TECHNICAL RESEARCH REPORT

Reduced Complexity Output Feedback Nonlinear HÄ Controllers and Relation to Certainty Equivalence

by J.S. Baras, N.S. Patel

T.R. 95-55

INSTITUTE FOR SYSTEMS RESEARCH

Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry

Reduced Complexity Output Feedback Nonlinear H_∞ Controllers and Relation to Certainty Equivalence*

John S. Baras† and Nital S. Patel
Dept. of Electrical Engineering and Institute for Systems Research
University of Maryland, College Park, MD 20742
baras@src.umd.edu   nsp@src.umd.edu

Abstract

In this paper, we consider the problem of construtcing reduced complexity controllers for output feedback nonlinear H_∞ control. We give sufficient conditions, under which the controllers so obtained, guarantee asymptotic stability of the closed-loop system when there are no exogenous inputs. The controllers obtained are non-optimal in general. However, in case optimality holds, we show that these controllers are in fact the certainty equivalence controllers.

1 Introduction

Since, Whittle [5] first postulated the minimum stress estimate for the solution of a risk-sensitive stochastic optimal control problem, it has evolved into the certainty equivalence principle. The latter states that under appropriate conditions, an optimal output feedback controller can be obtained by inserting an estimate of the state into the corresponding state feedback law. In general, however the controller so obtained is non-optimal. The certainty equivalence property is known to hold for linear systems with a quadratic cost [1]. The recent interest in nonlinear H_∞ control has led researchers to examine whether, certainty equivalence could be carried over to nonlinear systems. If certainty equivalence were to hold, it would result in a tremendous reduction in the complexity of the problem. In a recent paper [4], sufficient conditions were given for certainty equivalence to hold in terms of a saddle point condition. Also, in [2], a simple example is given to demonstrate the non-optimal nature of the certainty equivalence controller.

In this preliminary paper, we will be considering the infinite time case, and will present sufficiency conditions for a reduced complexity controller to exist. These conditions apply for both optimal and non-optimal policies. In general, obtaining an optimal solution to the output feedback problem, involves solving an infinite dimensional dynamic programming equation [3]. Hence, one may be satisfied with a reduced complexity non-optimal policy, which guarantees asymptotic stability of the nominal closed-loop system. In the special case, we show that the policies so obtained are certainty equivalence policies. Furthermore, in doing so, we will be able to give an equivalent sufficiency condition for certainty equivalence which may be more tractable than the one given in [4]. We will be interested in establishing dissipativity results, since these guarantee under detectability assumptions, asymptotic stability of the closed-loop system when exogenous inputs are zero.

*This work was supported by the National Science Foundation Engineering Research Centers Program: NSFD CDR 8803012

†Martin Marietta Chair in Systems Engineering

2 Problem Statement

We consider the following the system:

$$\Sigma \begin{cases} x_{k+1} &= f(x_k, u_k, w_k)\ , \quad x_0 \in \boldsymbol{R}^n \\ y_{k+1} &= g(x_k, u_k, w_k) \\ z_{k+1} &= l(x_k, u_k, w_k)\ , \quad k = 0, 1, 2, \ldots \end{cases}$$

where, $x_k \in \boldsymbol{R}^n$ are the states, $y_k \in \boldsymbol{R}^t$ are the measurements, $u_k \in \boldsymbol{U} \subset \boldsymbol{R}^m$ are the controls, $z_k \in \boldsymbol{R}^q$ are the regulated outputs, and $w_k \in \boldsymbol{R}^r$ are the exogenous inputs. Furthermore, we assume that 0 is an equilibrium point of Σ, and $\boldsymbol{U}$ is compact. We denote the set of feasible policies as O, i.e. if $u \in O$, then $u_k = u(y_{1,k}, u_{0,k-1})$. We also assume that f, g, and l are continuous. The output feedback problem is, given $\gamma > 0$, find a control policy $u^* \in O$, so as to ensure that there exists a finite $\beta^{u^*}(x) \geq 0$, $\beta^{u^*}(0) = 0$, such that

$$\sup_{w \in l^2([0,\infty), \boldsymbol{R}^r)} \sup_{x_0 \in \boldsymbol{R}^n} \{p_0(x_0) + \sum_{i=0}^{\infty} \mid z_{i+1} \mid^2 - \gamma^2 \mid w_i \mid^2\} \leq \sup_{x \in \boldsymbol{R}^n} \{p_0(x) + \beta^{u^*}(x)\}.$$

where, $p_0 \in \mathcal{E}$, with $\mathcal{E}$ defined as

$$\mathcal{E} \stackrel{\triangle}{=} \{p \in C(\boldsymbol{R}^n) \mid p(x) \leq R \text{ for some finite } R \geq 0\}$$

We also assume that for such $u \in O$, Σ^u is z-detectable. Furthermore, define the following sup pairing

$$(p, q) \stackrel{\triangle}{=} \sup_{x \in \boldsymbol{R}^n} \{p(x) + q(x)\}$$

An information state based solution was recently obtained in [3]. The information state is defined by the following recursion

$$\begin{aligned} p_{k+1} &= H(p_k, u_k, y_{k+1}),\ k = 0, 1, \ldots \\ p_0 &\in \mathcal{E} \end{aligned}$$

where

$$H(p_k, u_k, y_{k+1})(x) \stackrel{\triangle}{=}$$

$$\sup_{\xi \in \boldsymbol{R}^n} \{p_k(\xi) + \sup_{w \in \boldsymbol{R}^r} (\mid l(\xi, u_k, w) \mid^2 - \gamma^2 \mid w \mid^2 \mid\ x = f(\xi, u_k, w), y_{k+1} = g(\xi, u_k, w))\}.$$

We introduce the function $\delta_x \in \mathcal{E}$, $\delta_x : \boldsymbol{R}^n \to \boldsymbol{R}^*$

$$\delta_x(\xi) \stackrel{\triangle}{=} \begin{cases} 0 & \text{if } \xi = x \\ -\infty & \text{else} \end{cases}$$

The problem is solved via dynamic programming, where the upper value function satisfies

$$M(p) = \inf_{u \in \boldsymbol{U}} \sup_{y \in \boldsymbol{R}^t} \{M(H(p, u, y))\} \tag{1}$$

for all $p \in \mathcal{E}$, with $M(p) \geq (p, 0)$, and $M(\delta_0) = 0$. In particular, $M(p)$ is the least possible worst case cost to go, given $p_0 = p$. We call M, the upper value function of the output feedback game. Then, we have the following result.

Theorem 1 ([3]) *Let $u^* \in O$, be such that $u_k^* = \bar{u}(p_k)$, where $\bar{u}(p_k)$ achieves the minimum in (1) for $p = p_k$. Then $u^* \in O$ is an optimal policy for the output feedback problem.*

Now, assuming $u(p)$, is a non-optimal policy, then there exists a function $W : \mathcal{E} \to \boldsymbol{R}$, $W(p) \geq (p, 0)$, and $W(\delta_x) = 0$, and W satisfies for all $p \in \mathcal{E}$

$$W(p) \geq \sup_{y \in \boldsymbol{R}^t} W(H(p, u(p), y))$$

We call such a W, a storage function for the output feedback policy u.

In the well known, state feedback case, we denote by V the upper value function of the state feedback game. Furthermore, $V \geq 0$, $V(0) = 0$, and V satisfies

$$V(x) = \inf_{u \in \boldsymbol{U}} \sup_{w \in \boldsymbol{R}^n} \{\mid l(x, u, w) \mid^2 - \gamma^2 \mid w \mid^2 + V(f(x, u, w))\}.$$

for all $x \in \boldsymbol{R}^n$. The policy u_F, such that $u_F(x) = u^*$, where $u^* \in U$ achieves the infimum in the above equation is an optimal state feedback policy. For non-optimal state feedback policies u, there exists a $U : \boldsymbol{R}^n \to \boldsymbol{R}$, with $U \geq 0$, $U(0) = 0$, and satisfies

$$U(x) \geq \sup_{w \in \boldsymbol{R}^n} \{\mid l(x, u(x), w) \mid^2 - \gamma^2 \mid w \mid^2 + U(f(x, u(x), w))\}.$$

for all $x \in \boldsymbol{R}^n$. We call such a U a storage function for the state feedback policy u.

¿From now on, we define $\mathcal{I} \subset O$, to be the set of output feedback policies which have the separated structure, i.e. depend only on the information state p_k. We call such policies, information state feedback policies.

3 Reduced Complexity Controllers

The dynamic programming equation (1), is infinite dimensional in general. Hence, this motivates us to search for reduced complexity control policies, which preserve the stability properties of the closed-loop system.

For a given x,$\xi \in \boldsymbol{R}^n$, and $u \in \boldsymbol{U}$, we define

$$\Omega(x, u, \xi) \stackrel{\triangle}{=} \{w \in \boldsymbol{R}^r \mid x = f(\xi, u, w)\}.$$

Then, we have the following result.

Lemma 1 *For any $\xi \in \boldsymbol{R}^n$, $u \in \boldsymbol{U}$, and a given function $h : \boldsymbol{R}^n \times \boldsymbol{R}^r \times \boldsymbol{R}^n \to \boldsymbol{R}$, we have*

$$\sup_{x \in \boldsymbol{R}^n} \sup_{w \in \Omega(x,u,\xi)} h(x, w, \xi) \leq \sup_{w \in \boldsymbol{R}^r} h(f(\xi, u, w), w, \xi)$$

Proof:
For any $\epsilon > 0$, there exists $x^\epsilon \in \boldsymbol{R}^n$, and $w^\epsilon \in \Omega(x^\epsilon, u, \xi)$ (i.e. with $x^\epsilon = f(\xi, u, w^\epsilon)$) such that

$$\begin{aligned} \sup_{x \in \boldsymbol{R}^n} \sup_{w \in \Omega(x,u,\xi)} h(x, w, \xi) &< h(x^\epsilon, w^\epsilon, \xi) + \epsilon \\ &= h(f(\xi, u, w^\epsilon), w^\epsilon, \xi) + \epsilon \\ &\leq \sup_{w \in \boldsymbol{R}^r} h(f(\xi, u, w), w, \xi) + \epsilon \end{aligned}$$

Since, $\epsilon > 0$ is arbitrary, the result follows.

$\square$

Define, $J_U^p : \boldsymbol{R}^n \times \boldsymbol{U} \to \boldsymbol{R}$ as

$$J_U^p(x,u) \stackrel{\triangle}{=} \{p(x) + \sup_{w \in \boldsymbol{R}^r} \{\mid l(x,u,w) \mid^2 - \gamma^2 \mid w \mid^2 + U(f(x,u,w))\}$$

We now state a basic result, which will be used repeatedly.

Lemma 2 *For any $u \in \boldsymbol{U}$, and $U : \boldsymbol{R}^n \to \boldsymbol{R}$, and $p_k \in \mathcal{E}$,*

$$\sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x,u) \geq \sup_{y \in \boldsymbol{R}^t} (H(p_k,u,y),U).$$

Proof:

$$\begin{aligned}
\sup_{y \in \boldsymbol{R}^t} (p_{k+1},U) &= \sup_{y \in \boldsymbol{R}^t} \sup_{x \in \boldsymbol{R}^n} \sup_{\xi \in \boldsymbol{R}^n} \{p_k(\xi) + \sup_{w \in \boldsymbol{R}^r} (\mid l(\xi,u,w) \mid^2 - \gamma^2 \mid w \mid^2 \mid x = f(\xi,u,w), \\
&\qquad\qquad y = g(\xi,u,w)) + U(x)\} \\
&\leq \sup_{x \in \boldsymbol{R}^n} \sup_{\xi \in \boldsymbol{R}^n} \{p_k(\xi) + \sup_{w \in \boldsymbol{R}^r} (\mid l(\xi,u,w) \mid^2 - \gamma^2 \mid w \mid^2 \mid x = f(\xi,u,w)) + U(x)\} \\
&= \sup_{\xi \in \boldsymbol{R}^n} \sup_{x \in \boldsymbol{R}^n} \sup_{w \in \Omega(\xi,u,x)} \{p_k(\xi) + \mid l(\xi,u,w) \mid^2 - \gamma^2 \mid w \mid^2 + U(x)\} \\
&\leq \sup_{\xi \in \boldsymbol{R}^n} \sup_{w \in \boldsymbol{R}^r} \{p_k(\xi) + \mid l(\xi,u,w) \mid^2 - \gamma^2 \mid w \mid^2 + U(f(\xi,u,w))\} \\
&= \sup_{\xi \in \boldsymbol{R}^n} J_U^p(\xi,u)
\end{aligned}$$

$\square$

We now state the main theorem, which gives a sufficient condition for the existence of dissipative reduced complexity policies.

Theorem 2 *Given $U : \boldsymbol{R}^n \to \boldsymbol{R}$, $U \geq 0$, and $U(0) = 0$. If for all $p_k \in \mathcal{E}$*

$$(p_k,U) \geq \inf_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x,u)$$

then $\hat{u}(p_k) \in \arg\min_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x,u)$, solves the output feedback problem, and the associated storage function is $W(p_k) = (p_k,U)$.

Proof:

$$(p_k,U) \quad \geq \quad \inf_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x,u)$$

$$
\begin{aligned}
&= \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x, \hat{u}(p_k)) \\
&\geq \sup_{y \in \boldsymbol{R}^t} (H(p_k, \hat{u}(p_k), y), U)
\end{aligned}
$$

Furthermore, $(p_k, U) \geq (p_k, 0)$, and $(\delta_0, U) = 0$. Hence, (p_k, U) is a storage function, and $\hat{u}$ is a (non-optimal) solution to the output feedback problem.

$\square$

Remark: We could have considered any $\hat{u}_k$ such that

$$
\sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x, u(x)) \geq \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x, \hat{u}_k).
$$

Corollary 1 (Certainty Equivalence) *Given $U \equiv V$, the upper value function of the state feedback game, and the optimal state feedback policy u_F. If for all $p_k \in \mathcal{E}$*

$$
(p_k, V) = \inf_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_V^{p_k}(x, u) \tag{2}
$$

then $u(p_k) = u_F(\hat{x})$, where $\hat{x} \in \arg\max_{x \in \boldsymbol{R}^n} \{p_k(x) + V(x)\}$, is an optimal control policy for the output feedback problem.

Proof:
Clearly (2) implies that

$$
\sup_{x \in \boldsymbol{R}^n} J_V^{p_k}(x, u_F(x)) = \sup_{x \in \boldsymbol{R}^n} \inf_{u \in \boldsymbol{U}} J_V^{p_k}(x, u) = \inf_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_V^{p_k}(x, u)
$$

Hence, a saddle point exists, and so for any $\hat{x} \in \arg\max_{x \in \boldsymbol{R}^n}(p_k(x) + V(x))$, and $\hat{u} = u_F(\hat{x})$,

$$
(p_k, V) = J_V^{p_k}(\hat{x}, \hat{u}) = \sup_{x \in \boldsymbol{R}^n} J_V^{p_k}(x, \hat{u}) \geq \sup_{y \in \boldsymbol{R}^t} (H(p_k, \hat{u}, y), V)
$$

Hence, $W(p_k) = (p_k, V)$ is a storage function, and $W(\delta_x) = V(x)$, the optimal cost of the state feedback game. Hence, the policy is optimal for the output feedback game.

$\square$

Remark: It is sufficient that the conditions in Theorem 2 and Corollary 1 hold only for all p_k, $k = 0, 1, \ldots$. If this is the case, then U need not be a storage function for the state feedback problem. It is only when we need the conditions to hold for $p_k \in \{\delta_x \mid x \in \boldsymbol{R}^n\}$ that U is forced to be a storage function.

In general, conditions for the optimal policy maybe difficult to establish. However, there may exist non-optimal state feedback policies such that their storage functions satisfy the conditions of Theorem 2. In that case, using such non-optimal policies will guarantee that the system is asymptotically stable whenever the exogenous inputs are zero.

We now characterize certainty equivalence in terms of the upper value function of the output feedback game. In [4],[2] it is shown that certainty equivalence holds if, for all $k \geq 0$

$$M(p_k) = (p_k, V) \tag{3}$$

Lemma 3 *Let $\bar{u} \in \mathcal{I}$, with W its storage function. Then*

$$W(p_k) \geq \inf_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x,u), \;\; k = 0, 1, \ldots$$

where, $U(x) \stackrel{\triangle}{=} W(\delta_x)$.

Proof:

$$\begin{aligned}
W(p_k) &\geq \sup_{x \in \boldsymbol{R}^n} \{p_k(x) + \sup_{w \in l^2([0,\infty), \boldsymbol{R}^r)} \sum_{i=k}^{\infty} \mid z_{i+1} \mid^2 - \gamma^2 \mid w_i \mid^2 \mid x_k = x\} \\
&= \sup_{x \in \boldsymbol{R}^n} \{p_k(x) + \sup_{w_k \in \boldsymbol{R}^r} (\mid l(x, \bar{u}(p_k), w_k) \mid^2 - \gamma^2 \mid w_k \mid^2 + \\
&\quad \sum_{i=k+1}^{\infty} \mid z_{i+1} \mid^2 - \gamma^2 \mid w_i \mid^2 \mid \;\; x_{k+1} = f(x, \bar{u}(p_k), w_k))\} \\
&= \sup_{x \in \boldsymbol{R}^n} \{p_k(x) + \sup_{w_k \in \boldsymbol{R}^r} (\mid l(x, \bar{u}(p_k), w_k) \mid^2 - \gamma^2 \mid w_k \mid^2 + \sup_{\xi \in \boldsymbol{R}^n} \{\delta_x(\xi) + \\
&\quad \sum_{i=k+1}^{\infty} \mid z_{i+1} \mid^2 - \gamma^2 \mid w_i \mid^2 \mid \;\; x_{k+1} = f(\xi, \bar{}(p_k), w_k))\} \\
&\geq \inf_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} \{p_k(x) + \sup_{w \in \boldsymbol{R}^r} (\mid l(x, \bar{u}(p_k), w) \mid^2 - \gamma^2 \mid w \mid^2 + U(f(x, \bar{u}(p_k), w)))\} \\
&= \inf_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x,u)
\end{aligned}$$

□

Theorem 3 (Unicity) *Let M be the upper value function of the output feedback game. If there exists a function $U : \boldsymbol{R}^n \to \boldsymbol{R}$, such that $M(p_k) = (p_k, U)$, for all $p_k \in \mathcal{E}$, then $U \equiv V$, the upper value function of the state feedback game.*

Proof:
We have

$$(p_k, U) = M(p_k) \geq \inf_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x,u).$$

Let $\hat{u}(p_k) \in \arg\min_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x,u)$. Then

$$\begin{aligned}
(p_k, U) &\geq \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x, \hat{u}(p_k)) \\
&\geq \sup_{y \in \boldsymbol{R}^t} (H(p_k, \hat{u}(p_k), y), U) \\
&= \sup_{y \in \boldsymbol{R}^t} M(H(p_k, \hat{u}(p_k), y))
\end{aligned}$$

Hence, $\hat{u}$ is an optimal policy since $(p_0, U) = M(p_0), \forall\ p_o \in \mathcal{E}$. Thus,

$$M(p_k) = \sup_{y \in \boldsymbol{R}^t} M(H(p_k, \hat{u}(p_k), y))$$

which implies that

$$(p_k, U) = \inf_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x, u).$$

Setting, $p_k = \delta_x$, we obtain

$$U(x) = \inf_{u \in \boldsymbol{U}} \sup_{w \in \boldsymbol{R}^r} \{\mid l(x, u, w) \mid^2 - \gamma^2 \mid w \mid^2 + U(f(x, u, w))\}.$$

Hence, $U \equiv V$.

$\square$

Corollary 2 *If there exists a p_k such that $M(p_k) \neq (p_k, V)$, then there exists no function $Y : \boldsymbol{R}^n \to \boldsymbol{R}$, such that $M(p) = (p, Y)$ for all $p \in \mathcal{E}$.*

Corollary 3 *Let W be a storage function for an (non-optimal) information state feedback policy $\bar{u} \in \mathcal{I}$, and let $W(p_k) = (p_k, U)$, $k \geq 0$. Then $\hat{u}(p_k) \in \arg\min_{u \in \boldsymbol{U}} \sup_{x \in \boldsymbol{R}^n} J_U^{p_k}(x, u)$ solves the output feedback problem with the storage function $W(p)$. Furthermore, is we insist that $W(\delta_x) = (\delta_x, U)$, $\forall\ x \in \boldsymbol{R}^n$, then U is a storage function for a (non-optimal) state feedback policy. Also, if $W \equiv M$, the upper value function of the output feedback game, then the controller is a certainty equivalence controller.*

Remark: It is clear from the proof of Theorem 3, that if (3) holds, then so does (2). However, (2) is a more tractable condition, since it does not involve the upper value function M, which is what we are trying to avoid having to compute in the first place.

4 Conclusion

In this preliminary paper, we have identified a strategy for generating reduced complexity output feedback policies. Sufficiency conditions have been stated, which guarantee asymptotic stability of the closed-loop system, in the absence of any exogenous inputs ($w \equiv 0$). In the optimal case, it is observed that the controller generated by such strategies reduces to the certainty equivalence controller.

References

[1] T. Basar and P. Bernhard. *H_∞-optimal control and related minimax design problems: A dynamic game approach.* Birkhauser, Boston, 1991.

[2] M.R. James. On the certainty equivalence principle for partially observed dynamic games. *IEEE Transactions on Automatic Control*, 39(11):2321–2324, 1994.

[3] M.R. James and J.S. Baras. Robust H_∞ output feedback control of nonlinear systems. *IEEE Transactions on Automatic Control*, 40(6):1007–1017, 1995.

[4] M.R. James, J.S. Baras, and R.J. Elliott. Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. *IEEE Transactions on Automatic Control*, 39(4):780–792, 1994.

[5] P. Whittle. Risk-sensitive linear/quadratic/gaussian control. *Adv. Appl. Prob.*, 13:764–777, 1981.