Article

A Game—Theoretic Model for a Stochastic Linear Quadratic Tracking Problem

Vasile Drăgan [1,2,†] , Ivan Ganchev Ivanov [3,†] and Ioan-Lucian Popa [4,5,*,†]

1   "Simion Stoilow" Institute of Mathematics, Romanian Academy, P.O. Box 1-764, 014700 Bucharest, Romania
2   Academy of the Romanian Scientists, Str. Ilfov, Nr. 3, 50044 Bucharest, Romania
3   Faculty of Economics and Business Administration, Sofia University "St. Kl. Ohridski",
    125 Tzarigradsko Chaussee Blvd., bl. 3, 1113 Sofia, Bulgaria
4   Department of Computing, Mathematics and Electronics, Faculty of Computing and Engineering,
    "1 Decembrie 1918" University of Alba Iulia, 510009 Abla Iulia, Romania
5   Faculty of Mathematics and Computer Science, Transilvania University of Braşov, Iuliu Maniu Street 50,
    500091 Braşov, Romania
*   Correspondence: lucian.popa@uab.ro
†   These authors contributed equally to this work.

Abstract: In this paper, we solve a stochastic linear quadratic tracking problem. The controlled dynamical system is modeled by a system of linear Itô differential equations subject to jump Markov perturbations. We consider the case when there are two decision-makers and each of them wants to minimize the deviation of a preferential output of the controlled dynamical system from a given reference signal. We assume that the two decision-makers do not cooperate. Under these conditions, we state the considered tracking problem as a problem of finding a Nash equilibrium strategy for a stochastic differential game. Explicit formulae of a Nash equilibrium strategy are provided. To this end, we use the solutions of two given terminal value problems (TVPs). The first TVP is associated with a hybrid system formed by two backward nonlinear differential equations coupled by two algebraic nonlinear equations. The second TVP is associated with a hybrid system formed by two backward linear differential equations coupled by two algebraic linear equations.

Keywords: linear quadratic tracking problem; stochastic linear differential games; Nash equilibria

MSC: 91A15; 49N10

1. Introduction

Tracking problems are often encountered in many applications and have received attention from the research community in the past few decades [1–5]. In the stochastic context, this problem was studied in [6,7] as well as in [8,9]. For stochastic systems with time delay, the linear quadratic tracking problem has been studied in [10,11]. Applications of tracking problems may be found in economic policy control [12], process control [13], networked control systems [14], control of mobile robots [15], spacecraft hovering [16], etc. Usually, a linear quadratic tracking problem requires minimization of the L_2-norm of the deviation of a signal generated by a controlled linear system from a reference signal. When there exists more than one decision maker and each of them wants to minimize the deviation of a preferential signal from a given reference signal, the optimal tracking problem may be stated as a problem of finding a Nash equilibrium strategy for a linear quadratic differential game. If the controlled system whose outputs have to track the given reference signal is described by linear stochastic differential equations, one obtains a problem of finding a Nash equilibrium strategy for a stochastic differential game. Lately, the stochastic differential games have attracted an increasing research interest see, e.g., [17–20]. Moreover, for Nash tracking game problems for continuous-time systems over finite intervals see [21] and the reference therein.

In the present work, we consider the case when the controlled system is described by a system of Itô differential equations with coefficients affected by a standard homogeneous Markov process with a finite number of states. We assume that there exist at least two decision-makers. The aim of the k-th decision-maker is to minimize the deviation of an output $z_k(\cdot)$ of the controlled system from a reference signal $r_k(\cdot)$. The class of admissible strategies consists of stochastic processes in an affine state feedback form.

In the derivation of the main results we consider two cases:

(a) the case with only one decision-maker;

(b) the case of two decision-makers.

The result derived from case (a) is used in the case with more than one decision-maker in order to obtain an optimal strategy. In case (b), we study the game theoretic model for two players, where each player wants to find the optimal admissible strategy minimizing the deviation of the controlled signal from the given reference.

We assume that the players do not cooperate. The reasons for which they are not cooperating may be caused by individual motivations or by physical reasons. We provide explicit formulae of a Nash equilibrium strategy. To this end, we use the solutions of two TVPs. The first TVP is associated with a hybrid system formed by two backward nonlinear differential equations coupled with two algebraic nonlinear equations. The second TVP is associated with a hybrid system formed by two backward linear differential equations coupled by two algebraic linear equations.

The paper is organized as follows: Section 2 includes the description of the mathematical model as well as the formulation of the tracking problem as a problem of finding a Nash equilibrium strategy for a stochastic differential game. The main results are derived in Section 3. First, in Section 3.1, we consider the case with one decision-maker. Further, in Section 3.2, we obtain explicit formulae for the optimal strategies in the case of two decision-makers. In Section 4, we briefly discuss two special cases: (i) the case when the controlled system does not contain controlled dependent terms in the diffusion part; (ii) the case when the aim of the decision-makers is to minimize the deviation of the controlled signals from a given final target without restrictions regarding the behavior of the transient states. In Section 5, we provide a numerical example that shows that the proposed procedure is feasible. Finally, in Section 6, we provide some conclusions and future research directions.

2. The Problem

Consider the controlled system having the state space representation described by

$$
\begin{aligned}
dx(t) = {} & [A_0(t,\eta_t)x(t) + B_1(t,\eta_t)u_1(t) + B_2(t,\eta_t)u_2(t)]dt \\
& + [A_1(t,\eta_t)x(t) + D_1(t,\eta_t)u_1(t) + D_2(t,\eta_t)u_2(t)]dw(t),
\end{aligned}
$$

$$
x(t_0) = x_0, \tag{1a}
$$

$$
z_k(t) = C_k(t,\eta_t)x(t), \quad k = 1,2, \tag{1b}
$$

$t \in [t_0, t_f] \subset [0, \infty)$, where $x(t) \in \mathbb{R}^n$ is the state vector at instance time t and $u_k(t) \in \mathbb{R}^{m_k}$, $k = 1,2$, are the vectors of control parameters. In (1), $\{w(t)\}_{t\geq 0}$, is a one-dimensional standard Wiener process defined on a given probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and $\{\eta_t\}_{t\geq 0}$ is a standard right continuous Markov process taking values in a finite set $\mathcal{N} = \{1, 2, \ldots, N\}$ and having the transition semigroup $\mathbb{P}(t) = e^{Qt}$, $t \geq 0$. The elements q_{ij} of the generator matrix $Q \in \mathbb{R}^{N \times N}$ satisfy

$$
q_{ij} \geq 0 \ \text{ if } \ i \neq j, \tag{2a}
$$

$$
\sum_{l=1}^{N} q_{il} = 0, \tag{2b}
$$

for all $(i,j) \in \mathcal{N} \times \mathcal{N}$. For more details regarding the properties of a Wiener process, one can see [22] or [23], whereas for more properties of a Markov process, we refer, for example, to [24,25].

Throughout the paper, we assume that $\{w(t)\}_{t\geq 0}$ and $\{\eta_t\}_{t\geq 0}$ are independent stochastic processes. The dependence of the coefficients of the system (1) upon the Markov process $\{\eta_t\}_{t\geq 0}$ highlights the fact that this system may be considered the mathematical model of some phenomena in which abrupt structural changes are likely to occur. Such variations may be due, for instance, to changes between different operation points, sensor or actuator failures, temporary loss of communication, and so on. For the readers' convenience, we refer to [9,26–28] for extensive discussions on the subject. As usual, we shall write $A_k(t,i)$, $k = 0,1$, $B_j(t,i)$, $C_j(t,i)$, $D_j(t,i)$, $j = 1,2$ instead of $A_k(t,\eta_t)$, and so on, whenever $\eta_t = i \in \mathcal{N}$. Assume that $t \to A_k(t,i) : [t_0,t_f] \to \mathbb{R}^{n\times n}$, $k = 0,1$, $t \to (B_j(t,i), C_j(t,i), D_j(t,i)) : [t_0,t_f] \to \mathbb{R}^{n\times m_j} \times \mathbb{R}^{n_{z_j}\times n} \times \mathbb{R}^{n\times m_j}$, $j = 1,2$ are continuous matrix-valued functions. The set $\mathcal{U}_k$ of the admissible controls available to the decision-makers $\mathcal{P}_k$, $k = 1,2$, consists of the stochastic processes in an affine state feedback form, i.e.,

$$u_k(t) = F_k(t,\eta_t)x(t) + \varphi_k(t,\eta_t) \tag{3}$$

where $t \to F_k(t,i) : [t_0,t_f] \to \mathbb{R}^{m_k\times n}$ and $t \to \varphi_k(t,i) : [t_0,t_f] \to \mathbb{R}^{m_k}$ are arbitrary continuous functions. Applying Theorem 1.1 in Chapter 5 from [22] (see also Section 1.12 from [9]) we obtain:

Corollary 1. *For each $x_0 \in \mathbb{R}^n$ and for each $u(\cdot) = (u_1(\cdot),u_2(\cdot)) \in \mathcal{U}_1 \times \mathcal{U}_2$, the initial value problem (IVP) (1a) has a unique solution $x_u(\cdot;t_0,x_0)$ which is a stochastic process with the properties*

(a)　*$x_u(\cdot;t_0,x_0)$ is a.s. continuous in every $t \in [t_0,t_f]$;*

(b)　*for each $t \in [t_0,t_f]$, $x_u(t;t_0,x_0)$ is $\mathcal{F}_t-$measurable, where $\mathcal{F}_t \subset \mathcal{F}$ is the $\sigma-$algebra generated by the random variables $w(s)$, η_s, $0 \leq s \leq t$;*

(c)　*$\sup_{t\in[t_0,t_f]} \mathbb{E}[|x_u(t;t_0,x_0)|^p] < \infty$, for all $p \geq 1$;*

(d)　*$x_u(t_0;t_0,x_0) = x_0$.*

Throughout this work, $\mathbb{E}[\cdot]$ stands for the mathematical expectation. Roughly speaking, the aim of the decision-maker $\mathcal{P}_k$ is to find a control law (or an admissible strategy) of type (3), $\tilde{u}_k(\cdot) \in \mathcal{U}_k$ which minimizes the deviation of the signal $z_k(\cdot)$ from a given reference signal $r_k(\cdot)$, when the other decision-maker $\mathcal{P}_l$ wants to minimize the deviation of the signal $z_l(\cdot)$, $(l \neq k)$ from another reference signal $r_l(\cdot)$.

Since this problem is an optimization problem with two objective functions, its solution can be viewed as an equilibrium strategy for a non-cooperative differential game with two players. For a rigorous mathematical setting of this optimization problem, let us introduce the following cost functions which are modeling the performance criterion for each player $(k = 1,2)$:

$$
\begin{aligned}
J_k(x_0;u_1(\cdot),u_2(\cdot)) =\, & \mathbb{E}[(z_k(t_f) - \zeta_k)^T G_k(\eta_{t_f})(z_k(t_f) - \zeta_k)] \\
& + \mathbb{E}\left[\int_{t_0}^{t_f} \{(z_k(t) - r_k(t))^T M_k(t,\eta_t)(z_k(t) - r_k(t)) \right. \\
& \left. + u_1^T(t) R_{k1}(t,\eta_t)u_1(t) + u_2^T(t) R_{k2}(t,\eta_t)u_2(t)\}dt\right].
\end{aligned}
\tag{4}
$$

Here $r_k(\cdot) : [t_0,t_f] \to \mathbb{R}^{n_{z_k}}$, $k = 1,2$ is the reference which must be tracked by the signal $z_k(\cdot)$, and $\zeta_k \in \mathbb{R}^{n_{z_k}}$, $k = 1,2$ is the target of the final value $z_k(t_f)$. The weight matrices involved in (4) are satisfying the assumption:

Hypothesis 1.
(a) $t \to (M_k(t,i),\ R_{k1}(t,i),\ R_{k2}(t,i)) : [t_0, t_f] \to \mathcal{S}_{n_{z_k}} \times \mathcal{S}_{m_1} \times \mathcal{S}_{m_2}, i \in \mathcal{N}$ *are continuous matrix-valued functions;*
(b) $M_k(t,i) \geq 0, R_{kl}(t,i) \geq 0, R_{kk}(t,i) > 0,$ *for all* $t \in [t_0, t_f], k, l = 1, 2, l \neq k, G_k(i) \geq 0,$ $i \in \mathcal{N}.$

Here and in the sequel, $\mathcal{S}_p \subset \mathbb{R}^{p \times p}$ denotes the subspace of symmetric matrices of size $p \times p, p \geq 1$.

Definition 1. *We say that the pair of admissible strategies* $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) \in \mathcal{U}_1 \times \mathcal{U}_2$ *achieve a Nash equilibrium for the differential game described by the controlled system (1), the performance criterion (4), and the admissible strategies of type (3) if*

$$J_1(x_0; \tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) \leq J_1(x_0; u_1(\cdot), \tilde{u}_2(\cdot)), \quad \text{for all} \ \ u_1 \in \mathcal{U}_1 \tag{5}$$

and

$$J_2(x_0; \tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) \leq J_2(x_0; \tilde{u}_1(\cdot), u_2(\cdot)), \quad \text{for all} \ \ u_2 \in \mathcal{U}_2. \tag{6}$$

In the next section, we shall derive explicit formulae of a Nash equilibrium strategy $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$ for the linear quadratic differential game described by (1), (3), and (4).

Remark 1.
(a) *We shall see that for the computation of the gain matrices of a Nash equilibrium strategy we need to know a priori the whole reference signal* $r_k(\cdot)$.
(b) *When* $M_k(t,i) = 0,$ *for all* $(t,i) \in [t_0, t_f] \times \mathcal{N}$, *then (4) reduces to*

$$J_k(x_0; u_1(\cdot), u_2(\cdot)) = \mathbb{E}[(z_k(t_f) - \zeta_k)^T G_k(\eta_{t_f})(z_k(t_f) - \zeta_k)]$$
$$+ \mathbb{E}\left[\int_{t_0}^{t_f} \{u_1^T(t) R_{k1}(t, \eta_t) u_1(t) + u_2^T(t) R_{k2}(t, \eta_t) u_2(t)\} dt\right], \quad k = 1, 2. \tag{7}$$

The performance criterion (4) could be replaced by one of the form (7), when the decision-maker is interested only by the minimization of the deviation of the final value $z_k(t_f)$ *from the target* ζ_k. *The term*

$$\mathbb{E}\left[\int_{t_0}^{t_f} \{u_1^T(t) R_{k1}(t, \eta_t) u_1(t) + u_2^T(t) R_{k2}(t, \eta_t) u_2(t)\} dt\right],$$

which appears both in (4) and (7), must be viewed as a penalization of the control effort.

3. The Main Results

3.1. The Case with Only One Decision Maker

In order to derive in an elegant way the state space representation of a pair of form (3) which satisfies (5) and (6), respectively, we study first the problem of tracking a reference signal in the case where there is only one decision maker.

We consider the optimal control problem described by the controlled system:

$$dx(t) = [A_0(t, \eta_t)x(t) + B(t, \eta_t)u(t) + g_0(t, \eta_t)]dt$$
$$+ [A_1(t, \eta_t)x(t) + D(t, \eta_t)u(t) + g_1(t, \eta_t)]dw(t),$$
$$x(t_0) = x_0, \tag{8a}$$
$$z(t) = C(t, \eta_t)x(t), \tag{8b}$$

$t \in [t_0, t_f]$ and the performance criterion

$$
\begin{aligned}
J(x_0; u(\cdot)) =& \mathbb{E}[(z(t_f) - \zeta)^T G(\eta_{t_f})(z(t_f) - \zeta)] \\
& + \mathbb{E}\left[\int_{t_0}^{t_f} \{(z(t) - r(t))^T M(t, \eta_t)(z(t) - r(t)) + u^T(t)R(t, \eta_t)u(t)\}dt\right].
\end{aligned}
\tag{9}
$$

Here, the stochastic processes $\{w(t)\}_{t \geq 0}$ and $\{\eta_t\}_{t \geq 0}$ have the same properties as in the case of system (1). In (8) and (9), $x(t) \in \mathbb{R}^n$ is the state vector at the instance time t and $u(t) \in \mathbb{R}^m$ is the vector of the control parameters.

In this subsection, the set $\mathcal{U}$ of admissible controls consists of stochastic processes of the form

$$
u(t) = F(t, \eta_t)x(t) + \varphi(t, \eta_t),
\tag{10}
$$

$t \to F(t, i) : [t_0, t_f] \to \mathbb{R}^{m \times n}$, $t \to \varphi(t, i) : [t_0, t_f] \to \mathbb{R}^m$ which are arbitrary continuous functions. The optimal control problem which we want to solve in this subsection consists in finding a control $\tilde{u}(\cdot)$ from $\mathcal{U}$ which minimizes the cost function (9) along the trajectories of the system (8) determined by all admissible controls of the form (10).

Regarding the coefficients of (8) and (9), we suppose:

Hypothesis 2.
(a) $t \to (A_0(t, i),\ A_1(t, i),\ B(t, i),\ D(t, i),\ C(t, i),\ M(t, i),\ R(t, i)) : [t_0, t_f] \to \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{n_z \times n} \times \mathcal{S}_{n_z} \times \mathcal{S}_m$, $t \to g_k(t, i) : [t_0, t_f] \to \mathbb{R}^n$, $k = 0, 1$, $t \to r(t) : [t_0, t_f] \to \mathbb{R}^{n_z}$ *are continuous matrix-valued functions;*
(b) $M(t, i) \geq 0, R(t, i) > 0,$ *for all* $t \in [t_0, t_f]$, $G(i) \geq 0, i \in \mathcal{N}$.

Let us consider the function $V : [t_0, t_f] \times \mathbb{R}^n \times \mathcal{N} \to \mathbb{R}$ defined by

$$
V(t, x, i) = x^T X(t, i)x - 2x^T \Psi(t, i) + \mu(t, i)
\tag{11}
$$

where $t \to (X(t, i), \Psi(t, i), \mu(t, i)) : [t_0, t_f] \to \mathcal{S}_n \times \mathbb{R}^n \times \mathbb{R}$, $i \in \mathcal{N}$ are continuous and differentiable functions. Applying the Itô formula, see for example Theorem 1.10.2 from [9] in the case of the function (11) and to the stochastic process $x(t)$ satisfying (8a), we obtain

$$
\mathbb{E}[V(t_f, x(t_f), \eta_{t_f}) - V(t_0, x(t_0), \eta_{t_0})|\eta_{t_0} = i]
\tag{12}
$$

$$
= \mathbb{E}\left[\int_{t_o}^{t_f} \begin{pmatrix} x(t) \\ 1 \\ u(t) \end{pmatrix}^T \begin{pmatrix} W_{11}(t, \eta_t) & W_{12}(t, \eta_t) & W_{13}(t, \eta_t) \\ W_{12}^T(t, \eta_t) & W_{22}(t, \eta_t) & W_{23}(t, \eta_t) \\ W_{13}^T(t, \eta_t) & W_{23}^T(t, \eta_t) & W_{33}(t, \eta_t) \end{pmatrix} \begin{pmatrix} x(t) \\ 1 \\ u(t) \end{pmatrix} dt | \eta_{t_0} = i\right]
$$

for all $i \in \mathcal{N}$, where

$$
\begin{aligned}
W_{11}(t, i) =& \dot{X}(t, i) + A_0^T(t, i)X(t, i) + X(t, i)A_0(t, i) \\
& + A_1^T(t, i)X(t, i)A_1(t, i) + \sum_{j=1}^{N} q_{ij}X(t, j)
\end{aligned}
\tag{13a}
$$

$$
\begin{aligned}
W_{12}(t, i) =& -\dot{\Psi}(t, i) - A_0^T(t, i)\Psi(t, i) + X(t, i)g_0(t, i) \\
& + A_1^T(t, i)X(t, i)g_1(t, i) + \sum_{j=1}^{N} q_{ij}\Psi(t, j)
\end{aligned}
\tag{13b}
$$

$$
W_{13}(t, i) = X(t, i)B(t, i) + A_1^T(t, i)X(t, i)D(t, i)
\tag{13c}
$$

$$
W_{22}(t, i) = \dot{\mu}(t, i) + \sum_{j=1}^{N} q_{ij}\mu(t, j) + g_1^T(t, i)X(t, i)g_1(t, i) - 2g_0^T(t, i)\Psi(t, i)
\tag{13d}
$$

$$
W_{23}(t, i) = -\Psi^T(t, i)B(t, i) + g_1^T(t, i)X(t, i)D(t, i)
\tag{13e}
$$

$$
W_{33}(t, i) = D^T(t, i)X(t, i)D(t, i)
\tag{13f}
$$

Taking the expectation in (12) and adding with (9) we obtain

$$
\begin{aligned}
J(x_0, &u(\cdot)) + \mathbb{E}[V(t_f, x(t_f), \eta_{t_f})] - \mathbb{E}[V(t_0, x_0, \eta_{t_0})] \\
&= \mathbb{E}\left[(x(t_f) - \zeta)^T G(\eta_{t_f})(z(t_f) - \zeta) \right] \\
&+ \mathbb{E}\left[\int_{t_o}^{t_f} \begin{pmatrix} x(t) \\ 1 \\ u(t) \end{pmatrix}^T \begin{pmatrix} \hat{W}_{11}(t, \eta_t) & \hat{W}_{12}(t, \eta_t) & \hat{W}_{13}(t, \eta_t) \\ \hat{W}_{12}^T(t, \eta_t) & \hat{W}_{22}(t, \eta_t) & \hat{W}_{23}(t, \eta_t) \\ \hat{W}_{13}^T(t, \eta_t) & \hat{W}_{23}^T(t, \eta_t) & \hat{W}_{33}(t, \eta_t) \end{pmatrix} \begin{pmatrix} x(t) \\ 1 \\ u(t) \end{pmatrix} dt \right]
\end{aligned}
\tag{14}
$$

where

$$\hat{W}_{11}(t,i) = W_{11}(t,i) + C^T(t,i)M(t,i)C(t,i) \tag{15a}$$

$$\hat{W}_{12}(t,i) = W_{12}(t,i) - C^T(t,i)M(t,i)r(t,i) \tag{15b}$$

$$\hat{W}_{13}(t,i) = W_{13}(t,i) \tag{15c}$$

$$\hat{W}_{22}(t,i) = W_{22}(t,i) + r^T(t,i)M(t,i)r(t,i) \tag{15d}$$

$$\hat{W}_{23}(t,i) = W_{23}(t,i) \tag{15e}$$

$$\hat{W}_{33}(t,i) = R(t,i) + D^T(t,i)X(t,i)D(t,i) \tag{15f}$$

Let

$$\mathbf{X}(\cdot) = (X(\cdot, 1), X(\cdot, 2), \dots, X(\cdot, N)),$$

$$\mathbf{\Psi}(\cdot) = (\Psi(\cdot, 1), \Psi(\cdot, 2), \dots, \Psi(\cdot, N)),$$

$$\boldsymbol{\mu}(\cdot) = (\mu(\cdot, 1), \mu(\cdot, 2), \dots, \mu(\cdot, N))^T$$

be the solutions of the following terminal value problem (TVP)

$$
\begin{aligned}
\dot{X}(t,i) &+ A_0^T(t,i)X(t,i) + X(t,i)A_0(t,i) + A_1^T(t,i)X(t,i)A_1(t,i) \\
&- (X(t,i)B(t,i) + A_1^T(t,i)X(t,i)D(t,i))(R(t,i) + D^T(t,i)X(t,i)D(t,i))^{-1} \\
&\cdot (B^T(t,i)X(t,i) + D^T(t,i)X(t,i)A_1(t,i)) \\
&+ C^T(t,i)M(t,i)C(t,i) + \sum_{j=1}^{N} q_{ij}X(t,j) = 0, \quad t_0 \leq t \leq t_f,
\end{aligned}
\tag{16a}
$$

$$X(t_f, i) = C^T(t_f, i)G(i)C(t_f, i), \quad i \in \mathcal{N}, \tag{16b}$$

$$
\dot{\Psi}(t,i) + (A_0(t,i) + B(t,i)\tilde{F}(t,i))^T \Psi(t,i) + \sum_{j=1}^{N} q_{ij}\Psi(t,j) + G(t, X(t,i), i) = 0
$$

$$\Psi(t_f, i) = C^T(t_f, i)G(i)\zeta \tag{17a}$$

$$
\begin{aligned}
\tilde{F}(t,i) &= -(R(t,i) + D^T(t,i)X(t,i)D(t,i))^{-1}(B^T(t,i)X(t,i) \\
&\quad + D^T(t,i)X(t,i)A_1(t,i))
\end{aligned}
\tag{17b}
$$

$$
\begin{aligned}
G(t, X(t,i), i) &= C^T(t,i)M(t,i)r(t) - X(t,i)g_0(t,i) - (A_1(t,i) \\
&\quad + D(t,i)\tilde{F}(t,i))^T X(t,i)g_1(t,i)
\end{aligned}
\tag{17c}
$$

$X(t,i), i \in \mathcal{N}$ being the components of the solution of the TVP (16),

$$\dot{\mu}(t,i) + \sum_{j=1}^{N} q_{ij}\mu(t,j) + h(t,i) = 0$$

$$\mu(t_f,i) = \zeta^T G(i)\zeta \tag{18a}$$

$$h(t,i) = g_1^T(t,i)X(t,i)g_1(t,i) + r^T(t)M(t,i)r(t) - 2g_0^T(t,i)\Psi(t,i) -$$
$$- (B^T(t,i)\Psi(t,i) - D^T(t,i)X(t,i)g_1(t,i))^T (R(t,i) + D^T(t,i)X(t,i)D(t,i))^{-1}$$
$$\cdot (B^T(t,i)\Psi(t,i) - D^T(t,i)X(t,i)g_1(t,i)) \tag{18b}$$

$X(t,i)$, $\Psi(t,i)$, $i \in \mathcal{N}$ being the components of the solution of the TVPs (16) and (17), respectively. The main properties of the solutions of the TVPs (16)–(18) are summarized in the following lemma.

Lemma 1. *Under the assumption **(H2)** the following hold:*

(i) *the unique solution $\mathbf{X}(\cdot) = (X(\cdot,1),\dots,X(\cdot,N))$ of the TVP (16) is defined on the whole interval $[t_0,t_f]$. Moreover, $X(t,i) = X^T(t,i) \geq 0$, for all $(t,i) \in [t_0,t_f] \times \mathcal{N}$;*

(ii) *the TVPs (17) and (18) have unique solutions $t \to \mathbf{\Psi}(t) = (\Psi(t,1),\Psi(t,2),\dots,\Psi(t,N))$: $[t_0,t_f] \to \mathbb{R}^n \times \mathbb{R}^n \times \dots \times \mathbb{R}^n$ and $t \to \boldsymbol{\mu}(t) = (\mu(t,1),\mu(t,2),\dots,\mu(t,N))^T : [t_0,t_f] \to \mathbb{R}^N$.*

Proof.

(i) Follows immediately applying Corollary 5.2.3 from [9] applied in the case of TVP (16).

(ii) The TVP (17) is associated with a linear nonhomogeneous differential equation with time-varying coefficients. Hence its solution is defined on the whole interval of a definition of its coefficients. According to (i) it follows that the coefficients of the differential Equation (17a) are defined on the whole interval $[t_0,t_f]$. Hence, its solution is also defined on the whole interval $[t_0,t_f]$. The conclusion regarding the definition of the solution of TVP (18) on the interval $[t_0,t_f]$ is obtained in the same way. □

Further, we consider the case when (11) is defined using the solutions of the TVPs (16)–(18). In this case, (13), (15), (16)–(18) allow us to reduce (14) to

$$J(x_0;u(\cdot)) = \mathbb{E}[V(t_0,x_0,\eta_{t_o})] + \mathbb{E}\left[\int_{t_0}^{t_f} (u(t) - \tilde{F}(t,\eta_t)x(t) - \tilde{\varphi}(t,\eta_t))^T\right.$$

$$\left. \cdot \hat{W}_{33}(t,\eta_t)(u(t) - \tilde{F}(t,\eta_t)x(t) - \tilde{\varphi}(t,\eta_t))dt\right] \tag{19}$$

for all $u(\cdot)$ of type (10), where $\hat{W}_{33}(t,i)$ are computed as in (15f), $\tilde{F}(t,i)$ is computed as in (17b) based on the solution of TVP (16), whereas

$$\tilde{\varphi}(t,i) := -(R(t,i) + D^T(t,i)X(t,i)D(t,i))^{-1}(D^T(t,i)X(t,i)g_1(t,i) - B^T(t,i)\Psi(t,i)) \tag{20}$$

for all $(t,i) \in [t_0,t_f] \times \mathcal{N}$.

Now we are in position to state and prove the main result of this subsection.

Theorem 1. *Assume that the assumption H2 is fulfilled. We consider the control law*

$$\tilde{u}(t) = \tilde{F}(t,\eta_t)\tilde{x}(t) + \tilde{\varphi}(t,\eta_t) \tag{21}$$

where $\tilde{F}(t,i)$ and $\tilde{\varphi}(t,i)$ are computed via (17b) and (20), respectively, based on the solution $\mathbf{X}(\cdot)$ and $\mathbf{\Psi}(\cdot)$ of TVPs (16) and (17) and $\tilde{x}(t)$ is the solution of the closed-loop system obtained when

coupling the control (21) to (8a). Under these conditions, the control (21) satisfies the following minimality condition

$$J(x_0; \tilde{u}(\cdot)) = \min_{u(\cdot) \in \mathcal{U}} J(x_0; u(\cdot)).$$

The minimal value of the cost function (9) in the class of the controls $\mathcal{U}$ of type (10) is given by

$$J(x_0; \tilde{u}(\cdot)) = x_0^T \mathbb{E}[X(t_0, \eta_{t_0})] x_0 - 2x_0^T \mathbb{E}[\Psi(t_0, \eta_{t_0})] + \mathbb{E}[\mu(t_0, \eta_{t_0})].$$

Proof. From (15f) we deduce via Lemma 1 (i) that under the assumption H2(b) we have $\hat{W}_{33}(t, \eta_t) > 0$, for all $t \in [t_0, t_f]$. The conclusion is obtained immediately from (19). $\square$

3.2. The Case of Two Decision-Makers

In this subsection, we shall use the result derived in Theorem 1 to obtain the state space representation of an equilibrium strategy of type (3) which satisfies (5). Let $k, l \in \{1, 2\}$ be fixed such that $l \neq k$. Let

$$\tilde{u}_j(t) = \tilde{F}_j(t, \eta_t) x(t) + \tilde{\varphi}_j(t, \eta_t), \quad j = 1, 2$$

be a candidate for a Nash equilibrium strategy. Taking $j = l$ we rewrite (1) and (4) as

$$
\begin{aligned}
dx(t) =& [(A_0(t, \eta_t) + B_l(t, \eta_t) \tilde{F}_l(t, \eta_t)) x(t) + B_k(t, \eta_t) u_k(t) + B_l(t, \eta_t) \tilde{\varphi}(t, \eta_t)] dt \\
& + [(A_1(t, \eta_t) + D_l(t, \eta_t) \tilde{F}_l(t, \eta_t)) x(t) \\
& + D_k(t, \eta_t) u_k(t) + D_l(t, \eta_t) \tilde{\varphi}_l(t, \eta_t)] dw(t) \\
x(t_0) =& x_0
\end{aligned}
\tag{22}
$$

$$
\begin{aligned}
J_{kl}(x_0; u_k(\cdot)) =& \mathbb{E}\left[(z_k(t_f) - \zeta_k)^T G_k(\eta_{t_f})(z_k(t_f) - \zeta_k) \right] \\
& + \mathbb{E}\left[\int_{t_0}^{t_f} \{ (z_k(t) - r_k(t))^T M_k(t, \eta_t)(z_k(t) - r_k(t)) \right. \\
& + (\tilde{F}_l(t, \eta_t) x(t) + \tilde{\varphi}_l(t, \eta_t))^T R_{kl}(t, \eta_t)(\tilde{F}_l(t, \eta_t) x(t) + \tilde{\varphi}(t, \eta_t)) \\
& \left. + u_k^T(t) R_{kk}(t, \eta_t) u_k(t) \} dt \right]
\end{aligned}
\tag{23}
$$

From Definition 1 it follows that $(\tilde{u}_k(\cdot), \tilde{u}_l(\cdot))$ is a Nash equilibrium strategy for the dynamic game described by (1), (3), and (4) if $\tilde{u}_k(\cdot)$ minimizes the cost (23) along with the trajectories of the system (22), determined by the controls of type

$$u_k(t) = F_k(t, \eta_t) x(t) + \varphi_k(t, \eta_t)$$

$$\tag{24}$$

with $F_k(\cdot, i)$, $\varphi_k(\cdot, i)$ arbitrary continuous functions.

In order to obtain the explicit formula of $\tilde{u}_k(\cdot)$ with these properties, we apply Theorem 1 specialized to the case of the optimal tracking problem described by the system (22) and the performance criterion (23). To this end, we shall rewrite TVPs (16)–(18) with the updates

$$A_0(t, i) \leftarrow A_0(t, i) + B_l(t, i) \tilde{F}_l(t, i), \quad A_1(t, i) \leftarrow A_1(t, i) + D_l(t, i) \tilde{F}_l(t, i),$$

$$B(t, i) \leftarrow B_k(t, i), \quad D(t, i) \leftarrow D_k(t, i), \quad g_0(t, i) \leftarrow B_l(t, i) \tilde{\varphi}_l(t, i),$$

$$g_1(t, i) \leftarrow D_l(t, i) \tilde{\varphi}_l(t, i), \quad C(t, i) \leftarrow \begin{pmatrix} C_k(t, i) \\ \tilde{F}_l(t, i) \end{pmatrix},$$

$$M(t, i) \leftarrow \begin{pmatrix} M_k(t, i) & 0 \\ 0 & R_{kl}(t, i) \end{pmatrix}, \quad R(t, i) \leftarrow R_{kk}(t, i),$$

$$G(i) \leftarrow \begin{pmatrix} G_k(i) & 0 \\ 0 & 0 \end{pmatrix}, \quad r(t) = \begin{pmatrix} r_k(t) \\ \varphi_l(t,i) \end{pmatrix}.$$

Thus, TVP (16) becomes

$$\begin{aligned}
&\dot{X}_k(t,i) + (A_0(t,i) + B_l(t,i)\tilde{F}_l(t,i))^T X_k(t,i) + X_k(t,i)(A_0(t,i) \\
&+ B_l(t,i)\tilde{F}_l(t,i)) + (A_1(t,i) + D_l(t,i)\tilde{F}_l(t,i))^T X_k(t,i)(A_1(t,i) + D_l(t,i)\tilde{F}_l(t,i)) \\
&+ \sum_{j=1}^{N} q_{ij} X_k(t,j) - [X_k(t,i)B_k(t,i) + (A_1(t,i) + D_l(t,i)\tilde{F}_l(t,i))^T X_k(t,i)D_k(t,i)] \\
&\cdot (R_{kk}(t,i) + D_k^T(t,i)X_k(t,i)D_k(t,i))^{-1} \\
&\cdot [B_k^T(t,i)X_k(t,i) + D_k^T(t,i)X_k(t,i)(A_1(t,i) + D_l(t,i)\tilde{F}_l(t,i))] \\
&+ C_k^T(t,i)M_k(t,i)C_k(t,i) + \tilde{F}_l^T(t,i)R_{kl}(t,i)\tilde{F}_l(t,i) = 0
\end{aligned} \tag{25a}$$

$$X_k(t_f,i) = C_k^T(t_f,i)G_k(i)C_k(t_f,i) \tag{25b}$$

$i \in \mathcal{N}, k = 1, 2, l = 3 - k$. The analogous of the feedback gains $\tilde{F}(t,i)$ associated with the solution of TVP (16), via (17b) becomes, in the case of TVP (25)

$$\begin{aligned}
\tilde{F}_k(t,i) = &- (R_{kk}(t,i) + D_k^T(t,i)X_k(t,i)D_k(t,i))^{-1} \\
&\cdot [B_k^T(t,i)X_k(t,i) + D_k^T(t,i)X_k(t,i)(A_1(t,i) + D_{3-k}(t,i)\tilde{F}_{3-k}(t,i))].
\end{aligned} \tag{26}$$

In the case of the tracking problem described by (22) and (24), the TVPs (17) and (18) take the form

$$\begin{aligned}
&\Psi_k(t,i) + (A_0(t,i) + B_1(t,i)\tilde{F}_1(t,i) + B_2(t,i)\tilde{F}_2(t,i))^T \Psi_k(t,i) \\
&+ \sum_{j=1}^{N} q_{ij} \Psi_k(t,j) + \mathcal{G}_k(t, X_k(t,i), i) = 0
\end{aligned} \tag{27a}$$

$$\Psi_k(t_f,i) = C_k^T(t_f,i)G_k(i)\zeta_k \tag{27b}$$

where

$$\begin{aligned}
\mathcal{G}(t, X_k(t,i), i) = &\, C_k^T(t,i)M_k(t,i)r_k(t) - [X_k(t,i)B_{3-k}(t,i) + \tilde{F}_{3-k}^T(t,i)R_{k,3-k}(t,i) \\
&+ (A_1(t,i) + D_1(t,i)\tilde{F}_1(t,i) + D_2(t,i)\tilde{F}_2(t,i))^T X_k(t,i)D_{3-k}(t,i)]\tilde{\varphi}_{3-k}(t,i)
\end{aligned} \tag{27c}$$

$$\dot{\mu}_k(t,i) + \sum_{j=1}^{N} q_{ij}\mu_k(t,j) + h_k(t,i) = 0 \tag{28a}$$

$$\mu_k(t_f,i) = \zeta_k^T G_k(i)\zeta_k \tag{28b}$$

$$\begin{aligned}
h_k(t,i) = &\, \tilde{\varphi}_{3-k}^T(t,i)D_{3-k}^T(t,i)X_k(t,i)D_{3-k}(t,i)\tilde{\varphi}_{3-k}(t,i) + r_k^T(t,i)M_k(t,i)r_k(t,i) \\
&+ \tilde{\varphi}_{3-k}^T(t,i)R_{k,3-k}(t,i)\tilde{\varphi}_{3-k}(t,i) - 2\tilde{\Psi}_k^T(t,i)B_{3-k}(t,i)\tilde{\varphi}_{3-k}(t,i) \\
&- (\tilde{\varphi}_{3-k}^T(t,i)D_{3-k}^T(t,i)X_k(t,i)D_k(t,i) - \tilde{\Psi}_k^T(t,i)B_k(t,i)
\end{aligned} \tag{28c}$$

$$(R_{kk}(t,i) + D_k^T(t,i)X_k(t,i)D_k(t,i))^{-1} \cdot (D_k^T(t,i)X_k(t,i)D_{3-k}(t,i)\tilde{\varphi}_{3-k}(t,i) - B_k^T(t,i)\tilde{\Psi}_k(t,i))$$

for all $i \in \mathcal{N}, k = 1, 2$. In this context (20) becomes

$$\begin{aligned}
\tilde{\varphi}_k(t,i) = &- (R_{kk}(t,i) + D_k^T(t,i)X_k(t,i)D_k(t,i))^{-1} \times (D_k^T(t,i)X_k(t,i)D_{3-k}(t,i)\tilde{\varphi}_{3-k}(t,i) \\
&- B_k^T(t,i)\tilde{\Psi}_k(t,i))
\end{aligned} \tag{29}$$

for all $i \in \mathcal{N}, k = 1, 2$.

Remark 2. *Although the TVP (25) is defined by a Riccati differential equation of type (16), we cannot be sure that the solution of this problem is defined on the whole interval* $[t_0, t_f]$, *because the domain of definition of its coefficients depends upon the domain of definition of the gain matrices* $F_l(\cdot, i)$.

In the following, we shall regard (25) and (26) as a TVP associated with a hybrid system of nonlinear differential equations and nonlinear algebraic equations

$$
\dot{X}_1(t,i) + (A_0(t,i) + B_2(t,i)F_2(t,i))^T X_1(t,i) + X_1(t,i)(A_0(t,i) + B_2(t,i)F_2(t,i)) + (A_1(t,i)
$$
$$
+ D_2(t,i)F_2(t,i))^T X_1(t,i)(A_1(t,i) + D_2(t,i)F_2(t,i))
$$
$$
+ \sum_{j=1}^N q_{ij} X_1(t,j) - [X_1(t,i)B_1(t,i) + (A_1(t,i) + D_2(t,i)F_2(t,i))^T X_1(t,i)D_1(t,i)]
$$
$$
\cdot (R_{11}(t,i) + D_1^T(t,i)X_1(t,i)D_1(t,i))^{-1}
$$
$$
\cdot [B_1^T(t,i)X_1(t,i) + D_1^T(t,i)X_1(t,i)(A_1(t,i) + D_2(t,i)F_2(t,i))]
$$
$$
+ F_2^T(t,i)R_{12}(t,i)F_2(t,i) + C_1^T(t,i)M_1(t,i)C_1(t,i) = 0 \tag{30a}
$$

$$
\dot{X}_2(t,i) + (A_0(t,i) + B_1(t,i)F_1(t,i))^T X_2(t,i) + X_2(t,i)(A_0(t,i)
$$
$$
+ B_1(t,i)F_1(t,i)) + (A_1(t,i) + D_1(t,i)F_1(t,i))^T X_2(t,i)(A_1(t,i) + D_1(t,i)F_1(t,i))
$$
$$
+ \sum_{j=1}^N q_{ij} X_2(t,j) - [X_2(t,i)B_2(t,i) + (A_1(t,i)
$$
$$
+ D_1(t,i)F_1(t,i))^T X_2(t,i)D_2(t,i)] (R_{22}(t,i) + D_2^T(t,i)X_2(t,i)D_2(t,i))^{-1}
$$
$$
\cdot [B_2^T(t,i)X_2(t,i) + D_2^T(t,i)X_2(t,i)(A_1(t,i) + D_1(t,i)F_1(t,i))]
$$
$$
+ F_1^T(t,i)R_{21}(t,i)F_1(t,i) + C_2^T(t,i)M_2(t,i)C_2(t,i) = 0 \tag{30b}
$$

$$
(R_{11}(t,i) + D_1^T(t,i)X_1(t,i)D_1(t,i))F_1(t,i) + D_1^T(t,i)X_1(t,i)D_2(t,i)F_2(t,i)
$$
$$
+ B_1^T(t,i)X_1(t,i) + D_1^T(t,i)X_1(t,i)A_1(t,i) = 0 \tag{30c}
$$
$$
D_2^T(t,i)X_2(t,i)D_1(t,i)F_1(t,i) + (R_{22}(t,i) + D_2^T(t,i)X_2(t,i)D_2(t,i))F_2(t,i)
$$
$$
+ B_2^T(t,i)X_2(t,i) + D_2^T(t,i)X_2(t,i)A_1(t,i) = 0 \tag{30d}
$$
$$
X_k(t_f,i) = C_k^T(t_f,i)G_k(i)C_k(t_f,i) \tag{30e}
$$

$i \in \mathcal{N}, k = 1, 2$. At the same time, (27) and (29) can be viewed as a TVP associated with a hybrid system formed by two backward linear differential equations and two algebraic linear equations, as

$$
\dot{\Psi}_1(t,i) + (A_0(t,i) + B_1(t,i)\tilde{F}_1(t,i) + B_2(t,i)\tilde{F}_2(t,i))^T \Psi_1(t,i)
$$
$$
+ \sum_{j=1}^N q_{ij} \Psi_1(t,j) - \mathbb{G}_{12}(t,i)\varphi_2(t,i) + C_1^T(t,i)M_1(t,i)r_1(t) = 0 \tag{31a}
$$

$$
\dot{\Psi}_2(t,i) + (A_0(t,i) + B_1(t,i)\tilde{F}_1(t,i) + B_2(t,i)\tilde{F}_2(t,i))^T \Psi_2(t,i)
$$
$$
+ \sum_{j=1}^N q_{ij} \Psi_2(t,j) - \mathbb{G}_{21}(t,i)\varphi_1(t,i) + C_2^T(t,i)M_2(t,i)r_2(t) = 0 \tag{31b}
$$

$$
(R_{11}(t,i) + D_1^T(t,i)\tilde{X}_1(t,i)D_1(t,i))\varphi_1(t,i) + D_1^T(t,i)\tilde{X}_1(t,i)D_2(t,i)\varphi_2(t,i)
$$
$$
+ B_1^T(t,i)\Psi_1(t,i) = 0 \tag{31c}
$$
$$
D_2^T(t,i)\tilde{X}_2(t,i)D_1(t,i)\varphi_1(t,i) + (R_{22}(t,i) + D_2^T(t,i)\tilde{X}_2(t,i)D_2(t,i))\varphi_2(t,i)
$$
$$
+ B_2^T(t,i)\Psi_2(t,i) = 0 \tag{31d}
$$
$$
\Psi_k(t_f,i) = C_k^T(t_f,i)G_k(i)\zeta_k, \tag{31e}
$$

$i \in \mathcal{N}, k = 1, 2$, where we denoted

$$
\begin{aligned}
\mathbb{G}_{12}(t,i) =& \tilde{X}_1(t,i)B_2(t,i) + \tilde{F}_2^T(t,i)R_{12}(t,i) + (A_1(t,i) + D_1(t,i)\tilde{F}_1(t,i) \\
& + D_2(t,i)\tilde{F}_2(t,i))^T \tilde{X}_1(t,i)D_2(t,i)
\end{aligned}
\tag{32a}
$$

$$
\begin{aligned}
\mathbb{G}_{21}(t,i) =& \tilde{X}_2(t,i)B_1(t,i) + \tilde{F}_1^T(t,i)R_{21}(t,i) + (A_1(t,i) + D_1(t,i)\tilde{F}_1(t,i) \\
& + D_2(t,i)\tilde{F}_2(t,i))^T \tilde{X}_2(t,i)D_1(t,i).
\end{aligned}
\tag{32b}
$$

In (31) and (32), $(\tilde{X}_1(t,i), \tilde{X}_2(t,i), \tilde{F}_1(t,i), \tilde{F}_2(t,i))$, $i \in \mathcal{N}$, is a solution of the TVP (30). Applying Theorem 1 in the case of the optimal tracking problems described by system (22) and the performance criterion (23) for $k = 1$ and $k = 2$, we obtain

Theorem 2. *Assume:*

(a) *the assumption (H1) is fulfilled;*

(b) *the solutions $(\tilde{X}_1(\cdot,i), \tilde{X}_2(\cdot,i), \tilde{F}_1(\cdot,i), \tilde{F}_2(\cdot,i))$, $i \in \mathcal{N}$, and $(\check{\Psi}_1(\cdot,i), \check{\Psi}_2(\cdot,i), \tilde{\varphi}_1(\cdot,i),$ $\tilde{\varphi}_2(\cdot,i))$, $i \in \mathcal{N}$ of the TVPs (30) and (31), respectively, are defined on the whole interval $[t_0, t_f]$.*

We set

$$
\tilde{u}_j(t) = \tilde{F}_j(t, \eta_t)\tilde{x}(t) + \tilde{\varphi}_j(t, \eta_t), \quad j = 1, 2
\tag{33}
$$

$\tilde{x}(\cdot)$ being the solution of the IVP obtained replacing (33) in (1). Under these conditions, $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$ is an equilibrium strategy for the differential game described by the controlled system (1), the performance criteria (4), and the family of the admissible strategies of type (3). The optimal values of the performance criteria are given by

$$
J_k(x_0; \tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) = x_0^T \mathbb{E}[\tilde{X}_k(t_0, \eta_{t_0})]x_0 - 2x_0^T \mathbb{E}[\check{\Psi}_k(t_o, \eta_{t_0})] + \mathbb{E}[\tilde{\mu}_k(t_0, \eta_{t_0})],
$$

$k = 1, 2$.

4. Several Special Cases

4.1. The Case without Control-Dependent Noise of the Diffusion Part of the Controlled System

We assume that the controlled system (1) is in the special form

$$
\begin{aligned}
dx(t) &= (A_0(t, \eta_t)x(t) + B_1(t, \eta_t)u_1(t) + B_2(t, \eta_t)u_2(t))dt + A_1(t, \eta_t)x(t)dw(t) \\
x(t_0) &= x_0.
\end{aligned}
\tag{34}
$$

In this case the TVPs (30) and (31), respectively, reduce to

$$
\begin{aligned}
& \dot{X}_1(t,i) + (A_0(t,i) - S_2(t,i)X_2(t,i))^T X_1(t,i) + X_1(t,i)(A_0(t,i) \\
& - S_2(t,i)X_2(t,i)) + A_1^T(t,i)X_1(t,i)A_1(t,i) - X_1(t,i)S_1(t,i)X_1(t,i) \\
& + X_2(t,i)S_{12}(t,i)X_2(t,i) + \sum_{j=1}^{N} q_{ij}X_1(t,j) + C_1^T(t,i)M_1(t,i)C_1(t,i) = 0
\end{aligned}
\tag{35a}
$$

$$
\begin{aligned}
& \dot{X}_2(t,i) + (A_0(t,i) - S_1(t,i)X_1(t,i))^T X_2(t,i) + X_2(t,i)(A_0(t,i) \\
& - S_1(t,i)X_1(t,i)) + A_1^T(t,i)X_2(t,i)A_1(t,i) - X_2(t,i)S_2(t,i)X_2(t,i) \\
& + X_1(t,i)S_{21}(t,i)X_1(t,i) + \sum_{j=1}^{N} q_{ij}X_2(t,j) + C_2^T(t,i)M_2(t,i)C_2(t,i) = 0
\end{aligned}
\tag{35b}
$$

$$
X_k(t_f, i) = C_k^T(t_f, i)G_k(i)C_k(t_f, i)
\tag{35c}
$$

$$
F_k(t, i) = -R_{kk}^{-1}(t, i)B_k^T(t, i)X_k(t, i)
\tag{35d}
$$

$i \in \mathcal{N}, k = 1, 2,$

$$\dot{\Psi}_1(t,i) + (A_0(t,i) - S_1(t,i)X_1(t,i) - S_2(t,i)X_2(t,i))^T \Psi_1(t,i)$$

$$+ \sum_{j=1}^{N} q_{ij} \Psi_1(t,j) + (X_1(t,i)S_2(t,i) - X_2(t,i)S_{12}(t,i))\Psi_2(t,i)$$

$$+ C_1^T(t,i)M_1(t,i)r_1(t) = 0 \tag{36a}$$

$$\dot{\Psi}_2(t,i) + (X_2(t,i)S_1(t,i) - X_1(t,i)S_{21}(t,i))\Psi_1(t,i) + (A_0(t,i)$$

$$- S_1(t,i)X_1(t,i) - S_2(t,i)X_2(t,i))^T \Psi_2(t,i) + \sum_{j=1}^{N} q_{ij}\Psi_2(t,j)$$

$$+ C_2^T(t,i)M_2(t,i)r_2(t) = 0 \tag{36b}$$

$$\Psi_k(t_f,i) = C_k^T(t_f,i)G_k(i)\zeta_k \tag{36c}$$

$$\varphi_k(t,i) = -R_{kk}^{-1}(t,i)B_k^T(t,i)\Psi_k(t,i) \tag{36d}$$

$i \in \mathcal{N}, k = 1, 2.$ In (35) and (36) we have denoted

$$S_j(t,i) \triangleq B_j(t,i)R_{jj}^{-1}(t,i)B_j^T(t,i)$$

$$S_{jk}(t,i) \triangleq B_k(t,i)R_{kk}^{-1}(t,i)R_{jk}(t,i)R_{kk}^{-1}(t,i)B_k^T(t,i)$$

$j = 1, 2, k = 3 - j.$ The TVP (28) becomes

$$\dot{\mu}_k(t,i) + \sum_{j=1}^{N} q_{ij}\mu_k(t,j) + \tilde{h}_k(t,i) = 0 \tag{37a}$$

$$\mu_k(t_f,i) = \zeta_k^T G_k(i)\zeta_k \tag{37b}$$

$$\tilde{h}_1(t,i) = r_1^T(t)M_1(t,i)r_1(t) + \check{\Psi}_2^T(t,i)S_{12}(t,i)\check{\Psi}_2(t,i)$$

$$- 2\check{\Psi}_1^T(t,i)S_2(t,i)\check{\Psi}_2(t,i) - \check{\Psi}_1^T(t,i)S_1(t,i)\check{\Psi}_1(t,i) \tag{37c}$$

$$\tilde{h}_2(t,i) = r_2^T(t)M_2(t,i)r_2(t) + \check{\Psi}_1^T(t,i)S_{21}(t,i)\check{\Psi}_1(t,i)$$

$$- 2\check{\Psi}_2^T(t,i)S_1(t,i)\check{\Psi}_1(t,i) - \check{\Psi}_2^T(t,i)S_2(t,i)\check{\Psi}_2(t,i) \tag{37d}$$

In (37), $(\check{\Psi}_1(\cdot,i), \check{\Psi}_2(\cdot,i)), i \in \mathcal{N}$ is the solution of the TVP (36). Applying the result derived in Theorem 2, we obtain

Corollary 2. *Assume:*

(a) *the assumption **(H1)** is fulfilled;*

(b) *the solution $(\check{\tilde{X}}_1(\cdot,i), \check{\tilde{X}}_2(\cdot,i)), i \in \mathcal{N},$ of the TVP (35a)–(35c) is defined on the whole interval $[t_0, t_f]$.*

We set

$$\check{\tilde{u}}_j(t) = -R_{jj}^{-1}(t,\eta_t)B_j^T(t,\eta_t)(\check{\tilde{X}}_j(t,\eta_t)\check{\tilde{x}}(t) + \check{\tilde{\Psi}}_j(t,\eta_t)), \tag{38}$$

$\check{\tilde{x}}(\cdot)$ *being the solution of the IVP obtained substituting (38) in (34). Under these conditions, $(\check{\tilde{u}}_1(\cdot), \check{\tilde{u}}_2(\cdot))$ is an equilibrium strategy for the differential game described by the controlled system (34), the performance criterion (4), and the admissible strategies of type (3). The optimal values of the performance criterion (4) are given by*

$$J_k(x_0; \check{\tilde{u}}_1(\cdot), \check{\tilde{u}}_2(\cdot)) = x_0^T \mathbb{E}[\check{\tilde{X}}_k(t_0, \eta_{t_0})]x_0 - 2x_0^T \mathbb{E}[\check{\tilde{\Psi}}_k(t_o, \eta_{t_0})] + \mathbb{E}[\check{\tilde{\mu}}_k(t_0, \eta_{t_0})],$$

$k = 1, 2, \check{\tilde{\Psi}}_k(\cdot,i), \check{\tilde{\mu}}_k(\cdot,i)$ *being the solutions of the TVPs (36) and (37), respectively.*

4.2. The Case when the Performance Criterion (4) Is Replaced by Performance Criterion of Type (7)

In this case, the aim of the decision-makers is to minimize the mean square of the deviations of the final value of the output $z_k(t_f)$ from the target ζ_k, $k = 1, 2$. If so, the equilibrium strategy is obtained solving the TVPs (30) and (31), respectively, when the controlled system is of type (1) or the TVPs (35) and (36), respectively, when the controlled system is of type (34). In both cases $M_k(t, i) = 0$, $(t, i) \in [t_0, t_f] \times \mathcal{N}$.

5. A Numerical Experiment

For the numerical experiment we considered the time invariant case of the system (34) with performance criterion (4) without Markovian jumping. We rewrite equations (35a) and (35b) of the form:

$$\dot{X}_1(t) = -\mathcal{R}_1(t, X_1(t), X_2(t)) = -(C_1^T M_1 C_1$$
$$+ (A_0 - S_2 X_2(t))^T X_1(t) + X_1(t)(A_0 - S_2 X_2(t))$$
$$+ A_1^T X_1(t) A_1 - X_1(t) S_1 X_1(t) + X_2(t) S_{12} X_2(t))$$
$$\dot{X}_2(t) = -\mathcal{R}_2(t, X_1(t), X_2(t)) = -(C_2^T M_2 C_2$$
$$+ (A_0 - S_1 X_1(t))^T X_2(t) + X_2(t)(A_0 - S_1 X_1(t))$$
$$+ A_1^T X_2(t) A_1 - X_2(t) S_2 X_2(t) + X_1(t) S_{21} X_1(t))$$
$$X_k(t_f) = C_k^T G_k C_k, \; k = 1, 2$$
$$F_k(t) = -R_{kk}^{-1} B_k^T X_k(t), \; k = 1, 2.$$

In this case, we have $M_k(t, i) = M_k \geq 0$, $R_{kk}(t, i) = R_{kk} > 0$, $G_k(i) = G_k \geq 0$, $R_{k\ell}(t, i) = R_{k\ell} \geq 0$, $r_k(t) = 0$, $0 \leq t \leq t_f$, $k, \ell = 1, 2$, $\ell \neq k$.

The matrix coefficients of the controlled system (1) are :

$$A_0 = \begin{pmatrix} -0.4 & 0 & -0.3 & 0.7 \\ -0.3 & -0.2 & 0 & -0.1 \\ 0.8 & -0.25 & -0.25 & 0 \\ -0.1 & -0.2 & 0 & -0.5 \end{pmatrix}, \; A_0 \in \mathbb{R}^{4 \times 4}$$

$$A_1 = \begin{pmatrix} 0 & 0.2 & 0 & -0.1 \\ 0 & 0 & -0.3 & 0.15 \\ -0.145 & 0.06 & -0.2 & 0 \\ 0 & -0.3 & 0 & 0.5 \end{pmatrix}, \; A_1 \in \mathbb{R}^{4 \times 4}$$

$$B_1 = \begin{pmatrix} 0.5 & 0 \\ 1 & 2.5 \\ -0.5 & 2 \\ 1 & 3 \end{pmatrix} \quad B_2 = \begin{pmatrix} 0 & 0.5 \\ -0.75 & -1.5 \\ 1 & -2 \\ 1 & 0 \end{pmatrix}, \; B_1, B_2 \in \mathbb{R}^{4 \times 2}$$

$$C_1 = \begin{pmatrix} 1.0 & -0.25 & 0.75 & -0.5 \\ 0.25 & 0.75 & -0.5 & -0.25 \end{pmatrix} \quad C_2 = \begin{pmatrix} 1.2 & -0.5 & 0.75 & -0.045 \\ 0.5 & 0.15 & 0.35 & 0.55 \end{pmatrix}, \; C_1, C_2 \in \mathbb{R}^{4 \times 4}$$

$D_1 = 0$, $D_2 = 0$.

The weight matrices for the performance criteria are of the form:

$$G_1 = \begin{pmatrix} 0.2778 & -0.094 \\ -0.094 & 0.166 \end{pmatrix} \quad G_2 = \begin{pmatrix} 0.2778 & -0.133 \\ -0.133 & 0.31 \end{pmatrix}, \; G_1, G_2 \in \mathbb{R}^{2 \times 2}$$

$$M_1 = \begin{pmatrix} 3.2 & 0.5 \\ 0.5 & 2.5 \end{pmatrix} \quad M_2 = \begin{pmatrix} 0.75 & 0.05 \\ 0.05 & 0.95 \end{pmatrix}, \; M_1, M_2 \in \mathbb{R}^{2 \times 2}$$

$$R_{11} = \begin{pmatrix} 0.8 & 0.3 \\ 0.3 & 1.5 \end{pmatrix} \quad R_{22} = \begin{pmatrix} 0.95 & 0.65 \\ 0.65 & 1.25 \end{pmatrix}$$

$$R_{12} = \begin{pmatrix} 0.6 & -0.3 \\ -0.3 & 1.2 \end{pmatrix} \quad R_{21} = \begin{pmatrix} 0.8 & -0.2 \\ -0.2 & 1.0 \end{pmatrix}, \quad R_{ij} \in \mathbb{R}^{2 \times 2} \, i, j = 1, 2$$

Moreover, $r_k(t) = 0, k = 1, 2$ and the targets are $\zeta_1 = [0.3; 0.8]$, $\zeta_2 = [0.6; 0.9]$, and $t \in [0, 1]$. The initial point x_0 is chosen to be $x_0 = [0.4; 0.01; 0.2; 0.25]; \in \mathbb{R}^{4 \times 1}$.

To compute $X_1(t)$, $X_2(t)$ we can use the Euler discretization method as $(k = 1, 2)$:

$$\tilde{X}_k(jh) = \tilde{X}_k((j+1)h) + h\mathcal{R}_k((j+1)h, X_1((j+1)h), X_2((j+1)h))$$

with $\tilde{X}_k(Nh) = C_k^T G_k, C_k$, $j = N-1, N-2, \ldots, 1, 0$, $N = [t_f / h]$, $k = 1, 2$.

We consider the following algorithm to compute the behavior of the controlled signals $\tilde{z}_k, k = 1, 2$.

Step 1. The aim of this step is to compute the gains matrices $\tilde{F}_k(jh)$ and $\tilde{\varphi}_k(jh), j = 0, 1, \ldots, N$, $k = 1, 2$.

Step 2. The aim is to compute $\mathbb{E}[|\tilde{z}_k(jh)|^2]$ for $j = 0, 1, \ldots .N$; $k = 1, 2$. We have:

$$\mathbb{E}[|\tilde{z}_k(jh)|^2] = Tr[C_k \Sigma(jh) C_k^T]$$

Consider two cases:

(A). The base variant using the above matrix coefficients. We have executed Step 1 and Step 2. The computed values $\mathbb{E}[|\tilde{z}_k(jh)|^2], k = 1, 2$ for the signals $\tilde{z}_k, k = 1, 2$ of the players are given on Figures 1 and 2 for the first player and the second player, respectively. Moreover, we have obtained the following values of $E[|z_k(1) - \zeta_k|^2]$ for the players $k = 1, 2$, i.e.,

$$E[|z_1(1) - \zeta_1|^2] = 0.5033, \quad E[|z_2(1) - \zeta_2|^2] = 1.0173. \tag{39}$$

(B). We want to compute the output of the closed-loop system $z_k(\cdot)$ by using a control law (other than the optimal one) $\tilde{u}_k(jh) = F_k(jh) x(jh) + \varphi_k(jh)$ with $k = 1, 2, j = 0, 1, \ldots, N-1$. For this, we take $F_k(jh)$ and $\varphi_k(jh)$ different from the optimal cases. We use the same matrix coefficients. After Step 1, we obtain the optimal values of $\tilde{F}_k(jh)$ and $\tilde{\varphi}_k(jh) \, k = 1, 2$. Then, we compute the different values $\tilde{F}_k^k(jh), \tilde{\varphi}_k^k(jh), k = 1, 2$ as follows $(j = 0, 1, \ldots, N-1)$:

$$\tilde{F}_1^1(jh) = (\tilde{F}_1(jh) + \tilde{F}_2(jh))/2, \quad \tilde{F}_2^2(jh) = (\tilde{F}_1(jh) - \tilde{F}_2(jh))/2,$$

$$\tilde{\varphi}_1^1(jh) = (\tilde{\varphi}_1(jh) - \tilde{\varphi}_2(jh))/2, \quad \tilde{\varphi}_2^2(jh) = (\tilde{\varphi}_1(jh) + \tilde{\varphi}_2(jh))/2.$$

The computations continue with Step 2 with $\tilde{F}_k^k(jh), \tilde{\varphi}_k^k(jh), k = 1, 2$. The computed values of $E[|z_k(1) - \zeta_k|^2], k = 1, 2$ are

$$E[|z_1(1) - \zeta_1|^2] = 8.5203, \quad E[|z_2(1) - \zeta_2|^2] = 4.5174. \tag{40}$$

One sees from (39) and (40) that the values of the obtained deviation from the target provided by the optimal control are better than the ones provided by another control.
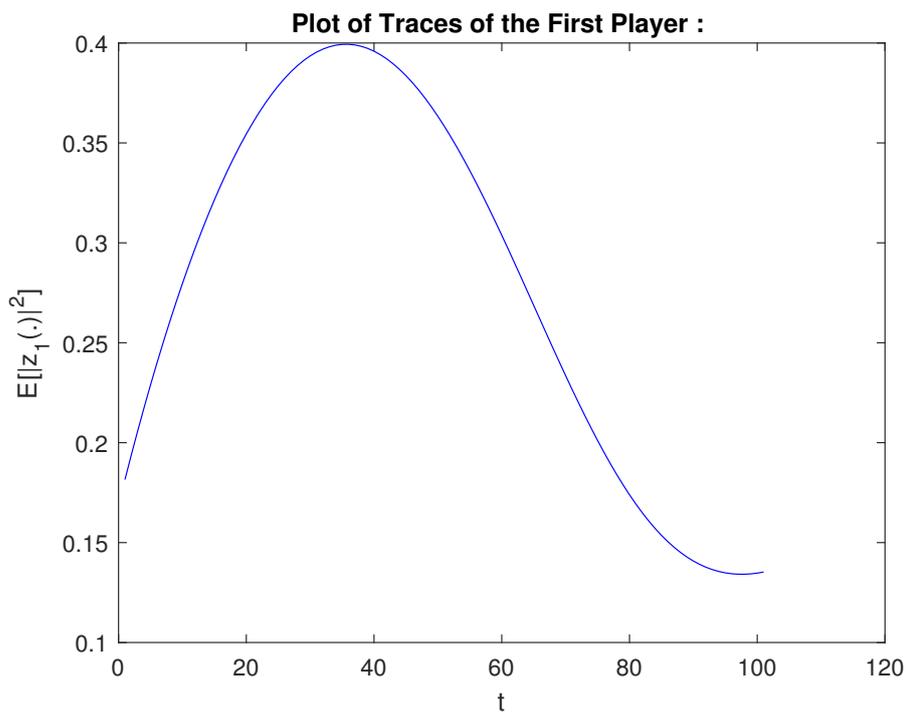
Figure 1. Plot of traces of the first player: $\mathbb{E}[|\tilde{z}_1(jh)|^2]$ in $[0, t_f] = [0, 1]$.
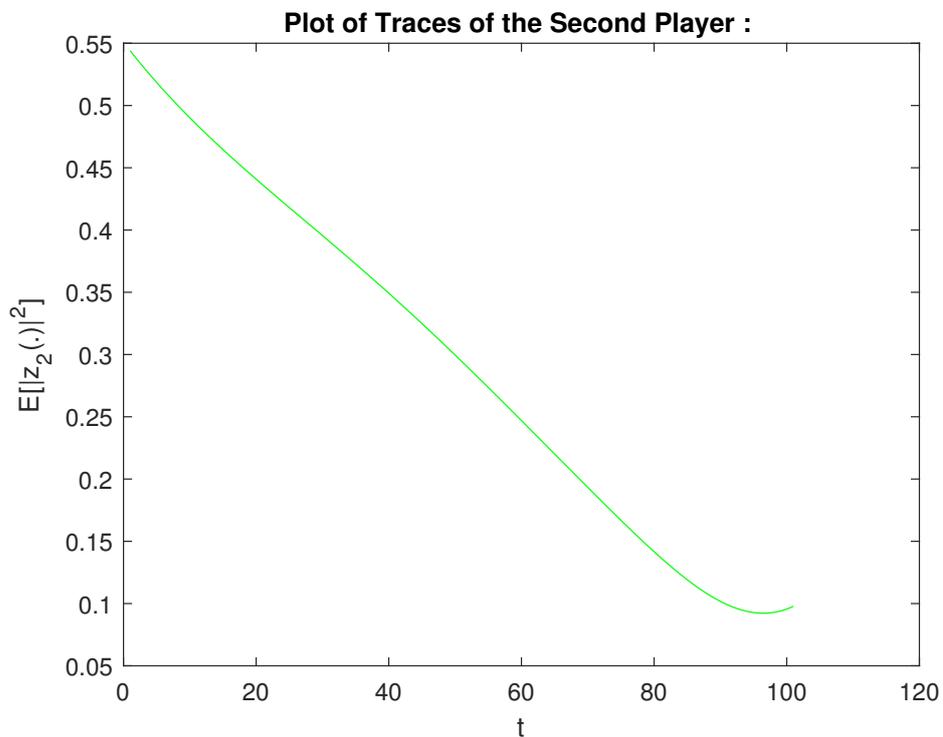
Figure 2. Plot of traces of the second player: $\mathbb{E}[|\tilde{z}_2(jh)|^2]$ in $[0, t_f] = [0, 1]$.

6. Conclusions

In this work, we studied the problem of the minimization of the deviation of some outputs of a controlled dynamical system from some given reference signals. We considered the case where the dynamical system is controlled by two decision-makers which are not cooperating. One of the two decision-makers wants to minimize the deviation of a

preferential output of the dynamical system from a given reference signal, whereas the other decision-maker wants to minimize the deviation of another output of the same dynamical controlled system from another reference signal. This problem was viewed as a problem of designing a Nash equilibrium strategy for an affine quadratic differential game with two players. Since it was supposed that the controlled dynamical system is subject to multiplicative white noise perturbations and Markovian jumping, we must find a Nash equilibrium strategy for a stochastic affine quadratic differential game. We have obtained explicit formulae of the equilibrium strategy. To this end, the solutions of two TVPs were involved. The first TVP is associated with a hybrid system formed by two nonlinear backward differential equations and two nonlinear algebraic equations, namely, the TVP (30). The second TVP is associated with a hybrid system formed by two backward linear differential equations coupled with two affine matrix algebraic equations, that is TVP (31). The first TVP is the same as that involved in the description of the Nash equilibrium strategy for an LQ differential game. The second TVP takes into consideration the reference signals $r_k(\cdot)$ together with the final targets $\zeta_k, k = 1, 2$.

There are few directions that can be considered as future research:

- Direct extensions from this article can be considered as follows: the case when two or more players (with different cost functionals) are willing to cooperate or the case when $t_f \to \infty$ for the tracking problem associated with a controlled system of type (1).
- Anther direction of future research can consider the case of a tracking problem with preview in the case when the controlled dynamical system is affected by state multiplicative and/or control multiplicative white noise perturbations. To our knowledge, this case was not yet considered in the existing literature. Some results in this direction have been reported, for example in [2,6,7], for the case of only one decision-maker and [29,30] for the case with more than one decision-maker.
- Finally, other directions can consider the case of linear quadratic tracking problem with a delay component (for one or more players) for Itô stochastic systems. Some results in this direction have been reported for examples in [10,11].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Axelband, E. The structure of the optimal tracking problem for distributed-parameter systems. *IEEE Trans. Autom. Control* **1968**, *13*, 50–56. [CrossRef]
2. Cohen, A., Shaked, U. Linear discrete-time H_∞-optimal tracking with preview. *IEEE Trans. Autom. Control* **1997** *42*, 270–276.
3. Emami-Naeini, A.; Franklin, G. Deadbeat control and tracking of discrete-time systems. *IEEE Trans. Autom. Control* **1982**, *27*, 176–181. [CrossRef]
4. Liu, D.; Liu, X. Optimal and minimum-energy optimal tracking of discrete linear time-varying systems. *Automatica* **1995**, *31*, 1407–1419. [CrossRef]
5. Shaked, U.; de Souza, C.E. Continuous-time tracking problems in an H_∞ setting: A game theory approach. *IEEE Trans. Autom. Control* **1995**, *40*, 841–852. [CrossRef]
6. Gershon, E.; Shaked, U.; Yaesh, I. H_∞ tracking of linear systems with stochastic uncertainties and preview. *IFAC Proc. Vol.* **2002** *35*, 407–412. [CrossRef]
7. Gershon, E.; Limebeer, D.J.N.; Shaked, U.; Yaesh, I. Stochastic H_∞ tracking with preview for state-multiplicative systems. *IEEE Trans. Autom. Control* **2004**, *49*, 2061–2068. [CrossRef]
8. Dragan, V.; Morozan, T. Discrete-time Riccati type equations and the tracking problem. *ICIC Express Lett.* **2008** *2*, 109–116.
9. Dragan, V.; Morozan, T.; Stoica, A.M. *Mathematical Methods in Robust Control of Linear Stochastic Systems*; Springer: New York, NY, USA, 2013.
10. Han, C.; Wang, W. Optimal LQ tracking control for continuous-time systems with pointwise time-varying input delay. *Int. J. Control. Autom. Syst.* **2017**, *15*, 2243–2252. [CrossRef]

11.   Jin, N.; Liu, S.; Zhang, H. Tracking Problem for Itô Stochastic System with Input Delay. In Proceedings of the Chinese Control Conference (CCC), Guangzhou, China, 27–30 July 2019; pp. 1370–1374.

12.   Pindyck, R. An application of the linear quadratic tracking problem to economic stabilization policy. *IEEE Trans. Autom. Control* **1972**, *17*, 287–300. [CrossRef]

13.   Alba-Florest, R.; Barbieri, E. Real-time infinite horizon linear quadratic tracking controller for vibration quenching in flexible beams. In Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Taipei, Taiwa, 8–11 October 2006; pp. 38–43.

14.   Wang, Y.-L.; Yang, G.-H. Robust H_∞ model reference tracking control for networked control systems with communication constraints. *Int. J. Control. Autom. Syst.* **2009**, *7*, 992–1000. [CrossRef]

15.   Ou, M.; Li, S.; Wang, C. Finite-time tracking control for multiple non-holonomic mobile robots based on visual servoing. *Int. J. Control* **2013**, *86*, 2175–2188. [CrossRef]

16.   Fu, Y.-M.; Lu, Y.; Zhang, T.; Zhang, M.-R.; Li, C.-J. Trajectory tracking problem for Markov jump systems with Itô stochastic disturbance and its application in orbit manoeuvring. *IMA J. Math. Control. Inf.* **2018**, 35, 1201–1216. [CrossRef]

17.   Basar, T. Nash equilibria of risk-sensitive nonlinear stochastic differential games. *J. Optim. Theory Appliations* **1999**, *100*, 479–498. [CrossRef]

18.   Buckdahn, R.; Cardaliaguet, P.; Rainer, C. Nash equilibrium payoffs for nonzero-sum stochastic differential games. *SIAM J. Control. Optim.* **2004**, *43*, 624–642. [CrossRef]

19.   Sun, H.Y.; Li, M.; Zhang, W.H. Linear quadratic stochastic differential game: Infinite time case. *ICIC Express Lett.* **2010**, *5*, 1449–1454.

20.   Sun, H.; Yan, L.; Li, L. Linear quadratic stochastic differential games with Markov jumps and multiplicative noise: Infinite time case. *Int. J. Innov. Comput. Inf. Control* **2015**, *11*, 348–361.

21.   Nakura, G. Nash tracking game with preview by stare feedback for linear continuous-time systems, In Proceedings of the 50th ISCIE International Symposium on Stochastic Systems Theory and Its Applications, Kyoto, Japan, 1–2 November 2018; pp. 49–55.

22.   Fridman, A. *Stochastic Differential Equations and Applications*; Academic: New York, NY, USA, 1975; Volume I.

23.   Oksendal, B. *Stoch. Differ. Equations*; Springer: Berlin/Heidelberg, Germany , 2003.

24.   Chung, K.L. *Markov Chains with Stationary Transition Probabilities*; Springer: Berlin/Heidelberg, Germany, 1967.

25.   Doob, J.L. *Stochastic Processes*; Wiley: New York, NY, USA, 1967.

26.   Boukas, E.R. *Stochastic Switching Systems: Analysis and Design*; Birkhäuser: Boston, MA, USA, 2005.

27.   Costa, O.L.V.; Fragoso, M.D.; Marques, R.P. *Discrete-Time Markov Jump Linear Systems*; Series: Probability and Its Applications; Springer: London, UK, 2005.

28.   Costa, O.L.V.; Fragoso, M.D.; Todorov, M.G. *Continuous-Time Markov Jump Linear Systems*; Series: Probability and Its Applications; Springer: Berlin/Heidelberg, Germany, 2013.

29.   Nakura, G. Soft-Constrained Nash Tracking Game with Preview by State Feedback for Linear Continuous-Time Markovian Jump Systems. In Proceedings of the 2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), Nara, Japan, 11–14 September 2018; pp. 450–455.

30.   Nakura, G. Nash Tracking Game with Preview by State Feedback for Linear Continuous-Time Markovian Jump Systems. In Proceedings of the 50th ISCIE International Symposium on Stochastic Systems Theory and Its Applications, Kyoto, Japan, 1–2 November 2018; pp. 56–63.