

AD-A145 212

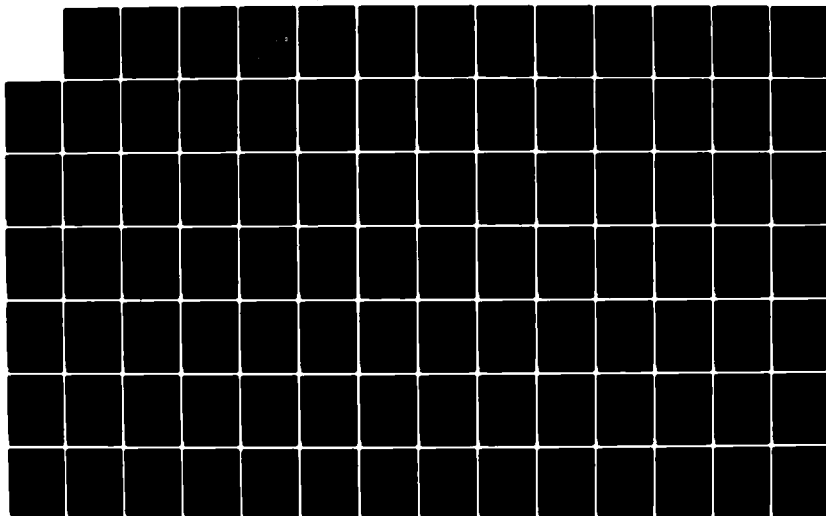
ROBUST TEAM-OPTIMAL AND LEADER-FOLLOWER POLICIES FOR
DECISION MAKING IN C... (U) ILLINOIS UNIV AT URBANA
T BASAR ET AL. JUL 84 N00014-82-K-0469

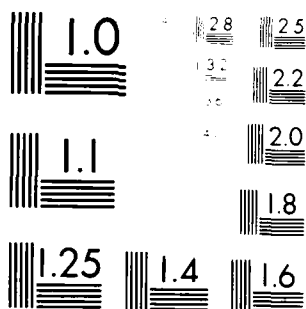
1/3

UNCLASSIFIED

F/G 12/1

NL





Resolution Test Chart
 (See Note 1 on Page 1 of 2)

AD-A145 212

DOCUMENTATION PAGE

1. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS None	
2. SECURITY CLASSIFICATION AUTHORITY N/A		3. DISTRIBUTION/AVAILABILITY OF REPORT	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A			
PERFORMING ORGANIZATION REPORT NUMBER(S)		5. MONITORING ORGANIZATION REPORT NUMBER(S)	
4. NAME OF PERFORMING ORGANIZATION		6b. OFFICE SYMBOL (If applicable) N/A	
7a. NAME OF MONITORING ORGANIZATION		7b. ADDRESS (City, State and ZIP Code)	
c. ADDRESS (City, State and ZIP Code)		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	
NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research		8b. OFFICE SYMBOL (If applicable)	
c. ADDRESS (City, State and ZIP Code) Department of the Navy ONR Arlington, Virginia 22217		10. SOURCE OF FUNDING NOS.	
1. TITLE (Include Security Classification) Robust Team-Optimal and Leader-Follower Policies for Decision Making in C3 Systems		PROGRAM ELEMENT NO.	
		PROJECT NO.	
		TASK NO.	
		WORK UNIT NO.	
2. PERSONAL AUTHOR(S) Tamer Başar, Jose B. Cruz, Jr.		11. DISTRIBUTION STATEMENT A Approved for public release Distribution Unlimited	
3a. TYPE OF REPORT Final Progress Report		13b. TIME COVERED FROM 5/15/82 TO 5/14/84	
14. DATE OF REPORT (Yr., Mo., Day) July 1984		15. PAGE COUNT	
3. SUPPLEMENTARY NOTATION Final Report on ONR Contract N00014-82-K-0469			
7. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB. GR.	
9. ABSTRACT (Continue on reverse if necessary and identify by block number)			
This report summarizes the findings of a two-year research program. We have been concerned with fundamental research underlying some aspects of the Command, Control, and Communications systems field. Our starting point has been that C₃ problems and other large scale decentralized decision and control problems which are traditionally modeled as team problems can be investigated much more effectively as multiple-goal multiple decision maker problems, especially under uncertainty, and we have developed a framework which has permitted us to perform a sensitivity analysis on team-optimal and leader-follower policies in such systems. Our general model involves hierarchies in decision making, informational decentralization, uncertainty in the available information, constraints on information transmission capabilities between different levels of hierarchies, and possible discrepancies between the perceptions of different decision makers of the common goal. The research program has comprised two main tasks, and in			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS ☐		21. ABSTRACT SECURITY CLASSIFICATION	
22a. NAME OF RESPONSIBLE INDIVIDUAL DTIC FILE COPY		22b. TELEPHONE NUMBER (Include Area Code)	
		22c. OFFICE SYMBOL	

addition to the investigation of the sensitivity and robustness of team-optimal and leader-follower strategies to variations in perceptions of the nominal goal (s), we have developed robust coordinator strategies (under uncertainty with regard to the environment and the perceived goals of the subordinate decision makers) which force the subordinate decision makers to act as a team as though their goals were exactly the same.

(12)

ROBUST TEAM-OPTIMAL AND LEADER-FOLLOWER POLICIES
FOR DECISION MAKING IN C^3 SYSTEMS

FINAL REPORT to ONR
for the period
May 15, 1982-May 14, 1984

Co-Principal Investigators:

Tamer Başar
J. B. Cruz, Jr.

DTIC
SELECTE
AUG 3 0 1984
H

Decision and Control Laboratory
Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Avenue
Urbana, Illinois 61801



DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

May 14, 1984

Association For		
DTIS GRA&I	☒	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By	<i>ltr</i>	
Distribution/		
Availability Codes		
Dist	Avail and/or	Special
A1		

TABLE OF CONTENTS

	Page
1. Introduction.....	1
2. Research Progress.....	3
3. Publications Supported by the Contract.....	9
4. Copies of Publications.....	11

1. INTRODUCTION

In this project our thesis has been that Command, Control and Communications (C^3) problems and other large scale decentralized decision and control problems which are traditionally modeled as team problems can be investigated much more effectively as multiple-goal multiple decision maker problems, especially under uncertainty, and our goal has been to develop a framework which would permit us to perform sensitivity analysis on team-optimal and leader-follower policies in such systems. Our general model has involved hierarchies in decision making, informational decentralization, uncertainty in the available information, constraints on information transmission-capabilities between different levels of hierarchies, and possible discrepancies between the perceptions of different decision makers of the common goal.

We have first noted that in order to model large scale decentralized decision and control problems as team problems, the following three conditions should be satisfied:

(i) All decision makers have exactly the same perception of an existing common goal, and quantify this perception in exactly the same way;

(ii) All decision makers have access to a common (probabilistic) description of the uncertainties inherent in the decision environment;

(iii) All decision makers adopt (or have access to) exactly the same (mathematical) model of the underlying system that characterizes the decision environment and possible paths of evolution of the decision processes, and have access to the same relevant information with regard to the interactions among, and capabilities of, different decision makers.

Any existing discrepancy between the perceptions of the decision makers with regard to any one of the foregoing three conditions leads to a decision problem which can no longer be treated as a team problem. Optimal decision rules derived by totally ignoring (or overlooking) this aspect of the problem in general lead to outcomes which are extremely sensitive even to small variations in the perceptions of the decision makers from the common nominal model.

Motivated by these considerations, we have undertaken, within the scope of this project, the tasks of (i) developing a methodology for performing a sensitivity analysis on deterministic and stochastic team and multi-person decision problems in the face of deviations (of the perceptions) from a common nominal model, (ii) using this methodology, to obtain minimally sensitive decision rules for different classes of deterministic and stochastic models, and (iii) investigating the role of information and hierarchy in the derivation of such policies.

We believe that in the scope of this project significant contributions have been made to the state of knowledge on these tasks, as documented in the references attached to the report. In the next section we provide brief descriptions of our research accomplishments in these and related areas, keyed to the reference list which constitutes Section 3 of this report.

2. RESEARCH PROGRESS

In this section, we briefly outline some of the new results obtained in the scope of this project, in five categories; full details can be found in the references given, which are attached (in full) to this report.

Category A: In the first three references listed in Category A, we have studied the sensitivity of leader-follower policies in two-agent deterministic decision problems to variations in the values of some parameters describing the objective functionals, for both Stackelberg and team problems. In [A1] we introduce an appropriate sensitivity function and introduce the notion of a "robust" incentive scheme for the leader as one that minimizes, in addition to the usual (standard) Stackelberg performance index, this sensitivity function. Such an approach has applications in decision problems wherein the leader does not know the exact values of some parameters characterizing the follower's cost functional, and seeks to "robustify" his optimum policy in the presence of deviations from the nominal values. In [A1] we also provide an in-depth analysis of such incentive design problems, obtain some explicit results for general convex cost functionals, and present some illustrative examples. In [A2] and [A3], on the other hand, we study a general class of "nominal" team problems with two agents and with a hierarchical decision structure, where we also allow one of the decision makers to have a slightly different perception of the overall team goal, with this slight variation not known by the other agent who is assumed to occupy the hierarchically dominant position. This leading agent is assumed to have access to dynamic information, and his role is to announce a robust policy (incentive scheme) which would lead to achievement of the overall team goal in spite of

the slight variations in the other agent's perception of that goal. In the paper we obtain such robust policies for the leading agent, for general cost functionals with convex structure, and also show that in some special cases this robust feature of the incentive scheme is maintained regardless of the magnitude and nature of the variations.

In [A5] and [A6] we study a version of the problem of [A1] in a stochastic context, which is a challenging dynamic optimization problem. More specifically, we consider a class of stochastic incentive decision problems in which the leader has access to the control value of the follower and to private as well as common information on the unknown state of nature. The follower's cost function depends on a finite number of parameters whose values are not known accurately by the leader, and in spite of this parametric uncertainty the leader seeks a policy which would induce the desired behavior (in a stochastic equivalence sense) on the follower. In the paper, we obtain such appealing policies for the leader, which are smooth, induce the desired behavior at the nominal values of these parameters, and furthermore make the follower's optimal reaction either minimally sensitive or totally insensitive to variations in the values of these parameters from the nominals. The general solution is determined by some orthogonality relations in some appropriately constructed (probability) measure spaces and leads to particularly simple incentive policies which have no counterparts in deterministic problems.

In [A6] we address a class of nominally stochastic team problems and explore the impact of the additional degree of freedom brought in by the team nature of the problem on the sensitivity properties of different team solutions. Finally, in [A7], which is currently under preparation, we provide a general survey of these results with possible extensions and applications.

Category B: In the first two papers listed in Category B, we study a class of decision problems in which this time the probabilistic description of the stochastic variables is perceived differently by different agents. We first show that when the decision makers have different probabilistic models of the stochastic environment, the resulting decision problem is a nonzero-sum (multi-criteria) stochastic game, even if the decision makers have a single common goal quantified in exactly the same way (say by a cost functional). Hence, even in team problems the corresponding solution concept (team-optimal solution) will have to be modified when discrepancies exist in the perception of the agents of the probabilistic model of the decision process. The currently available theory of nonzero-sum stochastic games was not applicable to such problems, and a brand new theory had to be developed. This is what we have accomplished in these papers, (cf. [B1], [B2] and [B4]), for two agent problems with static information patterns. We introduce the concept of "stable equilibrium solutions" for decision problems with multiple probabilistic models, and obtain sufficient conditions for existence and uniqueness of such equilibria (under a symmetric mode of decision making) when the objective functionals are quadratic and the decision spaces are general inner-product spaces. Furthermore, for the special case of Gaussian distributions in both discrete and continuous-time problems, we present in [B2] and [B4] some explicit stable equilibrium policies.

While [B1] and [B2] treat stochastic multimodeling under a symmetric mode of decision making, [B3] and part of [B4] deal with the case of asymmetric mode of decision making. Here again, even in team problems, a discrepancy in the perceptions of the decision makers of the underlying probabilistic model

leads to a stochastic nonzero-sum game, which is of the type not treated before in the literature. We develop a general equilibrium theory for such problems in [B3] and [B4], and analyze the special case of Gaussian distributions in greater depth. One of the important findings of this analysis is that while the equilibrium solution for the symmetric case is linear, this is no longer true when the mode of decision making is asymmetric. In this latter case, the unique equilibrium solution is nonlinear even under Gaussian distributions, when discrepancies exist. In the limiting case as discrepancies disappear, this nonlinear solution degenerates into a linear one (and so does the best linear one), thus displaying the existence of a bifurcation phenomenon. Reference [B5], which is currently under preparation, extends these results to decision problems with a larger number of agents and different types of hierarchies.

Category C: In paper [C1], we continue our earlier work on Stackelberg dynamic games and consider a subclass of such problems in which the leader has informational advantage over the follower, in the sense that the leader can observe the follower's actions at each stage (before he (the leader) acts) either perfectly or partially. Under a feedback Stackelberg solution concept which takes this informational advantage into account, we have studied derivation of optimal affine policies and have investigated the conditions under which such a solution coincides with the global Stackelberg solution. A second set of results obtained in [C1] involves an analysis of existence and derivation of causal real-time implementable global Stackelberg solution in dynamic games when the leader is allowed to use memory policies.

Category D: In the two papers listed in Category D, we have addressed the important problem of developing a general equilibrium theory for discrete and continuous-time dynamic games with varying (symmetrical and asymmetrical) modes of play, i.e. for games in which the solution concept itself and the leadership is determined by past actions of the players and the outcome of some (stochastic) process. In [D1] we study stochastic systems with structural and modal uncertainties described by a finite state jump process, and introduce a new concept of equilibrium (which we call "strong equilibrium") which encompasses both the feedback Nash and feedback Stackelberg solution concepts for the special cases of deterministic discrete-time games with symmetrical and asymmetrical modes of play, respectively. This new equilibrium concept also provides a convenient framework for the introduction of a feedback Stackelberg solution concept in deterministic differential games. For the general class of stochastic nonzero-sum dynamic games with structural and modal uncertainties, and under the feedback closed-loop information, we obtain the optimality conditions in both discrete and continuous time. We also study certain special cases, which are further discussed in [D2] along with some illustrative examples.

Category E: In paper [E1] we extend the currently available theory of dynamic games in a new direction, so as to encompass games with state equations of order higher than one. We first show that the standard state augmentation technique is not applicable in a game context, and therefore a new technique has to be developed which is tailored to the underlying information pattern. In the paper we develop such a technique whereby we obtain informationally unique Nash equilibrium solution to a class of dynamic games of order higher than one, and with random disturbances in the state equation.

In the invited paper [E2], we first provide a brief review of some of the recent results on dynamic games, in particular with regard to memory strategies, and then discuss potential applications of the techniques developed in this context to large scale systems design, optimization and coordination. We propose a number of design criteria for coordinator policies in interconnected systems, and provide recipes to obtain such policies with good sensitivity properties.

3. PUBLICATIONS SUPPORTED BY THE CONTRACT

A. *Sensitivity Considerations in Team-Stackelberg Equilibrium Solutions*

- [A1] D. H. Cansever and T. Başar, "A Minimum Sensitivity Approach to Incentive Design Problems," Large Scale Systems, vol. 5, December 1984, pp. 233-244. A preliminary version has been presented as a regular paper at the 21st IEEE Conference on Decision and Control, Orlando, FL, December 8-10, 1982, and appeared in the Conference Proceedings.
- [A2] D. H. Cansever, T. Başar and J. B. Cruz, Jr., "Robustness of Incentive Policies in Team Problems with Discrepancies in Goal Perceptions," in Dynamic Modelling and Control Economies, (Proc. 4th IFAC/IFORS Symposium, Washington, DC, June 17-19, 1983), 1984.
- [A3] T. Başar, J. B. Cruz, Jr. and D. H. Cansever, "A Minimum Sensitivity Incentive Control Approach to Team Problems," Proc. 6th MIT/ONR Conference on Command Control and Communications, Cambridge, MA, July 1983.
- [A4] T. Başar and D. H. Cansever, "Robust Incentive Policies for Stochastic Decision Problems in the Presence of Parametric Uncertainty," in Proc. IFAC World Congress, July 1984, Budapest, Hungary (to appear).
- [A5] D. H. Cansever and T. Başar, "Optimum or Near-Optimum Incentive Policies for Stochastic Decision Problems in the Presence of Parametric Uncertainty," submitted to Automatica, February 1984.
- [A6] D. H. Cansever and T. Başar, "Near-Optimum Incentive Policies in Stochastic Team Problems with Discrepancies in Goal Perceptions," in Proc. 23rd CDC, December 1984 (to appear).
- [A7] T. Başar, "Dynamic Games and Incentives," invited contribution to the Workshop on "Systems and Optimization" at THT, the Netherlands (April 16-18, 1984); volume including the paper will be published by Springer-Verlag in 1984, (still under preparation).

B. *An Equilibrium Theory when Discrepancies are present in the Perception of Probabilistic Models*

- [B1] T. Başar, "An Equilibrium Theory for Multi-Person Multi-Criteria Stochastic Decision Problems with Multiple Subjective Probability Measures," in Proc. 9th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, (Prague, Czechoslovakia, June 28-July 2, 1982), 1983.
- [B2] T. Başar, "Decision Problems with Multiple Probabilistic Models," Proc. 1983 American Control Conference, San Francisco, CA, pp. 1091-1096, June 1983.
- [B3] T. Başar, "An Equilibrium Theory for Multi-Person Decision Making with Multiple Probabilistic Models: Part II. Asymmetric Mode of Decision Making," Proc. 22nd IEEE Conference on Decision and Control, San Antonio, TX, 1983.
- [B4] T. Başar, "An Equilibrium Theory for Multi-Person Decision Making with Multiple Probabilistic Models," IEEE Transactions on Automatic Control, February 1985 (to appear as a regular paper).
- [B5] T. Başar, "Stochastic Multimodeling in a Game-Theoretic Framework with Applications in Economics," invited contribution to the 2nd Viennese Workshop on Economic Applications of Control Theory, volume including the paper to be published by Springer-Verlag in 1984, (manuscript still under preparation).

C. *Incentive Decision Problems with Multiple Stages*

- [C1] Y. P. Zheng, T. Başar and J. B. Cruz, Jr., "Incentive Stackelberg Strategies for Deterministic Multi-Stage Decision Processes," IEEE Transactions on Systems, Man and Cybernetics, January 1984. This has also been presented as an invited paper at the 21st CDC, Orlando, FL, December 8-10, 1982, and has appeared in the Conference Proceedings.

D. *Continuous-Time Decision Problems and Feedback Stackelberg Equilibria*

- [D1] T. Başar and A. Haurie, "Feedback Equilibria in Differential Games with Structural and Modal Uncertainties," in Advances in Large Scale Systems, (J. B. Cruz, Jr., ed.), vol. 1, JAI Press, Inc., Connecticut, 1984.
- [D2] T. Başar and A. Haurie, "Feedback Stackelberg Solution in Continuous-Time Multi-Level Games," Proc. 21st CDC, Orlando, FL, December 8-10, 1982.

E. *Equilibria in Dynamic Games*

- [E1] T. Başar, "Informational Uniqueness of Closed-Loop Nash Equilibria for a Class of Nonstandard Dynamic Games," J. Optimization Theory and Applications, 1985 (to appear).
- [E2] T. Başar, "Theory of Dynamic Games and Its Applications in Large Scale Systems Design and Optimization," in Proc. IFAC World Congress, July 1984, Budapest, Hungary (to appear).

A minimum sensitivity approach to incentive design problems*

Derya H. CANSEVER and Tamer BAŞAR

Decision and Control Laboratory and Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801, U.S.A.

Received 2 December 1982

Revised 1 May 1983

Communicated by A.P. Sage

In this paper we introduce the notion of 'robust' incentive schemes in multi-agent decision problems with a hierarchical decision structure, and discuss the derivation of such policies by minimizing, in addition to the usual (standard) Stackelberg performance indices, an appropriate sensitivity function. Such an approach has applications in decision problems wherein the leader does not know the exact values of some parameters characterizing the follower's cost functional, and seeks to robustify his optimum policy in the presence of deviations from the nominal values. An in-depth analysis of such incentive design problems is provided, and optimum robust incentive schemes are derived for general cost functionals with a convex structure. The results are then applied to an incentive design problem arising in economics, leading to some meaningful robust incentive policies.

1. Introduction

Optimum incentive design problems constitute a promising and mathematically challenging class of decision problems in economics and operations research [6-8], and have also recently attracted the attention of control theorists [2-5, 9-11] because of the close relationship with Stackelberg games [1, 4, 12]. Viewed as a dynamic Stackelberg game, an optimum design problem involves a hierarchy in decision-making and an information structure that allows the decision-maker at the top of the hierarchy (to be called the leader) to acquire (perfect or partial) information on the actions of the other decision-maker(s) (the so-called follower(s)). This available information enables the leader to design a policy (called an incentive scheme) that forces the follower(s) to a desired (from the leader's point of view) behavior. The utmost goal sought by the leader may be described in the form of maximization of a utility function (or minimization of a cost function), or it may be some set of Utopic points in an appropriate space, determined according to some criterion. The incentive scheme may be in the form of a threat strategy [9] or some smooth policy with appealing regularity conditions, as discussed in [5]. This latter reference, in particular, shows that in the case of two decision-maker problems, and when the follower's cost function is strictly convex, there exists an optimal affine strategy for the leader, which is affine in the dynamic information. This, however, does not imply that the solution will be unique; in fact, there will exist, in general, a multitude of optimal incentive schemes.

Existence of multiple solutions to incentive design problems prompts a further design criterion according to which a further selection can be made. In this paper we introduce such a selection criterion under which *robust* incentive policies can be obtained for two-agent deterministic problems with strictly

* This work was supported in part by the Office of Naval Research under Contract N00014-82-K-0469, in part by the Joint Services Electronics Program under Contract N00014-79-C-0424, and in part by the Electric Energy Systems Division, Department of Energy, under Contract DE-AC01-81RA-50658 with Dynamic Systems, P.O. Box 423, Urbana, IL 61801. A preliminary version of this paper was presented at the 21st IEEE Conference on Decision and Control, Orlando, FL, December 1982.

convex cost functionals. More specifically, we assume that the leader does not know the exact value of a parameter that characterizes the follower's cost function, and thereby his optimal response function. He may take a nominal value for this unknown parameter (assuming that such a candidate exists) and seek an optimal incentive scheme that is not only optimal for that chosen value but is also least sensitive to deviations from the nominal. We call such incentive policies *robust*, and develop in this paper a new approach for the derivation of such policies by introducing an appropriate sensitivity function.

The problem is formulated in precise mathematical terms in Section 2. Section 3 introduces sensitivity functions relevant to the context and derives robust affine policies for a general class of problems with convex cost functionals. A special class of problems with separable and singular cost functions are treated in Section 4, leading to some explicit, analytic solutions. Section 5 deals with an extension of these results to a larger class of incentive schemes which also include nonlinear policies. In Section 6, the results of Sections 3 and 4 are applied to a problem arising in microeconomics, and it is shown that the unique robust affine incentive policy for the leader bears a very meaningful economic interpretation. The concluding remarks of Section 7 end the paper.

2. Problem formulation

Consider a two-person deterministic dynamic game in normal form, described by the cost functionals $J_1(\gamma_1, \gamma_2)$ and $J_2(\gamma_1, \gamma_2, \alpha)$ for player 1 (the leader) and player 2 (the follower), respectively. Here, the strategies γ_1 and γ_2 belong to a priori specified strategy spaces Γ_1 and Γ_2 , respectively, and $\alpha \in \mathbb{A} \subset \mathbb{R}$ denotes a parameter on which the follower's cost functional depends. Let us denote the decision variables of the leader and the follower by $u \in U \equiv \mathbb{R}^n$ and $v \in V \equiv \mathbb{R}^m$, respectively. In this paper, we will assume that the follower has open-loop information, and hence take $\Gamma_2 \equiv V$. By an abuse of notation, we also let $J_1(u, v)$ and $J_2(u, v, \alpha)$ denote the cost functionals over the product space $U \times V$, for each $\alpha \in \mathbb{A}$. Let us further assume that:

- (i) $J_2(u, v, \alpha)$ is strictly convex¹ and twice continuously differentiable on $U \times V$, with $\alpha \in \mathbb{A}$;
- (ii) the leader has perfect access to the follower's action v ; and
- (iii) the leader is uncertain about the actual value of $\alpha \in \mathbb{A}$; however, he designs his strategy according to a nominal value of α , say $\alpha^* \in \mathbb{A}$, keeping in mind that α^* may not be the actual value.

Under this setup, the problem faced by the leader is twofold:

- (1) To design a Stackelberg strategy $\gamma_1^* \in \Gamma_1$ which, by also taking into account rational reactions of the follower, leads to a desired value of J_1 , which may be its global minimum over $U \times V$; let us denote such a value by J_1^* . More precisely, one of the objectives of the leader is to find a $\gamma_1^* \in \Gamma_1$ such that

$$J_1(\gamma_1^*, v) = J_1^*, \quad \text{for all } v \in R_\alpha(\gamma_1^*), \quad (1a)$$

where $R_\alpha(\gamma_1)$ denotes the optimal reaction set of the follower and is defined for each fixed $\alpha \in \mathbb{A}$, by

$$R_\alpha(\gamma_1) = \{v^* \in V : J_2(\gamma_1, v^*, \alpha) \leq J_2(\gamma_1, v, \alpha), \forall v \in V\}. \quad (1b)$$

Under an additional technical restriction which will be delineated in the sequel, it has been shown in [5] that, for each fixed $\alpha \in \mathbb{A}$, there exists an affine strategy for the leader which leads to J_1^* . More precisely, with α fixed at a nominal value α^* , there exists an $n \times m$ matrix $Q(\alpha^*)$, satisfying the relation:²

¹ This restriction will be relaxed later in Section 4.

² All partial derivatives of a scalar with respect to a vector are taken to be row vectors, whereas all other vectors are column vectors – a convention that we adopt throughout this paper.

$$\left. \frac{\partial J_2}{\partial u} Q(\alpha^*) + \frac{\partial J_2}{\partial v} \right|_{u=u^*, v=v^*} = 0, \quad (2)$$

so that the incentive strategy,

$$u = \gamma_1(v) = u^* + Q(\alpha^*)(v - v^*), \quad (3)$$

forces the follower to $v = v^*$, provided that $\partial J_2(u, v, \alpha)/\partial u$ evaluated at $(u = u^*, v = v^*)$ does not vanish within an ε -neighborhood of α^* , where (u^*, v^*) minimizes $J_1(u, v)$ over $U \times V$.

In general, equation (2) defines a class J of $(n \times m)$ matrices which force the follower to $v = v^*$. We also note that there exist other strategies for the leader, which attain J_1^* , and they need not be of the affine form as in (3).

(2) Since the leader does not have perfect knowledge of α , the actual value may be different from α^* in which case the leader will end up with an inferior performance, because γ_1^* so defined is optimal only at $\alpha = \alpha^*$. Therefore, it is highly desirable for the leader to have a robustness property associated with his optimal strategy. More precisely, the leader would like to have the sensitivity of the realized value of his cost functional, against the variations in α about its nominal value α^* , be as small as possible. This property may be induced by making use of the intrinsic nonuniqueness of the solutions of (2), and also by introducing nonlinear strategies which satisfy (1a), as will be shown in the sections to follow.

3. Introduction of a sensitivity function and derivation of robust solutions

Let $\gamma_1 \in \Gamma_1$ be an incentive strategy for the leader, and $v_a \in R_a(\gamma_1)$. Towards the goal set in Section 2, and as a measure of the sensitivity of $J_1(\gamma_1, v_a)$ with respect to deviations of α from its nominal value α^* , let us introduce the total derivative of $J_1(\gamma_1, v_a)$, with respect to α , evaluated at $\alpha = \alpha^*$, $u = u^*$ and $v = v^*$; more precisely, let us confine ourselves to affine strategies of the form (3) and, by abuse of terminology, let us define the *first-order sensitivity function* of $J_1(\gamma_1, v_a)$ with respect to α , and at $\alpha = \alpha^*$ as

$$I_1(\alpha^*) \triangleq \frac{dJ_1(\gamma_1, v_a)}{d\alpha} = \left(\frac{\partial J_1(u, v)}{\partial u} Q(\alpha) + \frac{\partial J_1(u, v)}{\partial v} \right) \left(\frac{dv}{d\alpha} \right) \bigg|_{\substack{\alpha=\alpha^* \\ u=u^* \\ v=v^*}}. \quad (4)$$

However, since the pair (u^*, v^*) globally minimizes J_1 on the product space $U \times V$, $I_1(\alpha^*)$ vanishes. Then, the next term in the Taylor expansion of $J_1(\gamma_1, v_a)$ around α^* , which we call the *second-order sensitivity function* of $J_1(\gamma_1, v_a)$ with respect to α , will have to be considered. Denoting it by $I_2(\alpha^*)$, and suppressing the arguments, we have:

$$\begin{aligned} I_2(\alpha^*) &\triangleq \frac{d^2 J_1(\gamma_1, v_a)}{d\alpha^2} \bigg|_{\substack{\alpha=\alpha^* \\ u=u^* \\ v=v^*}} \\ &= \left(\frac{dv}{d\alpha} \right) \left(Q' \frac{\partial^2 J_1}{\partial u^2} Q + Q' \frac{\partial^2 J_1}{\partial v \partial u} + \frac{\partial^2 J_1}{\partial u \partial v} Q + \frac{\partial^2 J_1}{\partial v^2} \right) \left(\frac{dv}{d\alpha} \right) \bigg|_{\substack{\alpha=\alpha^* \\ u=u^* \\ v=v^*}}. \end{aligned} \quad (5)$$

In order to find an expression for $dv/d\alpha$, we note that the equation,

$$\left. \frac{\partial J_2}{\partial u} Q(\alpha^*) + \frac{\partial J_2}{\partial v} \right|_{\substack{u=\gamma_1(v_a) \\ v=v_a}} = 0, \quad (6)$$

is an identity for all $\alpha \in \mathcal{A}$, and it completely specifies the optimal response v_α of the follower when affine strategies γ_1 of the form (3) are used. The derivative of the above expression with respect to α would still vanish for all $\alpha \in \mathcal{A}$. We then have, with $Q \equiv Q(\alpha^*)$:

$$\left[Q' \frac{\partial^2 J_2}{\partial u^2} Q + Q' \frac{\partial^2 J_2}{\partial v \partial u} + \frac{\partial^2 J_2}{\partial u \partial v} Q + \frac{\partial^2 J_2}{\partial v^2} \right] \left(\frac{dv}{d\alpha} \right) + \frac{\partial^2 J_2}{\partial \alpha \partial u} Q + \frac{\partial^2 J_2}{\partial \alpha \partial v} = 0, \quad \forall \alpha \in \mathcal{A}. \quad (7)$$

Since $J_2(u, v, \alpha)$ is strictly convex in u and v for all $\alpha \in \mathcal{A}$, the coefficient matrix of $(dv/d\alpha)$ is positive definite, and thereby invertible.¹ Combining (5) and (7) we obtain the second-order sensitivity function of $J_1(\gamma_1, v_\alpha)$ with respect to α , at $\alpha = \alpha^*$, to be:

$$\begin{aligned} I_2(\alpha^*) = & \left[\left(Q' \frac{\partial^2 J_2}{\partial u^2} Q + Q' \frac{\partial^2 J_2}{\partial v \partial u} + \frac{\partial^2 J_2}{\partial u \partial v} Q + \frac{\partial^2 J_2}{\partial v^2} \right)^{-1} \left(\frac{\partial^2 J_2}{\partial \alpha \partial u} Q + \frac{\partial^2 J_2}{\partial \alpha \partial v} \right) \right] \\ & \times \left[Q' \frac{\partial^2 J_1}{\partial u^2} Q + Q' \frac{\partial^2 J_1}{\partial v \partial u} + \frac{\partial^2 J_1}{\partial u \partial v} Q + \frac{\partial^2 J_1}{\partial v^2} \right] \\ & \times \left[\left(Q' \frac{\partial^2 J_2}{\partial u^2} Q + Q' \frac{\partial^2 J_2}{\partial v \partial u} + \frac{\partial^2 J_2}{\partial u \partial v} Q + \frac{\partial^2 J_2}{\partial v^2} \right)^{-1} \left(\frac{\partial^2 J_2}{\partial \alpha \partial u} Q + \frac{\partial^2 J_2}{\partial \alpha \partial v} \right) \right] \bigg|_{\substack{\alpha = \alpha^* \\ u = u^* \\ v = v^*}} \end{aligned} \quad (8)$$

Remark 1. When the leader enforces his team solution, his objective will be to minimize $I_2(\alpha^*)$ over $\mathcal{A}$, since $I_1(\alpha^*)$ vanishes. On the other hand, there may be cases when he would prefer to enforce a point other than (u^*, v^*) , in which case $I_1(\alpha^*)$ will not necessarily be zero. However, when the nonvanishing $I_1(\alpha^*)$ does not depend on the choice of Q from the class $\mathcal{Q}$ of $n \times m$ matrices which, together with an affine strategy of the form (2), enforce the follower to the desired point, one still has to consider $I_2(\alpha^*)$ as a measure of obtaining minimum sensitive solutions. This point will be further elucidated in Section 6. $\square$

The problem now is to minimize $I_2(\alpha^*)$ over all $(n \times m)$ real matrices $Q(\alpha^*)$ subject to the constraint:

$$\frac{\partial J_2}{\partial u} Q(\alpha^*) + \frac{\partial J_2}{\partial v} \bigg|_{\substack{\alpha = \alpha^* \\ u = u^* \\ v = v^*}} = 0. \quad (9)$$

We first observe that $I_2(\alpha^*) \geq 0$, and hence a lower bound for $I_2(\alpha^*)$ is zero. We will now show that this lower bound is tight for a fairly large class of problems. Towards this end, let us note that this lower bound is reached when

$$\frac{\partial^2 J_2}{\partial \alpha \partial u} Q(\alpha^*) + \frac{\partial^2 J_2}{\partial \alpha \partial v} \bigg|_{\substack{\alpha = \alpha^* \\ u = u^* \\ v = v^*}} = 0, \quad (10)$$

subject to (9). These two equations may be combined and written as:

$$\left(\frac{\partial^2 J_2 / \partial \alpha \partial u}{\partial J_2 / \partial u} \right) Q(\alpha^*) + \left(\frac{\partial^2 J_2 / \partial \alpha \partial v}{\partial J_2 / \partial v} \right) \bigg|_{\substack{\alpha = \alpha^* \\ u = u^* \\ v = v^*}} = 0. \quad (11)$$

¹ For invertibility it is, of course, sufficient that the Hessian matrix of J_2 be full rank

Let q_i , $i = 1, 2, \dots, m$, denote the i th row of $Q(\alpha^*)$. Then, (11) represents a collection of m sets of linear algebraic equations, with each set consisting of two equations with n unknowns, which correspond to entries of q_i . Thus, in order to have at least one solution to (11), it is sufficient that:

$$\text{rank} \left(\frac{\partial^2 J_2 / \partial \alpha \partial u}{\partial J_2 / \partial u} \right) \bigg|_{\substack{\alpha = \alpha^* \\ u = u^* \\ v = v^*}} = 2. \quad (12)$$

For this requirement to be met, it is, of course, necessary that $\dim(u) \geq 2$. Suppose (12) is satisfied, and let $Q^0(\alpha^*)$ denote a solution to (11).⁴ Then, an affine strategy given by (3), with $Q(\alpha^*) = Q^0(\alpha^*)$, makes the term $dv/d\alpha$ defined by (7) vanish at the nominal solution point. By this token, the sensitivity function of order 3, i.e.

$$\frac{d^3 J_1(\gamma_1(v_\alpha), v_\alpha)}{d\alpha^3} \bigg|_{\substack{\alpha = \alpha^* \\ u = u^* \\ v = v^*}}$$

is annihilated at the nominal solution point, since it carries (by the chain rule) the product term $dv/d\alpha$. In other words, the third-order Taylor approximation of the effect of a perturbation in the value of α on $J_1(\gamma_1(v_\alpha), v_\alpha)$ is zero within an ε -neighborhood of the nominal solution point. Therefore, the affine strategy,

$$\gamma_1^0(v) = u^* + Q^0(\alpha^*)(v - v^*), \quad (13)$$

where $Q^0(\alpha^*)$ satisfies (11), has very appealing sensitivity properties.

In the preceding discussion, we confined ourselves to a scalar parameter α . We now relax this condition, by letting $\alpha \in \mathbb{R}^r$. In this case $dv/d\alpha$ becomes an $(r \times m)$ matrix. If all the entries of $dv/d\alpha$ vanish at the nominal solution point, so will the entries of the sensitivity functions

$$\frac{d^p J_1(\gamma(v_\alpha), v_\alpha)}{d\alpha^p} \bigg|_{\substack{\alpha = \alpha^* \\ u = u^* \\ v = v^*}}, \quad p = 1, 2, 3.$$

This is the case if $Q(\alpha^*)$ satisfies (11) with $\partial^2 J_2 / \partial \alpha \partial u$ and $\partial^2 J_2 / \partial \alpha \partial v$ being $(r \times n)$ and $(r \times m)$ matrices, respectively. A sufficient condition for this (replacing (12)) is

$$\text{rank} \left(\frac{\partial^2 J_2 / \partial \alpha \partial u}{\partial J_2 / \partial u} \right) \bigg|_{\substack{\alpha = \alpha^* \\ u = u^* \\ v = v^*}} = r + 1. \quad (14)$$

We now summarize these results in the following proposition.

Proposition 1. Let (14) be satisfied. Then (i) there exists at least one $Q(\alpha^*)$ (denoted $Q^0(\alpha^*)$) satisfying (11), and (ii) the affine strategy,

$$\gamma_1(v) = u^* + Q^0(\alpha^*)(v - v^*),$$

induces the follower to play $v = v^*$ when $\alpha = \alpha^*$, and makes the first-, second- and third-order derivatives of $J_1(\gamma_1(v_\alpha), v_\alpha)$ with respect to the r -vector α vanish at the solution point. $\square$

⁴ When (11) admits more than one solution, each of them fulfills the robustness criterion set above. In this case, one can pursue further analysis to minimize higher order sensitivity functionals within the class of these robust solutions.

4. Singular incentive problems

In the previous section the least sensitive incentive design problem was solved by choosing $Q(\alpha^*)$ such that $dv/d\alpha$ is annihilated at the nominal solution point. The approach adopted in that section fails to be applicable when the cost functional of the follower is separable, as

$$J_2(u, v, \alpha) = g_1(u, v) + g_2(v, \alpha), \quad (15)$$

in which case the term $\partial^2 J_2 / \partial \alpha \partial u$ vanishes. A partial remedy for this class of problems can be worked out by choosing the entries of $Q(\alpha^*)$ as large as possible, subject to (9) and some self-imposed bounds [13]. Among this class, however, the so-called singular incentive problems are of particular interest. On the one hand, they lead to analytical inner point solutions which are quite appealing (see below the example of Section 6); on the other hand, they provide a convenient setting for relaxing the assumption on the strict convexity of J_2 on the product space $U \times V$.

Towards this goal, let us assume that the leader's control affects the follower's cost functional linearly. By analogy with singular control, this class of problems will be referred to as singular incentive problems. In a singular incentive problem, the follower's cost functional is not convex on the product space $U \times V$, and therefore the previous theory on linear incentive problems is not directly applicable; however, an extension is possible, as we elucidate in the sequel. First note that when the leader announces an affine strategy of the form (3), the follower's cost functional J_2 becomes a function of only v , for a given $\alpha \in \mathbb{A}$. Accordingly, if the functional $J_2(u' + Q(\alpha^*)(v - v'), v, \alpha)$ is strictly convex on V for each $\alpha \in \mathbb{A}$, and if equation (2) admits a nonempty family $\mathcal{J}$ of solutions at the desired point (u', v') , then the least-sensitive incentive design problem becomes meaningful for this particular class of cost functionals. More precisely, for a given $u' \in U$, let

$$\frac{d^2 J_2(u, v, \alpha)}{dv^2} = \frac{\partial^2 J_2(u, v, \alpha)}{\partial v^2} + \frac{\partial^2 J_2(u, v, \alpha)}{\partial v \partial u} Q(\alpha^*) + Q'(\alpha^*) \frac{\partial^2 J_2(u, v, \alpha)}{\partial v \partial u} > 0, \quad (16)$$

for all $v \in V$, $\alpha \in \mathbb{A}$ and $Q(\alpha^*) \in \mathcal{Q} \subseteq \mathcal{Q}$, where $\mathcal{Q}$ is assumed to be nonvoid. Under these assumptions, the optimal reaction of the follower to any announced strategy of the form (3), where $Q(\alpha^*) \in \mathcal{Q}$, is $v = v'$, which is confirmed by (2) as a first-order necessary condition, and (16) provides a second-order sufficient condition for the optimality of $v = v'$.

As we indicated above, for this class of incentive design problems it is possible to obtain an inner-point solution analytically because of the particular structure of the sensitivity function. Towards this end, let us assume that $n = 2$, $m = 1$, and let q_1 and q_2 denote the entries of $Q(\alpha^*)$. Using (9), $I_2(\alpha^*)$ becomes:

$$\begin{aligned} I_2(\alpha^*) = & \left[\frac{\partial^2 J_1}{\partial u_1^2} (\delta q_2 + \xi)^2 - 2 \frac{\partial^2 J_1}{\partial u_1 \partial v} [\delta q_2 + \xi] + 2 \frac{\partial^2 J_1}{\partial u_2 \partial v} q_2 - 2 \frac{\partial^2 J_1}{\partial u_1 \partial u_2} q_2 (\delta q_2 + \xi) \right. \\ & \left. + \frac{\partial^2 J_1}{\partial u_2^2} q_2^2 + \frac{\partial^2 J_1}{\partial v^2} \right] \left[\frac{\partial^2 J_2}{\partial \alpha \partial v} / \left(\frac{\partial^2 J_2}{\partial u_2 \partial v} q_2 - \frac{\partial^2 J_2}{\partial u_1 \partial v} (\delta q_2 + \xi) + \frac{\partial^2 J_2}{\partial v^2} \right) \right]^2 \bigg|_{\substack{\alpha = \alpha^* \\ u = u' \\ v = v'}} \end{aligned} \quad (17)$$

where δ and ξ are defined by:

$$\delta \triangleq \frac{\partial J_2}{\partial u_2} / \frac{\partial J_2}{\partial u_1} \bigg|_{\substack{\alpha = \alpha^* \\ u = u' \\ v = v'}}, \quad \xi \triangleq \frac{\partial J_2}{\partial v} / \frac{\partial J_2}{\partial u_1} \bigg|_{\substack{\alpha = \alpha^* \\ u = u' \\ v = v'}}. \quad (18)$$

The derivative of $I_2(\alpha^*)$ with respect to q_2 vanishes at

$$q_2^* = \left\{ \left[\xi^2 \frac{\partial^2 J_1}{\partial u_1^2} - 2\xi \frac{\partial^2 J_1}{\partial u_1 \partial v} + \frac{\partial^2 J_1}{\partial v^2} \right] \left[\frac{\partial^2 J_2}{\partial u_2 \partial v} - \delta \frac{\partial^2 J_2}{\partial u_1 \partial v} \right] - \left[\delta \xi \frac{\partial^2 J_1}{\partial u_1^2} - \delta \frac{\partial^2 J_1}{\partial u_1 \partial v} + \frac{\partial^2 J_1}{\partial u_2 \partial v} - \xi \frac{\partial^2 J_1}{\partial u_1 \partial u_2} \right] \left[\frac{\partial^2 J_2}{\partial v^2} - \xi \frac{\partial^2 J_2}{\partial u_1 \partial v} \right] \right\} / \left\{ \left[\delta^2 \frac{\partial^2 J_1}{\partial u_1^2} - 2\delta \frac{\partial^2 J_1}{\partial u_1 \partial u_2} + \frac{\partial^2 J_1}{\partial u_2^2} \right] \left[\frac{\partial^2 J_2}{\partial v^2} - \xi \frac{\partial^2 J_2}{\partial u_1 \partial v} \right] - \left[\delta \xi \frac{\partial^2 J_1}{\partial u_1^2} - \delta \frac{\partial^2 J_1}{\partial u_1 \partial v} + \frac{\partial^2 J_1}{\partial u_2 \partial v} - \xi \frac{\partial^2 J_1}{\partial u_1 \partial u_2} \right] \left[\frac{\partial^2 J_2}{\partial u_2 \partial v} - \delta \frac{\partial^2 J_2}{\partial u_1 \partial v} \right] \right\} \Bigg|_{\substack{\alpha = \alpha^* \\ u = u^* \\ v = v^*}} \quad (19)$$

provided that the denominator of (19) is not zero. We then readily have the following proposition.

Proposition 2. *The minimum-sensitive (robust) linear incentive design problem formulated in this section admits a unique solution (q_1^*, q_2^*) , where q_2^* is given by (19) and q_1^* is obtained through the linear constraint (9) which relates q_1 to q_2 , provided that the denominator of (19) does not vanish and $(q_1^*, q_2^*) \in \mathcal{Q}$.*

Proof. This result follows from the following four properties of the function $F(q_2) = I_2(\alpha^*)$:

(i) There exists a finite number M such that:

$$\lim_{q_2 \rightarrow -\infty} F(q_2) = \lim_{q_2 \rightarrow +\infty} F(q_2) = M.$$

(ii) $F(q_2) \rightarrow +\infty$ as:

$$q_2 \rightarrow \left(\frac{\partial^2 J_2}{\partial u_1 \partial v} \xi - \frac{\partial^2 J_2}{\partial v^2} \right) / \left[\frac{\partial^2 J_2}{\partial u_2 \partial v} - \frac{\partial^2 J_2}{\partial u_1 \partial v} \delta \right] \triangleq \bar{q}_2.$$

(iii) $F(q_2)$ is continuous, except at $q_2 = \bar{q}_2$.

(iv) $F(q_2)$ has a single stationary point q_2^* .

These four properties readily lead to the conclusion that q_2^* is the unique minimizing solution for F . $\square$

5. Use of nonlinear strategies in sensitivity considerations

The analysis of previous sections was confined to the class of linear strategies. Although this class is rich enough to provide optimal least sensitive solutions to incentive design problems, use of nonlinear strategies may provide additional degrees of freedom, especially when $Q(\alpha^*)$ is determined uniquely through (9). Specifically, when $\dim u = 1$ the constraint (9) determines $Q(\alpha^*)$ uniquely, and hence the set $\mathcal{Q}$ is a singleton and it does not allow sensitivity considerations. However, if the leader is permitted to enlarge his strategy space by including a suitable nonlinear term in his control, he may have extra degrees of freedom to reduce the sensitivity of his performance to changes in the uncertain parameter in the follower's cost functional. Towards this end, let us assume that the strategy space of the leader is the set of all mappings from V onto U , and twice continuously differentiable at the nominal solution point. For $n = m = 1$, $I_2(\alpha^*)$ becomes:

$$I_2(\alpha^*) = \left| \left(\frac{\partial^2 J_2}{\partial u^2} \left(\frac{\partial \gamma_1}{\partial v} \right)^2 + 2 \frac{\partial^2 J_1}{\partial u \partial v} \frac{\partial \gamma_1}{\partial v} + \frac{\partial^2 J_1}{\partial v^2} \right) \right. \\ \left. \times \left(\frac{\partial^2 J_2}{\partial u \partial \alpha} \frac{\partial \gamma_1}{\partial v} + \frac{\partial^2 J_2}{\partial u \partial \alpha} \right) / \left(\frac{\partial^2 J_2}{\partial u^2} \left(\frac{\partial \gamma_1}{\partial v} \right)^2 + 2 \frac{\partial^2 J_2}{\partial u \partial v} \frac{\partial \gamma_1}{\partial v} + \frac{\partial J_2}{\partial u} \frac{\partial^2 \gamma_1}{\partial v^2} + \frac{\partial^2 J_2}{\partial v^2} \right)^2 \right|_{\substack{\alpha = \alpha^* \\ u = u' \\ v = v'}} \quad (20a)$$

Here, $\partial \gamma_1 / \partial v$ is completely determined from:

$$\left. \frac{\partial J_2}{\partial u} \frac{\partial \gamma_1}{\partial v} + \frac{\partial J_2}{\partial v} \right|_{\substack{\alpha = \alpha^* \\ u = u' \\ v = v'}} = 0. \quad (20b)$$

The optimization of $I_2(\alpha^*)$ may require $\partial^2 \gamma_1 / \partial v^2$ to take arbitrarily large values. Such a strategy may give rise to arbitrarily large values of u for finite values of v . Since the affordability and credibility of such a strategy is questionable, it is necessary to impose bounds on $\partial^2 \gamma_1 / \partial v^2$ such as

$$\left| \frac{\partial^2 \gamma_1}{\partial v^2} \right| \leq k, \quad k \in \mathbb{R}^+. \quad (21)$$

Under such a constraint, the minimizing argument of $I_2(\alpha^*)$ is given by:

$$\frac{\partial^2 \gamma_1}{\partial v^2} = \text{sgn} \left(\frac{\partial J_2}{\partial u} \right) k, \quad (22)$$

where

$$\text{sgn}(x) = \begin{cases} +1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0. \end{cases} \quad (23)$$

Here, all the partials are evaluated at the nominal point. These results then lead to the following proposition.

Proposition 3. *A representation of the least sensitive incentive strategy within the class of policies which are twice continuously differentiable at the nominal solution point is given by:*

$$\gamma_1^*(v) = u' - \left[\frac{\partial J_2}{\partial v} / \frac{\partial J_2}{\partial u} \right] (v - v') + \frac{1}{2} \text{sgn} \left(\frac{\partial J_2}{\partial u} \right) k (v - v')^2, \quad (24)$$

where all the partials are evaluated at the nominal solution point. $\square$

6. An example from economics

In this section we discuss an example from microeconomics, which illustrates some of the results obtained in this paper, especially those on singular incentive problems. Let us consider a duopolistic market where the leader, P1, produces two goods, X and Y, and the follower produces a single good, Z. All three goods are substitutable and they are sold at the same price p which is assumed to satisfy the linear demand relation:

$$x + y + z = d_0 - d_1 p, \quad (25)$$

where x , y and z represent the quantities to be produced from each good X , Y and Z , respectively; d_0 and d_1 are positive constants. It is also assumed that each firm has a general positive, increasing and twice continuously differentiable cost function $l_i(\cdot)$.

Then, the profit functions of the firms, which are to be maximized, become:

$$\pi_1 = \left(\frac{d_0 - (x + y + z)}{d_1} \right) (x + y) - l_1(x, y), \quad (26a)$$

$$x, y, z \geq 0,$$

$$\pi_2 = \left(\frac{d_0 - (x + y + z)}{d_1} \right) z - l_2(c, z), \quad (26b)$$

where c is a positive parameter, whose role will be clarified in the sequel. Furthermore, we let $\gamma_1(z)$ denote a two-dimensional policy vector of firm 1, i.e. $(x, y)' = \gamma_1(z)$, whereas firm 2's policy is the static vector z .

For this problem it can be shown that there exists a sequence of closed-loop policies for P1 which force the follower to $z = 0$ [1] – a result which is valid for any demand relation in which the price is a strictly decreasing function of z . However, for the leader, the limiting strategy is not well defined since it requires the gain vector $Q(\alpha^*)$ to have unbounded elements (i.e. infinite threat). In general, such a strategy is neither credible for the follower, nor affordable to the leader. In its stead, an alternative is to assume that P1 imposes a suitable point of equilibrium, compatible with the duopolistic nature of the problem.

Towards this end, let the objective functional considered by the leader be

$$J_1 = \nu \pi_1 + (1 - \nu) \pi_2, \quad 0 < \nu < 1, \quad (27)$$

where ν is large enough to provide the leader better profit, but not so large to lead to a noncredible incentive scheme. A possible upper bound for ν may be the one which provides the follower a profit comparable with what he would make in a Nash equilibrium case. The objective functional of the follower is still π_2 . Let us assume, for simplicity in the analysis to follow, that $\nu = \frac{1}{2}$, in which case the leader can roughly guarantee two-thirds of the market.⁵ To complete the formulation of the problem, let us assume that the parameter in the cost of production of Z , namely c , is uncertain to the leader. However, he knows a nominal value of c , say c^* , around which c may vary.⁶ The goal of the leader is to design a strategy, using his strategic variables x and y , which will enforce the follower to the maximizing arguments of J_1 when $c = c^*$; and in addition he seeks a strategy under which his profit function is least sensitive to variations in c . Let us also assume that the leader confines himself to affine strategies. This problem is within the scope of singular incentive problems discussed in Section 4, with, however, two exceptions. One of them is that the desired solution point is not the team solution of the leader's profit function π_1 , for which the least-sensitivity property is sought. In this case, the first-order sensitivity function $I_1(\alpha^*)$ does not vanish. Nevertheless, it is a straightforward task to show that $I_1(\alpha^*)$ is invariant under the choice of $Q(\alpha^*) \in \mathcal{Z}$. Hence, the quantity we seek to minimize is still $I_2(\alpha^*)$.

The other distinct feature of this example is that the decisions of the agents are restricted to be non-negative, in contrast with the previous theory which requires the decision spaces to be appropriate dimensional Euclidian spaces. However, it will be clear from the analysis in the sequel that for the problem treated in this section, the above restriction does not invalidate the use of the previous theory.

⁵ This ratio of course depends on the relative structures and magnitudes of the cost functions l_1 and l_2 . When both firms have the same kind of cost function, this ratio will be exactly two-thirds.

⁶ Note that c plays the role of α introduced in the previous sections.

Let us assume that $l_1(x, y)$ and $l_2(z, c^*)$ in (26) are such that J_1 is strictly concave on $X \times Y \times Z$, and let (x^*, y^*, z^*) denote the triplet which globally maximizes J_1 .

Although we assumed that the leader should be satisfied with the profit corresponding to the product level (x^*, y^*, z^*) , the optimal reaction of the follower to (x^*, y^*) is of course different from z^* , since his ultimate goal is to maximize π_2 , not $\pi_1 + \pi_2$, wherein z^* maximizes the latter given that $x = x^*$, $y = y^*$ and $c = c^*$. In order to enforce the desired triplet (x^*, y^*, z^*) , it is assumed that P1 announces an affine policy of the form:

$$\gamma_1(z) = [x^* + q_1(z - z^*), y^* + q_2(z - z^*)]' \quad (28)$$

where q_1 represents the coefficient which indicates how P1 would modify x if $z \neq z^*$, and similarly, q_2 relates the change in z to y . The affine strategy (36) will induce the follower to produce the amount z^* when q_1 and q_2 satisfy the constraint (2), and in this duopoly problem it takes the form:

$$q_1 + q_2 = (x^* + y^*)/z^* \quad (29)$$

When the uncertain parameter c takes its nominal value c^* , any pair (q_1, q_2) satisfying (29) will induce the desired outcome. We will explore this nonuniqueness by choosing the pair (q_1^*, q_2^*) which renders the profit π_1 of P1 least sensitive to variations in the uncertain parameter c about its nominal value c^* . To gain more insight into the problem, we will now assume a specific form for the cost functions $l_1(x, y)$ and $l_2(z, c)$. Towards this goal, we let

$$l_1 = \frac{1}{2}c_1x^2 + \frac{1}{2}c_2y^2; \quad l_2 = \frac{1}{2}cz^2 \quad (30a, 30b)$$

Here, $c > 0$ is the uncertain parameter, with a nominal value c^* , which determines the cost of producing z units of good Z . Likewise, c_1 and c_2 determine the cost the leader acquires when he produces x and y units of X and Y , respectively. When the uncertain parameter c takes its nominal value c^* , the affine strategy (28), with (q_1, q_2) satisfying

$$q_1 = q_2 = \frac{c_1c^* + c_2c^*}{c_1c_2} \quad (31)$$

induces the follower to choose $z = z^*$. Among this class, the pair (q_1^*, q_2^*) which renders the leader's profit least sensitive to variations in c around c^* is computed from (19) and (31) to be:

$$q_1^* = c^*/c_1; \quad q_2^* = c^*/c_2 \quad (32a, 32b)$$

Here, it is seen that in order to reach a robust incentive scheme, the leader should allocate the incentive among the goods he produces as inversely proportional to their respective costs of production.

In the preceding example, the decisions of the agents were restricted to belong to $\mathbb{R}^+$. Owing to the smoothness properties of the profit functions and incentive strategies, incremental variations in the value of c around c^* (which was the framework for the preceding analysis) cannot drive the equilibrium triplet (x^*, y^*, z^*) too far to violate this restriction. In fact, the extreme case occurs when c becomes too big to discourage P2 from producing anything. When his competitor leaves the market, the leader has, of course, no motivation to implement his cooperative strategy (28). For the sake of completeness, let us assume that the leader sticks to (28), with (q_1, q_2) satisfying (29). When $z = 0$, we have

$$x + y = 0$$

for all such pairs. When they satisfy the equation of the least sensitive pair (32), x and y individually vanish for $z = 0$; hence, the possibility for a negative x or y is avoided, even for this extreme case.

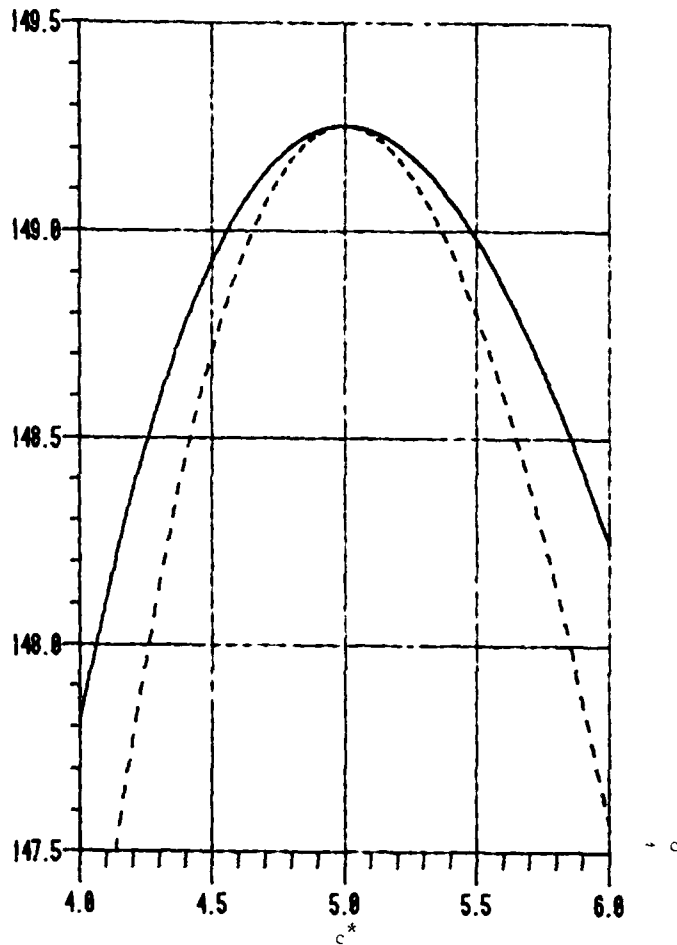


Fig. 1. A comparison of optimal profit functions of P1 under nonincremental variations in the value of the parameter c around its nominal value c^* .

As an illustration of the robustness properties of (32) the profit incurred by P1 is plotted in Fig. 1 against the possible values of the uncertain parameter c , for a given set of values of the parameters of the game. The solid line represents the profit made by P1 when the robust strategy (28), (32) is used. The dashed line stands for P1's profit when he allocates all the incentive to a single good. When the uncertain parameter assumes its nominal value $c^* = 5.0$, both incentive schemes yield the desired level of profit for P1. When c deviates from its nominal value, however, the profit value decreases, as is to be expected, but when the pair (q_1, q_2) is chosen according to (32), the incurred profit is less affected by such variations in c around c^* . Hence, the robustness property prevails not only for incremental variations around the nominal value, but also in a reasonably large neighborhood of the nominal.

7. Conclusion

In this paper we have introduced a minimum sensitivity approach towards the solution of deterministic incentive design problems, which leads to optimum robust incentive policies that are least

sensitive to the variations in the value of a parameter vector (from a nominal) characterizing the follower's cost function. Possible extensions of this minimum sensitivity approach are to the class of problems in which the leader has partial dynamic information (as in [5] and [9]) and to the class of stochastic incentive problems discussed in [11]. For an extension to the latter class of problems see [14].

References

- [1] T. Başar and H. Selbuz, Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems, *IEEE Trans. Automat. Control* AC-24 (1979) 166-179.
- [2] Y.C. Ho, P.B. Luh and G.J. Olsder, A control theoretic view on incentives, *Automatica* 18 (1982) 167-180.
- [3] M. Salman and J.B. Cruz, Jr., An incentive model of duopoly with government coordination, *Automatica* 17 (1981) 821-830.
- [4] T. Başar and J.B. Cruz, Jr., Concepts and methods in multi-person coordination and control, in: S.G. Tzafestas (Ed.), *Optimization and Control of Dynamic Operational Research Models* (North-Holland, Amsterdam, 1982) 351-387.
- [5] Y.P. Zheng and T. Başar, Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: A geometric approach, *Internat. J. Control* 35 (1982) 997-1011.
- [6] T. Groves and M. Loeb, Incentives in a divisionalized firm, *Management Sci.* 25 (1979) 221-230.
- [7] L. Hurwicz and L. Shapiro, Incentive structures maximizing residual gain under incomplete information, *Bell J. Econom.* 9 (1978) 180-191.
- [8] L.P. Jennergren, On the design of incentives in business firms - a survey of some recent research, *Management Sci.* 26 (7) (1980) 180-201.
- [9] T. Başar, A general theory for Stackelberg games with partial state information, *Large Scale Systems* 3 (1982) 47-56.
- [10] Y.C. Ho, P.B. Luh and R. Muralidharan, Information structure, Stackelberg games and incentive controllability, *IEEE Trans. Automat. Control* AC-26 (1981) 454-460.
- [11] T. Başar, Affine incentive schemes for stochastic systems with dynamic information, *Proc. of the 1982 American Control Conference*, Arlington, Virginia (June 1982) 127-132; also in *SIAM J. Control and Optimization* (March 1984).
- [12] B. Tolwinski, Closed-loop Stackelberg solution to multi-stage linear-quadratic game, *J. Optimization Theory Appl.* 34 (4) (1981) 485-501.
- [13] D.H. Cansever and T. Başar, A minimum sensitivity approach to incentive design problems, *Proc. of the 21st IEEE Conference on Decision and Control* (December 1982) 158-163.
- [14] T. Başar and D.H. Cansever, Robust incentive policies for stochastic decision problems in the presence of parametric uncertainty, to appear in *Proc. IX IFAC World Congress*, Budapest, Hungary (July 1984).

A MINIMUM SENSITIVITY APPROACH TO INCENTIVE DESIGN PROBLEMS*

Derya H. Cansever and Tamer Başar
Coordinated Science Laboratory
University of Illinois
Urbana, Illinois 61801

Abstract

In this paper we introduce the notion of "robust" incentive schemes in multi-agent decision problems with a hierarchical decision structure, and discuss derivation of such policies by minimizing, in addition to the usual (standard) Stackelberg performance indices, an appropriate sensitivity function. Such an approach has applications in decision problems wherein the leader does not know the exact values of some parameters characterizing the follower's cost functional, and seeks to robustify his optimum policy in the presence of deviations from the nominal values. An in-depth analysis of such incentive design problems is provided, and some concrete analytical results are obtained for general cost functionals with a convex structure. The results are then applied to an incentive design problem arising in economics, leading to some meaningful robust incentive policies.

I. Introduction

Optimum incentive design problems constitute a promising and mathematically challenging class of decision problems in economics and operations research ([6]-[8]), and have also recently attracted the attention of control theorists ([2]-[5], [9]-[11]) because of the close relationship with Stackelberg games ([1], [4], [12]). Viewed as a dynamic Stackelberg game, an optimum design problem involves a hierarchy in decision making and an information structure that allows the decision maker at the top of the hierarchy (to be called the leader) to acquire (perfect or partial) information on the actions of the other decision maker(s) (the so-called follower(s)). This available information enables the leader to design a policy (called an incentive scheme) that forces the follower(s) to a desired (from the leader's point of view) behavior. The utmost goal sought by the leader may be described in the form of maximization of a utility function (or minimization of a cost function), or it may be some set of utopic points in an appropriate space, determined according to some criterion. The incentive scheme may be in the form of a threat strategy [9] or some smooth policy with appealing regularity conditions, as discussed in [5]. This latter reference [5], in particular, shows that in the case of two decision-maker problems, and when the follower's cost function is strictly convex, there exists an optimal affine strategy for the leader, which is affine in the dynamic information. This, however, does not imply that the solution will be unique; in fact, there will exist, in general, a multitude of optimal incentive schemes.

Existence of multiple solutions to incentive design problems prompts a further design criterion according to which a further selection can be made. In this paper, we introduce such a selection criterion

*This work was supported in part by the Joint Services Electronics Program under Contract N00014-79-C-0001, and in part by the Dept. of Energy, Electric Energy Systems Division, under Contract DE-AC01-81RA-50654 with Dynamic Systems, POB #13, Urbana, IL 61801.

under which robust incentive policies can be obtained for two-agent deterministic problems with strictly convex cost functionals. More specifically, we assume that the leader does not know the exact value of a parameter that characterizes the follower's cost function, and thereby his optimal response function. He may take a nominal value for this unknown parameter (assuming that such a candidate exists) and seek an optimal incentive scheme that is not only optimal for that chosen value, but is also least sensitive to deviations from the nominal. We call such incentive policies robust, and develop in this paper a new approach for the derivation of such policies by introducing an appropriate sensitivity function.

The problem is formulated in precise mathematical terms in Section II. Section III introduces the first and second order sensitivity functions and discusses derivation of robust affine policies for a general class of problems with convex cost functionals. Some special classes of problems with separable and singular cost functions are treated in Section III leading to some explicit, analytic solutions. Section V deals with an extension of these results to a larger class of incentive schemes which also include nonlinear policies. In Section VI, the results of Sections III and IV are applied to a problem arising in micro-economics, and it is shown that the unique robust affine incentive policy for the leader bears a very meaningful economic interpretation. The concluding remarks of Section VII end the paper.

II. Problem Formulation

Consider a two-person deterministic dynamic game in normal form, described by the cost functionals $J_1(y_1, y_2)$ and $J_2(y_1, y_2, \lambda)$ for player 1 (the leader) and player 2 (the follower), respectively. Here, the strategies y_1 and y_2 belong to a priori specified strategy spaces Γ_1 and Γ_2 , respectively, and $\lambda \in \mathbb{A} \subset \mathbb{R}$ denotes a parameter on which the follower's cost functional depends. Let us denote the decision variables of the leader and of the follower by $u \in U$ and $v \in V$, respectively, where U and V are appropriate subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, and they represent the decision spaces of the leader and the follower, respectively. In this paper, we will assume that the follower has open-loop information, and hence take $V = \Gamma_2$. By an abuse of notation, we also let $J_1(u, v)$ and $J_2(u, v, \lambda)$ denote the cost functionals over the product space $U \times V$, for each $\lambda \in \mathbb{A}$. Let us further assume that

- (i) $J_2(u, v, \lambda)$ is strictly convex¹⁾ and twice continuously differentiable on $U \times V$, with $\lambda \in \mathbb{A}$.
- (ii) The leader has perfect access to the follower's action v .
- (iii) The leader is uncertain about the actual value of $\lambda \in \mathbb{A}$. However, he designs his strategy according to a nominal value of λ , say $\lambda^* \in \mathbb{A}$, keeping in mind that λ^* may not be the actual value.

¹⁾ This restriction will be relaxed later in §4.1.

Under this setup, the problem faced by the leader is twofold:

a) To design a Stackelberg strategy $\gamma_1^* \in \Gamma_1$ which, by also taking into account rational reactions of the follower, leads to a desired value of J_1 , which may be its global minimum over $U \times V$; let us denote such a value by J_1^* . More precisely, one of the objectives of the leader is to find a $\gamma_1^* \in \Gamma_1$ such that

$$J_1(\gamma_1^*, v) = J_1^* \text{ for all } v \in R_{\gamma_1^*}(\gamma_1^*), \quad (1a)$$

where $R_{\gamma_1^*}(\gamma_1^*)$ denotes the optimal reaction set of the follower and it is defined for each fixed $\alpha \in A$ by

$$R_{\gamma_1^*}(\gamma_1^*) = \{v^* \in V : J_2(\gamma_1^*, v^*, \alpha) \leq J_2(\gamma_1^*, v, \alpha), \forall v \in V\}. \quad (1b)$$

Under an additional technical restriction which will be delineated in the sequel, it has been shown in [5] that, for each fixed $\alpha \in A$, there exists an affine strategy for the leader which leads to J_1^* . More precisely, with α fixed at a nominal value α^* , there exists an $n \times m$ matrix $Q(\alpha^*)$, satisfying the relation

$$\frac{\partial J_2}{\partial u} Q(\alpha^*) + \frac{\partial J_2}{\partial v} = 0 \quad (2)$$

so that the incentive strategy

$$u = \gamma_1(v) = u^c + Q(\alpha^*)(v - v^c) \quad (3)$$

forces the follower to $v = v^c$, provided that $\frac{\partial J_2(u, v, \alpha)}{\partial u}$ evaluated at $(u = u^c, v = v^c)$ does not vanish within an ε -neighborhood of α^* , where (u^c, v^c) minimizes $J_1(u, v)$ over $U \times V$.

In general, equation (2) defines a class Q of $(n \times m)$ matrices which force the follower to $v = v^c$. We also note that there exist other strategies for the leader, which attain J_1^* , and they need not be of the affine form as in (3).

b) Since the leader does not have perfect knowledge of α , the actual value may be different from α^* , in which case the leader will end up with an inferior performance, because γ_1^* so defined is optimal only at $\alpha = \alpha^*$. Therefore, it is highly desirable for the leader to have a robustness property associated with his optimal strategy. More precisely, the leader would like to have the sensitivity of the realized value of his cost functional, against the variations in α about its nominal value α^* , be as small as possible. This property may be induced by making use of the intrinsic nonuniqueness of the solutions of (2), and also by introducing nonlinear strategies which satisfy (1a), as it will be shown in the sections to follow.

III. Introduction of a Sensitivity Function and Derivation of Robust Solutions

Let $\gamma_1 \in \Gamma_1$ be an incentive strategy for the leader, and $v_{\gamma_1} \in R_{\gamma_1}(\gamma_1)$. Towards the goal set in Section II, and as a measure of the sensitivity of $J_1(\gamma_1, v_{\gamma_1})$ with respect to deviations of α from its nominal value α^* , let us introduce the total derivative of $J_1(\gamma_1, v_{\gamma_1})$ with respect to α , evaluated at $\alpha = \alpha^*$, $u = u^c$, and $v = v^c$; more precisely, let us confine ourselves to affine strategies of the form (3) and, by abuse of terminology, let us define the first order sensitivity function of $J_1(\gamma_1, v_{\gamma_1})$ with respect to α , and at $\alpha = \alpha^*$, as

$$I_1(\alpha^*) = \frac{dJ_1(\gamma_1, v_{\gamma_1})}{d\alpha} = \frac{\partial J_1(u, v)}{\partial u} Q(\alpha^*) + \frac{\partial J_1(u, v)}{\partial v} \left(\frac{dv}{d\alpha} \right) \quad (4)$$

(u = u^c
v = v^c)

However, if the leader enforces the pair (u^c, v^c) which

globally minimizes J_1 on the produce space $U \times V$, $I_1(\alpha)$ identically vanishes for all $\alpha \in A$. In this case, the next term in the Taylor expansion of $J_1(\gamma_1, v_{\gamma_1})$ around α^* , which we call the second order sensitivity function of $J_1(\gamma_1, v_{\gamma_1})$ with respect to α , will have to be considered. Denoting it by $I_2(\alpha)$, and suppressing the arguments, we have

$$I_2(\alpha^*) = \frac{d^2 J_1(\gamma_1(v_{\gamma_1}), v_{\gamma_1})}{d\alpha^2} \bigg|_{\substack{\alpha=\alpha^* \\ u=u^c \\ v=v^c}} \\ = \frac{dv}{d\alpha} \left(Q' \frac{\partial^2 J_1}{\partial u^2} Q + Q' \frac{\partial^2 J_1}{\partial u \partial v} Q + \frac{\partial^2 J_1}{\partial v^2} \frac{dv}{d\alpha} \right) \bigg|_{\substack{\alpha=\alpha^* \\ u=u^c \\ v=v^c}} \quad (5)$$

In order to find an expression for $\frac{dv}{d\alpha}$, we note that the equation

$$\frac{\partial J_2}{\partial u} Q(\alpha^*) + \frac{\partial J_2}{\partial v} \bigg|_{\substack{u=\gamma_1(v_{\gamma_1}) \\ v=v^c}} = 0 \quad (6)$$

is an identity for all $\alpha \in A$, and it completely specifies the optimal response v_{α} of the follower when affine strategies γ_1 of the form (3) are used. The derivative of the above expression with respect to α would still vanish for all $\alpha \in A$. We then have, with $Q = Q(\alpha^*)$,

$$\left[Q' \frac{\partial^2 J_2}{\partial u^2} Q + Q' \frac{\partial^2 J_2}{\partial v \partial u} + \frac{\partial^2 J_2}{\partial u \partial v} Q + \frac{\partial^2 J_2}{\partial v^2} \right] \left(\frac{dv}{d\alpha} \right)' + Q' \frac{\partial^2 J_2}{\partial \alpha \partial u} + \frac{\partial^2 J_2}{\partial \alpha \partial v} = 0 \quad \forall \alpha \in A. \quad (7)$$

Since $J_2(u, v, \alpha)$ is strictly convex in u and v for all $\alpha \in A$, the coefficient matrix of $\left(\frac{dv}{d\alpha} \right)'$ is positive

definite, and thereby invertible.²⁾ Combining (5) and (7) we obtain the second-order sensitivity function of $J_1(\gamma_1, v_{\gamma_1})$ with respect to α , at $\alpha = \alpha^*$, to be

$$I_2(\alpha^*) = \left[\left(Q' \frac{\partial^2 J_2}{\partial u^2} Q + Q' \frac{\partial^2 J_2}{\partial v \partial u} + \frac{\partial^2 J_2}{\partial u \partial v} Q + \frac{\partial^2 J_2}{\partial v^2} \right)^{-1} \left(Q' \frac{\partial^2 J_2}{\partial \alpha \partial u} + \frac{\partial^2 J_2}{\partial \alpha \partial v} \right)' \right] \left[0' \frac{\partial^2 J_1}{\partial u^2} Q + Q' \frac{\partial^2 J_1}{\partial v \partial u} + \frac{\partial^2 J_1}{\partial u \partial v} Q + \frac{\partial^2 J_1}{\partial v^2} \right] \\ + \left[\left(Q' \frac{\partial^2 J_2}{\partial u^2} Q + Q' \frac{\partial^2 J_2}{\partial v \partial u} + \frac{\partial^2 J_2}{\partial u \partial v} Q + \frac{\partial^2 J_2}{\partial v^2} \right)^{-1} \left(Q' \frac{\partial^2 J_2}{\partial \alpha \partial u} + \frac{\partial^2 J_2}{\partial \alpha \partial v} \right)' \right] \left[\left(Q' \frac{\partial^2 J_2}{\partial u^2} Q + Q' \frac{\partial^2 J_2}{\partial v \partial u} + \frac{\partial^2 J_2}{\partial u \partial v} Q + \frac{\partial^2 J_2}{\partial v^2} \right)^{-1} \left(Q' \frac{\partial^2 J_2}{\partial \alpha \partial u} + \frac{\partial^2 J_2}{\partial \alpha \partial v} \right)' \right] \quad (8)$$

Remark 1. When the leader enforces his team solution, his objective will be to minimize $I_2(\alpha^*)$ over Q subject to (2), since $I_1(\alpha^*)$ vanishes. On the other hand, there may be cases when he would prefer to enforce a point other than (u^c, v^c) , in which case $I_1(\alpha^*)$ will not necessarily be zero. However, when the nonvanishing $I_1(\alpha^*)$ does not depend on the choice of Q from the class Q of $n \times m$ matrices which, together with an affine strategy of the form (2), enforce the follower to the desired point, one still has to consider $I_1(\alpha^*)$ as a measure of obtaining minimum sensitive solutions. This point will be further elucidated in Section VI.

The problem now is to minimize $I_2(\alpha^*)$ over all $(n \times m)$ real matrices $Q(\alpha^*)$ subject to the constraint

²⁾ For invertibility it is, of course, sufficient that the Hessian matrix of J_2 be full rank.

$$\frac{\partial J_2}{\partial u} Q(\alpha^*) + \frac{\partial J_2}{\partial v} \bigg|_{\substack{u=u^* \\ v=v^* \\ z=\alpha^*}} = 0. \quad (9)$$

The solution to this optimization problem may dictate some of the entries of $Q(\alpha^*)$ to take arbitrarily large values which corresponds to high gain feedback. However, if U is a bounded set, the value of such an optimizing strategy may not belong to the set of permissible controls. It is therefore necessary, also in view of the fact that high gain may not be desirable, to impose bounds on the entries of $Q(\alpha^*)$, of the form

$$|q_{ij}| \leq k_{ij}, \quad i=1, \dots, n; \quad j=1, \dots, m \quad (10a)$$

where

$$Q(\alpha^*) = \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1m} \\ q_{21} & \dots & \dots & \dots \\ \vdots & \dots & \dots & \dots \\ q_{n1} & \dots & \dots & q_{nm} \end{bmatrix}. \quad (10b)$$

Now, let us first assume that there exists an inner point of the set defined by (10a), which minimizes $I_2(\alpha^*)$ subject to (9). Then, the set of first order necessary conditions for optimality are given in the proposition to follow.

Proposition 1: A set of first order necessary conditions for an inner-point solution Q^* (satisfying (10a) with strict inequality) of the optimization problem formulated in this section is the existence of a vector of Lagrange multipliers $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)'$, $\lambda_i \in \mathbb{R}$, $i=1, 2, \dots, m$, such that

$$\nabla_{Q^*} I_2(\alpha^*) = (\lambda \frac{\partial J_2}{\partial u})' \bigg|_{\substack{\alpha=\alpha^* \\ Q=Q^*}} \quad (11a)$$

and

$$\frac{\partial J_2}{\partial u} Q^*(\alpha) + \frac{\partial J_2}{\partial v} \bigg|_{\substack{\alpha=\alpha^* \\ u=u^* \\ v=v^*}} = 0. \quad (11b)$$

where

$$\nabla_{Q^*} I_2(\alpha^*) = \begin{bmatrix} \frac{\partial I_2}{\partial q_{11}} & \frac{\partial I_2}{\partial q_{12}} & \dots & \frac{\partial I_2}{\partial q_{1m}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial I_2}{\partial q_{n1}} & \dots & \dots & \frac{\partial I_2}{\partial q_{nm}} \end{bmatrix} \bigg|_{\substack{\alpha=\alpha^* \\ Q=Q^*}}$$

provided that $\frac{\partial J_2}{\partial u}$ evaluated at $(u=u^*, v=v^*)$ does not vanish within an ϵ -neighborhood of α^* .

Proof: Let us assume that there exists an inner-point minimizing solution $Q^* \in Q$. Then, any matrix in the constraint set can be written as

$$Q = Q^* + \epsilon \Delta$$

where $\epsilon \in \mathbb{R}$ and $\Delta \in D \subset \mathbb{R}_{nm}$, where $\mathbb{R}_{nm}$ denotes the set of all $n \times m$ matrices with real entries and D is characterized by $D = \{\Delta \in \mathbb{R}_{nm} : \frac{\partial J_2}{\partial u} \Delta = 0\}$. Since D is a vector space, we should have

$$\frac{d}{d\epsilon} I_2(\alpha^*, Q^* + \epsilon \Delta) \bigg|_{\epsilon=0} = \nabla_{Q^*} I_2(\alpha^*, Q^*) \Delta = 0.$$

Hence the columns of $\nabla_{Q^*} I_2(\alpha^*)$ must be linearly independence on $\frac{\partial J_2}{\partial u}$, which implies the existence of a set of scalars $\lambda_1, \dots, \lambda_m$ such that (11a) is satisfied at an optimizing point.

When a solution satisfying (11) cannot be found, one

has to seek solutions which are on the boundary. In some special, but sufficiently general cases, such solutions can be obtained analytically, as we will elucidate in the next section.

IV. Solutions to Some Special Cases

In this section we relax the hypotheses of Proposition 1 in two different directions for some special cases, and obtain some explicit solutions. We restrict attention primarily to two classes, viz problems with separable cost functionals and the so-called singular incentive problems.

1. Separable cost functionals

Consider the class of problems in which the uncertain parameter α affects the cost functional of the follower only through v , i.e. $J_2(u, v, \alpha)$ is separable as

$$J_2(u, v, \alpha) = g_1(u, v) + g_2(v, \alpha) \quad (12)$$

in which case the term $\frac{\partial^2 J_2}{\partial \alpha \partial u}$ in (8) vanishes.

Let us assume, for the sake of simplicity in analysis, that $n=2$, $m=1$. Then, under the equality constraint (9), it is possible to write one component of $Q(\alpha^*)$ in terms of the other component, and hence express $I_2(\alpha^*)$ in terms of either q_1 or q_2 . We observe that in the limit as q_1 (or q_2) $\rightarrow \infty$, $I_2(\alpha^*)$ tends to its absolute minimum value zero. This, however, violates the constraint (10a), and therefore we have to look for boundary solutions.

Solving for q_1 in terms of q_2 from (9) and substituting into $I_2(\alpha^*)$, we arrive at the optimization problem.³⁾

$$\begin{aligned} \min_{|q_2| \leq k_2^2} I_2(\alpha^*) = & \left[\frac{\partial^2 J_1}{\partial u_1^2} (\delta q_2 + \epsilon)^2 - 2 \frac{\partial^2 J_1}{\partial u_1 \partial v} (\delta q_2 + \epsilon) + 2 \frac{\partial^2 J_1}{\partial u_2 \partial v} q_2 \right. \\ & - 2 \frac{\partial^2 J_1}{\partial u_1 \partial u_2} q_2 (\delta q_2 + \epsilon) + \frac{\partial^2 J_1}{\partial u_2^2} q_2^2 + \frac{\partial^2 J_1}{\partial v^2} \bigg] \\ & + \left[\frac{\partial^2 J_2}{\partial \alpha \partial v} / \left(-\frac{\partial^2 J_2}{\partial u_1^2} (\delta q_2 + \epsilon)^2 - 2 \frac{\partial^2 J_2}{\partial u_1 \partial u_2} q_2 (\delta q_2 + \epsilon) + \frac{\partial^2 J_2}{\partial u_2^2} q_2^2 \right. \right. \\ & \left. \left. - \frac{\partial^2 J_2}{\partial u_1 \partial v} (\delta q_2 + \epsilon) + \frac{\partial^2 J_2}{\partial u_2 \partial v} q_2 + \frac{\partial^2 J_2}{\partial v^2} \right) \right] \bigg|_{\substack{\alpha=\alpha^* \\ u=u^* \\ v=v^*}} \quad (13) \end{aligned}$$

where

$$\delta \triangleq \frac{\partial J_2}{\partial u_2} / \frac{\partial J_2}{\partial u_1}, \quad \epsilon \triangleq \frac{\partial J_2}{\partial v} / \frac{\partial J_2}{\partial u_1} \bigg|_{\substack{\alpha=\alpha^* \\ u=u^* \\ v=v^*}} \quad (14)$$

This is equivalent to maximizing the denominator of the second product term of (13) under the constraint $|q_2| \leq k_2^2$. Since the denominator is strictly convex in q_2 , the minimizing solution is

$$q_2^* = -\text{sgn}(q_2^0) k_2^2 \quad (15)$$

where q_2^0 is the value of q_2 at which the derivative of the denominator with respect to q_2 vanishes, namely

$$q_2^0 = \left(2 \frac{\partial^2 J_2}{\partial u_1 \partial u_2} \epsilon + \frac{\partial^2 J_2}{\partial u_1 \partial v} \right) / \left(-2 \frac{\partial^2 J_2}{\partial u_1^2} \epsilon - \frac{\partial^2 J_2}{\partial u_2 \partial v} \right) - \frac{\partial^2 J_2}{\partial u_1^2} \epsilon$$

³⁾ Here we assume that k_2 is sufficiently large so that, so long as the solution to (13) is finite, q_1 solved from (9) satisfies the given constraint.

$$-2 \frac{\partial^2 J_2}{\partial u_1 \partial u_2} + \frac{\partial^2 J_2}{\partial u_2^2}, \quad (16)$$

and

$$\text{sgn } q_2^0 = \begin{cases} +1 & \text{if } q_2^0 \geq 0 \\ -1 & \text{if } q_2^0 < 0 \end{cases} \quad (17)$$

Then, q_1^* is obtained via the constraint (9). If the constraint is not compatible, one should reformulate $I_2(a^*)$ in terms of q_1 and repeat the same procedure. This result can be extended to cases when $n > 2$. The following algorithm can be used to obtain $Q^*(a^*)$ which minimizes $I_2(a^*)$:

An algorithm to obtain the minimizing argument $Q^*(a^*)$ of $I_2(a^*)$:

(i) Compute

$$Q^0 = (q_1^0, q_2^0, \dots, q_n^0) = - \left[\frac{\partial^2 J_2}{\partial u^2} \right]^{-1} \frac{\partial^2 J_2}{\partial u \partial v} \bigg|_{\substack{a=a^* \\ u=u^0 \\ v=v^0}}$$

(ii) Set

$$q_1^* = -\text{sgn}(q_1^0) k_1^2, \quad i=1, 2, \dots, n$$

check for satisfaction of the constraint (9).

(iii) If (9) is not satisfied, set

$$q_1^* = -\text{sgn}(q_1^0) k_1^2, \quad \text{for all } i=1, \dots, n, \text{ but } i \neq j \text{ for some } j.$$

Search for $q_j^* \leq k_j^2$ satisfying (9); $j=1, \dots, n$.

(iv) If (9) is not satisfied, set

$$q_1^* = -\text{sgn}(q_1^0) k_1^2 \quad \forall i=1, \dots, n; \text{ but } i \neq j, i \neq l, \text{ for some } j, l.$$

Search for $q_j^* \leq k_j^2$ and $q_l^* \leq k_l^2$ satisfying (9); $j, l=1, 2, \dots, n$.

Stop when (9) is satisfied.

This algorithm can be extended to the case when $m \geq 1$ or functionals $J_2(u, v, x)$ which are separable as in (12).

1. Singular incentive problems

An interesting case occurs when the leader's control affects the follower's cost functional linearly. In this case, the follower's cost functional is not convex on the product space $U \times V$, and therefore the previous theory on linear incentive problems is not directly applicable; however, an extension is possible as we elucidate in the sequel. Towards this end, first note that when the leader announces an affine strategy of the form (3), the follower's cost functional J_2 becomes a function of only v , for a given $a \in A$. Accordingly, if the functional $J_2(u^t + Q(a)(v - v^t), v, a)$ is strictly convex on V for each $a \in A$, and if equation (2) admits a nonempty family Q of solutions at the desired point (u^t, v^t) , then the least-sensitive incentive design problem becomes meaningful for this particular class of cost functionals. More precisely, for a given $u^t \in U$, let

$$\frac{\partial^2 J_2(u, v, a)}{\partial v^2} = \frac{\partial^2 J_2(u, v, a)}{\partial v^2} - \frac{\partial^2 J_2(u, v, a)}{\partial v \partial u} Q(a^*) > 0 \quad (18)$$

for all $v \in V$, $a \in A$ and $Q(a^*) \in Q \subseteq Q$, where Q is assumed to be nonvoid. Under these assumptions, the optimal

reaction of the follower to any announced strategy of the form (3), where $Q(a^*) \in Q$, is $v = v^t$, which is confirmed by (2) as a first order necessary condition, and (18) provides a second order sufficient condition for the optimality of $v = v^t$.

For this class of incentive design problems, it is possible to obtain an inner-point solution analytically, because of the particular structure of the sensitivity function. Towards this end, let us assume that the follower's cost function is linear in u and separable as in (12). Let us further assume that $n=2$, $m=1$.

In this case, constructing q_1 in terms of q_2 as in §IV.1, $I_2(a^*)$ becomes

$$I_2(a^*) = \left[\frac{\partial^2 J_1}{\partial u_1^2} (\delta q_2 + \xi)^2 - 2 \frac{\partial^2 J_1}{\partial u_1 \partial v} (\delta q_2 + \xi) + 2 \frac{\partial^2 J_1}{\partial u_2 \partial v} q_2 - 2 \frac{\partial^2 J_1}{\partial u_1 \partial u_2} q_2 (\delta q_2 + \xi) + \frac{\partial^2 J_1}{\partial u_2^2} q_2^2 + \frac{\partial^2 J_1}{\partial v^2} \left(\frac{\partial^2 J_2}{\partial a \partial v} / \left(\frac{\partial^2 J_2}{\partial u_2 \partial v} q_2 - \frac{\partial^2 J_2}{\partial u_1 \partial v} (\delta q_2 + \xi) + \frac{\partial^2 J_2}{\partial v^2} \right) \right)^2 \right] \bigg|_{\substack{a=a^* \\ u=u^0 \\ v=v^0}} \quad (19a)$$

where δ and ξ are defined by (14).

The derivative of $I_2(a^*)$ with respect to q_2 vanishes at

$$q_2^* = \left\{ \left[\xi^2 \frac{\partial^2 J_1}{\partial u_1^2} - 2\xi \frac{\partial^2 J_1}{\partial u_1 \partial v} + \frac{\partial^2 J_1}{\partial v^2} \right] \left[\frac{\partial^2 J_2}{\partial u_2 \partial v} - \delta \frac{\partial^2 J_2}{\partial u_1 \partial v} \right] - \left[\delta \xi \frac{\partial^2 J_1}{\partial u_1^2} - \delta \frac{\partial^2 J_1}{\partial u_1 \partial v} + \frac{\partial^2 J_1}{\partial u_2 \partial v} - \xi \frac{\partial^2 J_1}{\partial u_1 \partial u_2} \right] \left[\frac{\partial^2 J_2}{\partial v^2} - \xi \frac{\partial^2 J_2}{\partial u_1 \partial v} \right] / \left[\left(\delta^2 \frac{\partial^2 J_1}{\partial u_1^2} - 2\delta \frac{\partial^2 J_1}{\partial u_1 \partial u_2} + \frac{\partial^2 J_1}{\partial u_2^2} \right) \left(\frac{\partial^2 J_2}{\partial v^2} - \xi \frac{\partial^2 J_2}{\partial u_1 \partial v} \right) - \left(\delta \xi \frac{\partial^2 J_1}{\partial u_1^2} - \delta \frac{\partial^2 J_1}{\partial u_1 \partial v} + \frac{\partial^2 J_1}{\partial u_2 \partial v} \right)^2 \right] \right\} \bigg|_{\substack{a=a^* \\ u=u^0 \\ v=v^0}} \quad (19b)$$

provided that the denominator of (19) does not vanish, and k_i^2 , $i=1, 2$, are sufficiently large. Furthermore, this is the unique stationary point of $I_2(a^*)$. This result leads to the following proposition.

Proposition 2: The minimum sensitive (robust) linear incentive design problem formulated in this section admits a unique solution (q_1^*, q_2^*) where q_2^* is given by (19b), and q_1^* is obtained through the linear constraint (9) which relates q_1 and q_2 , provided that the denominator of (19a) does not vanish, and $(q_1^*, q_2^*) \in Q$.

Proof: This result follows from the following four properties of the function $F(q_2) = I_2(a^*)$:

(i) There exists a finite number M such that

$$\lim_{q_2 \rightarrow -\infty} F(q_2) = \lim_{q_2 \rightarrow +\infty} F(q_2) = M.$$

$$(ii) F(q_2) \rightarrow -\infty \text{ as } q_2 \rightarrow \left(\frac{\partial^2 J_1}{\partial u_1 \partial v} - \xi \frac{\partial^2 J_1}{\partial v^2} \right) / \left(\frac{\partial^2 J_1}{\partial u_2 \partial v} - \delta \frac{\partial^2 J_1}{\partial u_1 \partial v} \right)$$

$$-\frac{\partial^2 J_2}{\partial u_1 \partial v} \Big|_{u=u^*, v=v^*} \frac{\partial y_1}{\partial v}.$$

(iii) $F(q_2)$ is continuous, except at $q_2 = \bar{q}_2$.

(iv) $F(q_2)$ has a single stationary point q_2^* .

The four properties readily lead to the conclusion that q_2^* is the unique minimizing solution for F .

V. Use of Nonlinear Strategies in Sensitivity Considerations

The analysis of previous sections was confined to the class of linear strategies. Although this class is big enough to provide optimal solutions to incentive problems [5], the use of nonlinear strategies may provide additional degrees of freedom when dealing with sensitivity problems. For example, when $n=m=1$, the constraint (9) determines $Q(u^*)$ uniquely, except for generic cases where optimum linear strategies do not exist. Hence in this case, the set Q is a singleton and it does not allow sensitivity considerations. However, if the leader is permitted to enlarge his strategy space by including a suitable nonlinear term in his control, he may have extra degrees of freedom to reduce the sensitivity of his performance to changes in the uncertain parameter in the follower's cost functional. Towards this end, let us assume that the strategy space of the leader is the set of all mappings from V onto U , twice continuously differentiable with respect to v , and with bounded first and second derivatives. For the case $n=m=1$, and with separable $J_2(u, v, y)$, the second order sensitivity function takes the form

$$I_2(u^*) = \left(\frac{\partial^2 J_2}{\partial u^2} \left(\frac{\partial y_1}{\partial v} \right)^2 + 2 \frac{\partial^2 J_2}{\partial u \partial v} \frac{\partial y_1}{\partial v} + \frac{\partial^2 J_2}{\partial v^2} \right) \left(\frac{\partial^2 J_2}{\partial v \partial u} / \frac{\partial^2 J_2}{\partial u^2} \left(\frac{\partial y_1}{\partial v} \right)^2 + 2 \frac{\partial^2 J_2}{\partial u \partial v} \left(\frac{\partial y_1}{\partial v} \right) - \frac{\partial J_2}{\partial u} \frac{\partial^2 y_1}{\partial v^2} + \frac{\partial^2 J_2}{\partial v^2} \right) \Big|_{\substack{u=u^* \\ v=v^*}} \quad (20)$$

The problem is to minimize $I_2(u^*)$ over $\frac{\partial y_1}{\partial v} \Big|_{\substack{u=u^* \\ v=v^*}}$ and

and $\frac{\partial^2 y_1}{\partial v^2} \Big|_{\substack{u=u^* \\ v=v^*}}$ subject to the constraints

$$\frac{\partial J_2}{\partial u} \frac{\partial y_1}{\partial v} + \frac{\partial J_2}{\partial v} = 0 \quad (21)$$

$$\left| \frac{\partial y_1}{\partial v} \right| \leq k_1^2 \quad (22a)$$

$$\left| \frac{\partial^2 y_1}{\partial v^2} \right| \leq k_2^2 \quad (22b)$$

at $u=u^*$, $v=v^*$, and $y=y^*$. Here, $\frac{\partial y_1}{\partial v}$ is completely determined by the constraint (9). Carrying out this minimization, one obtains

$$\frac{\partial y_1}{\partial v} = - \left(\frac{\partial J_2}{\partial v} \right) / \left(\frac{\partial J_2}{\partial u} \right) \quad (23)$$

and

$$\frac{\partial^2 y_1}{\partial v^2} = \text{sgn} \left(\frac{\partial J_2}{\partial u} \right) k_1^2 \quad (24)$$

where the partials are evaluated at $u=u^*$, $v=v^*$, and $y=y^*$. It is also assumed here that k_1^2 is sufficiently

large. These results lead to the following proposition.

Proposition 3: A representation of the least sensitive optimal incentive strategy within the class of twice continuously differentiable policies with bounded first and second derivatives with respect to v is given by

$$y_1^*(v) = \left[1 - \frac{1}{2} \text{sgn} \left(\frac{\partial J_2}{\partial u} \right) k_1^2 \right] u^* - \left[\frac{\partial J_2}{\partial v} / \frac{\partial J_2}{\partial u} \right] (v - v^*) + \frac{1}{2} \text{sgn} \left(\frac{\partial J_2}{\partial u} \right) k_1^2 (v - v^* + 1)^2 u^*. \quad (25)$$

Proof: Follows from the previous discussion.

The case when $n \geq 2$ can be treated in a similar manner. Components of $\frac{\partial y_1}{\partial v}$ are obtained via the algorithm presented in Section IV, and $\frac{\partial^2 y_1}{\partial v^2}$ is given by

$$\frac{\partial^2 y_1}{\partial v^2} = \text{sgn} \left(\frac{\partial J_2}{\partial u_1} \right) k_{12}^2 \Big|_{\substack{u=u^* \\ v=v^*}} \quad i=1, 2, \dots, n. \quad (26)$$

Remark: Since $\frac{\partial J_2}{\partial u_1} \frac{\partial^2 y_1}{\partial v^2} > 0$ $i=1, 2, \dots, n$ at the nominal point, the enlarging of the strategy space of the leader definitely improves his sensitivity function. One could introduce further improvements by allowing for higher order continuously differentiable strategies.

VI. An Example from Economics

In this section we discuss an example from microeconomics, which illustrates some of the results obtained in this paper, especially the ones on singular incentive problems. Let us consider a duopolistic market model in which two firms compete; the leader, P_1 , produces the goods X and Y , and the follower produces the good Z . All three goods are substitutable within an ϵ -neighborhood of the equilibrium point of the market and they are sold at the same price p which is assumed to satisfy the linear demand relation

$$x + y + z = d_0 - d_1 p \quad (27)$$

where x , y , and z represent the quantities to be produced from each good X , Y , and Z , respectively. d_0 and d_1 are positive constants. It is also assumed that firms have a logarithmic cost function, compatible with the appealing hypothesis that cost per unit of production decreases as the level of production increases.

Then, the profit functions of the firms, which are to be maximized, become

$$\pi_1 = \left(\frac{d_0 - (x+y+z)}{d_1} \right) (x+y) - c_1 \ln(x+1) - c_2 \ln(y+1) - L_1 \quad (28a)$$

$$\pi_2 = \left(\frac{d_0 - (x+y+z)}{d_1} \right) z - c_3 \ln(z+1) - L_2 \quad (28b)$$

$x, y, z \geq 0$.

Here $c_i > 0$, $i=1, 2, 3$ are parameters reflecting the differences in the cost of productions of X , Y , and Z , respectively. It is assumed that c_i , $i=1, 2, 3$ are constant within an ϵ -neighborhood of the equilibrium point of the market. $L_i > 0$, $i=1, 2$ are fixed costs of the firms.

For this problem there exists a sequence of closed-loop policies for P_1 which force the follower to $z=0$ [1]—a result which is valid for any demand relation in which the price is a strictly decreasing function of z . However, for the leader the limiting

strategy is not well-defined since it requires the gain vector $Q(\alpha^*)$ to have unbounded elements (i.e., infinite threat). In general, such a strategy is neither credible for the follower, nor affordable for the leader. In its stead, an alternative is to assume that P1 imposes a suitable point of equilibrium, compatible with the duopolistic nature of the problem.

Towards this end, let the objective functional considered by the leader be

$$J_1 = v\pi_1 + (1-v)\pi_2, \quad 0 < v < 1 \quad (29)$$

where v is large enough to provide the leader better profit and not so large to lead to a noncredible incentive scheme. A possible upper bound for v may be the one which provides the follower a profit comparable with what he or she would make in a Nash equilibrium case. The objective functional of the follower is still π_2 . Let us assume, for simplicity in the analysis to follow, that $v = 1/2$, in which case the leader can roughly guarantee two-thirds of the market. To complete the formulation of the problem, let us assume that the parameter in the cost of production of 2, namely c_3 , is uncertain for the leader. However, he knows a nominal value of c_3 , say c_3^* , around which c_3 can vary. The goal of the leader is to design a strategy, using his strategic variables x and y , which will enforce the follower to the maximizing arguments of J_1 when $c_3 = c_3^*$; and in addition he seeks a strategy under which his profit function is least sensitive to variations in c_3 . Let us also assume that the leader confines himself to affine strategies. This problem is within the scope of singular incentive problems discussed in Section IV.2 with the sole exception that the enforced point is not the team solution of his profit function π_1 for which he desires the least-sensitivity property. In this case, his first order sensitivity function $I_1(\alpha^*)$, as given by (4), does not vanish. However, it can be shown that $I_1(\alpha^*)$ is invariant under the choice of $Q(\alpha^*)$ which takes its values in the constraint defined by (2). Hence, in this problem, for the sensitivity analysis, one still has to consider $I_2(\alpha^*)$.

Using the objective function (29) with $v = 1/2$, it can be shown that the desired equilibrium point of the leader in terms of the parameters of the problem is given as follows, under certain not totally restrictive conditions:

$$x = \frac{(d_0 + 6) + v(d_0 + 6)^2 - 8(1 + \frac{c_2}{c_1} + \frac{c_3^*}{c_1})c_1 d_1}{4(1 + \frac{c_2}{c_1} + \frac{c_3^*}{c_1})} - 1 \quad (30a)$$

$$y^* = \frac{c_2}{c_1} (x^* + 1) - 1 \quad (30b)$$

$$z^* = \frac{c_3^*}{c_1} (x^* + 1) - 1. \quad (30c)$$

The class of affine strategies which yields x^* , y^* , and z^* when the uncertain parameter c_3 takes the nominal value c_3^* , is represented by

$$y_1(x) = [x^* + q_1(z - z^*), y^* + q_2(z - z^*)]' \quad (31)$$

where q_1 represents the coefficient which indicates how P1 would modify x if $z \neq z^*$, and similarly, q_2 relates the change in z to y . When the constraint (2) is computed for this problem, it turns out that it forces q_1 and q_2 to lie in the linear manifold defined by

$$q_1 + q_2 = \frac{d_3 x^* - x^* y^* + d_1 c_1}{x^* y^*} = 1. \quad (32)$$

Now, any pair (q_1, q_2) satisfying (32) and (18) would induce P2 to produce the amount z^* when the uncertain parameter c_3 takes the nominal value c_3^* . From this set $Q(\alpha^*)$, the pair (q_1^*, q_2^*) which minimizes the sensitivity function of the leader is computed using (20), and is given by

$$q_1^* = \frac{\theta c_1}{c_1 + c_2}, \quad q_2^* = \frac{\theta c_2}{c_1 + c_2} \quad (33)$$

which indicates that, in a least sensitive (robust) incentive scheme, the leader should allocate the incentive among the goods he produces according to their respective costs of production.

VII. Conclusion

In this paper, we have introduced a minimum sensitivity approach towards the solution of deterministic incentive design problems, which leads to robust incentive policies that are least sensitive to variations in the value of a parameter (from a nominal) characterizing the follower's cost function. Even though this approach has been developed with respect to a single parameter, it is possible to envision natural (conceptual) extensions to the multi-parameter situation in which case the second order sensitivity function $I_2(\alpha^*)$ will be defined as an appropriate norm of the matrix whose elements are the total derivatives of $J_1(y_1(v_2), v_2)$ with respect to the components of the vector α . Yet other possible extensions of this minimum sensitivity approach are to the class of problems in which the leader has partial dynamic information (as in [5], [9]) and to the class of stochastic incentive problems discussed in [11]. Such extensions are currently under study, and will be reported in forthcoming papers.

References

- [1] T. Başar and H. Selbuz, "Closed-Loop Stackelberg Strategies with Applications in the Optimal Control of Multilevel Systems," *IEEE Trans. on Automatic Control*, Vol. AC-24, April 1979, pp. 166-179.
- [2] Y. C. Ho, P. B. Luh, and G. J. Olsder, "A Control Theoretic View on Incentives," *Proc. 4th Intl. Conf. on Analysis and Optimization of Systems*, Springer, 1980.
- [3] M. Salman and J. B. Cruz, Jr., "An Incentive Model of Duopoly with Government Coordination," *Automatica*, Vol. 17, 1981, pp. 821-830.
- [4] T. Başar and J. B. Cruz, Jr., "Concepts and Methods in Multi-person Coordination and Control," in *Optimization and Control of Dynamic Operational Research Models*, S. G. Tzafestas, ed., North Holland, 1982.
- [5] Y. P. Zheng and T. Başar, "Existence and Derivation of Optimal Affine Incentive Schemes for Stackelberg Games with Partial Information: A Geometric Approach," *Int. J. of Control*, to appear.
- [6] T. Groves and M. Loeb, "Incentives in a Divisionalized Firm," *Management Science*, Vol. 25, 1979, pp. 221-230.
- [7] L. Hurwicz and L. Shapiro, "Incentive Structures Maximizing Residual Gain Under Incomplete Information," *Bell J. of Economics*, Vol. 9, 1978, pp. 180-191.
- [8] L. P. Jennergren, "On the Design of Incentives in Business Firms--A Survey of Some Recent Research," *Management Science*, Vol. 26, No. 8, 1980, pp. 180-201.
- [9] T. Başar, "A General Theory for Stackelberg Games with Partial State Information," *Large Scale Systems*, to be published. An earlier version has appeared in *Proc. of 4th Intl. Conf. on the Analysis and Optimization of Systems*, Springer, 1980.
- [10] Y. C. Ho, P. B. Luh, and R. Muralidharan, "Information Structure, Stackelberg Games, and Incentive Controllability," *IEEE Trans. on Automatic Control*, Vol. AC-26, April 1981, pp. 454-460.
- [11] T. Başar, "Affine Incentive Schemes for Stochastic Systems with Dynamic Information," *Proc. ACC*, June 1982.
- [12] B. Tolwinski, "Closed-Loop Stackelberg Solution to Multi-Stage Linear-Quadratic Game," *J. Optimiz. Theory Appl.*, September 1981.

ROBUSTNESS OF INCENTIVE POLICIES IN TEAM PROBLEMS WITH DISCREPANCIES IN GOAL PERCEPTIONS

Derya Cansever, Tamer Başar and Jose B. Cruz, Jr.

Department of Electrical Engineering and Coordinated Science Laboratory
University of Illinois, 1101 W. Springfield Avenue, Urbana, Illinois 61801 USA

Abstract. In this paper we analyze a class of two-agent team decision problems with a hierarchical decision structure, wherein one of the decision makers may have a slightly different perception of the overall team goal, with this slight variation not known by the other agent who is assumed to occupy the hierarchically dominant position. The leading agent has access to dynamic information and his role is to announce such a policy (incentive scheme) which would lead to achievement of the overall team goal, in spite of the slight variations in the other agent's perception of that goal, which are not known or predictable by him. We may call a policy with such an additional feature a robust incentive policy. We obtain, in the paper, robust policies for the leading agent, for a general cost functional with convex structure, which are least sensitive to variations in the following agent's perception of the team goal. In some special cases, we show that the robust feature of the incentive scheme is maintained regardless of the magnitude and nature of the variations, and illustrate the theory with two applications examples arising in microeconomics and armament limitation and control.

Keywords. Economics; game theory; incentives, multicriteria decision problems; robustness of decision policies.

I. INTRODUCTION

The main characteristic of team decision problems is the presence of several decision makers with a common goal, this common goal being quantified in a common objective functional which is to be optimized jointly (but possibly in a decentralized fashion) by all decision makers. An underlying stipulation in research on team theory has been the assumption that all agents perceive the common goal in exactly the same way, and face exactly the same mathematical optimization problem [Marschak and Radner (1972)]. In this paper we relax this basic assumption and allow (in the context of two-agent problems) one agent to have a somewhat different perception of the common goal and to quantify it in a slightly different way. Furthermore, we will assume that the other agent is not informed of the existence of this discrepancy in the perception of the common goal, but is able to monitor the decision of the former by occupying a higher (dominant) position in the decision process. The problem we address to is the design of a suitable strategy for the agent who occupies the hierarchically superior position and who still adopts the original team objective functional as his own, such that the change in the minimum value of the team cost because of the discrepancy in the perceptions of the common goal is kept to a minimum. Ideally, the hierarchically superior member of the team would

seek not to be affected by this discrepancy, if this is at all possible.

We will approach this problem using optimum incentive design schemes [Ho, Luh and Olsder (1982) and Zheng and Başar (1982)], which involve a hierarchy in decision making and a suitable information structure for the decision maker at the top of the hierarchy, that allows him to design a policy which in its turn induces the other decision maker with a different objective functional to behave in a desired manner. Recently in [Cansever and Başar (1982)], optimal incentive schemes have been used, within the context of Stackelberg games, to minimize the effect of changes in the parameters of the follower's cost functional on the leader's optimum cost value, by simultaneously achieving a desired goal. Here, we direct our attention to problems which are nominally team, and derive incentive schemes that are least sensitive to deviations of the hierarchically inferior decision maker's perceptions of the uncertain parameters. The fact that the underlying goal is common (that is, the nominal optimization problem is a team problem - a property that may be destroyed in the decision process) can be exploited to obtain very appealing robust strategies, as we will show in the sections to follow.

The problem is formulated in section II. In section III we introduce sensitivity functions

and obtain robust affine strategies for a general class of convex cost functionals. In section IV the results of the previous section are applied to a problem arising in microeconomics. Section V deals with the generalization of the above results to the multiparameter case, and section VI illustrates the basic idea of robust policies for the multiparameter case using a model from armament limitation and control.

II. PROBLEM FORMULATION

Consider a two-person deterministic team decision problem in normal form, described by the cost functional $J(Y_1, Y_2, \alpha)$, where $Y_1 \in \Gamma_1$ denotes the strategy of DM1 (1'sth decision maker) and $\alpha \in A \subset \mathbb{R}$ is a parameter on which the cost functional depends. Let $u \in U = \mathbb{R}^n$, $v \in V = \mathbb{R}^m$ denote the decision variables of DM1 and DM2, respectively, and assume that $\Gamma_1 = \{Y_1: V \rightarrow U\}$, $\Gamma_2 = V$; i.e. DM1 has access to the decision value of DM2. DM1 also knows the precise value of the parameter α (say α^0), whereas DM2 perceives its value differently (say $\alpha^+ \in A$), which in turn gives rise to a different cost functional from his point of view, namely $J_1(Y_1, Y_2, \alpha^+) \neq J(Y_1, Y_2, \alpha^0)$. Furthermore, DM1 does not know the exact value perceived by DM2, but his ultimate goal is to see that the lowest possible value is attained for $J(Y_1, Y_2, \alpha^0)$. The decision structure of the problem is assumed to be hierarchical, in the sense that DM1 is the dominant decision maker and has the power and means of declaring his policy in advance and enforcing it on the other DM. Hence, while DM2 is faced with the problem of minimizing $J(Y_1(v), v, \alpha^+)$ over $v \in V$, DM1 wishes to choose a $Y_1 \in \Gamma_1$ (in total ignorance of α^+) that would eventually lead to a minimum value for $J(Y_1(v), v, \alpha^0)$.

By an abuse of notation, let $J(u, v, \alpha)$ denote the cost functional on the product space $U \times V$, for each $\alpha \in A$, and assume that this functional is strictly convex on $U \times V$, for each $\alpha \in A$, is twice continuously differentiable in its first two arguments and continuously differentiable in its third argument. Furthermore, let us denote the unique minimum of $J(u, v, \alpha^0)$ by $(u^c, v^c) \in U \times V$. Restricting DM1 to affine policies in Γ_1 , we first note that the policy

$$Y_1(v) = u^c + P(v - v^c), \quad (1)$$

where P is an $(n \times m)$ -matrix, has the appealing property that if DM2's perception of α is α^0 , then $\min_{v \in V} J(Y_1(v), v, \alpha^0)$ leads to the desired

value $v^c \in V$ for any matrix P . If $\alpha^+ \neq \alpha^0$, however, the problem ceases to be a cooperative one since the problem faced by DM2 is

$$\min_{v \in V} J(Y_1(v), v, \alpha^+) \quad (2a)$$

whose minimizing solution (say $v^+ \in V$) satisfies (and is uniquely determined by) the equation

$$\nabla_u J(u^+, v^+, \alpha^+) P + \nabla_v J(u^+, v^+, \alpha^+) = 0 \quad (2b)$$

where $u^+ = Y_1(v^+) = u^c + P(v^+ - v^c)$ and is not necessarily the same as u^c . The problem we address to, in the sequel, is whether it is possible to choose a robust policy Y_1 (by choosing P appropriately) so that either $u^+ = u^c$ and $v^+ = v^c$, or the discrepancies will be small whenever α^+ is close to α^0 . In other words, we seek either total insensitivity or minimum sensitivity of the optimum value of $J(u, v, \alpha^0)$ to variations in the perception of DM2 (of α) by a proper choice of Y_1 .

III. INTRODUCTION OF A SENSITIVITY FUNCTION AND DERIVATION OF ROBUST SOLUTIONS

As a measure of the sensitivity of $J(u, v, \alpha^0)$ with respect to deviations of α^+ from its nominal value α^0 , let us introduce the total derivative of $J(u, v, \alpha^0)$ with respect to its third argument, which we call the first order sensitivity function of $J(u, v, \alpha)$ with respect to α :

$$\begin{aligned} I_1(\alpha^0) &= dJ(u^c, v^c, \alpha^0)/d\alpha \\ &= [\nabla_u J(u^c, v^c, \alpha^0)P + \nabla_v J(u^c, v^c, \alpha^0)] \cdot \left(\frac{dv^c}{d\alpha}\right) \end{aligned} \quad (3)$$

The first product term of the above expression vanishes at the nominal solution point. However, this is not necessarily the case if $\alpha \neq \alpha^0$. In order to find an expression for $\frac{dv}{d\alpha}$, let us first note that the equation that determines the optimum response of DM2 to (1), for a general α :

$$\nabla_u J(u, v, \alpha) P + \nabla_v J(u, v, \alpha) = 0 \quad (4)$$

is an identity for all $\alpha \in A$, and hence its derivative with respect to α still vanishes for all $\alpha \in A$. Such a consideration readily leads to

$$(\nabla_u J(u, v, \alpha) P + \nabla_v J(u, v, \alpha)) \cdot \left(\frac{dv}{d\alpha}\right) = -\frac{d}{d\alpha} [\nabla_u J(u, v, \alpha) P + \nabla_v J(u, v, \alpha)] \cdot (u, v)$$

from which

$$\frac{dv}{d\alpha} \cdot (\nabla_u J(u, v, \alpha) P + \nabla_v J(u, v, \alpha)) = -\frac{d}{d\alpha} [\nabla_u J(u, v, \alpha) P + \nabla_v J(u, v, \alpha)] \cdot (u, v) \quad (5)$$

Now, if we choose P such that

$$\nabla_u J(u^c, v^c, \alpha^0) P + \nabla_v J(u^c, v^c, \alpha^0) = 0 \quad (6)$$

then $\frac{dv}{d\alpha}$ vanishes at the nominal solution point. By the same token, the n th order sensitivity function of J with respect to α ,

$$I_n(\alpha) \triangleq \frac{d^n J(u, v, \alpha)}{d\alpha^n} \bigg|_{\substack{\alpha = \alpha^0 \\ u = u^c \\ v = v^c}}, \quad n=1,2,3$$

would carry, by chain rule, $\frac{dv}{da}$ as a product term; hence the sensitivity functions of order 1, 2 and 3 vanish at the nominal solution point. This situation, in turn, implies that when DM2's perception a^* stays within an ϵ -neighborhood of its nominal value a^0 , the 3rd order Taylor approximation of the effect of this discrepancy is zero. Therefore, when the class of matrices P defined by (6) is not empty, affine strategies

$$y_1 = u^c + P(v-v^c), \quad P \in \mathcal{F} \quad (7)$$

carry very appealing sensitivity properties.

Let us now assume that $J(u,v,a)$ is linear in a ; more precisely,

$$J(u,v,a) = g(u,v) + a h(u,v) \quad (8)$$

where $\nabla_u h(u^c, v^c)$ is not identically zero. Then $\nabla_{au} J(u,v,a)$ and $\nabla_{av} J(u,v,a)$ become independent of a , and it becomes possible to choose P such that sensitivity functionals of all orders vanish, for all values of $a \in A$. Such a choice would be a P satisfying

$$\nabla_u h(u^c, v^c) P + \nabla_v h(u^c, v^c) = 0, \quad (9)$$

which always exists since $\nabla_u h(u^c, v^c) \neq 0$. In this case, DML can induce DM2 to choose $y_2 = v^c$, independent of his perception of a , that is even if $a^* \neq a^0$.

Some more insight can be gained into this result by taking a slightly different approach. Consider the team decision problem described by the cost functional

$$J(u, v_1, v_2, a^0) = g(u, v_1) - \alpha h(u, v_2), \quad \alpha \neq 0, \quad (10)$$

where both v_1 and v_2 belong to V . Let $g(u, v_1)$ and $h(u, v_2)$ both be strictly convex on $U \times V$ and further assume that

$$J(u, v_1, v_2, \alpha) = g(u, v_1) + \alpha h(u, v_2)$$

is strictly convex on $U \times V \times V$ for each $\alpha \in A$. We now view the team decision problem (10) as a Stackelberg game where DML, the leader, with cost functional $J(u, v_1, v_2, \alpha^0)$, faces two hypothetical followers DM^s and DM^h with cost functionals $g(u, v_1)$ and $h(u, v_2)$, respectively. The problem faced by DML is to devise a strategy y_1 which would induce DM^s and DM^h to play $v_1 = v^c$, $v_2 = v^c$, simultaneously, where v^c minimizes, together with u^c , the function $J(u, v, \alpha^0)$ given by (8). Naturally, when $v_1 = v_2 = v^c$, the realized value of y_1 should be u^c (which is the side condition); where the pair (u^c, v^c, v^c) jointly minimizes $J(u, v_1, v_2, \alpha^0)$ over $U \times V \times V$. Let us assume that DML observes the decisions of DM^s and DM^h. Let us further assume that DML adopts an affine strategy of the form

$$y_1(v_1, v_2) = u^c + P_1(v_1 - v^c) + P_2(v_2 - v^c), \quad (11)$$

where P_1 and P_2 are $n \times m$ matrices satisfying

$$\nabla_u g(u^c, v^c) P_1 + \nabla_{v_1} g(u^c, v^c) = 0 \quad (12a)$$

and

$$\nabla_u h(u^c, v^c) P_2 + \nabla_{v_2} h(u^c, v^c) = 0 \quad (12b)$$

Note that because of strict convexity of $g(u, v_1)$ and $h(u, v_2)$, these are the necessary and sufficient conditions for $v_1 = v^c$ and $v_2 = v^c$ to minimize $g(u, v_1)$ and $h(u, v_2)$, respectively, when u is given by (11).

Now, since (u^c, v^c) minimizes (8) when $a = a^0$, we first have

$$\nabla_u g(u^c, v^c) = -\alpha^0 \nabla_u h(u^c, v^c) \quad (13a)$$

$$\nabla_{v_1} g(u^c, v^c) = -\alpha^0 \nabla_{v_1} h(u^c, v^c) \quad (13b)$$

and substituting these into (12a) we obtain

$$-\alpha^0 [\nabla_u h(u^c, v^c) P_1 + \nabla_{v_1} h(u^c, v^c)] = 0 \quad (14)$$

which, when compared with (12b), leads to the conclusion that P_1 satisfies the same equation as P_2 does, which is precisely (9). Furthermore, a solution always exists since $\nabla_u h(u^c, v^c) \neq 0$, which was one of our hypotheses in the formulation.

Now, let us suppose that one single player, say DM2, chooses a linear combination of $g(u, v)$ and $h(u, v)$ as his cost functional, with decision space V . More precisely, let

$$J_2(u, v) = k_1 g(u, v) + k_2 h(u, v)$$

where $k_1, k_2 \in \mathbb{R}$ are such that $J_2(u, v)$ is strictly convex on $U \times V$. From the previous discussion, it follows that if DML announces a strategy of the form

$$y_1(v) = u^c + P(v - v^c)$$

where P satisfies (9) or (12a), DM2 is induced to choose $v = v^c$, independent of k_1 and k_2 , as long as $J_2(u, v)$ remains strictly convex on $U \times V$, which establishes the desired result. It is interesting to observe that using sensitivity analysis, the optimal P was found to satisfy

$$\nabla_u h(u^c, v^c) P + \nabla_{v_1} h(u^c, v^c) = 0,$$

whereas the previous discussion shows that P also satisfies

$$\nabla_u g(u^c, v^c) P + \nabla_{v_1} g(u^c, v^c) = 0,$$

as long as $h(u, v)$ and $g(u, v)$ are both strictly convex on $U \times V$.

IV. AN EXAMPLE FROM MICROECONOMICS

In this section we discuss an example from microeconomics, which illustrates some of the results obtained in the previous sections. Let us consider a duopolistic market consisting of two firms DML and DM2, who produce the goods X and Y , respectively. Duopolistic

4

Assuming that the profit functions π_1 and π_2 are identical, and $\alpha^0=1$ (meaning equal shares in the market), we have $P \geq 1$, and (21) becomes $z = x^0 + y^0 + 2(y - y^0)$.

V. EXTENSION TO THE MULTIPARAMETER CASE

In the previous sections, we have restricted our discussion to the case $\alpha \in \mathbb{R}$. When $\alpha \in \mathbb{R}^r$, the first order sensitivity function $I_1(\alpha^0)$ becomes a $(1 \times r)$ vector given by

$$J_1(\alpha^0) = dJ(u^c, v^c, \alpha^0)/d\alpha = (\nabla_u J, P + \nabla_v J) \left(\frac{dv}{d\alpha} \right) \quad (22)$$

where

$$\left(\frac{dv}{d\alpha} \right)^T = [\nabla_{au} J, P + \nabla_{av} J] [P^{-1} \nabla_{uu} J, P + \nabla_{vu} J + \nabla_{uv} J, P + \nabla_{vv} J]^{-1} \quad (23)$$

with $\alpha = \alpha^0$, $u = u^c$, $v = v^c$.

Now, if

$$\nabla_{uv} J(u^c, v^c, \alpha^0) \in R \quad (\nabla_{uu} J(u^c, v^c, \alpha^0)) \quad (24)$$

(where R denotes the range) then it is possible to choose a P such that $\frac{dv}{d\alpha} = 0$ at the nominal solution point. In this case, sensitivity functions of orders 1, 2 and 3 vanish at the nominal solution point; hence, affine strategies have very appealing sensitivity properties in the multiparameter case, too. When condition (24) is not satisfied, one has to minimize a suitable norm of the leading sensitivity function with respect to the $(n \times m)$ matrix P . Since $I_1(\alpha^0)$ vanishes at the nominal solution point, one has to consider the second-order sensitivity functional $I_2(\alpha^0)$ which is an $(r \times r)$ nonnegative definite matrix. A suitable norm for minimization is, in this case, $\text{Tr}(I_2(\alpha^0))$. We are now faced with an unconstrained optimization problem on P , for which a closed-form solution does not in general exist, however, numerically it is a feasible problem. We will illustrate some of these ideas, in the next section, by solving an example that involves arms race between two nations.

VI. AN EXAMPLE FROM THE PROBLEM OF ARMAMENT LIMITATION AND CONTROL

In their papers on armament race and control [Simaan and Cruz (1975a) and Simaan and Cruz (1975b)], Simaan and Cruz have modeled the arms race problem as a noncooperative differential game between two nations. A salient feature of this model is that, when the respective cost functionals are taken to be quadratic in the decision variables, the resulting optimal state trajectory yields a discretized version of the armament model proposed earlier by Richardson [Richardson (1960)]. We will consider here the case when the two nations, DM1 and DM2, have agreed to reduce their respective armament

expenditures. Such a situation inevitably requires the presence of an element of cooperation between DM1 and DM2, since any significant departure from the armament level jointly agreed upon may eventually lead to the original high armament expenditure. Towards the formulation of this problem, let us assume that the goals of DMs can be represented by two objective functionals $J_i(x_1, x_2, u_1, u_2)$, $i=1,2$, wherein DM1 aims to optimize J_1 . In order to incorporate the cooperation element discussed above, we will adopt the Pareto optimal equilibrium concept, which will be realized [Schmitendorf and Leitmann (1974)] if the DMs jointly optimize

$$J(x_1, x_2, u_1, u_2) = k_1 J_1(x_1, x_2, u_1, u_2) + k_2 J_2(x_1, x_2, u_1, u_2) \quad (25)$$

where $k_i \in \mathbb{R}_+$; $u_i \in \mathbb{R}_+$ and $x_i \in \mathbb{R}_+$ denote DM1 and DM2's armament investments, respectively, and x_i represents the armament level of DM1, $i=1,2$, which further satisfies

$$x_i = f_i(x_{i0}, u_i), \quad i=1,2, \quad (26)$$

where $f_i: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a continuous function of x_{i0} and u_i , and is strictly increasing in its second argument. Here, x_{i0} denotes the initial armament level of DM1.

In order to obtain some explicit results, let us adopt the quadratic objective functional model proposed by Simaan and Cruz [Simaan and Cruz (1975a) and Simaan and Cruz (1975b)], because of its analytical tractability and other appealing features in relation with other existing models; namely, let

$$J_i(x_1, x_2, u_1, u_2) = \frac{1}{2} (x_i(u_i - v_i))^2 + q_i^0 (x_i - s_i x_j - v_i)^2; \quad (27)$$

and

$$x_i = f_i(x_{i0}, u_i) = s_i x_{i0} + u_i, \quad i=1,2; \quad j=1,2; \quad i \neq j \quad (28)$$

where

$$x_i > 0, \quad q_i^0 > 0, \quad s_i > 0, \quad 0 < s_i \leq 1, \quad i=1,2.$$

Here, x_{i0} denotes the given initial armament level of DM1 and equation (27) reveals that each DM wants to reduce the gap that exists between his armament level and a linear function of the other DM's armament level, and at the same time to minimize his expenditure. We refer to [Simaan and Cruz (1975b)] for an elaborated interpretation of (27). Under this set-up, there exists a unique pair (u_1^*, u_2^*) , minimizing $J(x_1, x_2, u_1, u_2)$ as a function of $\frac{x_1}{k_1} = x_0$, which corresponds

to the pair (u^c, v^c) in the general discussion of sections II and V.

As it may be the case, one of the DMs, say DM2, may deviate from u_2^* . The reason behind such a move may be that DM2 totally ignores the cooperation, and minimizes his own objective functional. Assuming that each DM can monitor the decisions of his adversary, this situation would immediately give rise to a Nash equilibrium with high

armament expenditures. Since we have assumed that each DM desires to reduce his expenditures while maintaining a certain balance of powers, such a unilateral and large deviation will be unlikely. In its stead, we will assume that DM2 may have an incentive to perform a relatively small deviation from the Pareto equilibrium point, being motivated by one of the following three considerations:

i) DM2 may decide to promote his relative importance in the agreement, which is reflected by an increase of the value of α from α^0 to α^+ , without informing DM1, while DM1 still uses the value α^0 in his objective functional;

ii) DM2 may develop a different perception of the values of one or more coefficients in the team objective functional without informing DM1. Let us assume, for instance, that DM2 has decided to place higher priority and emphasis to reducing the gap between his armament level and the linear functional of DM1's armament level than to minimizing his expenditure; more precisely, that he has decided to increase the value of Q_2 from Q_2^0 to Q_2^+ .

iii) Both (i) and (ii) may be present.

We now analyze these three cases separately.

Case (i). This is similar to the analysis of section IV. The optimal strategy for DM1, which leads to (u_1^*, u_2^*) as final outcome, independent of possible deviations in DM2's perception of α^0 , is given by

$$\gamma_1^{(i)}(u_2) = u_1^* + P^{(i)}(u_2 - u_2^*) \quad (29)$$

where

$$P^{(i)} = \frac{S_2 Q_2 \{ \beta_2 x_{20} + u_1^* - S_2 (\beta_1 x_{10} + u_1^*) - v_2 \}}{R_2 (u_2^* - u_2) + Q_2 (\beta_2 x_{20} + u_1^* - S_2 (\beta_1 x_{10} + u_1^*) - v_2)} \quad (30)$$

Case (ii). This situation corresponds to objective functionals affine in the uncertain parameter (cf. eq. (3)), in which case the analysis of section III prevails. Hence, there exists an optimum robust strategy realizing the team solution independent of DM2's different perceptions of Q_2 , and such a strategy is given by

$$\gamma_1^{(ii)}(u_2) = u_1^* + \frac{1}{S_2 \beta_1} (u_2 - u_2^*) \quad (31)$$

Case (iii). Here, condition (21) is not satisfied. Hence, within the class of affine policies, there does not exist any element which makes the cost that DM1 incurs completely insensitive to discrepancies in DM2's perceptions in more than one parameter. In order to overcome this difficulty, we adopt, as in section V, the scalarized sensitivity function $\gamma_1^{(iii)}(u_2)$, and minimize it subject to the constraint that the strategy of DM1 satisfies

$$\gamma_1^{(iii)}(u_2) = u_1^* + P^{(iii)}(u_2 - u_2^*) \quad (32)$$

This problem can be shown to admit a unique solution which can be obtained explicitly. Hence, when DM1 is uncertain about DM2's perception of both α and Q_2 , there still exists an affine strategy which minimizes an appropriate scalar function representing the sensitivity of DM1's incurred cost with respect to deviations in these coefficients from their nominal values, and such a strategy is given by

$$\gamma_1^{(iii)}(u_2) = u_1^* + P^{(iii)}(u_2 - u_2^*)$$

In the preceding analysis, $P^{(i)}$ is the same coefficient as DM1 would have used in his strategy in a Stackelberg game with DM2 being the follower and DM1 enforcing the point (u_1^*, u_2^*) . On the other hand, in case (ii), by announcing a strategy of the form (31), DM1 makes DM2's objective functional independent of the uncertain coefficient Q_2 . Therefore, DM2's discrepancies do not affect the team solution anymore. However, when the number of uncertain coefficients is large as compared with the dimension of DMs' decision vectors there still exists a compromise, which is to minimize the cumulative effect of variations of uncertain parameters around their nominal value: $\gamma_1^{(iii)}(u_2)$ is designed to perform such a compromise.

REFERENCES

- Marschak, J. and R. Radner (1972). *In Economic Theory of Teams*. Yale University Press, New Haven, CT.
- Canavese, D. H. and T. Basar (1982). A minimum sensitivity approach to incentive design problems. *Proc. 21st IEEE Conf. on Decision and Control*, Orlando, FL.
- Bo, Y. C., P. H. Luh and G. J. Olsder (1982). A control theoretic view on incentives. *Automatica*, Vol. 18, No. 2, 167-180.
- Richardson, L. F. (1960). *In Arms and Insecurity*. The Boxwood Press, Chicago, IL.
- Salman, M. and J. B. Cruz, Jr. (1981). An incentive model of duopoly with government coordination. *Automatica*, Vol. 17, No. 17, 821-830.
- Schmitendorf, W. E. and G. Leitmann (1974). A simple derivation of necessary conditions for Pareto optimality. *IEEE Trans. on Automatic Control*, Vol. AC-19, 601-602.
- Simaan, M. and J. B. Cruz, Jr. (1975a). Formulation of Richardson's model of arms race from a differential game viewpoint. *The Review of Economic Studies*, Vol. XLII, 67-77.
- Simaan, M. and J. B. Cruz, Jr. (1975b). Nash equilibrium strategies for the problem of armament race and control. *Management Science*, Vol. 22, 96-105.
- Zheng, Y. P. and T. Basar (1982). Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: a geometric approach. *International Journal of Control*, Vol. 35, No. 6, 497-511.

ACKNOWLEDGMENT

This work was supported in part by the Office of Naval Research under Contract N00014-82-D-0007 and in part by the Electric Energy Systems Division, Department of Energy under Contract DE-AC01-82-OR21401 with Dynamic Systems, P.O. Box 401, Urbana, IL 61801.

A MINIMUM SENSITIVITY INCENTIVE CONTROL APPROACH TO TEAM PROBLEMS*

Tamer Başar, Jose B. Cruz, Jr., and Derva Cansever
Department of Electrical Engineering and
Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Ave.
Urbana, Illinois 61801

ABSTRACT

In this paper we analyze a class of two-agent team decision problems with a hierarchical decision structure, wherein one of the decision makers may have a slightly different perception of the overall team goal, with this slight variation not known by the other agent who is assumed to occupy the hierarchically dominant position. The leading agent has access to dynamic information and his role is to announce such a policy (incentive scheme) which would lead to achievement of the overall team goal, in spite of the slight variations in the other agent's perception of that goal, which are not known or predictable by him. We may call a policy with such an additional feature a "minimum sensitivity" incentive policy. We obtain, in the paper, "minimum sensitivity" policies for the leading agent, for a general cost functional with convex structure, which are least sensitive to variations in the following agent's perception of the team goal. In some special cases, we show that the robust feature of the incentive scheme is maintained regardless of the magnitude and nature of the variations, and illustrate the theory with an example arising in armament limitation and control.

I. INTRODUCTION

The main characteristic of team decision problems is the presence of several decision makers with a common objective functional which is to be optimized jointly (but possibly in a decentralized fashion) by all decision makers. An underlying stipulation in research on team theory has been the assumption that all agents perceive the common goal in exactly the same way, and face exactly the same mathematical optimization problem [Marschak and Radner (1972)]. In this paper we relax this basic assumption and allow (in the context of two-agent problems) one agent to have a somewhat different perception of the common goal and to quantify it in a slightly different way. Furthermore, we will assume that the other agent is not informed of the existence of this discrepancy in the perception of the common goal, but is able to monitor the decision of the former by occupying a higher (dominant) position in the decision process. The problem we address to is the design of a suitable strategy for the agent who occupies the hierarchically superior position and who still adopts the original team objective functional as his own, such that the change in the minimum value of the team cost because of the discrepancy in the perceptions of the common goal is kept to a minimum. Ideally, the hierarchically superior member of the team would seek not to be affected by this discrepancy, if this is at all possible.

We will approach this problem using optimum incentive design schemes [Ho, Luh, and Disder (1982) and

Zheng and Başar (1982)], which involve a hierarchy in decision making and a suitable information structure for the decision maker at the top of the hierarchy, that allows him to design a policy which in its turn induces the other decision maker with a different objective functional to behave in a desired manner. Recently in [Cansever and Başar (1982)], optimal incentive schemes have been used, within the context of Stackelberg games, to minimize the effect of changes in the parameters of the follower's cost functional on the leader's optimum cost value, by simultaneously achieving a desired goal. Here, we direct our attention to problems which are nominally team, and derive incentive schemes that are least sensitive to deviations in the hierarchically inferior decision maker's perceptions of the uncertain parameters. The fact that the underlying goal is common (that is, the nominal optimization problem is a team problem—a property that may be destroyed in the decision process) can be exploited to obtain very appealing minimum sensitivity strategies, as we will show in the sections to follow.

The problem is formulated in Section II. In Section III we introduce sensitivity functions and obtain robust affine strategies for a general class of convex cost functionals. In Section IV we provide a geometrical interpretation for total insensitivity when the objective functional is affine in the unknown parameter. Section V deals with the generalization of some of these results to the multiparameter case, and Section VI illustrates the basic ideas developed in this paper using a model from armament limitation and control. Concluding remarks of Section VII end the paper.

II. PROBLEM FORMULATION

Consider a two-person deterministic team decision problem in normal form, described by the cost functional $J(v_1, v_2, x)$, where $v_i \in \Gamma_i$ denotes the strategy of DMi (i'th decision maker) and $x \in \mathcal{X}$ is a parameter on which the cost functional depends. Let $u \in U = \mathbb{R}^n$, $v \in V = \mathbb{R}^m$ denote the decision variables of DM1 and DM2, respectively, and assume that $\Gamma_1 = \{v \in V : v = u, u \in U\}$, $\Gamma_2 = V$; i.e. DM1 has access to the decision value of DM2. DM1 also knows the precise value of the parameter x (say x^0), whereas DM2 perceives its value differently (say $\hat{x} \in \mathcal{X}$), which in turn gives rise to a different cost functional from his point of view, namely, $J(\cdot, \cdot, \hat{x}) = J(\cdot, \cdot, x^0) + \Delta J(\cdot, \cdot)$. Furthermore, DM1 does not know the exact value perceived by DM2, but his ultimate goal is to see that the lowest possible value is attained for $J(v_1, v_2, x^0)$. The decision structure of the problem is assumed to be hierarchical, in the sense that DM1 is the dominant decision maker and has the power and means of declaring his policy in advance and enforcing it on the other DM. Hence, while DM2 is faced with the problem of minimizing $J(v_2, v_1, \hat{x})$ over $v_2 \in V$, DM1 wishes to choose $v_1 \in \Gamma_1$ (in total ignorance of $\hat{x}$) that would eventually lead to a minimum value for $J(v_1, v_2, x^0)$.

By an abuse of notation, let $J(v_1, v_2, x)$ denote the cost functional on the product space $U \times V$, for each

*This work was supported by the Office of Naval Research under Contract N00014-82-K-0469.

$x \in A$, and assume that this functional is strictly convex on $U \times V$, for each $x \in A$, is twice continuously differentiable in its first two arguments and continuously differentiable in its third argument. Furthermore, let us denote the unique minimum of $J(u, v, x^0)$ by $(u^*, v^*) \in U \times V$. Restricting DM1 to affine policies in π_1 , we first note that the policy

$$\pi_1(v) = u^* + P(v - v^*), \quad (1)$$

where P is an $(n \times m)$ -matrix, has the appealing property that if DM2's perception of x is x^0 , then $\min_{v \in V} J(\pi_1(v), v, x^0)$ leads to the desired value $v^* \in V$ for any matrix P . If $x^* \neq x^0$, however, the problem ceases to be a cooperative one since the problem faced by DM2 is

$$\min_{v \in V} J(\pi_1(v), v, x^*) \quad (2a)$$

whose minimizing solution (say $v^+ \in V$) satisfies (and is uniquely determined by) the equation

$$J_u(u^*, v^+, x^*)P + J_v(u^*, v^+, x^*) = 0 \quad (2b)^1$$

where $u^+ = \pi_1(v^+) = u^* + P(v^+ - v^*)$ and is not necessarily the same as u^* . The problem we address, in the sequel, is whether it is possible to choose a robust policy π_1 (by choosing P appropriately) so that either $u^+ = u^*$ and $v^+ = v^*$, or the discrepancies will be small whenever x^+ is close to x^0 ; in other words, we seek either total insensitivity or minimum sensitivity of the optimum value of $J(u, v, x^0)$ to variations in the perception of DM2 (of x) by a proper choice of π_1 .

III. INTRODUCTION OF A SENSITIVITY FUNCTION AND DERIVATION OF MINIMUM SENSITIVITY SOLUTIONS

As a measure of the sensitivity of $J(u, v, x^0)$ with respect to deviations in the perception of DM2 of x from its nominal value x^0 , let us introduce the total derivative of $J(u, v, x^0)$ with respect to x , when $u = u^+$, $v = v^+$, satisfying (2b), and at the point $x^+ = x^0$. We call this function the *first-order sensitivity function* of $J(u, v, x^0)$ with respect to x , at $x = x^0$, in view of (1) and the optimal response of DM2 as characterized (uniquely) by (2b):

$$\begin{aligned} I_1(x^0) &= dJ(u^+, v^+, x^0)/dx^+_{x^+=x^0} \\ &= (dJ/dv^+)(dv^+/dx^+)_{x^+=x^0} \\ &= [J_u(u^+, v^+, x^0)P + J_v(u^+, v^+, x^0)]v^*_1, \end{aligned} \quad (3)$$

where

$$v^*_1 = dv^+/dx^+_{x^+=x^0}$$

and is determined from (2b). To obtain an expression for v^*_1 , we note that (2b) is in fact an identity for all $x^+ \in A$, since it uniquely determines the optimal response of DM2 to the announced policy (1) of DM1, with his perceived value for x being x^+ . Hence, differentiating (2b) with respect to x^+ , and evaluating the resulting expression at $x^+ = x^0$, we obtain

$$[P'J_{uu}P + P'J_{uv} + J_{uv}P + J_{vv}]v^*_1 + [J_{xu}P + J_{xv}] = 0 \quad (4)$$

whereby

$$v^*_1 = -[P'J_{uu}P + P'J_{uv} + J_{uv}P + J_{vv}]^{-1}[J_{xu}P + J_{xv}] \quad (5)$$

¹Here J_u and J_v are row vectors of dimensions $1 \times n$ and $1 \times m$, respectively, denoting the partial derivatives with respect to the corresponding decision variables.

where the arguments are evaluated at $u = u^*$, $v = v^*$, $x = x^0$. Note that the required inverse in (5) exists under the initial hypothesis that J is strictly convex in (u, v) for all $x \in A$.

Now, since the pair (u^*, v^*) globally minimizes $J(u, v, x^0)$, we already know that

$$J_u(u^*, v^*, x^0) = 0, \quad J_v(u^*, v^*, x^0) = 0, \quad (6)$$

in view of which the first product term of (3) and hence $I_1(x^0)$ vanish. Then, the dominating term in the Taylor expansion of $J(u^*, v^+, x^0)$ around $x^+ = x^0$ is determined by the second-order sensitivity function:

$$\begin{aligned} I_2(x^0) &= d^2J(u^+, v^+, x^0)/dx^{+2}_{x^+=x^0} \\ &= [(dv^+/dx^+)(d^2J/dv^{+2})(dv^+/dx^+) \\ &\quad + (dJ/dv^+)(d^2v^+/dx^{+2})]_{x^+=x^0} \\ &= v^{*T}_1 [P'J_{uu}P + P'J_{uv} + J_{uv}P + J_{vv}]v^*_1 + [J_{xu}P + J_{xv}]v^*_1. \end{aligned} \quad (7)$$

Since the second term is zero, in view of (6), $I_2(x^0)$ vanishes if and only if $v^*_1 = 0$, which requires from (5) that there exist a P satisfying

$$J_{xu}(u^*, v^*, x^0)P + J_{xv}(u^*, v^*, x^0) = 0 \quad (8)$$

A sufficient condition for this is, of course,

$$J_{xu}(u^*, v^*, x^0) \neq 0, \quad (9)$$

which is also necessary if the second term in (8) does not vanish (at least one component is nonzero).

When v^*_1 vanishes, not only the second-order sensitivity function, but also the third-order sensitivity function

$$I_3(x^0) = d^3J(u^+, v^+, x^0)/dx^{+3}_{x^+=x^0} \quad (10)$$

vanishes, because it carries (by chain rule of differentiation) only terms that involve either v^*_1 or $[J_{xu}P + J_{xv}]$ as products. Hence, under the condition that (8) admits at least one solution, and when DM1 employs the corresponding policy (1), if DM2's perception x^+ (of x) stays within an ϵ -neighborhood of its nominal value x^0 , the 3rd order Taylor approximation of the effect of this discrepancy is zero. We now summarize this appealing feature of the linear policy (1) in the following proposition.

Proposition 1: Let condition (9) be satisfied, and let P^* denote a solution to (8). Then, if DM1 employs the policy

$$\pi_1^*(v) = u^* + P^*(v - v^*), \quad (11a)$$

and the unique minimizing decision of DM2 is

$$v^+(x^+) = \arg \min_{v \in R^m} J(\pi_1^*(v), v, x^+), \quad (11b)$$

$J(\pi_1^*(v^+), v^+, x^0)$ agrees with $J(u^*, v^*, x^0)$ to third order in x^+ when it lies in a sufficiently small neighborhood of x^0 . Equivalently, the discrepancy in costs is of fourth order.

When the objective function $J(u, v, x)$ is affine in x , we can obtain more explicit results. Specifically, let

$$J(u, v, x) = z(u, v) + xh(u, v) \quad (12a)$$

where g and h are continuously differentiable in their arguments,

$$h_u(u^c, v^c) \neq 0, \quad (12b)$$

and J is strictly convex in (u, v) for all $\lambda \in A$. Then, (9) reads

$$h_u(u^c, v^c)P + h_v(u^c, v^c) = 0, \quad (13)$$

a solution to which always exists because (12b) becomes equivalent to (9). Hence v_1^c , as given by (5) [evaluated at $\lambda = \lambda^0$] is zero. This, in turn, implies through an iterative verification that the vector $dnv^+(x^+)/dx^+$, where $v^+(x^+)$ is given by (11b), vanishes at $x^+ = \lambda^0$, for all $n = 1, 2, \dots$, simply because the second term in (4)

$$J_{uu}P + J_{uv} = h_u(u, v)P + h_v(u, v)$$

is not explicitly dependent on λ . Since the n th order sensitivity function

$$I^n(\lambda^0) = d^n J(u^+, v^+, \lambda^0) / d\lambda^{+n} \quad (14)$$

carries only $(dv^+(x^+)/dx^+)^i$, $i = 1, 2, \dots, n$ as product terms, which are all zero whenever P is chosen to satisfy (13), it follows that sensitivity functions of all orders vanish, at $\lambda^+ = \lambda^0$. Hence,

Proposition 2: When the objective function is given by (11a), under the condition (12b), let P^* be any solution of (13). Then, if (11a) is employed by DM1, the response of DM2 (i.e., (11b)) is independent of λ^+ , and $v^+ = v^c$. Hence, $J(u^+, v^+, \lambda^0) = J(u^c, v^c, \lambda^0)$ for all $\lambda^+ \in A$, that is the overall performance is independent of the perception of DM2 regarding the value of λ . $\square$

In the next section we provide a geometric interpretation of this appealing feature of the linear policy when the cost function is an affine function of the parameter λ .

IV. GEOMETRIC INTERPRETATION OF TOTAL INSENSITIVITY WHEN THE OBJECTIVE FUNCTIONAL IS AFFINE IN A PARAMETER

Let the objective function J be as given by (12a) with h satisfying condition (12b). Since J is strictly convex, the team solution (u^c, v^c) when $\lambda = \lambda^0$ is obtained (uniquely) from

$$g_u(u^c, v^c) + \lambda^0 h_u(u^c, v^c) = 0 \quad (15a)$$

$$g_v(u^c, v^c) + \lambda^0 h_v(u^c, v^c) = 0. \quad (15b)$$

Postmultiplying (15a) by P , adding this to (15b), and taking the transpose, we have

$$(P'g_u' + g_v') + \lambda^0 (P'h_u' + h_v') = 0. \quad (16)$$

Pictorially, the vectors $(P'g_u' + g_v')$ and $(P'h_u' + h_v')$ are oppositely oriented when λ^0 is a positive scalar. Clearly, λ^0 is the ratio of the magnitude of the vector $(P'g_u' + g_v')$ to the magnitude of the vector $(P'h_u' + h_v')$. If DM1 chooses P such that (13) is satisfied, then the magnitudes of both vectors become zero. In this case, if λ^0 is replaced by $\lambda \neq \lambda^0$ in (16), the equation would still hold, and (u^c, v^c) satisfies

$$(P'g_u' + g_v') + \lambda (P'h_u' + h_v') = 0. \quad (17)$$

Since (17) is the condition used by DM2 to optimize v (see also (2b)), he will choose $v = v^c$, no matter what his perceived value of λ is. Thus, DM1 achieves the

team-optimal solution for minimizing $J(u, v, \lambda^0)$ by choosing P such that P' transforms the vector h_u^0 to $(-h_v^0)$:

$$P'h_u^0(u^c, v^c) = -h_v^0(u^c, v^c). \quad (18a)$$

This same choice of P' transforms g_u^0 to $-g_v^0$:

$$P'g_u^0(u^c, v^c) = -g_v^0(u^c, v^c). \quad (18b)$$

V. EXTENSION TO THE MULTIPARAMETER CASE

In the previous sections, we have restricted our discussion to the case $\lambda \in \mathbb{A} \subset \mathbb{R}$. When $\mathbb{A} \subset \mathbb{R}^r$, the first-order sensitivity function $I_1(\lambda^0)$ becomes a $(r \times r)$ vector given by

$$I_1(\lambda^0) = dJ(u^+, v^+, \lambda^0) / d\lambda^+ \quad (19a)$$

where

$$v_1^c = -(P'J_{uu}P + P'J_{vu} + J_{uv}P + J_{vv})^{-1} (J_{u\lambda}P + J_{v\lambda}). \quad (19b)$$

and the arguments are evaluated at $u = u^c$, $v = v^c$, $\lambda = \lambda^0$. Note that $I_1(\lambda^0) = 0$, in view of (6). Furthermore, $v_1^c = 0$ (zero matrix) if

$$J_{v\lambda}(u^c, v^c, \lambda^0) \in \mathcal{R}(J_{uu}(u^c, v^c, \lambda^0)) \quad (20)$$

(where $\mathcal{R}$ denotes the range) since then it is possible to find an $(n \times m)$ matrix P to make the second product term of (19b) zero. In this case, sensitivity functions of orders 1, 2, and 3 vanish at the nominal solution point; hence, affine policies have very appealing sensitivity properties also in the multiparameter case. When condition (20) is not satisfied, however, one has to minimize a suitable norm of the leading sensitivity function with respect to the $(n \times m)$ matrix P . This is, in general, the second-order sensitivity functional $I_2(\lambda^0)$ which is an $(r \times r)$ nonnegative definite matrix. A suitable norm for minimization is, in this case, $\text{Tr } I_2(\lambda^0)$. We are now faced with an unconstrained optimization problem on P , for which a closed-form solution does not in general exist; however, numerically it is a feasible problem.

When the objective function J is affine in the parameter vector $\lambda \in \mathbb{R}^r$, a total insensitivity result could be established under certain conditions, by a direct extension of the discussion of Section IV. Towards this end, let

$$J(u, v, \lambda) = g(u, v) + \lambda'h(u, v) \quad (21)$$

where $g: U \times V \rightarrow \mathbb{R}$, $h: U \times V \rightarrow \mathbb{R}^r$, J is strictly convex and continuously differentiable in (u, v) . Then, the optimality conditions for $\lambda = \lambda^0$ are

$$g_u(u^c, v^c) + \lambda^0' h_u(u^c, v^c) = 0 \quad (22a)$$

$$g_v(u^c, v^c) + \lambda^0' h_v(u^c, v^c) = 0 \quad (22b)$$

where h_u (respectively, h_v) is an $r \times n$ (respectively, $r \times m$) matrix. The optimal response of DM2, under the policy (1) for DM1, is determined uniquely from (for a general λ)

$$(P'g_u' + g_v') + \sum_{i=1}^r \lambda_i (P'h_{ui}' + h_{vi}') = 0 \quad (23)$$

where subscript i denotes the i 'th component of the corresponding vector. Now, let us assume that there exists an $n \times n$ matrix P satisfying simultaneously

$$h_{1i}(u^c, v^c)P + h_{2i}(u^c, v^c) = 0, \quad i=1, \dots, r. \quad (24)$$

Under this condition, the second term in (23) vanishes at $v=v^c$, for all $u \in AC R^r$, and furthermore the first term also vanishes in view of (22a)-(22b), by basically following the argument of Section IV. Hence, under this particular choice of P , $v=v^c$ is the unique solution to (23) for all values of u ; that is, the optimal response of DM2 is independent of his perception of the value of u , provided that strict convexity of J is preserved. To summarize,

Proposition 3: When the objective function is given by (21), let there exist a solution to (24), to be denoted P^* . Then, if the policy

$$v^*(v) = u^c + P^*(v - v^c)$$

is employed by DM1, the response of DM2 (which is (11b) with $u \in R^r$) is independent of u^* , and $v^* = v^c$. Consequently, $J(u^*, v^*, u^c) = J(u^c, v^c, u^c)$ for all $u \in AC R^r$.

In the next section we consider an example that involves arms race between two nations, which serves to illustrate some of the ideas generated in this and the previous sections. Another example from micro-economics can be found in [Cansever, Başar, and Cruz (1983)].

VI. AN EXAMPLE FROM THE PROBLEM OF ARMAMENT LIMITATION AND CONTROL

In their papers on armament race and control [Simaan and Cruz (1975a) and Simaan and Cruz (1975b)], Simaan and Cruz have modeled the arms race problem as a noncooperative differential game between two nations. A salient feature of this model is that, when the respective cost functionals are taken to be quadratic in the decision variables, the resulting optimal state trajectory yields a discretized version of the armament model proposed earlier by Richardson [Richardson (1960)]. We will consider here the case when the two nations, DM1 and DM2, have agreed to reduce their respective armament expenditures. Such a situation inevitably requires the presence of an element of cooperation between DM1 and DM2, since any significant departure from the armament level jointly agreed upon may eventually lead to the original high armament expenditure. Towards the formulation of this problem, let us assume that the goals of the DM's can be represented by two objective functionals $J_i(x_1, x_2, u_1, u_2)$, $i=1,2$, wherein DM1 aims to optimize J_1 . In order to incorporate the cooperation element discussed above, we will adopt the Pareto optimal equilibrium concept, which will be realized [Schmitendorf and Leitmann (1974)] if the DM's jointly optimize

$$J(x_1, x_2, u_1, u_2) = k_1 J_1(x_1, x_2, u_1, u_2) + k_2 J_2(x_1, x_2, u_1, u_2) \quad (25)$$

where $k_i \in R_+$, $u_i \in R_+$, and $u_i \in R_+$ denote DM1 and DM2's²⁾ armament investments, respectively, and x_i represents the armament level of DM1, $i=1,2$, which further satisfies

$$x_i = f_i(x_{i0}, u_i), \quad i=1,2, \quad (26)$$

where $f_i: R_+ \times R_+ \rightarrow R_+$ is a continuous function of x_{i0} and u_i , and is strictly increasing in its second argument. Here, x_{i0} denotes the initial armament level of DM1.

In order to obtain some explicit results, let us adopt the quadratic objective functional model proposed by Simaan and Cruz [Simaan and Cruz (1975a) and Simaan and Cruz (1975b)], because of its analytical tractability and other appealing features in relation with other existing models; namely, let

$$J_i(x_1, x_2, u_1, u_2) = \frac{1}{2} R_i(u_i - w_i)^2 + Q_i^0(x_i - S_i x_j - v_i)^2 \quad (27)$$

and

$$x_i = f_i(x_{i0}, u_i) = \beta_i x_{i0} + u_i, \quad i=1,2; j=1,2; i \neq j \quad (28)$$

where

$$R_i > 0, \quad Q_i^0 \geq 0, \quad S_i \geq 0, \quad 0 \leq \beta_i \leq 1, \quad i=1,2.$$

Here, x_{i0} denotes the given initial armament level of DM1 and expression (27) reveals the fact that each DM wants to reduce the gap that exists between his armament level and a linear function of the other DM's armament level, and at the same time wishes to minimize his expenditure. We refer to [Simaan and Cruz (1975b)] for an elaborated interpretation of (27). Under this set up, there exists a unique pair (u_1^*, u_2^*) , minimizing $J(x_1, x_2, u_1, u_2)$ as a function of $\frac{x_2}{x_1} = \beta^2$, which corresponds to the pair (u^c, v^c) in the general discussion of Sections II and V.

As it may be the case, one of the DM's, say DM2, may deviate from u_2^* . The reason behind such a move may be that DM2 totally ignores the cooperation, and minimizes his own objective functional. Assuming that each DM can monitor the decisions of his adversary, this situation would immediately give rise to a Nash equilibrium with high armament expenditures. Since we have assumed that each DM desires to reduce his expenditures while maintaining a certain balance of powers, such a unilateral and large deviation will be unlikely. In its stead, we will assume that DM2 may have an incentive to perform a relatively small deviation from the Pareto equilibrium point, being motivated by one of the following three considerations:

- i) DM2 may decide to promote his relative importance in the agreement, which is reflected by an increase in the value of β from β^0 to β^* , without informing DM1, while DM1 still uses the value β^0 in his objective functional;
- ii) DM2 may develop a different perception of the values of one or more coefficients in the team objective functional without informing DM1. Let us assume, for instance, that DM2 has decided to place higher priority and emphasis on reducing the gap between his armament level and the linear functional of DM1's armament level than on minimizing his expenditure; more precisely, that he has decided to increase the value of Q_2 from Q_2^0 to Q_2^* .
- iii) Both i) and ii) may be present.

We now analyze these three cases separately.

Case i). This is similar to the analysis of Section IV. The optimal strategy for DM1, which leads to (u_1^*, u_2^*) as final outcome, independent of possible deviations in DM2's perception of β^0 , is given by

²⁾ R_+ denotes the positive real line.

$$p^{(i)}(u_2) = u_1^t + p^{(i)}(u_2 - u_1^t) \quad (29)$$

where

$$p^{(i)} = \frac{S_2 Q_2 (S_2 x_{20} + u_1^t - S_2 (-x_1 x_{10} + u_1^t) - v_2)}{R_2 (u_2^t - w_2) + \lambda_2 (-x_2 x_{20} + u_2^t - S_2 (-x_1 x_{10} + u_1^t) - v_2)} \quad (30)$$

Case 1b. The solution here again follows from the analysis of Section IV. Hence, there exists an optimum insensitivity strategy realizing the team solution independent of DM2's different perceptions of Q_2 , and such a strategy is given by

$$p^{(ii)}(u_2) = u_1^t + \frac{1}{S_2 - 1} (u_2 - u_1^t). \quad (31)$$

Case 1c. This case involves multiparameters where condition (20) is not satisfied. Hence, within the class of affine policies, there does not exist any element which makes the cost of DM1 completely insensitive to discrepancies in DM2's perceptions in more than one parameter. In order to overcome this difficulty, we adopt, as discussed in Section V, the scalarized sensitivity function $\text{Tr } I_2(v^0, Q_2^0)$, and minimize it subject to the constraint that the strategy of DM1 is given by

$$p^{(iii)}(u_2) = u_1^t + p^{(iii)}(u_2 - u_1^t). \quad (32)$$

This problem can be shown to admit a unique solution which can be obtained explicitly. Hence, when DM1 is uncertain about DM2's perception of both λ and λ_1 , there still exists an affine strategy which minimizes an appropriate scalar function representing the sensitivity of DM1's incurred cost with respect to deviations in these coefficients from their nominal values, and such a strategy is given by (32).

In the preceding analysis, $p^{(i)}$ is the same coefficient as DM1 would have used in his strategy in a Stackelberg game with DM2 being the follower and DM1 enforcing the point (u_1^t, u_2^t) . On the other hand, in case 1b, by announcing a strategy of the form (31), DM1 makes DM2's objective functional independent of the uncertain coefficient Q_2 . Therefore, DM2's discrepancies do not affect the team solution anymore. However, when the number of uncertain coefficients is large as compared with the dimension of DMs' decision vectors, there still exists a compromise, which is to minimize the cumulative effect of variations of uncertain parameters around their nominal values: $p^{(iii)}(u_2)$ is designed to perform such a compromise.

VII. CONCLUDING REMARKS

In this paper we have introduced the notion of optimum minimum sensitivity incentive policies in team decision problems wherein one member of the team has a somewhat different perception of the common goal than the other one, and we have derived explicit incentive policies which render the incurred value of the team objective functional least sensitive to, and in some cases even independent of, the discrepancies described above.

One field where this notion finds application is the military Command, Control, and Communications (C3) systems area. Here, there exist multiple decision makers (DM's) and multiple hierarchies in decision making, and the role of each DM is not only to issue orders to be executed by the DM's occupying the lower levels of hierarchy, but also to coordinate the actions

of his subordinates. There is an underlying goal, or objective, which involves a successful completion of a mission or task (such as multi-object tracking and fire control), and this goal is determined by the DM's at the top of the hierarchy in rather general terms (i.e., not in fine detail), which is then transmitted to the relevant DM's at the lower levels.

Hence, in a general framework, a C3 system involves a team of DM's who act in an uncertain environment, and who have limitations on control and communication capabilities. However, realistically, this is not strictly a team problem, because, in an uncertain environment, it is unlikely that every DM will develop precisely the same perception of the ultimate goal in every fine detail. In fact, in order to model C3 systems as team problems, it is absolutely necessary that all DM's have exactly the same perception of an existing common goal and quantify this perception in exactly the same way. Any discrepancy that exists between the perceptions of the DM's on the underlying common goal will lead to a decision problem which cannot be treated as a team problem, and optimal decision rules derived by totally ignoring (or overlooking) this aspect of the problem are apt to lead to outcomes which are extremely sensitive even to small variations in the perceptions with regard to real underlying goals of the mission. The approach developed in this paper remedies this deficiency because it takes into account the possibility that the DMs' perceptions of the "team goal" may deviate from the nominal set by the highest level decision making unit.

Two possible extensions of the general approach of this paper are to dynamic multi-stage decision problems and to stochastic team problems. In the latter case a natural source of discrepancy is the a priori statistical information which is normally assumed to be shared by the DM's. A recent reference [Başar (1983)] addresses the question of existence of suitable equilibrium solutions for such problems when there is discrepancy in the subjective probability measures characterizing the probability space. Derivation of minimum sensitivity incentive policies in this context is currently under study.

REFERENCES

- Başar, T. (1983). Decision problems with multiple probabilistic models. Proc. American Control Conf., San Francisco, CA, pp. 1091-1096.
- Cansever, D. H. and T. Başar (1982). A minimum sensitivity approach to incentive design problems. Proc. 21st IEEE Conf. on Decision and Control, Orlando, FL, pp. 158-163.
- Cansever, D. H., T. Başar, and J. B. Cruz, Jr. (1983). Robustness of incentive policies in team problems with discrepancies in goal perceptions. Proc. 4th IFAC/IFORS Conf. on the Modelling and Control of National Economies, Washington, D.C., pp. 175-180.
- Ho, Y. C., P. H. Lun, and J. L. Dister (1982). A control theoretic view on incentives. Automatica, Vol. 18, No. 2, pp. 157-180.
- Marschak, J. and R. Radner (1972). Economic Theory of Teams, Yale University Press, New Haven, CT.
- Richardson, L. F. (1960). Arms and Insecurity. The Boxwood Press, pp. 120, 121.

Salman, M. and J. B. Cruz, Jr. (1981). An incentive model of duopoly with government coordination. Automatica, Vol. 17, pp. 811-830.

Schmitendorf, W. E. and G. Leitmann (1974). A simple derivation of necessary conditions for Pareto optimality. IEEE Trans. on Automatic Control, Vol. AC-19, pp. 601-602.

Simaan, M. and J. B. Cruz, Jr. (1975a). Formulation of Richardson's model of arms race from a differential game viewpoint. The Review of Economic Studies, Vol. XLII, pp. 67-77.

Simaan, M. and J. B. Cruz, Jr. (1975b). Nash equilibrium strategies for the problem of armament race and control. Management Science, Vol. 22, pp. 96-105.

Zheng, Y. P. and T. Başar (1982). Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: A geometric approach. International Journal of Control, Vol. 35, No. 6, pp. 997-1011.

ROBUST INCENTIVE POLICIES FOR STOCHASTIC DECISION PROBLEMS IN THE PRESENCE OF
PARAMETRIC UNCERTAINTY†

Tamer Başar and Derya H. Cansever

Decision and Control Laboratory, Coordinated Science Laboratory, University of Illinois,
1101 W. Springfield Avenue, Urbana, Illinois 61801, U.S.A.

Abstract. In this paper we consider a general class of stochastic incentive decision problems in which the leader has access to the control value of the follower and to private as well as common information on the unknown state of nature. The follower's cost function depends on a finite number of parameters whose values are not known accurately by the leader, and in spite of this parametric uncertainty the leader seeks a policy which would induce the desired behavior on the follower. We obtain such robust policies for the leader, which are smooth, induce the desired behavior at the nominal values of these parameters, and furthermore make the follower's optimal reaction either minimally sensitive or totally insensitive to variations in the values of these parameters from the nominals. The general solution is determined by some orthogonality relations in some appropriately constructed (probability) measure spaces, and leads to particularly simple incentive policies. The features presented here are intrinsic to stochastic decision problems and have no counterparts in deterministic incentive problems.

Keywords. Stochastic systems; economic systems; team theory; decision theory; game theory; optimization.

1. INTRODUCTION

In this paper we consider a general class of stochastic Stackelberg game problems (equivalently, incentive decision problems, in our context) in which the leader has access to the control value of the follower and to private as well as common information on the unknown state of nature, whereas the follower has access to only common information which is shared by the leader. It is further assumed that the follower's cost function (which is strictly convex, but not necessarily quadratic) depends on a number of parameters whose values are not known accurately by the leader. The objective is to obtain robust incentive policies (decision rules) for the leader which would induce the desired (by the leader) behavior on the follower at the nominal values of these parameters, and be minimally sensitive to deviations in the values of these parameters.

For a rough mathematical (symbolic) description of the problem, let U and V be the decision spaces of the leader and the follower, respectively, $x \in X$ denote the random state of nature, and $z \in Z$ and $y \in Y$ denote the common (to both DMs) and private (to the leader) information related to x . We are assuming at this point that X , Z and Y are endowed with sufficiently rich topology so that probability measures can be defined on their subsets. Let Γ_f denote the class of all mappings from Z into V (satisfying some smoothness conditions to be delineated later), and Γ be the class of all mappings from $V \times Z \times Y$ into U . Furthermore let Γ^s be the class of mappings from $Z \times Y$ into U . We will denote generic elements of Γ^s , Γ_f and Γ by u , v and γ , respectively, so that $u = u(z, y)$, $v = v(z)$,

$\gamma = \gamma(u, z, y)$. Let us assume that there is a point (u^c, v^c) in $\Gamma^s \times \Gamma_f$ which is most desirable by the leader (such as a system trajectory or control trajectory) and he seeks to determine a policy $\gamma \in \Gamma$ which would force the follower to such an action that the resultant outcome (u, v) in the product space $\Gamma^s \times \Gamma_f$ is sufficiently close to (u^c, v^c) , by also taking into consideration the fact that the cost function of the follower is not known accurately.

Let $L(u, v, x, \alpha)$ denote the loss function of the follower, depending on a set of parameter values $\alpha \in A$. Let

$$J(u, v, \alpha) = E\{L(u(z, y), v(z), x, \alpha)\}$$

where E is the expectation over the statistics of the random variables x , z and y . (Note that we have abused the notation here and have used the same notation for both random variables (or vectors) and their realizations.) Likewise we introduce

$$J(\gamma, v, \alpha) = E\{L(\gamma(v(z), z, y), v(z), x, \alpha)\}$$

which we call the cost function of the follower. We assume that, for each $\alpha \in A$ and $x \in X$, $L(u, v, x, \alpha)$ is strictly convex in the pair $(u, v) \in \Gamma^s \times \Gamma_f$, and furthermore, Γ is structured so that for each $v \in \Gamma_f$ and $\alpha \in A$, $J(\gamma, v, \alpha)$ is strictly convex in v .

Under these assumptions, to each $v \in \Gamma_f$, and for fixed $\alpha \in A$, there corresponds a unique element of Γ_f , called v_γ , which will minimize $J(v, v, \alpha)$ over $v \in \Gamma_f$. [We are assuming here that, as a rational decision maker, the follower chooses his optimal policy (or rational reaction) by minimizing the function $J(\gamma, v, \alpha)$ where α is the true value characterizing his cost functional.] Hence, we have the unique correspondence (for each fixed $\alpha \in A$)

$$v \xrightarrow{S'_\alpha} v_\gamma$$

where the mapping S'_α depends explicitly on α and involves a minimization operation.

†Research reported herein was supported in part by the Office of Naval Research under Contract N00014-82-K-0469; and in part by the U.S. Department of Energy, Electric Energy Systems Division, under Contract DE-AC01-81RA-50658, with Dynamic Systems, P. O. Box 423, Urbana, Illinois 61801.

Let us further note that this unique relationship necessarily yields a unique element in Γ^s , given by

$$u(z, y) = \gamma(v(z), z, y),$$

and hence we can associate, with each $\gamma \in \Gamma$, a unique pair (u, v) :

$$\gamma \xrightarrow{\beta_2} (u, v)$$

where the mapping is denoted, in this case, by β_2 .

This is the familiar input-output relationship of system theory (with the mapping β_2 being much more complicated than the one normally encountered in system theory) and hence we can call Γ the input space and $U \times V$ the output space. In terms of the familiar jargon of system theory we can rephrase the problem posed in the earlier part of this section as follows:

Problem A

For a given nominal value $\alpha^* \in A$, find an input (control) γ which would drive the system output to a desired value (u^c, v^c) . [Note that the desired output is in fact a stochastic variable (or process); hence what we pose here is akin to stochastic controllability.]

The solution to Problem A has in fact been obtained in [1]; an important feature is that it is generally nonunique, even in the class of policies (for the leader) which are linear in v . This then prompts a second question (additional design criterion) related to the robustness of the "optimal" γ :

Problem B

Among the control inputs which solve Problem A, which one (ones) leads (lead) to outputs that are least sensitive to variations in the value of α from the nominal value α^* ?

In this paper we obtain a complete solution to this problem by using some ideas originally introduced in [2] for the deterministic version of the problem. The features of the solution for the stochastic problem are, however, inherently different from those of the deterministic problem, which will be pointed out as we go along.

In Section II we will present a complete solution to the scalar problem, while extension to the vector case has been included in the fuller version [3]. The theory developed in Section II is illustrated in Section III via a numerical example motivated by a problem that arises in the control of large organizations [5]. Section IV discusses possible extensions to other classes of related problems.

II. MAIN RESULTS

In this section we assume that the spaces U, V, X, Y, Z are 1-dimensional Euclidean, and α is a single parameter with nominal value (as perceived by the leader) α^* . The actual value of α may be in a small neighborhood of α^* , and we assume that $L(u, v, x, y)$ is strictly convex in (u, v) , and twice continuously differentiable in u, v and α , when α is restricted to lie in this neighborhood. Furthermore, we assume that every $\gamma \in \Gamma$ is continuously differentiable in v . The random variables x, y, z are jointly second-order random variables, and v , in addition to being continuously differentiable in v , is measurable in z and y , and u, v are also measurable in their arguments. Furthermore, if L is measurable in x , the expectation of L , written as

$$E\{L(v(z), z, y), v(z), x, \alpha)\} \quad (2.1)$$

is well-defined for all $\alpha \in \Gamma$, $y \in \Gamma_y$, $v \in \Gamma_v$ and for every x in a neighborhood of α^* . We finally assume that a desirable point in the product space $\Gamma^s \times \Gamma_v$, as chosen by the leader, is (u^c, v^c) .

It indeed takes the value α^* , the results of [1], when specialized to this scalar case, indicate that

this point (u^c, v^c) is induced by the leader by a linear (in v) policy

$$\gamma(v, z, y) = u^c(z, y) - Q(z, y) [v - v^c(z)] \quad (2.2)$$

whenever

$$E\{L_u(u^c(z, y), v^c(z), x, \alpha^*)\} \neq 0 \quad (2.3)$$

with positive probability in (y, z) . Since we have a simpler (scalar) problem here, the equation for Q can be obtained directly. Substituting (2.2) into (2.1) we have

$$E\{E\{L(u^c - Q[v - v^c], v, x, \alpha^*)\}\} \quad (2.4)$$

which is strictly convex in v since L was strictly convex in (u, v) and γ is linear in v . Hence (2.4) admits a unique minimum in Γ_v , which is obtained by differentiating the inner expression with respect to v , for fixed z , and setting equal to zero:

$$E\{L_u Q - L_v\} = 0$$

$$E\{Q(z, y) E\{L_u(u^c - Q[v - v^c], v, x, \alpha^*)\}\} \quad (2.5)$$

This is the equation that the minimizing v would satisfy (this is also sufficient because of strict convexity), and it is in general nonlinear in v . However, we do not need to solve this for v , but simply find a Q such that its unique solution is v^c —which would then turn yield $\gamma(v^c, z, y) = u^c$, in view of (2.2), and thereby the desired solution would be induced. Now, substituting $v = v^c$ in (2.5) we obtain

$$E\{Q(z, y) F(z, y)\} = G(z) \quad (2.6)$$

where

$$F(z, y) = E\{L_u(u^c(z, y), v^c(z), x, \alpha^*)\} \quad (2.7)$$

$$G(z) = E\{L_v(u^c(z, y), v^c(z), x, \alpha^*)\}$$

If this had been a deterministic problem, the solution to (2.6) would be unique, thus outruling any possibility of obtaining optimal policies that satisfy other design criteria. For the stochastic problem, however, (2.6) admits infinitely many solutions, with a family of such solutions being (which turns out to be sufficiently rich):

$$Q(z, y) = g(z, y) G(z) / [E\{g(z, y) F(z, y)\}] \quad (2.8)$$

where g is any function measurable in (z, y) satisfying the condition

$$E\{g(z, y) F(z, y)\} \neq 0, \quad z \in \mathbb{R} \quad (2.9)$$

Verification of this result is by direct substitution of (2.8) into (2.6). We now summarize this result below:

Theorem 1. Problem A formulated in Section I admits, for the scalar version, infinitely many linear (in v) solutions, with one such family given by (2.2), (2.8), under condition (2.9).

1) Here L_u is the partial derivative of L with respect to u , and $x/y, z$ denotes the conditional expectation over the statistics of x given (y, z) .

2) Note that the existence of infinitely many solutions is mainly due to the fact that the leader was also allowed to acquire private information y (be it correlated or uncorrelated with the common information z). If this had not been the case, then, g , being only a function of z , would cancel out in (2.8) thus leading to a unique solution 0.

In view of this nonuniqueness feature of the solution to Problem A, Problem B becomes relevant which we address to in the sequel. Towards this end, let us assume that the leader adopts the policy (2.2) with Q chosen as in (2.8) and g being arbitrary (but satisfying (2.9)). For any such g , this is an optimal policy inducing (u^c, v^c) , provided that $x = x^*$. If $x \neq x^*$, however, the follower's reaction to y will no longer be $v = v^c$. In fact, substituting (2.2) into (2.4) with x^* replaced by a general x , and differentiating the resulting expression with respect to v , we obtain (to replace (2.5))

$$\begin{aligned} E \{ Q(z, y) E \{ L_{uu}(u^c - Q[v_1 - v^c], v_1, x, z) \} \\ y/z \quad x/y, z \quad (2.10) \\ = E \{ L_v(u^c - Q[v_1 - v^c], v_1, x, z) \} \\ x, y/z \end{aligned}$$

which admits a unique solution $v_1(z)$ when x lies in a neighborhood of x^* (because of strict convexity of L). This solution is not obtainable explicitly, unless we specify a structure for L (such as quadratic); however, we in fact do not need an explicit expression for $v_1(z)$, as the following discussion reveals.

The solution $v_1(z)$ to (2.10) will, in general, depend on different choices of Q out of the family (2.8)-(2.9). What Problem B alludes to, is a choice which will render the difference $|v_1(z) - v^c(z)|$ sufficiently small (in norm) whenever x is close to x^* . Note that, if $v_1(z)$ is close to $v^c(z) = v_{1^*}(z)$, then $u_1(z, y) = \gamma(v_1, z, y)$ will be close to $u^c(z, y) = u_{1^*}(z, y)$, because of continuity properties of γ . Hence, as a measure of the closeness of $v_1(z)$ to $v^c(z)$, we now take the first order term dv_1/dx and evaluate it at $x = x^*$.

Since (2.10) is an identity for all x of interest, we could differentiate it with respect to x (for each fixed z), to obtain the equality (at $x = x^*$):

$$\begin{aligned} E \{ Q(z, y) (dv^c/dx) E \{ L_{uv}(u^c, v^c, x, x^*) \} \\ y/z \quad x/y, z \\ - Q^c(z, y) (dv^c/dx) E \{ L_{uu}(u^c, v^c, x, x^*) \} \\ x/y, z \\ - Q(z, y) E \{ L_{ua}(u^c, v^c, x, x^*) \} \\ x/y, z \\ = E \{ (dv^c/dx) L_{vv}(u^c, v^c, x, x^*) \\ x, y/z \\ - Q(z, y) (dv^c/dx) L_{vu}(u^c, v^c, x, x^*) \\ + L_{v_1}(u^c, v^c, x, x^*) \} \end{aligned}$$

where

$$dv^c(z)/dx = dv_1(z)/dx \Big|_{x=x^*} \quad (2.11)$$

Since dv^c/dx is z -measurable, we can easily solve for it to obtain

$$\frac{dv^c}{dx} = \frac{E \{ L_{ua}Q - L_{va} \} \quad x, y/z}{E \{ L_{uu}Q^2 - 2L_{uv}Q + L_{vv} \} \quad x, y/z} \quad (2.12)$$

where $L_{ua}, L_{va}, L_{uu}, L_{uv}$ all have (u^c, v^c, x^*) as their arguments. Note that since L was strictly convex in (u, v) , the denominator of (2.12) is always positive and hence dv^c/dx , which we will henceforth call the first order sensitivity function, is well-defined.

Now, the first objective is to make this expression identically zero, by an appropriate choice of $g(z, y)$ in (2.3). Substituting (2.3) into the numerator of (2.12) and reshuffling some terms, we obtain the condition

$$E \{ g(z, y) f_1(z, y) \} = 0 \quad (2.13) \quad y/z$$

where

$$\begin{aligned} f_1(z, y) &= F(z, y) E \{ L_{va}(u^c, v^c, x, x^*) \} \\ &\quad x, y/z \\ &- G(z) E \{ L_{ua}(u^c, v^c, x, x^*) \} \quad x/y, z \end{aligned} \quad (2.14)$$

Now let

$$\Gamma^s = \{ g \in \Gamma^s : g \text{ satisfies (2.9)} \} \quad (2.15)$$

Then it follows from the above discussion that if there exists a $g \in \Gamma^s$, satisfying (2.13), the corresponding policy for the leader renders the first-order sensitivity function zero, i.e., to first-order v_1 (and consequently u_1) becomes insensitive to variations in the value of x , from the nominal value x^* . But, such a $g \in \Gamma^s$ exists generically—choose any random variable that is orthogonal to $f_1 \in \Gamma^s$ under the conditional probability measure $F(y/z)$, but not orthogonal to $F(z, y)$, which will be possible as long as F is not linearly dependent on f_1 under $F(y/z)$. A sufficient condition for this to be true is, for every $k \in \Gamma^s$,

$$E \{ L_{ua}(u^c, v^c, x, x^*) \} \neq k(z) E \{ L_{va}(u^c, v^c, x, x^*) \} \quad x/y, z \quad (2.16)$$

A candidate solution to (2.13) is then,

$$g(z, y) = g_0(z) - g_1(y) \quad (2.17a)$$

where

$$g_0(z) = \frac{E \{ g_1(y) f_1(z, y) \} \quad y/z}{E \{ f_1(z, y) \} \quad y/z} \quad (2.17b)$$

which can easily be shown to satisfy (2.13) for any y -measurable $g_1(y)$. Hence, we in fact have a family of solutions, parameterized by g_1 . Note that to assume $F(y/z) \neq 0$ here is not restrictive because if that is not the case (2.13) is trivially satisfied (by, for example, choosing g to be a constant).

Theorem 2. Let (2.3) and (2.16) be satisfied. Then, the first order sensitivity function dv^c/dx can be made identically zero by an incentive policy of the form (2.2) where Q is given by (2.3) and $g \in \Gamma^s$ satisfies (2.13), with one family of candidates being (2.17).

Proof. The proof follows by construction, from the discussion preceding the theorem.

An expression for the second-order sensitivity function

$$\frac{d^2 v^c(z)}{dx^2} = \frac{d^2 v_1(z)}{dx^2} \Big|_{x=x^*} \quad (2.18)$$

can be obtained by following the lines that led to (2.12). In this case (2.13) is replaced by (using the fact that dv^c/dx can be made equal to zero)

$$\frac{d^2 v^c}{dx^2} = \frac{E \{ L_{uaa}Q - L_{vaa} \} \quad x, y/z}{E \{ L_{uu}Q^2 - 2L_{uv}Q + L_{vv} \} \quad x, y/z}$$

which can be made zero if and only if $g \in \Gamma^s$ in (2.3) satisfies (as a counterpart of (2.13):

$$E \{ g(z, y) f_2(z, y) \} = 0 \quad (2.19) \quad y/z$$

where

$$\begin{aligned} f_2(z, y) &= F(z, y) E \{ L_{vaa}(u^c, v^c, x, x^*) \} \\ &\quad x, y/z \\ &- G(z) E \{ L_{uaa}(u^c, v^c, x, x^*) \} \quad x/y, z \end{aligned} \quad (2.20)$$

Hence, for both first and second-order sensitivity functions to be identically zero, it is sufficient to find a $g \in \Gamma^s$ which is orthogonal to both $f_1 \in \Gamma^s$ and $f_2 \in \Gamma^s$ under the measure $F(y/z)$. But this is generically possible because, for fixed $z \in R$, Γ^s can be made a pre-Hilbert space under the inner product

$$g(z,y) = 11z^2 - 22zy + 60 \quad (3.12)$$

which is unique up to a multiplicative term, which may be a function of z , within the class of functions affine in y . The corresponding $Q(z,y)$ becomes

$$Q(z,y) = (11z^2 - 22zy + 60)/18. \quad (3.13)$$

In order to illustrate the robustness properties of (3.13), the optimum reaction of the follower is plotted against B_1 in Fig. 1. The solid line represents the follower's optimum reaction when $g(z,y) = \text{constant}$, with a corresponding $Q(z,y) = \frac{1}{18}$. The dashed line represents the follower's optimum reaction to the robust strategy (3.13) when B_1 varies about $B_1^* = 1$. At $B_1 = B_1^*$, we have $v = v^* = 0.1$ for both strategies. However, when B_1 varies about 1, the optimum reaction induced by (3.13) is considerably closer to v^* compared with the case $Q = \frac{1}{18}$. [In this discussion, as well as in the figure, we have taken $z = 0.3$; though the behavior is similar for other values of z .]

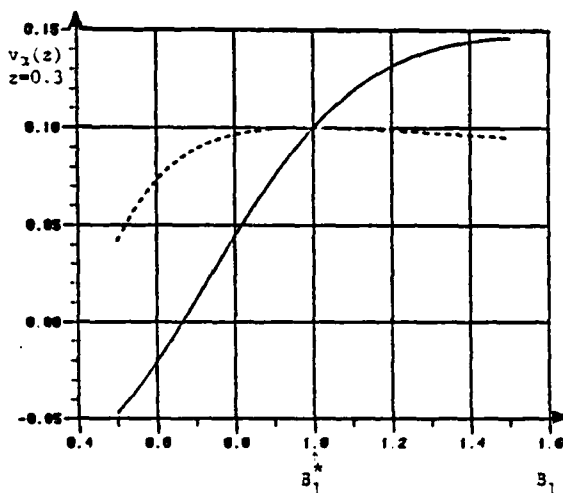


Fig. 1. Optimum response of the follower to leader's incentive policy for different values of B_1 .

Solid line: without the insensitivity term.
Dashed line: with the insensitivity term.

One would be tempted to think that the robustness properties of $Q(z,y)$ would be further enhanced if we chose $g(z,y)$ to be also orthogonal to $f_2(z,y)$ of (2.19), which in this case becomes

$$f_2(z,y) = (-z + y)z \quad (3.14)$$

However, since $-z + y$ is a linear combination of $f_1(z,y)$ and $F(z,y)$, any $g(z,y)$ orthogonal to f_1 and F is also orthogonal to F , therefore, such a $g(z,y)$ would not satisfy (2.9), thus ruling out the possibility of obtaining more robust strategies using the theory developed in Section II.

IV. EXTENSIONS

In this section we discuss some extensions of the foregoing results, which have been discussed in [3].

1) Higher (than 1) order sensitivity functions have been considered and it has been shown that for some classes of loss functions these can all be made identically zero by a proper choice of g^{ITS} in (2.3), which implies that in these cases the linear (in v) incentive policy (2.8) is not only locally insensitive, but also globally, i.e., it is a robust policy.

2) Theorems 2 and 3 are extended to the case when the parameter αA is a vector, and it is shown that the statements of these two theorems remain (basically) intact, as long as α is finite-dimensional. In this case (2.13) and (2.19) are each replaced by $p = \dim(\alpha)$ equations, one for each component of α .

3) These results are also extended to the case when U and V are finite dimensional Euclidean spaces.

REFERENCES

1. Başar, T. (1984). Affine incentive schemes for stochastic systems with dynamic information. SIAM J. Control and Optimization. Has also appeared in the Proc. 1982 American Control Conference, Arlington, VA, 127-132, June 1982.
2. Cansever, D. H. and T. Başar (1982). A minimum sensitivity approach to incentive design problems. Proc. 21st IEEE Conference on Decision and Control, Orlando, FL, 158-163. Also to appear in Large Scale Systems, 1984.
3. Başar, T. and D. H. Cansever. Robust incentive policies for stochastic decision problems in the presence of parametric uncertainty. Submitted for publication.
4. Wiedmann, J. (1980). Linear Operators in Hilbert Spaces. Springer-Verlag, New York.
5. Arrow, K. (1964). Control in large organizations. In Management Sci., Vol. 10, 3, 397-408.
6. Jennergren, L. P. (1980). On the design of incentives in business firms - a survey of some research. In Management Sci., Vol. 26, 2, 180-201.
7. Harris, M. and R. M. Townsend (1981). Resource allocation under asymmetric information. Econometrica, Vol. 49, 1, 33-64.

OPTIMUM OR NEAR-OPTIMUM INCENTIVE POLICIES FOR STOCHASTIC DECISION PROBLEMS
IN THE PRESENCE OF PARAMETRIC UNCERTAINTY[†]

by

Derya H. Cansever and Tamer Başar
Decision and Control Laboratory
Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Avenue
Urbana, Illinois 61801, U.S.A.

Address for Correspondence

Professor Tamer Başar
Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Avenue
Urbana, Illinois 61801, U.S.A.

[†]Research reported herein was supported in part by the Office of Naval Research under Contract N00014-84-K-0469; and in part by the U.S. Department of Energy, Electric Energy Systems Division, under Contract DE-AC01-81RA-50658, with Dynamic Systems, P. O. Box 423, Urbana, Illinois 61801.

ABSTRACT

In this paper we consider a general class of stochastic incentive decision problems in which the leader has access to the control value of the follower and to private as well as common information on the unknown state of nature. The follower's cost function depends on a finite number of parameters whose values are not known accurately by the leader, and in spite of this parametric uncertainty the leader seeks a policy which would induce the desired behavior on the follower. We obtain such policies for the leader, which are smooth, induce the desired behavior at the nominal values of these parameters, and furthermore make the follower's optimal reaction either minimally sensitive or totally insensitive to variations in the values of these parameters from the nominals. The general solution is determined by some orthogonality relations in some appropriately constructed (probability) measure spaces, and leads to particularly simple incentive policies. The features presented here are intrinsic to stochastic decision problems and have no counterparts in deterministic incentive problems.

Keywords: Stochastic systems; economic systems; team theory; decision theory; game theory; optimization; Stackelberg games.

I. INTRODUCTION

In this paper we consider a general class of stochastic Stackelberg game problems (equivalently, incentive decision problems, in our context) in which the leader has access to the control value of the follower and to private as well as common information on the unknown state of nature, whereas the follower has access to only common information which is shared by the leader. It is further assumed that the follower's cost function (which is strictly convex, but not necessarily quadratic) depends on a number of parameters whose values are not known accurately by the leader.

As it has been noted in the seminal paper by Harsanyi [1], the class of games with uncertain cost functions is in fact one of the three main prototypes in which a game with incomplete information can arise. Incomplete information is interpreted as lack of full information on the part of the players about the normal form of the game, and the other two cases which create games with incomplete information are:

- (i) Some or all of the players may not know the state of the nature, or the outcome of its evolution as a function of their decisions;
- (ii) The players may not know their own, or the other players' strategy spaces.

It is shown in [1] that the two cases cited above can be represented as uncertainties in the cost functions, so that games with any type of incomplete information can be treated as games wherein players are uncertain about their own or some other players' cost functions. In our Stackelberg game problem with two players, we will assume that both players have exact knowledge of their own cost functional, except for the state of the nature x , about which

players have imperfect information, with the follower's information on x being nested into the one of the leader's. We will also assume that the leader's lack of knowledge about the follower's cost function can be modeled as a finite dimensional parameter vector α . Under the assumption that each player knows his cost functional exactly, the asymmetry in the knowledge of α can affect the leader's cost function only through the follower's decision variable, where the latter is the outcome of an optimization problem, given the actual values of α . Under the adopted information structure and solution concept, follower's knowledge on the cost function of the leader is irrelevant to the analysis to follow.

We let the leader have a prior estimate of α , denoted by α^* , which will henceforth be referred to as α 's nominal value. Let us suppose for the time being that the leader knows the actual value of α ; then the outcome of the game with the extra information from the part of the leader is called the first best solution. The objective is to obtain "near-optimal" incentive policies (decision rules) for the leader which would induce the first best solution, characterized by a desired behavior (by the leader) on the follower when the nominal and actual values of α coincide, and such that this outcome will be minimally sensitive to deviations of the follower's perception of α from its nominal value.

The problem will be formulated in mathematical terms in Section II. In Section III, we will present a complete solution to the scalar problem, which possesses the main characteristics of the issue, while allowing better lucidity to the presentation. In Section IV, these results will be generalized to the case where the decision variables and uncertain parameters take values in finite-dimensional Euclidean spaces. The theory developed will be illustrated

via a numerical example motivated by a problem that arises in the control of large organizations [6], in Section V. Concluding remarks of Section VI end the paper.

II. PROBLEM FORMULATION

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be an underlying probability space on which three random variables x , y and z are defined; where $x \in X$ denotes the random state of the nature, and $z \in Z$ and $y \in Y$ denote, respectively, common (to both players) and private (to the leader) information related to x . We also let U and V be the decision spaces of the leader and the follower, respectively. We are assuming at this point that X , Z and Y are endowed with sufficiently rich topology so that probability measures can be defined on their subsets. Let Γ_f denote the class of all mappings from Z into V (satisfying some smoothness conditions to be delineated later), and Γ be the class of all mappings from $V \times Z \times Y$ into U . Furthermore let Γ^S be the class of mappings from $Z \times Y$ into U . We will denote generic elements of Γ^S , Γ_f and Γ by u , v and γ , respectively, so that $u = u(z, y)$, $v = v(z)$, $\gamma = \gamma(u, z, y)$. Let us assume that there is a point (u^t, v^t) in $\Gamma^S \times \Gamma_f$ which is most desirable by the leader (such as a system trajectory or control trajectory) and he seeks to determine a policy $\gamma \in \Gamma$ which would force the follower to such an action that the resultant outcome (u, v) in the product space $\Gamma^S \times \Gamma_f$ is sufficiently close to (u^t, v^t) , the first-best solution, by also taking into consideration the fact that the cost function of the follower is not known accurately.

Let $L(u, v, x, \alpha)$ denote the loss function of the follower, depending on a set of parameter values $\alpha \in A$. Let

$$J(u, v, \alpha) = E\{L(u(z, y), v(z), x, \alpha)\}$$

where E is the expectation over the statistics of the random variables x , z and y . (Note that we have abused the notation here and have used the same notation

for both random variables (or vectors) and their realizations.) Likewise we introduce

$$J(\gamma, v, \alpha) = E\{L(\gamma(v(z), z, y), v(z), x, \alpha)\}$$

which we call the cost function of the follower. We assume that, for each $\alpha \in A$ and $x \in X$, $L(u, v, x, \alpha)$ is strictly convex in the pair $(u, v) \in \Gamma^S \times \Gamma_f$, and furthermore, Γ is structured so that for each $\gamma \in \Gamma$ and $\alpha \in A$, $J(\gamma, v, \alpha)$ is strictly convex in v .

Under these assumptions, to each $\gamma \in \Gamma$, and for fixed $\alpha \in A$, there corresponds a unique element of Γ_f , called v_γ , which will minimize $J(\gamma, v, \alpha)$ over $v \in \Gamma_f$. [We are assuming here that, as a rational decision maker, the follower chooses his optimal policy (or rational reaction) by minimizing the function $J(\gamma, v, \alpha)$ where α is the true value characterizing his cost functional.] Hence, we have the unique correspondence (for each fixed $\alpha \in A$)

$$\gamma \xrightarrow{S'_\alpha} v$$

where the mapping S'_α depends explicitly on α and involves a minimization operation.

Let us further note that this unique relationship necessarily yields a unique element in Γ^S , given by

$$u(z, y) = \gamma(v(z), z, y) \quad ,$$

and hence we can associate, with each $\gamma \in \Gamma$, a unique pair (u, v) :

$$\gamma \xrightarrow{S_\alpha} (u, v)$$

where the mapping is denoted, in this case, by S_α .

This is the familiar input-output relationship of system theory (with the mapping S_{α} being much more complicated than the one normally encountered in system theory) and hence we can call Γ the input space $U \times V$ the output space. In terms of the familiar jargon of system theory we can rephrase the problem posed in the earlier part of this section as follows:

Problem A

For a given nominal value $\alpha^* \in A$, find an input (control) γ which would drive the system output to a desired value (u^t, v^t) . [Note that the desired output is in fact a stochastic variable (or process); hence what we pose here is akin to stochastic controllability.]

The solution to Problem A has in fact been obtained in [2]; an important feature is that it is generally nonunique, even in the class of policies (for the leader) which are linear in v . This then prompts a second question (additional design criterion) related to the sensitivity properties of the "optimal" γ :

Problem B

Among the control inputs which solve Problem A, which one (ones) leads (lead) to outputs that are least sensitive to variations in the value of α from the nominal value α^* ?

In this paper we obtain a complete solution to this problem by using some ideas originally introduced in [3] for the deterministic version of the problem. The features of the solution for the stochastic problem are, however, inherently different from those of the deterministic problem, which will be pointed out as we go along.

III. MAIN RESULTS FOR THE SCALAR PROBLEM

In this section we assume that the spaces U, V, X, Z are 1-dimensional Euclidean, and α is a single parameter with nominal value (as perceived by the leader) α^* . The actual value of α may be in a small neighborhood of α^* , and we assume that $L(u, v, x, \alpha)$ is strictly convex in (u, v) , and twice continuously differentiable in u, v and α , when α is restricted to lie in this neighborhood. Furthermore, we assume that every $\gamma \in \Gamma$ is continuously differentiable in v . The random variables x, y, z are jointly second-order random variables, and γ , in addition to being continuously differentiable in v , is measurable in z and y , and u, v are also measurable in their arguments. Furthermore, if L is measurable in x , the expectation of L , written as

$$E\{L(\gamma(v(z), z, y), v(z), x, \alpha)\} \quad , \quad (3.1)$$

is well-defined for all $\gamma \in \Gamma$, $u \in \Gamma^S$, $v \in \Gamma_F$, and for every α in a neighborhood of α^* . We finally assume that a desirable point in the product space $\Gamma^S \times \Gamma_F$, as chosen by the leader, is (u^t, v^t) .

If α indeed takes the value α^* , the results of [2], when specialized to this scalar case, indicate that this point (u^t, v^t) is induced by the leader by a linear (in v) policy

$$\gamma(v, z, y) = u^t(z, y) - Q(z, y) [v - v^t(z)] \quad (3.2)$$

whenever

$$E_{x/y, z} \{L_u(u^t(z, y), v^t(z), x, \alpha^*)\} \neq 0 \quad (3.3)$$

¹⁾ Here L_u is the partial derivative of L with respect to u , and $E_{x/y, z}$ denotes the conditional expectation over the statistics of x given (y, z) .

with positive probability in (y,z) . Since we have a simpler (scalar) problem here, the equation for Q can be obtained directly. Substituting (3.2) into (3.1) we have

$$E_z \{ E_{x,y/z} \{ L(u^t - Q[v - v^t], v, x, \alpha^*) \} \} \quad (3.4)$$

which is strictly convex in v since L was strictly convex in (u,v) and γ is linear in v . Hence (3.4) admits a unique minimum in Γ_f , which is obtained by differentiating the inner expression with respect to v , for fixed z , and setting equal to zero:

$$E_{x,y/z} \{ L_u Q - L_v \} = 0$$

—

$$\begin{aligned} E_{y/z} \{ Q(z,y) E_{x/y,z} \{ L_u(u^t - Q[v - v^t], v, x, \alpha^*) \} \} \\ = E_{x,y/z} \{ L_v(u^t - Q[v - v^t], v, x, \alpha^*) \} \end{aligned} \quad (3.5)$$

This is the equation that the minimizing v would satisfy (this is also sufficient because of strict convexity), and it is in general nonlinear in v . However, we do not need to solve this for v , but simply find a Q such that its unique solution is v^t — which would in turn yield $\gamma(v^t, z, y) = u^t$, in view of (3.2), and thereby the desired solution would be induced. Now, substituting $v = v^t$ in (3.5) we obtain

$$E_{y/z} \{ Q(z,y) F(z,y) \} = G(z) \quad (3.6)$$

where

$$F(z,y) \triangleq E_{x/y,z} \{L_u(u^t(z,y), v^t(z), x, u^*)\}$$

$$G(z) \triangleq E_{x,y/z} \{L_v(u^t(z,y), v^t(z), x, u^*)\}.$$

If this had been a deterministic problem, the solution to (3.6) would be unique, thus outruling any possibility of obtaining optimal policies that satisfy other design criteria. For the stochastic problem, however, (3.6) admits infinitely many solutions, with a family of such solutions being (which turns out to be sufficiently rich):

$$Q(z,y) = g(z,y)G(z)/[E_{y/z} \{g(z,y)F(z,y)\}] \quad (3.8)$$

where g is any function measurable in (z,y) satisfying the condition

$$E_{y/z} \{g(z,y)F(z,y)\} \neq 0, \quad z \in \mathbb{R}. \quad (3.9)$$

Verification of this result is by direct substitution of (3.8) into (3.6).

Now, let $\mathcal{F}_z$ denote the sigma-algebra generated by z , and $\mathcal{F}_{zVY}$ be the sigma algebra generated by z and y . The nonuniqueness of $Q(z,y)$, as described by (3.8), stems from the fact that it is not measurable with respect to $\mathcal{F}_z$, the sigma algebra generated by the information acquired by the follower. More precisely, since $\mathcal{F}_{zVY} \supset \mathcal{F}_z$ every atom²⁾ of $\mathcal{F}_z$ is a union of atoms of $\mathcal{F}_{zVY}$ and the collection of atoms of $\mathcal{F}_z$ gives a coarser partition of Ω than the corresponding collection from $\mathcal{F}_{zVY}$ [4]. Considering (3.6), the defining

²⁾ A is called an atom of a sigma algebra $\mathcal{F}$ if $A \in \mathcal{F}$, and no subset of A belongs to $\mathcal{F}$ other than A itself and the empty set.

equation for $Q(z,y)$, the $\mathcal{F}_{zV_y}$ -measurable function $Q(z,y)F(z,y)$, assumes a single value on an atom of $\mathcal{F}_{zV_y}$. As each atom of $\mathcal{F}_z$ is composed of a union of atoms of $\mathcal{F}_{zV_y}$, $Q(z,y)F(z,y)$ will in general take more than one value on every atom of $\mathcal{F}_z$. When the conditional expectation is taken, $Q(z,y)F(z,y)$ is averaged over the atoms of $\mathcal{F}_z$ to yield $G(t)$ for given $F(z,y)$. Clearly, infinitely many functions may yield the same average, such as the family characterized by (3.8). We note that the nonuniqueness is further pronounced by the presence of the $\mathcal{F}_{zV_y}$ -measurable function $F(z,y)$. Should $F(z,y)$ be $\mathcal{F}_z$ -measurable, then the nonuniqueness of $Q(z,y)$ as in (3.8) drops out when the expected value of $Q(z,y)$ conditioned on $\mathcal{F}_z$ is computed. However, when $F(z,y)$ is not $\mathcal{F}_z$ -measurable, then this nonuniqueness is genuine, in the sense that it remains to hold true even after the expected value of $Q(z,y)$ conditioned on $\mathcal{F}_z$ is computed, an operation performed by the follower when computing his optimum reaction to an announced strategy.

We now summarize the result on the nonuniqueness of $Q(z,y)$ below:

Theorem 1. Problem A formulated in Section I admits for the scalar version, infinitely many linear (in v) solutions, with one such family given by (3.2), (3.8), under condition (3.9). □

In view of this nonuniqueness feature of the solution to Problem A, Problem B becomes relevant which we address in the sequel. Towards this end, let us assume that the leader adopts the policy (3.2) with Q chosen as in (3.8) and g being arbitrary (but satisfying (3.9)). For any such g , this is an optimal policy inducing (u^t, v^t) , provided that $x = x^*$. If $x \neq x^*$, however, the follower's reaction to γ will no longer be $v = v^t$. In fact, substituting (3.2) into (3.4) with x^* replaced by a general x , and differentiating the resulting expression with respect to v , we obtain (to replace (3.5))

$$\begin{aligned}
& E_{y/z} \{Q(z,y) E_{x/y,z} \{L_u(u^t - Q[v_\alpha - v^t], v_\alpha, x, \alpha)\}\} \\
& = E_{x,y/z} \{L_v(u^t - Q[v_\alpha - v^t], v_\alpha, x, \alpha)\}
\end{aligned} \tag{3.10}$$

which admits a unique solution $v_\alpha(z)$ when α lies in a neighborhood of α^* (because of strict convexity of L). This solution is not obtainable explicitly, unless we specify a structure for L (such as quadratic); however, we in fact do not need an explicit expression for $v_\alpha(z)$, as the following discussion reveals.

The solution $v_\alpha(z)$ to (3.10) will, in general, depend on different choices of Q out of the family (3.8)-(3.9). What Problem B alludes to, is a choice which will render the difference $|v_\alpha(z) - v^t(z)|$ sufficiently small (in norm) whenever α is close to α^* . Note that, if $v_\alpha(z)$ is close to $v^t(z) = v_{\alpha^*}(z)$, then $u_\alpha(z,y) = \gamma(v_\alpha, z, y)$ will be close to $u^t(z,y) = u_{\alpha^*}(z,y)$, because of continuity properties of γ . Hence, as a measure of the closeness of $v_\alpha(z)$ to $v^t(z)$, we now take the first order term $dv_\alpha/d\alpha$ and evaluate it at $\alpha = \alpha^*$.

Since (3.10) is an identity for all α of interest, we could differentiate it with respect to α (for each fixed z), to obtain the equality (at $\alpha = \alpha^*$):

$$\begin{aligned}
& E_{y/z} \{Q(z,y) (dv^t/d\alpha) E_{x/y,z} \{L_{uv}(u^t, v^t, x, \alpha^*)\} - Q^2(z,y) (dv^t/d\alpha)\} \\
& \quad E_{x/y,z} \{L_{uu}(u^t, v^t, x, \alpha^*)\} + Q(z,y) E_{x/y,z} \{L_{ua}(u^t, v^t, x, \alpha^*)\} \\
& = E_{x,y/z} \{(dv^t/d\alpha) L_{vv}(u^t, v^t, x, \alpha^*) - Q(z,y) (dv^t/d\alpha) L_{vu}(u^t, v^t, x, \alpha^*) \\
& \quad + L_{vx}(u^t, v^t, x, \alpha^*)\}
\end{aligned}$$

where

$$dv^t(z)/d\alpha = \left. dv_{\alpha}(z)/d\alpha \right|_{\alpha=\alpha^*} \tag{3.11}$$

Since $dv^t/d\alpha$ is F_z -measurable, we can easily solve for it to obtain

$$\frac{dv^t}{d\alpha} = \frac{E_{x,y/z} \{L_{u\alpha}Q - L_{v\alpha}\}}{E_{x,y/z} \{L_{uu}Q^2 - 2L_{uv}Q + L_{vv}\}} \quad (3.12)$$

where $L_{u\alpha}$, $L_{v\alpha}$, L_{uu} , L_{uv} all have (u^t, v^t, α^*) as their arguments. Note that since L was strictly convex in (u, v) , the denominator of (3.12) is always positive and hence $dv^t/d\alpha$, which we will henceforth call the first order sensitivity function, is well-defined.

Now, the first objective is to make this expression identically zero, by an appropriate choice of $g(z, y)$ in (3.8). Substituting (3.8) into the numerator of (3.12) and reshuffling some terms, we obtain the condition

$$E_{y/z} \{g(z, y)f_1(z, y)\} = 0 \quad (3.13)$$

where

$$\begin{aligned} f_1(z, y) \triangleq & F(z, y) E_{x,y/z} \{L_{v\alpha}(u^t, v^t, x, \alpha^*)\} \\ & - G(z) E_{x/y,z} \{L_{u\alpha}(u^t, v^t, x, \alpha^*)\} \end{aligned} \quad (3.14)$$

Now let

$$\tau^{s-} = \{g \in \tau^s : g \text{ satisfies (3.9)}\} \quad (3.15)$$

Then it follows from the above discussion that if there exists a $g \in \tau^{s-}$, satisfying (3.13), the corresponding policy for the leader renders the first-order sensitivity function zero, i.e., to first-order v_1 (and consequently u_1) becomes insensitive to variations in the value of α , from the nominal value α^* .

But, such a $g \in \Gamma^S$ exists generically—choose any random variable that is orthogonal to $f_1 \in \Gamma^S$ under the conditional probability measure $P(y/z)$, but not orthogonal to $F(z,y)$, which will be possible as long as F is not linearly dependent on $f_1(z,y)$. This holds true if and only if we have, for every $k(\cdot) \in \Gamma_f$,

$$E_{x/y,z} \{L_{u\alpha}(u^t, v^t, x, \alpha^*)\} \neq k(z) E_{x/y,z} \{L_u(u^t, v^t, x, \alpha^*)\} \quad (3.16)$$

A candidate solution to (3.13) is then,

$$g(z,y) = E_{y/z} \{g_1(z,y)f_1(z,y)\} - g_1(z,y) E_{y/z} \{f_1(z,y)\} \quad (3.17)$$

which can easily be shown to satisfy (3.13) for any F_{zV_y} -measurable $g_1(z,y)$.

Hence, we in fact have a family of solutions parameterized by $g_1(z,y)$:

Theorem 2. Let (3.3) and (3.16) be satisfied. Then, the first order sensitivity function $dv^t/d\alpha$ can be made identically zero by an incentive policy of the form (3.2) where Q is given by (3.8) and $g \in \Gamma^{S^-}$ satisfies (3.13), with one family of candidates being (3.17).

Proof. The proof follows by construction, from the discussion preceding the theorem. □

It is shown in Appendix A that an expression for the second-order sensitivity function

$$\frac{d^2 v^t(z)}{d\alpha^2} = \frac{d^2 v_\alpha(z)}{d\alpha^2} \Big|_{\alpha=\alpha^*} \quad (3.18)$$

is given by

$$\frac{d^2 v^t}{d\alpha^2} = E_{x,y/z} \{L_{u\alpha\alpha} Q - L_{v\alpha\alpha}\} / E_{x,y/z} \{L_{uu} Q^2 - 2L_{uv} Q + L_{vv}\} \quad (3.19)$$

which can be made zero if and only if $g(z,y) \in \Gamma^{S-}$ in (3.8) satisfies [as a counterpart of (3.13)]:

$$E_{y/z} \{g(z,y) f_2(z,y)\} = 0 \quad (3.20)$$

where

$$\begin{aligned} f_2(z,y) &\triangleq F(z,y) E_{x,y/z} \{L_{v\alpha\alpha}(u^t, v^t, x, \alpha^*)\} \\ &\quad - G(z) E_{x/y,z} \{L_{u\alpha\alpha}(u^t, v^t, x, \alpha^*)\} \end{aligned} \quad (3.21)$$

Hence, for both first and second-order sensitivity functions to be identically zero, it is sufficient to find a $g \in \Gamma^{S-}$ which is orthogonal to both $f_1 \in \Gamma^S$ and $f_2 \in \Gamma^S$ under the measure $P(y/z)$. But this is generically possible because, for fixed $z \in \mathbb{R}$, Γ^{S-} can be made a pre-Hilbert space under the inner product

$$\langle g, f \rangle \triangleq \int_{\mathbb{R}} g(z,y) f(z,y) dP(y/z) \quad ,$$

with $g, f \in \Gamma^{S-}$, where $P(y/z)$ is a probability measure. In order to insure that there exists a $g \in \Gamma^{S-}$ orthogonal to both f_1 and f_2 , we have to assume, in addition to (3.16), the validity of the condition

$$E_{x/y,z} \{L_{u\alpha\alpha}(u^*, v^*, x, \alpha^*)\} \neq k(z) E_{x/y,z} \{L_u(u^t, v^t, x, \alpha^*)\}, \quad \forall k(\cdot) \in \Gamma_c \quad (3.22)$$

This then leads to the following theorem:

Theorem 3. Let conditions (3.3), (3.16) and (3.22) be satisfied. Then, for the scalar stochastic incentive problem of this section, there exists an incentive policy for the leader which induces the follower to play $v = v^t$ when $\alpha = \alpha^*$ (the nominal value) and furthermore makes the sensitivity functions of orders 1 and 2 (i.e., $dv^t/d\alpha$ and $d^2v^t/d\alpha^2$) identically zero a.e. $P(y/z)$. Such a policy is given by (3.8) where $g \in \Gamma^{S^-}$ satisfies (3.13) and (3.20).

Proof. (3.16) and (3.22) guarantee that f_1 and f_2 do not linearly depend on $F(z, y)$. Without loss of generality, let us assume that f_1 and f_2 are linearly independent. Then, there exists an orthonormal system (e_1, e_2) in Γ^{S^-} spanning the same subspace as (f_1, f_2) [5]. Now an $e_3 \in \Gamma^{S^-}$ orthogonal to both e_1 and e_2 can be constructed using Gram-Schmidt orthogonalization procedure, a.e., $P(y/z)$. This e_3 is the desired g . $\square$

The lines that led to the proof of Theorem 3 suggest that higher order sensitivity functions can be annihilated using the same approach. Towards this end, let us assume that $L(u, v, \alpha)$ is N times differentiable in α , where N is an arbitrary large positive integer. Let N denote the index set $\{1, 2, \dots, N\}$. Then, the n 'th order sensitivity function is defined as

$$\frac{d^n v^t(z)}{d\alpha^n} = \frac{d^n v_\alpha(z)}{d\alpha^n} \Big|_{\alpha=\alpha^*} \quad n \in N \quad (3.23)$$

Our objective is to annihilate the above expression for all $n \in N$.

This problem alludes to rendering the N 'th order Taylor approximation of $v_\alpha(z)$ sufficiently close to $v^t(z)$. Indeed, let the true value of α be

$$\alpha = \alpha^* + \varepsilon \quad (3.24)$$

where ϵ is a sufficiently small real number. The N 'th order Taylor approximation of $v_\alpha(z)$ about α^* is

$$v_\alpha(z) = v^t(z) + \sum_{n=1}^N \frac{\epsilon^n}{n!} \frac{d^n v_\alpha(z)}{d\alpha^n} + o(\epsilon^N) \quad (3.25)$$

where $v^t(z)$ is the first-best solution of the incentive control problem. To make $v_\alpha(z)$ as close as possible to the first-best solution, we will choose $Q(z,y)$ such that (3.23) vanishes for all $n \in \mathbb{N}$. It is shown in Appendix A that

$$\frac{d^n v^t}{d\alpha^n} = \frac{E_{x,y/z} \{L_{u\alpha n} Q - L_{v\alpha n}\}}{E_{x,y/z} \{L_{uu} Q^2 - 2L_{uv} Q + L_{vv}\}}, \quad (3.26)^3$$

this expression can be made zero, if and only if $g \in \Gamma^S$ in (3.8) satisfies

$$E_{y/z} \{g(z,y) f_n(z,y)\} = 0 \quad (3.27)$$

where

$$f_n(z,y) \triangleq F(z,y) E_{x,y/z} \{L_{v\alpha n}(u^t, v^t, x, \alpha^*)\} - G(z) E_{x/y,z} \{L_{u\alpha n}(u^t, v^t, x, \alpha^*)\}, \quad (3.28)$$

and

$$E_{x/y,z} \{L_{u\alpha n}(u^*, v^*, x, \alpha^*)\} \neq k(z) E_{x/y,z} \{L_u(u^t, v^t, x, \alpha^*)\}, \quad \forall k(\cdot) \in \Gamma_f \quad (3.29)$$

where (3.28) is a necessary and sufficient condition for the linear independence of f_n from $F(z,y)$, and f_n is the counterpart of (3.14) and (3.21) specialized for the n 'th order sensitivity function. The fact that Γ_S^- is a pre-Hilbert space enables us to prove the following theorem.

³⁾ Here $L_{\alpha n}$ is the n 'th order partial derivative of L with respect to α .

Theorem 4. Let conditions (3.3) and (3.29) be satisfied for all $n \in \mathbb{N}$. Further let

$$\alpha = \alpha^* + \epsilon$$

where ϵ is sufficiently small. Then, there exists an incentive policy for the leader which induces

$$v_\alpha(z) = v^t(z) + o(\epsilon^N) \quad \text{a.e. } P(y/z)$$

where $v^t(z)$ is the first-best solution of the problem, and N is an arbitrarily large finite positive integer. Such a policy is given by (3.8) where $g \in \Gamma^{S^-}$ satisfies (3.27) for all $n \in \mathbb{N}$.

Proof. Under (3.29), f_n is linearly independent from $F(z, y)$, so that (3.9) is not violated. Let S^N be the subspace of Γ^{S^-} spanned by $\{f_n, n \in \mathbb{N}\}$. Using the Gram-Schmidt orthogonalization procedure, one can construct an $e_{N+1} \in \Gamma^{S^-}$ orthogonal to S^N a.e. $P(y/z)$. This e_{N+1} is the desired g . $\square$

IV. GENERALIZATIONS TO THE VECTOR CASE

In the previous section, we confined our analysis to the case where A , U and V are one-dimensional, mainly not to obscure the main ideas with cumbersome notation. In this section, we will let these spaces be finite-dimensional Euclidean, and show that the results of Section III can be generalized to the vector case as well. The first step towards this goal is to let A be a subset of $\mathbb{R}^r$, and obtain a counterpart of Theorem 4 for this vector-parameter case. Now, let the actual value of the parameter α be related to α^* through

$$\alpha = \alpha^* + \varepsilon, \quad \varepsilon \in \mathbb{R}^r, \quad \alpha \in A. \quad (4.1)$$

Then the N 'th order Taylor approximation of $v_\alpha(z)$ around α^* for ε sufficiently small is

$$v_\alpha(z) = v^t(z) + \sum_{n=1}^N \frac{D_\varepsilon^n \{v_\alpha(z)\}}{n!} + o(\|\varepsilon\|^N) \quad (4.2)$$

where

$$D_\varepsilon^n v_\alpha(z) = (\varepsilon_1 D_1 + \dots + \varepsilon_r D_r)^n \{v_\alpha(z)\} \quad (4.3)$$

and D_i is the partial differential operator with respect to the i 'th component of the r -dimensional vector α , acting on $v_\alpha(z)$. Similarly, ε_i is the i 'th component of ε . We, therefore, have to find a Q orthogonal to $D_\varepsilon^n v_\alpha(t)$ for all permissible $\varepsilon \in \mathbb{R}^r$ and $n \in \mathbb{N}$. This is accomplished if we take Q to be orthogonal to the set of vectors defined by

$$D_1^{j_1} \dots D_r^{j_r} \{v_\alpha(z)\} \quad (4.4a)$$

$$\forall j_i \in \mathbb{Z}^+ \text{ such that } j_1 + \dots + j_r = n \quad n \in \mathbb{N}, \quad (4.4b)$$

where $\mathbb{Z}^+$ is the set of positive integers.

Let us assume that $L(u,v,x,\alpha)$ is sufficiently smooth so that the expressions given in the sequel are well-defined. An expression for the i 'th component of the vector valued n 'th order sensitivity function is then given by

$$\{D_1^{j_1} \dots D_r^{j_r} v_i(z)\}_i = E_{x,y/z} \{D_1^{j_1} \dots D_r^{j_r} \{L_u\} Q - D_1^{j_1} \dots D_r^{j_r} \{L_v\}\} / \quad (4.5)$$

$$E_{x,y/z} \{L_{uu} Q^2 - 2L_{uv} Q + L_{vv}\}, \forall j_i \in \mathbb{Z}^+ \text{ such that } j_1 + \dots + j_r = n.$$

For a finite N , there will be a finite number of these vectors in (4.2).

It is therefore possible to find a Q orthogonal to every term in $\sum_{n=1}^N \frac{D^n \{v_i(z)\}}{n!}$ under some linear independence conditions to be delineated in the

sequel. As a counterpart of (3.27), we should have

$$E_{y/z} \{g(z,y) f_n(z,y)\} = 0, \quad n \in \mathbb{N} \quad (4.6)$$

where the components of the vector valued function f_n are given by

$$\{f_n(z,y)\}_i = F(z,y) E_{x,y/z} \{D_1^{j_1} \dots D_r^{j_r} \{L_v(u^t, v^t, x, x^*)\}\} \quad (4.7)$$

$$- G(z) E_{x/y,z} \{D_1^{j_1} \dots D_r^{j_r} \{L_u(u^t, v^t, x, x^*)\}\}, \forall j_i \in \mathbb{Z}^+ \text{ such that } j_1 + \dots + j_r = n.$$

It is possible to choose such a g provided that $\{f_n(z,y)\}_1$ is linearly independent from $F(z,y)$. This condition is characterized by

$$\begin{aligned} E_{x/y,z} \{L_u(u^t, v^t, x, \alpha^*)\} \neq k(z) \quad E_{x/y,z} \{D_1^{j_1} \dots D_r^{j_r} \{L_u(u^t, v^t, x, \alpha^*)\}\} \\ \forall k(\cdot) \in \Gamma_f; \forall j_i \in \mathbb{Z}^+ \text{ such that } j_1 + \dots + j_r = n; \forall n \in \mathbb{N} \end{aligned} \quad (4.8)$$

We can now state the counterpart of Theorem 4 for this case where α is a vector. Its proof is along lines similar to the one of the previous theorem, and is therefore omitted.

Theorem 5. Let conditions (3.3) and (4.8) be satisfied. Let also

$$\alpha = \alpha^* + \varepsilon$$

where $\varepsilon \in \mathbb{R}^r$ is sufficiently small. Then, there exists an incentive policy for the leader which induces

$$v_\lambda(z) = v^t(z) + o(\|\varepsilon\|^N) \quad \text{a.e. } \mathcal{P}(y/z)$$

where $v^t(z)$ is the first-best solution of the problem, and N is an arbitrarily large finite positive integer. Such a policy is given by (3.8) where $g \in \Gamma^{S-}$ satisfies (4.6). □

We now let U and V be identical to $\mathbb{R}^m$ and $\mathbb{R}^l$, respectively. In this case, the defining equation for the $(m \times l)$ matrix $Q(z,y)$:

$$E_{y/z} \{F(z,y)Q(z,y)\} = G(z) \quad (4.9)$$

can be rewritten as a set of l equations given by

$$E_{y/z} \{F(z,y)Q_i(z,y)\} = G_i, \quad i \in I, \quad (4.10)$$

$$I \stackrel{\Delta}{=} \{1, \dots, l\}, \quad (4.11)$$

where $Q_i(z,y)$ is the i 'th column of $Q(z,y)$, and $G_i(z)$ is the i 'th element of the $(l \times 1)$ vector $G(z)$. This equation alludes first to a Euclidean inner product between two vectors F and Q_i with entries in Γ^S , and then to an average of the resulting quantity over the atoms of $\mathcal{F}_z$ to yield $G_i(z)$. Now, if we arbitrarily assign $m-1$ elements of $Q_i(z,y)$, and perform the operations required by (4.10) on these arbitrarily assigned elements and the corresponding entries of $F(z,y)$, and transfer the resulting $\mathcal{F}_z$ -measurable function to the right side of (4.10), we are left with an equation analogous to (3.6), determining the remaining entry of $Q_i(z,y)$. If the corresponding entry of $F(z,y)$ is nonzero, the remaining entry of $Q_i(z,y)$ admits an infinity of solutions characterized by an equation identical to (3.8). We now summarize this result below.

Theorem 6. Let $F(z,y)$ be different from the zero vector with positive probability in (y,z) . Then, any $(m-1)$ elements of each column of $Q(z,y)$ can be arbitrarily assigned, provided that the corresponding entry of F for the remaining element of that column is nonzero. There exists an infinity of solutions for the remaining entry, characterized by (3.8), with F and G of (3.8) being properly identified. c

This result enables us to characterize the family of solutions to $Q(z,y)$ when U and V are finite-dimensional vector spaces. We now require that l -vector $v_z(t)$ be as close as possible to the first best solution $v^t(z)$ when the actual parameter α is described by

$$\alpha = \alpha^* + \epsilon$$

for $\varepsilon \in \mathbb{R}^r$ being sufficiently small. Then, for each component of $v_\alpha(z)$, denoted by $v_\alpha^p(z)$, where p belongs to the index set $\mathcal{L}$, there is a Taylor expansion described by (4.2). In order to annihilate all the sensitivity functionals up to order N in the expansion of $v_\alpha(z)$, we need to choose Q orthogonal to the finite set of functions defined by

$$D_1^{j_1} \dots D_r^{j_r} \{v_\alpha^p(z)\}$$

$$\forall j_i \in \mathbb{Z}^+ \text{ such that } j_1 + \dots + j_r = n, n \in \mathbb{N}, p \in \mathcal{L}.$$

Let $L(u, v, x, \alpha)$ be sufficiently smooth. Then, an expression for the components of the numerator of the n 'th order sensitivity function is given by

$$\text{num}\{D_1^{j_1} \dots D_r^{j_r} v_\alpha^p(z)\}_i = E_{x, y/z} \{D_1^{j_1} \dots D_r^{j_r} \{F(z, y)\} Q_p - D_1^{j_1} \dots D_r^{j_r} \{G_p(z)\}\} \quad (4.12)$$

$$\forall j_i \in \mathbb{Z}^+ \text{ such that } j_1 + \dots + j_r = n; p \in \mathcal{L}.$$

We know from Theorem 6 that $m-1$ elements of the m -vector Q_p can be arbitrarily assigned. Without loss of generality, let us make them equal to zero, remaining with a nonzero and F_{zV_y} -measurable element Q_{ps} , for some positive integer s , less than or equal to m . This Q_{ps} is characterized by

$$Q_{ps}(z, y) = g_{ps}(z, y) G_p(z) / [E_{y/z} \{g_{ps}(z, y) F_s(z, y)\}] \quad (4.13)$$

where F_s is the s 'th element of the $(1 \times m)$ -vector $F(z, y)$, and $g_{ps}(z, y)$ is a F_{zV_y} -measurable scalar valued function satisfying

$$E_{y/z} \{g_{ps}(z, y) F_s(z, y)\} \neq 0, z \in \mathbb{R} \quad (4.14)$$

Now, as a counterpart of (4.6)-(4.7), we will choose this g_{ps} to be orthogonal to the elements of the vector-valued function $f_n(z,y)$ whose components are given by

$$\{f_n^p(z,y)\}_i = F_s(z,y)[D_1^{j_1} \dots D_r^{j_r}\{G_p(z)\}] - G_p(z)[D_1^{j_1} \dots D_r^{j_r}\{F_s(z,y)\}] \quad (4.15)$$

$\forall j_i \in \mathbb{Z}^+$ such that $j_1 + \dots + j_r = n$, $\forall n \in \mathbb{N}$; for each $p \in \mathcal{L}$, and for some s .

From the previous results, we know that it is possible to find a g_{ps} orthogonal to the family of functions described by (4.15), for each $p \in \mathcal{L}$, provided that some linear independence conditions hold. These conditions in this general case become

$$F_s(z,y) \neq k(z)[D_1^{j_1} \dots D_r^{j_r}\{F_s(z,y)\}] \quad (4.16)$$

$\forall k(\cdot) \in \Gamma_f$, $\forall j_i \in \mathbb{Z}$ such that $j_1 + \dots + j_r = n$, $\forall n \in \mathbb{N}$; for at least one $s \in \{1, \dots, m\}$.

When the decision space of the leader is finite dimensional Euclidean, the requirement that

$$F(z,y) \stackrel{\Delta}{=} E_{x/y,z} \{L_u(u^t, v^t, x, \alpha^*)\}$$

should be linear independent of its partial derivatives with respect to α has been relaxed to hold for at least one component of F . We now summarize the foregoing discussion in the theorem to follow.

Theorem 7. Let $F(z,y)$ be different from the zero vector with positive probability in (z,y) , and (4.16) hold for at least one $s \in \{1, \dots, m\}$. Let also

$$\alpha = \alpha^* + \epsilon$$

where $\varepsilon \in \mathbb{R}^r$ is sufficiently small. Then, there exists an incentive policy for the leader which induces

$$v_\alpha(z) = v^t(z) + o(\|\varepsilon\|^N) \text{ a.e. } P(y/z), \quad v_\alpha(z) \in \mathbb{R}^l$$

where $v^t(z)$ is the first-best solution of the problem, and N is an arbitrarily large finite positive integer. Such a policy is given by (4.13) where g_{ps} is orthogonal to the functions defined by (4.15). □

where $\epsilon \in \mathbb{R}^r$ is sufficiently small. Then, there exists an incentive policy for the leader which induces

$$v_\alpha(z) = v^t(z) + o(\|\epsilon\|^N) \text{ a.e. } \mathcal{P}(y/z), \quad v_\alpha(z) \in \mathbb{R}^l$$

where $v^t(z)$ is the first-best solution of the problem, and N is an arbitrarily large finite positive integer. Such a policy is given by (4.13) where g_{ps} is orthogonal to the functions defined by (4.15). $\square$

V. AN ILLUSTRATIVE EXAMPLE

To illustrate the theory of the previous sections, we now consider an incentive design problem in a divisionalized firm [6], [7]. Many large firms are organized into a multidivisional structure with interdependencies among them. For instance, when divisions compete in the same market, or when one division supplies another one with goods or services, conflict of interest among divisions arise, and the corporate center needs to coordinate their decisions in order to maximize the firm's overall profit. In our example, we focus into the coordination of a division director (the follower) by the corporate center (the leader). Let x_0 be a random variable representing the actual state of the firm, consisting of an aggregation of its profit level, its market share, the market value of the shareowner's equity, and the like. The leader and the follower have access to a noisy measurement of x_0 , denoted by z , through the internal reports provided by the staff of the corporate center. The leader has also access to a private measurement of x_0 , denoted by y , provided by an independent market research firm. The decision variable of the follower is its effort level v , which is also observed by the leader, and the decision variable of the leader is the amount of centrally allocated scarce resources and is denoted by u . We assume that the state evolves according to

$$x_1 = x_0 + B_1 u + B_2 v \quad (5.1)$$

where B_1 and B_2 represent the technology of transforming resources and labor into production. Under this setup, the objective of the follower is to minimize

$$E_{y/z} \{L\} = E_{y/z} \{S(T - (x_0 + B_1 u + B_2 v))^2 + Rv^2\} \quad (5.2)$$

over v , where S and R are scalar weighting the regulation of x_1 and the disutility of effort, respectively. The parameter T represents a target for the state of division, summarizing the division's and his director's best interests for the future; but this target is not necessarily consistent with the overall objective of the divisionalized firm. Furthermore, this target is set by the division director, and is not accurately known by the headquarters of the divisionalized firm; but we assume here that the headquarters have an a prior estimate of this target, denoted by T^* . An alternative interpretation is that this T^* may represent a desired target by the headquarters for that particular division, while the division's director may perceive a different target T , whose actual value is known only by himself, and he performs his optimization according to that value.

Let the pair (u^t, v^t) denote the optimizing arguments of the objective function of the leader, which is different from (5.2) — this difference being mainly due to a discrepancy in the perception of the value of T (the leader perceives it as T^*).⁴⁾ The aim of the leader is to induce the pair (u^t, v^t) in this decision problem, in spite of the uncertainty he is faced with in the value of T . He will realize this goal by letting his decision variable u to be a function of the follower's effort level.

This problem is a version of the principal-agent problem [8], where in this case the principal can observe the agent's effort level but is uncertain about his true objective. For a numerical illustration of this problem, we will assume some specific values for the parameters of the cost functional. More precisely, let L be (with $B_1 = B_2 = 1$, $S = K = 1/2$)

$$L = \frac{1}{2} (T - (x_0 + u + v))^2 + \frac{1}{2} v^2 \quad (5.3)$$

⁴⁾ We should note that this discrepancy is not the only factor that contributes to the difference between the two objective functions, but is the most pronounced one.

where T is the uncertain parameter to the leader, with a nominal value $T^* = 2$. We further take x_0 to be nominally distributed with mean one and unit variance, and the common observation to be given by

$$z = x_0 + w_1, \quad w_1 \sim N(0,1) \quad (5.4)$$

In addition to z , the leader has access to

$$y = x_0 + w_2, \quad w_2 \sim N(0,1) \quad (5.5)$$

where x_0 , w_1 and w_2 are mutually independent. Let the pair (u^t, v^t) given by

$$u^t = \frac{z}{2} + \frac{y}{3} + \frac{1}{3} \quad (5.6)$$

$$v^t = \frac{z}{2} + \frac{1}{3} \quad (5.7)$$

optimize the objective functional of the leader.⁵⁾ The leader seeks to induce the follower to choose $v=v^t$ when $T=T^*$, and to make sure that v is sufficiently close, and if possible equal, to v^t when T is different from T^* . He will realize this aim using a strategy of the form (3.2). As a counterpart of (3.6), $Q(z,y)$ is defined in this case to be the set of solutions of

$$\frac{E \{Q(z,y)(2y + 4z - 3)\}}{y/z} = (13z - 2)/2 \quad (5.8)$$

As in (3.8), a family of such solutions is characterized by

$$Q(z,y) = \frac{(13z - 2)g(z,y)}{2 \frac{E \{g(z,y)(2y + 4z - 3)\}}{y/z}} \quad (5.9)$$

where $g(z,y)$ is any $F_{z,y}$ -measurable function satisfying the condition

⁵⁾ It is always possible to find a loss functional for the leader, quadratic in u , v and x , for which (5.6) and (5.7) provide a global minimum.

$$E_{y/z} \{g(z,y)(2y + 4z - 3)\} \neq 0 \quad (5.10)$$

Any such $Q(z,y)$ will induce $v=v^t$, when $T=T^*$. However, when $T \neq T^*$, Q is made optimum by choosing $g(z,y)$ orthogonal to $f_1(z,y)$ defined by (3.14); equivalently,

$$E_{y/z} \{g(z,y)[[13z - 2] + (2y + 4z - 3)]\} = 0 \quad (5.11)$$

A possible solution to this is given by

$$g(z,y) = -3z^2 + 6zy + 4y - 5z + 10 \quad (5.12)$$

with the corresponding $Q(z,y)$ being

$$Q(z,y) = (-3z^2 + 6zy + 4y - 5z + 10)/12 \quad (5.13)$$

Now, the optimum reaction of the follower to an affine strategy (3.2) can be computed to yield, for all values of T :

$$\begin{aligned} v_z(z) = & \{T(1 - E[Q/z]) + E[Qx_0/z] - E[x_0/z] + E[u^t Q/z] - E[u^t/z] \\ & + E[Q^2/z]v^t - E[Q/z]v^t\} / \{E[Q^2/z] - 2E[Q/z] + 2\} \end{aligned} \quad (5.14)$$

when $Q(z,y)$ is chosen as in (5.13), we have

$$E[Q(z,y)/z] \equiv 1, \quad \forall z \in \mathbb{R} \quad (5.15)$$

so that the uncertain term T drops out from the optimal reaction of the follower. Since (5.13) satisfies (5.8), and the optimum reaction is independent from T , the first-best solution (u^t, v^t) is induced for all values of the uncertain parameter T . To illustrate the impact of the additional information-induced optimum strategy (5.13), the optimum reaction of the follower is plotted against

T for three different values of z ($z = -0.5, 0.2$ and 4.0). In Figs. 1-3 the solid lines represent the follower's optimum reaction when g is an arbitrary F_z -measurable function. Note that, for any F_z -measurable g , we have from (5.8)

$$Q = (13z - 2)/2(5z - 2) \quad . \quad (5.16)$$

On the other hand, the dashed lines represent the follower's optimum reaction to the optimum strategy (5.13) when T varies about $T^* = 2$. At $T = T^* = 2$, we have $v = v^t$ under both strategies. However, when T varies about T^* , the optimum reaction induced by the F_z -measurable Q departs from the first-best solution linearly, while under the optimum strategy it does not depart from v^t for any value of T .

The optimum policy (5.13) has been able to track to v^t independent of T , mainly because of the linearity of the optimum reaction (5.14) with respect to T . If the uncertain parameter were B_1 , which is the technology of transforming resources into production, then since the follower's optimum reaction is not linear in B_1 , the outcome of the corresponding near-optimum strategy would be very close to v^t in a certain neighborhood of B_1^* , but considerable departures would be observed when B_1^* is too far away from the actual value. We refer to [9] for an illustrative example of this kind where the uncertain parameter has been taken to be B_1 .

T for three different values of z ($z = -0.5, 0.2$ and 4.0). In Figs. 1-3 the solid lines represent the follower's optimum reaction when g is an arbitrary F_z -measurable function. Note that, for any F_z -measurable g , we have from (5.8)

$$Q = (13z - 2)/2(5z - 2) \quad . \quad (5.16)$$

On the other hand, the dashed lines represent the follower's optimum reaction to the optimum strategy (5.13) when T varies about $T^* = 2$. At $T = T^* = 2$, we have $v = v^t$ under both strategies. However, when T varies about T^* , the optimum reaction induced by the F_z -measurable Q departs from the first-best solution linearly, while under the optimum strategy it does not depart from v^t for any value of T .

The optimum policy (5.13) has been able to track to v^t independent of T , mainly because of the linearity of the optimum reaction (5.14) with respect to T . If the uncertain parameter were B_1 , which is the technology of transforming resources into production, then since the follower's optimum reaction is not linear in B_1 , the outcome of the corresponding near-optimum strategy would be very close to v^t in a certain neighborhood of B_1^* , but considerable departures would be observed when B_1^* is too far away from the actual value. We refer to [9] for an illustrative example of this kind where the uncertain parameter has been taken to be B_1 .

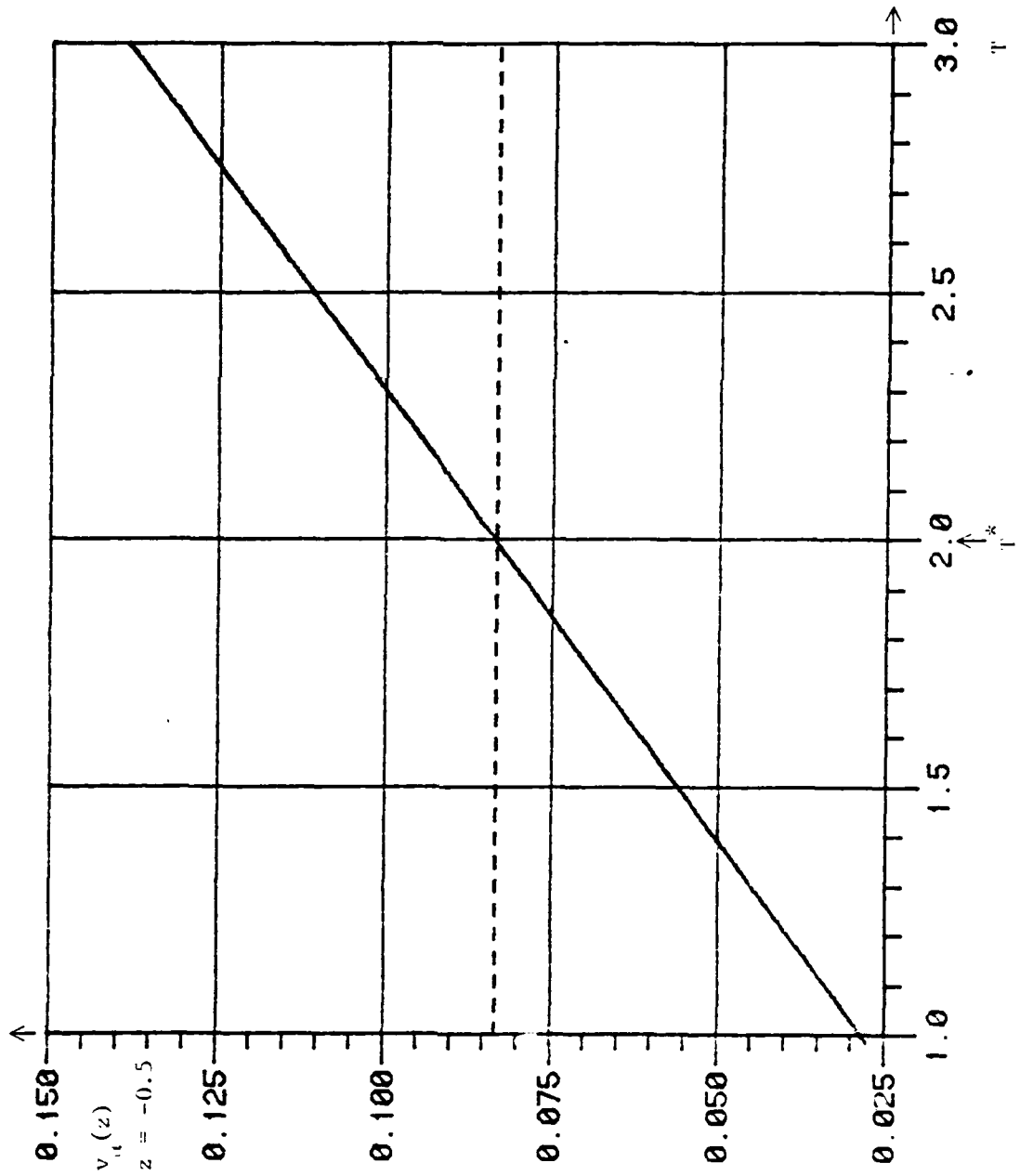


FIG. 13. Optimum response (5.14) of the follower to the leader's optimal policy (5.13) (dashed line) and F_Z -measurable policy (5.16) (solid line) for different values of T .
[Here $\gamma = 0.5$.]

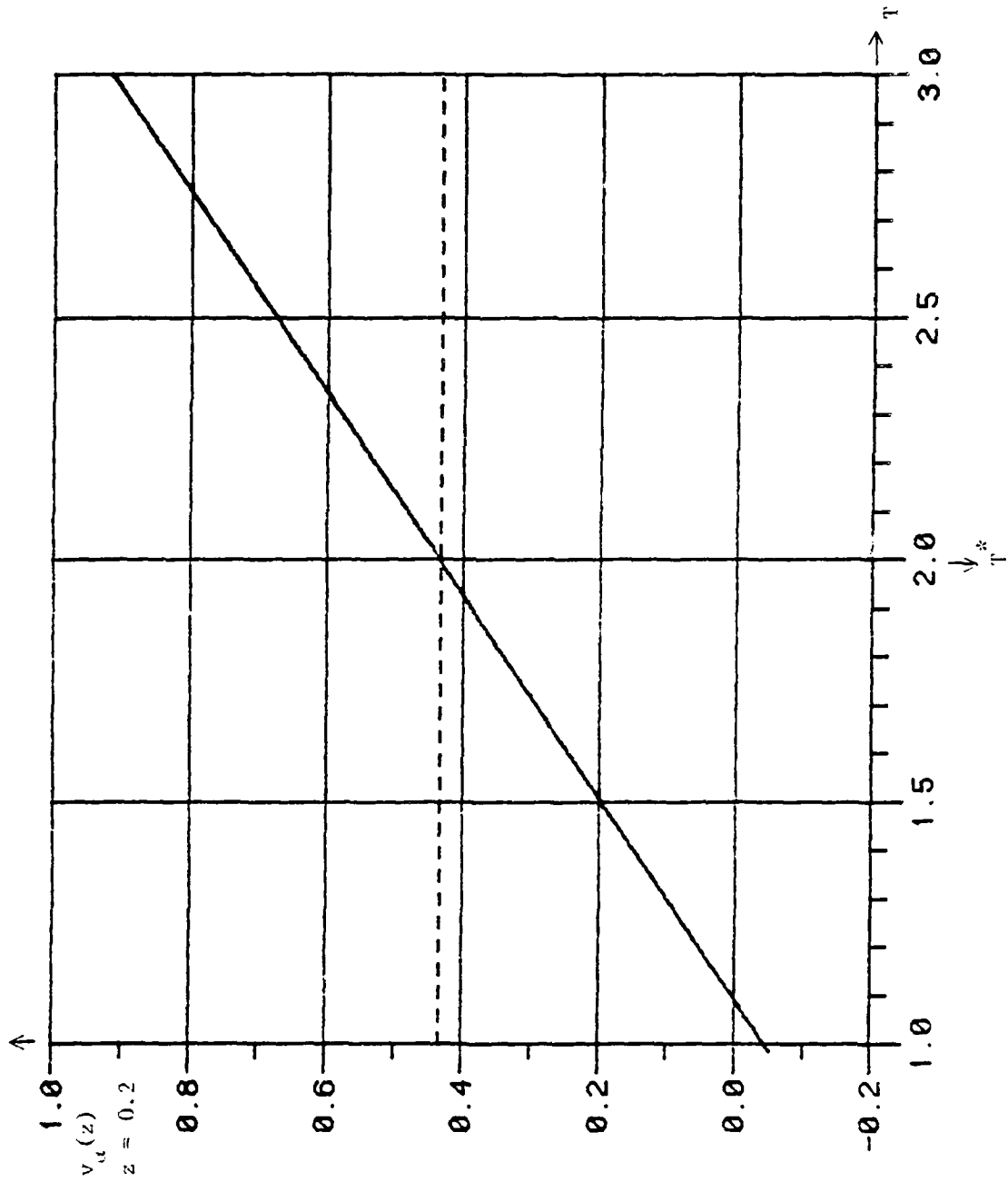


Fig. 2: Optimum response (5.14) of the follower to the leader's optimal policy (5.13) (dashed line) and F_z -measurable policy (5.16) (solid line) for different values of T .
[Here $z = 0.2$.]

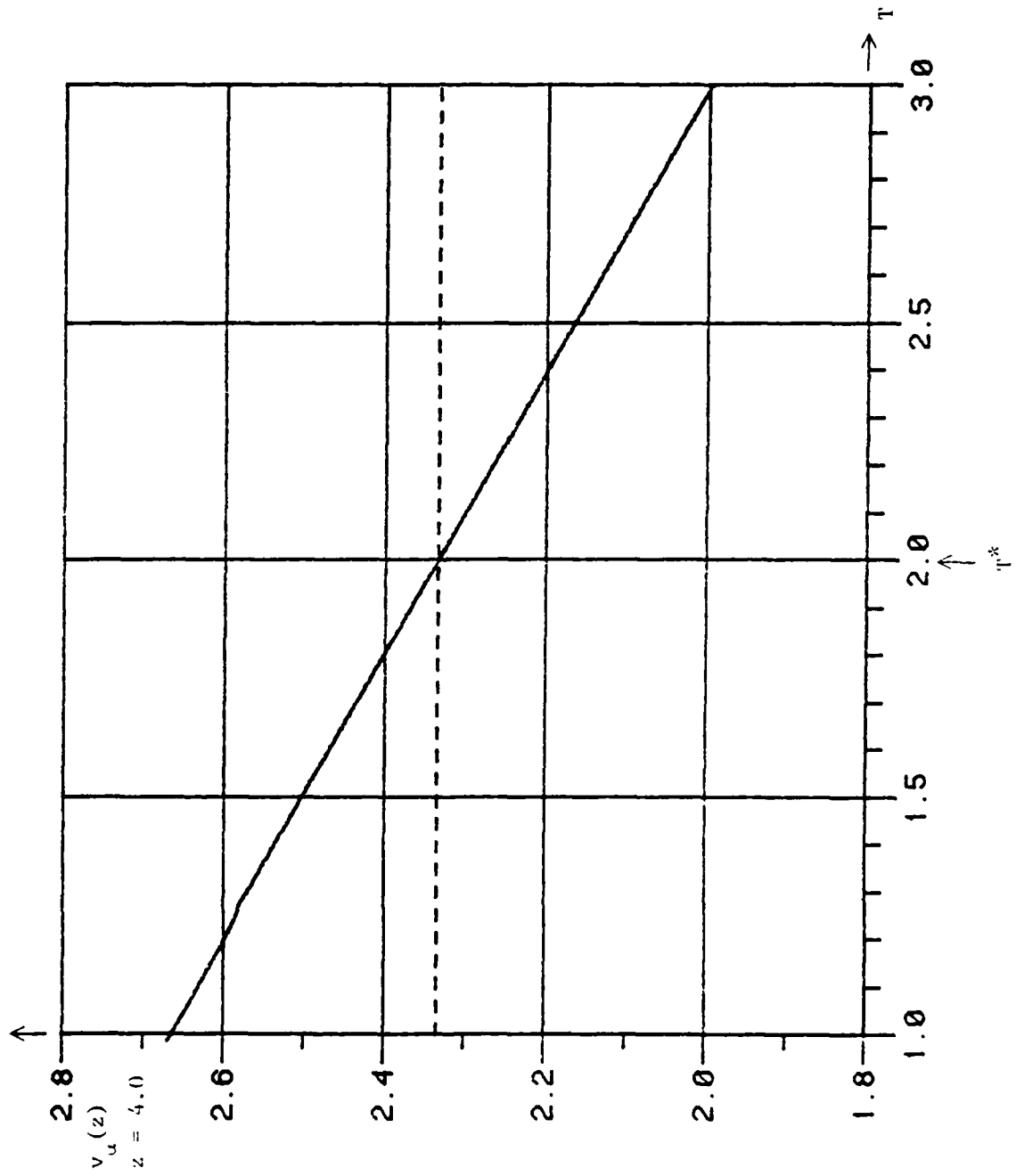


Fig. 3: Optimum response (5.14) of the follower to the leader's optimal policy (5.13) (dashed line) and F_z -measurable policy (5.16) (solid line) for $z = 4.0$. [Here $z = 4.0$.]

VI. CONCLUDING REMARKS

In this paper we have obtained, in the context of multi-person decision making, coordinator (leader) strategies which render the leader's performance index insensitive or minimally sensitive to variations in the parameters of the follower's objective functional, under some linear independence conditions. This appealing property is intrinsic to stochastic decision problems, and has no counterpart in deterministic incentive problems of the type, say, discussed in [3]. We have achieved this by basically exploiting the redundancy present in the leader's dynamic information, his private information and the ensuing fact that the leader's decision variable is not measurable with respect to the follower's information field.

A possible extension of the general approach of this paper would be to continuous-time decision problems with open-loop information, formulated in a Hilbert-space setting, along the lines of [10]. Yet another extension would be to a multistage decision problem wherein the leader uses closed-loop information to compute his robust strategy. Derivation of optimum or near-optimum strategies in these contexts is currently under study.

REFERENCES

1. Harsanyi, J. C. (1968). Games with incomplete information played by "Bayesian" players. Parts I, II and III. Management Science, Vol. 14, 3, 5, 7, 153-182, 320-334 and 486-502.
2. Başar, T. (1984). Affine incentive schemes for stochastic systems with dynamic information. SIAM J. Control and Optimization, March. Has also appeared in the Proc. 1982 American Control Conference, Arlington, VA, 127-132, June 1982.
3. Cansever, D. H. and T. Başar (1982). A minimum sensitivity approach to incentive design problems. Proc. 21st IEEE Conference on Decision and Control, Orlando, FL, 158-163. Also to appear in Large Scale Systems, 1984.
4. Wong, E. (1971). Stochastic Processes in Information and Dynamical Systems. McGraw-Hill, New York.
5. Widemann, J. (1980). Linear Operators in Hilbert Spaces. Springer-Verlag, New York.
6. Arrow, K. (1964). Control in large organizations. In Management Science, Vol. 10, 3, 397-408.
7. Jennergren, L. P. (1980). On the design of incentives in business firms - a survey of some research. In Management Science, Vol. 26, 2, 180-201.
8. Harris, M. and R. M. Townsend (1981). Resource allocation under asymmetric information. Econometrica, Vol. 49, 1, 33-64.
9. Başar, T. and D. H. Cansever (1984). Robust incentive policies for stochastic decision problems in the presence of parametric uncertainty. To appear in the Proc. IFAC World Congress, Budapest, Hungary.
10. Bagchi, A. and T. Başar (1981). Stackelberg strategies in linear-quadratic stochastic differential games. JOTA, Vol. 35, 3, 443-464.

ADDITIONAL BIBLIOGRAPHY

1. Ho, Y. C. (1980). Team decision theory and information structures. Proc. IEEE, Vol. 68, 6, 644-654.
2. Chang, T. S. and Y. C. Ho (1981). Incentive problems: A class of stochastic Stackelberg closed-loop dynamic games. Syst. Control Lett., Vol. 1, 1, 16-21.
3. Myerson, K. B. (1983). Mechanism design by an informed principal. Econometrica, Vol. 51, 6, 1767-1798.
4. Radner, R. Monitoring cooperative agreements in a repeated principal-agent relationship. Econometrica, Vol. 49, 5, 1127-1148.
5. Crawford, V. P. and J. Sobel (1982). Strategic information transmission. Econometrica, Vol. 50, 6, 1431-1452.

APPENDIX A

In this appendix, we will derive an expression for $\frac{dv^n}{d\alpha^n}$ evaluated at $u=u^t$, $v=v^t$ and $\alpha=\alpha^*$. We first differentiate both sides of (3.11) with respect to α to obtain

$$\begin{aligned} E_{y/z} \left\{ Q \left[\left(\frac{d^2 v^t}{d\alpha^2} \right) E_{x/y,z} \{ L_{uv} \} + \left(\frac{dv^t}{d\alpha} \right) E_{x/y,z} \{ L_{uv\alpha} \} \right] - Q^2 \left[\left(\frac{d^2 v^t}{d\alpha^2} \right) E_{x/y,z} \{ L_{uu} \} \right. \right. \\ \left. \left. + \left(\frac{dv^t}{d\alpha} \right) E_{x/y,z} \{ L_{u\alpha\alpha} \} \right] + Q E_{x/y,z} \{ L_{u\alpha\alpha} \} \right\} = E_{x,y/z} \left\{ \left(\frac{d^2 v^t}{d\alpha^2} \right) L_{vv} + \left(\frac{dv^t}{d\alpha} \right) L_{v\alpha\alpha} \right. \\ \left. - Q \left[\left(\frac{d^2 v^t}{d\alpha^2} \right) L_{vu} + \frac{dv^t}{d\alpha} L_{v\alpha\alpha} \right] + L_{v\alpha\alpha} \right\}. \end{aligned} \quad (A.1)$$

Since we seek a $Q(z,y)$ orthogonal to both $\frac{dv^t}{d\alpha}$ and $\frac{d^2 v^t}{d\alpha^2}$, we can consider $\frac{dv^t}{d\alpha}$ already to be annihilated in (A.1), and pull out the F_z -measurable $\frac{d^2 v^t}{d\alpha^2}$ from the conditional expectation. Using this and the smoothing property of conditional expectation, we readily obtain

$$\frac{d^2 v^t}{d\alpha^2} = E_{x,y/z} \{ L_{u\alpha\alpha} Q - L_{v\alpha\alpha} \} / E_{x,y/z} \{ L_{uu} Q^2 - 2L_{uv} Q + L_{vv} \}. \quad (A.2)$$

Then, an induction type of argument yields

$$\frac{d^n v^t}{d\alpha^n} = E_{x,y/z} \{ L_{u\alpha^n} Q - L_{v\alpha^n} \} / E_{x,y/z} \{ L_{uu} Q^2 - 2L_{uv} Q + L_{vv} \}. \quad (A.3)$$

March 1984

FOR 23rd CDC, ONLY

NEAR-OPTIMUM INCENTIVE POLICIES IN STOCHASTIC TEAM PROBLEMS
WITH DISCREPANCIES IN GOAL PERCEPTIONS[†]

by

Derya H. Cansever and Tamer Başar
Decision and Control Laboratory
Coordinated Science Laboratory and the Department
of Electrical and Computer Engineering
University of Illinois
1101 W. Springfield Avenue
Urbana, Illinois 61801

ABSTRACT

In this paper, we consider a two-agent stochastic team decision problem with a hierarchical decision structure in a general Hilbert space setting. One of the agents has a different perception of the common team objective functional, as quantified in terms of a finite dimensional parameter vector. The other hierarchically superior agent, uninformed about this discrepancy, but endowed with a suitable information structure, designs a near-optimal incentive policy such that the incurred value of the original team functional is arbitrarily close to its global optimum, in spite of the existing discrepancy. The general solution is determined by some orthogonality relations in some appropriately constructed probability measure spaces, and leads to particularly simple incentive policies.

Address for Correspondence:

Professor Tamer Başar, Coordinated Science Laboratory
University of Illinois, 1101 W. Springfield Avenue
Urbana, Illinois 61801 ; Tel: 217/333-3607

[†] Research reported herein was supported in part by the Office of Naval Research under Contract N00014-82-K-0469 and in part by the U. S. Air Force Office of Scientific Research under Contract AFOSR-84-0054.

I. INTRODUCTION

In this paper, we consider a general class of two-agent stochastic dynamic decision problems wherein both of the agents jointly optimize a given objective functional, which gives rise to a so-called team decision problem [1]. Within the spirit of [2], we will relax the basic assumption of a team decision problem that the agents perceive the common goal in exactly the same way, by allowing one of the agents to have a somewhat different perception of the common objective, and quantifying this discrepancy in terms of an objective functional which differs from the original objective functional up to a finite-dimensional parameter vector α . We further assume that the other agent is not informed of this discrepancy, but is able to monitor the decision of the former by assuming a hierarchically superior position in the decision process. The problem we address in the sequel is derivation of near-optimal decision policies for the hierarchically superior agent such that the variation in the decision value of the agent who is faced with a discrepancy in the common objective functional is kept to a certain minimum.

We will approach this problem using the notion of near optimum stochastic incentive policies [3]. Here the hierarchically superior agent (decision maker), henceforth referred to as DM1, incarnated with a suitable information structure to be delineated in the sequel, induces the other agent, henceforth referred to as DM2, to behave in a desired manner, while reducing the effects of the discrepancies to an arbitrarily small value under certain conditions. The next section presents a precise mathematical formulation for the problem, while Section III presents the main results. Proof of the main theorem has only been outlined in this paper, but a fuller version will be provided at the Conference in Las Vegas.

II. PROBLEM FORMULATION

Let $(\Omega, \mathcal{F}, P)$ be an underlying probability space, on which three correlated random variables x , y and z are defined. Here, $x \in \mathbb{R}^n$ denotes the random state of the nature, and $z \in \mathbb{R}^m$ is the common (to both agents) and $y \in \mathbb{R}^p$ private (to DM1) information related to x . Let U and V be given real separable Hilbert spaces denoting the decision spaces of DM1 and DM2, respectively, and let Γ_{1z} and Γ_{2z} denote the corresponding policy spaces, characterized for each fixed $z \in \mathbb{R}^m$ by

$$\Gamma_{2z} = \{\text{measurable } \gamma_2 : \mathbb{R}^m \rightarrow V, \text{ such that } \langle \gamma_2(z), \gamma_2(z) \rangle_V < \infty\} \quad (2.1)$$

$$\Gamma_{1z} = \{\text{measurable } \gamma_1 : V \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow U, \text{ such that } E_{y/z} \{ \langle \gamma_1[\gamma_2(z), z, y], \gamma_1[\gamma_2(z), z, y] \rangle_U < \infty, \forall \gamma_2 \in \Gamma_{2z} \} \} \quad (2.2)$$

We also let $\Gamma_{1z}^S \subset \Gamma_{1z}$ indicate the set of all "static" policies for DM1, defined by

$$\Gamma_{1z}^S = \{\text{measurable } \gamma_1 : \mathbb{R}^m \times \mathbb{R}^p \rightarrow U, \text{ such that } E_{y/z} \langle \gamma_1(z, y), \gamma_1(z, y) \rangle_U < \infty \}. \quad (2.3)$$

These transformations are restricted by the (implicit) condition that the expectations are well-defined. Here, $\langle \cdot, \cdot \rangle_U$ and $\langle \cdot, \cdot \rangle_V$ denote the inner products associated with U and V , respectively, and the Hilbert spaces Γ_{1z} and Γ_{2z} have their own derived inner products.

We now introduce the objective functional $L : \mathbb{R}^n \times U \times V \times \mathbb{A} \rightarrow \mathbb{R}$ which describes the team decision problem, and where $\mathbb{A}$ is a subset of $\mathbb{R}$, on which a parameter α takes values. Here, the parameter α characterizes the above mentioned

discrepancy. For a fixed $z \in \mathbb{R}^m$ and for all $\alpha \in A$, L belongs to $L_2[\Omega, \mathcal{P}(z, y/z)]$, the Hilbert space of square-integrable random variables under the conditional probability measure $\mathcal{P}(x, y/z)$. We further let the team objective functional be

$$J : \mathbb{R}^m \times \Gamma_1 \times \Gamma_2 \times A \rightarrow \mathbb{R}$$

where

$$J(z, \gamma_1, \gamma_2, \alpha) = E_{x, y/z} \{L(x, u, v, \alpha) | u = \gamma_1(v, z, y), v = \gamma_2(z)\}, \forall \alpha \in A, z \in \mathbb{R}^m. \quad (2.4)$$

We assume that $L(x, u, v, \alpha)$ is strictly convex on $U \times V$ for each $x \in \mathbb{R}^m$ and $\alpha \in A$, and is continuous in all of its arguments. Now, for a fixed $\alpha \in A$, say α^* , and fixed $z \in \mathbb{R}^m$ (the realized value of this random variable), let

$\{u_{z,y}^t = \gamma_1^t(z, y), v_z^t = \gamma_2^t(z)\}$ denote a unique pair in $\Gamma_{1z}^s \times \Gamma_{2z}$ that globally minimizes $J(z, \gamma_1, \gamma_2, \alpha^*)$, where $u_{z,y}^t$ is square integrable under the conditional probability measure $\mathcal{P}(y/z)$ derived from $\mathcal{P}(x, y/z)$, with $u_{z,y}^t$ and v_z^t are continuous in z, y and z , respectively. With this pair, one obtains the expected minimum value of the team decision problem, conditioned on the common observation z , at $\alpha = \alpha^*$, which we denote by $J_{\alpha^*, z}^t$. We assume that the hierarchically superior DM1 perceives this team decision problem at $\alpha = \alpha^*$, while DM2 has a different perception of α , say α^+ , and this discrepancy gives rise to a different objective functional for DM2, $J(z, \gamma_1, \gamma_2, \alpha^+)$; with this discrepancy, the decision problem ceases to be a team one.

By virtue of the information structure of the problem, and of his hierarchically superior position, DM1 is allowed to announce his policy $\gamma_1(\gamma_2(z), z, y)$ in advance, and to implement it. Then, DM2 computes his optimum reaction $v_z^{\gamma_1(\alpha^+)}$, for each announced γ_1 , given z and his perception α^+ .

Since α^+ is different from α^* , the parameter value of the team decision problem that DM1 sticks to, and since DM1 does not

know what DM2's perception α^+ is, we generally have $v_z^{\gamma_1(\alpha^+)} \neq v_z^t$,

$\forall \gamma_1 \in \Gamma_{1z}$. The goal adopted in this paper is to find a near-optimal $\gamma_1 \in \Gamma_{1z}$ such that $v_z^{\gamma_1(\alpha^+)}$ is arbitrarily close, or equal to v_z^t , (depending on the structure of L) when α^+ is within a certain neighborhood of α^* . By continuity of γ_1 with respect to v , the resulting $u_{z,y} \in \Gamma_{1z}^s$ will be arbitrarily close to $u_{z,y}^t$, and as a consequence of this near-optimal policy, the incurred team cost will be arbitrarily close to $J_{\alpha^*,z}^t$, in spite of the existing discrepancy regarding the team objective functional.

III. MAIN RESULTS

In this section, we will assume that α^+ is in a neighborhood of α^* , such that the Taylor series of $v_z^{\gamma_1}(\alpha^+)$ about α^* converges for a given γ_1 to be made precise in the sequel. We also assume that $L(x, u, v, \alpha)$ is Fréchet analytic [4] in u , v , and analytic in α , for each $\alpha \in A$, where A is identified as the above neighborhood. The random variables x , z and y are jointly second-order random variables and L is measurable with respect to (w.r.t.) the sigma-field $\mathcal{F}_x$ generated by x , so that

$$E \{ L(\gamma_1(v_z^{\gamma_1}(\alpha), z, y), v_z^{\gamma_1}(\alpha), z, \alpha) | z \} \quad (3.1)$$

is well-defined for all $\gamma_1 \in \Gamma_{1z}$, $v_z^{\gamma_1}(\alpha) \equiv \gamma_2 \in \Gamma_{2z}$ and $\alpha \in A, z \in \mathbb{R}^m$.

Towards the goal set in the previous section, we will now let DM1 adopt an affine (in v) policy [5-7], given by¹⁾

$$\gamma_1(v, z, y) = u_{z,y}^t - Q_{z,y}(v - v_z^t) \quad (3.2)$$

where for each fixed z and y , $Q_{z,y}(\cdot) \in \mathcal{L}(V, U)$ is a bounded linear operator mapping V into U , and for each fixed $\gamma_2(z) \in \Gamma_{2z}$, the resulting $u_{z,y}^t - Q_{z,y}(v - v_z^t)$ is an element of Γ_{1z} , with $v = \gamma_2(z)$.

¹⁾ If the underlying decision problem is a Stackelberg game, a necessary condition for the existence of an affine policy inducing (u^t, v^t) is that $\nabla_u L$ should be nonvanishing at the nominal point [5,3]. In our problem, this condition is not met. However, since (u^t, v^t) is naturally induced at the nominal value of α , the condition that $\nabla_u L$ be nonvanishing is not required, and in fact is not met in a team decision problem with a hierarchical structure.

We note in passing, that this γ_1 enjoys the desirable property of

$$\gamma_1(v, z, y) = u_{z, y}^t \quad \text{when } \alpha^+ = \alpha^* \quad (3.3)$$

so that the only restrictions on $Q_{z, y}(v)$ are that it should be an element of $\mathcal{L}(V, U)$ for each fixed z and y and $\forall v \in V$, and it should be square integrable under $P(y/z)$ for a fixed $v \in V$, for fixed $z \in \mathbb{R}^m$. Equation (3.3) follows from the fact that

$$v_z^{\gamma_1}(\alpha) = v_z^t \quad \text{when } \alpha^+ = \alpha^*, \quad \forall Q_{z, y}(\cdot) \in \mathcal{L}(V, U) \quad (3.4)$$

We now let

$$\alpha^+ = \alpha^* + h_\alpha, \quad h_\alpha \in \mathbb{R} \text{ such that } \alpha^+ \in A \quad (3.5)$$

Then, (3.4) no longer holds, and given the common observation z , DM2 faces the optimization problem of minimizing

$$E_{x, y/z} \{L(x, u^t - Q_{z, y}(v - v^t), v, x, \alpha^+)\} \quad (3.6)$$

over $v \in V$. This optimization problem, also called the optimum reaction of DM2, can be characterized by

$$D_v \{ E_{x, y/z} \{L(x, u^t - Q_{z, y}(v - v^t), v, x, \alpha^+)\} \} (h_v) = 0, \quad \forall h_v \in V \quad (3.7)$$

where $D_v(\cdot)$ is the Fréchet differential operator, $D_v(\cdot) \in \mathcal{L}(V, \mathbb{R}) \equiv V^*$, where V^* is the topological dual of V [8]. Since L is Fréchet analytic and is an element of $L_2(\Omega, P(y/z))$, it is not hard to show that the expectations and differentiations are interchangeable, so that (3.7) rewrites as

$$E_{x, y/z} \{ D_v(L) - Q_{z, y}^* \{ D_u(L) \} \} (h_v) = 0, \quad \forall h_v \in V \quad (3.8)$$

Where $D_u(\cdot)$ is the Fréchet differential operator w.r.t. $u \in U$, with $D_u(\cdot) \in \mathcal{L}(U, \mathbb{R}) \equiv U^*$, and $Q_{z,y}^*(\cdot)$ is the adjoint of $Q_{z,y}(\cdot)$, with $Q_{z,y}^*(\cdot) \in \mathcal{L}(U^*, V^*)$, for each fixed z, y .

For the characterization of a near-optimum policy, we will need the Taylor expansion of $v_z(\alpha^+) = v_z(\alpha^* + h_\alpha)$ about α^* . Towards this goal, we note that (3.8), the necessary condition for optimality, is in fact an identity for all $\alpha \in \mathbb{A}$. Therefore, when we take the (ordinary) derivative of (3.8) with respect to α it will still be equal to zero. We then have

$$\begin{aligned} E_{x,y/z} \{ d_\alpha D_v L + D_v^2(L) d_\alpha v - Q^*(D_u D_v L) d_\alpha v + (Q^*)^2 (D_u^2 L) d_\alpha v - Q^*(D_v D_u L) d_\alpha v \\ - Q^*(d_\alpha D_u L) \} (h_v, h_A) = 0 \quad \forall h_v \in V, \quad \forall h_A \in \mathbb{A} \end{aligned} \quad (3.9)$$

where $D_v^2(\cdot) \in \mathcal{L}(V \times V; \mathbb{R})$, $D_u^2 \in \mathcal{L}(U \times U; \mathbb{R})$, $(Q^*)^2 \in \mathcal{L}(U^* \times U^*; V^*)$ and $d_\alpha(\cdot)$ represents the operator which takes the ordinary derivative with respect to the scalar α . Rearranging (3.9), we obtain

$$\begin{aligned} E_{x,y/z} \{ [D_v^2 L - Q^*(D_u D_v L) - Q^*(D_v D_u L) + (Q^*)^2 (D_u^2 L)] d_\alpha v + d_\alpha D_v L \\ - Q^*(d_\alpha D_u L) \} (h_v, h_A) = 0, \quad \forall h_v \in V, \quad \forall h_A \in \mathbb{A}. \end{aligned} \quad (3.10)$$

We now observe that $(D_v^2 - Q^* D_u D_v - Q^* D_v D_u + (Q^*)^2 D_u^2)(\cdot)$ is a strictly positive operator, due to the strict convexity of L on $U \times V$, and therefore, if we can find a $Q_{z,y} \in \mathcal{L}(V, U)$ such that its adjoint satisfies at $\alpha = \alpha^*$

$$E_{x,y/z} \{ d_\alpha D_v L - Q^*(d_\alpha D_u L) \} (h_v, h_A) = 0 \quad \forall h_v \in V, \quad \forall h_A \in \mathbb{A}, \quad (3.11)$$

AD-A145 212

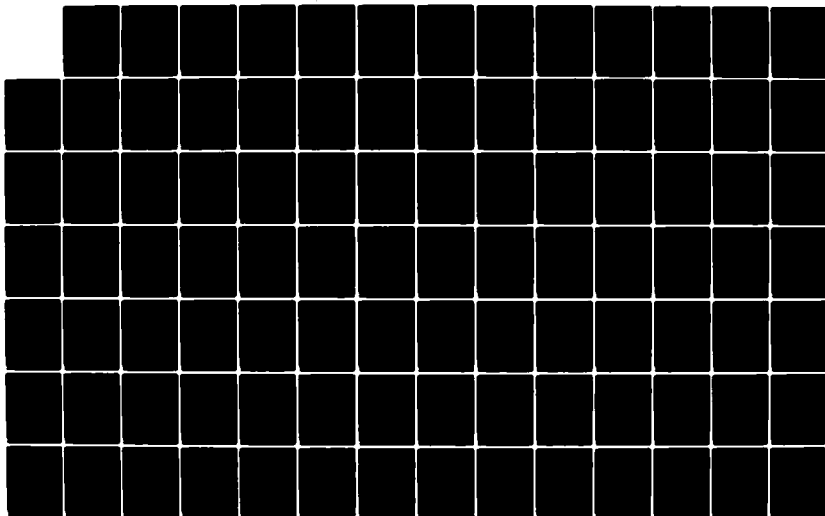
ROBUST TEAM-OPTIMAL AND LEADER-FOLLOWER POLICIES FOR
DECISION MAKING IN C... (U) ILLINOIS UNIV AT URBANA
T BASAR ET AL. JUL 84 N00014-82-K-0469

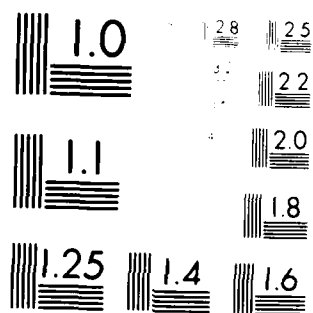
2/3

UNCLASSIFIED

F/G 12/1

NL





M. RUCOFF, RES. LAB. NBS, GAITHERSBURG, MD.

we will necessarily have

$$\left. \begin{aligned} d_{\alpha} v &= 0 \\ \alpha &= \alpha^* \end{aligned} \right| \quad (3.12)$$

Now, the Taylor series expansion of $v_z(\alpha^+)$ about α^* is

$$v_z(\alpha^+) = v_z(\alpha^*) + \sum_{n=1}^{\infty} \frac{1}{n!} d_{\alpha}^n(v_z(\alpha^*)) h_{\alpha}^n, \quad \forall h_{\alpha} \text{ such that } \alpha^* + h_{\alpha} = \alpha^+ \in A. \quad (3.13)$$

If we can realize (3.12), $v_z(\alpha^+)$ will match $v_z(\alpha^*) = v_z^t$ up to first order, $\forall \alpha^+ \in A$; and the existence of such an expansion is guaranteed by the Fréchet analyticity of L , as it will be clear in the sequel. Towards the realization of (3.13), we recall that $Q_{z,y}(\cdot)$ is a linear bounded and square integrable operator under $P(y/z)$, mapping V into U . Likewise, its adjoint Q^* is a linear operator with the above properties, mapping U^* into V^* . One such $Q_{z,y}^*(\cdot) \in \mathcal{L}(U^*, V^*)$ is

$$Q_{z,y}^*(\cdot) = \langle g_{z,y}, \cdot \rangle_{U^*} \ell_z / \int_{x,y/z} E \{ \langle g_{z,y}, f_{z,y} \rangle_{U^*} \}, \quad (3.14)$$

where $g_{z,y} \in U^*$, $\ell_z \in V^*$ and $f_{z,y} \in U^*$ satisfying

$$\int_{x,y/z} E \{ \langle g_{z,y}, f_{z,y} \rangle_{U^*} \} \neq 0, \quad (3.15)$$

and $g_{z,y}$, $f_{z,y}$ and ℓ_z are nonvanishing and bounded functionals, which are also weakly continuous in z , y and z , respectively. Furthermore, for each square integrable $u^* \in U^*$, and fixed $z \in \mathbb{R}^m$, $\langle g_{z,y}, u^* \rangle_{U^*}$ and $\langle f_{z,y}, u^* \rangle_{U^*}$ are elements of $L_2(\Omega, P(y/z))$. When we plug (3.14) into the right-hand side of (3.11), the goal will be accomplished if we can equate the following quantity to zero:

$$\begin{aligned}
& E_{x,y/z} \{d_{\alpha} D_v L - \langle g_{z,y}, d_{\alpha} D_u L \rangle_{u^*} z\} / E_{x,y/z} \{\langle g_{z,y}, f_{z,y} \rangle_{u^*}\} \\
&= E_{x,y/z} \{d_{\alpha} D_v L - \langle g_{z,y}, d_{\alpha} D_u L \rangle_{u^*} z\} / [\langle g_{z,y}, f_{z,y} \rangle_{u^*}] \\
&= E_{x,y/z} \{d_{\alpha} D_v L \langle g_{z,y}, f_{z,y} \rangle_{u^*} - z \langle g_{z,y}, d_{\alpha} D_u L \rangle_{u^*}\} / E_{x,y/z} \{\langle g_{z,y}, f_{z,y} \rangle_{u^*}\} \\
&= E_{x/z} \{[d_{\alpha} D_v L - z] E_{x,y/z} \{\langle g_{z,y}, f_{z,y} - d_{\alpha} D_u L \rangle_{u^*}\}\} / E_{x,y/z} \{\langle g_{z,y}, f_{z,y} \rangle_{u^*}\} \\
&= 0 .
\end{aligned} \tag{3.16}$$

For this, it is sufficient to choose a nonvanishing $g_{z,y}$ such that

$$E_{x,y/z} \{\langle g_{z,y}, f_{z,y} - d_{\alpha} D_u L \rangle_{u^*}\} = 0 . \tag{3.17}$$

To realize (3.17), we pick two bounded nonvanishing elements f_1 and f_2 of U^* , weakly continuous in z and y , which are also square integrable under $P(y/z)$, and let

$$g_{z,y} = \langle f_1, f_{z,y} - d_{\alpha} D_u L \rangle_{u^*} f_2 - E_{x,y/z} \{\langle f_1, f_{z,y} - d_{\alpha} D_u L \rangle_{u^*}\} f_2 . \tag{3.18}$$

It is easy to verify that a $g_{z,y}$ characterized by (3.18) satisfies (3.17) without violating (3.15) if

$$d_{\alpha} D_u L \neq 0 . \tag{3.19}$$

We will now show that $Q_{z,y}^*(\cdot)$ characterized by (3.14) and (3.18) leads to a well-defined policy. In these equations we had complete freedom over the

choice of $f_{z,y}$, l_z , f_1 and f_2 . The only term to study is $d_\alpha D_u L$. Since L is Fréchet analytic, $d_\alpha D_u L$ is a bounded linear functional for each fixed z and y . The following lemma is useful in establishing the fact that it is also square integrable under $P(y/z)$.

Lemma 1: The n 'th order Fréchet derivative of L w.r.t. u is square-integrable under $P(y/z)$, for $n=1,2,\dots,N < \infty$.

Proof: Since L is Fréchet analytic in U in a domain D , there is an open sphere S of radius r on which L is locally bounded, say by $M(u)$. For each $h \in U$ we have [4,p. 161]

$$\|D_u^n L(u)(h)\| \leq n! \frac{M(u) \|h\|^n}{r^n}.$$

Now, since $M(u)$, being a local bound on L , is (square) integrable under $P(y/z)$, by a Corollary to the Dominated Convergence Theorem, [10,p. 126], $D_u^n L$ is also square integrable under $P(y/z)$. $\square$

In the above discussion, L was evaluated at $u_{z,y}^t$, v_z^t and α^* , which makes it a function of z and y ; in fact, it is continuous in these random variables under the hypothesis cited while formulating the problem. The next step is to prove that the proposed incentive policy given by (3.2), (3.14) and (3.18) is a well-defined one, which is guaranteed if $Q_{z,y}^*$ is measurable in z and y . The following lemma is useful in establishing an even stronger regularity property for $Q_{z,y}^*$.

Lemma 2: The n 'th order Fréchet derivative of L w.r.t. u is weakly continuous in z and y , for $n=1,2,\dots,N < \infty$.

Proof: We will first consider the dependence on z . Since $L^z(u)$ is a continuous function of z , for a given u ($u_{z,y}^t$ in this case), and given z_0 (e.g. the realized observation), for each $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\|z - z_0\| < \delta \Rightarrow \|M^z(u) - M^{z_0}(u)\| < \varepsilon$$

where $M^{z_0}(u)$ is the local bound on L . Using the upper bound given in the proof of Lemma 1 on the Fréchet derivative $D_u^n L^z(u)(h)$, there exists a $\delta' = \frac{\delta r^n}{n! \|h\|^n}$ such that for each $\varepsilon > 0$, we have

$$\|z - z_0\| < \delta' \Rightarrow \|D_u^n L^z(u)(h) - D_u^n L^{z_0}(u)(h)\| < \varepsilon, \forall h \in U.$$

The similar proof applies to the y -dependence of $D_u^n L(u)(h)$. □

We now summarize the preceding discussion below.

Proposition 1: Let (3.19) be satisfied. Then, the optimum reaction of DM2, $v_z(\alpha^+)$, can be matched to the team-optimum decision v_z^t up to first order in the Taylor expansion of $v_z(\alpha^+)$, $\forall \alpha^+ \in A$, using an affine incentive policy (3.2), where the adjoint of $Q_{z,y}(\cdot)$ satisfies (3.14), with $g_{z,y}$ given by (3.18).

Proof: A policy with the above properties has been obtained by construction prior the statement of the proposition. The fact that this policy is square integrable under $P(y/z)$ is guaranteed by Lemma 1. Lemma 2 implies that $Q_{z,y}^*$ is weakly continuous in z and y , a stronger property than measurability.

Hence, this policy is an element of Γ_{1z} . □

We now focus attention to the next term in the Taylor expansion of $v_z(\alpha^+)$ about α^* , $d_{\alpha}^2 v_z(\alpha)$. Taking the second derivative of (3.8) w.r.t. α , and also by considering the fact that $d_{\alpha} v$ can be made equal to zero at $\alpha = \alpha^*$,

we obtain (at $\alpha = \alpha^*$)

$$\begin{aligned} E_{x,y/z} \{ [D_v^2 L - Q^*[D_u D_v L] - Q^*[D_v D_u L] + (Q^*)^2 [D_u^2 L]] d_\alpha^2 v + d_\alpha^2 D_v L \\ - Q^*(d_\alpha^2 D_u L) \} (h_v, h_A^1, h_A^2) = 0, \quad \forall h_v \in V, \quad \forall h_A^1, h_A^2 \in A \end{aligned} \quad (3.20)$$

Now, if $Q_{z,y}^*(\cdot)$ satisfies

$$E_{x,y/z} \{ d_\alpha^2 D_v L - Q^*(d_\alpha^2 D_u L) \} (h_v, h_A^1, h_A^2) = 0, \quad \forall h_v \in V, \quad \forall h_A^1, h_A^2 \in A \quad (3.21)$$

and simultaneously (3.21), both at $\alpha = \alpha^*$, then the optimum reaction of DM2 will match the team-optimum decision v_z^t up to the second order in the Taylor expansion of $v_z(\alpha^*)$. Adopting the representation (3.14) for $Q_{z,y}^*(\cdot)$, as a counterpart of (3.17) we obtain

$$E_{x,y/z} \{ \langle g_{z,y}, f_{z,y} - d_\alpha^2 D_u L \rangle_u \} = 0 \quad (3.22)$$

which, together with (3.17), yields a characterization for $g(z,y)$ annihilating the first and second order terms in the Taylor expansion of $v_z(\alpha^*)$. To show that (3.17) and (3.22) can be simultaneously satisfied, we define a new inner product on U^* ,

$$\langle \cdot, \cdot \rangle = E_{x,y/z} \{ \langle \cdot, \cdot \rangle_{u^*} \} \quad (3.23)$$

Now, let S^2 be a subspace of V^* generated by the pair $(f_{z,y} - d_\alpha D_u L, f_{z,y} - d_\alpha^2 D_u L)$. One can construct a vector $e_{z,y}$ orthogonal to S^2 under $\langle \cdot, \cdot \rangle$, by Gram-Schmidt orthogonalization procedure, and equate this $e_{z,y}$ to $g_{z,y}$. With this $g_{z,y}$, (3.17) and (3.22) will be simultaneously satisfied. As a

counterpart of (3.19), we also need

$$d_{\alpha}^2 D_u L \neq 0 \quad . \quad (3.24)$$

This discussion then leads to the following proposition which will be proven rigorously in the final version of the paper, to be presented at CDC 84.

Proposition 2: Let (3.19) and (3.24) be satisfied. Then, the optimum reaction of DM2, $v_z(\alpha^+)$, can be matched to the team-optimum decision v_z^t up to second order in the Taylor expansion of $v_z(\alpha^+)$, $\forall \alpha^+ \in A$, using an affine policy (3.2), where the adjoint of $Q_{z,y}(\cdot)$ satisfies (3.14), with $g_{z,y}$ being orthogonal to the subspace S^2 . $\square$

Now, it is clear that this procedure can be extended to annihilate $d_{\alpha}^n v$, $n=1, \dots, N$ where N is an arbitrarily large positive integer, provided that we have

$$d_{\alpha}^n D_u L \neq 0, \quad n=1, \dots, N \quad . \quad (3.25)$$

To realize the above assertion, it is sufficient to choose a $g_{z,y}$ orthogonal to the subspace S^N generated by $\{f_{z,y} - d_{\alpha}^n D_u L\}$, $n=1, \dots, N$.

Let A be a bounded interval of the real line, and let $\bar{A}$ denote a closed set contained in A . Then, by Weirstrass Theorem [9], the Taylor expansion of $v_z(\alpha^+)$ around α^* is a uniform approximation of $v_z(\alpha^+)$. On the other hand, under (3.25), we can annihilate the term in the Taylor expansion of $v_z(\alpha^+)$ up to N' th order. This then leads to the following theorem.

Theorem 1: Let (3.25) be satisfied, and the value of $\alpha^+ \in \bar{A}$ be uncertain to DM1. Then, there exists an affine incentive policy $\gamma_1^0 \in \Gamma_{1z}$ for DM1 such that the optimum response of DM2 to this policy, $v_z^{\gamma_1^0}(\alpha^+)$, satisfies

$$\|v_z^{\gamma_1^0}(\alpha^+) - v_z^t\|_V < \varepsilon$$

where ε is an arbitrarily small positive number. Such a policy is represented by

$$\gamma_1^0(v, y, z) = u_{z,y}^t - Q_{z,y}^0(v_z - v_z^t)$$

where $Q_{z,y}^0(\cdot) \in \mathcal{L}(V, U)$, and its adjoint satisfies (3.14), with $g_{z,y}$ being orthogonal to the subspace S^N . This γ_1^0 is called a near-optimal policy for DML.

□

Of course, if $v_z(\alpha^+)$ has an exact representation in terms of a finite number of powers in h_A , then this near-optimal policy becomes an optimal one.

Remark: The private information y guarantees the existence of a nonzero $g_{z,y}$ orthogonal to the subspace S^N generated by $\{f_{z,y} - d_{\alpha}^n D_u L\}$, under the inner product $\langle \cdot, \cdot \rangle$ defined in (3.23) provided that (3.25) holds. If the underlying decision space U is infinite dimensional, without the private information y , one can still find a g_z orthogonal to S^N under the inner product $\langle \cdot, \cdot \rangle_{u^*}$. In this case, the linear operator $Q_z^*(\cdot)$ is characterized by

$$Q_z^*(\cdot) = \langle g_z, \cdot \rangle_{u^*} \ell_z / \langle g_z, f_z \rangle_{u^*} \quad (3.26)$$

However, when U is finite dimensional, say of dimension M , with inner product defined by the scalar multiplication of two vectors, no nonzero $g_z \in U^*$ can be orthogonal to more than $M-1$ linearly independent vectors, so that Theorem 1 would not hold, true in general. The private information y induces the inner product $\langle \cdot, \cdot \rangle$, under which it is possible to find a $g_{z,y}$ orthogonal to S^N , regardless of the dimensionality of U , subject to some regularity conditions on L .

□

IV. EXTENSIONS

The final version of the paper, to be presented at CDC 84, will include rigorous proofs of Proposition 2 and Theorem 1 given above, and will also discuss the case where $\alpha \in \mathbb{A}$ is a vector. It will be shown that the results presented in this paper remain basically intact when α is a finite-dimensional vector.

The final version of the paper will also include an infinite-horizon LQG team decision problem with discounted cost functional, with an existing discrepancy between the agents about the discount factor. Some numerical studies will supplement the theoretical results.

V. REFERENCES

- [1] J. Marchak and R. Radner, Economic Theory of Teams, Yale University Press, New Haven, CT, 1972.
- [2] D. Cansever, T. Başar and J. B. Cruz, Jr., "Robustness of Incentive Policies in Team Problems with Discrepancies in Goal Perceptions," Proc. 4th IFAC/IFORS/IIASA Conf. on Modelling and Control of National Economies, 1983.
- [3] T. Başar and D. Cansever, "Robust Incentive Policies for Stochastic Decision Problems in the Presence of Parametric Uncertainty." To appear in the Proc. IFAC World Congress, Budapest, Hungary, 1984.
- [4] A. V. Balakrishnan, Applied Functional Analysis, Springer-Verlag, New York, 1976.
- [5] T. Başar, "Affine Incentive Schemes for Stochastic Systems with Dynamic Information," SIAM J. Control and Optimization, March 1984.
- [6] Y. P. Zheng and T. Başar, "Existence and Derivations of Optimal Affine Incentive Schemes for Stackelberg Games with Partial Information: A Geometric Approach," International Journal of Control, Vol. 35, No. 6, pp. 997-1011, 1982.
- [7] T. S. Chang and Y. C. Ho, "Incentive Problems: A Class of Stochastic Stackelberg Closed-Loop Dynamic Games," Syst. Control Lett., Vol. 1, No. 1, pp. 16-21, 1981.
- [8] R. F. Curtain and A. J. Prichard, Functional Analysis in Modern Applied Mathematics, Academic Press, New York, 1977.
- [9] W. Rudin, Principles of Mathematical Analysis, McGraw-Hill, New York, 1976.
- [10] M. Loève, Probability Theory, Van Nostrand Reinhold Co., New York, 1963.

TRANSACTIONS
of the
NINTH PRAGUE CONFERENCE
on

INFORMATION THEORY,
STATISTICAL DECISION FUNCTIONS,
RANDOM PROCESSES

held at

Prague, from June 26 to July 2, 1982

SEPARATUM



ACADEMIA
PUBLISHING HOUSE
OF THE
CZECHOSLOVAK ACADEMY OF SCIENCES

PRAGUE 1983

AN EQUILIBRIUM THEORY FOR MULTI-PERSON
MULTI-CRITERIA STOCHASTIC DECISION PROBLEMS
WITH MULTIPLE SUBJECTIVE PROBABILITY MEASURES⁺

Tamer Başar
Urbana

Key Words: Stochastic multi-person decision problems, noncooperative equilibrium theory, subjective probability measures, team theory.

ABSTRACT

An equilibrium theory is developed for multi-person multi-criteria stochastic decision problems wherein the decision makers have different subjective probability measures on the uncertain quantities. Particular attention is devoted to existence and uniqueness of stable equilibria in such problems, when the loss functionals are (locally) quadratic and the subjective probability measures are Gaussian.

INTRODUCTION AND PROBLEM FORMULATION

Consider the class of two-person two-criteria stochastic decision problems with loss functionals $L_1(x, u_1, u_2)$ and $L_2(x, u_1, u_2)$ for DM1 (first decision maker) and DM2, respectively, where u_1, u_2 denote the decision variables [of DM1 and DM2, respectively] belonging to some prescribed Hilbert spaces U_1 and U_2 , and $x \in X$ stands for the state of Nature. Let $y_1 \in Y_1$ and $y_2 \in Y_2$ be two stochastic variables, which are correlated with x and denote the measurements available to DM1 and DM2, respectively, so that u_i will be chosen as a measurable function of y_i , $i = 1, 2$, i.e. $u_i = v_i(y_i)$, where v_i belongs to a policy space Γ_i which will be delineated in the sequel. The sets X, Y_1 and Y_2 are assumed to be structured appropriately, so that each is a well-defined Hilbert space.

So far we have adopted the standard decision-theoretic framework (see e.g. Ferguson (1967)); we depart, however, from this standard formulation in the

⁺ This work was supported in part by the Joint Services Electronics Program under Contract N00014-79-C-0424, and in part by the Office of Naval Research.

description of the underlying probability space. Let $(\Omega, \mathcal{F})$ be a measurable space to which the triple (x, y_1, y_2) belongs; then, we assume that the decision makers have different (not necessarily the same) subjective probability measures ϕ_1 and ϕ_2 [for DM1 and DM2, respectively] on this measurable space $(\Omega, \mathcal{F})$ and let the random variables (x, y_1, y_2) have finite second moments under both ϕ_1 and ϕ_2 . Furthermore, we take Γ_i to be the Banach space of all measurable mappings $v_i : Y_i \rightarrow U_i$, with the additional property that $v_i(y_i)$, viewed as a random variable, has finite second moments.

Let $z = (x, y_1, y_2)$, and introduce, for each pair $(v_1, v_2) \in \Gamma_1 \times \Gamma_2$, the quantity

$$J_i(v_1, v_2) = \int L_i(x, v_1(y_1), v_2(y_2)) \phi_i(dz) \quad (1)$$

as the expected cost function of DMi corresponding to the decision rules (v_1, v_2) and under DMi's subjective probability measure ϕ_i . [Here, we implicitly assume that L_i is integrable under ϕ_i .] We should note at this point that even in team problems (with $L_1 \equiv L_2$) the decision makers will have different expected cost functions whenever ϕ_1 and ϕ_2 do not match, since then a common probability space will not exist.

Definition 1

A pair of policies $(v_1^0, v_2^0) \in \Gamma_1 \times \Gamma_2$ constitutes an equilibrium solution to the decision problem formulated above if

$$J_1(v_1^0, v_2^0) \leq J_1(v_1, v_2^0), \quad \forall v_1 \in \Gamma_1 \quad (2a)$$

$$J_2(v_1^0, v_2^0) \leq J_2(v_1^0, v_2), \quad \forall v_2 \in \Gamma_2 \quad (2b)$$

□

Definition 2

An equilibrium solution (v_1^0, v_2^0) is a locally stable equilibrium solution if there exists an $\varepsilon > 0$ and an open neighborhood $N_\varepsilon(v_1^0, v_2^0) \equiv \Gamma_1 \times \Gamma_2$ of (v_1^0, v_2^0) so that for all $(v_1^{(0)}, v_2^{(0)}) \in N_\varepsilon$,

$$\lim_{n \rightarrow \infty} v_i^{(n)} = v_i^0, \quad \text{in } \Gamma_i, \quad i = 1, 2,$$

where

$$v_1^{(n)} = \arg \min_{\Gamma_1} J_1(v_1, v_2^{(n-1)})$$

$$v_2^{(n)} = \arg \min_{\Gamma_2} J_2(v_1^{(n-1)}, v_2), \quad n=1, 2, \dots \quad \square$$

Definition 3

A locally stable equilibrium solution (v_1^0, v_2^0) is (globally) stable if $N_\varepsilon(v_1^0, v_2^0) = \Gamma_1 \times \Gamma_2$ in Definition 2. □

Our objective in this paper is to obtain conditions on L_1 , L_2 and the probability measures ϕ_1 , ϕ_2 , under which the decision problem formulated above will admit a locally or globally stable equilibrium solution. We will, in particular, consider the class of problems in which L_1 and L_2 are quadratic in the decision variables u_1 and u_2 , and also specialize our treatment to the case of jointly Gaussian distributions. The special case of $\phi_1 \equiv \phi_2$ has earlier been treated in Başar (1975) and Başar (1978), where conditions, independent of the probabilistic structure of the problem, have been obtained for stable equilibrium solutions. The present paper discusses nontrivial extensions of these results to the case $\phi_1 \neq \phi_2$, and only outlines the method of approach and the solution because of space limitations.

QUADRATIC PROBLEMS AND GENERAL CONDITIONS

FOR EXISTENCE OF A STABLE EQUILIBRIUM

Let L_1 and L_2 be defined by

$$L_1(x, u, v) = \frac{1}{2} \langle u_1, u_1 \rangle + \frac{1}{2} \langle u_2, D_{22}^1 u_2 \rangle - \langle u_1, F_1^1 x \rangle - \langle u_2, F_2^1 x \rangle - \langle u_1, D_{12}^1 u_2 \rangle \quad (3a)$$

$$L_2(x, u, v) = \frac{1}{2} \langle u_1, D_{11}^2 u_1 \rangle + \frac{1}{2} \langle u_2, u_2 \rangle - \langle u_2, D_{21}^2 u_1 \rangle - \langle u_1, F_1^2 x \rangle - \langle u_2, F_2^2 x \rangle \quad (3b)$$

where $D_{22}^1: U_2 \rightarrow U_2$ and $D_{11}^2: U_1 \rightarrow U_1$ are strongly positive operators, and we do not differentiate between inner products defined on different Hilbert spaces. Let $E^1[\mu(z)|y_1]$ denote the expectation of a z -measurable random variable $\mu(z)$ conditioned on the random variable y_1 and under the probability measure ϕ_1 . The following two results now follow readily from the analyses of Başar (1975) and Başar (1978).

Proposition 1

A pair of policies $(v_1^0, v_2^0) \in \Gamma_1 \times \Gamma_2$ constitutes an equilibrium solution to the two DM decision problem with quadratic loss functionals (3) if and only if it satisfies the pair of equations

$$v_1^0(y_1) = D_{12}^1 E^1 [v_2^0(y_2)|y_1] + F_1^1 E^1 [x|y_1] \quad (4a)$$

$$v_2^0(y_2) = D_{21}^2 E^2 [v_1^0(y_1)|y_2] + F_2^2 E^2 [x|y_2] \quad (4b)$$

□

Proposition 2

A pair of policies $(v_1^{(0)}, v_2^{(0)}) \in \Gamma_1 \times \Gamma_2$ constitutes a stable equilibrium solution if, for all $(v_1^{(n)}, v_2^{(n)}) \in \Gamma_1 \times \Gamma_2$,

$$v_1^0(y_1) = \lim_{n \rightarrow \infty} v_1^{(n)}(y_1) \quad \text{a.e. } \phi_1$$

where

$$\begin{aligned} \gamma_i^{(n)}(y_i) = & D_{ij}^1 D_{ji}^j E^i [E^j [\gamma_i^{(n-1)}(y_i) | y_j] | y_i] \\ & + D_{ij}^1 F_j^j E^i [E^j [x | y_j] | y_i] + F_i^i E[x | y_i], \end{aligned} \quad (5)$$

$$i, j = 1, 2; j \neq i; n = 1, 2, \dots$$

Furthermore, such an equilibrium solution is necessarily unique. $\square$

Let us introduce linear operators $\mathcal{F}_i: \Gamma_i \rightarrow \Gamma_i$ $i=1, 2$, by

$$\mathcal{F}_i(\gamma) = D_{ij}^1 D_{ji}^j E^i [E^j [\gamma(y_i) | y_j] | y_i], \quad j \neq i; i, j = 1, 2. \quad (6)$$

Then, in view of Proposition 2, the quadratic decision problem will admit a unique stable equilibrium solution if, and only if, $\mathcal{F}_1$ and $\mathcal{F}_2$ are contraction mappings, i.e. there exists a constant ρ , $0 < \rho < 1$, such that

$$\|\mathcal{F}_i\| \triangleq \sup_{\gamma \in \Gamma_i} \{ \langle \gamma(y_i), \mathcal{F}_i \gamma(y_i) \rangle / \langle \gamma(y_i), \gamma(y_i) \rangle \} < \rho, \quad i=1, 2, \quad (7)$$

where $\langle \cdot \rangle$ denotes the inner product on Γ_1 or Γ_2 . Since $\|\mathcal{F}_i\| \leq \|D_{ij}^1 D_{ji}^j\| \|E^i [E^j [\cdot | y_j] | y_i]\|$ by using a well-known property of linear operators defined on Banach spaces, a set of sufficient conditions for $\mathcal{F}_i$ to be a contraction mapping is existence of a pair (ρ_1, ρ_2) , $0 < \rho_1, \rho_2 \leq 1$, $\min(\rho_1, \rho_2) < 1$, such that

$$1) \quad \|D_{ij}^1 D_{ji}^j\| \leq \rho_1 \quad (8a)$$

$$2) \quad \|E^i [E^j [\cdot | y_j] | y_i]\| \leq \rho_2, \quad (8b)$$

which is a complete separation (in terms of sufficient conditions) of the deterministic and stochastic parts of the system.

Now, if the decision problem is a team problem with a common loss functional $L = L_1 = L_2$ [which requires $D_{22}^1 = 1$, $D_{12}^1 = D_{21}^{2*}$, $F_1^1 = F_1^2$, $F_2^1 = F_2^2$], and if L is strictly convex in the pair (u_1, u_2) , (8a) is always satisfied with $\rho_1 < 1$. If, furthermore, the subjective probabilities ϕ_1 and ϕ_2 are the same, the second part of the linear operator $\mathcal{F}_i$ becomes a projection operator, thus leading to satisfaction of the second condition (8b) with $\rho_2 = 1$. Hence, for the strictly convex quadratic team problem with $\phi_1 = \phi_2$, there exists a unique stable equilibrium solution (the so-called team-optimal solution), irrespective of the underlying common probability distribution — a result which is already well-established in the literature (see Radner (1962), Başar (1978)). However, for team problems with $\phi_1 \neq \phi_2$, such a result no longer holds true, because the second part of the operator $\mathcal{F}_i$ is not necessarily a projection operator, i.e. we may not be able to find a ρ_2 , $0 < \rho_2 \leq 1$, to satisfy (8b). The general condition then is (8b), which places some restrictions on the probability measures ϕ_1 and ϕ_2 .

To investigate this question somewhat further, let us assume that $Y_1 = \mathbb{R}^{m_1}$, $Y_2 = \mathbb{R}^{m_2}$, and that θ_1 admits a probability density function (with respect to the Lebesgue measure) denoted $p^1(x, y_1, y_2)$. By an abuse of notation, let us denote the marginal and conditional densities that involve y_1 and y_2 by $p^1(y_1)$ and $p^1(y_j | y_1)$, respectively. Then for $j \neq i=1, 2$,

$$\begin{aligned} E^1[E^j[v(y_1)|y_j]|y_1]^2 &= \int p^1(y_1|y_j) \int v(y_1') p(y_1'|y_j) dy_1' dy_j \dots^2 \\ &\leq \int dy_j dy_1 p^1(y_j) p^j(y_1|y_j) \langle v(y_1), v(y_1) \rangle \\ &= \int F_1(y_1) \langle v(y_1), v(y_1) \rangle p^1(y_1) dy_1 \end{aligned} \quad (9a)$$

where

$$\begin{aligned} F_1(y_1) &\triangleq E^1\{[p^j(y_1|y_j)/p^1(y_1|y_j)]|y_1\} \\ &= \int dy_j p^1(y_j|y_1) p^j(y_1|y_j)/p^1(y_1|y_j), \end{aligned} \quad (9b)$$

and in arriving at the inequality we have made repeated use of the Cauchy-Buniakowski inequality. Now, under the condition

$$F_i(y_i) \leq 1 \quad \forall y_i \in \mathbb{R}^{m_i}, \quad i=1, 2, \quad (10)$$

(9a) can be rewritten as

$$E^1[E^j[v(y_1)|y_j]|y_1]^2 \leq \langle v(y_1), v(y_1) \rangle$$

thereby satisfying (8b) with $\rho_2=1$. Also note that it will be sufficient for (8a) and (8b) to be satisfied for only one i ($i=1$ or 2), since if, for example, $\rho_1 \leq \rho < 1$, (5) admits a well defined limit for $i=1$ as $n \rightarrow \infty$, which implies through (4b) that $\lim_n \gamma_2^{(n)}$ is also well-defined. This then leads to the following Corollary to Proposition 2.

Corollary 1

If conditions (8a) and (10) are satisfied for $i=1$ or 2 , with $0 < \rho_i < 1$, the quadratic decision problem [with Y_i taken as Euclidean spaces] admits a unique stable equilibrium solution. $\square$

Remark 1

If $\theta_1 = \theta_2$, $F_i(y_i) = 1 \quad \forall y_i \in \mathbb{R}^{m_i}$, and hence (10) is always satisfied. $\square$

JOINTLY GAUSSIAN DISTRIBUTIONS AND DERIVATION OF EXPLICIT SOLUTIONS

To explore the extent of the restrictions imposed by condition (10) on the probabilistic structure of the problem, we now further assume that the random vectors are jointly Gaussian distributed, with mean zero and covariances

$$\text{cov}(y_1, y_2) = \Sigma_y^i = \begin{pmatrix} \Sigma_{y_1}^i & \Sigma_{y_1 y_2}^i \\ \Sigma_{y_2 y_1}^i & \Sigma_{y_2}^i \end{pmatrix} > 0, \text{ under } \varphi_i. \quad (11)$$

Then, straightforward manipulations lead to

$$F_i(y_i) = \sqrt{k_i} \exp \left\{ -\frac{1}{2} y_i' K_i y_i \right\} \quad (12)$$

where

$$k_i = \det \Sigma_{y_j}^j \det \Sigma_{y_i}^i / [\det \Sigma_y^j \det \Sigma_{y_j}^i \det [A_i' D_i^{-1} A_i + \Sigma_{y_j}^{i-1}]] \quad (13a)$$

$$A_i = \Sigma_{y_i y_j}^j \Sigma_{y_j}^{j-1} \quad (13b)$$

$$D_i = \Sigma_{y_i}^j - \Sigma_{y_i y_j}^j \Sigma_{y_j}^{j-1} \Sigma_{y_j y_i}^j > 0 \quad (13c)$$

$$K_i = D_i^{-1} - \Sigma_{y_i}^{i-1} - D_i^{-1} A_i [A_i' D_i^{-1} A_i + \Sigma_{y_j}^{i-1}]^{-1} A_i' D_i^{-1} \quad (13d)$$

$j \neq i$

For (12) to be no greater than unity uniformly in $y_i \in \mathbb{R}^{m_i}$, we will have to require K_i to be a nonnegative definite matrix, which is equivalent to the matrix inequality

$$I - D_i^{1/2} \Sigma_{y_i}^{i-1} D_i^{1/2} \geq D_i^{-1/2} A_i [A_i' D_i^{-1} A_i + \Sigma_{y_j}^{i-1}]^{-1} A_i' D_i^{-1/2}. \quad (14)$$

Now, in addition to (11), let us assume that x is also a zero-mean Gaussian random vector on $\mathbb{R}^n$, so that

$$\text{cov}(x, y_1, y_2) = \text{cov}(x, y) = \Sigma^i = \begin{pmatrix} \Sigma_x^i & \Sigma_{xy}^i \\ \Sigma_{yx}^i & \Sigma_y^i \end{pmatrix} > 0 \text{ under } \varphi_i, \quad (15)$$

and that $U^i = \mathbb{R}^{r_i}$, for some integer r_i , $i=1,2$. Then, we have the following theorem for Gaussian decision systems.

Theorem 1

Let (14) and the strict inequality

$$k_i \|D_{ij}^i D_{ji}^j\| < 1 \quad (16)$$

be satisfied for at least one $i=1,2$. Then, the quadratic Gaussian decision problem formulated above admits a unique stable equilibrium solution (v_1^0, v_2^0) which is linear in (y_1, y_2) and is given by

$$y_i^0(y_i) = L_i y_i, \quad i=1,2, \quad (17)$$

where (L_1, L_2) constitutes the unique solution to the Liapunov-type matrix equations

$$L_1 - D_{12}^1 D_{21}^2 L_1 \Sigma_{y_1 y_2}^2 \Sigma_{y_2}^{2-1} \Sigma_{y_2 y_1}^1 \Sigma_{y_1}^{1-1} - D_{12}^1 F_{2xy_2}^2 \Sigma_{y_2}^{2-1} \Sigma_{y_2 y_1}^1 \Sigma_{y_1}^{1-1} - F_{1xy_1}^1 \Sigma_{y_1}^{1-1} = 0 \quad (18a)$$

$$L_2 - D_{21}^2 D_{12}^1 L_2 \Sigma_{y_2 y_1}^1 \Sigma_{y_1}^{1-1} \Sigma_{y_1 y_2}^2 \Sigma_{y_2}^{2-1} - D_{21}^2 F_{1xy_1}^1 \Sigma_{y_1}^{1-1} \Sigma_{y_1 y_2}^2 \Sigma_{y_2}^{2-1} - F_{2xy_2}^2 \Sigma_{y_2}^{2-1} = 0 \quad (18b)$$

Proof

The first part (i.e. existence and uniqueness) follows from Corollary 1 and the discussion that precedes (14), also in view of the original contraction mapping inequality (7). The second part follows by noting that if $(y_1^{(o)}, y_2^{(o)})$ are taken to be linear in (y_1, y_2) in (5), all terms of the sequence are linear and hence the limit (y_1^o, y_2^o) in Proposition 2 is linear. Denoting the coefficient gain matrices by (L_1, L_2) , we readily arrive at (18a)-(18b) through straightforward manipulations. $\square$

Conditions for existence of an equilibrium solution are of course less restrictive than those under which the statement of Theorem 1 is valid. In fact, the solution depicted in Theorem 1 will constitute an equilibrium solution whenever there exists a pair (L_1, L_2) satisfying (18a)-(18b). A sufficient condition for this (which is less restrictive than (14) and (16)), is provided in the following proposition, whose proof follows readily from the proof of the second part of Theorem 3 of Başar (1975).

Proposition 3

The quadratic Gaussian decision problem of Theorem 1 admits an equilibrium solution (not necessarily stable) given by (17)-(18), if for at least one $i=1,2$,

$$|\lambda_{\max} (D_{ij}^1 D_{ji}^j)| < 1 \quad (19a)$$

$$|\lambda_{\max} (\Sigma_{y_i y_j}^j \Sigma_{y_j}^{j-1} \Sigma_{y_j y_i}^i \Sigma_{y_i}^{i-1})| \leq 1, \quad (19b)$$

where $\lambda_{\max} (A)$ denotes the eigenvalue of A which is maximum in absolute value. $\square$

For the purpose of illustrating the various conditions of existence obtained above, we now consider a family of scalar team problems, with the decision makers having different subjective probabilities on the uncertain quantities. To be more specific, let $D_{22}^1 = D_{11}^2 = 1$, $D_{12}^1 = D_{21}^2 = d$, $|d| < 1$, $F_1^1 = F_1^2 = f_1$, $F_2^1 = F_2^2 = f_2$, and

$$\Sigma_y^1 = \begin{pmatrix} \sigma_1^1 & \sigma_{12}^1 \\ \sigma_{12}^1 & \sigma_2^1 \end{pmatrix}$$

Then, conditions (14) and (16) are satisfied if either

$$\sigma_1^1 |\sigma_1^2| \geq \sigma_2^1 |\sigma_2^2| \quad ; \quad |d| < 1/k_1 \quad (20)$$

or

$$\sigma_2^2 |\sigma_2^1| \geq \sigma_1^2 |\sigma_1^1| \quad ; \quad |d| < 1/k_2 \quad (21)$$

are satisfied, where

$$k_i = \sigma_1^1 \sigma_2^2 |[\sigma_1^j \sigma_1^j - (\sigma_{ij}^j)^2 (1 - \sigma_j^1 |\sigma_j^j|)]|; i, j=1, 2; i \neq j,$$

which are the conditions for existence of a stable equilibrium solution.

If we are interested only in existence of equilibrium solutions (not necessarily stable), the conditions (20)-(21) can be relaxed. The conditions, in this case, follow from (19a)-(19b) to be $|d| < 1$ and either $\sigma_2^2 \sigma_1^1 \leq \sigma_1^2 \sigma_2^1$ or $\sigma_1^2 \sigma_1^1 \leq \sigma_2^2 \sigma_2^1$ which are always satisfied, provided that the loss function is strictly convex in (u_1, u_2) . Hence, the conclusion is that even if the subjective probabilities are different, the scalar Gaussian team problem with strictly convex loss functional admits an equilibrium solution; this solution, however, is not necessarily stable and additional conditions (such as (20) or (21)) have to be imposed to insure stability.

CONCLUDING REMARKS

The applicability of the general approach of this paper is not restricted to the class of quadratic two-person stochastic decision problems analyzed here in considerable depth, but can readily be extended to multi-person stochastic decision problems in which the decision makers have different subjective probabilities on the uncertain quantities governing the decision process. Extensions are also possible to nonquadratic loss functionals in which case we investigate existence and uniqueness of locally stable equilibrium solutions. Because of space limitations, we have not been able to discuss such extensions in the present paper.

REFERENCES

- Başar, T. (1975): Equilibrium solutions in two-person quadratic decision problems with static information structures, IEEE Trans. Automatic Control, Vol. AC-20, No. 3, 320-328.
- (1978): Decentralized multicriteria optimization of linear stochastic systems, IEEE Trans. Automatic Control, Vol. AC-23, No. 2, 233-243.
- Ferguson, T.S. (1967): Mathematical Statistics, New York, Academic.
- Radner, R. (1962): Team-decision problems, Annals Math. Statist., Vol. 33, No. 3, 857-881.

University of Illinois
Coordinated Science Laboratory
1101 W. Springfield Avenue
Urbana, Illinois 61801 USA

[B2]

DECISION PROBLEMS WITH MULTIPLE PROBABILISTIC MODELS

FA8 - 8:30

Tamer Başar
Department of Electrical Engineering
and Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Avenue
Urbana, Illinois 61801

Abstract

This paper discusses stochastic multi-agent team problems wherein the decision makers have different probabilistic models of the underlying decision process. A suitable equilibrium solution concept is introduced for such decision problems which exhibit probabilistic multi-modeling, and the existence, uniqueness and stability properties of this equilibrium solution are studied under static information patterns. The special case of Gaussian distributions is studied in some depth, and some explicit equilibrium policies are derived for both discrete and continuous-time team problems.

1. Introduction

A team is defined as a group of agents who work together in a coordinated effort, in a possibly hostile and uncertain environment, in order to achieve a common goal. In achieving this goal, the members of the team do not necessarily acquire the same information, and hence they have to operate in a decentralized mode of decision making. The scientific approach to formulation and analysis of team problems have involved (i) a quantification of the underlying common goal in the form of a (mathematical) objective function which is sought to be optimized jointly by the agents, and (ii) a modeling of the uncertain environment and the possible measurements made by the agents on the environment in the form of a probability space together with an appropriate information structure [2,3,4,7]. The underlying stipulation here has been the existence of a probability space that is common to all the agents, so that through their priors all members of the team "see the world" in exactly the same way.

One question that readily comes into mind at this point is the robustness of such a mathematical model, and the "optimum" solutions it produces, to slight variations in the underlying assumptions. In particular, what if the agents perceive the outside world in slightly different ways? Would the solution obtained under the assumption of common prior probability measures change drastically if there are discrepancies in the decision makers (DMs)' perceptions of the probabilistic description of the outside world? In order to be able to answer these queries satisfactorily and effectively, we need a theory of equilibrium for decision problems in which the DMs have different probabilistic models of the system; such a general theory will clearly subsume the currently available results on teams which use a common probability space.

Consider a static team decision problem, formulated in the standard manner as in [7], with the only difference being in the underlying probability space. In particular, assume that the DMs assign different subjective probabilities to the uncertain events, in which case there will not exist a common probability space, thereby leading to a different expected (average) cost function for each DM. Hence, once we relax the assumption of

existence of a common probability space, the team problem is no longer a stochastic optimization problem with a single objective functional, and we inevitably have to treat it as a nonzero-sum stochastic game [5,8,12]. Furthermore, even though the original team decision problem with a common probability space will admit the same team-optimal solution(s) regardless of the mode of decision making (that is, regardless of whether the roles of the DMs are symmetric or whether there is a hierarchy and dominance in decision making), this feature ceases to hold true when there exists a discrepancy between the perceived probability measures. When there are only two members, for example, two possibilities emerge in the presence of discrepancies: the totally symmetric roles, corresponding to the Nash equilibrium solution, and the hierarchical mode, corresponding to the Stackelberg equilibrium solution.

Motivated by these considerations, we treat in this paper a general class of two-person stochastic team problems which can be viewed as static stochastic nonzero-sum games with the DMs having different subjective probability measures. Adopting the symmetric mode of decision making, we introduce the so-called "stable equilibrium solution" concept for such problems, and develop a general theory when the objective functionals are quadratic and the decision spaces are appropriate Hilbert spaces. Such a formulation includes both finite-dimensional (discrete) and continuous-time decision problems, and involves arbitrary probability measures which are, though, restricted ~~and motivated~~ by the conditions of existence and uniqueness developed in the paper. The special case of Gaussian distributions is studied in considerable depth, and some explicit solutions are obtained with appealing features.

In the next section (§2) we provide a precise problem formulation, and introduce the solution concept adopted in this paper. Section 3 develops general conditions for existence and uniqueness of a stable equilibrium solution, and elucidates the extent of the restrictions imposed on the problem by these conditions. Section 4 deals with the special class of Gaussian distributions, verifies the existence of unique linear stable equilibrium solutions and provides explicit expressions for these solutions. Furthermore, some special cases, such as the finite-dimensional and continuous-time problems are also discussed in this section. Because of space limitations we do not provide verification of the major results in this paper; detailed proofs can be obtained from the author upon request.

2. Mathematical Formulation and Some Basic Results

Probability Spaces

Let $\Omega = \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \triangleq X \times Y_1 \times Y_2$, $\mathcal{B}$ denote the Borel field of subsets of Ω , and $\mathcal{B}^k$ denote the Borel field of subsets of $\mathbb{R}^k$, $k=n, m_1, m_2$. Let $\mathcal{P}$ denote the set of all probability measures on $(\Omega, \mathcal{B})$ with finite second moments, and for each $P \in \mathcal{P}$ denote the corresponding marginal measures on $\mathbb{R}^n$, $\mathbb{R}^{m_1}$ and $\mathbb{R}^{m_2}$ by P_X , P_{Y_1} and P_{Y_2} , respectively. Furthermore, let the collection of all such probability measures be denoted by $\mathcal{F}_X$, $\mathcal{F}_{Y_1}$ and $\mathcal{F}_{Y_2}$, respectively. Then, for each $P \in \mathcal{P}$, the

This work was supported in part by the Office of Naval Research under Contract N00014-82-K-0469 and in part by the Joint Services Electronics Program under Contract N00014-79-C-0424.

vector $z = (x, y_1, y_2)'$, taking values in Ω , becomes a well-defined random vector on $(\Omega, \mathcal{B}, P)$, and likewise x is a random vector on $(\mathbb{R}^n, \mathcal{B}^n, P_x)$ and y_1 is a random vector on $(\mathbb{R}^{m_1}, \mathcal{B}^{m_1}, P_{y_1})$.

Here, x denotes the unknown state of Nature, and y_i denotes an observation of DMI (i 'th decision maker) which is correlated with x . We now choose two elements out of $\mathcal{F}$, P^1 and P^2 , which denote the subjective probabilities assigned to z by DMI and DM2, respectively. For technical reasons, we place some further restrictions on the choices of P^1 and P^2 through the marginals $P_{y_j}^i$, in particular we assume that

Condition (1). $P_{y_2}^1$ and $P_{y_1}^2$ are absolutely continuous [1] with respect to $P_{y_2}^2$ and $P_{y_1}^1$, respectively; that is, using the standard notation in probability theory,

$$P_{y_2}^1 \ll P_{y_2}^2, \quad P_{y_1}^2 \ll P_{y_1}^1. \quad (1)$$

Condition (2). The nonincreasing sequence of numbers $\{a_k^i, k > 0\}$ defined by

$$a_k^i = P_{y_i}^i(\{\xi: g^i(\xi) \geq k\}), \quad k > 0, \quad i=1,2, \quad (2a)$$

where $g^i(\xi)$ is the Radon-Nikodym derivative [1]

$$g^i(\xi) = dP_{y_i}^i / dP_{y_i}^1, \quad (2b)$$

has the property that

$$a_k^i = 0 \quad \text{for all } k > K^i, \text{ for some } K^i < \infty. \quad (2c)$$

Decision and Policy Spaces

The decision variable of DMI will be denoted by u_i which belongs to a real Hilbert space U_i with inner product $(\cdot, \cdot)_i$. Permissible policies (decision rules) for DMI are measurable mappings

$$y_i: \mathbb{R}^{m_i} \rightarrow U_i \quad (3)$$

satisfying the square-integrability conditions

$$\|y_i\|_{L^2}^2 = \int_{\mathbb{R}^{m_i}} \|y_i(\xi)\|_i^2 dP_{y_i}^i(\xi) < \infty, \quad j=1,2, \quad (4)$$

where $\|\cdot\|_i$ is the natural norm derived from $(\cdot, \cdot)_i$. Note that the condition (4) requires that the permissible policies of each DM have bounded second-order moments under both probability measures.

Let Γ_i denote the space of all permissible policies (3) of DMI satisfying (4), and which is equipped with the metric

$$d_i(\gamma, \beta) = \left(\int_{\mathbb{R}^{m_i}} \|\gamma(\xi) - \beta(\xi)\|_i^2 dP_{y_i}^i(\xi) \right)^{1/2}, \quad \gamma, \beta \in \Gamma_i. \quad (5)$$

Then, we have the following basic result on the topological structure of Γ_i :

Lemma 1. If the underlying probability measures satisfy the conditions (1) and (2), Γ_i equipped with the metric (5) is a Banach space. $\square$

Furthermore, introduce the inner-product $\langle \cdot, \cdot \rangle_i$ on elements of Γ_i by

$$\langle \gamma, \beta \rangle_i = \int_{\mathbb{R}^{m_i}} (\gamma(\xi), \beta(\xi))_i dP_{y_i}^i(\xi) \quad (6)$$

which makes Γ_i a Hilbert space. We are now in a position to introduce the cost functionals for the two DMs.

Cost Functionals

Let $D_{ij}: U_j \rightarrow U_i$ and $F_j: X \rightarrow U_i$ ($i \neq j$, $i, j=1,2$) be bounded linear operators with $D_{12} = D_{21}^*$. Furthermore, let $E^i[\cdot | z]$ denote the mathematical expectation of a $\mathcal{Z}$ -measurable random variable $\cdot | z$ taking values in U_i

conditioned on the random variable y_i , and under the probability measure P^i , i.e.

$$E^i[\cdot | z] = \int_{\mathbb{R}^{m_i}} \cdot P_{y_i}^i(dz | y_i) \quad (7)$$

where the second term of the integrand is the conditional probability measure derived from P^i . Then, for each pair $(\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2$, we have a quadratic cost functional for each DM, defined for DMI by

$$\begin{aligned} J_1(\gamma_1, \gamma_2) &= \frac{1}{2} \langle \gamma_1, \gamma_1 \rangle_1 + \frac{1}{2} \int_{\mathbb{R}^{m_1}} (\gamma_1(\xi), \gamma_2(\xi))_1 dP_{y_1}^1(\xi) \\ &\quad - \langle \gamma_1, E^1[F_1 x | y_1] \rangle_1 - \int_{\mathbb{R}^{m_1}} (\gamma_1(\xi), F_2 x)_1 dP_{y_1}^1(\xi) \\ &\quad - \langle \gamma_1, E^1[D_{12}^* \gamma_2 | y_1] \rangle_1, \end{aligned} \quad (8a)$$

which is derived from a common strictly convex quadratic team cost functional. The strict convexity requirement is met if we choose D_{12} to satisfy

$$\|D_{12} D_{12}^*\|_1 \leq \bar{\rho} < 1. \quad (8b)$$

Equilibrium Solution

Since the expected cost functionals (8a), together with the policy spaces, provide a normal (strategic) form description, regardless of the presence of multiple probability measures, the standard definition of non-cooperative (Nash) equilibrium [5] (which we adopt as our solution concept) remains intact, as given below.

Definition 1.

A pair of policies $(\gamma_1^0, \gamma_2^0) \in \Gamma_1 \times \Gamma_2$ constitutes a Nash equilibrium solution if

$$J_1(\gamma_1^0, \gamma_2^0) \leq J_1(\gamma_1, \gamma_2^0), \quad \forall \gamma_1 \in \Gamma_1 \quad (9a)$$

$$J_2(\gamma_1^0, \gamma_2^0) \leq J_2(\gamma_1^0, \gamma_2), \quad \forall \gamma_2 \in \Gamma_2. \quad (9b)$$

Definition 2.

A Nash equilibrium solution (γ_1^0, γ_2^0) is locally stable if there exists an $\epsilon > 0$ and an open neighborhood $N_\epsilon(\gamma_1^0, \gamma_2^0) \subset \Gamma_1 \times \Gamma_2$ of (γ_1^0, γ_2^0) such that for all $(\gamma_1^{(0)}, \gamma_2^{(0)}) \in N_\epsilon$,

$$\lim_{k \rightarrow \infty} \gamma_i^{(k)} = \gamma_i^0, \quad \text{in } \Gamma_i, \quad i=1,2, \quad (10)$$

where

$$\gamma_1^{(k)} = \arg \min_{\Gamma_1} J_1(\gamma_1, \gamma_2^{(k-1)}) \quad (11a)$$

$$\gamma_2^{(k)} = \arg \min_{\Gamma_2} J_2(\gamma_1^{(k-1)}, \gamma_2) \quad k=1,2,\dots \quad (11b)$$

Definition 3.

A locally stable Nash equilibrium solution (γ_1^0, γ_2^0) is (globally) stable if $N_\epsilon(\gamma_1^0, \gamma_2^0) = \Gamma_1 \times \Gamma_2$ in Definition 2.

3. General Conditions for a Stable Equilibrium Solution

We now obtain some general conditions for existence of stable equilibrium solutions. Because of space limitations we simply give the main results without verification.

Proposition 1.

A pair of policies $(\gamma_1^0, \gamma_2^0) \in \Gamma_1 \times \Gamma_2$ constitutes an equilibrium solution to the decision problem of §2.1, and only if, it satisfies the pair of equations

$$\gamma_1^0 = D_{12} E^1[\gamma_2^0 | \gamma_1^0] + F_1[x | \gamma_1^0] \quad (12)$$

$$y_2^0 = D_{21} E^2[y_1^0(y_1)|y_2] + F_2 E^2[x|y_2] \quad (13)$$

Proposition 2.

A pair of policies $(\gamma_1^0, \gamma_2^0) \in \Gamma_1 \times \Gamma_2$ constitutes a stable equilibrium solution if, and only if, for all $(\gamma_1^{(0)}, \gamma_2^{(0)}) \in \Gamma_1 \times \Gamma_2$,

$$y_i^0(y_i) = \lim_{k \rightarrow \infty} y_i^{(k)}(y_i) \quad \text{in } \Gamma_i, \quad (14)$$

where $y_i^{(k)}$ is given recursively by

$$y_i^{(k)}(y_i) = D_{ij} D_{ij}^* E^i[E^j[y_i^{(k-1)}(y_i)|y_j]|y_i] + D_{ij} F_j E^i[E^j[x|y_j]|y_i] + F_i E[x|y_i], \quad j=1,2; j \neq i; k=1,2,\dots \quad (15)$$

Furthermore, such a stable equilibrium solution is necessarily unique.

Let us introduce linear operators $S_i: \Gamma_i \rightarrow \Gamma_i$, $i=1,2$, by

$$S_i(\gamma) = D_{ij} D_{ij}^* E^i[E^j[\gamma(y_j)|y_i]|y_i], \quad j \neq i; i,j=1,2. \quad (16)$$

Note that S_i indeed maps Γ_i into Γ_i , because the conditional expectation $E^j[D_{ij} \gamma(y_j)|y_i]$ maps Γ_j into Γ_j ($j \neq i$) when the probability measures satisfy conditions (1) and (2), and every element of Γ_i is square-integrable under P_i^1 and P_i^2 .

Furthermore, let us introduce the notation $\langle\langle S \rangle\rangle_i$ to denote the norm of a linear bounded operator $S: \Gamma_i \rightarrow \Gamma_i$; which is defined by

$$\langle\langle S \rangle\rangle_i = \sup_{\gamma \in \Gamma_i} [\langle S\gamma, S\gamma \rangle_i / \langle \gamma, \gamma \rangle_i]^{1/2}, \quad (17a)$$

and $r_i(S)$ to denote the spectral radius of S , which is defined by [13]

$$r_i(S) = \lim_{k \rightarrow \infty} \langle\langle S^k \rangle\rangle_i^{1/k} \quad (17b)$$

where S^k denotes the k 'th power of S . Finally, let us introduce the linear operator

$$\bar{P}_{i|i} \triangleq E^i[E^j[\cdot|y_j]|y_i], \quad (18)$$

which maps Γ_i into itself. Then, the following proposition whose proof depends on a contraction mapping argument, provides a set of necessary and sufficient conditions for existence of the unique equilibrium solution alluded to in Proposition 2.

Proposition 3.

When the probability measures satisfy conditions (1) and (2), the decision problem of section 2 admits a unique stable equilibrium solution if, and only if, there exists α^i , $0 < \alpha^i < 1$, such that

$$r_i(S_i) = r_i(D_{ij} D_{ij}^* \bar{P}_{i|i}) \leq \alpha^i, \quad (19)$$

for at least one $i=1,2$.

The next proposition provides a set of stronger but more versatile conditions under which a unique stable equilibrium solution exists.

Proposition 4.

A set of sufficient conditions for (19) to hold true for at least one i , and thereby for a unique stable equilibrium solution to exist, is the existence of a pair of positive scalars (α_1, α_2) , with

$$\alpha_1 \alpha_2 < 1, \quad (20a)$$

such that

$$r_i(\bar{P}_{i|i}) \leq \alpha_i, \quad (20b)$$

for at least one $i=1,2$ where $\bar{P}$ is introduced by (8b). Furthermore, a sufficient condition for (20b) is

$$\langle\langle \bar{P}_{i|i} \rangle\rangle_i \leq \alpha_i \quad (20c)$$

This result provides a partial separation (in terms of sufficient conditions) of the deterministic and stochastic parts of the system. Now, if the subjective probability measures assigned to the pair (y_1, y_2) by the two DM's are equivalent, $\bar{P}_{i|i}$ becomes a projection operator, thus leading to satisfaction of (20b) with $\alpha_1 = \alpha_2 = 1$, and thereby satisfaction of (20a) since $\alpha_1 \alpha_2 = 1$. Hence, as a corollary to Proposition 4, we obtain the following result which is known in different contexts [7,8,9].

Corollary 1.

For the strictly convex quadratic team problem with equivalent subjective probability measures assigned by the two DM's to (y_1, y_2) , there exists a unique stable equilibrium solution (the so-called team-optimal solution), irrespective of the underlying common probability measure.

For team problems with $P^1 \neq P^2$, a result along the lines of Corollary 1 does not in general hold, because the operator $\bar{P}_{i|i}$ is not necessarily a projection operator, i.e. we may not be able to find $\alpha_i, 0 < \alpha_i \leq 1$, to satisfy (20c) [or (20b)]. Then, the general condition is (19) [or the stronger one, (20a)] which places some restrictions on the parameters of the cost functional, as well as the probability measures P^1 and P^2 . To delineate the extent of these restrictions, we now study condition (20c) somewhat further and obtain the following sufficient condition.

Corollary 2.

For a given α_i , inequality (20c) is satisfied if

$$P_{y_i}^1 \left(\int_{\Gamma_j} g^i(\xi) \int_{\Gamma_j} g^j(\eta) P_{y_j|y_i}^j(d\eta|\xi) d\xi > \alpha_i^2 \right) = 0 \quad i=1,2, j \neq i, \quad (21)$$

where $g^i(\xi)$ and $g^j(\eta)$ are the Radon-Nikodym derivatives, as defined by (2b).

4. Jointly Gaussian Distributions and Derivation of Explicit Solutions

When the subjective probability measures of the two DMs are equivalent, one special class of problems that admit closed-form solutions is that with Gaussian distributions, and equilibrium solutions in these cases have been shown to be affine functions of the observations, as documented in the literature for quadratic team problems defined on Euclidean spaces [7], quadratic nonzero-sum stochastic games on Euclidean spaces [8], and quadratic continuous-time stochastic team problems [9]. In this section, we study possible extensions of these results to the case when discrepancies exist between the subjective Gaussian distributions, as reflected in the covariances of the random vectors (y_1, y_2) .

Hence, let us now assume that (y_1, y_2) are Gaussian random vectors under both P^1 and P^2 , with

$$\text{mean}(y_1, y_2) = 0, \quad (22)$$

$$\text{covariance}(y_1, y_2) = \Sigma_y^i = \begin{pmatrix} \Sigma_{y_1}^i & \Sigma_{y_1 y_2}^i \\ \Sigma_{y_2 y_1}^i & \Sigma_{y_2}^i \end{pmatrix} > 0 \quad (23)$$

These probability distributions clearly satisfy the absolute continuity condition (1) of section 2, and

also satisfy the uniform boundedness condition (2) whenever

$$W_i = \tau_{y_i}^{j-1} - \tau_{y_i}^{i-1} \geq 0, \quad i, j=1, 2; \quad j \neq i. \quad (24)$$

Using standard properties of Gaussian distributions, we obtain

$$g^i(y_i) E^j [g^j(y_j) | y_i] = q^i \exp \left\{ -\frac{1}{2} y_i^T D_i y_i \right\} \quad (25)$$

where

$$D_i = M_{ii}^j - M_{ij}^j B_j^{-1} M_{ji}^j - \tau_{y_i}^{i-1}, \quad j \neq i \quad (26a)$$

$$\begin{pmatrix} M_{11}^i & M_{12}^i \\ M_{21}^i & M_{22}^i \end{pmatrix} \cong \begin{pmatrix} \tau_{y_1}^i & \tau_{y_1 y_2}^i \\ \tau_{y_1 y_2}^i & \tau_{y_2}^i \end{pmatrix} \quad (26b)$$

$$B_j = M_{jj}^j + W_j \quad (26c)$$

$$q^i = [w_1 w_2 \det(\tau_{y_i}^j) \det(B_j^{-1}) \det(\tau_{y_i}^{i-1})]^{1/2} \quad (26d)$$

$$w_i = \det \tau_{y_i}^i / \det \tau_{y_i}^j. \quad (26e)$$

For the condition of Corollary 1 (i.e. (21)) to be satisfied for a given finite ρ_i , it is necessary that expression (25) be uniformly bounded in y_i , which holds true if, and only if,

$$D_i \geq 0. \quad (27)$$

Under this restriction, condition (20d) becomes

$$\sqrt{q^i} \leq \rho_i,$$

and referring back to Proposition 4, we obtain a sufficient condition for existence of a unique stable equilibrium solution to be

$$\|D_{ij} D_{ij}^* \|_i^2 q^i < 1, \quad (28)$$

for at least one $i=1, 2$.

In order to complete the solution of the Gaussian decision problem, let us further assume that x is also Gaussian distributed, with

mean $\langle x \rangle = 0$

$$\text{covariance } (x, y_1, y_2) = \text{cov}(x, y) = \tau^i = \begin{pmatrix} \tau_x^i & \tau_{xy}^i \\ \tau_{yx}^i & \tau_y^i \end{pmatrix} > 0.$$

Then, the following theorem summarizes the complete solution.

Theorem 1.

Let

$$(i) \quad W_i \geq 0, \quad i=1, 2, \quad (29a)$$

and the following conditions hold for at least one $i=1, 2$:

$$(ii) \quad D_i \geq 0, \quad (29b)$$

$$(iii) \quad \|D_{ij} D_{ij}^* \|_i^2 < 1/q^i, \quad i, j=1, 2; \quad j \neq i. \quad (29c)$$

Then, the quadratic Gaussian decision problem formulated in this section admits a unique stable equilibrium solution $(\bar{q}, \bar{q})$, where $\bar{q}_i^0 = \bar{q}_i^0(y_i)$ are linear in y_i , and are given by

$$\bar{q}_i^0(y_i) = L_i y_i, \quad i=1, 2. \quad (30)$$

Here, $L_i: \mathbb{R}^{m_i} \rightarrow \mathbb{U}^i$ are bounded linear operators, constituting the unique solution to the linear operator equations

$$\begin{aligned} L_i y_i - D_{ij} D_{ij}^* L_j y_j \tau_{y_j}^{j-1} \tau_{y_j y_i}^{i-1} \tau_{y_i}^{i-1} \\ - D_{ij} F_{ij}^j \tau_{x y_j}^j \tau_{y_j}^{j-1} \tau_{y_j y_i}^{i-1} \tau_{y_i}^{i-1} \\ - F_i \tau_{x y_i}^i \tau_{y_i}^{i-1} y_i = 0, \quad \forall y_i \in \mathbb{R}^{m_i}, \quad i=1, 2. \end{aligned} \quad (31)$$

In the statement of Theorem 1, the condition (29a) places some restrictions on the second moments of the underlying distributions (in case a discrepancy exists), which may however be relaxed if we are willing to consider equilibrium policies in a more restricted space. More specifically, satisfaction of (29a) ensures that regardless of what initial set of policies the DMs start the infinite recursion (14) with, every element of this series is well-defined, and under (29b)-(29c) it will converge to a unique limit which is affine; in other words, even if the DMs start with nonlinear policies, the end result will be an affine equilibrium solution. The condition (29a) is restrictive, because we require (without imposing any constraints on the policy spaces) the series generated by (14) to be well-defined even with nonlinear starting conditions. However, if we restrict ourselves to affine policies from the outset, under Gaussian distributions elements of the series (14) will always be well-defined (without requiring (29a)) and will converge to the equilibrium solution provided that (29b)-(29c) holds for at least one $i=1, 2$. This line of reasoning then leads to the following result:

Proposition 5.

Let $\bar{\tau}_i^a$ be the class of all linear policies in the form (30), with $L_i: \mathbb{R}^{m_i} \rightarrow \mathbb{U}^i$ a bounded linear operator, $i=1, 2$. On $\bar{\tau}_i^a \times \bar{\tau}_i^a$, the statement of Theorem 1 is valid even if (29a) does not hold true.

Finite-dimensional Decision Spaces

We have so far obtained conditions for existence and uniqueness of stable equilibrium solutions, so that the recursive relations (11) converge for all possible starting points, either in $\bar{\tau}_1 \times \bar{\tau}_2$ or in $\bar{\tau}_i^a \times \bar{\tau}_i^a$. If we are interested only in existence of equilibrium solutions (cf. Definition 1), however, the corresponding conditions will clearly be less restrictive. For a further elaboration on this point, consider, for example, the class of decision problems wherein all linear operators are matrices. A stable equilibrium will exist, in this case, under the conditions (replacing (29c))

$$\lambda_{\max} (D_{ij} D_{ij}^*)^2 \cdot [\det(B_j) \det(\tau_{y_j}^i) / [\det(\tau_{y_j}^i) \det(\tau_{y_j}^i)]]^{1/2} < 1, \quad i, j=1, 2; \quad j \neq i, \quad (32)$$

where $\lambda_{\max}(A)$ denotes the eigenvalue of the square matrix A which is maximum in absolute value. Furthermore, this unique stable equilibrium solution will be given by (30), where L_i is now a matrix (of appropriate dimensions) satisfying (31) with the multiplying y_i 's left out. Such a solution is definitely also an equilibrium solution (cf. Definition 1), and as such it exists whenever (31) is solvable. A sufficient condition for solvability of (the finite-dimensional version of) (31), which is less restrictive than (32), is given below as Proposition 6.

Proposition 6.

When the decision spaces are finite dimensional, the quadratic Gaussian decision problem admits an equilibrium solution (not necessarily stable) given by (30)-(31), if, for at least one $i=1, 2$, the following inequality holds

$$\max_{i,j} (D_{ij} D_{ij}^*) \leq \max_{i,j} \left(\frac{z_i^1}{y_i y_j} \frac{z_j^{-1}}{y_j y_i} \frac{z_i^{-1}}{y_i y_j} \right) < 1. \quad (33)$$

For the purpose of illustrating the various conditions of existence (and uniqueness) obtained above and in the preceding sections, let us now consider a family of scalar quadratic Gaussian team problems, with the DM's having different subjective probabilities on the uncertain quantities. To be more specific, let $D_{11} = \alpha$, $|\alpha| < 1$, $F_1 = f_1$, $F_2 = f_2$, $n=m_1=m_2=1$, and

$$z_y^1 = \begin{pmatrix} a & e \\ e & nb \end{pmatrix}; \quad z_y^2 = \begin{pmatrix} ua & c \\ c & b \end{pmatrix}, \quad nab > e^2, \quad uab > c^2. \quad (34)$$

Firstly, condition (1) of section 2 on absolute continuity of various measures is clearly satisfied, because all probability measures are Gaussian. Secondly, condition (2) of section 2 is satisfied if, and only if, both

$$0 < u \leq 1, \quad 0 < n \leq 1. \quad (35)$$

This is condition (29a) of Theorem 1. For condition (29b), we evaluate D_1 :

$$D_1 = (uab - e^2)(1-u) [ua^2 - e^2 - (1-u)ab] / \{a[u^2a^2 - (1-u)(uab - e^2)]\} \geq 0 \quad (36a)$$

$$D_2 = (nab - e^2)(1-n) [nb^2 - e^2 - (1-n)ab] / \{b[n^2b^2 - (1-n)(nab - e^2)]\} \geq 0, \quad (36b)$$

and require either (36a) or (36b) to be satisfied. Finally, condition (29c), whose counterpart in this context is (32), dictates either

$$ua^2 |z| < n[u^2a^2 - (1-u)(uab - e^2)] \quad (37a)$$

or

$$nb^2 |z| < u[n^2b^2 - (1-n)(nab - e^2)] \quad (37b)$$

provided that the terms on the right-hand-side are positive (if not, then the inequalities will accordingly change direction).

The set of values for a, b, e, u, n that satisfy (34)-(37) is clearly not empty. To gain some further insight into these conditions, let us consider the class of team decision problems in which the discrepancies between the DM's perceptions of the variances of different Gaussian random variables is relatively small, that is there exist sufficiently small $\epsilon_1 > 0$ and $\epsilon_2 > 0$ such that $u = 1 - \epsilon_1$, $n = 1 - \epsilon_2$, and furthermore $e = c$, and $|c|$ is considerably smaller than both a and b . Note that, when $\epsilon_1 = \epsilon_2 = 0$, conditions (35)-(37) are all satisfied (note that $|z| < 1$ because of strict convexity of the objective functional), regardless of the relative magnitudes of e and c . Hence, when the discrepancy is only in the perceptions of the correlation between y_1 and y_2 , the scalar quadratic Gaussian team problem always admits a stable equilibrium solution. Now, for nonzero, but positive, and sufficiently small ϵ_1 , the dominating term in (36a) will be

$$D_1 \sim \epsilon_1 (uab - e^2) (ua^2 - e^2) / u^2 a^3$$

which is positive, in view of (34) and the initial hypothesis that $|a/c| \gg 1$. Likewise, D_2 is positive, whenever $0 < \epsilon_2 < 1$, and $|b/e| \gg 1$. Furthermore, given a $\bar{e}$, $0 < \bar{e} < 1$, we can always find ϵ_1 and ϵ_2 , both in $(0, 1)$, so that both (37a) and (37b) are satisfied whenever $|z| < \bar{e}$. Hence, the conclusion is that when the deviations of the perceptions of the DMs from the common Gaussian probability measures are incremental (and satisfying (35)), the linear equilibrium solution of the Gaussian scalar team problem retains its stability property (but, of course, at a different (possibly close, in norm) equilibrium point).

Finally, if our interest lies only in the existence of an equilibrium solution (not necessarily stable), the condition that replaces (29a)-(29c) is (33) which, in our case, is independent of i and reads:

$$|a|^2 |ce/ab| < 1. \quad (38)$$

This condition is clearly much less restrictive than (35)-(37), and is satisfied whenever $a \gg |c|$ and $b \gg |e|$ which are reasonable restrictions in (34).

Infinite-dimensional Decision Spaces

As another illustration of Theorem 1, for infinite-dimensional decision spaces, we consider in the sequel a class of stochastic Gaussian team problems defined in continuous time. More specifically, let $L_1 = L_2 = L_0(0, T)$, the Hilbert space of all scalar-valued Lebesgue-integrable functions on the bounded interval $[0, T]$, endowed with the standard inner product $\int_0^T u(t)v(t)dt$, for $u, v \in L_2$.

Furthermore, let $Y_1 = Y_2 = \mathbb{R}$, and the Gaussian statistics have zero mean, and variances as given in (34). Let $D_{12} = D_{21}^*$ be the Fredholm operator

$$D_{12} u = \int_0^T K(t, s) u(s) ds \quad (39)$$

where $K(t, s)$ is a continuous kernel on $0 \leq t, s \leq T$, and finally let $F_i = f_i(t)$, $i=1, 2$, which are continuous functions on $[0, T]$.

Now, conditions (29a) and (29b) depend only on the probabilistic structure, and are therefore again given by (35) and (36), respectively. For (29c), however, we have to obtain the counterpart of (37), by simply replacing $|z|$ with the norm of the operators $D_{12} D_{12}^*$ and $D_{12}^* D_{12}$, respectively. Since

$$D_{12}^* u = \int_0^T K(s, t) u(s) ds,$$

the self-adjoint operator $D_{12} D_{12}^*$ is given by

$$D_{12} D_{12}^* u = \int_0^T \int_0^T K(t, \tau) K(s, \tau) u(s) ds = \int_0^T K(t, s) u(s) ds,$$

where

$$K(t, s) = \int_0^T K(t, \tau) K(s, \tau) d\tau. \quad (40a)$$

Let

$$\lambda = \left(\int_0^T \int_0^T K(t, s)^2 dt ds \right)^{1/2}. \quad (40b)$$

Then,

$$\begin{aligned} \|D_{12} D_{12}^* u\|^2 &= \int_0^T \left| \int_0^T K(t, s) u(s) ds \right|^2 dt \\ &\leq \int_0^T \int_0^T K(t, s)^2 ds \left| \int_0^T u(s)^2 ds \right| dt \\ &= \lambda^2 \|u\|^2 \end{aligned}$$

where the second step follows from the Cauchy-Buniakowski inequality. Hence,

$$\|D_{12} D_{12}^*\| \leq \lambda,$$

and because of symmetry $D_{12}^* D_{12}$ is also bounded in norm by the same quantity. This then leads to the following counterpart of (37): A sufficient condition for satisfaction of (29c) is either

$$ua^2 \lambda^2 < n[u^2a^2 - (1-u)(uab - e^2)] \quad (41a)$$

or

$$nb^2 \lambda^2 < u[n^2b^2 - (1-n)(nab - e^2)] \quad (41b)$$

provided that the terms on the right-hand-side are positive, where λ is defined by (40a)-(40b).

Hence, under (35) and either (36a) and (41a) or (36b) and (41b), the continuous-time static decision problem formulated above admits a unique stable equilibrium solution, and this solution is given by (from Theorem 1):

$$k_i(t, v_i) = k_i(t) v_i, \quad i=1,2, \quad (42)$$

where $k_i(t)$ are continuous functions on $[0, T]$, satisfying

$$k_1(t) = \left(\frac{c_1}{\lambda} \right) \int_0^T k_1(s) ds - \left(\frac{c_1}{\lambda} \right) \int_0^T K(t,s) f_1(s) ds - \left(\frac{1}{\lambda} \right) f_1(t) = 0 \quad (43a)$$

$$k_2(t) = \left(\frac{c_2}{\lambda} \right) \int_0^T k_2(s) ds - \left(\frac{1}{\lambda} \right) \int_0^T K(t,s) f_2(s) ds - \left(\frac{1}{\lambda} \right) f_2(t) = 0. \quad (43b)$$

Note that $k_i(t)$ above stands for operator L_i in (31), and we have already shown that a unique solution to both (43a) and (43b) exist in $L_2[0, T]$, under (35) and either (36a) and (41a) or (36b) and (41b), and this solution is also continuous.

Finally, if our interest lies only in the existence of a unique linear equilibrium solution in the class of linear policies (not necessarily stable), the required condition is unique solvability of the integral equations (43a)-(43b), for which a sufficient condition is [6]

$$(\partial \partial / \partial \lambda) \lambda < 1 \quad (44)$$

where λ is defined by (40b).

5. Concluding Remarks

In the preceding sections, we have developed an equilibrium theory for two-person quadratic team decision problems with static information patterns, wherein the decision makers (DMs) do not necessarily have the same perception of the underlying probability space; that is, our formulation allows for discrepancies in the way different DMs perceive the probability space. As indicated earlier, when such discrepancies exist team problems have to be analyzed in the framework of nonzero-sum games, and in such a framework the Nash solution concept is the most suitable equilibrium concept if the DMs occupy symmetric (non-hierarchical) positions in the decision process. When the equilibrium policies satisfy the further requirement of stability, this solution becomes very appealing because, in order to arrive at equilibrium (as a consequence of an infinite number of response iterations), each DM does not have to know the subjective probability measures perceived by the other DM, but has to know only the policy adopted by the other DM at the most recent step of the iteration. Under the stipulation that each DM chooses his policy at each stage by responding optimally to the other DM's policy, we have derived in this paper a set of conditions that insure existence of unique limits to such iterations, and thereby existence of a unique stable equilibrium solution.

The analyses of section 4 have shown that when the underlying probability distributions belong to a Gaussian class, conditions of existence and uniqueness, as well as the stable equilibrium solution itself, can be obtained explicitly, with the latter being affine in the available static information. For two special cases, namely when the decision spaces are finite dimensional or when the decision problem is defined in continuous-time with a specific cost structure, we have obtained analytic expressions for the gain operators, and have also further delineated the existence conditions. The

general Hilbert-space framework adopted in this paper and the general solution presented in section 4 (Theorem 1) applies to other models also, such as the one similar to the continuous-time team problem treated in [9] but with the DMs having different probability models. It is expected that some explicit results (closed-form solutions) can also be obtained in this case, but this point has not been pursued in this paper and is left for future research.

One source of motivation for the research reported in this paper has been (as discussed in section 1) the desire to investigate the sensitivity and robustness of team-optimal solutions (in stochastic teams) to independent variations in the perceptions of the DMs of the underlying probability space (and, in particular, the probability measure). The analysis of this paper indeed provides a framework for such a study when the roles of the DMs are symmetric, since an equilibrium theory (of the "Nash" type) has been established (in terms of existence, uniqueness and derivation of stable solutions) within an " ϵ -neighborhood" of the team-optimal solution. Some further work is needed in order to determine the "satisfiability" of the several existence conditions obtained in the paper, when the region of interest is an ϵ -neighborhood of a common probability space, and to further extend the analysis to an investigation of sensitivity and robustness properties of team solutions (obtained under the stipulation of existence of a common underlying probability space) in this ϵ -neighborhood.

Extensions of the analyses of this paper to intrinsic nonzero-sum games with symmetric and asymmetric modes of decision making under different subjective probabilities can be found in [10] and [11], respectively.

References

- [1] M. Loève, *Probability Theory*, New York: Van Nostrand Reinhold, 1963.
- [2] J. Marschak and R. Radner, *Economic Theory of Teams*, New Haven, CT: Yale University Press, 1972.
- [3] Y. C. Ho, "Team decision theory and information structures," *Proc. IEEE*, vol. 68, no. 6, pp. 644-654, June 1980.
- [4] Y. C. Ho and K. C. Chu, "Team decision theory and information structures in optimal control problems - Part I," *IEEE Transactions on Automatic Control*, vol. AC-17, no. 1, pp. 15-22, February 1972.
- [5] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, London: Academic Press, 1982.
- [6] L. A. Ljusternik and V. J. Sobolev, *Elements of Functional Analysis*, New York: Ungar, 1961.
- [7] R. Radner, "Team-decision problems," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 857-881, 1962.
- [8] T. Basar, "Decentralized multicriteria optimization of linear systems," *IEEE Transactions on Automatic Control*, vol. AC-23, no. 2, pp. 233-243, April 1978.
- [9] A. Bagchi and T. Basar, "Team decision theory for linear continuous-time systems," *IEEE Transactions on Automatic Control*, vol. AC-25, no. 6, pp. 1154-1161, December 1980.
- [10] T. Basar, "An equilibrium theory for multi-person decision making with multiple probabilistic models: Part I. Symmetric mode of decision making," submitted for publication (November 1982).
- [11] T. Basar, "An equilibrium theory for multi-person decision making with multiple probabilistic models: Part II. Asymmetric mode of decision making," submitted for publication (February 1983).
- [12] G. Papavasiliopoulos, "Solution of some stochastic quadratic Nash and leader-follower games," *SIAM J. Optimiz. Contr.*, vol. 19, pp. 651-666, September 1981.
- [13] L. W. Kantorovich and G. P. Akilov, *Functional Analysis*, 2nd edition, Moscow: Nauka Publishing Co., 1977; English translation of the 1st edition: *Functional Analysis in Normed Spaces*, New York: MacMillan, 1964.

AN EQUILIBRIUM THEORY FOR MULTI-PERSON DECISION MAKING WITH MULTIPLE PROBABILISTIC MODELS

Part II. Asymmetric Mode of Decision Making

Ismar Basar
Department of Electrical Engineering
and Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Avenue
Urbana, Illinois 61801

Abstract

This paper develops an equilibrium theory for two-person two-criteria stochastic decision problems with static information patterns and an asymmetric mode of decision making, wherein the decision makers have different probabilistic models of the underlying process. The objective functions are quadratic and the decision spaces are general inner-product spaces. Firstly, a set of sufficient conditions is obtained for existence and uniqueness of Stackelberg equilibria, and a uniformly convergent iterative scheme is developed, whereby the equilibrium policy of the leader can be obtained by evaluating a number of conditional expectations. When the probability measures are Gaussian, the equilibrium solution is shown to be generically nonlinear, with the linear structure prevailing only in some special cases which are delineated in the paper.

1. Introduction

In [1], we have presented a theory of equilibrium for team decision making when the decision makers have different perceptions of the underlying probability measures and when the mode of decision making is symmetric. These results have also been extended to stochastic nonzero-sum games, again under the symmetric mode of decision making, in [2]. It has been shown in these two references that a stable equilibrium solution exists under a reasonable set of conditions which place some restrictions on the probabilistic and nonprobabilistic parts of the description of the decision problem. Furthermore, the solution was obtainable by successive approximation, which was shown to lead to affine equilibrium solutions when the underlying statistics were Gaussian.

As it has been pointed out in [1], even though the mode of decision making is irrelevant in stochastic team problems with a common probability space for both DMs, it becomes an important factor in the derivation of equilibrium solution when there is a discrepancy in the DMs' perceptions of the probability measures. Hence, even in team problems, derivation of equilibrium solution requires separate (and possibly different) analyses in the two cases corresponding to symmetric and asymmetric modes of decision making. Having provided a complete analysis of the former in [1] and [2], here we direct attention to an investigation of the nature, existence, uniqueness and derivation of equilibrium solution under the latter mode. Towards this end, we start with a more general class of problems, as in [2], which subsumes stochastic team problems as a special class, and we formulate the hierarchical decision problem in general Hilbert spaces and under a general probabilistic description. As for a solution concept we adopt that of Stackelberg equilibrium which is the natural counterpart of the (Nash)

equilibrium solution of [1] in the present context.

A problem formulation and precise delineation of a set of conditions which lead to a meaningful description are provided in Section 2, where we also rely on the more detailed exposition of [1] and [2]. A derivation of unique Stackelberg equilibrium solution, proof of existence and uniqueness, and some elucidation of the required conditions [cf. Condition 5] occupies us through Section 3. Perhaps a most surprising by-product of this analysis, as contrasted with [1], is in the characterization of the equilibrium solution for the special case of jointly Gaussian distributions: The solution is generically nonlinear, and contains summation of terms which involve products of linear functions of measurements with exponential terms (whose exponents are quadratic in the measurements). A full description of the solution for this class (of Gaussian distributions) is given in Section 4 where it is also shown that for some special cases the solution is still affine in the measurements. One of these cases corresponds to the special class of Gaussian distributions where the DMs' perceptions of the marginal probability distributions of each others measurements are identical, but the correlations between the measurement variables could be perceived differently. The section also contains some discussion on stochastic team problems, treated as a special case.

The paper ends with the concluding remarks of Section 5. For proofs of some of the results given this paper the reader is referred to the more complete version [3].

2. Mathematical Formulation and Basic Definitions

Problem formulation with the exception of the definition of equilibrium, is analogous to that of [1], and hence the presentation below will be brief.

Using the notation of [1], we let x (unknown state of Nature) be a random vector on (R^n, B^n, P_x) , y_i (observation of DMI) be a random vector on $(R^{n_i}, B^{n_i}, P_{y_i})$, and P be a class of probability measures for $z = (x', y_1', y_2')$. Choosing two elements out of P (to be denoted P^1 and P^2), we assume that

Condition 1. $P^1 \ll P^2$, $P^2 \ll P^1$ (absolute continuity)

Condition 2. The Radon-Nikodym derivative

$$g^i(z) = \frac{dP^i}{dP^1} \quad i=1,2 \quad (1)$$

is uniformly bounded a.e. P^i , $i=1,2$.

The decision variable u_i of DMI belongs to a real separable Hilbert space U_i with inner product $(\cdot, \cdot)$. Permissible policies (decision rules) for DMI are measurable mappings

$$\pi_i: R^{n_i} \rightarrow U_i$$

satisfying

$$P_j^1(\xi) \left(\int_1 P_j^1(d\xi) \right) = 0, \quad j=1,2.$$

Let π_1 be the space of all such policies, with the probability measures satisfying conditions 1) and 2). Then, it is a Hilbert space under the inner product

$$\langle \gamma, \beta \rangle_1 = \int_1 (\gamma(\xi), \beta(\xi))_1 P_1^1(d\xi). \quad (2)$$

Now finally, let us introduce a cost function for each DM, by

$$J_1(\gamma_1, \gamma_2) = \frac{1}{2} \langle \gamma_1, \gamma_1 \rangle_1 + \frac{1}{2} \int_1 (\gamma_j(\xi), D_{jj}^1 \gamma_j(\xi))_j P_j^1(d\xi) - \langle \gamma_1, E^1[F_1^1 x | \gamma_1] \rangle_1 - \int_{X \times Y_2} (\gamma_j(\xi), F_j^1 x)_j P^1(dx, \gamma_1, d\xi) - \langle \gamma_1, E^1[D_{1j}^1 \gamma_j(\gamma_2) | \gamma_1] \rangle_1 \quad (3)$$

where D_{jj}^1 , F_1^1 , F_j^1 , D_{1j}^1 are linear operators, $X = R^n$, $Y_1 = R^{m_1}$, and

$$E^1[-(z) | \gamma_1] = \int_1 -(z) P_1^1(dz | \gamma_1)$$

Here we assume that the mode of decision making is asymmetric with one DM dominating the decision process. If there is such a hierarchy, which permits one decision maker (say DM1) to announce and enforce his policy on the other DM, the relevant equilibrium solution is the leader-follower (Stackelberg) solution defined below.

Definition 1: A pair of policies $(\gamma_1^*, \gamma_2^*) \in \pi_1 \times \pi_2$ constitutes a leader-follower (Stackelberg) equilibrium solution to the decision problem formulated above, and with unique follower responses, if there exists a unique mapping $T_2: \pi_1 \rightarrow \pi_2$ satisfying

$$T_2(\gamma_1) = \gamma_2^* \quad \text{for } \gamma_1 \in \pi_1, \quad (4)$$

and furthermore

$$J_1(\gamma_1, T_2(\gamma_1)) \leq J_1(\gamma_1, \gamma_2) \quad \text{for } \gamma_1 \in \pi_1, \quad (5)$$

with $\gamma_2^* = T_2(\gamma_1^*)$.

1. General Sufficient Conditions for a Stackelberg Equilibrium Solution

In this section we obtain some general sufficient conditions for existence of a Stackelberg equilibrium solution, and provide a complete characterization of the solution. Subsequently we consider some special cases with some further structure imposed on the cost functionals and the probability measures.

Firstly we obtain an expression for DM2's reaction $T_2: \pi_1 \rightarrow \pi_2$, as defined by (4), using Proposition 1 of [1]:

$$T_2(\gamma_1) = \gamma_2^* = D_{21}^1 E^2[\gamma_1(\gamma_1) | \gamma_2] + F_2^2 E^2[x | \gamma_2] \quad (6)$$

where, for notation, we refer the reader to [1]. Note that the uniqueness assumption of Def. 1 is satisfied in this case. Hence, the derivation of the leader's Stackelberg policy $\gamma_1^* \in \pi_1$ involves (in view of (5)) the minimization of J_1 over π_1 after γ_2^* given by (6) is substituted in. This substitution yields

$$J_1(\gamma_1) = J_1(\gamma_1, \gamma_2^*) = \frac{1}{2} \langle \gamma_1, \gamma_1 \rangle_1 + \frac{1}{2} \int_1 (F_1^1 E^2[x | \gamma_2]) - D_{21}^1 E^2[\gamma_1(\gamma_1) | \gamma_2] \cdot D_{12}^1 D_{21}^1 E^2[\gamma_1(\gamma_1) | \gamma_2]$$

$$+ D_{22}^1 F_2^2 E^2[x | \gamma_2])_2 P_2^1(d\xi) - \langle \gamma_1, E^1[F_1^1 x | \gamma_1] \rangle_1 \quad (7) + \int_{X \times Y_2} (D_{21}^1 E^2[\gamma_1(\gamma_1) | \gamma_2] + F_2^2 E^2[x | \gamma_2])_2 P^1(dx, \gamma_1, d\xi) - \langle \gamma_1, E^1[D_{12}^1 D_{21}^1 E^2[\gamma_1(\gamma_1) | \gamma_2] | \gamma_1] \rangle_1 + E^1[D_{12}^1 F_2^2 E^2[x | \gamma_2] | \gamma_1] \rangle_1,$$

where we have deleted the subscript 1 in γ_1 in order to simplify the notation. Now, since π_1 is a linear space, and J is the sum of terms homogeneous of degree zero, one and two (maximum), any minimizing solution $\gamma_1^* \in \pi_1$ will have to satisfy

$$\Delta J(\gamma; h) = J(\gamma+h) - J(\gamma) = \delta J(\gamma; h) + \delta^2 J(\gamma; h) \geq 0 \quad \text{for } h \in \pi_1, \quad (8)$$

where $\delta J(\gamma; h)$ is the Gateaux variation of $J(\gamma)$ of degree 1. Extensive manipulations, details of which are given in [3], lead to the following expressions for δJ and $\delta^2 J$:

$$\delta J(\gamma; h) = \langle h, \gamma \rangle_1 - \int_1 (h(\gamma_1), (E^2 \gamma)(\gamma_1))_1 P_1^1(d\gamma_1) - \int_1 (h(\gamma_1), E^2(\gamma_1))_1 P_1^1(d\gamma_1) \quad (9)$$

$$\delta^2 J(\gamma; h) = \frac{1}{2} \langle h, h \rangle_1 + \frac{1}{2} \int_1 (h(\gamma_1), g^1(\gamma_1) E^2[g^2(\gamma_1) D_{21}^1 D_{12}^1] D_{21}^1 E^2[h(\gamma) | \gamma_1] | \gamma_1])_1 P_1^1(d\gamma_1) - \langle h, D_{21}^1 D_{12}^1 E^2[\gamma_1 | \gamma_1] \rangle_1 \quad (10)$$

where $E: \pi_1 \rightarrow \pi_1$ and $E \in \pi_1$ are defined by

$$(E\gamma)(\gamma_1) = (D_{21}^1 D_{12}^1 P_1^1 - D_{21}^1 D_{12}^1 P_1^1)(\gamma_1) - D_{21}^1 D_{22}^1 D_{21}^1 g^1(\gamma_1) E^2[g^2(\gamma_1) E^2[\gamma(\gamma_1) | \gamma_2] | \gamma_1] \quad (11a)$$

$$E(\gamma_1) = F_1^1 E[x | \gamma_1] - D_{21}^1 D_{22}^1 F_2^2 g^1(\gamma_1) E^2[g^2(\gamma_1) E^2[x | \gamma_2] | \gamma_1] - D_{21}^1 F_2^2 E^1[E^2[x | \gamma_2] | \gamma_1] + D_{21}^1 F_2^2 g^1(\gamma_1) E^2[g^2(\gamma_1) E^1[x | \gamma_2] | \gamma_1] \quad (11b)$$

$P_{1|1}: U_1 \rightarrow U_1$ is a linear operator given by

$$P_{1|1}(\gamma_1) = E^1[E^2[\gamma(\gamma_1) | \gamma_2] | \gamma_1] \quad (12)$$

U_1 is the space of γ_1 -measurable random variables taking values in U_1 , and $g^i(\xi)$ are the Radon-Nikodym derivatives given by (1). Note that $P_{1|1}$ is related to $\bar{P}_{1|1}$ defined by (18) in [1] by

$$P_{1|1}(\gamma(\gamma_1)) = (\bar{P}_{1|1})(\gamma_1)$$

where the latter (which is a mapping from π_1 into π_1) has been used in (10) and will also be used in the sequel whenever needed.

Now, since (8) is also equivalent to

$$\left. \begin{aligned} \delta J(\gamma; h) &= 0 & \text{for } h \in \pi_1 \\ \delta^2 J(\gamma; h) &\geq 0 & \text{for } h \in \pi_1 \end{aligned} \right\} \quad (13)$$

a Stackelberg solution $\gamma_1^* \in \pi_1$ will exist for the leader if, and only if,

(1) (10) is nonnegative definite, and (from (9)):

$$(11) \quad (y_1) = (\bar{E}y_1) - \bar{E}(y_1) = 0, \text{ a.e. } P_{y_1}^1. \quad (14)$$

Since the first of these conditions does not depend on the optimal solution is solely determined by (14), we

$$\begin{aligned} (y_1) &= D_{11}^1 D_{21}^2 E^1 [E^2 (y_1) (y_2) | y_1] \\ &+ D_{21}^1 D_{12}^2 g^1(y_1) E^2 [g^2(y_2) E^1 [(y_1) (y_2) | y_1] \\ &+ D_{21}^1 F_{12}^1 g^1(y_1) E^2 [g^2(y_2) E^1 [x (y_2) | y_1] \\ &- D_{21}^1 D_{22}^2 D_{21}^2 g^1(y_1) E^2 [g^2(y_2) E^1 [(y_1) (y_2) | y_1] \\ &+ F_{12}^1 E^1 [x (y_1) | y_1] - D_{21}^1 D_{22}^2 F_{12}^1 g^1(y_1) E^2 [g^2(y_2) E^2 [x (y_2) | y_1] \\ &+ D_{12}^1 F_{21}^2 E^1 [E^2 [x (y_2) | y_1] \end{aligned} \quad (15)$$

where we have utilized the fact that the adjoint of P_{11}^1 is a linear operator $P_{11}^1: U_1 \rightarrow U_1$, given by (see [3])

$$\begin{aligned} P_{11}^1 (y_1) &= \frac{\int_{y_1} \int_{y_2} (dx dy_2) P_{y_1 y_2}^1 (dy_1 x dy_2)}{\int_{y_1} \int_{y_2} P_{y_1 y_2}^1 (dy_2) P_{y_1}^1 (dy_1)} \\ &= g^1(y_1) E^1 [g^2(y_2) E^1 [(y_1) (y_2) | y_1] \end{aligned} \quad (16)$$

Furthermore, condition (1) can be rewritten as

$$\begin{aligned} A &= I - \frac{1}{2} D_{11}^1 D_{21}^2 D_{21}^2 K + K^* \\ &- D_{11}^1 D_{21}^2 \bar{P}_{11}^1 - D_{21}^1 D_{12}^2 \bar{P}_{11}^1 \geq 0 \end{aligned} \quad (17)$$

where $I: U_1 \rightarrow U_1$ is the identity operator, and $K: U_1 \rightarrow U_1$ is defined by

$$K(y_1) = g^1(y_1) E^1 [g^2(y_2) E^2 [(y_1) (y_2) | y_1] \quad (18)$$

We now summarize these results in the following proposition:

Proposition 1: Under Conditions 1) and 2), the decision problem with multiple probability measures admits a Stackelberg equilibrium solution if, and only if, A is nonnegative definite and (15) admits a solution in $\bar{E}$.

Equation (15) will, in general, not admit a closed-form solution, even if all random variables are jointly Gaussian distributed (see Section 4); therefore, we will have to resort to numerical computations which will involve a recursion of some type. Hence, in analyzing the conditions of existence of a solution to (15) we may also require that such a numerical scheme be globally convergent (or stable, in the sense defined in [1]). One appealing scheme whereby a unique solution to (15) (or, equivalently, (14)) can be obtained is the recursion

$$(k) (y_1) = (\bar{E}^{(k-1)})(y_1) - \bar{E}(y_1), \quad k=1,2,\dots \quad (19)$$

where (0) is chosen as an arbitrary element of U_1 . If the limit $\lim_{k \rightarrow \infty} (k) \bar{E}^{(k)}$ exists in U_1 , for all such initial choices, then $\bar{E}$ will necessarily constitute a solution to (15). A sufficient condition for this is the following:

Proposition 2: In addition to the conditions of Prop. 1, assume that there exists a scalar $\alpha, 0 < \alpha < 1$, such that

$$r(\bar{E}) \leq \alpha \quad (20)$$

where $r(\bar{E})$ is the spectral radius of $\bar{E}$. Then, the decision problem admits a unique Stackelberg equilibrium solution $(y^s, T_2(y^s))$, where $y^s \in U_1$ is the limit of the iterative scheme (19), and T_2 is the affine operator (6).

We now elaborate on (20), so as to bring it to a form which separates out the contributions from the deterministic and probabilistic components of the problem. Towards this end, let us first note that using (18) in (11a):

$$r(\bar{E}) = r(D_{12}^1 D_{21}^2 \bar{P}_{11}^1 + D_{21}^1 D_{12}^2 \bar{P}_{11}^1 - D_{21}^1 D_{22}^2 D_{21}^2 K) \quad (21)$$

and utilizing the inequality relationship between the spectral radius and norm of an operator (see [3]) this can be bounded from above by

$$\leq \|D_{12}^1 D_{21}^2 \bar{P}_{11}^1 + D_{21}^1 D_{12}^2 \bar{P}_{11}^1 - D_{21}^1 D_{22}^2 D_{21}^2 K\|_1$$

where $\|\cdot\|_1$ is the operator norm as defined in [1, (17a)]. Using the standard (triangle inequality) property of norms, this can further be bounded from above by

$$\leq \|D_{12}^1 D_{21}^2 \bar{P}_{11}^1 + D_{21}^1 D_{12}^2 \bar{P}_{11}^1\|_1 + \|D_{21}^1 D_{22}^2 D_{21}^2 K\|_1$$

Now since both $D_{12}^1 D_{21}^2 \bar{P}_{11}^1$ and K map a Hilbert space (U_1) into itself, using the norm inequality for products of linear operators, we further have

$$\leq \|D_{12}^1 D_{21}^2 \bar{P}_{11}^1 + D_{21}^1 D_{12}^2 \bar{P}_{11}^1\|_1 + \|D_{21}^1 D_{22}^2 D_{21}^2\|_1 \|K\|_1$$

$$= \|D_{12}^1 D_{21}^2 \bar{P}_{11}^1 + D_{21}^1 D_{12}^2 \bar{P}_{11}^1\|_1 + \|D_{21}^1 D_{22}^2 D_{21}^2\|_1 \|K\|_1$$

where the equality follows because (1) the spectral radius and norm of a self-adjoint linear operator are equal [5, p. 514], (2) norm of a "non-self-adjoint" linear operator K is equal to the square root of the spectral radius of the self-adjoint operator K^*K (see [3]). Finally, the latter is bounded from above by [3]

$$\begin{aligned} r(\bar{E}) &\leq 2 \|D_{12}^1 D_{21}^2 D_{21}^2 \bar{P}_{11}^1\|_1^{1/2} [r(\bar{P}_{11}^1 P_{11}^1)]^{1/2} \\ &+ r(D_{21}^1 D_{22}^2 D_{21}^2) [r(K^*K)]^{1/2} \end{aligned} \quad (22)$$

Now, let us assume the following:

Condition 3): There exist four positive scalars $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, satisfying

$$\alpha_1 \alpha_2 \alpha_3 + \alpha_3 \alpha_4 = 1 \quad (23)$$

such that

$$r(D_{12}^1 D_{21}^2 D_{21}^2 \bar{P}_{11}^1) \leq \alpha_1^2, \quad r(D_{21}^1 D_{22}^2 D_{21}^2) \leq \alpha_3^2 \quad (24a)$$

$$r(\bar{P}_{11}^1 P_{11}^1) \leq \alpha_2^2, \quad r(K^*K) \leq \alpha_4^2 \quad (24b)$$

Then, we have

Theorem 1: Under Conditions 1) - 3) of §2 and Condition 3) given above, the decision problem admits a unique Stackelberg equilibrium solution $(y^s, T_2(y^s))$, where $y^s \in U_1$ is the limit of the iterative scheme (19), and T_2 is given by (6).

Proof: The result follows from Prop. 1 and the discussion and derivation that leads to Condition 3), provided we show that the given three conditions subsume (17), i.e. nonnegativity of operator A . We now verify that Condition 3) in fact implies that A is a strongly positive operator. First note that A is self-adjoint, because K commutes with $D_{12}^1 D_{21}^2 D_{21}^2$. Hence, we can write down the inequality

$$r(A-I) \leq \frac{1}{2} r(D_{12}^1 D_{21}^2 D_{21}^2 (K+K^*))$$

$$+ r(D_{12}^1 D_{21}^2 \bar{P}_{11} + D_{21}^{2*} D_{12}^{1*} P_{11}^*)$$

Then, using the line of arguments that led to (22) from (21), and the spectral radius inequality for the product of two self-adjoint operators, we obtain the bound

$$\begin{aligned} r(A-I) &\leq \frac{1}{2} r(D_{12}^1 D_{21}^2 \bar{P}_{11} + D_{21}^{2*} D_{12}^{1*} P_{11}^*) \\ &\quad + r[(D_{12}^1 D_{21}^2 D_{21}^{2*} D_{12}^{1*})]^{1/2} [r(\bar{P}_{11} P_{11}^*)]^{1/2} \\ &\leq \frac{1}{2} \rho_3 r(K+K^*) + \rho_{1-2} \end{aligned}$$

But note that

$$r(K+K^*) = \sup_{\langle y, y \rangle_1 = 1} [\langle y, (K+K^*)y \rangle_1] = 2 \sup_{\langle y, y \rangle_1 = 1} [\langle y, Ky \rangle_1 / \langle y, y \rangle_1]$$

and since, from the Cauchy-Schwarz inequality of inner products,

$$|\langle y, Ky \rangle_1| \leq \langle y, y \rangle_1^{1/2} \langle Ky, Ky \rangle_1^{1/2}$$

we have

$$\begin{aligned} r(K+K^*) &\leq 2 \sup_{\langle y, y \rangle_1 = 1} [\langle Ky, Ky \rangle_1 / \langle y, y \rangle_1]^{1/2} = 2 \sup_{\langle y, y \rangle_1 = 1} [\langle y, K^* Ky \rangle_1 / \langle y, y \rangle_1]^{1/2} \\ &= 2 [r(K^* K)]^{1/2} \leq 2 \rho_4 \end{aligned}$$

Thus,

$$r(A-I) \leq \rho_3 + \rho_{1-2} \leq \rho_1$$

implying that the spectrum of the self-adjoint operator $A-I$ is uniformly in the unit sphere. Hence, A is strongly positive.

A special class of problems of interest are the strictly convex team problems for which

$$D_{12}^1 = I, D_{12}^2 = D_{21}^{1*}, F_1^1 = F_2^1, F_1^2 = F_2^2 \text{ and } D_{12}^1 D_{12}^{1*} > I$$

For such problems eq. (15) simplifies to

$$\begin{aligned} v(y_1) &= D_{12}^1 D_{12}^{1*} [E^1 \{E^2 [v(y_1) | y_2] + g^1(y_1) E^2 [g^2(y_2) | y_1] \\ &\quad + E^1 [v(y_1) | y_2] - E^2 [v(y_1) | y_2] | y_1] \\ &\quad + F_1^1 E^1 [x | y_1] + D_{12}^1 F_2^1 g^1(y_1) E^2 [g^2(y_2) | y_1] \\ &\quad + E^1 [x | y_1] | y_1] + D_{12}^1 F_2^1 E^1 [E^2 [x | y_2] | y_1] \end{aligned} \quad (25)$$

and in Condition (3) inequalities (24a) are replaced by the single inequality

$$[r(D_{12}^1 D_{12}^{1*} D_{12}^2 D_{12}^{2*})]^{1/2} = \langle D_{12}^1 D_{12}^{1*} \rangle_1 = \bar{\rho} \leq \rho_1 = \rho_3$$

where $\bar{\rho}_1$ can be taken to be less than one. Hence, (23) reads

$$(2\rho_2 + \bar{\rho}_4) < 1/\bar{\rho} \quad (26)$$

We now summarize these results as a corollary to Theorem 1.

Corollary 1: Under Conditions (1)-(2) of §2, and (26) given above, the strictly convex quadratic team problem, with multiple probability measures and asymmetric mode of decision making, admits a unique Stackelberg equilibrium solution $(\bar{v}, \bar{T}_2)$, where $\bar{v}$ is the limit of the iterative scheme (19) with

$$E^1(y_1) = D_{12}^1 D_{12}^{1*} (P_{11} + P_{11}^*) (y_1)$$

$$- g^1(y_1) E^2 [g^2(y_2) | y_1] + g^1(y_1) [y_2] | y_1]$$

and $\bar{T}_2$ is given by (6).

Remark 1: When the original problem is a Stackelberg game, but the probability measures are identical, a study of the original condition (20) reveals the inequality

$$\begin{aligned} r(\bar{Z}) &= r(D_{12}^1 D_{21}^2 + D_{21}^{2*} D_{12}^{1*} - D_{21}^{2*} D_{22}^2 D_{21}^2 \bar{P}_{11}) \\ &\leq r(D_{12}^1 D_{21}^2 + D_{21}^{2*} D_{12}^{1*} - D_{21}^{2*} D_{22}^2 D_{21}^2) \leq \rho < 1. \end{aligned}$$

This is the existence condition for the standard stochastic Stackelberg game to admit a unique solution, which corroborates the earlier result obtained in [4].

4. Jointly Gaussian Distributions

In decision and control theory, one appealing class of probability distributions has been the Gaussian distribution, because it leads to closed-form solutions in most cases. Indeed, even for the class of nonstandard multi-criteria stochastic decision problems with multiple probabilistic models, we have observed in [1] and [2] that when both subjective probabilities are Gaussian, the unique stable Nash equilibrium solution can be obtained in closed form and is affine in the observations. The question now is whether this appealing feature also extends to the model treated in this paper, where the solution concept is Stackelberg instead of Nash. The main conclusion of this section is that when the subjective Gaussian probability distributions are different, there is in general no counterpart of the results of [1,2] in the present context; that is the unique equilibrium solution will, in general, not be affine. However, for some special cases which will be delineated in the sequel, the unique solution will still be affine.

Towards this end, let us adopt the model and notation of [1, Section 4], and assume validity of (1, (29a)). Furthermore, to simplify the discussion to follow, let us take the mean values of the Gaussian distributions to be zero. Hence, let

$$\text{mean } (y_1, y_2) = (0, 0) \text{ under both } P^1 \text{ and } P^2 \quad (27)$$

$$\text{covariance } (y_1, y_2) = \Sigma_y = \begin{pmatrix} \Sigma_{y_1}^1 & \Sigma_{y_1 y_2}^1 \\ \Sigma_{y_2 y_1}^1 & \Sigma_{y_2}^1 \end{pmatrix} = 0 \quad (28)$$

under P^i , $i=1,2$, and assume

$$W_{ij} = \Sigma_{y_i}^{j-1} - \Sigma_{y_i}^{i-1} \geq 0, \quad i,j=1,2; j \neq i \quad (29)$$

Then, the decision problem will admit a unique linear solution, if and only if, equation (15) is satisfied by the decision rule

$$v(y_1) = A y_1 \quad (30)$$

for some linear bounded operator $A: R^{n_1} \rightarrow U_1$. Hence, using (15), A should be the solution of (by pulling A out of the conditional expectations)

$$\begin{aligned} A y_1 &= D_{12}^1 D_{21}^2 A E^1 [E^2 [v(y_1) | y_2] | y_1] \\ &\quad + D_{12}^1 D_{21}^2 A g^1(y_1) E^2 [g^2(y_2) | y_1] \\ &\quad + D_{12}^1 D_{21}^2 A g^1(y_1) E^2 [g^2(y_2) | y_1] \end{aligned} \quad (31a)$$

$$\begin{aligned}
& + F_{11}^1 E^1[x|y_1] + D_{21}^{2*} F_{22}^1 g^1(y_1) E^2[g^2(y_2) E^1[x|y_2] | y_1] \\
& - D_{21}^{2*} D_{22}^1 F_{22}^1 g^1(y_1) E^2[g^2(y_2) E^2[x|y_2] | y_1] \\
& - D_{11}^1 F_{22}^1 E^1[E^2[x|y_2] | y_1] , \forall y_1 \in \mathbb{R}^{m_1} .
\end{aligned}$$

Since the random variables are jointly Gaussian under both measures,

$$E^i[y_k | y_i] = S_{ki}^i y_i , \quad k, i = 1, 2, i=1, 2 \quad (31b)$$

$$E^i[x | y_i] = S_{0i}^i y_i , \quad i=1, 2 \quad (31c)$$

for some matrices S_{ki}^i and S_{0i}^i . In view of this, (31a) can be rewritten as

$$\begin{aligned}
Ay_1 &= (D_{12}^1 D_{21}^2 AS_{12}^2 S_{21}^1 + F_{11}^1 S_{01}^1 + D_{12}^1 F_{22}^2 S_{02}^2 S_{21}^1) y_1 \\
&+ (D_{21}^{2*} D_{12}^1 AS_{12}^2 - D_{21}^{2*} D_{22}^1 D_{21}^2 AS_{12}^2 \\
&+ D_{21}^{2*} F_{22}^1 S_{02}^2 - D_{21}^{2*} D_{22}^1 F_{22}^2 S_{02}^2) g^1(y_1) E^2[g^2(y_2) y_2 | y_1] .
\end{aligned} \quad (32)$$

This then leads to the following Proposition:

Proposition 3: Let (29) and Condition (5) be satisfied, and either $P_{y_1}^1 = P_{y_1}^{1*}$ or $P_{y_2}^1 = P_{y_2}^{1*}$. Then, the quadratic

Gaussian decision problem with asymmetric mode of decision making (as formulated in this section) admits a linear (Stackelberg) equilibrium solution, if and only if,

(i) there exists a bounded linear operator A:

$\mathbb{R}^{m_1} \rightarrow \mathbb{R}^{m_1}$ satisfying

$$A = D_{12}^1 D_{21}^2 AS_{12}^2 S_{21}^1 - F_{11}^1 S_{01}^1 - D_{12}^1 F_{22}^2 S_{02}^2 S_{21}^1 , \quad (33a)$$

and

(ii) this solution also satisfies

$$\begin{aligned}
D_{12}^1 D_{21}^2 AS_{12}^2 - D_{21}^{2*} D_{12}^1 D_{21}^2 AS_{12}^2 - D_{21}^{2*} F_{22}^1 S_{02}^2 \\
- D_{21}^{2*} D_{22}^1 F_{22}^2 S_{02}^2 = 0 .
\end{aligned} \quad (33b)$$

Proof: See [3].

Remark 2: A sufficient condition for (33a) to admit a unique solution in the Banach space of linear bounded operators mapping $\mathbb{R}^{m_1}$ into $\mathbb{U}_1$ is

$$\|(D_{12}^1 D_{21}^2 D_{12}^1 D_{21}^2)^{-1} \text{Tr} [S_{12}^2 S_{21}^1 S_{12}^2] \| < 1 ,$$

which is clearly satisfied under Condition 3.

The conditions of Prop. 3 are clearly non-void; because, given the unique solution of (33a), it may be possible to find $F_{11}^1, F_{22}^2, S_{02}^1$ and S_{02}^2 so that (33b) is satisfied. However, it should also be clear that satisfaction of (33b) places some severe restrictions on the parameters of the problem, which in general will not be met. Hence, it is fair to say that, if either $P_{y_1}^1 \neq P_{y_1}^{1*}$ or $P_{y_2}^1 \neq P_{y_2}^{1*}$, generically the problem does not admit a linear equilibrium solution, even if it is a team problem; that is:

Corollary 1: If either $P_{y_1}^1 \neq P_{y_1}^{1*}$ or $P_{y_2}^1 \neq P_{y_2}^{1*}$ (or both),

the quadratic Gaussian decision problem does not admit (generically) a linear (Stackelberg) equilibrium solution. The unique solution, which exists under (29) and Condition 3, is nonlinear.

The conditions of the preceding Corollary involve only the marginal distributions of y_1 and y_2 ; in the complement of these conditions we can derive the following linear solution:

Proposition 4: For the quadratic Gaussian decision problem, let both $P_{y_1}^1 = P_{y_1}^{1*}$ and $P_{y_1}^2 = P_{y_1}^{2*}$ (but not necessarily

ii $P_{y_2}^1 = P_{y_2}^{1*}$ and even $P_{y_1}^1 = P_{y_1}^{1*}$). Then, if

$$2[(D_{12}^1 D_{21}^2 D_{21}^1 D_{12}^1)^{1/2} + (D_{21}^{2*} D_{22}^1 D_{21}^2)^{1/2}] < 1 \quad (34)$$

the problem admits a unique Stackelberg equilibrium solution for DM1 (the leader) which is linear in y_1 :

$$g^1(y_1) = Ay_1 \quad (35a)$$

where A: $\mathbb{R}^{m_1} \rightarrow \mathbb{U}_1$ is the unique bounded linear operator solving

$$\begin{aligned}
Ay_1 &= (D_{12}^1 D_{21}^2 AS_{12}^2 S_{21}^1 + D_{21}^{2*} D_{12}^1 AS_{12}^2 S_{21}^1 \\
&- D_{21}^{2*} D_{22}^1 D_{21}^2 AS_{12}^2 S_{21}^1 + F_{11}^1 S_{01}^1 + D_{12}^1 F_{22}^2 S_{02}^2 S_{21}^1 \\
&+ D_{21}^{2*} F_{22}^1 S_{02}^2 - D_{21}^{2*} D_{22}^1 F_{22}^2 S_{02}^2) y_1 , \forall y_1 \in \mathbb{R}^{m_1} .
\end{aligned} \quad (35b)$$

and S_{ki}^i are defined by (30a)-(30c), and S_{0i}^i is defined by $E^i[x|y_i] = S_{0i}^i y_i$.

Proof: When $P_{y_1}^1 = P_{y_1}^{1*}$ and $P_{y_1}^2 = P_{y_1}^{2*}$, $g^1(y_1) = g^2(y_2) = 1$ and

hence Conditions 1 and 2 of Theorem 1 are always satisfied (see [1]), and in Condition 3, $\lambda = 1$. Then, (34) is the counterpart of (23), and hence existence and uniqueness follow from Theorem 1. Linearity, on the other hand, follows by noting that if we start iteration (19) with $v_1^{(0)} = 0$, since $g^1(y_1) = g^2(y_2) = 1$ every term will be linear in y_1 (see also [22]), and hence the limit which exists by Theorem 1 will be linear. Then, substituting $g^1(y_1) = Ay_1$ in (15), we obtain (35b), by simply letting $g^1(y_1) = g^2(y_2) = 1$ in (32).

Remark 3: For the special case of strictly convex team problems (cf. Corollary 1), (34) is replaced by

$$3 \ll D_{12}^1 D_{21}^2 \gg_1 < 1 , \quad (36a)$$

and (35b) simplifies to (cf. (25))

$$\begin{aligned}
Ay_1 &= (D_{12}^1 D_{21}^2 A(S_{12}^2 S_{21}^1 + S_{12}^2 S_{21}^1 - S_{12}^2 S_{21}^1) + F_{11}^1 S_{01}^1 \\
&+ D_{12}^1 F_{22}^2 (S_{02}^2 - S_{02}^2) S_{21}^1 + D_{12}^1 F_{22}^2 S_{02}^2 S_{21}^1) y_1 .
\end{aligned} \quad (36b)$$

When there is a discrepancy between the DMs' perceptions of the variances of either y_1 or y_2 , Prop. 4 will not hold, and the problem will admit (generically) a nonlinear equilibrium solution, as proven earlier in Prop. 3 and Corollary 3. In this case, an explicit closed-form solution cannot be obtained; however, an approximate solution can be derived by using the iteration (19) which, for the Gaussian problem, becomes

$$\begin{aligned}
y^{(k+1)}(y_1) &= D_{12}^1 D_{21}^2 E^1[E^2(y^{(k)}(y_1) | y_2) | y_1] - D_{21}^{2*} D_{12}^1 D_{21}^2 E^1[E^2(y^{(k)}(y_1) | y_2) | y_1] \\
&- E^2[g^2(y_2) E^1(y^{(k)}(y_1) | y_2) | y_1] - D_{21}^{2*} D_{22}^1 D_{21}^2 E^1[E^2(y^{(k)}(y_1) | y_2) | y_1] \\
&- E^2[g^2(y_2) E^2(y^{(k)}(y_1) | y_2) | y_1] + F_{11}^1 S_{01}^1 \\
&+ D_{12}^1 F_{22}^2 S_{02}^2 S_{21}^1 y_1 + (D_{21}^{2*} F_{22}^1 S_{02}^2 - D_{21}^{2*} D_{22}^1 F_{22}^2 S_{02}^2) \\
&- g^1(y_1) E^2[g^2(y_2) y_2 | y_1] .
\end{aligned} \quad (37)$$

If we start this iteration with $v^{(0)}(y_1)=0$, or any linear function of v_1 , at every iteration we obtain linear combinations of terms of the type $A^{(k)}y_1$ and $B^{(k)}y_1 \exp \{-\frac{1}{2}v_1^T V^{(k)} y_1\}$, where $A^{(k)}$ and $B^{(k)}$ are linear operators, and $V^{(k)} \geq 0$ is an $m \times m$ matrix. Since this is a successive approximation technique under condition 5, even stopping the iteration after a finite number of terms will provide a solution sufficiently close to the unique optimum. Hence, generically, a suboptimal policy for DM1, which is sufficiently close to the unique solution of (15), will be of the form

$$v_1(y_1) = A^{(N)}y_1 + \sum_{i \leq N} B^{(i)}y_1 \exp \{-\frac{1}{2}v_1^T V^{(i)} y_1\},$$

where N is a sufficiently large integer (related to the number of iterations taken in (37)), and $A^{(N)}$, $B^{(i)}$, $V^{(i)}$ are generated via the iteration (47). Note that as $N \rightarrow \infty$ this solution will uniformly converge to the unique optimum.

Yet another suboptimal solution can be obtained by restricting DM1's policies, at the outset, to linear functions of y_1 , i.e. to the form (30) where A is a variable linear operator. DM2's response to any such policy will also be linear (in y_2), thus making T_2 in (6) a linear operator. Then, the problem faced by DM1 is minimization of (7) with $v(y_1)=Ay_1$, over all linear bounded operators A . The solution of this minimization problem will provide DM1 with a linear policy that is (in general) inferior to the limiting solution of (37), unless, of course, $g^*(y_1)=g^*(y_2)=1$ in which case the two solutions will be the same (satisfying (36b)). We do not pursue here the details of the derivation of the best linear solution for the general case (as outlined above).

Furthermore, it is possible to work out the various conditions obtained for the special cases of finite dimensional problems (especially the scalar team problem) and continuous-time problems, and write down the equilibrium solution explicitly whenever it is linear. Such an analysis would routinely follow the lines of the discussion of [1, Section 4], and hence it will not be included here mainly because of space limitations.

5. Conclusions

This paper has presented an equilibrium theory for two-person quadratic decision problems with static information patterns, wherein the decision makers (DMs) do not necessarily have the same perception of the underlying probability space, and there is a hierarchy in decision making. As indicated earlier in Section 1, when such discrepancies exist, even team problems have to be analyzed in the framework of nonzero-sum stochastic games, and because of the presence of hierarchy the Stackelberg solution becomes the most meaningful equilibrium concept for such decision problems.

Section 3 of the paper has provided a set of sufficient conditions for existence and uniqueness of an equilibrium solution for a general decision problem defined on Hilbert decision spaces and with arbitrary multiple probabilistic description. These conditions also ensure that the solution (more precisely, the equilibrium policy of the leader) can be obtained as the limit of an infinite sequence which involves conditional expectations under two different probability measures. This sequence is structurally different from its counterpart in the case when the mode of decision making is symmetric [1], even for team problems, and it contains Radon-Nikodym derivatives of the two probability measures as multiplying factors (which were absent in the end result of [1]).

This different structure has led, in Section 4, to a seemingly surprising (unexpected result for the special case of Gaussian distributions - the unique

equilibrium solution being generically nonlinear in the measurements. This constitutes the first nonlinear solution reported in the literature for a quadratic Gaussian static game or team problem. It should be noted that we have not given a closed-form expression for this nonlinear solution, but have provided a recursive scheme which generates admissible policies that come arbitrarily close to the optimum solution. Furthermore, under some assumptions on the relative structures of the probability measures perceived by the DMs, we have shown that the unique solution is linear in the static measurements.

Possible extensions of this study could be carried out along the lines discussed in some detail in Section 5 of [1]. Particularly, one of the issues that requires immediate attention is an analysis of the existence conditions of this paper, and the structure of the equilibrium solution, when the discrepancies between the perceived probability measures are sufficiently small, such as the probability measures being within an ϵ -neighborhood of a common nominal one. Indeed, such an analysis and the ensuing results will provide the right framework for a further analysis that involves an investigation of sensitivity and robustness properties of team solutions (obtained under the common probability measure alluded to above) in this ϵ -neighborhood. These questions are currently under study, and results along these lines will be reported in the future.

References

- [1] T. Başar, "Decision problems with multiple probabilistic models," *Proc. 1983 American Control Conference*, San Francisco, June 1983, pp. 1091-1096.
- [2] T. Başar, "An equilibrium theory for multi-person decision making with multiple probabilistic models: Part I. Symmetric mode of decision making," University of Illinois, Decision and Control Lab, Internal Report, 1983.
- [3] T. Başar, "An equilibrium theory for multi-person decision making with multiple probabilistic models: Part II. Asymmetric mode of decision making," University of Illinois, Decision and Control Lab, Internal Report, 1983.
- [4] T. Başar, "Hierarchical decision making under uncertainty," in *Dynamic Optimization and Mathematical Economics*, P. T. Liu (ed.), Plenum Press, New York, 1979, pp. 205-221.
- [5] L. W. Kantorovich and G. P. Akhilov, *Functional Analysis in Normed Spaces*, New York: McMillan, 1964.

November 1982
Revision July 1983

AN EQUILIBRIUM THEORY FOR MULTI-PERSON DECISION MAKING
WITH MULTIPLE PROBABILISTIC MODELS¹

by

Tamer Başar
Department of Electrical Engineering
and Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Avenue
Urbana, Illinois 61801

¹Research reported in this paper was supported in part by the Office of Naval Research under Contract N00014-82-K-0469 and in part by the Joint Services Electronics Program under Contract N00014-79-C-0424.

Abstract

This paper develops an equilibrium theory for two-person two-criteria stochastic decision problems with static information patterns, wherein the decision makers (DM's) have different probabilistic models of the underlying process, the objective functionals are quadratic and the decision spaces are general inner-product spaces. Under two different modes of decision making (viz. symmetric and asymmetric), sufficient conditions are obtained for the existence and uniqueness of equilibrium solutions (stable in the former case), and in each case a uniformly convergent iterative scheme is developed whereby the equilibrium policies of the DM's can be obtained by evaluating a number of conditional expectations. When the probability measures are Gaussian, the equilibrium solution is linear under the symmetric mode of decision making, whereas it is generically nonlinear in the asymmetric case, with the linear structure prevailing only in some special cases which are delineated in the paper.

1. Introduction

A *team* is defined as a group of agents who work together in a coordinated effort, in a possibly hostile and uncertain environment, in order to achieve a common goal. In achieving this goal, the members of the team do not necessarily acquire the same information, and hence they have to operate in a decentralized mode of decision making. The scientific approach to formulation and analysis of team problems has involved (i) a quantification of the underlying common goal in the form of a (mathematical) objective function which is sought to be optimized jointly by the agents, and (ii) a modeling of the uncertain environment and the possible measurements made by the agents on this environment in the form of a probability space together with an appropriate information structure [14,7,15,16]. The underlying stipulation here has been the existence of a probability space that is common to all the agents, so that through their priors all members of the team "see the world" in exactly the same way.

One question that readily comes into mind at this point is the robustness of such a mathematical model, and the "optimum" solutions it produces, to slight variations in the underlying assumptions. In particular, what if the agents perceive the outside world in slightly different ways? Would the solution obtained under the assumption of common prior probability measures change drastically if there are discrepancies in the agents' perceptions of the probabilistic description of the outside world? In order to be able to answer these queries satisfactorily and effectively, we need a theory of equilibrium for decision problems in which the decision makers (DM's) have different probabilistic models of the system; such a general theory will clearly subsume the currently available results on teams which use a common probability space.

Consider a static team decision problem, formulated in the standard manner as in [7], with the only difference being in the underlying probability space. In particular, assume that the DM's assign different subjective probabilities to the uncertain events, in which case there will not exist a common probability space, thereby leading to a different expected (average) cost function for each DM. Hence, once we relax the assumption of existence of a common probability space, the team problem is no longer a stochastic optimization problem with a single objective functional, and we inevitably have to treat it as a nonzero-sum stochastic game [5,8,12]. Furthermore, even though the original team decision problem with a common probability space will admit the same team-optimal solution(s) regardless of the mode of decision making (that is, regardless of whether the roles of the DM's are symmetric or whether there is a hierarchy and dominance in decision making), this feature ceases to hold true when there exists a discrepancy between the perceived probability measures. When there are only two members, for example, two possibilities emerge in the presence of discrepancies: the totally symmetric roles, corresponding to the Nash equilibrium solution, and the hierarchical mode, corresponding to the Stackelberg equilibrium solution.

Motivated by these considerations, we treat in this paper a more general (than team) class of two-person stochastic decision problems which can be viewed as static stochastic nonzero-sum games with the DM's having different subjective probability measures. Adopting both the symmetric and asymmetric modes of decision making, we develop in each case a general theory of equilibrium when the objective functionals are quadratic and the

decision spaces are appropriate Hilbert spaces. Such a formulation includes both finite-dimensional (discrete) and continuous-time decision problems, and involves arbitrary probability measures which are, though, restricted *a posteriori* by the conditions of existence and uniqueness developed in the paper. The special case of Gaussian distributions is studied in considerable depth, and some explicit solutions are obtained with appealing features.

The organization of the paper is as follows. The next section (§2) provides a precise problem formulation, and introduces the two solution concepts adopted in the paper. Section 3 develops general conditions for existence and uniqueness of a stable equilibrium solution under the symmetric mode of decision making, and elucidates the extent of the restrictions imposed on the problem by these conditions. Section 4 presents a counterpart of the results of Section 3 under the asymmetric mode of decision making, with the mathematical machinery used being inherently different from that of §3. Section 5 deals with the special class of Gaussian distributions, under both symmetric and asymmetric modes of decision making. In the former case it is shown that the unique stable equilibrium solution is affine in the measurements and can be obtained explicitly. In the latter case, however, the solution is generically nonlinear, and contains summation of terms which involve products of linear functions of measurements with exponential terms (whose exponents are quadratic in the measurements). The section also contains some discussion on finite-dimensional and continuous-time problems, treated as special cases. Section 6 is devoted to discussions on possible extensions of these results in different directions, provides some interpretation of the general approach and results, and includes some concluding remarks. The paper ends with five Appendices which include results used in the main body of the paper.

2. Mathematical Formulation and Some Basic Results

2.1. *Probability Spaces*

Let $\Omega = \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \triangleq X \times Y_1 \times Y_2$, $\mathcal{B}$ denote the Borel field of subsets of Ω , and $\mathcal{B}^k$ denote the Borel field of subsets of $\mathbb{R}^k$, $k = n, m_1, m_2$. Let $\mathcal{P}$ denote the set of all probability measures on $(\Omega, \mathcal{B})$ with finite second moments, and for each $P \in \mathcal{P}$ denote the corresponding marginal measures on $\mathcal{B}^n$, $\mathcal{B}^{m_1}$ and $\mathcal{B}^{m_2}$ by P_x , P_{y_1} and P_{y_2} , respectively. Furthermore, let the collection of all such probability measures be denoted by $\mathcal{P}_x$, $\mathcal{P}_{y_1}$ and $\mathcal{P}_{y_2}$, respectively. Then, for each $P \in \mathcal{P}$, the vector $z = (x', y_1', y_2')'$, taking values in Ω , becomes a well-defined random vector on $(\Omega, \mathcal{B}, P)$, and likewise x is a random vector on $(\mathbb{R}^n, \mathcal{B}^n, P_x)$ and y_i is a random vector on $(\mathbb{R}^{m_i}, \mathcal{B}^{m_i}, P_{y_i})$.

Here, x denotes the unknown state of Nature, and y_i denotes an observation of DMi (i'th decision maker), which is correlated with x . We now choose two elements out of $\mathcal{P}$, P^1 and P^2 , which denote the subjective probabilities assigned to z by DM1 and DM2, respectively. For technical reasons, we place some further restrictions on the choices of P^1 and P^2 through the marginals $P_{y_j}^i$; in particular we assume that

Condition (1). $P_{y_2}^1$ and $P_{y_1}^2$ are absolutely continuous [1] with respect to $P_{y_2}^2$ and $P_{y_1}^1$, respectively; that is, using the standard notation in probability theory,

$$P_{y_2}^1 \ll P_{y_2}^2, \quad P_{y_1}^2 \ll P_{y_1}^1. \quad (1)$$

Condition (2). The Radon-Nikodym (R-N) derivative [1]

$$g^i(\xi) = dP_{y_i}^j / dP_{y_i}^i, \quad j \neq i, \quad (2)$$

is uniformly bounded a.e. $P_{y_i}^i$, $i=1,2$.

The necessity of these two conditions in the formulation of our problem will be made clear in the sequel. We should note, however, that for the special case when P^1 is equivalent to P^2 , both of these conditions are satisfied (in the latter case the bound is equal to 1) and we have the standard decision theoretic framework [2] with a single probability space.

2.2. Decision and Policy Spaces

The decision variable of DMi will be denoted by u_i which belongs to a real separable Hilbert space U_i with inner product $(\cdot, \cdot)_i$. Permissible policies (decision rules) for DMi are measurable mappings

$$\gamma_i: \mathbb{R}^{m_i} \rightarrow U_i, \quad \int \|\gamma_i(\xi)\|_i^2 P_{y_i}^i(d\xi) < \infty \quad (3)$$

where $\|\cdot\|_i$ is the natural norm derived from $(\cdot, \cdot)_i$. Let Γ_i denote the space of all such policies, which is further equipped with the inner product

$$\langle \gamma, \beta \rangle_i = \int_{Y_i} (\gamma(\xi), \beta(\xi))_i P_{y_i}^i(d\xi) \quad (4)$$

Then, we have the following two results the first of which is standard [3] and the second one involves a change of measures using the R-N derivative.

Lemma 1. Γ_i is a Hilbert space. □

Lemma 2. If Conditions (1) and (2) are satisfied, every element of Γ_i has bounded second-order moments also under $P_{y_i}^j$, $j \neq i$.

2.3. Cost Functionals

Let $D_{jj}^i: U_j \rightarrow U_j$ ($i \neq j$, $i, j=1, 2$) be strongly positive⁺ bounded linear operators, and $F_j^i: X \rightarrow U_j$ be bounded linear operators for all $i, j=1, 2$. Furthermore, let $E^i[u^i(z)|y_i]$ denote the mathematical expectation of a

⁺That is, there exists $\alpha > 0$ such that $(u, D_{jj}^i u)_j \geq \alpha(u, u)_j$ for all $u \in U_j$.

z -measurable random variable $\mu^i(z)$ taking values in U_i conditioned on the random variable y_i , and under the probability measure P^i , i.e.

$$E^i[\mu(z)|y_i] = \int_{\Omega} \mu(z) P^i_{|y_i}(dz|y_i) \quad (5)$$

where the second term of the integrand is the conditional probability measure derived from P^i . Then, for each pair $(\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2$, we have a quadratic expected cost functional for each DM, defined for DMi by

$$J_i(\gamma_1, \gamma_2) = \frac{1}{2} \langle \gamma_i, \gamma_i \rangle_i + \frac{1}{2} \int_{Y_j} (\gamma_j(\xi), D_{jj}^i \gamma_j(\xi))_j P^i_{Y_j}(d\xi) - \langle \gamma_i, E^i[F_i^i x | y_i] \rangle_i \\ - \int_{X \times Y_j} (\gamma_j(\xi), F_{ij}^i x)_j P^i(dx, Y_i, d\xi) - \langle \gamma_i, E^i[D_{ij}^i \gamma_j(y_j) | y_i] \rangle_i \quad (6)$$

every term of which can be shown to be finite, in view of Lemmas 1 and 2. Note that in the absence of *Conditions (1) and (2)*, J_i is not necessarily finite and hence the problem is not well defined.

It is worth mentioning here that J_i describes a most general type of quadratic cost functional which is strictly convex in u_i , and that the formulation here covers also the cases of team problems

$$(D_{jj}^i = I, D_{12}^1 = D_{21}^{2*}, F_i^1 = F_i^2, i, j=1, 2, i \neq j) \text{ and } \underline{\text{zero-sum games}}$$

$(D_{jj}^i = -I, D_{12}^1 = -D_{21}^{2*}, F_i^1 = -F_i^2, i, j=1, 2, i \neq j)$. But even in these "single loss-functional" problems, the DM's will have inherently different expected cost functions whenever P^1 and P^2 are different, since then a common probability space does not exist. This forces us to formulate the problem as a multi-criteria optimization problem and introduce equilibrium solution concepts that would be appropriate in this framework.

A superscript () designates the adjoint of a given linear operator defined on a Hilbert space, and I designates the identity operator.

2.4. Equilibrium Solution Under the Symmetric Mode of Decision Making

Since the expected cost functionals (6), together with the policy spaces, provide a normal (strategic) form description, regardless of the presence of multiple probability measures, the standard definition of noncooperative (Nash) equilibrium [5] remains intact, which is the most reasonable solution concept here under the symmetric mode of decision making.

Definition 1. A pair of policies $(\gamma_1^0, \gamma_2^0) \in \Gamma_1 \times \Gamma_2$ constitutes a Nash equilibrium solution if

$$J_1(\gamma_1^0, \gamma_2^0) \leq J_1(\gamma_1, \gamma_2^0) \quad , \quad J_2(\gamma_1^0, \gamma_2^0) \leq J_2(\gamma_1^0, \gamma_2) \quad , \quad \forall \gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2 \quad . \quad (7)$$

Definition 2. A Nash equilibrium solution (γ_1^0, γ_2^0) is stable if for all $(\gamma_1^{(0)}, \gamma_2^{(0)}) \in \Gamma_1 \times \Gamma_2$,

$$\lim_{k \rightarrow \infty} \gamma_i^{(k)} = \gamma_i^0 \quad , \quad \text{in } \Gamma_i \quad , \quad i=1,2, \quad (8)$$

where

$$\gamma_1^{(k)} = \arg \min_{\Gamma_1} J_1(\gamma_1, \gamma_2^{(k-1)}) \quad (9a)$$

$$\gamma_2^{(k)} = \arg \min_{\Gamma_2} J_2(\gamma_1^{(k-1)}, \gamma_2) \quad , \quad k=1,2, \dots \quad (9b)$$

Remark 1. The notion of stable equilibrium makes particular sense (and is of paramount importance) in decision problems wherein the DM's have different priors on the uncertain quantities, because it is determined as the outcome of a natural iterative process. In this process, each DM responds optimally (using his priors) to the most recent decision (policy) of the other DM, with the priors on which this decision is based being irrelevant. In other words, even though the computation of the Nash equilibrium solution will depend on the different prior probability

measures perceived by two DM's, in the iterative procedure that leads to this equilibrium each DM has to know only his own prior and the other one's announced policy at the previous step. For an earlier utilization of this concept in a deterministic setting we refer the reader to [28]. $\square$

2.5. Equilibrium Solution Under an Asymmetric Mode of Decision Making

In the case of the asymmetric mode there is a hierarchy in decision making, which permits one DM (say DM1—*leader*) to announce and enforce his policy on the other DM (*follower*). The relevant solution concept here is the leader-follower (Stackelberg) solution which is introduced below.

Definition 3. A pair of policies $(\gamma_1^s, \gamma_2^s) \in \Gamma_1 \times \Gamma_2$ constitutes a leader-follower (Stackelberg) equilibrium solution with unique follower responses, if there exists a unique mapping $T_2: \Gamma_1 \rightarrow \Gamma_2$ satisfying

$$J_2(\gamma_1, T_2[\gamma_1]) \leq J_2(\gamma_1, \gamma_2) \quad , \quad \forall (\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2 \quad (10)$$

and furthermore

$$J_1(\gamma_1^s, T_2[\gamma_1^s]) \leq J_1(\gamma_1, T_2[\gamma_1]) \quad , \quad \forall \gamma_1 \in \Gamma_1 \quad (11)$$

with

$$\gamma_2^s = T_2[\gamma_1^s] \quad .$$

Remark 2. The uniqueness condition on T_2 is satisfied in our case, because J_2 is strictly convex (and quadratic) in γ_2 . $\square$

Remark 3. The solution introduced above may not, at first glance, appear to be an equilibrium solution, because of the strict ordering of the DM's. However, it can be shown, by following an argument first developed in [17], that the Stackelberg solution can be viewed as the so-called "strong equilibrium" of a decision problem with a modified (dynamic) information pattern [see Appendix E]. $\square$

3. General Conditions for a Stable Equilibrium Solution Under the Symmetric Mode

We now obtain some general conditions for existence of stable equilibrium solutions under the symmetric mode of decision making, and also consider some special cases when the probability measures of both DM's are absolutely continuous with respect to the Lebesgue measure (i.e. when densities exist). Firstly we have Proposition 1. A pair of policies $(\gamma_1^0, \gamma_2^0) \in \Gamma_1 \times \Gamma_2$ constitutes a Nash equilibrium solution to the decision problem of §2, if, and only if, it satisfies the pair of equations (under the notation of (5)):

$$\gamma_1^0(y_1) = D_{12}^1 E^1[\gamma_2^0(y_2)|y_1] + F_1^1 E^1[x|y_1] \quad (12a)$$

$$\gamma_2^0(y_2) = D_{21}^2 E^2[\gamma_1^0(y_1)|y_2] + F_2^2 E^2[x|y_2]. \quad (12b)$$

Proof. This result follows from a simple minimization of the two quadratic forms $J_1(\gamma_1, \gamma_2^0)$ and $J_2(\gamma_1^0, \gamma_2)$ on the two Hilbert spaces Γ_1 and Γ_2 , respectively, and by virtue of the fact that these two quadratic forms are positive definite in the relevant variables. □

By the same argument used in the proof of Proposition 1, relations (9a) and (9b) in Def. 2 can equivalently be written as

$$\gamma_1^{(k)} = D_{12}^1 E^1[\gamma_2^{(k-1)}(y_2)|y_1] + F_1^1 E^1[x|y_1] \quad (13a)$$

$$\gamma_2^{(k)} = D_{21}^2 E^2[\gamma_1^{(k-1)}(y_1)|y_2] + F_2^2 E^2[x|y_2] \quad , \quad k=1,2, \dots \quad (13b)$$

Now, substituting (13b) into (13a), and also (13a) into (13b), by appropriately matching the superscripts, we arrive at the following two recursive relations:

$$\begin{aligned} \gamma_i^{(k)}(y_i) = & D_{ij}^i D_{ji}^j E^i[E^j[\gamma_i^{(k-2)}(y_i)|y_j]|y_i] + F_i^i E^i[x|y_i] \quad , \\ & + D_{ij}^i F_j^j E^i[E^j[x|y_j]|y_i] \quad , \quad j,i=1,2; \quad j \neq i; \quad k=2,4,\dots \text{ or } k=3,5,\dots \end{aligned} \quad (14)$$

Note that if the recursive scheme (14) converges for even values of k , it also converges (to the same limit) for odd values of k [this follows from expressions (13a)-(13b)]. Hence, we confine attention only to even values of k and obtain the following result as a direct consequence of the foregoing analysis:

Proposition 2. A pair of policies $(\gamma_1^0, \gamma_2^0) \in \Gamma_1 \times \Gamma_2$ constitutes a stable Nash equilibrium solution if, and only if, for all $(\gamma_1^{(0)}, \gamma_2^{(0)}) \in \Gamma_1 \times \Gamma_2$,

$$\gamma_i^0(y_i) = \lim_{\bar{k} \rightarrow \infty} \gamma_i^{(2\bar{k})}(y_i) \quad \text{in } \Gamma_i, \quad (15)$$

where $\gamma_i^{(2\bar{k})}$, $\bar{k}=1,2,\dots$, is given recursively by (14). Furthermore, such a stable equilibrium solution is necessarily unique. $\square$

Let us now introduce linear operators $S_i: \Gamma_i \rightarrow \Gamma_i$, $i=1,2$, by

$$S_i(\gamma) = D_{ij}^i D_{ji}^j E^i[E^j[\gamma(y_i)|y_j], y_i], \quad j \neq i; \quad i,j=1,2. \quad (16)$$

Note that S_i indeed maps Γ_i into Γ_i , because the conditional expectation $E^j[D_{ji}^j \gamma(y_i)|y_j]$ maps Γ_i into Γ_j ($j \neq i$) when the probability measures satisfy Conditions (1) and (2), and every element of Γ_i is square-integrable under both $P_{y_i}^i$ and $P_{y_i}^j$ (cf. Lemma 2).

Furthermore, let us introduce the notation $\langle\langle S \rangle\rangle_i$ to denote the norm of a linear bounded operator $S: \Gamma_i \rightarrow \Gamma_i$, which is defined by

$$\langle\langle S \rangle\rangle_i = \sup_{\gamma \in \Gamma_i} [\langle S\gamma, S\gamma \rangle_i / \langle \gamma, \gamma \rangle_i]^{1/2}, \quad (17a)$$

and $r_i(S)$ to denote the spectral radius of S , which is defined by [see Appendix A]

$$r_i(S) = \lim_{k \rightarrow \infty} \sup [\langle\langle S^k \rangle\rangle_i]^{1/k} \quad (17b)$$

where S^k denotes the k 'th power of S . Finally, let us introduce the linear operators

$$\bar{D}^i \triangleq D_{ij}^i D_{ji}^j \quad (18a)$$

and

$$\bar{P}_{i|i} \triangleq E^i[E^j[\cdot | y_j] y_i] \quad (18b)$$

both of which map Γ_i into itself (the former also maps U_i into itself). Then, the following Proposition, whose proof depends on a contraction mapping argument (see Appendix B), provides a set of necessary and sufficient conditions for existence of the unique equilibrium solution alluded to in Prop. 2.

Theorem 1. (i) Under Conditions (1) and (2), the decision problem of Section 2 admits a unique stable Nash equilibrium solution given by (15) if, and only if, there exists, for at least one $i=1,2$, a ρ^i , $0 < \rho^i < 1$, such that

$$r_i(S_i) = r_i(\bar{D}^i \bar{P}_{i|i}) \leq \rho^i \quad (19)$$

(ii) A set of sufficient conditions for (19) to hold true is the existence of a pair of positive scalars (ρ_1^i, ρ_2^i) , such that

$$\rho_1^i \rho_2^i < 1, \quad r_i(\bar{D}^i) \leq \rho_1^i, \quad r_i(\bar{P}_{i|i}) \leq \rho_2^i \quad (20a)$$

Furthermore, a set of sufficient conditions for the latter two is

$$\langle \langle \bar{D}^i \rangle \rangle_i \equiv \|\bar{D}^i\|_i \leq \rho_1^i, \quad \langle \langle \bar{P}_{i|i} \rangle \rangle_i \leq \rho_2^i \quad (20b)$$

where $\|\cdot\|_i$ denotes the operator norm on U_i , as a counterpart of (17a).

Proof. See Appendix B. □

Part (ii) of Thm. 1 provides a partial separation (in terms of sufficient conditions) of the deterministic and stochastic parts of the system. Now, if the decision problem is a team problem with a common loss functional [which requires $D_{22}^1 = I$, $D_{12}^1 = D_{21}^{2*}$, $F_1^1 = F_1^2$ and $F_2^1 = F_2^2$], and if team cost is strictly

convex in the pair (u_1, u_2) [which is true if and only if $\|D_{12}^1 D_{12}^{1*}\|_1 = \bar{\rho} < 1$], it follows that the first inequality holds with $\rho_1^1 = \rho_1^2 < 1$. If, furthermore, the subjective probability measures assigned to the pair (y_1, y_2) by the two DM's are equivalent, $\bar{P}_{i|i}$ becomes the product of two projection operators, thus leading to satisfaction of the second inequality in (20b) with $\rho_2^1 = \rho_2^2 = 1$, and thereby to satisfaction of (20a). Hence, as a corollary to the second part of Prop. 3, we obtain the following result which is known in different contexts [7,8,9].[†]

Corollary 1. For the strictly convex quadratic team problem with equivalent subjective probability measures assigned by the two DM's to (y_1, y_2) , there exists a unique stable equilibrium solution (the so-called team-optimal solution), irrespective of the underlying common probability measure. $\square$

For team problems with $P^1 \neq P^2$, a result along the lines of Corollary 1 does not in general hold, because the operator $P_{i|i}$ is not necessarily the product of two projection operators. Then, the general condition is (19) [or the stronger one, (20a)], which places some restrictions on the parameters of the cost functional, as well as the probability measures P^1 and P^2 . To delineate the extent of these restrictions, we now study the second inequality of (20b) somewhat further and obtain the following sufficient condition.

Corollary 2. For a given ρ_2^1 , the second inequality of (20b) is satisfied if the expression

$$g^i(y_i) E^j [g^j(y_j) | y_i] = g^i(y_i) \int_{Y_j} g^j(n) P_{y_j|y_i}^j (dn | \xi = y_i) \quad (21a)$$

is uniformly bounded from above by $(\rho_2^1)^2$ a.e. $P_{y_i}^i$. Furthermore, if the probability measures P^1 and P^2 are absolutely continuous with respect to the Lebesgue measure, this condition can be expressed equivalently in terms of the probability densities $p^i(y_i, y_j)$ as follows:

[†]This result is slightly more general than the related ones that can be found in [7,8,9], since here P_X^1 is allowed to be different from P_X^2 , though still a restriction is imposed on these (indirectly) via the equivalence between P_{y_1, y_2}^1 and P_{y_1, y_2}^2 .

$$\frac{p_i^j(y_i)}{p_{y_i}^i(y_i)} \int_{Y_j} [p_{y_j}^i(n) p_{y_j|y_i}^j(n|y_i) / p_{y_j}^j(n)] dn \leq (\varepsilon_2^i)^2 \quad (21b)$$

Proof. For (21a) see Appendix C; (21b) follows readily from (21a). $\square$

4. General Sufficient Conditions for a Stackelberg Equilibrium Solution

We now turn our attention to the asymmetric mode of decision making, obtain some general sufficient conditions for existence of a Stackelberg equilibrium solution, and provide a complete characterization of the solution. Subsequently we consider some special cases with some further structure imposed on the cost functions and the probability measures.

Firstly we obtain an expression for DM2's unique reaction $T_2: \Gamma_1 \rightarrow \Gamma_2$, as defined by (10), using Prop. 1:

$$T_2[y_1] = y_2^0(y_1) = D_{21}^2 E^2[y_1(y_1)|y_2] + F_2^2 E^2[x|y_2] \quad (22)$$

Hence, the derivation of the leader's Stackelberg policy $\gamma_1^s \in \Gamma_1$ involves (in view of (11)) the minimization of J_1 over Γ_1 after γ_2^0 given by (22) is substituted in. This substitution yields

$$\begin{aligned} \tilde{J}(\gamma) \triangleq J_1(\gamma, \gamma_2^0) &= \frac{1}{2} \langle \gamma, \gamma \rangle_1 + \frac{1}{2} \int_{Y_2} (F_2^2 E^2[x|y_2] \\ &+ D_{21}^2 E^2[\gamma(y_1)|y_2], D_{22}^1 D_{21}^2 E^2[\gamma(y_1)|y_2] + D_{22}^1 F_2^2 E^2[x|y_2])_2 P_{y_2}^1(d\xi) \\ &- \langle \gamma, E^1[F_1^1 x|y_1] \rangle_1 + \int_{X \times Y_2} (D_{21}^2 E^2[\gamma(y_1)|y_2] + F_2^2 E^2[x|y_2], F_2^1 x)_2 \\ &\quad \cdot P^1(dx, Y_1, d\xi) \\ &- \langle \gamma, E^1[D_{12}^1 D_{21}^2 E^2[\gamma(y_1)|y_2]|y_1] + E^1[D_{12}^1 F_2^2 E^2[x|y_2]|y_1] \rangle_1, \end{aligned} \quad (23)$$

where we have deleted the subscript 1 in γ_1 in order to simplify the notation.

Now, since Γ_1 is a linear space, and $\tilde{J}$ is the sum of terms homogeneous of degree zero, one and two (maximum), any minimizing solution $\gamma_1^s \in \Gamma_1$ will have to satisfy

$$\Delta \tilde{J}(\gamma; h) = \tilde{J}(\gamma+h) - \tilde{J}(\gamma) = \delta \tilde{J}(\gamma; h) + \delta^2 \tilde{J}(\gamma; h) \geq 0 \quad \forall h \in \Gamma_1, \quad (24)$$

where $\delta^1 \tilde{J}(\gamma; h)$ is the Gateaux variation of $\tilde{J}(\gamma)$ of degree 1. Extensive manipulations, details of which are given in Appendix D (subsection 1), lead to the following expressions for $\delta \tilde{J}$ and $\delta^2 \tilde{J}$:

$$\begin{aligned} \delta \tilde{J}(\gamma; h) = & \langle h, \gamma \rangle_1 - \int_{Y_1} (h(y_1), (\Xi \gamma)(y_1))_1 P_{y_1}^1(dy_1) \\ & - \int_{Y_1} (h(y_1), \beta(y_1))_1 P_{y_1}^1(dy_1) \end{aligned} \quad (25)$$

$$\begin{aligned} \delta^2 \tilde{J}(\gamma; h) = & \frac{1}{2} \langle h, h \rangle_1 + \frac{1}{2} \int_{Y_1} (h(y_1), g^1(y_1) E^2[g^2(y_2) D_{21}^{2*} D_{22}^1 D_{21}^2 \\ & \cdot E^2[h(\xi)|y_2||y_1])_1 P_{y_1}^1(dy_1) - \langle h, D_{12}^1 D_{21}^2 \bar{P}_1 | h \rangle_1 \end{aligned} \quad (26)$$

where $\Xi: \Gamma_1 \rightarrow \Gamma_1$ and $\beta \in \Gamma_1$ are defined by

$$\begin{aligned} (\Xi \gamma)(y_1) = & (D_{12}^1 D_{21}^2 P_{1|1} + D_{21}^{2*} D_{12}^1 P_{1|1}^*) \gamma(y_1) \\ & - D_{21}^{2*} D_{22}^1 D_{21}^2 g^1(y_1) E^2[g^2(y_2) E^2[\gamma(y_1)|y_2||y_1]] \end{aligned} \quad (27a)$$

$$\begin{aligned} \beta(y_1) = & F_1^1 E[x|y_1] - D_{21}^{2*} D_{22}^1 F_2^2 g^1(y_1) E^2[g^2(y_2) E^2[x|y_2||y_1]] \\ & - D_{12}^1 F_2^2 E^1[E^2[x|y_2||y_1]] + D_{21}^{2*} F_2^1 g^1(y_1) E^2[g^2(y_2) E^1[x|y_2||y_1]]. \end{aligned} \quad (27b)$$

$P_{1|1}: \mathcal{U}_1 \rightarrow \mathcal{U}_1$ is a linear operator given by

$$P_{1|1} \gamma(y_1) = E^1[E^2[\gamma(y_1)|y_2||y_1]], \quad (28)$$

[†] Here $\delta^1 \tilde{J}$ is written simply as $\delta \tilde{J}$.

U_1 is the space of y_1 -measurable random variables taking values in U_1 , and $g^1(\xi)$ are the R-N derivatives (2). Note that $P_{1|1}$ is related to $\bar{P}_{1|1}$ defined by (18b) by

$$P_{1|1}[\gamma(y_1)] = (\bar{P}_{1|1}\gamma)(y_1)$$

where the latter (which is a mapping from Γ_1 into Γ_1) has been used in (26) and will also be used in the sequel whenever needed.

Now, since (24) is also equivalent to

$$\left. \begin{aligned} \delta \tilde{J}(\gamma, h) &= 0 & \forall h \in \Gamma_1 \\ \delta^2 \tilde{J}(\gamma, h) &\geq 0 & \forall h \in \Gamma_1 \end{aligned} \right\}, \quad (29)$$

a Stackelberg solution $\gamma \in \Gamma_1$ will exist for the leader if, and only if,

(i) (26) is nonnegative definite,

and (from (25)):

$$(ii) \gamma(y_1) - (E\gamma)(y_1) - \beta(y_1) = 0, \quad \text{a.e. } P_{y_1}^1. \quad (30)$$

Since the first of these conditions does not depend on γ , the optimal solution is solely determined by (30), which can be rewritten as

$$\begin{aligned} \gamma(y_1) &= D_{12}^1 D_{21}^2 E^1[E^2[\gamma(y_1)|y_2]|y_1] + D_{21}^{2*} D_{12}^{1*} g^1(y_1) E^2[g^2(y_2) E^1[\gamma(y_1)|y_2]|y_1] \\ &+ D_{21}^{2*} F_2^1 g^1(y_1) E^2[g^2(y_2) E^1[x|y_2]|y_1] - D_{21}^{2*} D_{22}^1 D_{21}^2 g^1(y_1) E^2[g^2(y_2) E^2[\gamma(y_1)|y_2]|y_1] \\ &+ F_1^1 E^1[x|y_1] - D_{21}^{2*} D_{22}^1 F_2^2 g^1(y_1) E^2[g^2(y_2) E^2[x|y_2]|y_1] + D_{12}^1 F_2^2 E^1[E^2[x|y_2]|y_1], \end{aligned} \quad (31)$$

where we have utilized the fact that the adjoint of $P_{1|1}$ is a linear operator

$P_{1|1}^*$: $U_1 \rightarrow U_1$, given by [see Appendix D, subsection 2]

$$P_{1,1}^*(y_1) = \int_{Y_1} \int_{Y_2} \frac{P_{y_1 y_2}^1(dx dy_2) P_{y_1 y_2}^2(dy_1 dx dy_2)}{P_{y_2}^2(dy_2) P_{y_1}^1(dy_1)} = g^1(y_1) E^2[g^2(y_2) E^1[\gamma(y_1)|y_2]|y_1] \quad (32)$$

Furthermore, condition (i) can be rewritten as

$$A \triangleq I + \frac{1}{2} D_{21}^{2*} D_{22}^1 D_{21}^2 (K+K^*) - D_{12}^1 D_{21}^2 \bar{P}_{1/1} - D_{21}^{2*} D_{12}^1 P_{1/1}^* \geq 0 \quad (33)$$

where $I: \Gamma_1 \rightarrow \Gamma_1$ is the identity operator, and $K: \Gamma_1 \rightarrow \Gamma_1$ is defined by

$$(KY)(y_1) = g^1(y_1) E^2[g^2(y_2) E^2[\gamma(y_1)|y_2]|y_1] \quad (34)$$

We now summarize these results in the following proposition:

Proposition 3. Under Conditions (1) and (2), the decision problem with multiple probability measures admits a Stackelberg equilibrium solution if, and only if, A is nonnegative definite and (31) admits a solution in Γ_1 .

Equation (31) will, in general, not admit a closed-form solution, even if all random variables are jointly Gaussian distributed (see §5.3); therefore, we will have to resort to numerical computations which will involve a recursion of some type. Hence, in analyzing the conditions of existence of a solution to (31) we may also require that such a numerical scheme be globally convergent (or stable). One appealing scheme whereby a unique solution to (31) [or, equivalently, (30)] can be obtained is the recursion

$$\gamma^{(k)}(y_1) = (\bar{E}\gamma^{(k-1)})(y_1) + \beta(y_1), \quad k=1,2,\dots \quad (35)$$

where $\gamma^{(0)}$ is chosen as an arbitrary element of Γ_1 . If the limit $\lim_{k \rightarrow \infty} \gamma^{(k)} \triangleq \gamma^s$ exists in Γ_1 , for all such initial choices, then γ^s will necessarily constitute a solution to (31). A sufficient condition for this readily follows from

Lemma B.1, which we give below as Prop. 4.

Proposition 4. In addition to the conditions of Prop. 3, assume that there exists a scalar ρ , $0 < \rho < 1$, such that

$$r(\tilde{Z}) \leq \rho \quad (36)$$

where $r(\tilde{Z})$ is the spectral radius of $\tilde{Z}$. Then, the decision problem admits a unique Stackelberg equilibrium solution $(\gamma^S, T_2[\gamma^S])$, where $\gamma^S \in \Gamma_1$ is the limit of the iterative scheme (35), and T_2 is the affine operator (22). $\square$

We now further elaborate on (36), so as to bring it to a form which separates out the contributions from the deterministic and probabilistic components of the problem. [Here, we are seeking sufficient conditions which would constitute the counterpart of (20) in this context]. Towards this end, let us first note that using (34) in (25a):

$$r(\tilde{Z}) = r(D_{12}^1 D_{21}^2 \bar{P}_{1|1} + D_{21}^{2*} D_{12}^{1*} \bar{P}_{1|1}^* - D_{21}^{2*} D_{22}^1 D_{21}^2 K) \quad (37)$$

and utilizing the inequality relationship between the spectral radius and norm of an operator (see Appendix A, Lemma A.1) this can be bounded from above by

$$\leq \langle\langle D_{12}^1 D_{21}^2 \bar{P}_{1|1} + D_{21}^{2*} D_{12}^{1*} \bar{P}_{1|1}^* - D_{21}^{2*} D_{22}^1 D_{21}^2 K \rangle\rangle_1$$

where $\langle\langle \cdot \rangle\rangle_1$ is the operator norm as defined in (17a). Using the standard (triangle inequality) property of norms, this can further be bounded from above by

$$\leq \langle\langle D_{12}^1 D_{21}^2 \bar{P}_{1|1} + D_{21}^{2*} D_{12}^{1*} \bar{P}_{1|1}^* \rangle\rangle_1 + \langle\langle D_{21}^{2*} D_{22}^1 D_{21}^2 K \rangle\rangle_1$$

Now since both $D_{21}^{2*} D_{22}^1 D_{21}^2$ and K map a Hilbert space (Γ_1) into itself, using the norm inequality for products of linear operators, we further have

$$\begin{aligned} &\leq \langle\langle D_{12}^1 D_{21}^2 \bar{P}_{1|1} + D_{21}^{2*} D_{12}^{1*} \bar{P}_{1|1}^* \rangle\rangle_1 + \langle\langle D_{21}^{2*} D_{22}^1 D_{21}^2 \rangle\rangle_1 \langle\langle K \rangle\rangle_1 \\ &= r(D_{12}^1 D_{21}^2 \bar{P}_{1|1} + D_{21}^{2*} D_{12}^{1*} \bar{P}_{1|1}^*) + r(D_{21}^{2*} D_{22}^1 D_{21}^2) [r(K^* K)]^{1/2} \end{aligned}$$

where the equality follows because (i) the spectral radius and norm of a self-adjoint linear operator are equal [13,p.514], (ii) norm of a "non-self-adjoint" linear operator K is equal to the square root of the spectral radius of the self-adjoint operator K^*K (see Appendix A, Lemma A.1). Finally, using the result of Lemma A.2 (Appendix A), the latter is bounded from above by

$$r(\Xi) \leq 2[(D_{12}^1 D_{21}^2 D_{21}^{2*} D_{12}^{1*})]^{1/2} [r(\bar{P}_1^* \bar{P}_1)]^{1/2} + r(D_{21}^{2*} D_{22}^1 D_{21}^2) [r(K^*K)]^{1/2}. \quad (38)$$

Now, let us assume the following:

Condition (3). There exist four positive scalars $\rho_1, \rho_2, \rho_3, \rho_4$, satisfying

$$2\rho_1\rho_2 + \rho_3\rho_4 < 1 \quad (39)$$

such that

$$r(D_{12}^1 D_{21}^2 D_{21}^{2*} D_{12}^{1*}) \leq (\rho_1)^2, \quad r(D_{21}^{2*} D_{22}^1 D_{21}^2) \leq \rho_3 \quad (40a)$$

$$r(\bar{P}_1^* \bar{P}_1) \leq (\rho_2)^2, \quad r(K^*K) \leq (\rho_4)^2. \quad (40b)$$

Then, we have

Theorem 2. Under *Conditions (1)-(2)* of §2 and *Condition (3)* given above, the decision problem admits a unique Stackelberg equilibrium solution $(\gamma^S, T_2[\gamma^S])$, where $\gamma^S \in \Gamma_1$ is the limit of the iterative scheme (35), and T_2 is given by (22).

Proof. The result follows from Prop. 4 and the discussion and derivation that leads to *Condition (3)*, provided we show that the given three conditions subsume (33), i.e. nonnegativity of operator A . We now verify that *Condition (3)* in fact implies that A is a strongly positive operator. First note that A is self-adjoint, because K commutes with $D_{21}^{2*} D_{22}^1 D_{21}^2$. Hence, using Lemma A.3 (Appendix A), we can write down the inequality

$$r(A-I) \leq \frac{1}{2} r(D_{21}^{2*} D_{22}^1 D_{21}^2 (K+K^*)) + r(D_{12}^1 D_{21}^2 \bar{P}_1^* \bar{P}_1 + D_{21}^{2*} D_{12}^1 P_1^* P_1).$$

Then, using the line of arguments that led to (38) from (37), and the spectral radius inequality for the product of two self-adjoint operators, we obtain the bound

$$\begin{aligned} r(A-I) &\leq \frac{1}{2} r(D_{21}^2 D_{22}^1 D_{21}^2) r(K+K^*) + r[(D_{12}^1 D_{21}^2 D_{21}^{2*} D_{12}^{1*})]^{1/2} [r(\bar{P}_{11}^* \bar{P}_{11})]^{1/2} \\ &\leq \frac{1}{2} \rho_3 r(K+K^*) + \rho_1 \rho_2 . \end{aligned}$$

But note that

$$r(K+K^*) = \sup_{\gamma \in \Gamma_1} [\langle \gamma, (K+K^*) \gamma \rangle_1 | \langle \gamma, \gamma \rangle_1 |] = 2 \sup_{\gamma \in \Gamma_1} [\langle \gamma, K \gamma \rangle_1 | \langle \gamma, \gamma \rangle_1 |] ,$$

and since, from the Cauchy-Schwarz inequality of inner products,

$$|\langle \gamma, K \gamma \rangle_1|^2 \leq |\langle \gamma, \gamma \rangle_1| |\langle K \gamma, K \gamma \rangle_1| ,$$

$$\text{we have } r(K+K^*) \leq 2 \sup_{\gamma \in \Gamma_1} [\langle K \gamma, K \gamma \rangle_1 | \langle \gamma, \gamma \rangle_1 |]^{1/2}$$

$$= 2 \sup_{\gamma \in \Gamma_1} [\langle \gamma, K^* K \gamma \rangle_1 | \langle \gamma, \gamma \rangle_1 |]^{1/2} = 2 [r(K^* K)]^{1/2} \leq 2 \rho_4 .$$

$$\text{Thus, } r(A-I) \leq \rho_3 \rho_4 + \rho_1 \rho_2 < 1 ,$$

implying that the spectrum of the self-adjoint operator $A-I$ is uniformly in the unit sphere. Hence, A is strongly positive. $\square$

For the special class of strictly convex team problems (cf. §3) with multiple probability measures, several simplifications can be made. In this case eq. (31) simplifies to

$$\begin{aligned} \gamma(y_1) &= D_{12}^1 D_{12}^{1*} \{ E^1[E^2[\gamma(y_1)|y_2] + g^1(y_1)E^2[g^2(y_2) \{ E^1[\gamma(y_1)|y_2] \\ &\quad - E^2[\gamma(y_1)|y_2] \} |y_1]] + F_1^1 E^1[x|y_1] \} \\ &\quad + D_{12}^1 F_2^1 g^1(y_1) E^2[g^2(y_2) \{ E^1[x|y_2] - E^2[x|y_2] \} |y_1] + D_{12}^1 F_2^1 E^1[E^2[x|y_2]|y_1] , \end{aligned} \quad (41)$$

and in Condition (3) inequalities (40a) are replaced by the single inequality

$$[r(D_{12}^1 D_{12}^{1*} D_{12}^1 D_{12}^{1*})]^{1/2} = \langle\langle D_{12}^1 D_{12}^{1*} \rangle\rangle_1 = \bar{\rho} \leq \rho_1 = \rho_3$$

where ρ_1 can be taken to be less than one. Hence, (39) reads

$$(2\rho_2 + \rho_4) < 1/\bar{\rho} \quad (42)$$

We now summarize these results as a corollary to Thm. 2:

Corollary 3. Under Conditions (1)-(2) of §2, and (42) given above, the strictly convex quadratic team problem, with multiple probability measures and asymmetric mode of decision making, admits a unique Stackelberg equilibrium solution

$(\gamma^s, T_2[\gamma^s])$, where $\gamma^s \in \Gamma_1$ is the limit of the iterative scheme (35) with

$$(\bar{z}\gamma)(y_1) = D_{12}^1 D_{12}^{1*} [(p_{1|1} + p_{1|1}^*)\gamma(y_1) - g^1(y_1)E^2[g^2(y_2)E^2[\gamma(y_1)|y_2]|y_1]],$$

and T_2 is given by (22). □

Remark 3. When the original problem is a Stackelberg game, but the probability measures are identical, a study of the original condition (36) reveals the inequality

$$r(\bar{z}) \leq r(D_{12}^1 D_{21}^2 + D_{21}^2 D_{12}^{1*} - D_{21}^2 D_{22}^1 D_{21}^2) \leq \rho < 1.$$

This is the existence condition associated with the standard stochastic Stackelberg game, which corroborates the earlier result obtained in [25]. □

We now conclude this section by presenting the counterpart of Corollary 2 in the present context, which provides a set of (simpler) sufficient conditions for (40b) to be satisfied:

Corollary 4. For a given pair (ρ_2, ρ_4) , the first and second inequalities of (40b) are satisfied if, respectively,

$$g^1(y_1) \int_{Y_2} g^2(\eta) P_{y_2|y_1}^2(d\eta|\xi = y_1) \quad (43a)$$

$$\text{and} \quad g^1(y_1) \int_{Y_2} |g^2(\eta)|^2 P_{y_2|y_1}^2(d\eta|\xi = y_1) \int_{Y_1} g^1(b) P_{y_1|y_2}^2(db|y_2 = \eta) \quad (43b)$$

are uniformly bounded from above by $(\rho_2)^2$ and $(\rho_4)^2$.

Furthermore, if probability densities exist (with respect to the Lebesgue measure), these conditions can be expressed in terms of the corresponding probability density functions $p_{y_1 y_2}^i(\cdot)$ as follows:

$$\frac{p_{y_1}^2(y_1)}{p_{y_1}^1(y_1)} \int_{Y_2} \frac{p_{y_2}^1(n)}{p_{y_2}^2(n)} p_{y_2|y_1}^2(n|y_1) dn \leq (\rho_2)^2 \quad (44a)$$

$$\frac{p_{y_1}^2(y_1)}{p_{y_1}^1(y_1)} \int_{Y_2} \left| \frac{p_{y_2}^1(n)}{p_{y_2}^2(n)} \right|^2 p_{y_2|y_1}^2(n|y_1) dn \int_{Y_1} \frac{p_{y_1}^2(b)}{p_{y_1}^1(b)} p_{y_1|y_2}^2(b|n) db \leq (\rho_4)^2 \quad (44b)$$

Proof. For (43a)-(43b) see Appendix C; (44a)-(44b), however, follow readily from (43a)-(43b). □

5. Jointly Gaussian Distributions

In decision and control theory, one appealing class of probability distributions is the Gaussian distribution, because it leads to tractable problems admitting, in most cases, closed-form solutions. Indeed when the probability measures of the two DM's are identical and Gaussian, equilibrium solutions have been shown to be affine functions of the observations for (i) quadratic stochastic team problems defined on Euclidean spaces [7], (ii) quadratic stochastic Nash games on Euclidean spaces [8], (iii) quadratic continuous-time stochastic team problems [9], (iv) quadratic stochastic Stackelberg games on Euclidean spaces [25], and (v) quadratic continuous-time stochastic Stackelberg games [26]. In this section, we investigate possible extensions of this appealing structural feature to the case when discrepancies exist between the subjective Gaussian distributions, as reflected in the covariances of the random vectors (y_1, y_2) . We could also have included discrepancies in the perceptions of the mean values, but such a more general treatment does not contribute substantially to the qualitative nature of the results obtained in the sequel, and besides it makes the expressions notationally cumbersome. Interested reader could find relevant expressions for the nonzero mean case in [27].

We first introduce notation and terminology, and delineate *Conditions (1) and (2)* of §2 (§5.1). Then, we study the case of symmetric mode of decision making in §5.2, and show that the unique equilibrium solution of Thm. 1 is linear. Finally in §5.3 we treat the case of asymmetric mode of decision making, and show that (in contradistinction with the result of §5.2) the unique Stackelberg solution of Thm. 2 is generically nonlinear.

5.1. *Notation and Terminology*

Let (x, y_1, y_2) be zero-mean Gaussian random vectors under both P^1 and P^2 , with

$$\text{covariance } (y_1, y_2) = \Sigma_y^i = \begin{pmatrix} \Sigma_{y_1}^i & \Sigma_{y_1 y_2}^i \\ \Sigma_{y_2 y_1}^i & \Sigma_{y_2}^i \end{pmatrix} > 0, \text{ under } P^i; \quad (45a)$$

$$\text{covariance } (x, y_1, y_2) = \text{cov}(x, y) = \Sigma^i = \begin{pmatrix} \Sigma_x^i & \Sigma_{xy}^i \\ \Sigma_{yx}^i & \Sigma_y^i \end{pmatrix} > 0 \text{ under } P^i. \quad (45b)$$

These probability distributions clearly satisfy the absolute continuity condition (*Condition (1)*) of Thms. 1 and 2. Furthermore, since

$$g^i(\xi) = (\det \Sigma_{y_i}^i / \det \Sigma_{y_i}^j) \exp \left\{ -\frac{1}{2} \xi^T W_i \xi \right\} \quad (46a)$$

$$W_i \triangleq \Sigma_{y_i}^{j-1} - \Sigma_{y_i}^{i-1}, \quad j \neq i, \quad (46b)$$

the uniform boundedness condition (*Condition (2)*) of Thms. 1 and 2 is satisfied whenever

$$W_i \geq 0, \quad i=1,2. \quad (47)$$

After making these observations, let us introduce the additional notation

$$N_i \triangleq M_{ii}^j - M_{ij}^j B_j^{-1} M_{ji}^j - \Sigma_{y_i}^{i-1}, \quad j \neq i \quad (48a)$$

$$B_j = M_{jj}^j + W_j \quad (48b)$$

$$\begin{pmatrix} M_{11}^i & M_{12}^i \\ M_{21}^i & M_{22}^i \end{pmatrix} \triangleq \begin{pmatrix} \Sigma_{y_1}^i & \Sigma_{y_1 y_2}^i \\ \Sigma_{y_2 y_1}^i & \Sigma_{y_2}^i \end{pmatrix}^{-1} \quad (48c)$$

$$q^i \triangleq [\det \Sigma_{y_i}^i \cdot \det \Sigma_{y_j}^j / \det \Sigma_{y_j}^i \cdot \det B_j \cdot \det \Sigma_y^i]^{1/2} \quad (48d)$$

in terms of which we evaluate (21a) [using standard properties of Gaussian distributions] to be

$$g^i(y_1) E^j[g^j(y_j) | y_1] = q^i \exp \left\{ -\frac{1}{2} y_1^T D_i y_1 \right\} \quad (49)$$

We are now in a position to specialize the results of Thms. 1 and 2 to Gaussian distributions and obtain some explicit results.

5.2. Symmetric Mode of Decision Making

In order to apply Thm. 1 to the Gaussian decision problem formulated above, we first explore the satisfaction of various conditions given there. We have already shown above that *Condition (1)* is always satisfied and *Condition (2)* is satisfied whenever $W_i \geq 0$. For the remaining condition we study inequalities (20b). The second of these is satisfied, for a given ρ_2^i , if (using (21a)) expression (49) is uniformly bounded in y_i , and this bound is no greater than ρ_2^i . For uniform boundedness of (49) it is necessary and sufficient that

$$N_i \geq 0, \quad (50a)$$

under which the latter condition becomes

$$q^i \leq (\rho_2^i)^2.$$

Hence going back to (20a), the condition

$$\|D_{ij}^i D_{ji}^j\|_i^2 < 1/q^i, \text{ for at least one } i=1,2, \quad (50b)$$

becomes sufficient for (19). We are now in a position to state and prove the following theorem:

Theorem 3. Let (47) hold for $i=1,2$, and (50a)-(50b) hold for at least one i . Then, the quadratic Gaussian decision problem formulated in this section admits a unique stable Nash equilibrium solution (γ_1^0, γ_2^0) , where $u_i^0 = \gamma_i^0(y_i)$ are linear in y_i , and are given by

$$\gamma_i^0(y_i) = L_i y_i \quad i=1,2. \quad (51)$$

Here, $L_i: \mathbb{R}^{m_i} \rightarrow U^i$ are bounded linear operators, constituting the unique solution to the linear operator equations

$$L_i y_i - D_{ij}^i D_{ji}^j L_i y_j = D_{ij}^i D_{ji}^j L_i y_j - D_{ij}^i D_{ji}^j L_i y_j, \quad (52)$$

$$- D_{ij}^i F_{jxy}^j L_i y_j = F_{ixy}^i L_i y_i = 0, \quad \forall y_i \in \mathbb{R}^{m_i}, \quad i=1,2.$$

Proof. The existence and uniqueness of the solution follows from Thm. 1, Corollary 2, and the discussion that precedes the statement of the theorem. The linearity of this unique solution, on the other hand, follows by noting that if the pair $(\gamma_1^{(0)}, \gamma_2^{(0)})$ is taken to be linear in (y_1, y_2) in (14), all the terms of the sequence are linear, and hence the limit (which exists as already proven) is linear. Hence, choosing γ_i as in (51), where $L_i: \mathbb{R}^{m_i} \rightarrow U_i$ are bounded linear operators, substituting this into (14) and requiring it to hold for all $y_i \in \mathbb{R}^{m_i}$ (since all probability measures are Gaussian) leads to the unique relations (52). $\square$

Remark 4. Thm. 3 above extends the result of Thm. 2 of [8] on quadratic Gaussian games to the case when a common probability space does not exist and the decision spaces are not necessarily finite dimensional, and shows that the appealing linear structure prevails when there exists a discrepancy in the perceptions of the two DM's of the underlying probability measures. The existence and uniqueness conditions here are, however, more restrictive than those of [8], and also involve the probabilistic structure (see (50b)). Expression (21a) in the most general case (and (49) for the special Gaussian case) is not uniformly (in y^i) bounded by 1, unless $g^i(y_i) = g^j(y_j) = 1$ a.e. $P_{y_i}^i$ and $P_{y_j}^j$ (which corresponds to the case of equivalent probability measures), since R-N derivatives (if different from 1) will be both smaller and larger than unity on sets of nonzero measure. This then implies, in view of (47), and from (49), that $q_i^i \geq 1$, $i=1,2$, with the inequality being strict if $P_{y_i}^i$ is not equivalent to $P_{y_i}^j$ for a least one $i=1,2$, $j \neq i$. In such a case, even team problems, a stable equilibrium solution may not exist, particularly if $1/q_i < \|D_{ij}^i D_{ji}^j\|_i^2 < 1$ for at least one $i=1,2$; $j \neq i$. This indicates, in general, the presence of a strong coupling between probabilistic and deterministic elements of the problem in terms of existence conditions. However, if the discrepancy between perceptions of the DM's on the probability measures (measured in terms of R-N derivatives) is sufficiently small, one would expect q_i to be sufficiently close to unity, which ensures satisfaction of condition (50c) for a fairly general

class of quadratic strictly convex Gaussian team problems (since, $\|D_{ij}^j D_{ji}^j\|_i = \|D_{ji}^j D_{ij}^i\|_j = \bar{\epsilon} < 1$, for such team problems). For further discussion on this point we refer the reader to [10].

In the statement of Thm. 1, the condition (47) places some severe restrictions on the second moments of the underlying distributions (in case a discrepancy exists), which may however be relaxed if we are willing to consider equilibrium policies in a more restricted space. More specifically, satisfaction of (47) ensures that regardless of what initial set of policies the DM's start the infinite recursion (15) with, every element of this series is well-defined, and under (50a)-(50b) it will converge to a unique limit which is linear; in other words, even if the DM's start with nonlinear policies, the end result will be a linear equilibrium solution. The condition (47) is restrictive, because we require (without imposing any constraints on the policy spaces) the series generated by (15) to be well-defined even with nonlinear starting conditions. However, if we restrict the team agents to linear policies from the outset, under Gaussian distributions (and following the argument used in the proof of Thm. 1) elements of the series (15) will be well-defined (without requiring (47)) and will converge to the equilibrium solution provided that (50a)-(50b) hold for at least one $i=1,2$. This line of reasoning then leads to the following result which we give without a proof.

Proposition 5. Let Γ_i^a be the class of all linear policies in the form (51), with $L_i: \mathbb{R}^{m_i} \rightarrow U^i$ a bounded linear operator, $i=1,2$. On $\Gamma_1^a \times \Gamma_2^a$, the statement of Thm. 1 is valid even if (47) does not hold true.

We now interpret these results in the context of two examples one of which is a scalar team problem and the other one is a continuous-time team problem, both with multiple subjective Gaussian probabilities.

Example 1. Consider a family of scalar Gaussian team problems, with

$D_{22}^1 = D_{11}^2 = 1$, $D_{12}^1 = D_{21}^2 = \bar{x}$, $|\bar{x}| < 1$, $F_1^1 = f_1$, $F_2^2 = f_2$, $n=m_1=m_2=1$, and

$$\Sigma_y^1 = \begin{pmatrix} a & e \\ e & \eta b \end{pmatrix} ; \quad \Sigma_y^2 = \begin{pmatrix} \mu a & \beta \\ \beta & b \end{pmatrix} , \quad \mu a b > e^2, \quad \mu a b > \beta^2. \quad (53)$$

To investigate the applicability of Thm. 1 to this class of problems, let us first observe that condition (47) is satisfied if, and only if, both

$$0 < \mu \leq 1, \quad 0 < \eta \leq 1. \quad (54)$$

For condition (50a), we evaluate N_1 and require it to be nonnegative for either $i=1$, or $i=2$:

$$N_1 = (\mu ab - c^2)(1-\mu)[\mu a^2 - c^2 - (1-\mu)ab] / \{a[\mu^2 a^2 - (1-\mu)(\mu ab - c^2)]\} \geq 0 \quad (55a)$$

or

$$N_2 = (\eta ab - c^2)(1-\eta)[\eta b^2 - c^2 - (1-\eta)ab] / \{b[\eta^2 b^2 - (1-\eta)(\eta ab - c^2)]\} \geq 0, \quad (55b)$$

Finally, condition (50b) dictates either

$$\mu a^2 |a|^2 < \eta[\mu^2 a^2 - (1-\mu)(\mu ab - c^2)] \quad (56a)$$

or

$$\eta b^2 |a|^2 < \mu[\eta^2 b^2 - (1-\eta)(\eta ab - c^2)] \quad (56b)$$

provided that the terms on the right-hand-side are positive (if not, then the inequalities will accordingly change direction).

The set of values for $a, b, c, \varepsilon, \mu, \eta$ that satisfy (54)-(56) is clearly not empty. To gain some further insight into these conditions, let us consider the class of team decision problems in which the discrepancies between the DMs' perceptions of the variances of different Gaussian random variables is relatively small, that is there exist sufficiently small $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ such that $\mu = 1 - \varepsilon_1$, $\eta = 1 - \varepsilon_2$, and furthermore $c = 0$, and $|c|$ is considerably smaller than both a and b . Note that, when $\varepsilon_1 = \varepsilon_2 = 0$, conditions (54)-(56) are all satisfied (note that $|a| < 1$ because of strict convexity of the objective functional) regardless of the relative magnitudes of ε and η . Hence, when the discrepancy is only in the perceptions of the correlation between y_1 and y_2 , the scalar quadratic Gaussian team problem always admits a stable equilibrium solution. Now, for nonzero, but positive, and sufficiently small ε_1 , the dominating term in (55a) is

$$N_1 \sim \epsilon_1 (\mu ab - c^2) (\mu a^2 - c^2) / \mu^2 a^3$$

which is positive, in view of (53) and the initial hypothesis that $|a/c| \gg 1$. Likewise, D_2 is positive whenever $0 < \epsilon_2 < 1$ and $|b/c| \gg 1$. Furthermore, given a $\bar{d}$, $0 < \bar{d} < 1$, we can always find ϵ_1 and ϵ_2 , both in $(0,1)$, so that both (56a) and (56b) are satisfied whenever $|\bar{d}| < \bar{d}$. Hence, the conclusion is that when the deviations of the perceptions of the DM's from the common Gaussian probability measures are incremental (and satisfying (54)), the linear equilibrium solution of the Gaussian scalar team problem retains its stability property (but, of course, at a different (possibly close, in norm) equilibrium point). $\square$

Example 2. As a second illustration of Thm. 1, for infinite-dimensional decision spaces we consider here a class of stochastic Gaussian team problems defined in continuous time. More specifically, let $U_1 = U_2 = \mathcal{L}_2(0,T)$, the Hilbert space of all scalar-valued Lebesgue-integrable functions on the bounded interval $[0,T]$, endowed with the standard inner product $\int_0^T u(t)v(t)dt$, for $u, v \in \mathcal{L}_2$. Furthermore, let Y_1 and $Y_2 = \mathbb{R}$, and the Gaussian statistics have zero mean, and variances be as given in (53). Let $D_{11}^2 = D_{22}^1 = I$, the identity operator on $\mathcal{L}_2$, and $D_{12}^1 = D_{21}^{2*}$ be the Fredholm operator

$$D_{12}^1 u = \int_0^T K(t,s)u(s)ds \quad (57)$$

where $K(t,s)$ is a continuous kernel on $0 \leq t, s \leq T$, and finally let $F_i^1 = f_i(t)$, $i=1,2$, which are continuous functions on $[0,T]$.

Now, conditions (47a) and (50a) depend only on the probabilistic structure, and are therefore again given by (54) and (55), respectively. For (50b), however, we have to obtain the counterpart of (56), by simply replacing $|\bar{d}|^2$ with the norm of the operators $D_{12}^1 D_{12}^{1*}$ and $D_{12}^{1*} D_{12}^1$, respectively. Since $D_{12}^{1*} u = \int_0^T K(s,t)u(s)ds$, the self-adjoint operator $D_{12}^1 D_{12}^{1*}$ is given by

$$D_{12}^1 D_{12}^{1*} u = \int_0^T \int_0^T K(t, \tau) K(s, \tau) u(s) ds d\tau = \int_0^T K(t, s) u(s) ds,$$

where
$$K(t, s) \triangleq \int_0^T K(t, \tau) K(s, \tau) d\tau. \quad (58a)$$

Let
$$\lambda = \left\{ \int_0^T \int_0^T |K(t, s)|^2 dt ds \right\}^{1/2}. \quad (58b)$$

Then,
$$\|D_{12}^1 D_{12}^{1*} u\|^2 \triangleq \int_0^T \left| \int_0^T K(t, s) u(s) ds \right|^2 dt \leq \int_0^T \left[\int_0^T |K(t, s)|^2 ds \right] \left[\int_0^T |u(s)|^2 ds \right] dt = \lambda^2 \|u\|^2$$

where the second step follows from the Cauchy-Schwarz inequality. Hence,

$$\|D_{12}^1 D_{12}^{1*}\| \leq \lambda,$$

and because of symmetry $D_{12}^{1*} D_{12}^1$ is also bounded in norm by the same quantity.

This then leads to the following counterpart of (56): A sufficient condition for satisfaction of (50b) is either

$$\mu a^2 \lambda < \eta [u^2 a^2 - (1-\mu)(\mu ab - c^2)] \quad (59a)$$

or

$$\eta b^2 \lambda < \mu [\eta^2 b^2 - (1-\eta)(\eta ab - e^2)] \quad (59b)$$

provided that the terms on the right-hand-side are positive, where λ is defined by (58a)-(58b).

Hence, under (54) and either (55a) and (59a) or (55b) and (59b), the continuous-time static decision problem formulated above admits a unique stable equilibrium solution, and this solution is given by (from Thm. 3):

$$\gamma_i^0(t, y_i) = k_i(t) y_i, \quad i=1,2. \quad (60)$$

where $k_i(t)$ are continuous functions on $[0, T]$, satisfying

$$k_1(t) - \left(\frac{ec}{ab}\right) \int_0^T K(t,s)k_1(s)ds - (\sigma_{xy_2}^2 e/ab) \int_0^T K(t,s)f_1(s)ds - (\sigma_{xy_1}^1/a)f_1(t)=0 \quad (61a)$$

$$k_2(t) - \left(\frac{ec}{ab}\right) \int_0^T K(s,t)k_2(s)ds - (\sigma_{xy_1}^1 c/ab) \int_0^T K(s,t)f_2(s)ds - (\sigma_{xy_2}^2/b)f_2(t)=0. \quad (61b)$$

Note that $k_1(t)$ above stands for operator L_1 in (52), and we have already shown that a unique solution to both (61a) and (61b) exist in $L_2 [0,T]$, under (54) and either (55a) and (59a) or (55b) and (59b), and this solution is also continuous.

Finally, if our interest lies only in the existence of a unique linear equilibrium solution (not necessarily stable), the required condition is unique solvability of the integral equations (61a)-(61b), for which a sufficient condition is [6]

$$(ec/ab) \lambda < 1$$

where λ is defined by (58b). □

5.3. Asymmetric Mode of Decision Making

To obtain the counterpart of the results of §5.2 under the asymmetric mode of decision making, we first investigate the possibility for the unique solution of Thm. 2 to be linear. Towards this end we first observe that the decision problem will admit a unique linear solution if, and only if, equation (31) is satisfied by the decision rule

$$\gamma(y_1) = Ay_1 \quad (62)$$

for some linear bounded operator $A: \mathbb{R}^{m_1} \rightarrow U_1$. Hence, using (31), A should be the solution of (by pulling A out of the conditional expectations)

$$\begin{aligned} Ay_1 = & D_{12}^1 D_{21}^2 A E^1[E^2[y_1|y_2]|y_1] + D_{21}^{2*} D_{12}^{1*} A g^1(y_1) E^2[g^2(y_2) E^1[y_1|y_2]|y_1] \\ & - D_{21}^{2*} D_{22}^1 D_{21}^2 A g^1(y_1) E^2[g^2(y_2) E^2[y_1|y_2]|y_1] \\ & + F_1^1 E^1[x|y_1] + D_{21}^{2*} F_2^1 g^1(y_1) E^2[g^2(y_2) E^1[x|y_2]|y_1] \\ & - D_{21}^{2*} D_{22}^1 F_2^2 g^1(y_1) E^2[g^2(y_2) E^2[x|y_2]|y_1] - D_{12}^1 F_2^2 E^1[E^2[x|y_2]|y_1], \quad \forall y_1 \in \mathbb{R}^{m_1}. \end{aligned} \quad (63a)$$

Since the random variables are jointly Gaussian under both measures,

$$E^1[y_k|y_\ell] = S_{k\ell}^1 y_\ell, \quad k \neq \ell: \quad i, k, \ell = 1, 2 \quad (63b)$$

$$E^1[x|y_\ell] = S_{0\ell}^1 y_\ell, \quad i, \ell = 1, 2, \quad (63c)$$

for some matrices $S_{k\ell}^1$ and $S_{0\ell}^1$. In view of this, (63a) can be rewritten as

$$\begin{aligned} Ay_1 = & (D_{12}^1 D_{21}^2 AS_{12}^2 S_{21}^1 + F_1^1 S_{01}^1 + D_{12}^1 F_2^2 S_{02}^2 S_{21}^1) y_1 + [D_{21}^{2*} D_{12}^{1*} AS_{12}^1 - D_{21}^{2*} D_{22}^1 D_{21}^2 AS_{12}^2 \\ & + D_{21}^{2*} F_2^1 S_{02}^1 - D_{21}^{2*} D_{22}^1 F_2^2 S_{02}^2] g^1(y_1) E^2[g^2(y_2) y_2 | y_1]. \end{aligned} \quad (6)$$

This then leads to the following Proposition:

Proposition 6. Let (47) and Condition (3) be satisfied, and either $P_{y_1}^1 \neq P_{y_1}^2$ or $P_{y_2}^1 \neq P_{y_2}^2$. Then, the quadratic Gaussian decision problem with asymmetric mode of decision making admits a linear (Stackelberg) equilibrium solution if, and only if,

(i) there exists a bounded linear operator $A: \mathbb{R}^m \rightarrow U_1$ satisfying

$$A = D_{12}^1 D_{21}^2 AS_{12}^2 S_{21}^1 + F_1^1 S_{01}^1 + D_{12}^1 F_2^2 S_{02}^2 S_{21}^1, \quad (65a)$$

and (ii) this solution also satisfies

$$D_{21}^{2*} D_{12}^{1*} AS_{12}^1 - D_{21}^{2*} D_{22}^1 D_{21}^2 AS_{12}^2 + D_{21}^{2*} F_2^1 S_{02}^1 - D_{21}^{2*} D_{22}^1 F_2^2 S_{02}^2 = 0. \quad (65b)$$

Proof. Since the "if" part is obvious in view of Thm. 2, we verify only

the "only if" part of the proposition. [In what follows we adopt the notation

$S \geq 0$ to imply that the nonnegative definite matrix S has at least one positive eigenvalue.] The proof proceeds by showing for three exclusive

(and exhaustive) cases that $f(y_1) = g^1(y_1) E^2[g^2(y_2) y_2 | y_1]$ is a nonlinear function of y_1 .

$$(a) \quad P_{y_2}^2 = P_{y_2}^1, \text{ and } P_{y_1}^1 \neq P_{y_1}^2.$$

Here, $g^2(y_2) = 1$, and $g^1(y_1) = c_1 \exp\{-\frac{1}{2} y_1^T W_1 y_1\}$, where $W_1 \geq 0$, and

$c_1 > 0$ is a constant. Hence, $f(y_1) = g^1(y_1) S_{21}^2 y_1$ which is nonlinear since $W_1 \geq 0$.

$$(b) \quad P_{y_2}^2 \neq P_{y_2}^1, \quad P_{y_1}^1 = P_{y_1}^2.$$

Here, $g^1(y_1)=1$, and $g^2(y_2)=c_2 \exp \{-\frac{1}{2} y_2^T W_2 y_2\}$, where $W_2 \geq 0$, and $c_2 > 0$ is a constant. In this case, f can be evaluated to be

$$f(y_1) = c(V+W_2)^{-1} V S_{21}^2 y_1 \exp \{-\frac{1}{2} y_1^T B y_1\}$$

where $V \triangleq E^2\{(y_2 - S_{21}^2 y_1)(y_2 - S_{21}^2 y_1)^T\}$

$$B = S_{21}^2 V S_{21}^2 - S_{21}^2 V (V+W_2)^{-1} V S_{21}^2 \geq 0,$$

and c is a constant. Since $W_2 \geq 0$, B has at least one positive eigenvalue, and hence $f(y_1)$ is again nonlinear in y_1 .

$$(c) \quad P_{y_2}^2 \neq P_{y_2}^1 \text{ and } P_{y_1}^1 \neq P_{y_1}^2.$$

In this case, following the same lines as above, we find

$$f(y_1) = \bar{c}(V+W_2)^{-1} V S_{21}^2 y_1 \exp \{-\frac{1}{2} y_1^T (B+W_1) y_1\}$$

which is nonlinear since both $B \geq 0$, $W_1 \geq 0$.

Hence, in view of the preceding analysis, a necessary condition for existence of a solution to (64) is that the last term should vanish (i.e. (65b)) for an A that solves (65a). $\square$

Remark 5. A sufficient condition for (65a) to admit a unique solution in the Banach space of linear bounded operators mapping $\mathbb{R}^{m_1}$ into U_1 is

$$r(D_{12}^1 D_{21}^2 D_{21}^{2*} D_{12}^{1*}) \operatorname{Tr} \{S_{12}^2 S_{21}^1 S_{21}^{1*} S_{12}^{2*}\} < 1,$$

which is clearly satisfied under Condition (3). $\square$

The conditions of Prop. 6 are clearly non-void; because, given the unique solution of (65a), it may be possible to find F_2^1 , F_2^2 , S_{02}^1 and S_{02}^2 so that (65b) is satisfied. However, it should also be clear that satisfaction of (65b) places some severe restrictions on the parameters of the problem, which in general will not be met. Hence, it is fair to say that, if either $P_{y_1}^1 \neq P_{y_1}^2$ or $P_{y_2}^1 \neq P_{y_2}^2$, generically the problem does not admit a linear equilibrium solution, even if it is a team problem; that is:

Corollary 5. If either $P_{y_1}^1 \neq P_{y_1}^2$ or $P_{y_2}^1 \neq P_{y_2}^2$ (or both), the quadratic Gaussian decision problem does not admit (generically) a linear Stackelberg equilibrium solution. The unique solution, which exists under (47) and *Condition (3)*, is nonlinear. $\square$

The conditions of the preceding Corollary involve only the marginal distributions of y_1 and y_2 ; in the complement of these conditions we can derive the following linear solution:

Proposition 7. For the quadratic Gaussian decision problem, let both $P_{y_1}^1 = P_{y_1}^2$ and $P_{y_1}^1 = P_{y_2}^2$ (but not necessarily $P_{y_1 y_2}^1 = P_{y_1 y_2}^2$, and even $P_{y_1 y_2}^1 = P_{y_1 y_2}^2$). Then, if

$$2[r(D_{12}^1 D_{21}^2 D_{21}^{2*} D_{12}^{1*})]^{1/2} + [r(D_{21}^{2*} D_{22}^1 D_{21}^2)]^{1/2} < 1 \quad (66)$$

the problem admits a unique Stackelberg equilibrium solution for DM1 (the leader) which is linear in y_1 :

$$y_1^s(y_1) = Ay_1 \quad (67a)$$

where $A: \mathbb{R}^2 \rightarrow U_1$ is the unique bounded linear operator solving

$$\begin{aligned} Ay_1 = & (D_{12}^1 D_{21}^2 AS_{12}^2 S_{21}^1 + D_{21}^{2*} D_{12}^{1*} AS_{12}^1 S_{21}^2 - D_{21}^{2*} D_{22}^1 D_{21}^2 AS_{12}^2 S_{21}^2 + F_{10}^1 S_{01}^1 + D_{12}^1 F_{20}^2 S_{02}^2 S_{21}^1 \\ & + D_{21}^{2*} F_{20}^1 S_{02}^1 S_{21}^2 - D_{21}^{2*} D_{22}^1 F_{20}^2 S_{02}^2 S_{21}^2) y_1, \quad \forall y_1 \in \mathbb{R}^{m_1}, \end{aligned} \quad (67b)$$

and $S_{k\ell}^i$ are defined by (31b)-(31c), and S_{0i}^i is defined by $E^i[x|y_i] = S_{0i}^i y_i$.

Proof. When $P_{y_1}^1 = P_{y_1}^2$ and $P_{y_2}^1 = P_{y_2}^2$, $g^1(y_1) = g^2(y_2) = 1$ and hence *Conditions (1) and (2)* of Thm. 2 are always satisfied, and in *Condition (3)*, $\rho_2 = \rho_4 = 1$. Then, (66) is the counter part of (39), and hence existence and uniqueness follow from Thm. 2. Linearity, on the other hand, follows by noting that if we start iteration (35) with $y_1^{(0)} = 0$, since $g^1(y_1) = g^2(y_2) = 1$ every term will be linear in y_1 (see also (64)), and hence the limit (which exists by Thm. 2) will be linear. Then, substituting $y_1^1 = Ay_1$ in (31), we obtain (67b), by simply letting $g^1(y_1) = g^2(y_2) = 1$ in (64). $\square$

When there is a discrepancy between the DM's perceptions of the variances of either y_1 or y_2 , Prop. 7 will not hold, and the problem will admit (generically) a nonlinear equilibrium solution, as proven earlier in Prop. 6 and Corollary 5. In this case, an explicit closed-form solution cannot be obtained; however, an approximate solution can be derived by using the iteration (35) which, for the Gaussian problem, becomes

$$\begin{aligned}
 \gamma^{(k+1)}(y_1) = & D_{12}^1 D_{21}^2 E^1 [E^2 [\gamma^{(k)}(y_1) | y_2] | y_1] + D_{21}^{2*} D_{12}^{1*} g^1(y_1) \\
 & \cdot E^2 [g^2(y_2) E^1 [\gamma^{(k)}(y_1) | y_2] | y_1] - D_{21}^{2*} D_{22}^1 D_{21}^2 g^1(y_1) \\
 & \cdot E^2 [g^2(y_2) E^2 [\gamma^{(k)}(y_1) | y_2] | y_1] + (F_1^1 S_{01}^1 + D_{12}^1 F_2^2 S_{02}^2 S_{21}^1) y_1 \\
 & + (D_{21}^{2*} F_2^1 S_{02}^1 - D_{21}^{2*} D_{22}^1 F_2^2 S_{02}^2) g^1(y_1) E^2 [g^2(y_2) y_2 | y_1] \quad (68)
 \end{aligned}$$

If we start this iteration with $\gamma^{(0)}(y_1)=0$, or any linear function of y_1 , at every iteration we obtain linear combinations of terms of the type $A^{(k)} y_1$ and $B^{(k)} y_1 \exp \{-\frac{1}{2} y_1^T V^{(k)} y_1\}$, where $A^{(k)}$ and $B^{(k)}$ are linear operators, and $V^{(k)} \succeq 0$ is an $m_1 \times m_1$ matrix. Since this is a successive approximation technique under Condition (6), even stopping the iteration after a finite number of terms will provide a solution sufficiently close to the unique optimum. Hence, generically, a suboptimal policy for DM1, which is sufficiently close to the unique solution of (31), will be of the form

$$\gamma_1(y_1) = A^{(N)} y_1 + \sum_{\ell \leq N} B^{(\ell)} y_1 \exp \{-\frac{1}{2} y_1^T V^{(\ell)} y_1\}$$

where N is a sufficiently large integer (related to the number of iterations taken in (68)), and $A^{(N)}$, $B^{(\ell)}$, $V^{(\ell)}$ are generated via the iteration (68). Note that as $N \rightarrow \infty$ this solution will uniformly converge to the unique optimum.

Yet another suboptimal solution can be obtained by restricting DM1's policies, at the outset, to linear functions of y_1 , i.e. to the form (62) where A is a variable linear operator. DM2's response to any such policy will

also be linear (in y_2), thus making T_2 in (10) a linear operator. Then, the problem faced by DMI is minimization of (11), with $\gamma(y_1) = Ay_1$, over all linear bounded operators A . The solution of this minimization problem will provide DMI with a linear policy that is (in general) inferior to the limiting solution of (68), unless, of course, $g^1(y_1) = g^2(y_2) = 1$ in which case the two solutions will be the same (satisfying (67b)). We do not pursue here the details of the derivation of the best linear solution for the general case (as outlined above).

Furthermore, it is possible to work out the various conditions for the special cases of the scalar and continuous-time team problems (formulated as in Examples 1 and 2) and write down the equilibrium solution explicitly whenever it is linear. Such an analysis would routinely follow the lines of the discussion of Examples 1 and 2, and hence will not be included here mainly because of space limitations.

6. Discussion of Possible Extensions, and Concluding Remarks

In the preceding sections, we have developed an equilibrium theory for two-person quadratic decision problems with static information patterns, wherein the decision makers (DM's) do not necessarily have the same perception of the underlying probability space; that is, our formulation allows for discrepancies in the way different DM's perceive the probability space. As indicated earlier, when such discrepancies exist, even team problems have to be analyzed in the framework of nonzero-sum stochastic games, and in such a framework the Nash solution concept is the most suitable equilibrium concept if the DM's occupy symmetric (non-hierarchical) positions in the decision process, and the Stackelberg solution concept becomes more meaningful if there is a hierarchy in decision making.

Section 3 of the paper has provided a set of sufficient conditions for existence and uniqueness of Nash equilibrium in the case of symmetric mode of decision making, with the additional feature that it be stable. This is an appealing feature of the solution because, in order to arrive at equilibrium (as a consequence of an infinite number of response iterations), each DM does not have to know the subjective probability measures perceived by the other DM, but has to know only the policy adopted by the other DM at the most recent step of the iteration.

In Section 4 we have presented a counterpart of the results of §3 under the asymmetric mode of decision making. The conditions derived ensure that the equilibrium policy of the leader can be obtained as the limit of an infinite sequence which involves conditional expectations under two different probability measures. This sequence [(35), (27)] is structurally different from its counterpart in §3 (see (14)), even for team problems, and it contains R-N derivatives of the two probability measures as multiplying factors (which are absent in (14)).

In Section 5 we have shown that when the underlying probability distributions belong to a Gaussian class, the Nash equilibrium solution will be linear (affine, if mean values are nonzero) in the available static measurements, with the gain operator

satisfying a Lyapunov-type operator equation (cf. Thm. 3). This solution and the associated existence conditions have been studied further in the context of two examples which involve scalar and continuous-time stochastic team problems with multiple probability models. In developing a counterpart of Thm. 3 for asymmetric mode of decision making, we have arrived at a seemingly surprising (unexpected) result — the unique Stackelberg equilibrium solution being generically nonlinear in the measurements (even under Gaussian multiple probability measures). This constitutes the first unique nonlinear solution reported in the literature for a quadratic Gaussian static game or team problem.[†] It should be noted that we have not given a closed-form expression for this nonlinear solution, but have instead provided a recursive scheme which generates admissible policies that come arbitrarily close to the optimum solution.

Several extensions of the results presented in this paper seem to be possible. Firstly, we should note that the general Hilbert-space framework adopted in this paper and the general solutions presented for the Gaussian problems in Section 5 (Thms. 3 and 4) apply to other models also, such as the ones similar to the continuous-time team problem treated in [9] and the Stackelberg problem of [26], but with the DM's having different probability models. It is expected that some explicit results (closed-form solutions) can also be obtained in these cases, but this point has not been pursued in this paper and is left for future research.

Another possible extension of the results of this paper would be to the class of problems in which the random state of nature (i.e. x) as well as the measurements (y_i) are stochastic processes. The general theories of Sections 3 and 4 could easily be extended so as to encompass this class of problems also, provided that the problem is set up under the right mathematical assumptions. In particular, if the random

[†]Reference [12] also reports on existence of nonlinear (Nash) solutions for quadratic Gaussian nonzero-sum games, but there the nonlinear solution is one of many solutions one of which is linear, and is due to nonunique intersection of reaction functions (which disappears under appropriate conditions).

variables are taken to be Hilbert space valued weak random variables, with the inner product satisfying some continuity and boundedness conditions [11], Thms. 1-4 directly apply to this more general class of decision problems, when interpreted in the right framework. Furthermore, extensions to dynamic (multi-stage) problems is also possible, by adopting the framework of (say) [8] for the linear-quadratic-Gaussian problem. Then, the unique Nash equilibrium solution under the one-step-delay observation sharing pattern can be obtained by basically following the approach of [8] and utilizing in the recursive derivation Thm. 3 of this paper instead of Thm. 2 of [8]. Details of this derivation are, however, rather involved, and will be reported elsewhere.

Regarding the Nash equilibrium solution, yet another possible extension would be to multiple decision-maker problems with more than two (say, N) DM's. Even though the definition of Nash equilibrium (cf. Def. 1) admits a natural (unique) extension to such problems, that of stable equilibrium (cf. Def. 2) does not extend in a unique way. One viable alternative is to assume that each DM reacts optimally to the set of most recent policies of all the other DM's, which leads to a set of N relations similar to (9). In this case, (12) will be replaced by N equations with the right-hand-side expressions involving $N-1$ policies of different DM's. However, the line of reasoning that took us from (13) to (14) does not have a counterpart if $N > 2$, and in general it is not possible to obtain N recursion relations each of which involves only one DM's policies at consecutive stages. Then, the counterpart of (13) will have to be treated as a "multi-valued" operator equation, in which context an existence and uniqueness result will have to be established. This seems to be a challenging problem whose solution requires somewhat different mathematical techniques than the ones employed in this paper.

One source of motivation for the research reported in this paper has been (as discussed in Section 1) the desire to investigate the sensitivity and robustness of team-optimal solutions (in stochastic teams) to independent variations in the perceptions of the DM's of the underlying probability space (and, in particular, the

probability measure). The analysis of this paper indeed provides a framework for such a study when the roles of the DM's are either symmetric or asymmetric, since an equilibrium theory has been established in both cases within an " ϵ -neighborhood" of the team-optimal solution. Some further work is needed in order to determine the "satisfiability" of the several existence conditions obtained in the paper when the region of interest is an ϵ -neighborhood of a common probability space, and to further extend the analysis to an investigation of sensitivity and robustness properties of team solutions (obtained under the stipulation of existence of a common underlying probability space) in this ϵ -neighborhood.

An aspect of the decision problem studied here, which is worth bringing forth, is that the subjective probability measures perceived by each DM is fixed in advance and the DM's do not attempt to change their subjective priors during the course of the decision process. Hence, in this sense, the problem treated here is categorically different from the class of problems treated in [18]-[21], where the objective was for the DM's to arrive at a common (consistent) set of probabilistic descriptions of the unknown variables. In the symmetric mode, there is, however, an implicit learning process built in the recursive process that leads to the stable equilibrium decision rules for each DM, since the DM's do not necessarily have access to each other's perception of the priors.

Yet another aspect of the problem treated in this paper is that the general formulation could be viewed as a multi-modeling in multiple-decision maker problems; however, as opposed to the singular perturbations approach of [22]-[24], here the multi-modeling is in the probabilistic description of the decision problem, with each DM having a different probabilistic model of the "rest of the world."

Appendix A

In this appendix we state a number of results concerning the spectral radii of linear bounded operators.

Let $A: \Gamma \rightarrow \Gamma$ and $B: \Gamma \rightarrow \Gamma$ be two linear bounded operators where Γ is a Hilbert space equipped with the inner product $\langle \cdot, \cdot \rangle$. Then the spectral radius of A is defined by

$$r(A) = \limsup_{k \rightarrow \infty} [\langle A^k \rangle]^{1/k} \quad (A-1)$$

where $\langle A \rangle$ is the norm of A , given by

$$\langle A \rangle = \sup_{\gamma \in \Gamma} [\langle A\gamma, A\gamma \rangle / \langle \gamma, \gamma \rangle]^{1/2} \quad (A-2)$$

For self-adjoint operators there is an equivalence between the spectral radius and norm of an operator; specifically, if A is self-adjoint,

$$r(A) = \langle A \rangle = \sup_{\gamma \in \Gamma} \{ |\langle \gamma, A\gamma \rangle| / \langle \gamma, \gamma \rangle \} \quad (A-3)$$

[see [13], p. 514]. However, for operators which are not self-adjoint, such an equivalence does not exist, and one can only provide bounds on $r(A)$:

Lemma A.1.

For any linear bounded operator A ,

$$r(A) \leq \langle A \rangle = [r(A^*A)]^{1/2}$$

Proof. Since A belongs to a Banach algebra, $\langle A^k \rangle \leq \langle A \rangle^k$ and hence

$$r(A) \leq \limsup_{k \rightarrow \infty} \{ \langle A \rangle^k \}^{1/k} = \langle A \rangle$$

$$\text{Furthermore, } \langle A \rangle = \sup_{\gamma \in \Gamma} [\langle \gamma, A^*A\gamma \rangle / \langle \gamma, \gamma \rangle]^{1/2}$$

which is $[r(A^*A)]^{1/2}$ by (A-3) because A^*A is self-adjoint. □

Lemma A.2. Let A and B be two linear bounded operators which commute. Then,

$$(i) \quad r(AB + A^* B^*) \leq 2[r(AA^*)r(B^* B)]^{1/2} = 2[r(A^* A)r(BB^*)]^{1/2}$$

$$(ii) \quad r(AB) \leq r(A)r(B)$$

Proof. (i) Since $AB + A^* B^*$ is self-adjoint, using (A-3)

$$r(AB + A^* B^*) = \sup_{\gamma \in \Gamma} \{ |\langle \gamma, (AB + A^* B^*) \gamma \rangle| / \langle \gamma, \gamma \rangle \} = 2 \sup_{\gamma \in \Gamma} \{ |\langle A\gamma, B^* \gamma \rangle| / \langle \gamma, \gamma \rangle \}$$

where the equality has followed since A and B commute. Using Cauchy-Schwarz [3] inequality, this expression can be bounded from above by

$$\leq 2 \sup_{\gamma \in \Gamma} \left\{ \frac{|\langle A\gamma, A\gamma \rangle|^{1/2} |\langle B^* \gamma, B^* \gamma \rangle|^{1/2}}{\langle \gamma, \gamma \rangle} \right\}$$

and performing individual supremization we further obtain the bound

$$\begin{aligned} &\leq 2 \sup_{\gamma \in \Gamma} \left[\frac{|\langle A\gamma, A\gamma \rangle|}{\langle \gamma, \gamma \rangle} \right]^{1/2} \sup_{\gamma \in \Gamma} \left[\frac{|\langle B^* \gamma, B^* \gamma \rangle|}{\langle \gamma, \gamma \rangle} \right]^{1/2} \\ &= 2 \langle\langle A \rangle\rangle \langle\langle B^* \rangle\rangle = 2[r(A^* A)r(BB^*)]^{1/2} \end{aligned}$$

where the last line has followed from Lemma A.1. Note that this expression can be written in different ways because $r(A^* A) = r(AA^*)$, $r(BB^*) = r(B^* B)$.

(ii) Firstly note that

$$r(AB) = \limsup_{k \rightarrow \infty} [\langle\langle (AB)^k \rangle\rangle]^{1/k} = \limsup_{k \rightarrow \infty} [\langle\langle A^k B^k \rangle\rangle]^{1/k} \quad (*)$$

where the last equality has followed because A and B commute. Now, since A, B belong to a Banach algebra, $\langle\langle A^k B^k \rangle\rangle \leq \langle\langle A^k \rangle\rangle \langle\langle B^k \rangle\rangle$ for every k $\Rightarrow$

$$\langle\langle A^k B^k \rangle\rangle^{1/k} \leq [\langle\langle A^k \rangle\rangle \langle\langle B^k \rangle\rangle]^{1/k} = [\langle\langle A^k \rangle\rangle]^{1/k} [\langle\langle B^k \rangle\rangle]^{1/k} \text{ for every } k$$

and taking lim sup of both sides, and using (*)

$$r(AB) \leq \limsup_{k \rightarrow \infty} [\langle\langle A^k \rangle\rangle]^{1/k} [\langle\langle B^k \rangle\rangle]^{1/k} \leq r(A)r(B)$$

which proves the desired result. □

Lemma A.3. Let A and B be both self-adjoint. Then,

$$r(A + B) \leq r(A) + r(B)$$

Proof. This follows from (A-3) and the triangle inequality applied to norm $\langle\langle\cdot\rangle\rangle$. $\square$

Appendix B

Proof of Theorem 1

Let us first recall the following result from functional analysis (see, for example [13, Chapter XIII, Theorem 3]).

Lemma B.1. Let S be a linear bounded operator mapping a Hilber space Γ into itself, and consider the equation

$$\gamma = S\gamma + \mu \quad (B-1)$$

defined on Γ . Furthermore, consider the "successive approximation"

$$\gamma^{(k+1)} = \mu + S\gamma^{(k)}, \quad k=0,1,\dots \quad (B-2)$$

to the solution of (B-1). Then, the sequence generated by (B-2) converges to a unique element of Γ , for any starting point $\gamma^{(0)} \in \Gamma$, which is further a solution of (B-1), if, and only if, the spectral radius of S is less than unity, i.e. there exists a ρ , $0 < \rho < 1$, such that

$$r(S) \leq \rho \quad (B-3)$$

Now, applying this Lemma to our problem, we identify S with either S_1 or S_2 (given by (16)), Γ with Γ_1 or Γ_2 , the successive approximation (B-2) with (14), and condition (B-3) above with (19) for either $i=1$ or 2 . Then, the statement of Thm. 1 (i) readily follows from the preceding Lemma, in view of Prop. 2.

Furthermore, since S_i can be written as the product of two commuting operators, using Lemma A.2(ii) we obtain

$$r_i(S_i) = r_i(\bar{D}^i \bar{P}_{i|i}) \leq r_i(\bar{D}^i) r_i(\bar{P}_{i|i}).$$

Under (20a) this can be bounded from above by $\rho_1^i \rho_2^i = \rho^i < 1$, thereby ensuring (19). On the other hand, since the spectral radius of a bounded linear operator is bounded from above by its norm [13], and that $\|\bar{D}^i\|_i = \langle \langle \bar{D}^i \rangle \rangle_i$ because $\bar{D}^i$ also maps U_i into U_i (in addition to being a mapping from Γ_i into itself), (20b) follows. This completes the proof of Thm. 1. $\square$

Appendix C

1. Proof of Corollary 2 (Section 3)

Here we verify that the second inequality of (20b) is implied by the condition that (21a) is uniformly bounded by ρ_2^i . Towards this end, we first have, for each $\gamma \in \Gamma_i$, from the Cauchy-Buniakowski (Schwarz) inequality [3] applied to Γ_i :

$$\begin{aligned} \|\bar{P}_i^i\|_i^2 &= \left\| \int_{Y_j} P_{y_j|y_i}^i (dn|y_i) \int_{Y_i} \gamma(\xi) P_{y_i|y_j}^j (d\xi|n) \right\|_i^2 \leq \left\| \int_{Y_i} \gamma(\xi) P_{y_i|y_j}^j (d\xi|n) \right\|_i^2 \\ &= \int_{Y_j} \left(\int_{Y_i} \gamma(\xi) P_{y_i|y_j}^j (d\xi|n) \right) \cdot \int_{Y_i} \gamma(\xi) P_{y_i|y_j}^j (d\xi|n) \cdot P_{y_j}^j (dn) g^j(n) \end{aligned}$$

where the last equality has involved a change of measures, using the R-N derivative $g^j(n)$. Now, again using the Cauchy-Schwarz inequality, this expression can be bounded from above by

$$\begin{aligned} &\leq \int_{Y_j} \int_{Y_i} (\gamma(\xi), \gamma(\xi))_i P_{y_i|y_j}^j (d\xi|n) g^j(n) P_{y_i}^j (dn) \\ &= \int_{Y_i} (\gamma(\xi), \gamma(\xi))_i P_{y_i|y_j}^j (d\xi) g^i(\xi) \int_{Y_j} P_{y_j|y_i}^j (dn|\xi) g^j(n) \quad , \end{aligned}$$

where the last equality has followed from Bayes Theorem. It now readily follows that under the condition of Corollary 2, the last expression is bounded from above by $\leq \rho_2^i \|\gamma\|_i^2$, thus proving the desired result for $i=1,2$. $\square$

2. Proof of Corollary 4 (Section 4)

The fact that uniform boundedness of (43a) (by $(\rho_2)^2$) implies the first inequality of (40b) follows readily from the proof given above, since the spectral

radius of $\bar{P}_{1|1}^*$ $\bar{P}_{1|1}$ is equal to the square of the norm of $\bar{P}_{1|1}$. Now, to verify that uniform boundedness of (43b) implies the second inequality of (40b) we follow basically the same line of reasoning, but the details of the proof are more involved. Towards this end we first note that for each $\gamma \in \Gamma_1$,

$$\begin{aligned} \|K\gamma\|_1^2 &= \|g^1(\xi) \int_{Y_2} P_{y_2|y_1}^2(d\eta|y_1=\xi) g^2(\eta) \int_{Y_1} P_{y_1|y_2}^2(db|y_2=\eta) \gamma(b)\|_1^2 \\ &= \|\sqrt{g^1(\xi)} \int_{Y_2} P_{y_2|y_1}^2(d\eta|y_1=\xi) g^2(\eta) \int_{Y_1} P_{y_1|y_2}^2(db|y_2=\eta) \gamma(b)\|_2^2 \\ &\leq \|\sqrt{g^1(\xi) g^2(\eta)} \int_{Y_1} P_{y_1|y_2}^2(db|y_2=\eta) \gamma(b)\|_2^2, \end{aligned}$$

where the second equality follows from a change of measures, and the last bound follows from the Cauchy-Schwarz inequality. It should be pointed out that here we have abused the notation and have used $\|\cdot\|_2$ to mean

$$\|m(\xi, \eta)\|_2 = \left\{ \int_{Y_1 Y_2} (m(\xi, \eta), m(\xi, \eta))_1 P_{y_1 y_2}^2(d\xi x d\eta) \right\}^{1/2}$$

where m is a $y_1 y_2$ -measurable random variable taking values in U_1 ; hence, the sub-index "2" indicates that the probability space is the one determined by the subjective probability measure of DM2.

Now, the latter bound can further be bounded above by

$$\leq \int_{Y_1 Y_2} g^1(\xi) |g^2(\eta)|^2 P_{y_1 y_2}^2(d\xi x d\eta) \int_{Y_1} P_{y_1|y_2}^2(db|y_2=\eta) (\gamma(b), \gamma(b))_1$$

since (i)

$$\begin{aligned} &\left(\int_{Y_1} P_{y_1|y_2}^2(db|y_2=\eta) \gamma(b) \right), \int_{Y_1} P_{y_1|y_2}^2(db|y_2=\eta) \gamma(b) \Big|_1 \\ &\leq \int_{Y_1} P_{y_1|y_2}^2(db|y_2=\eta) (\gamma(b), \gamma(b))_1 \end{aligned}$$

by the Cauchy-Schwarz inequality (because $P_{y_1|y_2}^2$ is also a probability measure), and (ii) $g^1(\xi) |g^2(\eta)|^2 \geq 0$. Hence, by interchanging the variables ξ and η ,

$$\begin{aligned} \|K_Y\|_1^2 &\leq \int_{Y_1} \int_{Y_2} \int_{Y_1} (\gamma(\xi), \gamma(\xi))_1 g^1(b) |g^2(\eta)|^2 P_{y_1 y_2}^2 (db x d\eta) \frac{P_{y_2|y_1}^2 (d\eta | y_1 = \xi) P_{y_1}^2 (d\xi)}{P_{y_2}^2 (d\eta)} \\ &= \int_{Y_1} P_{y_1}^1 (d\xi) (\gamma(\xi), \gamma(\xi))_1 \int_{Y_2} \int_{Y_1} g^1(\xi) g^1(b) |g^2(\eta)|^2 \\ &\quad \cdot P_{y_1|y_2}^2 (db | y_2 = \eta) P_{y_2|y_1}^2 (d\eta | y_1 = \xi) \end{aligned}$$

and under (43b) this can be bounded above by

$$\leq \int_{Y_1} P_{y_1}^1 (d\xi) (\gamma(\xi), \gamma(\xi))_1 \rho_4^2 = \rho_4^2 \| \gamma \|_1^2,$$

which completes the proof. =

Appendix D

1. Derivation of First and Second Gateaux Variations [(25)-(26)]

Starting with the expression for $\tilde{J}$ as given by (23), we first obtain

$$\begin{aligned} \Delta \tilde{J}(\gamma; h) &\triangleq \tilde{J}(\gamma+h) - \tilde{J}(\gamma) \\ &= \frac{1}{2} \langle h, \gamma \rangle_1 + \frac{1}{2} \langle \gamma, h \rangle_1 + \frac{1}{2} \langle h, h \rangle_1 \\ &\quad + \frac{1}{2} \int_{Y_2} \{ (F_2^2 E^2[x|y_2] + D_{21}^2 E^2[\gamma(y_1)|y_2], D_{22}^1 D_{21}^2 E^2[h(y_1)|y_2])_2 \\ &\quad + (D_{21}^2 E^2[h(y_1)|y_2], D_{22}^1 F_2^2 E^2[x|y_2] + D_{22}^1 D_{21}^2 E^2[\gamma(y_1)|y_2])_2 \\ &\quad + (D_{21}^2 E^2[h(y_1)|y_2], D_{22}^1 D_{21}^2 E^2[h(y_1)|y_2])_2 \} P_{y_2}^1 (d\xi) \end{aligned}$$

$$\begin{aligned}
& - \langle h, E^1[F_1^1 x | y_1] \rangle_1 - \int_{X \times Y_2} (D_{21}^2 E^2[h(y_1) | y_2], F_2^1 x)_2 P^1(dx, Y_1, d\xi) \\
& - \langle h, E^1[D_{12}^1 D_{21}^2 E^2[\gamma(y_1) | y_2] | y_1] + E^1[D_{12}^1 F_2^2 E^2[x | y_2] | y_1] \rangle_1 \\
& - \langle \gamma, E^1[D_{12}^1 D_{21}^2 E^2[h(y_1) | y_2] | y_1] \rangle_1 - \langle h, D_{12}^1 D_{21}^2 E^1[E^2[h(y_1) | y_2] | y_1] \rangle_1 \\
& \equiv \delta \tilde{J}(\gamma; h) + \delta^2 \tilde{J}(\gamma; h) .
\end{aligned}$$

Now, since $\delta \tilde{J}(\gamma; h)$ is homogeneous of degree one, and $\delta^2 \tilde{J}(\gamma; h)$ is homogeneous of degree two, $\Delta \tilde{J}(\gamma; h)$ admits a unique decomposition with the corresponding expressions being (after some simplification)

$$\begin{aligned}
\delta \tilde{J}(\gamma; h) &= \langle h, \gamma \rangle_1 + \int_{Y_2} (E^2[h(y_1) | y_2], D_{21}^{2*} D_{22}^1 \{F_2^2 E^2[x | y_2] \\
&+ D_{21}^2 E^2[\gamma(y_1) | y_2]\})_1 P_{Y_2}^1(d\xi) - \langle h, E^1[F_1^1 x | y_1] \rangle_1 \\
&- \int_{X \times Y_2} (E^2[h(y_1) | y_2], D_{21}^{2*} F_2^1 x)_1 P^1(dx, Y_1, d\xi) \quad (D-1) \\
&- \langle h, D_{12}^1 D_{21}^2 P_{1|1} \gamma(y_1) + D_{12}^1 F_2^2 E^1[E^2[x | y_2] | y_1] \rangle_1 - \langle \bar{P}_{1|1} h, D_{21}^{2*} D_{12}^1 \gamma \rangle_1
\end{aligned}$$

$$\begin{aligned}
\delta^2 \tilde{J}(\gamma; h) &= \frac{1}{2} \langle h, h \rangle_1 + \frac{1}{2} \int_{Y_2} (E^2[h(y_1) | y_2], D_{21}^{2*} D_{22}^1 D_{21}^2 \\
&\cdot E^2[h(y_1) | y_2])_1 P_{Y_2}^1(d\xi) - \langle h, D_{12}^1 D_{21}^2 \bar{P}_{1|1} h \rangle_1 \quad (D-2)
\end{aligned}$$

where we have used some properties of adjoint operators under inner products, and the notation introduced in (28); we have also made use of the fact that the bounded linear operator $D_{12}^1 D_{21}^2: \mathcal{U}_1 \rightarrow \mathcal{U}_1$ commutes with the double conditional expectation operator $\bar{P}_{1|1}$ (or $P_{1|1}$).

We now prove a lemma which will be used in simplifying these expressions further.

Lemma D.1. For $h(\cdot) \in \mathcal{H}_1$, $f(\cdot) \in \mathcal{H}_2$,

$$\begin{aligned} \int_{Y_2} (E^2[h(y_1)|y_2=\xi], f(\xi))_1 P_{y_2}^1(d\xi) &= \\ \int_{Y_2} (h(n), g^1(n) E^2[g^2(y_2)f(y_2)|y_1=n])_1 P_{y_1}^1(dn) & \\ \triangleq \langle h, g^1(y_1) E^2[g^2(y_2)f(y_2)|y_1] \rangle_1 & \end{aligned} \quad (D-3)$$

where $g^i(\cdot)$ are given by (2).

Proof. The proof follows from the following set of equalities where we are allowed to change orders of integration because $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces of random variables well defined under both measures:

$$\begin{aligned} \int_{Y_2} (E^2[h(y_1)|y_2=\xi], f(\xi))_1 P_{y_2}^1(d\xi) &= \int_{Y_2} \left(\int_{Y_1} h(n) P_{y_1|y_2}^2(dn|\xi), f(\xi) \right)_1 P_{y_2}^1(d\xi) \\ &= \int_{Y_2} \int_{Y_1} (h(n), f(\xi))_1 P_{y_1|y_2}^2(dn|\xi) P_{y_2}^1(d\xi) \\ &= \int_{Y_2} \int_{Y_1} (h(n), f(\xi))_1 P_{y_2|y_1}^2(d\xi|y_1=n) g^1(n) g^2(\xi) P_{y_1}^1(dn) \end{aligned}$$

where, in the next to the last line, we have used continuity property of inner product in pulling out $P_{y_1|y_2}^2(dy_1|\xi)$. Now, pulling the integration over Y_2 into the inner product, we further obtain

$$\begin{aligned} &= \int_{Y_1} P_{y_1}^1(dn) [(h(n), \int_{Y_2} g^2(\xi) f(\xi) P_{y_2|y_1}^2(d\xi|y_1=n))_1 g^1(n)] \\ &= \int_{Y_1} P_{y_1}^1(dn) (h(n), g^1(n) E^2[g^2(y_2)f(y_2)|y_1=n])_1 \end{aligned}$$

which is the desired result. $\square$

Now, using (D-3) in (D-2) we obtain

$$\hat{J}(\gamma; h) = \frac{1}{2} \langle h, h \rangle_1 + \frac{1}{2} \int_{Y_1} (h(n), g^1(n) D_{21}^2 D_{22}^1 D_{21}^2 E^2[g^2(y_2) E^2[h(y_1)|y_2] | y_1=n])_1 P_{y_1}^1(dn)$$

$$-\frac{1}{2} \langle h, D_{12}^1 D_{21}^2 \bar{P}_1 | h \rangle_1 - \frac{1}{2} \langle h, D_{21}^{2*} D_{12}^1 \bar{P}_1^* | h \rangle_1$$

which verifies (26).

To verify (25), we apply the result of Lemma D.1 to (D-1) to obtain

$$\begin{aligned} \delta J(\gamma; h) &= \langle h, \gamma \rangle_1 + \langle h, g^1(y_1) D_{21}^{2*} D_{22}^1 \{ F_2^2 E^2[g^2(y_2) E^2[x|y_2]|y_1] \\ &\quad + D_{21}^2 E^2[g^2(y_2) E^2[\gamma(y_1)|y_2]|y_1] \} \rangle_1 - \langle h, F_1^1 E^1[x|y_1] \rangle_1 \\ &\quad - \langle h, g^1(y_1) D_{21}^{2*} F_2^1 E^2[g^2(y_2) E^1[x|y_2]|y_1] \rangle_1 - \langle h, (D_{12}^1 D_{21}^2 \bar{P}_1 | 1 + D_{21}^{2*} D_{12}^1 \bar{P}_1^* | 1) \gamma \rangle_1 \\ &\quad - \langle h, D_{12}^1 F_2^1 E^1[E^2[x|y_2]|y_1] \rangle_1 \equiv \langle h, \gamma \rangle_1 - \langle h, \tilde{z} \gamma \rangle_1 - \langle h, \beta \rangle_1 \end{aligned}$$

where $\tilde{z}$ and β are defined by (27a) and (27b), respectively. This then completes the verification of (25) and (26).

2. Derivation of an Expression for $P_{1|1}^*$, the adjoint of $P_{1|1}$

Firstly note that

$$\begin{aligned} \int_{Y_1} (P_{1|1}^* \gamma(y_1), h(y_1))_1 P_{y_1}^1(dy_1) &= \int_{Y_1} (\gamma(y_1), P_{1|1} h(y_1))_1 P_{y_1}^1(dy_1) \\ &\equiv E^1[(\gamma(y_1), E^1[E^2[h(y_1)|y_2]|y_1])_1] = E^1[(\gamma(y_1), E^2[h(y_1)|y_2])_1] \end{aligned}$$

where we have used the smoothing property of conditional expectation under the probability measure $P_{y_1}^1$. Now, a further conditioning under $P_{y_1|y_2}^1$ yields

$$= E^1[(E^1[\gamma(y_1)|y_2], E^2[h(y_1)|y_2])_1],$$

and using (D-3) [cf. Lemma D-1] this becomes equivalent to

$$= E^1[(g^1(y_1) E^2[g^2(y_2) E^1[\gamma(y_1)|y_2]|y_1], h(y_1))_1],$$

thus proving (32). The first expression in (32) follows by routine manipulations.

Appendix E

In this appendix we show that the Stackelberg solution satisfying (10)-(11) is indeed an equilibrium solution--the so-called strong equilibrium of a decision problem with a modified (dynamic) information pattern. Towards this end, let us replace the original decision problem with one in which the decision (action) variables are $\gamma_1 \in \Gamma_1$ and $\gamma_2 \in \Gamma_2$, for DM1 and DM2, respectively, and the information pattern is dynamic (for DM2), with DM2 having access to the decision γ_1 of DM1. Let $\tilde{U}_1$ and $\tilde{U}_2$ denote the strategy spaces of DM1 and DM2, respectively, under this new information pattern; furthermore denote their generic elements by β_1 and β_2 , respectively. Now, since DM1 has static information, all permissible policies β_1 will be constant mappings: $\beta_1: \Gamma_1 \rightarrow \Gamma_1$, and hence $\tilde{U}_1 = \Gamma_1$. For DM2, on the other hand, all permissible policies will be measurable mappings $\beta_2: \Gamma_1 \rightarrow \Gamma_2$. Finally, let $\tilde{J}_i: \tilde{U}_1 \times \tilde{U}_2 \rightarrow \mathbb{R}$ be the cost function of DMi, satisfying the boundary condition

$$\tilde{J}_i(\beta_1, \beta_2) = J_i(\gamma_1, \gamma_2) \quad , \quad \forall \beta_1 \equiv \gamma_1 \in \Gamma_1 \equiv \tilde{U}_1 \quad , \quad (E-1)$$

where $\gamma_2 \in \Gamma_2$ is uniquely defined for each $\gamma_1 \in \Gamma_1$ by

$$\gamma_2 = \beta_2(\gamma_1) \quad \text{in } \Gamma_2 \quad . \quad (E-2)$$

Now, let $(\gamma_1^s, \gamma_2^s) \in \Gamma_1 \times \Gamma_2$ be a Stackelberg solution to the original decision problem with the unique mapping T_2 satisfying (10). Note that $T_2 \in \tilde{U}_2$, and hence relabelling T_2 as β_2^s , and γ_1^s as β_1^s , in (10) and (11), we obtain in view of (E-1)-(E-2)

$$\begin{aligned} \tilde{J}_1(\beta_1^s, \beta_2^s) &\leq \tilde{J}_1(\beta_1, \beta_2^s) & \forall \beta_1 \in \tilde{U}_1 \\ \tilde{J}_2(\beta_1, \beta_2^s) &\leq \tilde{J}_2(\beta_1, \beta_2) & \forall (\beta_1, \beta_2) \in \tilde{U}_1 \times \tilde{U}_2, \end{aligned}$$

which clearly indicate that $(\beta_1^s, \beta_2^s) \in \tilde{U}_1 \times \tilde{U}_2$ is a noncooperative Nash equilibrium. This is, in fact, a stronger equilibrium (called "strong equilibrium" [17]) because the second inequality is satisfied not only for $\beta_1 = \beta_1^s$, but for all $\beta_1 \in \tilde{U}_1$.

References

- [1] M. Loève, Probability Theory, New York: Van Nostrand Reinhold, 1963.
- [2] T. S. Ferguson, Mathematical Statistics, New York: Academic Press, 1967.
- [3] R. B. Ash, Real Analysis and Probability, New York: Academic Press, 1972.
- [4] T. Başar, "Equilibrium solutions in two-person quadratic decision problems with static information structures," IEEE Transactions on Automatic Control, vol. AC-20, no. 3, pp. 320-328, June 1975.
- [5] T. Başar and G. J. Olsder, Dynamic Noncooperative Game Theory, London: Academic, 1982.
- [6] L. A. Liusternik and V. J. Sobolev, Elements of Functional Analysis, New York: Ungar, 1961.
- [7] R. Radner, "Team-decision problems," Ann. Math Statist., vol. 33, no. 3, pp. 857-881, 1962.
- [8] T. Başar, "Decentralized multicriteria optimization of linear systems," IEEE Transactions on Automatic Control, vol. AC-23, no. 2, pp. 233-243, April 1978.
- [9] A. Bagchi and T. Başar, "Team decision theory for linear continuous-time systems," IEEE Transactions on Automatic Control, vol. AC-25, no. 6, pp. 1154-1161, December 1980.
- [10] T. Başar, "Decision Problems with multiple probabilistic models," Proc. 1983 American Control Conference, San Francisco, CA. pp. 1091-1096, June 1983.
- [11] A. V. Balakrishnan, Applied Functional Analysis, New York: Springer-Verlag, 1976.
- [12] G. Papavassilopoulos, "Solution of some stochastic quadratic Nash and leader-follower games," SIAM J. Optimiz. Contr., vol. 19, pp. 651-666, September 1981.
- [13] L. W. Kantorovich and G. P. Akilov, Functional Analysis, 2nd edition, Moscow: Nauka Publishing Co., 1977; English translation of the 1st edition: Functional Analysis in Normed Spaces, New York: MacMillan, 1964.
- [14] J. Marschak and R. Radner, Economic Theory of Teams, New Haven, CT: Yale University Press, 1972.
- [15] Y. C. Ho, "Team decision theory and information structures," Proc. IEEE, vol. 68, no. 6, pp. 644-654, June 1980.

- [16] Y. C. Ho and K. C. Chu, "Team decision theory and information structures in optimal control problems - Part I," IEEE Transactions on Automatic Control, vol. AC-17, no. 1, pp. 15-22, February 1972.
- [17] T. Başar and A. Haurie, "Feedback equilibria in differential games with structural and modal uncertainties," in Advances in Large Scale Systems, vol. 1, J. B. Cruz, Jr. (edt.), JAI Press, Inc., Conn., 1983.
- [18] J. C. Harsanyi, "Games with incomplete information played by "Bayesian" players: Part I. The Basic Model," Management Science, vol. 14, no. 3, pp. 159-182, November 1967.
- [19] J. C. Harsanyi, "Games with incomplete information played by "Bayesian" players: Part II. Bayesian Equilibrium Points," Management Science, vol. 14, no. 5, pp. 320-334, January 1968.
- [20] J. C. Harsanyi, "Games with incomplete information played by "Bayesian" players Part III. The Basic probability distribution of the game," Management Science, vol. 14, no. 7, pp. 486-502, March 1968.
- [21] M. Aoki, "On decentralized linear stochastic control problems with quadratic cost," IEEE Transactions on Automatic Control, vol. AC-18, pp. 243-250, June 1973.
- [22] H. K. Khalil and P. V. Kokotovic, "Control strategies for decision makers using different models of the same system," IEEE Transactions on Automatic Control, vol. AC-23, no. 2, pp. 289-298, April 1978.
- [23] V. Saksena and J. B. Cruz, Jr., "A multi-model approach to stochastic Nash games," Automatica, vol. 18, no. 3, pp. 295-306, May 1982.
- [24] V. Saksena and T. Başar, "A multi-model approach to stochastic team problems," Automatica, vol. 18, no. 6, November 1982.
- [25] T. Başar, "Hierarchical decision making under uncertainty," in Dynamic Optimization and Mathematical Economics, P. T. Liu (edt.), Plenum Press, NY, pp. 205-221, 1980.
- [26] A. Bagchi and T. Başar, "Stackelberg strategies in linear-quadratic stochastic differential games," J. Optimiz. Theory and Appl., vol. 35, no. 3, pp. 443-464, November 1981.
- [27] T. Başar, "An equilibrium theory for multi-person decision making with multiple probabilistic models: Symmetric and Asymmetric modes of decision making," Technical report, Coordinated Science Laboratory, University of Illinois, 1983.
- [28] J. H. Case and G. Kimeldorf, "On Nash equilibrium points and games of imperfect information," J. Optimiz. Theory and Appl., vol. 9, no. 5, pp. 302-323, May 1972.

Stackelberg Strategies and Incentives in Multiperson Deterministic Decision Problems

YING-PING ZHENG, TAMER BASAR, FELLOW, IEEE, AND JOSE B. CRUZ, JR., FELLOW, IEEE

Abstract—In this paper discrete and continuous-time two-person decision problems with a hierarchical decision structure are studied and applicability and appropriateness of a function-space approach in the derivation of causal real-time implementable optimal Stackelberg (incentive) strategies under various information patterns are discussed. Results on existence and derivation of incentive strategies for dynamic games formulated in abstract inner-product spaces, in the absence of any causality restriction on the leader's policies, are first presented and then these results are extended (and specialized) in two major directions: 1) discrete-time dynamic games with informational advantage to the leader at each stage of the decision process, which involves partial observation of the follower's decisions; and derivation of multistage incentive strategies for the leader under a feedback Stackelberg solution adapted to the feedback information pattern; and 2) derivation of causal, physically realizable optimum affine Stackelberg policies for both discrete and continuous-time problems, in terms of the gradients of the cost functionals evaluated at the optimum (achievable) operating point (which is in some cases the globally minimizing solution of the leader's cost functional). The paper is concluded with some applications of the theory to important special cases, some extensions to infinite-horizon problems, and some numerical examples that further illustrate these results.

I INTRODUCTION AND A GENERAL DESCRIPTION OF THE STACKELBERG PROBLEM

A. General Introduction

THE PRESENCE of multiple decisionmakers is a common phenomenon in many large-scale decision problems, especially if they involve humanistic and socioeconomic elements. The decisionmakers may have noncommensurable, and at times conflicting, preferences, or they may have basically the same goal but may wish to decentralize the decisionmaking process in order to alleviate the heavy burden of acquiring, transmitting, and processing the excessive amount of information needed for a centralized control (Athans [1], Başar and Cruz [5]). In either case, the decisionmakers (or players, in the terminology of game theory) would have different objective func-

tions, and acquire possibly different information in the decisionmaking process. Furthermore, there would be an order in which the decisionmakers act and/or announce their policies, and this order would either be fixed (predetermined) or determined as a consequence of the players' actions. All these factors contribute to the concept of solution to be adopted for a general multiperson decisionmaking problem, and have to be taken into account before the derivation of the solution process.

There is a growing variety of solution concepts in dynamic game theory (such as team solution, Pareto solution, Nash solution, etc.; see Başar and Cruz [5], Başar and Olsder [7]), and among these the Stackelberg (or leader-follower) strategies (Cruz [12], [13]) have recently attracted more and more attention, in both the control and economics literatures. This concept was first introduced by Von Stackelberg [26] for a class of static decision problems arising in economics. Then its dynamic version was presented in a control theoretic framework by Chen and Cruz [11] and Simaan and Cruz [24], [25]. This solution concept is especially suitable for hierarchical multilevel decision problems wherein the decisionmakers hold nonsymmetric roles in the decisionmaking process. One of the players, called the leader, occupies a higher decision level, and this superior position enables him to announce his strategy in advance and enforce it on the other players. By taking into account the optimal responses of the followers, which may be determined as the solution of some other multiperson decision problem under a specific solution concept relevant to that problem (see [18], [3]), the leader seeks the policy that leads to a most favorable outcome for him.

Such situations arise in many real world problems. In a large organization, the headquarters decisionmaker (the leader) cannot dictate every subdivision's (the follower's) task in fine detail; in its stead, it simply announces and executes appropriate strategies (policies), such as the resource allocation strategy, penalty or reward policy, the profit-sharing policy, etc., so as to induce the subdivisions to work in accordance with the interests of the entire organization ([2], [15], [16], [19], [22]). Some recent investigations have been devoted to the study and construction of such leader-follower strategies in special types of organizations. For example, a standard and efficient way for a government (the leader) to solve the water pollution prob-

Manuscript received May 21, 1983; revised June 15, 1983. The work of T. Başar and J. B. Cruz, Jr. was supported in part by the Joint Services Electronics Program under Contract N00014-79-C-0424 and in part by the Office of Naval Research under Contract N00014-82-K-0469.

Y. P. Zheng was with Coordinated Science Laboratory, University of Illinois, 1101 W. Springfield Avenue, Urbana, IL 61801. He is now with the Institute of Automation, Academia Sinica, Beijing, People's Republic of China.

T. Başar and J. B. Cruz, Jr. are with the Decision and Control Laboratory, Coordinated Science Laboratory, University of Illinois, 1101 W. Springfield Avenue, Urbana, IL 61801.

lem is to design some subsidy programs or penalty policies to encourage or induce the chemical plants (the followers) to act cooperatively. A utility company (the leader) may use a price strategy (or a pricing strategy) to induce the customers (the followers) to consume the utility resource more reasonably ([18], [20]). In a market with both free competition and government adjustment, the government (the leader) may design a strategy of adjusting the effective income of the potential buyers of the commodity so as to induce the competing duopolistic firms to cooperate and achieve a Pareto-optimal solution [23]. All of these problems can be studied in the framework of Stackelberg dynamic game theory, thus making this new field very promising in applications.

B. General Description of the Stackelberg Problem

To be more precise in our description of a Stackelberg game and the related solution concepts, let us now consider a two-person dynamic game problem with a hierarchical decision structure under which player 1 acts as the leader and player 2 as the follower. The state $x(\cdot)$ of the underlying decision process evolves according to either (in continuous time)

$$\dot{x}(t) = f(t, x(t), u(t), v(t)), \quad t \in [0, T] \quad (1)$$

or (in discrete time)

$$x(k+1) = f(k, x(k), u(k), v(k)), \\ k = 0, 1, \dots, N-1, \quad (2)$$

where $u(\cdot)$ is the leader's decision variable and $v(\cdot)$ is the follower's, and they are either time-functions or time-series, belonging to the corresponding Hilbert spaces.

$$u(\cdot) \in L_2^m[0, T], \quad v(\cdot) \in L_2^n[0, T], \\ \text{in continuous time;}$$

or

$$u(\cdot) \in l_2^m[0, N-1], \quad v(\cdot) \in l_2^n[0, N-1], \\ \text{in discrete time.}$$

Generically, let us denote the decision variables of the leader and the follower by u and v , respectively, and the decision spaces by U and V . Furthermore, let $X = L_2^s[0, T]$ or $l_2^s[0, N-1]$ denote the state space for the process, where $x(\cdot)$ belongs, and let $Y_1 \subseteq X \times V$ and $Y_2 \subseteq X \times U$ denote the information (observation) spaces of the leader and the follower, respectively. A permissible policy (strategy) $\gamma_i \in \Gamma_i$ for player i is a Borel-measurable mapping from his observation space into his decision space, satisfying some additional regularity conditions like causality, Lipschitz continuity, etc. that will be delineated later in proper contexts. One underlying assumption here is that, with $x_0 \in \mathbb{R}^n$ fixed, to each $(\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2$, there corresponds a unique state trajectory $x(\cdot) \in X$ and a unique pair of cost values $\{J_1(\gamma_1, \gamma_2), J_2(\gamma_1, \gamma_2)\}$.

The Stackelberg game problem involves, in a nutshell, determination of a leader's policy $\gamma_1^* \in \Gamma_1$ satisfying

$$J_1(\gamma_1^*, T(\gamma_1^*)) = \min_{\gamma_1} J_1(\gamma_1, T(\gamma_1)) \quad (3)$$

where $T: \Gamma_1 \rightarrow \Gamma_2$ is the unique rational response mapping of the follower, i.e.,

$$T(\gamma_1) = \arg \min_{\gamma_2 \in \Gamma_2} J_2(\gamma_1, \gamma_2) \quad (4)$$

where we tacitly assumed existence of a unique solution to (4).

Even though this definition is valid for all types of information available to the players (i.e., for arbitrary Y_1 and Y_2), the derivation of the solution will depend to a great extent on the underlying information structure, as to be elucidated in the sequel.

1) *Open-Loop Information Structure*: The players' information comprises only the *a priori* information, e.g., the structural parameters of the problem and the initial conditions. In this case, strategies and the decision variables coincide, and are chosen as time-functions from the beginning.

Necessary conditions for the open-loop solution of Stackelberg dynamic game problems can be obtained without any conceptual difficulties, although it is rather difficult to solve them analytically or even numerically [13], [7, ch. 7].

2) *Closed-Loop Information Structure*: Here the leader is assumed to acquire state information with memory, i.e., elements of Y_1 are given by $y_1(t) = \{x(\tau), \tau \leq t\}$ or $y_1(k) = \{x(i), i = 0, \dots, k\}$, thus leading to policies in the form $u(t) = \gamma_1(t; x(\tau), \tau \leq t)$ or $u_1(k) = \gamma_1(k; x(k), x(k-1), \dots, x(0))$. The follower, on the other hand, could acquire closed-loop or open-loop information. Any direct approach towards the solution of such dynamic Stackelberg games meets with formidable difficulties, since the optimization problem (4) is "structurally" dependent on the structure of leader's strategy, that is, the follower faces an optimization problem "parameterized" by the structure of γ_1 . Such "nonclassical" optimal control problems and indirect ways of obtaining the solution have been discussed in many papers: see [21], [8], [9], [6], [3], [27], and it has been shown that in certain cases the leader can achieve the global minimum value of his cost function J_1 . This feature has also been established by an "indirect method" [4] under two conditions: 1) the leader can detect the follower's action (detectability); and 2) by choosing an appropriate strategy the leader is able to threaten the follower by severe punishment in case of any deviation from the desired solution trajectory (enforceability). It has been shown, moreover, that the closed-loop information is rich enough to allow for the solution to satisfy additional design specifications. One such specification involves a "robustness" feature: that is, in case of a deviation from an optimal path, not to punish the follower indefinitely at all future stages, but rather use an effective threat policy which would carry a punitive action role for only a few (two or three) stages. This aspect of the problem and its solution has been discussed in some recent papers in the literature, see e.g., [27].

3) *Feedback Strategies and the Feedback Stackelberg Solution Concept*: A subclass of closed-loop strategies com-

prises those policies that depend only on the current value of the state without memory. That is

$$y_1(t) = \{x(t)\} \quad \text{or} \quad y_1(k) = \{x(k)\}.$$

Under such feedback information pattern for the leader, the Stackelberg solution is still very difficult to obtain, and in fact, in most cases, it will not even exist if the initial state $x(0)$ is taken a variable and a solution is sought for all $x(0) \in \mathbb{R}^n$. A way to circumvent this difficulty, in the case of discrete-time problems, is to require that any subproblem-to-go is also an optimal process in the Stackelberg sense [25]. This permits the adoption of a dynamic programming type approach which involves the solution of static Stackelberg games at each stage (and in retrograde time). In comparison with the Stackelberg solution, the feedback Stackelberg solution gives only a suboptimal solution; though it has the advantage of being simpler in structure, computationally feasible and implementable. Furthermore, it has better robustness properties against noise and disturbance, since the leader can update on his policy at each stage of the decision process.

4) *Incentive Strategies:* As we have indicated in case 2), the leader may expect to achieve the global optimum of his cost function (the team solution), as though the follower was cooperating with him, provided that he has memory, can detect follower's actions and can announce and implement enforceable policies. In order to investigate the Stackelberg problem from this viewpoint, we include the decision $v(\cdot)$ of the follower directly in the information $y_1(\cdot)$ available to the leader, thereby allowing the leader to adopt a strategy in the form

$$u(\cdot) = \gamma_1(x(\cdot), v(\cdot)) \quad (5)$$

which explicitly displays the dependence of the leader's decision variable u on the follower's, v . Such a dependence is not necessarily instantaneous, and may involve delays; furthermore, y_1 may carry only partial information on v , such as the one obtained through the present and past values of the state. Whatever the nature of the dependence is, such a structure (as in (5)) is called an incentive strategy, because it displays the extent of the leader's power in enforcing a certain action on the follower through a punishment or reward scheme and by utilizing the information acquired through y_1 ([18], [28]).

C. Outline of the Following Sections

This paper is devoted to an extensive discussion and derivation of closed-loop Stackelberg strategies and incentive policies in dynamic decision problems of the types introduced above, and an elaboration on their properties. In the next section we first discuss the incentive decision problem when the leader's permissible strategies are of the form (5), in abstract inner-product spaces, and present some general results on the existence and derivation of linear incentive policies. These results are then extended in Sections III and IV in two different directions. In Section III we treat the discrete-time Stackelberg problem with

dynamic informational advantage to the leader at each stage of the game, and under the feedback Stackelberg solution concept. General conditions are obtained for existence of a solution and for this solution to coincide with the global Stackelberg solution. In Section IV, we extend the results of Section II to derivation of *causal* incentive schemes and construction of real-time closed-loop Stackelberg strategies from a normal-form description, in both discrete and continuous time. Some applications to important special cases with illustrative numerical examples are given in Section V.

II SOME GENERAL RESULTS ON EXISTENCE AND DERIVATION OF OPTIMAL INCENTIVE STRATEGIES

In this section we consider an abstract reformulation of the dynamic game problem of Section I-B with the leader allowed to have a partial measurement of the follower's decision variable v . Towards this end, let U and V be Hilbert spaces, with elements u and v , respectively, and the cost functional J_i of player i ($i = 1, 2$) be a mapping from $U \times V$ into $\mathbb{R}$. In this reformulation, the dynamic nature of the decision process is suppressed, $\Gamma_2 \equiv V$, and Γ_1 is the class of all Borel-measurable mappings from Y into U , where Y is a Hilbert space comprising observations of the form

$$y = Nr$$

where $N: V \rightarrow Y$ is a linear operator with full range in Y . The case when N is invertible is known as the perfect information case; otherwise we say that the leader has only partial information on the actions of the follower.

A. Perfect Information Case

Let $(u^d, v^d) \in U \times V$ be a desirable solution from the point of view of the leader—this point could, for example, be chosen as the global minimizer of the leader's cost function $J_1(u, v)$ over $U \times V$, if such a solution exists. Then, an optimal incentive policy for the leader is one that forces the follower to choose the decision v^d , by making the incurred cost corresponding to $v \neq v^d$ sufficiently large; in other words, for a given incentive strategy γ_1 to be implementable it should satisfy the strict inequality

$$J_2(u = \gamma_1(v), v) > J_2(u^d, v^d), \quad \text{for all } v \neq v^d, \quad v \in V \quad (6)$$

together with the side condition

$$\gamma_1(v^d) = u^d. \quad (7)$$

To formalize this concept, we introduce the set

$$\Omega_d = \{(u, v) \in U \times V: J_2(u, v) \leq J_2(u^d, v^d)\} \quad (8)$$

and immediately arrive at the following result.

Proposition 1: A desired decision pair $(u^d, v^d) \in U \times V$ can be induced by an incentive strategy $\gamma_1 \in \Gamma_1$, if to each $v \in V$, $v \neq v^d$, there corresponds a $u = \gamma_1(v) \in U$ such that $(u, v) \notin \Omega_d$. A strategy that accomplishes this is the

so-called (discontinuous) threat policy given by

$$\gamma_1(v) = \begin{cases} u^d, & \text{if } v = v^d \\ \text{any } u \text{ such that } (u, v) \notin \Omega_d, & \text{if } v \neq v^d. \end{cases} \quad (9)$$

Remark 1: The preceding proposition provides a sufficient condition for existence of an optimal incentive strategy. This condition is also necessary if we make an additional behavioral assumption on the follower, which is that on the boundary of Ω_d (which is his indifference curve) he chooses points that are detrimental to the leader.

The next proposition shows that the hypothesis of Proposition 1 is satisfied for an important class of problems.

Proposition 2: If $J_2(u, v)$ is continuous and strictly convex on $U \times V$, any desired decision pair $(u^d, v^d) \in U \times V$ is inducible by an appropriate incentive strategy.

Proof: Since $J_2(u, v)$ is continuous and strictly convex, the set Ω_d is closed and strictly convex. We now prove the proposition by contradiction. Assume that there exists a $(u^d, v^d) \in U \times V$ which cannot be induced by an appropriate incentive strategy. That is there exists a $\bar{v} \in V$, $\bar{v} \neq v^d$, such that $(\bar{u}, \bar{v}) \in \Omega_d$ for every $\bar{u} \in U$. Let $v_\alpha = \alpha v^d + (1 - \alpha)\bar{v}$, $0 < \alpha < 1$; then $(\bar{u}, v_\alpha) = \alpha(u^d, v^d) + (1 - \alpha)(\bar{u}, \bar{v}) \in \Omega_d$, where $u_\alpha = (1/1 - \alpha)(\bar{u} - \alpha u^d) \in U$. When $\alpha \rightarrow 1$, $(\bar{u}, v_\alpha) \rightarrow (\bar{u}, v^d)$ and hence the limit point $(\bar{u}, v^d)$ belongs to Ω_d for every $\bar{u} \in U$. In particular, if $\bar{u}$ is chosen as $u^d + u_0$ and $u^d - u_0$ ($u_0 \in U$), the convex combination $(u^d, v^d) = (1/2)(u^d + u_0, v^d) + 1/2(u^d - u_0, v^d)$ should be an inner point of the strictly convex set Ω_d . This is contradictory to the fact that (u^d, v^d) is a boundary point of Ω_d and this completes the proof.

Incentive policies that induce the pair (u^d, v^d) , under the hypotheses of Proposition 2 are not only of the type (9), but could also be continuous and even continuously differentiable. However, if we further restrict the class of incentive strategies to affine ones (because of their simple structure), we have to impose an additional restriction on J_2 , as elucidated in the Proposition 3 below, whose proof can be found in [28].

Proposition 3: For an incentive Stackelberg game, let $J_2(u, v)$ be strictly convex and Fréchet differentiable on $U \times V$, and its gradient with respect to u , evaluated at the desired decision point $(u^d, v^d) \in U \times V$, does not vanish, i.e.,

$$\nabla_u J_2(u^d, v^d) \neq 0. \quad (10)$$

Then, the desired decision pair can be induced by an affine incentive strategy

$$\gamma_1(v) = u^d - Q(v - v^d) \quad (11)$$

where $Q: V \rightarrow U$ is a linear operator whose adjoint $Q^*: U \rightarrow V$ satisfies the equation

$$\nabla_u J_2(u^d, v^d) = Q^* \nabla_u J_2(u^d, v^d) \quad (12)$$

which admits at least one solution under (10).

It should be noted that whenever a global minimum to $J_2(u, v)$ exists on $U \times V$ (say, (u^*, v^*)), by letting (u^d, v^d)

$= (u^*, v^*)$ above in (11) and (12), the leader can force the follower to minimize collectively the leader's cost functional $J_1(u, v)$.

B. Partial Dynamic Information

If the leader does not have access to v , he cannot necessarily enforce an arbitrary decision pair $(u^d, v^d) \in U \times V$ on the follower, and, in particular, (u^d, v^d) is in general not achievable. In fact, achievable solution pairs will be elements of the product space $U \times Y$, with the best achievable performance for the leader being [4].

$$\min_{U \times Y} \bar{J}_1(u, y) \quad (13)$$

where

$$\bar{J}_1(u, y) = J_1(u, v^*(u, y)) \quad (14)$$

$$v^*(u, y) = \arg \left\{ \min_{v \in V} J_2(u, v) \text{ subject to } Nv = y \right\}. \quad (15)$$

Here we have tacitly assumed that in (15) the argument is unique for every $(u, y) \in U \times Y$, which in fact holds whenever $J_2(u, v)$ is strictly convex on $U \times V$ [28, Lemma 2]. Further introducing

$$\bar{J}_2(u, y) = J_2(u, v^*(u, y))$$

it can be shown [28] that strict convexity of $J_2(u, v)$ implies strict convexity of $\bar{J}_2(u, y)$ on $U \times Y$, and hence the incentive problem with partial dynamic information becomes equivalent to one with perfect information, with $(u, y) \in U \times Y$ being the decision variables and $\bar{J}_i(u, y)$, $i = 1, 2$, the cost functionals. Propositions 1-3 apply directly to this transformed, or so-called "projected" problem, provided that the desirable solution pair (u^d, y^d) is chosen out of $U \times Y$. In this context, a direct application of Proposition 3 leads to affine optimal incentive policies

$$\gamma_1(y) = u^d - Q(y - y^d) \quad (16)$$

where $Q^*: U \rightarrow Y$ satisfies

$$\nabla_y \bar{J}_2(u^d, y^d) = Q^* \nabla_u \bar{J}_2(u^d, y^d) \quad (17)$$

provided that $\bar{J}_2(u, y)$ is Fréchet differentiable on $U \times Y$ and $\nabla_u \bar{J}_2(u^d, y^d) \neq 0$.

Obviously, the operator Q^* in either (12) or (17) is not uniquely defined. Thus, there exist several candidates for the solution of the incentive problem at our disposal to satisfy some additional requirements. Some possible ways of constructing the operator Q^* , with application examples and other details on these approaches can be found in [28]. Yet another possible selection criterion based on sensitivity considerations has been presented and discussed in [10].

III. THE FEEDBACK STACKELBERG GAME WITH STAGewise INFORMATION ADVANTAGE TO THE LEADER

As one application of the general results presented in the previous section, we consider here a feedback dynamic game in discrete-time, as described by the state evolution (2) and with player i 's cost function given by $J^i(\lambda, u, v)$.

where

$$\begin{aligned} J_i^N(x(k), u_k^{N-1}, v_k^{N-1}) \\ = \sum_{j=k}^{N-1} g_i(j, x(j), u(j), v(j)) + g_i(N, x(N)) \\ u_k^{N-1} = \{u(k), u(k+1), \dots, u(N-1)\}, \\ v_k^{N-1} = \{v(k), v(k+1), \dots, v(N-1)\} \\ u = u_0^{N-1}, \quad v = v_0^{N-1}, \\ u(k) \in \mathbb{R}^{m_1}, \quad v(k) \in \mathbb{R}^{m_2}, \quad x(k) \in \mathbb{R}^n, \\ k = 0, 1, \dots, N-1. \end{aligned} \quad (18)$$

We endow the leader with such an information pattern that permits him to use incentive strategies under partial observation of the follower's current actions; that is, letting

$$\begin{aligned} y_1(k) = y(k) = N_k(x(k), v(k)) \in Y_k \subseteq \mathbb{R}^p; \\ N_k: \mathbb{R}^n \times \mathbb{R}^{m_2} \rightarrow \mathbb{R}^p \end{aligned} \quad (19)$$

we assume that permissible policies for the leader are Borel measurable mappings

$$\gamma_1(k; \cdot): \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^{m_1} \quad (20)$$

so that

$$u(k) = \gamma_1[k; x(k), y(k)]. \quad (21)$$

For the follower, on the other hand, we assume that only feedback state information is available, i.e.,

$$v(k) = \gamma_2[k; x(k)]. \quad (22)$$

What we envisage here is a decision making process wherein the leader is dominant only stagewise, not only by announcing his policy ahead of the follower but also by incorporating partial information on the follower's current action in his incentive strategy. More precisely, the rules that underlie the game are as follows: At each stage $k = 0, 1, \dots, N-1$, the leader announces his strategy $u(k) = \gamma_1[k; x(k), y(k)]$ first, to which the follower reacts by minimizing his stagewise cost function. This then determines the values of $y(k)$, $u(k)$, $v(k)$ and $x(k+1)$ in terms of $x(k)$, and transition to the next stage takes place. Of course, while making decisions at each stage, the players will have to anticipate their future moves and arrive at their policies accordingly. At each stage a dynamic Stackelberg game (incentive) problem of the type discussed in Section II is solved, with the leader, not only announcing his policy ahead of the follower, but also having informational advantage (partial information on the follower's decision). We call such a game a "feedback Stackelberg game with informational advantage to the leader" and the associated solution concept the "feedback Stackelberg solution with informational advantage to the leader" (FSIA). Note that this solution concept coincides with the standard feedback Stackelberg concept (cf. [24], [25]) in the case $N_k(x(k), v(k))$ is independent of $v(k)$.

We now discuss derivation of the FSIA for the finite horizon multistage decision process formulated in this section.

Let us consider the last step decision problem starting from $x(N-1)$, with only $u(N-1)$ and $v(N-1)$ to be determined (the problem $(N-1)$). Following the discussion of Section II-B, the best response of the follower to fixed values of $x(N-1)$, $u(N-1)$ and $v(N-1) = N_{N-1}(x(N-1), u(N-1))$ will be

$$\begin{aligned} v(N-1) \\ = \arg \left\{ \min_{v \in V(N-1)} J_2^{N-1}(x(N-1), u(N-1), v(N-1)), \right. \\ \left. v_{N-1} \in N_{N-1}^{-1}(x(N-1), y(N-1)) \right\} \\ \triangleq v_{N-1}^*(N-1, u(N-1), y(N-1)). \end{aligned} \quad (23)$$

where $N_{N-1}^{-1}(x, y) = \{v \in \mathbb{R}^{m_2}: N_{N-1}(x, v) = y\}$, thus, leading to the "projected" cost functionals

$$\begin{aligned} \tilde{J}_i^{N-1}[x(N-1), u(N-1), y(N-1)] \\ \triangleq J_i^{N-1}[x(N-1), u(N-1), v_{N-1}^*(x(N-1), \\ \times u(N-1), y(N-1))] \quad (i = 1, 2). \end{aligned} \quad (24)$$

Therefore, the lowest cost value the leader can hope to attain is

$$\begin{aligned} J_1^{N-1}[x(N-1)] \triangleq \min_{u \in U(N-1)} \tilde{J}_1^{N-1} \\ \cdot [x(N-1), u(N-1), y(N-1)] \end{aligned} \quad (25)$$

Let us assume that, for each $x(N-1) \in \mathbb{R}^n$, there exists a unique solution $(u'(N-1), y'(N-1))$ to (25). (If the solution to (25) is not unique, we adopt one of the possible solutions according to some other consideration of preference (for the leader), see [4] for a discussion on this point.)

Now introduce the counterpart of set (8), in this context, which will depend explicitly on $x(N-1)$:

$$\begin{aligned} \Omega_{N-1}(x(N-1)) = \{u, y \in \mathbb{R}^{m_1} \times Y_{N-1} | \tilde{J}_2^{N-1} \\ \cdot (x(N-1), u, y) \leq \tilde{J}_2^{N-1}(x(N-1), u'(N-1), \\ \cdot y'(N-1))\} \end{aligned} \quad (26)$$

and let $\Omega_{N-1}^c(x(N-1))$ denote its complement. Then, we have

Definition 1. For problem $(N-1)$, a state $x(N-1)$ is called *incentive controllable* if either Y_{N-1} is a singleton or for any $y \in Y_{N-1}$, $y \neq y'(N-1)$, there exists $u \in \mathbb{R}^{m_1}$ such that $(u, y) \in \Omega_{N-1}(x(N-1))$. Furthermore, if all states $x(N-1) \in \mathbb{R}^n$ are incentive controllable, then the problem $(N-1)$ is called *completely incentive controllable*.

Now, an existence result follows immediately from Proposition 2.1:

Proposition 4: Assume that problem $(N-1)$ is completely incentive controllable. Then for each x_{N-1} there exists an incentive strategy $u(N-1) = \gamma_1[N-1, x(N-1), y(N-1)]$ which forces the follower to take the decision $v(N-1) = v_{N-1}^*(x(N-1), u'(N-1), y'(N-1))$, with the realized cost value for the leader being the minimum value of $\tilde{J}_1^{N-1}$, i.e., $J_1^{N-1}(x(N-1))$.

Remark 2: If $y(N-1) = v(N-1)$, or $v(N-1) = \mathcal{V}_2^{-1}(x(N-1), y(N-1))$ exists uniquely the leader has complete access to the follower's decision and the problem becomes one with perfect dynamic information (see Section II-A). In this case the attainable lower bound for an incentive controllable state x_{N-1} is exactly the team solution:

$$I_1^{N-1}(x(N-1)) = \min_{u(N-1), v(N-1)} J_1^{N-1}(x(N-1), u(N-1), v(N-1)) \quad (27)$$

which is obviously the absolute lower bound.

The result of Proposition 4 can now be applied recursively by simply replacing J_1^k with the cost-to-go function $\tilde{J}_1^k$ to be introduced below and by appropriately redefining $I_1^k(x(k))$. Towards this end, let

$$\tilde{J}_1^k(x(k), u(k), v(k)) \triangleq g_1^k(k, x(k), u(k), v(k)) + I_1^{k+1}(x(k+1)) \quad (28)$$

where

$$x(k+1) = f(k, x(k), u(k), v(k)) \quad (29)$$

and

$$\begin{aligned} I_1^{k+1}(x(k+1)) &= \min_{u, v} \tilde{J}_1^{k+1}(x(k+1), u, v) \\ &\triangleq \tilde{J}_1^{k+1}(x(k+1), u'(k+1), y'(k+1)) \end{aligned} \quad (30)$$

$$I_2^{k+1}(x(k+1)) = \tilde{J}_2^{k+1}(x(k+1), u'(k+1), y'(k+1)). \quad (31)$$

Construct

$$\begin{aligned} v_k^*(x(k), u(k), y(k)) &= \arg \min_{v \in \mathcal{V}_2^{-1}(x(k), y(k))} \tilde{J}_2^k(x(k), u(k), v) \\ &\quad \cdot v \in \mathcal{V}_2^{-1}(x(k), y(k)) \end{aligned} \quad (32)$$

$$\begin{aligned} \tilde{J}_i^k(x(k), u(k), y(k)) &= \tilde{J}_i^k(x(k), u(k), v_k^*(x(k), u(k), y(k))) \\ &\quad (i = 1, 2) \end{aligned} \quad (33)$$

$$\begin{aligned} (u'(k; x(k)), y'(k; x(k))) &= \arg \min_{u, y} \tilde{J}_1^k(x(k), u, y) \end{aligned} \quad (34)$$

$$I_i^k(x(k)) = \tilde{J}_i^k(x(k), u'(k), y'(k)) \quad (i = 1, 2) \quad (35)$$

$$\begin{aligned} \Omega_k(x(k)) &= \{(u, y) \in \mathbb{R}^m \times Y_k | \tilde{J}_2^k(x(k), u, y) \\ &\leq \tilde{J}_2^k(x(k), u'(k), y'(k))\}. \end{aligned} \quad (36)$$

Then, the problem considered at stage k has the projected cost functions $\tilde{J}_1^k$ and $\tilde{J}_2^k$, with a lower bound on the former given by $I_1^k(x(k))$. The following (recursive) definition now paves the way for Proposition 5, the generalization of Proposition 4.

Definition 2: The $(N-k)$ -stage problem (k) is called *completely incentive controllable*, if

1) the corresponding problem $(k+1)$ is completely incentive controllable; and 2) the equivalent one-stage incentive problem (28)–(29) is completely incentive controllable in the sense of Definition 1.

Proposition 5: For a completely incentive controllable problem (k) , and for each starting state x_k , there exists an optimal incentive strategy

$$\begin{aligned} u^*(k) &= \gamma_1^*[k; x(k), y(k)] \\ &= \begin{cases} u'(k; x(k)), & \text{when } y(k) = y'(k; x(k)) \\ u \in \mathbb{R}^m \text{ such that } (u, y) \in \Omega_k^c(x(k)), & \text{when } y(k) \neq y'(k; x(k)). \end{cases} \end{aligned} \quad (37)$$

that forces the follower to take the decision $v(k) = v_k^*(x(k), u'(k), y'(k))$, with the realized cost value for the leader being $I_1^k(x(k))$. This constitutes a FSIA solution for the dynamic game problem considered in this section.

Remark 3: Equations (28)–(35) constitute the recurrence relations between $I_1^k(x(k))$ and $I_1^{k+1}(x(k+1))$, $(i = 1, 2)$. This is the generalized optimality principle for the feedback Stackelberg game problem with informational advantage to the leader, under the assumption of the complete incentive controllability.

We now put some more structure on the underlying spaces and functionals, in order to obtain some specific results. The first set of such restraints and the main result that ensues are the following.

Proposition 6: The feedback Stackelberg game is completely incentive controllable if for each $x(k) \in \mathbb{R}^n$, and $k = 0, \dots, N-1$, Y_k is a vector space and $\tilde{J}_2^k(x(k), u, v)$ is continuous and strictly convex in the pair $(u, v) \in \mathbb{R}^m \times Y_k$.

Proof: Verification of this result involves a repeated application of Proposition 2 in a routine way, and is therefore omitted.

Corollary 1: When we construct the sequence $\{\tilde{J}_i^k(x(k), u(k), v(k))\}$ according to relations (28)–(35), and recursively from $k = N-1$ backwards, if all $\tilde{J}_2^k(x(k), u(k), v(k))$ are continuous, strictly convex in $u(k)$ and $y(k)$ for all $x(k) \in \mathbb{R}^n$ and $k = 0$, then the problem always admits a FSIA solution, with one such optimal incentive strategy given by (37).

The conditions of this corollary (and of Proposition 6) are actually satisfied for a class of problems of practical importance. Consider, for example, the following set of sufficient conditions:

- 1) $g_2(N; x)$ is convex in $x \in \mathbb{R}^n$;
- 2) $g_2(k; x, u, v)$ is decomposable in the form: $g_2(k; x, u, v) = P_2(k; x) + q_2(k; u, v)$, where $P_2(k; x)$ is convex in x and $q_2(k; u, v)$ is strictly convex in (u, v) ;
- 3) $f(k; x, u, v)$ and $N_2(x, v)$ are affine in their arguments;
- 4) $u'(k; x)$ and $y'(k; x)$ are affine in x ;
- 5) $v_k^*(x, u, y)$ is affine in x, u and y .

These guarantee satisfaction of the hypotheses of the corollary. One such special class is the linear quadratic

problem where

$$\begin{aligned} g_2(N; x) &= \langle x, Q_N x \rangle \quad (Q_N \geq 0) \\ g_2(k; x, u, v) &= \langle x, Q_k x \rangle + \langle u, R_k u \rangle + \langle v, S_k v \rangle \\ &\quad (Q_k \geq 0, R_k > 0, S_k > 0) \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes appropriate inner products in vector spaces:

$$\begin{aligned} f(k; x, u, v) &= A_k x + B_k u + C_k v \\ N_k(x, v) &= N_k v \end{aligned}$$

where A_k, B_k, C_k, N_k are matrices of appropriate dimensions, with N_k being of full rank.

It readily follows from Proposition 3 that in this case the FSIA solutions are not only of the type (37), but can also be taken to be affine, in which case

$$\begin{aligned} u^i(k) &= L_1(k)x(k), \quad y^i(k) = L_2(k)x(k) \\ v^i(k; x, u, y) &= M_1(k)x + M_2(k)u(k) + M_3(k)y(k) \\ u^*(k) &= \gamma_1^*[k; x(k), y(k)] \\ &= L_1(k)x(k) - Q_1(k)[y(k) - L_2(k)x(k)] \end{aligned}$$

with capital letters denoting matrices of appropriate dimensions, and $Q_1(k)$ being a gain matrix whose transpose satisfies a gradient equation of the type (12), for each $k \geq 0$. Explicit expressions for these matrices can be obtained by basically solving (28)–(35), recursively, and by noting that J^* and I^* are quadratic functionals for each $k \geq 0$.

Remark 4: The preceding results find natural extensions to the class of problems wherein the control and measurement spaces are arbitrary (infinite dimensional) Banach spaces, instead of being finite dimensional. Particularly, for the linear-quadratic problem discussed above, the same affine structure prevails provided that we interpret the inner-products appropriately and replace all matrices with linear operators. Such a result, then, would be applicable to continuous-time dynamic games in which the decision makers have access to sampled information and the feedback solution is defined in between different sampled subintervals.

Remark 5: Under the conditions of Proposition 4 and Corollary 3.1, and when the leader has perfect access to the follower's decision variable at each stage, the affine FSIA solution has also a robust feature in the sense that its truncated version constitutes a FSIA solution to a dynamic feedback game of shorter duration, defined on the interval $[k, N-1]$, for any $k > 0$. This result is a direct consequence of the fact that the trajectory corresponding to the original FSIA solution satisfies the principle of optimality (being the team solution from the leader's point of view) and the leader's affine FSIA strategy employs only current state information.

IV. DERIVATION OF CAUSAL STACKELBERG SOLUTIONS TO DISCRETE AND CONTINUOUS-TIME DYNAMIC GAMES

In this section we turn our attention to the global Stackelberg solution in both discrete and continuous-time

dynamic games of the type introduced in §II-B, and under the closed-loop information pattern. Here, the leader will not have any stagewise informational advantage over the follower, but he will still dominate the decision process by announcing his strategy ahead of time and enforcing it on the follower, in accordance with the solution concept (3) (4). Furthermore, because of its appealing features, we restrict attention to those strategies for the leader that are linear in the dynamic part of the information, and also assume, without any loss of generality, that the follower employs only open-loop policies (which does not lead to any degradation in his performance (see, e.g., [7])).

Let J_1 and J_2 be appropriate cost functionals for players 1 and 2, which, for fixed initial state $x_0 \in \mathbb{R}^n$, can always be rewritten (by elimination of the state variable) as functions of solely the decision variables $(u, v) \in U \times V$ (see Section I-B for notation). Since every discrete or continuous-time dynamic game can be expressed in this form, the analyses and results of Section II are directly applicable here provided that the corresponding optimal incentive strategy for the leader is permissible, i.e., it is causal and satisfies the additional structural restrictions that may be imposed on elements of Γ_1 . Specifically, let us assume that:

1) Through the closed-loop state information, the leader is able to infer perfectly the past values of $v(\cdot)$, the decision function of the follower.

2) $J_2(u, v)$ is Fréchet-differentiable and strictly convex on $U \times V$.

3) A global minimum to $J_1(u, v)$ exists on $U \times V$, which we denote by $(u^*, v^*) \in U \times V$, and which is adopted as a desirable solution by the leader.

4) At this solution point,

$$\nabla_u J_2(u^*, v^*) = 0. \quad (38)$$

Then, we know from Proposition 3 that, in the absence of causality, every optimal affine Stackelberg solution can be written as

$$u = \gamma_1(v) = u^* - Q(v - v^*) \quad (39)$$

where the adjoint of Q , $Q^*: U \rightarrow V$, satisfies

$$\nabla_v J_2(u^*, v^*) = Q^* \nabla_u J_2(u^*, v^*). \quad (40)$$

Now, the real question here is whether we can find an operator Q whose adjoint satisfies (40) above, and which is further causal and leads to a policy γ_1 , as given by (39), belonging to a given closed-loop policy space Γ_1 . We show below that, under the closed-loop pattern and for both discrete and continuous-time problems satisfying appropriate structural assumptions, such a linear operator can be constructed.

Towards this end, we first introduce some notation. Let the inner-product of two elements $f(\cdot)$ and $g(\cdot)$ in $L^2[0, T]$ be defined by

$$(f, g) \triangleq \int_0^T f(t)g(t)dt = \int_0^T g(t)f(t)dt \quad (41)$$

and further introduce the notation

$$(f, g)_{\text{causal}} \triangleq \int_0^t f(\tau)g(\tau)d\tau = \int_0^t g(\tau)f(\tau)d\tau \quad (42)$$

where $0 \leq t_1 \leq t_2 \leq T$. Similarly, for discrete-time processes, the inner-product of $f(\cdot), g(\cdot) \in l_2^N[0, N]$ is defined by

$$\langle f, g \rangle \triangleq \sum_{i=0}^N f'(i)g(i) = \sum_{i=0}^N g'(i)f(i) \quad (43)$$

and furthermore

$$\langle f, g \rangle_{(k,l)} = \sum_{i=k}^l f'(i)g(i) = \sum_{i=k}^l g'(i)f(i) \quad (44)$$

where k, l are integers, $0 \leq k \leq l \leq N$. Now introduce

$$\phi(\cdot) \triangleq \nabla_u J_2(u, v), \quad \Psi(\cdot) \triangleq \nabla_v J_2(u, v) \quad (45)$$

where the gradients are evaluated at some specific values of $u \in U$ and $v \in V$ that will be clear from the context.

To reveal a property of $\phi(\cdot)$ and $\Psi(\cdot)$ which is vital in the construction of affine incentive strategies (cf. (39)), let us consider the variation in J_2 resulting from, for example, an infinitesimal variation $\delta u(\cdot)$ in $u(\cdot)$:

$$\begin{aligned} \delta J_2 &= \langle \nabla_u J_2(u, v), \delta u(\cdot) \rangle = \int_0^T \phi'(t) \delta u(t) dt \\ &= \sum_{i=0}^N \phi'(i) \delta u(i). \end{aligned} \quad (46)$$

Thus, the value of $\phi(\cdot)$ at time t simply represents the local sensitivity of J_2 with respect to $u(t)$, in other words, the ability of the leader to influence J_2 by changing his decision variable $u(t)$ at time t . Likewise, the time function $\Psi(\cdot) = \nabla_v J_2(u, v)$ represents the follower's ability to influence his cost functional J_2 by changing $v(t)$, the value of $v(\cdot)$ at time t . Hence, they can be referred to as "sensitivity functions" representing the sensitivity of J_2 to the players' actions, which may be taken as a measure of the players' control ability in the related optimization problems.

Of course, when we speak of "changing" or "influence" as above, we use these terms in the meaning of "infinitesimal variations" or the "first order approximation." Thus, they make sense only in a small neighborhood of a specific point $(u, v) \in U \times V$.

Hence, in the absence of a causality restriction, the results of Section II admit an explicit "physical interpretation." The only condition for existence of an affine incentive solution to the Stackelberg dynamic game is that the sensitivity of J_2 with respect to $u(\cdot)$ should not be zero (cf. Proposition 3, and also (38)). That is, whenever the leader is able to influence the follower's cost-functional (infinitesimally), he can always force the follower to choose the prescribed value for his decision variable.

Now, when the leader is faced with the additional constraint that his control at time t cannot depend on the "future values" of $v(\cdot)$, the operator Q in the incentive strategy (39) should be a causal operator (or equivalently, Q^* satisfying (40) should be anti-causal). Moreover, if the leader needs a nonzero time duration ϵ to infer the necessary information on $v(t)$ from the current observation, the control $u(t)$ can only depend on the value of $v(\tau)$ for

$\tau \leq t - \epsilon$; such an operator Q will be called ϵ -strong causal (and Q^* - ϵ -strong anticausal).

Let us first introduce

$$\Phi(t) \triangleq \langle \phi, \phi \rangle_{(t,T)} \text{ or } \langle \phi, \phi \rangle_{(t,N)}, \quad (47)$$

and

$$\psi(t) = \langle \Psi, \Psi \rangle_{(t,T)} \text{ or } \langle \Psi, \Psi \rangle_{(t,N)}. \quad (48)$$

If $\Phi(t) = 0$, i.e., $\phi(\tau) = 0, \tau \geq t$ (recall that, for functions $f(t), g(t)$ in Hilbert space $L_2^m[t_0, t_f]$, $f(\cdot) = g(\cdot)$ means $f(t) = g(t)$ for almost all $t \in [0, T]$, except perhaps on some set of measure zero; this fact has to be noted throughout the paper), then the leader cannot control the situation during $\tau \in [t, T]$. If, concurrently, $\psi(t) \neq 0$, the follower can change the value of $J_2(u, v)$ by infinitesimal variations in $v(\cdot)$. Thus, it is intuitively evident that the leader may not be able to enforce any desired decision pair (u^d, v^d) by a causal incentive strategy, because he cannot respond effectively to the variation in the follower's decision, even though he may be able to detect it.

To put the above intuitive reasoning into precise form, we first prove for continuous-time systems the following result.

Lemma 1: For any $\phi(\cdot) \in U = L_2^{m_1}[t_0, t_f]$ and $\Psi(\cdot) \in V = L_2^{m_2}[t_0, t_f]$, a set of sufficient conditions for existence of an anticausal bounded linear operator $Q^*: U \rightarrow V$ satisfying (40) which can be rewritten as $Q^*\phi = \Psi$, is the following:

- For all $t \in [0, T]$, $\Psi(t) \neq 0$ implies $\Phi(t) \neq 0$. (Let t_ϕ be the smallest time such that $\Phi(t_\phi) = 0$, and t_ψ be the smallest time making $\psi(t) = 0$; then the condition says $t_\phi \geq t_\psi$.)
- When $t_\phi = t_\psi$, the following integral exists and remains bounded

$$\int_0^{t_\phi} \frac{\Psi'(t)\Psi(t)}{\Phi(t)} dt. \quad (49)$$

(This second condition means that the follower's "control ability," measured in terms of $\Psi(t)$, cannot be much stronger than the leader's, at the point $t_\phi = t_\psi$ when they concurrently lose their control ability.)

Proof: The lemma can be proved simply by giving one of the possible solutions for operator Q^* , which is

$$Q^*[f(\tau)](t) \triangleq \begin{cases} \int_{t_\phi}^t \frac{\Psi(\tau)\phi'(\tau)f(\tau)}{\Phi(\tau)} d\tau & (t < t_\phi) \\ 0 & (t \geq t_\phi) \end{cases} \quad (50)$$

Thus, it is an integral operator with kernel

$$R(t, \tau) = \begin{cases} \frac{\Psi(t)\phi'(\tau)}{\Phi(t)} & (\tau, t \in [t_\phi, T]) \\ 0 & (\text{otherwise}) \end{cases} \quad (51)$$

(see, e.g., [29, p. 67]). Note that

$$\begin{aligned} \|R\|^2 &\triangleq \int_0^T \int_0^T \text{Tr} [R(t, \tau) R'(t, \tau)] d\tau dt \\ &= \int_{t_0}^{t_0} \int_{t_0}^{t_0} \text{Tr} \frac{\Psi(t) \Phi'(t) \Phi(\tau) \Psi'(\tau)}{\Phi^2(t)} d\tau dt \\ &= \int_0^{t_0} \frac{\Psi'(t) \Psi(t)}{\Phi^2(t)} \int_{t_0}^{t_0} \Phi'(\tau) \Phi(\tau) d\tau dt \\ &= \int_{t_0}^{t_0} \frac{\Psi'(t) \Psi(t)}{\Phi(t)} dt < \infty \end{aligned} \quad (52)$$

and therefore Q^* is well-defined and bounded, with $\|Q^*\|^2 \leq \|R\|^2$. It is anticausal, since the value of $Q^*[f(\tau)]$ at time t depends only on the values of $f(\tau)$ for $\tau \in [t, t_0]$. Finally it is straightforward to verify that $Q^*[\phi(\cdot)](t) = \Psi(t)$, except perhaps at times t belonging to a set of measure zero.

Remark 6: The operator Q , being the adjoint of the anticausal operator Q^* , is a causal operator. The adjoint of (50) can readily be computed to be

$$\begin{aligned} Q[g(t)](\tau) &= \int_0^T R'(t, \tau) g(t) dt \\ &= \int_0^T \frac{\Phi(\tau) \Psi'(t)}{\Phi(t)} g(t) dt. \end{aligned} \quad (53)$$

The lemma can easily be generalized to the case when $Q^*: U \rightarrow V$ is required to be an ϵ -strong anticausal operator. The sufficient conditions become, in this case the following:

i) whenever $\langle \Psi, \Psi \rangle_{(t, T)} \neq 0$, we must have $\Phi(t + \epsilon) = \langle \Phi, \Phi \rangle_{(t, T)} \neq 0$, that is $t_0 \geq t_\Psi + \epsilon$. This implies that $\Psi(t) = 0$ for all $t \geq T - \epsilon$; and

ii) when $t_0 = t_\Psi + \epsilon$, the following integral exists and remains bounded:

$$\int_0^{t_0} \frac{\Psi'(t) \Psi(t)}{\Phi(t + \epsilon)} dt. \quad (54)$$

In this case the kernel corresponding to (51) is

$$R(t, \tau) = \begin{cases} \frac{\Psi(t) \Phi'(\tau)}{\Phi(t + \epsilon)} & (t < t_0 - \epsilon, t + \epsilon < \tau < t_0) \\ 0 & (\text{otherwise}). \end{cases} \quad (55)$$

Moreover, the counterparts of (50) and (53), in this case, are, respectively,

$$\begin{aligned} Q^*[f(\tau)](t) &= \begin{cases} \int_{t_0}^{t_0} \frac{\Psi(t) \Phi'(\tau)}{\Phi(t + \epsilon)} f(\tau) d\tau & (t < t_0 - \epsilon) \\ 0 & (\text{otherwise}) \end{cases} \\ & \quad (56) \end{aligned}$$

and

$$Q[g(t)](\tau) = \begin{cases} \int_0^T \frac{\Phi(\tau) \Psi'(t)}{\Phi(t + \epsilon)} g(t) dt & (\epsilon \leq \tau < t_0) \\ 0 & (\tau < \epsilon \text{ or } \epsilon \geq t_0). \end{cases} \quad (57)$$

The counterpart of Lemma 1 for the discrete-time case is much simpler, since Condition b is then implied by Condition a.

Lemma 2: For any $\phi(\cdot) \in U = l_2^m[0, N-1]$, $\Psi(\cdot) \in V = l_2^n[0, N-1]$, a sufficient condition for existence of a one-step-strong anticausal linear operator $Q^*: U \rightarrow V$ such that $Q^*\phi = \Psi$ is that

iii) whenever $\psi(k) = \sum_{i=k}^{N-1} \Psi'(i) \Psi(i) \neq 0$, we must have $\Phi(k+1) = \sum_{i=k+1}^{N-1} \Phi'(i) \Phi(i) \neq 0$; that is, $t_0 \geq t_\Psi + 1$.

The proof of this lemma is similar to that of Lemma 1 and is therefore omitted. The corresponding linear operators are

$$Q^*[f(i)](k) = \begin{cases} \sum_{i=k+1}^{N-1} \frac{\Psi(k) \Phi'(i)}{\Phi(k+1)} f(i) & (k < t_0 - 1) \\ 0 & (k \geq t_0 - 1) \end{cases} \quad (58)$$

and

$$Q[g(k)](i) = \begin{cases} \sum_{k=0}^{i-1} \frac{\Phi(i) \Psi'(k)}{\Phi(k+1)} g(k) & (1 \leq i < t_0) \\ 0 & (i = 0 \text{ or } i \geq t_0) \end{cases} \quad (59)$$

which are one-step strong anticausal and one-step strong causal, respectively. The general conclusion we derive from these two lemmas is the following:

Proposition 7: For the general Stackelberg dynamic game problem (Section I-B) with $U = L_2^m[0, T]$, $V = L_2^n[0, T]$ or $U = l_2^m[0, N-1]$, $V = l_2^n[0, N-1]$, in addition to the assumptions 1)-4) made in this section, let conditions a) and b) of Lemma 1 (or correspondingly, condition iii) of Lemma 2) be satisfied, and let the leader have perfect access to the past values of the follower's control variable (by possibly inferring these values perfectly through the observation of the state). Then the operator Q defining the affine incentive strategy

$$u = \gamma_1(r) = u^t - Q[r - r^t] \quad (60)$$

where

$$(u^t, r^t) = \arg \min_{(u, r) \in U \times V} J_1(u, r), \quad (61)$$

can be chosen as a causal operator (correspondingly, one-step strong-causal operator). One of its possible forms is

Here t_0 and t_Ψ are the discrete-time counterparts of those introduced in Lemma 1.

given by (53) (or, correspondingly, by (59)), and it provides a global Stackelberg solution to the problem.

Remark 7: For the discrete-time case, a very simple and useful version of the sufficient conditions is that at the last decision stage

$$\phi(N) \neq 0 \quad \text{and} \quad \Psi(N) = 0.$$

The causal incentive solution obtained above offers us a possible way of constructing the "closed-loop" solution to the Stackelberg dynamic game problem, provided that $v(\cdot)$ can be reconstructed from the observed state information in real-time. If, for example, there is a causal operator H such that

$$v(\cdot) = Hx(\cdot) \quad (62)$$

then we have the closed-loop solution

$$u(\cdot) = u^+(\cdot) - Q[Hx(\cdot) - v^+(\cdot)] \quad (63)$$

which is physically realizable. More specific derivations along this line are provided in the next section.

We should note that when the ϵ -strong causality conditions i) and ii) are taken instead of a) and b), the statement of Proposition 6 can be modified in a straightforward manner, which then says that affine ϵ -strong causal solutions exist. These may be used in realizing the optimum Stackelberg strategy of the leader, with an ϵ -delay in the reconstruction of $v(\cdot)$ from the state observation $x(\cdot)$.

Finally, we should remark that the results of this section, in particular those of Lemmas 1, 2, and Proposition 7, can be extended to the case when the leader has only partial state information and/or partial dynamic information on the follower's actions, without much difficulty and with only minor modifications. This extension involves, basically, the derivation of an achievable desirable solution (u^*, v^*) to replace (61) (cf. Section II-B), "projected" cost functional $J_2(u, v)$ for the follower, and rewording of Lemmas 1, 2 and Proposition 7 in terms of this new notation. We do not pursue this point here; see, however, the specific problem solved in Section V-B.

V. APPLICATIONS AND EXAMPLES

In this section, the concepts and results presented in Sections II and IV will be applied to some special cases of practical interest. Some numerical examples will be given to show the applicability of the theory and the general approach.

A. Causal Stackelberg Solution to Discrete-Time Linear Quadratic Dynamic Game Problems

One of the important subclasses of problems widely discussed in the literature (see e.g., [9], [27]) is the discrete-time dynamic Stackelberg game with linear-state equation

$$x(k+1) = A(k)x(k) + B(k)u(k) + C(k)v(k) \quad (k = 0, 1, \dots, N-1) \quad (64)$$

and quadratic cost-functionals

$$J_i = x'(N)Q(N)x(N) + \sum_{j=0}^{N-1} \{x'(j)Q_i(j)x(j) + u'(j)R_i(j)u(j) + v'(j)S_i(j)v(j)\} \quad (65)$$

where $i = 1, 2$, refer to the leader and the follower, respectively.

The approach presented in the previous section can be used in obtaining a causal solution to this problem under the closed-loop information pattern. Here we give only a numerical example to illustrate the method.

Example 1:

$$x(k+1) = x(k) + u(k) + v(k) \quad (k = 0, 1, \dots, N-2)$$

$$x(N) = x(N-1) + u(N-1)$$

$$J_1 = x^2(N) + \sum_{k=0}^{N-1} (x^2(k) + 2u^2(k) + v^2(k))$$

$$J_2 = x^2(N) + \sum_{k=0}^{N-1} (x^2(k) + u^2(k) + 3v^2(k)).$$

In accordance with the method presented in Sections II and IV, first the team solution of minimizing J_2 is obtained from the standard Riccati recurrence relations, which involves the value function $x'(k)P(k)x(k)$, the corresponding team optimal controls $u^*(k)$ and $v^*(k)$, and optimal trajectory $x^*(\cdot)$. Then the gradients $\phi(k) = \nabla_{u(k)} J_1^*$ and $\Psi(k) = \nabla_{v(k)} J_1^*$ at the desired team solution are derived from the dynamic equations as

$$\begin{aligned} \phi(k) &= \sum_{j=k+1}^N x^*(j) + u^*(k) \quad (k = 0, 1, \dots, N-1) \\ \Psi(k) &= \sum_{j=k+1}^N x^*(j) + 3v^*(k) \quad (k = 0, 1, \dots, N-2), \\ \Psi(N-1) &= 0. \end{aligned}$$

Note that $\Phi(k) = \sum_{i=k+1}^N \phi^2(i)$, and from (60) and (59) the optimal Stackelberg strategy for the leader is

$$\begin{aligned} z(t; x) &= u^*(t) + \phi(t) \sum_{k=t}^{N-1} \frac{\Psi(k) [v(k) - v^*(k)]}{\Phi(k+1)} \\ &= P(t)x(t) + \phi(t) \sum_{k=t}^{N-1} \frac{\Psi(k)}{\Phi(k+1)} \\ &\quad \cdot [x(k+1) - x(k) - u^*(k) - v^*(k)], \\ &u(0) = u^*(0). \end{aligned}$$

The values of these coefficients for the case $N = 4$ are

TABLE I

k	0	1	2	3	4
$P(k)$	1.4576075	1.4592593	1.4761905	1.6666667	1.0000000
$\hat{u}^+(k), \hat{x}^+(k)$	-0.2288037	-0.2296296	-0.2380953	-0.3333333	—
$\hat{v}^+(k), \hat{x}^+(k)$	-0.4576074	-0.4592593	-0.4761905	0	—
$\hat{x}^+(k), \hat{x}^+(k-1)$	—	0.3135888	0.3111111	0.2857143	0.6666667
$\hat{x}^+(k), \hat{x}(0)$	1	0.3135888	0.097561	0.0278746	0.018543
$\hat{u}(k), \hat{x}(0)$	0.2288037	0.0720093	0.0132298	0.0092915	—
$\hat{v}(k), \hat{x}(0)$	-0.9152148	-0.2880372	-0.0929152	0	—
$\hat{u}(k), \hat{x}^2(0)$	0.0581624	0.0058112	0.006259	0.0000963	—

listed in Table I, with the corresponding policies being

$$u(1) = u^+(1) - 11.340855(v(0) - v^+(0))$$

$$u(2) = u^+(2) - 3.6583393$$

$$-(v(0) - v^+(0) - 1.068982(v(1) - v^+(1)))$$

$$u(3) = u^+(3) - 1.4633326$$

$$-(v(0) - v^+(0) - 0.4275919(v(1) - v^+(1)))$$

$$-10.003726(v(2) - v^+(2)).$$

Here

$$v(0) - v^+(0) = x(1) - x(0) - u^+(0) - v^+(0)$$

$$v(1) - v^+(1) = x(2) - x(1) - u^+(1) - v^+(1)$$

$$v(2) - v^+(2) = x(3) - x(2) - u^+(2) - v^+(2).$$

This is a causal closed-loop Stackelberg solution which achieves the globally optimal team solution.

B. The Linear Quadratic Infinite-Time Stackelberg Problem

Consider the continuous-time problem formulated by

$$\dot{x} = Ax + Bu + Cv \quad x(t) \in \mathbb{R}^n, \quad u(t) \in \mathbb{R}^m, \\ v(t) \in \mathbb{R}^m, \quad x(0) = x_0, \quad t \geq 0; \quad (66)$$

$$J_i = \int_0^\infty (x'Q_i x + u'R_i u + v'S_i v) dt \quad (i = 1, 2) \quad (67)$$

where B and C have full-column rank m_1 and m_2 , respectively, $(A, (B; C))$ is controllable, $(Q_i^{\frac{1}{2}}, A)$ is observable, $Q_i \geq 0$, $R_i > 0$, $S_i > 0$, $R_2 \geq 0$. The team solution that minimizes J_1 is

$$u^* = -R_1^{-1}B'Px^*, \quad v^* = -S_1^{-1}C'Px^*, \quad J_1^* = x_0'Px_0 \quad (68)$$

where P is the unique positive definite solution of the

algebraic Riccati equation

$$P[BR_1^{-1}B' + CS_1^{-1}C']P - A'P - PA - Q_1 = 0 \quad (69)$$

and the optimal trajectory x^* satisfies

$$\dot{x}^* = A_c x^* \triangleq [A - (BR_1^{-1}B' + CS_1^{-1}C')P]x^*, \\ x^*(0) = x_0. \quad (70)$$

We now attempt to solve this problem under two different causal-functional dependences for the leader's policy, viz $\gamma_1: V \rightarrow U$; $\gamma_1(v) = u^* - Q(v - v^*)$ and $\gamma_1: X \rightarrow U$; $\gamma_1(x) = u^* - Q(x - x^*)$, where Q is, in each case, a linear causal operator.

In the former case, we first calculate the gradients of J_2 with respect to u and v (see Appendix A) and arrive at

$$\phi(t) = \nabla_u J_2(u^*, v^*) = 2Mx^*(t) \quad (71)$$

$$\Psi(t) = \nabla_v J_2(u^*, v^*) = 2Nx^*(t) \quad (72)$$

where

$$M = (B'I_0 - R_2R_1^{-1}B'P) \quad (73)$$

$$N = (C'I_0 - S_2S_1^{-1}C'P) \quad (74)$$

and I_0 is the solution of the matrix equation

$$A'I_0 + I_0A + Q_2 = 0. \quad (75)$$

If a constant matrix gain solution is desired, then we must have

$$Q'M = N. \quad (76)$$

Unless $\text{Range } M' \supset \text{Range } N'$, such a Q does not exist, and hence the problem does not admit a solution. However, if we also allow dependence on the initial state x_0 , affine causal solutions exist provided that for all $x_0 \neq 0$, $Me^{-A}x_0 \neq 0$, which is equivalent to the requirement that (M, A) be observable. In this case an optimal affine

incentive scheme is

$$u(t) = u^* - Q(v - v^*)$$

$$= u^*(t) - \int_0^t \frac{\phi(\sigma)\Psi'(\sigma)[v(\sigma) - v^*(\sigma)]}{\int_0^\infty \phi'(s)\phi(s) ds} d\sigma. \quad (77)$$

Next, we seek a solution in the form $u = u^* - Q(x - x^*)$, where Q is causal. By using the approach outlined in Section II-B and taking the entire trajectory x as the leader's observed information, we have, uniquely,

$$v^*(u, x) = C^+(\dot{x} - Ax - Bu) \quad (78)$$

where $C^+ \triangleq (C'C)^{-1}C'$ is the pseudo-inverse of C . Note that in this case the absolute lower bound given by (68) is attainable, since the operator N of Section II-B is invertible (C being a matrix of full-column rank). Now, projecting the problem into $U \times X$ where (u, x) belongs, we obtain

$$\bar{J}_2(u, x) = \langle x, Q_2 x \rangle + \langle u, R_2 x \rangle$$

$$+ \langle (\dot{x} - Ax - Bu), \bar{C}(\dot{x} - Ax - Bu) \rangle \quad (79)$$

where $\bar{C} \triangleq C^+ S_2 C^+$. The gradients $\nabla_u \bar{J}_2$ and $\nabla_x \bar{J}_2$ at (u^*, x^*) can be evaluated as (see Appendix B)

$$\phi(t) = \nabla_u \bar{J}_2(u^*, x^*) = 2Mx^*(t) \quad (80)$$

$$\Psi(t) = \nabla_x \bar{J}_2(u^*, x^*) = 2Nx^*(t) \quad (81)$$

where

$$M = (B'C^+ S_2 S_1^{-1} C' - R_2 R_1^{-1} B')P \quad (82)$$

$$N = Q_2 + C^+ S_2^{-1} C' P A_c + A' C^+ S_2 S_1^{-1} C' P. \quad (83)$$

The conclusion we arrive at here is almost the same as in the case (73)–(74). When $\text{Range } M' \supset \text{Range } N'$, there exists a constant gain solution $u = u^* - Q(x - x^*)$ with Q satisfying $Q'M = N$. Otherwise, provided that (M, A_c) is observable, there exists an affine causal solution, depending on $x_0 \neq 0$, given by

$$u(t) = u^* - Q(x - x^*)$$

$$= u^*(t) - \int_0^t \frac{\phi(\sigma)\Psi'(\sigma)(x(\sigma) - x^*(\sigma))}{\int_0^\infty \phi'(s)\phi(s) ds} d\sigma, \quad (84)$$

where $\phi(\cdot)$ and $\Psi(\cdot)$ are given by (80) and (81), respectively.

We now provide a numerical example to illustrate these results.

Example 2:

$$\dot{x} = 2x + u + v, \quad x(0) = x_0 \quad t \in [0, \infty)$$

$$J_1 = \int_0^\infty (6x^2 + u^2 + v^2) dt$$

$$J_2 = \int_0^\infty (qx^2 + rv^2) dt, \quad q > 0, r > 0.$$

The team solution is

$$u^* = -3x^*, \quad v^* = -3x^*, \quad x^*(t) = e^{-4t}x_0,$$

since from $2P^2 - 4P - 6 = 0$ we have $P = 3$ and $A = -4$. From (75), $I_0 = (1/2)q$. From (73)–(74), $M = (1/2)q$, $N = (1/2)q - 3r$. That is, $\phi(t) = qx^*(t)$, $\Psi(t) = (q - 6r)x^*(t)$. The optimal Stackelberg strategy is

$$u(t) = u^*(t) - Q[v - v^*](t)$$

where the operator Q is either $Q = N/M = 1 - 6r/q$ or

$$Q[g(t)](\tau) = \phi(\tau) \int_0^\tau \frac{\Psi(t)g(t)}{\int_t^\infty \phi^2(s) ds} dt$$

$$= 8\left(1 - \frac{6r}{q}\right) \int_0^\tau e^{-4(\tau-t)} g(t) dt.$$

This solution can be implemented by a first-order block with transfer function

$$\bar{W}(s) = \frac{U(s) - U^*(s)}{V(s) - V^*(s)} = 8\left(1 - \frac{6r}{q}\right) \frac{1}{s+4},$$

where s is the Laplace variable.

On the other hand, from (82) and (83)

$$M = 3r/2, \quad N = (q - 6r)/2.$$

Thus, the optimum Stackelberg strategy in case of x -dependence is

$$u(t) = u^*(t) - Q[x - x^*](t),$$

where the operator Q is either $Q = N/M = q/3r - 2$ or

$$Q[g(t)](\tau) = 3re^{-4\tau}x_0 \int_0^\tau \frac{(q - 6r)e^{-4t}x_0 g(t)}{\int_t^\infty 9r^2 x_0^2 e^{-8s} ds} dt$$

$$= \int_0^\tau \left(\frac{q}{3r} - 2\right) e^{-4(\tau-t)} g(t) dt$$

which may be implemented in the frequency domain by

$$W(s) = \frac{U(s) - U^*(s)}{X(s) - X^*(s)} = 8\left(\frac{q}{3r} - 2\right) \frac{1}{s+4}.$$

C. The Linear Quadratic Finite-Time Closed-Loop Stackelberg Problem

In this subsection we provide an example illustrating the results obtained in Section IV when $t_0 < \infty$.

Example 3:

Consider the problem with the specifics

$$\dot{x} = 2x + u + b(t)v, \quad t \in [0, 1],$$

$$J_1 = 4x^2(1) + \int_0^1 (6x^2 + u^2 + v^2) dt,$$

$$J_2 = 2x^2(1) + \int_0^1 (qx^2 + rv^2) dt, \quad (q > 0, r > 0)$$

where the time-varying gain $b(t)$ is a continuous bounded function; furthermore, when $t \rightarrow 1$, $b(t) \rightarrow 0$ with the order of magnitude being

$$b(t) = b_0 e^{-\alpha/(1-t)}, \quad \alpha > 1, \quad t \rightarrow 1, \quad b_0 > 0.$$

AD-A145 212

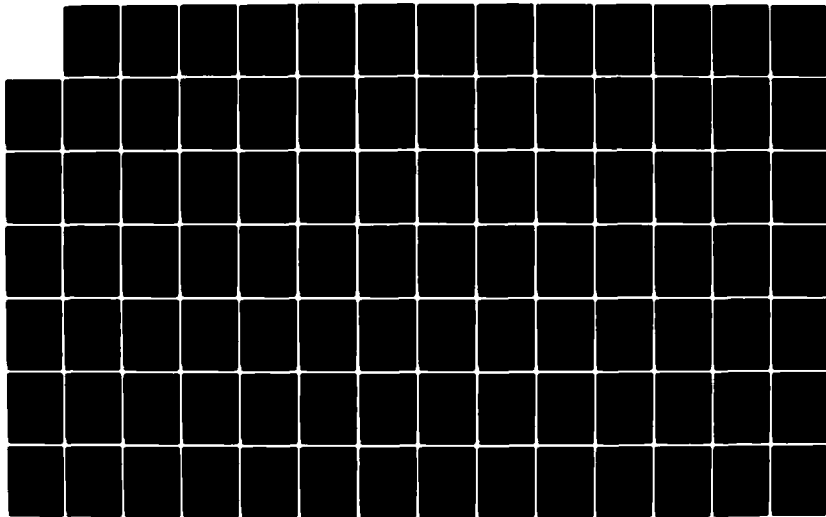
ROBUST TEAM OPTIMAL AND LEADER-FOLLOWER POLICIES FOR
DECISION MAKING IN C... (U) ILLINOIS UNIV AT URBANA
T BASAR ET AL. JUL 84 N00014-82-K-0469

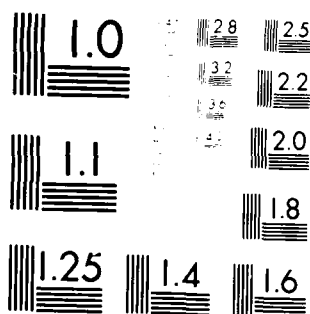
3/3

UNCLASSIFIED

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART
 NATIONAL BUREAU OF STANDARDS-1963-A

The team solution for J_1 is

$$u^*(t) = -2P(t)x^*(t), \quad v^* = -2b(t)P(t)x^*(t)$$

$$\dot{x}^* = (2 - 2(1+b)^2)Px^* \Rightarrow x^*(\tau)$$

$$= \left[\exp \left(- \int_0^\tau 2P(b^2 + 2b) dt \right) \right] x_0$$

$$\frac{dP}{dt} + 4P - 2P^2(1+b^2) + 3 = 0, \quad P(1) = 2.$$

Both $x^*(t)$ and $P(t)$ are bounded continuous functions on $[0, 1]$.

$$\nabla_u J_2^* = \phi(\tau) = 2q \int_\tau^1 e^{2(1-\tau)t} x^*(t) dt + 4x^*(1)e^{2(1-\tau)}$$

$$\begin{aligned} \nabla_v J_2^* = \Psi(\tau) &= 2qb(\tau) \int_\tau^1 e^{2(1-\tau)t} x^*(t) dt \\ &+ 4x^*(1)e^{2(1-\tau)b(\tau)} + 2v^*(\tau). \end{aligned}$$

When $\tau \rightarrow 1$, $\phi(\tau) \rightarrow \phi(1) = 4x^*(1)$, thus $t_0 = 1$. Furthermore,

$$\Phi(t) = \int_t^1 \phi^2(\tau) d\tau = 4x^*(1)(1-t) + 0(1-t).$$

Therefore, condition (49) is satisfied:

$$\int_t^1 \frac{\Psi^2(\tau)}{\Phi(\tau)} d\tau < \infty$$

and by Lemma 1, the operator

$$Q[g(\cdot)](\tau) = \int_0^\tau \frac{\phi(\tau)\Psi(t)}{\Phi(t)} g(t) dt$$

is linear, bounded, and can be used in the construction of the Stackelberg strategy $u = u^* - Q(x - x^*)$.

VI. CONCLUDING REMARKS

In this paper we have discussed derivation of closed-loop Stackelberg strategies and incentive policies for a general class of dynamic decision problems with a hierarchical decision structure, in both discrete and continuous time. The first set of results involve discrete-time dynamic games in which the leader has informational advantage over the follower, in the sense that he can observe the follower's actions at each stage (before he acts) either perfectly or partially. Under a feedback Stackelberg solution concept that takes this informational advantage into account, we have studied derivation of optimal affine policies. Furthermore, we have investigated the conditions under which such a solution coincides with the global Stackelberg solution (cf. Section III).

A second set of results presented in this paper has involved an analysis of existence and derivation of causal real-time implementable global Stackelberg solutions in dynamic games wherein the leader is allowed to use memory policies. In this context, we have treated both discrete-time and continuous-time problems, and using a function space approach we have solved certain special cases both analytically and numerically (Sections II, IV, and V).

APPENDIX A

Derivation of (71)-(72):

Since $x(t) = e^{A_1 t} x_0 + \int_0^t e^{A_1(t-\tau)} (Bu(\tau) + Cv(\tau)) d\tau$,

$$\langle \delta x, Q_2 x \rangle$$

$$= \int_0^\infty \delta x'(t) Q_2 x(t) dt$$

$$= \int_0^\infty \int_0^t (\delta v'(\tau) C' + \delta u'(\tau) B') e^{A_1(t-\tau)} Q_2 x(t) d\tau dt$$

$$= \int_0^\infty d\tau \int_\tau^\infty [\delta v'(\tau) C' + \delta u'(\tau) B'] e^{A_1(t-\tau)} Q_2 x(t) dt$$

Therefore, for variations δu and δv we have

$$\frac{1}{2} \delta J_2 = \langle \delta x, Q_2 x \rangle + \langle \delta u, R_2 u \rangle + \langle \delta v, S_2 v \rangle$$

$$= \langle \delta u, R_2 u \rangle + \int_\tau^\infty B' e^{A_1(t-\tau)} Q_2 x(t) dt$$

$$+ \langle \delta v, S_2 v \rangle + \int_\tau^\infty C' e^{A_1(t-\tau)} Q_2 x(t) dt$$

$$\frac{1}{2} \nabla_u J_2 = R_2 u + \int_\tau^\infty B' e^{A_1(t-\tau)} Q_2 x(t) dt$$

$$\frac{1}{2} \nabla_v J_2 = S_2 v + \int_\tau^\infty C' e^{A_1(t-\tau)} Q_2 x(t) dt.$$

When $u = u^*$, $v = v^*$,

$$\frac{1}{2} \phi(\tau) = \frac{1}{2} \nabla_u J_2(u^*, v^*)$$

$$= R_2 u^* + B' e^{-A_1 \tau} \left[\int_\tau^\infty e^{A_1 t} Q_2 e^{-A_1 t} dt \right] x_0$$

$$\frac{1}{2} \Psi(\tau) = \frac{1}{2} \nabla_v J_2(u^*, v^*)$$

$$= S_2 v^* + C' e^{-A_1 \tau} \left[\int_\tau^\infty e^{A_1 t} Q_2 e^{-A_1 t} dt \right] x_0.$$

Since

$$I_\tau = \int_\tau^\infty e^{A_1 t} Q_2 e^{-A_1 t} dt$$

$$= e^{A_1 t} A_1^{-1} \Big|_\tau^\infty - \int_\tau^\infty A_1' e^{A_1 t} Q_2 e^{-A_1 t} A_1^{-1} dt$$

$$= A_1^{-1} e^{A_1 t} Q_2 e^{-A_1 t} \Big|_\tau^\infty - \int_\tau^\infty A_1' e^{A_1 t} Q_2 e^{-A_1 t} A_1^{-1} dt,$$

it follows that

$$A_1' I_\tau + I_\tau A_1 + e^{A_1 \tau} Q_2 e^{-A_1 \tau} = 0.$$

Let I_0 satisfy

$$A_1' I_0 + I_0 A_1 + Q_2 = 0;$$

then

$$I_\tau = e^{A_1 \tau} I_0 e^{-A_1 \tau}.$$

Therefore,

$$\frac{1}{2} \phi(t) = -R_2^{-1} B' P x^*(t) + B' I_0 x^*(t)$$

$$\frac{1}{2} \Psi(t) = -S_2^{-1} C' P x^*(t) + C' I_0 x^*(t)$$

and relations (73) and (74) follow.

APPENDIX B

Derivation of Gradients (80)–(81): Note that

$$\begin{aligned} \bar{J}_2(u, x) &= \langle x, Q_2 x \rangle + \langle u, R_2 u \rangle \\ &\quad + \langle (\dot{x} - Ax - Bu), \bar{C}(\dot{x} - Ax - Bu) \rangle, \end{aligned}$$

and consider only those x and u with their values and variations δx and δu satisfying

$$\begin{aligned} x(\infty) &= \dot{x}(\infty) = 0 \\ \delta x(0) &= \delta x(\infty) = \delta \dot{x}(\infty) = 0 \\ u(\infty) &= \delta u(\infty) = 0. \end{aligned}$$

We have, for variations δx and δu :

$$\begin{aligned} \delta \langle \dot{x}, \bar{C} \dot{x} \rangle &= 2 \langle \delta \dot{x}, \bar{C} \dot{x} \rangle = 2 \int_0^\infty \delta \dot{x}(t) \bar{C} \dot{x}(t) dt \\ &= 2 \delta x' \bar{C} \dot{x}|_0^\infty - 2 \int_0^\infty \delta \dot{x}' \bar{C} \dot{x} dt \\ &= -2 \int_0^\infty \delta \dot{x}' \bar{C} \dot{x} dt \end{aligned}$$

$$\nabla_x \langle \dot{x}, \bar{C} \dot{x} \rangle = -2 \bar{C} \dot{x}$$

$$\begin{aligned} \delta \langle \dot{x}, \bar{C} Bu \rangle &= \int_0^\infty \delta \dot{x}' \bar{C} Bu dt + \int_0^\infty \dot{x}' \bar{C} B \delta u dt \\ &= \int_0^\infty \dot{x}' \bar{C} B \delta u dt + \delta x' \bar{C} Bu|_0^\infty - \int_0^\infty \delta \dot{x}' \bar{C} Bu dt \end{aligned}$$

$$\nabla_x \langle \dot{x}, \bar{C} Bu \rangle = -\bar{C} Bu$$

$$\nabla_u \langle \dot{x}, \bar{C} Bu \rangle = B' \bar{C} \dot{x}.$$

Therefore,

$$\begin{aligned} \frac{1}{2} \nabla_u \bar{J}_2 &= R_2 u - B' \bar{C} \dot{x} + B' \bar{C} Ax + B' \bar{C} Bu \\ &= R_2 u - B' \bar{C} C v \\ \frac{1}{2} \nabla_x \bar{J}_2 &= Q_2 x - \bar{C} \dot{x} + A' \bar{C} Ax + \bar{C} A \dot{x} \\ &\quad - A' \bar{C} \dot{x} + \bar{C} Bu + A' \bar{C} Bu \\ &= Q_2 x - \bar{C} C v - A' \bar{C} C v \\ &= Q_2 x - C' S_2 \dot{x} - A' C' S_2 v. \end{aligned}$$

At point (u^*, x^*) , these expressions become equal to

$$\begin{aligned} \frac{1}{2} \phi(t) &= \frac{1}{2} \nabla_u \bar{J}_2(u^*, x^*) \\ &= (-R_2 R_1^{-1} B' P x^* + B' \bar{C} C S_1^{-1} C' P x^*)(t) \\ &= (B' \bar{C} C S_1^{-1} C' - R_2 R_1^{-1} B') P x^*(t) \\ &\triangleq M(t) x^*(t) \\ \frac{1}{2} \Psi(t) &= \frac{1}{2} \nabla_x \bar{J}_2(u^*, x^*) \\ &= (Q_2 x^* + C' S_2 S_1^{-1} C' P x^* \\ &\quad + A' C' S_2 S_1^{-1} C' P x^*)(t) \\ &= (Q_2 x^* + C' S_2 S_1^{-1} C' P A, x^* \\ &\quad + A' C' S_2 S_1^{-1} C' P x^*)(t) \end{aligned}$$

$$\begin{aligned} &= (Q_2 + C' S_2 S_1^{-1} C' P A, \\ &\quad + A' C' S_2 S_1^{-1} C' P) x^*(t) = N(t) x^*(t). \end{aligned}$$

This then completes the verification.

REFERENCES

- [1] M. Athans, "On large-scale systems and decentralized control," *IEEE Trans. Automat. Contr.*, vol. AC-23, 1978.
- [2] K. J. Arrow and R. Radner, "Allocation of resources in large teams," *Econometrica*, vol. 47, 1979.
- [3] T. Başar, "Equilibrium strategies in dynamic games with multilevels of hierarchy," *Automatica*, vol. 17, pp. 749–754, 1981.
- [4] ———, "A general theory for Stackelberg games with partial state information," *Large Scale Systems*, vol. 3, pp. 47–56, 1982.
- [5] T. Başar and J. B. Cruz, Jr., "Concepts and methods in multi-person Coordination and Control," in *Optimization and Control of Dynamic Operational Research Models*, S. G. Tzafestas, Ed. Amsterdam: North Holland, 1982, Ch. 11, pp. 351–387.
- [6] T. Başar and G. J. Olsder, "Team-optimal closed-loop Stackelberg strategies in hierarchical control problems," *Automatica*, vol. 16, pp. 409–414, 1980.
- [7] ———, *Dynamic Noncooperative Game Theory*. London: Academic, 1982.
- [8] T. Başar and H. Selbuz, "A new approach for derivation of closed-loop Stackelberg strategies," *Proc. 19th IEEE Conf. Decision and Control*, San Diego, CA, 1979, pp. 1113–1118.
- [9] ———, "Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems," *IEEE Trans. Automat. Contr.*, vol. AC-24, pp. 166–179, 1979.
- [10] D. Cansever and T. Başar, "A minimum sensitivity approach to incentive design problems," in *Proc. 21st IEEE Conf. Decision and Control*, Orlando, FL, 1982, pp. 158–163.
- [11] C. I. Chen and J. B. Cruz, Jr., "Stackelberg solutions for two-person games with biased information patterns," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 791–798, 1972.
- [12] J. B. Cruz, Jr., "Survey of Nash and Stackelberg equilibrium strategies in dynamic games," *Annals of Economic and Social Management*, vol. 4, pp. 339–344, 1975.
- [13] ———, "Leader-follower strategies for multilevel systems," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 244–255, 1978.
- [14] B. F. Gardner, Jr. and J. B. Cruz, Jr., "Feedback Stackelberg strategies for a two-player game," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 270–271, 1977.
- [15] T. Groves, "Incentives in teams," *Econometrica*, vol. 41, pp. 617–631, 1973.
- [16] T. Groves and M. Loeb, "Incentives in a divisionalized firm," *Management Sci.*, vol. 25, pp. 221–230, 1979.
- [17] Y. C. Ho, P. B. Luh, and R. Muralidharan, "Information structure, Stackelberg games, and incentive controllability," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 454–460, 1981.
- [18] Y. C. Ho, P. B. Luh, and G. J. Olsder, "A control-theoretic view on incentives," *Automatica*, vol. 18, pp. 167–179, 1982.
- [19] L. P. Jennergren, "On the design of incentives in business firms: A survey of some research," *Management Sci.*, vol. 26, pp. 180–201, 1980.
- [20] P. B. Luh, Y. C. Ho, and R. Muralidharan, "Load adaptive pricing: an emerging tool for electric utilities," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 320–328, 1982.
- [21] G. P. Papavasiliopoulos and J. B. Cruz, Jr., "Nonclassical control problems and Stackelberg games," *IEEE Trans. Automat. Contr.*, vol. AC-24, pp. 155–166, 1979.
- [22] R. Radner, "Monitoring cooperative agreements in a repeated principal-agent relationship," *Econometrica*, vol. 49, pp. 1127–1148, 1981.
- [23] M. Salman and J. B. Cruz, Jr., "An incentive model of duopoly with government coordination," *Automatica*, vol. 17, pp. 821–829, 1981.
- [24] M. Simaan and J. B. Cruz, Jr., "On the Stackelberg strategy in nonzero-sum games," *J. Optimiz. Theory Appl.*, vol. 11, pp. 833–855, 1973.
- [25] ———, "Additional aspects of the Stackelberg strategy in nonzero-sum games," *J. Optimiz. Theory Appl.*, vol. 11, pp. 613–626, 1973.

- [26] H. Von Stackelberg, *Marktform und Gleichgewicht*. Vienna: Springer-Verlag. Eng. transl., *The Theory of the Market Economy*. Oxford: Oxford University Press, 1952.
- [27] B. Tolwinski, "Closed-loop Stackelberg solution to multi-stage linear-quadratic game," *J. Optimization Theory and Applications*, vol. 34, pp. 485-501, 1981.
- [28] Y. P. Zheng and T. Başar, "Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: A geometric approach," *Int. J. Control*, vol. 35, no. 6, pp. 997-1011, 1982.
- [29] A. V. Balakrishnan, *Applied Functional Analysis*. New York: Springer-Verlag, 1976.

FEEDBACK EQUILIBRIA IN DIFFERENTIAL GAMES
WITH STRUCTURAL AND MODAL UNCERTAINTIES

Tamer Başar
Department of Electrical Engineering
and Coordinated Science Laboratory
University of Illinois
Urbana, Illinois 61801 USA

Alain Haurie
Ecole des Hautes Etudes Commerciales
Montreal, Quebec, H3T1V6 Canada

Contents

0. Abstract and Acknowledgements.....	1
1. Introduction.....	1
2. Strong and Weak Equilibrium Solutions, and the Relationship Between Stackelberg and Nash Equilibria.....	5
3. Stochastic Dynamic Games with Structural and Modal Uncertainties Described by Jump Processes.....	14
3.1 Problem Formulation.....	14
3.2 The Concept of Strong Equilibrium.....	17
3.3 Derivation of Strong Equilibria.....	18
4. Feedback Equilibria for Differential Games with Jump Disturbances.....	22
4.1 The Controlled Stochastic System.....	22
4.2 $G(\delta)$ -Differential Games.....	27
4.3 Hamilton-Jacobi Equations.....	34
5. Some Special Cases.....	40
5.1 Fixed Mode of Play.....	41
5.2 Differential Games with the Mode of Play Determined by a Markov Jump Process.....	49
6. Concluding Remarks.....	53
7. Appendix I.....	56
8. Appendix II.....	60
References.....	63
Notation.....	65

Abstract

In this study we introduce a general definition of an equilibrium concept (called "strong equilibrium") for both discrete and continuous time dynamic games and under varying (symmetrical and asymmetrical) modes of play. The underlying system is stochastic, with structural and modal uncertainties determined by a finite state jump process. The new equilibrium concept encompasses both the feedback Nash and feedback Stackelberg solution concepts for the special cases of deterministic discrete-time games with symmetrical and asymmetrical modes of play, respectively, and it also provides a convenient framework for the introduction of a feedback Stackelberg solution concept in deterministic differential games. For the general class of stochastic nonzero-sum games with structural and modal uncertainties, and under the feedback closed-loop information, we obtain the optimality conditions in both discrete and continuous time. Certain special cases are also studied, and the intrinsic relationship between information patterns and possible definitions of value in nonzero-sum differential games is clarified.

Acknowledgements: Research of T. Başar was supported in part by the Office of Naval Research under Contract N00014-82-K-0469 and in part by the Joint Services Electronics Program under Contract N00014-79-C-0424. Research of A. Haurie was supported by FCAC-Quebec and SSHRC under Grant 410-81-0722. The first author would also like to acknowledge the support of "Direction de la Recherche-HEC" for his visit to "Ecole des Hautes Etudes Commerciales" in May 1980 during which time the research which led to the results reported herein was initiated. Both authors are indebted to R. Rishel for enlightening discussions on the optimum control of systems with jump disturbances and regarding possible extensions to differential games.

1. Introduction

Stackelberg solution concepts arise in games with asymmetrical modes of play, with one of the players, called the leader, having the ability and power to announce his move (or policy) first, leaving to the other players the possibility to react.

This solution concept, first introduced by Von Stackelberg [1] in the realm of the economic theory of imperfect competition, attracted the attention of control theorists concerned with hierarchical systems, with the first set of related results documented in the works of Chen and Cruz [11], Simaan and Cruz [2a, 2b], and Castanon [12]. It was soon discovered that the derivation of the closed-loop Stackelberg solution (corresponding to the case when all players acquire closed-loop state information) involved an extremely challenging (nonclassical) class of optimization problems, and that it would generally not lend itself to a dynamic programming approach--the absence of "tenet of transition" precluding the possibility of applying the standard backward induction procedure. For an up-to-date account of these aspects of the Stackelberg problem, and in view of the recent developments on the solvability of the closed-loop Stackelberg game by also allowing for memory policies, we refer the reader to [3], and also to [13].

As an alternative to the closed-loop Stackelberg solution, Simaan and Cruz introduced in [2] the concept of feedback Stackelberg solution where the leadership is defined and implemented sequentially, in the spirit of the dynamic programming approach. In other words, the feedback Stackelberg solution is defined recursively, by solving a

static Stackelberg game at every stage of the decision process and by proceeding in retrograde time. It was defined only for discrete-time systems, and the concept was hastily discredited when it appeared that the leader, in such a game, was not assured of a better payoff than what he could obtain if he were playing according to the Nash equilibrium solution concept.

Our objective, here, is to provide a new interpretation of the feedback Stackelberg solution, which permits us to relate the Nash equilibrium, the closed-loop Stackelberg solution and the feedback Stackelberg solution as various manifestations of a central concept in dynamic games, which we call "strong (feedback) equilibrium". Using this new "unified" concept, and adopting an approach similar to that used by Friedman [10] for the definition of differential games (i.e. continuous-time dynamic games), we will then be able to extend the concept of a feedback Stackelberg solution to the case of systems described by ordinary differential equations. Furthermore, the leadership is often a changing "gift", and hence it may rotate between different players according to some pre-chosen deterministic or random rule. Using our new extended solution concept, we will be able to model situations where the mode of play is randomly evolving in time, possibly dependent on the past moves (actions) of the players as observed through the current value of the state.

In section 2 the so-called "single-act Stackelberg game" is described, using the fundamental concept of an extensive form of a two-person game. It is then shown that the Stackelberg solution concept is in fact an equilibrium solution (so-called strong equilibrium) associated with a peculiar information structure. This "single-act game" incorporates all the essential ingredients of the most general structure for

a two-person game: First a chance player acts and decides on the system to be controlled and the mode of play (i.e. whether or not there will be asymmetry in the information pattern, and in case of an asymmetry who will be the leader); then, the game is played according to the rules set by the chance move.

In section 3 this basic structure, as well as the concept of strong equilibrium, is extended to a multi-stage decision framework, where the players have access to the current value of the state of the dynamic system and the outcome of the finite-state jump process which characterizes the underlying dynamic system and the current mode of play. For this class of stochastic dynamic games, we obtain a set of recursive equations which completely characterizes the strong equilibrium solution. Two special cases of this, corresponding to the fixed asymmetric and symmetric modes of play, are the feedback Stackelberg and the feedback Nash solutions, respectively.

In Section 4 the set-up of Section 3 is extended to the continuous time. Here, we consider the class of two-person differential games in which the mode of play and the structure of the underlying system are determined, at each point in time, as the outcome of a finite state jump process evolving in continuous time and depending on the current value of the state. Both players observe this outcome and the current value of the state, and they play the specific game chosen by the chance mechanism, using feedback control laws. Of course, in doing this, the players have to anticipate the future moves and possible realizations of the jump process that determine

the future rules of the game. For such a stochastic differential game, it is not possible to introduce the concept of an equilibrium directly; however, by introducing a sequence of $G(\delta)$ -games which are discretized (in time) versions of the original differential game, and adopting a generalized definition of a strategy on "supergames" "à la Friedman" [10], we are able to provide a definition of what we call a "strong equilibrium" (as opposed to "weak equilibrium" which is also elucidated in the text). For the special case of deterministic differential games with a fixed asymmetric mode of play, this new concept provides a natural counterpart of the discrete-time feedback Stackelberg solution.

After introducing the general "strong equilibrium" solution, we also derive in section 4 the Hamilton-Jacobi equation associated with the optimal feedback solutions. Then, in section 5, we treat some special cases, viz. the purely deterministic differential game with a fixed asymmetric mode of play, and the linear-quadratic differential game in which the jump process determines (independent of the current value of the state) only the mode of the play. For the former case we show that, in general, the feedback Stackelberg solution is different from the feedback Nash equilibrium solution, and in this context we clarify and extend a result obtained in [10] with regard to limit points of equilibria of $G(\delta)$ -games under asymmetric modes of play.

Section 6 includes some discussions and concluding remarks, and the two appendices provide derivations of some of the results used in the main body.

2. Strong and Weak Equilibrium Solutions, and the Relationship Between Stackelberg and Nash Equilibria

Let $U_k \subseteq \mathbb{R}^{m_k}$ be a measurable space to which the control variable u_k of player k (P_k) belongs ($k=1,2$). Let $J_k(u_1, u_2)$, $J_k: U_1 \times U_2 \rightarrow \mathbb{R}$, be a real-valued function denoting the cost functional of P_k . Stipulating an asymmetry in the roles of the players, let P_1 be the leader, announcing his control (constant policy) $u_1^0 \in U_1$ first, to which P_2 reacts optimally by minimizing $J_2(u_1^0, u_2)$ over $u_2 \in U_2$. Let us assume that the reaction set

$$R_2(u_1^0) = \{u_2^0 \in U_2 : J_2(u_1^0, u_2^0) = \min_{u_2 \in U_2} J_2(u_1^0, u_2)\} \quad (2.1)$$

is a singleton, so that there exists a unique mapping $T_2: U_1 \rightarrow U_2$ with the property

$$J_2(u_1, T_2(u_1)) = \min_{u_2 \in U_2} J_2(u_1, u_2), \quad \forall u_1 \in U_1. \quad (2.2)$$

Then, we call a pair $(u_1^*, u_2^*) \in U_1 \times U_2$ a Stackelberg solution [1,2,3] for the static game, with P_1 as the leader, if

$$u_1^* = \arg \min_{u_1 \in U_1} J_1(u_1, T_2(u_1)) \quad (2.3)$$

$$u_2^* = T_2(u_1^*).$$

A Stackelberg solution with P_2 as the leader can be defined analogously, by interchanging the roles of the players. Furthermore, a pair (u_1^N, u_2^N) is called a Nash equilibrium solution [3,4] for a static game in which the

roles of the players are symmetric, if it satisfies the pair of inequalities

$$J_1(u_1^N, u_2^N) \leq J_1(u_1, u_2^N) \quad \forall u_1 \in U_1 \quad (2.4a)$$

$$J_2(u_1^N, u_2^N) \leq J_2(u_1^N, u_2) \quad \forall u_2 \in U_2 \quad (2.4b)$$

For dynamic games the same definitions of Stackelberg and Nash equilibrium solutions apply, provided that we have a normal form description of the game [3], in which case u_k and U_k will have to be interpreted the strategy and strategy space, respectively, of P_k . Moreover, static Stackelberg games can be viewed as special types of dynamic Nash game wherein the Stackelberg solution concept coincides with a particular type of Nash equilibrium solution, as to be elucidated in the sequel.

Consider the static two person game $\{J_k, U_k; k=1,2\}$ introduced earlier, with P_1 acting as the leader. Introduce a 2-stage nonzero sum single act dynamic game $\{\tilde{J}_k, \tilde{U}_k; k=1,2\}$ whose extensive form description is as follows:

State equations:

$$x_2 = (u_1', 0_{m_2}') \quad (2.5a)$$

$$x_3 = (0_{m_1}', u_2') \quad (2.5b)$$

Strategies: Constant mapping $\gamma_1: \rightarrow U_1$ for P_1 [i.e., $u_1 = U_1$]; measurable mapping $\gamma_2(x_2)$, $\gamma_2: \mathbb{R}^m \rightarrow U_2$ [$m=m_1+m_2$], for P_2 , with the corresponding space denoted $\mathcal{U}_2$.

Cost functions: $\tilde{J}_k = J_k(\tilde{x}_2, \tilde{x}_3)$, where

$$\tilde{x}_2 \triangleq [I_{m_1}, 0_{m_1 \times m_2}] x_2 \quad (2.6)$$

$$\tilde{x}_3 \triangleq [0_{m_2 \times m_1}, I_{m_2}] x_3$$

Now note that to each $\gamma_2 \in \mathcal{U}_2$ there corresponds a unique $\tilde{\gamma}_2: \mathbb{R}^{m_1} \rightarrow U_2$, satisfying $\tilde{\gamma}_2(\tilde{x}_2) \equiv \gamma_2(x_2)$, so that $\tilde{J}_k(u_1, \gamma_2(u_1)) \equiv J_k(u_1, \tilde{\gamma}_2(u_1))$. Hence, if $(u_1^*, u_2^*) \in U_1 \times U_2$ is a Stackelberg solution with P1 as the leader, relations (2.2) and (2.3) imply that

$$\left. \begin{aligned} J_1(u_1^*, T_2(u_1^*)) &\leq J_1(u_1, T_2(u_1)) \quad \forall u_1 \in U_1 \\ J_2(u_1^*, T_2(u_1^*)) &\leq J_2(u_1^*, \tilde{\gamma}_2(u_1^*)) \quad , \quad \tilde{\gamma}_2: U_1 \rightarrow U_2 \\ u_2^* &= T_2(u_1^*) \end{aligned} \right\} \quad (2.7)$$

$\Leftrightarrow$ there exists a $\gamma_2^* \in \mathcal{U}_2$, $\gamma_2^*([u_1^*, 0_{m_2}^T]^T) = u_2^*$, such that

$$\left. \begin{aligned} \tilde{J}_1(u_1^*, \gamma_2^*) &\leq \tilde{J}_1(u_1, \gamma_2^*) \quad , \quad \forall u_1 \in U_1 \\ \tilde{J}_2(u_1^*, \gamma_2^*) &\leq \tilde{J}_2(u_1^*, \gamma_2^*) \quad , \quad \forall \gamma_2 \in \mathcal{U}_2 \end{aligned} \right\} \quad (2.8)$$

[Note that $\gamma_2^*(x_2) = T_2(x_2)$].

Therefore, the conclusion is that, if $(u_1^*, u_2^* = T_2(u_1^*))$ constitutes a Stackelberg solution to the static game $\{J_k, U_k; k=1,2\}$, with P1 as the leader, the strategy pair $(u_1^*, \gamma_2^*) \in \mathcal{U}_1 \times \mathcal{U}_2$, with $\gamma_2^*(x_2) = T_2(x_2)$, $\forall x_2 \in U_2$,

constitutes a Nash equilibrium solution to the 2-stage dynamic game $\{\tilde{J}_k, u_k; k=1,2\}$. However, the converse statement is not true; that is, every Nash solution of $\{\tilde{J}_k, u_k; k=1,2\}$ (satisfying inequalities (2.8)) does not correspond to a Stackelberg solution of the static game, mainly because of informational nonuniqueness [3,5,6]. A delayed commitment (feedback) Nash equilibrium solution $(u_1^0, \gamma_2^0) \in U_1 \times U_2$, for the dynamic game, on the other hand, satisfies the pair of equalities [3]

$$\begin{aligned} \tilde{J}_1(u_1^0, \gamma_2^0(x_2^0)) &\leq \tilde{J}_1(u_1, \gamma_2^0(x_2)) \quad , \quad \forall u_1 \in U_1 \\ \tilde{J}_2(u_1, \gamma_2(x_2)) &\leq \tilde{J}_2(u_1, \gamma_2(x_2)) \quad , \quad \forall u_1 \in U_1, \gamma_2 \in U_2, \end{aligned} \quad (2.9)$$

which can equivalently be written as

$$\begin{aligned} J_1(u_1^0, \tilde{\gamma}_2^0(u_1^0)) &\leq J_1(u_1, \tilde{\gamma}_2^0(u_1)) \quad , \quad \forall u_1 \in U_1 \\ J_2(u_1, \tilde{\gamma}_2(u_1)) &\leq J_2(u_1, \tilde{\gamma}_2(u_1)) \quad , \quad \forall u_1 \in U_1, \tilde{\gamma}_2: U_1 \rightarrow U_2, \end{aligned} \quad (2.10)$$

where $\tilde{\gamma}_2(u_1) = \gamma_2([u_1', o_{m_2}'])$.

Hence, $(u_1^0, u_2^0 = \tilde{\gamma}_2^0(u_1^0))$ constitutes a Stackelberg solution to $\{J_k, U_k; k=1,2\}$ with P1 as the leader. We call such an equilibrium, when viewed as a (feedback) Nash equilibrium of a dynamic game with hierarchical decision structure, the "strong equilibrium," as opposed to any (informationally nonunique Nash) solution (u_1^*, γ_2^*) satisfying (2.8), which we call a "weak equilibrium."

We now have the following result, which follows from the preceding discussion and analysis.

Proposition 2.1.

If $\{J_1, J_2; U_1, U_2\}$ is a static game admitting a unique Stackelberg solution with P_k as the leader, there exists a single-act 2-stage dynamic feedback game $\{\tilde{J}_1, \tilde{J}_2; \mathcal{U}_k, U_k, k \neq k\}$ which admits a unique feedback Nash (synonymously, delayed commitment type or strong) equilibrium solution. Furthermore, there is a unique correspondence between these two solutions. $\square$

Hence, every Stackelberg solution of a static two-person game can be viewed as a strong (Nash) equilibrium solution of a particular dynamic feedback game with perfect state information. This correspondence can in fact be extended to the feedback Stackelberg solution of dynamic feedback games [2,3], by identifying it as the feedback Nash equilibrium solution of related dynamic feedback games with twice as many stages. Towards this end, consider the N stage dynamic feedback game with state dynamics

$$\begin{aligned} x(n+1) &= f_n[x(n), u_1(n), u_2(n)] \quad ; \quad n=0, 1, \dots, N-1 \\ x(n) \in X \subseteq \mathbb{R}^{m_0}, \quad u_k(n) \in U_{k,n} \subseteq \mathbb{R}^{m_k}, \quad k=1, 2, \end{aligned} \quad (2.11)$$

and cost functionals

$$J_k = q_k[x_N] + \sum_{n=0}^{N-1} g_{k,n}[x(n), u_1(n), u_2(n)] \quad ; \quad k=1, 2, \quad (2.12)$$

where $f_n, q_k, g_{k,n}$ are mappings of appropriate specifications. Controls are allowed to depend only on the current value of the state, so that admissible policies $\gamma_{k,n} \in \mathcal{U}_{k,n}$ for P_k , at stage n , are $\gamma_{k,n}(x(n))$, $\gamma_{k,n}: X \rightarrow U_{k,n}$, satisfying certain measurability requirements.

Consider a stagewise asymmetric mode of play whereby one of the players (say, P1) announces his strategy and moves before the other player does, at each stage. The relevant solution concept in this case is the so-called feedback Stackelberg solution [2,3] which is defined recursively in retrograde time and involves the solution of a static Stackelberg game with P1 as the leader at each stage. By basically following the arguments that led to Proposition 2.1, it is not difficult to see that the feedback Stackelberg solution of the N-stage game (2.11)-(2.12) corresponds uniquely to the feedback Nash equilibrium solution of a 2N-stage dynamic feedback game (with perfect state information) defined as follows:

State equation:

$$y(s+1) = F_s[y(s), \bar{u}_1(s), \bar{u}_2(s)]; \quad y(0) = \begin{bmatrix} x(0) \\ 0_{m_1} \end{bmatrix} \quad (2.13)$$

where

$$F_s[y, \bar{u}_1, \bar{u}_2] = \begin{cases} y(s) + [0_{m_0}, \bar{u}_1'(s)]', & s \text{ even} \\ [f', \frac{s-1}{2} [(I_{m_0}, 0_{m_0 \times m_1})y(s), (0_{m_0 \times m_1}, I_{m_1})y(s), \bar{u}_2(s), 0_{m_1}]], & s \text{ odd} \end{cases} \quad (2.14)$$

$$\left. \begin{aligned} \bar{u}_1(s) &\equiv u_1\left(\frac{s}{2}\right), & s \text{ even} \\ \bar{u}_2(s) &\equiv u_2\left(\frac{s-1}{2}\right), & s \text{ odd} \end{aligned} \right\} \quad (2.15a)$$

$$y(s) \triangleq \begin{cases} [x'(\frac{s}{2}), o_{m_1}']', & s \text{ even} \\ [x'(\frac{s-1}{2}), u_1'(\frac{s-1}{2})]', & s \text{ odd} \end{cases}; \quad y(s) \in Y \subseteq \mathbb{R}^{m_0+m_1}. \quad (2.15b)$$

Cost functionals:

$$\tilde{J}_k = \tilde{q}_k[y(2N)] + \sum_{s=0}^{2N-1} \tilde{g}_{k,s}[y(s), \bar{u}_1(s), \bar{u}_2(s)], \quad k=1,2, \quad (2.16)$$

where

$$\tilde{q}_k[y(2N)] = q_k[(I_{m_0}, 0_{m_0 \times m_1})y(2N)] \quad (2.17)$$

$$\tilde{g}_{1,s}[y, \bar{u}_1, \bar{u}_2] = \begin{cases} g_{k, \frac{s-1}{2}}[(I_{m_0}, 0_{m_0 \times m_1})y(s), \bar{u}_1(s), \bar{u}_2(s)]; & s \text{ odd} \\ 0 & ; s \text{ even} \end{cases} \quad (2.18)$$

$$\tilde{g}_{2,s}[y, \bar{u}_1, \bar{u}_2] = \begin{cases} g_{k,s}[(I_{m_0}, 0_{m_0 \times m_1})y(s), (0_{m_0 \times m_1}, I_{m_1})y(s), \bar{u}_2(s)] & ; s \text{ odd} \\ 0 & ; s \text{ even} \end{cases} \quad (2.19)$$

Admissible control laws: $\tilde{\gamma}_{1,s}(y(s)), \tilde{\gamma}_{1,s}: Y \rightarrow U_1$, for P1; s even (2.20)

$\tilde{\gamma}_{2,s}(y(s)), \tilde{\gamma}_{2,s}: Y \rightarrow U_2$, for P2; s odd.

Let $\tilde{\Gamma}_{1,s}$ and $\tilde{\Gamma}_{2,s}$ denote the corresponding strategy spaces. The following proposition now follows as a counterpart of Proposition 2.1 for this more general class of dynamic games.

Proposition 2.2.

If $\{J_k, \Gamma_{k,n}; k=1,2; n=0,1,\dots,N-1\}$ is an N -stage dynamic feedback game as defined by (2.11)-(2.12), admitting a unique feedback Stackelberg solution (γ_1^*, γ_2^*) with P_1 as the leader (at each stage), there exists a $2N$ -stage dynamic feedback game $\{\tilde{J}_k, \tilde{\Gamma}_{k,s}; k=1,2; s=0,1,\dots,2N-2\}$ defined by (2.13)-(2.20), which admits a unique feedback Nash (strong) equilibrium solution $(\tilde{\gamma}_1^*, \tilde{\gamma}_2^*)$. Furthermore, there is a unique correspondence between these two solutions, given by

$$\left. \begin{aligned} \tilde{\gamma}_{1,2n}^*([x(n), \circ_{m_1}]) &= \gamma_{1,n}^*(x(n)) \\ \tilde{\gamma}_{2,2n+1}^*([x(n), \gamma_{1,n}^*(x(n))]) &= \gamma_{2,n}^*(x(n)) \end{aligned} \right\} \quad (2.21)$$

$n=0,1,\dots,N-1.$ □

Remark 2.1.

Since every feedback Nash (synonymously, strong) equilibrium solution is a ("weak") Nash equilibrium solution [3,6], the feedback Stackelberg solution of the original dynamic game is also a Nash equilibrium solution of the $2N$ -stage dynamic game constructed prior to Proposition 2.2. However, we cannot claim a unique correspondence between the two games in the framework of Nash equilibria, because there exists informationally nonunique weak equilibria in the latter case. [Note that the further restriction to delayed commitment strategies (feedback Nash equilibria) eliminates this informational nonuniqueness, as discussed

extensively in [3], and leads to strong equilibria]. The important conclusion here, though, is that the feedback Stackelberg solution in dynamic feedback games is indeed an equilibrium solution, which is readily seen by reformulating the problem in an appropriate framework. $\square$

3. Stochastic Dynamic Games with Structural and Modal Uncertainties Described by Jump Processes

3.1. Problem Formulation

Having settled the problem of identifying the feedback Stackelberg solution as an equilibrium solution concept, we now turn to introducing and solving a general class of such problems in which both the structure of the system dynamics (i.e., transitions from one state to another) and the mode of play (i.e., whether the roles of the players are asymmetric or not, and in case of asymmetry which player becomes the leader) are uncertain and are determined by the outcome of a finite state jump process.

More specifically, consider the N-stage stochastic dynamic game, with state

$$y(n) = [x(n)', r(n)']' \in X \times I \quad (3.1)$$

at stage $n \in \bar{N} = \{0, 1, \dots, N-1\}$, where $X \subseteq \mathbb{R}^{m_0}$, $I = S_1 + S_2 + \mathcal{N}$, the sets S_1 , S_2 and $\mathcal{N}$ being finite and disjoint. If $r(n) \in S_k$, then there is asymmetry in the roles of the players and P_k acts as the leader at stage n , whereas, if $r(n) \in \mathcal{N}$, there is no asymmetry and the players choose their controls in accordance with the Nash equilibrium solution concept at stage n .

The control of P_k at stage n is denoted $u_k(n) \in U_{k,n}$, and the probability of transition from state $y(n) = (x(n) = x, r(n) = i)$ to state

$y(n+1) = (x(n+1) \in dX \subseteq X, h(n+1) = j)$, as a result of feedback controls

$u_k(n) = \gamma_{k,n}[x(n), r(n)] = u_k$; $k=1,2$, is given by

$$Q(dx, j; x, i, u_1, u_2) = P \{x(n+1) \in dX, r(n+1) = j | x(n) = x, r(n) = i, u_1(n) = u_1, u_2(n) = u_2\} \quad (3.2)$$

where $Q \geq 0$ and

$$\sum_{j \in I} \int_X Q(dX, j; x, i, u_1, u_2) = 1, \quad \forall i \in I, u_k \in U_{k,n}.$$

The control law of P_k at stage n , $\gamma_{k,n}[x(n), r(n)]$, is a measurable mapping $\gamma_{k,n}: X \times I \rightarrow U_{k,n}$, and the cost functional of P_k is $J_{k,0}$, where

$$J_{k,p}(\gamma_1^p, \gamma_2^p; x^0, i^0) = E_{\substack{x(p) = x^p \\ r(p) = i^p}} \{ q_{k,N}[x(N)] + \sum_{n=p}^{N-1} g_{k,n} [x(n), r(n), u_1(n), u_2(n)] \} \quad (3.3)$$

where $E\{\cdot\}$ is the expectation operation with respect to the probability measures that govern the transition probabilities (3.2), with $\gamma_{k,n} \in \Gamma_{k,n}$ ($k=1,2; n=0, \dots, N-1$) fixed, and γ_k^p denotes the set of policies $\{\gamma_{k,n}, n=p, p+1, \dots, N-1\}$.

Note that this is a stochastic dynamic game with perfect state information for both players, but not a standard one because the mode of play at each stage is determined by the outcome of a jump process $\{r(\cdot)\}$, which in turn is affected by the past decisions through (3.2). However, since the state of the game involves the current value of $r(\cdot)$, both players know the mode of play at the current stage (i.e., whether there is asymmetry or symmetry in decision making at that stage, and in the former case which player acts as the leader), and therefore, the equilibrium solution is well-defined stagewise. Hence, the game is played as follows:

Stage 0: Both players observe the value of $x(0)=x^0$ and the outcome of the random variable $r(0)=r^0$. If $r^0 \in S_k$, first P_k chooses his control u_k^0 (and announces it) and then $\bar{P}_k$ reacts to that by announcing his control $u_{\bar{k}}^0$; if $r^0 \in \eta$, both players choose their controls (u_1^0 and u_2^0) simultaneously. Then, a transition to a new state $y(1)=[x(1)', r(1)']$ takes place under the stationary transition probability $Q(dx, j; x^0, r^0, u_1^0, u_2^0)$, and a cost of $g_{k,0}[x^0, r^0, u_1^0, u_2^0]$ is incurred by P_k .

• • • • • • • • •
• • • • • • • • •

Stage n: Both players observe the value of $x(n)=x^n$ and the outcome of the random variable $r(n)=r^n$. If $r^n \in S_k$, first P_k chooses his control u_k^n (and announces it) and then $\bar{P}_k$ reacts to that by announcing his control $u_{\bar{k}}^n$, (i.e., P_k has an informational advantage over $\bar{P}_k$); if $r^n \in \eta$, however, both players choose their controls (u_1^n, u_2^n) simultaneously. Then, a transition to a new state $y(n+1) = [x(n+1)', r(n+1)']$ takes place under the stationary transition probability $Q(dx, j; x^n, r^n, u_1^n, u_2^n)$, and an additional cost of $g_{k,n}[x^n, r^n, u_1^n, u_2^n]$ is incurred by P_k .

- - - - - - - - -
- - - - - - - - -

Stage N-1: Both players observe the value of $x(N-1)=x^{N-1}$ and the outcome of the random variable $r(N-1)=r^{N-1}$. If $r^{N-1} \in S_k$, first P_k chooses his control u_k^{N-1} (and announces it) and then $\bar{P}_k$ reacts to that by announcing his control $u_{\bar{k}}^{N-1}$; if $r^{N-1} \in \eta$, both players choose their controls (u_1^{N-1}, u_2^{N-1}) simultaneously. Then, a transition to $x(N)$ takes place under the stationary marginal transition probability $\sum_j Q(dx, j; x^{N-1}, r^{N-1}, u_1^{N-1}, u_2^{N-1})$, and an

additional cost of $g_{k,N}[x(N)] + g_{k,N-1}[x^{N-1}, r^{N-1}, u_1^{N-1}, u_2^{N-1}]$ is incurred by P_k , so that the total cost adds up (for each sample path) to $q_{k,N}[x(N)] + \sum_{n=0}^{N-1} g_{k,n}[x^{N-1}, r^{N-1}, u_1^{N-1}, u_2^{N-1}]$, for P_k .

3.2. The Concept of Strong Equilibrium

Even though the stochastic game and the moves of the players for a particular realization are delineated in forward time, the equilibrium solution is defined in retrograde time. Towards this end, we first consider a single stage game which comprises only the last stage of the stochastic game of §3.1, with cost function $J_{k,N-1}$, for P_k . Here, if the outcome of the random variable $r(N-1)$ belongs to S_k , then the players choose their controls (as functions of $x(N-1) \in X$ which is arbitrary) according to the Stackelberg solution concept, with P_k acting as the leader; if, however, $r(N-1)$ belongs to the index set $\mathcal{N}$, the players determine their equilibrium controls according to the Nash solution concept and for all values of $x(N-1)$. Assuming that these solutions are unique in each case (or that there is mutual agreement between the players as to which pair of controls to adopt in case of nonunique equilibria), there will be a unique pair of expected cost-to-go values transferred to stage $N-1$ in terms of the stationary transition probability $Q(\cdot)$.

Next, we consider the 2-stage dynamic game problem with cost functions $J_{k,N-2}, k=1,2$, and with the policies $\gamma_{k,N-1}^*[x(N-1), r(N-1)], k=1,2$, at stage $N-1$ being fixed as determined above. Then, this is again basically a single stage game, with the mode of play depending on the outcome of

$r(N-2)$, as at stage $N-1$; and stipulating existence of a unique equilibrium for each element of I and for all $x(N-2) \in X$, we obtain unique equilibrium policies $\gamma_{k,N-2}^*[x(N-2), r(N-2)]$, $k=1,2$, leading to a unique pair of expected cost-to-go values to be transferred to stage $N-2$.

If this procedure is followed up inductively, up to the initial stage $n=0$, we obtain a pair of N -tuple policies $\{\gamma_{1,N-1}^*, \dots, \gamma_{1,0}^*; \gamma_{2,N-1}^*, \dots, \gamma_{2,0}^*\} \triangleq \{\gamma_1^0, \gamma_2^0\}$ which we call a "strong equilibrium" for the stochastic dynamic game of §3.1. Note that this is indeed an equilibrium solution, since it can be shown by following the arguments and the procedure of Section 2 that it is related to the feedback Nash equilibrium solution of a dynamic game with twice as many stages [see Appendix I]. In fact, when $\mathcal{N}=\emptyset$ and $S_2=\emptyset$, strong equilibrium is identical with the stochastic feedback Stackelberg equilibrium with P_1 as the leader [3], and when $S_k=\emptyset$, $k=1,2$, it coincides with the concept of feedback (delayed commitment type) Nash equilibrium in stochastic dynamic games (which we have also called "strong equilibrium" in Proposition 2.2).

3.3. Derivation of Strong Equilibria

A set of necessary and sufficient conditions for a strong equilibrium solution of the stochastic dynamic game of §3.1 can be obtained by basically following the procedure outlined in §3.2, in the spirit of dynamic programming since, for each fixed pair of controls, the state $\{y(n)\}$ is a Markov process, and furthermore the controls are restricted to be Markov (feedback) controls depending only on the

current value of the state. The result is the dynamic programming type equations given below in Proposition 3.1 in terms of the optimum (strong equilibrium) expected cost-to-go functions $V_{k,n}(x,i), k=1,2; n \in \bar{N}$.

Proposition 3.1.

Let $\{u_1^* = \gamma_1^*(\cdot), u_2^* = \gamma_2^*(\cdot)\}, (\gamma_1^*, \gamma_2^*) \in \Gamma_1 \times \Gamma_2$, denote a strong equilibrium solution for the stochastic dynamic game of §3.1, and $J_k^*(x^0, i^0)$ denote the corresponding cost to P_k when the initial state is $x(0) = x^0$, $r(0) = i^0$. Then, it is necessary and sufficient that the following relations are satisfied:

$$J_k^*(x^0, i^0) = V_{k,0}(x^0, i^0) \quad , \quad k=1,2, \quad (3.4)$$

where $V_{k,n}(x,i)$ is recursively defined by

$$V_{k,n}(x,i) = \min_{u_k \in U_{k,n}} \left\{ \int \sum_{j \in I} V_{k,n+1}(\xi, j) Q[d\xi, j; x, i, \pi_k(u_k, T_{k,n}(u_k, x, i))] + g_{k,n}[x, i, \pi_k(u_k, T_{k,n}(u_k, x, i))] \right\}, \quad (3.5a)$$

$i \in S_k$

$$V_{k,n}(x,i) = \int \sum_{j \in I} V_{k,n+1}(\xi, j) Q[d\xi, j; x, i, \pi_k(T_{k,n}[\gamma_{k,n}^*(x,i), x, i], \gamma_{k,n}^*(x,i)) + g_{k,n}[x, i, \pi_k(T_{k,n}[\gamma_{k,n}^*(x,i), x, i], \gamma_{k,n}^*(x,i))] \quad , \quad i \in S_{\bar{k}} \quad (3.5b)$$

$$\begin{aligned}
 V_{k,n}(x,i) = \min_{u \in U_{k,n}} \{ \int \sum_{j \in I} V_{k,n+1}(\xi,j) Q[d\xi,j; x,i, \pi_k(u_k, \gamma_{k,n}^*(x,i))] \\
 k=1,2 \quad (3.5c) \\
 + g_{k,n}[x,i, \pi_k(u_k, \gamma_{k,n}^*(x,i))], i \in \eta,
 \end{aligned}$$

with the boundary condition

$$V_{k,N}(x,i) = q_{k,N}[x(N)] \quad , \quad \forall i \in I \quad , \quad k=1,2. \quad (3.6)$$

Here, $T_{k,n}(u_k, x, i)$ is defined by

$$\begin{aligned}
 T_{k,n}(u_k, x, i) = \arg \min_{u_k \in U_{k,n}} \{ \int \sum_{j \in I} V_{k,n+1}(\xi,j) Q[d\xi,j; x,i, u_1, u_2] \\
 (3.7) \\
 + g_{k,n}[x,i, u_1, u_2] \}, i \in S_k,
 \end{aligned}$$

$$\pi(u_k, T_{k,n}) \triangleq \begin{cases} u_1, T_{2,n} & \text{if } k=1 \\ T_{1,n}, u_2 & \text{if } k=2 \end{cases} \quad (3.8)$$

and

$$\gamma_{k,n}^*(x,i) = \begin{cases} \text{argument of (3.5a) for all } x \in X, \text{ if } i \in S_k \\ T_{k,n}[\gamma_{k,n}^*(x,i), x, i] \quad , \quad \text{if } i \in S_k^- \\ \text{argument of (3.5c) for all } x \in X \quad , \text{ if } i \in \eta. \end{cases} \quad (3.9)$$

The strong equilibrium solution is unique whenever (3.7) and (3.9) are uniquely defined. □

Remark 3.1.

One special class of problems whose strong equilibrium solution can be determined explicitly is the class characterized by a linear state equation, quadratic cost functionals and with the transition probabilities (3.2) independent of the state and controls, i.e., $Q(dx, j; dx, i, u_1, u_2) = \lambda_{ij}$, where λ_{ij} 's are constants. In this case the equilibrium strategies will be linear functions of the current value of the state, with the multiplying gain matrices depending on the outcome of the Markov jump process $\{r(\cdot)\}$; exact expressions can readily be determined from (3.5) by recursive evaluation. □

4. Feedback Equilibria for Differential Games with Jump Disturbances

4.1. The Controlled Stochastic System

Consider a stochastic system of the form

$$\dot{x} = f^{r(t)}(t, x, u_1, u_2), \quad x \in X, \quad u_k \in U_k, \quad k=1,2 \quad (4.1)$$

with an initial condition

$$x(0) = x^0, \quad r(0) = r^0. \quad (4.2)$$

In (4.1), $r(t)$ is a finite-state stochastic jump process and the RHS changes from $f^i(t, x, u_1, u_2)$ to $f^j(t, x, u_1, u_2)$ as $r(t)$ jumps from i to j . In (4.2), r^0 is a random variable determining the initial state of the process $r(t)$.

The state x belongs to $X \subseteq \mathbb{R}^{m_0}$, and the control u_k takes values in $U_k \subseteq \mathbb{R}^{m_k}$, ($k=1,2$). At a fixed terminal time T , there are bounded functions

$$q_k^i(x), \quad q_k^i: X \rightarrow \mathbb{R} \quad k=1,2, \quad i \in I \quad (4.3)$$

which are continuously differentiable, with bounded derivatives in x , and they determine the respective terminal costs incurred by the two players $k=1,2$, if $r(T)=i$ and $x(T)=x$. Let I denote the state set of $r(t)$. For each i in I , let $f^i(t, x, u_1, u_2)$,

$$f^i: [0, T] \times X \times U_1 \times U_2 \rightarrow \mathbb{R}^{m_0}$$

be a continuous bounded function, continuously differentiable with bounded partial derivatives in x , u_1 and u_2 . Let $\mathcal{U}_k, k=1,2$ be two classes of admissible control laws $u_k^i(t,x)$ with values in U_k , defined on $I \times [0,T] \times X$ such that $u_k^i(t,x)$ is piecewise continuous in t , continuously differentiable with bounded derivatives in x .

In order to introduce the controlled stochastic process, we suppose that a measurable space $(\Omega, \mathcal{F})$ is given, called the sample space. We consider a function $y(t,\omega)$,

$$y : [0, T] \times \Omega \rightarrow X \times I$$

$$y(t,\omega) = (x(t,\omega), r(t,\omega))',$$

which is measurable* w.r.t. $\mathcal{B}_{[0,T]} \times \mathcal{F}$.

Let $\mathcal{F}_t = \sigma\{y(s, \cdot) | s \leq t\}$ be the σ -fields generated by past observations of y up to time t .

We now assume

A 1. The behavior of the system under any admissible control law $u \in \mathcal{U}_1 \times \mathcal{U}_2$ is completely described by a probability measure ϕ^u on $(\Omega, \mathcal{F}_T)$.

* Assuming a measurable state space $(X \times I, \mathcal{X}_I)$ and denoting by $\mathcal{B}_{[0,T]}$ the Borel σ -field on $[0,T]$.

Thus, the process

$$y^u = (y(t, \cdot), \mathfrak{F}_t, \theta^u), \quad t \in [0, T]$$

is a well-defined stochastic process.

A 2. For any control law $u \in \mathcal{U}_1 \times \mathcal{U}_2$ and almost any ω in Ω , there exists a piecewise constant function $\mu_\omega(t, x)$

$$\mu_\omega : [0, T] \times X \rightarrow I, \quad ,$$

such that $y(t, \omega) = [x'(t, \omega), r'(t, \omega)]'$ satisfies the following equations

$$\dot{x} = f_{\mu_\omega(t, x)}(t, x, u_1(t, x), u_2(t, x)) \quad (4.4)$$

$$r(t, \omega) = \mu_\omega(t, x) \quad (4.5)$$

where $r(t, \omega)$ is a step function with a finite number of jumps on $[0, T]$.

A 3. For each (t, x) in $[0, T] \times X$, there exists a matrix with elements $\{\lambda_{ij}(t, x)\}$ if which are real-valued, continuous, bounded functions, such that

$$\lambda_{ij}(t, x) \geq 0, \quad \forall t \in [0, T], \quad x \in X, \quad i \neq j \quad (4.6)$$

$$\sum_{j \in I} \lambda_{ij}(t, x) = 0, \quad \forall t \in [0, T], \quad x \in X$$

and such that for any admissible control law $\tilde{u} \in \mathcal{U}_1 \times \mathcal{U}_2$

$$\begin{aligned} P^{\tilde{u}} [\mu(t+h, x(t+h))=j \mid \mu(t, x(t))=i, x(t)=x] \\ = \lambda_{ij}(t, x)h + o(h; \tilde{u}, x) \end{aligned} \quad (4.7)$$

$$P^{\tilde{u}} [\mu(t+h, x(t+h))=i | \mu(t, x(t))=i, x(t)=x] =$$

$$1 + \lambda_{ii}(t, x)h + o(h; \tilde{u}, x) \quad (4.8)$$

where $o(h; \tilde{u}, x)$ is a quantity such that

$$\lim_{h \downarrow 0} \frac{o(h; \tilde{u}, x)}{h} = 0 \quad (4.9)$$

uniformly for all x in X and $\tilde{u}$ in $\mathcal{U}_1 \times \mathcal{U}_2$.

The assumption A1 allows the modeling of the system as a controlled probability space. Assumption A2 describes the assumed relationship between $x(t)$ and $r(t)$ through the differential system (4.1). Assumption A3 is a conditional Markov assumption on the jump process.

For an admissible control $u \in \mathcal{U}_1 \times \mathcal{U}_2$, $u \triangleq (u_1^i(t, x), u_2^i(t, x))$ let $V_k^i(t, x)$ denote the corresponding values of the conditional expectation:

$$V_k^i(t, x) = E[q_k^{r(T)}(x(T)) | x(t)=x, r(t)=i] \quad , \quad k=1,2 \quad *$$

which will henceforth be referred to as the cost-to-go function for player k . As regards this function, we now state two lemmas which will be used later in our analysis.

Lemma 4.1: The cost-to-go functions $V_k^i(t, x)$, $k=1,2$ associated with an admissible control law $\tilde{u} \in \mathcal{U}_1 \times \mathcal{U}_2$ satisfy the system of partial differential equations

*Note that we have only terminal cost function for both players, which simplifies the mathematical derivation to be given in the sequel, without bringing in any real loss of generality.

$$\begin{aligned}
& \frac{\partial}{\partial t} v_k^i(t, x) + \frac{\partial}{\partial x} v_k^i(t, x) f^i(t, x, u_1(t, x), u_2(t, x)) \\
& + \sum_{j \in I} \lambda_{ij}(t, x) v_k^j(t, x) = 0 \quad , \quad i \in I
\end{aligned} \tag{4.10}$$

with boundary conditions

$$v_k^i(T, x) = q_k^i(x) \quad , \quad i \in I \quad , \quad x \in X \quad , \quad k=1,2.$$

Proof: See Appendix II. □

Lemma 4.2: Maximum principle for the partial differential operator:

If $h_k^i(t, x)$, $i \in I$, $k \in \{1, 2\}$, is a bounded continuous function on $[0, T] \times X$, is continuously differentiable in x and piecewise continuously differentiable in t , such that

$$\begin{aligned}
& h_k^i(T, x) \leq 0 \\
& \frac{\partial}{\partial t} h_k^i(t, x) + \frac{\partial}{\partial x} h_k^i(t, x) f^i(t, x, u_1(t, x), u_2(t, x)) + \\
& \sum_{j \in I} \lambda_{ij}(t, x) h_k^j(t, x) \geq 0 \quad ,
\end{aligned}$$

then

$$h_k^i(t, x) \leq 0 \quad \text{on} \quad [0, T] \times X.$$

Proof: Is similar to the proof of Theorem 3 in [7]. □

4.2. $G(\delta)$ - Differential Games

Consider a fixed partition of the set I of possible values of $r(t)$ into three subsets

$$I = S_1 + S_2 + \mathcal{N} \quad (4.11)$$

Consider also a partition of $[0, T]$ into N subintervals of length $\delta = \frac{T}{N}$

$$[0, T] = [0, t_1) \cup [t_1, t_2) \cup \dots \cup [t_{N-1}, T],$$

and on each subinterval $[t_\ell, t_{\ell+1})$ introduce the classes $\tilde{U}_k^{\delta, \ell}$ of controls

$$u_k^{\delta, \ell}(\cdot) : [t_\ell, t_{\ell+1}) \rightarrow U_k, \quad k=1, 2.$$

A δ -strategy for player k is a vector

$$\gamma_k^\delta = (\gamma_k^{\delta, \ell})_{\ell=0, 1, \dots, N-1}$$

where each component $\gamma_k^{\delta, \ell}$ is a function

$$\gamma_k^{\delta, \ell} : X \times I \times \tilde{U}_k^{\delta, \ell} \rightarrow \tilde{U}_k^{\delta, \ell} \quad (4.12)$$

which associates a unique control $u_k^{\delta, \ell}$ in $\tilde{U}_k^{\delta, \ell}$,

$$u_k^{\delta, \ell}(\cdot) = \gamma_k^{\delta, \ell}(x(t_\ell), r(t_\ell), u_k^{\delta, \ell}(\cdot)) \quad ,$$

with each state $(x'(t_\ell), r'(t_\ell))'$ observed at sampled time t_ℓ and each control $u_k^{\delta, \ell}(\cdot)$ in $\tilde{U}_k^{\delta, \ell}$ (as in §3, $\bar{k}$ stands for 2 if $k=1$, and 1 if $k=2$), such that the following condition holds

$$r(t_\ell) = i \in \mathcal{N} \cup S_k \Rightarrow \gamma_k^{\delta, \ell}(x(t_\ell), r(t_\ell), u_k^{\delta, \ell}(\cdot)) = \hat{\gamma}_k^{\delta, \ell}(x(t_\ell), r(t_\ell)) \quad (4.13)$$

where $\hat{\gamma}_k^{\delta, \ell} : X \times I \rightarrow \tilde{U}_k^{\delta, \ell}$.

This last condition ensures that, if $r(t_\ell)$ is in $\mathcal{N}$ or S_k , P_k cannot adjust the control $u_k^{\delta, \ell}(\cdot)$ he will use on the time interval $[t_\ell, t_{\ell+1})$ to the control $u_k^{\delta, \ell}(\cdot)$ chosen by his opponent.

Given a δ -strategy pair $\gamma^\delta = (\gamma_1^\delta, \gamma_2^\delta)$, the game is played as follows (analogously to the discrete-time game discussed in Section 3).

1. The two players observe $x(0)=x^0$ and the initial value $r(0)=i^0$. If $i^0 \in S_1$ player 1 is the leader and he has to move first; if $i^0 \in S_2$ player 2 is the leader and moves first; if $i^0 \in \mathcal{N}$ the two players move simultaneously.
2. At any sampled time t_ℓ , if player k is the leader (i.e., $r(t_\ell) \in S_k$), or if there is no leader (i.e., if $r(t_\ell) \in \mathcal{N}$), then this player moves first by choosing his control $u_k^{\delta, \ell}(\cdot)$ according to the mapping $\hat{\gamma}_k^{\delta, \ell}(x(t_\ell), r(t_\ell))$. If $r(t_\ell) \in S_k$, then the other player, $P_{\bar{k}}$, is the follower and he chooses his control $u_{\bar{k}}^{\delta, \ell}(\cdot)$ according to the mapping $\hat{\gamma}_{\bar{k}}^{\delta, \ell}(x(t_\ell), r(t_\ell), u_k^{\delta, \ell}(\cdot))$.
3. Once the controls $(u_k^{\delta, \ell}(\cdot))_{k=1,2}$ have been determined, the system evolves from $(t_\ell, x(t_\ell), r(t_\ell))$ to $(t_{\ell+1}, x(t_{\ell+1}), r(t_{\ell+1}))$. Again the leadership is determined by the value of $i^{\ell+1}$ and the game is played as described above.

Associated with a δ -strategy pair $\gamma^\delta = (\gamma_1^\delta, \gamma_2^\delta)$ are thus defined two cost functions

$$J_k(t_\ell, \bar{x}, i; \gamma_1^\delta, \gamma_2^\delta) = E[q_k^{r(T)}(x(T)) | x(t_\ell) = \bar{x}, r(t_\ell) = i] \quad (4.14)$$

$k=1,2$

defined for each sampled time t_ℓ and each possible state $(x(t_\ell)=\bar{x}, r(t_\ell)=i)$.

For each triple (t_ℓ, ξ, i) the costs (4.14) define a game in normal form. This class of games on $[0, T] \times X \times I$ will be called the $G(\delta)$ -game associated with the dynamical system (4.1)-(4.7). A $G(\delta)$ -game has exactly the same structure as the multistage game introduced in Section 3. Such a game is thus defined for each $\delta \in \Delta$ where $\Delta \triangleq \{T, \frac{T}{2}, \frac{T}{3}, \dots\}$. A strategy for player k is defined as a sequence

$$\gamma_k = \{\gamma_k^\delta\}_{\substack{\delta \rightarrow 0 \\ \delta \in \Delta}} \quad (4.15)$$

Furthermore, a strategy pair $\gamma = (\gamma_1, \gamma_2)$ is playable on $[0, T] \times X \times I$ if the limit (4.16) exists for each $k=1, 2, t_\ell \in [0, T], \xi \in X, i \in I$

$$\lim_{\substack{\delta \rightarrow 0 \\ \delta \in \Delta \\ t_\ell \rightarrow t}} J_k(t_\ell, \xi, i; \gamma_1^\delta, \gamma_2^\delta) \triangleq J_k(t, \xi, i; \gamma_1, \gamma_2) \quad (4.16)$$

Definition 4.1: The Differential Game associated with the dynamical system (4.1)-(4.7) is the family $G = \{G(\delta)\}$ of all $G(\delta)$ -games having cost functions (4.14), with $\delta \in \Delta$. □

We now introduce the concept of a "pure-feedback strategy" for such a game. Towards this end, let there be given two functions

$$\tilde{\gamma}_k(\cdot) : [0, T] \times X \times I \times U_k^- \rightarrow U_k \quad k=1, 2$$

such that

$$i \in \mathcal{N} \cup S_k \Rightarrow \tilde{\gamma}_k(t, x, i, u_k^-) \equiv \tilde{\gamma}(t, x, i), \quad (4.17)$$

and which are continuous in t , bounded, continuously differentiable in x and u_k , with bounded derivatives. The relation (4.17) must be linked to (4.13) as it expresses a similar rule: when i is such that P_k is not the follower, the function $\tilde{\gamma}_k(\cdot)$ does not depend on u_k .

It is possible to associate a control law $u \in \mathcal{U}_1 \times \mathcal{U}_2$ with the pair $(\tilde{\gamma}_1(\cdot), \tilde{\gamma}_2(\cdot))$ by defining

$$\left. \begin{aligned} u_k^i(t, x) &\triangleq \tilde{\gamma}_k(t, x, i) \text{ if } i \in \mathcal{N} \cup S_k \\ u_k^i(t, x) &\triangleq \tilde{\gamma}_k(t, x, i, \tilde{\gamma}_{\bar{k}}(t, x, i)) \text{ if } i \in S_{\bar{k}} \end{aligned} \right\} \quad (4.18)$$

We can also associate a whole family of δ -strategies with the pair $(\tilde{\gamma}_1(\cdot), \tilde{\gamma}_2(\cdot))$ by proceeding as follows:

$$\left. \begin{aligned} \forall \delta \in \Delta, \delta = \frac{T}{N}, \forall l \in \{0, 1, \dots, N-1\} \\ \gamma_k^{\delta, l}(x(t_l), r(t_l), u_k^{\delta, l}(\cdot)) = u_k^{\delta, l}(\cdot) \\ \text{with} \\ u_k^{\delta, l}(t) = \tilde{\gamma}_k(t, x(t_l), r(t_l), u_k^{\delta, l}(t)) \quad , \quad t \in [t_l, t_{l+1}) \\ k=1, 2 \quad , \quad \bar{k}=1, 2 \quad , \quad k \neq \bar{k}. \end{aligned} \right\} \quad (4.19)$$

The controls in (4.19) are well defined since the maps $\tilde{\gamma}_1(\cdot)$ and $\tilde{\gamma}_2(\cdot)$ satisfy (4.17).

Therefore, a strategy pair $\gamma = (\gamma_1, \gamma_2)$ for the differential game G can be associated with $(\tilde{\gamma}_1(\cdot), \tilde{\gamma}_2(\cdot))$. Thus, we pose

Definition 4.2: A strategy pair $\gamma = (\gamma_1, \gamma_2)$ is called a pure-feedback strategy for G if there exist two functions $\tilde{\gamma}_1(\cdot)$ and $\tilde{\gamma}_2(\cdot)$ satisfying (4.17) and defining through (4.18) an admissible control law, such that $\gamma = (\gamma_1, \gamma_2)$ can be associated with $(\tilde{\gamma}_1(\cdot), \tilde{\gamma}_2(\cdot))$. $\square$

We are now in a position to introduce the concepts of weak and strong equilibria (as counterparts of those introduced in Sections 2 and 3) for the differential game formulated above. We first have

Definition 4.3: A pure-feedback strategy pair $\gamma^* = (\gamma_1^*, \gamma_2^*)$ constitutes a weak equilibrium over $[0, T] \times X \times I$, if it is playable on $[0, T] \times X \times I$ and if, for every γ_k such that $\pi_k(\gamma_k, \gamma_k^*)$ is pure-feedback and playable on $[0, T] \times X \times I$, and for every $\tau \in [0, T]$, $\xi \in X$, $i \in I$, the following holds:

$$J_k(\tau, \xi, i; \gamma_1^*, \gamma_2^*) \leq J_k(\tau, \xi, i; \pi_k(\gamma_k, \gamma_k^*)) \quad , \quad k=1,2. \quad (4.20)$$

$\square$

As noted in Section 3, the class of weak equilibria is very rich indeed, since it involves "informational nonuniqueness." A stronger concept that is free of informational nonuniqueness is that of a strong equilibrium which was introduced in Section 3 for discrete-time (multistage) dynamic games. In order to introduce a similar concept in a differential game we have to use a limiting process as follows.

Assume that, at time $t \in [0, T]$, the value of $r(t)$ is observed to be in S_k , implying that P_k is the leader. Given the state $(x(t), r(t))$ and a weak equilibrium pair $\gamma^* = (\gamma_1^*, \gamma_2^*)$, we define an ϵ -deviation with value $\bar{u}_k$, for the leader, as a pure-feedback strategy $\gamma_{k; \epsilon, \bar{u}_k}^D$ such that $\pi_k(\gamma_{k; \epsilon, \bar{u}_k}^D, \gamma_k^*)$ is playable on $[0, T] \times X \times I$ and satisfies

$$\forall \tau > t + \varepsilon, \quad \forall \xi \in X, \quad \forall i \in I$$

$$J_k(\tau, \xi, i; \pi_k^D(\gamma_k; \varepsilon, \bar{u}_k, \gamma_k^*)) = J_k(\tau, \xi, i; \gamma_1^*, \gamma_2^*) \quad , \quad k'=1,2 \quad (4.21)$$

and

$$\forall \tau \in [t, t+\varepsilon], \text{ if } r(s) \equiv r(t) \text{ for } s \in [t, \tau], \text{ then } u_k(\tau) = \bar{u}_k \quad (4.22)$$

We will call an ε -reaction with value $\bar{u}_k$, for the follower, a pure-feedback strategy $\gamma_k^R; \varepsilon, \bar{u}_k$ such that

$$\pi_k^D(\gamma_k; \varepsilon, \bar{u}_k, \gamma_k^R; \varepsilon, \bar{u}_k) \text{ is playable on } [0, T] \times X \times I$$

and satisfies

$$\forall \tau > t + \varepsilon, \quad \forall \xi \in X, \quad \forall i \in I$$

$$J_k(\tau, \xi, i; \pi_k^D(\gamma_k; \varepsilon, \bar{u}_k, \gamma_k^R; \varepsilon, \bar{u}_k)) = J_k(\tau, \xi, i; \gamma_1^*, \gamma_2^*), \quad (4.23)$$

and

$$\forall t \in [t, t+\varepsilon], \text{ if } r(s) \equiv r(t) \text{ for } s \in [t, \tau], \text{ then } u_k(t) \equiv \bar{u}_k \quad . \quad (4.24)$$

Remark 4.1: The notions of ε -deviation and ε -reaction are defined locally.

They correspond to a temporary perturbation of the equilibrium strategy pair $\gamma^* = (\gamma_1^*, \gamma_2^*)$. For a length of time less than ε the leader deviates from his equilibrium strategy, by playing $\bar{u}_k$. On the same time interval the follower responds by playing $\bar{u}_k$. After time $t+\varepsilon$, or earlier if there is a jump in $r(\cdot)$, the original equilibrium strategy resumes. $\square$

We can now introduce

Definition 4.4: A pure-feedback strategy pair $\gamma^* = (\gamma_1^*, \gamma_2^*)$ is a strong equilibrium for the differential game G if

- (i) it is a weak equilibrium,
- (ii) for $k=1,2$, for any $(t, \xi, i) \in [0, T] \times X \times S_k$, $\bar{u}_k \in U_k$ and $\bar{u}_k^- \in U_k^-$, there exists $\varepsilon' > 0$ such that for all ε , $0 < \varepsilon < \varepsilon'$, the ε -deviation with value $\bar{u}_k$ and the ε -reaction with value $\bar{u}_k^-$ are well-defined pure-feedback strategies, and the following holds:

$$J_k^D(t, \xi, i; \pi_k(\gamma_k^D; \varepsilon, \bar{u}_k, \gamma_k^*)) \leq J_k^R(t, \xi, i; \pi_k(\gamma_k^D; \varepsilon, \bar{u}_k, \gamma_k^R; \varepsilon, \bar{u}_k^-)) + o(\varepsilon) \quad (4.25)$$

with

$$\lim_{\varepsilon \rightarrow 0} \frac{o(\varepsilon)}{\varepsilon} = 0$$

Remark 4.2: In a strong equilibrium, the strategy γ_k^* is the best response, locally, by P_k , when he is the follower, to a temporary deviation from the equilibrium by the leader (P_k). □

Remark 4.3: It should be noted that the concept of "strong equilibrium" (as well as that of weak equilibrium) is a limiting property of a G game, since a strong equilibrium strategy pair $\gamma^* = (\gamma_1^*, \gamma_2^*)$ does not necessarily provide a strong equilibrium to every $G(\delta)$ game, for $\delta > 0$, in the sense of §3.2. But it turns out that this is a more convenient and versatile definition for verifying the validity of the results to be given in Thms. 4.1 and 4.2 in the sequel.

4.3 Hamilton-Jacobi Equations

We are now in a position to obtain the Hamilton Jacobi equations associated with the weak and strong equilibrium solutions of the differential game G . We first make two additional assumptions concerning the existence and admissibility of solutions.

A 4. The differential game G admits at least one strong equilibrium in the class of pure-feedback strategies. $\square$

A 5. For $k=1,2$, given any $(t, \xi, i) \in [0, T] \times X \times I$, and any $\bar{u}_k \in U_k$, there exists $h > 0$, $\epsilon > 0$, and a playable pure-feedback strategy pair $\pi_k(\gamma_k, \gamma_k^*)$ which gives rise to an admissible control law $u \in \mathcal{U}_1 \times \mathcal{U}_2$, such that

$$u_k^i(s, x) \equiv \bar{u}_k(s) \text{ for } s \in [t, t+h], \quad k=1,2,$$

for all x in an ϵ -neighborhood of ξ . $\square$

Theorem 4.1: Assume that A1-A5 hold true. A necessary and sufficient condition for a pure-feedback strategy pair $(\tilde{\gamma}_1^*(\cdot), \tilde{\gamma}_2^*(\cdot))$ to provide a weak equilibrium for the differential game G is that for each $i \in I$, $k \in \{1,2\}$, the cost-to-go functions $V_k^i(t, x)$ satisfy the following partial differential equations

$$\begin{aligned} \min_{u_k \in U_k} \{ & \frac{\partial}{\partial t} V_k^i(t, x) + \frac{\partial}{\partial x} V_k^i(t, x) f^i(t, x, \pi_k(u_k, \tilde{\gamma}_k^*(t, x, i, u_k))) \\ & + \sum_{j \in I} \lambda_{ij}(t, x) V_k^j(t, x) \} = \end{aligned}$$

$$\begin{aligned}
& \frac{\partial}{\partial t} v_k^i(t, x) + \frac{\partial}{\partial x} v_k^i(t, x) f^i(t, x, u_1^{i*}(t, x), \tilde{u}_2^{i*}(t, x)) \\
& + \sum_{j \in I} \lambda_{ij}(t, x) v_k^j(t, x) = 0,
\end{aligned} \tag{4.26}$$

where $k=1,2$, $\bar{k}=1,2$, $k \neq \bar{k}$, and $u^* \in \mathcal{U}_1 \times \mathcal{U}_2$ is the feedback control law generated by $(\tilde{\gamma}_1^*(\cdot), \tilde{\gamma}_2^*(\cdot))$. The boundary conditions are

$$v_k^i(T, x) = q_k^i(x), \quad i \in I, \quad x \in X, \quad k=1,2. \tag{4.27}$$

Proof of Necessity: Let u_k be a point in U_k and u be the control law defined by

$$\left\{ \begin{array}{l} u_k^i(s, x') = u_k \\ u_k^i(s, x') = \gamma_k^*(s, x', i, u_k) \end{array} \right\} \quad \begin{array}{l} \text{if } t \leq s \leq t+h \text{ and } x' \text{ is in an} \\ \epsilon\text{-neighborhood of } x \end{array}$$

$$u_l^i(s, x') = u_l^{i*}(s, x'), \quad l=1,2, \text{ otherwise.}$$

The control u is admissible according to A5. Furthermore, it is the control law associated with a playable strategy pair $\pi_k(\gamma_k, \gamma_k^*)$ which is also defined by two mappings

$$\left\{ \begin{array}{l} \tilde{\gamma}_k(s, x', i, u_k) \equiv u_k \\ \tilde{\gamma}_k(v, \xi, j, u_k) = \tilde{\gamma}_k^*(v, \xi, j, u_k) \\ \tilde{\gamma}_k(v, \xi, j, u_k) = \tilde{\gamma}_k^*(v, \xi, j, u_k) \end{array} \right\} \quad \begin{array}{l} \text{if } t \leq s \leq t+h, x' \text{ is in an} \\ \epsilon\text{-neighborhood of } x \\ \text{elsewhere} \\ \text{everywhere.} \end{array}$$

Let $\tilde{V}_k^i(t, x)$ denote the expected cost-to-go corresponding to u . Because of the equilibrium property, $\pi_k(\gamma_k, \gamma_k^*)$ is such that

$$J_k(\pi_k(\gamma_k, \gamma_k^*)) \geq J_k(\gamma_1^*, \gamma_2^*)$$

i.e.

$$\tilde{V}_k^i(t, x) \geq V_k^i(t, x) \quad .$$

By an argument similar to the one used by Rishel in the proof of Theorem 4 of [7], we obtain, when $h \rightarrow 0$,

$$\begin{aligned} 0 \leq \frac{\partial}{\partial t} V_k^i(t, x) + \frac{\partial}{\partial x} V_k^i(t, x) f^i(t, x, \pi_k(u_k, \gamma_k^*(t, x, i, u_k))) \\ + \sum_{j \in I} \lambda_{ij}(t, x) V_k^j(t, x) \quad . \end{aligned} \quad (4.28)$$

Combining (4.28) with (4.10) of Lemma 4.1 gives (4.25).

Proof of Sufficiency: Let $u_k^i(t, x)$ be a control law for player k and define $u_k^i(t, x) = \gamma_k^*(t, x, i, u_k^i(t, x))$. If the pair $\pi_k(u_k^i(t, x), u_k^i(t, x))$ constitutes an admissible control law $u \in \mathcal{U}_1 \times \mathcal{U}_2$, then it defines an expected cost-to-go $\tilde{V}_k^i(t, x)$ which satisfies the inequality

$$\tilde{V}_k^i(t, x) \geq V_k^i(t, x).$$

The proof of this inequality is a direct consequence of Lemma 4.2, as in Rishel's Theorem 4 [7]. □

Theorem 4.2: Assume that A1-A5 hold true. A necessary and sufficient condition for a pure-feedback strategy pair $(\tilde{\gamma}_1^*(\cdot), \tilde{\gamma}_2^*(\cdot))$ to provide a strong equilibrium is that for each $i \in I$, $k \in \{1, 2\}$, the cost-to-go functions $V_k^i(t, x)$ satisfy the set of partial differential equations (4.26), together with the boundary condition (4.27), and that the following holds:

$$\forall (t, x) \in [0, T] \times X, \forall i \in S_k, \forall u_k \in U_k, \forall \bar{u}_k \in \bar{U}_k$$

$$\frac{\partial}{\partial x} V_k^i(t, x) f^i(t, x, \gamma_k(u_k, \tilde{\gamma}_k^*(t, x, i, u_k))) \quad (4.29)$$

$$\leq \frac{\partial}{\partial x} V_k^i(t, x) f^i(t, x, \pi_k(u_k, \bar{u}_k))$$

$$k=1, 2, \quad \bar{k}=1, 2, \quad \bar{k} \neq k.$$

Proof of Necessity: Consider, at a point $(t, x) \in [0, T] \times X$, and for $i \in S_k$, an ε -deviation with value $\bar{u}_k$ for player k and an ε -reaction with value $\bar{u}_{\bar{k}}$ for player $\bar{k}$.

Let $\bar{x}^D(t+s; x, i)$ and $\bar{x}^{DR}(t+s; x, i)$, $s \geq 0$, denote the trajectories emanating from (t, x) , with $r(t) \equiv i$, and generated, respectively, by $\pi_k^D(\gamma_k; \varepsilon, \bar{u}_k, \gamma_k^*)$ and $\pi_k^D(\gamma_k; \varepsilon, \bar{u}_k, \gamma_k^R; \varepsilon, \bar{u}_{\bar{k}})$.

According to Definition 4.4, we must have, for ε sufficiently small,

$$\begin{aligned}
& J_k^-(t, x, i; \pi_k(\gamma_k^D; \varepsilon, \bar{u}_k, \gamma_k^*), \bar{\gamma}_k^*) = V_k^i(t+\varepsilon, \bar{x}^D(t+\varepsilon; x, i)) \\
& + \int_t^{t+\varepsilon} \lambda_{ij}(t, \bar{x}^D(t+s; x, i)) V_k^i(t, \bar{x}^D(t+s; x, i)) ds \\
& \leq J_k^-(t, x, i; \pi_k(\gamma_k^D; \varepsilon, \bar{u}_k, \gamma_k^R; \varepsilon, \bar{u}_k), \bar{\gamma}_k^R) + o(\varepsilon) = \\
& V_k^i(t+\varepsilon, \bar{x}^{DR}(t+\varepsilon; x, i)) + \int_t^{t+\varepsilon} \lambda_{ij}(t, \bar{x}^{DR}(t+s; x, i)) \\
& V_k^i(t, \bar{x}^{DR}(t+s; x, i)) ds + o(\varepsilon).
\end{aligned} \tag{4.30}$$

Using the differentiability of $V_k^i(\cdot)$ at (t, x) and taking the limit of (4.30) when ε tends to zero we get (4.29).

Proof of Sufficiency: We have

$$\begin{aligned}
& J_k^-(t, x, i; \pi_k(\gamma_k^D; \varepsilon, \bar{u}_k, \gamma_k^*), \bar{\gamma}_k^*) - \\
& J_k^-(t, x, i; \pi_k(\gamma_k^D; \varepsilon, \bar{u}_k, \gamma_k^R; \varepsilon, \bar{u}_k), \bar{\gamma}_k^R) = \\
& V_k^i(t+\varepsilon, \bar{x}^D(t+\varepsilon; x, i)) - V_k^i(t+\varepsilon, \bar{x}^{DR}(t+\varepsilon; x, i)) \\
& + \int_t^{t+\varepsilon} \lambda_{ij}(t, \bar{x}^D(t+s; x, i)) V_k^i(t, \bar{x}^D(t+s; x, i)) \\
& - \lambda_{ij}(t, \bar{x}^{DR}(t+s; x, i)) V_k^i(t, \bar{x}^{DR}(t+s; x, i)) \\
& = \frac{\partial}{\partial x} V_k^i(t, x) (f^i(t, x, \pi_k(u_k, \gamma_k^*(t, x, i, u_k))) \\
& - f^i(t, x, \pi_k(u_k, \bar{u}_k))) + o(\varepsilon).
\end{aligned}$$

Therefore, if (4.29) holds true, condition (4.24) is also satisfied.

□

5. Some Special Cases

In this section we study the dynamic programming equation associated with the strong equilibrium solution (cf. Thm. 4.2) in some detail and for some special cases. For future reference, we first rewrite the dynamic programming equation satisfied by the strong equilibrium solution $\{u_1^*(t,x), u_2^*(t,x)\}$, for the case when the cost functions are given as

$$J_k = E\{q_k^i(x(T)) + \int_0^T g_k^i(t,x,u_1,u_2)dt\} \quad : \quad (5.1)$$

$$-\frac{\partial}{\partial t} V_k^i(t,x) - \sum_{j \in I} \lambda_{ij}(t,x) V_k^j(t,x) =$$

$$\left\{ \begin{array}{l} \min_{u_k} [\frac{\partial}{\partial x} V_k^i(t,x) \cdot f^i(t,x,\pi_k(u_k, u_k^{i*}(t,x))) + g_k^i(t,x,\pi_k(u_k, u_k^{i*}(t,x)))], \\ i \in \mathcal{N} ; \end{array} \right. \quad (5.2)$$

$$= \left\{ \begin{array}{l} \min_{u_k} [\frac{\partial}{\partial x} V_k^i(t,x) \cdot f^i(t,x,\pi_k(u_k, T_k^i(t,x,u_k, \frac{\partial}{\partial x} V_k^i))) \\ + g_k^i(t,x,\pi_k(u_k, T_k^i(t,x,u_k, \frac{\partial}{\partial x} V_k^i)))], i \in S_k ; \end{array} \right. \quad (5.3)$$

$$\left\{ \begin{array}{l} [\frac{\partial}{\partial x} V_k^i(t,x) \cdot f^i(t,x,\pi_k(u_k^{i*}, T_k^i(t,x,u_k^{i*}, \frac{\partial}{\partial x} V_k^i))) \\ + g_k^i(t,x,\pi_k(u_k^{i*}, T_k^i(t,x,u_k^{i*}, \frac{\partial}{\partial x} V_k^i)))], i \in S_k^- ; \end{array} \right. \quad (5.4)$$

$$V_k^i(T,x) = q_k^i(x) \quad , \quad i \in I \quad , \quad x \in X \quad , \quad k=1,2 \quad , \quad (5.5)$$

where

$$T_k^i(t, x, u_k, \frac{\partial}{\partial x} V_k^i) \triangleq \arg \min_{u_k^-} [\frac{\partial}{\partial x} V_k^i(t, x) \cdot f^i(t, x, \pi(u_k, u_k^-)) + g_k^i(t, x, \pi(u_k, u_k^-))] \quad (5.6)$$

$$u_k^{i*} = \begin{cases} \arg \text{ of (5.2) } , & i \in \mathcal{I} \\ \arg \text{ of (5.3) } , & i \in S_k \\ T_k^i(t, x, u_k^{i*}, \frac{\partial}{\partial x} V_k^i) , & i \in S_k^- \end{cases} \quad (5.7)$$

Now, in the two subsections to follow, we consider two special cases, viz. the case when the mode of play is fixed (there is no chance variable) and the case of linear-quadratic differential games with the mode of play determined by a Markov jump process.

5.1. Deterministic Differential Games with a Fixed Mode of Play

Assuming that there are no chance moves, we now differentiate between two prototypes, viz. the case when one of the players, say P_1 , is always the leader (i.e., he has informational advantage, which is though only incremental), and the case of symmetric mode of play, which corresponds to the Nash equilibrium solution. For the latter, $S_1 = S_2 = \emptyset$, $\mathcal{I} = \{1, 2\}$, $\lambda_{ij} = 0 \forall ij$, and the dynamic programming equations associated with a strong equilibrium solution $u^* = (u_1^*(t, x), u_2^*(t, x))$ are easily obtainable from (5.2) to be

$$\begin{aligned} \frac{\partial}{\partial t} v_k(t, x) = - \min_{u_k} [& \frac{\partial}{\partial x} v_k(t, x) \cdot f(t, x, \pi_k(u_k, u_k^*(t, x))) \\ & + g_k(t, x, \pi_k(u_k, u_k^*(t, x)))], \quad k=1,2 \end{aligned} \quad (5.8)$$

$$v_k(T, x) = q_k(x) \quad (5.9)$$

where $u_k^*(t, x) = \arg \text{ RHS of (5.8) }.$

These relations characterize the so-called "feedback Nash equilibrium solution" [3,9], and in this case the concepts of weak and strong equilibria coincide.

In the former case, however, $\eta = \emptyset$, $S_1 = \{1\}$, $S_2 = \emptyset$, $\lambda_{ij} = 0 \forall ij$, and the dynamic programming equations associated with a strong equilibrium solution $u^* = (u_1^*(t, x), u_2^*(t, x))$ read [from (5.3)-(5.7)]:

$$\begin{aligned} \frac{\partial}{\partial t} v_1(t, x) = - \min_{u_1} [& \frac{\partial}{\partial x} v_1(t, x) \cdot f(t, x, u_1, T_2(t, x; u_1, \frac{\partial v_2}{\partial x})) \\ & + g_1(t, x, u_1, T_2(t, x; u_1, \frac{\partial v_2}{\partial x}))] \end{aligned} \quad (5.10)$$

$$\begin{aligned} \frac{\partial}{\partial t} v_2(t, x) = - \frac{\partial}{\partial x} v_2(t, x) \cdot f(t, x, u_1^*(t, x), u_2^*(t, x)) \\ - g_2(t, x, u_1^*(t, x), u_2^*(t, x)) \end{aligned} \quad (5.11)$$

$$v_k(T, x) = q_k(x) \quad (5.12)$$

where

$$T_2(t, x; u_1, \frac{\partial V_2}{\partial x}) = \arg \min_{u_2} [\frac{\partial V_2}{\partial x} \cdot f(t, x, u_1, u_2) + g_2(t, x, u_1, u_2)] \quad (5.13)$$

$$u_1^* = \arg \text{ RHS of (5.10)} \quad (5.14a)$$

$$u_2^* = T_2(t, x; u_1^*(t, x), \frac{\partial}{\partial x} V_1(t, x)) \quad (5.14b)$$

The solution (u_1^*, u_2^*) satisfying the above relations may be called the continuous-time feedback Stackelberg solution, because it is the natural counterpart of the discrete-time feedback Stackelberg solution (well-established in the literature for discrete-time dynamic games [2,3]) in the continuous-time domain, that is for differential games. Here, the leader has only an incremental informational advantage over the follower, at each instant of time, and he cannot announce his strategy ahead of time as in the case of the standard Stackelberg problem. Furthermore, it should be noted that weak and strong equilibria do not necessarily coincide here, and there exist in general infinitely many weak equilibria because of "informationally nonuniqueness".

Since the asymmetry in the roles of the players in a continuous-time feedback Stackelberg solution is only incremental, one may be led to the conclusion that the feedback Stackelberg solution should coincide with (or be very close to) the feedback Nash solution. Such an implication is also evident in the analysis of Friedman in Chapter 8 of his book [10, p. 290], where he defines the "stable equilibrium value" of an N-person deterministic differential game as (using our terminology, naturally extended to N-person games) the limit (as $\delta > 0$) of the N-tuple cost values

associated with strong equilibria of $G(\delta)$ differential games, independent of the nature of the asymmetry in the roles of the players (i.e., independent of the order in which the players announce their decisions stagewise).

Friedman also shows that for linear-quadratic differential games such a value exists whenever T is sufficiently small, which is the "Nash" value associated with the feedback Nash equilibrium solution [10, Thm.8.7.1].

However, such a result does not hold for more general classes of games, as can be observed by comparing the conditions (5.8) and (5.10)-(5.13); in particular, if T_2 defined by (5.13) is functionally dependent on u_1 [which is not the case in the strictly linear-quadratic problem considered by Friedman], the solutions of (5.8)-(5.9) and (5.10)-(5.12) will in general be different[†], which implies that the strong equilibrium solution under an asymmetric mode of play (i.e., the feedback Stackelberg solution) is in general different from the strong equilibrium solution under a symmetric mode of play (i.e., the feedback Nash solution). To illustrate this point, let us consider again the linear-quadratic structure, but somewhat more general than that of [10, Thm.8.7.1]:

$$f = Ax - B_1 u_1 - B_2 u_2 \quad (5.15a)$$

$$g_k = \frac{1}{2} (x' Q_k x + u_k' u_k - 2u_k' R_{kk} u_k) \quad , \quad k=1,2 \quad (5.15b)$$

$$q_k = \frac{1}{2} x' C_k x \quad ; \quad Q_k \geq 0, C_k \geq 0 \quad (5.15c)$$

[†] Assuming, of course, that neither $J_1 \equiv -J_2$, nor $J_1 \equiv J_2$, i.e., the underlying problem is not a team or a zero-sum game.

where capital letters denote matrices of appropriate dimensions.

Then,

$$\begin{aligned}
 T_2 = \arg \min_{u_2} \{ & \frac{\partial V_2}{\partial x} [Ax - B_1 u_1 - B_2 u_2] + \frac{1}{2} x' Q_2 x + \frac{1}{2} u_2' u_2 \\
 & - u_2' R_{21} u_1 \} \\
 = & R_{21} u_1 + \left(\frac{\partial V_2}{\partial x} B_2 \right)' \quad . \quad (5.16)
 \end{aligned}$$

Substituting this into the RHS of (5.10) and evaluating the minimizing u_1 we obtain

$$\begin{aligned}
 u_1^* = & (I - R_{12} R_{21} - R_{21}' R_{12}')^{-1} [R_{12} B_2' \left(\frac{\partial V_2}{\partial x} \right)' + (B_1' + R_{21}' B_2') \left(\frac{\partial V_2}{\partial x} \right)'] \\
 \underline{u} = & K_1 \left(\frac{\partial V_1}{\partial x} \right)' - K_2 \left(\frac{\partial V_2}{\partial x} \right)' \quad , \quad (5.17)
 \end{aligned}$$

assuming that the required inverse exists. The corresponding u_2^* is then [from (5.16) and (5.17)]

$$\begin{aligned}
 u_2^* = & - R_{21} K_1 \left(\frac{\partial V_1}{\partial x} \right)' + (B_2' - R_{21} K_2) \left(\frac{\partial V_2}{\partial x} \right)' \\
 \underline{u} = & L_1 \left(\frac{\partial V_1}{\partial x} \right)' - L_2 \left(\frac{\partial V_2}{\partial x} \right)' \quad . \quad (5.18)
 \end{aligned}$$

Resubstituting (5.17) and (5.18) into (5.10) and (5.11), we finally obtain the coupled set of partial differential equations:

$$\begin{aligned}
\frac{\partial V_1}{\partial t} &= - \frac{\partial V_1}{\partial x} [Ax + \bar{B}_1 \left(\frac{\partial V_1}{\partial x} \right)' + \bar{B}_2 \left(\frac{\partial V_2}{\partial x} \right)'] - \frac{1}{2} x' Q_1 x \\
&\quad - \frac{1}{2} [K_1 \left(\frac{\partial V_1}{\partial x} \right)' + K_2 \left(\frac{\partial V_2}{\partial x} \right)'] [K_1 \frac{\partial V_1}{\partial x} + K_2 \frac{\partial V_2}{\partial x}] \\
&\quad + [K_1 \left(\frac{\partial V_1}{\partial x} \right)' + K_2 \left(\frac{\partial V_2}{\partial x} \right)']' R_{12} [L_1 \left(\frac{\partial V_1}{\partial x} \right)' + L_2 \left(\frac{\partial V_2}{\partial x} \right)'] \\
\frac{\partial V_2}{\partial t} &= - \frac{\partial V_2}{\partial x} [Ax + \bar{B}_1 \left(\frac{\partial V_1}{\partial x} \right)' + \bar{B}_2 \left(\frac{\partial V_2}{\partial x} \right)'] - \frac{1}{2} x' Q_2 x \\
&\quad - \frac{1}{2} [L_1 \left(\frac{\partial V_1}{\partial x} \right)' + L_2 \left(\frac{\partial V_2}{\partial x} \right)']' [L_1 \left(\frac{\partial V_1}{\partial x} \right)' + L_2 \left(\frac{\partial V_2}{\partial x} \right)'] \\
&\quad + [L_1 \left(\frac{\partial V_1}{\partial x} \right)' + L_2 \left(\frac{\partial V_2}{\partial x} \right)'] R_{21} [K_1 \left(\frac{\partial V_1}{\partial x} \right)' + K_2 \left(\frac{\partial V_2}{\partial x} \right)']
\end{aligned} \tag{5.19}$$

$$V_k(T, x) = \frac{1}{2} x' C_k x$$

$$\bar{B}_1 \triangleq B_1 K_1 + B_2 L_1 \quad ; \quad \bar{B}_2 \triangleq B_1 K_2 + B_2 L_2 \tag{5.20}$$

For the feedback Nash equilibrium solution, however, the relevant set of PDE's can be derived using (5.8), and found to be in the same form as (5.19) but with K_1 , K_2 , L_1 , L_2 , $\bar{B}_1$ and $\bar{B}_2$, respectively replaced by the "hat'ted" quantities

$$\begin{aligned}
\hat{K}_1 &= - (I - R_{12} R_{21})^{-1} B_1' \quad ; \quad \hat{K}_2 = - (I - R_{12} R_{21})^{-1} R_{12} B_2' \\
\hat{L}_1 &= - (I - R_{21} R_{12})^{-1} R_{21} B_1' \quad ; \quad \hat{L}_2 = - (I - R_{21} R_{12})^{-1} B_2' \\
\hat{\bar{B}}_1 &= B_1 \hat{K}_1 + B_2 \hat{L}_1 \quad ; \quad \hat{\bar{B}}_2 = B_1 \hat{K}_2 + B_2 \hat{L}_2
\end{aligned}$$

Note that if the cross terms in the cost functions are absent (i.e., $R_{12}=0$, $R_{21}=0$), we have the simple relations

$$K_1 = \hat{K}_1 = -B_1' \quad ; \quad K_2 = \hat{K}_2 = 0$$

$$L_1 = \hat{L}_1 = 0 \quad ; \quad L_2 = \hat{L}_2 = -B_2'$$

which imply that the two sets of PDE's become identical, thus admitting the same set of solutions. This then corroborates Friedman's result mentioned earlier. If the cross terms are not absent, however, the two sets of PDE's are intrinsically different, and admit different sets of solutions. Hence, even in linear quadratic games with generalized quadratic cost functionals, the feedback Stackelberg and Nash solutions may be different (or, in Friedman's terminology, a stable equilibrium value may not exist, no matter how small T is).

We now conclude this subsection by reporting a result on the existence and structure of the feedback Stackelberg solution of the linear-quadratic differential game described by (5.15).

Proposition 5.1. If T is sufficiently small and the matrix inverse in (5.17) exists, the linear-quadratic differential game described by (5.15), and with P_1 as the leader, admits a feedback Stackelberg solution given by

$$u_1^*(t, x) = - (K_1 P_1 + K_2 P_2) x \quad (5.21a)$$

$$u_2^*(t, x) = - (L_1 P_1 + L_2 P_2) x \quad , \quad (5.21b)$$

where $\{P_1(t), P_2(t)\}$ are symmetric solutions of the coupled set of Riccati equations

$$\begin{aligned} \dot{P}_1 = & -P_1 L - L' P_1 - Q_1 - (K_1 P_1 + K_2 P_2)' (K_1 P_1 + K_2 P_2) \\ & + (L_1 P_1 + L_2 P_2)' R_{12}' (K_1 P_1 + K_2 P_2) \end{aligned} \quad (5.22a)$$

$$+ (K_1 P_1 + K_2 P_2)' R_{12} (L_1 P_1 + L_2 P_2) \quad ; \quad P_1(T) = C_1$$

$$\begin{aligned} \dot{P}_2 = & -P_2 L - L' P_2 - Q_2 - (L_1 P_1 + L_2 P_2)' (L_1 P_1 + L_2 P_2) \\ & + (L_1 P_1 + L_2 P_2)' R_{21}' (K_1 P_1 + K_2 P_2) \end{aligned} \quad (5.22b)$$

$$+ (K_1 P_1 + K_2 P_2)' R_{21} (L_1 P_1 + L_2 P_2) \quad ; \quad P_2(T) = C_2$$

$$L \triangleq A + \bar{B}_1 P_1 + \bar{B}_2 P_2 \quad (5.23)$$

Proof: This result follows by substituting $V_k = \frac{1}{2} x' P_k x$ into the PDE's (5.19) and observing that (5.22) imply satisfaction of (5.19) by such a quadratic cost-to-go function. Existence of a (unique) solution to (5.22) when T is sufficiently small follows from a standard property of ordinary differential equations with continuous right-hand-sides. $\square$

Remark 5.1: Proposition 5.1 has a natural counterpart in the context of feedback Nash equilibria, simply with K_k , L_k and $\bar{B}_k$ replaced by their corresponding "hat'ted" versions introduced earlier. The solution to the resulting set of Riccati equations will not be the same as the solution to (5.22), unless $R_{21} = 0$, $R_{12} = 0$. $\square$

5.2. Linear-Quadratic Differential Games with the Mode of Play Determined by Markov Jump Processes

Consider the class of deterministic differential games wherein the mode of play is determined by the output of a 3-state Markov jump process with constant parameters $\lambda_{ij}(t,x) = \lambda_{ij}$, with the three possible states corresponding to the three different modes of play: Stackelberg with P1 as leader ($i=1$), Stackelberg with P2 as leader ($i=2$), and Nash equilibrium ($i=3$). Hence, in terms of the terminology we have adopted for our general formulation, $I = \{1,2,3\}$, $S_1 = \{1\}$, $S_2 = \{2\}$, $\mathcal{N} = \{3\}$, and f^r , q_k^r and g_k^r are independent of r . Then, the related Hamilton-Jacobi equation that yields the strong equilibria is (5.2)-(5.7), with λ_{ij} independent of (t,x) and f^i , g_k^i , g_k^i not depending on i . We now study these equations in more detail for the special class of linear-quadratic differential games described by (5.15) but without the cross terms in control (i.e., $R_{12} = 0$, $R_{21} = 0$).

First, we evaluate

$$\begin{aligned} T_k^i(t,x,u_k, \frac{\partial V_k^i}{\partial x}) &= \arg \min_{u_k} \left[\frac{\partial}{\partial x} V_k^i \cdot (Ax - B_1 u_1 - B_2 u_2) \right. \\ &\quad \left. + \frac{1}{2} x' Q_k x + \frac{1}{2} u_k' u_k \right. \\ &\quad \left. - u_k' R_{kk} u_k \right] \\ &= R_{kk} u_k + B_k' \left(\frac{\partial}{\partial x} V_k^i \right)' \end{aligned}$$

and then substitute this into (5.2)-(5.5) to obtain, after performing the minimizations,

$$\begin{aligned}
& - \frac{\partial}{\partial t} V_k^i(t, x) - \sum_{j=1}^3 \lambda_{ij} V_k^j(t, x) \\
& = \frac{\partial V_k^i}{\partial x} \{ Ax + \bar{B}_1^i \left(\frac{\partial V_k^i}{\partial x} \right)' + \bar{B}_2^i \left(\frac{\partial V_k^i}{\partial x} \right)' \} + \frac{1}{2} x' Q_k x \\
& \quad + \frac{1}{2} \| N_{kk}^i \left(\frac{\partial V_k^i}{\partial x} \right)' + N_{kk}^i \left(\frac{\partial V_k^i}{\partial x} \right)' \|^2 \\
& \quad - [N_{kk}^i \left(\frac{\partial V_k^i}{\partial x} \right)' + N_{kk}^i \left(\frac{\partial V_k^i}{\partial x} \right)']' R_{kk} [N_{kk}^i \left(\frac{\partial V_k^i}{\partial x} \right)' + N_{kk}^i \left(\frac{\partial V_k^i}{\partial x} \right)']; \quad (5.24)
\end{aligned}$$

$$V_k^i(t, x) = \frac{1}{2} x' C_k x, \quad k=1, 2; \quad i=1, 2, 3, \quad (5.25)$$

where

$$N_{kk}^i = \begin{cases} - (I - R_{kk} \bar{R}_{kk} - R_{kk}' R_{kk}')^{-1} (B_k' + R_{kk}' \bar{B}_k') & i=k \\ R_{kk} \bar{N}_{kk}^k - B_k' & i=\bar{k} \\ - (I - R_{kk} \bar{R}_{kk})^{-1} B_k' & i=3 \end{cases} \quad (5.26a)$$

$$N_{kk}^i = \begin{cases} - (I - R_{kk} \bar{R}_{kk} - R_{kk}' R_{kk}')^{-1} R_{kk} \bar{B}_k' & ; \quad i=k \\ R_{kk} \bar{N}_{kk}^k & ; \quad i=\bar{k} \\ - (I - R_{kk} \bar{R}_{kk})^{-1} R_{kk} \bar{B}_k' & ; \quad i=3 \end{cases} \quad (5.26b)$$

$$\bar{B}_k^i = B_k N_{kk}^i + B_k \bar{N}_{kk}^i \quad (5.26c)$$

and $\|\cdot\|$ denotes the Euclidean norm.

Now, trying out a solution to (5.24) in the form

$$V_k^i(t, x) = \frac{1}{2} x' P_k^i(t) x \quad ; \quad P_k^i(T) = C_k$$

we obtain the following differential equations for the symmetric matrices

$P_k^i(\cdot)$, $k=1,2$; $i=1,2,3$:

$$\begin{aligned} \dot{P}_k^i = & - P_k^i L^i - L^{i'} P_k^i - \sum_{j=1}^3 \lambda_{ij} P_k^j - Q_k \\ & - (N_{kk}^i P_k^i + N_{kk}^{\bar{i} P_k^i})' (N_{kk}^i P_k^i + N_{kk}^{\bar{i} P_k^i}) \\ & + (N_{kk}^i P_k^i + N_{kk}^{\bar{i} P_k^i})' R_{kk} (N_{kk}^i P_k^i + N_{kk}^{\bar{i} P_k^i}) \\ & + (N_{kk}^i P_k^i + N_{kk}^{\bar{i} P_k^i})' R_{kk}' (N_{kk}^i P_k^i + N_{kk}^{\bar{i} P_k^i}) \quad ; \quad P_k^i(T) = C_k \end{aligned} \quad (5.27)$$

where

$$L^i = A + B_1^i P_1^i + B_2^i P_2^i \quad (5.28)$$

Assuming that a solution to (5.27) exists, the feedback control laws associated with the strong equilibrium solution are given by

$$u_k^{i*}(t, x) = - (N_{kk}^i P_k^i + N_{kk}^{\bar{i} P_k^i}) x \quad , \quad k=1,2 \quad (5.29)$$

which depend on the observed values of $x(t)$ and $r(t)$, denoted by x and i , respectively. The following proposition now summarizes the result.

Proposition 5.2: If the inverses in (5.26) exist and the coupled set of differential equations (5.27) admits a symmetric solution in the interval $[0, T]$, the linear quadratic differential game, wherein the modes of play are determined by the outcome of the Markov jump process $\{r(\cdot)\}$, admits a strong equilibrium solution given by (5.29). $\square$

Remark 5.2: If the cross terms in (5.15b) are absent, i.e., $R_{21} = 0$, $R_{12} = 0$, it follows from (5.26a) and (5.26b) that N_{kk}^i and $N_{kk}^{\bar{i}}$ are independent of i , which implies that the set of equations (5.27) are also independent of i . Hence, the solution (5.29) does not depend on the observed value i of the Markov jump process $\{r(\cdot)\}$. Furthermore, since P_k^j is independent of j and $\sum_{j=1}^3 \lambda_{ij} = 0$, (5.27) becomes the same equation set as (5.22) with $R_{12} = R_{21} = 0$. The implication then is that the strong equilibrium solution, for this special case, is identical with the feedback Nash equilibrium (or equivalently, the feedback Stackelberg equilibrium) solution. $\square$

6. Concluding Remarks

The objectives of this study have been two-fold: Firstly, to provide a general definition of an equilibrium solution for discrete-time dynamic games which would encompass both feedback Nash and feedback Stackelberg solutions and also be extendable to games in which both the underlying system dynamics and the mode of play are determined (nondeterministically) as the outcome of a finite state stochastic jump process; this has been accomplished in Sections 2 and 3 which also contain the optimality equations for such games. Secondly, to introduce a feedback Stackelberg equilibrium concept for (continuous-time) differential games with a fixed asymmetric mode of play, and to obtain the associated optimality conditions; this has been achieved in Sections 4 and 5 by formulating a general stochastic differential game with structural and modal uncertainties and by associating the (strong) equilibrium solution with the limiting solution of a sequence of discretized ($G(\delta)$) games. This has led to an indirect derivation of a pair of Hamilton-Jacobi equations, which characterizes the set of optimality conditions for the differential game.

An important by-product of our analysis is the observation that the feedback Nash solution is not in general the same as the feedback Stackelberg solution, in continuous-time dynamic games, unless the state equation and the cost functionals are of particular forms, as discussed in Section 5. Hence, the mode of play is a crucial factor in the characterization of equilibria in differential games, and if, for example, the Nash equilibrium is defined as the limit of the equilibrium solutions of a sequence of discretized (in time) games, the end result will very much depend on the information structure to be

adopted in the solution of these games [i.e., whether or not one player has informational advantage over the other in terms of observing his actions]. In other words, in nonzero-sum differential games the equilibrium value cannot in general be defined independently of the information structures of the sub-games -- quite contrary to what has been a common practice in zero-sum differential games.

In the context of zero-sum differential games, it is possible to define the (saddle-point) value in several different ways, as the limit of saddle-point equilibria of a sequence of discrete-time sub-games with various information structures [see, [10],[14]-[17] for such different definitions, the actual difference lying in the information allowed to the players in the discretized games]. Such different limiting procedures all lead to the same numerical "value" for the original differential game, even though the existence conditions (for the limit) are different under different schemes. Motivated by this result (which is an inherent property of the saddle-point value in zero-sum differential games), Friedman has attempted in [10, Chapter 8] to introduce the concept of "stable (Nash) equilibrium" in nonzero-sum differential games by relating it to the limit of equilibria of discretized games. In each such discretized game, Friedman has adopted a strictly hierarchical mode of play at each stage, with the players moving in a predetermined sequential manner, and he has defined the "value" as one which is attained in the limit (as the discretized games converge to the original continuous-time game) independently of the order in which the players act at each stage of the discretized games. Furthermore, he has shown that such a definition is meaningful in the case of a special class of linear-quadratic games, since

the feedback Nash value is then independent of the order of play. For more general types of nonzero-sum differential games, however, this will no longer be true, as can be concluded from our analyses and results of Sections 4 and 5. Our contention is that the equilibrium value will in general be dependent on the mode of play adopted for the discretized versions of the continuous-time nonzero-sum game, and hence the stable equilibrium, as defined by Friedman, will exist only in a very restricted class of problems.

Our indirect definition of "strong equilibrium" in Section 4, as well as the derivation of the associated Hamilton-Jacobi equation, have bypassed the seemingly difficult task of proving existence of a limit to the sequence of $G(\delta)$ games and identifying the solution of the Hamilton-Jacobi equation as the limit of the equilibrium values of these games. This is a challenging task that needs to be undertaken in the future in order to complete the theory presented here. An extension of our analysis to general M -player nonzero-sum games vested with a variety of modes of play, however, is rather routine, and one may expect to arrive at similar qualitative conclusions from such an extended analysis; we have chosen not to do it here in order not to bury the essential ideas in notational complexity. An application of the theory of Section 4 and some of the results of Section 5 to certain problems arising in economics can be found in [18].

7. Appendix I

In this appendix we formulate a 2N-stage stochastic dynamic game whose feedback Nash equilibrium solution coincides with the strong equilibrium solution of §3.2. In this regard, this result can be viewed as an extension of Proposition 2.2 to stochastic dynamic games with variable modes of play, as formulated in §3.1.

Consider the following state equation and cost functionals:

State Equation: $z(s+1) = F_s[z(s), \bar{u}_1(s), \bar{u}_2(s), w(s)]$; $z(0) = [x^0, r^0, o^1, o^1]'$

where

$$F_s[z, \bar{u}_1, \bar{u}_2] = \begin{cases} z(s) + [\sigma_{m_0}', \tau_k(\bar{u}_k(s), o_{m_k}')']' ; & s \text{ even} \\ & \bar{r}(s) \in S_k \\ z(s) ; & s \text{ even, } \bar{r}(k) \in \mathcal{N} \\ \bar{m}_0 = m_0 + 1, \\ [w'(s), o_{m_1}', o_{m_2}']' ; & s \text{ odd} \end{cases}$$

and for s odd,

$$\underline{P}(w(s) \in (dX, i) | z(s), \bar{u}_1(s), \bar{u}_2(s)) = Q[dX, i; \bar{y}(s), \bar{r}(s), \bar{u}_1(s), \bar{u}_2(s)];$$

$$\bar{r}(s) \in \mathcal{N}$$

$$Q[dX, i; \bar{y}(s), \bar{r}(s), \tau_k(\bar{u}_k(s-1), \bar{u}_k(s))];$$

$$\bar{r}(s) \in S_k$$

with $\bar{y}(s) = (I_{m_0}, 0_{m_0 \times (m_1+m_2+1)})z(s)$

$$\bar{r}(s) = (o_{m_0}', 1, o_{m_1}', o_{m_2}')z(s) \equiv \begin{cases} r(\frac{s}{2}) & ; \quad s \text{ even} \\ r(\frac{s-1}{2}) & ; \quad s \text{ odd.} \end{cases}$$

$$\hat{u}_k(s-1) = (0_{m_k \times m_0}, \pi_k(0_{m_k \times m_k}, I_{m_k}))z(s)$$

Note that

$$z(s) \triangleq \begin{cases} [x'(\frac{s}{2}), r'(\frac{s}{2}), o_{m_1}', o_{m_2}']', & s \text{ even} \\ [x'(\frac{s-1}{2}), r(\frac{s-1}{2}), \pi_k(u_k(\frac{s-1}{2}), o_{m_k}')]', & s \text{ odd, } r(\frac{s-1}{2}) \in S_k \\ [x'(\frac{s-1}{2}), r(\frac{s-1}{2}), o_{m_1}', o_{m_2}']', & s \text{ odd, } r(\frac{s-1}{2}) \in \bar{S}_k \end{cases}$$

and

$$\bar{u}_k(s) = \begin{cases} u_k(\frac{s}{2}) & s \text{ even, } r(\frac{s}{2}) \in S_k \\ u_k(\frac{s-1}{2}) & s \text{ odd, } r(\frac{s-1}{2}) \in \bar{S}_k \end{cases}$$

Cost Functionals: $k=1,2$

$$\tilde{J}_k = E \left[\tilde{q}_k[z(2N)] + \sum_{s=0}^{2N-1} g_{k,s}[z(s), \bar{u}_1(s), \bar{u}_2(s)] \right], \quad k=1,2,$$

$\begin{matrix} x(0)=x_0^0 \\ r(0)=r \end{matrix}$

where

$$\begin{aligned}
 \bar{q}_k[z(2N)] &= q_k[(I_{m_0}, 0_{m_0 \times (m_1+m_2+1)})z(2N)] \\
 \bar{g}_{k,s}[z, \bar{u}_1, \bar{u}_2] &= \begin{cases} g_{k, \frac{s-1}{2}}[(I_{m_0}, 0_{m_0 \times (m_1+m_2+1)})z(s), \bar{u}_1(s), \bar{u}_2(s)], & s \text{ odd,} \\ & \bar{r}(s) \in S_k \\ \\ g_{k, \frac{s-1}{2}}[(I_{m_0}, 0_{m_0 \times (m_1+m_2+1)})z(s), \pi_k(\hat{u}_k(s-1), \bar{u}_k(s))], & \\ & s \text{ odd, } \bar{r}(s) \in S_k^- \\ \\ 0 & , \quad s \text{ even} \end{cases}
 \end{aligned}$$

Admissible Control Laws:

$$\text{For } P_k \text{ at stage } s: \begin{cases} \bar{\gamma}_{k,s}(z(s)), \bar{\gamma}_{k,s}: Z \rightarrow U_k; \text{ if } \bar{r}(s) \in S_k, s \text{ even,} \\ \\ \text{or } F(s) \in S_k^-, s \text{ odd} \\ \\ \text{void; otherwise.} \end{cases}$$

Let $\bar{\Gamma}_{k,s}$ denote the corresponding strategy space of P_k . Then, the counterpart (extension) of Proposition 2.2 in this case would be the following:

Proposition 7.1:

If $\{J_k, \Gamma_{k,n}; k=1,2; n=0,1,\dots,N-1\}$ is an N -stage stochastic dynamic feedback game as defined in §3.1-§3.2, admitting a unique strong equilibrium solution (γ_1^*, γ_2^*) , there exists a $2N$ -stage stochastic dynamic feedback game $\{\tilde{J}_k, \tilde{\Gamma}_{k,s}; k=1,2; s=0,1,\dots, 2N-2\}$ as defined above, which admits a unique feedback Nash (strong) equilibrium solution $(\tilde{\gamma}_1^*, \tilde{\gamma}_2^*)$. Moreover, there is a unique correspondence between these two solutions, given by

$$\tilde{\gamma}_{k,s}^*([x(\frac{s}{2}), i, o_{m_1}, o_{m_2}]) = \gamma_{k, \frac{s}{2}}^*(x(\frac{s}{2}), i); i \in S_k, s \text{ even}$$

$$\begin{aligned} \tilde{\gamma}_{k,s}^*[x(\frac{s-1}{2}), i, \pi_k(o_{m_k}, \gamma_{k, \frac{s-1}{2}}^*(x(\frac{s-1}{2}), i))] \\ = \gamma_{k, \frac{s-1}{2}}^*(x(\frac{s-1}{2}), i); i \in S_k^-, s \text{ odd} \end{aligned}$$

$$\tilde{\gamma}_{k,s}^*([x(\frac{s-1}{2}), i, o_{m_1}, o_{m_2}]) = \gamma_{k, \frac{s-1}{2}}^*(x(\frac{s-1}{2}), i); i \in \eta, s \text{ even.}$$

$$s=0,1,\dots,2N-2 \quad ; \quad k=1,2.$$

□

8. Appendix II

In this appendix we provide a proof for Lemma 4.1 of Section 4, which follows the lines of the proof given in [7] for Theorems 1 and 2 (see also [8]), with the only difference being that we have to account for the dependence of λ_{ij} on t and x (whereas in [7] and [8] λ_{ij} 's were taken as constants).

Let $u = (u_1^i(t, x), u_2^i(t, x)) \in \mathcal{U}_1 \times \mathcal{U}_2$ be an admissible control law, (t, x) be a point at which a jump to $r(t, x) = i$ has occurred and $x^i[s; t, x]$ denote the corresponding state trajectory ($s \geq t$). Then, the probability density of the event that the first jump of r after time t is from i to j and occurs at time $s \geq t$ is given by

$$\lambda_{ij}(s, x^i[s; t, x]) \exp \left[\int_t^s \lambda_{ii}(\sigma, x^i[\sigma; t, x]) d\sigma \right] \quad (8.1a)$$

and the probability of the event that there are no jumps in the interval $(t, s]$ is

$$\exp \left[\int_t^s \lambda_{ii}(\sigma, x^i[\sigma; t, x]) d\sigma \right], \quad (8.1b)$$

which follow directly from the transition probabilities (4.7) and (4.8).

If there are no jumps after time t (until the terminal time T) the cost-to-go for P_k is clearly $q_k^i(x^i[T; t, x])$, whereas if a jump from i to j has occurred at time s , $t < s < T$, the cost-to-go from s onwards is (by definition) $v_k^j(s, x^i[s; t, x])$. Hence, using (8.1a) and (8.1b), v_k^i satisfies the integral equation

$$\begin{aligned}
V_k^i(t, x) = & q_k^i(x^i[T; t, x]) \exp \left[\int_t^T \lambda_{ii}(\sigma, x[\sigma; t, x]) d\sigma \right] \\
& + \sum_{j \neq i} \int_t^T \lambda_{ij}(s, x^i[s; t, x]) \exp \left[\int_t^s \lambda_{ii}(\sigma, x^i[\sigma; t, x]) d\sigma \right] \\
& \cdot V_k^j(s, x^i[s; t, x]) ds
\end{aligned} \quad (8.2)$$

Using (8.2), we can now compute

$$\begin{aligned}
& \int_t^T \lambda_{ii}(v, x^i[v; t, x]) V_k^i(v, x^i[v; t, x]) dv = \\
& q_k^i(x^i[T; t, x]) \int_t^T \lambda_{ii}(v, x^i[v; t, x]) \exp \left[\int_v^T \lambda_{ii}(\sigma, x^i[\sigma; t, x]) d\sigma \right] dv \\
& + \sum_{j \neq i} \int_t^T \int_v^T \lambda_{ii}(v, x^i[v; t, x]) \lambda_{ij}(s, x^i[s; t, x]) \\
& \cdot \exp \left[\int_v^s \lambda_{ii}(\sigma, x^i[\sigma; t, x]) d\sigma \right] V_k^j(s, x^i[s; t, x]) ds dv
\end{aligned} \quad (8.3)$$

By an interchange of the order of integration in the double integral in (8.3) we get for this term

$$\begin{aligned}
& \sum_{j \neq i} \int_t^T \lambda_{ij}(s, x^i[s; t, x]) V_k^j(s, x^i[s; t, x]) \\
& \cdot \int_t^s \lambda_{ii}(v, x^i[v; t, x]) \exp \left[\int_v^s \lambda_{ii}(\sigma, x^i[\sigma; t, x]) d\sigma \right] dv ds
\end{aligned}$$

which, after integration with respect to v , yields

$$\sum_{j \neq i} \int_t^T \lambda_{ij}(s, x^i[s; t, x]) (-1 + \exp \int_t^s \lambda_{ii}(\sigma; x^i[\sigma; t, x]) d\sigma) V_k^j(s, x^i[s; t, x]) ds, \quad (8.4)$$

while the first term on the RHS of (8.3) is equal to

$$q_k^i(x^i[T; t, x]) (-1 + \exp \int_t^T \lambda_{ii}(\sigma; x^i[\sigma; t, x]) d\sigma). \quad (8.5)$$

Therefore, by (8.4), (8.5) and using again (8.2) we find that (8.3) is equal to

$$V_k^i(t, x) - q_k^i(x^i[T; t, x]) - \sum_{j \neq i} \int_t^T \lambda_{ij}(s, x^i[s; t, x]) V_k^j(s, x^i[s; t, x]) ds \quad (8.6)$$

which finally yields the integral equation satisfied by $V_k^i(t, x)$:

$$V_k^i(t, x) = q_k^i(x^i[T; t, x]) + \sum_{j \neq i} \int_t^T \lambda_{ij}(s, x^i[s; t, x]) V_k^j(s, x^i[s; t, x]) ds. \quad (8.7)$$

This implies that $V_k^i(t, x)$ is piecewise continuously differentiable in t .

We thus have, by (8.7),

$$\frac{d}{ds} V_k^i(s, x^i[s; t, x]) = \frac{\partial V_k^i}{\partial s} + \frac{\partial V_k^i}{\partial x} \frac{dx^i(s)}{ds} = - \sum_j \lambda_{ij}(s, x^i[s; t, x]) V_k^j(s, x^i[s; t, x])$$

Evaluating this at $s = t$ we obtain (4.10). □

It should be noted that Rishel's Theorem 3 of [7] can also be directly extended to this class of stochastic systems, yielding a maximum principle for the partial differential operator of the form (4.10). □

References

- [1] Von Stackelberg, H., Marktform und Gleichgewicht, Springer, Vienna, 1934.
- [2a] Simaan, M. and J. B. Cruz, Jr., "On the Stackelberg strategy in nonzero-sum games," Journal of Optimization Theory and Applications, 11, no. 5, 1973a, 533-555.
- [2b] Simaan, M. and J. B. Cruz, Jr., "Additional aspects of the Stackelberg strategy in nonzero-sum games," Journal of Optimization Theory and Applications, 11, no. 6, 1973b, 613-626.
- [3] Başar, T. and G. J. Olsder, Dynamic Noncooperative Game Theory, Academic Press, London/New York, 1982.
- [4] Nash, J., "Noncooperative games," Annals of Mathematics, 54, no. 2, 1951, 286-295.
- [5] Başar, T., "A counter example in linear-quadratic games: Existence of nonlinear Nash solutions," Journal of Optimization Theory and Applications, 14, no. 4, 1974, 425-430.
- [6] Başar, T., "On the uniqueness of the Nash solution in linear-quadratic differential games," International Journal of Game Theory, 5, no. 2/3, 1976a, 65-90.
- [7] Rishel, R., "Control of systems with jump Markov disturbances," IEEE Transactions on Automatic Control, AC-20, no. 2, 1975, 241-244.
- [8] Rishel, R., "Dynamic programming and minimum principles for systems with jump Markov disturbances," SIAM J. Control, 13, no. 2, 1975, 338-371.
- [9a] Starr, A. W. and Y. C. Ho, "Nonzero-sum differential games," Journal of Optimization Theory and Applications, 3, no. 3, 1969a, 184-206.
- [9b] Starr, A. W. and Y. C. Ho, "Further properties of nonzero-sum differential games," Journal of Optimization Theory and Applications, 3, 1969b, 207-219.
- [10] Friedman, A., Differential Games, John Wiley, New York, 1971.
- [11] Chen, C. I. and J. B. Cruz, Jr., "Stackelberg solution for two-person games with biased information patterns," IEEE Transactions on Automatic Control, AC-17, no. 5, 1972, 791-798.

- [12] Castanon, D. A., "Equilibria in Stochastic Dynamic Games of Stackelberg Type," Ph.D. Thesis, M.I.T., Electronics Systems Laboratory, Cambridge, Massachusetts, 1976.
- [13] Papavassilopoulos, G. P. and J. B. Cruz, Jr., "Nonclassical control problems and Stackelberg games," IEEE Transactions on Automatic Control, AC-24, no. 2, 1979a, 155-166.
- [14] Elliott, R. J. and N. J. Kalton, "The existence of value in differential games of pursuit and evasion," Journal of Differential Equations, 12, 1972, 504-523.
- [15] Elliott, R. J., "Feedback strategies in deterministic differential games," in P. Hagedorn, H. W. Knobloch and G. J. Olsder, eds., Differential Games and Applications, Lecture Notes in Control and Information Sciences, 3, Springer-Verlag, Berlin, 1977, 136-142.
- [16] Fleming, W. H., "The convergence problem for differential games," Journal of Mathematical Analysis and Applications, 3, 1961, 102-116.
- [17] Varaiya, P. P. and J. G. Lin, "Existence of saddle points in differential games," SIAM Journal on Control, 7, 1969, 141-157.
- [18] Başar, T. and A. Haurie, "Feedback Stackelberg solution in continuous-time multi-level games," Proceedings of the 21st IEEE Conference on Decision and Control, Orlando, Florida, 1982.

Notation, Terminology and Abbreviations

I_m	:	($m \times m$)-dimensional identity matrix
$O_{m_0 \times m_1}$	:	($m_0 \times m_1$)-dimensional zero matrix ; $O_{m \times m} \triangleq O_m$
o_m	:	Zero vector of dimension m .
R^m	:	m -dimensional real line (Euclidean space)
P_k	:	Player k
U_k	:	Control set of P_k
$U_{k,n}$	:	Control set of P_k at stage n , for the discrete-time game
U_k	:	Set of open-loop controls of P_k
u_k	:	Open or closed-loop control of P_k
$\mathcal{U}_k$	:	Set of closed-loop feedback controls of P_k
$\mathcal{U}_{k,n}$	:	Set of closed-loop feedback controls of P_k at stage n , for the discrete-time game
Γ_k	:	Strategy space of P_k
γ_k	:	Strategy of P_k
$\pi_k(a,b)$	=	(a,b) if $k=1$ (b,a) if $k=2$
$\bar{k}$	=	1 if $k=2$ 2 if $k=1$
w.r.t.	:	with respect to
$[0,T]$	:	Time interval on which the differential game is defined
X	:	State space
S_k	:	Set of indices corresponding to the states of the jump process with asymmetric mode of play under P_k 's leadership
η	:	Set of indices corresponding to the states of the jump process with symmetric mode of play
$I \triangleq S_1 + S_2 + \eta$		
ϕ	:	empty set
RHS (LHS)	:	right hand side (left hand side)

FEEDBACK STACKELBERG SOLUTION IN CONTINUOUS-TIME MULTI-LEVEL GAMES

Tamer Basar
University of Illinois
1101 West Springfield Avenue
Urbana, Illinois 61801, U.S.A.

and Alain Haurie
Ecole des Hautes Etudes Commerciales
5255 avenue Decelles
Montréal, Québec H3T 1V6, Canada

Abstract: This paper introduces a "Feedback Stackelberg" solution concept for continuous-time multi-level dynamic optimization problem, and discusses its appropriateness for decision making and optimization in the presence of hierarchy.

Acknowledgements: Research of T. Basar was supported in part by the Office of Naval Research (N00014-82-K-0469) and in part by the Joint Services Electronics Program (N00014-79-C-0424). Research of A. Haurie was supported in part by SSHRC-Canada (410-81-0722R1) and in part by FCAC-Québec.

1. INTRODUCTION

This paper introduces a "generalized feedback Stackelberg solution" concept for continuous-time multi-level dynamic optimization problems. The basic idea is to relate this solution concept to a class of feedback equilibria with particular information structures.

A complete and formal discussion of these topics is provided in Ref. [1]. The present paper is devoted to a quick presentation of the main results proved in [1] and a first discussion of the possible use of this solution concept in the modeling of economics imperfect competition problems.

In section 2 the Stackelberg solution concept is discussed in the realm of extensive games. It is shown on simple game trees that the Stackelberg solution concept is in fact an equilibrium solution associated with a peculiar information structure. The feedback Stackelberg solution (FSS) concept for multi-stage game is then discussed as a natural extension of this feedback equilibrium solution.

With this new interpretation of the FSS, a natural extension is to let the leadership change with time. This leads to the concept of "generalized feedback Stackelberg solution" presented in section 3 for a continuous time system with jump disturbances affecting the structure and the mode of play of the game. The link between the multistage and differential game formulation is more fully studied in [1] using the 5-game approach proposed by Friedman ([2]).

This new solution concepts permits one to formulate a new class of imperfect competition models, where the market structure changes with dominant firms risking to loose their leadership. This is discussed in section 4.

2. THE STACKELBERG SOLUTION AS AN EQUILIBRIUM

Consider a static two player game defined by the cost functionals $J_k: U_1 \times U_2 \rightarrow \mathbb{R}$, $k=1,2$, where U_k is the set of admissible controls for player k (denoted P_k hereinafter).

If P_1 is the leader, he announces his control $u_1^0 \in U_1$ first, to which P_2 reacts optimally by minimizing $J_2(u_1^0, u_2)$ over $u_2 \in U_2$. Assume there is a unique mapping $T_2: U_1 \rightarrow U_2$ such that

$$J_2(u_1, T_2(u_1)) = \min_{u_2 \in U_2} J_2(u_1, u_2), \quad \forall u_1 \in U_1.$$

Then a pair $(u_1^*, u_2^*) \in U_1 \times U_2$ is called a Stackelberg solution for the static game with P_1 the leader if

$$u_1^* = \arg \min_{u_1 \in U_1} J_1(u_1, T_2(u_1))$$

$$u_2^* = T_2(u_1^*).$$

A diagrammatic representation of such a game is given in Figure 1. The game is also defined as the following matrix game.

μ_2 μ_1	0	1
0	1,1	2,0
1	1,2	0,4

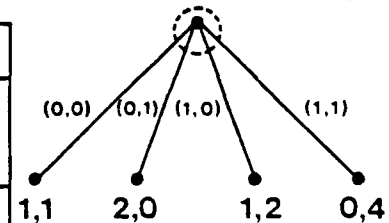


Fig. 1. Static Game

Here one has

$$T_2(0) = 1, \quad T_2(1) = 0.$$

Therefore $\min_{u_1 \in U_1} J_1(u_1, T_2(u_1)) = \min\{2, 1\} = 1$

and the pair $(u_1^*, u_2^*) = (1, 0)$ is the Stackelberg solution with P_1 the leader.

Now consider the extensive game defined by the tree and the information structure shown on Figure 2. (For a most recent presentation of games in extensive form see Ref. [3]). Next to Figure 2 the matrix game formulation is given. This game in extensive form describes the sequence of two moves: first P_1 moves, then P_2 moves knowing the control chosen by P_1 . Therefore the decision of P_2 is now described by a mapping $\gamma_2: U_1 \rightarrow U_2$. There are four of them. $(0,0)$ is the mapping such that $\gamma_2(0)=0$, $\gamma_2(1)=0$, $(0,1)$ is the mapping such that $\gamma_2(0)=0$, $\gamma_2(1)=1$ etc...

μ_2	(0,0)	(0,1)	(1,0)	(1,1)
μ_1				
0	1,1	1,1	2,0	2,0
1	1,2	0,4	1,2	0,4

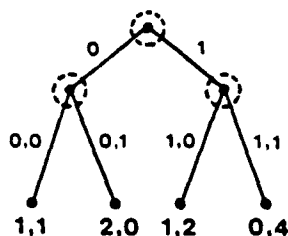


Fig. 2: Game in extensive form

The circled cost pair (1,2) in the bimatrix game is clearly a Nash equilibrium. The associated pair of controls (u_1^*, u_2^*) defined by $u_1^* = 1$, $u_2^* = \gamma_2^*(1) = 0$ is also the Stackelberg solution previously obtained. Therefore there is a possibility to interpret the Stackelberg solution as a particular equilibrium in a suitably defined game in extensive form. Notice that there may be many other equilibria for the game in Fig. 2 due to informationally non uniqueness (see Ref. [3]). What has been illustrated on this simple example could be generalized to more general games such as the multistage games considered by Simaan and Cruz in Ref. [4a,b]. One can prove the following proposition (see [1]):

Proposition 1: <<If an N-stage dynamic feedback game admits a unique feedback Stackelberg solution (in the sense of Ref. [4] with P1 the leader, there exists a 2N-stage dynamic feedback game which admits a unique feedback Nash equilibrium with a unique correspondence between these two solutions.>>

This result permits one to relate the feedback Stackelberg solution to the normal form description of the game and therefore to give a "good definition" of this solution concept with regard to the theory of games. Furthermore the interpretation of leadership as an asymmetry in the information structure leads to the consideration of dynamical systems with varying leadership. This idea will be further pursued in the next section which deals with continuous time systems.

3. CONTINUOUS TIME FEEDBACK STACKELBERG SOLUTION

Consider for $t \in [0, T]$ a stochastic system of the form

$$\dot{x} = f^r(t, x, u_1, u_2), \quad x \in X, \quad u_k \in U_k, \quad k=1,2 \quad (1)$$

with initial condition

$$x(0) = x^0, \quad r(0) = r^0.$$

The state x belongs to $X \subset \mathbb{R}^n$, and the control u_k takes values in $U_k \subset \mathbb{R}^{m_k}$, ($k=1,2$). In (1) $r(t)$ is a finite-state stochastic jump process which takes values in I . The RHS of (1) changes from $f^i(\cdot)$ to $f^j(\cdot)$ as $r(t)$ jumps from i to j . For each $i \in I$, f^i is C^1 w.r.t all its arguments. Let U_k , $k=1,2$, be two classes of admissible control laws $u_k^i(t, x)$ with value in U_k , defined on $I \times [0, T] \times X$ such that $u_k^i(\cdot)$ is piecewise continuous in t , C^1 in x .

The relationship between the processes $x(t)$ and $r(t)$ is assumed to be such that for any control law and almost any ω in the sample space (Ω, F) there exists a piecewise constant function $\omega: [0, T] \times X \rightarrow I$ such that $(x(t, \omega), r(t, \omega))$ satisfies

$$\dot{x} = f^{\omega(t, x)}(t, x, u_1(t, x), u_2(t, x))$$

$$r(t, \omega) = \mu(t, x).$$

It is also assumed that for any admissible control law $\tilde{u} \in U_1 \times U_2$ the following conditional Markov property holds

$$P^{\tilde{u}}[\mu(t+h, x(t+h)) = j | \mu(t, x(t)) = i, x(t) = x] = \lambda_{ij}^1(t, x)h + o(h) \quad (2)$$

$$P^{\tilde{u}}[\mu(t+h, x(t+h)) = i | \mu(t, x(t)) = i, x(t) = x] = 1 + \lambda_{ii}^1(t, x)h + o(h) \quad (3)$$

where $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$ uniformly in x and $\tilde{u}$.

For any admissible control law $\tilde{u}$ let $V_k^1(t, x)$ denote the corresponding value of the conditional expectation

$$V_k^1(t, x) = E_{\tilde{u}}[q_k^r(T)(x(T)) | x(t) = x, r(t) = i] \quad (4)$$

where $q_k^1(\cdot)$, $i \in I$, $k=1,2$ are C^1 functions.

The system (1)-(4) is a slight generalization of the one studied by Rishel in [5]. The process $r(t)$ models random structural changes in the dynamical system to control. The two controls u_1 and u_2 correspond respectively to P1 and P2, and one assumes that the mode of play of the dynamic game is determined by the outcome of $r(t)$. More precisely I is assumed to be partitioned into three subsets $I = S_1 + S_2 + N$. If $r(t) \in S_k$, it is said that P_k is the leader at time t . If $r(t) \in N$, there is no leader at time t . Consider two classes Γ_k , ($k=1,2$) of functions $\gamma_k: [0, T] \times X \times I \rightarrow U_k$,

where $\bar{k}=2$ if $k=1$, $\bar{k}=1$ if $k=2$, and such that

$$i \in NUS_k \Rightarrow \gamma_k(t, x, i, u_{\bar{k}}) = \hat{\gamma}_k(t, x, i). \quad (5)$$

The relation (5) expresses the fact that when P_k is not the follower the function γ_k cannot depend on $u_{\bar{k}}$.

The functions γ_k are assumed to be C^1 in x and $u_{\bar{k}}$.

It is possible to associate a control law $\tilde{u} \in U_1 \times U_2$ with a pair $(\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2$ by defining

$$\left. \begin{aligned} u_k^i(t, x) &\triangleq \hat{\gamma}_k(t, x, i) \text{ if } i \in NUS_k \\ u_k^i(t, x) &\triangleq \gamma_k(t, x, i, \hat{\gamma}_{\bar{k}}(t, x, i)) \text{ if } i \in S_{\bar{k}} \end{aligned} \right\} \quad (6)$$

A pair $(\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2$ is called a pure-feedback strategy (PFS) pair if the control law defined through (6) is admissible. Associated with a PFS are two cost functions

$$J_k(t, x, i; \gamma_1, \gamma_2) \triangleq E_{\tilde{u}}[q_k^r(T)(x(T)) | x(t) = x, r(t) = i] \quad (7)$$

where $\tilde{u}$ satisfies (6).

A PFS (γ_1^*, γ_2^*) is said to be a weak-equilibrium if for every $t \in [0, T]$, $x \in X$, $i \in I$ the following holds:

$$\begin{aligned} J_1(\tau, \xi, i; \gamma_1^*, \gamma_2^*) &\leq J_1(\tau, \xi, i; \gamma_1, \gamma_2^*) \\ J_2(\tau, \xi, i; \gamma_1^*, \gamma_2^*) &\leq J_2(\tau, \xi, i; \gamma_1^*, \gamma_2) \end{aligned} \quad (8)$$

for any PFS pair (γ_1, γ_2^*) or (γ_1^*, γ_2) .

This is the usual definition of a feedback Nash equilibrium. Now, by definition of γ_k^* , P_k has the opportunity to react "locally" at t to the choice of control made by P_k if this is the leader ($r(t) \in S_k^-$).

Define the ε -deviation of P_k at (t, x, i) with value $\bar{u}_k^D$ as a PFS $\gamma_{k; \varepsilon, \bar{u}_k^D}^D$ such that

$\forall t \in [t, t+\varepsilon]$, if $r(s) = r(t)$ for $s \in [t, t+\varepsilon]$ then $u_k(\tau) = \bar{u}_k^D$, and which coincides with γ_k^* otherwise.

Define the ε -reaction of P_k at (t, x, i) with value $\bar{u}_k^R$ as a PFS $\gamma_{k; \varepsilon, \bar{u}_k^R}^R$ such that

$\forall t \in [t, t+\varepsilon]$, if $r(s) = r(t)$ for $s \in [t, t+\varepsilon]$ then $u_k(\tau) = \bar{u}_k^R$ and which coincides with γ_k^* otherwise.

Definition: A PFS pair (γ_1^*, γ_2^*) is a strong equilibrium (or a generalized feedback Stackelberg solution (GFSS)) if: (i) it is a weak equilibrium; (ii) for any $(t, \xi, i) \in [0, T] \times X \times S_k^-$, $\bar{u}_k^D \in U_k^-$, $\bar{u}_k^R \in U_k^-$, the ε -deviation $\gamma_{k; \varepsilon, \bar{u}_k^D}^D$ and the ε -reaction $\gamma_{k; \varepsilon, \bar{u}_k^R}^R$ are well defined for ε small enough and the following holds:

$$\begin{aligned} J_k(t, \xi, i; \gamma_{k; \varepsilon, \bar{u}_k^D}^D, \gamma_k^*) &\leq \\ J_k(t, \xi, i; \gamma_{k; \varepsilon, \bar{u}_k^R}^R, \gamma_{k; \varepsilon, \bar{u}_k^D}^D) &+ o(\varepsilon) \end{aligned} \quad (9)$$

with $\lim_{\varepsilon \rightarrow 0} \frac{o(\varepsilon)}{\varepsilon} = 0$.

Remark 1: When P_k is the follower he has the possibility to adjust to the choice of control by P_k . In a GFSS this adjustment is "locally optimal".

Remark 2: In Ref [1] the differential game is more precisely defined using the 5-game approach of Friedman. It is shown that a PFS may be defined as a sequence of 5-strategies used for $G(5)$ games having the structure of multistage games.

The main result, proved in [1] is the following Hamilton-Jacobi equation characterization of a GFSS.

Proposition: Under assumption A1-A5 (specified in [1]), a necessary and sufficient condition for a PFS pair (γ_1^*, γ_2^*) , generating the control law $\bar{u}^*$, to be a GFSS is that for each $i \in I$, $k \in 1, 2$, the cost-to-go functions $V_k^i(t, x)$ defined by (4) (6) satisfy the following partial differential equations:

$$\begin{aligned} \min_{u_k \in U_k} \left\{ \frac{\partial}{\partial t} V_k^i(t, x) + \frac{\partial}{\partial x} V_k^i(t, x) f^i(t, x, u_k, \gamma_k^*(t, x, i, u_k)) \right\} + \\ \sum_{j \in I} \lambda_{ij}(t, x) V_k^j(t, x) = \frac{\partial}{\partial t} V_k^i(t, x) + \\ \frac{\partial}{\partial x} V_k^i(t, x) f^i(t, x, u_1^*(t, x), u_2^*(t, x)) + \\ \sum_{j \in I} \lambda_{ij}(t, x) V_k^j(t, x) \end{aligned} \quad (10)$$

with the boundary conditions

$$V_k^i(t, x) = q_k^i(x) \quad i \in I, x \in X \quad (11)$$

and with the additional condition that

$$V(t, x) \in [0, T] \times X, \quad \forall i \in S_k^-, \quad \forall u_k \in U_k, \quad \forall u_{-k} \in U_{-k}$$

$$\begin{aligned} \frac{\partial}{\partial x} V_k^i(t, x) f^i(t, x, u_k, \gamma_k^*(t, x, i, u_k)) \leq \\ \frac{\partial}{\partial x} V_k^i(t, x) f^i(t, x, u_k, u_{-k}) \end{aligned} \quad (12)$$

$k=1, 2, \quad \bar{k}=1, 2, \quad k \neq \bar{k}$.

Remark 3: In the particular case where there is no uncertainty, with one player P_k the leader, the concept of GFSS happens to be the continuous time counterpart of the feedback Stackelberg solution concept proposed in [4] for multistage systems.

4. APPLICATION TO A NEW CLASS OF IMPERFECT COMPETITION MODELS.

4.1 Dominant firms and leadership

Von Stackelberg initially proposed the solution concept associated with his name, as a description of the economic warfare between two firms in a situation of duopoly (Ref. [6]). He insisted on the fact that, in most cases, both firms would desire to become the "leader", which would lead to an impossible equilibrium.

Some markets are characterized by a "dominant firm" which is usually the "price leader", and a "competitive fringe" composed of small firms acting as "price takers". The Stackelberg solution concept is applicable to such a situation. Recently several attempts have been made to model such markets using a differential game approach. Gilbert [7] proposes, for example, an interesting model of an "OPEC-like" cartel exploiting an exhaustible resource. The proposed model makes use of the so-called open loop Stackelberg solution, and it is therefore assumed that the cartel announces its production path for the whole future. Such an assumption is difficult to accept without criticism since such a behaviour is seldom observed (certainly not from OPEC members). What is observed, in the case of a market dominated by a firm or a cartel, is only the fact that the leader has to announce first his current decision. Non-OPEC members (e.g. Canada, England, etc...) know, at each instant of time the price set by OPEC members and they can adjust to this information. What they don't know (and the recent impact on Canada and Mexico economies of the oil-glut shows it) is the path of oil extraction or the path of oil price for the future.

We contend that the generalized feedback Stackelberg solution, with its interpretation as an equilibrium associated with a particular information structure, should be more appropriate for the analysis of such imperfect markets.

It may happen that the Nash-feedback solution could leave one player better off than the FSS with this player the leader. However this is perfectly normal if, being the "leader", only means that a sort of informational disadvantage is associated with the dominant position. The leader has to act first and thus gives some extra information to the follower.

Furthermore, by letting the leadership be the outcome of a random process, the displacement of leadership from one player to another can be modelled. This would eliminate some difficulties associated with the desire for each player to become the leader.

In the next sub-section a model of competition through advertising is proposed as an illustration of the modeling permitted by GFSS.

4.2 Profit maximization through advertising

Two firms are competing on a given market through their advertising expenditures. It is assumed that the total advertising made by both firms together creates the demand for the product on the market, while the relative values of advertising per dollar of sale at a given time determine the market share obtained by each firm. It is finally assumed that the marginal production costs are constant for both firms.

The state of this system at time t is supposed to be described by the pair $(x(t), r(t))$ where $x(t) \geq 0$ is the total demand for the marketed product, expressed in dollars of sale, and $r(t) \in \{0, 1, 2\}$ is an indicator of the presence of a dominant firm ($r(t) = k \neq 0$ means that firm k is dominant and acts as a leader at time t ; $r(t) = 0$ means that there is no leader).

The control of each firm k at time t is given by $u_k(t)$, the advertising expenditure per dollar of sale. So, if the firm k has a sales level of $x_k(t)$ at time t , then its total advertising expenditure will be

$$a_k(t) = u_k(t) x_k(t) \quad (13)$$

In order to describe the sharing of the market at time t , consider a function $g^j: R^2 \rightarrow [0, 1]$ defined for each value $j \in \{0, 1, 2\}$ and such that

$$\left. \begin{aligned} x_1(t) &= g^{r(t)}(u_1(t), u_2(t)) x(t) \\ x_2(t) &= (1 - g^{r(t)}(u_1(t), u_2(t))) x(t) \end{aligned} \right\} \quad (14)$$

Notice that it is assumed that the market structure (i.e. existence of a dominant firm) can affect the market sharing mechanism.

The process $r(t)$ is assumed to be a Markov chain with continuous jump rate functions $\lambda_{ij}(t)$, $i, j = 0, 1, 2$, while the sales evolution is described by

$$\frac{dx(t)}{dt} = (\alpha G^{r(t)}(u_1(t), u_2(t)) - \beta) x(t) \triangleq \alpha [u_1(t) g^{r(t)}(u_1(t), u_2(t)) + u_2(t) (1 - g^{r(t)}(u_1(t), u_2(t)))] x(t) - \beta x(t) \quad (15)$$

Notice that, according to (13)-(14) the first term in the R.H.S. of (15) corresponds to $\alpha(a_1(t) + a_2(t))$, and therefore is proportional to the total advertising made at time t . The constants α and β are positive. The equation (15) shows an exponential decay of the sales in the absence of advertising.

At time t , the rate of profit of firm k is given by

$$\pi_k(t) = (1 - c_k - u_k(t)) x_k(t) \quad (16)$$

where c_k is the constant production cost. Each firm wants to maximize its total expected profit over a fixed time period $[0, T]$.

The equations (13)-(16) describe a dynamical system having the properties assumed in section 3.

Using the necessary and sufficient conditions stated in section 3 one is able to characterize a GFSS (γ_1^*, γ_2^*) by solving the Hamilton Jacobi equations

$$-\frac{\partial}{\partial t} V_k^j(t, x) - \sum_{i=0}^2 \lambda_{ji}(t) V_k^i(t, x) =$$

$$(a) \quad \max_{u_k} \left[\frac{\partial}{\partial x} V_k^j(t, x) (\alpha G^j([u_k, \tilde{u}_k^j(t, x)]) - \beta) x + (1 - c_k - u_k) g_k^j([u_k, \tilde{u}_k^j(t, x)]) x \right] \text{ if } j \neq k$$

$$(b) \quad \max_{u_k} \left[\frac{\partial}{\partial x} V_k^j(t, x) (\alpha G^j([u_k, T_k^j(t, x, \frac{\partial}{\partial x} V_k^j(t, x), u_k)]) - \beta) x + (1 - c_k - u_k) g_k^j([u_k, T_k^j(t, x, \frac{\partial}{\partial x} V_k^j(t, x), u_k)]) x \right] \text{ if } j = k$$

where

$$T_k^j(t, x, \frac{\partial}{\partial x} V_k^j(t, x), u_k) = \arg \max_{u_k} \left[\frac{\partial}{\partial x} V_k^j(t, x) \right.$$

$$\left. (\alpha G^j([u_k, u_k]) - \beta) x + (1 - c_k - u_k) g_k^j([u_k, u_k]) x \right].$$

Here one has used the notations g_k^j for the function which is equal to g^j if $k=1$ or to $1-g^j$ if $k=2$. Also $[u_k, u_k]$ stands for the pair (u_1, u_2) if $k=1$ or 2. The

control laws $\tilde{u}_k^*(t, x)$ are the control laws generated by (γ_1^*, γ_2^*) as shown in section 3.

A solution to these equations can be found in the class of strategies defining control laws which are independent of x . To show this assume that the functions V_k^j have the particular affine form

$$V_k^j(t, x) = \ell_k^j(t) + m_k^j(t) x \quad (17)$$

and therefore satisfy

$$\frac{\partial}{\partial x} V_k^j(t, x) = m_k^j(t) \quad (18)$$

substituting in the R.H.S. of the Hamilton-Jacobi equations one sees that x factorizes in these expressions leading to the following conditions on $u_k, k=1, 2$

$$u_k = \arg \max_{u_k} \left[m_k^j(t) (\alpha G^j([u_k, \tilde{u}_k^j(t, x)]) - \beta) + (1 - c_k - u_k) g_k^j([u_k, \tilde{u}_k^j(t, x)]) \right] \text{ if } j \neq k$$

$$u_k = \arg \max_{u_k} \left[m_k^j(t) (\alpha G^j([u_k, T_k^j(t, u_k)]) - \beta) + (1 - c_k - u_k) g_k^j([u_k, T_k^j(t, u_k)]) \right] \text{ if } j = k$$

with

$$T_k^j(t, u_k) = \arg \max_{u_k} \left[m_k^j(t) (\alpha G^j([u_k, u_k]) - \beta) + (1 - c_k - u_k) g_k^j([u_k, u_k]) \right].$$

It appears that a solution to these conditions can be found in the class of control laws

$$\tilde{u}_k^j(t, x) = \tilde{u}_k^j(t) \quad j=0, 1, 2; \quad k=1, 2.$$

It would be now straightforward to verify that such a control law is compatible with the affine structure (17) for the functions $V_k^j(t, x)$, yielding eventually to

$\ell_k^j(t) \equiv 0$, $j=0, 1, 2$; $k=1, 2$, and $m_k^j(t)$, $j=0, 1, 2$; $k=1, 2$ obtained as the solution of a set of non linear differential equations.

This characterization of an equilibrium is obviously reminiscent of Refs. [8]-[11] dealing with a similar structure in various economic models based on an optimal control or differential game paradigm.

When the time horizon T tends to infinity, and if the profits are discounted, it is possible to obtain a solution of the Hamilton-Jacobi equations in the class of stationary strategies. In this particular model this would lead to a strong equilibrium obtained in the class of piecewise constant control laws

$$\bar{u}^{*1}(t, x) = u^{*1} \quad i = 0, 1, 2.$$

A consequence of the linear structure of the functions V_k^j will then be that the leader will always be better-off by using his leadership position (and announcing his current control) than by playing according to the usual feedback Nash solution.

A detailed proof of this result is left for a forthcoming paper.

5. CONCLUSION

In this paper, the feedback Stackelberg solution has been given a precise definition as a particular equilibrium in a game of extensive form having a special information structure. By using a limiting process à la Friedman a similar definition is given for a FSS in a differential game with jump disturbances.

In Ref. [1] the case of Linear-Quadratic differential games is fully treated and it is shown that, in the presence of cross terms in the cost functions involving the controls of both players, the FSS is different from the usual feedback Nash equilibrium.

REFERENCES

- [1] T. Basar and A. Haurie, "Feedback Equilibria in Differential Games with Structural and Modal Uncertainties", in J.B. Cruz edit. Advances in Large Scale Systems, Vol. 1, JAI Press Corp., 1983.
- [2] Friedman, A., Differential Games, John Wiley, New York, 1971.
- [3] Basar, T. and G.J. Olsder, Dynamic Noncooperative Game Theory, Academic Press, London/New York, 1982.
- [4a] Simaan, M. and J.B. Cruz, Jr., "On the Stackelberg Strategy in Nonzero-sum games", Journal of Optimization Theory and Applications, 11, No. 5.
- [4b] Simaan, M. and J.B. Cruz, Jr., "Additional Aspects of the Stackelberg Strategy in nonzero-sum games", Journal of Optimization Theory and Applications, 11, No. 6, 1973, 613-626.
- [5] Rishel, R., "Control of Systems with Jump Markov Disturbances", IEEE Transactions on Automatic Control, AC-20, No. 2, 1975, 241-244.
- [6] Von Stackelberg, H., Marktform und Gleichgewicht, Springer, Vienna, 1934.
- [7] R.J. Gilbert, "Dominant Firm Policy in a market for an exhaustible resource", The Bell Journal of Economics, 1979.
- [8] D.G. Luenberger, "A Nonlinear Economic Control Problem with a Linear Feedback Solution", IEEE Transactions on Automatic Control, Vol. AC-20, April 1975.
- [9] A. Haurie and M.P. Polis, "Comments on A Non Linear Economic Control Problem with a Linear Feedback Solution", IEEE Transactions on Automatic Control, Vol. AC-21, Dec. 1976.
- [10] G. Leitmann and W. Schmitendorf, "Profit Maximization Through Advertising: A Nonzero-sum of Differential Game Approach", IEEE Transactions on Automatic Control, Vol. AC-23, August 1978.
- [11] J.R. Reinganum, "A Class of Differential Games for Which the Closed-loop and Open-loop Nash Equilibria Coincide", Journal of Optimization and Applications, Vol. 36, February 1982.

January 1984

To appear in the Journal of Optimization and Applications, 1985.

51

Informational Uniqueness of Closed-Loop Nash Equilibria
for a Class of Nonstandard Dynamic Games¹⁾

*Dedicated to Professor George Leitmann on the
occasion of his 60th birthday.*

by

TAMER BAŞAR²⁾

Decision and Control Laboratory
Department of Electrical Engineering
and Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Avenue
Urbana, Illinois 61801

¹⁾ Research that led to this paper was supported in part by the Office of Naval Research under Contract N00014-82-K-0469 and in part by the U.S. Air Force under Grant AFOSR-84-0054.

²⁾ Professor of Electrical Engineering and Research Professor in the Coordinated Science Laboratory, University of Illinois, 1101 W. Springfield Avenue, Urbana, Illinois 61801.

Abstract. This paper discusses an extension of the currently available theory of noncooperative dynamic games to game models whose state equations are of order higher than one. In a discrete-time framework it first elucidates the reasons why the theory developed for first-order systems is not applicable to higher order systems, and then presents a general procedure to obtain informationally unique Nash equilibrium solution in the presence of random disturbances. A numerical example solved in the paper illustrates the general approach.

Key Words. Dynamic games, noncooperative differential games, Nash equilibrium solution, uniqueness of equilibria, second-order systems, stochastic dynamics, closed-loop information pattern.

1. Introduction

It is by now an established fact in the theory of dynamic games that redundant information leads to nonuniqueness of noncooperative (Nash) equilibria, giving rise to so-called informationally nonunique equilibrium solutions (Ref. 1). This is true in the case of both differential and discrete-time deterministic games, under, for example, the closed-loop information pattern for at least one player, which includes memory (that is, knowledge of past values of the state variable). One way of removing informational nonuniqueness in deterministic dynamic games with closed-loop information patterns is to "robustify" the state equation by including a zero-mean additive noise term, such as (in the case of a discrete-time state equation with two players)

$$x_{k+1} = \tilde{f}(k, x_k, u_k, v_k) + w_k \quad (1.1)$$

where w_k is the noise term accounting for the inaccuracies in the modelling. Here, x_n , u_n , v_n are the state variable and the control variables of players 1 and 2, respectively, and $\{w_k\}$ is a sequence of i.i.d. random vectors of the same dimension as x_k (say, n) and with probability distribution that assigns positive probability mass to every open subset of $\mathbb{R}^n$. With the inclusion of such a term in the state equation we know that informational nonuniqueness disappears and the Nash equilibrium solution becomes more meaningful (Ref. 1).

Another method to remove informational nonuniqueness in Nash equilibria is to restrict the solution concept further to "delayed commitment" type Nash equilibria, which leads to the so-called feedback Nash equilibrium, and this is free of informational nonuniqueness (Ref. 1). But both these methods, as discussed above, and the verification of informational nonuniqueness of equilibria assume that the dynamics are described by first-order (differential or difference) equations. However, this is not an exhaustive class of dynamic games, since

there exist models in both engineering and economics, whose dynamics are initially described by equations of second or higher orders. In this paper, we study a class of such dynamic games described in the discrete-time domain and present a number of results which shed light on the properties (such as existence, uniqueness, solvability) of Nash equilibrium in dynamic games whose initial dynamics are not of first order.

In the next section, we will provide a mathematical formulation of a class of two-person nonzero-sum dynamic games whose state equations (in discrete-time) are of second order, and information pattern is closed-loop for both players. Using this class as a prototype model we will show why it would not be possible to reformulate such higher order state dynamics as first order equations and utilize the currently available theory of dynamic games. Then, in Section 3 we will present and discuss a procedure which would iteratively obtain the Nash equilibrium solution of the class of problems formulated in Section 2, and verify that it is informationally nonunique. This procedure will then be illustrated on a numerical example in Section 4, which will lead to a unique Nash equilibrium solution. The paper ends with the concluding remarks of Section 5.

2. A Class of Games with Second-Order Dynamics

Assume that a game process evolves according to the second-order difference equation

$$\begin{aligned} x_{k+2} &= f(k, x_{k+1}, x_k, u_k, v_k) + w_k \\ x_0, x_1 &\text{ given, } k=0,1,\dots \end{aligned} \quad (2.1)$$

where x_k is the n -dimensional state variable at the discrete time instant k , u_k and v_k are the r_1 and r_2 dimensional control variables of Players 1 and 2, respectively, and $\{w_k\}$ is a zero-mean i.i.d. (independent identically distributed) sequence of n -dimensional random variables with a probability distribution that assigns positive probability mass to every open subset of $\mathbb{R}^n$.

Let us take the information pattern for both players to be closed-loop:

$$\eta_k = \{x_k, x_{k-1}, \dots, x_1, x_0\}, \quad k=0, 1, \dots, \quad (2.2)$$

which generates a finite dimensional space N_k for each k . The strategies of Players 1 and 2 are, respectively,

$$\gamma_k : N_k \rightarrow \mathbb{R}^{r_1} ; \beta_k : N_k \rightarrow \mathbb{R}^{r_2} \quad (2.3)$$

at time instant k , which are Borel-measurable functions. Hence,

$$u_k = \gamma_k(\eta_k) ; v_k = \beta_k(\eta_k) . \quad (2.4)$$

Finally, the cost function of player i is taken to be

$$J_i(\gamma, \beta) = E\{q^i(x_k) + \sum_{k=0}^{K-1} g^i(k, x_k, u_k, v_k)\} \quad (2.5)$$

where u_k and v_k are given by (2.4),

$$\gamma \triangleq \{\gamma_0, \dots, \gamma_{K-1}\}, \beta \triangleq \{\beta_0, \dots, \beta_{K-1}\},$$

K is some positive integer (could be "infinite"), q^i and g^i are functionals which are continuously differentiable in their arguments (except k), and E denotes the expectation operator over the prior statistics of $\{w_k\}_{k=0}^{K-2}$.

Let us recall that a pair (γ^*, β^*) , belonging to an admissible class, is a Nash equilibrium if and only if the pair of inequalities

$$J_1(\gamma^*, \beta^*) \leq J_1(\gamma, \beta^*) \quad (2.6a)$$

$$J_2(\gamma^*, \beta^*) \leq J_2(\gamma^*, \beta) \quad (2.6b)$$

hold for all admissible γ and β .

Now, in order to utilize the available theory for first-order systems in order to obtain the solution of this class of dynamic games, one would immediately be tempted to reformulate this problem by introducing a $2n$ -dimensional variable

$$y_k = (y_k^1, y_k^2)' \quad (2.7)$$

where

$$y_k^1 = x_k \quad (2.8a)$$

$$y_k^2 = x_{k+1} \quad (2.8b)$$

and thus increasing the dimension of the state by a factor of two. Then, y_k satisfies the first-order difference equation:

$$\begin{cases} y_{k+1}^1 = y_k^2 & (2.9a) \\ y_{k+1}^2 = f(k, y_k^1, y_k^2, u_k, v_k) + w_k & (2.9b) \end{cases}$$

$$y_{k+1} = \bar{f}(k, y_k, u_k, v_k) + \begin{bmatrix} 0 \\ I \end{bmatrix} w_k \quad (2.10)$$

where the definition of $\bar{f}$ should be obvious. Furthermore, the cost functions could be expressed in terms of the new state variable:

$$J_1 = E\{\bar{q}^1(y_K) + \sum_{k=0}^{K-1} \bar{g}^1(k, y_k, u_k, v_k)\} \quad (2.11)$$

for some appropriate $\bar{q}^1$ and $\bar{g}^1$.

Even though the above formulation appears to be in the standard (first-order) form of a dynamic game, there are in fact two pitfalls (in disguise) which render the available theory inapplicable:

1) Because of the original information structure, the controls are not allowed to depend on all components of y , but only on the first block component y^1 , that is

$$u_k = \gamma_k(y_k^1, y_{k-1}^1, \dots, y_0^1)$$

and similarly for v_k . Hence, the original perfect state information problem has been turned (through reformulation) into a stochastic dynamic game with partial state information. Such problems are in general very difficult to solve, and currently there is no general theory which would be applicable in this framework.

2) The noise term in (2.10) does not directly affect all components of the state vector, and hence the sufficient (and generically necessary) condition for informationally unique equilibrium (cf. Ref. 1) is not satisfied. This indicates that, even though we were able to obtain a solution for the first-order model (stochastic dynamic game with partial state information) formulated in this section, we would not be able to conclude that there was no other solution which would also constitute a Nash equilibrium.

Thereby, we abandon the above first-order model, and seek to develop a method of derivation for the second-order model originally formulated in Section 2.

3. A Direct Iterative Method of Derivation and Existence of Informationally Unique Equilibrium

What we intend to show in this section is that the problem formulated in Section 2 indeed admits a solution which can be obtained (at least in principle) by a careful iterative argument. It turns out that there is no informational nonuniqueness, and the solution depends, in general, not only on the current values of the state but also on the entire past history.

Towards this end, let us first assume, without any loss of generality, that $g^i(K-1, \dots)$ in (2.5) depends only on x_{K-1} but not on u_{K-1} and v_{K-1} . Furthermore, let $(\bar{\gamma}, \bar{\beta})$ be a Nash equilibrium solution. With the dynamics evolved up to time $K-2$ under this set of equilibrium solutions, let us isolate the game from that point onwards and see how the dependence of $(\bar{\gamma}_{K-2}, \bar{\beta}_{K-2})$ is on the previously adopted policies. This new (reduced) game will have cost functions

$$J_i^{K-2}(\gamma_{K-2}, \beta_{K-2}) = E\{q^i(x_K) + g^i(K-1, x_{K-1}) + g^i(K-2, x_{K-2}, u_{K-2}, v_{K-2})\} \quad i=1,2 \quad (3.1)$$

in which x_K and x_{K-1} can be expressed through (2.1) in terms of $x_{K-2}, u_{K-2}, v_{K-2}, x_{K-3}, \bar{u}_{K-3}, \bar{v}_{K-3}, w_{K-2}$ and w_{K-3} . Such a substitution then leads to, for some h_{K-2}^i whose exact form will not be given here but can easily be determined,

$$J_1^{K-2}(\gamma_{K-2}, \beta_{K-2}) = E(h_{K-2}^1(x_{K-2}, u_{K-2}, v_{K-2}, x_{K-3}, \bar{u}_{K-3}, \bar{v}_{K-3}, w_{K-2}, w_{K-3})) \quad (3.2)$$

where

$$\bar{u}_{K-3} = \bar{\gamma}_{K-3}(\eta_{K-3}) ; \bar{v}_{K-3} = \bar{\beta}_{K-3}(\eta_{K-3})$$

and expectation is over the statistics of w_{K-2} and w_{K-3} .

Since x_{K-2} and x_{K-3} do not depend on w_{K-2} and w_{K-3} , we could first take expectation of (3.2) over w_{K-2} and w_{K-3} , and then minimize the resulting expression over u_{K-2} for $i=1$ and with v_{K-2} fixed, and over v_{K-2} for $i=2$ with u_{K-2} fixed. This leads to expressions of the form

$$\left. \begin{aligned} u_{K-2} &= \varphi_{K-2}^1(v_{K-2}, p_{K-2}) \\ v_{K-2} &= \varphi_{K-2}^2(u_{K-2}, p_{K-2}) \end{aligned} \right\} \quad (3.3)$$

where

$$p_{K-2} = (x_{K-2}, x_{K-3}, \bar{u}_{K-3}, \bar{v}_{K-3}) \quad (3.4)$$

Let us assume that the minimization problems above have, in fact, led to unique solutions, and let us further assume that the set of simultaneous equations (3.3) admits a unique solution:

$$\left. \begin{aligned} \hat{u}_{K-2} &= \hat{\gamma}_{K-2}(p_{K-2}) \\ \hat{v}_{K-2} &= \hat{\beta}_{K-2}(p_{K-2}) \end{aligned} \right\} \quad (3.5)$$

[Note that this uniqueness is "structural," but not necessarily "informational," since we are solving basically a static problem.]

Hence, the bottom line of this analysis and discussion is that if $(\bar{\gamma}, \bar{\beta})$ is a Nash equilibrium solution, we necessarily have the relationship

$$\left. \begin{aligned} \bar{\gamma}_{K-2}(\eta_{K-2}) &\equiv \hat{\gamma}_{K-2}(p_{K-2}) \\ \bar{\beta}_{K-2}(\eta_{K-2}) &\equiv \hat{\beta}_{K-2}(p_{K-2}) \end{aligned} \right\} \quad (3.6)$$

where $\bar{\gamma}, \bar{\beta}$ are as determined above (by (3.5)), p_{K-2} is given by (3.4) and

$$\left. \begin{aligned} \bar{u}_{K-3} &= \bar{\gamma}_{K-3}(\eta_{K-3}) \\ \bar{v}_{K-3} &= \bar{\beta}_{K-3}(\eta_{K-3}) \end{aligned} \right\} \quad (3.7)$$

Now the question is whether this solution is "informationally unique."

It would have been informationally nonunique if we were able to express x_{K-2} and/or x_{K-3} in terms of the values of the state variables at earlier stages [see the argument in Ref. 2 for the case of a first-order model]. But this is not possible here because of the presence of the noise term in the state equation (2.1). Hence, the structural form (3.6) is informationally unique at stage K-2. Of course, to complete the description of $(\bar{\gamma}_{K-2}, \bar{\beta}_{K-2})$ we still need expressions for $(\bar{\gamma}_{K-3}, \bar{\beta}_{K-3})$, which we do next.

We now substitute (3.6) into (3.2) to obtain

$$J_i^{K-2}(\bar{\gamma}_{K-2}, \bar{\beta}_{K-2}) = \alpha_{K-2}^i(x_{K-2}, x_{K-3}, \bar{u}_{K-3}, \bar{v}_{K-3}) \quad (3.8)$$

where α_{K-2}^i is some function with the given arguments. To determine the static stochastic game at stage K-3, we first start with

$$J_i^{K-3}(\gamma_{K-3}, \beta_{K-3}) = E\{\alpha_{K-2}^i(\cdot) + g^i(K-3, x_{K-3}, u_{K-3}, v_{K-3})\} \quad (3.9)$$

and then express x_{K-2} in terms of $x_{K-3}, x_{K-4}, \bar{u}_{K-4}, \bar{v}_{K-4}, w_{K-4}$, to obtain

$$J_i^{K-3}(\gamma_{K-3}, \beta_{K-3}) = E\{h_{K-3}^i(x_{K-3}, u_{K-3}, v_{K-3}, \bar{u}_{K-3}, \bar{v}_{K-3}, x_{K-4}, \bar{u}_{K-4}, \bar{v}_{K-4}, w_{K-4})\} \quad (3.10)$$

where

$$\bar{u}_{K-4} = \bar{\gamma}_{K-4}(\eta_{K-4}) ; \bar{v}_{K-4} = \bar{\beta}_{K-4}(\eta_{K-4})$$

and expectation is over the statistics of w_{K-4} . Note the presence of also $\bar{u}_{K-3}$ and $\bar{v}_{K-3}$ in (3.10) which come from α_{K-2}^i .

Now we solve for the Nash equilibrium solution of the static game described by (3.10) after averaging over w_{K-4} . Only u_{K-3} and v_{K-3} are the variables here, and assuming that a unique Nash equilibrium solution to this static game exists, it will be in the form,

$$\left. \begin{aligned} u_{K-3} &= \tilde{\gamma}_{K-3}(p_{K-3}, \bar{u}_{K-3}, \bar{v}_{K-3}) \\ v_{K-3} &= \tilde{\beta}_{K-3}(p_{K-3}, \bar{u}_{K-3}, \bar{v}_{K-3}) \end{aligned} \right\} \quad (3.11)$$

where

$$p_{K-3} = (x_{K-3}, x_{K-4}, \bar{u}_{K-4}, \bar{v}_{K-4}) \quad (3.12)$$

Because of consistency, we should have $u_{K-3} = \bar{u}_{K-3}$, $v_{K-3} = \bar{v}_{K-3}$, and hence using this in (3.11) and solving for $(\bar{u}_{K-3}, \bar{v}_{K-3})$ we obtain, for some appropriate $\tilde{\gamma}_{K-3}, \tilde{\beta}_{K-3}$,

$$\left. \begin{aligned} \bar{u}_{K-3} &= \tilde{\gamma}_{K-3}(p_{K-3}) \\ \bar{v}_{K-3} &= \tilde{\beta}_{K-3}(p_{K-3}) \end{aligned} \right\} \quad (3.13)$$

which replaces (3.5). Therefore, if $(\bar{\gamma}, \bar{\beta})$ is a Nash equilibrium solution, at stage K-3 we necessarily have the relationship

$$\left. \begin{aligned} \bar{\gamma}_{K-3}(\gamma_{K-3}) &= \tilde{\gamma}_{K-3}(p_{K-3}) \\ \bar{\beta}_{K-3}(\gamma_{K-3}) &= \tilde{\beta}_{K-3}(p_{K-3}) \end{aligned} \right\} \quad (3.14)$$

Furthermore, this is the informationally unique Nash equilibrium because of the argument made in the derivation at stage K-2. The conclusion then is that the Nash equilibrium solution at stage K-3 depends, in general, on x_{N-3} , x_{N-2} , and the values of state at earlier stages, through $\bar{\gamma}_{K-4}$ and $\bar{\beta}_{K-4}$. This, in turn, implies for stage K-2, using (3.6), that $\bar{\gamma}_{K-2}$ and $\bar{\beta}_{K-2}$ depend, in general, on $x_{K-2}, x_{K-3}, x_{K-4}$ and the values of the state at earlier stages through $\bar{\gamma}_{K-4}$ and $\bar{\beta}_{K-4}$.

This procedure, applied at stage K-3, can be followed up for other stages in a descending order and iteratively, to obtain finally $(\bar{\gamma}_0, \bar{\beta}_0)$, the pair of equilibrium strategies at time K=0. Then, moving in forward time, the entire set of equilibrium strategies can be determined by using relations like (3.14) and (3.6) through recursive substitution. This entire procedure will be illustrated in the next section, using a numerical example.

The conclusions that can be drawn at this point are that: (i) the dynamic game of Section 2 generically admits an equilibrium solution; (ii) this equilibrium solution can be derived (in principle) using an iterative procedure which sweeps the time interval twice (once forwards and once backwards) — in that sense the method is in the spirit of the one employed in [3] in a different context; (iii) the equilibrium solution does not exhibit informational nonuniqueness.

4. A Numerical Example

As an illustration of the procedure presented in the previous section, we consider here a 4-stage dynamic game with scalar dynamics

$$x_{k+2} = 2x_{k+1} - x_k + u_k + v_k + w_k$$

where x_0 and x_1 are specified a priori, and $\{w_k\}$ are zero-mean i.i.d. Gaussian random variables. The information structure is closed-loop perfect state for both players and the cost functions are

$$\begin{aligned} J_1 &= E\{[x_4]^2 + \sum_{k=0}^2 [u_k]^2\} \\ J_2 &= E\{[x_4]^2 + \sum_{k=0}^2 [v_k]^2\} \end{aligned}$$

Let $\{\bar{u}_0 = \bar{\gamma}_0(x_0), \bar{u}_1 = \bar{\gamma}_1(x_0, x_1), \bar{u}_2 = \bar{\gamma}_2(x_0, x_1, x_2)$

$$\bar{v}_0 = \bar{\beta}_0(x_0), \bar{v}_1 = \bar{\beta}_1(x_0, x_1), \bar{v}_2 = \bar{\beta}_2(x_0, x_1, x_2)\}$$

be a Nash equilibrium solution for this game. Then, following the procedure of Section 3, we have:

Stage 2: Holding $\bar{u}_0, \bar{u}_1, \bar{v}_0, \bar{v}_1$ fixed as given, for $\{\bar{u}_2, \bar{v}_2\}$ to be in equilibrium with these it is necessary and sufficient (sufficiency follows from strict convexity of J_1 and J_2) that

$$\left. \begin{aligned} \frac{\partial}{\partial u_2} \{E\{[x_4]^2 + [u_2]^2\} &= 0 \\ \frac{\partial}{\partial v_2} \{E\{[x_4]^2 + [v_2]^2\} &= 0 \end{aligned} \right\} \quad (4.1)$$

at $(\bar{u}_2, \bar{v}_2)$.

Since, $x_3 = 2x_2 - x_1 + \bar{u}_1 + \bar{v}_1 + w_1$

$$x_4 = 2x_3 - x_2 + u_2 + v_2 + w_2 = 3x_2 - 2x_1 + 2\bar{u}_1 + 2\bar{v}_1 + u_2 + v_2 + w_2 + 2w_1,$$

using the fact that $\{w_1, w_2\}$ is an independent sequence, we obtain for (4.1)

$$u_2 = -(1/2)[3x_2 - 2x_1 - 2\bar{u}_1 - 2\bar{v}_1] - (1/2)v_2$$

$$v_2 = -(1/2)[3x_2 - 2x_1 - 2\bar{u}_1 - 2\bar{v}_1] - (1/2)u_2,$$

and solving for $\{u_2, v_2\}$ we further obtain the unique relationship (which is the counterpart of (3.5)):

$$\left. \begin{aligned} \hat{u}_2 &= \bar{\gamma}_2(x_1, x_2, \bar{u}_1, \bar{v}_1) = -(1/3)[3x_2 - 2x_1 - 2\bar{u}_1 - 2\bar{v}_1] \\ \hat{v}_2 &= \bar{\beta}_2(x_1, x_2, \bar{u}_1, \bar{v}_1) = -(1/3)[3x_2 - 2x_1 - 2\bar{u}_1 - 2\bar{v}_1] \end{aligned} \right\} \quad (4.2)$$

Hence, if $(\bar{\gamma}, \bar{\beta})$ is a Nash equilibrium, we necessarily have at stage $k=2$,

$$\left. \begin{aligned} \bar{\gamma}_2(x_2, x_1, x_0) &\equiv \bar{\gamma}_2(x_1, x_2, \bar{u}_1, \bar{v}_1) \\ \bar{\beta}_2(x_2, x_1, x_0) &\equiv \bar{\beta}_2(x_1, x_2, \bar{u}_1, \bar{v}_1) \end{aligned} \right\} \quad (4.3)$$

and this is also a unique representation because we cannot express x_2 in terms of x_1 and x_0 without introducing an error. Note that we still have to determine $\{\bar{u}_1, \bar{v}_1\}$ to complete the description of $\bar{\gamma}_2$ and $\bar{\beta}_2$.

Stage 1. Here we take u_2 and v_2 as given by (4.2) and hold $(\bar{u}_0, \bar{v}_0)$ fixed. Then, to obtain expressions for $\bar{\gamma}_1$ and $\bar{\beta}_1$, we substitute (4.2) into the state equation for x_4 and x_3 and x_2 , and differentiate the resulting expressions for J_1 and J_2 with respect to u_1 and v_1 , respectively. This leads to expressions for u_1 and v_1 in terms of $x_1, x_0, \bar{u}_1, \bar{v}_1, \bar{u}_0, \bar{v}_0$ [this is the counterpart of (3.11)], and requiring consistency in the solution we let $u_1 = \bar{u}_1, v_1 = \bar{v}_1$, to obtain the unique solution (as counterpart of (3.13))

$$\left. \begin{aligned} \hat{u}_1 &= \bar{\gamma}_1(x_1, x_0, \bar{u}_0, \bar{v}_0) = - (6/43)[(4/3)x_1 - x_0 + \bar{u}_0 + \bar{v}_0] \\ \hat{v}_1 &= \bar{\beta}_1(x_1, x_0, \bar{u}_0, \bar{v}_0) = - (6/43)[(4/3)x_1 - x_0 + \bar{u}_0 + \bar{v}_0] \end{aligned} \right\} \quad (4.4)$$

The uniqueness here follows again from strict convexity of J_1 and J_2 . Hence, if $(\bar{\gamma}, \bar{\beta})$ is a Nash equilibrium solution, we necessarily have

$$\begin{aligned}\bar{\gamma}_1(x_1, x_0) &\equiv \tilde{\gamma}_1(x_1, x_0, \bar{u}_0, \bar{v}_0) \\ \bar{\beta}_1(x_1, x_0) &\equiv \tilde{\beta}_1(x_1, x_0, \bar{u}_0, \bar{v}_0)\end{aligned}\tag{4.5}$$

and this is a unique representation because x_1 does not depend on x_0 .

Note that still this is not in final form because of dependence on the Nash equilibrium policies at stage 0, $(\bar{u}_0, \bar{v}_0)$. Though, one conclusion we can arrive at here is that if $(\bar{u}_0, \bar{v}_0)$ is unique, then $(\bar{\gamma}_1, \bar{\beta}_1)$ will be unique via (4.5) and (4.4), and in turn $(\bar{\gamma}_2, \bar{\beta}_2)$ will be unique via (4.3) and (4.2).

Stage 0. To obtain the expressions for $(\bar{u}_0, \bar{v}_0)$, we take (u_1, v_1) as given by (4.4), and (u_2, v_2) as given by (4.2), with $(\bar{u}_1, \bar{v}_1)$ substituted from (4.4); then we evaluate x_4 , u_2 and v_2 to obtain

$$\begin{aligned}x_4 &= (4/43)x_1 - (3/43)x_0 - (40/43)[\bar{u}_0 + \bar{v}_0] + u_0 + v_0 + w_2 + 2w_1 + w_0 \\ u_2 = v_2 &= -(68/43)x_1 + (51/43)x_0 - u_0 - v_0 - w_0 - (8/43)[\bar{u}_0 + \bar{v}_0]\end{aligned}$$

Then, performing the optimizations

$$\min_{u_0} E\{[x_4]^2 + [u_2]^2 + [u_1]^2 + [u_0]^2\}$$

$$\min_{v_0} E\{[x_4]^2 + [v_2]^2 + [v_1]^2 + [v_0]^2\}$$

we obtain the unique relationships (note that u_1 does not depend on u_0 , and v_1 does not depend on v_0):

$$\left. \begin{aligned}2(x_4 - u_2) + 2u_0 &= 0 \\ 2(x_4 - v_2) + 2v_0 &= 0\end{aligned} \right\}$$

$$\begin{aligned}(72/43)x_1 - (54/43)x_0 + 3u_0 + 2v_0 - (32/43)[\bar{u}_0 + \bar{v}_0] &= 0 \\ (72/43)x_1 - (54/43)x_0 + 3v_0 + 2u_0 - (32/43)[\bar{u}_0 + \bar{v}_0] &= 0\end{aligned}$$

For consistency, setting $u_0 = \bar{u}_0$, $v_0 = \bar{v}_0$, we obtain the unique solution of this set of simultaneous equations to be

$$\left. \begin{aligned} u_0^* &= \gamma_0^*(x_0, x_1) = (6/61)(3x_0 - 4x_1) \\ v_0^* &= \beta_0^*(x_0, x_1) = (6/61)(3x_0 - 4x_1) \end{aligned} \right\} \quad (4.6)$$

which are the unique Nash policies at stage 0.

This, then, completes derivation in retrograde time. We now have to sweep the stages in forward time to complete the expressions for the Nash policies at other stages. Towards this end, we first use (4.6) in (4.4) for $(\bar{u}_0, \bar{v}_0)$, to arrive at

$$\left. \begin{aligned} u_1^* &= \gamma_1^*(x_0, x_1) = [150/(43)(61)][x_0 - 4x_1] \\ v_1^* &= \beta_1^*(x_0, x_1) = [150/(43)(61)][x_0 - 4x_1] \end{aligned} \right\} \quad (4.7)$$

Finally, using (4.7) in (4.2) for $(\bar{u}_1, \bar{v}_1)$ we obtain

$$\left. \begin{aligned} u_2^* &= \gamma_2^*(x_0, x_1, x_2) = -x_2 + (2/2623)[(2023/3)x_1 + 50x_0] \\ v_2^* &= \beta_2^*(x_0, x_1, x_2) = -x_2 + (2/2623)[(2023/3)x_1 + 50x_0] \end{aligned} \right\} \quad (4.8)$$

Hence, $\gamma^* = (\gamma_0^*, \gamma_1^*, \gamma_2^*)$; $\beta^* = (\beta_0^*, \beta_1^*, \beta_2^*)$, as given by (4.6)-(4.8) constitutes the unique Nash equilibrium solution for the dynamic game (with second-order state equation) formulated in this section. It should be noted that the equilibrium policies use complete memory — not only the current values of the state. A second observation to be made is that the unique solution for the linear-quadratic game is linear in the available state information. Some scrutiny reveals that this is in fact a property that is shared by all linear-quadratic dynamics games (that is, for games for which, in the framework of the general formulation of Section 2, f is linear, and g^i, q^i are quadratic). In other words, for linear-quadratic dynamic games with second (or higher) order

dynamics, and with the additive noise satisfying the properties elucidated in Section 2, the Nash equilibrium solution will generically exist, will be unique and linear, and will depend not only on the current value of the state but also on memory. A precise verification of this result is notationally cumbersome, but it follows from an otherwise routine application of the procedure of Section 3 to linear-quadratic games.

5. Concluding Remarks

In this paper we have addressed a class of noncooperative dynamic games which have not been treated before and to which the currently available theory does not apply. This class involves discrete-time game models with state dynamics of second order, and with additive noise disturbance satisfying certain regularity conditions. It has been shown, through an intricate set of arguments, that the Nash equilibrium solution will be informationally unique, whenever it exists, and can be obtained by an iterative procedure which sweeps the time interval twice. In the case of linear-quadratic games and under the closed-loop information pattern for all players, the informationally unique Nash solution is linear in the current and past values of the state, and can be obtained explicitly.

An obvious, but challenging, extension of the theory presented here would be to continuous-time problems with second or higher order dynamics and subject to additive noise with independent increments. It is anticipated that the procedure developed here for the discrete time problem would have a natural counterpart in this case, with the structural properties of the equilibrium solution as presented here remaining intact.

6. References

- [1] T. Başar and G. J. Olsder, Dynamic Noncooperative Game Theory, Academic Press, London/New York, 1982.
- [2] T. Başar, "On the Uniqueness of the Nash Solution in Linear-Quadratic Differential Games," *International Journal of Game Theory*, Vol. 5, No. 2/3, pp. 65-90, 1976.
- [3] T. Başar, "Decentralized Multicriteria Optimization of Linear Stochastic Systems," *Special Issue of IEEE Transactions on Automatic Control on "Decentralized Control and Large Scale Systems,"* Vol. AC-23, No. 2, pp. 233-243, April 1978.

THEORY OF DYNAMIC GAMES AND ITS APPLICATIONS IN LARGE SCALE SYSTEMS DESIGN AND OPTIMIZATION"

Tamer Başar

Decision and Control Laboratory, Coordinated Science Laboratory, University of Illinois,
1101 W. Springfield Avenue, Urbana, Illinois 61801, U.S.A.

Abstract. In this paper we first provide a brief review of some of the recent results on dynamic games, in particular with regard to memory strategies, and then discuss potential applications of the techniques developed in this context to large scale systems design, optimization and coordination. Incorporation of memory strategies in the optimization of interconnected systems enables one to enforce satisfiability of more than one criterion and to treat coordination problems wherein the lower levels have different perceptions of the underlying model and goals. A number of such design criteria are introduced, and recipes are given to obtain coordinator policies with good sensitivity properties.

Keywords. Large scale systems, interconnected systems, dynamic games, coordination, minimum sensitivity analysis.

1. INTRODUCTION

This paper discusses the role of memory strategies in the optimization and robustness considerations of interconnected systems controlled by several decision makers and under possibly different performance measures. It is a known fact in control theory that the use of control laws which incorporate also the past values of the state vector lead to better system performance in terms of robustness, and minimum sensitivity to changes in the nominal values of the system parameters and to external disturbances. Even in linear deterministic systems with partially unknown parameters, or with parameters which are slowly varying (drifting), use of higher order compensators (instead of pure feedback control laws) lead to more acceptable controller designs (Kokotovic, et al, 1968; Sundarajan and Cruz, 1972) with better sensitivity properties in case of small deviations from the nominally adopted model. Inclusion of also the past values of the state variables in the controller design procedure brings in redundancy in information which, if used judiciously, leads to considerable improvement in overall system performance.

Parallel to the centralized case, incorporation of memory in controller design also gains paramount importance in the optimization, decentralized control and coordination of interconnected systems. The objective of this paper is to discuss this role of memory policies in multi-station multi-goal optimization and control problems, and to convey the reader the important message that sensitivity analysis finds a natural place and extension in such systems. The paper is of tutorial nature and emphasizes more the underlying concepts and methodology rather than derivation of specific results. However, the reader could use the recipes given in the paper to obtain robust, minimally

sensitive control policies for interconnected systems with specific structures.

Crucial in this development, is the role of memory policies in dynamic games, particularly with regard to existence and uniqueness of various types of equilibrium. For this reason, the next section provides a brief description of and a perspective on dynamic games, with emphasis on recent developments and informational redundancy. Section 3 discusses ways of utilizing informational redundancy in obtaining robust coordinator policies with appealing sensitivity properties in the face of modelling inaccuracies. Section 4 introduces a different type of discrepancy and evaluates memory policies from that perspective. The paper ends with the concluding remarks of Section 5.

2. A GAME-THEORETIC MODEL AND SOME SALIENT CHARACTERISTICS

To fix the ideas and to have a unifying framework to work in, let us consider an n -dimensional system described in continuous time and controlled by three different stations:

$$\dot{x} = f(x, u, v, w, t) ; x(t_0) = x_0 ; t \in [t_0, t_f] \quad (2.1)$$

Here u, v, w are the control vectors of stations or suboptimally, decision makers or players in the terminology of game theory (1, 2 and 3, respectively), taking values in some Euclidean spaces of appropriate dimensions and satisfying some smoothness conditions, such as piecewise continuity in t . Furthermore, f is some possibly vector-valued parameter, representing the modelling uncertainty from the viewpoint of the stations. For the present we assume that f depends on x, t , and suppress the dependence on u, v, w in (2.1).

Research on this paper was supported in part by the Office of Naval Research, Grant N00014-80-1-0700, and in part by the U.S. Department of Energy, Office of Biological and Environmental Research, Grant DE-AC02-80-OR21400.

The author would like to thank Dr. J. B. Clewley for his helpful comments on an earlier draft of this paper, and Dr. R. S. Sutton for his helpful comments on the manuscript.

$$u(t) = u_1(t, x_1(t)), \quad v(t) = v_2(t, x_2(t)), \\ w(t) = w_3(t, x_3(t)),$$

where we assume that u_1 is sufficiently smooth so that (2.1) admits a unique solution for each triple (u_1, v_2, w_3) .

Now, if the stations have possibly different objective functions, say

$$J_i(u_1, v_2, w_3) = h_i(x(t_f)) + \int_{t_0}^{t_f} g_i[x(t), u(t), v(t), w(t)] dt \quad (2.2) \\ i=1,2,3,$$

where t_f denotes the terminal time, a relevant non-cooperative solution concept in this context is the so-called Nash equilibrium $y^* = (y_1^*, y_2^*, y_3^*)$ which satisfies

$$J_i(y^*) \leq J_i(y_i^*) \text{ for all admissible } y_i \quad (2.3) \\ i=1,2,3$$

where

$$y_i^* = \begin{cases} (u_1^*, v_2^*, w_3^*) & i=1 \\ (u_1^*, v_2^*, w_3^*) & i=2 \\ (u_1^*, v_2^*, w_3^*) & i=3 \end{cases}$$

For the case when all J_i 's are the same function, and y_i^* 's are equivalent, this problem becomes equivalent to a centralized optimal control problem (with the three stations considered as a single station), and (2.3) represents in this case a weaker version of the (team) optimality condition.

One of the important results of dynamic game theory says that if the information pattern $\pi = (\pi_1, \pi_2, \pi_3)$ incorporates some memory (i.e., redundant state information) for at least one station, then in general the Nash solution y^* is nonunique (in fact infinitely many), leading to infinitely many possible equilibrium cost triples $(J_i(y^*), i=1,2,3)$ whenever J_i 's are different [see Başar, 1977; Başar and Olsder, 1982]. Hence it seems that, at least at the outset, presence of redundancy in the available information leads to ambiguity in the solution of the game problem, making the entire analysis totally worthless. Fortunately, however, this is not the entire story, and the situation could be salvaged by introducing an element of coordination into the model. In fact, it will turn out that this nonuniqueness of equilibria (so-called "informational nonuniqueness") is a blessing in disguise, which can be used to considerable advantage if the model is set up properly.

Towards that end, let us now endow one of the stations, say station 1, with additional power or authority, to see that (again in a decentralized framework) an acceptable set of controllers are chosen by all stations, satisfying, perhaps, some pre-set conditions. With such an asymmetry in the roles of the stations (decision makers) we call station 1 the leader (or the coordinator) and stations 2 and 3 the followers, adopting the terminology of Stackelberg games. Here, station 1 announces a policy (a control law $u_1(\cdot)$) to which the other stations respond by minimizing their cost functions J_2 and J_3 , for stations 2 and 3, respectively. Anticipating these responses, the leader decides on a policy (which could also be called the "coordinating policy") which leads to an overall satisfactory performance. Such a policy, provided that it incorporates memory, could lead to a well-posed optimization problem at the lower level and also remove the inefficiency of Nash equilibria. Let us now elaborate on these two points.

For each arbitrarily announced memory policy of the leader, the two stations at the lower level will be faced with a Nash game which could still be ill-posed, be due to ambiguity in the dynamic Nash equilibria, when the control laws involve

memory. However, if the memory policy of the leader is chosen judiciously, it could break the Nash game into two independent (decoupled) optimization problems (one for each station), thereby eliminating the informational nonuniqueness associated with Nash equilibria. This is possible because whenever the leader's policy incorporates memory, he is in general able to "control" the structure of the cost functions faced by the other two stations without incurring degradation in his own performance [Başar and Selbuz, 1979; Ho, Luh and Muralidharan, 1981].

(ii) Utilization of an appropriate memory strategy by the leader could also lead (together with the other two stations' optimal response policies) to a Pareto efficient solution, which could be defined, for example, as the minimization of a convex combination of the two stations' cost functions. This property has been established in the literature in different frameworks, when at the lower level there are two stations [see Başar, 1980] or a single station [Başar and Selbuz, 1979; Papavassilopoulos and Cruz, 1979; Tolwinski, 1981].

Optimum coordinating memory policies referred to above are not in general unique, even though each one leads to the same performance from the coordinating station's point of view. Hence, the issue here is not that there is still ambiguity in the proposed solution (or that the problem with the coordinator is ill-posed), but that this nonuniqueness brings in additional degrees of freedom which could be used to our advantage in improving the overall performance of the system in case of inaccuracies in system modeling and/or goal perceptions. The former type of inaccuracy is discussed in the next section, and the latter type is discussed in Section 4.

3. COORDINATING POLICIES WITH OPTIMUM SENSITIVITY PROPERTIES: MODELLING INACCURACIES

Let us now go back to the general model (2.1) where λ is an unknown parameter vector, with the uncertainty being in a small neighborhood of a nominal value λ^0 . Now, with λ fixed at λ^0 , and known by all parties, let us denote the class of all optimal coordinator policies u_1 by $\mathcal{U}^0$ which will in general have infinitely many elements when policies are allowed to depend also on the past values of the state. Hence, for all $u_1 \in \mathcal{U}^0$, the dynamic Nash game described by the cost functionals (2.2):

$$J_2(u_1, v_2, w_3), J_3(u_1, v_2, w_3) \quad (3.1)$$

and state equation

$$\dot{x} = f(x, u_1, v_2, w_3, \lambda^0); x(t_0) = x_0, \quad t \in [t_0, t_f] \quad (3.2)$$

will be well-defined, and its solution will be the same regardless of which element is chosen out of $\mathcal{U}^0$. Therefore, as far as the nominal model goes, $\mathcal{U}^0$ constitutes the solution set for the coordinator. [Note that this discussion also pertains to the case when J_2 and J_3 are identical, that is when we have a team problem at the lower level.] Now, we are faced with the secondary (but important) problem of introducing an additional optimality criterion on $\mathcal{U}^0$ so as to absorb the freedom embodied in the choice of optimal policies out of $\mathcal{U}^0$. This optimality criterion is introduced as follows:

We have earlier mentioned that λ will in general take values in a neighborhood of λ^0 , and hence the system model is actually not unique but belongs to a certain class of models. When faced with the constraint that know the actual model, and the lower level controllers have the same access to this information, the coordinator has to choose a policy from the optimal set $\mathcal{U}^0$ that is insensitive to

that the coordinator may broadcast the nominal value x_0 to the lower level controllers, but because of noise in the transmission the actual value received by them may show a small variation from x_0 , say $x_0 + \Delta x = \bar{x}$. Let us now pursue this discussion somewhat further and study the effect of such a deviation from the nominal model on the nominal equilibrium solution, when both lower level controllers perceive the underlying model as described by

$$\dot{x} = f(x, \bar{v}_1(\bar{\tau}_1), v, w, \bar{z}); x(t_0) = x_0 \quad (3.3)$$

where $\bar{v}_1 = v_1^0 + \Delta v_1$, and $\bar{\tau}_1 \in \bar{\tau}_1^0$ is an announced policy of the coordinator. (Note that here both controllers use the same (albeit different from the nominal) model. The case when $\bar{\tau}_1$ is different for two controllers can also be studied, but the analysis is more involved.)

With the cost functions as described by (3.1), and system model by (3.3), we have a two-person Nash game whose solution can be obtained, under the given information pattern, using the available theory [Basar and Olsder, 1982], which will in general be nonunique if the dynamic information at the lower level involves memory and J_2 is intrinsically different from J_3 . However, if $\bar{\tau}_1$ is chosen judiciously, the lower level Nash game will be well-defined informationally, admitting a unique solution for each adopted information pattern, as discussed earlier in Section 2. Hence, we assume that $\bar{\tau}_1$ is in fact such a class of policies for the coordinator (this is indeed a rich class, in general), which will trivially be the case if the lower level problem is a team (i.e., $J_2 \equiv J_3$). The solution at this lower level will in general depend on different choices for $\bar{\tau}_1 \in \bar{\tau}_1^0$ and the value $\bar{x}$. To indicate this implicit dependence, let us denote the optimal solution at the lower level by $(\bar{v}_2(\bar{\tau}_1, \bar{x}, \bar{z}), \bar{v}_3(\bar{\tau}_1, \bar{x}, \bar{z}))$. When this pair of policies is substituted into the state equation (3.2) for the nominal model, and the resulting equation solved, we arrive at a trajectory which again will vary with $\bar{\tau}_1$ and $\bar{x}$, with the dependence on $\bar{x}$ being only through $\bar{v}_2$ and $\bar{v}_3$. Finally, when this trajectory, together with $\bar{v}_1, \bar{v}_2, \bar{v}_3$ are substituted into J_1 , the coordinator's cost functional, the resulting expression for J_1 will depend on $\bar{\tau}_1$ and the choice $\bar{x}$ out of $\bar{\tau}_1^0$; let us denote this by

$$F(\bar{\tau}_1, \bar{x}) = J_1(\bar{v}_1(\bar{\tau}_1), \bar{v}_2(\bar{\tau}_1, \bar{x}, \bar{z}), \bar{v}_3(\bar{\tau}_1, \bar{x}, \bar{z})). \quad (3.4)$$

One important property of F is that, when $\bar{x} = x_0$, it is independent of $\bar{\tau}_1$; that is, $F(\bar{\tau}_1, x_0)$ is a constant over $\bar{\tau}_1^0$. [Because, as indicated earlier, when the same nominal model is used at all levels, every policy of the coordinator out of $\bar{\tau}_1^0$ leads to the same performance and the same trajectory.]

Since F varies with $\bar{\tau}_1$ over $\bar{\tau}_1^0$, a natural question to ask here is "What is the best choice out of $\bar{\tau}_1^0$ which renders the sensitivity of optimum J_1 to changes in the value of $\bar{x}$ from the nominal value x_0 minimum?" Since, by construction,

$$F(\bar{\tau}_1, \bar{x}) \geq F(\bar{\tau}_1, x_0), \quad \forall \bar{\tau}_1 \in \bar{\tau}_1^0, \quad (3.5)$$

this question can be rephrased in mathematical terms as

$$\min_{\bar{\tau}_1 \in \bar{\tau}_1^0} F(\bar{\tau}_1, \bar{x}), \quad \bar{\tau}_1 \in N_{\bar{\tau}_1^0}(x_0) \quad (3.6)$$

where $N_{\bar{\tau}_1^0}(x_0)$ denotes a given ϵ -neighborhood of x_0 . Since ϵ is not fixed, (3.6) is still ambiguous, and therefore, we have to adopt either a worst case analysis:

$$\min_{\bar{\tau}_1 \in \bar{\tau}_1^0} \sup_{\bar{x} \in N_{\bar{\tau}_1^0}(x_0)} F(\bar{\tau}_1, \bar{x}) \quad (3.7)$$

or an "incremental deviation sensitivity" approach which is more in line with the approach taken in the references cited in Section 1 for single station control problems.

In this latter approach we minimize a weighted first-order or second-order sensitivity function associated with F and with regard to the parameter $\bar{x}$ around the nominal value x_0 . If $\bar{x} \in \mathbb{R}^n$, then the first order differential is:

$$\partial F(\bar{\tau}_1, \bar{x}) = \sum_{i=1}^n [dF(\bar{\tau}_1, \bar{x})/d\bar{x}_i]_{\bar{x}=x_0} (\bar{x}_i - x_{i0}) \quad (3.8)$$

Generally this expression will either be zero, or be nonzero but independent of $\bar{\tau}_1$, so that we will have to consider the next leading term:

$$\partial^2 F(\bar{\tau}_1, \bar{x}) = \sum_{j=1}^n \sum_{i=1}^n [d^2 F(\bar{\tau}_1, \bar{x})/d\bar{x}_i d\bar{x}_j]_{\bar{x}=x_0} (\bar{x}_i - x_{i0})(\bar{x}_j - x_{j0}) \quad (3.9)$$

which will explicitly depend on $\bar{\tau}_1$. Then, a meaningful criterion will be the minimization of a suitable norm of the nonnegative definite matrix

$$\Delta(\bar{\tau}_1) = [d^2 F(\bar{\tau}_1, \bar{x})/d\bar{x}_i d\bar{x}_j]_{i,j=1,\dots,n}, \quad \bar{x} = x_0 \quad (3.10a)$$

over $\bar{\tau}_1 \in \bar{\tau}_1^0$. Let

$$\bar{\tau}_1^* = \arg \min_{\bar{\tau}_1 \in \bar{\tau}_1^0} \|\Delta(\bar{\tau}_1)\| \quad (3.10b)$$

where $\|\cdot\|$ denotes a suitable matrix norm. Feasibility of this minimization problem will, of course, depend on the general structures of F and $\bar{\tau}_1^0$, which will in turn depend on the structures of the cost functions J_1, J_2, J_3 , and the function f which characterizes the state equation; specific results could be obtained by assuming specific structures for f, h_1 and g_1 in (2.1)-(2.2). What is true independent of what the structures of these solutions are, however, is the fact that $\bar{\tau}_1^*$ as defined by (3.10b) serves two important roles as an optimum coordinator strategy:

(i) It ensures that when the nominal model is adopted by all stations, a certain acceptable performance is attained at the upper level, and the optimization problem faced by the lower level stations is well-posed and even decoupled.

(ii) If the lower level stations deviate from the nominal model when computing their optimum response controls, the effect of this deviation on the performance at the upper level is minimal.

This, of course, is all possible provided that the coordinator is allowed to use memory policies.

In the next section, we will observe a similar effective role played by coordinator policies in the context of a different type of system inaccuracy.

4. COORDINATING POLICIES WITH OPTIMUM SENSITIVITY PROPERTIES: INACCURACIES IN GOAL PERCEPTIONS

Consider the problem formulation of Section 2, but with $J_1 = J_2 = J_3 \equiv J$ (nominally), and $x = x^0$ a fixed value known by all stations. This is then, basically an optimal control problem, and under appropriate convexity conditions the optimizing control policies $(\bar{v}_1(\bar{\tau}_1), \bar{v}_2(\bar{\tau}_1), \bar{v}_3(\bar{\tau}_1))$ will be unique as far as their open-loop values go, but nonunique otherwise — in other words, when the underlying information pattern is dynamic, we have equivalence classes $\bar{v}_i^0, \bar{v}_j^0$ with the property that for any $\bar{v}_i^0, \bar{v}_j^0 \in \bar{v}_i^0, \bar{v}_j^0$

$$J(\bar{v}_i^0, \bar{v}_j^0) = \min_{\bar{v}_i^0, \bar{v}_j^0 \in \bar{v}_i^0, \bar{v}_j^0} J(\bar{v}_i^0, \bar{v}_j^0) \quad (3.11)$$

Hence, here optimality does not imply freedom in the choice of optimal control laws, because of restrictions in the available information, which can be used to determine an

robustifying the overall performance against inaccuracies or discrepancies in perceptions of the stations — provided that one station is given a superior role in coordinating the policies of the other stations. In the previous section we have allowed for inaccuracies in the modeling of the state equation, and have argued that it is possible to find an optimally coordinating policy which also renders the overall performance minimally sensitive or insensitive to deviations in the state model from the nominal. Note that if the nominal problem is taken as a team problem (with identical cost functions for all stations), derivation of minimally sensitive policies requires consideration of a dynamic game, because the inaccuracies in modeling cannot be handled in the framework of team problems. Likewise, when the inaccuracy is in the goal perception of the lower-level stations (which will be the topic in this section), derivation of minimally sensitive policies asks for a game theoretic analysis.

Towards this end, let $\hat{s}$ be a parameter with nominal value s^0 , affecting directly only the cost functions but not the state equation; that is, $J_i = J_i(\gamma_1, \gamma_2, \gamma_3, \hat{s})$, with the further property that nominally,

$$J_i(\gamma_1, \gamma_2, \gamma_3, s^0) \equiv J(\gamma_1, \gamma_2, \gamma_3), \quad i=1,2,3 \quad (4.2)$$

with J being the team cost function used in (4.1). Now, if the lower level stations had the same (common) perception of the common cost function (which corresponds to $\hat{s} = s^0$), every triple in $\Gamma_1^0 \times \Gamma_2^0 \times \Gamma_3^0$ would constitute an optimal solution. If the lower level stations have a different perception of J , this discrepancy being quantified in the value of $\hat{s}$ (say $\hat{s}^2$ for station 2 and $\hat{s}^3$ for station 3), the optimization problems at the lower levels do not necessarily admit solutions in $\Gamma_2^0 \times \Gamma_3^0$, that is

$$\arg \min_{\Gamma_2^0} J_2(\gamma_1, \gamma_2, \gamma_3, \hat{s}^2) \neq \emptyset \quad (4.2a)$$

$$\arg \min_{\Gamma_3^0} J_3(\gamma_1, \gamma_2, \gamma_3, \hat{s}^3) \neq \emptyset \quad (4.2b)$$

with $(\gamma_1, \gamma_2, \gamma_3) \in \Gamma_1^0 \times \Gamma_2^0 \times \Gamma_3^0$. Hence, a relevant question here is whether there exists a $\gamma_1^* \in \Gamma_1^0$ with the properties that

$$1) \min_{\Gamma_2^0 \times \Gamma_3^0} J_2(\gamma_1^*, \gamma_2, \gamma_3, \hat{s}^i) \text{ is independent of}$$

i ($i=1,2,3$) for γ_1^* in an open neighborhood of γ_1^* .

$$2) \text{ If } \gamma_1^* = \arg \min_{\Gamma_1^0} J_1(\gamma_1, \gamma_2, \gamma_3, \hat{s}^i), i=2,3,$$

$J_1(\gamma_1^*, \gamma_2^0, \gamma_3^0)$ is sufficiently close to $J(\gamma_1^*, \gamma_2^0, \gamma_3^0)$ given by (4.1), when $\hat{s}^2$ and $\hat{s}^3$ are in a small neighborhood of s^0 . In other words, we wish to find that $\gamma_1^* \in \Gamma_1^0$ which protects the overall system performance (and maintains optimality) against incremental deviations in the goal perceptions of the lower level stations from the nominal team cost function (as quantified in the value of $\hat{s}$).

This again asks for a minimum sensitivity analysis which could be carried out following the lines of argument of Section 3. Let $\bar{\Gamma}_1$ be a subset of Γ_1^0 which comprises coordinator policies having property 1) above. For each $\gamma_1 \in \bar{\Gamma}_1$, let

$$\gamma_2^*(\gamma_1) = \arg \min_{\Gamma_2^0} J_2(\gamma_1, \gamma_2, \gamma_3, \hat{s}^i), \quad i=2,3, \quad (4.3)$$

and let

$$F(\gamma_1, \hat{s}) = J_1(\gamma_1, \gamma_2^*(\gamma_1), \gamma_3^0, \hat{s}), \quad (4.4a)$$

$$\gamma_3^*(\gamma_1) = \arg \min_{\Gamma_3^0} J_3(\gamma_1, \gamma_2^*(\gamma_1), \gamma_3, \hat{s}), \quad (4.4b)$$

When $\hat{s} = s^0 = (s^0, \dots, s^0)$, we clearly have, for all $\gamma_1 \in \bar{\Gamma}_1$,

$$F(\gamma_1, s^0) = J(\gamma_1, \gamma_2^0, \gamma_3^0)$$

where the latter has been defined by (4.1), and furthermore

$$F(\gamma_1, \hat{s}) \geq F(\gamma_1, s^0), \quad \forall \gamma_1 \in \bar{\Gamma}_1, \quad (4.5)$$

Hence, in order to satisfy property (ii) above, $\gamma_1^* \in \bar{\Gamma}_1$ has to be chosen so as to minimize the leading terms in the expansion of $F(\gamma_1, \hat{s}) - F(\gamma_1, s^0)$ around $\hat{s}$. If $\hat{s} \in \mathbb{R}^n$, the first and second order terms are

$$\delta F(\gamma_1, \hat{s}) = \sum_{i=1}^{2n} [dF(\gamma_1, \hat{s}) / d\hat{s}_i]_{\hat{s}=s^0} (\hat{s}_i - s_i^0) \quad (4.6)$$

and

$$\delta^2 F(\gamma_1, \hat{s}) = \sum_{j=1}^{2n} \sum_{i=1}^{2n} [d^2 F(\gamma_1, \hat{s}) / d\hat{s}_j d\hat{s}_i]_{\hat{s}=s^0} (\hat{s}_j - s_j^0) (\hat{s}_i - s_i^0), \quad (4.7)$$

respectively. The first of these is in general independent of γ_1 (as in (3.8)), but the second one does depend on γ_1 , thus making the problem of minimizing a suitable norm of the matrix

$$A(\gamma_1) = [d^2 F(\gamma_1, \hat{s}) / d\hat{s}_j d\hat{s}_i]_{\hat{s}=s^0}$$

over $\bar{\Gamma}_1$ a meaningful one. The desired $\gamma_1^* \in \bar{\Gamma}_1$ is then the argument of $\min_{\bar{\Gamma}_1} \|A(\gamma_1)\|$, which carries with it appealing sensitivity properties, as a coordinating policy for station 1.

Such an analysis has been carried out for some special cases and has been found to lead to some robust policies with simple structures (Cansever and Basar, 1982; Cansever, Basar and Cruz, 1983). The special cases treated in both references pertain to the situation when there is only one station at the lower level and the nominal problem is defined in discrete time, with two stages. In the former the cost functionals at the upper and lower levels are taken to be nominally different, whereas in the latter reference they are taken to be nominally the same, which leads to higher degrees of freedom in the choice of leader policies and could result in total insensitivity of the overall performance to deviations in the goal perceptions of the lower level.

In the continuous-time model and with more than one station at the lower level (as adopted here) the analysis will be more complicated, but it will basically follow the general steps outlined above. Its feasibility and computational tractability will depend to a great extent on the specific structures of f and J_i 's and their dependence on $\hat{s}$. It is also possible to envision situations where the uncertainty is both in the model (i.e., state equation, as in Section 3) and cost functionals. Minimum sensitivity coordinating policies can be obtained for that more general case also, at least in principle, by following the general lines of this and the previous section.

5. CONCLUDING REMARKS

This paper has discussed the general role of memory strategies in the coordination of interconnected systems when there is deterministic uncertainty in the modelling and in the quantification of objectives, which is perceived differently at different stations. The proposed approach has been a sensitivity analysis in a game-theoretic framework, which takes uncertainty to lie in a sufficiently small neighborhood of a nominal model and cost criteria, subject to space limitations. We have not included any explicit derivation of robust coordinating policies.

but the recipe given here could be used in the derivation of such policies for a large class of problems.

If either the nominal parameter values are not known or the discrepancies are not infinitesimal, one has to adopt a worst case analysis and solve coordination problems for the worst possible model and objective from the coordinator's point of view. This approach has not been adopted here because memory strategies do not play as significant role in such problems as they do in the ones discussed in this paper.

REFERENCES

- [1] Başar, T. (1977). Informationally nonunique equilibrium solutions in differential games. SIAM J. on Control, 15, 4, 636-660.
- [2] Başar, T. (1980). Optimum coordination of linear interconnected systems. Large Scale Systems, 1, 1, 17-27.
- [3] Başar, T. and G. J. Olsder (1982). Dynamic Noncooperative Game Theory, Academic Press, London/New York.
- [4] Başar, T. and H. Selbuz (1979). Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems. IEEE Trans. on Automatic Control, AC-24, 2, 166-179.
- [5] Cansever, D. H. and T. Başar (1982). A minimum sensitivity approach to incentive design problems. Proc. 21st IEEE Conf. on Decision and Control, 158-163.
- [6] Cansever, D. H., T. Başar and J. B. Cruz, Jr. (1983). Robustness of incentive policies in team problems with discrepancies in goal perceptions. Proc. 4th IFAC/IFORS/IIASA Conf. on Modelling and Control of National Economies.
- [7] Ho, Y. C., P. B. Luh and R. Muralidharan (1981). Information structure, Stackelberg games and incentive controllability. IEEE Trans. on Automatic Control, AC-26, 2, 454-490.
- [8] Kokotovic, P. V., J. B. Cruz, Jr., J. E. Heller and P. Sannuti (1968). Synthesis of optimally sensitive systems. Proc. IEEE, 56, 3, 1318-1324.
- [9] Papavassilopoulos, G. P. and J. B. Cruz, Jr. (1979). Nonclassical control problems and Stackelberg games. IEEE Trans. on Automatic Control, AC-24, 2, 155-166.
- [10] Sundararajan, N. and J. B. Cruz, Jr. (1972). Sensitivity reduction in time-varying linear and nonlinear systems. Int. J. of Control, 15, 3, 937-943.
- [11] Tolwinski, B. (1961). Closed-loop Stackelberg solution to multistage linear quadratic game. J. Optim. Theor. Appl.

END