ISSN 0005-1179, Automation and Remote Control, 2017, Vol. 78, No. 8, pp. 1438–1448. © Pleiades Publishing, Ltd., 2017.
Original Russian Text © V.D. Konakov, A.R. Markova, 2017, published in Avtomatika i Telemekhanika, 2017, No. 8, pp. 100–112.

STOCHASTIC SYSTEMS

Nonlinear Trend Exclusion Procedure for Models Defined by Stochastic Differential and Difference Equations

V. D. Konakov[*,a] **and A. R. Markova**[*,b]

[*] *National Research University Higher School of Economics, Moscow, Russia*
e-mail: [a] *VKonakov@hse.ru,* [b] *anna.r.markova@gmail.com*

Received September 19, 2016

Abstract—We consider a diffusion process and its approximation with a Markov chain whose trends contain a nonlinear unbounded component. The usual parametrix method is inapplicable here since the trend is unbounded. We present a procedure that lets us exclude a nonlinear growing trend and pass to a stochastic differential equation with bounded drift and diffusion coefficients. A similar procedure is also considered for a Markov chain.

Keywords: stochastic differential equation, diffusion process, Markov chains, parametrix method.

DOI: 10.1134/S0005117917080057

1. INTRODUCTION

This work is a continuation of our previous paper [1] that considered the case of a linear trend. It is known that the parametrix method is usually applied for stochastic differential equations (SDE) with bounded drift and diffusion coefficients (see [2–4]). However, in practice diffusions with unbounded coefficients occur quite often. The main goal of this work is to obtain a procedure that reduces the original equation with unbounded drift to an equation with bounded drift. The idea of this procedure is to consider a new process and, using Ito's formula, write down the stochastic differential for that process. We consider an auxiliary ordinary differential equation (ODE) that results by removing the Brownian component and the bounded part of the drift in the original SDE. The new process is constructed as the preimage of the original process under the action of backward phase flow of this ODE. A similar procedure is also given for Markov chains.

To apply Ito's formula, we need smoothness with respect to initial conditions and initial time moment. Therefore, we need the corresponding smoothness conditions for an unbounded trend component [5]. We note that to apply the procedure introduced in this work conditions of the existence and uniqueness theorem for ODE solutions must hold not only locally but also on the entire interval $[0, T]$. We also note that a local approach could be developed under weaker conditions.

2. CONDITIONS AND MAIN RESULT

Consider the following diffusion model:

$$\begin{aligned} dY_t &= \{F(t, Y_t) + m(t, Y_t)\}\, dt + \sigma(t, Y_t) dW_t, \\ Y_0 &= x_0 \in \mathbb{R}^d, \quad 0 \leqslant t \leqslant T, \end{aligned} \tag{1}$$

where $F(t, y)$ and $m(t, y)$ are d-dimensional vector functions, $\sigma(t, y)$ is a $(d \times d)$ diffusion matrix, W_t is a standard d-dimensional Wiener process.

In what follows we denote by $F_*(t, \mathbf{x})$ the derivative with respect to the spatial variable $\mathbf{x}$ for a fixed t. Thus, $F_*(t, \mathbf{x})$ is a linear operator from $\mathbb{R}^d$ to $\mathbb{R}^d$. We will use standard notation $D_x^\nu F(t, x)$ for higher order derivatives.

We introduce the following *assumptions*.

A1 (uniform ellipticity). Matrix $a := \sigma\sigma^*$ is uniformly elliptic, i.e., there exists $\Lambda \geqslant 1$ such that $\forall (t, x, \xi) \in [0, T] \times (\mathbb{R}^d)^2$

$$\Lambda^{-1} |\xi|^2 \leqslant \langle a(t, x)\xi, \xi \rangle \leqslant \Lambda |\xi|^2 .$$

A2 (smoothness of the trend component). Function $F(t, x) \in C^2 \left([0, T] \times \mathbb{R}^d \right)$ and

$$\max_{t \in [0,T]} \|D_x^\nu F(t, x)\| \leqslant M_F, \quad |\nu| = 1, 2.$$

Here $\|\cdot\|$ is the Euclidean norm of the vector (for $|\nu| = 1$) or matrix (for $|\nu| = 2$).

A3 (boundedness of the trend component). Components of the vector function $m(t, x)$ are continuous and bounded on $(t, x) \in [0, T] \times \mathbb{R}^d$.

Consider an ordinary differential equation corresponding to Eq. (1),

$$\frac{dx}{dt} = F(t, x) \tag{2}$$

and transformation $G : [0, T] \times \mathbb{R}^d \to [0, T] \times \mathbb{R}^d$ defined as

$$G(t, \mathbf{x}) = (t, \mathbf{g}(t; t_0, \mathbf{x})) ,$$

where

$$\mathbf{g}(t; t_0, \mathbf{x}) = (\mathbf{g}_1(t; t_0, \mathbf{x}), \ldots, g_d(t; t_0, \mathbf{x}))$$

is the solution of ODE (2) satisfying the initial condition $\mathbf{g}(t_0; t_0, \mathbf{x}) = \mathbf{x}$.

Now assumption **A2** and the two theorems formulated below imply that G is a diffeomorphism of class C^2. In particular, this gives us the necessary smoothness and lets us apply Ito's lemma to the process $\tilde{Y}_t$ introduced below.

Theorem A [5, p. 200]. *Suppose that the right-hand side of Eq.* (2) *is* $r \geqslant 1$ *times continuously differentiable in some neighborhood of the point* (t_0, x_0). *Then solution* $\phi(t)$ *with initial condition* $\phi(t_0) = x$ *is a differentiable function from class* C^r *jointly with respect to variables* t *and* x *when* t *and* x *vary in some* (*perhaps smaller*) *neighborhood of the point* (t_0, x_0):

$$F \in C^r \Rightarrow \phi \in C^r \quad \textit{for} \quad r \geqslant 1.$$

Note that under the assumptions of Theorem A continuous differentiability with respect to t_0 follows from Theorem 5.2.1 from [6].

Theorem B [7, p. 392]. *Suppose that a function* $F(t, x)$, *continuous in* t, *satisfies for* $t \in [a, b]$ *and* $x \in \mathbb{B}$ (where $\mathbb{B}$ *is a Banach space*) *the following conditions:*

$$\|F(t, x)\| \leqslant M_1 + M_0 \|x\| ,$$
$$\|F(t, x_2) - F(t, x_1)\| \leqslant M_2 \|x_2 - x_1\| ,$$

where M_i, $i = 0, 1, 2$, *are positive constants. Then for any* $x_0 \in B$ *and* $t_0 \in [a, b]$ *differential Eq.* (2) *has on the entire interval* $[a, b]$ *a unique solution* $x = \phi(t)$ *satisfying the initial condition* $\phi(t_0) = x_0$.

Let now $t_0 = 0$. Consider a new stochastic process

$$\tilde{Y}_t = \mathbf{g}^{-1}(0; t, Y_t), \quad \tilde{Y}_0 = Y_0 = x_0, \tag{3}$$

where Y_t is a solution of Eq. (1). We now formulate the main result of this work.

Theorem. *Suppose that assumptions* **A1**–**A3** *hold, and* Y_t *is a solution of SDE* (1) *with, generally speaking, unbounded trend. Then process* $\tilde{Y}_t$ *defined in* (3) *is a diffusion process with stochastic differential*

$$d\tilde{Y}_t = \tilde{m}\left(t, \tilde{Y}_t\right) dt + \tilde{\sigma}\left(t, \tilde{Y}_t\right) dW_t,$$

where

$$\tilde{m}\,(t, y) = \mathbf{g}_*^{-1}(t; 0, y) \left\{ m(t, \mathbf{g}(t; 0, y)) + \frac{1}{2} \sum_{i,j=1}^{d} c_{ij}(t, 0, y) \sum_{p=1}^{d} \sigma_{ip}(t, \mathbf{g}(t; 0, y)) \sigma_{jp}(t, \mathbf{g}(t; 0, y)) \right\},$$

$$\tilde{\sigma}(t, y) = \mathbf{g}_*^{-1}(t; 0, y) \sigma(t, \mathbf{g}(t; 0, y)).$$

The vectors $c_i j$ *are defined in* (12), *moreover, all trend components* $\tilde{m}(t, y)$ *are bounded.*

Proof of Theorem. The derivative $\mathbf{g}_*(t; t_0, \mathbf{x}) := \|z_{ik}(t; t_0, \mathbf{x})\|$, $z_{ik}(t; t_0, \mathbf{x}) = \frac{\partial \mathbf{g}_i(t; t_0, \mathbf{x})}{\partial x_k}$, $i, k = 1, \ldots, d$, of the solution of Eq. (2) with initial condition $\mathbf{x}$ satisfies the variational equation with initial condition $\mathbf{g}_*(t_0; t_0, \mathbf{x}) = \mathbf{I}$, where $\mathbf{I}$ is the unit matrix:

$$\begin{gathered} \frac{\partial}{\partial t} \mathbf{g}(t; t_0, \mathbf{x}) = F(t, \mathbf{g}(t; t_0, \mathbf{x})), \\ \frac{\partial}{\partial t} \mathbf{g}_*(t; t_0, \mathbf{x}) = F_*(t, \mathbf{g}(t; t_0, \mathbf{x})) \mathbf{g}_*(t; t_0, \mathbf{x}), \\ \mathbf{g}(t_0; t_0, \mathbf{x}) = x, \mathbf{g}_*(t_0; t_0, \mathbf{x}) = x_* = \mathbf{I} \end{gathered} \tag{4}$$

(for more details see [5, p. 225]). Moreover,

$$\frac{\partial}{\partial t_0} \mathbf{g}(t; t_0, \mathbf{x}) = -\mathbf{g}_*(t; t_0, \mathbf{x}) F(t_0, \mathbf{x}). \tag{5}$$

The proof of (5) is contained, for example, in [7]. It is clear that $\mathbf{g}_*(t; t_0, \mathbf{x})$ is the matrix of the differential at point $\mathbf{x}$ for the direct (t_0, t)-transform defined by the forward phase flow. Similarly, the inverse matrix $\mathbf{g}_*^{-1}(t; t_0, \mathbf{x})$ is the matrix of the differential for the (t, t_0)-inverse transform defined by the backward phase flow. Relation (5) can be interpreted geometrically as follows. The derivative of the solution with respect to the original time moment is the vector image of the tangent vector $-F(t_0, \mathbf{x})$ under the action of the differential at point $\mathbf{x}$ for the direct (t_0, t)-transform defined by the forward phase flow. The matrix of the differential for the inverse transform equals the inverse matrix of the differential of the transform defined by the forward phase flow. Symmetrically to (5) we get that

$$\frac{\partial \mathbf{g}^{-1}(t_0; t, \mathbf{y})}{\partial t} = -\mathbf{g}_*^{-1}(t; t_0, \mathbf{g}^{-1}(t_0; t, \mathbf{y})) F(t, \mathbf{y}). \tag{6}$$

Note that in the linear case $F(t, \mathbf{x}) = b(t)\mathbf{x}$, and using notation from [1] we get that

$$F(t, \mathbf{g}(t; t_0, \mathbf{x})) = b(t) \mathbf{g}(t; t_0, \mathbf{x}), \quad F_*(t, \mathbf{g}(t; t_0, \mathbf{x})) = b(t), \quad \mathbf{g}_*(t; t_0, \mathbf{x}) = \Phi(t),$$

where $\Phi(t)$ is the fundamental matrix corresponding to system (2) which in this case will be a linear system.

Now (4) and (6) imply relations for the linear case:

$$\Phi'(t) = b(t)\Phi(t), \quad \left[\Phi^{-1}(t)\right]' = -\left[\Phi(t)\right]^{-1} b(t).$$

We denote $\Psi(t,\mathbf{y}) := \mathbf{g}^{-1}(0;t,\mathbf{y})$ and apply the d-dimensional Ito's formula to process $\tilde{Y}_t = \Psi(t,Y_t)$. In order to apply Ito's formula, we need to compute the first and second derivatives with respect to the spatial variable $\mathbf{y}$ and first derivatives with respect to the time t. Differentiating the identity

$$\mathbf{g}_i\left(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y})\right) = y_i, \tag{7}$$

we get that

$$\frac{\partial}{\partial y_k}\mathbf{g}_i(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y})) = \sum_{l=1}^{d}\frac{\partial \mathbf{g}_i(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))}{\partial \mathbf{g}_l^{-1}(0;t,\mathbf{y})}\frac{\partial \mathbf{g}_l^{-1}(0;t,\mathbf{y})}{\partial y_k} = \frac{\partial y_i}{\partial y_k} = \delta_{ik},$$

and, consequently,

$$\mathbf{g}_*^{-1}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y})) = \left\|\frac{\partial \mathbf{g}_i^{-1}(0;t,\mathbf{y})}{\partial y_k}\right\| := \left\|z^{ik}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))\right\|. \tag{8}$$

Now (4) and Liouville's theorem [5] imply that for some $C > 1$

$$C^{-1} \leqslant \det \mathbf{g}_*(t;t_0,\mathbf{x}) = \exp\int_0^t \operatorname{trace}[F_*(s,\mathbf{g}(s;t_0,\mathbf{x}))]ds \leqslant C \tag{9}$$

and by Theorem 8.65 from [8]

$$\|\mathbf{g}_*(t;t_0,\mathbf{x})\| \leqslant C(d,T). \tag{10}$$

Now (8)–(10) imply that all elements $z^{ik}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))$ of the inverse matrix $\mathbf{g}_*^{-1}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))$ are bounded, and since functions $m(t,Y_t)$ and $\sigma(t,Y_t)$ are bounded then functions $\tilde{m}(t,Y_t) := \mathbf{g}_*^{-1}(t;0,\mathbf{g}^{-1}(0;t,Y_t))m(t,Y_t)$ and $\tilde{\sigma}(t,Y_t) := \mathbf{g}_*^{-1}(t;0,\mathbf{g}^{-1}(0;t,Y_t))\sigma(t,Y_t)$ are also bounded. Now (7) implies that

$$\frac{\partial^2}{\partial y_j\partial y_k}\mathbf{g}_i(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y})) = \sum_{l=1}^{d}\frac{\partial}{\partial y_j}\left[\frac{\partial \mathbf{g}_i(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))}{\partial \mathbf{g}_l^{-1}(0;t,\mathbf{y})}\right]\frac{\partial \mathbf{g}_l^{-1}(0;t,\mathbf{y})}{\partial y_k}$$
$$+\sum_{l=1}^{d}\frac{\partial \mathbf{g}_i(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))}{\partial \mathbf{g}_l^{-1}(0;t,\mathbf{y})}\frac{\partial^2 \mathbf{g}_l^{-1}(0;t,\mathbf{y})}{\partial y_j\partial y_k} = 0,$$

or, in our notation,

$$\frac{\partial^2}{\partial y_j\partial y_k}\mathbf{g}_i(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y})) = \sum_{l,p=1}^{d}\frac{\partial z_{il}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))}{\partial \mathbf{g}_p^{-1}(0;t,\mathbf{y})}z^{pj}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))$$
$$\times z^{lk}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y})) + \sum_{l=1}^{d} z_{il}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))\frac{\partial^2 \mathbf{g}_l^{-1}(0;t,\mathbf{y})}{\partial y_j\partial y_k} = 0. \tag{11}$$

Differentiating both parts of (4) with respect to $\mathbf{x}$ and applying **A2** and Gronwall's inequality [8, Theorem 8.65], we get that uniformly in $(t,\mathbf{y}) \in [0,T]\times\mathbb{R}^d$ and for $1 \leqslant i,l,p \leqslant d$

$$\left|\frac{\partial z_{il}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))}{\partial \mathbf{g}_p^{-1}(0;t,\mathbf{y})}\right| \leqslant C,$$

where the constant C depends only on T and M_F, which follows from assumption **A2**. Consequently,

$$c^i_{jk}(t,0,\mathbf{g}^{-1}(0;t,\mathbf{y})) := \sum_{l,p=1}^{d} \frac{\partial z_{il}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))}{\partial \mathbf{g}_p^{-1}(0;t,\mathbf{y})} z^{pj}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y}))$$
$$\times z^{lk}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y})) \leqslant C_1 < \infty.$$

We get from (11) that

$$\begin{pmatrix} \dfrac{\partial^2 \mathbf{g}_1^{-1}(0;t,\mathbf{y})}{\partial y_j \partial y_k} \\ \vdots \\ \dfrac{\partial^2 \mathbf{g}_d^{-1}(0;t,\mathbf{y})}{\partial y_j \partial y_k} \end{pmatrix} = -\mathbf{g}_*^{-1}(t;0,\mathbf{g}^{-1}(0;t,\mathbf{y})) c_{jk}(t,0,\mathbf{g}^{-1}(0;t,\mathbf{y})), \tag{12}$$

where

$$c_{jk}(t,0,\mathbf{g}^{-1}(0;t,\mathbf{y})) = \begin{pmatrix} c^1_{jk}(t,0,\mathbf{g}^{-1}(0;t,\mathbf{y})) \\ \vdots \\ c^d_{jk}(t,0,\mathbf{g}^{-1}(0;t,\mathbf{y})) \end{pmatrix},$$

and, consequently, all components of the vector in the left-hand side of (12) are bounded. Applying the d-dimensional Ito's formula, we get that

$$d\tilde{Y}_{t,k} = \frac{\partial \mathbf{g}_k^{-1}(0;t,Y_t)}{\partial t} dt + \sum_{i=1}^{d} \frac{\partial \mathbf{g}_k^{-1}(0;t,Y_t)}{\partial y_i} dY_{t,i} + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2 \mathbf{g}_k^{-1}(0;t,Y_t)}{\partial y_i \partial y_j} dY_{t,i} dY_{t,j}, \quad k = 1,\ldots,d.$$

For $d\tilde{Y}_t = (d\tilde{Y}_{t,1},\ldots,d\tilde{Y}_{t,d})^{\mathbf{T}}$ from (1), (6) and (12) we get that

$$\begin{aligned} d\tilde{Y}_t &= -\mathbf{g}_*^{-1}\left(t;0,\mathbf{g}^{-1}(0;t,Y_t)\right) F(t,Y_t)dt \\ &+ \mathbf{g}_*^{-1}\left(t;0,\mathbf{g}^{-1}(0;t,Y_t)\right) [(F(t,Y_t) + m(t,Y_t))dt + \sigma(t,Y_t)dW_t] \\ &+ \frac{1}{2}\mathbf{g}_*^{-1}(t;0,\mathbf{g}^{-1}(0;t,Y_t)) \sum_{i,j=1}^{d} c_{ij}(t,0,\mathbf{g}^{-1}(0;t,Y_t)) \sum_{p=1}^{d} [\sigma_{ip}(t,Y_t)\sigma_{jp}(t,Y_t)]dt \\ &= \mathbf{g}_*^{-1}(t;0,\tilde{Y}_t)\left\{ m(t,\mathbf{g}(t;0,\tilde{Y}_t)) + \frac{1}{2}\sum_{i,j=1}^{d} c_{ij}(t,0,\tilde{Y}_t)\sum_{p=1}^{d}[\sigma_{ip}(t,\mathbf{g}(t;0,\tilde{Y}_t))\sigma_{jp}(t,\mathbf{g}(t;0,\tilde{Y}_t))]\right\} dt \\ &+ \mathbf{g}_*^{-1}(t;0,\tilde{Y}_t)\sigma(t,\mathbf{g}(t;0,\tilde{Y}_t))dW_t := \tilde{m}(t,\tilde{Y}_t)dt + \tilde{\sigma}(t,\tilde{Y}_t)dW_t, \end{aligned} \tag{13}$$

where

$$\tilde{m}(t,\tilde{Y}_t) = \mathbf{g}_*^{-1}(t;0,\tilde{Y}_t)\left\{ m(t,\mathbf{g}(t;0,\tilde{Y}_t)) \right.$$
$$\left. + \frac{1}{2}\sum_{i,j=1}^{d} c_{ij}(t,0,\tilde{Y}_t)\sum_{p=1}^{d}\left[\sigma_{ip}(t,\mathbf{g}(t;0,\tilde{Y}_t))\sigma_{jp}(t,\mathbf{g}(t;0,\tilde{Y}_t))\right]\right\},$$
$$\tilde{\sigma}(t,\tilde{Y}_t) := \mathbf{g}_*^{-1}(t;0,\tilde{Y}_t)\sigma(t,\mathbf{g}(t;0,\tilde{Y}_t)).$$

We get that process $\tilde{Y}_t$ satisfies Ito's SDE (13) with *bounded coefficients* $\tilde{m}(t,y)$ and $\tilde{\sigma}(t,y)$.

This completes the proof of the theorem.

In the linear case, all vectors $c_{jk} = 0$ since $\mathbf{g}_*(t;t_0,\mathbf{x}) = \Phi(t)$ does not depend on the initial condition $\mathbf{x}$.

3. EXAMPLE

Consider nonlinear unbounded trend exclusion in the stochastic differential equation

$$dY_t = \sqrt{Y_t^2 + 1}dt + \sigma dW_t, \quad Y_0 = x_0, \quad \sigma > 0, \quad t \in [0, T]. \tag{14}$$

In this example, $F(t, x) = \sqrt{x^2 + 1}$, $a = \sigma^2$, $m(t, x) = 0$, and assumptions **A1**–**A3** hold. The differential equation

$$\frac{dx(t)}{dt} = \sqrt{x^2(t) + 1}$$

can be solved by separating the variables

$$\int\limits_{x_0}^{x} \frac{dy}{\sqrt{y^2 + 1}} = \ln\left(y + \sqrt{y^2 + 1}\right)\Big|_{x_0}^{x} = \ln\left(\frac{x + \sqrt{x^2 + 1}}{x_0 + \sqrt{x_0^2 + 1}}\right) = \int\limits_{t_0}^{t} ds = t - t_0.$$

Resolving the latter equation with respect to x, we get that

$$x = x(t; t_0, x_0) = \frac{A^2 - 1}{2A}, \quad A = e^{t - t_0}\left(x_0 + \sqrt{x_0^2 + 1}\right).$$

And vice versa, by given $(t, x(t))$ with the backward phase flow we find that

$$x_0 + \sqrt{x_0^2 + 1} = e^{t_0 - t}\left(x + \sqrt{x^2 + 1}\right),$$
$$x_0(t_0; t, x(t)) = \frac{B^2 - 1}{2B}, \quad B = e^{t_0 - t}\left(x + \sqrt{x^2 + 1}\right).$$

In what follows $t_0 = 0$, and the nonnegative stochastic process $\tilde{Y}_t$ is defined as follows:

$$\tilde{Y}_t = e^{-t}\left(Y_t + \sqrt{Y_t^2 + 1}\right), \quad Y_t = \frac{e^{2t}\tilde{Y}_t^2 - 1}{2e^t\tilde{Y}_t},$$

where Y_t is a solution of SDE (14). By Ito's formula we find the stochastic differential process $\tilde{Y}_t$:

$$\begin{aligned} d\tilde{Y}_t &= -e^{-t}\left(Y_t + \sqrt{Y_t^2 + 1}\right)dt + e^{-t}\left(1 + \frac{Y_t}{\sqrt{Y_t^2 + 1}}\right)\left(\sqrt{Y_t^2 + 1}dt + \sigma dW_t\right) \\ &+ \frac{1}{2}e^{-t}\frac{\sigma^2}{(1 + Y_t^2)^{3/2}}dt = \frac{1}{2}e^{-t}\frac{\sigma^2}{(1 + Y_t^2)^{3/2}}dt + e^{-t}\sigma\left(1 + \frac{Y_t}{\sqrt{Y_t^2 + 1}}\right)dW_t \\ &= \frac{\sigma^2}{4}\frac{e^{2t}\tilde{Y}_t^3}{\left(e^{2t}\tilde{Y}_t^2 + 1\right)^3}dt + \sigma\frac{2e^t\tilde{Y}_t^2}{e^{2t}\tilde{Y}_t^2 + 1}dW_t. \end{aligned}$$

The trend and diffusion of process $\tilde{Y}_t$ are bounded in $[0, T] \times \mathbb{R}$:

$$\tilde{m}(t, x) = \frac{\sigma^2}{4}\frac{e^{2t}x^3}{(e^{2t}x^2 + 1)^3}, \quad \tilde{\sigma}(t, x) = \frac{2\sigma e^t x^2}{e^{2t}x^2 + 1}, \quad t \in [0, T].$$

4. MARKOV CHAINS

Suppose that assumption **A1**–**A3** hold.

By Hadamard's global inversion theorem, $G : \mathbb{R}^d \to \mathbb{R}^d : x \to x + hF(t, x)$ is a C^2 diffeomorphism for every $t \in [0, T]$ and sufficiently small h.

We consider a (generally speaking, non-uniform) partition $0 = t_0^n < t_1^n < \ldots < t_n^n = T$ and Markov chain

$$X(t_{k+1}^n) = X(t_k^n) + h_k^n \left\{F(t_k^n, X(t_k^n)) + m(t_k^n, X(t_k^n))\right\} + \sqrt{h_k^n}\sigma\left(t_k^n, X(t_k^n)\right)\varepsilon(t_{k+1}^n), \tag{15}$$
$$X(0) = x \in \mathbb{R}^d, \quad k = 0, \ldots, n-1, \quad h_k^n = t_{k+1}^n - t_k^n.$$

A4. There exists a constant $C > 1$ such that

$$C^{-1} \leqslant \frac{h_k^n}{h_l^n} \leqslant C \quad \text{for} \quad n \geqslant 1 \quad \text{and} \quad 1 \leqslant k, l \leqslant n,$$
$$\lim_{n\to\infty} h_1^n = 0.$$

Consider the difference equation

$$\frac{\widehat{\mathbf{g}}(t_{k+1}^n; 0, x) - \widehat{\mathbf{g}}(t_k^n; 0, x)}{h_k^n} = F(t_k^n, \widehat{\mathbf{g}}(t_k^n; 0, x)), \quad \widehat{\mathbf{g}}(0; 0, x) = x, \tag{16}$$

where $\widehat{\mathbf{g}}(t_k^n; 0, x)$ are Euler broken lines corresponding to the partition $0 = t_0^n < t_1^n < \ldots < t_n^n = T$, that start at point x at time moment 0 and are constructed for the solution $\mathbf{g}(t; 0, x)$ of equation $\dot{\mathbf{x}} = F(t, \mathbf{x})$, $\mathbf{g}(0; 0, x) = x$. If we let $\widehat{X}(t_k^n) := \widehat{\mathbf{g}}^{-1}(0; t_k^n, X(t_k^n))$ and replace x with $\widehat{X}(t_k^n)$ in (16), we will get that

$$\frac{\widehat{\mathbf{g}}(t_{k+1}^n; 0, \widehat{X}(t_k^n)) - \widehat{\mathbf{g}}(t_k^n; 0, \widehat{X}(t_k^n))}{h_k^n} = F(t_k^n, \widehat{\mathbf{g}}(t_k^n; 0, \widehat{X}(t_k^n))),$$
$$\widehat{X}(t_k^n) = \widehat{\mathbf{g}}(0; 0, \widehat{X}(t_k^n)).$$

Then (15) can be rewritten as

$$\widehat{\mathbf{g}}\left(t_{k+1}^n; 0, \widehat{X}(t_{k+1}^n)\right) = \widehat{\mathbf{g}}\left(t_{k+1}^n; 0, \widehat{X}(t_k^n)\right) + h_k^n m\left(t_k^n, \widehat{\mathbf{g}}(t_k^n; 0, \widehat{X}(t_k^n))\right)$$
$$+ \sqrt{h_k^n}\sigma\left(t_k^n, \widehat{\mathbf{g}}(t_k^n; 0, \widehat{X}(t_k^n))\right)\varepsilon(t_{k+1}^n).$$

Iterating (16), we get that

$$\widehat{\mathbf{g}}(t_{k+1}^n; 0, x) := L_{t_{k+1}^n}^n(x),$$

where $L_t^n(x)$ are Euler broken lines for equations $\dot{\mathbf{x}} = F(t, \mathbf{x})$ starting at point x and corresponding to the partition $0 = t_0^n < t_1^n < \ldots < t_n^n = T$:

$$L_{t_1^n}^n(x) = x + h_1^n F(0, x),$$
$$\ldots$$
$$L_{t_{k+1}^n}^n(x) = L_{t_k^n}^n(x) + h_{k+1}^n F(t_k^n, L_{t_k^n}^n(x)).$$

Iterating the latter equation, we get that

$$\widehat{\mathbf{g}}(t_{k+1}^n; 0, x) = x + \sum_{j=0}^{k} h_{j+1}^n F(t_j^n, \widehat{\mathbf{g}}(t_j^n; 0, x)), \tag{17}$$
$$\widehat{\mathbf{g}}(t_k^n; 0, x) = (\widehat{\mathbf{g}}_1(t_k^n; 0, x), \ldots, \widehat{\mathbf{g}}_d(t_k^n; 0, x))^{\mathbf{T}}.$$

We introduce the Jacobi matrix $\widehat{\mathbf{g}}_*(t_k^n;0,x) := \|\widehat{z}_{ij}(t_k^n;0,x)\|$, $\widehat{z}_{ij}(t_k^n;0,x) := \frac{\partial \widehat{\mathbf{g}}_i(t_k^n;0,x)}{\partial x_j}$. The inverse matrix is given by $\widehat{\mathbf{g}}_*^{-1}(t_k^n,0,y) := \|\widehat{z}^{\,ij}(t_k^n;0,y)\|$, $\widehat{z}^{\,ij}(t_k^n;0,y) := \frac{\partial \widehat{\mathbf{g}}_i^{-1}(0;t_k^n,y)}{\partial y_j}$. Differentiating (17) with respect to x, we get that

$$\widehat{\mathbf{g}}_*(t_{k+1}^n;0,x) = I + \sum_{j=0}^{k} h_{j+1}^n F_*(t_j^n, \widehat{\mathbf{g}}(t_j^n;0,x))\widehat{\mathbf{g}}_*(t_j^n;0,x), \tag{18}$$

$$\|\widehat{\mathbf{g}}_*(t_{k+1}^n;0,x)\| \leqslant \sqrt{d} + K\sum_{j=0}^{k} h_{j+1}^n \left\|\widehat{\mathbf{g}}_*(t_j^n;0,x)\right\|. \tag{19}$$

To estimate the left-hand side of (19), we use the following lemma.

Lemma (discrete Gronwall's lemma [8]). *Let $\{y_n\}$ and $\{g_n\}$ be two nonnegative sequences, and let c be a nonnegative constant. Suppose that*

$$y_{k+1} \leqslant c + \sum_{j=0}^{k} g_j y_j, \quad k \geqslant 0,$$

then

$$y_{k+1} \leqslant c + \prod_{j=0}^{k}(1+g_j) \leqslant c\exp\left(\sum_{j=0}^{k} g_j\right).$$

We use this assumption with $y_k = \|\widehat{\mathbf{g}}_*(t_k^n;0,x)\|$, $c = \sqrt{d}$, $g_j = Kh_{j+1}^n$, $j = 0,\dots,k$. We get from (19) that

$$\|\widehat{\mathbf{g}}_*(t_{k+1}^n;0,x)\| \leqslant \sqrt{d}\exp(KT), \quad k = 0,\dots,n-1.$$

We can now check by substitution that Eq. (18) has an explicit solution

$$\widehat{\mathbf{g}}_*(t_m^n;0,x) = \prod_{j=m-1}^{0}[I + h_{j+1}^n F_*(t_j^n, \widehat{\mathbf{g}}(t_j^n;0,x))] \tag{20}$$

(where we let $\prod_{j=-1}^{0}[I + h_{j+1}^n F_*(jh, \widehat{\theta}_{jh}(x))] = I$). Indeed, substituting (20) into (18), we get that

$$\begin{aligned} &I + \sum_{j=0}^{k}\left\{[I + h_{j+1}^n F_*(t_j^n, \widehat{\mathbf{g}}(t_j^n;0,x)) - I]\prod_{l=j-1}^{0}[I + h_{l+1}^n F_*(t_l^n, \widehat{\mathbf{g}}(t_l^n;0,x))]\right\} \\ &= I + \sum_{j=0}^{k}\left\{\prod_{l=j}^{0}[I + h_{l+1}^n F_*(t_l^n, \widehat{\mathbf{g}}(t_l^n;0,x))] - \prod_{l=j-1}^{0}[I + h_{l+1}^n F_*(t_l^n, \widehat{\mathbf{g}}(t_l^n;0,x))]\right\} \\ &= I + \prod_{l=k}^{0}[I + h_{l+1}^n F_*(t_l^n, \widehat{\mathbf{g}}(t_l^n;0,x))] - I = \widehat{\mathbf{g}}_*(t_{k+1}^n;0,x). \end{aligned}$$

If every matrix $I + h_{l+1}^n F_*(t_l^n, \widehat{\mathbf{g}}(t_l^n;0,x))$, $l = 0,1,\dots,k$, is invertible then $\widehat{\mathbf{g}}_*(t_{k+1}^n;0,x)$ is also invertible:

$$\widehat{\mathbf{g}}_*^{-1}(t_{k+1}^n;0,x) = \prod_{j=0}^{k}[I + h_{j+1}^n F_*(t_j^n, \widehat{\mathbf{g}}(t_j^n;0,x))]^{-1}.$$

By assumptions **A2** and **A4** $\|h^n_{j+1}F_*(t^n_j,\widehat{\mathbf{g}}(t^n_j;0,x))\| \leqslant h^n_{j+1}K < \frac{1}{2}$ for sufficiently large n, so each matrix $I + h^n_{j+1}F_*(t^n_j,\widehat{\mathbf{g}}(t^n_j;0,x))$, $j = 0,1,\dots,k$, is invertible. Consequently,

$$\max_{0\leqslant m\leqslant n}\left\|\widehat{\mathbf{g}}_*^{-1}(t^n_m;0,x)\right\| \leqslant \max_{0\leqslant m\leqslant n}\prod_{j=0}^{m-1}\left\|[I + h^n_{j+1}F_*(t^n_j,\widehat{\mathbf{g}}(t^n_j;0,x))]^{-1}\right\|$$

$$\leqslant \max_{1\leqslant m\leqslant n}\prod_{j=0}^{m-1}\frac{1}{1-\left\|h^n_{j+1}F_*(t^n_j,\widehat{\mathbf{g}}(t^n_j;0,x))\right\|}$$

$$\leqslant \max_{1\leqslant m\leqslant n}\prod_{j=0}^{m-1}\left(1+\frac{Kh^n_{j+1}}{1-Kh^n_{j+1}}\right) \leqslant (1+2Kh^n_{j+1})^n \leqslant C(K,T)$$

for sufficiently large n. Passing in (18) to the limit $n\to\infty$, we get $\widehat{\mathbf{g}}_*(\phi^n(t);0,x)\to\mathbf{g}_*(t;0,\mathbf{x})$, $\phi^n(t) = \inf\{t^n_i : t^n_i \leqslant t < t^n_{i+1}\}$, where $\mathbf{g}_*(t;0,\mathbf{x})$ is a solution of equation

$$\mathbf{g}_*(t;0,\mathbf{x}) = I + \int_0^t F_*(u,\mathbf{g}(u;0,\mathbf{x}))\mathbf{g}_*(u;0,\mathbf{x})du,$$

or

$$\frac{\partial}{\partial t}\mathbf{g}_*(t;0,\mathbf{x}) = F_*(t,\mathbf{g}(t;0,\mathbf{x}))\mathbf{g}_*(t;0,\mathbf{x}), \mathbf{g}_*(0;0,\mathbf{x}) = \mathbf{I}.$$

As we have noted above, $x\to x+hF(t,x)$ is a C^2-diffeomorphism for every $t\in[0,T]$ and sufficiently small h. This implies that $\widehat{\mathbf{g}}^{-1}(0;t^n_k,y)$ is uniquely determined by and smoothly depends on y. Consequently, the mapping $\mathbb{R}^d\to\mathbb{R}^d : x\to\widehat{\mathbf{g}}(t^n_k;0,x)$ is a C^2-diffeomorphism. The differential of this mapping is given by the Jacobi matrix $\widehat{\mathbf{g}}_*(t^n_k;0,x) := \|\widehat{z}_{ij}(t^n_k;0,x)\|$, $\widehat{z}_{ij}(t^n_k;0,x) = \frac{\partial\widehat{\mathbf{g}}_i(t^n_k;0,x)}{\partial x_j}$. The differential of the inverse mapping is given by matrix $\widehat{\mathbf{g}}_*^{-1}(0;t^n_k,y) := \|\widehat{z}^{\,ij}(0;t^n_k,y)\|$, $\widehat{z}^{\,ij}(0;t^n_k,y) = \frac{\partial\widehat{\mathbf{g}}^{-1}_{*i}(0;t^n_k,y)}{\partial y_j}$. We define

$$\widetilde{X}(t^n_k) = \widehat{\mathbf{g}}^{-1}(0;t^n_k,X(t^n_k)).$$

Then we have that

$$\widetilde{X}(t^n_{k+1}) - \widetilde{X}(t^n_k) = \widehat{\mathbf{g}}^{-1}(0;t^n_{k+1},X(t^n_{k+1})) - \widehat{\mathbf{g}}^{-1}(0;t^n_{k+1},\widehat{\mathbf{g}}(h^n_k;t^n_k,X(t^n_k)))$$

$$= \left(\int_0^1 \widehat{\mathbf{g}}_*^{-1}(0;t^n_{k+1},\Psi_u(\widetilde{X}(t^n_k),\varepsilon(t^n_k)))du\right)(X(t^n_{k+1}) - \widehat{\mathbf{g}}(h^n_k;t^n_k,X(t^n_k)))$$

$$= h^n_k\left(\int_0^1 \widehat{\mathbf{g}}_*^{-1}(0;t^n_{k+1},\Psi_u(\widetilde{X}(t^n_k),\varepsilon(t^n_k)))du\right)m(t^n_k,\widehat{\mathbf{g}}(t^n_k;0,\widetilde{X}(t^n_k)))$$

$$+ \sqrt{h^n_k}\left(\int_0^1 \widehat{\mathbf{g}}_*^{-1}(0;t^n_{k+1},\Psi_u(\widetilde{X}(t^n_k),\varepsilon(t^n_k)))du\right)\sigma(t^n_k,\widehat{\mathbf{g}}(t^n_k;0,\widetilde{X}(t^n_k)))\varepsilon(t^n_{k+1}),$$

where

$$\Psi_u(\widetilde{X}(t^n_k),\varepsilon(t^n_k)) := \widehat{\mathbf{g}}(h^n_k;t^n_k,X(t^n_k)) + u[X(t^n_{k+1}) - \widehat{\mathbf{g}}(h^n_k;t^n_k,X(t^n_k))]$$

$$= X(t^n_k) + h^n_kF(t^n_k,X(t^n_k)) + uh^n_km(h^n_k,X(t^n_k)) + u\sqrt{h^n_k}\sigma(h^n_k,X(t^n_k))\varepsilon(t^n_{k+1})$$

$$= \widehat{\mathbf{g}}(t^n_k;0,\widetilde{X}(t^n_k)) + h^n_kF(t^n_k,\widehat{\mathbf{g}}(t^n_k;0,\widetilde{X}(t^n_k))) + uh^n_km(t^n_k,\widehat{\mathbf{g}}(t^n_k;0,\widetilde{X}(t^n_k)))$$

$$+u\sqrt{h^n_k}\sigma(t^n_k,\widehat{\mathbf{g}}(t^n_k;0,\widetilde{X}(t^n_k)))\varepsilon(t^n_{k+1}).$$

We denote now

$$\widetilde{m}(t_k^n, \widetilde{X}(t_k^n), \varepsilon(t_{k+1}^n)) = \left(\int_0^1 \widehat{\mathbf{g}}_*^{-1}(0; t_{k+1}^n, \Psi_u(\widetilde{X}(t_k^n), \varepsilon(t_{k+1}^n))) du \right) m(t_k^n, \widehat{\mathbf{g}}(t_k^n; 0, \widetilde{X}(t_k^n))), \tag{21}$$

$$\widetilde{\sigma}(t_k^n, \widetilde{X}(t_k^n), \varepsilon(t_{k+1}^n)) = \left(\int_0^1 \widehat{\mathbf{g}}_*^{-1}(0; t_{k+1}^n, \Psi_u(\widetilde{X}(t_k^n), \varepsilon(t_{k+1}^n))) du \right) \sigma(t_k^n, \widehat{\mathbf{g}}(t_k^n; 0, \widetilde{X}(t_k^n))). \tag{22}$$

This leads to the Markov chain

$$\widetilde{X}(t_{k+1}^n) = \widetilde{X}(t_k^n) + h_k^n \widetilde{m}(t_k^n, \widetilde{X}(t_k^n), \varepsilon(t_{k+1}^n)) + \sqrt{h_k^n} \widetilde{\sigma}(t_k^n, \widetilde{X}(t_k^n), \varepsilon(t_{k+1}^n)) \varepsilon(t_{k+1}^n) \tag{23}$$

with bounded coefficients $\widetilde{m}(t_k^n, \widetilde{X}(t_k^n), \varepsilon(t_{k+1}^n))$ and $\widetilde{\sigma}(t_k^n, \widetilde{X}(t_k^n), \varepsilon(t_{k+1}^n))$.

Remark. Markov chain (23), obtained as a result of the above-mentioned trend exclusion procedure, differs from the original Markov chain (15). The difference is that now both trend and diffusion depend on innovations $\varepsilon(t_{k+1}^n)$. This more general form of Markov chains was studied in [9]. The authors of [9] considered the following class of Markov chains:

$$X(t_{k+1}^n) = \Psi_k \left(X(t_k^n), \frac{\varepsilon_{k+1}}{\sqrt{n}}, \frac{1}{n} \right),$$

where $\Psi_k : \mathbb{R}^d \times \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}^d$ are smooth functions, and ε_k, $k \in \mathbb{N}$, is a sequence of independent centered random vectors (innovations). In that work, they consider the case when $t_k^n = \frac{k}{n}$ and study the convergence $Ef(X(t)) \to Ef(Y(t))$, where $Y(t)$ is a continuous time Markov process. The proof is based on a new version of the Malliavin calculus proposed by the authors. As it usually happens with such an approach, they assume infinite differentiability of the functions Ψ_k. In the present work, this corresponds to infinite differentiability of functions F, m, and σ in the model (15). The main result of [9] was obtained under assumptions (1.9)–(1.11) from [9], which, as it is easy to see, does not hold for the model (15) since function F is unbounded. It suffices to let $\alpha = \gamma = 0$ and $\beta = 1$ in (1.10) from [9] to see that this condition does not hold. On the contrary, for the dynamics (23) one can show a whole class of models with infinitely differentiable F, m, and σ for which these conditions hold, which lets us apply the results of [9].

In the linear case

$$F(t,x) = b(t)x, F_*(t_j^n, \cdot) = b(t_j^n)$$

and

$$\widehat{\mathbf{g}}_*^{-1}(t_{k+1}^n; 0, \Psi_u(\widetilde{X}(t_k^n))) = \prod_{j=0}^{k} [I + h_{j+1}^n b(t_j^n)]^{-1}. \tag{24}$$

If we substitute (24) into (21) and (22), we see that (23) in the case of a uniform partition coincides with the equation obtained in [1].

5. CONCLUSION

In this work, we have shown a procedure that lets us reduce the original equation with unbounded trend to an equation with bounded trend. A similar procedure is also given for Markov chains. This work is a continuation of our previous work [1] on a modification of the parametrix method. In what follows we plan to consider a procedure that lets one reduce an original equation with unbounded diffusion to an equation with bounded diffusion.

ACKNOWLEDGMENTS

The study has been funded by the Russian Academic Excellence Project "5–100."

REFERENCES

1. Konakov, V.D. and Markova, A.R., Linear Trend Exclusion for Models Defined with Stochastic Differential and Difference Equations, *Autom. Remote Control*, 2015, vol. 76, no. 10, pp. 1771–1783.
2. Il'in, A.M., Kalashnikov, A.S., and Oleinik, O.A., Second Order Linear Equations of Parabolic Type, *Usp. Mat. Nauk*, 1962, vol. 17, no. 3, pp. 3–146.
3. Friedman, A., *Partial Differential Equations of Parabolic Type*, Englewood Cliffs: Prentice Hall, 1964.
4. Konakov, V.D., *The Parametrix Method for Diffusions and Markov Chains, Preprint*, Moscow: Izd. Popech. Soveta Mekhaniko-Matematicheskogo Fakul'teta MGU, Ser. WP BRP "STI," 2012.
5. Arnol'd, V.I., *Obyknovennye differentsial'nye uravneniya* (Ordinary Differential Equations), Moscow: Nauka, 1971.
6. Bibikov, Yu.N., *Kurs obyknovennykh differentsial'nykh uravnenii* (A Course in Ordinary Differential Equations), Moscow: Vysshaya Shkola, 1991.
7. Daletskii, Yu.L. and Krein, M.G., *Ustoichivost' reshenii differentsial'nykh uravnenii v banakhovom prostranstve* (Stability of Solutions for Differential Equations in a Banach Space), Moscow: Nauka, 1970.
8. Kelley, W. and Peterson, A., *The Theory of Differential Equations, Classical and Qualitative*, Upper Saddle River: Prentice Hall, 2004.
9. Bally, V. and Rey C., Approximation of Markov Ssemigroups in Total Variation Distance, *Electron. J. Probab.*, 2016, no. 12, pp. 1–44.

This paper was recommended for publication by A.V. Nazin, a member of the Editorial Board