# Certainty Equivalence Principle and Minimax Team Problems

Pierre BERNHARD [*], Naira HOVAKIMYAN [†]

August 18, 1998

**Abstract**

We consider a team problem, with two decision makers for simplicity, where the uncertainties are dealt with in a minimax fashion rather than in a stochastic framework. We do not assume that the players exchange information at any time. Thus new ideas are necessary to investigate that situation. In contrast with the classical literature we do not use necessary conditions, but investigate to what extent ideas comming from the (nonlinear) minimax certainty equivalence theory allow one to conclude here. We are led to the introduction of a "partial team" problem, where one of the decision makers has perfect state information. We then investigate the full team problem, but the main result concerning it is shown to be still rather weak. We nevertheless apply it to the linear quadratic case, where it yields an original result.

## 1 Introduction

The origin of this research was in non linear robust control, as explained in [4]. There we show that the natural nonlinear equivalent to the (linear) $\mathcal{H}_\infty$-optimal control approach is *not* the so-called "non linear $\mathcal{H}_\infty$" control problem, but a minimax team problem. By this we mean a team problem where the uncertainties are dealt with in a minimax fashion, looking for a guaranteed outcome, rather than in a stochastic fashion, looking for a mean outcome.

As far as we know, explicit results published for stochastic dynamic team problems all involve either a form of imbeddedness of the informations available to the players, as in [10, 7], or other very special features (see [1] for

[*]University of Nice-Sophia Antipolis/I3S and INRIA-Sophia Antipolis, France
[†]Institute of Mechanics, National Academy of Sciences of Armenia

instance). In the case of minimax team problems, often investigated under the umbrella of Nash equilibria, the results we are aware of almost always assume a dicrete time model with a one step delay information sharing pattern, see e.g. [8, 14], except for some asymptotic results for weak coupling, see [15].

Here we attempt to deal with non imbedded information, without any information sharing. While the result obtained is rather weak, as stressed by the section dealing with an abstract analysis of that result, yet it allows us to give an original result for the continuous time linear quadratic problem. It would clearly be a simple matter to derive the equivalent result for the discrete time equivalent.

As a matter of fact, our basic derivation applies to a nonlinear setting, since it uses the basically non linear minimax certainty equivalence principle. But practically there is little hope to get tractable results other than to the LQ problem, or to some simple low dimensional nonlinear problems as was done with the basic principle.

# 2 The system considered

## 2.1 Dynamics and cost-function

Consider a team of two decision makers, whom we call *players* for short, each controlling different actions and having access to different informations. There is a common pay-off for both players, which has to be minimized by them. (In [4] the case is considered, when their dynamics are separated with respect to all variables.) To be more precise let $x(t)$ be the state variable of the system and $u_1, u_2, w$ denote correspondingly the control variables for each of the players and the disturbance, in terms of which the dynamic equation of system's evolution over time in the nonlinear general setup can be presented as:

$$\dot{x} = f(t, x, u_1, u_2, w)\,, \qquad x(t_0) = x^0\,, \tag{1}$$

where $t \in [t_0, +\infty), x(t) \in \mathbb{R}^n, u_i \in \mathbb{R}^{m_i}, i = 1, 2, w \in \mathbb{R}^l$. The disturbance variable $w$ is dealt with by considering the "guaranteed performance level", leading to a formulation in terms of a minimax or dynamic game problem.

The controls of the players and the disturbance obey the following restrictions:

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in \mathsf{U} := \mathsf{U}_1 \times \mathsf{U}_2, \quad w \in \mathsf{W},$$

where the $\mathsf{U}_1$, $\mathsf{U}_2$, $\mathsf{W}$ are compact convex sets in appropriate spaces. The sets $\mathcal{U}_i$ and $\mathcal{W}$ of admissible open-loop controls $u_i(\cdot)$ and $w(\cdot)$ will contain all measurable functions from $[t_0, \infty)$ into $\mathsf{U}_i$ and $\mathsf{W}$ respectively.

Under the necessary regularity assumptions (specified below) we shall denote for a given initial time $t_0 \in \mathbb{R}$ by $x(.) = S(t_0, x^0, u_1(.), u_2(.), w(.))$ the unique (Cauchy) solution of the system (1). (The first argument in $S$ will often be omitted.)

We shall consider the following performance index:

$$J = M(x(T)) + \int_{t_0}^{T} L(t, x, u_1, u_2, w)\, dt + N(x^0)\,, \tag{2}$$

where $L, M$ and $N$ are given differentiable functions from the appropriate spaces into $\mathbb{R}$. Adequate regularity and growth assumptions are assumed to hold on $f$ and $L$ to insure existence and unicity of the solution of (1) and existence of (2).

**Regularity assumptions** We shall assume that the functions $f, L$ are of class $C^1$ and a growth condition holds on $f$, that guarantees the existence of a unique solution $S$ to (1) over $[t_0, T]$ for any $(U, W) \in (\mathcal{U}, \mathcal{W})$.

The precise formulations of the problems depend upon information structures and will be given below in the following sections.

## 2.2 The classical game problem formulation and its solution

Let us introduce the standard problem in perfect information, that is with admissible strategies in state feedback, i.e. of the form:

$$u_i = \varphi_i(t, x), \quad i = 1, 2,$$

and recall the classical Hamilton-Jacobi-Isaacs solution [11, 3]. Given the initial time and state $(t_0, x^0)$ determine, if it exists, the Isaacs' upper value function:

$$V(t_0, x^0) = \min_{u_1} \min_{u_2} \max_{w} \tilde{J}\,, \tag{3}$$

where $\tilde{J} = J - N(x^0)$.

**Proposition 1.** *If there exists a $C^1$ function $V : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}$, solution of the partial differential equation*

$$-\frac{\partial V}{\partial t} = \min_{u_1 \in \mathsf{U}_1} \min_{u_2 \in \mathsf{U}_2} \max_{w \in \mathsf{W}} H(t, x, \frac{\partial V}{\partial x}, u_1, u_2, w) \tag{4}$$

*with boundary condition:*

$$\forall x,\, V(T, x) = M(x),$$

*where*

$$H(t, x, \mu, u_1, u_2, w) = L + \langle \mu, f \rangle \ ,$$

*is the Hamiltonian of the system (the angled brackets* $\langle \cdot, \cdot \rangle$ *denote the scalar product in* $\mathbb{R}^n$*), then the upper value of the game (3) is* $V(t_0, x^0)$*. Moreover, if the Hamiltonian has a saddle point in* $(u, w)$ *for all* $(x, \mu)$*, and if there exist admissible strategies*

$$u(t) = \varphi^*(t, x(t)) = \begin{pmatrix} \varphi_1^*(t, x(t)) \\ \varphi_2^*(t, x(t)) \end{pmatrix}, \quad w(t) = \psi^*(t, x(t)) \tag{5}$$

*which are a saddle point of* $H(t, x, \partial V/\partial x, u_1, u_2, w)$*, then they are optimal.*

Hereafter, $\varphi^*$, $\psi^*$, together with $V$, will be referred to as the *Isaacs solution.*

**Standing assumption** In the sequel, we shall allways assume that the Isaacs solution exists and is unique.

# 3 The state feedback partial team problem

## 3.1 Statement of the problem

In the problem investigated in this section the players have different informations about the evolution of the system over time: we shall suppose that the first player (indicated by subindex 1) can measure only :

$$y_1 = h_1(t, x, w)\,, \tag{6}$$

while the second one (indicated by subindex 2) has access to exact and instantaneous state measurement.

For any function $a(\cdot) : t \to a(t)$, we shall use the notation $a^\tau$ for its restriction to $[t_0, \tau]$. Notice, that with a mild abuse of notations we may write causality of $S$ as $S^\tau(x^0, u_1, u_2, w) = S^\tau(x^0, u_1^\tau, u_2^\tau, w^\tau)$.

The admissible strategies are thus of the form:

$$u_1(t) = \varphi_1(t, y_1^t)\,, \quad u_2(t) = \varphi_2(t, x(t)). \tag{7}$$

For arbitrary initial conditions $(t_0, x^0)$ call disturbances the pairs $\omega := (x^0, w) \in \Omega := \mathbb{R}^n \times \mathcal{W}$, and define $X^0 := \mathbb{R}^n$. We shall consider several information structures beyond (7), where $x^0 \in X^0$ is not known to the players. This is why we have added the "initial cost" $N(x^0)$ in (2).

**The state feedback partial team problem** is the following: *Under the information structure (7) find optimal controls for the minimizing players, guaranteeing*

$$\min_{\varphi_1}\min_{\varphi_2}\max_{\omega} J(t_0, x^0, \varphi_1, \varphi_2, w)$$

*J being given by (2).*

### 3.2 Solution via the minimax certainty equivalence principle

We shall consider the system driven by $u_2(t) = \varphi_2^*(t, x)$. If this strategy is fixed, then the problem for player 1 is a classical partial information minimax control problem, as in [2] or [6]. Observe also that the corresponding full state information minimax control problem for player 1 has $\varphi_1^*$ and $V$ as its unique solution.

Let the trajectories of this system be denoted by $x(\cdot) = S_1(x_0, u_1, w) = S_1(u_1, \omega)$, and denote by

$$\Omega_1^\tau(u_1^\tau, y_1^\tau) = \{\omega \in \Omega \,|\, h_1^\tau(\cdot, S_1(u_1, \omega), w(\cdot)) = y_1^\tau\}$$

the set of $\omega$'s which are compatible with the past observations of the first player. It is clear that $\forall t, \Omega_1^t \in \Omega$.

We introduce, following [2] or [6], the *auxiliary problem* : for every $\tau \in [t_0, T]$ and for any fixed $u_1^\tau$ and $y_1^\tau$, we define a maximisation problem in $\omega$ as follows. Let

$$\dot{x} = f(t, x, u_1^\tau(t), \varphi_2^*(t, x), w)\,, \quad x(t_0) = x^0\,.$$

For $t \in [t_0, \tau]$, the state $x$ is now a function of $\omega$, (and more specifically of its restriction $\omega^\tau$ to $[t_0, \tau]$). Let

$$G(\tau, u_1^\tau, \omega) = V(\tau, x(\tau)) + \int_{t_0}^{\tau} L(t, x(t), u_1(t), \varphi_2^*(t, x(t)), w(t))dt + N(x^0),$$

The auxiliary problem is to characterize

$$\widehat{\Omega}_1^t = \arg\max_{\omega \in \Omega_1^t(u_1^\tau, y_1^\tau)} G(t, u_1, \omega)\,,$$

and

$$\widehat{X}_1(t) = \{\hat{x}_1(t) \mid \hat{x}_1(\cdot) = S_1(u_1^t, \hat{\omega})\,, \text{and } \hat{\omega} \in \widehat{\Omega}_1^t\}.$$

Following [2, 6], we assume :

**Crucial Assumption.** For all pairs $(\omega, u_1, u_2)$ and for all $t \in [t_0, T]$, suppose that $\widehat{X}_1$ is a singleton [6].

Let therefore $\widehat{X}_1 = \{\hat{x}_1(t)\}$. (This does not necessarily imply that $\widehat{\Omega}_1^t$ is a singleton.)

**Theorem 1.** *Under the Crucial Assumption above, the pair of optimal controls $\varphi_1^*(t, \hat{x}_1(t))$, $\varphi_2^*(t, x(t))$ ($\varphi_i^*(t, x(t))$ being defined by (5)) solves the partial team problem. Moreover*

$$\max_{\omega \in \Omega} J(t_0, \varphi_1^*(t, \hat{x}_1(t)), \varphi_2^*(t, x(t)), \omega) = \max_{x^0 \in X^0} [V(t_0, x^0) + N(x^0)]. \quad (8)$$

*Proof*. Apply the certainty equivalence principle of [2] to the problem obtained replacing everywhere $u_2$ by $\varphi_2^*(t, x)$ in the original problem (1)(2)(6). This yields (8). Now, notice that the value (8) is equal to the Isaacs value, corresponding to the full information case, and (on the basis of uniqueness of $\hat{x}_1(t)$) can be presented as:

$$\max_{\omega \in \Omega} J(t_0, \varphi_1^*(t, \hat{x}_1), \varphi_2^*(t, x), \omega) = V(t_0, \hat{x}_1^0) + N(\hat{x}_1^0). \quad (9)$$

Hereafter we shall call the strategy $\hat{u}_1(t) = \varphi_1^*(t, \hat{x}_1(t))$ the partial team strategy of the first player.

It is clear that a similar result can be proved for the second player, supposing that the first player has complete knowledge of the system's evolution in time, whereas the second one has access only to an output map, other than (6):

$$y_2 = h_2(t, x, w) \quad (10)$$

The partial team strategy of the second player, solving his or her partial team problem, will be:

$$\hat{u}_2(t) = \varphi_2^*(t, \hat{x}_2(t)), \quad (11)$$

$\hat{x}_2$ being generated by a similar auxiliary problem, corresponding to measurement equation (10), well posed for player 2. The value of the game in this case is also equal to Isaacs' value and can be presented as:

$$\max_{\omega \in \Omega} J(t_0, \varphi_1^*(t, x), \varphi_2^*(t, \hat{x}_2), \omega) = V(t_0, \hat{x}_2^0) + N(\hat{x}_2^0). \quad (12)$$

Denote the worst disturbance in this case by $\widehat{\Omega}_2 = (\hat{x}_2^0, \hat{w}_2)$ and by $\widehat{\Omega}_2^t$ the set of worst disturbances compatible with the observations of the second player.

Notice that $\hat{x}_1(t)$ and $\hat{x}_2(t)$ are the estimates of the *same* state variable $x(t)$ from the viewpoint of two players-minimizers 1 and 2, depending upon their informations, given by (6) and (10) respectively. As well $\widehat{\Omega}_1$ and $\widehat{\Omega}_2$ are the worst disturbances in the corresponding contexts.

One can conclude on the base of (8), (9), (12), that

$$\hat{x}_1^0 = \hat{x}_2^0 = \arg\max_x [V(t_0, x) + N(x)].$$

# 4 The full team problem

## 4.1 A direct analysis

Consider now the case when both players have different imperfect measurements of the state variable (6), (10), i.e. the admissible strategies are of the form:

$$u_1(t) = \varphi_1(t, y_1^t), \quad u_2(t) = \varphi_2(t, y_2^t). \tag{13}$$

Write, for short, $\hat{\varphi}_1$ for $\varphi_1^*(t, \hat{x}_1)$ and $\hat{\varphi}_2$ for $\varphi_2^*(t, \hat{x}_2)$.

**The full team problem** is the following: *Under the information structure (6),(10),(13), find optimal controls for players-minimizers, guaranteeing*

$$\min_{\varphi_1} \min_{\varphi_2} \max_{\omega} J(t_0, x_0, \varphi_1, \varphi_2, w)$$

*J being defined by (2).*

We need:

**Crucial Assumption.** The Crucial Assumption of the partial team problem for player 1 and the corresponding one for player 2 hold, and in addition, the functional $\omega \mapsto J(\hat{\varphi}_1, \hat{\varphi}_2, \omega)$ has a unique local and global maximum (e.g., it is quasiconcave).

**Technical Assumptions.**

- The value function in (4) and the optimal strategies $\varphi^*$ are continuously differentiable with respect to $x$,
- the optimal controls $u_i = \varphi_i^*(t, x(t))$ are interior to the sets $\mathsf{U}_i$ along an optimal trajectory,

- under the the partial-team strategies, the "worst case" estimates $\hat{x}_1$, $\hat{x}_2$ satisfy filter-like equations as in [9] :

$$\dot{\hat{x}}_i = \hat{f}_i(t, \hat{x}_i, y_i) \tag{14}$$

with $\hat{f}_i$ continuously differentiable in $\hat{x}_i$.

**Remark 1.** *Such filter like equations were derived in [9] for the case where the disturbance variable is split into a dynamics disturbance, say $w$, and additive measurement disturbance(s), say $y_i = h_i(t,x) + v_i$, and the integrand in the performance index is a sum $L = L_0(t,x,u,w) + K(t,x,u,v_1,v_2)$. A similar approach can be used to derive filter equations in the more general setting of this paper, but they would lack the appealing character of Didinsky's filter of [9].*

**Theorem 2.** *Under the above two assumptions the pair of control laws $\hat{\varphi}_1 = \varphi_1^*(t,\hat{x}_1)$, $\hat{\varphi}_2 = \varphi_2^*(t,\hat{x}_2)$ solves the full team problem. Moreover*

$$\max_{\omega\in\Omega} J(t_0, \hat{\varphi}_1, \hat{\varphi}_2, \omega) = \max_{x^0\in X^0} [V(t_0, x^0) + N(x^0)] .$$

**Proof** Recall first that the "filter equations" (14) for $i = 1, 2$ have the same initial state

$$\hat{x}^0 = \arg\max_x [V(t_0, x) + N(x)] .$$

Consider the following system of differential equations:

$$\left\{\begin{array}{ll} \dot{x} = f(t, x, \varphi_1^*(t,\hat{x}_1), \varphi_2^*(t,\hat{x}_2), w) , & x(t_0) = x^0 , \\ \dot{\hat{x}}_1 = \hat{f}_1(t, \hat{x}_1, y_1) , & \hat{x}_1(t_0) = \hat{x}^0 , \\ \dot{\hat{x}}_2 = \hat{f}_2(t, \hat{x}_2, y_2) , & \hat{x}_2(t_0) = \hat{x}^0 . \end{array}\right. \tag{15}$$

the following performance index:

$$J(t_0, \omega) = M(T, x(T)) + \int_{t_0}^{T} L(t, x, \hat{\varphi}_1, \hat{\varphi}_2, w)dt + N(x^0) \tag{16}$$

and the goal

$$\max_{\omega} J(t_0, \omega). \tag{17}$$

The system (15), (16), (17) presents a classical optimal control problem with respect to the control variable $w$ and the initial state $x^0$. Under the Technical Assumption it can be investigated via Pontryagin's maximum principle [12].

The following notations are accepted below:

$$\frac{\partial \varphi_i^*}{\partial x} = \frac{\partial \varphi_i^*}{\partial x}(t,x)\,, \qquad \frac{\partial \hat{\varphi}_i}{\partial \hat{x}_i} = \frac{\partial \varphi_i^*}{\partial x}(t,\hat{x}_i)\,.$$

Let us write the Hamiltonian and necessary conditions:

$$H(t,x,\hat{x}_1,\hat{x}_2,\lambda,\hat{\lambda}_1,\hat{\lambda}_2,w) = L + \lambda^t f + \hat{\lambda}_1^t \hat{f}_1 + \hat{\lambda}_2^t \hat{f}_2\,,$$

where $\lambda, \hat{\lambda}_1, \hat{\lambda}_2$ are the corresponding adjoint variables for $x, \hat{x}_1, \hat{x}_2$.

We write the adjoint equations for the maximization problem (17):

$$\begin{cases} -\dot{\lambda}^t = \frac{\partial L}{\partial x} + \lambda^t \frac{\partial f}{\partial x} + \hat{\lambda}_1^t \frac{\partial \hat{f}_1}{\partial y_1}\frac{\partial h_1}{\partial x} + \hat{\lambda}_2^t \frac{\partial \hat{f}_2}{\partial y_2}\frac{\partial h_2}{\partial x}\,, \\ -\dot{\hat{\lambda}}_1^t = (\frac{\partial L}{\partial u_1} + \lambda^t \frac{\partial f}{\partial u_1})\frac{\partial \hat{\varphi}_1}{\partial \hat{x}_1} + \hat{\lambda}_1^t \frac{\partial \hat{f}_1}{\partial \hat{x}_1}\,, \\ -\dot{\hat{\lambda}}_2^t = (\frac{\partial L}{\partial u_2} + \lambda^t \frac{\partial f}{\partial u_2})\frac{\partial \hat{\varphi}_2}{\partial \hat{x}_2} + \hat{\lambda}_2^t \frac{\partial \hat{f}_2}{\partial \hat{x}_2}\,, \end{cases} \tag{18}$$

and the following boundary conditions:

$$\begin{cases} \lambda(t_0) + \frac{\partial N}{\partial x}(x^0) = 0\,, \\ \lambda(T) - \frac{\partial M}{\partial x}(T, x(T)) = 0\,, \\ \hat{\lambda}_i(T) = 0. \end{cases} \tag{19}$$

Pontryagin's principle [12] claims, that the optimal control $\hat{w}$, solving the problem (15), (16) (17) yields:

$$\max_w H(x,\hat{x}_1,\hat{x}_2,\lambda,\hat{\lambda}_1,\hat{\lambda}_2,w) = H(x,\hat{x}_1,\hat{x}_2,\lambda,\hat{\lambda}_1,\hat{\lambda}_2,\hat{w}). \tag{20}$$

Denote by $x^*(\cdot)$ the optimal trajectory, i.e. the trajectory generated from $x_0$ by the strategies $\varphi_i^*$ and $\psi^*$. Then we claim:

**Proposition 2.** *The following solves Pontryagin's necessary conditions (18) to (20):*

$$x(t) = \hat{x}_1(t) = \hat{x}_2(t) = x^*(t)\,, \tag{21}$$

$$\lambda(t) = \frac{\partial V}{\partial x}(t, x^*(t))\,, \tag{22}$$

$$\hat{\lambda}_1(t) = \hat{\lambda}_2(t) = 0\,, \tag{23}$$

$$\hat{w}(t) = \psi^*(t, x^*(t))\,. \tag{24}$$

*Proof of the proposition* Observe that, if indeed the $\hat{\lambda}_i$'s are zero, then the first equations in (18) and (19) are the standard adjoint equations associated with the full information Isaacs equation, and because we have assumed $V$ to be $C^1$ in $x$, its gradient solves them along an optimal trajectory. Also, because we have assumed that the optimal strategies $\varphi_i^*$ take their values in the interior of the control sets $\mathsf{U}_i$, with the proposed solution,

$$\frac{\partial L}{\partial u_1} + \lambda^t \frac{\partial f}{\partial u_1} = 0\,, \qquad \frac{\partial L}{\partial u_2} + \lambda^t \frac{\partial f}{\partial u_2} = 0\,,$$

so that the differential equations for $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are homogeneous. And their boundary condition being 0, $\hat{\lambda}_i \equiv 0$ is indeed a solution. The rest follows easily: the proposed $w$ indeed maximizes the Hamiltonian.

As a consequence of the proposition, the disturbance generated by $\psi^*$ solves the optimisation problem (15),(16),(17). Therefore the worst possible performance index under the proposed strategies for the team players is again Isaac's value, and no causal controller can do better.

### 4.2 An abstract analysis

An abstract view of the last result is as follows. Let $J : \mathcal{U} \times \mathcal{W} \to \mathbb{R}$ be a performance index from two Banach spaces $\mathcal{U}$ (here $\mathcal{U}_1 \times \mathcal{U}_2$) and $\mathcal{W}$ into the real line, with Fréchet derivatives $\partial J/\partial u$ and $\partial J/\partial w$, the first one jointly continuous in $(u, w)$. Let $\Phi : \mathcal{W} \to \mathcal{U}$ be a set of admissible strategies for player 1, (here, causal strategies, or state feedbacks), and $\widetilde{\Phi}$ a subset of $\Phi$ (here partial information decentralized strategies of the form (13)).

We make the following hypotheses on $\Phi$:

- $\Phi$ is a vector space,
- $\Phi$ contains the constant maps, i.e. $\forall \bar{u} \in \mathcal{U}$, the map $w \mapsto \bar{u}$, $\forall w \in \mathcal{W}$ is in $\Phi$.

Assume there exists $\varphi^* \in \Phi$ solution of the problem

$$\min_{\varphi \in \Phi} \sup_{w \in \mathcal{W}} J(\varphi(w), w)\,,$$

and a *unique* $w^*$ solution of the related maximization problem:

$$J(\varphi^*(w^*), w^*) = \max_{w \in \mathcal{W}} J(\varphi^*(w), w)\,.$$

**Lemma 1.** *We have the following fact:*

$$\frac{\partial J}{\partial u}(\varphi^*(w^*), w^*) = 0\,.$$

*Proof* Let $G(\varphi, w) = J(\varphi(w), w)$, and $\bar{G}(\varphi) = \sup_{w\in\mathcal{W}} G(\varphi, w)$. Place on $\Phi$ the topology of the uniform convergence. It is clear that $G$ has a partial Fréchet derivative in $\varphi$, given by

$$\forall\psi\in\Phi\,,\quad \frac{\partial G}{\partial\varphi}(\varphi, w)\cdot\psi = \frac{\partial J}{\partial u}(\varphi(w), w)\cdot\psi(w)\,.$$

This partial derivative is continuous in $(\varphi, w)$ (still for the topology of the uniform convergence on $\Phi$). Hence, by Danskin's theorem [5], everytime the max in $w$ is reached at a unique point $\hat{w}$, there exists

$$\frac{d\bar{G}}{d\varphi}(\varphi) = \frac{\partial G}{\partial\varphi}(\varphi, \hat{w})\,.$$

Since $\varphi^*$ minimizes $\bar{G}$ over $\Phi$, and we assumed the maximizing $w^*$ to exist and be unique, it comes

$$0 = \frac{d\bar{G}}{d\varphi}(\varphi^*) = \frac{\partial G}{\partial\varphi}(\varphi^*, w^*)$$

Hence, $\forall\psi\in\Phi$,

$$\frac{\partial J}{\partial u}(\varphi^*(w^*), w^*)\cdot\psi(w^*) = 0\,.$$

But we have also assumed that constant maps were in $\Phi$. Hence

$$\forall u\in\mathcal{U}\,,\quad \frac{\partial J}{\partial u}(\varphi^*(w^*), w^*)\cdot u = 0\,,$$

whence the claim of the lemma.

We add now a technical assumption: $\varphi^*$ is differentiable.

Differentiating $J(\varphi^*(w), w)$ at $w^*$, we deduce easily that in addition to the above, $(\partial J/\partial w)(\varphi^*(w^*), w^*) = 0$.

Let now $\hat{\varphi}\in\widetilde{\Phi}$ (in our case our set of two partial team strategies together) be differentiable, and such that

- $\hat{\varphi}(w^*) = \varphi^*(w^*)$,
- $w\mapsto J(\hat{\varphi}(w), w)$ has a unique local and global maximum. (Say, is quasi-concave).

Then we have the proposition:

**Proposition 3.** *$\hat{\varphi}$ solves the problem*

$$\min_{\varphi\in\widetilde{\Phi}} \sup_{w\in\mathcal{W}} J(\varphi(w), w)\,.$$

*Proof* Clearly, $J(\hat{\varphi}(w), w)$ is stationary at $w^*$, which suffices to show that its max in $w$ is $J(\hat{\varphi}(w^*), w^*) = J(\varphi^*(w^*), w^*)$, which is the smallest possible $\sup_w J(\varphi(w), w)$ for $\varphi$ ranging over $\Phi$, and a fortiori for $\varphi$ ranging over the subset $\widetilde{\Phi}$.

This abstract version of our last result only stresses that it is indeed weak. One could *assume* that the payoff $J$ with the *open loop* controls $\hat{u}_i$ has only one local and global maximum in $\omega$, and then these open loop controls would constitute an optimal decentralized strategy. Hence, our result needs to be complemented by some kind of reciprocal, showing that if our "partial team" strategies do not succeed, none will, in some sense. Investigation of this matter is in progress.

# 5 The linear quadratic case

## 5.1 The system

Let the system with its observed outputs be

$$\begin{aligned} \dot{x} &= Ax + B_1u_1 + B_2u_2 + Dw\,, \\ y_1 &= C_1x + E_1w\,, \\ y_2 &= C_2x + E_2w\,. \end{aligned}$$

We shall assume that the pair $(A, D)$ is stabilizable, that

$$\begin{pmatrix} D \\ E_1 \\ E_2 \end{pmatrix} (D^t\; E_1^t\; E_2^t) = \begin{pmatrix} M & 0 & 0 \\ 0 & N_1 & 0 \\ 0 & 0 & N_2 \end{pmatrix}$$

with $N_1$ and $N_2$ positive definite.

Let the performance index be of the form

$$J = \|x(T)\|_X^2 + \int_0^T (\|x\|_Q^2 + \|u_1\|_{R_1}^2 + \|u_2\|_{R_2}^2 - \gamma^2\|w\|^2)\,dt - \gamma^2\|x^0\|_Y^2\,.$$

Assume finally that $Q \geq 0$ and that the pair $(Q^{1/2}, A)$ is completely reconstructible.

Using, for instance, the theory of [2], it is a simple matter to apply the above theory to that minimax team problem.

### 5.2 The Isaacs solution

The above system admits a state feedback minimax control of the form

$$\varphi^*(t,x) = \begin{pmatrix} -F_1(t)x \\ -F_2(t)x \end{pmatrix}$$

with $F_i = R_i^{-1} B_i^t P(t)$, $i = 1, 2$, if, firstly, the Riccati equation

$$\dot{P} + PA + A^t P - F_1^t R_1 F_1 - F_2^t R_2 F_2 + \gamma^{-2} PMP + Q = 0\,, \quad P(T) = X \tag{25}$$

has no conjugate point, defining the symmetric matrix function $P(\cdot)$ over $[0, T]$.

Then, the worst disturbance is given by $w(t) = \psi^*(t, x(t))$ with $\psi^*(t,x) = \gamma^{-2} D^t P(t) x$. The Value function for the game $\tilde{J}$ (i.e. without the initial term) is $V(x^0) = \|x^0\|^2_{P(0)}$, so that the minimax of the criterion exists if, secondly, $P(0) - \gamma^2 Y < 0$. It is known that this last condition can also be written in terms of $Z := Y^{-1}$ as $\rho(ZP(0)) < \gamma^2$.

In accordance with the above theory, the standing assumption will be that these conditions are fulfilled.

### 5.3 The partial team problem

Let us now examine our partial team problem. We set $u_2(t) = \varphi_2^*(t, x(t))$, so that the dynamics of the problem at hand are now

$$\dot{x} = A_1 x + B_1 u_1 + Dw\,,$$

with $A_1 := A - B_2 F_2$.

Also, the relevant criterion is

$$J_1 = \|x(T)\|_X^2 + \int_0^T (\|x\|^2_{Q_1} + \|u_1\|_{R_1} - \gamma^2 \|w\|^2)\, dt + \|x^0\|^2_Y\,,$$

with $Q_1 := Q + F_2^t R_2 F_2$.

To apply the theory of [2] to this problem, we should first define a matrix $P_1(t)$ solution of a game Riccati equation written with the sole control $u_1$, and $A$ and $Q$ replaced by $A_1$ and $Q_1$. We notice however that under the standing assumption that $P$ exists, this equation has a solution, viz. $P(\cdot)$ itself. Therefore, the two partial team problems and the full state feedback control problem share the same $P$ matrix. (Denoted $Z$ in [2]).

Let us further write the dual Riccati equation of that problem :

$$\dot{\Sigma}_1 = A_1\Sigma_1 + \Sigma_1 A_1^t - \Sigma_1 C_1^t N_1^{-1} C_1 \Sigma_1 + \gamma^{-2}\Sigma_1 Q_1 \Sigma_1 + M\,, \quad \Sigma_1(0) = Z\,. \tag{26}$$

The crucial assumption for this partial team problem is here that

1. The Riccati equation (26) has no conjugate point over $[0,T]$,
2. $\forall t \in [0,T]\,, \qquad \rho(\Sigma_1(t)P(t)) < \gamma^2$.

Then, the "worst" state $\hat{x}_1(t)$ is given as the solution of the filter equation

$$\dot{\hat{x}}_1 = A_P\hat{x}_1 + K_1(y_1 - C_1\hat{x}_1)\,, \quad \hat{x}_1(0) = 0\,, \tag{27}$$

with $A_P := A - B_1F_1 - B_2F_2 + \gamma^{-2}MP$ and $K_1 = (I - \gamma^{-2}\Sigma_1 P)^{-1}C_1^t N_1^{-1}$. The partial team minimax strategy is

$$u_1(t) = -F_1(t)\hat{x}_1(t)\,. \tag{28}$$

### 5.4 The full team problem

Let us now turn to the full team problem. the maximization problem to investigate has its state in $\mathbb{R}^{3n}$ and is as follows :

$$\begin{pmatrix} \dot{x} \\ \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{pmatrix} = \mathcal{A}\begin{pmatrix} x \\ \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} + \mathcal{D}w\,, \qquad \begin{pmatrix} x(0) \\ \hat{x}_1(0) \\ \hat{x}_2(0) \end{pmatrix} = \begin{pmatrix} x^0 \\ 0 \\ 0 \end{pmatrix}\,,$$

with

$$\mathcal{A} = \begin{pmatrix} A & -B_1F_1 & -B_2F_2 \\ K_1C_1 & A_P - K_1C_1 & 0 \\ K_2C_2 & 0 & A_P - K_2C_2 \end{pmatrix}\,, \quad \mathcal{D} = \begin{pmatrix} D \\ K_1E_1 \\ K_2E_2 \end{pmatrix}\,,$$

and the performance index to be maximized by $w$ is

$$G = \|x(T)\|^2 + \int_0^T (\|x\|_Q^2 + \|\hat{x}_1\|^2_{F_1^tR_1F_1} + \|\hat{x}_2\|^2_{F_2^tR_2F_2} - \gamma^2\|w\|^2)dt - \gamma^2\|x^0\|_Y^2.$$

This maximization problem has a (unique) solution if, firstly, the Riccati equation for the $3n \times 3n$ block matrix $\Pi = (\Pi_{ij})$, $i,j = 0,1,2$ :

$$\dot{\Pi} + \Pi\mathcal{A} + \mathcal{A}^t\Pi + \gamma^{-2}\Pi\mathcal{D}\mathcal{D}^t\Pi + \mathcal{Q} = 0\,, \quad \Pi(T) = \mathrm{diag}(X,0,0)\,, \tag{29}$$

with $\mathcal{Q} = \mathrm{diag}(Q, F_1^tR_1F_1, F_2^tR_2F_2)$, has no conjugate point over $[0,T]$, and, secondly $\Pi_{00} - \gamma^2 Y < 0$, or, equivalently, $\rho(Z\Pi_{00}(0)) < \gamma^2$.

Let us summarize:

**Theorem 3.** *If*

1. *The Riccati equation (25) has no conjugate point over* $[0, T]$*,*

2. *The Riccati equations (26) and the similar one reversing the indices 1 and 2 have no conjugate points over* $[0, T]$*,*

3. $\forall t \in [0, T] \quad \rho(\Sigma_i(t)P(t)) < \gamma^2$*,* $i = 1, 2$*,*

4. *The Riccati equation (29) has no conjugate point over* $[0, T]$*,*

5. $\rho(Z\Pi_{00}(0)) < \gamma^2$*,*

*then the strategies (28) where* $\hat{x}_1(\cdot)$ *is given by (27), and similarly mutatis mutandis for* $u_2$*, are team mimnimax optimal.*

(The condition on $\rho(ZP(0))$ has been omitted, because it is implied by conditions (3) at time zero.)

Writing explicitly the six blocks of Riccati equation (29) is no too nice to do. It is nevertheless nothing more than an exercise in elementary matrix calculus. The noteworthy point about it is that it is not difficult to check on those equations that

$$\Pi(t) \begin{pmatrix} I \\ I \\ I \end{pmatrix} = \begin{pmatrix} P(t) \\ 0 \\ 0 \end{pmatrix} .$$

As a consequence, one can check, as should be expected, that if $w$ is set to the worst disturbance, all three vectors $x$, $\hat{x}_1$ and $\hat{x}_2$ agree. (Also, absence of a conjugate point to (29) implies absence of a conjugate point to (25).)

# References


[1] R. Bansal and T. Başar : "Solutions to a class of linear-quadratic-gaussian (LQG) stochastic team problems with nonclassical information", *Systems & Control Letters*, **9**, pp 125–130, 1987.

[2] T. Başar and P. Bernhard: *$\mathcal{H}_\infty$-optimal control and related minimax control problems: a dynamic game approach*, Birkhauser, Boston, 1991. (Second edition: 1995.)

[3] P. Bernhard, *"Differential Games: Isaacs' equation", in Encyclopedia of Systems and Control,* M.Singh (Ed), Pergamon, Pergamon Press, 1987.

[4] P. Bernhard, N. Hovakimyan: "Nonlinear robust control and minimax team problems", *International Journal of Robust and Nonlinear Control.* Vol.9, pp.239-257, 1999.

[5] P. Bernhard and A. Rapaport: "On a theorem of Danskin with an application to a theorem of Von Neumann-Sion", *Nonlinear Analysis, Theory, Methods & Applications"*, **24**, pp1163–1181, 1995.

[6] P. Bernhard and A. Rapaport: " Min-Max certainty equivalence principle and differential games", *Internat. Journal on robust and nonlinear control*, **6**, pp 825–842, 1996.

[7] K-C. Chu : "Team decision theory and information structures in optimal control problems II" *IEEE-AC*, **17**, pp 22–28, 1972

[8] G. Didinsky and T. Başar : "Minimax decentralized controllers for discrete time linear systems", *Proceedings, 31st IEEE-CDC*, Tucson, USA, december 1992.

[9] G. Didinsky, T. Başar and P. Bernhard, *Structural properties of minimax controllers for a class of differential games arising in nonlinear $H^\infty$ control*, Systems & Control Letters 21, pp.433-441, 1993.

[10] Y-C. Ho and K-C Chu : "Team decision theory and information structures in optimal control problems I" *IEEE-AC*, **17**, pp 15–22, 1972

[11] R. Isaacs, *Differential Games.* Wiley, New York, 1965.

[12] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, E.F.Mischenko, *The Mathematical Theory of Optimal Processes*, John Wiley & Sons, New York, London, 1962.

[13] R. Radner "Team decision problems", *Annals of mathematical statistics*, **33**, pp 857–881, 1962.

[14] R. Srikant : "Relationship between decentralized controller design using $H_\infty$ and stochastic risk averse criteria", *IEEE-AC* **39**, 861–864, 1994

[15] R. Srikant and T. Başar : "Approximate Nash strategies and suboptimal team solutions in Markov chains with weakly coupled decision makers", *International Year Book on Game Theory and Applications* **1**, pp 65–82, 1993.