{n \times 1}, c \in \mathbb{R}$). Then

$$\bigcap_{i=1}^{n+1} \text{Spec}(A + b \tilde{l}_i) = \emptyset.$$

■

Remark III.5 Notice that the condition that the \tilde{l}_i 's should not satisfy any relation of the form $al = c$, is scale-invariant. ■

Remark III.6 One can show that for every n -tuple $\tilde{l}_1, \dots, \tilde{l}_r$, there exists a controllable pair (M_0, b_0) such that $\bigcap_{i=0}^r \text{Spec}(M_0 + b_0 \tilde{l}_i) \neq \emptyset$. This indicates that $n+1$ is the minimal number needed to fulfill the condition in Theorem III.3. ■

B. How to ensure that stability is preserved?

In Subsection III-A we proposed a time varying gain for the control law (10) such that $G = \{(A_0, b_0)\}$. Of course, the time varying part of the feedback could destroy the asymptotic stability of the controlled system, unless the time varying part is sufficiently small in norm. In this subsection we will investigate how small the time-varying part should be. We will do this using theory of stability radius [5].

Whereas most of our results concern single-input systems only, the analysis in this section can be done for multi-input multi-output systems as well. Therefore we consider systems of the following form:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k \end{aligned} \quad (16)$$

where $(A, B, C) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{p \times n}$, (A, B) is controllable, (A, C) is observable.

We assume that (16) is controlled by the state feedback

$$u_k = (F(A, B, C) + L_k C)x_k \quad (17)$$

where $F(A, B, C)$ is the optimal feedback:

$$F(A, B, C) = -(B^T P B + R)^{-1} B^T P A \quad (18)$$

with $R = R^T > 0$, P the unique symmetric positive definite solution of the Algebraic Riccati Equation (ARE):

$$\begin{aligned} P - A^T P A + A^T P B (B^T P B + R)^{-1} B^T P A - \\ C^T C = 0 \end{aligned} \quad (19)$$

and L_k yet to be constructed.

If $L_k \equiv 0$ then (17) minimizes the quadratic cost criterion [7], [8]:

$$\sum_{k=0}^{\infty} y_k^T y_k + u_k^T R u_k \quad (20)$$

The closed loop interconnection of (16) and (17) is:

$$\begin{aligned} x_{k+1} &= (A + BF(A, B) + BL_k C)x_k \\ y_k &= Cx_k \end{aligned} \quad (21)$$

We are interested in the relation between the norm of L_k and the stability of (21).

Consider the matrix $A + BF(A, B) + BL_k C$. This matrix can be decomposed as the sum of a stable, constant matrix

$$M := A + BF(A, B, C) \quad (22)$$

and a time varying matrix $BL_k C$. The matrix $BL_k C$ can be seen as a perturbation of M . Of course, if the norm of L_k is sufficiently small, the perturbed matrix will be stable. How small L_k should be can be derived from the structured stability radius of the

triple (A, B, C) , see [6]. Although the perturbation is real, we prefer to work with the complex stability radius rather than with the real stability radius. The following definition may be found in [6].

Definition III.7 Let $M \in \mathbb{C}^{n \times n}$ have its eigenvalues in the open unit disk, $(B, C) \in \mathbb{C}^{n \times m} \times \mathbb{C}^{p \times n}$, $\bar{\mathbb{C}}_1$ be the complement of the open unit disk. The real number

$$r_{\mathbb{C}} = \inf\{\|D\| \mid \text{Spec}(M + BDC) \cap \bar{\mathbb{C}}_1 \neq \emptyset\} \quad (23)$$

is called the complex stability radius of the stable matrix M with respect to perturbations of the given structure. ■

The reason that we are interested in the complex stability radius is that if we could ensure that $\|L_k\|$ is smaller and bounded away from $r_{\mathbb{C}}$, than the resulting time-varying system will still be exponentially stable, as the following proposition expresses.

Proposition III.8: Suppose that $(M, B, C) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{p \times n}$, $\text{Spec}\{M\} \subset \mathbb{C}_1$ and $L_k : \mathbb{N} \rightarrow \mathbb{R}^{m \times n}$. If $\|L_k\| < r_{\mathbb{C}}(M, B, C) - \epsilon$ for all k and for some $\epsilon > 0$, then the system

$$x_{k+1} = (M + BL_kC)x_k \quad (24)$$

is exponentially stable ■

We want to use the stability radius as an upper bound on the norm of the time varying feedback adaptively. Calculating the stability radius at each iteration is out of the question, the problem is now how to obtain at least a lower bound for $r_{\mathbb{C}}$.

We will show that such a lower bound may be obtained from the controller Algebraic Riccati Equation (19) that we will have to solve anyway. In fact (19) will provide an upper bound on the \mathcal{H}_{∞} norm of the controlled system (M, B, C) . Since $r_{\mathbb{C}}$ is just the inverse of the \mathcal{H}_{∞} norm, see [5], an upper bound on the latter will provide a lower bound on the first.

Theorem III.9: Consider the matrices A, B, C that specify (16), ARE (19), $F(A, B, C)$ defined by (18), and M defined by (22). Let γ be the \mathcal{H}_{∞} norm of the system

$$\begin{aligned} x_{k+1} &= Mx_k + Bu_k \\ y_k &= Cx_k \end{aligned} \quad (25)$$

If

$$\bar{\gamma}^2 I \geq B^T P B + R \quad (26)$$

then

$$\gamma \leq \bar{\gamma} \quad (27)$$

Corollary III.10: If $m = 1$ then

$$r_{\mathbb{C}}(M, B, C) \geq \frac{1}{\sqrt{b^T P b + r}} \quad (28)$$

C. How much does the time-varying gain add to the costs?

Now that we know how to design the time varying feedback so as to ensure that the only closed-loop unfalsified model is the true system, we would like to know how much the modification of the usual certainty equivalence design adds to the costs. It should be clear that by uniqueness of the optimal control law, the modified controller will not be optimal. However, we will see that we can approximate the optimal costs arbitrarily well by sufficiently scaling down the time varying part of the feedback. This result still concerns the behavior of the controlled system when it is controlled on the basis of a closed-loop unfalsified model. Recall from the introduction that this behavior may be arbitrarily bad without the modification.

Theorem III.11: Let the system (1) be controlled on the basis of a closed-loop unfalsified model (A, b) as:

$u_k = f(A, b)x_k + \alpha \tilde{l}_{s(k)} x_k = f(A, b)x_k + \alpha l_k x_k, \alpha > 0$ where $\{\tilde{l}_i\}$ satisfy the conditions of Lemma III.4 and $s(k)$ is defined by (11). The optimal costs (5) may approximated arbitrarily well by choosing α sufficiently small. ■

IV. APPLICATION TO ADAPTIVE LQ CONTROL

The analysis presented thus far was non-adaptive. In this section we will use the ideas developed in the previous sections to design an adaptive control scheme based on an LQ design and using a time-varying feedback on top of the usual certainty equivalent feedback. For the identification of the system parameters we will use a standard projection algorithm. To calculate the optimal feedback corresponding to the estimates we would have to solve the algebraic Riccati equation for each iteration, which is numerically unacceptable. Instead, we will use the Riccati Difference equation to approximate the solution of the algebraic equation. Since due to the time variations in the feedback we are going to identify (A_0, b_0) , we may hope that the solution of the difference equation will converge to the positive solution of the algebraic equation. Then there is the issue of stability. The time-variations in the feedback should be smaller than the stability radius. Since we do not know the stability radius we use an estimate for a lower bound of the stability radius of the estimate. This lower bound will be obtained from the solution of the Riccati Difference Equation using Corollary III.10. The interesting feature of the algorithm is that the Riccati Difference Equation is used for approximating the certainty equivalent part of the feedback as well as for obtaining a lower bound of the stability radius.

A. Identification algorithm

The identification procedure that we are going to use is the orthogonal projection algorithm [4]. The estimate of (A_0, b_0) at time k is denoted by (\hat{A}_k, \hat{b}_k) . Two crucial properties of the projection algorithm are:

Lemma IV.1: For the estimates generated by the projection algorithm we have:

$$\begin{aligned} \|(\hat{A}_{k+1}, \hat{b}_{k+1}) - (A_0, b_0)\| &\leq \|(\hat{A}_k, \hat{b}_k) - (A_0, b_0)\| \\ \lim_{k \rightarrow \infty} \|(\hat{A}_{k+1}, \hat{b}_{k+1}) - (\hat{A}_k, \hat{b}_k)\| &= 0 \end{aligned}$$

B. Solving ARE by means of the Riccati Difference Equation

We will now show that under mild conditions the solution of the Riccati Difference Equation converges to the positive solution of ARE. The results also hold for multivariable systems and a more general cost criterion. However, for simplicity we confine ourselves to the class of systems that we are considering in this paper. The Riccati Difference Equation corresponding to the sequence of estimates (\hat{A}_k, \hat{b}_k) is given by:

$$\begin{aligned} \hat{P}_{k+1} &= \hat{A}_k^T \hat{P}_k \hat{A}_k - \hat{A}_k^T \hat{P}_k \hat{b}_k (\hat{b}_k^T \hat{P}_k \hat{b}_k + r)^{-1} \hat{b}_k^T \hat{P}_k \hat{A}_k \\ &+ I \quad \text{where } \hat{P}_0 \geq 0 \end{aligned}$$

As a standing (technical) assumption for this section we will take:

Assumption IV.2 The sequence of estimates (\hat{A}_k, \hat{b}_k) is contained in a compact subset of the set of controllable pairs. ■

Theorem IV.3: Assume that the sequence of estimates (\hat{A}_k, \hat{b}_k) , satisfies Assumption IV.2. Then \hat{P}_k is a bounded sequence of nonnegative definite matrices. As a consequence $(\hat{A}_k, \hat{b}_k) \rightarrow (A_0, b_0)$ and \hat{P}_k converges to the unique symmetric positive definite solution of

$$P - A_0^T P A_0 + A_0^T P b_0 (b_0^T P b_0 + r)^{-1} b_0^T P A_0 - I = 0 \quad (29)$$

For the proof of Theorem IV.3 we consider the following fictitious system:

$$x_{k+1} = \hat{A}_{N-1-k} x_k + \hat{b}_{N-1-k} u_k \quad (30)$$

and the cost criterion:

$$J^N(u) = \sum_{k=0}^{N-1} x_k^T x_k + r u_k^2 + x_N^T \hat{P}_0 x_N \quad (31)$$

The reason that we are interested in (30) and (31) is that \hat{P}_N is just the optimal value function for the finite horizon problem defined by (31). This observation is a simple consequence of standard dynamic

programming. Next we prove that there exists a feedback strategy for the finite horizon problem such that the resulting costs are bounded independent of the horizon N . This then implies boundedness of the sequence \hat{P}_k . Boundedness of \hat{P}_k is essential, for if we would not have that, the time-varying part of the feedback could then vanish.

Theorem IV.4: Consider the system (30) and the finite horizon cost criterion (31). There exists a feedback that yields costs that are bounded independent of the horizon N . ■

The proof of the existence of a feedback that yields finite costs for the system (30) with respect to the cost criterion (31), relies on Lemma IV.1.

Remark IV.5 Similar results may be found in [13], however, the proofs are not provided there. ■

C. The adaptive control algorithm

In this section we will propose the adaptive algorithm. To be able to use Theorem IV.3, we will assume that we have some prior knowledge about the true system.

Assumption IV.6 The true system parameters (A_0, b_0) (1) belong to a known closed and convex set Υ of controllable models. ■

Assumption IV.6 is a well known condition in adaptive control algorithms. It is usually assumed to avoid non-controllable models in the identification procedure. In the present algorithm it seems that due to the time varying part of the feedback, the assumption is in fact superfluous. This point is presently under investigation.

In the algorithm presented below, the estimates of (A_0, b_0) are denoted by (\hat{A}_k, \hat{b}_k) and are generated by the projection algorithm. The time varying part of the feedback contains $\tilde{l}_1, \dots, \tilde{l}_{n+1}$ which are assumed to be uniformly distributed over the unit ball so that they do not satisfy the same linear affine relation. The solution of the Riccati Difference Equation is denoted by \hat{P}_k , the estimate for the lower bound of the stability radius $r_C(\hat{M}_k, \hat{b}_k, I)$, where $\hat{M}_k = \hat{A}_k + \hat{b}_k f(\hat{A}_k, \hat{b}_k)$, will be denoted by $\hat{r}(\hat{A}_k, \hat{b}_k)$ or \hat{r}_k for short. Finally, α is a scaling factor: $0 < \alpha < 1$.

Algorithm IV.7

Initialization: $\hat{P}_0 \geq 0, (\hat{A}_0, \hat{b}_0) \in \Upsilon$

Recursion:

$$\begin{aligned} (\hat{A}_{k+1}, \hat{b}_{k+1}) = \\ (\hat{A}_k, \hat{b}_k) + (x_{k+1} - \hat{A}_k x_k - \hat{b}_k u_k) \frac{(x_k^T, u_k)}{\|(x_k^T, u_k)\|^2} \end{aligned}$$

$$\hat{P}_k = I + \hat{A}_{k-1}^T \hat{P}_{k-1} \hat{A}_{k-1} -$$

$$\hat{A}_{k-1}^T \hat{P}_{k-1} \hat{b}_{k-1} (\hat{b}_{k-1}^T \hat{P}_{k-1} \hat{b}_{k-1} + r)^{-1} \hat{b}_{k-1}^T \hat{P}_{k-1} \hat{A}_{k-1}$$

$$\hat{r}_k = \frac{1}{\sqrt{\hat{b}_k^T \hat{P}_k \hat{b}_k + r}}$$

$$u_k = -(\hat{b}_k^T \hat{P}_k \hat{b}_k + r)^{-1} \hat{b}_k^T \hat{P}_k \hat{A}_k x_k + \alpha \hat{r}_k \tilde{l}_{s(k)} x_k$$

The behavior of the adaptively controlled system is characterized by the following theorem.

Theorem IV.8: Let the system (1) be controlled by Algorithm IV.7. Assume that for all k , $x_k \neq 0$.

1. $\lim_{k \rightarrow \infty} (\hat{A}_k, \hat{b}_k) = (A_0, b_0)$
2. $\lim_{k \rightarrow \infty} x_k = 0$
3. $\lim_{\alpha \downarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{\|x_N\|^2} \left(\sum_{k=N}^{\infty} x_k^T x_k + r u_k^2 - x_N^T P_0 x_N \right) = 0$

The first and the second part of Theorem IV.8 state that the true system will be identified and that the adaptively controlled system is asymptotically stable. The third part expresses that the normalized asymptotic costs are arbitrarily close to the normalized optimal costs provided α is sufficiently small.

V. CONCLUSIONS

For the class of single input/state systems we proposed an almost optimal LQ controller. On the basis of this controller we designed an adaptive algorithm. The crucial property of this algorithm is that asymptotically the value of the LQ cost criterion approaches the optimal costs by choosing a design parameter sufficiently small.

An interesting by-product of the analysis is that the solution of the algebraic Riccati equation associated with the true system may be obtained as the limit of the solution of the time-varying Riccati difference equation corresponding to the estimated models.

Extensions to the case that the state is not available through the output only may be found in [11]. We would like to obtain similar results for a more general weighting factor for the costs associated with the state, the Q matrix in (2), which is identity in the present paper.

VI. ACKNOWLEDGMENT

The authors would like to thank A.A. Stoorvogel, S. Weiland and H.J. Zwart for helpful discussions and D. Hinrichsen for providing some of the references.

REFERENCES

- [1] Y. BOERS (1994) *Adaptive LQ Regelaars met gesloten lus excitatie*. (In Dutch.) Masters thesis, University of Twente, the Netherlands.
- [2] V. BORKAR AND P. VARAIYA (1979). Adaptive control of Markov chains, I: Finite parameter set. *IEEE Trans. Aut. Contr. AC-24*, 953-957.
- [3] P.T. DE BRUIN (1992): *Adaptieve LQ regeling met neutrale zekerheidsequivalentie algoritmes*. Master thesis (in Dutch), Dept. Appl. Math. University of Twente, The Netherlands.
- [4] G.C. GOODWIN AND K.S. SIN (1984): *Adaptive Filtering Prediction and Control*. Prentice-Hall, Inc., Englewood Cliffs, N.J.
- [5] D. HINRICHSEN AND A.J. PRICHARD (1991): On the robustness of stable discrete time linear systems. *Progress in Systems and Control Theory*, vol. 7, Birkhäuser.
- [6] D. HINRICHSEN AND N.K. SON (1991): Stability Radii of Linear Discrete-Time Systems and Symplectic Pencils. *International Journal of Robust and Nonlinear Control*, vol. 1, 79-97.
- [7] T. KAILATH (1980). *Linear Systems*. Prentice-Hall, Englewood Cliffs N.J.
- [8] H. KWAKERNAAK AND R. SIVAN (1972): *Linear Optimal Control Systems*. John Wiley & Sons, Inc..
- [9] W. LIN, P.R. KUMAR AND T.I. SEIDMAN. (1985) Will the self-tuning approach work for general cost criteria? *Systems & Control Letters* 6, 77-86.
- [10] J.W. POLDERMAN (1989). Adaptive LQ Control: Conflict between Identification and Control. *Linear Algebra and its Applications* 122, 123, 124, 219-244.
- [11] J.W. POLDERMAN AND K. ARENT (1995). *Almost Optimal Adaptive LQ Control: SISO Case*. Memorandum, University of Twente, Dept. Appl. Math.
- [12] J.W. POLDERMAN (1988): *Adaptive Control & Identification: Conflict or Conflux ?* CWI Tract 67, Center for Mathematics and Computer Science, Amsterdam, The Netherlands.
- [13] C. SAMSON (1982): An Adaptive LQ controller for non-minimum phase systems. *Int. J. Contr.*, vol. 35, no. 1, pp. 1-28.
- [14] J.H. VAN SCHUPPEN (1994): Tuning of Gaussian Stochastic Control Systems. *IEEE Trans. Aut. Contr.*, vol. 39, pp. 2178-2190.