ISSN 0081-5438, Proceedings of the Steklov Institute of Mathematics, 2010, Vol. 271, pp. 125–137. © Pleiades Publishing, Ltd., 2010.
Original Russian Text © A.V. Kryazhimskii, Yu.S. Osipov, 2010, published in Trudy Matematicheskogo Instituta imeni V.A. Steklova, 2010, Vol. 271, pp. 134–147.

On Dynamical Regularization under Random Noise

A. V. Kryazhimskii[a,b] **and Yu. S. Osipov**[c]

Received February 2010

Abstract—We consider the problem of constructing a robust dynamic approximation of a time-varying input to a control system from the results of inaccurate observation of the states of the system. In contrast to the earlier studied cases in which the observation errors are assumed to be small in the metric sense, the errors in the present case are allowed to take, generally, large values and are subject to a certain probability distribution. The observation errors occurring at different instants are supposed to be statistically independent. Under the assumption that the expected values of the observation errors are small, we construct a dynamical algorithm for approximating the normal (minimal in the sense of the mean-square norm) input; the algorithm ensures an arbitrarily high level of the mean-square approximation accuracy with an arbitrarily high probability.

DOI: 10.1134/S0081543810040103

INTRODUCTION

In [1], Krasovskii and Subbotin proposed a general method for constructing feedback control laws that are robust with respect to observation errors; this method is known as a control procedure with a model (with a guide). Control procedures with a model remove possible instability of the basic feedbacks, which rely on the use of exact information about the current states of the control system. In the theory of feedback control, such procedures play the role of regularizing algorithms understood in the sense of the theory of ill-posed problems [2].

The control procedure with a model is related to the problem of stable tracking of motions, which is well known in engineering; more precisely, this procedure is implemented by solving a specially constructed (from the original problem of guaranteeing control) problem of this class. The problems of stable tracking of motions, just as the control procedures with a model, are traditionally aimed at removing the instability effect caused by small noises in the observation channel of a deterministic control system. In recent years, generalized statements in which both the control systems themselves and the observation noises contain stochastic elements have been addressed [3, 4]. When considering such statements, one uses the formalism of random processes (see [5, 6]) applied to control systems regulated by stochastic feedbacks. The methods of investigation of the corresponding random control processes are related to the studies on stochastic approximation of mixed strategies in the theory of positional differential games (see [7–10]). The present paper is largely based on these studies.

The purpose of the paper is to extend the scope of application of dynamic regularization methods to systems with uncertain inputs (see [11–16]). The methods of this class relate the control procedure with a model to regularizing algorithms in the theory of ill-posed problems and aim for the robust real-time reconstruction of current values of unobservable inputs of control systems from the available results of observation of their current states. To date, the methods of dynamic regularization have been developed under conditions of low deterministic noises in the observation channel. In the present paper, we allow for random observation noises that may generally take

[a] Steklov Mathematical Institute, Russian Academy of Sciences, ul. Gubkina 8, Moscow, 119991 Russia.
[b] International Institute for Applied Systems Analysis, Schlossplatz 1, Laxenburg, A-2361, Austria.
[c] Presidium of the Russian Academy of Sciences, Leninskii pr. 14, Moscow, 119991 Russia.

large values. Under the assumption that the observation noises occurring at different instants of time are mutually independent and identically distributed and that the norms of their expectations are small, we construct an algorithm for the dynamic approximation of a normal (minimal in the sense of the mean-square norm) input. This algorithm provides an arbitrarily high accuracy of the mean-square approximation with arbitrarily high probability and represents a modification of the "deterministic" method of regularized extremal shift proposed in [11]. The modification consists in the following: on a small time interval preceding the formation of the current value of the sought approximation, one collects sufficiently rich observation statistics and interprets their mean value as the value of a deterministic noise; on the basis of this value, a current value of approximation is generated by the basic deterministic method.

In Section 1, we describe the basic deterministic method. In Section 2, we modify this method as applied to the case of a stochastic observation noise.

Throughout the paper, the symbols $|\cdot|$ and $\langle\cdot,\cdot\rangle$ stand for the norm and the inner product in Euclidean space, respectively.

1. DETERMINISTIC SCHEME

Consider a dynamical system described by the following ordinary differential equation and initial condition in the space $\mathbb{R}^n$:

$$\dot{x}(t) = f_1(t, x(t)) + f_2(t, x(t))v(t), \qquad x(t_0) = x^0. \tag{1.1}$$

Here the time variable t runs through the bounded interval $[t_0, \vartheta]$, $x(t) \in \mathbb{R}^n$ is the state of the system at instant t, and $v(t) \in \mathbb{R}^r$ is the value of the input variable (input) of the system at this instant. The functions f_1 and f_2, which map $[t_0, \vartheta] \times \mathbb{R}^n$ to $\mathbb{R}^n$ and to the normed space of $r \times n$ matrices, respectively, are assumed to be continuous. The initial state $x^0 \in \mathbb{R}^n$ is assumed fixed.

We suppose that the values of the input are limited to a given convex compact set $V \subset \mathbb{R}^r$. An *admissible input* (for system (1.1)) is any Lebesgue measurable function $v(\cdot)\colon [t_0, \vartheta] \mapsto V$. Assume that for any admissible input $v(\cdot)$, any interval $[t_1, t_2] \subset [t_0, \vartheta)$, and any $x_1 \in \mathbb{R}^n$, the Cauchy problem

$$\dot{x}(t) = f_1(t, x(t)) + f_2(t, x(t))v(t), \qquad x(t_1) = x_1,$$

has a unique solution defined on $[t_1, t_2]$; this solution is understood in the sense of Carathéodory. For any admissible input $v(\cdot)$, the corresponding solution to the Cauchy problem (1.1) defined on $[t_0, \vartheta]$ is called a *motion* (of system (1.1)) generated by the admissible input $v(\cdot)$. The motion generated by some admissible input is called an *admissible motion*. We assume that the set of all admissible motions is uniformly bounded.

Since our analysis is restricted to the admissible motions of system (1.1), we can assume without loss of generality that the functions f_1 and f_2 are bounded. (If this is not so, we redefine the functions f_1 and f_2 outside the set of states that can occur during admissible motions.) Thus, we include the boundedness of the functions f_1 and f_2 into the number of initial assumptions.

In [11] and [12], we set and solved the following problem of stable approximation of an unobservable input in the real-time mode (the problem of dynamic regularization, or modeling, of an input). Suppose that at every instant $t \in [t_0, \vartheta]$ the result of measurement of the current state $x(t)$ of system (1.1) is available to an observer who controls this system; however, the measurement result is inaccurate, namely, it has the form $x(t) + \nu(t)$, where $\nu(t)$ is the value of an unknown measurement error that does not exceed a (small) positive value h: $|\nu(t)| \le h$. Below, any function $\nu(\cdot)\colon [t_0, \vartheta] \mapsto \mathbb{R}^n$ such that $|\nu(t)| \le h$ for all $t \in [t_0, \vartheta]$ is called a *measurement error* with margin of error h.

No information about the input generating an observed motion (except the a priori information that this input is admissible) is available to the observer. The task of the observer is as follows: using incoming information, provide a real-time approximation to the admissible input that generates the observed motion. The sought approximations are required to be robust with respect to observation errors, i.e., sufficiently accurate, for example, in the sense of mean-square deviation, under a sufficiently small upper bound h for the values of these errors.

This meaningful statement should be refined if the kernel of the matrix $f_2(t, x(t))$ calculated along the observed motion $x(\cdot)$ becomes nontrivial at certain instants t and therefore the admissible input that generates this motion cannot be uniquely identified in principle. In this case, in accordance with the standard technique used in the theory of ill-posed problems, the observer is prescribed to approximate one of such admissible inputs, say, the input with the least mean-square norm (the so-called normal input). Henceforth, we follow precisely this view of the problem.

Omitting the formal statement (formalism for problems of this type is presented in [12]), we describe its solution. It will be clear from the construction that the solution matches the above meaningful statement.

Let us introduce a refining definition. For any admissible motion $x(\cdot)$, an admissible input that generates this motion and has the least norm in the space $L^2([t_0,\vartheta],\mathbb{R}^r)$ is said to be *normal* for $x(\cdot)$ and is denoted by $v^*(\cdot\,|x(\cdot))$. Since the set V of admissible input values is convex and compact, a normal admissible input exists for any admissible motion and is unique as an element of the space $L^2([t_0,\vartheta],\mathbb{R}^r)$.

We will construct stable dynamic approximations of normal inputs for observed admissible motions according to the following algorithm.

Before the motion starts, the observer chooses a *time grid*, i.e., a finite family $(\tau_i)_{i=0}^N$ of instants in the interval $[t_0,\vartheta)$, where $\tau_0 = t_0$, $\tau_{i+1} > \tau_i$, $i = 0,\dots,N$, and $\tau_{N+1} = \vartheta$.

At every instant τ_i, $i = 0,1,\dots,N$, the observer uses the measurement result $x(\tau_i)+\nu(\tau_i)$ of the current state $x(\tau_i)$ of the system together with an auxiliary vector $y(\tau_i)\in\mathbb{R}^n$ produced by this time instant. This auxiliary vector is used to form a vector $u_i^{[v]}\in V$ that serves as a prediction of the values $v(t)$ of the sought normal input for $t\in[\tau_i,\tau_{i+1})$. The resulting piecewise constant function $u^{[v]}(\cdot)$ of the form

$$u^{[v]}(t) = u^{[v]}(\tau_i) = u_i^{[v]}, \qquad t\in[\tau_i,\tau_{i+1}), \quad i = 0,1,\dots,N, \tag{1.2}$$

provides the sought approximation to the admissible input $v^*(\cdot\,|x(\cdot))$ that is normal for the observed motion $x(\cdot)$. From the practical point of view, it is important that the approximation $u^{[v]}(\cdot)$ in this algorithmic scheme is generated dynamically in real-time mode: the values are assigned to the function $u^{[v]}(\cdot)$ during the observation process and are not recalculated later.

The values $y(\tau_i)$ of the auxiliary variable mentioned above are formed as the states of an auxiliary dynamical system, a *model*, that is described by the following differential equation and initial condition in $\mathbb{R}^n$:

$$\dot y(t) = f_1\big(t,u_i^{[x]}\big) + f_2\big(t,u_i^{[x]}\big)u_i^{[v]}, \qquad t\in[\tau_i,\tau_{i+1}), \quad i = 0,1,\dots,N, \qquad y(t_0) = x_0; \tag{1.3}$$

here $u_i^{[x]}\in\mathbb{R}^n$ and $u_i^{[v]}\in V$ are the values of control actions on the interval $[\tau_i,\tau_{i+1})$. For every $i = 1,\dots,N$, the control value $u_i^{[x]}$ traces the current measurement result, i.e.,

$$u_i^{[x]} = x(\tau_i)+\nu(\tau_i), \qquad i = 1,\dots,N, \tag{1.4}$$

while the control value $u_i^{[v]}\in V$ is defined as a solution to a linear–quadratic optimization problem constructed in accordance with the *regularized principle of Krasovskii's extremal shift* (see [1, 12]):

$$u_i^{[v]} = \arg\min\big\{\big\langle y(\tau_i)-(x(\tau_i)+\nu(\tau_i)), f_2\big(\tau_i,u_i^{[x]}\big)u^{[v]}\big\rangle+\alpha\big|u^{[v]}\big|^2\colon\ u^{[v]}\in V\big\}, \qquad i = 1,\dots,N; \tag{1.5}$$

here α is a positive *regularization parameter* fixed by the observer. For $i = 0$, the values $u_0^{[x]}$ and $u_0^{[v]}$ are chosen from the conditions

$$u_0^{[x]} = u^{[x]}(\tau_0) = x_0, \qquad u_0^{[v]} = u^{[v]}(\tau_0) \in V. \tag{1.6}$$

By analogy with (1.2), below we consider a time realization of the first control variable, i.e., a function $u^{[x]}(\cdot)\colon [t_0, \vartheta] \mapsto \mathbb{R}^n$, of the form

$$u^{[x]}(t) = u^{[x]}(\tau_i) = u_i^{[x]}, \qquad t \in [\tau_i, \tau_{i+1}), \quad i = 0, 1, \ldots, N. \tag{1.7}$$

By a *model process* corresponding to an admissible motion $x(\cdot)$ within the margin of measurement error h, we mean any triple $(u^{[x]}(\cdot), u^{[v]}(\cdot), y(\cdot))$ such that, for some measurement error $\nu(\cdot)$ with margin of error h, the function $u^{[x]}(\cdot)\colon [t_0, \vartheta] \mapsto \mathbb{R}^n$ is determined from (1.7) and (1.4), $u^{[v]}(\cdot)$ is an admissible input determined from (1.2) and (1.5), and $y(\cdot)$ is a (Carathéodory) solution on $[t_0, \vartheta]$ to the Cauchy problem (1.3); the function $u^{[v]}(\cdot)$ will be called the *output* of the model process. The latter definition stresses that we see the output of the algorithm (the approximation of a normal input for the observed motion $x(\cdot)$) as a realization of the control variable (1.2) in model (1.3) under the feedback control determined by the rule (1.4) of "tracking a measurement result" and by the rule (1.5) of regularized extremal shift.

Here we present a result on the mean-square dynamic approximation of the normal input; the approximation is uniform with respect to the classes $\mathcal{X}$ of admissible motions whose normal inputs form compact sets in $L^2([t_0, \vartheta], \mathbb{R}^r)$ (see [12]; in [11], the result is formulated for one-element classes $\mathcal{X}$). Below, we denote by $\bar{\tau}$ the maximal step of the time grid:

$$\bar{\tau} = \max\bigl\{|\tau_{i+1} - \tau_i|\colon\ i = 0, \ldots, N\bigr\}.$$

Theorem 1.1. *Let $\mathcal{X}$ be an arbitrary set of admissible motions of system* (1.1) *such that the family of normal inputs for all $x(\cdot) \in \mathcal{X}$ is a compact set in $L^2([t_0, \vartheta], \mathbb{R}^r)$. Then, for any $\varepsilon_1 > 0$, there exists an $\varepsilon_2 > 0$ such that if the margin of measurement error h, the maximal step $\bar{\tau}$ of the time grid, and the regularization parameter α satisfy the inequality*

$$h + \alpha + \frac{h}{\alpha} + \bar{\tau} \le \varepsilon_2, \tag{1.8}$$

then, for any admissible motion $x(\cdot) \in \mathcal{X}$, the mean-square deviation of the output $u^{[v]}(\cdot)$ of any model process corresponding to $x(\cdot)$ within the margin of measurement error h from the admissible input $v^*(\cdot\,|x(\cdot))$ normal for $x(\cdot)$ is not greater than ε_1:

$$\left(\int_{t_0}^{\vartheta} \bigl|u^{[v]}(t) - v^*(t|x(\cdot))\bigr|^2 \, dt \right)^{1/2} \le \varepsilon_1. \tag{1.9}$$

This result lays the base for the analysis carried out in the next section.

2. ROBUST DYNAMIC APPROXIMATION OF THE INPUT UNDER RANDOM OBSERVATION NOISE

Consider the problem of robust dynamic approximation of a normal input under the assumption that the values of the measurement errors are of random character and so may in general be large.

Suppose that at every instant $t \in [t_0, \vartheta]$, the result of observing the motion $x(t)$ of the system is $x(t) + \xi(t)$, where $\xi(t)$ is a random noise. Thus, we deal with a family $(\xi(t))_{t \in [t_0, \vartheta]}$ of n-dimensional random variables, a *random observation noise*. For every $t \in [t_0, \vartheta]$, the random variable $\xi(t)$ is assumed to be defined on its natural probability space $(\mathbb{R}^n, \mathcal{B}, p)$; i.e., this variable has the form

$\xi(t)(\omega) = \omega$, $\omega \in \mathbb{R}^n$; henceforth, $\mathcal{B}$ stands for the σ-algebra of Borel sets in $\mathbb{R}^n$ and p is a probability measure on $\mathcal{B}$. We assume that the random variables $\xi(t)$, $t \in [t_0, \vartheta]$, are pairwise independent, identically distributed, and have the same finite (small in the norm) mathematical expectation $\bar{\xi}$ and the same finite variance.

Under these conditions, to solve the problem, we modify the above-described method for controlling the model (1.3). The modification consists in the following: at every point τ_i of the time grid ($i = 1, \dots, N$), we form the values $u_i^{[x]}$ and $u_i^{[v]}$ of the control actions of the model by applying the above-described control law (1.4), (1.5) with an artificially synthesized value $\nu(\tau_i)$ of the measurement error. Namely, for every $i = 1, \dots, N$, we fix a family $(\tau_{ij})_{j=0}^{k}$ of pairwise distinct instants in the interval $[\tau_{i-1}, \tau_i)$, observe the states $x(\tau_{i1}), \dots, x(\tau_{ik})$ of the system at these instants, and, having obtained randomly perturbed results $x(\tau_{i1}) + \xi_{i1}, \dots, x(\tau_{ik}) + \xi_{ik}$ of observations ($\xi_{i1} = \xi(\tau_{i1}), \dots, \xi_{ik} = \xi(\tau_{ik})$), set

$$\nu(\tau_i) = \frac{(x(\tau_{i1}) + \xi_{i1}) + \ldots + (x(\tau_{ik}) + \xi_{ik})}{k} - x(\tau_i); \tag{2.1}$$

here k is a sufficiently large positive integer, called a *memory depth*. As before, we choose the values $u_0^{[x]}$ and $u_0^{[v]}$ from conditions (1.6).

Applying such a control procedure to the model results in a modified model process, which we call a model process with averaging. More precisely, by a *model process with averaging* corresponding to an admissible motion $x(\cdot)$ we mean an arbitrary triple $(u^{[x]}(\cdot), u^{[v]}(\cdot), y(\cdot))$ with the following property: there exist $\bar{\xi}_1 = (\xi_{11}, \dots, \xi_{1k}), \dots, \bar{\xi}_N = (\xi_{N1}, \dots, \xi_{Nk}) \in (\mathbb{R}^n)^k$ such that, for $\nu(\tau_i)$ determined from (2.1) ($i = 1, \dots, N$), the function $u^{[x]}(\cdot)\colon [t_0, \vartheta] \mapsto \mathbb{R}^n$ has the form (1.7), (1.4), $u^{[v]}(\cdot)$ is an admissible input defined by (1.2) and (1.5), and $y(\cdot)$ is a (Carathéodory) solution on $[t_0, \vartheta]$ to the Cauchy problem (1.3). The function $u^{[v]}(\cdot)$ will be called the *output* of this model process with averaging, the sequence $(\bar{\xi}_i)_{i=1}^N$ will be called a *sequence of observation errors* associated with this process, and the sequence $(\nu(\tau_i))_{i=1}^N$ will be called a *synthetic measurement error* associated with this process.

Remark 2.1. It is easy to see that if the associated synthetic measurement error $(\nu(\tau_i))_{i=1}^N$ satisfies the condition $|\nu(\tau_i)| \le h$, $i = 1, \dots, N$, for a model process with averaging corresponding to an admissible motion $x(\cdot)$, then such a process is also a model process corresponding to $x(\cdot)$ within the margin of measurement error h.

Take an arbitrary admissible motion $x(\cdot)$. Note that the values $u_i^{[x]}$ and $u_i^{[v]}$ of the control ($i = 1, \dots, N$) formed during model processes with averaging corresponding to the admissible motion $x(\cdot)$ depend on the values $\nu(\tau_i)$ (2.1) of synthesized measurement errors and hence on the values $\xi_{ij} = \xi(\tau_{ij})$ of random observation errors ($i = 1, \dots, N$, $j = 1, \dots, k$). Therefore, we can say that the model processes with averaging are generated by the trajectories of a certain random process that depends on $x(\cdot)$.

Let us define this random process; we will call it a *random process with averaging* corresponding to the admissible motion $x(\cdot)$. We define it conveniently as a random process with discrete time, taking the grid $(\tau_i)_{i=0}^N$ as a discrete time scale. This random process is determined by the initial probability (which is independent of $x(\cdot)$) and transition probabilities that depend on $x(\cdot)$.

We need to introduce some notation. For any $i = 0, 1, \dots, N-1$ and arbitrary $y_i \in \mathbb{R}^n$, $u_i^{[x]} \in \mathbb{R}^n$, and $u_i^{[v]} \in V$, we denote by $y_{i+1}\big(\cdot \,|\, y_i, u_i^{[x]}, u_i^{[v]}\big)$ a solution on $[\tau_i, \tau_{i+1}]$ to the Cauchy problem

$$\dot{y}(t) = f_1\big(t, u_i^{[x]}\big) + f_2\big(t, u_i^{[x]}\big) u_i^{[v]}, \qquad y(\tau_i) = y_i.$$

Further, for any $i = 1, \dots, N-1$ and any sequence $\bar{\xi}_i = (\xi_{i1}, \dots, \xi_{ik}) \in (\mathbb{R}^n)^k$ of observation noise values at instants $\tau_{i1}, \dots, \tau_{ik}$, we denote by $u_i^{[x]}(\bar{\xi}_i; x(\cdot))$ and $u_i^{[v]}(\bar{\xi}_i; x(\cdot))$ the vectors $u_i^{[x]}$ and $u_i^{[v]}$, respectively, that are defined by formulas (1.4) and (1.5) in which the value $\nu(\tau_i)$ of the measurement error is defined by the averaging formula (2.1).

Remark 2.2. It is obvious that the functions $\big(y_i, u_i^{[x]}, u_i^{[v]}\big) \mapsto y_{i+1}\big(\tau_{i+1}|y_i, u_i^{[x]}, u_i^{[v]}\big)$, $\bar{\xi}_i \mapsto u_i^{[x]}(\bar{\xi}_i; x(\cdot))$, and $\bar{\xi}_i \mapsto u_i^{[v]}(\bar{\xi}_i; x(\cdot))$ are continuous for every $i = 1, \dots, N-1$.

Now we define the initial probability. Since the values $u_0^{[x]}$ and $u_0^{[v]}$ of the control on the interval $[\tau_0, \tau_1)$ are fixed (see (1.6)), the state $y_1 = y(\tau_1)$ of the model at instant τ_1 is defined uniquely and does not depend on the observation results; more precisely,

$$y_1 = y_1\big(\cdot|x_0, u_0^{[x]}, u_0^{[v]}\big).$$

Under these conditions, the random values of the noise $\xi_{11} = \xi(\tau_{11}), \dots, \xi_{1k} = \xi(\tau_{1k})$, affecting the values of the control $u_1^{[x]}(\bar{\xi}_1; x(\cdot))$ and $u_1^{[v]}(\bar{\xi}_1; x(\cdot))$, where $\bar{\xi}_i = (\xi_{i1}, \dots, \xi_{ik}) \in (\mathbb{R}^n)^k$, define the possible states $y_2 = y(\tau_2)$ of the model at instant τ_2. Therefore, as the *initial probability space*, we take the product of k copies of the probability space $(\mathbb{R}^n, \mathcal{B}, p)$ of the observation noise, or, which is the same, the probability space $((\mathbb{R}^n)^k, \mathcal{B}^{(k)}, p^k)$; here $\mathcal{B}^{(k)}$ is the Borel σ-algebra on $(\mathbb{R}^n)^k$ and p^k is the product of k copies of the probability measure p; the latter product will serve as the initial probability measure r_1:

$$r_1 = p^k. \tag{2.2}$$

Let us define the transition probabilities. Since the states $y_2 = y(\tau_2), \dots, y_N = y(\tau_N) \in \mathbb{R}^n$ of the model that are formed during the random process represent n-dimensional random variables, we assume that the probability measures characterizing the distributions of these random variables are defined on the σ-algebra $\mathcal{B}$ of Borel subsets in $\mathbb{R}^n$. Thus, we have a Borel measurable space $(\mathbb{R}^n, \mathcal{B})$ of *states of the model.*

The formation of the random state $y_{i+1} = y(\tau_{i+1})$ of the model at instant τ_{i+1}, $i = 1, \dots, N-1$, depends on the model state $y_i = y(\tau_i)$ realized at instant τ_i and on the sequence $\bar{\xi}_i = (\xi_{i1}, \dots, \xi_{ik}) = (\xi(\tau_{i1}), \dots, \xi(\tau_{ik})) \in (\mathbb{R}^n)^k$ of observation errors realized at instants $\tau_{i1}, \dots, \tau_{ik} \in [\tau_{i-1}, \tau_i)$, which determine the control actions $u_i^{[x]}(\bar{\xi}_i; x(\cdot))$ and $u_i^{[v]}(\bar{\xi}_i; x(\cdot))$. Therefore, we will regard the pair $(y_i, \bar{\xi}_i)$, where $y_i \in \mathbb{R}^n$ and $\bar{\xi}_i = (\xi_{i1}, \dots, \xi_{ik}) \in (\mathbb{R}^n)^k$, as a full state of the random process at instant τ_i, $i = 1, \dots, N$. Thus, we define a measurable space (E, A) of *full states* as the product of the measurable spaces $(\mathbb{R}^n, \mathcal{B})$ and $((\mathbb{R}^n)^k, \mathcal{B}^{(k)})$ (which is obviously identified with the Borel space on $(\mathbb{R}^n)^{k+1}$):

$$(E, A) = (\mathbb{R}^n, \mathcal{B}) \times ((\mathbb{R}^n)^k, \mathcal{B}^{(k)}). \tag{2.3}$$

With an arbitrary sequence $\bar{\xi}_1 = (\xi_{11}, \dots, \xi_{1k}) \in (\mathbb{R}^n)^k$ of values of observation errors at instants $\tau_{11}, \dots, \tau_{1k} \in [\tau_0, \tau_1)$, we associate a probability $r_2(\cdot|\bar{\xi}_1; x(\cdot))$ describing a transition of the model from instant τ_1 to instant τ_2. The probability $r_2(\cdot|\bar{\xi}_1; x(\cdot))$ characterizes the distribution of full states $(y_2, \bar{\xi}_2) \in E$ that can appear at instant τ_2 provided that the sequence $\bar{\xi}_1 = (\xi_{11}, \dots, \xi_{1k})$ of values of observation errors is realized at instants $\tau_{11}, \dots, \tau_{1k}$. The state $y_2 = y(\tau_2)$ of the model is uniquely defined by the formula

$$y_2 = y_2\big(\cdot|y_1, u_1^{[x]}(\bar{\xi}_1; x(\cdot)), u_1^{[v]}(\bar{\xi}_1; x(\cdot))\big),$$

and the next sequence $\bar{\xi}_2 = (\xi_{21}, \dots, \xi_{2k})$ of values of observation errors realized at instants $\tau_{21}, \dots, \tau_{2k} \in [\tau_1, \tau_2)$ is random and independent of $\bar{\xi}_1$. Therefore, $r_2(\cdot|\bar{\xi}_1; x(\cdot))$ should be defined as the product of the probability measure concentrated at the point y_2 and the probability p^k, which defines the distribution of the random sequence $\bar{\xi}_2$.

Let us formalize the above reasoning. For any $y \in \mathbb{R}^n$, denote by $\delta(\cdot|y)$ the probability measure on $\mathcal{B}$ concentrated at the point y, i.e., $\delta(\{y\}|y) = 1$. For any $\bar{\xi}_1 \in (\mathbb{R}^n)^k$, we define the probability measure $r_2(\cdot|\bar\xi_1;x(\cdot))$ on the σ-algebra A as the product of the probability measures $\delta\big(\cdot|y_2\big(\cdot|y_1,u_1^{[x]}(\bar\xi_1;x(\cdot)),u_1^{[v]}(\bar\xi_1;x(\cdot))\big)\big)$ and p^k:

$$r_2(\cdot|\bar\xi_1;x(\cdot))=\delta\big(\cdot|y_2\big(\cdot|y_1,u_1^{[x]}(\bar\xi_1;x(\cdot)),u_1^{[v]}(\bar\xi_1;x(\cdot))\big)\big)\times p^k. \tag{2.4}$$

Thus, we have defined a function $r_2(\cdot|\cdot;x(\cdot))\colon (D,\bar\xi_1)\mapsto r_2(D|\bar\xi_1;x(\cdot))\colon A\times\mathbb{R}^n\mapsto[0,1]$.

Similarly (this time omitting meaningful explanations), for every $i=2,\dots,N-1$ and for any $(y_i,\bar\xi_i)\in E$, we define the probability measure

$$r_{i+1}(\cdot|y_i,\bar\xi_i;x(\cdot))=\delta\big(\cdot|y_{i+1}\big(\cdot|y_1,u_i^{[x]}(\bar\xi_i;x(\cdot)),u_i^{[v]}(\bar\xi_i;x(\cdot))\big)\big)\times p^k \tag{2.5}$$

on the σ-algebra A. Thus we define the functions $r_{i+1}(\cdot|\cdot;x(\cdot))\colon (D,(y_i,\bar\xi_i))\mapsto r_{i+1}(D|y_i,\bar\xi_i;x(\cdot))\colon A\times E\mapsto[0,1]$, $i=2,\dots,N-1$.

Lemma 2.1. *For an arbitrary admissible motion $x(\cdot)$, the function $r_2(\cdot|\cdot;x(\cdot))$ is a transition probability for the measurable spaces $((\mathbb{R}^n)^k,\mathcal{B}^{(k)})$ and (E,A), and the function $r_{i+1}(\cdot|\cdot;x(\cdot))$, $i=2,\dots,N-1$, is a transition probability for the measurable spaces (E,A) and (E,A).*

Proof. Let $x(\cdot)$ be an arbitrary admissible motion. Let us show that the function $r_2(\cdot|\cdot;x(\cdot))$ is a transition probability for the measurable spaces $((\mathbb{R}^n)^k,\mathcal{B}^{(k)})$ and (E,A). According to the definition of the transition probability (see, for example, [5, Sect. III.2]), to this end we should prove that the function $\bar\xi_1\mapsto r_2(D|\bar\xi_1;x(\cdot))$ is Borel measurable for arbitrary $D\in A$. Let us prove this. Let $D\in A$, a be an arbitrary number, and

$$X=\big\{\bar\xi_1\in(\mathbb{R}^n)^k\colon\ r_2(D|\bar\xi_1;x(\cdot))<a\big\}.$$

We have to show that $X\in\mathcal{B}$. Since the σ-algebra A is generated by the product $\mathcal{B}\times\mathcal{B}^{(k)}$ (see (2.3)), we may assume without loss of generality that $D=D_1\times D_2\in\mathcal{B}\times\mathcal{B}^{(k)}$. For an arbitrary $\bar\xi_1\in(\mathbb{R}^n)^k$, according to (2.4), we have

$$r_2(D|\bar\xi_1;x(\cdot))=\delta\big(D_1|y_2\big(\cdot|y_1,u_1^{[x]}(\bar\xi_1;x(\cdot)),u_1^{[v]}(\bar\xi_1;x(\cdot))\big)\big)\times p^k(D_2);$$

hence,

$$r_2(D|\bar\xi_1;x(\cdot))=\begin{cases}p^k(D_2) & \text{for}\ \ y_2\big(\cdot|y_1,u_1^{[x]}(\bar\xi_1;x(\cdot)),u_1^{[v]}(\bar\xi_1;x(\cdot))\big)\in D_1,\\ 0 & \text{for}\ \ y_2\big(\cdot|y_1,u_1^{[x]}(\bar\xi_1;x(\cdot)),u_1^{[v]}(\bar\xi_1;x(\cdot))\big)\notin D_1.\end{cases}$$

Therefore, $X=\varnothing$ if $a\le p(D_2)$, and

$$X=\big\{\bar\xi_1\in(\mathbb{R}^n)^k\colon\ y_2\big(\cdot|y_1,u_1^{[x]}(\bar\xi_1;x(\cdot)),u_1^{[v]}(\bar\xi_1;x(\cdot))\big)\in D_1\big\}$$

if $a>p(D_2)$. In the first case, the inclusion $X\in\mathcal{B}$ is obvious; in the second case, it holds because the mapping $\bar\xi_1\mapsto y_2\big(\tau_2|y_1,u_1^{[x]}(\bar\xi_1;x(\cdot)),u_1^{[v]}(\bar\xi_1;x(\cdot))\big)$ is continuous and, hence, Borel measurable (see Remark 2.2). Similarly one establishes that the function $r_{i+1}(\cdot|\cdot;x(\cdot))$ is a transition probability for the measurable spaces (E,A) and (E,A) for any $i=2,\dots,N-1$. The proof is complete.

We define a measurable space $(\mathcal{E},\mathcal{A})$ of *trajectories of random processes with averaging* as the product of the measurable space $((\mathbb{R}^n)^k,\mathcal{B}^{(k)})$, which carries the initial sequence of observation errors, and $N-1$ copies of the measurable space (E,A) of full states:

$$(\mathcal{E},\mathcal{A})=((\mathbb{R}^n)^k,\mathcal{B}^{(k)})\times\prod_2^N(E,A). \tag{2.6}$$

For an arbitrary admissible motion $x(\cdot)$, the initial probability r_1 and the transition probabilities $r_{i+1}(\cdot\,|\,\cdot;x(\cdot))$, $i=1,\dots,N-1$, define in a standard way a probability measure $r(\cdot;x(\cdot))$ that describes the distribution of the trajectories $q=(\bar\xi_1,(y_2,\bar\xi_2),\dots,(y_N,\bar\xi_N))$ of the random process with averaging corresponding to $x(\cdot)$. Namely, $r(\cdot;x(\cdot))$ is defined on the σ-algebra $\mathcal{A}$ of the space of trajectories and is characterized by the equality

$$r(D;x(\cdot))=\int\limits_{D_1} r_1(d\bar\xi_1;x(\cdot))\int\limits_{D_2} r_2(d(y_2,\bar\xi_2)|\bar\xi_1;x(\cdot)) \times\int\limits_{D_3} r_3(d(y_3,\bar\xi_3)|y_2,\bar\xi_2;x(\cdot))\dots\int\limits_{D_N} r_N(d(y_N,\bar\xi_N)|y_{N-1},\bar\xi_{N-1};x(\cdot)) \tag{2.7}$$

with $D=D_1\times D_2\times\dots\times D_N\in\mathcal{A}$, which holds for any $D_1,D_2,\dots,D_N\in\mathcal{B}^{(k)}\times A\times\dots\times A$ (see [5, Sect. V.1, Corollary 2]). For every $D\in\mathcal{A}$, the number $r(D;x(\cdot))$ is the probability that the trajectory $q=(\bar\xi_1,(y_2,\bar\xi_2),\dots,(y_N,\bar\xi_N))$ of the random process with averaging corresponding to $x(\cdot)$ is contained in the set D.

We say that a trajectory $q=(\bar\xi_1,(y_2,\bar\xi_2),\dots,(y_N,\bar\xi_N))\in\mathcal{E}$ is *associated* with a model process $(u^{[x]}(\cdot),u^{[v]}(\cdot),y(\cdot))$ with averaging corresponding to an admissible motion $x(\cdot)$ if

$$y_i=y(\tau_i),\qquad u_i^{[x]}(\bar\xi_i;x(\cdot))=u^{[x]}(\tau_i),\qquad u_i^{[v]}(\bar\xi_i;x(\cdot))=u^{[v]}(\tau_i),\qquad i=1,\dots,N,$$

and $(\bar\xi_i)_{i=1}^N$ is a sequence of observation errors associated with $(u^{[x]}(\cdot),u^{[v]}(\cdot),y(\cdot))$. We say that a trajectory $q\in\mathcal{E}$ is *associated* with an admissible motion $x(\cdot)$ if it is associated with some model process with averaging corresponding to $x(\cdot)$.

We are interested in the probability that the output $u^{[v]}(\cdot)$ of a model process $(u^{[x]}(\cdot),u^{[v]}(\cdot),y(\cdot))$ with averaging corresponding to an admissible motion $x(\cdot)$ of some class provides a mean-square approximation of the admissible normal input $v^*(\cdot|x(\cdot))$ for $x(\cdot)$ with a given accuracy ε_1, i.e., that estimate (1.9) holds.

We define this probability as follows. Let $\Phi(\varepsilon_1;x(\cdot))$ (with ε_1 an arbitrary positive number) be the set of all model processes $(u^{[x]}(\cdot),u^{[v]}(\cdot),y(\cdot))$ with averaging corresponding to an admissible motion $x(\cdot)$ for which inequality (1.9) holds, and let $\Phi^*(\varepsilon_1;x(\cdot))$ be the set of all trajectories $q=(\bar\xi_1,(y_2,\bar\xi_2),\dots,(y_N,\bar\xi_N))\in\mathcal{E}$ each of which is associated with some $(u^{[x]}(\cdot),u^{[v]}(\cdot),y(\cdot))\in\Phi(\varepsilon_1;x(\cdot))$. We will regard $r(\Phi^*(\varepsilon_1;x(\cdot));x(\cdot))$ as the probability that the model process $(u^{[x]}(\cdot),u^{[v]}(\cdot),y(\cdot))$ with averaging corresponding to $x(\cdot)$ belongs to the set $\Phi(\varepsilon_1;x(\cdot))$ (satisfies inequality (1.9)). We will also denote this quantity by $r(\Phi(\varepsilon_1;x(\cdot));x(\cdot))$.

Remark 2.3. Since the functions $\big(y_i,u_i^{[x]},u_i^{[v]}\big)\mapsto y_{i+1}\big(\tau_{i+1}|y_i,u_i^{[x]},u_i^{[v]}\big)$, $\bar\xi_i\mapsto u_i^{[x]}(\bar\xi_i;x(\cdot))$, and $\bar\xi_i\mapsto u_i^{[v]}(\bar\xi_i;x(\cdot))$ are continuous (see Remark 2.2) and the outputs $u^{[v]}(\cdot)$ of model processes with averaging are piecewise constant (see (1.2)), it follows that the set $\Phi^*(\varepsilon_1;x(\cdot))$ belongs to the σ-algebra $\mathcal{A}$; hence, the value $r(\Phi^*(\varepsilon_1;x(\cdot));x(\cdot))$ is well defined.

The following lemma states that the trajectories of a random process corresponding to an admissible motion $x(\cdot)$ are associated with $x(\cdot)$ with probability one. In other words, this random process, which is formally defined by means of the initial probability r_1 and the transition probabilities $r_{i+1}(\cdot\,|\,\cdot;x(\cdot))$, $i=1,\dots,N-1$, certainly does not generate trajectories that cannot be realized by model process with averaging corresponding to $x(\cdot)$.

Lemma 2.2. *Let $x(\cdot)$ be an admissible motion and $T(x(\cdot))$ be the set of all trajectories $q\in\mathcal{E}$ associated with $x(\cdot)$. Then $T(x(\cdot))\in\mathcal{A}$ and $r(T(x(\cdot));x(\cdot))=1$.*

Proof. It is obvious that $T(x(\cdot))=\bigcap_{i=1}^{N-1}T_i$, where

$$T_i=\big\{q=(\bar\xi_1,(y_2,\bar\xi_2),\dots,(y_N,\bar\xi_N))\colon\ y_{i+1}=y_{i+1}\big(\tau_{i+1}|y_i,u_i^{[x]}(\bar\xi_i;x(\cdot)),u_i^{[v]}(\bar\xi_i;x(\cdot))\big)\big\}.$$

Therefore, to prove the inclusion $T(x(\cdot)) \in \mathcal{A}$, it suffices to show that $T_i \in \mathcal{A}$ for $i = 1, \dots, N-1$. We apply induction. Let $i = 1$. Clearly,

$$T_1 = X_1 \times \prod_{2}^{N-1} E, \tag{2.8}$$

where

$$X_1 = \big\{(\bar{\xi}_1, (y_2, \bar{\xi}_2)) \in (\mathbb{R}^n)^k \times E\colon\ y_2 = y_2\big(\tau_2|y_1, u_1^{[x]}(\bar{\xi}_1; x(\cdot)), u_1^{[v]}(\bar{\xi}_1; x(\cdot))\big)\big\}. \tag{2.9}$$

Since the mapping $\bar{\xi}_1 \mapsto y_2\big(\tau_2|y_1, u_1^{[x]}(\bar{\xi}_1; x(\cdot)), u_1^{[v]}(\bar{\xi}_1; x(\cdot))\big)$ is continuous (see Remark 2.2), its graph is a closed (Borel) set in $(\mathbb{R}^n)^k \times \mathbb{R}^n$. Hence, X_1 is a Borel subset in $(\mathbb{R}^n)^k \times (\mathbb{R}^n \times (\mathbb{R}^n)^k) = (\mathbb{R}^n)^k \times E$ (see (2.3)). This, combined with (2.8) and the definition of the measurable space $(\mathcal{E}, \mathcal{A})$ (see (2.6)), implies that $T_1 \in \mathcal{A}$. The induction basis is established. The inductive step is performed in a similar way.

We have shown that $T(x(\cdot)) \in \mathcal{A}$. Let us prove that $r(T(x(\cdot)); x(\cdot)) = 1$. Suppose the contrary: $r(T(x(\cdot)); x(\cdot)) < 1$. Then

$$r(\mathcal{E} \setminus T(x(\cdot)); x(\cdot)) = r\left(\bigcup_{i=1}^{N-1} D_i\right) > 0;$$

here $D_i = \mathcal{E} \setminus T_i$; note that $D_i \in \mathcal{A}$ since $T_i \in \mathcal{A}$, $i = 1, \dots, N-1$. Therefore, we conclude that $r(D_1; x(\cdot)) + \ldots + r(D_{N-1}; x(\cdot)) > 0$. Thus, $r(D_i; x(\cdot)) > 0$ for some $i \in \{1, \dots, N-1\}$. Fix the minimal index i with this property. Consider the case of $i = 1$. So,

$$r(D_1; x(\cdot)) > 0. \tag{2.10}$$

By (2.8), we have

$$D_1 = [((\mathbb{R}^n)^k \times E) \setminus X_1] \times \prod_{2}^{N-1} E,$$

where X_1 is defined by (2.9). Let $g_1(\cdot)$ be the characteristic function of the set $((\mathbb{R}^n)^k \times E) \setminus X_1$ in $(\mathbb{R}^n)^k \times E$. It is clear that

$$g_1(\bar{\xi}_1, (y_2, \bar{\xi}_2)) = \begin{cases} 1 & \text{for}\ \ y_2 \neq y_2\big(\tau_2|y_1, u_1^{[x]}(\bar{\xi}_1; x(\cdot)), u_1^{[v]}(\bar{\xi}_1; x(\cdot))\big), \\ 0 & \text{for}\ \ y_2 = y_2\big(\tau_2|y_1, u_1^{[x]}(\bar{\xi}_1; x(\cdot)), u_1^{[v]}(\bar{\xi}_1; x(\cdot))\big), \end{cases} \tag{2.11}$$
$$(\bar{\xi}_1, (y_2, \bar{\xi}_2)) \in (\mathbb{R}^n)^k \times E.$$

According to the definition of the measurable space $(\mathcal{E}, \mathcal{A})$ (see (2.6)) and the probability measure $r(\cdot\,; x(\cdot))$ (see (2.7)), for the set D_1 we have

$$r(D_1; x(\cdot)) = \int\limits_{(\mathbb{R}^n)^k} r_1(d\bar{\xi}_1; x(\cdot)) \int\limits_{E} g_1(\bar{\xi}_1, (y_2, \bar{\xi}_2)) r_2(d(y_2, \bar{\xi}_2)|\bar{\xi}_1; x(\cdot)).$$

Taking into account the form of the probability measure $r_2(\cdot\,|\bar{\xi}_1; x(\cdot))$ (see (2.4)) and (2.11), we find that the inner integral is

$$\int\limits_{(\mathbb{R}^n)^k} g_1\big(\bar{\xi}_1, \big(y_2\big(\tau_2|y_1, u_1^{[x]}(\bar{\xi}_1; x(\cdot)), u_1^{[v]}(\bar{\xi}_1; x(\cdot))\big), \bar{\xi}_2\big)\big) p^k(d\xi_2) = 0$$

for every $\bar{\xi}_1 \in (\mathbb{R}^n)^k$. Hence, $r(D_1; x(\cdot)) = 0$. The contradiction with (2.10) completes the proof for the case of $i = 1$. The case of $i > 1$ is similar. The lemma is proved.

Let us estimate the probability $r(\Phi(\varepsilon_1; x(\cdot)); x(\cdot))$ from below.

By the law of large numbers, for any positive μ and ε, there exists a $k^*(\mu,\varepsilon) > 0$ such that the probability that the mean value of the identically distributed pairwise independent random variables $\xi_{i1} = \xi(\tau_{i1}), \ldots, \xi_{ik} = \xi(\tau_{ik})$ deviates from their common mathematical expectation $\bar\xi$ by at most μ is not less than $1-\varepsilon$ for all $k \geq k^*(\mu,\varepsilon)$:

$$p^k\left\{\bar\xi_i = (\xi_{i1},\ldots,\xi_{ik}) \in (\mathbb{R}^n)^k\colon \left|\frac{\xi_{i1}+\ldots+\xi_{i1}}{k} - \bar\xi\right| \leq \mu\right\} \geq 1-\varepsilon, \qquad i = 1,\ldots,N. \tag{2.12}$$

We fix this number $k^*(\mu,\varepsilon)$ for any positive μ and ε.

By assumption, the functions f_1 and f_2 are bounded. Introduce a constant $K > 0$ such that

$$|f_1(t,x) + f_2(t,x)v| \leq K, \qquad t \in [t_0,\vartheta], \quad x \in \mathbb{R}^n, \quad v \in V. \tag{2.13}$$

The following stochastic analog of Theorem 1.1 is the main result of the present study on the stable dynamic approximation of a normal input.

Theorem 2.1. *Suppose that*

(i) $\mathcal{X}$ *is an arbitrary set of admissible motions of system* (1.1) *such that the family of normal inputs for all* $x(\cdot) \in \mathcal{X}$ *is a compact set in* $L^2([t_0,\vartheta],\mathbb{R}^r)$;

(ii) ε_1 *is an arbitrary positive number and* $\varepsilon_2 > 0$ *is chosen, depending on* $\mathcal{X}$ *and* ε_1, *in the same way as in Theorem* 1.1;

(iii) *the margin of measurement error* h, *the maximal step* $\bar\tau$ *of the time grid, and the regularization parameter* α *satisfy the inequality*

$$h + \alpha + \frac{h}{\alpha} + \bar\tau \leq \varepsilon_2; \tag{2.14}$$

(iv) *the memory depth* k, *the maximal step* $\bar\tau$ *of the time grid, and the mathematical expectation* $\bar\xi$ *of observation errors satisfy the inequalities*

$$k \geq k^*(\mu,\varepsilon), \tag{2.15}$$

$$K\bar\tau + \mu + |\bar\xi| \leq h \tag{2.16}$$

for some $\varepsilon, \mu > 0$.

Then, for any admissible motion $x(\cdot) \in \mathcal{X}$, *the probability that the mean-square deviation of the output of the model process with averaging corresponding to* $x(\cdot)$ *from the admissible input normal for* $x(\cdot)$ *is at most* ε_1 *is not less than* $1 - N\varepsilon$:

$$r(\Phi(\varepsilon_1; x(\cdot)); x(\cdot)) \geq 1 - N\varepsilon. \tag{2.17}$$

Before proving the theorem, we comment on its substance.

Theorem 2.1 suggests an algorithm that provides, with a given probability β arbitrarily close to one, a dynamic approximation of the normal input of any admissible motion $x(\cdot) \in \mathcal{X}$ with prescribed accuracy ε_1 in the mean-square norm, where $\mathcal{X}$ is an arbitrary class satisfying condition (i) of the theorem.

At the first step of the algorithm, given an accuracy ε_1, one chooses ε_2 as indicated in condition (ii). Then one chooses sufficiently small values of the margin of measurement error h, the maximal step $\bar\tau$ of the time grid $(\tau_i)_{i=1}^N$, and the regularization parameter α so as to satisfy (2.14); note that the ratio h/α is also sufficiently small.

At the second step, according to (2.16), one evaluates an upper bound for the norm $|\bar\xi|$ of the mathematical expectation of random observation errors. Here one can also vary an auxiliary small positive parameter μ and modify the time grid $(\tau_i)_{i=1}^N$ by decreasing its maximal step $\bar\tau$ in order to satisfy (2.16). Note that the choice of the constraint on $|\bar\xi|$ is not under the control of the observer and, strictly speaking, cannot be a part of the algorithm. There is no need for such a constraint (which is only included for the sake of generality) if the observation errors are unbiased, i.e., if $\bar{\xi} = 0$.

At the third step, having fixed μ and the time grid $(\tau_i)_{i=1}^N$, we take $\varepsilon > 0$ so small that

$$1 - N\varepsilon \geq \beta.$$

At the fourth step, using the values of μ and ε, we choose the memory depth k involved in the averagings (see (2.1)) from condition (2.15).

This completes the adjustment of the procedure of dynamic approximation of a normal input. The procedure itself is implemented by the model process with averaging corresponding to an admissible motion $x(\cdot)$, which is generated in an arbitrary way within the set $\mathcal{X}$. The output of this model process provides a mean-square approximation to the normal input for the motion $x(\cdot)$ with probability at least β.

Proof of of Theorem 2.1. Let $x(\cdot)$ be an arbitrary admissible motion. Introduce the sets

$$D_i = \left\{ q = (\bar{\xi}_1, (y_2, \bar{\xi}_2), \ldots, (y_N, \bar{\xi}_N)) \in \mathcal{E} \colon\ \bar{\xi}_i = (\xi_{i1}, \ldots, \xi_{ik}),\ \left| \frac{\sum_{j=1}^k \xi_{ij}}{k} - \bar{\xi} \right| \leq \mu \right\}, \quad i = 1, \ldots, N. \tag{2.18}$$

It follows from the definition of the probability measure $r(\cdot\,; x(\cdot))$ (2.7) that

$$r(D_i; x(\cdot)) = \int\limits_{(\mathbb{R}^n)^k} r_1(d\bar{\xi}_1; x(\cdot)) \int\limits_E r_2(d(y_2, \bar{\xi}_2)|\bar{\xi}_1; x(\cdot)) \ldots \int\limits_E r_{i-1}(d(y_{i-1}, \bar{\xi}_{i-1})|y_{i-2}, \bar{\xi}_{i-2}; x(\cdot)) \times \int\limits_{\mathbb{R}^n \times B_i} r_i(d(y_i, \bar{\xi}_i)|y_{i-1}, \bar{\xi}_{i-1}; x(\cdot)) \tag{2.19}$$

for every $i \in \{1, \ldots, N\}$, where

$$B_i = \left\{ \bar{\xi}_i = (\xi_{i1}, \ldots, \xi_{ik}) \colon \left| \frac{\sum_{j=1}^k \xi_{ij}}{k} - \bar{\xi} \right| \leq \mu \right\}.$$

It follows from (2.15) that

$$p^k(B_i) \geq 1 - \varepsilon. \tag{2.20}$$

Since each integration with respect to the first variable in (2.19) is over the whole space $\mathbb{R}^n$, it follows from the form of the transition probabilities $r_j(\cdot\,; x(\cdot))$, $j = 1, \ldots, N-1$, (2.4), (2.5) that (2.19) reduces to

$$r(D_i; x(\cdot)) = \int\limits_{(\mathbb{R}^n)^k} p^k(d\bar{\xi}_1) \int\limits_{(\mathbb{R}^n)^k} p^k(d\bar{\xi}_2) \ldots \int\limits_{B^i} p^k(d\bar{\xi}_i) = p^k(B_i).$$

Hence, in view of (2.20), we have

$$r(D_i; x(\cdot)) \geq 1 - \varepsilon, \qquad i = 1, \ldots, N. \tag{2.21}$$

Set

$$D_j^0 = \bigcap_{i=1}^j D_i, \qquad j = 1, \ldots, N. \tag{2.22}$$

Let us show that

$$r\bigl(D_j^0; x(\cdot)\bigr) \geq 1 - j\varepsilon \tag{2.23}$$

for every $j = 1,\dots,N$. We apply induction. For $j = 1$, (2.23) is valid by virtue of (2.21). Suppose that (2.23) is valid for some index $j \in \{1,\dots,N-1\}$. Using (2.23) and the estimate $r(\mathcal{E}\setminus D_{j+1};x(\cdot)) \le \varepsilon$, which follows from (2.21), we obtain

$$r\big(D^0_{j+1};x(\cdot)\big) = r\big(D^0_j \cap D_{j+1};x(\cdot)\big) = r\big(D^0_j;x(\cdot)\big) - r\big(D^0_j \cap (\mathcal{E}\setminus D_{j+1});x(\cdot)\big)$$
$$\ge r\big(D^0_j;x(\cdot)\big) - r(\mathcal{E}\setminus D_{j+1};x(\cdot)) \ge 1 - j\varepsilon - \varepsilon.$$

Let $T(x(\cdot))$ be the set of all trajectories $q \in \mathcal{E}$ associated with $x(\cdot)$. By Lemma 2.2, we have $r(T(x(\cdot));x(\cdot)) = 1$. Therefore,

$$r\big(D^0_N \cap T(x(\cdot));x(\cdot)\big) = r(D^0;x(\cdot));$$

hence, in view of (2.23), for $j = N$ we have

$$r\big(D^0_N \cap T(x(\cdot));x(\cdot)\big) \ge 1 - N\varepsilon. \tag{2.24}$$

Recall that the probability $r(\Phi(\varepsilon_1;x(\cdot));x(\cdot))$ (see (2.17)) is equal to $r(\Phi^*(\varepsilon_1;x(\cdot));x(\cdot))$ by definition, where $\Phi^*(\varepsilon_1;x(\cdot))$ is the set of all trajectories $q = (\bar\xi_1,(y_2,\bar\xi_2),\dots,(y_N,\bar\xi_N)) \in \mathcal{E}$ each of which is associated with some $(u^{[x]}(\cdot),u^{[v]}(\cdot),y(\cdot)) \in \Phi(\varepsilon_1;x(\cdot))$. Therefore, taking into account (2.24), we conclude that to complete the proof, i.e., to prove estimate (2.17), it suffices to show that

$$D^0_N \cap T(x(\cdot)) \subset \Phi^*(\varepsilon_1;x(\cdot)). \tag{2.25}$$

Let us show this. Take an arbitrary trajectory

$$q = (\bar\xi_1,(y_2,\bar\xi_2),\dots,(y_N,\bar\xi_N)) \in D^0_N \cap T(x(\cdot)). \tag{2.26}$$

Since this trajectory is associated with $x(\cdot)$ ($q \in T(x(\cdot))$), it is associated with some model process $(u^{[x]}(\cdot),u^{[v]}(\cdot),y(\cdot))$ with averaging corresponding to $x(\cdot)$; in particular, the sequence $(\nu(\tau_i))_{i=1}^N$ defined as

$$\nu(\tau_i) = \frac{\sum_{j=1}^k (x(\tau_{ij}) + \xi_{ij})}{k} - x(\tau_i), \qquad i = 1,\dots,N$$

(see (2.1)), is associated with $(u^{[x]}(\cdot),u^{[v]}(\cdot),y(\cdot))$. Since $\tau_{ij} \in [\tau_{i-1},\tau_i)$, $i = 1,\dots,N$, $j = 1,\dots,k$, by (2.13) we have

$$|x(\tau_{ij}) - x(\tau_i)| \le K(\tau_i - \tau_{i-1}) \le K\bar\tau, \qquad i = 1,\dots,N, \quad j = 1,\dots,k.$$

Therefore, for every $i = 1,\dots,N$,

$$|\nu(\tau_i)| \le K\bar\tau + \left|\frac{\sum_{j=1}^k \xi_{ij}}{k}\right|. \tag{2.27}$$

Since $q \in D^0_N$, it follows (see (2.22)) that $q \in D_i$ for every $i = 1,\dots,N$. Hence, by the definition of the sets D_i, for every $i = 1,\dots,N$ we have

$$\left|\frac{\sum_{j=1}^k \xi_{ij}}{k} - \bar\xi\right| \le \mu.$$

From this estimate and (2.27) we find that

$$|\nu(\tau_i)| \le K\bar\tau + \mu + |\bar\xi| \le h$$

for every $i = 1,\dots,N$; the latter inequality is valid by assumption (iv) (see (2.16)). Hence, using Remark 2.1, we conclude that the model process $(u^{[x]}(\cdot),u^{[v]}(\cdot),y(\cdot))$ with averaging corresponding to the admissible motion $x(\cdot)$ is a model process corresponding to this admissible motion within the margin of measurement error h.

Recall that the margin of measurement error h, the maximal step $\bar{\tau}$ of the time grid, and the regularization parameter α satisfy inequality (2.14), in which ε_2 is chosen in the same way as in Theorem 1.1 (see assumptions (ii) and (iii) of the present theorem). Therefore, applying Theorem 1.1, we conclude that

$$\left(\int_{t_0}^{\vartheta} \left|u^{[v]}(t) - v^*(t|x(\cdot))\right|^2 dt\right)^{1/2} \leq \varepsilon_1.$$

Hence, $(u^{[x]}(\cdot), u^{[v]}(\cdot), y(\cdot)) \in \Phi(\varepsilon_1; x(\cdot))$. Since the trajectory q (2.26) is associated with the model process $(u^{[x]}(\cdot), u^{[v]}(\cdot), y(\cdot))$, we have $q \in \Phi^*(\varepsilon_1; x(\cdot))$. Due to the arbitrariness of the choice of the trajectory q within the set $D_N^0 \cap T(x(\cdot))$, the embedding (2.25) is proved. The proof of the theorem is complete.

ACKNOWLEDGMENTS

The first author acknowledges partial support of the Russian Foundation for Basic Research, project no. 09-01-00624-a.

REFERENCES

1. N. N. Krasovskii and A. I. Subbotin, *Positional Differential Games* (Nauka, Moscow, 1974); English transl.: *Game-Theoretical Control Problems* (Springer, New York, 1988).
2. A. N. Tikhonov and V. Ya. Arsenin, *Solutions of Ill-Posed Problems* (Nauka, Moscow, 1974; J. Wiley & Sons, New York, 1977).
3. A. N. Kotel'nikova, "A Problem of Game Stochastic Stabilization of a Controlled Object," Tr. Inst. Math. Mekh., Ural. Otd. Ross. Akad. Nauk **10** (2), 69–82 (2004) [Proc. Steklov Inst. Math., Suppl. 2, S76–S90 (2004)].
4. N. N. Krasovskii and A. N. Kotel'nikova, "One Problem on Stable Tracking of Motion," Tr. Inst. Math. Mekh., Ural. Otd. Ross. Akad. Nauk **12** (1), 142–157 (2006) [Proc. Steklov Inst. Math., Suppl. 1, S151–S167 (2006)].
5. J. Neveu, *Bases mathématiques du calcul des probabilités* (Masson, Paris, 1964); English transl.: *Mathematical Foundations of the Calculus of Probability* (Holden-Day, San Francisco, 1965).
6. R. S. Liptser and A. N. Shiryaev, *Statistics of Random Processes: Nonlinear Filtration and Related Questions* (Nauka, Moscow, 1974); English transl.: R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes* (Springer, New York, 1977, 1978), Vols. 1, 2.
7. N. N. Krasovskii, A. I. Subbotin, and V. F. Rossokhin, "Stochastic Strategies in Differential Games," Dokl. Akad. Nauk SSSR **220** (5), 1023–1026 (1975) [Sov. Math., Dokl. **16**, 198–202 (1975)].
8. N. N. Krasovskii and A. I. Subbotin, "Differential Games in Mixed Strategies," in *Problems of Analytical Mechanics, Stability Theory, and Control* (Nauka, Moscow, 1975), pp. 11–18 [in Russian].
9. N. N. Krasovskii, "Mixed Strategies in Differential Games," Dokl. Akad. Nauk SSSR **235** (3), 519–522 (1977) [Sov. Math., Dokl. **18**, 960–963 (1978)].
10. A. V. Kryazhimskii, "On Stochastic Approximation in Differential Games," Dokl. Akad. Nauk SSSR **241** (5), 1020–1023 (1978) [Sov. Math., Dokl. **19**, 955–959 (1978)].
11. A. V. Kryazhimskii and Yu. S. Osipov, "Modelling of a Control in a Dynamic System," Izv. Akad. Nauk SSSR, Ser. Tekh. Kibern., No. 2, 51–60 (1983) [Eng. Cybern. **21** (2), 38–47 (1984)].
12. Yu. S. Osipov and A. V. Kryazhimskii, *Inverse Problems for Ordinary Differential Equations*: *Dynamical Solutions* (Gordon and Breach, London, 1995).
13. Yu. S. Osipov, F. P. Vasil'ev, and M. M. Potapov, *Fundamentals of the Dynamic Regularization Method* (Mosk. Gos. Univ., Moscow, 1999) [in Russian].
14. V. I. Maksimov, *Problems of Dynamic Reconstruction of Inputs of Infinite-Dimensional Systems* (Inst. Mat. Mekh., Ural. Otd. Ross. Akad. Nauk, Yekaterinburg, 2000).
15. Yu. S. Osipov, A. V. Kryazhimskii, and V. I. Maksimov, "Dynamic Inverse Problems for Parabolic Systems," Diff. Uravn. **36** (5), 579–597 (2000) [Diff. Eqns. **36**, 643–661 (2000)].
16. A. Kryazhimskiy and V. Maksimov, "On Rough Inversion of a Dynamical System with a Disturbance," J. Inverse Ill-Posed Probl. **16**, 587–600 (2008).

Translated by I. Nikitin