Stochastic Differential Games and Intricacy of Information Structures

Tamer Başar

Abstract This chapter discusses, in both continuous time and discrete time, the issue of certainty equivalence in two-player zero-sum stochastic differential/dynamic games when the players have access to state information through a common noisy measurement channel. For the discrete-time case, the channel is also allowed to fail sporadically according to an independent Bernoulli process, leading to intermittent loss of measurements, where the players are allowed to observe past realizations of this process. A complete analysis of a parametrized two-stage stochastic dynamic game is conducted in terms of existence, uniqueness and characterization of saddle-point equilibria (SPE), which is shown to admit SPE of both certainty-equivalent (CE) and non-CE types, in different regions of the parameter space; for the latter, the SPE involves mixed strategies by the maximizer. The insight provided by the analysis of this game is used to obtain through an indirect approach SPE for three classes of differential/dynamic games: (i) linear-quadratic-Gaussian (LQG) zero-sum differential games with common noisy measurements, (ii) discrete-time LQG zero-sum dynamic games with common noisy measurements, and (iii) discrete-time LQG zero-sum dynamic games with intermittently missing perfect state measurements. In all cases CE is a generalized notion, requiring two separate filters for the players, even though they have a common communication channel. Discussions on extensions to other classes of stochastic games, including nonzero-sum stochastic games, and on the challenges that lie ahead conclude the chapter.

1 Introduction

In spite of decades long research activity on stochastic differential games, there still remain some outstanding fundamental questions on existence, uniqueness, and characterization of non-cooperative equilibria when players have access to noisy state information. Even in zero-sum games and with common measurement channel that feeds noisy state information to both players, derivation of saddle-point poli-

T. Başar (✉)

Department of Electrical and Computer Engineering, Coordinated Science Laboratory, University of Illinois, Urbana, IL, USA
e-mail: basar1@illinois.edu

cies is quite an intricate task, as first identified in Başar (1981). That paper also addressed the issue of whether saddle-point equilibria (SPE) in such games is of the certainty-equivalent (CE) type (Witsenhausen 1971a), that is whether the solution of a similarly structured game but with perfect state measurements for both players can be used in the construction of SPE for the stochastic game with noisy measurement, by simply replacing the state with an appropriately constructed conditional estimate. The answer was a "cautious conditional yes," in the sense that not all SPE are of the CE type, and when they are in both the construction of the conditional estimate and the derivation of conditions for existence many perils exist. This chapter picks up where Başar (1981) had left, and develops further insights into the intricacies and pitfalls in the derivation of SPE of the CE as well as non-CE types. It also provides a complete solution to a two-stage stochastic game of the linear-quadratic-Gaussian (LQG) type where the common measurement channel is not only noisy but also fails intermittently.

Research on stochastic differential games with noisy state measurements goes back to the 1960's, where two-person zero-sum games with linear dynamics and measurement equations, Gaussian statistics, and quadratic cost functions (that is, LQG games) were addressed when players have access to different measurements, within however some specific information structures (Behn and Ho 1968; Rhodes and Luenberger 1969; Willman 1969). A zero sum differential game where one player's information is nested in the other player's was considered in Ho (1974), and a class of zero-sum dynamic games where one player has noisy state information while the other one plays open loop was considered in Başar and Mintz (1973) which showed that the open-loop player's saddle-point strategy is *mixed*. A class of zero-sum stochastic games where the information structure is of the nonclassical type was considered in Başar and Mintz (1972), which showed that some zero-sum games could be tractable even though their team counterparts, as in Witsenhausen (1968), Bansal and Başar (1987), Ho (1980) are not; see also Başar (2008).

When a game is not of the zero-sum type, derivation of equilibria (which in this case would be Nash equilibria) is even more challenging, even when players have access to common noisy measurements, with or without delay, as discussed in Başar (1978a) where an indirect approach of the backward-forward type was developed and employed; see also Başar (1978b) for a different formulation and approach for derivation. Recently, a new class of discrete-time nonzero-sum games with *asymmetric information* was introduced in Nayyar and Başar (2012), where the evolution of the local state processes depends only on the global state and control actions and not on the current or past values of local states. For this class of games, it was possible to obtain a characterization of some Nash equilibria by *lifting* the game and converting it to a symmetric one, solving the symmetric one in terms of Markov equilibria, and then converting it back. Among many others, two other papers of relevance to stochastic nonzero-sum dynamic games are Altman et al. (2009) and Hespanha and Prandini (2001), and one of relevance to teams with delayed sharing patterns is Nayyar et al. (2011).

The paper is organized as follows. In the next section, we introduce LQG zero-sum stochastic differential/dynamic games (ZSDGs) with common noisy measurements, first in continuous time and then in discrete time, and for the latter we also

include the possibility of intermittent failure of the measurement channel (modeled through a Bernoulli process), leading to occasionally missing measurements. In the section we also introduce the concept of certainty equivalence, first in the context of the classical LQG optimal control problem and then generalized (in various ways) to the two classes of games formulated. In Sect. 3, we introduce a two-stage stochastic dynamic game, as a special case of the general discrete time LQG game of Sect. 2, which is solved completely for its SPE in both pure and mixed strategies, some of the CE type and others non-CE (see Theorem 1 for the complete solution). Analysis of the two-stage game allows us to develop insight into the intricate role information structures play in the characterization and existence of SPE for the more general ZS-DGs of Sect. 2, and what CE means in a game context. This insight is used in Sect. 4 in the derivation of generalized CE SPE for the continuous-time LQG ZSDG with noisy state measurements (see Theorem 2 for the penultimate result) as well as for the continuous-time LQG ZSDG with noisy state measurements and perfect state measurements with intermittent losses. The paper ends with a recap of the results of the paper and a discussion on extensions and open problems, in Sect. 5.

2 Zero-Sum Stochastic Differential and Discrete-Time Games with a Common Measurement Channel and Issue of Certainty Equivalence

2.1 Formulation of the Zero-Sum Stochastic Differential Game

We first consider the class of so-called Linear-Quadratic-Gaussian zero-sum differential games (LQG ZSDGs), where the two players' actions are inputs to a linear system driven also by a Wiener process, and the players have access to the system state through a common noisy measurement channel which is also linear in the state and the driving Wiener noise process. The objective function, to be minimized by one player and maximized by the other, is quadratic in the state and the actions of the two players.

For a precise mathematical formulation, let $\{x_t, y_t, t \geq 0\}$, be respectively the n-dimensional state and m-dimensional measurement processes, generated by

$$dx_t = (Ax_t + Bu_t + Dv_t)dt + Fdw_t, \quad t \geq 0, \tag{1}$$

$$dy_t = Hx_tdt + Gdw_t, \quad t \geq 0, \tag{2}$$

where $\{u_t, t \geq 0\}$ and $\{v_t, t \geq 0\}$ are respectively Player 1's and Player 2's controls (say of dimensions r_1 and r_2, respectively), nonanticipative with respect to the measurement process, and generated by measurable control policies $\{\gamma_t\}$ and $\{\mu_t\}$, respectively, that is

$$u_t = \gamma_t(y_{[0,t)}), \qquad v_t = \mu_t(y_{[0,t)}), \quad t \geq 0. \tag{3}$$

In (1) and (2), x_0 is a zero-mean Gaussian random vector with covariance Λ_0 (that is $x_0 \sim N(0, \Lambda_0)$), $\{w_t, t \geq 0\}$ is a vector-valued standard Wiener process independent of x_0, and A, B, D, F, H, G are constant[1] matrices of appropriate dimensions, with (to avoid singularity) $FF^T > 0$, $GG^T > 0$, and $FG^T = 0$, where the last condition assures that system and channel noises are independent. We let Γ and $\mathcal{M}$ denote the classes of admissible control policies for Player 1 and Player 2, respectively, with elements $\gamma := \{\gamma_t\}$ and $\mu := \{\mu_t\}$, as introduced earlier. The only restriction on these policies is that when (3) is used in (1), we have unique second-order stochastic process solutions to (1) and (2), with almost sure continuously differentiable sample paths. Measurability and uniform Lipschitz continuity will be sufficient for this purpose.

To complete the formulation of the differential game, we now introduce a quadratic performance index over a finite interval $[0, t_f]$:

$$J(\gamma, \mu) = E\left\{ |x_{t_f}|_{Q_f}^2 + \int_0^{t_f} \left[|x_t|_Q^2 + \lambda |u_t|^2 - |v_t|^2 \right] dt \,\Big|\, u = \gamma(\cdot), v = \mu(\cdot) \right\}, \quad (4)$$

where expectation $E\{\cdot\}$ is over the statistics of x_0 and $\{w_t\}$; further, $|x|_Q^2 := x^T Q x$, Q and Q_f are non-negative definite matrices, and $\lambda > 0$ is a scalar parameter. Note that any objective function with nonuniform positive weights on u and v can be brought into the form above by a simple rescaling and re-orientation of u and v and a corresponding transformation applied to B and D, and hence the structure in (4) does not entail any loss of generality as a quadratic performance index.

The problem of interest in the context of LQG ZSDGs is to find conditions for existence and characterization of saddle-point strategies, that is ($\gamma^* \in \Gamma, \mu^* \in \mathcal{M}$) such that

$$J(\gamma^*, \mu) \leq J(\gamma^*, \mu^*) \leq J(\gamma, \mu^*), \quad \forall \gamma \in \Gamma, \mu \in \mathcal{M}. \quad (5)$$

A question of particular interest in this case is whether the saddle-point equilibrium (SDE) has the *certainty equivalence* property, that is whether it can be obtained directly from the perfect state-feedback SPE of the corresponding deterministic differential game, by simply replacing the state by its "best estimate," as in the one-player version, the so-called LQG optimal control problem. This will be discussed later in the section.

If a saddle-point equilibrium (SDE) does not exist, then the next question is whether the upper value of the game is bounded, and whether there exists a control strategy for the minimizer that achieves it, that is existence of a $\bar{\gamma} \in \Gamma$ such that

$$\inf_{\gamma} \sup_{\mu} J(\gamma, \mu) = \sup_{\mu} J(\bar{\gamma}, \mu). \quad (6)$$

Note that the lower value of the game, $\sup_\mu \inf_\gamma J(\gamma, \mu)$, is always bounded away from zero, and hence its finiteness is not an issue.

---

[1]They are taken to be constant for simplicity in exposition; the main message of the paper and many of the expressions stand for the time-varying case as well, with some obvious modifications.

2.2 Formulation of the Discrete-Time Zero-Sum Stochastic Dynamic Game with Failing Channels

A variation on the LQG ZSDG is its discrete-time version, which will allow us also to introduce intermittent failure of the common measurement channel. The system equation (1) is now replaced by

$$x_{t+1} = Ax_t + Bu_t + Dv_t + Fw_t, \quad t = 0, 1, \ldots, \tag{7}$$

and the measurement equation (2) by

$$y_t = \beta_t(Hx_t + Gw_t), \quad t = 0, 1, \ldots, \tag{8}$$

where $x_0 \sim N(0, \Lambda_0)$; $\{w_t\}$ is a zero-mean Gaussian process, independent across time and of x_0, and with $E\{w_t w_t^T\} = I$, $\forall t \in [0, T-1] := \{0, 1, \ldots, T-1\}$; and $\{\beta_t\}$ is a Bernoulli process, independent across time and of x_0 and $\{w_t\}$, with Probability$(\beta_t = 0) = p$, $\forall t$. This essentially means that the channel that carries information on the state to the players fails with equal probability p at each stage, and these failures are statistically independent. A different expression for (8) which essentially captures the same would be

$$y_t = \beta_t Hx_t + Gw_t, \quad t = 0, 1, \ldots, \tag{9}$$

where what fails is the sensor that carries the state information to the channel and not the channel itself. In this case, when $\beta_t = 0$, then this means that the channel only carries pure noise, which of course is of no use to the controllers.

Now, if the players are aware of the failure of the channel or the sensor when it happens (which we assume to be the case), then what replaces (3) is

$$u_t = \gamma_t(y_{[0,t]}, \beta_{[0,t]}), \qquad v_t = \mu_t(y_{[0,t]}, \beta_{[0,t]}), \quad t = 0, 1, \ldots, \tag{10}$$

where $\{\gamma_t\}$ and $\{\mu_t\}$ are measurable control policies; let us again denote the spaces where they belong respectively by Γ and M.

The performance index replacing (4) for the discrete-time game, over the interval $\{0, 1, \ldots, T-1\}$ is[2]

$$J(\gamma, \mu) = E\left\{\sum_{t=0}^{T-1}\left[|x_{t+1}|_Q^2 + \lambda|u_t|^2 - |v_t|^2\right]dt \,\Big|\, u = \gamma(\cdot), v = \mu(\cdot)\right\}, \tag{11}$$

where the expectation is over the statistics of x_0, $\{w_t\}$ and $\{\beta_t\}$.

The goal is again the one specified in the case of the LQG ZSDG—to study existence and characterization of SPE (defined again by (5), appropriately interpreted),

---

[2]We are using "T" to denote the number of stages in the game; the same notation was used to denote *transpose*. These are such distinct usages that no confusion or ambiguity should arise.

boundedness of upper value if a SPE does not exist, and certainty-equivalence property of the SPE. We first recall below the certainty-equivalence property of the standard LQG optimal control problem, which is a special case of the LQG ZSDG obtained by leaving out the maximizer, that is by letting $D = 0$. We discuss only the continuous-time case; an analogous result holds for the discrete-time case (Witsenhausen 1971a; Yüksel and Başar 2013).

2.3 The LQG Optimal Control Problem and Certainty Equivalence

Consider the LQG optimal control problem, described by the linear state and measurement equations

$$dx_t = (Ax_t + Bu_t)dt + Fdw_t, \quad t \geq 0, \tag{12}$$

$$dy_t = Hx_t dt + Gdw_t, \quad t \geq 0, \tag{13}$$

and the quadratic cost function

$$J(\gamma) = E\left\{|x_{t_f}|^2_{Q_f} + \int_0^{t_f} \left[|x_t|^2_Q + \lambda|u_t|^2\right]dt \,\Big|\, u = \gamma(\cdot)\right\}, \tag{14}$$

where F and G satisfy the earlier conditions, and as before $\gamma \in \Gamma$.

It is a standard result in stochastic control (Fleming and Soner 1993) that there exists a unique $\gamma^* \in \Gamma$ that minimizes $J(\gamma)$ defined by (14), and $\gamma_t^*(y_{[0,t)})$ is linear in $y_{[0,t)}$. Specifically,

$$u^*(t) = \gamma_t^*(y_{[0,t)}) = \tilde{\gamma}_t(\hat{x}_t) = -\frac{1}{\lambda}B^T P(t)\hat{x}_t, \quad t \geq 0, \tag{15}$$

where P is the unique non-negative definite solution of the retrograde Riccati differential equation

$$\dot{P} + PA + A^T P - \frac{1}{\lambda}PBB^T P + Q = 0, \quad P(t_f) = Q_f, \tag{16}$$

where $\{\hat{x}_t\}$ is generated by the Kalman Filter:

$$d\hat{x}_t = (A\hat{x}_t + Bu_t)dt + K(t)(dy_t - H\hat{x}_t dt), \quad \hat{x}_0 = 0, t \geq 0, \tag{17}$$

$$K(t) = \Lambda(t)H^T\left[GG^T\right]^{-1} \tag{18}$$

with Λ being the unique non-negative definite solution of the forward Riccati differential equation

$$\dot{\Lambda} - A\Lambda - \Lambda A^T + \Lambda H^T\left[GG^T\right]^{-1}H\Lambda - FF^T = 0, \quad \Lambda(0) = \Lambda_0. \tag{19}$$

Note that this is a certainty-equivalent (CE) controller, because it has the structure of the optimal controller for the deterministic problem, that is $-\frac{1}{\lambda}B^T P(t)x_t$, with the state x_t replaced by its conditional mean, $E[x_t|y_{[0,t)}, u_{[0,t)}]$, which is given by (17). The controller gain $(-\frac{1}{\lambda}B^T P(t))$ is constructed independently of what the estimator does, while the estimation or filtering is also essentially an independent process with however the past values of the control taken as input to the Kalman filter. Hence, in a sense we have a *separation* of estimation and control, but not complete decoupling. In that sense, we can say that the controller has to cooperate with the estimator as the latter needs to have access to the output of the *control box* for the construction of the conditional mean. Of course, an alternative representation for (17) would be the one where the optimal controller is substituted in place of u:

$$d\hat{x}_t = \left(\left(A - \frac{1}{\lambda}BB^T P(t)\right)\hat{x}_t\right)dt + K(t)(dy_t - H\hat{x}_t dt), \quad \hat{x}_0 = 0, t \geq 0, \quad (20)$$

but in this representation also there is a need for collaboration or sharing of information, since the estimator has to have access to $P(\cdot)$ or the cost parameters that generate it. Hence, the solution to the LQG problem has an implicit cooperation built into it, but this does not create any problem or difficulty in this case, since the estimator and the controller are essentially a single unit.

2.4 The LQG ZSDG and Certainty Equivalence

Now we move on to the continuous-time (CT) LQG ZSDG, to obtain a CE SPE, along the lines of the LQG control problem discussed above. The corresponding deterministic LQ ZSDG, where both players have access to perfect state measurements, admits the state-feedback SPE (Başar and Olsder 1999):

$$u^*(t) = \gamma_t^*(x_{[0,t]}) = \tilde{\gamma}_t(x_t) = -\frac{1}{\lambda}B^T Z(t)x_t, \quad t \geq 0, \quad (21)$$

$$v^*(t) = \mu_t^*(x_{[0,t]}) = \tilde{\mu}_t(x_t) = D^T Z(t)x_t, \quad t \geq 0, \quad (22)$$

where Z is the unique non-negative definite continuously differentiable solution of the following Riccati differential equation (RDE) over the interval $[0, t_f]$:

$$\dot{Z} + A^T Z + ZA - Z\left(\frac{1}{\lambda}BB^T - DD^T\right)Z + Q = 0, \quad Z(t_f) = Q_f. \quad (23)$$

Existence of such a solution (equivalently nonexistence of a conjugate point in the interval $(0, t_f)$, or *no finite escape*) to the RDE (23) is also a necessary condition for existence of any SPE (Başar and Bernhard 1995), in the sense that even if any (or both) of the players use memory on the state, the condition above cannot be further relaxed. This conjugate-point condition translates, in this case, on a condition on λ, in the sense that there exists a critical value of λ, say λ^* (which will depend on the

parameters of the game and the length of the time horizon, and could actually be any value in $(0, \infty)$), so that for each $\lambda \in (0, \lambda^*)$, the pair (21)–(22) provides a SPE to the corresponding deterministic ZSDG.

Now, if a natural counterpart of the CE property of the solution to the LQG optimal control problem would hold for the LQG ZSDG, we would have as SPE:

$$u^*(t) = \gamma_t^*(y_{[0,t)}) = \tilde{\gamma}_t(\hat{x}_t) = -\frac{1}{\lambda} B^T Z(t) \hat{x}_t, \quad t \geq 0, \tag{24}$$

$$v^*(t) = \mu_t^*(y_{[0,t)}) = \tilde{\mu}_t(\hat{x}_t) = D^T Z(t) \hat{x}_t, \quad t \geq 0, \tag{25}$$

where

$$\hat{x}_t := E\big[x_t | y_{[0,t)}, \{u_s = \gamma_s^*(y_{[0,s)}), v_s = \mu_s^*(y_{[0,s)}), 0 \leq s < t\}\big]$$

is generated by (as counterpart of (20)):

$$d\hat{x}_t = \hat{A}\hat{x}_t dt + K(t)(dy_t - H\hat{x}_t dt), \quad \hat{x}_0 = 0, t \geq 0, \tag{26}$$

where

$$\hat{A} := A - \left(\frac{1}{\lambda} B B^T - D D^T\right) Z(t) \tag{27}$$

and K is the Kalman gain, given by (18), with Λ now solving

$$\dot{\Lambda} - \hat{A}\Lambda - \Lambda\hat{A}^T + \Lambda H^T \big[G G^T\big]^{-1} H \Lambda - F F^T = 0, \quad \Lambda(0) = \Lambda_0. \tag{28}$$

The question now is whether the strategy pair (γ^*, μ^*) above constitutes a SPE for the LQG ZSDG, that is whether it satisfies the pair of inequalities (5). We will address this question in Sect. 4, after discussing in the next section some of the intricacies certainty equivalence entails, within the context of a two-stage discrete-time stochastic dynamic game. But first, we provide in the subsection below the counterpart of the main result of this subsection for the discrete-time dynamic game.

2.5 The LQG Discrete-Time ZS Dynamic Game and Certainty Equivalence

Consider the discrete-time (DT) LQG ZS dynamic game (DG) formulated in Sect. 2.2, but with non-failing channels (that is, with $p = 0$). We provide here a candidate CE SPE for this game, by following the lines of the previous subsection, but in discrete time. First, the corresponding deterministic LQ ZSDG, where both players have access to perfect state measurements admits the state-feedback SPE (Başar and Olsder 1999) (as counterpart of (21)–(22))

$$u_t^* = \gamma_t^*(x_{[0,t]}) = \tilde{\gamma}_t(x_t) = -\frac{1}{\lambda} B^T Z_{t+1}(N_t^{-1})^{-1} A x_t, \quad t = 0, 1, \ldots, \quad (29)$$

$$v_t^* = \mu_t^*(x_{[0,t]}) = \tilde{\mu}_t(x_t) = D^T Z_{t+1}(N_t^{-1})^{-1} A x_t, \quad t = 0, 1, \ldots, \quad (30)$$

where

$$N_t = I + \left(\frac{1}{\lambda} B B^T - D D^T\right) Z_{t+1}, \quad t = 0, 1, \ldots, \quad (31)$$

and Z_t is a non-negative definite matrix, generated by the following discrete-time game Riccati equation (DTGRE):

$$Z_t = Q + A^T Z_{t+1}(N_t^{-1})^T A, \quad Z(T) = Q. \quad (32)$$

Under the additional condition

$$I - D^T Z_{t+1} D > 0, \quad t = 0, 1, \ldots, T - 1, \quad (33)$$

which also guarantees invertibility of N, the pair (29)–(30) constitutes a SPE. If, on the other hand, the matrix (33) has a negative eigenvalue for some t, then the upper value of the game is unbounded (Başar and Bernhard 1995). As in the CT conjugate point condition, the condition (33) translates into a condition on λ, in the sense that there exists a critical value of λ, say λ_c (which will depend on the parameters of the game and the number of stages in the game), so that for each $\lambda \in (0, \lambda_c)$, the pair (29)–(30) provides a SPE to the corresponding deterministic ZSDG.

Now, the counterpart of (24)–(25), as a candidate CE SPE, would be

$$u_t^* = \gamma_t^*(y_{[0,t]}) = \tilde{\gamma}_t(\hat{x}_{t|t}) = -\frac{1}{\lambda} B^T Z_{t+1}(N_t^{-1})^T A \hat{x}_{t|t}, \quad t = 0, 1, \ldots, \quad (34)$$

$$v_t^* = \mu_t^*(y_{[0,t]}) = \tilde{\mu}_t(\hat{x}_{t|t}) = D^T Z_{t+1}(N_t^{-1})^T A \hat{x}_{t|t}, \quad t = 0, 1, \ldots, \quad (35)$$

where

$$\hat{x}_{t|t} := E\big[x_t | y_{[0,t]}, \{u_s = \gamma_s^*(y_{[0,s]}), v_s = \mu_s^*(y_{[0,s]}), s = 0, \ldots, t - 1\}\big]$$

is generated by, with $\hat{x}_{0|-1} = 0$:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \Lambda_t H^T (H \Lambda_t H^T + G G^T)^{-1}(y_t - H \hat{x}_{t|t-1})$$
$$\hat{x}_{t+1|t} = (N_t)^{-1} A \hat{x}_{t|t-1} \quad (36)$$
$$+ (N_t)^{-1} A \Lambda_t H^T (H \Lambda_t H^T + G G^T)^{-1}(y_t - H \hat{x}_{t|t-1}),$$

where the sequence $\{\Lambda_t, t = 1, 2, \ldots, T\}$ is generated by

$$\Lambda_{t+1} = (N_t)^{-1} A \Lambda_t \big[I - H^T (H \Lambda_t H^T + G G^T)^{-1} H \Lambda_t\big] A^T ((N_t)^{-1})^T$$
$$+ F F^T, \quad (37)$$

with the initial condition Λ_0 being the covariance of x_0.

Now, if instead of the noisy channel, we have intermittent failure of a channel which otherwise carries perfect state information, modeled as in Sect. 2.2 but with *clean* transmission when the channel operates (failure being according to an independent Bernoulli process as before), then as a candidate CE SPE, the pair (34)–(35) is replaced by

$$u_t^* = \gamma_t^*(y_{[0,t]}) = \tilde{\gamma}_t(\zeta_t) = -\frac{1}{\lambda} B^T Z_{t+1} \left(N_t^{-1} \right)^T A \zeta_t, \quad t = 0, 1, \ldots, \quad (38)$$

$$v_t^* = \mu_t^*(y_{[0,t]}) = \tilde{\mu}_t(\zeta_t) = D^T Z_{t+1} \left(N_t^{-1} \right)^T A \zeta_t, \quad t = 0, 1, \ldots, \quad (39)$$

where the stochastic sequence $\{\zeta_t, t = 0, 1, \ldots\}$ is generated by

$$\zeta_t = \beta_t y_t + (1 - \beta_t) \left(I - \left(\frac{1}{\lambda} B B^T + D D^T \right) Z_{t+1} \left(N_t^{-1} \right)^T \right) A \zeta_{t-1}, \quad \zeta_0 = y_0. \quad (40)$$

We will explore in Sect. 4 whether these CE policies are in fact SPE policies, and under what conditions.

3 A Two-Stage Discrete-Time Game with Common Measurement

To explicitly demonstrate the fact that certainty equivalence generally fails in games (but holds in a restricted sense), we consider here a specific 2-stage version of the formulation (7), (8), (11):

$$x_2 = x_1 - u + w_1; \qquad x_1 = x_0 + 2v + w_0, \quad (41)$$

$$y_1 = \beta_1(x_1 + r_1); \qquad y_0 = \beta_0(x_0 + r_0), \quad (42)$$

$$J(\gamma, \mu) = E\left\{ (x_2)^2 + \lambda u^2 - v^2 | u = \gamma(\cdot), v = \mu(\cdot) \right\}, \quad (43)$$

with $u = \gamma(y_1, y_0; \beta_1, \beta_0)$, $v = \mu(y_0; \beta_0)$, where the random variables x_0, w_0, w_1, r_0, r_1 are independent, Gaussian, with zero mean and unit variance, and β_1, β_0 are independent Bernoulli random variables with Probability$(\beta_t = 0) = p$, for $t = 0, 1$.

3.1 Certainty-Equivalent SPE

The deterministic version of the game above, with $u = \gamma(x_1, x_0)$, $v = \mu(x_0)$, admits a unique saddle-point solution (Başar and Olsder 1999), given by

$$\gamma^*(x_1, x_0) = \frac{1}{1 + \lambda} x_1, \qquad \mu^*(x_0) = \frac{2\lambda}{1 - 3\lambda} x_0, \quad (44)$$

whenever

$$0 < \lambda < \frac{1}{3}, \tag{45}$$

and for $\lambda > 1/3$, the upper value is unbounded.

Now, a certainty-equivalent (CE) SPE for the original stochastic game, if exists, would be one obtained from the SPE of the related deterministic dynamic game, as above, by replacing x_0 and x_1 by their conditional means, which in the case of x_1 would require the SP policy at the earlier stage (that is, stage 0). Carrying this out, we have

$$\mu^*(y_0; \beta_0) = \frac{2\lambda}{1 - 3\lambda} E[x_0|y_0; \beta_0] = \frac{\lambda}{1 - 3\lambda} y_0, \tag{46}$$

and

$$
\begin{aligned}
\gamma^*(y_{[0,1]}; \beta_{[0,1]}) &= \frac{1}{1 + \lambda} E\big[x_1|y_1, y_0; \beta_1, \beta_0; \upsilon = \mu^*(y_0; \beta_0)\big] \\
&= \frac{1}{1 + \lambda}\bigg[\beta_1\bigg(\frac{2}{3}y_1 - \frac{3}{10}y_0 - \frac{6}{5}\mu^*(y_0; \beta_0)\bigg) \\
&\quad + \beta_0\bigg(-\frac{1}{15}y_1 + \frac{1}{2}y_0 + 2\mu^*(y_0; \beta_0)\bigg)\bigg].
\end{aligned}
\tag{47}
$$

Note that if the channel does not fail at all (that is, $\beta_0 = \beta_1 = 1$), then one can have a simpler expression for (47), given by:

$$\gamma^*(y_1, y_0) = \frac{3}{5(1 + \lambda)} y_1 + \frac{1}{5(1 - 3\lambda)} y_0. \tag{48}$$

3.2 Analysis for $p = 0$ for CE SPE and Beyond

We assume in this subsection that $p = 0$, in which case the CE SPE (whose SPE property is yet to be verified) is given by (46)–(48). It is easy to see that $J(\gamma^*, \upsilon)$, with γ^* as in (48) is unbounded in υ unless $\lambda < 3/25$, which means that the CE pair (46)–(48) cannot be a SPE for $\lambda \in [3/25, 1/3)$, even though the pair (44) was for the deterministic game. For the interval $\lambda \in (0, 3/25)$, however, the CE pair (46)–(48) is a SPE for the stochastic game, as it can easily be shown to satisfy the pair of inequalities (5). Further, for this case, since μ^* is the unique maximizer to $J(\gamma^*, \mu)$, and γ^* is the unique minimizer to $J(\gamma, \mu^*)$, it follows from the *ordered interchangeability* property of multiple SP equilibria (Başar and Olsder 1999) that the SPE is *unique*. Hence, for the parametrized stochastic dynamic game, a "restricted" CE property holds—*restricted* to only some values of the parameter.

Now the question is whether the parametrized stochastic game admits a SPE for $\lambda \in [3/25, 1/3)$. Clearly, it cannot be a CE SPE, that is the SPE of the deterministic version cannot be used to obtain it. Note that, for $\lambda \in [3/25, 1/3)$, if one picks

$\gamma(y_1, y_0) = [1/(1 + \lambda)]y_1$ in $J(\gamma, \mu)$, then the maximum of this function with respect to μ exists and is bounded, which means that the upper value of the stochastic game is bounded. Its lower value is also clearly bounded (simply pick $v = 0$).

Again working with the special case $p = 0$ (that is no failure of the noisy channels), we now claim that there indeed exists a SPE for $\lambda \in [3/25, 1/3)$, but it entails a mixed strategy for the maximizer (Player 2) and still a pure strategy for the minimizer (Player 1). These are:

$$v = \mu^*(y_0) = \frac{2\lambda}{1 - 3\lambda} E[x_0|y_0] + \xi = \frac{\lambda}{1 - 3\lambda} y_0 + \xi,$$

$$\xi \sim N(0, \sigma^2), \quad \sigma^2 = \frac{4 - 5\sqrt{1 - 3\lambda}}{8\sqrt{1 - 3\lambda}}, \tag{49}$$

and

$$u = \gamma^*(y_1, y_0) = \frac{1}{1 + \lambda} E\big[x_1|y_1, y_0, v = \mu^*(y_0)\big]$$

$$= \frac{2 - \sqrt{1 - 3\lambda}}{2(1 + \lambda)} y_1 + \frac{1}{4\sqrt{1 - 3\lambda}} y_0. \tag{50}$$

First note that $\sigma^2 > 0$ for $\lambda \in (3/25, 1/3)$, and $\sigma^2 = 0$ at $\lambda = 3/25$, and further that the policies (49)–(50) agree with (46)–(47) at $\lambda = 3/25$, and hence transition from CE SPE to the non-CE one is continuous at $\lambda = 3/25$. Now, derivation of (49)–(50) as a SPE uses the *conditional equalizer* property of the minimizer's policy (that is (50)). One constructs a policy γ for the minimizer, under which (that is, with $u = \gamma(y_0, y_1)$) the conditional cost

$$E\big\{(x_2)^2 + \lambda u^2 - v^2|y_0\big\}$$

becomes independent of v, and (50) is such a policy; it is in fact the unique such policy in the linear class. Hence, any choice of μ, broadened to include also mixed policies, would be a maximizing policy for Player 2, and (49) is one such policy. This establishes the left-hand-side inequality in (5). For the right-hand-side inequality, it suffices to show that (50) minimizes $J(\gamma, \mu^*)$; this is in fact a strictly convex LQG optimization problem, whose unique solution is (50). Because of this uniqueness, and *ordered interchangeability* of multiple SPE (Başar and Olsder 1999), the SPE (49)–(50) is *unique*.

3.3 The Case $p > 0$

We now turn to the case where the channels fail with positive probability, for which a candidate pair of SPE policies, based on CE, was given by (46)–(47). Their SPE property is yet to be shown, as well as the range of values of λ for which it is valid as

a SPE is yet to be determined, which we address in this subsection. Toward that end, let us first consider, as a benchmark, the special case with noise-free (but still failure prone) measurement channel. This would therefore correspond to the formulation where (42) is replaced by

$$y_1 = \beta x_1; \qquad y_0 = \beta_0 x_0.$$

The counterpart of (46)–(47) in this case would be (this can also be obtained directly from the perfect-state SPE):

$$v = \mu^*(y_0; \beta_0) = \frac{2\lambda}{1 - 3\lambda} y_0, \tag{51}$$

$$u = \gamma^*(y_{[0,1]}; \beta_{[0,1]}) = \beta_1 \frac{1}{1 + \lambda} y_1 + (1 - \beta_1) \frac{1}{1 + \lambda} \big[y_0 + 2\mu^*(y_0) \big]. \tag{52}$$

Now, if Player 2 employs (51), then the unique response of Player 1 will be $(1/(1 + \lambda))x_1$ for $\beta_1 = 1$, and $(1/(1 + \lambda))E[x_1|x_0] = (1/(1 + \lambda))(x_0 + 2\mu^*(x_0))$ if $\beta_1 = 0$ and $\beta_0 = 1$, which agrees with (52); if both β's are *zero*, then clearly Player 1's response will also be *zero*. Note that the responses by Player 1 in all these cases are unique.

If Player 1 employs (52), then the conditional cost (conditioned on y_0, β_0) seen and to be maximized by Player 2 is:

For $\beta_0 = 1$ (after some simplifications):

$$1 + p + (1 - p) \frac{\lambda}{1 + \lambda} \big[1 + (x_0 + 2v)^2 \big] + p \left[\frac{\lambda}{1 + \lambda} x_0 + 2v - \frac{2}{1 + \lambda} \mu^*(y_0) \right]^2 p$$

$$+ \frac{\lambda}{(1 + \lambda)^2} \big[x_0 + 2\mu^*(y_0) \big]^2 p - v^2 \tag{53}$$

and for $\beta_0 = 0$ (after some simplifications):

$$1 + p + \frac{\lambda}{1 + \lambda}(2 - p) + \frac{3\lambda + 4p - 1}{1 + \lambda} v^2. \tag{54}$$

Both (53) and (54) are strictly concave in v if and only if

$$p < \frac{1}{4} \quad \text{and} \quad \lambda < \frac{1 - 4p}{3}, \tag{55}$$

in which case the unique maximizing solution to (54) is $v^* = 0$, and likewise to (53) is $v^* = \mu^*(x_0) = (2\lambda/1 - 3\lambda)x_0$, both of which agree with (51). Hence, the CE pair (51)–(52) indeed provides a SPE if the condition (55) holds, that is the failure probability should be less than $1/4$, and the parameter λ should be less than a specific threshold, which decreases with increasing p. Note that if $p = 0$, we recover the earlier condition (45) for the deterministic game, where we know that if $\lambda > 1/3$, then the upper value is unbounded. The question is whether the same applies here, for $\lambda > (1 - 4p)/3$. This indeed is the case, as one can easily argue that

the concavity condition of (54) cannot be improved further as universally optimal choice for γ when Player 1 has access to x_1 but not x_0 has led to that conditional cost. Hence, the pair (51)–(52) is the complete set of SPE for the game of this subsection (with channel failure but no noise in the channel), and the condition (55) is tight.

We are now in a position to discuss the SPE of the original stochastic game of this section, to find the conditions (if any) under which the CE policies (46)–(47) are in SPE, and whether those conditions can be relaxed by employing structurally different policies (as in Sect. 3.2).

3.4 CE SPE and Beyond for the 2-Stage Game

To obtain the complete set of SPE for the original stochastic 2-stage game, our starting point will be the pair of CE policies (γ^*, μ^*) given by (47) and (46), and to find the region in the $\lambda - p$ space for which these policies constitute a SPE. Clearly, we would expect that region (if exists) to be no larger than the one described by (55) since that one corresponded to the noise-free channel.

Let us first consider the right-hand inequality of (5) for this game, with $\mu^*(y_0; \beta_0)$ given by (46). In terms of γ this is a strictly convex quadratic optimization problem, which one minimizes with respect to u after conditioning the cost on $y_{[0,1]}$ and $\beta_{[0,1]}$; the result is the unique solution given by (47). This part of the inequality does not bring in any additional restriction on λ and p, other than the condition $\lambda < 1/3$ needed in the expression for μ^*.

The left-hand inequality of (5) for this game is a bit more involved. We now pick γ^* as given by (47), and maximize the resulting cost over μ, which is equivalent to maximizing the conditional cost with respect to v where conditioning is with respect to y_0 and β_0. Even though this is also a quadratic optimization problem, existence and uniqueness of maximum are not guaranteed for all values of λ and p, and we have to find (necessary and sufficient) conditions for strict concavity (in v). Now, the conditional cost (conditioned on (y_0, β_0), and with $v = \mu(y_0, \beta_0)$) is:

For $\beta_0 = 1$ (after some simplifications):

$$
p\left[x_0 + 2v - \frac{1}{2(1-3\lambda)}y_0\right]^2 + (1-p)\left[\frac{2+5\lambda}{5(1+\lambda)}(x_0+2v) - \frac{1}{5(1-3\lambda)}y_0\right]^2
$$

$$
+ \lambda(1-p)\left[\frac{3}{5(1+\lambda)}(x_0+2v) + \frac{1}{5(1-3\lambda)}y_0\right]^2 - v^2
$$

$$
+ 2p + (1-p)\frac{50\lambda^2 + 88\lambda + 38}{25(1+\lambda)^2} + \frac{\lambda p}{4(1-3\lambda)^2}(y_0)^2, \tag{56}
$$

and for $\beta_0 = 0$ (after some simplifications):

$$(1-p)\left[\frac{2+5\lambda}{5(1+\lambda)}(x_0+2v)\right]^2 + \lambda(1-p)\left[\frac{3}{5(1+\lambda)}(x_0+2v)\right]^2 - v^2$$

$$+ p[x_0+2v]^2 + 2p + (1-p)\frac{50\lambda^2 + 88\lambda + 38}{25(1+\lambda)^2}. \tag{57}$$

Both (56) and (57) are strictly concave in v if and only if the coefficients of the quadratic terms in v (identical in the two cases) are negative, that is

$$(1-p)\left[\frac{4(2+5\lambda)^2}{25(1+\lambda)^2} + \frac{36\lambda}{25(1+\lambda)^2}\right]^2 + 4p - 1 < 0,$$

which is equivalent to

$$p < \frac{3}{28} \quad \text{and} \quad \lambda < \frac{3-28p}{25}. \tag{58}$$

We note that the upper bound on the failure probability p is precisely the condition that makes the upper bound on λ in (58) positive. Another point worth making is that we naturally would expect the conditions on p and λ as given above in (58) to be more stringent than the ones in (55), for the noise-free case. Clearly the condition on p is more restrictive, as $3/28 < 1/4$. For the bound on λ, it again immediately follows that

$$\frac{3-28p}{25} < \frac{1-4p}{3}, \tag{59}$$

whenever $p < 1$.

Now, to complete the verification of the SPE property of the pair (47) and (46), we still have to show that the unique maximizers of the strictly concave (under (58)) quadratic conditional costs (57) and (56) are given by (46). For the former, the result follows readily since its maximizer is $v = 0$. For the latter, a simple differentiation with respect to v, and using $E[x_0|y_0, \beta_0 = 1] = (1/2)y_0$, leads after some extensive calculations and simplifications to $v = [\lambda/(1-3\lambda)]y_0$, which is the same as (46).

Hence, the SPE for the 2-stage game of this section (with noisy channels and nonzero failure probabilities) is a CE SPE, but for a more restrictive set of values for p and λ (compare (58) with (55), as we have noted earlier). The question now is whether the gap can be closed by using non-CE policies, as was done in the failure-free case ($p = 0$). Clearly, the upper bound of the game is finite for the entire set of values of p and λ in (55) (simply substitute (52) into the cost for u, with additive noise in y_1 and y_0) and note that the presence of additive noise in the channels does not alter the required concavity condition, and hence we have a well-defined strictly concave quadratic maximization problem for v under the same condition (55).

As already mentioned, the region in the parameter space $\lambda - p$ that corresponds to the CE SPE (47) and (46) is smaller than the region corresponding to the SPE of the noise-free channel case, and the question now is whether region of existence of a SPE can be enlarged by transitioning to a pair of non-CE policies, as it was done in Sect. 3.4 for the case when $p = 0$. We will see that this is indeed the case, and

an equalizer policy for Player 1 does the job. Its derivation, however, is a bit more complicated than the one of Sect. 3.4 because the possibility of channel failures brings in an additional element of complexity (even though the basic idea is still the same). Let us first assume that $\beta_0 = 1$, and start with a general linear policy for Player 1:

$$u = \hat{\gamma}(y_1, y_0, \beta_1) = \alpha_1 y_1 + \alpha_0(\beta_1) y_0, \tag{60}$$

where α_1, $\alpha_0(\beta_1 = 1) =: \alpha_0^1$, $\alpha_0(\beta_1 = 0) =: \alpha_0^0$ are constant parameters yet to be determined.[3] They will be determined in such a way that with (60) used in $J(\gamma, v)$, the latter expression becomes independent of v (when conditioned on y_0). Skipping the details, the expression for

$$J(\hat{\gamma}, v) = E\big\{(x_1 - \hat{\gamma}(y_1, y_0, \beta_1) + w_1)^2 + \lambda\big(\hat{\gamma}(y_1, y_0, \beta_1)\big)^2 - v^2 | y_0, \beta_0 = 1\big\} \tag{61}$$

is

$$p\big[x_0 + 2v - \alpha_0^0 y_0\big]^2 + (1 - p)\big[(1 - \alpha_1)(x_0 + 2v) - \alpha_0^1 y_0\big]^2$$
$$+ \lambda(1 - p)\big[\alpha_1(x_0 + 2v) + \alpha_0^1 y_0\big]^2 - v^2 + (1 - p)\lambda\big(2(\alpha_1)^2 + \big(\alpha_0^1\big)^2 + 1\big)$$
$$+ 2p + (1 - p)\big((\alpha_1)^2 + (1 - \alpha_1)^2 + 1\big) + \lambda p\big(\alpha_0^0\big)^2 (y_0)^2 + 1. \tag{62}$$

This is a quadratic function of v; the coefficient of v^2 can be annihilated by choosing

$$\alpha_1 = \frac{1}{1 + \lambda}\left[1 - \frac{\sqrt{(1 - 4p - 3\lambda)(1 - p)}}{2(1 - p)}\right], \tag{63}$$

which is a well-defined expression provided that $4p + 3\lambda < 1$, and naturally (since $\lambda > 0$) also $p < 1/4$ which are identical to (55). For annihilation of the coefficient of v, on the other hand, we need the following relationship between α_0^1 and α_0^0:

$$2p\alpha_0^0 + \alpha_0^1\sqrt{(1 - 4p - 3\lambda)(1 - p)} = \frac{1}{4}. \tag{64}$$

Now, we have to show that these are best responses to some policy by Player 2, which will necessarily be a mixed strategy, as in Sect. 3.4. The process now is to assume that v has the form[4]

$$v = \hat{\mu}(y_0) = k_0 y_0 + \xi, \quad \xi \sim N\big(0, \sigma^2\big), \tag{65}$$

for some k_0 and σ^2; find the best response of Player 1 to this by minimizing $J(\gamma, \hat{\mu})$ with respect to γ (which, by strict convexity, will clearly be unique, and be in

---

[3] We have taken α_1 not to be dependent on β_1, because when $\beta_1 = 0$, $y_1 \equiv 0$, making α_1 superfluous.

[4] One can take any form here, since $\hat{\gamma}$ had annihilated v, but we take linear-plus-Gaussian in anticipation of $\hat{\gamma}$ to be in equilibrium with v.

the structural form (60)); require consistency with (63)–(64); and solve for k_0, σ^2, α_0^1, α_0^0. The outcome is the following set of unique expressions:

$$k_0 = \frac{\lambda}{1 - 3\lambda}, \qquad \sigma^2 = \frac{\sqrt{(1-p)}}{2\sqrt{(1-4p-3\lambda)}} - \frac{5}{8} \tag{66}$$

$$\alpha_0^0 = \frac{1}{2(1-3\lambda)}, \qquad \alpha_0^1 = \frac{\sqrt{(1-4p-3\lambda)(1-p)}}{4(1-p)(1-3\lambda)}. \tag{67}$$

To complete the construction of the SPE, we still have to find the conditions under which σ^2 as given above is well defined (that is, it is positive). The required condition (both necessary and sufficient, provided that (55) holds, which is a natural condition) is

$$4(1-p) > 5\sqrt{(1-4p-3\lambda)(1-p)} \quad \Leftrightarrow \quad \lambda > \max\left(0, \frac{3-28p}{25}\right). \tag{68}$$

Note that the (lower) bound on λ matches exactly the upper bound in (58), and hence non-CE SPE policies make up for the restriction brought in by the CE SPE.

To gain further insight (for purposes of establishing continuity) we can look at two limiting cases: (i) For $p = 0$, the non-CE solution matches exactly the one given in Sect. 3.2 for the failure-free case. (ii) At $\lambda = (3-28p)/25$, with $p > 3/28$, which is the boundary between the two regions corresponding CE and non-CE SPE, α_1 in (63) is $3/[5(1+\lambda)]$, which is exactly the coefficient of y_1 in (47) with $\beta_1 = \beta_0 = 1$; α_0^1 in (67) is $1/[5(1-3\lambda)]$, which is exactly the coefficient of y_0 in (47) with $\beta_1 = \beta_0 = 1$; and finally, α_0^0 in (67) (which does not depend on p) is exactly the coefficient of y_0 in (47) with $\beta_1 = 0$, $\beta_0 = 1$, and this one is for all λ satisfying all other conditions, and not only at the boundary.

The remaining case to analyze is $\beta_0 = 0$. The CE SPE in this case would be (from (46)–(47)):

$$\gamma^*(y_1; \beta_1) = \frac{2}{3(1+\lambda)} y_1, \qquad v^* = 0, \tag{69}$$

which is a valid one (that is the cost under γ^* is strictly concave in v) if and only if the multiplying term for v^2 is negative, that is

$$(1-p)\left[\frac{4(3\lambda+1)^2}{9(1+\lambda)^2} + \frac{16\lambda}{9(1+\lambda)^2}\right] - 1 + 4p < 0,$$

which simplifies to $\lambda < (5-32p)/27$, for which we need $p < 5/32$ (for positivity). To extend the solution to a larger region, we again have to look for an equalizer policy that annihilates v, and is also best response to $v = \xi \sim N(0, \sigma^2)$ for some σ^2. Following the same process as earlier, we start with $u = \alpha_1 y_1$, and compute the cost faced by Player 2, where the multiplying term for v^2 is:

$$4(1-p)\left[(1-\alpha_1)^2 + \lambda(\alpha_1)^2\right] - 1 + 4p.$$

Setting this equal to zero, and solving for α_1 we obtain the expression given by (63). Now, the best response by Player 1 to $v = \xi$ is

$$u = \frac{1}{1+\lambda} E[x_0 + 2\xi + w_0 | y_1] = \frac{1}{\lambda}\left(1 - \frac{1}{2}\sqrt{1 - 3\lambda}\right) y_1,$$

where we then invoke the multiplying term above to equal α_1 given by (63), which leads to the following unique expression for σ^2:

$$\sigma^2 = \frac{\sqrt{1-p}}{2\sqrt{1 - 4p - 3\lambda}} - \frac{3}{4}, \tag{70}$$

which is well-defined and positive provided that

$$\max\left(0, \frac{5 - 32p}{27}\right) < \lambda < \frac{1 - 4p}{3}. \tag{71}$$

Note that the lower bound on λ matches the upper bound in the case of the CE SPE, and that the SPE policy of Player 1 is continuous across the boundary $\lambda = (5 - 32p)/27$.

We now collect all this in the following theorem, which is the main result of this section.

Theorem 1 *The two stage discrete-time stochastic game formulated in this section admits a saddle-point equilibrium* (*SPE*) *provided that*

$$0 \le p < \frac{1}{4} \quad and \quad 0 < \lambda < \frac{1 - 4p}{3};$$

otherwise, the upper value of the game is unbounded. The SPE policies of the players, (γ^*, μ^*), *admit different characterizations in two different regions of the parameter space, and also depending on whether* $\beta_0 = 1$ *or* 0:

- *For* $\lambda \le (5 - 32p)/27$ *and* $p < 5/32$ *when* $\beta_0 = 0$, *and* $\lambda \le (3 - 28p)/25$ *and* $p < 3/28$ *when* $\beta_0 = 1$, γ^* *and* μ^* *are given by* (47) *and* (46), *respectively; this constitutes a certainty-equivalent* (*CE*) *SPE*.
- *For* $\max(0, (5 - 32p)/27) < \lambda < 1 - (4p/3)$ *and* $p < 1/4$ *when* $\beta_0 = 0$, *and* $\max(0, (3 - 28p)/25) < \lambda < 1 - (4p/3)$ *and* $p < 1/4$ *when* $\beta_0 = 1$, *the SPE policies are of the non-CE type, given by*

$$\gamma^*(y_1, y_0; \beta_1, \beta_0) = \alpha_1 y_1 + \left(\beta_1 \alpha_0^1 + (1 - \beta_1)\alpha_0^0\right) y_0 \tag{72}$$

$$\mu^*(y_0; \beta_0) = k_0 y_0 + \xi,$$

$$\xi \sim N\left(0, \sigma^2\right), \quad \sigma^2 = \frac{\sqrt{1-p}}{2\sqrt{1 - 45 - 3\lambda}} - \frac{3}{4} + \frac{1}{8}\beta_0, \tag{73}$$

where α_1 *is given by* (63), *and the pair* (α_0^0, α_0^1) *is given by* (67).

4 CE SPE of the LQG ZSDG in Continuous and Discrete Time

4.1 Various Approaches Toward Construction of SPE

For a two-person ZSDG (in normal or strategic form), with strategy spaces Γ (for Player 1, the minimizer) and $\mathcal{M}$ (for Player 2, the maximizer), with (expected) cost function J, defined on $\Gamma \times \mathcal{M}$, let us recall from (5) that a pair $(\gamma^* \in \Gamma, \mu^* \in M)$ is in SPE if

$$J(\gamma^*, \mu) \le J(\gamma^*, \mu^*) \le J(\gamma, \mu^*), \quad \forall \gamma \in \Gamma, \mu \in \mathcal{M}.$$

The general direct approach toward derivation of a SPE would be:

- Fix $\mu \in \mathcal{M}$ as an *arbitrary* policy for Player 2, and minimize $J(\gamma, \mu)$ with respect to γ on Γ.
- Fix $\gamma \in \Gamma$ as an *arbitrary* policy for Player 1, and maximize $J(\gamma, \mu)$ with respect to μ on $\mathcal{M}$.
- Look for a fixed point, which would then be a SPE.

Even though direct, this approach would entail a very complex process for *dynamic* games (in continuous or discrete time), even if they are of the linear-quadratic type. Unless the information structure is static, the optimization problems involved structurally depend on the selection of arbitrarily fixed policies, rendering the underlying optimization problems unwieldy.

An alternative (still direct) approach would be a recursive (*backward-forward*) one, applicable to discrete-time dynamic games with particular information structures, and possibly extendable to some classes of continuous-time ZSDGs:

- Proceed recursively at $t = T - 1, T - 2, \ldots$.
- At t, solve for SPE (if exists) of the 1-stage game by fixing in J policies for $t + 1, \ldots, T - 1$ at their optimum choices and for $0, \ldots, t - 1$ arbitrarily, with the former possibly depending on the optimum policies (yet to be determined) at $0, 1, \ldots, t$.

Such a construction is doable, but it is quite tedious (and depends on the specific information structure, and applies primarily to discrete-time games); for such a derivation, in a broader Nash equilibrium context, see Başar (1978a).

A third, *indirect* approach, entails expansion of information structures of the players, obtaining a SPE in the induced expanded (richer) policy spaces, and then projecting the solution (*contracting* it) back to the original policy spaces. Such an approach works when the SPE values of the two games (one on original policy spaces and the other one on the expanded ones) are the same, and this is generally the case if the *expansion* involves only past actions of the players. Hence, we have the following process:

- Endow both players with past actions, assuming that they already have access to the past measurements in terms of which the actions were generated.

- Any SPE to the original stochastic dynamic game (SDG) is also a SPE to the new one (but naturally not vice versa).
- Any two SPE of the new SDG are *ordered interchangeable*.
- Solve for some (conveniently constructed) SP policies for the new SDG, and find *representations* (Başar and Olsder 1999) in the original policy spaces.
- *Verify* for the original SDG that the policies arrived at are indeed in SPE (this step is a verification of existence, which is much simpler than verifying characterization).

A further justification of this indirect approach can be found in Başar (1981). In the next two subsections, we illustrate the approach on the two LQ ZSDGs introduced and discussed in Sects. 2.1, 2.2, 2.4, and 2.5. While doing this, we have to keep in mind the features we have observed within the context of the 2-stage ZS SDG of Sect. 3.

4.2 SPE Property of CE Policies of the LQG ZSDG

We turn here to the continuous-time LQG ZSDG of Sect. 2.1, for which the CE policies (24)–(25) were offered as a candidate SPE for the original SDG with noise in the common channel. We now investigate whether these policies are indeed in SPE for at least some region of the parameter space (as was the case for the 2-stage game of Sect. 3). Toward this end, we first enlarge the policy spaces of the players to include also past actions, that is, the players now have access to $(y_{[0,t)}, u_{[0,t)}, v_{[0,t)})$ at time t. Denote the corresponding expanded policy spaces for Players 1 and 2 by $\tilde{\Gamma}$ and $\tilde{\mathcal{M}}$, respectively. If $y_{[0,t)}$ was replaced by x_t (that is, the perfect state measurement case) and still allowing players to have access to past actions, the pair of policies (21)–(22) would still be in SPE (Başar and Olsder 1999), whose CE counterpart in $\tilde{\Gamma} \times \tilde{\mathcal{M}}$ would still be of the form (24)–(25), with however $\{\hat{x}_t, t \geq 0\}$ replaced by $\{\zeta_t, t \geq 0\}$, generated by

$$d\zeta_t = (A\zeta_t + Bu_t + Dv_t)dt + K(t)(dy_t - H\zeta_t dt), \quad \zeta_0 = 0, t \geq 0. \quad (74)$$

Note that the above is still the Kalman filter equation, but driven not only by the measurement but also by the past actions. Now, one can show using the ordered interchangeability property of multiple SPE policies that any pair of SP policies in $\Gamma \times \mathcal{M}$ also constitute a SPE in the expanded policy spaces $\tilde{\Gamma} \times \tilde{\mathcal{M}}$ (but not vice versa) (Başar and Olsder 1999; Başar 1981), and further that by some standard properties of the LQG control problem discussed in Sect. 2.3, the pair (24)–(25) indeed constitutes a SPE for the new SDG with expanded policy spaces, provided that the RDE (23) does not have a conjugate point in the interval $[0, t_f]$, which is exactly the condition of existence of SPE to the LQG ZSDG with perfect-state measurements.

Clearly, however, the SP policies above for the SDG with expanded policy spaces are not implementable even for that game, because they require cooperation on the

generation of the conditional mean of x, or that estimate ζ (as in (74)) to be generated by a third party, and supplied to the two antagonistic players, which is not realistic. To make it real-time implementable, and in line with the adversarial aspect of the game, we have to replace these policies with ones that allow players to run their own filters, driven by the common measurement (but not with actions of the players), as given below:

$$u^*(t) = \gamma_t^*(y_{[0,t)}) = \gamma_t^{CE}(z_t) = -\frac{1}{\lambda}B^T Z(t)z_t, \quad t \geq 0, \tag{75}$$

$$v^*(t) = \mu_t^*(y_{[0,t)}) = \mu_t^{CE}(\eta_t) = D^T Z(t)\eta_t, \quad t \geq 0, \tag{76}$$

where z and η are generated by (as counterpart of (74)):

$$dz_t = \hat{A}z_t dt + K(t)(dy_t - Hz_t dt), \quad z_0 = 0, t \geq 0, \tag{77}$$

$$d\eta_t = \hat{A}\eta_t dt + K(t)(dy_t - H\eta_t dt), \quad \eta_0 = 0, t \geq 0, \tag{78}$$

where

$$\hat{A} := A - \left(\frac{1}{\lambda}BB^T - DD^T\right)Z(t),$$

K is the Kalman gain, given again by (18), with Λ solving (28).

The policies (γ^{CE}, μ^{CE}) given by (75)–(76) constitute representations of the SP policies in the expanded policy spaces and now belong to $\Gamma \times \mathcal{M}$, and as such also constitute SPE for the original SDG, as argued earlier, provided that the response of Player 1 to (76) and that of Player 2 to (75) are well defined, leading to bounded costs. For the former, it can be shown easily (and in fact argued without any explicit computation) that

$$\min_{\gamma \in \Gamma} J\left(\gamma, \mu^{CE}\right) = J\left(\gamma^{CE}, \mu^{CE}\right),$$

and particularly that the quadratic function $J(u, \mu^{CE})$ is strictly convex in u. This establishes the right-hand-side of the SP inequality (5). For the left-hand-side inequality, on the other hand, we have the LQG optimal control problem

$$\max_{\mu \in \mathcal{M}} J\left(\gamma^{CE}, \mu\right),$$

with $2n$-dimensional differential constraints:

$$dx_t = (Ax + Dv_t)dt - \frac{1}{\lambda}B^T Z(t)z_t dt + F dw_t, \quad t \geq 0,$$

$$dz_t = \tilde{A}z_t dt + K(t)(dy_t - Hz_t dt), \quad z_0 = 0, t \geq 0.$$

The conjugate-point condition on (23) is not sufficient for this LQG optimal control problem to be well defined, as the cost $J(\gamma^{CE}, v)$ could be non-concave in v. Strict concavity here is in fact the only condition that would be needed for the pair

$(\gamma^{\mathrm{CE}}, \mu^{\mathrm{CE}})$ in (75)–(76) to constitute a SPE. Now note that $J(\gamma^{\mathrm{CE}}, v)$ can be written as

$$J(\gamma^{\mathrm{CE}}, v) = E\left\{ |x_{t_f}|^2_{Q_f} + \int_0^{t_f} \left[|x_t|^2_Q + \frac{1}{\lambda} |B^T Z(t) z_t|^2 - |v_t|^2 \right] dt \right\}$$

$$=: E\left\{ |m_{t_f}|^2_{\tilde{Q}_f} + \int_0^{t_f} \left[|m_t|^2_{\tilde{Q}} - |v_t|^2 \right] dt \right\},$$

where $m := (x^T z^T)^T$, $\tilde{Q}_f := \text{block diag}(Q_f, 0)$, and $\tilde{Q} := \text{block diag}(Q, (1/\lambda) \times Z B B^T Z)$. Further, m evolves according to

$$dm_t = \tilde{A} m_t dt + \tilde{D} v_t dt + \tilde{F} dw_t,$$

where $\tilde{D} := [D^T, 0]^T$, $\tilde{F} := [F^T, G^T K^T]^T$, and $\tilde{A}$ is a $2n \times 2n$ matrix, whose ij-th block is, for $i, j = 1, 2$: $[\tilde{A}]_{11} := A$, $[\tilde{A}]_{21} := KH$, $[\tilde{A}]_{12} := -(1/\lambda) B B^T Z$, and $[\tilde{A}]_{22} := \hat{A} - KH$.

The condition for strict concavity for this optimization problem, regardless of the nature of the information available to Player 2, is (Başar and Bernhard 1995) nonexistence of a conjugate point to the RDE below on the interval $[0, t_f]$:

$$\dot{S} + S\tilde{A} + \tilde{A}^T S + S\tilde{D}\tilde{D}^T S + \tilde{Q} = 0, \quad S(t_f) = \tilde{Q}_f. \tag{79}$$

We now collect all this in the theorem below.

Theorem 2 *The continuous-time LQG ZSDG of Sect.* 2.1 *admits a pure-strategy SPE provided that the RDEs* (23) *and* (79) *have well-defined nonnegative-definite solutions on the interval* $[0, t_f]$, *in which case the corresponding policies for the players, in SPE, are given by* (75)–(76). *These feature a* restricted certainty equivalence *property*.

Remark 1 A number of observations are in order here. First, the policies in SPE for the SDG are not simple CE versions of the SPE of the deterministic game, that is they are not the pair (24)–(25), even though they can be derived from the SPE of the deterministic game by endowing the players with two separate filter equations even though the players have access to a common measurement channel. Second, the condition of existence of a pure-strategy SPE for the SDG is more restrictive than its counterpart for the perfect-state version (or essentially equivalently the deterministic game). This would not be surprising in view of the results of Sect. 3 for perhaps the simplest stochastic dynamic game, where the gap between the two conditions (for existence of pure-strategy SPE in the games with perfect state and noisy state information) was completely covered by allowing for mixed strategies (for the maximizing player). It is quite plausible that the same would hold here, but derivation of such a mixed-strategy SPE for the continuous-time LQG ZSDG still remains a challenging task.

4.3 SPE Property of CE Policies of the LQG Discrete-Time ZSDG

We now proceed with an analysis that is the counterpart of the one above (in Sect. 4.2) for the discrete-time game of Sect. 2.2 and Sect. 2.5, not for the most general case, but for two scenarios: (i) when there is no failure of channels (that is $p = 0$, as in Sect. 3.2), and (ii) the channel provides perfect state measurement, but intermittently fails (as in Sect. 3.3). In both cases, we obtain restricted CE SPE. The derivation is a direct counterpart of the one in Sect. 4.2), and hence to avoid duplication we will just provide the basic results without providing details of the reasoning and the pathway.

Let us first discuss case (i). In Sect. 2.5, we had offered (34)–(35) as a candidate SPE pair for this scenario, but as we have argued in the previous subsection, having a single filter to be shared by both players is not a realistic situation, and hence we will have to introduce individualized compensators. In view of this, (34)–(35) will have to be modified as follows:

$$u_t^* = \gamma_t^*(y_{[0,t]}) = \gamma_t^{CE}(z_{t|t}) = -\frac{1}{\lambda} B^T Z_{t+1}\big(N_t^{-1}\big)^T A z_{t|t}, \quad t = 0, 1, \ldots, \quad (80)$$

$$v_t^* = \mu_t^*(y_{[0,t]}) = \mu_t^{CE}(\eta_t) = D^T Z_{t+1}\big(N_t^{-1}\big)^T A \eta_{t|t}, \quad t = 0, 1, \ldots, \quad (81)$$

where $z_{t|t}$ and $\eta_{t|t}$ are generated by, with $z_{0|-1} = 0$:

$$
\begin{aligned}
z_{t|t} &= z_{t|t-1} + \Lambda_t H^T \big(H \Lambda_t H^T + GG^T\big)^{-1}(y_t - H z_{t|t-1}) \\
z_{t+1|t} &= (N_t)^{-1} A z_{t|t-1} \\
&\quad + (N_t)^{-1} A \Lambda_t H^T \big(H \Lambda_t H^T + GG^T\big)^{-1}(y_t - H z_{t|t-1}),
\end{aligned}
\tag{82}
$$

and, with $\eta_{0|-1} = 0$,

$$
\begin{aligned}
\eta_{t|t} &= \eta_{t|t-1} + \Lambda_t H^T \big(H \Lambda_t H^T + GG^T\big)^{-1}(y_t - H \eta_{t|t-1}) \\
\eta_{t+1|t} &= (N_t)^{-1} A \eta_{t|t-1} \\
&\quad + (N_t)^{-1} A \Lambda_t H^T \big(H \Lambda_t H^T + GG^T\big)^{-1}(y_t - H \eta_{t|t-1}),
\end{aligned}
\tag{83}
$$

and the sequence $\{\Lambda_t, t = 1, 2, \ldots, T\}$ is as in (37). By going through similar arguments as in the previous subsection, the pair (80)–(81) provides a SPE, provided that (33) holds and the quadratic function $J(\gamma^{CE}, v)$ is strictly concave in v. An explicit condition can be obtained for the latter in terms of a $2n \times 2n$ discrete-time Riccati equation, which involves a recursive verification as in (33).

For case (ii), that is when $y_t = \beta_t x_t$, $t = 0, 1, \ldots, T - 1$, the starting point is the pair of policies (38)–(39), where as before we endow the players with two separate compensators, with states ζ^1 and ζ^2, generated by, for $i = 1, 2$,

$$\zeta_t^i = \beta_t y_t + (1 - \beta_t)\left(I - \left(\frac{1}{\lambda}BB^T + DD^T\right)Z_{t+1}(N_t^{-1})^T\right)A\zeta_{t-1}^i, \quad \zeta_0^i = y_0.$$

(84)

Hence, the players' CE policies become

$$u_t^* = \gamma_t^*(y_{[0,t]}) = \tilde{\gamma}_t^{CE}(\zeta_t) = -\frac{1}{\lambda}B^T Z_{t+1}(N_t^{-1})^{-1}A\zeta_t^1, \quad t = 0, 1, \ldots, \quad (85)$$

$$v_t^* = \mu_t^*(y_{[0,t]}) = \tilde{\mu}_t^{CE}(\zeta_t^2) = D^T Z_{t+1}(N_t^{-1})^{-1}A\zeta_t, \quad t = 0, 1, \ldots, \quad (86)$$

which, by an argument similar to the earlier case, are in SPE provided that (33) holds and the quadratic function $J(\tilde{\gamma}^{CE}, v)$ is strictly concave in v. As before, an explicit condition can be obtained for the latter in terms of a $2n \times 2n$ discrete-time Riccati equation, which involves a recursive verification as in (33).

In view of the complete set of results of Sect. 3 for a 2-stage version of this game, for case (ii), we would not expect a less stringent condition to be obtained (that is, there would not be any need to expand the policy spaces to include mixed strategies), whereas for case (i) the extra condition introduced in terms of strict concavity of $J(\gamma^{CE}, v)$ in v can be dispensed with by inclusion of mixed strategies. We do not pursue this any further here.

For the more general case, however, when the channel is noisy and failure probability is $p > 0$, still a restricted CE will hold, with z and η in (82)–(83) now incorporating the possibility of failures, as in the case of derivation of Kalman filters with missing measurements (Shi et al. 2010). Here also a strict concavity condition will be needed for the existence of a pure-strategy SPE, in addition to the one for $p = 0$, which however can be dispensed with by inclusion of mixed strategies.

5 Discussion, Extensions, and Conclusions

One important message that this chapter conveys (which applies to more general differential/dynamic games with similar information structures) is that in zero-sum stochastic differential/dynamic games (ZS SDGs) a restricted certainty equivalence (CE) applies if players have a common measurement channel, but the adversarial nature of the problem creates several caveats not allowing the standard notions of certainty equivalence or separation prominent in stochastic control problems (Witsenhausen 1971a, 1971b; Fleming and Soner 1993; Yüksel and Başar 2013) to find an immediate extension. Expansion of information structures to include also action information compatible with the original information, and without increasing payoff relevant information, appears to be a versatile tool in an indirect derivation of pure-strategy saddle-point equilibria (SPE), which however does not apply to derivation of mixed-strategy SPE, as it relies heavily on the ordered interchangeability property of multiple pure-strategy SPE. For the same reason, the indirect approach does not apply to nonzero-sum dynamic games; in fact, Nash equilibria of genuinely

nonzero-sum stochastic games (unless they are strategically equivalent to zero-sum games or team problems) never satisfy CE (Başar 1978b). Now, coming back to ZS SDGs, when a generalized CE SPE exists in some region of the parameter space, this is not the full story because the game may also admit mixed-strategy SPE outside that region, which however has to be obtained using a different approach–using notions of *annihilation* and *conditional equalization*, as it has been demonstrated in Sect. 3. Hence, expansion of strategy (policy) spaces from *pure* to *mixed* helps to recover the missing SPE.

We have deliberately confined our treatment in this paper to ZS SDGs with *symmetric information*, to be able to introduce a *restricted* (and in some sense *generalized*) notion of CE and to show that any attempt of directly extending CE from stochastic optimal control to games is a path full of pitfalls, even though the problem (of derivation of SPE) is still tractable, but using an indirect approach (that makes use of expansion of strategy spaces and ordered interchangeability property of multiple pure-strategy SPE). As indicated earlier, this approach does not extend to nonzero-sum games (NZSGs), because expansion of strategy spaces (through actions) leads to multiplicity of Nash equilibria, and in fact a continuum of them (Başar and Olsder 1999), and multiple Nash equilibria (NE) are not orderly interchangeable. Still, there is another approach to derivation of NE with nonredundant information, as briefly discussed in Sect. 1, provided that we have a discrete-time game, with complete sharing of information (that is, a common measurement channel) or sharing of observations with one step delay (Başar 1978a). The same approach would of course apply to ZSDGs as well (with one-step delayed sharing), but then the SPE will not be of the CE type. If there is no sharing of information (or with delay of two units or more), and players receive noisy state information through separate channels, then the problem remains to be challenging in both ZS and NZS settings, unless there is a specific structure of the system dynamics along with the information structure, as in Nayyar and Başar (2012).

Several fairly direct extensions of the results of this chapter are possible, all in the ZS setting. First, it is possible to introduce intermittent failure of the common measurement channel (2) in the continuous-time case, by mimicking (8):

$$dy_t = \beta_t(Hx_t dt + Gdw_t) \quad \text{or} \quad dy_t = \beta_t Hx_t dt + Gdw_t, \quad t \geq 0,$$

where β_t is an independent two-state Markov jump process (or a piecewise deterministic process) with a given rate (jumps are between $\beta_t = 1$ and $\beta_t = 0$), and both players observe realization of this process. The counterpart of the analysis for the discrete-time case could be carried over to this case also (for a related framework, see Pan and Başar 1995). A variant of this, in both discrete and continuous time, is the more challenging class of problems where the failure of the transmission of the common noisy measurement of the state to the players is governed by two independent Bernoulli processes with possibly different rates. Such ZS SDGs would involve primarily two scenarios: (i) the players are not aware of the failure of links corresponding to each other, and (ii) this information is available (that is players share explicitly or implicitly the failure information) but with one step delay.

Further extensions to (i) multi-player ZS SDGs (with teams playing against teams, where agents in each team do not have identical information), and (ii) nonzero-sum stochastic differential games (with particular type of asymmetric information among the players) constitute yet two other classes of challenging problems. In all these problems, including the ones discussed in Sect. 4, characterization of mixed-strategy SPE (as extension of the analysis of Sect. 3) or NE stand out as challenging but tractable avenues for future research.

References

Altman, E., Kambley, V., & Silva, A. (2009). Stochastic games with one step delay sharing information pattern with application to power control. In *Proc. internat. conf. game theory for networks (GameNets'09)*, Turkey, Istanbul (pp. 124–129).

Bansal, R., & Başar, T. (1987). Stochastic teams with nonclassical information revisited: when is an affine law optimal? *IEEE Transactions on Automatic Control*, *32*(6), 554–559.

Başar, T. (1978a). Two-criteria LQG decision problems with one-step delay observation sharing pattern. *Information and Control*, *38*(1), 21–50.

Başar, T. (1978b). Decentralized multicriteria optimization of linear stochastic systems. *IEEE Transactions on Automatic Control*, *23*(2), 233–243.

Başar, T. (1981). On the saddle-point solution of a class of stochastic differential games. *Journal of Optimization Theory and Applications*, *33*(4), 539–556.

Başar, T. (2008). Variations on the theme of the Witsenhausen counterexample. In *Proc. 47th IEEE conf. decision and control* (pp. 1614–1619).

Başar, T., & Bernhard, P. (1995). *Systems & control: foundations and applications series*. *H∞ optimal control and related minimax design problems: a dynamic game approach*. Boston: Birkhäuser.

Başar, T., & Mintz, M. (1972). On the existence of linear saddle-point strategies for a two-person zero-sum stochastic game. In *Proc. 1972 IEEE conf. decision and control* (pp. 188–192).

Başar, T., & Mintz, M. (1973). A multistage pursuit-evasion game that admits a Gaussian random process as a maximin control policy. *Stochastics*, *1*(1–4), 25–69.

Başar, T., & Olsder, G. J. (1999). *SIAM series in classics in applied mathematics*. *Dynamic non-cooperative game theory*. Philadelphia: Society for Industrial and Applied Mathematics.

Behn, R., & Ho, Y.-C. (1968). On a class of linear stochastic differential games. *IEEE Transactions on Automatic Control*, *13*(3), 227–240.

Fleming, W. H., & Soner, H. M. (1993). *Controlled Markov processes and viscosity solutions*. Berlin: Springer.

Hespanha, J., & Prandini, M. (2001). Nash equilibria in partial-information games on Markov chains. In *Proc. 40th IEEE conf. decision and control* (pp. 2102–2107).

Ho, Y.-C. (1974). On the minimax principle and zero-sum stochastic differential games. *Journal of Optimization Theory and Applications*, *13*(3), 343–361.

Ho, Y.-C. (1980). Team decision theory and information structures. *Proceedings of the IEEE*, *68*(6), 644–654.

Nayyar, A., & Başar, T. (2012). Dynamic stochastic games with asymmetric information. In *Proc. 51st IEEE conf. decision and control*, Maui, Hawaii (pp. 7145–7150).

Nayyar, A., Mahajan, A., & Teneketzis, D. (2011). Optimal control strategies in delayed sharing information structures. *IEEE Transactions on Automatic Control*, *57*(7), 1606–1620.

Pan, Z., & Başar, T. (1995). H^∞ control of Markovian jump systems and solutions to associated piecewise-deterministic differential games. In G. J. Olsder (Ed.), *Annals of dynamic games* (Vol. 2, pp. 61–94). Boston: Birkhäuser.

Rhodes, I., & Luenberger, D. (1969). Differential games with imperfect state information. *IEEE Transactions on Automatic Control*, *14*(1), 29–38.

Shi, L., Epstein, M., & Murray, R. M. (2010). Kalman filtering over a packet-dropping network: a probabilistic perspective. *IEEE Transactions on Automatic Control*, *55*(1), 594–604.

Willman, W. (1969). Formal solutions for a class of stochastic pursuit-evasion games. *IEEE Transactions on Automatic Control*, *14*(5), 504–509.

Witsenhausen, H. (1968). A counterexample in stochastic optimum control. *SIAM Journal on Control*, *6*, 131–147.

Witsenhausen, H. (1971a). Separation of estimation and control for discrete time systems. *Proceedings of the IEEE*, *59*(11), 1557–1566.

Witsenhausen, H. (1971b). On information structures, feedback and causality. *SIAM Journal on Control*, *9*(2), 149–160.

Yüksel, S., & Başar, T. (2013). *Systems & control: foundations and applications series. Stochastic networked control systems: stabilization and optimization under information constraints.* Boston: Birkhäuser.