A Differential Game with the Possibility of Early Termination

D. V. Khlopin[1]

Abstract—We consider a zero-sum differential game on a finite interval, in which the players not only control the system's trajectory but also influence the terminal time of the game. It is assumed that the early terminal time is an absolutely continuous random variable, and its density is given by bounded measurable functions of time assigned by both players (the intensities of the influence of each player on the termination of the game). The payoff function may depend both on the terminal time of the game together with the position of the system at this time and on the player who initiates the termination. The strategies are formalized by using nonanticipating càdlàg processes. The existence of the game value is shown under the Isaacs condition. For this, the original game is approximated by an auxiliary game based on a continuous-time Markov chain, which depends on the controls and intensities of the players. Based on the strategies optimal in this Markov game, a control procedure with a stochastic guide is proposed for the original game. It is shown that, under an unlimited increase in the number of points in the Markov game, this procedure leads to a near-optimal strategy in the original game.

Keywords: two-person zero-sum game, Dynkin game, differential game, stochastic guide, extremal shift, continuous-time Markov chain.

INTRODUCTION

This work is devoted to the development of methods for solving differential games with random terminal time. We consider a zero-sum differential game on a finite interval $[0; T]$, in which the players can influence in a nonanticipating way the density of the moment of early (before the time T) termination of the game, setting at each time, each on their own part, the intensity (conditional density) of the termination. The cost function depends both on the terminal time and the position of the system at that moment and on the player who initiates the early termination; thus, this function becomes a random variable. For this reason, the actions of the players are considered to be aimed at optimizing the expected value of the cost function. Each player still has the opportunity to rely on their knowledge of the position of the system, assigning their own control and intensity during the game, but it should also be noted that the game between the players can occur repeatedly and the actions of the players can change during such a series of games; in particular, their actions may be nondeterministic, resulting from certain random processes. That is why we assume that, on

[1]Krasovskii Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences, Yekaterinburg, 620108 Russia
e-mail: khlopin@imm.uran.ru

the one hand, the players can choose the distributions of their controls rather than controls per se. On the other hand, by collecting statistics on a sufficiently long series of games, players have the opportunity to recover the distribution used by the opponent with any predetermined accuracy; in particular, knowing nothing during the game about the intensity of early termination already assigned by the opponent, each of them can recover the probabilistic law governing it.

In connection with the above, we assume in this paper that the strategies are random processes, which depend in a nonanticipating way on the realized trajectory of the system and, in the case of a counterstrategy, on the actions of the opponent. A detailed definition of strategies and counterstrategies is given in Section 2; we only note here that they are in principle not new. For a continuous-time Markov chain, such general definitions can be found, for example, in [26], where more constructive special cases of the definitions are described in detail. For example, in our setting, each player can use the classical deterministic strategies: piecewise continuous program strategies or piecewise constant positional strategies. We also note that the requirement imposed in this paper that the strategies can employ only càdlàg processes is rather of a technical nature; to some extent it can be justified by the facts that in control theory piecewise continuous controls are everywhere dense, whereas in the theory of differential games, within the classical setting [20, 21, 30], piecewise constant positional controls of the players implement the value function with any predetermined accuracy.

Let us discuss similar settings studied in the literature. The general statement of a game in which a player controls the moment of exiting the process goes back to Dynkin's classic paper [10]. In such games (*Dynkin games, stopping games*), it is usually assumed that there is some conflict-control process and at least one of the players must choose by their actions a convenient time for stopping the entire game. Such game settings have numerous applications, in particular, in financial mathematics (see, for example, [1, 7, 15]); for them, the questions of the existence and description of the value of the game, the construction of optimal strategies, and the structure of such strategies are investigated. Starting with the classical works [5, 6], the dynamics of the game is usually described by a stochastic differential equation, but Lévy–Feller processes [4], as well as discrete-time (see, for example, [24, 29]) or continuous-time [11, 22] Markov chains, are also used. Finally, a differential game in which the distribution of the terminal time (and hence the intensity) is known to the players in advance was considered, for example, in [13, 23].

In this paper, for the game under consideration, we prove the existence of both the value and approximately optimal strategies that implement the value with any predetermined accuracy. The construction of the strategies is based on the technique of a stochastic guide; the idea of the technique is as follows. For any motion y of the original conflict-control system, a player uniquely reconstructs step-by-step a certain distribution $\mathbb{P}^Z$ of auxiliary motions z calculated by this player. At each time t_k of some partition, the player assigns for the following time interval a control $\bar{u}$ using the realized position $z(t_k)$, which is in fact a random variable, by the rule

$$\bar{u}(t) \equiv u^{\text{to } z(t_k)}(t_k, y(t_k)) \quad \forall t \in [t_k; t_{k+1}), \tag{0.1}$$

where $u^{\text{to } z}$ is a control that maximally shifts the position of the original system to the point z. This technique solves the original problem and goes back conceptually to the classical methods for differential games [20] proposed by Krasovskii and Subbotin. For example, the rule $\bar{u}^{\text{to } z}$ is fundamental to the extremal shift method, and the rule (0.1) with substitution $y \equiv z$ is used in the positional formalization of a differential game introduced by Krasovskii and Subbotin in [20]. The method of control with a guide proposed by the same authors originally presumed a one-to-one dependence of z on y, but later a nondeterministic rule was proposed for constructing z

in Krasovskii's paper [17]. In [18, 19] a stochastic guide was constructed with the use of the Itô differential equation. In [2, 3] the trajectory of a continuous-time Markov game is considered as z. In the present paper, z is also generated by an approximately optimal strategy in a continuous-time Markov game. In particular, the question of existence of the value in the original game is reduced to a similar fact for a Markov game, which was obtained in [14].

The paper is organized as follows. In Section 1 we formulate the original differential game and state (up to the definition of the strategies) Theorem 1, which guarantees the existence of the value in this game and is the main result of the paper. In Section 2 we give some basic definitions from the theory of random processes and introduce formalizations of the strategies used in the paper. In the next section, we extend the dynamics of the original game and represent the original objective functional as a discounted payoff. Section 4 is devoted to an auxiliary Markov game. In Section 5 the technique of a stochastic guide for the first player is considered: a random process (double game) is introduced, in which the first component corresponds to the state of the original game and the second component corresponds to the state of the stochastic guide; then, with the use of an extremal shift, the control of the first player is written in each component and a fictitious control is written for the second player in the Markov game. The last section is devoted to the proof of the theorem, which reduces to estimating the difference between the components of the random process, i.e., the trajectories of the original game and of the stochastic guide.

1. PROBLEM STATEMENT

1.1. Dynamics of the original game.
Consider a conflict-control system in $\mathbb{R}^d$

$$\frac{dy(t)}{dt} = f(t, y(t), u(t), v(t)), \quad y(0) = x_* \in \mathbb{R}^d, \quad t \in [0; T], \quad u(t) \in \mathbb{U}, \quad v(t) \in \mathbb{V}. \tag{1.1}$$

Assume that $\mathbb{U}$ and $\mathbb{V}$ are metric compact sets in some finite-dimensional Euclidean spaces and the function $f : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{U} \times \mathbb{V} \to \mathbb{R}^d$ is Lipschitz in all variables. We also require the saddle point condition in the small game (the Isaacs condition) [20, 21]: for all $x, w \in \mathbb{R}^d$,

$$\min_{u \in \mathbb{U}} \max_{v \in \mathbb{V}} \langle w, f(t, x, u, v) \rangle = \max_{v \in \mathbb{V}} \min_{u \in \mathbb{U}} \langle w, f(t, x, u, v) \rangle. \tag{1.2}$$

In addition to controlling the motion in $\mathbb{R}^d$, each player can influence the terminal time of the game. Let us fix closed intervals $[\varphi_-; \varphi_+], [\psi_-; \psi_+] \subset \mathbb{R}_+$, which are sets of parameters that control the intensity of game termination on the part of each player. Assume that the first and second players have chosen piecewise continuous functions $[0; T] \ni t \mapsto \varphi(t) \in [\varphi_-; \varphi_+]$ and $[0; T] \ni t \mapsto \psi(t) \in [\psi_-; \psi_+]$, respectively. Then, for almost all $t \in [0; T)$, the conditional probability of termination of the game on the interval $[t; t + \Delta t)$ if the game was still continuing at the time t is $(\varphi(t) + \psi(t))\Delta t + o(\Delta t)$. Solving the corresponding differential equation for the total intensity $\varphi + \psi$, we find that, for any time $\tau \in [0; T)$, the probability of termination of the game before the time τ is

$$\int_0^\tau (\varphi(t) + \psi(t)) e^{-\int_0^t (\varphi(s) + \psi(s))\, ds}\, dt; \tag{1.3}$$

in particular, for $\varphi(t) + \psi(t) \equiv C$, the terminal time of the game has the same distribution as $\min(\theta, T)$, where θ is distributed exponentially with parameter C. Using (1.3) and the corresponding differential equations for the probabilities of termination of the game before the moment $\tau < T$

on the initiative of the first and second player, respectively, we have for them the values

$$\int_0^\tau \varphi(t) e^{-\int_0^t (\varphi(s)+\psi(s))\,ds}\,dt, \qquad \int_0^\tau \psi(t) e^{-\int_0^t (\varphi(s)+\psi(s))\,ds}\,dt. \tag{1.4}$$

As a corollary, we find that, in the case of termination of the game at a time $t < T$, the player responsible for this is the first with probability $\varphi(t)/(\varphi(t) + \psi(t))$ and the second with probability $\psi(t)/(\varphi(t) + \psi(t))$. It remains to formalize formulas (1.3) and (1.4).

For this, we equip the square $[0;1]^2$ with the σ-algebra $\mathcal{B}([0;1]^2)$ of Borel sets and consider the Lebesgue measure λ on them. We obtain the probability space

$$\left([0;1]^2, \mathcal{B}([0;1]^2), \mathbb{P}_{[0;1]^2} \overset{\triangle}{=} \lambda\right),$$

which assigns a random number ω_i to each player. Now, for all Borel mappings $\varphi : [0;T) \to [\varphi_-;\varphi_+]$ and $\psi : [0;T) \to [\psi_-;\psi_+]$, we introduce random variables $\omega \mapsto \theta_1(\omega)$ and $\omega \mapsto \theta_2(\omega)$ by the following rule: for all $\omega = (\omega_1, \omega_2) \in [0;1]^2$,

$$\theta_1(\omega) \overset{\triangle}{=} \inf\left\{t \in [0;T) \mid \int_0^t \varphi(t)\,dt = -\ln\omega_1\right\} \cup \{T\}, \tag{1.5}$$

$$\theta_2(\omega) \overset{\triangle}{=} \inf\left\{t \in [0;T) \mid \int_0^t \psi(t)\,dt = -\ln\omega_2\right\} \cup \{T\}. \tag{1.6}$$

Note that, for any $\tau \in [0;T)$, the probability of the event $\theta_1 < \tau$ coincides with (1.3) in the case $\psi \equiv 0$. Similarly, the probability of the event $\theta_2 < \tau$ coincides with (1.3) in the case $\varphi \equiv 0$. Thus, the random variable θ_i coincides with the time of the possible termination of the game on the initiative of the ith player in the case of zero intensity of the opponent. Moreover, for any mappings $\varphi : [0;T) \to [\varphi_-;\varphi_+]$ and $\psi : [0;T) \to [\psi_-;\psi_+]$, the random variables θ_i are independent because the random variables ω_i are independent; hence, for the time of the early termination of the game

$$\theta_{\min} \overset{\triangle}{=} \min(\theta_1, \theta_2),$$

which is a random variable, the probability of the event $\theta_{\min} < \tau$ is found by formula (1.3), and the probabilities of the events $\theta_1 < \tau$ and $\theta_2 < \tau$ are found by formula (1.4), as required. In what follows, we suppose that (ω_1, ω_2) is chosen before the start of the game and is independent of the actions of the players, who have no information about (ω_1, ω_2) before the terminal time. Thus, the players know formulas (1.3) and (1.4) and know (or hypothesize) their intensity and/or the intensity of the opponent but have no other information about the terminal time $\theta_{\min}$.

For simplicity of notation, we introduce the sets of control parameters of the first and second players: $\bar{\mathbb{U}} \overset{\triangle}{=} \mathbb{U} \times [\varphi_-;\varphi_+]$ and $\bar{\mathbb{V}} \overset{\triangle}{=} \mathbb{V} \times [\psi_-;\psi_+]$.

1.2. Goals of the players and the value of the game. Let us define the goals of the players. Let functions $\sigma_0 : \mathbb{R}^d \to \mathbb{R}$, $\sigma_1 : [0;T] \times \mathbb{R}^d \to \mathbb{R}$, and $\sigma_2 : [0;T] \times \mathbb{R}^d \to \mathbb{R}$ be given. Assume that these functions are locally Lipschitz in x, t and have norms bounded by the number 1. Let $\sigma_0(y(T))$ be the payoff of the first player to the second if there has been no early termination, i.e., the game has terminated at the time T. Let the number $\sigma_1(\theta_1, y(\theta_1))$ be the payoff of the first

player to the second if the first player has initiated an early termination of the game (at a time θ_1). Similarly, $\sigma_2(\theta_2, y(\theta_2))$ is the payoff of the first player to the second if the termination of the game (at the time θ_2) has been initiated by the second player. Thus, almost always on $[0;1]^2$, for a known trajectory $y(\cdot)$ and known actions $(\bar{u}, \bar{v})(\cdot)$ of the players, the objective function is defined; it is the random variable

$$
J(y, \bar{u}, \bar{v}) \triangleq
\begin{cases}
\sigma_0(y(T)), & T = \theta_{\min}, \\
\sigma_1(\theta_1, y(\theta_1)), & \theta_1 = \theta_{\min}, \\
\sigma_2(\theta_2, y(\theta_2)), & \theta_2 = \theta_{\min}.
\end{cases}
$$

Since the players do not know the outcome $\omega \in [0;1]^2$, they can only optimize the expected value of the objective function J

$$
\mathbb{E} J(y, \bar{u}, \bar{v}),
$$

which, similarly to the trajectory $y(\cdot)$ of system (1.1), is uniquely recovered from the players' controls $\bar{u}(\cdot)$ and $\bar{v}(\cdot)$.

Formalizations of the strategies will be given in the following section; we only note that one of the players will be allowed to feed as their control a càdlàg (right continuous and having limit on the left) trajectory of some random process, which depends on the trajectory of (1.1) in a nonanticipating way. The other will be allowed to know not only the trajectory but also the control used by the opponent; as a consequence, the control of this player depends on these trajectory and control as a random process in a nonanticipating way. Let us introduce the necessary notation.

Let $\mathfrak{S}^{\mathrm{I}}$ and $\mathfrak{S}^{\mathrm{II}}$ be the sets of admissible strategies of the players considered as distributions of random processes that depend in a nonanticipating way on the trajectory $y(\cdot)$. Denote the sets of admissible counterstrategies of the players by $\mathfrak{Q}^{\mathrm{I}}_{x_*}$ and $\mathfrak{Q}^{\mathrm{II}}_{x_*}$, and denote by $\mathfrak{R}_{x_*}(\mathbb{P}^{\mathrm{I+II}})$ the set of all possible distributions $\mathbb{P}^{\mathrm{all}}_{x_*}$ of random processes in the game where the players adhere to a joint strategy $\mathbb{P}^{\mathrm{I+II}}$. Naturally, any stepwise strategies (in particular, program strategies) are admissible strategies, and any admissible strategy is a counterstrategy.

Depending on which of the players is informationally discriminated, i.e., which of them has in advance informed the opponent about their strategy, we obtain two values of the game:

$$
V^+ \triangleq \inf_{\mathbb{P}^{\mathrm{I}} \in \mathfrak{S}^{\mathrm{I}}_{x_*}} \sup_{\mathbb{Q}^{\mathrm{II}} \in \mathfrak{Q}^{\mathrm{II}}_{x_*}} \sup_{\mathbb{P}^{\mathrm{all}}_{x_*} \in \mathfrak{R}_{x_*}(\mathbb{Q}^{\mathrm{II}}[\mathbb{P}^{\mathrm{I}}])} \mathbb{E}^{\mathrm{all}}_{x_*} J(y, \bar{u}, \bar{v}),
$$

$$
V^- \triangleq \sup_{\mathbb{P}^{\mathrm{II}} \in \mathfrak{S}^{\mathrm{II}}_{x_*}} \inf_{\mathbb{Q}^{\mathrm{I}} \in \mathfrak{Q}^{\mathrm{I}}_{x_*}} \inf_{\mathbb{P}^{\mathrm{all}}_{x_*} \in \mathfrak{R}_{x_*}(\mathbb{Q}^{\mathrm{I}}[\mathbb{P}^{\mathrm{II}}])} \mathbb{E}^{\mathrm{all}}_{x_*} J(y, \bar{u}, \bar{v}).
$$

Theorem 1. *The equality $V^- = V^+$ is true. Moreover, for each positive ε, the players have admissible strategies $\mathbb{P}^{\mathrm{I}}_\varepsilon \in \mathfrak{S}^{\mathrm{I}}_{x_*}$ and $\mathbb{P}^{\mathrm{II}}_\varepsilon \in \mathfrak{S}^{\mathrm{II}}_{x_*}$ that guarantee*

$$
-\varepsilon + \sup_{\mathbb{Q}^{\mathrm{II}} \in \mathfrak{Q}^{\mathrm{II}}_{x_*}} \sup_{\mathbb{P}^{\mathrm{all}}_{x_*} \in \mathfrak{R}_{x_*}(\mathbb{Q}^{\mathrm{II}}[\mathbb{P}^{\mathrm{I}}_\varepsilon])} \mathbb{E}^{\mathrm{all}}_{x_*} J(y, \bar{u}, \bar{v}) \le V^+ = V^- \le \varepsilon + \inf_{\mathbb{Q}^{\mathrm{I}} \in \mathfrak{Q}^{\mathrm{I}}_{x_*}} \inf_{\mathbb{P}^{\mathrm{all}}_{x_*} \in \mathfrak{R}_{x_*}(\mathbb{Q}^{\mathrm{I}}[\mathbb{P}^{\mathrm{II}}_\varepsilon])} \mathbb{E}^{\mathrm{all}}_{x_*} J(y, \bar{u}, \bar{v}).
$$

This theorem will be proved in the last section. Before that we recall some definitions from the theory of random processes and give the definitions of strategies promised earlier.

2. STRATEGIES AS DISTRIBUTIONS OF RANDOM PROCESSES

2.1. General facts from the theory of random processes. Recall that, for any pair of probability spaces $(\Omega', \mathcal{F}', \mathbb{P}')$ and $(\Omega'', \mathcal{F}'', \mathbb{P}'')$, we can consider the σ-algebra $\mathcal{F}' \otimes \mathcal{F}''$ of subsets of $\Omega' \times \Omega''$ as the smallest σ-algebra containing all products of the form $A \times B$ for $A \in \mathcal{F}'$ and $B \in \mathcal{F}''$. On this algebra, we can uniquely reconstruct a probability $\mathbb{P}' \otimes \mathbb{P}''$ by the rule

$$(\mathbb{P}' \otimes \mathbb{P}'')(A \times B) = \mathbb{P}'(A)\mathbb{P}''(B) \quad \text{for all} \quad A \in \mathcal{F}' \text{ and } B \in \mathcal{F}''.$$

We obtain the product $(\Omega' \times \Omega'', \mathcal{F}' \otimes \mathcal{F}'', \mathbb{P}' \otimes \mathbb{P}'')$ of the original probability spaces. In addition, if there is a mapping g from a probability space $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ to a measurable space $(\Omega_2, \mathcal{F}_2)$ and this mapping is measurable, i.e., $g^{-1}(A) \in \mathcal{F}_1$ for all $A \in \mathcal{F}_2$, then this mapping specifies the *push-forward of the measure*, probability $\mathbb{P}_2$ for $(\Omega_2, \mathcal{F}_2)$, by the rule $\mathbb{P}_2(A) \stackrel{\triangle}{=} \mathbb{P}_1(g^{-1}(A))$ for all $A \in \mathcal{F}_2$. In particular, for the probability introduced above for $\Omega' \times \Omega''$, the marginal probabilities $\mathbb{P}'$ and $\mathbb{P}''$ for Ω' and Ω'' coincide with the push-forwards of the measure $\mathbb{P}' \otimes \mathbb{P}''$ by the mappings $(\omega', \omega'') \mapsto \omega'$ and $(\omega', \omega'') \mapsto \omega''$, respectively.

Fix a state space $\mathbb{X}$ as a closed subset of a finite-dimensional Euclidean space, and equip it with the σ-algebra of Borel sets $\mathcal{B}(\mathbb{X})$, thus obtaining the measurable space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$. Fix also a time interval $\mathcal{T}$.

Consider now the Skorokhod space $D(\mathcal{T}, \mathbb{X})$ [8, 16] of all possible càdlàg functions from $\mathcal{T}$ to $\mathbb{X}$ (functions from $\mathcal{T}$ to $\mathbb{X}$ that are right continuous and have limit on the left). This space is Polish; equipping it with the corresponding Borel σ-algebra $\mathcal{B}(D(\mathcal{T}, \mathbb{X}))$, we obtain a measurable space. On each Skorokhod space, we fix the default *canonical filtration*

$$\left(\mathcal{A}_t^{\mathbb{X}} \stackrel{\triangle}{=} \mathcal{B}\big(D((-\infty; t] \cap \mathcal{T}, \mathbb{X})\big) \otimes \{\varnothing, D\big((t; +\infty) \cap \mathcal{T}, \mathbb{X}\big)\}\right)_{t \in \mathcal{T}},$$

obtaining the *stochastic basis* $\big(D(\mathcal{T}, \mathbb{X}), \mathcal{B}(D(\mathcal{T}, \mathbb{X})), (\mathcal{A}_t^{\mathbb{X}})_{t \in \mathcal{T}}\big)$.

Following [25, Ch. IV, Sect. 1, Definition O2], we define a *random process with time interval $\mathcal{T}$ and state space* $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ as a system consisting of some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and some family $(X_t)_{t \in \mathcal{T}}$ of measurable mappings from Ω to $\mathbb{X}$. Each X_t is called the *state of the process at the time t*, and the mapping $t \mapsto X_t(\omega)$ is called the *trajectory of the process* corresponding to the outcome $\omega \in \Omega$. A random process is called a *càdlàg process* if $\mathbb{P}$-almost all its trajectories are such. For each process, we can consider the corresponding *first canonical process* [25], replacing each ω by the corresponding trajectory $t \mapsto X_t(\omega)$; in particular, all càdlàg processes with state space $\mathbb{X}$ and time interval $\mathcal{T}$ are uniquely recovered by some probability measure (*distribution of the process* [9, Sect. 17.1]) given on the measurable space $(D(\mathcal{T}, \mathbb{X}), \mathcal{B}(D(\mathcal{T}, \mathbb{X})))$. In what follows, we will specify most càdlàg processes in the form of first canonical processes, fixing only their distributions in an appropriate Skorokhod space.

We will say that $(X_t)_{t \in \mathcal{T}}$ is *adapted* to $(\mathcal{F}_t)_{t \in \mathcal{T}}$ if, for each $t \in \mathcal{T}$, the random variable $\omega \mapsto X_t(\omega)$ is $\mathcal{F}_t$-measurable. A random variable τ with values in $\mathcal{T}$ is called a *stopping time* (with respect to a filtration $(\mathcal{F}_t)_{t \in \mathcal{T}}$) if, for all $t \in \mathcal{T}$, the event $\{\tau \leq t\} = \{\omega \mid \tau(\omega) \leq t\}$ lies in $\mathcal{F}_t$. Note that if the canonical filtration $(\mathcal{A}_t^{\mathbb{X}})_{t \in \mathcal{T}}$ is taken as a filtration $(\mathcal{F}_t)_{t \in \mathcal{T}}$ for $\Omega \stackrel{\triangle}{=} D(\mathcal{T}, \mathbb{X})$, then any process given on this basis is $\mathcal{A}_t^{\mathbb{X}}$-measurable, and all possible stopping times τ are exactly the mapping $\tau : D(\mathcal{T}, \mathbb{X}) \mapsto \mathcal{T}$ such that, for any $x, x' \in D(\mathcal{T}, \mathbb{X})$, the equality $x(t) = x'(t)$ for all $\tau(x) \geq t$ implies $\tau(x) = \tau(x')$.

2.2. Strategy sets. Let $\mathcal{T}$ be an interval closed on the left, and let t_0 be its lower bound.

In addition to the Skorokhod space $D(\mathcal{T}, \mathbb{X})$, we consider some compact sets Υ_1 and Υ_2 in some finite-dimensional Euclidean spaces; these sets are the sets of control parameters of the players. Let $\Upsilon \overset{\triangle}{=} \Upsilon_1 \times \Upsilon_2$. *Program strategies of the first and second players* and their *joint program strategies* are all possible piecewise right continuous and having limit on the left mappings that are elements of $D(\mathcal{T}, \Upsilon_1)$, $D(\mathcal{T}, \Upsilon_2)$, and $D(\mathcal{T}, \Upsilon)$, respectively. As earlier, equip them with σ-algebras of Borel sets with the canonical filtration.

For the state space $\mathbb{X}$ and the set of control parameters Υ, we call a mapping

$$\mathcal{P}(\mathbb{X}) \times D(\mathcal{T}, \Upsilon) \ni (\varrho, \upsilon(\cdot)) \mapsto \mathbb{P}^G_{\upsilon(\cdot)}[\varrho] \in \mathcal{P}\big(D(\mathcal{T}, \mathbb{X})\big)$$

an *admissible dynamics* $\mathbb{P}^G$ with initial distribution $\varrho \in \mathcal{P}(\mathbb{X})$ if it satisfies the initial condition $\mathbb{P}^G_{\upsilon(\cdot)}[\varrho](X(t_0) \in A) \equiv \varrho(A)$ for all $A \in \mathcal{B}(\mathbb{X})$ and $\upsilon(\cdot) \in D(\mathcal{T}, \Upsilon)$ and is nonanticipating; i.e., for any $\upsilon'(\cdot), \upsilon''(\cdot) \in D(\mathcal{T}, \Upsilon)$ and time $t \in \mathcal{T}$, the equality $\upsilon'|_{[t_0;t)} = \upsilon''|_{[t_0;t)}$ implies $\mathbb{P}^G_{\upsilon'(\cdot)}[\varrho](A) = \mathbb{P}^G_{\upsilon''(\cdot)}[\varrho](A)$ for all $A \in \mathcal{A}^{\mathbb{X}}_t$. Note that any probability $\mathbb{P}^G_{\upsilon(\cdot)}[\varrho]$ is regular, since $D(\mathcal{T}, \mathbb{X})$ is a complete separable metric space.

A *strategy* $\mathbb{P}^i$ *of the ith player* is a mapping taking each element from $x \in D(\mathcal{T}, \mathbb{X})$ to a distribution $\mathbb{P}^i_{x(\cdot)}$ on $D(\mathcal{T}, \Upsilon_i)$ (and, hence, to a first canonical process $(\upsilon^i_{x(\cdot)}(t))_{t \in \mathcal{T}}$ with values in Υ_i) such that the nonanticipation condition holds for all $t \in \mathcal{T}$: if $x'(s) = x''(s)$ for all $s \in \mathcal{T}$ not greater than t, then $\mathbb{P}^i_{x'(\cdot)}(A) = \mathbb{P}^i_{x''(\cdot)}(A)$ for all $A \in \mathcal{A}^{\Upsilon_i}_t$. The notion of *joint strategy* $\mathbb{P}^{\mathrm{I+II}}$ of the players, which maps trajectories from $D(\mathcal{T}, \mathbb{X})$ to distributions on $D(\mathcal{T}, \Upsilon)$, is introduced similarly.

Note that any program strategy $\upsilon_i(\cdot) \in D(\mathcal{T}, \Upsilon_i)$, being a mapping $x(\cdot) \mapsto \upsilon_i(\cdot)$ from $D(\mathcal{T}, \mathbb{X})$ to $D(\mathcal{T}, \Upsilon_i)$ independent of the trajectory $x(\cdot)$, is a strategy of the ith player.

For any initial condition $\varrho \in \mathcal{P}(\mathbb{X})$, we call a random process $(X_t)_{t \in \mathcal{T}}$ with values in $\mathbb{X}$ a *realization of a joint strategy* $\mathbb{P}^{\mathrm{I+II}}$ *for an admissible dynamics* $\mathbb{P}^G$ if there is a process $(\upsilon_t, X_t)_{t \in \mathcal{T}}$ given by some distribution $\mathbb{P}^{\mathrm{all}}_\varrho$ on $D(\mathcal{T}, \Upsilon) \times D(\mathcal{T}, \mathbb{X})$ for which the mappings $\upsilon(\cdot) \mapsto \mathbb{P}^G_{\upsilon(\cdot)}[\varrho]$ and $x(\cdot) \mapsto \mathbb{P}^{\mathrm{I+II}}_{x(\cdot)}$ are the Radon–Nikodym derivatives of the marginal distributions $\mathbb{P}^\chi_\varrho$ and $\mathbb{P}^\Upsilon_\varrho$ (on $D(\mathcal{T}, \mathbb{X})$ and $D(\mathcal{T}, \Upsilon)$, respectively) with respect to $\mathbb{P}^{\mathrm{all}}_\varrho$; i.e.,

$$\mathbb{P}^\Upsilon_\varrho(A) \overset{\triangle}{=} \mathbb{P}^{\mathrm{all}}_\varrho(\upsilon. \in A) = \int_{D(\mathcal{T}, \mathbb{X})} \mathbb{P}^{\mathrm{I+II}}_{x(\cdot)}(A)\, \mathbb{P}^\chi(dx) \quad \forall A \in \mathcal{B}(D(\mathcal{T}, \Upsilon)),$$

$$\mathbb{P}^\chi_\varrho(C) \overset{\triangle}{=} \mathbb{P}^{\mathrm{all}}_\varrho(X. \in C) = \int_{D(\mathcal{T}, \Upsilon)} \mathbb{P}^G_{\upsilon(\cdot)}[\varrho](C)\, \mathbb{P}^\Upsilon(d\upsilon) \quad \forall C \in \mathcal{B}(D(\mathcal{T}, \mathbb{X})).$$

We will denote the set of all possible realizations for a strategy $\mathbb{P}^{\mathrm{I+II}}$ by $\mathfrak{R}_\varrho(\mathbb{P}^{\mathrm{I+II}})$.

We call a strategy $\mathbb{P}^i$ of the ith player *admissible* for an initial condition $\varrho \in \mathcal{P}(\mathbb{X})$ if the pair of strategies $(\mathbb{P}^i, \delta_{\upsilon_{3-i}})$ has the corresponding realization for any program strategy $\upsilon_{3-i} \in D(\mathcal{T}, \Upsilon_{3-i})$ of the opponent. In particular, the program strategies themselves are admissible; indeed, for any choice of program strategies (υ_1, υ_2) of the players, we can take $\delta_{(\upsilon_1, \upsilon_2)} \otimes \mathbb{P}^G_{(\upsilon_1, \upsilon_2)(\cdot)}[\varrho]$ as realizations of their joint strategy.

A *counterstrategy* $\mathbb{Q}^i[\mathbb{P}^{\mathrm{III}-i}]$ *of the ith player* is a mapping taking each admissible strategy of the opponent $\mathbb{P}^{\mathrm{III}-i}$ to some joint strategy $x(\cdot) \mapsto \mathbb{Q}^i[\mathbb{P}^{\mathrm{III}-i}]_{x(\cdot)}$ on $D(\mathcal{T}, \Upsilon_1 \times \Upsilon_2)$ (and, hence,

to a random process $(v_{x(\cdot)}(t))_{t \in \mathcal{T}}$ with values in $\Upsilon_1 \times \Upsilon_2)$ such that the corresponding marginal distribution on Υ_{3-i} coincides with $\mathbb{P}^{\mathrm{III}-i}_{x(\cdot)}$. A counterstrategy is called *admissible* if the strategy $x(\cdot) \mapsto \mathbb{Q}^i[\mathbb{P}^{\mathrm{III}-i}]_{x(\cdot)}$ is admissible for any admissible strategy of the opponent $\mathbb{P}^{\mathrm{III}-i}$. It is clear that any admissible strategy $\mathbb{P}^i$ is also an admissible counterstrategy because it assigns to a strategy of the opponent $\mathbb{P}^{\mathrm{III}-i}$ some joint strategy $x(\cdot) \mapsto \mathbb{P}^{\mathrm{I}}_{x(\cdot)} \otimes \mathbb{P}^{\mathrm{II}}_{x(\cdot)}$.

Denote by $\mathfrak{S}^{\mathrm{I}}_{x_*}$ and $\mathfrak{S}^{\mathrm{II}}_{x_*}$ the sets of admissible (for an initial condition $\varrho = \delta_{x_*}$) strategies of the first and second players. Similarly, denote by $\mathfrak{Q}^{\mathrm{I}}_{x_*}$ and $\mathfrak{Q}^{\mathrm{II}}_{x_*}$ the sets of admissible (for the same condition) counterstrategies of the players. The above implies the chain of embeddings $D(\mathcal{T}, \Upsilon_i) \hookrightarrow \mathfrak{S}^i_{x_*} \hookrightarrow \mathfrak{Q}^i_{x_*}$.

We construct a stepwise strategy, for example, for the first player, showing at the same time that any such strategy is also admissible. Consider some partition $\Delta = (t_k)_{k \in \mathbb{N}}$ of the time interval $\mathcal{T}$, and let some admissible (for example, program) strategy $\mathbb{P}^{I:k}$ of the first player be given for each $k \in \mathbb{N}$ on the interval $[t_{k-1}; t_k)$. In particular, for any program strategy $v_k \in D([t_{k-1}; t_k), \Upsilon_2)$ of the second player, there is a realization $\mathbb{P}^{\mathrm{all}:k}$ of their joint strategy $\delta_{v_k} \otimes \mathbb{P}^{I:k}$ on this interval. Now, each initial (at the time t_{k-1}) distribution ϱ_{k-1} corresponds to the probability $\mathbb{P}^{\mathrm{all}:k}_{\varrho_{k-1}}(X(\cdot) \in B)$ for all Borel subsets B of the set $D([t_{k-1}; t), \mathbb{X})$ and, hence, to the marginal probability $\varrho_k[\varrho](B) \overset{\triangle}{=} \mathbb{P}^{\chi:k}_{\varrho_{k-1}}(X(t_k) \in B)$ depending on ϱ. Thus, we have constructed a sequence of transition probabilities $\varrho_k : \mathcal{P}(\mathbb{X}) \to \mathcal{P}(\mathbb{X})$ depending on ϱ_0; after that, we can introduce a distribution $\mathbb{P}^{\mathrm{all}}_{\varrho_0}$ as the probability $\otimes_{k \in \mathbb{N}} \mathbb{P}^{\mathrm{all}:k}_{\varrho_{k-1}}$ on Borel subsets of $\prod_{k \in \mathbb{N}} D([t_{k-1}; t_k), \Upsilon \times \mathbb{X})$. Thus, the realization of the joint strategy $\mathbb{P}^{\mathrm{all}}_{\varrho_0}$ has been constructed, and any strategy constructed in this way (any *stepwise strategy*) is admissible.

2.3. Admissibility of the original dynamics. Let us show that the original game defines an admissible dynamics with $\mathbb{X} \overset{\triangle}{=} \mathbb{R}^{d+1}$, $\mathcal{T} \overset{\triangle}{=} [0; T]$, and $\Upsilon \overset{\triangle}{=} \bar{\mathbb{U}} \times \bar{\mathbb{V}}$. Indeed, if we fix program strategies, i.e., some càdlàg controls $\bar{u} = (u, \varphi)$ and $\bar{v} = (v, \psi)$, then the corresponding deterministic process $t \mapsto y(t)$ uniquely defines the Dirac measure $\mathbb{P}_{(u,v)}[\varrho] = \delta_y$ over the set of its trajectories from $D(\mathcal{T}, \mathbb{R}^d)$ under a Dirac initial condition ϱ. Due to their linearity, these distributions $\mathbb{P}_{(u,v)}[\varrho]$ are defined for all possible initial conditions ϱ. Further, each pair (φ, ψ) defines stopping times θ_i and, together with them, $\theta_{\min}$ by means of formulas (1.5), (1.6). Let a random variable s be zero if the game is not completed upon the expiry of the time of the game ($T = \theta_{\min}$), be $1 + \theta_{\min}$ if $\theta_1 = \theta_{\min} \in [0; T)$, and be $-1 - \theta_{\min}$ if $\theta_2 = \theta_{\min} \in [0; T)$. To each pair

$$(x(\cdot), s) \in D([0; T], \mathbb{R}^d) \times ([-T - 1; -1] \cup \{0\} \cup [1; T + 1])$$

we assign a trajectory $y_{(x(\cdot), s)}(\cdot) \in D([0; T], \mathbb{X})$ by the equalities

$$t \mapsto \big(t, x(\min(t, |s| - 1)), s1_{[0;t]}(|s| - 1)\big) \quad \text{for } |s| \neq 0 \quad \text{and} \quad t \mapsto (t, x(t), 0) \quad \text{otherwise.} \quad (2.1)$$

Then the push-forward of the measure $\mathbb{P}_{(u,v)}[\varrho] \otimes \mathbb{P}_{(\varphi,\psi)}$ by this mapping specifies a distribution $\mathbb{P}^G_{(\bar{u},\bar{v})}[\varrho]$ on $D(\mathcal{T}, \mathbb{X})$. In particular, the last coordinate now switches from the zero value to the value of the random variable s at the time of the early termination of the game. Now the distribution $\mathbb{P}^{\mathrm{all}}_{\varrho}$ on $D(\mathcal{T}, \bar{\mathbb{U}} \times \bar{\mathbb{V}}) \times D(\mathcal{T}, \mathbb{X})$ is uniquely recovered by the following rule: for all Borel subsets $B \subset D(\mathcal{T}, \bar{\mathbb{U}} \times \bar{\mathbb{V}})$ and $D \subset D(\mathcal{T}, \mathbb{X})$,

$$\mathbb{P}^{\mathrm{all}}_{\varrho}(B \times D) \overset{\triangle}{=} \mathbb{P}^G_{(\bar{u},\bar{v})}[\varrho](D)1_B(\bar{u}, \bar{v}). \quad (2.2)$$

Since a realization is constructed in the original game for an arbitrary joint program strategy, i.e., a pair of càdlàg controls $\bar{u} = (u, \varphi)$ and $\bar{v} = (v, \psi)$, it follows that the dynamics of the original game is admissible. Consequently, stepwise strategies in the original game are also admissible.

3. EXTENSION OF THE ORIGINAL GAME

Although the game process corresponding to the original game is completely defined, it will be convenient for us to extend its definition so that the process will be given on an unbounded time interval and in an extended state space. For this, we put $\mathcal{X} \triangleq \mathbb{R}_+ \times \mathbb{R}^{d+1}$ and define unit vectors of this space:

$$\pi_0 \triangleq (1, 0, \ldots, 0), \quad \pi_1 \triangleq (0, 1, \ldots, 0), \quad \ldots, \quad \pi_d \triangleq (0, 0, \ldots, 1, 0), \quad \pi_{d+1} \triangleq (0, 0, \ldots, 1).$$

Now, for any vector $z \in \mathcal{X}$, its ith coordinate can be written in the form $\pi_i z \triangleq \langle \pi_i, z \rangle$. Finally, denote by $|| \cdot ||_d$ the norm of a vector without the last coordinate: $||z||_d \triangleq ||z - (\pi_{d+1} z)\pi_{d+1}||$.

3.1. Required constants. Since the dynamics of the system is bounded, there exists a compact set $K_< \subset \mathbb{R}^d$ such that any solution y of system (1.1) with the initial condition from the unit ball centered at x_* stays within this set on the interval $[0; T]$. Increasing if necessary this set to $K_>$, we can assume that the distance of any solution of system (1.1) (with the initial condition x_*) from the boundary of $K_>$ is greater than 1 at any time from $[0; T]$. Now we can find a sufficiently large number $L \geq 2$ such that the functions $f|_{[0;T] \times K_> \times \mathbb{U} \times \mathbb{V}}$, $\sigma_0|_{K_>}$, $\sigma_1|_{[0;T] \times K_>}$, and $\sigma_2|_{[0;T] \times K_>}$ are Lipschitz with the constant L. We can also assume that the norm of $f|_{[0;T] \times K_> \times \mathbb{U} \times \mathbb{V}}$ and the numbers φ_+ and ψ_+ are not greater than $L - 1$. Recall that, by our assumptions, the functions σ_0, σ_1, and σ_2 do not exceed 1 in absolute value; for definiteness, we continue the mappings $\sigma_1|_{[0;T] \times \mathbb{R}^d}$ and $\sigma_2|_{[0;T] \times \mathbb{R}^d}$ to $[T; \infty) \times \mathbb{R}^d$ preserving both their Lipschitz constant and their norm in the uniform metric; this is possible by [28, Theorem 9.58].

Fix a smooth monotonically nonincreasing 2-Lipschitz scalar function $a : \mathbb{R} \to [0; 1]$ such that $a(0) = 1$, $a(1) = 0$, and the function $r \mapsto a(r - 1/2) - 1/2$ is odd. We also define $a_K(x) \triangleq 1 - a(\text{dist}(x; \mathbb{R}^d \setminus K_>))$ for all $x \in \mathbb{R}^d$.

Fix some positive number $h < \min(1, e^{-2}/T, 56L^2 T e^{-12LT}/d)$ such that $e^{-T/8h} \leq hT$ and $1/h \in \mathbb{N}$. Define

$$\gamma \triangleq \sqrt{-2hT \ln(hT)}, \quad \Gamma \triangleq T + \gamma = T + \sqrt{-2hT \ln(hT)}.$$

We have the inequalities

$$-\ln(hT) \geq 2, \quad 2\sqrt{hT} \leq \gamma, \quad \gamma < T/2, \quad h^2 d/2 \leq 7L^2 e^{-12LT} \gamma^2. \tag{3.1}$$

3.2. Extended dynamics of the original system. Recall that, according to the formalization proposed in the preceding section, the termination of the game on the initiative of the first or the second player at a time θ consists in shifting the position of the system along the additional $(d+1)$th coordinate by the vector $(1 + \theta)\pi_{d+1}$ or $-(1 + \theta)\pi_{d+1}$, respectively. Now the extended space of the original dynamic system (1.1) coincides with the product $\mathbb{R}^{d+1} \times \{s = 0\}$, and the set $\mathbb{R}^{d+1} \times \{s \geq 1\}$ contains all possible terminal positions of the system if the game has been terminated on the initiative of the first player; similarly, the set $\mathbb{R}^{d+1} \times \{s \leq -1\}$ contains all terminal positions of the system if the game has been terminated on the initiative of the second player. Consequently, we extend the dynamics of (1.1) to $\mathcal{X} = \mathbb{R}_+ \times \mathbb{R}^{d+1}$.

For this, to each $w = (t, x, s) \in \mathcal{X}$, $(u, \varphi) \in \bar{\mathbb{U}}$, and $(v, \psi) \in \bar{\mathbb{V}}$, we assign a vector

$$\hat{f}(w, \bar{u}, \bar{v}) \triangleq (1, 1_{[0;T)}(t)a_K(x)f(t, x, u, v), 0). \tag{3.2}$$

It is easy to verify that the function $\hat{f}$ is bounded in the norm by the number L and is $3L$-Lipschitz in each of the domains $[0; T) \times \mathbb{R}^{d+1}$ and $[T; \infty) \times \mathbb{R}^{d+1}$. In addition, $\hat{f}(w, \bar{u}, \bar{v})$ coincides with $(1, f(t, x, u, v), 0)$ if $w = (t, x, s) \in [0; T) \times K_< \times \{0\}$, $(u, \varphi) \in \bar{\mathbb{U}}$, and $(v, \psi) \in \bar{\mathbb{V}}$. Moreover, outside the set $\mathcal{X}_< \triangleq \mathbb{R}_+ \times K_> \times [-2 - T; 2 + T]$, the function $\hat{f}$ is zero in all the coordinates except for the zero coordinate.

Consider a conflict-control system

$$\frac{dy(t)}{dt} = \hat{f}(y(t), \bar{u}(t), \bar{v}(t)), \quad y(0) = (0, x_*, 0) \text{ for all } t \in \mathbb{R}_+, \quad y \in \mathcal{X}, \quad \bar{u}(t) \in \bar{\mathbb{U}}, \quad \bar{v}(t) \in \bar{\mathbb{V}}. \tag{3.3}$$

If the dynamics in this system is defined by (3.2) and the initial conditions are in the set $\mathcal{X}_<$, then the trajectories of (3.3) stay inside this set; in addition, for the original initial conditions and any controls of the players, the corresponding coordinates of the trajectories of (3.3) coincide with the coordinates of the trajectories of (1.1), and the time coincides with the zero coordinate of the solution of (3.3). Note also that, after the terminal time of the game, i.e., for $\pi_0 y(t) = t \geq T$ or for $|\pi_{d+1} y(t)| \geq 1$, only the zero coordinate of the solution of system (3.3), which corresponds to time, changes. Finally, by the boundedness of the right-hand side for any solution y with dynamics (3.2), we have

$$\|y(t) - y(t')\|_d \leq L(t - t') \quad \forall t' > t \geq 0. \tag{3.4}$$

In order to describe the jump in the last coordinate, we will require that the conditional probability of termination of the game on the initiative of the first (second) player during a small interval $[t; t + \Delta t)$ is $a(|\pi_{d+1}y|)1_{[0;T)}(\pi_0 y)\varphi(t)\Delta t + o(\Delta t)$ $(a(|\pi_{d+1}y|)1_{[0;T)}(\pi_0 y)\psi(t)\Delta t + o(\Delta t)$, respectively). Then the times of positive and negative jumps in the last coordinate, which are the stopping times θ_1 and θ_2, are described by the rules

$$\theta_1 \triangleq \inf\{t \mid \pi_{d+1}y(t) > \pi_{d+1}y(0)\} \cup \{T\}, \quad \theta_2 \triangleq \inf\{t \mid \pi_{d+1}y(t) < \pi_{d+1}y(0)\} \cup \{T\},$$

which in the case of the original initial conditions correspond to the rules (1.5) and (1.6) with the zero intensity of the opponent. In the general case, in view of the equality $t = \pi_0 y(t)$, the nondeterministic transition of the system corresponding to an early termination of the game and, hence, the value $\theta_{\min} = \min(\theta_1, \theta_2) < T$, can be described by the Lévy measure

$$\hat{\eta}(y, \bar{u}, \bar{v}; A) \triangleq a(|\pi_{d+1}y|)1_{[0;T)}(\pi_0 y)(\varphi\delta_{(1+\pi_0 y)\pi_{d+1}} + \psi\delta_{-(1+\pi_0 y)\pi_{d+1}})(A) \quad \forall A \in \mathcal{B}(\mathcal{X}),$$

depending on z, $\bar{u} = (u, \varphi)$, and $\bar{v} = (v, \psi)$.

Let us show that the dynamics extended in this way is admissible. Once again, for $\mathcal{T} \triangleq \mathbb{R}_+$, we assign to each pair of controls $\bar{u} = (u, \varphi)$, $\bar{v} = (v, \psi)$ the Dirac measures $\hat{\mathbb{P}}_{(\bar{u}, \bar{v})}[\varrho]$ supported by solutions y with dynamics (3.2) generated by these controls on the entire semiaxis. For the intensities (φ, ψ), we find the stopping times θ_1 and θ_2 and, hence, $\theta_{\min}$. As in the preceding section, we define for θ_1, θ_2 a random variable s and assign to each pair $(x(\cdot), s) \in D(\mathbb{R}_+, \mathbb{X}) \times ([-1 - T; -1] \cup \{0\} \cup [1; 1 + T])$ a path $y_{(x(\cdot), s)}(\cdot) \in D(\mathbb{R}_+, \mathcal{X})$ by the rule (2.1). Once again, the push-forward of the measure $\hat{\mathbb{P}}_{(\bar{u}, \bar{v})} \otimes \mathbb{P}_{(\varphi, \psi)}$ by this mapping specifies a distribution $\hat{\mathbb{P}}^G_{(\bar{u}, \bar{v})}[\varrho]$ on

$D(\mathcal{T}, \mathcal{X})$; then, the distribution $\hat{\mathbb{P}}_{\varrho}^{\text{all}}$ on $D(\mathcal{T}, \bar{\mathbb{U}} \times \bar{\mathbb{V}}) \times D(\mathcal{T}, \mathcal{X})$ is uniquely recovered by an analog of (2.2), i.e., by the rule

$$\hat{\mathbb{P}}_{\varrho}^{\text{all}}(B \times D) \triangleq \hat{\mathbb{P}}_{(\bar{u}, \bar{v})}^{G}[\varrho](D) 1_B(\bar{u}, \bar{v}) \quad \forall B \in \mathcal{B}(D(\mathcal{T}, \bar{\mathbb{U}} \times \bar{\mathbb{V}})), \quad D \in \mathcal{B}(D(\mathcal{T}, \mathcal{X})).$$

Thus, in the extended game, for an arbitrary pair of càdlàg controls $\bar{u} = (u, \varphi)$, $\bar{v} = (v, \psi)$, we have constructed their realization; hence, the extended dynamics is admissible. Therefore, as earlier, we have the admissibility of the program and stepwise strategies.

3.3. Payoff function in an integral form. We define a function $W : \mathcal{X} \to \mathbb{R}$ by the following rule: for any $z = (t, x, s) \in \mathcal{X}$,

$$W(z) \triangleq a_K(x)\big(a(|s|)\sigma_0(x) + a(1 - s)\sigma_1(s - 1, x) + a(s - 1)\sigma_2(1 - s, x)\big).$$

Note that the function W does not exceed 1 in absolute value, is independent of the zero coordinate, and is $3L$-Lipschitz because $L \geq 2$.

To any path $y \in D(\mathbb{R}_+, \mathcal{X})$, we assign a payoff

$$\bar{\sigma}(y) \triangleq \int_0^{\infty} h e^{h(T + 2\gamma - t)} 1_{[T + 2\gamma; \infty)}(\pi_0 y(t)) W(y(t)) \, dt. \tag{3.5}$$

Let us establish the following representation of the expected value of the payoff function: for any random process with distribution $\hat{\mathbb{P}}_{\bar{u}, \bar{v}}^{G}[\delta_{(0, x_0, 0)}] \in \mathcal{P}(D(\mathbb{R}_+, \mathcal{X}))$ of trajectories $t \mapsto y(t) = (t, x(t), s(t))$ with initial conditions $x(0) \in K_<$, $s(0) = 0$,

$$\mathbb{E}J(x, \bar{u}, \bar{v}) \triangleq \hat{\mathbb{E}}_{\bar{u}, \bar{v}}^{G} \bar{\sigma}(y) = \hat{\mathbb{E}}_{\bar{u}, \bar{v}}^{G} \int_0^{\infty} h e^{h(T + 2\gamma - t)} 1_{[T + 2\gamma; \infty)}(\pi_0 y(t)) W(y(t)) \, dt. \tag{3.6}$$

Indeed, consider an arbitrary trajectory $t \mapsto y(t) = (t, x(t), s(t))$ of the process. This trajectory is defined on the entire semiaxis, but, starting from some stopping time τ_T, all its coordinate except for the zero coordinate do not change anymore. This happens either upon the expiry of the time of the game, and then $\tau_T = T$ and $W(y(t)) = \sigma_0(x(T))$ for all $t \geq \tau_T$; or on the initiative of the first player, and then $\tau_T = \theta_1$ is the first time with $s(t) > 0$, and $z(\tau_T) = (\tau_T, x(\tau_T), \tau_T + 1)$ and $W(y(t)) = \sigma_1(\tau_T, x(\tau_T))$ for all $t \geq \tau_T$; or on the initiative of the second player, and then $\tau_T = \theta_2$ is the first time with $s(t) < 0$, and $W(y(t)) = \sigma_2(\tau_T, x(\tau_T))$ for all $t \geq \tau_T$. Since $e^{h(T + 2\gamma - t)} 1_{[T + 2\gamma; \infty)}(\pi_0 y(t)) W(y(t)) \equiv 0$ until the time $T + 2\gamma$ along the same trajectory $y(\cdot)$, it follows that $\bar{\sigma}(y)$ is exactly $\sigma_0(x(T))$, $\sigma_1(\theta_1, x(\theta_1))$, and $\sigma_2(\theta_2, x(\theta_2))$ in each of these cases. Thus, equality (3.6) is established.

Since the expected values of the payoff functions coincide in view of (3.6), the constructed game with the extended dynamics and the initial condition $(0, x_*, 0)$ coincides with the original game.

4. AUXILIARY MARKOV GAME

We now start to construct an approximating Markov game depending on the step h. Consider an integer lattice $\mathcal{Z} \triangleq h\mathbb{Z}^{d+2} \cap \mathcal{X}$ with step h that has dimension $d + 2$ and nonnegative zero coordinate. For the set of states, we take $\mathcal{Z}_< \triangleq \mathcal{Z} \cap \mathcal{X}_<$. The space of paths is the Skorokhod space $D(\mathbb{R}_+, \mathcal{Z}_<)$ equipped with the σ-algebra of Borel sets and the canonical filtration.

4.1. Randomized strategies in the Markov game.   We take randomized strategies as strategies of the players.

A *randomized strategy* $\bar{\mu}$ *of the first player* is a pair (μ, φ_μ) of mappings taking each $w \in \mathcal{Z}_<$ to a time-dependent probability measure $\mu[w]$ on Borel subsets of $\mathbb{U}$ and a Borel measurable function $\varphi_\mu[w]$ with values in $[\varphi_-; \varphi_+]$. A *randomized strategy* $\bar{\nu}$ *of the second player* is a pair (ν, ψ_ν) of mappings taking each $w \in \mathcal{Z}_<$ to a time-dependent probability measure $\nu[w]$ on Borel subsets of $\mathbb{V}$ and a Borel measurable function $\varphi_\nu[w]$ with values in $[\psi_-; \psi_+]$. Denote the sets of all randomized strategies of the first and second players by $\breve{\mathbb{U}}_\varpi$ and $\breve{\mathbb{V}}_\varpi$, respectively. If an element from $\breve{\mathbb{U}}_\varpi$ (or $\breve{\mathbb{V}}_\varpi$) is independent of time, we will call it a *stationary strategy*. Denote the sets of all stationary strategies of the first and second players by $\breve{\mathbb{U}}_\varsigma$ and $\breve{\mathbb{V}}_\varsigma$, respectively, and denote by $\mathbb{U}_\varsigma$ and $\mathbb{V}_\varsigma$ their projections, i.e., the mappings $w \mapsto \mu[w]$ and $w \mapsto \nu[w]$, respectively.

Define a mapping $\breve{f} : \mathcal{Z}_< \times \breve{\mathbb{U}}_\varsigma \times \breve{\mathbb{V}}_\varsigma \to \mathcal{Z}_<$ as follows: for any pair of strategies $(\bar{\mu}, \bar{\nu}) \in \breve{\mathbb{U}}_\varsigma \times \breve{\mathbb{V}}_\varsigma$ and a point $w \in \mathcal{Z}_<$,

$$\breve{f}(w, \bar{\mu}, \bar{\nu}) \triangleq \int_U \int_V \hat{f}(w, u, \varphi_\mu, v, \psi_\nu) \, \mu[w](du)\nu[w](dv).$$

We also define for any $i \in [0 : d+1]$ the projections $\pi_i^+ \breve{f}$ and $\pi_i^- \breve{f}$ of the mapping $\breve{f}$ by the rules

$$\pi_i^+ \breve{f}(w, \bar{\mu}, \bar{\nu}) \triangleq \int_U \int_V \max(0, \pi_i \hat{f}(w, u, \varphi_\mu[w], v, \psi_\nu[w])) \, \mu[w](du)\nu[w](dv),$$

$$\pi_i^- \breve{f}(w, \bar{\mu}, \bar{\nu}) \triangleq \int_U \int_V \max(0, -\pi_i \hat{f}(w, u, \varphi_\mu[w], v, \psi_\nu[w])) \, \mu[w](du)\nu[w](dv).$$

Note that we will substitute into $\breve{f}$ (and into other objects) values of time-dependent randomized strategies $\bar{\mu}$ and $\bar{\nu}$ in place of stationary randomized strategies, but in this case the argument responsible for time will be omitted for simplicity of notation.

4.2. Dynamics of the Markov game.   For all points $w \in \mathcal{Z}_<$ and randomized strategies $\bar{\mu}$ and $\bar{\nu}$ of the players, we define the Itô measure $\breve{\eta}(w, \bar{\mu}, \bar{\nu}; \cdot)$: for all Borel subsets $A \subset \mathcal{Z}_<$,

$$\breve{\eta}(w, \bar{\mu}, \bar{\nu}; A) \triangleq \frac{1}{h} \sum_{i=0}^{d} \left[\pi_i^+ \breve{f}(w, \bar{\mu}, \bar{\nu})\delta_{h\pi_i}(A) + \pi_i^- \breve{f}(w, \bar{\mu}, \bar{\nu})\delta_{-h\pi_i}(A) \right]$$

$$+ 1_{[0;\Gamma)}(\pi_0 w)a(|\pi_{d+1}w|)\left[\bar{\varphi}[w]\delta_{(\pi_0 w+1)\pi_{d+1}}(A) + \bar{\psi}[w]\delta_{-(\pi_0 w+1)\pi_{d+1}}(A)\right].$$

For this measure, one can easily verify the following equalities, which will be needed below in the construction of the Lévy–Khinchin generator: for all $w \in \mathcal{Z}_<$,

$$\int_{\mathcal{X}} [y - (\pi_{d+1}y)\pi_{d+1}] \, \breve{\eta}(w, \bar{\mu}, \bar{\nu}; dy) = \int_{\mathcal{X}} \sum_{i=0}^{d} \pi_i y \, \breve{\eta}(w, \bar{\mu}, \bar{\nu}; dy) = \breve{f}(w, \bar{\mu}, \bar{\nu}), \tag{4.1}$$

$$\int_{\mathcal{X}} \|y\|_d^2 \, \breve{\eta}(w, \bar{\mu}, \bar{\nu}; dy) = h \sum_{i=0}^{d} \left(\pi_i^+ \breve{f}(w, \bar{\mu}, \bar{\nu}) + \pi_i^- \breve{f}(w, \bar{\mu}, \bar{\nu}) \right). \tag{4.2}$$

The introduced measure $\check{\eta}$ corresponds to a continuous-time Markov chain given by the Kolmogorov matrix $(\bar{Q}_{wy}(\bar{\mu}, \bar{\nu}))_{w,y \in \mathcal{Z}_<}$: for all time-dependent randomized strategies $\bar{\mu}, \bar{\nu}$,

$$\bar{Q}_{wy}(\bar{\mu}, \bar{\nu}) \triangleq \begin{cases} 1/h, & y = w + h\pi_0, \\ \pi_i^+ \check{f}(w, \bar{\mu}, \bar{\nu})/h, & y = w + h\pi_i, i \in [1:d], \\ \pi_i^- \check{f}(w, \bar{\mu}, \bar{\nu})/h, & y = w - h\pi_i, i \in [1:d], \\ 1_{[0;\Gamma)}(\pi_0 w)a(|\pi_{d+1}w|)\varphi_\mu[w], & y = w + (1 + \pi_0 w)\pi_{d+1}, \\ 1_{[0;\Gamma)}(\pi_0 w)a(|\pi_{d+1}w|)\psi_\nu[w], & y = w - (1 + \pi_0 w)\pi_{d+1}, \\ -\sum_{i=0}^{d}(\pi_i^+ \check{f}(w, \bar{\mu}, \bar{\nu}) + \pi_i^- \check{f}(w, \bar{\mu}, \bar{\nu}))/h \\ \quad - 1_{[0;\Gamma)}(\pi_0 w)a(|\pi_{d+1}w|)(\varphi_\mu[w] + \psi_\nu[w]), & y = w, \\ 0, & \text{otherwise.} \end{cases}$$

Note that, for each $w \in \mathcal{Z}_<$, both the measure $\check{\eta}(w, \bar{\mu}, \bar{\nu}; \cdot)$ on the power set of $\mathcal{Z}_<$ and the operator corresponding to $\bar{Q}_{wy}$ linearly depend on $\bar{\mu}$ and $\bar{\nu}$. Now, for any mapping $q : \mathcal{Z}_< \to \mathbb{R}$, from the minimax theorem we have the equality

$$\min_{\bar{\mu} \in \check{\mathbb{U}}_\varsigma} \max_{\bar{\nu} \in \check{\mathbb{V}}_\varsigma} \sum_{y \in \mathcal{Z}_<} \bar{Q}_{wy}(\bar{\mu}, \bar{\nu})q(y) = \max_{\bar{\nu} \in \check{\mathbb{V}}_\varsigma} \min_{\bar{\mu} \in \check{\mathbb{U}}_\varsigma} \sum_{y \in \mathcal{Z}_<} \bar{Q}_{wy}(\bar{\mu}, \bar{\nu})q(y). \tag{4.3}$$

For any mapping $q : \mathcal{Z}_< \to \mathbb{R}$, we introduce (and fix) stationary strategies $\bar{\mu}^{\downarrow q}$ and $\bar{\nu}^{\uparrow q}$ implementing this minimax at each point.

4.3. Objective function and the value of the Markov game. Since all matrices $\bar{Q}_{wy}$, similarly to the lengths of the jumps, are uniformly bounded, all assumptions of [14, Remark 4.2(b); 26, Theorem 5.1; 27, Assumptions 2.1, 2.2] are satisfied. Now, in view of [14, Proposition 3.1(a); 27, Sect. 2.3], for each pair of time-dependent randomized strategies $(\bar{\mu}, \bar{\nu}) \in \check{\mathbb{U}}_\varpi \times \check{\mathbb{V}}_\varpi$ and initial conditions ϱ (a probability over $\mathcal{Z}_<$), there are a process $(\check{Y}(t))_{t \geq 0}$ generated by them and, hence, its distribution $\check{\mathbb{P}}_{\bar{\mu}, \bar{\nu}}^G[\varrho] \in \mathcal{P}(D(\mathbb{R}_+, \mathcal{Z}_<))$. Then each randomized strategy is admissible, and we can assign to it a process $(\check{Y}(t))_{t \geq 0}$ with distribution

$$\check{\mathbb{P}}_\varrho^{\text{all}} \triangleq \delta_{\bar{\mu}, \bar{\nu}} \otimes \check{\mathbb{P}}_{\bar{\mu}, \bar{\nu}}^G[\varrho] \in \mathcal{P}(D(\mathbb{R}_+, \check{\mathbb{U}}_\varpi \times \check{\mathbb{V}}_\varpi) \times D(\mathbb{R}_+, \mathcal{Z}_<)).$$

Once again, to each element of the path space $z \in D(\mathbb{R}_+, \mathcal{Z}_<) \subset D(\mathbb{R}_+, \mathcal{X}_<)$, we assign the payoff $\bar{\sigma}(z)$ by the rule (3.5). For each initial position $z_0 \in \mathcal{Z}_<$, the players at time 0 can provide by the choice of the randomized strategy one of the following values depending on who of them is informationally discriminated:

$$\check{\mathcal{V}}^+(z_0) = \inf_{\bar{\mu} \in \check{\mathbb{U}}_\varpi} \sup_{\bar{\nu} \in \check{\mathbb{V}}_\varpi} \check{\mathbb{E}}_{z_0}^{\text{all}} \int_0^\infty h e^{h(T+2\gamma-t)} 1_{[T+2\gamma;\infty)}(\pi_0 z(t)) W(z(t))\, dt,$$

$$\check{\mathcal{V}}^-(z_0) = \sup_{\bar{\nu} \in \check{\mathbb{V}}_\varpi} \inf_{\bar{\mu} \in \check{\mathbb{U}}_\varpi} \check{\mathbb{E}}_{z_0}^{\text{all}} \int_0^\infty h e^{h(T+2\gamma-t)} 1_{[T+2\gamma;\infty)}(\pi_0 z(t)) W(z(t))\, dt.$$

Moreover, as follows from [14, Theorem 5.1; 26, Theorem 2], the system of equations (see [14, (5.4); 26, (11)])

$$\inf_{\bar{\mu}[w]} \sup_{\bar{\nu}[w]} \sum_{y \in \mathcal{Z}_<} \bar{Q}_{wy}(\bar{\mu}, \bar{\nu})\check{\mathcal{V}}(y) = h(\check{\mathcal{V}}(w) - e^{h(T+2\gamma)} 1_{[T+2\gamma;\infty)}(\pi_0 w) W(w))$$

$$= \sup_{\bar{\nu}[w]} \inf_{\bar{\mu}[w]} \sum_{y \in \mathcal{Z}_<} \bar{Q}_{wy}(\bar{\mu}, \bar{\nu}) \check{\mathcal{V}}(y), \quad w \in \mathcal{Z}_<,$$

has a unique solution in view of equality (4.3), and this solution coincides with $\check{\mathcal{V}} \triangleq \check{\mathcal{V}}^- = \check{\mathcal{V}}^+$. Note that, by the construction of the matrix $\bar{Q}$, there are at most $2d + 4$ nonzero terms in each equation of the system. Moreover, in the case $\pi_0 y \geq \Gamma$, the value $\check{\mathcal{V}}(y)$ is calculated explicitly and equals $e^{h(T + 2\gamma - \pi_0 w)} W(w)$. Thus, this system reduces to a system of a finite number of equations.

4.4. Optimal stationary strategies. It follows from [26, (6); 14, (5.5)–(5.7)] that any stationary strategies $\bar{\mu}^{\mathrm{opt}}$ and $\bar{\nu}^{\mathrm{opt}}$ that solve for all $w \in \mathcal{Z}_<$ the system

$$h(\check{\mathcal{V}}(w) - e^{h(T + 2\gamma)} 1_{[T + 2\gamma; \infty)}(\pi_0 w) W(w)) = \max_{\bar{\nu} \in \check{\mathbb{V}}_\varsigma} \sum_{y \in \mathcal{Z}_<} \bar{Q}_{wy}(\bar{\mu}^{\mathrm{opt}}, \bar{\nu}) \check{\mathcal{V}}(y),$$

$$h(\check{\mathcal{V}}(w) - e^{h(T + 2\gamma)} 1_{[T + 2\gamma; \infty)}(\pi_0 w) W(w)) = \min_{\bar{\mu} \in \check{\mathbb{U}}_\varsigma} \sum_{y \in \mathcal{Z}_<} \bar{Q}_{wy}(\bar{\mu}, \bar{\nu}^{\mathrm{opt}}) \check{\mathcal{V}}(y)$$

are optimal in this problem. Thus, now we can take $\bar{\mu}^{\mathrm{opt}} \equiv \bar{\mu}^{\downarrow \check{\mathcal{V}}}$ and $\bar{\nu}^{\mathrm{opt}} \equiv \bar{\nu}^{\uparrow \check{\mathcal{V}}}$ as optimal strategies; moreover, we can choose them so that they take values among Dirac measures over $\hat{\mathbb{U}}$ and $\hat{\mathbb{V}}$, respectively. In this case, the second component in $\bar{\mu}^{\mathrm{opt}} = (u^{\mathrm{opt}}, \varphi^{\mathrm{opt}})$ is uniquely recovered by the rule

$$\varphi^{\mathrm{opt}}[w] = \begin{cases} \varphi_-, & \check{\mathcal{V}}(w + \pi_{d+1}(1 + \pi_0 w)) > \check{\mathcal{V}}(w), \\ \varphi_+, & \check{\mathcal{V}}(w + \pi_{d+1}(1 + \pi_0 w)) < \check{\mathcal{V}}(w) \end{cases}$$

if $|\pi_{d+1} w| < 1$, $\pi_0 w \in [0; \Gamma)$, and $\check{\mathcal{V}}(w + \pi_{d+1}(1 + \pi_0 w)) \neq \check{\mathcal{V}}(w)$ and can be arbitrary if at least one of these conditions does not hold. A similar property holds for $\bar{\nu}^{\mathrm{opt}} = (v^{\mathrm{opt}}, \psi^{\mathrm{opt}})$.

4.5. Estimates for the zero coordinate, which is independent of the players' actions. Fix some $z_0 \in \mathcal{Z}_<$ with $\pi_0 z_0 = 0$ and the Dirac measure ϱ supported at this point. Then, for arbitrary, possibly, time-dependent randomized strategies of the players $\bar{\mu}$ and $\bar{\nu}$, there exists a distribution $\check{\mathbb{P}}_{\delta_{z_0}}^{\mathrm{all}}$ of the process $(\check{Y}(t))_{t \geq 0}$ that linearly depends on this Dirac measure. For simplicity of notation, we assume until the end of this section that $\check{\mathbb{P}} \triangleq \check{\mathbb{P}}_{\delta_{z_0}}^{\mathrm{all}}$.

Note that, regardless of the players' actions, the process $\pi_0 \check{Y}(t)$ has independent increments, and the random variables $\pi_0(\check{Y}(t_0 + t) - \check{Y}(t))/h$ have Poisson distribution with parameter t/h; in particular, their expected value and dispersion are t/h. Then $\check{\mathbb{E}} |\pi_0 \check{Y}(t) - t|^2 = ht$; in particular, for $t = T + 2\gamma$,

$$\check{\mathbb{E}} |\pi_0 \check{Y}(T + 2\gamma) - (T + 2\gamma)|^2 \leq (T + 2\gamma) h \overset{(3.1)}{\leq} 2hT.$$

Finally, since $\pi_0 \check{Y}(t) - t$ is also a martingale, it follows by Doob's inequality that

$$\check{\mathbb{E}} \max_{t \in [0; T + 2\gamma]} |\pi_0 \check{Y}(t) - t|^2 \leq 4 \check{\mathbb{E}} |\pi_0 \check{Y}(T + 2\gamma) - (T + 2\gamma)|^2 \leq 8hT \overset{(3.1)}{\leq} 2\gamma^2. \tag{4.4}$$

Now, let us consider the random variable $\xi = \pi_0 \check{Y}(\Gamma)/h$, which has Poisson distribution with parameter Γ/h. According to the Chernoff bound (a variant of Markov's inequality), for any negative δ, we have

$$\check{\mathbb{P}}(\xi \leq \tfrac{T}{h}) \leq \frac{\check{\mathbb{E}} e^{\delta \xi}}{e^{\delta \frac{T}{h}}} = e^{\frac{\Gamma}{h}(e^\delta - 1)} e^{-\delta \frac{T}{h}}.$$

The equality sign here is due to the substitution of the value of the probability generating function (at the point δ) for the Poisson distribution with parameter Γ/h. Now, using the inequalities $e^\delta - 1 \leq \delta + \delta^2/2$ and $\delta < 0$ and substituting $\delta = -\gamma/\Gamma$, $T = \Gamma - \gamma$, and $\gamma = \sqrt{-2hT\ln(hT)}$, we finally obtain

$$\check{\mathbb{P}}(\pi_0 \check{Y}(\Gamma) \leq T) \leq e^{\frac{\Gamma}{h}(\frac{\gamma^2}{2\Gamma^2} - \frac{\gamma}{\Gamma}) + \frac{\gamma T}{h\Gamma}} = e^{\frac{\gamma}{h}(\frac{\gamma}{2\Gamma} - 1 + \frac{\Gamma - \gamma}{\Gamma})} = e^{-\frac{\gamma^2}{2h\Gamma}} = e^{\frac{T\ln(hT)}{\Gamma}} = e^{\frac{\ln(hT)}{1+\gamma/T}}.$$

Similarly, for $\check{\mathbb{P}}(\pi_0 \check{Y}(\Gamma + \gamma) \leq \Gamma)$ and $\delta = -\gamma/(\Gamma + \gamma)$, we have

$$\check{\mathbb{P}}(\pi_0 \check{Y}(\Gamma + \gamma) \leq \Gamma) \leq e^{\frac{\Gamma + \gamma}{h}(\frac{\gamma^2}{2(\Gamma+\gamma)^2} - \frac{\gamma}{\Gamma+\gamma}) + \frac{\gamma\Gamma}{h(\Gamma+\gamma)}} = e^{\frac{\gamma}{h}(\frac{\gamma}{2(\Gamma+\gamma)} - 1 + \frac{\Gamma}{\Gamma+\gamma})} = e^{\frac{\ln(hT)}{1+2\gamma/T}}.$$

Once again, using $\gamma = \sqrt{-2hT\ln(hT)} < T/2$ and $2\sqrt{hT} \leq \gamma$ (see (3.1)), we obtain

$$\check{\mathbb{P}}(\pi_0 \check{Y}(\Gamma) \leq T) + \check{\mathbb{P}}(\pi_0 \check{Y}(\Gamma + \gamma) \leq \Gamma) \leq e^{\frac{\ln(hT)}{1+\gamma/T}} + e^{\frac{\ln(hT)}{1+2\gamma/T}} < 2\sqrt{hT} \leq \gamma. \tag{4.5}$$

4.6. Estimates for the trajectories in the Markov game. In this section we derive an analog of the estimate [2, Lemma 12] on required classes of strategies.

Fix some pair of randomized strategies $\bar{\mu} = (\mu, \varphi_\mu)$, $\bar{\nu} = (\nu, \psi_\nu)$ and, hence, the distribution $\check{\mathbb{P}} \triangleq \check{\mathbb{P}}_\varrho^{\text{all}}$ and the Itô measure $\eta(z; \cdot) \triangleq \check{\eta}(z, \bar{\mu}, \bar{\nu}; \cdot)$. This Itô measure corresponds to a Lévy–Khinchin generator (see [16, (5.1); 27, (2.14)]), which takes each function $g \in C_c^2(\mathcal{Z}_<)$ to a mapping $x \mapsto \check{\Lambda}g(x)$ by the rule

$$\check{\Lambda}g(x) = \int_{\mathbb{X}} [g(x+y) - g(x)]\eta(x; dy) \quad \forall x \in \mathcal{Z}_<. \tag{4.6}$$

Note that all the numbers $\eta(x; \mathbb{X})$ are bounded uniformly in x, the supports of the measures $\eta(x; \cdot)$ are uniformly bounded, and $\eta(x; G) \equiv 0$ for some neighborhood G of zero. Thus, the conditions [27, Assumptions 1, 2] are satisfied, and we have Dynkin's formula [27, Proposition 2.3]:

$$\check{\mathbb{E}}\left(g(\hat{Y}(t)) - \int_r^t \check{\Lambda}g(\check{Y}(s))\, ds \,\Big|\, \check{Y}(r)\right) = \check{\mathbb{E}}g(\check{Y}(r)) \tag{4.7}$$

for all $g \in C_c^2(\mathcal{Z}_<)$ and all nonnegative t and r ($t \geq r$); i.e., the process

$$g(\check{Y}(t)) - \int_r^t \check{\Lambda}g(\check{Y}(s))\, ds \tag{4.8}$$

is a martingale. By the uniform boundedness of the supports of the Itô measure, Dynkin's formula (4.7) also holds for all $g \in C^2(\mathcal{X})$, since these functions can always be assumed to be zero outside a compact neighborhood of the union of all supports.

Now, let strategies $\bar{\mu} = (\mu, \varphi_\mu)$ and $\bar{\nu} = (\nu, \psi_\nu)$ be stepwise with some partition $(t_k)_{k\in\mathbb{N}}$. Since such a strategy is admissible, we once again we have a distribution $\check{\mathbb{P}}_\varrho^{\text{all}}$. Fix some time $t' \triangleq t_{k-1}$ from the partition and an element $w' \in \mathcal{X}$, which can depend on $\check{Y}(t')$. Now, on the interval $[t'; t'') \triangleq [t_{k-1}; t_k)$, the strategy $\bar{\mu} = (\mu, \varphi_\mu)$, $\bar{\nu} = (\nu, \psi_\nu)$, as a program strategy with a

possible dependence on $\check{Y}(\cdot)|_{[0;t')}$, also has its Lévy–Khinchin generator (4.6), for which Dynkin's formula (4.7) holds.

Denote by $\mathbb{E}_{y'}^{\text{all}}$ the conditional expected value with respect to $y' = \check{Y}(t')$. Applying the constructed generator (4.6) to the function $g(w) \stackrel{\triangle}{=} ||w - w'||_d^2$ on the interval $[t'; t'')$, we conclude that, according to (4.8),

$$
g(\check{Y}(t)) + \int_{t_k}^t \left[\sum_{i=0}^d 2\pi_i \check{f}(\check{Y}(s), \bar{\mu}, \bar{\nu}) \, \pi_i(\check{Y}(s) - w') - \int_{\mathcal{Z}_<} ||y||_d^2 \, \eta(\check{Y}(t); dy) \right] ds \tag{4.9}
$$

is a martingale; therefore, for all $t \in [t'; t'')$,

$$
\mathbb{E}_{y'}^{\text{all}} g(\check{Y}(t)) - g(y') = \mathbb{E}_{y'}^{\text{all}} \int_{t'}^t \left[\int_{\mathcal{Z}_<} ||y||_d^2 \, \eta(\check{Y}(s); dy) - \sum_{i=0}^d 2\pi_i \check{f}(\check{Y}(s), \bar{\mu}, \bar{\nu}) \, \pi_i(\check{Y}(s) - w') \right] ds.
$$

Thus,

$$
\mathbb{E}_{y'}^{\text{all}} g(\check{Y}(t)) - g(y') \le \mathbb{E}_{y'}^{\text{all}} \int_{t'}^t \left[2||\check{f}(\check{Y}(s), \bar{\mu}, \bar{\nu})||_d \sqrt{g(\check{Y}(s))} + \int_{\mathcal{Z}_<} ||y||_d^2 \, \eta(\check{Y}(s); dy) \right] ds. \tag{4.10}
$$

Define $M \stackrel{\triangle}{=} \sqrt{hL\sqrt{d+1} + L^2}$. Estimate the doubled product in (4.10) by the sum of squares. By the boundedness of (4.1) and (4.2) in the norm by the numbers L and $hL\sqrt{d+1}$, we have

$$
\mathbb{E}_{y'}^{\text{all}} \big(g(\check{Y}(t)) - g(y') \big) \le M^2(t - t') + \int_{t'}^t \mathbb{E}_{y'}^{\text{all}} g(\check{Y}(s)) \, ds.
$$

Hence, by Gronwall's inequality, we have

$$
\mathbb{E}_{y'}^{\text{all}} g(\check{Y}(t)) \le (M^2 + g(y'))e^{t-t'} - M^2 \tag{4.11}
$$

for all $t \in [t'; t'')$; i.e.,

$$
e^{-t} \mathbb{E}_{y'}^{\text{all}} (g(\check{Y}(t)) + M^2) \le e^{-t'} (g(y') + M^2) = e^{-t'} (g(\check{Y}(t')) + M^2).
$$

Thus, on any interval $[t'; t'') \stackrel{\triangle}{=} [t_{k-1}; t_k)$ of the partition, the process

$$
e^{-t} (||\check{Y}(t) - w'||^2 + M^2) \tag{4.12}
$$

becomes a supermartingale. Since a stepwise strategy is switched only at times specified in advance, adding over all intervals of the partition, we obtain the supermartingale (4.12), as well as inequality (4.11), for all $t \ge t'$. Moreover, for $w' = y' = \check{Y}(t')$, (4.11) implies

$$
\mathbb{E}_{y'}^{\text{all}} ||\check{Y}(t) - y'||_d \le M\sqrt{e^{t-t'} - 1}. \tag{4.13}
$$

By $||\check{f}||_d \le L$ and the upper bound $hL\sqrt{d+1}$ for the norm of (4.2), we find from (4.10) that

$$
\mathbb{E}_{y'}^{\text{all}} g(\check{Y}(t)) - g(y') \stackrel{(4.10)}{\le} \mathbb{E}_{y'}^{\text{all}} \int_{t'}^t \left[2L\sqrt{g(\check{Y}(s))} + hL\sqrt{d+1} \right] ds
$$

for all t from the partition interval $[t'; t'']$. Substituting estimate (4.13) here for $w' = y' = \check{Y}(t')$, we obtain

$$\check{\mathbb{E}}^{\mathrm{all}}_{y'} ||\check{Y}(t) - y'||^2_d \overset{(4.10)}{\leq} \check{\mathbb{E}}^{\mathrm{all}}_{y'} \int_{t'}^{t} \left[2L\sqrt{g(\check{Y}(s))} + hL\sqrt{d+1} \right] ds$$

$$\overset{(4.13)}{\leq} hL\sqrt{d+1}|t - t'| + 2L \int_{t'}^{t} M\sqrt{e^{s-t'} - 1}\, ds$$

$$= hL\sqrt{d+1}|t - t'| + 4LM(\sqrt{e^{t-t'} - 1} - \arctan\sqrt{e^{t-t'} - 1})$$

$$\leq hL\sqrt{d+1}|t - t'| + 4LM(e^{t-t'} - 1)^{3/2}/3, \tag{4.14}$$

where we have used $\arctan x \geq x - x^3/3$ in the last inequality. Thus, by (4.14), we have

$$\check{\mathbb{E}}^{\mathrm{all}}_{\check{Y}(t')} ||\check{Y}(t) - \check{Y}(t')||^2_d \leq hL\sqrt{d+1}|t - t'| + 4L\sqrt{hL\sqrt{d+1} + L^2}(e^{t-t'} - 1)^{3/2}/3 \tag{4.15}$$

for any $t \in [t'; t'')$, where $[t'; t'')$ is an arbitrary interval from the partition chosen for the stepwise strategy.

Finally, since the functions g under consideration are independent of the last coordinate, the calculations and the resulting inequalities are independent of the choice of φ_μ and φ_ν.

5. SCHEME OF THE GUIDE AND A DOUBLE GAME

We take the product $\mathbb{X} \overset{\triangle}{=} \mathcal{X}_< \times \mathcal{Z}_<$ for the state space of a new game; then, the state of paths is the Skorokhod space $D(\mathbb{R}_+, \mathcal{X}_< \times \mathcal{Z}_<)$. The component $\mathcal{X}_<$ contains the original game; the game on the component $\mathcal{Z}_<$ (except for the last coordinate) is calculated by the first player, and its position is the guide that helps to construct a strategy in $\mathcal{X}_<$. The sets of control parameters of the players are $\Upsilon_1 \overset{\triangle}{=} \mathbb{U} \times \check{\mathbb{U}}_\varsigma \times \mathbb{V}_\varsigma$ and $\Upsilon_2 \overset{\triangle}{=} \bar{\mathbb{V}}$. Then all possible joint instantaneous controls compose the set $\Upsilon \overset{\triangle}{=} \mathbb{U} \times \check{\mathbb{U}}_\varsigma \times \mathbb{V}_\varsigma \times \bar{\mathbb{V}}$. In order to describe an admissible dynamics of the double game, we construct for each program strategy $\upsilon(\cdot) \overset{\triangle}{=} (u, \mu, \varphi, \nu, v, \psi)(\cdot) \in D(\mathbb{R}_+, \Upsilon)$ a distribution $\widetilde{\mathbb{P}}^G_\upsilon[\varrho]$ that depends linearly on the initial condition $\varrho \in \mathcal{P}(\mathbb{X})$. Assuming a linear dependence on ϱ, we can restrict ourselves to the case where ϱ is a Dirac measure, supported at some point $(x, z) \in \mathcal{X}_< \times \mathcal{Z}_<$.

Let us first construct auxiliary processes $(\hat{Y}_t)_{t\geq 0}$ and $(\check{Y}_t)_{t\geq 0}$. For this we consider on $\mathcal{X}_<$ the dynamics $\hat{\mathbb{P}}^G_{(u,0,v,\psi)}[\delta_x](\cdot)$ of a differential game with zero intensity of the first player in the last coordinate, and for $\mathcal{Z}_<$ we introduce the dynamics $\check{\mathbb{P}}^G_{(\mu,\varphi,\nu,0)}[\delta_z](\cdot)$ of a Markov game with zero intensity of the second player in the last coordinate. Both dynamics are admissible, and there exist the corresponding processes $(\hat{Y}_t)_{t\geq 0}$ and $(\check{Y}_t)_{t\geq 0}$. Since $\pi_{d+1}\hat{Y}$ and $\pi_{d+1}\check{Y}$ are governed by the intensities ψ and φ, respectively, we can define $\hat{\theta}_2$ and $\check{\theta}_1$ as the stopping times of the game by virtue of each players and, hence, the stopping time $\theta_{\min} \overset{\triangle}{=} \min(\check{\theta}_1, \hat{\theta}_2)$. Further, each pair of trajectories $(\hat{Y}_t, \check{Y}_t)_{t\geq 0}$ together with the stopping times $\hat{\theta}_2$ and $\check{\theta}_1$ corresponds to a trajectory $(\hat{w}, \check{w})(\cdot) \in D([0; \infty), \mathbb{R}^2)$ according to the following rules: $(\hat{w}, \check{w}) \equiv (0, 0)$ for $\theta_{\min} \geq T$;

$$(\hat{w}, \check{w}) \equiv (\pi_{d+1}\hat{Y}_t)\left(1, \frac{1 + \pi_0 \check{Y}_{\hat{\theta}_2}}{1 + \hat{\theta}_2}\right) \quad \text{and} \quad (\hat{w}, \check{w}) \equiv (\pi_{d+1}\check{Y}_t)\left(\frac{1 + \check{\theta}_1}{1 + \pi_0 \check{Y}_{\check{\theta}_1}}, 1\right)$$

for $\theta_{\min} = \hat{\theta}_2 < T$ and $\theta_{\min} = \check{\theta}_1 < T$, respectively.

Finally, we define the process $(Y'(t), Y''(t))_{t \geq 0}$ with values in $\mathcal{X}_< \times \mathcal{Z}_<$ as the push-forward of the process $(\hat{Y}_t, \check{Y}_t, (\hat{w}, \check{w})(t))_{t \geq 0}$ by the nonanticipating mapping

$$(\hat{y}_t, \check{y}_t, (\hat{w}, \check{w})(t)) \mapsto \left(y'(t) \triangleq (\pi_0 \hat{y}_t, \ldots, \pi_d \hat{y}_t, \hat{w}(t)), y''(t) \triangleq (\pi_0 \check{y}_t, \ldots, \pi_d \check{y}_t, \check{w}(t)) \right).$$

The resulting process specifies a dynamics for each joint program strategy from $D(\mathbb{R}_+, \Upsilon)$. Thus, an admissible dynamics is constructed.

Let us mention some properties of the process. First, the dynamics in each component with respect to all coordinates except for the last one completely coincides with the dynamics described earlier for the extended original game and the Markov game. Second, in each component, the lengths of the jumps and their intensities up to the time $\min\{t \,|\, \pi_0 Y''(t) \geq \Gamma\} \cup \{T\}$ obey the same Lévy measures as in the mentioned games. Third, before the time T, the last coordinates in each component, being zero at the initial moment, become nonzero only simultaneously. Thus, the nondeterminacy of the component $Y'(t)$ is completely described by the nondeterminacy of the last coordinate of the component $Y''(t)$; hence, for the computation of the last coordinate of the component $Y''(t)$, we do not need to know the intensity of the second player during the game, since this coordinate can change only simultaneously with the termination of the game. Finally, this means that the control of the second player can be given in the format of the control of the original game, which has been done by the choice of $\Upsilon_1 = \mathbb{U} \times \check{\mathbb{U}}_\varsigma \times \mathbb{V}_\varsigma$, $\Upsilon_2 = \bar{\mathbb{V}}$.

5.1. Aiming. Constructing the strategy of the first player and the response of the second player imagined by the first player, we will need to aim at and deviate from positions on each component. Let us introduce required functions.

First note that, in view of (1.2), for any vectors $x, w \in \mathcal{X}$, there exist controls $u^{\mathrm{to}\ w}(x) \in \mathbb{U}$ and $v^{\mathrm{from}\ w}(x) \in \mathbb{V}$ such that

$$\min_{u \in \mathbb{U}} \max_{v \in \mathbb{V}} \langle x - w, \hat{f}(x, u, \varphi, v, \psi) \rangle = \max_{v \in \mathbb{V}} \min_{u \in \mathbb{U}} \langle x - w, \hat{f}(x, u, \varphi, v, \psi) \rangle$$

$$= \min_{u \in \mathbb{U}} \langle x - w, \hat{f}(x, u, \varphi, v^{\mathrm{from}\ w}(x), \psi) \rangle = \max_{v \in \mathbb{V}} \langle x - w, \hat{f}(x, u^{\mathrm{to}\ w}(x), \varphi, v, \psi) \rangle$$

for all $(\varphi, \psi) \in [\varphi_-; \varphi_+] \times [\psi_-; \psi_+]$. For any vector $w \in \mathcal{X}$, we introduce a control of the first player by the rule

$$\mathcal{X} \ni x \mapsto \bar{u}^{\mathrm{to}\ w}(x) \triangleq (u^{\mathrm{to}\ w}(x), \varphi^{\mathrm{opt}}[w]) \in \bar{\mathbb{U}}.$$

For all $x, w \in [0; T] \times \mathbb{R}^{d+1}$, since $\hat{f}$ is Lipschitz, we have

$$\langle x - w, \hat{f}(x, u^{\mathrm{to}\ w}(x), \varphi, v^{\mathrm{from}\ w}(x), \psi) - \hat{f}(w, u^{\mathrm{to}\ x}(w), \varphi, v^{\mathrm{from}\ x}(w), \psi) \rangle$$

$$= \min_{u \in \mathbb{U}, v' \in \mathbb{V}} \max_{v \in \mathbb{V}, u' \in \mathbb{U}} \langle x - w, \hat{f}(x, u, \varphi, v, \psi) - \hat{f}(w, u', \varphi, v', \psi) \rangle \leq 3L \|x - w\|_d^2,$$

where the inequality holds due to the substitution $u = u'$, $v' = v$.

In order to define the behavior of the second player on the component $\mathcal{Z}_<$, we assign for each $x \in \mathcal{X}$ a stationary randomized strategy of the second player in the form

$$\mathcal{Z}_< \ni w \mapsto \bar{\nu}^{\mathrm{from}\ x}[w] = (\delta_{v^{\mathrm{from}\ x}(w)}, \psi^{\mathrm{opt}}).$$

Then, for any stationary randomized strategy $\bar{\mu}$ of the first player, we have

$$\langle x - w, \check{f}(w, \bar{\mu}, \bar{\nu}^{\text{from } x})\rangle \geq \min_{\bar{\mu} \in \check{\mathbb{U}}_{\varsigma}} \langle x - w, \check{f}(w, \bar{\mu}, \bar{\nu}^{\text{from } x})\rangle$$

$$\overset{(1.2)}{=} \min_{\bar{u} \in \bar{\mathbb{U}}} \max_{\bar{v} \in \bar{\mathbb{V}}} \langle x - w, \hat{f}(w, \bar{u}, \bar{v})\rangle = \langle x - w, \check{f}(w, \bar{u}^{\text{to } x}, \bar{\nu}^{\text{from } x})\rangle.$$

Therefore, for all $\bar{v}$, $\bar{\mu}$, φ, and ψ in the case $x, w \in [0; T) \times \mathbb{R}^{d+1}$ we have

$$\langle x - w, \hat{f}(x, u^{\text{to } x}(w), \varphi, \bar{v}) - \check{f}(w, \bar{\mu}, \bar{\nu}^{\text{from } w}(x))\rangle$$

$$\leq \langle x - w, \hat{f}(x, u^{\text{to } x}(w), \varphi, v^{\text{from } x}(w), \psi) - \check{f}(w, \bar{\mu}^{\text{to } w}, \bar{\nu}^{\text{from } w})\rangle \leq 3L||x - w||_d^2.$$

Fix some pair of $(x', w') \in \mathcal{X} \times \mathcal{Z}_<$. Consider a mapping $R_{x',w'} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ independent of the $(d+1)$th coordinates of its arguments and given by the rule

$$\hat{R}_{x',w'}(x) = \sum_{i=0}^{d} (\pi_i x' - \pi_i w')(\pi_i x - \pi_i x') \quad \forall x \in \mathcal{X}. \tag{5.1}$$

To obtain an analog of (4.9), to each stationary strategy $\bar{\mu}$ we assign the generator (4.6) corresponding to the joint stationary strategy $(\bar{\mu}, \bar{\nu}^{\text{from } x'}(w'))$ for the function $g(w) \overset{\triangle}{=} R_{w',x'}(w)$:

$$\check{\Lambda}[\bar{\mu}, \bar{\nu}^{\text{from } x'}(w')]R_{w',x'}(w) = \sum_{i=0}^{d} (\pi_i w' - \pi_i x')\pi_i \hat{f}(w, \bar{\mu}, \bar{\nu}^{\text{from } x'}) = -\langle x' - w', \hat{f}(w, \bar{\mu}, \bar{\nu}^{\text{from } x'})\rangle.$$

Now, for $x, w, x', w' \in [0; T] \times \mathbb{R}^{d+1}$, since $\hat{f}$ is Lipschitz, we have

$$\check{\Lambda}[\bar{\mu}, \bar{\nu}^{\text{from } x'}(w')]R_{w',x'}(w) \leq -\langle x' - w', \hat{f}(w', \bar{\mu}, \bar{\nu}^{\text{from } x'})\rangle + 3L||w - w'||_d ||x' - w'||_d. \tag{5.2}$$

For any solution $t \mapsto y(t)$ of system (3.3) generated by the joint control $(\bar{u}, \bar{v})$, we have

$$\frac{dg(y(t))}{dt} - \left\langle \frac{\partial g(x)}{\partial x}\Big|_{x=y(t)}, \hat{f}(y(t), \bar{u}(t), \bar{v}(t))\right\rangle = 0$$

for any smooth function g. Hence, similarly to (4.9), for $\hat{\Lambda}[\bar{u}, \bar{v}]g(x) \overset{\triangle}{=} \left\langle \frac{\partial g(x)}{\partial x}|_{x=y(t)}, \hat{f}(x, \bar{u}, \bar{v})\right\rangle$, the process

$$g(Y'(t)) - \int_{t_0}^{t} \hat{\Lambda}[\bar{u}, \bar{v}]g(Y'(s)) \, ds \tag{5.3}$$

for all smooth functions $g : \mathcal{X} \to \mathbb{R}$ independent of the last coordinate becomes a martingale; in particular, we have Dynkin's formula (4.7). In addition, we have

$$\hat{\Lambda}[\bar{u}^{\text{to } w'}(x'), \bar{v}]R_{x',w'}(x) = \sum_{i=0}^{d} (\pi_i x' - \pi_i w')\pi_i \hat{f}(x, \bar{u}^{\text{to } w'}(x'), \bar{v})$$

$$= \langle x' - w', f(x, \bar{u}^{\text{to } w'}(x'), \bar{v})\rangle \leq \langle x' - w', f(x', \bar{u}^{\text{to } w'}(x'), \bar{v})\rangle + 3L||x - x'||_d ||x' - w'||_d \tag{5.4}$$

in the case $x, w, x', w' \in [0; T] \times \mathbb{R}^{d+1}$. Adding (5.4) and (5.2), we obtain for all such x, w, x', and w' and any $\bar{\mu} \in \check{\mathbb{U}}_{\varsigma}$, $\bar{v} \in \bar{\mathbb{V}}$ the estimate

$$\hat{\Lambda}[\bar{u}^{\text{to } w'}(x'), \bar{v}]R_{x',w'}(x) + \check{\Lambda}[\bar{\mu}, \bar{\nu}^{\text{from } x'}]R_{w',x'}(w) \leq \frac{3L}{2}\left(4||x' - w'||_d^2 + ||x - x'||_d^2 + ||w - w'||_d^2\right). \tag{5.5}$$

5.2. Construction of a guide for the first player. Suppose that the first player learns the position of the trajectory $x(\cdot)$ of the real game at some times $(t_k)_{k \in \mathbb{N}}$ given in advance. We assume that this sequence of times is increasing, $t_0 = 0$, and some t_k equals T. For this sequence of times, we define the fineness of the partition $r \stackrel{\triangle}{=} \max_{k \in \mathbb{N}, t_k \leq T} |t_k - t_{k-1}|$ and the mapping $\mathbb{R}_+ \ni t \mapsto \tau_\Delta(t) \stackrel{\triangle}{=} \max\{t_i \,|\, t_i \leq t\}$.

Let us describe a stepwise strategy of the first player for the chosen partition $(t_k)_{k \in \mathbb{N}}$.

Let the first player apply an optimal stationary strategy μ^{opt} at each time on the component $\mathcal{Z}_<$; in particular, let the first player use $\varphi^{\mathrm{opt}}[z(t)]$ as their intensity of termination of the game. Then this player uses $\varphi^{\mathrm{opt}}[z(t)]$ on the first component as well, i.e., in the real game.

Let the first player try to minimize the distance from $z(\cdot)$ to $x(\cdot)$ on the first, real, component as follows: on the interval $[t_{k-1}; t_k)$, this player uses a control that is best (at the time t_{k-1}) for the shift of x to $z(t_{k-1})$. Thus, the first player applies the following rule: $\mathbb{R}_+ \ni t \mapsto (u^{\mathrm{to}\ z(\tau_\Delta(t))}(x(\tau_\Delta(t)), \mu^{\mathrm{opt}}[z], \varphi^{\mathrm{opt}}[z]) \in \Upsilon_1 \stackrel{\triangle}{=} \mathbb{U} \times \check{\mathbb{U}}_\varsigma$. Further, within the subgame calculated by the first player on $\mathcal{Z}_<$, let the second player as imagined by the first try to maximize the distance from $z(\cdot)$ to $x(\cdot)$, applying a stationary strategy that is worst for the shift of z towards $x(\tau_\Delta(t))$. Then, however, the strategy $\mathbb{R}_+ \ni t \mapsto \nu^{\mathrm{from}\ x(\tau_\Delta(t))}$, which depends nonanticipatingly on x and stationarily on z, is used as ν. Thus, the following strategy of the first player is specified: at each time,

$$\widetilde{v}_t[x, z] = (u^{\mathrm{to}\ z(\tau_\Delta(t))}(x(\tau_\Delta(t))), \mu^{\mathrm{opt}}[z], \varphi^{\mathrm{opt}}[z], \nu^{\mathrm{from}\ x(\tau_\Delta(t))}) \in \Upsilon_1. \tag{5.6}$$

Since this strategy is stepwise, it is admissible.

Now, having all the required constructions, we can proceed to proving the theorem.

6. PROOF OF THEOREM 1

Assume that the second player has all information about the strategy of the first player; in particular, the first player knows both the rule (5.6) and the position x and z at each moment t and, hence, knows the trajectory of the double game of the first player realized in the course of the game. Fix an arbitrary counterstrategy of the second player, which assigns in a nonanticipating way to each trajectory $(x, z)(\cdot)$ of the double game a distribution $\widetilde{\mathbb{Q}}^{\mathrm{II}}[\widetilde{\mathbb{P}}^{\mathrm{I}}]_{(x,z)(\cdot)}$ on the Skorokhod space $D(\mathbb{R}_+, \mathbb{U} \times \check{\mathbb{U}}_\varsigma \times \mathbb{V}_\varsigma \times \bar{\mathbb{V}})$ such that its projection onto $D(\mathbb{R}_+, \mathbb{U} \times \check{\mathbb{U}}_\varsigma \times \mathbb{V}_\varsigma)$ coincides with the distribution $\widetilde{\mathbb{P}}^{\mathrm{I}}$ chosen by the first player together with the rule (5.6). In particular, for the counterstrategy $\widetilde{\mathbb{Q}}^{\mathrm{II}}$, we could take an arbitrary element $\mathbb{Q}^{\mathrm{II}}$ from $\mathfrak{Q}^{\mathrm{II}}$, i.e., a counterstrategy of the original game; in this case, the above mapping would not depend on z.

Fix initial conditions $(x_0 \stackrel{\triangle}{=} (0, x_*, 0), z_0 = (0, y_0, 0))$ such that $||x_* - y_0||^2 \leq 7L^2\gamma^2 e^{-12LT}$ and $y_0 \in h\mathbb{R}^d$. For this it is sufficient (see (3.1)) to choose as y_0 the element from $h\mathbb{Z}^d$ nearest to x_*. For the strategy $\mathbb{P}^{\mathrm{I}}$ of the first player, we take the described stepwise strategy and fix an arbitrary realization $\widetilde{\mathbb{P}}^{\mathrm{all}}_{\delta_{(x_0, z_0)}}$ of the resulting joint strategy $\widetilde{\mathbb{P}}^{\mathrm{I+II}}$. Thus, we have specified both the probability $\widetilde{\mathbb{P}}^{\chi}_{\delta_{(x_0, z_0)}}$, which tracks only the position of the system, and the realization $\hat{\mathbb{P}}^{\mathrm{all}}_{\delta_{x_0}}$ not containing the guide. In particular, if the second player chooses the counterstrategy $\mathbb{Q}^{\mathrm{II}} = \widetilde{\mathbb{Q}}^{\mathrm{II}}$ among the counterstrategies of the original game, the distribution $\hat{\mathbb{P}}^{\mathrm{all}}_{\delta_{x_0}}$ becomes a realization of some joint strategy from $\mathbb{Q}^{\mathrm{II}}[\mathbb{P}^{\mathrm{I}}]$.

Since the initial position is fixed, we will omit the corresponding lower indices.

6.1. Divergence of trajectories. Define a stopping time $\tau \stackrel{\triangle}{=} \min\{t \mid \pi_0 Y''(t) \geq T\} \cup \{T\}$ and consider two consecutive times t_{k-1} and t_k from the partition and, together with them, stopping times $t' \stackrel{\triangle}{=} \min(t_{k-1}, \tau)$ and $t'' \stackrel{\triangle}{=} \min(t_k, \tau)$. On the interval $[t'; t'')$, the first player implements a constant control $(\bar{u}^{\text{to } w'}(x'), \bar{\mu}^{\text{opt}}, \nu^{\text{from } x'})$, where $x' = Y'(t')$ and $w' = Y''(t')$. Disintegrating $\widetilde{\mathbb{P}}^\chi$ over $x' = Y'(t')$ and $w' = Y''(t')$, we consider for all $x' \in \mathcal{X}$ and $w' \in \mathcal{Z}_<$ the corresponding conditional expected value $\widetilde{\mathbb{E}}^\chi_{t', x', w'}$ of the difference $\|Y''(t'') - Y'(t'')\|^2_d - \|Y''(t') - Y'(t')\|^2_d$.

From the definitions of $R_{x', w'}$ and $R_{w', x'}$ (see (5.1)) and the chain of equalities

$$\|w - x\|^2_d - \|w - x + x' - w'\|^2_d = -2\sum_{i=0}^d \pi_i(x' - w')\,\pi_i(w - x) - \|x' - w'\|^2_d$$

$$= 2\sum_{i=0}^d \pi_i(x' - w')\,\pi_i(x - x' + w' - w) + \|x' - w'\|^2_d = 2R_{x', w'}(x) + 2R_{w', x'}(w) + \|x' - w'\|^2_d,$$

we have the inequality $\big(\|x - w\|^2_d - \|x' - w'\|^2_d\big)/2 \leq \|x - x'\|^2_d + \|w - w'\|^2_d + R_{x', w'}(x) + R_{w', x'}(w)$. Putting $S(x, w, x', w') \stackrel{\triangle}{=} \|x - x'\|^2_d + R_{x', w'}(x) + R_{w', x'}(w) + \|w - w'\|^2_d$, we get

$$\widetilde{\mathbb{E}}^\chi_{t', x', w'}\|Y'(t'') - Y''(t'')\|^2_d \leq \|x' - w'\|^2_d + 2\widetilde{\mathbb{E}}^\chi_{t', x', w'} S(Y'(t''), Y''(t''), x', w'). \tag{6.1}$$

Let us estimate the expected values of R and S. Substituting the stopping time t'', since (4.8) and (5.3) are martingales, we obtain

$$\widetilde{\mathbb{E}}^\chi_{t', x', w'} R_{x', w'}(Y'(t'')) = \widetilde{\mathbb{E}}^\chi_{t', x', w'} \int_{t'}^{t''} \hat{\Lambda}[\bar{u}^{\text{to } w'}(x'), \bar{v}] R_{x', w'}(Y'(s))\, ds,$$

$$\widetilde{\mathbb{E}}^\chi_{t', x', w'} R_{w', x'}(Y''(t'')) = \widetilde{\mathbb{E}}^\chi_{t', x', w'} \int_{t'}^{t''} \check{\Lambda}[\bar{\mu}^{\text{opt}}, \bar{v}^{\text{from } x'}(w')] R_{w', x'}(Y''(s))\, ds.$$

Adding and using inequality (5.5), we get

$$\widetilde{\mathbb{E}}^\chi_{t', x', w'}\big(R_{x', w'}(Y'(t'')) + R_{w', x'}(Y''(t''))\big)$$

$$= \widetilde{\mathbb{E}}^\chi_{t', x', w'} \int_{t'}^{t''} \left[\hat{\Lambda}[\bar{u}^{\text{to } w'}(x'), \bar{v}] R_{x', w'}(Y'(s)) + \check{\Lambda}[\bar{\mu}^{\text{opt}}, \bar{v}^{\text{from } x'}(w')] R_{w', x'}(Y''(s))\right] ds$$

$$\leq \frac{3L}{2} \int_{t'}^{t''} \widetilde{\mathbb{E}}^\chi_{t', x', w'}\Big(4\|x' - w'\|^2_d + \|Y'(s) - x'\|^2_d + \|Y''(s) - w'\|^2_d\Big)\, ds.$$

Thus, the following inequality holds for S:

$$\widetilde{\mathbb{E}}^\chi_{t', x', w'}\left[S(Y'(t''), Y''(t''), x', w') - \|Y'(t'') - x'\|^2_d - \|Y''(t'') - w'\|^2_d\right]$$

$$\leq 3L\widetilde{\mathbb{E}}^\chi_{t', x', w'} \int_{t'}^{t''} \left(2\|x' - w'\|^2_d + \frac{\|Y'(s) - x'\|^2_d + \|Y''(s) - w'\|^2_d}{2}\right) ds.$$

By the estimates (3.4) and (4.15), we have

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}||Y'(t) - x'||^2_d \overset{(3.4)}{\leq} L^2(t_{k-1} - t_k)^2,$$

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}||Y''(t) - w'||^2_d \overset{(4.15)}{\leq} L(t - t_{k-1})\Big(h\sqrt{d+1} + \frac{4M(e^{t-t_{k-1}} - 1)^{3/2}}{3(t - t_{k-1})}\Big)$$

for all $t \in [t_{k-1}; t_k]$, where $M^2 = hL\sqrt{d+1} + L^2$. For $\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}S(Y'(t''), Y''(t''), x', w')$, let us substitute the stopping time $t'' \leq t' + t_{k-1} - t_k$ as t and define the function $g(r) \overset{\triangle}{=} (h\sqrt{d+1} + 4M(e^r - 1)^{3/2}/3r + Lr)(1 + Lr)/6$, where $r = \max_{t_i \leq T}(t_i - t_{i-1})$. Then

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}S(Y'(t''), Y''(t''), x', w') \leq 6L(t'' - t')||x' - w'||^2_d + 6L(t_k - t_{k-1})g(r).$$

Now, returning to (6.1), we obtain

$$\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}||Y'(t'') - Y''(t'')||^2_d \leq (1 + 12L(t_k - t_{k-1}))\widetilde{\mathbb{E}}^{\chi}_{t',x',w'}||Y'(t') - Y''(t')||^2_d + 12L(t_k - t_{k-1})g(r).$$

Integrating the resulting estimate over all $t' = \min(t_{k-1}, \tau)$ and $(x', w') = (Y'(t'), Y''(t'))$, we get

$$\widetilde{\mathbb{E}}^{\mathrm{all}}\big[||Y'(t'') - Y''(t'')||^2_d - e^{12L(t_k - t_{k-1})}||Y'(t') - Y''(t')||^2_d\big] \leq 12L(t_k - t_{k-1})g(r);$$

now, substituting $t' = \min(t_{k-1}, \tau)$ and $t'' = \min(t_k, \tau)$, we obtain

$$\widetilde{\mathbb{E}}^{\mathrm{all}}||Y'(\min(t_k, \tau)) - Y''(\min(t_k, \tau))||^2_d$$
$$\leq e^{12L(t_k - t_{k-1})}\widetilde{\mathbb{E}}^{\mathrm{all}}||Y'(\min(t_{k-1}, \tau)) - Y''(\min(t_{k-1}, \tau))||^2_d + 12L(t_k - t_{k-1})g(r)e^{12L(t_k - t_{k-1})}.$$

Since this is true for all intervals $[t_{i-1}; t_i)$, multiplying by $e^{12L(t_k - t_i)}$, we also obtain

$$e^{12L(t_k - t_i)}\widetilde{\mathbb{E}}^{\mathrm{all}}||Y'(\min(t_i, \tau)) - Y''(\min(t_i, \tau))||^2_d$$
$$\leq e^{12L(t_k - t_{i-1})}\widetilde{\mathbb{E}}^{\mathrm{all}}||Y'(\min(t_{i-1}, \tau)) - Y''(\min(t_{i-1}, \tau))||^2_d + 12L(t_i - t_{i-1})g(r)e^{12L(t_k - t_{i-1})}.$$

Successively substituting these inequalities into each other, we get

$$\widetilde{\mathbb{E}}^{\mathrm{all}}||Y'(\min(t_k, \tau)) - Y''(\min(t_k, \tau))||^2_d$$
$$\leq e^{12Lt_k}\widetilde{\mathbb{E}}^{\mathrm{all}}||Y'(0) - Y''(0)||^2_d + 12Lg(r)\sum_{i=1}^{k}(t_i - t_{i-1})e^{12L(t_k - t_{i-1})};$$

replacing here the sum over lower rectangles for the integral, we obtain

$$\widetilde{\mathbb{E}}^{\mathrm{all}}||Y'(\min(t_k, \tau)) - Y''(\min(t_k, \tau))||^2_d \leq e^{12Lt_k}\widetilde{\mathbb{E}}^{\mathrm{all}}||Y'(0) - Y''(0)||^2_d + g(r)(e^{12Lt_k} - 1)$$

for all natural k. Now, from $||Y'(0) - Y''(0)||^2_d \leq 7L^2 e^{-12LT}\gamma^2$, since T belongs to the partition, we finally derive

$$\widetilde{\mathbb{E}}^{\mathrm{all}}_{t'}||Y'(\min(T, \tau)) - Y''(\min(T, \tau))||^2_d \leq 7L^2\gamma^2 + e^{12LT}g(r). \tag{6.2}$$

Recall that, as shown in (4.4),

$$\widetilde{\mathbb{E}}^{\mathrm{all}}|T - \tau|^2 \leq \widetilde{\mathbb{E}}^{\mathrm{all}}\max_{t\in[0;T+2h]}|\pi_0 Y'(t) - \pi_0 Y''(t)|^2 \leq 2\gamma^2. \tag{6.3}$$

Then the estimate (3.4) implies

$$\widetilde{\mathbb{E}}^{\mathrm{all}}||Y'(T) - Y'(\min(T,\tau))||_d \le 3\sqrt{2}L\gamma. \tag{6.4}$$

Further, recall that $e^{-t}(||\check{Y}(t) - w'||_d^2 + M^2)$ is a supermartingale for $M^2 = hL\sqrt{d+1} + L^2$ (see (4.12)). Then, for the stopping times $\min(T,\tau)$ and $T + 2\gamma$, we have the inequality

$$e^{-T-2\gamma}(1 + \widetilde{\mathbb{E}}^{\mathrm{all}}||Y''(T + 2\gamma) - Y''(\min(T,\tau))||_d^2/M^2) \le \widetilde{\mathbb{E}}^{\mathrm{all}}e^{-\min(T,\tau)}.$$

Estimating its right-hand side by means of

$$\widetilde{\mathbb{E}}^{\mathrm{all}}e^{-\min(T,\tau)} = e^{-T}\widetilde{\mathbb{E}}^{\mathrm{all}}[1 + e^{\max(0,T-\tau)} - 1] \le e^{-T} + \widetilde{\mathbb{E}}^{\mathrm{all}}\max(0, T - \tau) \overset{(6.3)}{\le} e^{-T} + \sqrt{2}\gamma$$

and putting $H(\gamma) \overset{\triangle}{=} (hL\sqrt{d+1} + L^2)\gamma e^{2\gamma}(2 + \sqrt{2}e^T)$, we obtain

$$e^{-T-2\gamma}(1 + \widetilde{\mathbb{E}}^{\mathrm{all}}||Y''(T + 2\gamma) - Y''(\min(T,\tau))||_d^2/M^2) \le e^{-T} + \sqrt{2}\gamma$$

and

$$\widetilde{\mathbb{E}}^{\mathrm{all}}||Y''(T + 2\gamma) - Y''(\min(T,\tau))||_d^2 \le M^2(e^{2\gamma} - 1 + \sqrt{2}e^{T+2\gamma}\gamma) \le H(\gamma).$$

Together with (6.2) and (6.4), since $\sqrt{7} + 3\sqrt{2} < 7$, we get

$$\widetilde{\mathbb{E}}^{\mathrm{all}}||Y'(T) - Y''(T + 2\gamma)||_d \le 7L\gamma + \sqrt{H(\gamma)} + e^{6LT}\sqrt{g(r)}. \tag{6.5}$$

6.2. Divergence of payoffs. Denote by A the event

$$(\pi_0 Y''(T + 2\gamma) > T + \gamma) \,\&\, (\pi_0 Y''(T + \gamma) > T) \,\&\, (\bar{\sigma}(\hat{Y}) \equiv \bar{\sigma}(Y'));$$

here $\hat{Y}(\cdot)$ is the trajectory of the original game. Note that $\hat{Y}(\cdot)$ can differ from $Y'(\cdot)$ only on $[T;\infty)$ and only because of a jump of Y' in the last coordinate. Jumps are impossible after the stopping time $\tau' = \min\{t \,|\, \pi_0 Y''(t) > T + \gamma\}$. Since the intensities of the players do not exceed L, we find that the probability of such a jump does not exceed

$$2L(\gamma + \widetilde{\mathbb{E}}^{\mathrm{all}}|\tau' - T|) < 2L\gamma + 2L\widetilde{\mathbb{E}}^{\mathrm{all}}\max_{t\in[0;T+2\gamma]}|\pi_0 Y''(t) - t| \overset{(6.3)}{\le} 2L(1 + \sqrt{2})\gamma;$$

consequently, the probability of the event $(\bar{\sigma}(\hat{Y}) \ne \bar{\sigma}(Y'))$ is not greater either. On the other hand, (4.5) implies $\widetilde{\mathbb{P}}^{\mathrm{all}}((\pi_0 Y''(T + 2\gamma) > T + \gamma) \,\&\, (\pi_0 Y''(T + \gamma) > T)) \ge 1 - \gamma$. Since $L \ge 2$ and the absolute values of the functions W and $\bar{\sigma}$ are bounded by 1, we have

$$\widetilde{\mathbb{P}}^{\mathrm{all}}(\text{ not } A)\,\widetilde{\mathbb{E}}^{\mathrm{all}}(|\bar{\sigma}(\hat{Y}) - \bar{\sigma}(Y'')|\,|\,\text{not } A) \le 2L(3 + 2\sqrt{2})\gamma \le 6L^2\gamma. \tag{6.6}$$

Consider a trajectory for which the event A is true. Recall that $\hat{Y}(t) = Y'(t) = Y'(t_T)$ and $Y''(t) = Y''(t_T)$ for all $t \ge t_T \overset{\triangle}{=} \sup\{t \,|\, \pi_0 Y''(t) < T + \gamma\} \cup \{T\}$. Now, for the trajectory of the process in the case of the event A, we have

$$\bar{\sigma}(Y'') = \int_0^\infty he^{h(T+2\gamma-t)}1_{[0;T+2\gamma)}(t)W(Y''(t))\,dt = W(Y''(T + 2\gamma)),$$

$$\bar{\sigma}(\hat{Y}) = \int_0^\infty he^{h(T+2\gamma-t)}1_{[0;T+2\gamma)}(t)W(Y'(t))\,dt = W(Y'(T + 2\gamma)).$$

In addition, in the case of A, a jump in the last coordinate may occur no more than once, and it occurs simultaneously in both components by the distance $1 + \pi_0 Y'$ and $1 + \pi_0 Y''$, respectively. Hence, it follows from (6.3) that

$$\widetilde{\mathbb{E}}^{\mathrm{all}} \max_{t \in [0; T+2\gamma]} |\pi_{d+1}(Y'(t) - Y''(t))|^2 \leq 2\gamma^2,$$

whence, in particular, $\widetilde{\mathbb{P}}^{\mathrm{all}}(A)\widetilde{\mathbb{E}}^{\mathrm{all}}\big(|\pi_{d+1}(Y'(t_T) - Y''(t_T))| \, \big| \, A\big) \leq \sqrt{2}\gamma \leq L\gamma$. Then in view of the $3L$-Lipschitzness of the mapping $y \mapsto W(y)$, it follows from (6.5) that

$$\widetilde{\mathbb{P}}^{\mathrm{all}}(A)\widetilde{\mathbb{E}}^{\mathrm{all}}\big(|\bar{\sigma}(\hat{Y}) - \bar{\sigma}(Y'')| \, \big| \, A\big) \leq 3L(8L\gamma + \sqrt{H(\gamma)} + e^{6LT}\sqrt{g(r)}).$$

Together with (6.6), this gives

$$\widetilde{\mathbb{E}}^{\mathrm{all}}|\bar{\sigma}(\hat{Y}) - \bar{\sigma}(Y'')| \leq 3L(10L\gamma + \sqrt{H(\gamma)} + e^{6LT}\sqrt{g(r)}). \tag{6.7}$$

Recall that, by the construction of the double game, Y'' is the process generated in the Markov game in which the first player uses on each interval $[t_{k-1}; t_k)$ the strategy $\bar{\mu}^{\mathrm{opt}} \equiv \bar{\mu}^{\downarrow\check{\mathcal{V}}}$. Then $\check{\mathcal{V}}(Y''(t))$ becomes a supermartingale on each of these intervals; i.e.,

$$\check{\mathbb{E}}^{\mathrm{all}}(\mathcal{V}(Y''(t_k))|Y''(t_{k-1})) \leq \mathcal{V}(Y''(t_{k-1})).$$

Uniting over all intervals, we find that $\check{\mathcal{V}}(z_0) \geq \widetilde{\mathbb{E}}^{\mathrm{all}}\bar{\sigma}(Y'')$; i.e., by (6.7),

$$\check{\mathcal{V}}(z_0) \geq \widetilde{\mathbb{E}}^{\mathrm{all}}\bar{\sigma}(\hat{Y}) - 30L^2\gamma - 3L\sqrt{H(\gamma)} - 3Le^{6LT}\sqrt{g(r)}$$

for any actions of the opponent $\bar{v}$. Since $\widetilde{\mathbb{E}}^{\mathrm{all}}\bar{\sigma}(\hat{Y})$ is exactly the objective function in the original game and the game with a guide within the double game is one of the methods to play in the original game for the first player in the case of discrimination of the first player by the second, we have shown that

$$V^+ = \hat{\mathcal{V}}^+(x_0) \leq \check{\mathcal{V}}(z_0) + 30L^2\gamma + 3L\sqrt{H(\gamma)} + 3Le^{6LT}\sqrt{g(r)}.$$

Similarly, using a guide, we construct a strategy for the second player with a symmetric estimate. Since $\gamma = \sqrt{2hT\ln(hT)}$, and the grid step h and partition fineness r can be chosen arbitrarily small, the equality $V^+ = V^-$ is also shown.

FUNDING

CONFLICT OF INTEREST

The author declares that he has no conflicts of interest.

REFERENCES

1. R. Amir, I. V. Evstigneev, and K. R. Schenk-Hoppé, "Asset market games of survival: A synthesis of evolutionary and dynamic games," Annals Finance **9** (2), 121–144 (2013). doi 10.1007/s10436-012-0210-5

2. Y. Averboukh, "Approximate solutions of continuous-time stochastic games," SIAM J. Control Optim. **54** (5), 2629–2649 (2016). doi 10.1137/16M1062247

3. Y. Averboukh, "Approximate public-signal correlated equilibria for nonzero-sum differential games," SIAM J. Control Optim. **57** (1), 743–772 (2019). doi 10.1137/17M1161403

4. A. Basu and L. Stettner, "Zero-sum Markov games with impulse controls," SIAM J. Control Optim. **58** (1), 580–604 (2020). doi 10.1137/18M1229365

5. A. Bensoussan and A. Friedman, "Nonlinear variational inequalities and differential games with stopping times," J. Functional Anal. **16** (3), 305–352 (1974). doi 10.1016/0022-1236(74)90076-7

6. A. Bensoussan and A. Friedman, "Nonzero-sum stochastic differential games with stopping times and free boundary problems," Trans. Amer. Math. Soc. **231** (2), 275–327 (1977). doi 10.1090/S0002-9947-1977-0453082-7

7. T. R. Bielecki, S. Crépey, M. Jeanblanc, and M. Rutkowski, "Arbitrage pricing of defaultable game options with applications to convertible bonds," Quantitative Finance **8** (8), 795–810 (2008). doi 10.1080/14697680701401083

8. P. Billingsley, *Convergence of Probability Measures* (Wiley, New York, 1968; Nauka, Moscow, 1977).

9. A. A. Borovkov, *Probability Theory* (URSS, Moscow, 1999; Springer, London, 2013).

10. E. B. Dynkin, "Game variant of a problem on optimal stopping," Soviet Math. Dokl. **10**, 270–274 (1969).

11. E. Ekström and G. Peskir, "Optimal stopping games for Markov processes," SIAM J. Control Optim. **47** (2), 684–702 (2008). doi 10.1137/060673916

12. F. Gensbittel and C. Grün, "Zero-sum stopping games with asymmetric information," Math. Oper. Res. **44** (1), 277–302 (2019). doi 10.1287/moor.2017.0924

13. E. Gromova, A. Malakhova, and A. Palestini, "Payoff distribution in a multi-company extraction game with uncertain duration," Math. **6** (9), article no. 165 (2018). doi 10.3390/math6090165

14. X. Guo and O. Hernández-Lerma, "Zero-sum continuous-time Markov games with unbounded transition and discounted payoff rates," Bernoulli **11** (6), 1009–1029 (2005). doi 10.3150/bj/1137421638

15. S. Hamadéne, "Mixed zero-sum stochastic differential game and American game options," SIAM J. Control Optim. **45** (2), 496–518 (2006). doi 10.1137/S036301290444280X

16. V. N. Kolokoltsov, *Markov Processes, Semigroups and Generators* (De Gryuter, Berlin, 2011), Ser. De Gruyter Studies in Mathematics, Vol. 38.

17. N. N. Krasovskii, "An approach–evasion game with a stochastic guide," Dokl. Akad. Nauk SSSR **237** (5), 1020–1023 (1977).

18. N. N. Krasovskii and A. N. Kotel'nikova, "On a differential interception game," Proc. Steklov Inst. Math. **268** (1), 161–206 (2010). doi: 10.1134/S008154381001013X

19. N. N. Krasovskii and A. N. Kotel'nikova, "Stochastic guide for a time-delay object in a positional differential game," Proc. Steklov Inst. Math. **277** (Suppl. 1), S145–S151 (2012). doi 10.1134/S0081543812050148

20. N. N. Krasovskii and A. I. Subbotin, *Positional Differential Games* (Nauka, Moscow, 1974) [in Russian].

21. N. N. Krasovskii and A. I. Subbotin, *Game-Theoretical Control Problems* (Springer, New York, 1988).

22. R. Laraki and E. Solan, "The value of zero-sum stopping games in continuous time," SIAM J. Control Optim. **43** (5), 1913–1922 (2005). doi 10.1137/S0363012903429025

23. J. Marin-Solano and E. Shevkoplyas, "Non-constant discounting and differential games with random time horizon," Automatica **47** (12), 2626–2638 (2011). doi 10.1016/j.automatica.2011.09.010

24. V. Mazalov, "Dynamic games with optimal stopping," in *Game Theory and Applications*, Ed. by L. A. Petrosjan and V. V. Mazalov (Nova Science, New York, 1996), Vol. 2, pp. 37–46.

25. P.-A. Meier, *Probability and Potentials* (Mir, Moscow, 1973) [in Russian].

26. A. Neyman, "Continuous-time stochastic games," Games Econ. Behavior **104**, 92–130 (2017). doi 10.1016/j.geb.2017.02.004

27. T. Prieto-Rumeau and O. Hernández-Lerma, *Selected Topics on Continuous-Time Controlled Markov Chains and Markov Games* (Imperial College Press, London, 2012), Ser. ICP Advanced Texts in Mathematics, Vol. 5.

28. R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis* (Springer, Berlin, 2009).

29. S. Sorin and G. Vigeral, "Reversibility and oscillations in zero-sum discounted stochastic games," J. Dyn. Games **2** (1), 103–115 (2015). doi 10.3934/jdg.2015.2.103

30. A. I. Subbotin, *Generalized Solutions of First-Order Partial Differential Equations*: *The Dynamical Optimization Perspective* (Birkhäuser, Boston, 1995; Inst. Komp. Issled., Moscow–Izhevsk, 2003).

Translated by E. Vasil'eva