Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Variations on the Theme of the Witsenhausen Counterexample

Tamer Başar

Abstract— **This is a semi-tutorial paper that places Witsenhausen's celebrated 1968 counterexample within a broad class of dynamic decision problems with nonclassical information, which includes stochastic linear-quadratic Gaussian (LQG) teams as well as LQG zero-sum stochastic games. For a fixed (nonclassical) information structure, there are instances (depending on the structure of the objective function) when linear policies are optimal and other instances (including Witsenhausen's counterexample) when the optimal policies are nonlinear. The paper discusses these instances, optimality as well as saddle-point property (in the case of zero-sum games) of linear policies, and implications of these results for general multi-stage decision problems with specific information structures. It also discusses possible extensions to nonzero-sum stochastic dynamic games where the solution concept is Nash equilibrium.**

I. Introduction

Since the work of Feldbaum [15] and his contemporaries, it has been recognized that in stochastic optimum control problems control generally has a *dual* role. One of these is the "action" role, which is what one would have in deterministic optimum control, where the controller tries to move the system state toward a desired target value while optimizing a given performance index. The second role is the *probing* one, where the controller, recognizing that use of "higher quality" information will generally lead to better performance, will attempt to *shape* the signals carried to future stages in such a way that the information content of the measurement received by the controller at future stages will be enhanced. This *probing* role of control is in general conflicting with its *action* role, and hence an optimal (stochastic) controller is one that achieves an "optimal tradeoff" between these two apparently conflicting objectives.

Stochastic optimum control problems which do not exhibit a "dual" role for control are known as *neutral* problems, where the conditional probability distribution of the state vector given past and present measurements, past control actions, and past control laws (or policies) is independent of the control laws [23]. The main implication of this property is that the "quality" of the information carried to future stages cannot be affected by the choice of the control policies in the past, thus allowing for a two step derivation of the optimum controller: First determine the conditional probability distribution (*cpd*) of the state, and express the expected cost in terms of this quantity and the control (yet to be determined), and subsequently minimize the new expected cost function over all control laws as functions of the *cpd*, which provides *sufficient statistics* for the stochastic control problem. This property of neutral stochastic control problems is known as the *separation* of estimation and control (or loosely referred to as the *separation principle*). The standard Linear-Quadratic-Gaussian (LQG) problem is a prime example of a neutral stochastic control problem where in fact an even stronger version of the separation principle applies, where the minimization-relevant part of the expected cost as a function of the *cdp* depends only on the conditional mean. Hence, in this case, the minimization problem faced by the controller is identical to the one where all random quantities are replaced by their mean values, which justifies the coining of the word *certainty equivalence*.

If a stochastic control problem is not neutral, then the *dual* role of control becomes dominant. Such problem formulations arise in various contexts, such as when control has limited or no memory, when system dynamics depend on some parameters which are correlated across stages (a setting that arises in, for instance, stochastic adaptive control), or stochastic teams where different control stations have access to different but correlated measurement channels. The well-known counterexample of Witsenhausen [22] provides perhaps the simplest model (two-stage, scalar, LQG, but memoryless controller) that depicts eloquently the conflicting role of the control at the first stage, between "action now" and "maximum information transmission" to the next stage for the benefit of "action then." Even though explicit solutions to some nonneutral stochastic control and team problems have been obtained in the past (see, for instance, [17], [9]), a general theory is still lacking on this 40th anniversary of the appearance of *Witsenhausen's Counterexample*.

This paper uses this occasion to review some existing results on problems which can be viewed as belonging to a *neighborhood* of the Witsenhausen counterexample, with some of these problems being tractable while others not.

II. Two-Stage LQG Problems with Nonclassical Information

Consider the following two-stage stochastic control or equivalently two-agent dynamic stochastic team problem, where all quantities are scalar:

A Gaussian random variable, x, with mean zero and variance σ_x^2 is to be transformed into another random variable, $u_0 = \gamma_0(x)$, which is transmitted over a Gaussian channel, $y = u_0 + w$, with zero-mean additive Gaussian noise w of variance σ_w^2, the output, y, of which is to be further transformed into another random variable, $u_1 = \gamma_1(y)$. The objective is to choose the transformations γ_0 and γ_1 in such

Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois, 1308 West Main Street, Urbana, IL 61801-2307, USA. Tel: (+1) 217 333-3607; Fax: (+1) 217 244-1653; Email: tbasar@control.csl.uiuc.edu.

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

a way that a performance index, $Q(x, u_0, u_1)$, quadratic in x, u_0, and u_1, is minimized in the average sense. That is, we seek the pair $\gamma^* := (\gamma_0^*, \gamma_1^*)$ such that

$$J(\gamma^*) = \min_{\gamma} J(\gamma) =: J^* \qquad (1)$$

where

$$J(\gamma) = E\left[Q(x, \gamma_0(x), \gamma_1(y))\right] \qquad (2)$$

with expectation, $E[\cdot]$, taken over the statistics of x and w, which are assumed to be independent. Furthermore, the minimization is over the space of all Borel-measurable maps, that is both policies (decision rules) γ_0 and γ_1 are taken to be Borel-measurable maps of the real line onto itself.

This is a stochastic decision problem with *nonclassical information*, because the information to be used by the decision rule, γ_1, of the second agent depends on the action, u_0, of the first agent (and thereby on the decision rule of the first agent), but the second agent does not have access to the information of the first agent (that is, x). If we view it as a single-agent problem where the agent acts twice, then it is one where the agent is memoryless, that is it does not remember what it had observed at the earlier stage. As such, these problems belong to the realm of inherently difficult decision problems for which a systematic solution process does not exist, one of the main reasons being that due to loss in memory, a sequential decomposition is not possible [23], [24]. We now consider different instances of this class of problems, corresponding to different choices of the performance index Q, some of which admit explicit, relatively simpler solutions while some others do not. Hence the message will be that it is not only the nonclassical nature of the information structure but also the structure of the performance index that contributes to the difficulty in solving these problems.

A. Witsenhausen's counterexample

Now let the quadratic performance index Q be picked as

$$Q_W(x, u_0, u_1) = k_0(u_0 - x)^2 + (u_1 - u_0)^2 , \qquad (3)$$

where $k_0 > 0$ is a given parameter. Note that here the first agent wants to stay as close to x as possible, while the second agent wants to stay as close to the action of the first agent, u_0, as possible. This can also be viewed as a standard discrete-time two-stage LQG optimal control problem, with state equations

$$x_1 = x_0 + v_0 , \qquad x_2 = x_1 - v_1$$

measurement equations

$$y_0 = x_0 , \qquad y_1 = x_1 + w_1 ,$$

and memoryless controls

$$v_0 = \mu_0(y_0) , \qquad v_1 = \mu_1(y_1) ,$$

where μ_0 and μ_1 are the instantaneous measurement output control policies at stages 0 and 1, respectively. This becomes equivalent to the earlier formulation in view of the correspondences

$$u_0 = x_0 + v_0 , \quad u_1 = v_1 , \quad x = x_0 , \quad w = w_1 , \quad y = y_1 ,$$

if we pick the cost function as

$$\tilde{Q}(x_2, v_0) = (x_2)^2 + k_0(v_0)^2 \equiv Q_W(x_1 - v_0, x_1, x_1 - x_2) .$$

Witsenhausen has shown in a 1968 paper [22] that the optimal solution to this problem exists, but the optimum decision rules (μ_0 and μ_1, or equivalently γ_0 and γ_1) are not linear. For the latter, he has shown that there exist nonlinear policies which outperform the best linear ones.[1] A class of such nonlinear policies introduced by Witsenhausen, and further improved upon by Bansal and Başar [1] is

$$\begin{aligned} u_0 &= \gamma_o(x) = \epsilon\,\mathrm{sgn}(x) + \lambda x , \\ u_1 &= \gamma_1(y) = E[\epsilon\,\mathrm{sgn}(x) + \lambda x | y] \end{aligned}$$

where ϵ and λ are parameters to be optimized over (in [22] the values are picked as $\lambda = 0$ and $\epsilon = \sigma_x$, and some asymptotics are studied). Clearly, if $\epsilon = 0$, this class of decision rules will be linear, since $E[\lambda x|y]$ will be linear for each λ, however when $\epsilon \neq 0$, the decision rules at both stages will be nonlinear. To give some indication of how much can be gained by taking $\epsilon \neq 0$, let us consider the case with parameter values $k_0 = 0.1, \sigma_x^2 = 10, \sigma_w^2 = 1$; then the best linear policy at stage *zero* has the gain $\lambda_{\text{opt}} = -0.1127$, with the corresponding value of J being 0.900. If however ϵ is picked to be 2, the corresponding value of J (for the same choice of λ which is clearly not optimal and can be further improved upon) is 0.5203, which registers a substantial improvement over the best linear solution. For another scenario, let us take $k_0 = 0.01, \sigma_x^2 = 80, \sigma_w^2 = 1$; in this case the best linear policy at stage *zero* has the gain $\lambda_{\text{opt}} = 0.01006$, with the corresponding value of J being 0.7920, whereas for the same value of λ, picking $\epsilon = 5$ leads to a value of $J = 0.3309$. Further numerical results can be found in [1], which also shows that if $\lambda = 0$, $\epsilon = \sigma_x\sqrt{2/\pi}$ and $k_0\sigma_x^2 = 1$, as $k_0 \to 0$ the bound on asymptotic performance becomes $(1 - (2/\pi)) = 0.363$.

B. Generalized Gaussian test channel

Now consider a different choice for Q:

$$Q_{\text{TC}}(x, u_0, u_1) = k_0(u_0)^2 + (u_1 - x)^2 , \qquad (4)$$

where again $k_0 > 0$. Note that here the second agent's objective is to estimate the random variable x in the minimum mean square (MMS) sense, using a measurement that is transmitted over a Gaussian channel where the input to the channel is *shaped* by the first agent who has access to x and has a soft constraint ($k_0E[(u_0)^2]$) on its action. The version of this problem where the soft constraint is replaced by a hard power constraint, $E[(u_0)^2] \leq k$, is known as the *Gaussian Test Channel* (GTC), and in this context γ_0 is the *encoder* and γ_1 the *decoder*, whose optimal choice is clearly

[1] As of today, closed-form expressions for the optimal nonlinear policies are not available, and their characterization is not known.

the conditional mean of x given y, that is $E[x|y]$. The best encoder for the GTC can be shown to be linear (a scaled version of the source output, x), which in turn leads to a linear optimal decoder. The approach here (as we will discuss further below for a more general, soft-constrained version), which is in fact the only approach known to apply here, is to obtain bounds on the performance using an inequality from information theory involving channel capacity [25] and rate distortion function [14], and then showing that the bound can be achieved using linear policies.

Now, consider the more general version of (4):

$$Q_{\text{GTC}}(x, u_0, u_1) = k_0(u_0)^2 + (u_1 - x)^2 + b_0 u_0 x , \quad (5)$$

where b_0 is a parameter. Let

$$E[(u_0)^2] =: \alpha \quad \text{and} \quad E[(u_1 - x)^2] =: \beta .$$

Then, with J defined as before, by (2), and with γ_0 and γ_1 constrained as above, we have the inequalities

$$\begin{aligned} \inf_{\gamma} J(\gamma) &\geq k_0\alpha + \beta + \inf_{\gamma_0} b_0 E[\gamma_0(x)x] \\ &\geq k_0\alpha + \beta - |b_0|\sigma_x\sqrt{\alpha} \quad (6) \end{aligned}$$

where the second one follows from the Cauchy-Schwarz inequality.

Now, by the data processing theorem [25], in a linear configuration the mutual information between two random variables closer to each other is no smaller than the mutual information between two random variables farther apart. In our case, this translates to

$$I(x; y) \geq I(x; u_1) \quad (7)$$

where $I(\cdot;\cdot)$ stands for mutual information. For each fixed $\alpha > 0$, $I(x; y)$ is bounded from above by the capacity of the channel, $C(\alpha)$, which is known for the Gaussian channel to be [16]

$$C(\alpha) = \frac{1}{2}\log(1 + (\alpha/\sigma_w^2)\,) .$$

Further, for each fixed β, the quantity $I(x; u_1)$ is bounded from below by the rate distortion function, $R(\beta)$, for which the expression is [14]

$$R(\beta) = \frac{1}{2}\log(\sigma_x^2/\beta) .$$

In view of (7), we have

$$\frac{1}{2}\log(1 + (\alpha/\sigma_w^2)\,) = C(\alpha) \geq R(\beta) = \frac{1}{2}\log(\sigma_x^2/\beta)$$

which leads to the following bound on β:

$$\beta \geq \sigma_w^2\sigma_x^2/(\alpha + \sigma_w^2)$$

which is tight with

$$\gamma_0(x) = -\text{sgn}(b_0)\frac{\sqrt{\alpha}}{\sigma_x}x \quad (8)$$

Substitution of this in (6) leads to

$$\inf_{\gamma} J(\gamma) \geq k_0\alpha + \sigma_w^2\sigma_x^2/(\alpha + \sigma_w^2) - |b_0|\sigma_x\sqrt{\alpha} \quad (9)$$

Let α^* be the positive value of α that minimizes the bound in (9), which exists and is unique. It is a solution of the polynomial equation

$$[2k_0\sqrt{\alpha} - |b_0|\sigma_x]\,[\alpha + \sigma_x^2]^2 = 2\sigma_w^2\sigma_x^2\sqrt{\alpha} . \quad (10)$$

Then, when Q is in the structural form (5), the solution to (2) exists, **is linear**, and is given by

$$\gamma_0^*(x) = -\text{sgn}(b_0)\frac{\sqrt{\alpha^*}}{\sigma_x}x , \quad (11)$$

$$\gamma_1^*(y) = E[x|y] = -\frac{\text{sgn}(b_0)\sigma_x\sqrt{\alpha^*}}{\alpha^* + \sigma_w^2}y . \quad (12)$$

Remark: The main difference between the two problems of Subsection II-A and Subsection II-B is that Q in the former has a product term between the decision rules of the two agents while in the latter it does not. Hence, it is not only the nonclassical nature of the information structure but also the structure of the performance index that determines whether linear policies are optimal in LQG multi-stage decision problems. ◇

III. Conflicting Objectives: A Zero-Sum Stochastic Game

To bring in a further perspective on the general problem formulated in the previous section, we consider now a game situation with the first term in (3) now negative:

$$Q_G(x, u_0, u_1) = -k_0(u_0 - x)^2 + (u_1 - u_0)^2 , \quad (13)$$

where again $k_0 > 0$, and the information structure is the same as before. Here the two agents (players) have opposing objectives: while the second agent still wants to minimize the expected value of Q_G, the first agent wants to maximize it. This is then a zero-sum stochastic game (with dynamic information), and an appropriate solution concept in this case is that of a *saddle point* [13]. That is, with J defined as before, we are looking for a pair of decision rules (γ_0^*, γ_1^*), such that the following pair of saddle-point inequalities hold: for all Borel measurable functions γ_0 and γ_1,

$$J(\gamma_0, \gamma_1^*) \leq J(\gamma_0^*, \gamma_1^*) \leq J(\gamma_0^*, \gamma_1) \quad (14)$$

Now it turns out ([11], [12]) that the *solution* here can be obtained explicitly and is linear, and hence this problem where there is a conflict in objectives is in some sense *easier* than the Witsenhausen counterexample where the objectives were aligned. First note that if γ_0 is linear in x, say $\gamma_0(x) = \alpha x$ for some parameter α,[2] then the γ_1 that minimizes J is also linear (and unique), being the conditional mean of αx given y. Hence,

$$\gamma_0(x) = \alpha x \quad \Rightarrow \quad \gamma_1(y) = \frac{\alpha^2\sigma_x^2}{\alpha^2\sigma_x^2 + \sigma_w^2}y$$

Conversely, if γ_1 is linear in y, say $\gamma_1(y) = \lambda y$ for some parameter λ, then provided that

$$k_0 > (\lambda - 1)^2 \quad (15)$$

[2]This α is not related to the one in Subsection II-B.

which makes Q_G strictly concave in u_0,[3] the γ_0 that maximizes J is also linear (and unique). Hence,

$$\gamma_1(y) = \lambda y \quad \Rightarrow \quad \gamma_0(x) = -\frac{k_0}{k_0 - (\lambda - 1)^2} x$$

For these policies to constitute a saddle point, we have to find a pair (α, λ) which simultaneously solve

$$\lambda = \frac{\alpha^2 \sigma_x^2}{\alpha^2 \sigma_x^2 + \sigma_w^2}$$
$$\alpha = -\frac{k_0}{k_0 - (\lambda - 1)^2}$$

while satisfying the constraint (15). Note that from the first equation above, we have $0 < \lambda < 1$, and hence if $k_0 > 1$ the constraint (15) is satisfied whenever the simultaneous equations above admit a solution for λ in the interval $(0, 1)$. Exploring this further, substituting for α from the second equation above into the first equation, and rearranging, we obtain the following equation involving a 5th-order polynomial in λ one of whose roots is the value sought:

$$f(\lambda) := (\sigma_w^2/\sigma_x^2)\lambda \left[k_0 - (1 - \lambda)^2\right]^2 - k_0^2(1 - \lambda) = 0$$

Since $f(0) < 0, f(1) > 0$, and f is a strictly increasing function on $(0, 1)$, it has a unique root in the interval of interest, which we denote by λ^*. Hence, provided that $k_0 > 1$, the stochastic game of this section admits a saddle-point solution which is linear in the measurements of the two players. Further, this saddle-point solution is unique, which follows from the *ordered interchangeability* property [13] of multiple saddle points, since the optimum response of each player to an announced linear policy of the other player is unique, as already shown above.

The development above holds even if $k_0 \leq 1$. Then, the constraint (15) will have to be enforced, and we have to look for a root of $f(\lambda)$ in the interval $(1 - \sqrt{k_0}, 1)$. One can again see that $f(1 - \sqrt{k_0}) < 0, f(1) > 0$, and f is a strictly increasing function in $(1 - \sqrt{k_0}, 1)$, and thus has a unique root in the interval of interest. Hence, the game has a unique saddle-point solution as long as $k_0 > 0$, and the saddle-point policies are linear:

$$\gamma_0^*(x) = -\frac{k_0}{k_0 - (\lambda^* - 1)^2} x\,, \quad \gamma_1^*(y) = \lambda^* y$$

where λ^* is the unique solution of the polynomial equation $f(\lambda) = 0$ in the interval $(\max(0, 1 - \sqrt{k_0}), 1)$.

IV. Extensions to the Basic Model

The basic model of Section II can be extended in different directions, including:

(A) u_0 having access to noise corrupted version of x, that is $u_0 = \gamma_0(x + v)$, where v is a zero mean Gaussian random variable, independent of x and w.

(B) Vector-valued variables, where x, w, y, u_0, u_1, as well as v introduced above, are no longer scalar valued, and could have different dimensions, with for example, in the setting of Section II,

$$y = Cu_0 + Dw\,, \quad u_0 = \gamma_0(z)\,, \quad u_1 = \gamma_1(y)\,,$$

where $z := Fx + Gv$, C, D, F, G are matrices of appropriate dimensions, and the independent Gaussian zero-mean random vectors x, v, w have specified covariance matrices. The performance index is still given by (2), and the criterion is (1), with the appropriate interpretation for vector-valued variables.

(C) Stochastic LQG teams where in the formulation above different components of u_0 and different components of u_1 are controlled by different agents who have access to individual compoments of z and y, respectively. That is, assuming that the dimensions of u_0, z, u_1 and y are the same, say n,

$$[u_0]_i = \gamma_{0i}(z_i)\,, \quad [u_1]_i = \gamma_{1i}(y_i)\,, \quad i = 1, \ldots, n\,. \qquad (16)$$

(D) Multiple agents with different objectives, where in the setting above, there are n performance indices, $Q_i(x, u_0, u_1)$, $i = 1, \ldots, n$, with corresponding expected costs, $J_i(\gamma)$. An appropriate solution concept here is that of noncooperative Nash equilibrium: an n-tuple $\gamma^* = (\gamma_{0i}^*, \gamma_{1i}^*)$, $i = 1, \ldots, n$, is in Nash equilibrium if for all $(\gamma_{0i}, \gamma_{1i})$ and all $i = 1, \ldots, n$,

$$J_i(\gamma^*) \leq J_i((\gamma_{0i}, \gamma_{1i}); \gamma_{-i}^*)\,, \qquad (17)$$

where γ_{-i}^* stands for γ^* with only the i-th one left out.

(E) Multiple (higher than two) stages, with information carried from one stage to the next being limited.

We now discuss briefly the extent of the difficulties in obtaining the solution to problems that fall in these five different classes.

A. Noise corrupted initial state

When u_0 has access to $z = x + v$, instead of x, the results of Sections II and III hold structurally. If there is a cross term between u_0 and u_1 in Q, the same difficulties as in the Witsenhausen counterexample arise. When the cross term is absent, however, the optimal solution again exists and is linear. That is, for some positive α, which is again the solution of a polynomial equation (as in (10)), we have [2]

$$u_0 = \gamma_0^*(z) = \alpha z\,, \quad u_1 = \gamma_1^*(y) = \frac{\alpha \sigma_x^2}{\alpha^2(\sigma_x^2 + \sigma_v^2) + \sigma_w^2}\, y$$

The proof again follows information-theoretic arguments [2], and the same structural result holds even if w is correlated with z.

Now, for the stochastic game problem of Section III, again the same structural result holds: there exists a saddle-point solution, which is linear for both players. Following the analysis of Section III, the counterpart of the polynomial function f in this case is

$$f(\lambda) := \frac{\sigma_w^2}{\sigma_x^2}\left[1 + \frac{\sigma_v^2}{\sigma_x^2}\right] \lambda \left[k_0 - (1 - \lambda)^2\right]^2 - k_0^2(1 - \lambda)$$

[3]If the condition (15) is not satisfied, then for the given γ_1, the player who chooses u_0 can make the value of J arbitrarily large.

which again has a unique root in the interval $(1-\sqrt{k_0}, 1)$, since $f(1-\sqrt{k_0}) < 0, f(1) > 0$, and f is a strictly increasing function over this interval. Letting this unique solution be λ^*, the unique saddle-point solution, which is linear, is:

$$\gamma_0^*(z) = -\frac{k_0}{k_0 - (\lambda^* - 1)^2} \frac{\sigma_x^2}{\sigma_x^2 + \sigma_v^2} z\,, \quad \gamma_1^*(y) = \lambda^* y\,.$$

Remark: Clearly in these problems *certainty equivalence* does not hold, that is the solution is not of the type where one first solves the deterministic version of the problem and then replaces the random variables with their conditional mean values at the solution point. On the other hand, one would normally expect certainty equivalence to hold if the information structure is of the classical type, that is (in this case) the agent acting at the second stage has access to not only his private information y but also z. This is indeed the case with team problems (which then become standard LQG stochastic control problems), but not necessarily for stochastic zero-sum games which feature many pitfalls; for details we refer to [8].

B. Vector-valued variables

For vector-valued variables, as introduced earlier in this section, both problems of Section II, with and without cross terms between u_0 and u_1, present difficulties due to the nonclassical nature of the information structure, and still remain unsolved today. Linear decision rules are not generally optimal for the vector version of the Gaussian test channel [18], unless u_0 and u_1 are still scalar, and so are x and v, but y is vector-valued [2]. This can be viewed as the transmission of a garbled version (z) of a Gaussian message (x) over a number of noisy (Gaussian) channels under a quadratic fidelity criterion. Optimum solution again consists of linearly transforming z to a certain optimum power level α^* (by the first agent, $\gamma_0^*(z)$) and then optimally decoding it at the receiving end by using a linear transformation (by the second agent, $\gamma_1^*(y)$), with α^* obtained as the unique (relevant) solution of a fifth-order polynomial equation [2].

Now, regarding the vector version of the zero-sum stochastic game of Section III, its saddle-point solution is still linear (and unique), and the analysis of Section III readily applies at the conceptual level, with some of the details, however, being more involved; see [11].

C. Stochastic LQG teams

The stochastic LQG team problem formulated earlier in this section (with decentralized information) features all the complexities of the centralized vector-valued one (centralized at each end, but still nonclassical from one stage to the next), unless there is a forward channel that informs agents at the front end on the garbled decentralized information received at the other (back) end. This is known as *quasi-classical* information, with (16) replaced by

$$[u_0]_i = \gamma_{0i}(z_i)\,, \;\; [u_1]_i = \gamma_{1i}(y_i, z)\,, \quad i = 1, \ldots, n\,, \quad (18)$$

where z denotes the collection of measurements of the agents at stage *zero*. The problem is still minimization of the expected value of a general quadratic function $Q(x, u_0, u_1)$. The quasi-classical nature of the information allows for a sequential decomposition, where at the front end and subsequently at the back end static quadratic Gaussian teams are solved. At each stage, one makes use of a result due to Radner [20] to establish linearity of (team-)optimum decision rules. Hence, with quasi-classical information (enabled through the feedforward transmission of measurements from back stages, but no sharing of information among agents occupying the same level), LQG teams, strictly convex in the decision variables, admit linear optimum decision rules, that is optimum choices for γ_{0i} and γ_{1i} in (18) are linear in their arguments.

D. Stochastic Nash games

Consider now the same structure (with quasi-classical information structure) as in Subsection IV-C above, but with different agents having different performance indices, in which case the noncooperative Nash equilibrium as defined by (17) becomes relevant. These games, when the performance index for each player is quadratic, again admit unique linear solutions, but the sequential decomposition in this case is more involved than in the case of stochastic teams; see [7]. It is important to mention that for *uniqueness* it is necessary that the actions are not feedforwarded but only the measurements are, whereas in the stochastic team problem the distinction between the two types of information structures is not consequential. In other words, in game situations there should not be any redundancy in information received by the players, as that will lead to consequential nonuniqueness. In the derivation of the unique linear Nash equilibrium, a key result that is used at each stage is an extension of Radner's result on quadratic teams referred to in the preceding subsection to quadratic Gaussian static Nash games [4], which says that decision rules in Nash equilibrium are linear in the measurements of individual players, whenever an equilibrium exists; see also [5], [6].

E. Multi-stage LQG with nonclassical information

We have already introduced in the previous two subsections some multi-stage formulations, where however the nonclassical information was relaxed to a quasi-classical one where an agent following another one in the decision tree has access to the measurements used by the other agent, but agents operating at the same level of the decision tree need not share information. There are, however, multi-stage decision/control problems that are tractable in spite of the nonclassical nature of the information structure. One such class of problems arise in joint controller-sensor design, as discussed in [3]. Consider the scalar discrete-time plant

$$x_{n+1} = \rho_n x_n + u_n + v_n\,, \;\; n = 0, 1, \ldots$$

along with the scalar measurement

$$y_n = h_n + w_n\,, \;\; n = 0, 1, \ldots\,,$$

where $\{v_n\}$ and $\{w_n\}$ are i.i.d. Gaussian random variables, with zero mean and independent of each other as well as of the Gaussian initial state x_0. The variable u_n is the control, allowed to depend on the present and past values of y, and h_n is another decision variable (the sensor structure), which has to be designed as a function of the current value of the state, x_n, and possibly also of the past values of y, and this design has to be picked optimally, along with the control, so as to minimize the expected value of a stage-additive quadratic cost function. This is a dynamic decision problem that features nonclassical information because u_n and h_n can be seen as the actions of two agents with the decision of one affecting the information of the other, who however do not share information. Employing the Gaussian test channel sequentially, as well as sequential rate distortion theory [14], this nonclassical stochastic control problem can be shown to admit a linear optimal solution (for both u_n and h_n) [3]. Its continuous-time version (again scalar) also admits a linear optimal solution [10], where now the continuous-time Gaussian test channel with feedback is employed [19]. None of these results admit easy extensions to multivariable systems, when optimum solutions (if they exist) will in general not be linear.

For another type of a joint controller/sensor design under constraints on the observation alphabet and power, which also employs sequential rate distortion theory, we refer to [21]. Yet another class of tractable/solvable dynamic stochastic optimization problems with nonclassical information (that is of *non-neutral* type) can be found in [9], where the tool this time is the powerful machinery of saddle-point equilibria. The approach there is to relate a single-objective dynamic optimization problem to a sequence of nested zero-sum games. The problem arises as a macro-economic model of credibility and monetary policy, and involves *active learning*.

V. Conclusions

Written on the 40th anniversary of the appearance of the celebrated counterexample of Witsenhausen in stochastic control, this paper has discussed a number of problems in stochastic dynamic decision making which are variations on the theme of the counterexample. These different problem formulations have shown that it is not only the nonclassical nature of the information structure, but also the structure of the performance index and the coupling of different decision variables that are responsible for intractability of some of these decision problems. While some are intractable, yet there are others with still nonclassical information which are tractable, both as teams as well as games. The paper has identified these tractable problems, and has also discussed extensions to zero-sum as well as nonzero-sum games, identifying also the challenges encountered.

References

[1] R. Bansal and T. Başar, "Stochastic team problems with nonclassical information revisited: When is an affine law optimal?" *IEEE Trans. Automat. Contr.*, AC-32(6):554-559, June 1987.

[2] R. Bansal and T. Başar, "Solutions to a class of linear-quadratic-Gaussian (LQG) stochastic team problems with nonclassical information," *Systems Contr. Letters*, 9(2):125-130, 1987.

[3] R. Bansal and T. Başar, "Simultaneous design of measurement and control strategies in stochastic systems with feedback," *Automatica*, 45(5):679-694, September 1989.

[4] T. Başar, "An extension of Radner's theorem to nonzero-sum games," *Proc. 12th Annual Allerton Conf. Circuit and System Theory*, pp. 199–206, October 1974.

[5] T. Başar, "Equilibrium solutions in two-person quadratic decision problems with static information structures," *IEEE Trans. Automat. Contr.*, AC-20(3):320-328, June 1975.

[6] T. Başar, Decentralized multicriteria optimization of linear stochastic systems, *IEEE Trans. Automat. Contr.*, AC-23(2):233–243, April 1978.

[7] T. Başar, "Two-criteria LQG decision problems with one-step delay observation sharing pattern," *Information and Control*, 38(1):21-50, July 1978.

[8] T. Başar, "On the saddle-point solution of a class of stochastic differential games," *Journal on Optimization Theory and Applications*, 33(4):539-556, April 1981.

[9] T. Başar, "Solutions to a class of nonstandard stochastic control problems with active learning," *IEEE Trans. Automat. Contr.*, AC-33(12):1122-1129, December 1988.

[10] T. Başar and R. Bansal, "Optimum design of measurement channels and control policies for linear-quadratic stochastic systems," *European J. Operations Research*, 73:226-236, 1994.

[11] T. Başar and M. Mintz, "Minimax estimation under generalized quadratic loss," *Proc. 1971 IEEE Conf. Decision Contr.*, pp. 456-461, December 1971.

[12] T. Başar and M. Mintz, "On the existence of linear saddle-point strategies for a two-person zero-sum stochastic game," *Proc. 1972 IEEE Conf. Decision Contr.*, pp. 188-192, December 1972.

[13] T. Başar and G.J. Olsder, *Dynamic Noncooperative Game Theory*, SIAM Classics in Applied Mathematics, Philadelphia, 1999.

[14] T. Berger, *Rate Distortion Theory*, Prentice Hall, Englewood-Cliffs, NJ, 1971.

[15] A. A. Feldbaum, "Dual control theory," *Automation and Remote Control*, Part I: 21(9):874–880, April 1961; Part II: 21(11):1033–1039, May 1961.

[16] R.G. Gallagher, *Information Theory and Reliable Communication*, Wiley, New York, NY, 1968.

[17] Y.C. Ho, "Team decision theory and information structures," *Proc. IEEE*, 68:644-654, June 1980.

[18] Y.C. Ho, M.P. Kastner, and E. Wong, "Teams, signaling, and information theory," *IEEE Trans. Automat. Contr.*, AC-23(4):305-311, April 1978.

[19] T.T. Kadota, M. Zakai, and J. Ziv, "Capacity of a continuous memoryless channel with feedback," *IEEE Trans. Inform. Theory*, IT-17:372-378, 1971.

[20] R. Radner, "Team decision problems," *Ann. Math. Statist.*, 33:857-881, 1962.

[21] S. Tatikonda, A. Sahai, and S. Mitter, "Control of LQG systems under communication constraints," *Proc. American Contr. Conf.*, San Diego, CA, pp. 2778-2782, June 1999.

[22] H.S. Witsenhausen, "A counterexample in stochastic optimal control," *SIAM J. Contr.*, 6(1):131-147, 1968.

[23] H.S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proc. IEEE*, 59(11):1557-1566, 1971.

[24] H.S. Witsenhausen, "The intrinsic model for discrete stochastic control: Some open problems," *Lecture Notes in Econ. and Math. Syst.*, 107:322-335, 1974.

[25] A.D. Wyner, "Another look at the coding theorem of information theory: A tutorial," *Proc. IEEE*, 58:894-913, June 1970.