Chapter 8
Linear-Quadratic Gaussian Dynamic Games with a Control-Sharing Information Pattern

Meir Pachter

Abstract A "zero-sum" Linear-Quadratic Gaussian Dynamic Game (LQGDG) where the players have partial information is considered. Specifically, the players' initial state information and their measurements are private information, but each player is able to observe his antagonist's past inputs: the protagonists' past controls is shared information. Although this is a game with partial information, the control-sharing information pattern renders the game amenable to solution by the method of dynamic programming. The correct solution of LQGDGs with a control-sharing information pattern is obtained in closed-form.

Keywords Linear quadratic Gaussian dynamic game • Partial information

MSC Codes: 91A25, 93C41, 49N70

8.1 Introduction

The complete solution of Linear-Quadratic Gaussian Dynamic Games (LQGDGs) has been a longstanding goal of the controls and games communities. That LQGDGs with a nonclassical information pattern can be problematic has been amply illustrated in Witsenhausen's seminal paper (Witsenhausen 1968)—see also Pachter and Pham (2014). Control theorists have traditionally emphasized control theoretic aspects and the backward induction/dynamic programming solution method, which however is not applicable to dynamic games with partial information—one notable exception notwithstanding, being the game with partial information that will be discussed herein. And game theorists have focused on information economics, that is, the role of information in games, but for the most part, discrete games. The state of affairs concerning dynamic games with partial information is not

M. Pachter (✉)
Department of Electrical and Computer Engineering, Air Force Institute of Technology, AFIT/ENG, 2950 Hobson Way, Bldg. 640, Wright-Patterson A.F.B., OH 45433, USA
e-mail: meir.pachter@afit.edu

F. Thuijsman, F. Wagener (eds.), *Advances in Dynamic and Evolutionary Games*, Annals of the International Society of Dynamic Games 14,
DOI 10.1007/978-3-319-28014-1_8

satisfactory. In this respect, the situation is not much different now than it was in 1971 when Witsenhausen made a similar observation (Witsenhausen 1971). In this article a careful analysis of dynamic games with partial information is undertaken. We exclusively focus on LQGDGs, which are more readily amenable to analysis. Indeed, Linear-Quadratic Dynamic Games (LQDGs) with perfect information stand out as far as applications of the theory of dynamic games are concerned: a canonical instance of an application of the theory of LQDGs can be found in Ho et al. (1965) where it has been shown that its solution yields the Proportional Navigation (PN) guidance law which is universally used in Air-to-Air missiles. Furthermore, the theory of LQDGs has been successfully applied to the synthesis of H_∞ control laws (Basar and Bernhard 2008). The theory of LQDGs with perfect information has therefore received a great deal of attention (Basar and Olsder 1995; Engwerda 2005; Pachter and Pham 2010). In these works, the concepts of state, and state feedback, are emphasized and the solution method entails backward induction, a.k.a., Dynamic Programming (DP).

Concerning informational issues in LQGDGs: In previous work Radner (1962) and Pachter and Pham (2013) a static Linear-Quadratic Gaussian (LQG) team problem was addressed and a static "zero-sum" LQG game with partial information was analyzed in Pachter (2013). In this article a dynamic "zero-sum" LQG game, that is, a LQGDG, where the players have partial information, is addressed. The information pattern is as follows. The players' initial state information and their measurements are private information, but each player is able to observe the antagonist's past inputs: the protagonists' past controls is shared information. This information pattern has previously been discussed by Aoki (1973), and in the context of a team decision problem, this information pattern has also been discussed in Sandell and Athans (1974). However, Aoki (1973) took "a wrong turn": as so often happens in the literature of games with partial information, one is tempted to assume the players will try to second guess the opponents' private information, say, their measurements. The vicious cycle of second guessing the opponent's measurements leads to a mirror gallery like setting and to a dead end. This point is discussed in Sect. 8.4. Concerning reference Sandell and Athans (1974) where a decentralized dynamic team problem with a control-sharing information pattern is considered: It is argued that an infinite amount of information is contained in a real number, which, in theory, is correct. And since the control information is shared, then at least in a cooperative control/team setting, a player/agent could in principle encode in the controls information about to be sent to his partner his private information, for example, his measurements history. This is due to the fact that the controls which are about to be communicated can be modified slightly to encode the measurements information of the protagonists without significantly disturbing the control, and consequently, have a barely noticeable effect on the value of the game. One then falls back on the solution of the LQG cooperative control problem with a one-step-delay shared information pattern (Kurtaran and Sivan 1974). However, this scheme has no place in an antagonistic scenario, a.k.a., "zero-sum" LQGDG as discussed in our paper, and also does not properly model a decentralized control scenario. Moreover, this scheme totally depends on the players' ability of obtaining a noiseless observation of the broadcast control and as such, exhibits a lack of robustness to measurement error and is not a viable proposition. In the end, it is acknowledged in Sandell and Athans (1974) that the control-sharing information pattern leads to a stochastic control problem that is ill posed and it is stated that "the need for future work on this problem is obvious". Unfortunately, the analysis of LQGDGs with a control-sharing information pattern presented in Aoki (1973) and Sandell and Athans (1974) is patently incorrect. In this paper LQGDGs with a control-sharing information pattern are revisited. A careful analysis reveals that although this is a game with partial information, the control-sharing information pattern renders the game amenable to solution by the method of DP. It is shown that the solution of the LQGDG with a control-sharing information pattern is similar in structure to the solution of the LQG optimal control problem in so far as the principle of certainty equivalence/decomposition holds. A correct closed-form solution of a LQGDG with a control-sharing information pattern is obtained.

The paper is organized as follows. The LQGDG problem statement and the attendant control-sharing information pattern are presented in Sect. 8.2. The state estimation algorithm required for the solution of LQGDGs with a control-sharing information pattern is developed in Sect. 8.3. The analysis of Linear-Quadratic Gaussian Games with a control-sharing information pattern is anchored in Sect. 8.4 where the end-game is solved and the solution of the LQGDG with a control-sharing information pattern is obtained in Sect. 8.5 using the method of backward induction/DP. The results are summarized in Sect. 8.6, followed by concluding remarks in Sect. 8.7. For the sake of completeness, the solution of the baseline deterministic LQDG game with perfect information (Pachter and Pham 2010) is included in the Appendix. The somewhat lengthy exposition could perhaps be excused in light of Witsenhausen's observation when discussing LQG control (Witsenhausen 1971): "The most confused derivations of the correct results are also among the shortest".

8.2 Linear Quadratic Gaussian Dynamic Game

Two player Linear Quadratic Gaussian Dynamic Games (LQGDGs) are considered. The players are designated P and E, and the game is specified as follows.

Dynamics: Linear

$$x_{k+1} = A_k x_k + B_k u_k + C_k v_k + \Gamma_k w_k,\ x_0 \equiv x_0,\ \ k = 0, \ldots, N-1 \qquad (8.1)$$

At decision time k the controls of players P and E are u_k and v_k, respectively. The process noise $w_k \sim \mathcal{N}(0, Q_p),\ \ k = 0, \ldots, N-1$. The planning horizon is N.

Measurements: Linear

The N measurements of player P are:

At time $k = 0$, $\overline{x}_0^{(P)}$— player P believes that the initial state

$$x_0 \sim \mathcal{N}(\overline{x}_0^{(P)}, P_0^{(P)}), \tag{8.2}$$

and thereafter he takes the measurements

$$z_{k+1}^{(P)} = H_{k+1}^{(P)} x_{k+1} + v_{k+1}^{(P)}, \;\; v_{k+1}^{(P)} \sim \mathcal{N}(0, R_m^{(P)}), \; k = 0, \ldots, N-2 \tag{8.3}$$

The N measurements of player E are:
At time $k = 0$, $\overline{x}_0^{(E)}$— player E believes that the initial state

$$x_0 \sim \mathcal{N}(\overline{x}_0^{E)}, P_0^{(E)}), \tag{8.4}$$

and thereafter he takes the measurements

$$z_{k+1}^{(E)} = H_{k+1}^{(E)} x_{k+1} + v_{k+1}^{(E)}, \;\; v_{k+1}^{(E)} \sim \mathcal{N}(0, R_m^{(E)}), \; k = 0, \ldots, N-2 \tag{8.5}$$

Cost/Payoff Function: Quadratic

We confine our attention to the antagonistic "zero-sum" game scenario where the respective P and E players strive to minimize and maximize the cost/payoff function

$$J = x_N^T Q_F x_N + \sum_{k=0}^{N-1} [x_{k+1}^T Q_{k+1} x_{k+1} + u_k^T R_k^{(P)} u_k - v_k^T R_k^{(E)} v_k] \rightarrow min_{\{u_k\}_{k=0}^{N-1}} max_{\{v_k\}_{k=0}^{N-1}} \tag{8.6}$$

Specifically, players P and E minimize and maximize their expected cost/payoff $E(J \mid \cdot)$, conditional on their private information. The expectation operator is liberally used in the dynamics game literature but oftentimes it is not clearly stated with respect to which random variables the expectation is calculated and on which random variables the expectation is conditional. This tends to mask the fact that what appear to be "zero-sum" games are in fact nonzero-sum games. Upon considering "zero-sum" games with partial information, the illusion is then created that a zero-sum game is considered. One then tends to rely on the uniqueness of the saddle point value and the interchangeability of non-unique optimal saddle point strategies in zero-sum games. This argument is flawed because, as previously discussed, in "zero-sum" games with partial information the P and E players calculate their respective cost and payoff conditional on their private information, as is correctly done in this paper; that's why I put the term zero-sum in quotation marks. Thus, although high powered mathematics is oftentimes used, serious conceptual errors make the "results" not applicable. Contrary to statements sometimes encountered in the literature, in "zero-sum" games with partial information one cannot look for a saddle point solution and the correct solution concept is a Nash equilibrium, that is, Person by Person Satisfactory (PBPS) solution. In this paper a unique Nash equilibrium is provided and the P and E players' value functions are calculated.

Information pattern

1. Public information
   (a) Problem parameters: A_k, B_k, C_k, $H_k^{(P)}$, $H_k^{(E)}$, Q_p, Q_k, Q_F, $R_k^{(P)}$, $R_k^{(E)}$, $R_m^{(P)}$, $R_m^{(E)}$.
   (b) Prior information: $P_0^{(P)}$, $P_0^{(E)}$.
2. Private information
   At decision time $k = 0$ the prior information of player P is $\bar{x}_0^{(P)}$.
   At decision time $k = 0$ the prior information of player E is $\bar{x}_0^{(E)}$.
   At decision time $1 \leq k \leq N-1$ the information of player P are his measurements $\bar{x}_0^{(P)}, z_1^{(P)}, \ldots, z_k^{(P)}$ and ownship control history $u_0, \ldots, u_{k-1}$.
   At decision time $1 \leq k \leq N-1$ the information of player E are his measurements $\bar{x}_0^{(E)}, z_1^{(E)}, \ldots, z_k^{(E)}$ and ownship control history $v_0, \ldots, v_{k-1}$.

Sufficient statistics

The sufficient statistics of player P at decision time $k = 0$: $x_0 \sim \mathcal{N}(\bar{x}_0^{(P)}, P_0^{(P)})$.
The sufficient statistics of player E at decision time $k = 0$: $x_0 \sim \mathcal{N}(\bar{x}_0^{(E)}, P_0^{(E)})$.
The sufficient statistics of player P at decision time $1 \leq k \leq N-1$: The p.d.f. $f_k^{(P)}(\cdot)$ of the physical state x_k, as calculated by player P using his private information.
The sufficient statistics of player E at decision time $1 \leq k \leq N-1$: The p.d.f. $f_k^{(E)}(\cdot)$ of the physical state x_k, as calculated by player E using his private information.

Remark. In the static LQGDG (Pachter 2013) where $N = 1$ the respective sufficient statistics of P and E are $\bar{x}_0^{(P)}$ and $\bar{x}_0^{(E)}$.

8.2.1 Problem Statement

The LQGDG (8.1)–(8.6) is considered and it is assumed that the players' information sets are augmented as follows: At decision time k, $k = 1, \ldots, N-1$, player P is endowed with the additional information regarding the control history $v_0, \ldots, v_{k-1}$ of player E. Thus, player P observed the past inputs $v_0, \ldots, v_{k-1}$ of player E. Similarly, at decision time k, $k = 1, \ldots, N-1$, player E is endowed with the additional information regarding the control history $u_0, \ldots, u_{k-1}$ of player P. Thus, player E observed the past inputs $u_0, \ldots, u_{k-1}$ of player P.

The information pattern considered herein is referred to as the *control-sharing* information pattern. The dynamics and the measurement equations are linear, and the cost/payoff function is quadratic, but the information pattern is not classical.

Strictly speaking, the information pattern is not partially nested because E's measurements, which he used to form his controls, are not known to P, and vice versa, P's measurements, which he used to form his controls, are not known to E. However, this is now a moot point because the information pattern is s.t. the control history of player E is known to player P, and vice versa, the control history of player P is known to player E. This, and the fact that player P and player E, each separately, perceive the initial state x_0 to be Gaussian, causes the state estimation problem faced by the players at decision time k to be Linear and Gaussian (LG). Hence, at decision time k, the knowledge of the complete control history $u_0, v_0, \ldots, u_{k-1}, v_{k-1}$ and their private measurement records makes it possible for both players to separately apply the linear Kalman filtering algorithm: based on his private measurements record, each player runs a Kalman filter using his measurements and the complete input history, and separately obtains an estimate of the state x_k—strictly speaking, the p.d.f. of x_k is separately obtained by each player. Thus, players P and E perceive the current state x_k to be Gaussian distributed. Having run their respective Kalman filters, at time k player P believes that the state

$$x_k \sim \mathscr{N}(\overline{x}_k^{(P)}, P_k^{(P)}) , \quad \forall\, k, \;\; N-1 \geq k \geq 0 \tag{8.7}$$

and player E believes that the state

$$x_k \sim \mathscr{N}(\overline{x}_k^{(E)}, P_k^{(E)}) , \quad \forall\, k, \;\; N-1 \geq k \geq 0 \tag{8.8}$$

but they are also aware that their state estimates are correlated—see Sect. 8.3.

Since the LQGDG (8.1)–(8.6) is LG, the P and E players' separately calculated sufficient statistics are given by Eqs. (8.7) and (8.8), and their controls will be determined by their optimal strategies according to $u_k^* = (\gamma_k^{(P)}(\overline{x}_k^{(P)}, P_k^{(P)}))^*$ and $v_k^* = (\gamma_k^{(E)}(\overline{x}_k^{(E)}, P_k^{(E)}))^*$. In fact, we shall show that in LQGDGs with a control-sharing information pattern the optimal strategies are of the form

$$u_k^* = (\gamma_k^{(P)}(\overline{x}_k^{(P)}))^* , \tag{8.9}$$

$$v_k^* = (\gamma_k^{(E)}(\overline{x}_k^{(E)}))^* , \quad \forall\, k, \;\; 0 \leq k \leq N-1 \tag{8.10}$$

and are linear.

8.3 Kalman Filtering

LQGDGs with a control-sharing information pattern are Linear-Gaussian (LG) and consequently at decision time k each player can separately calculate his estimate of the physical state x_k using a *linear* Kalman Filter (KF). Player P runs the KF

$$(\bar{x}_k^{(P)})^- = A\bar{x}_{k-1}^{(P)} + Bu_{k-1} + Cv_{k-1}, \;\; \bar{x}_0^{(P)} \equiv \bar{x}_0^{(P)} \tag{8.11}$$

$$(P_k^{(P)})^- = AP_{k-1}^{(P)}A^T + \Gamma Q_p \Gamma^T, \;\; P_0^{(P)} \equiv P_0^{(P)} \tag{8.12}$$

$$K_k^{(P)} = (P_k^{(P)})^-(H^{(P)})^T[H^{(P)}(P_k^{(P)})^-(H^{(P)})^T + R_m^{(P)}]^{-1} \tag{8.13}$$

$$\bar{x}_k^{(P)} = (\bar{x}_k^{(P)})^- + K_k^{(P)}[z_k^{(P)} - H^{(P)}(\bar{x}_k^{(P)})^-] \tag{8.14}$$

$$P_k^{(P)} = (I - K_k^{(P)}H^{(P)})(P_k^{(P)})^- \tag{8.15}$$

and so at decision time k player P obtains his estimate $\bar{x}_k^{(P)}$ of the state x_k. Similarly, player E runs the KF

$$(\bar{x}_k^{(E)})^- = A\bar{x}_{k-1}^{(E)} + Bu_{k-1} + Cv_{k-1}, \;\; \bar{x}_0^{(E)} \equiv \bar{x}_0^{(E)} \tag{8.16}$$

$$(P_k^{(E)})^- = AP_{k-1}^{(E)}A^T + \Gamma Q_p \Gamma^T, \;\; P_0^{(E)} \equiv P_0^{(E)} \tag{8.17}$$

$$K_k^{(E)} = (P_k^{(E)})^-(H^{(E)})^T[H^{(E)}(P_k^{(E)})^-(H^{(E)})^T + R_m^{(E)}]^{-1} \tag{8.18}$$

$$\bar{x}_k^{(E)} = (\bar{x}_k^{(E)})^- + K_k^{(E)}[z_k^{(E)} - H^{(E)}(\bar{x}_k^{(E)})^-] \tag{8.19}$$

$$P_k^{(E)} = (I - K_k^{(E)}H^{(E)})(P_k^{(E)})^- \tag{8.20}$$

and so at decision time k player E obtains his estimate $\bar{x}_k^{(E)}$ of the state x_k. The P and E players can calculate their respective state estimation error covariances and Kalman gains $P_k^{(P)}$, $K_{k+1}^{(P)}$, $P_k^{(E)}$ and $K_{k+1}^{(E)}$ ahead of time and off line.

In LQGDGs with a control-sharing information pattern the players' sufficient statistic is their state estimate; the latter is the argument of their strategy functions (8.9) and (8.10). Hence, in the process of countering E's action, P must compute the *statistics* of E's state estimate $\bar{x}_k^{(E)}$, and, vice versa, while planning his move, E must compute the *statistics* of P's state estimate $\bar{x}_k^{(P)}$. Momentarily assume the point of view of player P: As far as P is concerned, the unknown to him state estimate of player E at time k, $\bar{x}_k^{(E)}$, is a random variable (and consequently E's input at time k is a random variable). Similarly, player E will consider the unknown to him state estimate of player P at time k, $\bar{x}_k^{(P)}$, to be a random variable (and consequently P's input at time k is a random variable). Hence, in the LQGDG with a control-sharing information pattern, at time k player P will estimate E's state estimate $\bar{x}_k^{(E)}$ using his calculated ownship state estimate $\bar{x}_k^{(P)}$, and vice versa, player E will estimate P's state estimate $\bar{x}_k^{(P)}$ using his calculated ownship state estimate $\bar{x}_k^{(E)}$. Thus, in the LQGDG with a control-sharing information pattern and with his state estimate $\bar{x}_k^{(P)}$ at time k in hand, player P calculates the statistics of E's state estimate $\bar{x}_k^{(E)}$, conditional on the public and private information available to him at time k. Similarly, having obtained at time k his state estimate $\bar{x}_k^{(E)}$, player E calculates the statistics of the state estimate $\bar{x}_k^{(P)}$ of player P, conditional on the public and private information available to him at time k. Let's start at decision time $k = 0$.

Player P models his measurement/estimate $\overline{x}_0^{(P)}$ of the initial state x_0 as

$$\overline{x}_0^{(P)} = x_0 + e_0^{(P)} , \tag{8.21}$$

where x_0 is the true physical state and $e_0^{(P)}$ is player P's measurement/estimation error, whose statistics, in view of Eq. (8.2), are $e_0^{(P)} \sim \mathcal{N}(0, P_0^{(P)})$. In addition, player P models player E's measurement $\overline{x}_0^{(E)}$ of the initial state x_0 as

$$\overline{x}_0^{(E)} = x_0 + e_0^{(E)} , \tag{8.22}$$

where, as before, x_0 is the true physical state and $e_0^{(E)}$ is player E's measurement/estimation error, whose statistics, which are known to P—see Eq. (8.4)—are $e_0^{(E)} \sim \mathcal{N}(0, P_0^{(E)})$. The Gaussian random variables $e_0^{(P)}$ and $e_0^{(E)}$ are *independent*—by hypothesis. From player P's point of view, $\overline{x}_0^{(P)}$ is known but $\overline{x}_0^{(E)}$ is a random variable. Subtracting Eq. (8.21) from Eq. (8.22), at time $k = 0$ player P concludes that as far as he is concerned, player E's measurement upon which he will decide, according to Eq. (8.10), on his optimal control v_0^*, is the random variable

$$\overline{x}_0^{(E)} = \overline{x}_0^{(P)} + e_0^{(E)} - e_0^{(P)} , \tag{8.23}$$

In other words, as far as P is concerned, E's estimate $\overline{x}_0^{(E)}$ of the initial state x_0 is the Gaussian random variable

$$\overline{x}_0^{(E)} \sim \mathcal{N}(\overline{x}_0^{(P)}, P_0^{(P)} + P_0^{(E)}) \tag{8.24}$$

Thus, player P has used his measurement/private information $\overline{x}_0^{(P)}$ and the public information $P_0^{(P)}$ and $P_0^{(E)}$ to calculate the statistics of the sufficient statistic $\overline{x}_0^{(E)}$ of player E, which is the argument of E's strategy function $\gamma_0^{(E)}(\cdot)$; the latter, along with P's control u_0, will feature in player's P cost functional. Similarly, as far as player E is concerned, at time $k = 0$ the statistics of the sufficient statistic $\overline{x}_0^{(P)}$ of player P are

$$\overline{x}_0^{(P)} \sim \mathcal{N}(\overline{x}_0^{(E)}, P_0^{(P)} + P_0^{(E)}) \tag{8.25}$$

Similar to the case where $k = 0$, as far as player P is concerned the state estimate of player E at decision time $k \geq 1$ is the random variable

$$\overline{x}_k^{(E)} = \overline{x}_k^{(P)} + e_k^{(E)} - e_k^{(P)},$$

that is, at decision time k player P believes that the state estimate $\overline{x}_k^{(E)}$ of player E is

$$\overline{x}_k^{(E)} \sim \mathcal{N}(\overline{x}_k^{(P)}, P_k^{(E,P)}) \tag{8.26}$$

where the covariance matrix

$$P_k^{(E,P)} \equiv E(\ (e_k^{(E)} - e_k^{(P)})(e_k^{(E)} - e_k^{(P)})^T\ )$$
$$= P_k^{(P)} + P_k^{(E)} - E(\ e_k^{(P)}(e_k^{(E)})^T\ ) - (E(\ e_k^{(P)}(e_k^{(E)})^T\ ))^T$$

At the decision time instants $k = 1, \ldots, N-1$ the P and E players' respective state estimation errors $e_k^{(P)}$ and $e_k^{(E)}$ are now correlated—this is caused by the process dynamics being driven in part by process noise.

Similarly, as far as he is concerned, player E believes that at decision time k the state estimate $\bar{x}_k^{(P)}$ of player P is the random variable

$$\bar{x}_k^{(P)} \sim \mathcal{N}(\bar{x}_k^{(E)}, P_k^{(E,P)}) \tag{8.27}$$

Concerning decision time $k \geq 1$: Let the covariance matrix

$$\tilde{P}_k^{(E,P)} \equiv E(\ e_k^{(P)}(e_k^{(E)})^T\ ) \tag{8.28}$$

It can be shown that the recursion for the correlation matrix $\tilde{P}_k^{(E,P)}$ is

$$\tilde{P}_{k+1}^{(E,P)} = (I - K_{k+1}^{(P)}H^{(P)})(A\tilde{P}_k^{(P,E)}A^T + \Gamma Q_p \Gamma^T)(I - K_{k+1}^{(E)}H^{(E)})^T,\ \tilde{P}_0^{(P,E)} = 0,$$
$$k = 0, \ldots, N-1 \tag{8.29}$$

In summary, at decision time $k = 0, \ldots, N-1$ player P believes that the statistics of E's estimate $\bar{x}_k^{(E)}$ of the state x_k are given by Eq. (8.26) and player E believes that the statistics of P's estimate $\bar{x}_k^{(P)}$ of the state x_k are given by Eq. (8.27) where

$$P_k^{(E,P)} = P_k^{(P)} + P_k^{(E)} - \tilde{P}_k^{(E,P)} - (\tilde{P}_k^{(E,P)})^T$$

The KF covariance matrices $P_k^{(P)}$, $P_k^{(E)}$ and $\tilde{P}_k^{(E,P)}$ are calculated ahead of time by solving the respective recursions (8.12), (8.13), (8.15); (8.17), (8.18), (8.20); and (8.29).

Finally, since in LQGDGs with a control-sharing information pattern the sufficient statistic is the players' state estimate, then upon employing the method of Dynamic Programming, at decision time k player P must project ahead the *estimate* of the physical state x_{k+1} that the Kalman filtering algorithm will provide at time $k+1$. It can be shown that at time k player P believes that the future state x_{k+1} at time $k+1$ will be the Gaussian random variable

$$\bar{x}_{k+1}^{(P)} = A\bar{x}_k^{(P)} + Bu_k + C\gamma_k^{(E)}(\bar{x}_k^{(P)} + e_k^{(P)} - e_k^{(E)}) + K_{k+1}^{(P)}(H^{(P)}\Gamma w_k$$
$$+ v_{k+1}^{(P)} - H^{(P)}Ae_k^{(P)}) \tag{8.30}$$

Similarly, at decision time k player E's estimate of the state x_{k+1} at time $k+1$ will be the Gaussian random variable

$$\bar{x}_{k+1}^{(E)} = A\bar{x}_k^{(E)} + B\gamma_k^{(P)}(\bar{x}_k^{(E)} + e_k^{(E)} - e_k^{(P)}) + Cv_k + K_{k+1}^{(E)}(H^{(E)}\Gamma w_k + v_{k+1}^{(E)} - H^{(E)}Ae_k^{(E)}) \quad (8.31)$$

8.4 End Game

In the best tradition of backward induction/Dynamic Programming, the terminal stage of the game, namely, the players' decision time $k = N-1$ is analyzed first. In the end game the cost/payoff function is

$$\begin{aligned} J_{N-1}(u_{N-1}, v_{N-1}; x_{N-1}) &= x_N^T Q_F x_N + x_N^T Q x_N + u_{N-1}^T R_u u_{N-1} - v_{N-1}^T R_v v_{N-1} \\ &= x_N^T (Q + Q_F) x_N + u_{N-1}^T R_u u_{N-1} - v_{N-1}^T R_v v_{N-1} \end{aligned}$$

It is convenient to momentarily set $Q_F := Q + Q_F$ whereupon the terminal cost/payoff

$$J_{N-1}(u_{N-1}, v_{N-1}; x_{N-1}) = x_N^T Q_F x_N + u_{N-1}^T R_u u_{N-1} - v_{N-1}^T R_v v_{N-1} \quad (8.32)$$

The players' sufficient statistics in this LG game are the expectation of the physical state and the covariance of the state's estimation error: having run their respective Kalman filters during the time interval $[1, N-1]$, at decision time $N-1$ the information available to player P is $(\bar{x}_{N-1}^{(P)}, P_{N-1}^{(P)})$ and the information of player E is $(\bar{x}_{N-1}^{(E)}, P_{N-1}^{(E)})$. In other words, at decision time $N-1$ player P believes the physical state x_{N-1} to be $x_{N-1} \sim \mathcal{N}(\bar{x}_{N-1}^{(P)}, P_{N-1}^{(P)})$ whereas player E believes the physical state x_{N-1} to be specified as $x_{N-1} \sim \mathcal{N}(\bar{x}_{N-1}^{(E)}, P_{N-1}^{(E)})$. This is tantamount to stipulating that players P and E took separate measurements of the state x_{N-1}. The quality of the players' "instruments" used to take the measurements and also the degree of correlation of the players' measurement errors is public knowledge—we refer to the measurement error covariances $P_{N-1}^{(E)}$, $P_{N-1}^{(E)}$ and $\tilde{P}_{N-1}^{(E,P)}$. At the same time, the recorded measurements $\bar{x}_{N-1}^{(P)}$and $\bar{x}_{N-1}^{(E)}$ are the private information of the respective players P and E: the "measurement" $\bar{x}_{N-1}^{(E)}$ recorded by player E is his private information and is not shared with player P. Thus, player P has partial information. Similarly, the "measurement" $\bar{x}_{N-1}^{(P)}$ recorded by player P is his private information and is not shared with player E, so also player E has partial information.

To gain a better appreciation of the informational issues in games with partial information, it is instructive to briefly digress and employ an "appealing" approach which is familiar to workers in deterministic control and which, unfortunately, is an approach sometimes employed in stochastic games. We now intentionally "take a wrong turn" which quickly leads us to a dead end. A correct analysis of the informational situation at hand follows.

Consider the following flawed argument: At time $N-1$ the state information available to player P is $x_{N-1} \sim \mathcal{N}(\bar{x}^{(P)}_{N-1}, P^{(P)}_{N-1})$ and thus Player P calculates the expectation of his cost function

$$\begin{aligned}\bar{J}^{(P)}_{N-1}(u_{N-1}, v_{N-1}; \bar{x}^{(P)}_{N-1}, P^{(P)}_{N-1}) &\equiv E_{x_{N-1}}\left(J(u_{N-1}, v_{N-1}; x_{N-1}) \mid \bar{x}^{(P)}_{N-1}, P^{(P)}_{N-1}\right)\\ &= (\bar{x}^{(P)})^T_{N-1} A^T Q_F A \bar{x}^{(P)}_{N-1} + Trace(A^T Q_F A P^{(P)}_{N-1})\\ &+ u^T_{N-1}(R_u + B^T Q_F B) u_{N-1} - v^T_{N-1}(R_v - C^T Q_F C) v_{N-1}\\ &+ 2u^T_{N-1} B^T Q_F A \bar{x}^{(P)}_{N-1} + 2v^T_{N-1} C^T Q_F A \bar{x}^{(P)}_{N-1}\\ &+ 2u^T_{N-1} B^T Q_F C v_{N-1} + Trace(\Gamma^T Q_F \Gamma Q_p)\end{aligned} \tag{8.33}$$

At the same time the state information available to player E is $x_{N-1} \sim \mathcal{N}(\bar{x}^{(E)}_{N-1}, P^{(E)}_{N-1})$ and Player E calculates the expectation of his payoff function

$$\begin{aligned}\bar{J}^{(E)}_{N-1}(u_{N-1}, v_{N-1}; \bar{x}^{(E)}_{N-1}, P^{(E)}_{N-1}) &\equiv E_{x_{N-1}}\left(J(u_{N-1}, v_{N-1}; x_{N-1}) \mid \bar{x}^{(E)}_{N-1}, P^{(E)}_{N-1}\right)\\ &= (\bar{x}^{(E)})^T_{N-1} A^T Q_F A \bar{x}^{(E)}_{N-1} + Trace(A^T Q_F A P^{(E)}_{N-1})\\ &+ u^T_{N-1}(R_u + B^T Q_F B) u_{N-1} - v^T_{N-1}(R_v - C^T Q_F C) v_{N-1}\\ &+ 2u^T_{N-1} B^T Q_F A \bar{x}^{(E)}_{N-1} + 2v^T_{N-1} C^T Q_F A \bar{x}^{(E)}_{N-1}\\ &+ 2u^T_{N-1} B^T Q_F C v_{N-1} + Trace(\Gamma^T Q_F \Gamma Q_p)\end{aligned} \tag{8.34}$$

Now Player P's optimization, that is, the differentiation of his *deterministic* cost function (8.33), yields the relationship

$$(R_u + B^T Q_F B) u_{N-1} + B^T Q_F C v_{N-1} = -B^T Q_F A \bar{x}^{(P)}_{N-1} \tag{8.35}$$

and Player E's optimization, that is, the differentiation of his *deterministic* payoff function (8.34), yields the relationship

$$C^T Q_F B u_{N-1} - (R_v - C^T Q_F C) v_{N-1} = -C^T Q_F A \bar{x}^{(E)}_{N-1} \tag{8.36}$$

Have obtained two equations in the players' optimal *controls*, namely, the two unknowns u^*_{N-1} and v^*_{N-1}, which players P and E must *separately* solve in order to calculate their respective optimal controls. However player P cannot solve the set of two equations (8.35) and (8.36) because he does *not* know the "measurement" $\bar{x}^{(E)}_{N-1}$ of E, and player E cannot solve this set of two equations because he does *not* know the "measurement" $\bar{x}^{(P)}_{N-1}$ of P—both players have reached a dead end and it would appear that all that's left to do is try to guess and outguess the opponent's "measurement". This state of affairs is caused by the players having partial information. This approach brings on the much maligned infinite regress in reciprocal reasoning! Unfortunately, this flawed approach is not foreign to the literature on dynamic stochastic games and it leads to erroneous "results"—see Aoki (1973) where, using this flawed argument, the LQGDG with a shared-control information pattern was "solved" and complicated "strategies" were computed.

We now change course and undertake a correct analysis of our LQGDG with a shared-control information pattern. To this end, it is imperative that one thinks in *strategic* terms. The strategies available to player P are mappings $f : R^n \to R^{m_u}$ from his information set into his actions set; thus, the action of player P is $u_{N-1} = f(\overline{x}^{(P)}_{N-1}, P^{(P)}_{N-1})$. Similarly, the strategies available to player E are mappings $g : R^n \to R^{m_v}$ from his information set into his actions set—thus, the action of player E is $v_{N-1} = g(\overline{x}^{(E)}_{N-1}, P^{(E)}_{N-1})$. However, we'll show in the sequel that it suffices to consider P and E strategies of the form (8.9) and (8.10), respectively.

It is now important to realize that from player P's vantage point, the action v_{N-1} of player E is a *random* variable. This is so because as far as player P is concerned the measurement $\overline{x}^{(E)}_{N-1}$ of player E used in (8.10) to form his control v_{N-1} is a random variable. Similarly, from player E's vantage point, the action u_{N-1} of player P is also a function of a random variable, $\overline{x}^{(P)}_{N-1}$.

Consider the decision process of player P whose private information is $\overline{x}^{(P)}_{N-1}$. He operates against the *strategy* $g(\cdot)$ of player E. Therefore, from player P's perspective, the random variables at work are x_{N-1} *and* $\overline{x}^{(E)}_{N-1}$. At decision time $k = N - 1$ player P is confronted with a *stochastic* optimization problem and he calculates the *expectation* of the cost function (8.32), conditional on his private information $\overline{x}^{(P)}_{N-1}$,

$$\overline{J}^{(P)}(u_{N-1}, g(\cdot); \overline{x}^{(P)}_{N-1}) \equiv E_{x_{N-1}, \overline{x}^{(E)}_{N-1}}(J(u_{N-1}, g(\overline{x}^{(E)}_{N-1}); x_{N-1}) \mid \overline{x}^{(P)}_{N-1}) \quad (8.37)$$

By correctly using in the calculation of his expected cost (8.37) player's E *strategy function* $g(\overline{x}^{(E)}_{N-1})$ rather than, as before, player E's *control* v_{N-1}, player P has eliminated the possibility of an infinite regress in reciprocal reasoning. This is so because P now has all the information to be able, in principle, to calculate the said expectation. Thus, player P calculates his expected cost

$$\begin{aligned}
\overline{J}^{(P)}(u_{N-1}, g(\cdot); \overline{x}^{(P)}_{N-1}) = {} & (\overline{x}^{(P)}_{N-1})^T A^T Q_F A \overline{x}^{(P)}_{N-1} + Trace(A^T Q_F A P^{(P)}_{N-1}) \\
& + u^T_{N-1}(R_u + B^T Q_F B) u_{N-1} + 2u^T_{N-1} B^T Q_F A \overline{x}^{(P)}_{N-1} \\
& + 2E_{x_{N-1}, \overline{x}^{(E)}_{N-1}}(g^T(\overline{x}^{(E)}_{N-1}) C^T Q_F A x_{N-1} \mid \overline{x}^{(P)}_{N-1}) \\
& - E_{\overline{x}^{(E)}_{N-1}}(g^T(\overline{x}^{(E)}_{N-1})(R_v - C^T Q_F C) g(\overline{x}^{(E)}_{N-1}) \mid \overline{x}^{(P)}_{N-1}) \\
& + 2u^T_{N-1} B^T Q_F C E_{\overline{x}^{(E)}_{N-1}}(g(\overline{x}^{(E)}_{N-1}) \mid \overline{x}^{(P)}_{N-1}) \\
& + Trace(\Gamma^T Q_F \Gamma Q_p) \qquad (8.38)
\end{aligned}$$

Player P calculates the expectations with respect to the random variable $\bar{x}_{N-1}^{(E)}$ which features in Eq. (8.38), cognizant that it is $\bar{x}_{N-1}^{(E)} \sim \mathscr{N}(\bar{x}_{N-1}^{(P)}, P_{N-1}^{(E,P)})$. In this game with partial information, player P is using his measurement/private information $\bar{x}_{N-1}^{(P)}$ and the public information to estimate the sufficient statistic $\bar{x}_{N-1}^{(E)}$ of player E, which is the argument of E's strategy function $g(\cdot)$; the latter features in player's P cost functional (8.38) and thus enters the calculation of P's cost.

The careful analysis of the optimization problem at hand leads to a Fredholm equation of the second kind of the convolution type with a kernel which is a Gaussian function; the unknown functions are the players' optimal strategies. Taking the point of view of player E yields a similar Fredholm integral equation in the players' optimal strategies. The solution of the set of two Fredholm equations yields the optimal strategies of players P and E. The optimal strategies turn out to be linear after all! The reader is referred to reference Pachter (2013) for the complete derivation.

8.5 Dynamic Programming

We consider the LQGDG (8.1)–(8.6) with a control-sharing information pattern as in Aoki (1973). The planning horizon $N \geq 2$.

8.5.1 Sufficient Statistics

The initial state information and the measurements of players P and E are their private information but their past controls are shared information. Even though the players have partial information because the initial state information and their measurements are not shared, from the point of view of both players P and E, the control system is nevertheless Linear Gaussian (LG). This is so because at decision time k their respective adversary's information state components $v_0, \ldots, v_{k-1}$ and $u_0, \ldots, u_{k-1}$ are *not* random variables with unknown p.d.f.s but are known to the players: The LQGDG with a control-sharing information pattern is LG and therefore the conditions for the P-player's information state to be Gaussian hold and at decision time k the sufficient statistics of P and E are $\bar{x}_k^{(P)}$ and $\bar{x}_k^{(E)}$, respectively. Furthermore, as far as player P is concerned, at time k the sufficient statistic $\bar{x}_k^{(E)}$ of player E is the random variable $\bar{x}_k^{(E)} \sim \mathscr{N}(\bar{x}_k^{(P)}, P_k^{(E,P)})$ and he uses this information in the calculation of his cost-to-go/value function at time k. Similarly, player E considers the sufficient statistic $\bar{x}_k^{(P)}$ of player P to be $\bar{x}_k^{(P)} \sim \mathscr{N}(\bar{x}_k^{(E)}, P_k^{(E,P)})$ and player E uses this information in the calculation of his cost-to-go/value function at time k.

8.5.2 Analysis

The analysis is along the lines of the analysis of the static LQG game with partial information (Pachter 2013) and the analysis of the end game in Sect. 8.4 where $k = N - 1$. We shall require

Proposition 1. *The value functions of players P and E are quadratic in their respective sufficient statistics* $\bar{x}_k^{(P)}$ *and* $\bar{x}_k^{(E)}$*, that is*

$$V_k^{(P)}(\bar{x}_k^{(P)}, P_k^{(P)}; P_k^{(E)}, \tilde{P}_k^{(E,P)}) = (\bar{x}_k^{(P)})^T \Pi_k \bar{x}_k^{(P)} + c_k^{(P)}(P_k^{(P)}; P_k^{(E)}, \tilde{P}_k^{(E,P)}),\ \ k = 0, 1, \ldots, N-1,$$
$$V_k^{(E)}(\bar{x}_k^{(E)}, P_k^{(E)}; P_k^{(P)}, \tilde{P}_k^{(E,P)}) = (\bar{x}_k^{(E)})^T \Pi_k \bar{x}_k^{(E)} + c_k^{(E)}(P_k^{(E)}; P_k^{(P)}, \tilde{P}_k^{(E,P)}),\ \ k = 0, 1, \ldots, N-1$$

where

Π_k *are* $n \times n$ *real symmetric matrices and the scalars* $c_k^{(P)}, c_k^{(E)} \in R^1,\ k = 0, \ldots, N.$

□

Similar to the *correct* approach outlined in Sect. 8.4 we calculate the value functions by taking the expectations over the relevant random variables.

$$\begin{aligned}
V_k^{(P)}(\bar{x}_k^{(P)}, P_k^{(P)}; P_k^{(E)}, \tilde{P}_k^{(E,P)}) = &\ min_{u_k}\{u_k^T[R_u + B^T(Q + \Pi_{k+1})B]u_k \\
&\ +2u_k^T B^T(Q + \Pi_{k+1})(A\bar{x}_k^{(P)} \\
&+ CE_{\tilde{w}}\ (\gamma_k^{(E)}(\bar{x}_k^{(P)} + \tilde{w})))\} + (\bar{x}_k^{(P)})^T A^T(Q + \Pi_{k+1})A\bar{x}_k^{(P)} \\
&- E_{\tilde{w}}\ ((\gamma_k^{(E)}(\bar{x}_k^{(P)} + \tilde{w}))^T \\
&\ [R_v - C^T(Q + \Pi_{k+1})C]\gamma_k^{(E)}(\bar{x}_k^{(P)} + \tilde{w})) \\
&+ 2E_{\tilde{w}}\ ((\gamma_k^{(E)}(\bar{x}_k^{(P)} + \tilde{w}))^T)C^T(Q + \Pi_{k+1})A\bar{x}_k^{(P)} \\
&- 2E_{e_k^{(E)}, e_k^{(P)}}\ ((\gamma_k^{(E)}(\bar{x}_k^{(P)} + e_k^{(E)} - e_k^{(P)}))^T C^T(Q \\
&+ \Pi_{k+1}K_{k+1}^{(P)}H^{(P)})Ae_k^{(P)}) + Trace(A^T QAP_k^{(P)}) \\
&+ Trace(\Gamma^T Q\Gamma Q_p) + Trace((K_{k+1}^{(P)})^T \Pi_{k+1}K_{k+1}^{(P)}R_m^{(P)}) \\
&+ Trace(\Gamma^T(H^{(P)})^T(K_{k+1}^{(P)})^T \Pi_{k+1}K_{k+1}^{(P)}H^{(P)}\Gamma Q_p) \\
&+ Trace(A^T(H^{(P)})^T(K_{k+1}^{(P)})^T \Pi_{k+1}K_{k+1}^{(P)}H^{(P)}AP_k^{(P)}) \\
&+ c_{k+1}^{(P)}(P_{k+1}^{(P)}; P_{k+1}^{(E)}, \tilde{P}_{k+1}^{(E,P)}),
\end{aligned} \tag{8.39}$$

$$
\begin{aligned}
V_k^{(E)}(\overline{x}_k^{(E)}, P_k^{(E)}; P_k^{(P)}, \tilde{P}_k^{(E,P)}) = & \, max_{v_k}\{-v_k^T[R_v - C^T(Q+\Pi_{k+1})C]v_k \\
& +2v_k^T C^T(Q+\Pi_{k+1})(A\overline{x}_k^{(E)} \\
& + BE_{\tilde{w}}\,(\gamma_k^{(P)}(\overline{x}_k^{(E)} - \tilde{w})))\} + (\overline{x}_k^{(E)})^T A^T(Q+\Pi_{k+1})A\overline{x}_k^{(E)} \\
& + E_{\tilde{w}}\,((\gamma_k^{(P)}(\overline{x}_k^{(E)} - \tilde{w}))^T \\
& [R_u + B^T(Q+\Pi_{k+1})B]\gamma_k^{(P)}(\overline{x}_k^{(E)} - \tilde{w})) \\
& + 2E_{\tilde{w}}\,((\gamma_k^{(P)}(\overline{x}_k^{(E)} - \tilde{w}))^T)B^T(Q+\Pi_{k+1})A\overline{x}_k^{(E)} \\
& - 2E_{e_k^{(E)}, e_k^{(P)}}\,((\gamma_k^{(P)}(\overline{x}_k^{(E)} - e_k^{(E)} + e_k^{(P)}))^T B^T(Q \\
& + \Pi_{k+1}K_{k+1}^{(E)}H^{(E)})Ae_k^{(E)}) + Trace(A^T QAP_k^{(E)}) \\
& + Trace(\Gamma^T Q\Gamma Q_p) + Trace((K_{k+1}^{(E)})^T\Pi_{k+1}K_{k+1}^{(E)}R_m^{(E)}) \\
& + Trace(\Gamma^T(H^{(E)})^T(K_{k+1}^{(E)})^T P_{k+1}K_{k+1}^{(E)}H^{(E)}\Gamma Q_p) \\
& + Trace(A^T(H^{(E)})^T(K_{k+1}^{(E)})^T\Pi_{k+1}K_{k+1}^{(E)}H^{(E)}AP_k^{(E)}) \\
& + c_{k+1}^{(E)}(P_{k+1}^{(P)}; P_{k+1}^{(E)}, \tilde{P}_{k+1}^{(E,P)})
\end{aligned}
\tag{8.40}
$$

where the random variable $\tilde{w} \equiv e_k^{(P)} - e_k^{(E)} \sim \mathcal{N}(0, P_k^{(E,P)})$.

8.5.3 Optimization

Consider the minimization problem faced by P at decision time $0 \leq k \leq N-2$: Differentiating the RHS of Eq. (8.39) in his *control* u_k he obtains the optimality condition

$$
u_k^* = -[R_u + B^T(Q+\Pi_{k+1})B]^{-1}B^T(Q+\Pi_{k+1})(A\overline{x}_k^{(P)} + CE_{\tilde{w}}(\,(\gamma_k^{(E)}(\overline{x}_k^{(P)} + \tilde{w})\,)),
$$
$$
k = 0, 1, \ldots, N-1 \tag{8.41}
$$

and similarly, upon differentiating the RHS of Eq. (8.40) in v_k player E obtains

$$
v_k^* = [R_v - C^T(Q+\Pi_{k+1})C]^{-1}C^T(Q+\Pi_{k+1})(A\overline{x}_k^{(E)} + BE_{\tilde{w}}(\,(\gamma_k^{(P)}(\overline{x}_k^{(E)} - \tilde{w})\,)),
$$
$$
k = 0, 1, \ldots, N-1 \tag{8.42}
$$

Player P has obtained an expression for his optimal control u_k^* where $\overline{x}_k^{(E)}$ does not feature and u_k^* is a function of the parameter $\overline{x}_k^{(P)}$ only. However, the strategy function $\gamma_k^{(E)}(\cdot)$ of player E features in this equation. Indeed, the strategic relationship holds

$$(\gamma_k^{(P)}(\overline{x}_k^{(P)}))^* = -[R_u + B^T(Q + \Pi_{k+1})B]^{-1}B^T(Q + \Pi_{k+1})(A\overline{x}_k^{(P)} + CE_{\tilde{w}}(\,((\gamma_k^{(E)}(\overline{x}_k^{(P)} + \tilde{w}))^*\,)), \quad k = 0, 1, \ldots, N-1 \quad (8.43)$$

We have obtained an expression for P's optimal strategy function $(\gamma_k^{(P)}(\overline{x}_k^{(P)}))^*$ in terms of the strategy $\gamma_k^{(E)}(\cdot)$ of player E. Payer P obtained a linear relationship which directly ties together the as yet unknown optimal strategies $(\gamma_k^{(P)}(\overline{x}_k^{(P)}))^*$ and $(\gamma_k^{(E)}(\overline{x}_k^{(E)}))^*$ of players P and E. Similarly, also player E obtains a linear relationship among the players' optimal *strategies*:

$$(\gamma_k^{(E)}(\overline{x}_k^{(E)}))^* = [R_v - C^T(Q + \Pi_{k+1})C]^{-1}C^T(Q + \Pi_{k+1})(A\overline{x}_k^{(E)} + BE_{\tilde{w}}(\,((\gamma_k^{(P)}(\overline{x}_k^{(E)} - \tilde{w}))^*\,)), \quad k = 0, 1, \ldots, N-1 \quad (8.44)$$

Equations (8.43) and (8.44) constitute a linear system of Fredholm integral equations of the second kind in the players' optimal *strategies* $(\gamma_k^{(P)}(\overline{x}_k^{(P)}))^*$ and $(\gamma_k^{(E)}(\overline{x}_k^{(E)}))^*$. Similar to the analysis in reference Pachter (2013), the solution of the linear system of Fredholm integral equations of the second kind, Eqs. (8.43) and (8.44), yields the optimal strategies which are linear in the players' sufficient statistics, namely

$$\gamma_k^{(P)}(\overline{x}_k^{(P)}) = F_k^{(P)} \cdot \overline{x}_k^{(P)}, \quad \gamma_k^{(E)}(\overline{x}_k^{(E)}) = F_k^{(E)} \cdot \overline{x}_k^{(E)}$$

and the formulae for the optimal gains

$$(F_k^{(P)})^* = -S_B^{-1}(Q + \Pi_{k+1})B^T(Q + \Pi_{k+1})\{I + C[R_v - C^T(Q + \Pi_{k+1})C]^{-1}C^T(Q + \Pi_{k+1})\}A, \qquad k = 0, \ldots, N-1 \quad (8.45)$$

$$(F_k^{(E)})^* = -S_C^{-1}(Q + \Pi_{k+1})C^T(Q + \Pi_{k+1})\{I - B[R_u + B^T(Q + \Pi_{k+1})B]^{-1}B^T(Q + \Pi_{k+1})\}A, \qquad k = 0, \ldots, N-1 \quad (8.46)$$

The control system is Linear - Gaussian (LG) and therefore the players' information states are Gaussian, time consistency in this dynamic game is provided by the application of the method of Dynamic Programming (DP) where the DP state is the information state, and, by construction, the strategies are Person-By-Person-Satisfactory (PBPS), so in the LQGDG with a control-sharing information pattern, a Nash equilibrium is obtained—as was also the case in the static LQG game with partial information (Pachter 2013).

8.5.4 Value Functions

The parameters which specify the statistics of the random variables in the LQGDG do not feature in the formulae (8.45) and (8.46) for the players' optimal strategies and consequently an inspection of the DP equations (8.39) and (8.40) tells us that the matrices Π_k won't be a function of the said parameters; in other words, the matrices Π_k are exclusively determined by the deterministic plant's parameters A, B, C, Q, Q_F, $R_c^{(P)}$ and $R_c^{(E)}$. Hence $\Pi_k = P_k$, where P_k is the solution of the Riccati equation (8.57) derived for the *deterministic* LQDG discussed in the Appendix. Upon defining $P_k := P_k + Q$, the optimal gains correspond to the optimal gains in the deterministic LQDG, Eqs. (8.61) and (8.62) in the Appendix and the players' optimal gains are

$$(F_k^{(P)})^* = -S_B^{-1}(P_{k+1}+Q)B^T(P_{k+1}+Q)\{I + C[R_v - C^T(P_{k+1}+Q)C]^{-1}C^T(P_{k+1} +Q)\}A \tag{8.47}$$

$$(F_k^{(E)})^* = -S_C^{-1}(P_{k+1}+Q)C^T(P_{k+1}+Q)\{I - B[R_u + B^T(P_{k+1}+Q)B]^{-1}B^T(P_{k+1} +Q)\}A \tag{8.48}$$

The recursions for the scalars $c_k^{(P)}$ and $c_k^{(E)}$ are obtained from the respective DP equations (8.39) and (8.40):

$$\begin{aligned} c_k^{(P)} &= c_{k+1}^{(P)} \\ &+ 2\,Trace((F_k^{(P)}))^{*T}C^T(Q + P_{k+1}K_{k+1}^{(P)}H^{(P)})A(P_k^{(P)} - \tilde{P}_k^{(P,E)}) \\ &+ Trace(A^TQAP_k^{(P)}) + Trace(\Gamma^TQ\Gamma Q_p) \\ &+ Trace(\Gamma^T(H^{(P)})^T(K_{k+1}^{(P)})^TP_{k+1}K_{k+1}^{(P)}H^{(P)}\Gamma Q_p) \\ &+ Trace((K_{k+1}^{(P)})^TP_{k+1}K_{k+1}^{(P)}R_m^{(P)}) \\ &+ Trace(A^T(H^{(P)})^T(K_{k+1}^{(P)})^TP_{k+1}K_{k+1}^{(P)}H^{(P)}AP_k^{(P)}),\ k = N-2,\dots,0 \end{aligned} \tag{8.49}$$

and for $k = N-1$ we use the end-game equation

$$\begin{aligned} c_{N-1}^{(P)}(P_{N-1}^{(P)}; P_{N-1}^{(E)}, \tilde{P}_{N-1}^{(P,E)}) = Trace(\, & A^TQ_FAP_{N-1}^{(P)} + 2(P_{N-1}^{(P)} \\ & - \tilde{P}_{N-1}^{(P,E)}P_{N-1}^{(P)})A^TQ_FC(F_{N-1}^{(E)})^* - ((F_{N-1}^{(E)})^*)^T(R_v \\ & - C^TQ_FC)(F_{N-1}^{(E)})^*P_{N-1}^{(E,P)} + \Gamma^TQ_F\Gamma Q_p\,) \end{aligned} \tag{8.50}$$

Similarly,

$$\begin{aligned}
c_k^{(E)} &= c_{k+1}^{(E)} \\
&+ 2\, Trace((F_k^{(E)})^*)^T B^T (Q + P_{k+1} K_{k+1}^{(E)} H^{(E)}) A (P_k^{(E)} - \tilde{P}_k^{(E,P)}) \\
&+ Trace(A^T Q A P_k^{(E)}) + Trace(\Gamma^T Q \Gamma Q_p) \\
&+ Trace(\Gamma^T (H^{(E)})^T (K_{k+1}^{(E)})^T P_{k+1} K_{k+1}^{(E)} H^{(E)} \Gamma Q_p) \\
&+ Trace((K_{k+1}^{(E)})^T P_{k+1} K_{k+1}^{(E)} R_m^{(E)}) \\
&+ Trace(A^T (H^{(E)})^T (K_{k+1}^{(E)})^T P_{k+1} K_{k+1}^{(E)} H^{(E)} A P_k^{(E)}),\ k = N-2, \ldots, 0
\end{aligned} \tag{8.51}$$

and for $k = N-1$ we use the end-game equation

$$\begin{aligned}
c_{N-1}^{(E)}(P_{N-1}^{(E)}; P_{N-1}^{(P)}, \tilde{P}_{N-1}^{(E,P)}) = Trace(\, & A^T Q_F A P_{N-1}^{(E)} + ((F_{N-1}^{(P)})^*)^T (R_u \\
& + B^T Q_F B)(F_{N-1}^{(P)})^* P_{N-1}^{(E,P)} + 2((F_{N-1}^{(P)})^*)^T B^T Q_F A (P_{N-1}^{(E)} \\
& - \tilde{P}_{N-1}^{(E,P)} + \Gamma^T Q_F \Gamma Q_p)\,)
\end{aligned} \tag{8.52}$$

Remark. Only the parameters A, B, C, Q, Q_F, $R_c^{(P)}$ and $R_c^{(E)}$ feature in the Riccati equation for P_k, as if the game would be the deterministic LQDG. The players' measurement matrices, the process noise parameters and the measurement noise covariances do not feature in Eq. (8.57). However the solution P_k of the Riccati equation (8.57) and the LQGDG's measurements—related parameters $H^{(P)}$, $H^{(E)}$, the process noise parameters , $R_m^{(P)}$ and $R_m^{(E)}$, and the Kalman gains, all enter the recursions for the "intercepts" $c^{(P)}$ and $c^{(E)}$.

8.6 Main Result

The analysis of the LQGDG with a control-sharing information pattern is summarized in the following

Theorem 1. *Consider the LQGDG (8.1)–(8.6) with the information pattern:*

1. *The P and E players' prior information is given in Eqs. (8.2) and (8.4), respectively. The prior information $\overline{x}_0^{(P)}$ and $\overline{x}_0^{(E)}$ is private information of the respective P and E players and it is not shared among the P and E players. The covariances $P_0^{(P)}$ and $P_0^{(E)}$ are finite and are public information.*
2. *At decision time $1 \le k \le N-1$ the measurements of player P and player E are $z_k^{(P)}$ and $z_k^{(E)}$ and their measurement equations are Eqs. (8.3) and (8.5), respectively. At decision time $1 \le k \le N-1$ the respective measurement records*

$Z_k^{(P)} = \{z_1^{(P)}, \ldots, z_k^{(P)}\}$ and $Z_k^{(E)} = \{z_1^{(E)}, \ldots, z_k^{(E)}\}$ are the private information of players P and E and the measurements are not shared among the P and E players.

3. *At decision time $k = 1, \ldots, N-1$ the P and E players have complete recall of their respective ownship control histories $U_k = \{u_0, \ldots, u_{k-1}\}$ and $V_k = \{v_0, \ldots, v_{k-1}\}$.*
4. *The players observe their opponent's moves: at decision time $1 \le k \le N-1$ the control history $V_k = \{v_0, \ldots, v_{k-1}\}$ of player E is known to player P and, similarly, player E knows the control history $U_k = \{u_0, \ldots, u_{k-1}\}$ of player P.*

The players obtain their private state estimates $\bar{x}_k^{(P)}$ and $\bar{x}_k^{(E)}$ by running two *separate* Kalman Filters (KFs) in parallel driven by their private prior information and their separate measurements: Player P initialized his KF (8.11)–(8.15) with his prior information $(\bar{x}_0^{(P)}, P_0^{(P)})$ and uses his measurements $z_k^{(P)}$. Similarly, player E initialized his KF (8.16)–(8.20) with his prior information $(\bar{x}_0^{(E)}, P_0^{(E)})$ and uses his measurements $z_k^{(E)}$. Both players use the shared complete input history.

The players reuse the state feedback optimal strategies derived for the deterministic LQDG as provided by Theorem A1: In Eq. (8.61) player P sets $x_k := \bar{x}_k^{(P)}$ and in Eq. (8.62) player E sets $x_k := \bar{x}_k^{(E)}$.

A Nash equilibrium for the "zero-sum" LQGDG with a control-sharing information pattern is established. The value functions of players P and E are

$$V_k^{(P)}(\bar{x}_k^{(P)}, P_k^{(P)}; P_k^{(E)}, \tilde{P}_k^{(E,P)}) = (\bar{x}_k^{(P)})^T P_k \bar{x}_k^{(P)} + c_k^{(P)}$$
$$V_k^{(E)}(\bar{x}_k^{(E)}, P_k^{(E)}; P_k^{(P)}, \tilde{P}_k^{(E,P)}) = (\bar{x}_k^{(E)})^T P_k \bar{x}_k^{(E)} + c_k^{(E)}$$

where the matrices P_k are the solution of the Riccati equation (8.57). The "intercepts" $c_k^{(P)}$ and $c_k^{(E)}$ are obtained by solving the respective scalar recursions [(8.49), (8.50), (8.48)] and [(8.51), (8.52), (8.47)]. The covariance matrices $P_k^{(P)}$, $P_k^{(E)}$ and $\tilde{P}_k^{(E,P)}$ exclusively feature in the intercepts' recursions. The matrices $\tilde{P}_k^{(E,P)}$ are given by the solution of the Lyapunov-like linear matrix equation (8.29). The control Riccati equation (8.57), the KF Riccati equation (8.12), (8.13), (8.15) of player P, the KF Riccati equation (8.17), (8.18), (8.20) of player E , and the Lyapunov-like linear matrix equation (8.29) can all be solved ahead of time and off line. Once the three Riccati equations and the Lyapunov equation have been solved, the value functions' "intercepts" $c_k^{(P)}$ and $c_k^{(E)}$ are also obtained off line. □

8.7 Conclusion

Linear-Quadratic Gaussian Dynamic Games with a control-sharing information pattern have been considered. The players' initial state information and their measurements are private information, but each player is able to observe his antagonist's past inputs: the protagonists' past controls is shared information. Although this is a game with partial information, the control-sharing information pattern renders the game amenable to solution by the method of DP and a Nash equilibrium for the "zero-sum" LQGDG is established. The attendant optimal strategies of the LQGDG with a control-sharing information pattern are linear and certainty equivalence holds. The linearity of the optimal strategies has not been artificially imposed from the outset but follows from the LQG nature of the optimization problem at hand, courtesy of the control-sharing information pattern. The correct solution of LQGDGs with a control-sharing information pattern is obtained in closed-form.

Appendix: Linear-Quadratic Dynamic Game

The solution of Linear-Quadratic Dynamic Games (LQDG) with perfect information, a.k.a., deterministic LQDGs, was derived in Pachter and Pham (2010, Theorem 2.1). The Schur complement concept (Fuzhen 2005) was used to invert a blocked $(m_u + m_v) \times (m_u + m_v)$ matrix which contains four blocks, its two diagonal blocks being a $m_u \times m_u$ matrix and a $m_v \times m_v$ matrix. We further improve on the results of Pachter and Pham (2010) by noting that a matrix with four blocks has *two* Schur complements, say S_B and S_C. This allows one to obtain *explicit* and *symmetric* formulae for the P and E players' optimal strategies, thus yielding the *complete* solution of the deterministic LQDG. These results are used in this paper and for the sake of completeness, the closed form solution of the perfect information/deterministic zero-sum LQDG is included herein.

The linear dynamics are

$$x_{k+1} = Ax_k + Bu_k + Cv_k \ , \quad x_0 \equiv x_0 \ , \quad k = 0, 1, \ldots, N-1 \tag{8.53}$$

Payer P is the minimizer and his control $u_k \in R^{m_u}$. Player E is the maximizer and his control $v_k \in R^{m_v}$. The planning horizon is N. The cost/payoff functional is quadratic:

$$J(\{u_k\}_{k=0}^{N-1}, \{v_k\}_{k=0}^{N-1}; x_0) = x_N^T Q_F x_N + \sum_{k=0}^{N-1} (x_{k+1}^T Q x_{k+1} + u_k^T R_u u_k - v_k^T R_v v_k) \tag{8.54}$$

and Q and Q_F are real symmetric matrices. The players' control effort weighting matrices R_u and R_v are typically real symmetric and positive definite. Oftentimes it is stipulated that also the state penalty matrices Q and Q_F be positive definite, or, at least, positive semi-definite; these assumptions can be relaxed. The following holds.

Theorem A1. *A necessary and sufficient condition for the existence of a solution to the deterministic zero-sum LQDG (8.53) and (8.54) is*

$$R_u + B^T P_k B > 0 \tag{8.55}$$

and

$$R_v > C^T P_k C \tag{8.56}$$

$\forall\ k = 1, \dots, N-1$, *where the real, symmetric matrices* P_k *are the solution of the Riccati difference equation*

$$\begin{aligned} P_{k+1} &= A^T\{P_k - P_k[BS_B^{-1}(P_k)B^T + BS_B^{-1}(P_k)B^T P_k C(R_v \\ &\quad - C^T P_k C)^{-1} C^T + C(R_v - C^T P_k C)^{-1} C^T P_k B S_B^{-1}(P_k) B^T \\ &\quad + C(R_v - C^T P_k C)^{-1} C^T P_k B S_B^{-1}(P_k) B^T P_k C(R_v \\ &\quad - C^T P_k C)^{-1} C^T + C(C^T P_k C - R_v)^{-1} C^T] P_k\} A + Q\ , \\ P_0 &= Q + Q_F\ ,\ \ k = 0, \dots, N-1 \end{aligned} \tag{8.57}$$

In Eq. (8.57), the first Schur complement matrix function

$$S_B(P_k) \equiv B^T P_k B + R_u + B^T P_k C(R_v - C^T P_k C)^{-1} C^T P_k B$$

In addition, the problem's parameters must satisfy the conditions

$$R_u + B^T(Q + Q_F)B > 0 \tag{8.58}$$

and

$$R_v > C^T(Q + Q_F)C \tag{8.59}$$

The value of the LQDG is

$$V_0(x_0) = x_0^T(P_N - Q)x_0 \tag{8.60}$$

The players' optimal strategies are the linear state feedback control laws

$$\begin{aligned} u_k^*(x_k) &= -S_B^{-1}(P_{N-k-1})B^T[I + P_{N-k-1}C(R_v \\ &\quad - C^T P_{N-k-1} C)^{-1} C^T] P_{N-k-1} A \cdot x_k, \end{aligned} \tag{8.61}$$

$$\begin{aligned} v_k^*(x_k) &= -S_C^{-1}(P_{N-k-1})C^T[I - P_{N-k-1}B(R_u \\ &\quad + B^T P_{N-k-1} B)^{-1} B^T] P_{N-k-1} A \cdot x_k \end{aligned} \tag{8.62}$$

In Eq. (8.62) the second Schur complement matrix function

$$\begin{aligned} S_C(P_{k+1}) &\equiv -\{R_v - C^T(Q + P_{k+1})C + C^T(Q \\ &\quad + P_{k+1})B[B^T(Q + P_{k+1})B + R_u]^{-1} B^T(Q + P_{k+1})C\} \end{aligned} \qquad \square$$

References

Aoki M (1973) On decentralized linear stochastic control. IEEE Trans Autom Control 18(3): 243–250

Basar T, Bernhard P (2008) H^{∞} - optimal control and related minimax design problems: a dynamic game approach. Birkhauser, Boston

Basar T, Olsder GJ (1995) Dynamic non-cooperative game theory. Academic, New York

Engwerda J (2005) LQ dynamic optimization and differential games. Wiley, New York

Fuzhen Z (2005) The Schur complement and its applications. Springer, Berlin

Ho YC, Bryson AE, Baron S (1965) Differential games and optimal pursuit-evasion strategies. IEEE Trans Autom Control 10(4):385–389

Kurtaran BZ, Sivan R (1974) Linear-quadratic-Gaussian control with one-step-delay sharing pattern. IEEE Trans Autom Control 13(5):571–574

Pachter M (2013) Static linear-quadratic Gaussian games. In: Krivan V, Zaccour G (eds) Advances in dynamic games. Birkhauser, Boston, pp 85–105

Pachter M, Pham K (2010) Discrete-time linear-quadratic dynamic games. J Optim Theory Appl 146(1):151–179

Pachter M, Pham K (2013) Static teams and stochastic games. In: Pardalos PM, et al (eds) Dynamics of information systems: algorithmic approaches. Springer, New York, pp 147–176

Pachter M, Pham K (2014) Informational issues in decentralized control. In: Vogiatzis C, et al (eds) Dynamics of information systems: computational and mathematical challenges. Springer, New York, pp 45–76

Radner R (1962) Static teams and stochastic games. Ann Math Stat 33(3):857–867

Sandell NR, Athans M (1974) Solution of some nonclassical LQG stochastic decision problems. IEEE Trans Autom Control 19(2):108–116

Witsenhausen HS (1968) A counterexample in stochastic optimal control. SIAM J Control 6(1):131–147

Witsenhausen HS (1971) Separation of estimation and control for discrete time systems. Proc IEEE 59(11):1557–1566