Structural Properties of Minimax Policies for a Class of Differential Games Arising in Nonlinear H^{∞}-Control and Filtering *†

Garry Didinsky and Tamer Başar‡
Decision and Control Laboratory
Coordinated Science Laboratory and the
Department of Electrical and Computer Engineering
University of Illinois
1308 West Main Street
Urbana, Illinois 61801 / USA

Pierre Bernhard
INRIA-Sophia Antipolis
2004, Route des Lucioles
06902 Sophia Antipolis Cedex / France

Abstract

This paper introduces, in precise mathematical terms, two properties (named, *certainty equivalence* and *generalized certainty equivalence*) that nonlinear minimax controller problems might possess. The *certainty equivalence* is a generalization of the one introduced earlier in [2] and [3], which applies to problems where the "worst-case disturbance" may not be unique (but the worst-case state trajectory is). The *generalized certainty equivalence*, on the other hand, extends this to accommodate nonunique worst-case state trajectories, and leads to the construction of controllers that guarantee a bounded upper value for the underlying game. The paper also shows that for a large class of games (and under certain conditions) certainty-equivalent (as well as generalized certainty-equivalent) controllers admit (infinite-dimensional) estimator (Kalman-filter) structures, where the estimator gain depends on the state of the estimator. These results are then applied to the nonlinear minimax filtering problem, which is treated here as a special case of the general control problem.

1 Introduction

During the last few years, several authors have obtained results on various nonlinear extensions of the linear H^{∞} theory [3, 4, 5]. These results employed established game-theoretic methods used

*This work was supported in part by the National Science Foundation under Grant ECS91-13153, and in part by the Joint Services Electronics Program under Contract N00014-84-C-0149.

†Paper to be presented at the 32nd CDC, held in San Antonio, Texas, December 1993.

‡All correspondence should be addressed to Prof. Tamer Başar, Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, Illinois 61801, USA; Fax: 217-244-1653; E-mail: tbasar@black.csl.uiuc.edu.

for linear systems, where the H^∞ optimal control problem is treated as a minimax game with a soft-constrained kernel, and the controller and disturbance acting as minimizing and maximizing players, respectively.

It was shown in [2] (and later in [3]) that if there exists a full-state information saddle-point controller, and if a set of truncated optimization problems admit unique solutions, then the measurement feedback game admits a saddle-point controller, which satisfies (and can be computed through) a certainty-equivalence principle.

In [4] and [5], a somewhat different direction was followed. The authors considered games where the system dynamics are affine in control and disturbance, and the costs are quadratic, which implies existence of a saddle-point whenever the upper value is bounded. Furthermore, they take the controller to be a full-state information controller, with the state variable replaced by a state estimator. A byproduct of their investigation is a set of necessary and sufficient conditions for the proposed controller to be locally asymptotically stable.

The current paper contributes to this literature by extending the certainty-equivalence principle, originally developed in [2] and [3], to a more general class of minimax control problems where the unicity of the worst-case disturbance is replaced by that of the corresponding current "worst-case state". To deal with problems where also the worst-case trajectory might be nonunique, the generalized certainty-equivalence property is introduced, in which existence of a "directional" function leads to the construction of controllers that guarantee a bounded upper value for the underlying game. These generalizations come at the expense of having to introduce some additional assumptions on the regularity of the "cost-to-come" function, which however is probably not a very restrictive condition given the other regularity hypotheses one has to make for the analysis to go through.

The paper furthermore shows that for a class of games – a class that includes games with quadratic kernels and affine dynamics – under certain conditions the certainty-equivalent, as well as the generalized certainty-equivalent, controllers admit an estimator (Kalman filter) structure where the gain depends on the state of the estimator. Although this certainty-equivalent controller is infinite dimensional, it corroborates the estimator-controller structure used in [4].

The paper is organized as follows. The problem formulation is given in Section 2. In Section 3, *certainty-equivalence property* is defined, and a sufficient condition is given for a certainty-equivalent controller to be minimax. In Section 4, a *generalized certainty-equivalence property* is defined and sufficient conditions are given for a generalized certainty-equivalent controller to exist. In Section 5, the structure for the estimator (observer) of the certainty-equivalent, as well as the generalized certainty-equivalent, controller is given. Section 6 specializes results of the previous sections to the special case of nonlinear minimax filtering, and Section 7 provides some concluding remarks.

2 Statement of the problem

Consider a minimax control problem with n-dimensional state dynamics described by the following nonlinear differential equation, defined over a time interval $[0, T]$:

$$\dot{x} = f(t, x, u, w), \qquad x(0) = x_0. \tag{2.1}$$

Here $u \in \mathcal{U} \subset I\!R^m$ is a control vector and $w \in \mathcal{W} \subset I\!R^l$ may be thought of as a disturbance. The controller, choosing u, has access to only a noise corrupted output y in $I\!R^p$, given by

$$y = h(t, x) + v \tag{2.2}$$

where $v \in I\!R^p$ is part of the general disturbance, with $v(\cdot)$ belonging to some specified set $\mathcal{V}$ of admissible functions, say square integrable ($\mathcal{L}_2$), or simply measurable. Since the state of the system is not known, we shall also let x_0 be part of the disturbance, and for convenience write

$$(x_0, w(\cdot)) =: \omega \in \Omega = I\!R^n \times \mathcal{L}_2\left([0, T], I\!R^l\right)$$

(Again, $\mathcal{L}_2$ is an arbitrary choice here, and could very well have been replaced by the class of measurable functions.)

To complete the description, we must specify what the admissible *control strategies* (denoted $\mu \in \mathcal{M}$) are. We shall let $\mathcal{M}$ be the set of all causal maps from time functions in $I\!R^p$ to time functions in $I\!R^m$ *such that* the differential equation (2.1) with $\mu_t(y_{[0,t]})$ substituted for u has a solution for every ω in Ω and every admissible $v(\cdot)$. By causality, the value of a strategy μ at time t, μ_t, can be computed as a function of the control history up to that time, in view of which we shall adopt a notation such as $\mu_t(y_{[0,t]}, u_{[0,t)})$. It will be clear from the context that in such cases, it is assumed that $u_{[0,t]} = \mu_{[0,t]}(y_{[0,t]})$.

Next we introduce the performance index for the problem. Suppose that a cost function J is given as in (2.3) below:

$$J(x_0, u(\cdot), v(\cdot), w(\cdot)) = q_T(x(T)) + \int_0^T (q(t,x,u,w) - r(t,v))\, dt - q_0(x_0). \tag{2.3}$$

Assume that the functions q, q_T, q_0, and r, and the earlier defined f and h are assumed to be of class C^1 jointly in all their arguments; we take, without any loss of generality, $\min_v r(t,v) = 0$, for all $t \in [0,T]$. When defined on the controller-disturbance space $\mathcal{M} \times \Omega \times \mathcal{V}$, we will refer to (2.3) again as a cost function, and denote it by $J(\mu, \omega, v)$ by a slight abuse of notation. Then, the objective is to find a *minimax strategy* for the controller, that is a control policy $\mu^* \in \mathcal{M}$ that minimizes the supremum of J over all possible disturbances:

$$\sup_{\substack{\omega\in\Omega \\ v\in\mathcal{V}}} J(\mu^*, \omega, v) = \min_{\mu\in\mathcal{M}} \sup_{\substack{\omega\in\Omega \\ v\in\mathcal{V}}} J(\mu, \omega, v). \tag{2.4}$$

3 A Certainty Equivalence Principle

We will start this section by considering games for which a property, which may be called *certainty equivalence*, holds. Before we define what we mean exactly by certainty equivalence, let us introduce the cost-to-go (or upper-value) function, $V(\cdot,\cdot) : I\!R \times I\!R^n \longrightarrow R$, associated with the full-state information game.

Assumption 3.1. The game defined in section 2, but with full state measurement, has an upper value V for every initial time and state, which is C^1 in these variables. It is achieved by a state feedback controller $u(t) = \mu_t^F(x(t))$. □

Under Assumption 3.1, V satisfies Isaacs' equation [1]:

$$-\frac{\partial V}{\partial t} = \inf_u \sup_w \left\{ \frac{\partial V}{\partial x} f(t,x,u,w) + q(t,x,u,w) \right\}; \quad V(T,x) = q_T(x), \tag{3.1}$$

and further, the RHS of (3.1) becomes equivalent to

$$\sup_w \left\{ \frac{\partial V}{\partial x} f(t,x,\mu_t^F(x),w) + q(t,x,\mu_t^F(x),w) \right\}.$$

(Note that we do not need to impose the Isaacs' condition, i.e. that the *inf-sup* above be a saddle point. This is why V here is an *upper value*, and not necessarily a value function.)

Let us also define the truncated kernels,

$$\begin{aligned}&L^s(u_{[0,s]}, y_{[0,s]}, \omega_{[0,s]}) = \\ &\quad V(s, x(s)) + \int_0^s \{q\,(t, x, u, w) - r(t, y - h(t, x))\}\, dt - q_o\,(x_o)\,,\end{aligned} \tag{3.2}$$

and introduce the notation

$$L^s_\mu(y_{[0,s]}, \omega_{[0,s]}) := L^s(\mu_{[0,s]}(y_{[0,s]}), y_{[0,s]}, \omega_{[0,s]})$$

Note that in (3.2), we have substituted the expression $y - h(t, x)$ for v in r. That way, we readily restrict the perturbations to those that are consistent with the past control and measurement histories.

We next make the definition of a “certainty-equivalent” controller precise.

Definition 3.1. *Let $\hat{\mu} \in \mathcal{M}$ be a controller generated by the relationship*

$$\hat{\mu}_t(y_{[0,t]}, u_{[0,t)}) = \mu^F_t(\hat{x}(t)), \quad t \in [0, T], \tag{3.3}$$

where $\hat{x}(t)$ is the state of the system corresponding to a worst-case disturbance $\omega_{[0,t]}$ which is obtained by maximizing $L^t_{\hat{\mu}}(y_{[0,t]}, \omega_{[0,t]})$, i.e.

$$\hat{\omega}_{[0,t]} := \arg\max_{\omega\in\Omega} L^t_{\hat{\mu}}(y_{[0,t]}, \omega_{[0,t]}) \tag{3.4}$$

If the maximum in (3.4) is achieved (as stipulated) for each frozen $t \in [0, T]$, then the controller $\hat{\mu}$ is called a certainty-equivalent controller *(or* policy*).* □

Remark 3.1: Note that the procedure underlying Definition 3.1 involves a recursive construction in forward time. It will lead to “certainty-equivalent” controllers which are generally not finite dimensional, because the worst-case state may not be obtained as the solution to an ordinary differential equation. □

Definition 3.2. *If there exists a certainty-equivalent controller $\hat{\mu}$ which is also* **minimax***, then we say that* certainty equivalence holds *for the underlying game.* □

Remark 3.2: In the case when the full-state information problem has a saddle point and $\hat{\omega}_{[0,t]}$ is unique for every time instant t and measurement output $y_{[0,T]}$, it has been shown in [2] and [3] that certainty-equivalence property (as defined here) holds. □

It is useful at this point to recall a result from [2], [3] :

Theorem 3.0. *If for every* $(u_{[0,T]}, y_{[0,T]})$ *there exists a* t *such that the truncated kernel* $L^t(u_{[0,T]}, y_{[0,T]}, \omega)$ *has an infinite supremum in* ω*, then there is no controller* $\mu \in \mathcal{M}$ *that keeps the criterion* $\sup_{\omega,v} J(\mu, \omega, v)$ *finite.* □

In the view of this theorem, we shall henceforth consider only the cases where there exists a $\mu \in \mathcal{M}$ such that this supremum is finite.

Before we present a set of less stringent (than before) sufficient conditions for the certainty-equivalence property to hold, we first introduce the *"cost-to-come" function*: Given an admissible controller policy $\mu_{[0,T]}$ and an output measurement history $y_{[0,T]}$, generating a well-defined $u_{[0,T]} = \mu(y_{[0,T]})$, and given a point $\xi \in I\!R^n$, we denote by $\Omega^s_\mu(y_{[0,T]}, \xi)$ the set of restrictions $\omega_{[0,s]}$ of ω such that together with that particular $u_{[0,T]}$, it generates a trajectory with terminal value $x(s) = \xi$. Although it will be convenient to write it as we did here, it should be clear that Ω^s_μ depends only on the restrictions to $[0, s]$ of both μ and y, or simply of u. Then, the *"cost-to-come" function* is defined as:

$$W_\mu(s, \xi; y_{[0,T]}) := \sup_{\omega_{[o,s]} \in \Omega^s_\mu(y_{[0,T]}, \xi)} \left\{ \int_0^s \left(q\left(t, x, u, w\right) - r(t, y - h(t, x)) \right) dt - q_o\left(x_o\right) \right\} \tag{3.5}$$

where $u(t)$ is the control action dictated by strategy μ at time t, given $y_{[0,T]}$. As with Ω^s_μ, the *cost-to-come function* $W_\mu(s, \xi, y_{[0,T]})$ actually depends only on the restriction $y_{[0,s]}$ of y.

Assumption 3.2. Given a certainty-equivalent controller $\hat{\mu}$, $W_{\hat{\mu}}$ is bounded and is C^1 (jointly continuously differentiable in t and x). □

Under Assumption 3.2, $W_{\hat{\mu}}$ satisfies the following partial differential equation, where u stands for $\hat{\mu}_t(y_{[0,T]})$, and $\partial W_{\hat{\mu}}/\partial t$ and $\partial W_{\hat{\mu}}/\partial x$ stand for partial derivatives of $W_{\hat{\mu}}$ with respect to its first and second arguments respectively:

$$\begin{aligned} \frac{\partial W_{\hat{\mu}}}{\partial t} &= \sup_w \left\{ -\frac{\partial W_{\hat{\mu}}}{\partial x} f(t, x, u, w) + q(t, x, u, w) \right\} \\ &\quad -r(t, y - h(t, x)) \\ W_{\hat{\mu}}(0, x) &= -q_o(x) \end{aligned} \tag{3.6}$$

Assumption 3.3. The set of states $x(t)$ that are consistent with certainty-equivalent strategies is all of $I\!R^n$.[1] □

[1]This is plausible because all initial states are allowed in Ω. In fact, it suffices for our purposes that it be an open set X. Then the $\max_{x \in I\!R^n}$ operators in the sequel would be replaced by $\max_{x \in X}$. But we see no point in pursuing this level of generality.

Under Assumption 3.3, the certainty-equivalence property implies that the state generated by a "certainty-equivalent" controller satisfies

$$\hat{x}(t) \in \arg\max_{x \in I\!R^n} \left\{ V(t,x) + W_{\hat{\mu}}(t,x;y_{[0,t]}) \right\} =: \hat{X}_{\hat{\mu}}(t, y_{[0,t]}). \tag{3.7}$$

We finally introduce the following crucial assumption:

Assumption 3.4. The sets $\hat{X}_{\hat{\mu}}(t, y_{[0,t]})$, as defined by (3.7), are singletons for every $t \in [0,T]$ and $y_{[0,t]} \in \mathcal{L}_2([0,t], I\!R^p)$. □

The main theorem of this section, which provides a set of sufficient conditions for the certainty-equivalence principle to hold, is stated next:

Theorem 3.1. *Let a certainty-equivalent controller $\hat{\mu}$ exist, leading to satisfaction of Assumptions 3.2-3.4. Further, let Assumption 3.1 be satisfied. Then, certainty equivalence holds.*

Proof: Given an output measurement $y_{[0,T]} \in \mathcal{L}_2[0,T]$, define

$$G(t) := \max_x \left\{ V(t,x) + W_{\hat{\mu}}(t,x;y_{[0,T]}) \right\}. \tag{3.8}$$

Danskin's theorem for differentiation [6] implies the existence of a time derivative:

$$\dot{G}(t) := \frac{dG}{dt} = \frac{\partial V}{\partial t}(t,\hat{x}) + \frac{\partial W_{\hat{\mu}}}{\partial t}(t,\hat{x};y_{[0,T]}), \tag{3.9}$$

and hence, the first-order condition for maximization in (3.8) reads

$$\frac{\partial V}{\partial x}(t,\hat{x}) + \frac{\partial W_{\hat{\mu}}}{\partial x}(t,\hat{x};y_{[0,T]}) = 0. \tag{3.10}$$

Substituting this into (3.6), and using (3.3) to replace $\hat{\mu}$ by $\mu^F(\hat{x}(t))$, and placing the resulting expression for $\partial W_{\hat{\mu}}$ in (3.9) above, leads to

$$\begin{aligned} \dot{G}(t) &= \frac{\partial V}{\partial t}(t,\hat{x}) + \sup_w \left\{ \frac{\partial V}{\partial x}(t,\hat{x}) f(t,\hat{x},\mu^F(\hat{x}),w) \right. \\ &\qquad \left. + q(t,\hat{x},\mu^F(\hat{x}),w) \right\} - r(t, y - h(t,\hat{x})) \\ &= -r(t, y(t) - h(t,\hat{x})) \le 0 \end{aligned} \tag{3.11}$$

Therefore

$$G(T) \le G(0) = \max_x \{ V(0,x) - q_o(x) \} \tag{3.12}$$

where the right-hand side of the inequality is bounded by our earlier assumption on $\hat{X}_{\hat{\mu}}(0)$. But notice that since $V(T,x) = q_T(x)$, for all $y_{[0,T]}$,

$$G(T) := \sup_{\omega\in\Omega} J(\hat{\mu}(y_{[0,T]}), \omega, y - h).$$

Also, the quantity on the right-hand side of (3.12) is the value of the corresponding full-state information game, and hence, the certainty-equivalent controller is minimax. □

Remark 3.3: It might be useful to point out where the unicity of $\hat{x}$ is crucial in the proof above. If $\hat{x}$ were not unique, then the expression for the directional time derivative of G would still be in the same form, but with an operator $\max_{\hat{x}\in\hat{X}}$ in front of it, which applies only to those $\hat{x}$'s which do not appear as arguments of μ^F. Then, it might happen that the $\hat{x}$ in $\hat{X}$ for which this maximum is achieved would not coincide with the one used as argument of μ^F. Consequently, we cannot use Isaacs' equation (3.1) to complete the proof above. □

Remark 3.4: It should be pointed out that if $\mu^F(x)$ is the unique minimum in (3.1), then under Assumption 3.4, the certainty-equivalent controller is unique. □

4 Suboptimal Controllers

Ideas underlying Theorem 3.1 can be extended to yield less restrictive conditions for the existence of a minimax controller. One can think of the cost-to-go function, $V(t,x)$, as a "direction" function, i.e., as an "estimated" worst future value of the game that helps to orient the controller in the direction that best counteracts future worst-case trends. With that in mind, we can generalize the idea of cost-to-go, or that of the "directional" function, to a "bounding" function $U(t,x)$ that satisfies, for some nonnegative function $p : [0,T] \times I\!\!R^n \longrightarrow I\!\!R$, C^1 in its second argument, the partial differential equation:

$$-\frac{\partial U}{\partial t} = \min_u \sup_w \left\{ \frac{\partial U}{\partial x} f(t,x,u,w) + q(t,x,u,w) \right\} + p(t,x), \tag{4.1}$$

with a boundary condition $U(T,x) \geq q_T(x)$. Note that this is in fact the Isaacs' equation for a new full-state information game, where the incremental cost has been increased form q to $q+p$, and terminal cost from $q_T(x)$ to $U(T,x)$.

This interpretation immediately leads to the bound $U(t,x) \geq V(t,x)$, for all t and x. Let the full-state minimax controller in (4.1) (assuming that it exists) be denoted again by $\mu_t^F(x)$, which now clearly depends on the choice of p.

Along with this new cost-to-go function, let us redefine the truncated kernels (3.2) as

$$\begin{array}{l} L^s(u_{[0,s]}, y_{[0,s]}, \omega_{[0,s]}) = \\ U(s, x(s)) + \int_0^s \{q(t, x, u, w) - r(t, y - h(t, x))\}\, dt - q_o(x_o)\,, \end{array} \tag{4.2}$$

and, in accordance with it, introduce the notation

$$L^s_\mu(y_{[0,s]}, \omega_{[0,s]}) := L^s(\mu_{[0,s]}(y_{[0,s]}), y_{[0,s]}, \omega_{[0,s]})$$

Let us further introduce a definition and some assumptions paralleling those introduced in the previous section:

Definition 4.1. *Let $\hat{\mu} \in \mathcal{M}$ be a controller defined as in Definition 3.1, with only L^t replaced by its expression given by (4.2). Then, it is called a generalized certainty-equivalent controller (or policy). If furthermore it guarantees a bounded upper value for the underlying game, we say that generalized certainty equivalence holds.* □

Assumption 4.1. Given a generalized certainty-equivalent controller $\hat{\mu}$, $W_{\hat{\mu}}$ is bounded and is C^1 (jointly continuously differentiable in t and x). □

Under Assumption 4.1, $W_{\hat{\mu}}$ satisfies the partial differential equation

$$\begin{array}{rcl} \dfrac{\partial W_{\hat{\mu}}}{\partial t} & = & \sup\limits_w \left\{ -\dfrac{\partial W_{\hat{\mu}}}{\partial x} f(t, x, u, w) + q(t, x, u, w) \right\} \\ & & -r(t, y - h(t, x)) \\ W_{\hat{\mu}}(0, x) & = & -q_o(x) \end{array} \tag{4.3}$$

Assumption 4.2. The set of states $x(t)$ that are consistent with generalized certainty-equivalent strategies is all of $I\!R^n$. □

Under Assumption 4.2, the generalized certainty-equivalence property implies that the state generated by a "generalized certainty-equivalent" controller satisfies

$$\hat{x}(t) \in \arg\max_{x \in I\!R^n} \left\{ U(t, x) + W_{\hat{\mu}}(t, x; y_{[0,t]}) \right\} =: \hat{X}_{\hat{\mu}}(t, y_{[0,t]}). \tag{4.4}$$

This is similar to the corresponding result obtained in Section 3, but here $\hat{x}$ does not have to be unique.

We now present the following theorem, which is the counterpart of Theorem 3.1.

Theorem 4.1. *Let a generalized certainty-equivalent controller $\hat{\mu}$ exist, leading to satisfaction of Assumptions 4.1-4.2. Then, generalized certainty equivalence holds if there exists a mapping $\hat{x} : [0, T] \times \mathcal{L}_2[0, T] \longrightarrow I\!R^n$, such that*

1. $\hat{x}(t, y_{[0,t]}) \in \hat{X}_{\hat{\mu}}(t, y_{[0,t]})$,

2. for all $x \in \hat{X}_{\hat{\mu}}(t, y_{[0,t]})$,

$$\frac{\partial U}{\partial t}(t,x) + \sup_w \left\{ \frac{\partial U}{\partial x}(t,x) f(t,x,\mu_t^F(\hat{x}),w) + q(t,x,\mu_t^F(\hat{x}),w) \right\} \leq 0. \tag{4.5}$$

Moreover, the controller $\mu_t^F(\hat{x})$ *guarantees the bounded upper value*

$$\max_x \{U(0,x) - q_o(x)\}. \tag{4.6}$$

Proof: The proof is similar to that of Theorem 3.1, except that V is now replaced by U. (See also Remark 3.3.) Also, note that the upper bound on the value of the game, $\max_x \{U(0,x) - q_o(x)\}$, is finite by hypothesis 1 of the theorem. □

Remark 4.1: A sufficient condition for the generalized certainty-equivalent controller to be a minimax controller is for the upper bound (4.6) to equal the upper value of the original game. □

5 Minimax and suboptimal observers

In the previous two sections, certainty-equivalence or generalized certainty-equivalence properties have enabled us to express minimax or suboptimal controllers as functions of the worst-case state $\hat{x}(t)$. We will now turn our attention to computing the worst-case trajectory in a recursive way.

We shall need the following assumptions, where we work with U (introduced in section 4) instead of V which is the true full-state information cost-to-go function, since the former captures situations (as described in section 4) where the latter would not be applicable.

Assumption 5.1. $\hat{x}$ is continuously differentiable in t, for all $y_{[0,T]} \in \mathcal{L}_2[0,T]$. □

Assumption 5.2. $U(t,x)$ and $W_{\hat{\mu}}(t,x;y_{[0,T]})$ are twice jointly continuously differentiable in t and x, for all $y_{[0,T]} \in \mathcal{L}_2[0,T]$. □

Assumption 5.3. Given any t and $y_{[0,T]}$, the following inequality holds:

$$[U + W_{\hat{\mu}}]_{xx}(t, \hat{x}(t)) < 0. \tag{5.7}$$

□

Assumption 5.4. The arguments $\mu^F(x)$ and $\nu^F(x)$ of the minimax in (4.1) (resp. (3.1)) and the argument $\hat{w}$ of the maximum in (4.3) (resp. (3.6)) are unique. □

The first-order necessary condition for $\hat{x}(t)$ to be optimal is

$$[U + W_{\hat{\mu}}]_x\,(t, \hat{x}(t)) = 0, \quad \forall t \in [0, T]. \tag{5.8}$$

Thanks to Assumptions 5.1 and 5.2, we may apply the implicit function theorem to differentiate $\hat{x}$:

$$\dot{\hat{x}} = -\,[U + W_{\hat{\mu}}]_{xx}^{-1}\,[U + W_{\hat{\mu}}]'_{xt}\,(t, \hat{x})\,. \tag{5.9}$$

To find $[U + W_{\hat{\mu}}]_{xt}\,(t, \hat{x})$, we will differentiate (4.1) and (4.3) with respect to x. Using Danskin's theorem, thanks to Assumption 5.4, differentiation leads to

$$-\,\frac{\partial^2 U}{\partial x \partial t} \;=\; f'\frac{\partial^2 U}{\partial x^2} + \frac{\partial U}{\partial x}\frac{\partial f}{\partial x} + \frac{\partial q}{\partial x} + \frac{\partial p}{\partial x}, \tag{5.10}$$

$$\frac{\partial^2 W_{\hat{\mu}}}{\partial x \partial t} \;=\; -f'\frac{\partial^2 W_{\hat{\mu}}}{\partial x^2} - \frac{\partial W_{\hat{\mu}}}{\partial x}\frac{\partial f}{\partial x} + \frac{\partial q}{\partial x} - \frac{\partial r}{\partial v}\frac{\partial h}{\partial x}. \tag{5.11}$$

In the first equation above, all functions are evaluated at $(t, x, \mu_t^F(x), \nu_t^F(x))$, while in the second, the arguments are $(t, x, \mu_t^F(\hat{x}(t)), \hat{w})$. Now notice that at $\hat{x}$, $(W_{\hat{\mu}})_x = -U_x$, so that $\hat{w}$ coincides with ν^F. As a result we may subtract the former from the latter, and evaluate at $\hat{x}(t)$, to get

$$\begin{aligned}[U + W_{\hat{\mu}}]'_{xt}\,(t, \hat{x}) = -\,[U + W_{\hat{\mu}}]_{xx}\,(t, \hat{x})\, f(t, \hat{x}, \mu_t^F(\hat{x}), \nu_t^F(\hat{x}))) \\ + \left(\tfrac{\partial h}{\partial x}\right)' \left(\tfrac{\partial r}{\partial v}\right)' - \left(\tfrac{\partial p}{\partial x}\right)',\end{aligned} \tag{5.12}$$

which together with (5.9) leads to the (observer) equation:

$$\begin{aligned}&\dot{\hat{x}} = f(t, \hat{x}, \mu_t^F(\hat{x}), \nu_t^F(\hat{x})) \\ &-[U + W_{\hat{\mu}}]_{xx}^{-1}\,(t, \hat{x}) \left[\left(\tfrac{\partial h}{\partial x}\right)' \left(\tfrac{\partial r}{\partial v}(t, y - h(t, \hat{x}))\right)' - \left(\tfrac{\partial p}{\partial x}\right)' \right].\end{aligned} \tag{5.13}$$

Theorem 5.1. *If generalized certainty equivalence and Assumptions 5.1-5.4 hold, then the worst-case state $\hat{x}(t)$ used in the generalized certainty-equivalent controller of Theorem 4.1 is generated by equation (5.13) above.* □

Remark 5.1: Quite naturally, there is a similar theorem for the certainty-equivalent controller of Theorem 3.1, which is now generated by equation (5.13) with U replaced by V and p identically zero. □

Remark 5.2: If Assumption 5.1 does not hold, but the worst-case trajectory is piecewise continuously differentiable, then a natural counterpart of Theorem 5.1 holds with (5.13) valid on subintervals of $[0, T]$ where $\hat{x}$ is continuously differentiable. □

Remark 5.3: It should be noted that controller (5.13) has the estimator (Kalman-filter) form, which is already known to be optimal for linear H^∞ problems [7]. Unlike the linear case, here the controller is infinite dimensional, because $[U + W_{\hat{\mu}}]_{xx}(t, \hat{x})$ cannot, to the best of our knowledge, be obtained as the solution of a finite dimensional differential equation; nor can it be precomputed off line as in the linear-quadratic case.

Still, an immediate significance of this result is that, under appropriate regularity hypotheses, the minimax controller lies in the class of controllers of the form

$$\begin{cases} u_t = \mu_t(y_{[0,T]}) = \mu_t^F(\hat{x}_t), \\ \dot{\hat{x}} = f(t, \hat{x}, \mu_t^F(\hat{x}), \nu_t^F(\hat{x})) \\ \qquad -M(t, \hat{x}) \left[\left(\frac{\partial h}{\partial x} \right)' \left(\frac{\partial r}{\partial v}(t, y - h(t, \hat{x})) \right)' - m(t, \hat{x}) \right], \end{cases} \tag{5.14}$$

where M and m are appropriate dimensional functions. Then, alternative search methods could be employed to find "good" finite dimensional controllers from this restricted class of controllers (see for example [4]). □

6 Minimax Filtering

Consider now a minimax filtering problem with n-dimensional state dynamics described by the following nonlinear differential equation, defined over a time interval $[0, T]$:

$$\dot{x} = f(t, x, w), \qquad x(0) = x_0, \tag{6.1}$$

Here, just as in the control problem formulation, $w \in \mathcal{W} \subset I\!R^l$ and $x_0 \in I\!R^n$ may be thought of as disturbances, and for convenience, we may write

$$(x_0, w(\cdot)) =: \omega \in \Omega = I\!R^n \times \mathcal{L}_2\left([0, T], I\!R^l\right).$$

There is a noise corrupted output y in $I\!R^p$, given by

$$y = h(t, x) + v. \tag{6.2}$$

As in Section 2, here $v \in I\!R^p$ is part of the general disturbance, with $v(\cdot)$ belonging to some specified set $\mathcal{V}$ of admissible functions, say square integrable ($\mathcal{L}_2$), or simply measurable.

A minimax filtering policy will be chosen from the set of all admissible filtering strategies, denoted $\mathcal{D}$, which is the set of all causal maps from time functions $(y_{[0,T]})$ in $I\!R^p$ to time functions in $I\!R^n$, according to cost (error) function J, given in (6.3) below:

$$J(x_0, \bar{x}(\cdot), v(\cdot), w(\cdot)) = \int_0^T [q(t, x - \bar{x}, w) - r(t, v)]\, dt - q_0(x_0). \tag{6.3}$$

Here, $\bar{x} \in \bar{X} \subset I\!R^n$ is a filtering vector, which is determined by an admissible filtering policy $\delta \in \mathcal{D}$. A strategy $\delta^* \in \mathcal{D}$ will be called a *minimax strategy* for the filter, if it satisfies

$$\sup_{\substack{\omega \in \Omega \\ v \in \mathcal{V}}} J(\delta^*, \omega, v) = \min_{\delta \in \mathcal{D}} \sup_{\substack{\omega \in \Omega \\ v \in \mathcal{V}}} J(\delta, \omega, v). \tag{6.4}$$

In addition to all the earlier hypotheses that q, q_0, r, f and h are in class C^1 jointly in all their arguments, and that $\min_v r(t, v) = 0$, for all $t \in [0, T]$, it is assumed here that $q(t, \cdot, 0)$ and $-q(t, 0, \cdot)$ are positive definite for all $t \in [0, T]$. Here, by saying $q(\cdot)$ is positive definite, we mean $q(x) \geq 0$ with equality occurring only when $x = 0$. Hence, $q(t, \cdot, w)$ could be a norm for all t and w, designating for example a pointwise estimation error.

It is easy to see that the minimax filtering problem can be considered as a special case of the minimax control problem of Section 2. Under the set up of the filtering problem, the cost-to-go function, V, can be easily verified to be identically zero and $\delta_t^F(x) = x$ is a FSI minimax filtering policy. This then takes us to the following definition of a "certainty-equivalence" filter.

Definition 6.1. *Let $\hat{\delta} \in \mathcal{D}$ be a filter generated by the relationship*

$$\hat{\delta}_t(y_{[0,t]}, \bar{x}_{[0,t)}) = \hat{x}(t), \quad t \in [0, T], \tag{6.5}$$

where $\hat{x}(t)$ is a worst-case state of the system, i.e.,

$$\hat{x}(t) \in \arg\max_x W_{\hat{\delta}}(t, x; y_{[0,T]}), \tag{6.6}$$

where $W_{\hat{\delta}}$ is the cost-to-come function associated with policy $\hat{\mu}$. If the maximum above is achieved (as stipulated) for each frozen $t \in [0, T]$, then the filter $\hat{\mu}$ is called a certainty-equivalent filter *(or* policy*).* □

Definition 6.2. *If there exists a certainty-equivalent filter $\hat{\delta}$ which is also* **minimax**, *then we say that* certainty equivalence holds *for the underlying game.* □

Let us recall that, if Assumption 3.2 holds, $W_{\hat{\delta}}$ satisfies the following partial differential equation:

$$\begin{aligned} \frac{\partial W_{\hat{\delta}}}{\partial t} &= \sup_w \left\{ -\frac{\partial W_{\hat{\delta}}}{\partial x} f(t,x,w) + q(t, x-\bar{x}, w) \right\} \\ &\quad -r(t, y-h(t,x)) \\ W_{\hat{\mu}}(0,x) &= -q_o(x) \end{aligned} \tag{6.7}$$

where $\bar{x}$ stands for $\hat{\delta}_t(y_{[0,T]})$.

The results of Sections 3 and 5 are then directly applicable here for the filtering problem, and this leads us to the following theorem:

Theorem 6.1. *Let a certainty-equivalent filter $\hat{\delta}$ exist, leading to satisfaction of Assumptions 3.2 and 3.4. Then, certainty equivalence holds for the filtering problem. Moreover, if Assumptions 5.1-5.4 hold with $U \equiv 0$, then the worst-case state $\hat{x}(t)$ which characterizes the certainty-equivalent filter is generated by*

$$\begin{aligned} \dot{\hat{x}} &= f(t,\hat{x},0) \\ &\quad - \left[W_{\hat{\delta}}\right]_{xx}^{-1}(t,\hat{x}) \left[\left(\tfrac{\partial h}{\partial x}\right)' \left(\tfrac{\partial r}{\partial v}(t, y-h(t,\hat{x}))\right)' + \tfrac{\partial q}{\partial x}(t,0,0) \right] \end{aligned} \tag{6.8}$$

with $\hat{x}(0) = \max_x(-q_o(x))$. □

Remark 6.1: It should be noted that the filter given above in (6.8) has the familiar form of an extended Kalman filter [8]. □

Remark 6.2: For the LQ problem (more precisely, the H^∞-filtering problem), the filter given above in (6.8) becomes identical to that given in [9]. However, in the nonlinear case, (6.8) represents an infinite dimensional filter, because $\left[W_{\hat{\delta}}\right]_{xx}$ cannot, to the best of our knowledge, be obtained as the solution of a finite dimensional differential equation. □

7 Conclusion

This paper has introduced, in precise mathematical terms, two properties – certainty equivalence and generalized certainty equivalence – that differential games encountered in nonlinear H^∞-control and H^∞-filtering may possess. The former allows the designer to choose a minimax policy to be the optimal full-state information policy with the state vector replaced by an appropriate worst-case state $\hat{x}$. The latter generalizes the certainty-equivalence property by allowing a suboptimal full-state information controller to be a function of a more conservative (but more readily computable) cost-to-go function $U(t, x)$, and hence provides a more general set of sufficient conditions for the existence of a suboptimal controller that guarantees bounded upper value of the game.

The paper has also shown that under appropriate regularity assumptions, the (generalized) certainty-equivalent controller or filter has the familiar estimator (Kalman-filter) structure. Unlike the linear case, the nonlinear policy that comes out of this structure is infinite-dimensional. This possible drawback may be overcome by use of alternate methods of search among finite-dimensional policies possessing the same estimator structure, but most possibly at the expense of loss of performance.

This paper has only scratched the surface as many more questions remain to be answered. First, it is still not clear how "optimal" a (generalized) certainty-equivalent policy is in the context of the original nonlinear problem. Second, it is not known whether, given any policy, it is possible to find a generalized certainty-equivalent controller that will yield no worse performance.

Finally, not much is known on the structure of the minimax policies if the "worst state" is not unique, in either controller or the filtering problem.

References

[1] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory.* London/New York: Academic Press, 1982. Second printing 1989.

[2] P. Bernhard, "A certainty equivalence principle and its applications to continuous-time, sampled data, and discrete time H^∞ optimal control." INRIA Report #1347, August 1990.

[3] T. Başar and P. Bernhard, *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach.* Boston, MA: Birkhäuser, 1991.

[4] A. Isidori and A. Astolfi, "Disturbance attenuation and H_∞- control via measurement feedback in nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 37, pp. 1283–1293, September 1992.

[5] A. J. Van der Schaft, "A state-space approach to nonlinear H_∞ control," *Syst. Contr. Lett.*, vol. 16, pp. 1–8, 1991.

[6] J. M. Danskin, *The Theory of Max Min.* Berlin: Springer, 1967.

[7] J. Doyle, K. Glover, P. Khargonekar, and B. Francis, "State-space solutions to standard H_2 and H_∞ control problems," vol. AC- 34, no. 8, pp. 831–847, 1989.

[8] A. H. Jazwinski, *Stochastic Processes and Filtering Theory.* New York: Academic, 1970.

[9] P. P. Khargonekar and K. N. Nagpal and K. R. Poolla, "H^∞ control with Transients," *SIAM Journal on Control and Optimization*, vol. 29, no. 6, pp. 1373–1393, 1991.