Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 1–11

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 1–11

ON COUPLING MOMENT INTEGRABILITY FOR TIME-INHOMOGENEOUS MARKOV CHAINS

UDC 519.21

V. V. GOLOMOZIY AND N. V. KARTASHOV

Abstract. In this paper, we find the conditions under which the expectation of the first coupling moment for two independent, discrete, time-inhomogeneous Markov chains will be finite. We consider discrete chains with a phase space $\{0, 1, \dots\}$ and as the coupling moment we understand the first moment of visiting zero state by the both chains at the same time.

Анотація. В даній роботі знаходяться умови за яких гарантовано існування математичного сподівання моменту склеювання для двох незалежних, дискретних, неоднорідних за часом Марківських ланцюгів. Розглядаються дискретні ланцюги з фазовим простором $\{0, 1, \dots\}$ та під моментом склеювання розуміється перший момент одночасного потрапляння в нуль обох ланцюгів.

Аннотация. В даной работа рассматриваются условия при которых гарантировано существует конечное математическое ожидание момента склеивания для двух независимых, дискретных, неоднородных по времени цепей Маркова. Рассматриваются дискретные цепи с пространством состояний $\{0, 1, \dots\}$ и под моментом склеивания мы понимаем первый момент одновременного попадания в нулевое состояние обеих цепей.

1. Introduction

The problem of finiteness for the moment of simultaneous hitting for two chains into certain set (or simultaneous renewal of two renewal processes) play a crucial role in evaluation of the stability estimates using coupling method. Similar estimates one can find in the authors' works [4, 5]. The problem of stability for a time-inhomogeneous Markov chain is investigated there using a coupling method as a key method of the research. Similar problems, but for the homogeneous Markov chains, are also considered in the work [7].

The key question for the stability estimate evaluation in these papers is how we can estimate the expectation for the moment of simultaneous hitting for two Markov chains. The coupling setup can be found in the following work [5].

The problem of integrability and finiteness for the coupling moment can be reduced to the problem of integrability and finiteness for the moment of simultaneous hitting into certain set or to the problem of finiteness for the moment of simultaneous renewal. Similar task is considered in the Lindvall's book [14]. It worth to mention, that this monograph is a classical book on the coupling method. There introduced different types of coupling: week coupling, maximal coupling, Ornshtein coupling, Mineka coupling and so on. Another famous book on the coupling method is a Torrison's work [15].

The coupling method is also used in many other works. The first works on coupling method are [1, 12, 13]. An example of how the coupling method is used to establish stability estimates for time-homogeneous chain with different initial distributions is proposed in [2].

2000 *Mathematics Subject Classification.* Primary 60J45; Secondary 60A05, 60K05.

Key words and phrases. Coupling theory, coupling method, maximal coupling, discrete Markov chains, stability of distributions, дискретні ланцюги Маркова, стійкість розподілів, метод склеювання, теорія склеювання.

However, the problem of coupling only the same homogeneous Markov chain and related problems were considered in books mentioned above. In particular, the theorem about integrability of the coupling moment in the book by Lindvall [14] had been proved for two copies of the same time-homogeneous Markov chain whith different initial distributions. In the investigation of stability there arises the necessity to extend coupling moment for different, not necessarily homogeneous Markov chains. So, well-known classical Lindvall's and Thorrison's results do not work in this case. Meantime, it is important to note that the main theorem of this article uses the same proof schema as Lindvall's theorem 4.2 [14, p. 27].

The paper [9] is devoted to the investigation of such problem as integrability of the coupling moment for two different Markov chains. In this work the estimates for the expectation of a coupling moment for two different time-homogeneous Markov chains starting with a random delays are presented. The conditions under which these estimates were obtained are the strong aperiodicity ($g_1^1+g_1^2 > 0$) and the finiteness of second renewal moments.

The maximal coupling for two time-inhomogeneous chains is considered in other author's papers [10, 11].

In the current paper these results extended to the time-inhomogeneous case. It is important that in this case the fundamental principle of independence of the renewal times does not hold true anymore. Instead, the conditional independence should be considered given the fixed moments of the previous renewal process.

The main theorem of this paper gives the general conditions which guarantee the integrability of the coupling moment. They are the condition of the separation from a zero for renewal probabilities (in the time-homogeneous case this condition automatically holds true for the non-periodic renewal distribution with a finite mean) and the uniform integrability of the renewal distributions. It is interesting that similarity of the condition can be noted for homogeneous and inhomogeneous case. In particular, for the time-homogeneous case, an estimate similar to the one from the work [9] is derived in a principal different way.

2. Dependence of renewal moments for time-inhomogeneous Markov chain

The fundamental fact defining the proof schema in the time-inhomogeneous case is that elements of a renewal sequence are not independent and the distribution of the $k+1$-st renewal moment is completely defined by the k-th renewal value.

Let's examine an example that leads to the renewal sequence generated by the time-inhomogeneous Markov chain.

Consider some time-inhomogeneous discrete Markov chain $(X_t, t \geq 0)$ with a phase space $\{0, 1, 2, \dots\}$. Its transition probabilities are defined in the following way:

$$\mathsf{P}\{X_{t+1} = j \mid X_t = i\} = P_t(i,j) = p_{ij}^{(t)}, \qquad t \geq 0. \tag{1}$$

In the zero moment of time the chain is in the zero state. Let's introduce the following notation:

$$\begin{aligned} \theta_1 &= \inf\{t > 0 \colon X_t = 0\} \\ \theta_2 &= \inf\{t > \theta_1 \colon X_t = 0\} \\ &\dots \\ \theta_m &= \inf\{t > \theta_{m-1} \colon X_t = 0\}, \qquad m > 1, \end{aligned} \tag{2}$$

where θ_1 is time of the first returning to zero, θ_2 is time between first and second zero hitting, and so on. In this case $\tau_k = \sum_{j=1}^{k} \theta_k$ is the k-th hitting moment.

The sequence $\{\theta_m, m \geq 1\}$ is a renewal sequence generated by the time-inhomogeneous Markov chain X_t. In general case, for the chain starting from a non-zero state we may consider an initial delay θ_0. It is time that a chain takes till hitting zero for the first time.

Let's now investigate a problem of dependence for the θ_m variables. In the homogeneous case, these variables are independent. But if the chain is time-inhomogeneous there is dependence between θ_m's. Let's see an example below.

The random variable θ_1 has a following distribution:

$$\begin{aligned} pr\{\theta_1 = k\} &= \mathsf{P}\{X_k = 0, X_{k-1} \neq 0, \dots, X_1 \neq 0, X_0 = 0\} \\ &= \sum_{i_0=0, i_1\neq 0, i_2\neq 0, \dots, i_{k-1}\neq 0, i_k=0} \prod_{j=0}^{k-1} p^{(j)}_{i_j i_{j+1}}. \end{aligned} \tag{3}$$

So, we can see that a distribution potentially depends from all X_t, $t \leq k$.

The distribution of the random variable θ_2 is as follows

$$\begin{aligned} \mathsf{P}\{\theta_2 = k\} &= \sum_{j=1}^{k-1} \mathsf{P}\{\theta_2 = k, \theta_1 = j\} \\ &= \sum_{j} \mathsf{P}\{X_k = 0, X_{k-1} \neq 0, X_{j+1} \neq 0, X_j = 0, \\ &\qquad\qquad X_{j-1} \neq 0, \dots, X_1 \neq 0, X_0 = 0\}. \end{aligned} \tag{4}$$

Note, that for each term in the last sum, the following equality holds true:

$$\begin{aligned} &\sum \mathsf{P}\{X_k = 0, X_{k-1} \neq 0, X_{j+1} \neq 0 \mid X_j = 0\}\, \mathsf{P}\{\theta_1 = j\} \\ &\quad = \sum \mathsf{P}\{X_k = 0, X_{k-1} \neq 0, X_{j+1} \neq 0 \mid X_j = 0\}\, \mathsf{P}\{\tau_1 = j\}. \end{aligned}$$

So, the distribution of the random variable θ_2 depends on the variable τ_1 and all X_t, $t > \tau_1$. We'll show that this situation holds true for the other θ_m as well.

Let us now consider

$$\mathsf{P}\{\theta_m = k\} = \sum \mathsf{P}\{X_k = 0, X_{k-1} \neq 0, \dots, X_{j+1} \neq 0 \mid X_j = 0\}\, \mathsf{P}\{\tau_{m-1} = j\}. \tag{5}$$

So the distribution of the θ_m depends on probabilities $p^{(t)}_{ij}$ where $t \geq \tau_{m-1}$. In other words, in order to write down a distribution for the θ_m, one should know the value of the variable τ_{m-1} but now necessarily the values of variables $\theta_1, \dots, \theta_{m-1}$. Moreover, under fixed τ_{m-1} the distribution of θ_m does not depend on the values $\theta_1, \dots, \theta_{m-1}$.

Now we have:

$$\begin{aligned} &\mathsf{P}\{\theta_m = i, \theta_{m-1} = j \mid \tau_{m-1} = t\} \\ &\quad = \mathsf{P}\{\theta_m = i, \theta_{m-1} = j \mid X_t = 0, X_l = 0, \text{exactly } m-2 \text{ times}, l < m-1\} \\ &\quad = \mathsf{P}\{X_k = 0, k \in \{i, t, t-j\}, X_k \neq 0 \text{ otherwise}, A\}\, \mathsf{P}^{-1}(A) \\ &\quad = \mathsf{P}\{X_i = 0, X_l \neq 0, \\ &\qquad\qquad l = t+1, \dots, i-1 \mid X_t = 0, X_{t-1} \neq 0, \dots, X_{t-j} = 0, X_{t-j-1} \neq 0, A\} \\ &\qquad \times \mathsf{P}\{X_t = 0, X_{t-1} \neq 0, \dots, X_{t-j} = 0, X_{t-j-1} \neq 0 \mid A\} \\ &\quad = \mathsf{P}\{X_i = 0, X_l \neq 0, l = t+1, \dots, i-1 \mid X_t = 0\}\, \mathsf{P}\{\theta_{m-1} = j \mid \tau_{m-1} = t\} \\ &\quad = \mathsf{P}\{X_i = 0, X_l \neq 0, l = i-1, \dots, t+1 \mid X_t = 0, B\}\, \mathsf{P}\{\theta_{m-1} = j \mid \tau_{m-1} = t\} \\ &\quad = \mathsf{P}\{\theta_m = i \mid \tau_{m-1} = t\}\, \mathsf{P}\{\theta_{m-1} = j \mid \tau_{m-1} = t\}, \end{aligned}$$

where the set $A = \{X_t = 0, X_l = 0, \text{ exactly } m-2 \text{ times}, l < m-1\} = \{\tau_{m-1} = t\}$, $B = \{\text{exactly } m-1 \text{ zero hittings happened till time } t-1\}$.

So we have proved that

$$\mathsf{P}\{\theta_m = i, \theta_{m-1} = j \mid \tau_{m-1} = t\} = \mathsf{P}\{\theta_m = i \mid \tau_{m-1} = t\}\, \mathsf{P}\{\theta_{m-1} = j \mid \tau_{m-1} = t\}, \tag{6}$$

which means that variables θ_m and θ_{m-1} are conditionally independent given τ_{m-1}.

Let us also note, that formula (5) implies that the distribution of the θ_m is parameterized by only one parameter j (values of a τ_{m-1}), and does not depend on index m. So we can write:

$$g_n^j = \mathsf{P}\{\theta_m = n \mid \tau_{m-1} = j\}.$$

This fact leads us to consideration of the random variables $\theta(t)$ which have the same distribution as $(g_n^t)_{n\geq 0}$. This variables can be handled as moments of the first after time t returning to zero, if we know that a chain is in the zero state at the moment t.

3. Key definitions

In this section and further on we'll consider two time-inhomogeneous Markov chains $(X_t^1, t \geq 0)$ and $(X_t^2, t \geq 0)$ defined on a phase space $E = \{0, 1, \ldots\}$. The chains are defined by their transition probabilities on the s-th step $P_s(x, A, 1)$, $P_s(x, A, 2)$ for chains X_t^1, X_t^2 respectively. Let's define transition probabilities for $n > 0$ steps:

$$P^{t,n}(x, A, l) = \left(\prod_{k=0}^{n-1} P_{t+k}\right)(x, A, l). \tag{7}$$

Having this set of transition probabilities and the initial conditions $\mu^l(\cdot)$ we can build a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where both chains (X_t^l), $l \in \{1, 2\}$, are defined and

$$\mathbb{P}\{X_s^l \in A\} = \int_E \mu^l(dx) P^{0,s}(x, A, l), \qquad \mathbb{P}\{X_{s+1}^l \in A \mid X_s^l = x\} = P_s(x, A, l).$$

Let's define renewal intervals θ_k^l, $l \in \{1, 2\}$:

$$\theta_0^l = \inf\{t \geq 0 \colon X_t = 0\}, \qquad \theta_m^l = \inf\{t > \theta_{m-1} \colon X_t = 0\}, \quad m > 1, \tag{8}$$

which are defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The classes of variables $\{\theta_k^1\}_{k\geq 0}$ and $\{\theta_k^2\}_{k\geq 0}$ are independent. θ_k^l for each $l \in \{1, 2\}$ and $k > 0$ have only positive integer values while θ_0^l take non-negative integers. Let's define renewal sequences in the following way:

$$\tau_n^l = \sum_{k=0}^{n} \theta_k^l, \qquad l \in \{1, 2\}. \tag{9}$$

We will assume that neighboring variables inside each class are conditionally independent giving τ. In other words, for each k, t, l the following equality holds true:

$$\mathsf{E}\left[f\left(\theta_k^l\right) g\left(\theta_{k+1}^l\right) \mid \tau_k^l\right] = \mathsf{E}\left[f\left(\theta_k^l\right) \mid \tau_k^l\right] \mathsf{E}\left[g\left(\theta_{k+1}^l\right) \mid \tau_k^l\right], \tag{10}$$

for any bounded Borel functions f and g.

Let's introduce a definition for the conditional distribution of the θ_k^l variable (please, note that this distribution does not depend on k):

$$g_n^{t,l} = \mathsf{P}\left\{\theta_k^l = n \;\middle|\; \tau_{k-1} = t\right\}, \qquad l =\in \{1, 2\},\ n \geq 0, \tag{11}$$

and we assume that $g_0^{t,l} = \mathsf{P}\{\theta_k^l = 0 \mid \tau_{k-1} = t\} = 0$. The variables θ_k^l, $k \geq 1$ will be interpreted as renewal steps and θ_0^l as a delay.

We'll say that $T > 0$ is a coupling (or simultaneously hitting) moment if:

$$T = \min\left\{t > 0 \colon \exists m, n \colon t = \tau_m^1 = \tau_n^2\right\}. \tag{12}$$

Our goal is to find conditions which guarantee $T < \infty$ a.s. and $\mathsf{E}[T] < \infty$.

By $u_n^{(t,l)}$ we define a renewal sequence for the process τ^l. In other words, $u_n^{(t,l)}$ is a probability of a renewal at the moment $t+n$ having renewal at the moment t. Formally $u_n^{(t,l)}$ can defined in a following way:

$$u_0^{(t,l)} = 1, \qquad u_n^{(t,l)} = \sum_{k=0}^{n} u_k^{(t,l)} g_{n-k}^{t+k,l}. \tag{13}$$

4. Formal definition of the $\theta^l(t)$ variable

As we've seen before, the distribution of the $k+1$-st renewal interval is completely defined by the value of the τ_k variable, i.e. by the moment of the previous renewal and does not depend on the index k. That's why we have introduced the notations $g_n^{t,l}$ and $u_n^{(t,l)}$. Our goal is to define random variables $\theta^l(t)$ in such a way that $g_n^{t,l}$ be a distribution for $\theta^l(t)$.

For simplicity we'll omit index l in this section.

Assume X_t is some time-inhomogeneous Markov chain with transition probabilities on the t-th step equal to $P_t(x, A)$. As before, let's define:

$$P^{t,n}(x, A) = \left(\prod_{k=0}^{n-1} P_{t+k}\right)(x, A),$$

transition probability for the time from t to $t+n$.

For each t we define probability space $(\Omega_t, \mathcal{F}_t, \mathbb{P}_t)$ as a canonical space for the Markov chain X_{t+n} which starts at zero. Let's note that

$$\theta(t) = \min\{j > 0 \colon X_{t+j} = 0\}, \tag{14}$$

and $g_n^t = \mathbb{P}_t\{\theta(t) = n\}$ is the distribution of the variable $\theta(t)$. Then,

$$g_n^t = \int_{(E\setminus\{0\})^{n-1}} P_t(0, dx_0) P_{t+1}(x_0, dx_1) \dots P_{t+n-1}(x_{n-1}, \{0\}). \tag{15}$$

As in the previous section let's define $\theta^l(t)$ as a moment of the first hitting zero state for the chain $(X_{t+k}^l, k \geq 0)$ which starts from zero. Then a variable $\theta^l(t)$ has the distribution $(g_n^{t,l})_{n\geq 0}$.

Let's define an overshoot:

$$D_n(t) = \min\{j \geq 0 \colon X_{t+n+j} = 0\}. \tag{16}$$

The variable $D_n(t)$ should be understood as a time that has left till hitting $\{0\}$ after moment $t+n$ having $X_t = 0$. Note that variables $D_n(t)$ and $\theta(t)$ are defined on the common probability space $(\Omega_t, \mathcal{F}_t, \mathbb{P}_t)$.

The following lemma is a key in proving the main theorem (the proof will be given later):

Lemma 4.1. *If a distribution family g_n^t (or, a family of random variables $\theta(t)$) is uniformly integrable then for each $\rho \in (0,1)$ there exists a constant $C = C(\rho) \geq 0$, such that for each t the following inequality holds true:*

$$\mathbb{E}_t[D_n(t)] \leq \rho n + C.$$

5. Main theorem

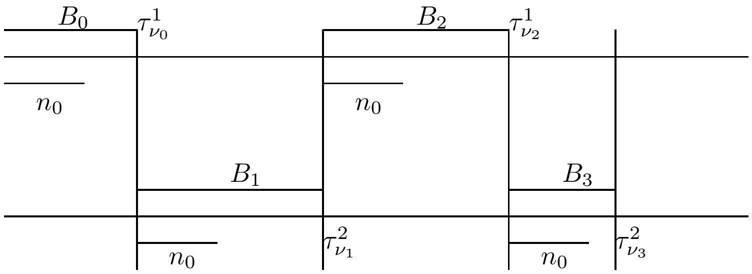
Theorem 5.1. *Assume that (in notations introduced before):*

1) *The set of random variables $\theta^l(t)$ is uniformly integrable (or, in other words, the family of distributions $g_n^{t,l}$ is uniformly integrable).*
2) *There exists a constat $\gamma > 0$ and a positive integer $n_0 > 0$ such that for all t, l and $n \geq n_0$: $u_n^{(t,l)} \geq \gamma$.*

Then the coupling moment is integrable: $\mathsf{E}[T] < \infty$.

6. Setup for the proof of the theorem 5.1

Following the Lindvall approach (see. [14, p. 27]) let's define the following random variables:



$$\nu_0 := \min\left\{j \geq 1 \colon \tau^1_j > n_0\right\},$$
$$B_0 := \tau^1_{\nu_0},$$
$$\nu_1 := \min\left\{j \geq \nu_0 \colon \tau^2_j - \tau^1_{\nu_0} > n_0, \text{ or } \tau^2_j - \tau^1_{\nu_0} = 0\right\},$$
$$B_1 := \tau^2_{\nu_1} - \tau^1_{\nu_0},$$

and further on

$$\nu_{2m} := \min\left\{j \geq \nu_{2m-1} \colon \tau^1_j - \tau^2_{\nu_{2m-1}} > n_0, \text{ or } \tau^1_j - \tau^2_{\nu_{2m-1}} = 0\right\},$$
$$B_{2m} := \tau^1_{\nu_{2m}} - \tau^2_{\nu_{2m-1}},$$
$$\nu_{2m+1} := \min\left\{j \geq \nu_{2m} \colon \tau^2_j - \tau^1_{\nu_{2m}} > n_0\,, \text{ or } \tau^2_j - \tau^1_{\nu_{2m}=0}\right\},$$
$$B_{2m+1} := \tau^2_{\nu_{2m+1}} - \tau^1_{\nu_{2m}}.$$

ν_k is called as coupling trials. Let's define $\tau = \min\{n \geq 1 \colon B_n = 0\}$ and a sequence of sigma-fields $\mathfrak{B}_n$, $n \geq 0$ in the following way:

$$\mathfrak{B}_n = \sigma\left[B_k, \nu_k, \tau^l_j, k \leq n, j \leq \nu_n\right].$$

Let's also define random variables: $D^{k,l}_n = \min\{j \colon \exists m, \tau^l_m = \tau^l_k + n + j\}$.

7. The proof of the theorem 5.1

At the beginning we assume that $\theta^2_0 = 0$.

The following inequality is true:

$$T \leq \theta^1_0 + \sum_{n=0}^{\tau} B_n = \theta^1_0 + \sum_{n \geq 0} B_n \mathbb{1}_{\tau \geq n}. \tag{17}$$

According to the lemma 8.4 for each $n \geq 0$, $\rho \in (0,1)$ the following inequality holds true:

$$\mathsf{E}[B_n \mid \mathfrak{B}_{n-1}] \leq \rho B_{n-1} + C, \tag{18}$$

which implies that

$$\begin{aligned}
\mathsf{E}[B_k \mathbb{1}_{\tau \geq k} \mid \mathfrak{B}_{k-1}] &= \mathsf{E}\left[B_k \prod_{n=0}^{k-1} \mathbb{1}_{B_k \neq 0} \,\middle|\, \mathfrak{B}_{k-1}\right] = \mathbb{1}_{\tau \geq n}\, \mathsf{E}[B_n \mid \mathfrak{B}_{n-1}] \\
&\leq \mathbb{1}_{\tau \geq n}(\rho B_{n-1} + C) = \rho B_{n-1} \mathbb{1}_{\tau \geq n} + C \mathbb{1}_{\tau \geq n} \\
&\leq \rho B_{n-1} \mathbb{1}_{\tau \geq n-1} + C \mathbb{1}_{\tau \geq n},
\end{aligned}$$

where the latest equality follows from the relation $\{\tau \geq n\} \subset \{\tau \geq n-1\}$ and so $\mathbb{1}_{\tau \geq n} \leq \mathbb{1}_{\tau \geq n-1}$.

So we've proved the following inequality:

$$\mathsf{E}[B_n \mathbb{1}_{\tau \geq n}] \leq \rho\, \mathsf{E}\left[B_{n-1} \mathbb{1}_{\tau \geq n-1}\right] + C\, \mathsf{P}\{\tau \geq n\}. \tag{19}$$

It follows from lemma 8.5 that

$$\mathsf{P}\{\tau \geq n\} \leq (1-\gamma)^n.$$

Let's define $a_n = \mathsf{E}[B_n \mathbb{1}_{\tau \geq n}]$. Then (19) implies:

$$a_n \leq \rho a_{n-1} + C(1-\gamma)^n \leq C \sum_{k=0}^{n} \rho^k (1-\gamma)^{n-k} \leq Cn \max(\rho, (1-\gamma))^n.$$

Note, since ρ is arbitrary, we can choose it be equal to $(1-\gamma)$. In this case

$$a_n \leq Cn(1-\gamma)^n.$$

So

$$\mathsf{E}[T] \leq \mathsf{E}\left[\theta_0^1\right] + \sum_{n \geq 0} a_n \leq \mathsf{E}\left[\theta_0^1\right] + \frac{C}{\gamma^2} < \infty. \tag{20}$$

Recall our assumption $\theta_0^2 = 0$. Now we will get rid of it. Let's define as T' a coupling moment for the processes with the following delays:

$$\theta_0'^1 = \max\left(\theta_0^1, \theta_0^2\right) - \min\left(\theta_0^1, \theta_0^2\right),$$

$$\theta_0'^2 = 0.$$

Note that $T = T' + \min(\theta_0^1, \theta_0^2)$. So

$$\mathsf{E}[T] \leq \mathsf{E}\left[\min\left(\theta_0^1, \theta_0^2\right)\right] + \mathsf{E}[T'] < \infty.$$

Note that

$$\mathsf{E}[T'] \leq \mathsf{E}\left[\theta_0'^1\right] + \frac{C}{\gamma^2},$$

or

$$\mathsf{E}[T] \leq \mathsf{E}\left[\max\left(\theta_0^1, \theta_0^2\right)\right] + \frac{C}{\gamma^2}.$$

8. Auxiliary lemmas

Lemma 8.1. *Let $x_n^{(t)}$, $y_n^{(t)}$ be some inhomogeneous sequences of real numbers, $u_n^{(t)}$ be some inhomogeneous renewal sequence defined by the formula (13): $g_0^{(t)} = 0$, for all t. Assume the following conditions are true*

$$x_n^{(t)} = \sum_{k=0}^{n} g_k^{(t)} x_{n-k}^{(t+k)} + y_n^{(t)}, \tag{21}$$

$$x_n^0 \geq \sum_{k=0}^{n} g_k^{(t)} x_{n-k}^0 + y_n^{(t)}. \tag{22}$$

Then for any t, n:

$$x_n^{(t)} \leq x_n^0.$$

Proof. Let's show that

$$x_n^{(t)} = \sum_{k=0}^{n} u_k^{(t)} y_{n-k}^{(t+k)}. \tag{23}$$

We'll do this by induction:

For the $n = 0$: $x_0^{(t)} = g_0^{(t)} x_0^{(t)} + y_0^{(t)} = y_0^{(t)} = u_0^{(t)} y_0^{(t)}$.

Assuming the statement holds true for all $k \leq n$, lets prove it for the $n+1$.

$$x_{n+1}^{(t)} = \sum_{k=0}^{n+1} g_k^{(t)} x_{n+1-k}^{(t+k)} + y_{n+1}^{(t)} = \sum_{k=1}^{n+1} g_k^{(t)} \sum_{j=0}^{n+1-k} u_j^{(t+k)} y_{n+1-k-j}^{(t+k+j)} + g_0^{(t)} x_{n+1}^{(t)} + y_{n+1}^{(t)}$$
$$= \sum_{k=1}^{n+1} u_k^{(t)} y_{n-k}^{(t+k)} + y_n^{(t)} = \sum_{k=0}^{n+1} u_k^{(t)} y_{n-k}^{(t+k)}.$$

Then for any t, n:

$$y_n^{(t)} \leq x_n^{(0)} - \sum_{k=0}^{n} g_k^{(t)} x_{n-k}^{0},$$
$$x_n^{(t)} \leq \sum_{k=0}^{n} u_k^{(t)} x_{n-k}^{(0)} - \sum_{k=0}^{n} u_k^{(t)} \sum_{j=0}^{n-k} g_j^{(t+k)} x_{n-k-j}^{(0)}. \tag{24}$$

Let us consider the second term

$$\sum_{k=0}^{n} u_k^{(t)} \sum_{j=0}^{n-k} g_j^{(t+k)} x_{n-k-j}^{(0)} = x_0^0 \sum_{k=0}^{n} u_k^{(t)} g_{n-k}^{(t+k)} + x_1^0 \sum_{k=0}^{n-1} u_k^{(t)} g_{n-1-k}^{(t+k)} + \ldots + x_n^0 u_0^{(t)} g_0^{(t)}$$
$$= \sum_{k=0}^{n-1} x_k^0 u_{n-k}^{(t)} = \sum_{k=1}^{n} u_k^{(t)} x_{n-k}^0.$$

Applying the last relation to the (24) we derive:

$$x_n^{(t)} \leq \sum_{k=0}^{n} u_k^{(t)} x_{n-k}^{(0)} - \sum_{k=1}^{n} u_k^{(t)} x_{n-k}^0 = u_0^{(t)} x_n^0 = x_n^0.$$ □

Lemma 8.2. *Assume A is a some set defined by the variables $\tau_{\nu_k}^l$, ν_k, $k < n$. Then:*

$$\mathsf{E}\left[D_{k+n_0}^{m,l} \mid B_{n+1} = k, \tau_{\nu_n}^l = t, \nu_n = m, A\right] = \mathbb{E}_t\left[D_{k+n_0}^l(t)\right].$$

Proof. Let's denote $t + k + n_0 = q$. Then:

$$\mathsf{P}\left\{D_{k+n_0}^{m,l} = r, B_{n+1} = k, \tau_{\nu_n}^l = t, \nu_n = m, A\right\}$$
$$= \mathbb{P}\left\{X_{q+r}^l = 0, X_{q+s}^l \neq 0, s = 0, \ldots, r-1, X_t^l = 0, \tau_{\nu_n}^l = t, \nu_n = m, B_{n+1} = k, A\right\}$$
$$= \left(\int_{(E\setminus 0)^r} P^{t,k+n_0}(0, dx_0, l) P_q(x_0, dx_1, l) \ldots P_{q+r-1}(x_{r-1}, dx_r, l) P_{q+r}(x_r, 0, l)\right)$$
$$\times \mathbb{P}\left\{X_t = 0, \tau_{\nu_n}^l = t, \nu_n = m, B_{n+1} = k, A\right\}$$
$$= \mathbb{P}_t\left\{D_{k+n_0}^l(t) = r\right\} \mathbb{P}\left\{\tau_{\nu_n}^l = t, \nu_n = m, B_{n+1} = k, A\right\}.$$ □

Lemma 8.3.

$$\mathsf{E}[B_{2n} \mid \mathfrak{B}_{2n-1}] = \sum_{t,k} \mathbb{E}_t\left[D_{k+n_0}^1(t)\right] \mathbb{1}_{\tau_{\nu_{2n-2}}^1 = t} \mathbb{1}_{B_{2n-1} = k},$$
$$\mathsf{E}[B_{2n+1} \mid \mathfrak{B}_{2n}] = \sum_{t,k} \mathbb{E}_t\left[D_{k+n_0}^2(t)\right] \mathbb{1}_{\tau_{\nu_{2n-1}}^2 = t} \mathbb{1}_{B_{2n} = k}.$$

Proof. At the beginning we should note that the sigma-field $\mathfrak{B}_m$ is generated by the finite amount of random variables, and each of them takes only no more than countable number of values. So, for each m, $\mathfrak{B}_m$ is generated by the finite number of events.

Let us define a set of events $\{A_n(i), i \in \mathfrak{I}_n\}$ as $A_n(i) = \{\tau^l_{\nu_k} = t_{lk}, \nu_k = n_k, k \leq n\}$ and note that $\mathfrak{I}_n$ is a countable set. Let's add the following notation

$$C_n(s,t,m,k) = \left\{\tau^2_k = t, \tau^1_m = s, \nu_{2n-1} = k, \nu_{2n-2} = m, A_{2n-3}(i)\right\}.$$

Note that it follows from the definition of B_{2n} that

$$B_{2n} = D^{\nu_{2n-2},1}_{B_{2n-1}+n_0} + n_0, \tag{25}$$

which implies

$$\begin{aligned}
&\mathsf{E}[B_{2n} - n_0 \mid \mathfrak{B}_{2n-1}] \\
&\quad = \sum_{s<t,m,k,i\in\mathfrak{I}_{2n-3}} \mathsf{E}\left[D^{m,1}_{t-s+n_0} \,\middle|\, C_n(s,t,m,k), A_{2n-3}(i)\right] \nVdash_{C_n(s,t,m,k)} \nVdash_{A_{2n-3}(i)}.
\end{aligned}$$

Using lemma 8.2 we derive that the last term is equal

$$\begin{aligned}
&\sum_{s<t,m,k,i\in\mathfrak{I}_{2n-3}} \mathbb{E}_s\left[D^1_{t-s+n_0}(s)\right] \nVdash_{C_n(s,t,m,k)} \nVdash_{A_{2n-3}(i)} \\
&\quad = \sum_{s<t,m,k} \mathbb{E}_s\left[D^1_{t-s+n_0}(s)\right] \nVdash_{C_n(s,t,m,k)} = \sum_{s<t} \mathbb{E}_s\left[D^1_{t-s+n_0}(s)\right] \nVdash_{\tau^2_{\nu_{2n-1}}=t} \nVdash_{\tau^1_{\nu_{2n-2}}=s} \\
&\quad = \sum_{s,k} \mathbb{E}_s\left[D^1_{k+n_0}\right] \nVdash_{\tau^1_{\nu_{2n-2}}=s} \nVdash_{B_{2n-1}=k},
\end{aligned}$$

where we used the following equality $B_{2n-1} = \tau^2_{\nu_{2n-1}} - \tau^1_{\nu_{2n-2}}$ in the last relation.

The corresponding statement for $\mathsf{E}[B_{2n+1} \mid \mathfrak{B}_{2n}]$ can be derived in a similar way. □

Lemma 8.4. *Assuming the conditions of the theorem 5.1 holds true for each $\rho \in (0,1)$ there exists a constant $C \in (0,\infty)$, that for every $n \geq 0$ a following inequality is true*

$$\mathsf{E}[B_n \mid \mathfrak{B}_{n-1}] \leq \rho B_{n-1} + C.$$

Proof. Using lemmas 8.3 and 4.1 we will get

$$\begin{aligned}
\mathsf{E}[B_{2n} \mid \mathfrak{B}_{2n-1}] &= \sum_{t,k} \mathbb{E}_t\left[D^1_{k+n_0}(t)\right] \nVdash_{\tau^1_{\nu_{2n-2}}=t} \nVdash_{B_{2n-1}=k} \\
&\leq \sum_{t,k} (\rho(k+n_0) + C) \nVdash_{\tau^1_{\nu_{2n-2}}=t} \nVdash_{B_{2n-1}=k} = \rho B_{2n-1} + C'.
\end{aligned}$$

The same statement holds true for the $\mathsf{E}[B_{2n+1} \mid \mathfrak{B}_{2n}]$. □

Lemma 8.5. *The following inequality is true*

$$\mathsf{P}\{\tau > n\} \leq (1-\gamma)^n.$$

Proof. Recall that $\tau = \min(n \colon B_n = 0)$. An event $\{\tau > n\} = \{\prod_{k=0}^n B_k \neq 0\}$.

$$\begin{aligned}
\mathsf{E}\left[\nVdash_{\prod_{k=0}^n B_k \neq 0}\right] &= \mathsf{E}\left[\nVdash_{\prod_{k=0}^{n-1} B_k \neq 0} \mathsf{E}\left[\nVdash_{B_n \neq 0} \mid \mathfrak{B}_{n-1}\right]\right] \\
&= \mathsf{E}\left[\nVdash_{\prod_{k=0}^{n-1} B_k \neq 0}\right] \mathsf{P}\left\{\theta^l_\eta > B_n + B_{n-1}\right\} \\
&\leq \mathsf{E}\left[\nVdash_{\prod_{k=0}^{n-1} B_k \neq 0}\right] \mathsf{P}\left\{\theta^l_\eta > n_0\right\} \leq \mathsf{E}\left[\nVdash_{\prod_{k=0}^{n-1} B_k \neq 0}\right] (1-\gamma) \leq (1-\gamma)^n,
\end{aligned}$$

where η is a number of the next after B_{n-1} renewal in the l-th series. □

9. The proof of the lemma 4.1

Let's consider the random variable $D_n(t)\nVdash_{\theta(t)=j}$, $j \le n$. By the direct calculation it is easy to verify that

$$\mathbb{P}_t\{D_n(t) = k, \theta(t) = j\} = \mathbb{P}_t\{\theta(t) = j\}\mathbb{P}_{t+j}\{D_{n-j}(t+j) = k\}. \tag{26}$$

The following inequality holds true:

$$D_n(t)\nVdash_{\theta(t)>n} = (\theta(t) - n)\nVdash_{\theta(t)>n}. \tag{27}$$

Then, having in mind inequalities (26) and (27) we'll get

$$\begin{aligned}
\mathbb{E}_t[D_n(t)] &= \sum_{j=1}^{n} \mathbb{E}_t\left[D_n(t)\nVdash_{\theta(t)=j}\right] + \mathbb{E}_t\left[D_n(t)\nVdash_{\theta(t)>n}\right] \\
&= \sum_{j=1}^{n}\left(\sum_{k=0}^{\infty} k\mathbb{P}_t\{D_n(t) = k, \theta(t) = j\}\right) + \mathbb{E}_t\left[(\theta(t) - n)\nVdash_{\theta(t)>n}\right] \\
&= \sum_{j=1}^{n} \mathbb{P}_t\{\theta(t) = j\}\left(\sum_{k=0}^{\infty} k\mathbb{P}_{t+j}\{D_{n-j}(t+j) = k\}\right) + \mathbb{E}_t[(\theta(t) - n)\nVdash_{\theta(t)>n}] \\
&= \sum_{j=1}^{n} g_j^t \mathbb{E}_{t+j}[D_{n-j}(t+j)] + \mathbb{E}_t\left[(\theta(t) - n)\nVdash_{\theta(t)>n}\right].
\end{aligned}$$

So we have the following equality

$$\mathbb{E}_t[D_n(t)] = \sum_{j=1}^{n} g_j^t \mathbb{E}_{t+j}[D_{n-j}(t+j)] + \mathbb{E}_t\left[(\theta(t) - n)\nVdash_{\theta(t)>n}\right]. \tag{28}$$

After that we'll use the lemma 8.1. Let's define:

$$x_n^{(t)} = \mathbb{E}_t[D_n(t)],$$
$$y_n^{(t)} = \mathbb{E}_t\left[\nVdash_{\theta(t)>n}(\theta(t) - n)\right],$$

then (28) implies the condition (21).

We define as

$$x_n^0 = \rho n + C.$$

Let's proof that the condition (22) of the lemma 8.1 holds true. For doing that we should show, that for any $\rho \in (0, 1)$ there exists such $C = C(\rho)$, that

$$\rho n + C \ge \sum_{j=0}^{n} g_j^t(\rho(n-j) + C) + \sum_{j>n}(j-n)g_j^t, \tag{29}$$

We'll derive the following from the statement (29)

$$\begin{aligned}
(29) \Leftrightarrow \rho n + C &\ge n\rho\sum_{j=0}^{n} g_j^t + C\sum_{j=0}^{n} g_j^t - \rho\sum_{j=0}^{n} jg_j^t + \sum_{j>n} jg_j^t - nG_n^t \\
\Leftrightarrow n\rho G_n^t + CG_n^t &\ge \mathbb{E}_t\left[\theta(t)\nVdash_{\theta(t)>n}\right] - \rho\mathbb{E}_t\left[\theta(t)\nVdash_{\theta(t)\le n}\right] - nG_n^t \\
\Leftrightarrow n(\rho+1)G_n^t + CG_n^t + \rho\mathbb{E}_t\left[\theta(t)\nVdash_{\theta(t)\le n}\right] &\ge \mathbb{E}_t\left[\theta(t)\nVdash_{\theta(t)>n}\right] \\
\Leftrightarrow n(\rho+1)G_n^t + CG_n^t + \rho\mathbb{E}_t\left[\theta(t)\right] &\ge (1+\rho)\mathbb{E}_t\left[\theta(t)\nVdash_{\theta(t)>n}\right],
\end{aligned} \tag{30}$$

So, the inequalities (29) are equivalent to (30). Note that, in the case of $G_n^t = 0$ the equality (30) holds true automatically. Assume than $G_n^t > 0$. But $\mathbb{E}_t[\theta(t)] \ge 1$ and the uniform integrability implies that there is a number n_0, such that for all $t > 0, n \ge n_0$: $\mathbb{E}_t[\theta(t)\nVdash_{\theta(t)>n}] \le \rho/(1+\rho)$. The constant C we'll choose in the way to satisfy (30) for $n \le n_0$.

Let's show now, that C could be chosen disregarding of t. For $\varepsilon = \rho/(1+\rho)$ we'll find such $\delta > 0$, that for each set A, such that $\mathbb{P}(A) < \delta$ it follows that $\mathbb{E}_t[\theta(t)\mathbb{1}_A] < \varepsilon$. It is possible, since $\theta(t)$ are uniformly integrable. Let's define then

$$C := \frac{(1+\rho)sup_t\mathbb{E}_t[\theta(t)] - \rho}{\delta}.$$

Now having $G_n^t < \delta$ inequality (30) holds true automatically. In the case of $G_n^t \geq \delta$, we'll get:

$$\begin{aligned} n(\rho+1)G_n^t + CG_n^t + \rho\mathbb{E}_t[\theta(t)] &> (1+\rho)\sup_t \mathbb{E}_t[\theta(t)] \geq (1+\rho)\mathbb{E}_t[\theta(t)] \\ &\geq (1+\rho)\mathbb{E}_t\left[\theta(t)\mathbb{1}_{\theta(t)>n}\right]. \end{aligned}$$

References

1. W. Doeblin, *Expose de la theorie des chaines simples constantes de Markov a un nomber fini d'estats*, Mathematique de l'Union Interbalkanique **2** (1938), 77–105.
2. R. Douc, G. Fort, and A. Guillin, *Subgeometric rates of convergence of f-ergodic strong Markov processes*, Stochastic Processes and their Applications **119** (2009), no. 3, 897–923.
3. W. Feller, *An Intoduction to Probability Theory and its Applications*, vol. 1, 1966, John Wiley & Sons, New York.
4. V. V. Golomoziy, *Stability of inhomogeneous Markov chains*, Bulletin of Taras Shevchenko National University of Kyiv. Series: Physics and Mathematics **4** (2009), 10–15. (Ukrainian)
5. V. V. Golomoziy, *A subgeometric estimate of the stability for time-homogeneous Markov chains*, Theory of Probability and Mathematical Statistics **81** (2010), 31–46.
6. N. V. Kartashov, *Exponential asymptotic for Markov renewal matrix*, Asymptotic problems for stochastic processes, Preprints of Institute of Mathematics of UNAS, no. 77-24, Kyiv, 1977, pp. 2–43. (Russian)
7. N. V. Kartashov, *Strong Stable Markov Chains*, VSP, Utrecht, The Netherlands, 1996.
8. N. V. Kartashov, *Boundedness, limits and stability of solutions of a perturbation of a nonhomogeneous renewal equation on a semiaxis*, Theory of Probability and Mathematical Statistics **81** (2009), 65-75.
9. N. V. Kartashov and V. V. Golomoziy, *The mean coupling time for independent discrete renewal processes*, Theory of Probability and Mathematical Statistics **84** (2012), 78–85.
10. N. V. Kartashov and V. V. Golomoziy, *Maximal coupling procedure and stability of discrete Markov chains. I*, Theory of Probability and Mathematical Statistics **86** (2012), 81–92.
11. N. V. Kartashov and V. V. Golomoziy, *Maximal coupling procedure and stability of discrete Markov chains. II*, Theory of Probability and Mathematical Statistics. (to appear)
12. I. N. Kovalenko and N. Yu. Kuznetsov, *Construction of the embedded renewal process for substantially multidimensional processes in queueing theory and its application for obtaining limit theorems*, Preprints of UNAS, no. 80 2-12, Institute of Cybernetics, Kyiv, 1980. (Russian)
13. T. Lindvall, *On coupling of discrete renewal sequences*, Z. Wahrsch. Verw. Gebiete **48** (1979), 57–70.
14. T. Lindvall, *Lectures on the Coupling Method*, John Wiley and Sons, 1991.
15. H. Thorisson, *Coupling, Stationarity, and Regeneration*, Springer, New York, 2000.

Department of Probability Theory, Statistics and Actuarial Mathematics, Faculty of Mechanics and Mathematics, Taras Shevchenko National University of Kyiv 64/13, Volodymyrska Street, City of Kyiv, Ukraine, 01601

E-mail address: mailtower@gmail.com, mykartashov@gmail.com

Department of Probability Theory, Statistics and Actuarial Mathematics, Faculty of Mechanics and Mathematics, Taras Shevchenko National University of Kyiv 64/13, Volodymyrska Street, City of Kyiv, Ukraine, 01601

E-mail address: mykartashov@gmail.com

Received 01/11/2011

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 12–20

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 12–20

SEMI-MARKOV APPROACH TO THE PROBLEM OF DELAYED REFLECTION OF DIFFUSION MARKOV PROCESSES

UDC 519.21

B. P. HARLAMOV

Abstract. An one-dimensional diffusion process with positive values, reflecting from zero, is considered. All the variants of reflecting with preservation of the semi-Markov property are described. This property is characterized by a family of Laplace images of times from the first hitting of zero up to the first hitting of a level r for any $r > 0$. The parameter $C(\lambda)$ of this family is used for construction of a time change, transforming a process with instantaneous reflection to the process with delayed reflection.

Анотація. Розглядається одновимірний дифузійний процес з додатними значеннями, що відбивається від точки нуль. Описано всі варіанти відбиття із збереженням напівмарковської властивості процесу, що характеризується сімейством перетворень Лапласа моментів від першого досягнення нуля до першого досягнення заданого рівня r для всіх $r > 0$. Параметр-функція $C(\lambda)$ цього сімейства використовується для побудови заміни часу, що перетворює процес із миттєвим відбиттям у процес із уповільненим відбиттям.

Аннотация. Рассматривается одномерный диффузионный процесс с положительными значениями, отражающийся от точки 0. Описываются все варианты отражения с сохранением полумарковского свойства процесса, которое характеризуется семейством преобразований Лапласа времен от первого достижения нуля до первого достижения заданного уровня r для всех $r > 0$. Параметр-функция $C(\lambda)$ этого семейства используется для вывода характеристик замены времени, превращающей процесс с мгновенным отражением в процесс с замедленным отражением.

1. Introduction

Apparently Gihman and Skorokhod were the first who investigated reflection with delaying of one-dimensional Markov diffusion processes ([1, p. 197]). They applied a method of stochastic integral equations which takes into account preserving the Markov property while reflecting. However there exist examples of interaction between a process and a boundary of its range of values, which can be interpreted like reflection, when the Markov property is being lost, although the property of continuous semi-Markov processes is preserved. Here is a simple example.

Let $w(t)$, $t \ge 0$, be Wiener process. Let us consider on the segment $[a, b]$, $a < w(0) < b$, the truncated process

$$\overline{w}(t) = \begin{cases} b, & w(t) \ge b \\ w(t), & a < w(t) < b \\ a, & w(t) \le a \end{cases}$$

for all $t \ge 0$. It is clear that this process is not Markov. However it remains to be continuous semi-Markov [4]: the Markov property is fulfilled with respect to the first exit time from any open interval inside the segment, and also that from any one-sided neighborhood of any end of the segment.

2000 *Mathematics Subject Classification.* Primary 60J25, 60J60.

Key words and phrases. Diffusion, Markov, continuous semi-Markov, reflection, delaying, first exit time, transition function, Laplace transformation, time change.

This work is supported by grant RFBR 12-01-00457-a.

The semi-Markov approach to the problem of reflection consists in solution of the following task: to determine a semi-Markov transition function for the process at a boundary point for the process preserving its diffusion form inside its open range of values, i. e. that up to the first exit time from the region and any time when it leaves the boundary. A more specific task to find reflection, preserving a global Markov property, is reduced to a problem to find a subclass of Markov reflected processes in the class of all the semi-Markov ones. Tasks of such a kind are important for applications where one takes into account interaction of diffusion particles with a boundary of a container, leading to a dynamic equilibrium of the system (see, e. g. [7]).

In paper [3] all class of semi-Markov characteristics of reflection for a given locally Markov diffusion process is described. In paper [5] conditions for a semi-Markov characteristic to give a globally Markov process are found. In the present paper we continue to investigate processes with semi-Markov reflection. The aim of investigation is to find formulae, characterizing a time change, transforming a process with instantaneous reflection into the process with delaying reflection,

In paper [6] while analyzing a two-dimensional diffusion process in a neighborhood of a flat screen a time change in a tangential component of the process with respect to a normal component time run is factually treated. This splitting of the process on two components makes the situation easier to be understood, but at the same time it masks the true mechanism of transformation. In fact the time change could be learned on the initial stage of semi-Markov approach to the problem of reflection. In the present paper this shortcoming of our first paper on this theme is removed.

2. Semi-Markov transition function on a boundary

We will consider a diffusion process $X(t)$ on the half-line $t \geq 0$ with one boundary at zero. We assume that the process does not go to infinity and from any positive initial point it hits zero with probability one. For example, it could be a diffusion Markov process with a negative drift and bounded local variance. We had substantiated above why it is expedient to consider semi-Markov reflection. Semi-Markov approach permits to consider from unit point of view an operation of instantaneous reflection as well as an operation of truncation.

In frames of semi-Markov models of reflection it is natural to assume that $X(t)$ is a semi-Markov process of diffusion type. Let (P_x), $x \geq 0$, be a consistent family of measures of the process, depending on initial points of trajectories. On interval $(0, \infty)$ semi-Markov transition generating functions of the process

$$g_{(a,b)}(\lambda, x) := \mathbb{E}_x \left(e^{-\lambda \sigma_{(a,b)}}; X(\sigma_{(a,b)}) = a \right);$$
$$h_{(a,b)}(\lambda, x) := \mathbb{E}_x \left(e^{-\lambda \sigma_{(a,b)}}; X(\sigma_{(a,b)}) = b \right),$$

$a < x < b$, satisfy the differential equation

$$\frac{1}{2} f'' + A(x) f' - B(\lambda, x) f = 0,$$

with boundary conditions

$$g_{(a,b)}(\lambda, a+) = h_{(a,b)}(\lambda, b-) = 1, \qquad g_{(a,b)}(\lambda, b-) = h_{(a,b)}(\lambda, a+) = 0.$$

The coefficients of the equation are assumed to be piece-wise continuous functions of $x > 0$, and for any x function $B(\lambda, x)$ is non-negative and has completely monotone partial derivative with respect to λ. First of all reflection of the process from point $x = 0$ means addition of this point to the range of values of the process. Further all the semi-closed intervals $[0, r)$ are considered what the process can only exit from open boundary. Corresponding semi-Markov transition generating functions are denoted as $h_{[0,r)}(\lambda, x)$ with main distinction from exit from an open set $h_{[0,r)}(\lambda, 0) > 0$. Function

$K(\lambda,r):=h_{[0,r)}(\lambda,0)$ plays an important role for description of properties of reflected processes. Using semi-Markov properties of the process, we obtain

$$h_{[0,r)}(\lambda,x)=h_{(0,r)}(\lambda,x)+g_{(0,r)}(\lambda,x)K(\lambda,r),$$

and also

$$K(\lambda,r)=K(\lambda,r-\varepsilon)(h_{(0,r)}(\lambda,r-\varepsilon)+g_{(0,r)}(\lambda,r-\varepsilon)K(\lambda,r)).$$

Assuming that there exist derivatives with respect to the second argument we have

$$\begin{aligned}g_{(a,b)}(\lambda,x)&=1+g'_{(a,b)}(\lambda,a+)(x-a)+o(x-a),\\ g_{(a,b)}(\lambda,x)&=-g'_{(a,b)}(\lambda,b-)(b-x)+o(b-x),\\ h_{(a,b)}(\lambda,x)&=h'_{(a,b)}(\lambda,a+)(x-a)+o(x-a),\\ h_{(a,b)}(\lambda,x)&=1-h'_{(a,b)}(\lambda,b-)(b-x)+o(b-x),\end{aligned}$$

and obtain the differential equation

$$K'(\lambda,r)+K(\lambda,r)h'_{(0,r)}(\lambda,r-)+K^2(\lambda,r)g'_{(0,r)}(\lambda,r-)=0.$$

Its general solution is

$$K(\lambda,r)=\frac{h'_{(0,r)}(\lambda,0+)}{C(\lambda)-g'_{(0,r)}(\lambda,0+)},$$

where arbitrary constant $C(\lambda)$ can depend on λ. In order for $K(\lambda,r)$ to be a Laplace transform it is sufficient that function $C(\lambda)$ to be non-decreasing, $C(0)=0$, and its derivative to be a completely monotone function [5]. Under our assumptions it is fair

$$K(\lambda,r)=1-C(\lambda)r+o(r),\qquad r\to 0.$$

Our next task is to learn a time change in the process with instantaneous reflection which derives the process with delayed reflection.

3. Time change with respect to time run under instantaneous reflection

Let us denote θ_t the shift operator on the set of trajectories; σ_Δ the operator of the first exit time from set Δ. For any Markov times τ_1, τ_2 (with respect to the natural filtration) on set $\{\tau_1<\infty\}$ let us determine the following operation

$$\tau_1\dot{+}\tau_2:=\tau_1+\tau_2\circ\theta_{\tau_1}.$$

It is known [4], that for any open (in relative topology) sets Δ_1, Δ_2, if $\Delta_1\subset\Delta_2$, then

$$\sigma_{\Delta_2}=\sigma_{\Delta_1}\dot{+}\sigma_{\Delta_2}.$$

In this case $\sigma_\Delta(\xi)=0$, if $\xi(0)\notin\Delta$.

Let us introduce special denotations for some first exit times and their combinations, and that for random intervals as $\varepsilon>0$

$$\begin{gathered}\alpha:=\sigma_{[0,\varepsilon)},\qquad \beta:=\sigma_{(0,\infty)},\qquad \gamma(0):=\beta,\\ \gamma:=\alpha\dot{+}\beta,\qquad \gamma(n):=\gamma(n-1)\dot{+}\gamma,\qquad n\ge 1,\\ b(0):=[0,\beta),\qquad a(n):=\big[\gamma(n-1),\gamma(n-1)\dot{+}\alpha\big),\qquad b(n)=\big[\gamma(n-1)\dot{+}\alpha,\gamma_n\big).\end{gathered}$$

The random times α, $\gamma(n)$, and intervals $a(n)$, $b(n)$, $n=1,2,\dots$, depend on ε. In some cases we will denote this dependence by the lower index.

Let us remark that sequence $(\gamma(n))$ forms moments of jumps of a renewal process. Besides if $X(t)>0$ then for any $t>0$ there exist $\varepsilon>0$, and $n\ge 1$ such that $t\in b_\varepsilon(n)$. It implies that for $\varepsilon\to 0$ random set $\bigcup_{k=1}^\infty b_\varepsilon(k)$ covers all the set of positive values of process X with probability one. On share of supplementary set (a limit of set $\bigcup_{k=1}^\infty a_\varepsilon(k)$) there remain possible intervals of constancy and also a discontinuum of points (closed set, equivalent to continuum, without any intervals, [2, p. 158]), consisted of zeros of process X. The linear measure of it can be more than or equal to 0. This measure is included as a component in a measure of delaying while reflecting.

It is known ([4, p. 111]) that continuous homogeneous semi-Markov process is a Markov process if and only if it does not contain intrinsic intervals of constancy (it can have an interval of terminal stopping). This does not imply that a process with delayed deflection cannot be globally Markov. Its delaying is exceptionally at the expense of the discontinuum. A process without intervals of constancy at zero, and with the linear measure of the discontinuum of zeros which equals to zero is said to be a process with instantaneous reflection.

We will construct a non-decreasing sequence of continuous non-decreasing functions $V_\varepsilon(t)$, $t \geq 0$, converging to some limit $V(t)$ as $\varepsilon \to 0$ uniformly on every bounded interval.

Let $X(0) > 0$, and $V_\varepsilon(t) = t$ on interval $b(0)$, and $V_\varepsilon(t) = \beta$ on interval $a(1)$. On interval $b(1)$ the process V_ε increases linearly with a coefficient 1. On interval $a(2)$ function V_ε is constant. Then it increases with coefficient 1 on interval $b(2)$, and so on, being constancy on intervals $a(k)$, increasing with coefficient 1 on intervals $b(k)$. Noting that if $\varepsilon_1 > \varepsilon_2$, for any iterval $a_{\varepsilon_2}(k)$ there exists n such that $a_{\varepsilon_2}(k) \subset a_{\varepsilon_1}(n)$, we convince ourself that the sequence of constructed functions does not decrease, bounded and consequently tends to a limit.

Let us define a process with instantaneous reflecting obtained from the original process X as a process, obtained after elimination of all its intervals of constancy at zero, and contraction of a linear measure of its discontinuum of zeros to zero. This process can be represented as a limit (in Skorokhod metric) of a sequence of processes $X_\varepsilon(t)$, determined for all t by formula

$$X_\varepsilon(t) = X\left(V_\varepsilon^{-1}(t)\right),$$

where $V_\varepsilon^{-1}(y)$ is defined as the first hitting time of the process $V_\varepsilon(t)$ to a level y. Hence $X_\varepsilon(t)$ has jumps of value ε at the first hitting time to zero and its iterations. Let us denote the process with instantaneous reflecting as $X_0(t)$, and the map $X \mapsto X_0$ as ϕ_V. Such a process is measurable (with respect to the original sigma-algebra of subsets) and continuous. Let $P_x^0 = P_x \circ \phi_V^{-1}$ be the induced measure of this process.

Then it is clear that V is an inverse time change transforming the process X_0 into the process X, i.e. $X = X_0 \circ V$. In this case for any open interval $\Delta = (a, b)$, $0 < a < b$, or $\Delta = [0, r)$, $r > 0$, it is fair

$$\sigma_\Delta(X_0 \circ V) = V^{-1}(\sigma_\Delta(X_0)).$$

The function V^{-1} we call a direct time change, which corresponds to every "intrinsic" Markov time of the original process (in given case $X_0(t)$) the analogous time of the transformed process.

Remark, that for $\varepsilon_1 > \varepsilon_2$ the set $\{\gamma_{\varepsilon_1}(n), n = 0, 1, 2, \dots\}$ is a subset of the set $\{\gamma_{\varepsilon_2}(n), n = 0, 1, 2, \dots\}$. That is why every Markov time $\gamma_\varepsilon(n)$ is a Markov regeneration time of the process V, what permits in principle to calculate finite-dimensional distributions of this process. On the other hand this process is synonymously characterized by its inverse, i.e. the process

$$V^{-1}(y) := \inf\{t \geq 0 \colon V(t) \geq y\}, \qquad y > 0.$$

This process is more convenient to deal with because Laplace transform of its value at a point y can be found as a limit of a sequence of easy calculable Laplace images of values $V_\varepsilon^{-1}(y)$.

Theorem 1. *A direct time change $V^{-1}(y)$, mapping a process with instantaneous reflection into a process with delayed reflection satisfy the relation*

$$\mathbb{E}_0 \exp\left(-\lambda V^{-1}(y)\right) = \mathbb{E}_0 \exp(-\lambda y - C(\lambda) W(y)), \tag{1}$$

where $W^{-1}(t)$ is a non-decreasing process with independent increments for which

$$\mathbb{E}_0 \exp\left(-\lambda W^{-1}(t)\right) = \exp\left(g'_{(0,\infty)}(\lambda, 0+)t\right). \tag{2}$$

Short proof. Without loss of generality we suppose that $X(0) = 0$. Let $N_\varepsilon(t) = n$ if and only if

$$\sum_{k=1}^{n-1} |b(k)| < t \le \sum_{k=1}^{n} |b(k)|$$

($|a(k)|$ and $|b(k)|$ are lengths of intervals $a(k)$, $b(k)$). Then

$$\begin{aligned}\mathbb{E}_0 \exp\left(-\lambda V^{-1}(y)\right) &= \lim_{\varepsilon\to 0} \mathbb{E}_0 \exp\left(-\lambda V_\varepsilon^{-1}(y)\right)\\ &= \lim_{\varepsilon\to 0} \mathbb{E}_0 \left(-\lambda y - \lambda \sum_{k=1}^{N_\varepsilon(y)} |a(k)|\right).\end{aligned}$$

We have

$$\begin{aligned}
&\mathbb{E}_0 \exp\left(-\lambda\left(V_\varepsilon^{-1}(y) - y\right)\right) = \mathbb{E}_0 \exp\left(-\lambda \sum_{k=1}^{N_\varepsilon(y)} |a(k)|\right)\\
&\quad= \sum_{n=0}^{\infty} \mathbb{E}_0 \exp\left(-\lambda \sum_{k=1}^{n} \alpha\circ\theta_{\gamma(k-1)}; N_\varepsilon(t) = n\right)\\
&\quad= P_\varepsilon(\beta \ge y) + \sum_{n=1}^{\infty} \mathbb{E}_0\left(\exp\left(-\lambda\sum_{k=1}^{n}\alpha\circ\theta_{\gamma(k-1)}\right); \sum_{k=1}^{n-1}|b(k)| < y \le \sum_{k=1}^{n}|b(k)|\right)\\
&\quad= P_\varepsilon(\beta \ge y)\\
&\qquad+ \sum_{n=1}^{\infty} \mathbb{E}_0\Bigg(\exp\left(-\lambda\alpha - \lambda\sum_{k=2}^{n}\alpha\circ\theta_{\gamma(k-1)}\right); \beta\circ\theta_\alpha + \sum_{k=2}^{n-1}\beta\circ\theta_\alpha\circ\theta_{\gamma(k-1)}\\
&\qquad\qquad< y \le \beta\circ\theta_\alpha + \sum_{k=2}^{n}\beta\circ\theta_\alpha\circ\theta_{\gamma(k-1)}\Bigg)\\
&\quad= P_\varepsilon(\beta \ge y) + \sum_{n=1}^{\infty}\int_0^y \mathbb{E}_0\Bigg(\exp\left(-\lambda\alpha - \lambda\sum_{k=2}^{n}\alpha\circ\theta_{\gamma(k-1)}\right); \beta\circ\theta_\alpha \in dx,\\
&\qquad\qquad \sum_{k=2}^{n-1}\beta\circ\theta_\alpha\circ\theta_{\gamma(k-1)} < y - x \le \sum_{k=2}^{n}\beta\circ\theta_\alpha\circ\theta_{\gamma(k-1)}\Bigg)\\
&\quad= P_\varepsilon(\beta \ge y)\\
&\qquad+ \sum_{n=1}^{\infty}\int_0^y \mathbb{E}_0\left(e^{-\lambda\alpha}; \beta\circ\theta_\alpha \in dx\right)\\
&\qquad\qquad\times \mathbb{E}_0\Bigg(\exp\left(-\lambda\sum_{k=2}^{n}\alpha\circ\theta_{\gamma(k-2)}\right);\\
&\qquad\qquad \sum_{k=2}^{n-1}\beta\circ\theta_\alpha\circ\theta_{\gamma(k-2)} < y - x \le \sum_{k=2}^{n}\beta\circ\theta_\alpha\circ\theta_{\gamma(k-2)}\Bigg)
\end{aligned}$$

$$
\begin{aligned}
&= P_\varepsilon(\beta \geq y) \\
&\quad + \sum_{n=1}^{\infty} \int_0^y P_\varepsilon(\beta \in dx) \mathbb{E}_0 \left(e^{-\lambda\alpha}\right) \\
&\qquad \times \mathbb{E}_0 \left(\exp\left(-\lambda \sum_{k=1}^{n-1} \alpha \circ \theta_{\gamma(k-1)} \right); \right. \\
&\qquad\qquad \left. \sum_{k=1}^{n-2} \beta \circ \theta_\alpha \circ \theta_{\gamma(k-1)} < y - x \leq \sum_{k=1}^{n-1} \beta \circ \theta_\alpha \circ \theta_{\gamma(k-1)} \right) \\
&= P_\varepsilon(\beta \geq y) \\
&\quad + \int_0^y P_\varepsilon(\beta \in dx) \mathbb{E}_0 \left(e^{-\lambda\alpha}\right) \sum_{n=0}^{\infty} \mathbb{E}_0 \left(\exp\left(-\lambda \sum_{k=1}^{n} \alpha \circ \theta_{\gamma(k-1)} \right); N_\varepsilon(y-x) = n \right) \\
&= P_\varepsilon(\beta \geq y) + \int_0^y P_\varepsilon(\beta \in dx) \mathbb{E}_0 \left(e^{-\lambda\alpha}\right) \mathbb{E}_0 \exp\left(-\lambda \left(V_\varepsilon^{-1}(y-x) - (y-x) \right) \right).
\end{aligned}
$$

Let us denote $Z(y) := \mathbb{E}_0 \exp(-\lambda(V_\varepsilon^{-1}(y) - y))$, $F(x) := P_x(\beta < x)$, $\overline{F}(x) := 1 - F(x)$, $A := \mathbb{E}_0(e^{-\lambda\alpha})$. We obtain an integral equation

$$
Z(y) = \overline{F}(x) + A \int_0^y Z(y-x)\, dF(x),
$$

with a solution which can be written as follows

$$
Z(y) = \sum_{n=0}^{\infty} A^n \left(F^{(n)}(y) - F^{(n+1)}(y) \right),
$$

where $F^{(n)}$ is n-times convolution of distribution F. Let us consider a sequence of independent and identically distributed random values $|b(n)|$, $n = 1, 2, \dots$. Let P_ε^* is the distribution of a renewal process $N_\varepsilon(y)$ with this sequence of lengths of intervals, and $\mathbb{E}_\varepsilon^*$ is the corresponding expectation. Then

$$
\mathbb{E}_\varepsilon^* A^{N_\varepsilon(y)} = \sum_{n=0}^{\infty} A^n P_\varepsilon^*(N_\varepsilon(y) = n) = \sum_{n=0}^{\infty} A^n \left(F^{(n)}(y) - F^{(n+1)}(y) \right),
$$

Thus

$$
\mathbb{E}_0 \exp\left(-\lambda V_\varepsilon^{-1}(y) \right) = e^{-\lambda y} \mathbb{E}_\varepsilon^* \left(\mathbb{E}_0 e^{-\lambda\alpha} \right)^{N_\varepsilon(y)}.
$$

On the other hand it is clear that there exists a version of the process $N_\varepsilon(y)$, measurable with respect to the basic sigma-algebra, and adapted to the natural filtration of the original process, and having identical distribution with respect to measure P_0. Preserving denotations we can write

$$
\mathbb{E}_\varepsilon^* \left(\mathbb{E}_0 e^{-\lambda\alpha} \right)^{N_\varepsilon(y)} = \mathbb{E}_0 \left(\mathbb{E}_0 e^{-\lambda\alpha} \right)^{N_\varepsilon(y)}.
$$

Moreover, measures P_0 and P_0^0 coincide on sigma-algebra F^*, generated by all the random values $\beta^\varepsilon \circ \theta_{\alpha^\varepsilon} \circ \theta_{\gamma(k)^\varepsilon}$, $\varepsilon > 0$, $k = 1, 2, \dots$. From here

$$
\mathbb{E}_0 \left(\mathbb{E}_0 e^{-\lambda\alpha} \right)^{N_\varepsilon(y)} = \mathbb{E}_0^0 \left(\mathbb{E}_0 e^{-\lambda\alpha} \right)^{N_\varepsilon(y)}.
$$

Taking into account that α depends on ε and using our former denotations we can write

$$
\mathbb{E}_0 e^{-\lambda\alpha} = K(\lambda, \varepsilon) = 1 - C(\lambda)\varepsilon + o(\varepsilon).
$$

We will show that the process $W_\varepsilon(y) := \varepsilon N_\varepsilon(y)$ tends weakly to a limit $W(y)$ as $\varepsilon \to 0$, which is an inverse process with independent increments with known parameters, and measurable with respect to sigma-algebra F^*. Actually, the process $W_\varepsilon(y)$ does not decrease and is characterized completely by the process $W_\varepsilon^{-1}(t)$. The latter has independent positive jumps on the lattice with a pitch ε. Hence it is a process with independent increments. Evidently a limit of a sequence of such processes, if it exists, is a process with independent increments too. Its existence follows from evaluation of Laplace transform of its increment. We have

$$\begin{aligned}\mathbb{E}_0^0 e^{-\lambda W_\varepsilon^{-1}(t)} &= \mathbb{E}_0^0 \exp\left(-\lambda \sum_{k=1}^{[t/\varepsilon]} |b(k)|\right) = \left(\mathbb{E}_\varepsilon e^{-\lambda\beta}\right)^{[t/\varepsilon]} \\ &= \left(1 + g'_{(0,\infty)}(\lambda,0)\varepsilon + o(\varepsilon)\right)^{[t/\varepsilon]} \to e^{g'_{(0,\infty)}(\lambda,0)t}, \qquad \varepsilon \to 0.\end{aligned}$$

Using the sufficient condition of weak convergence of processes in terms of convergence of their points of the first exit from open sets ([4], p. 287), we obtain

$$\mathbb{E}_0 \exp\left(-\lambda V^{-1}(y)\right) = \mathbb{E}_0^0 \exp\left(-\lambda y - C(\lambda) W(y)\right),$$

what can be considered as description of the direct time change in terms of the process with instantaneous reflection and the main characteristic of delaying, function $C(\lambda)$. $\square$

We use this formula for deriving the Laplace transform of a difference between the first exit times from an one-sided neighborhood of the boundary point for processes with delayed and instantaneous reflection.

Denote

$$\begin{gathered}\beta^r := \sigma_{(0,r)}, \qquad \gamma^r(0) = 0, \\ \gamma^r := \alpha \dot{+} \beta^r, \qquad \gamma^r(n) := \gamma^r(n-1) \dot{+} \gamma^r, \qquad n \geq 1, \\ b^r(n) = \left[\gamma^r(n-1) \dot{+} \alpha, \gamma^r(n)\right), \qquad n \geq 1, \\ M_\varepsilon^r := \inf\{n \geq 0 \colon X(\gamma^r(n)) \geq r\}.\end{gathered}$$

Hence

$$\begin{gathered}P_0(M_\varepsilon^r = n) = P_0(X(\gamma^r(1)) = 0, \qquad \dots, \qquad X(\gamma^r(n-1)) = 0, \\ X(\gamma^r(n-1)) = r) = (p(\varepsilon, r))^{n-1}(1 - p(\varepsilon, r)),\end{gathered}$$

where $p(\varepsilon, r) := P_0(X(\gamma^r(1)) = 0)$.

Theorem 2. *A difference between the first exit times from a semi-closed interval* $[0, r)$ *for processes with delayed and instantaneous reflection obeys to the relation*

$$\mathbb{E}_0 \exp\left(-\lambda\left(\sigma_{[0,r)} - \sigma^0_{[0,r)}\right)\right) = \frac{-G'_{(0,r)}(0+)}{C(\lambda) - G'_{(0,r)}(0+)}, \tag{3}$$

where $G_{(0,r)}(x) = g_{(0,r)}(0, x)$.

Short proof. Let $X(0) = 0$. Then evidently, $\sigma_{[0,r)} = \gamma^r_{M_\varepsilon^r}$ for any $\varepsilon < r$. On the other hand, it is clear, that $\gamma^r = \gamma$ on the set $\{X(\gamma^r) = 0\}$, and by induction we conclude that

$$\gamma^r(n) = \gamma(n) \quad \text{ on the set } \quad \bigcap_{k=1}^n \{X(\gamma^r(k)) = 0\}.$$

From here

$$\begin{aligned}\gamma^r(n) I(M_\varepsilon^r = n) &= (\gamma^r(n-1) \dot{+} \gamma^r) I\left(\bigcap_{k=1}^{n-1} \{X(\gamma^r(k)) = 0\}\right) \cap \{X(\gamma^r(n)) = r\} \\ &= (\gamma(n-1) \dot{+} \gamma^r) I(M_\varepsilon^r = n).\end{aligned}$$

Let us denote $\sigma^0_{[0,r)}$ the first exit time from interval $[0,r)$ of the process with instantaneous reflection (formally it means $\sigma^0_{[0,r)} = V(\sigma_{[0,r)})$). Then $V^{-1}\big(\sigma^0_{[0,r)}\big) = \sigma_{[0,r)}$, and from formula (1) it follows that

$$\begin{aligned}
&\mathbb{E}_0 \exp\Big(-\lambda\Big(\sigma_{[0,r)} - \sigma^0_{[0,r)}\Big)\Big) = \mathbb{E}^0_0 \exp\Big(-C(\lambda)W\Big(\sigma^0_{[0,r)}\Big)\Big) \\
&\quad = \mathbb{E}^0_0 \exp\big(-C(\lambda)W(\gamma^r(M^r_\varepsilon))\big) = \sum_{n=1}^\infty \mathbb{E}^0_0\left(\exp\big(-C(\lambda)W(\gamma^r(n))\big); M^r_\varepsilon = n\right) \\
&\quad = \sum_{n=1}^\infty \mathbb{E}^0_0 \exp\left(-C(\lambda)W\left(\gamma(n-1)\dot{+}\gamma^r\right); M^r_\varepsilon = n\right) \\
&\quad = \sum_{n=1}^\infty \mathbb{E}^0_0\left(\exp(-C(\lambda)W\left(\sum_{k=1}^{n-1}(|a(k)| + |b^r(k)|) + |a(n)| + |b^r(n)|\right); M^r_\varepsilon = n\right).
\end{aligned}$$

Taking into account P^0_0-almost sure convergence $\sum_{k=1}^{M^r_\varepsilon}(|a(k)| \to 0$ as $\varepsilon \to 0$, we have

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\sum_{n=1}^\infty \mathbb{E}^0_0\left(\exp(-C(\lambda)W\left(\sum_{k=1}^{n-1}(|a(k)| + |b^r(k)|) + |a(n)| + |b^r(n)|\right); M^r_\varepsilon = n\right) \\
&\quad = \lim_{\varepsilon\to 0}\mathbb{E}^0_0\left(\exp(-C(\lambda)W\left(\sum_{k=1}^{M^r_\varepsilon-1}|b^r(k)| + |b^r(M^r_\varepsilon)|\right)\right) \\
&\quad = \lim_{\varepsilon\to 0}\mathbb{E}^0_0\left(\exp(-C(\lambda)\varepsilon N_\varepsilon\left(\sum_{k=1}^{M^r_\varepsilon-1}|b^r(k)| + |b^r(M^r_\varepsilon)|\right)\right).
\end{aligned}$$

From the definition of the process $N_\varepsilon(t)$ it folows that

$$N_\varepsilon\left(\sum_{k=1}^{n-1}|b^r(k)| + |b^r(n)|\right) = n, \qquad n = 1, 2 \ldots.$$

Consecuently

$$\begin{aligned}
\mathbb{E}_0 \exp\Big(-\lambda\Big(\sigma_{[0,r)} - \sigma^0_{[0,r)}\Big)\Big) &= \lim_{\varepsilon\to 0}\mathbb{E}^0_0 \exp\left(-C(\lambda)\varepsilon M^r_\varepsilon\right) \\
&= \lim_{\varepsilon\to 0}\sum_{n=1}^\infty e^{-C(\lambda)\varepsilon n}(p(\varepsilon,r))^{n-1}(1-p(\varepsilon,r)) \\
&= \lim_{\varepsilon\to 0} e^{-C(\lambda)\varepsilon}\frac{1-p(\varepsilon,r)}{1-e^{-C(\lambda)\varepsilon}p(\varepsilon,r)}.
\end{aligned}$$

and taking into account that

$$\begin{aligned}
p(\varepsilon,r) &= P_0(X(\gamma^r_\varepsilon) = 0) = P_0(X(\alpha_\varepsilon\dot{+}\beta^r_\varepsilon) = 0) = P_0(X(\beta^r_\varepsilon)\circ\theta_{\alpha_\varepsilon} = 0) \\
&= P_\varepsilon(X(\beta^r_\varepsilon) = 0) := G_{(0,r)}(\varepsilon),
\end{aligned}$$

and that the last expression (the partial case $g_{(0,r)}(\lambda,\varepsilon)$ for $\lambda = 0$) has an asymptotic $G_{(0,r)}(\varepsilon) = 1 + G'_{(0,r)}(0+)\varepsilon + o(\varepsilon)$, we obtain at last

$$\mathbb{E}_0 \exp\Big(-\lambda\Big(\sigma_{[0,r)} - \sigma^0_{[0,r)}\Big)\Big) = \frac{-G'_{(0,r)}(0+)}{C(\lambda) - G'_{(0,r)}(0+)}. \qquad \square$$

It is interesting to note that for a linear function $C(\lambda) = k\lambda$, when a reflecting locally Markov process is globally Markov [5], the difference between the first exit times from a semi-closed interval $[0,r)$ for processes with delayed and instantaneous reflection has the exponential distribution with parameter $-G'_{(0,r)}(0+)/k$.

4. Example

Let us consider the standard Wiener process truncated in its negative values

$$\overline{w}(t) = \begin{cases} 0, & w(t) \le 0, \\ w(t), & w(t) > 0. \end{cases}$$

In frames of the semi-Markov model of reflection it is characterized by the function

$$K(\lambda, r) = \frac{h'_{(0,r)}(\lambda, 0+)}{C(\lambda) - g'_{(0,r)}(\lambda, 0+)} = \frac{\sqrt{2\lambda}/\sinh r\sqrt{2\lambda}}{C(\lambda) - \sqrt{2\lambda}\cosh r\sqrt{2\lambda}/\sinh r\sqrt{2\lambda}}.$$

Taking into account the origin of this process one can write

$$K(\lambda, r) = \mathbb{E}_0^w \exp\left(-\lambda \sigma_{(-\infty,r)}\right) = \exp\left(-r\sqrt{2\lambda}\right).$$

Comparing derivatives at zero of these two representations of the same function, we obtain $C(\lambda) = \sqrt{2\lambda}$. Now we can obtain the main characteristic of delay of this process under reflection (including lengths of all the intervals of constancy) from the first hitting time of the level 0 up to the first hitting time of the level r:

$$\mathbb{E}_0 \exp\left(-\lambda\left(\sigma_{[0,r)} - \sigma^0_{[0,r)}\right)\right) = \frac{1/r}{\sqrt{2\lambda} + 1/r},$$

what relates to tabulated values of Laplace transforms, and here is not exposed.

References

1. I. I. Gihman and A. V. Skorokhod, *Stochastic Differential Equation*, “Naukova dumka”, Kiev, 1968. (Russian)
2. F. Hausdorff, *Theory of Sets*, “KomKniga”, Moscow, 2006. (Russian)
3. B. P. Harlamov, *Diffusion process with delay on edges of a segment*, Zapiski Nauchnyh Seminarov POMI **351** (2007), 284–297. (Russian)
4. B. P. Harlamov, *Continuous Semi-Markov Processes*, ISTE & Wiley, London, 2008.
5. B. P. Harlamov, *On a Markov diffusion process with delayed reflection on boundaries of a segment*, Zapiski Nauchnyh Seminarov POMI **368** (2009), 231–255. (Russian)
6. S. S. Rasova and B. P. Harlamov, *On moving of Brownian particles along a delaying skreen*, Zapiski Nauchnyh Seminarov POMI **396** (2011), 175–194. (Russian)
7. B. P. Harlamov, *Stochastic model of gas capillary chromatography*, Communication in Statistics—Simulation and Computation **41** (2012), no. 7, 1023–1031.

Institute of Problems of Mechanical Engineering, RAS, Saint-Petersburg, Russia
E-mail address: b.p.harlamov@gmail.com

Received 30/01/2013

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 21–29

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 21–29

ON A CONJECTURE OF ERDÖS ABOUT ADDITIVE FUNCTIONS

UDC 519.21

KARL-HEINZ INDLEKOFER

Abstract. For a real-valued additive function $f\colon \mathbb{N}\to\mathbb{R}$ and for each $n\in\mathbb{N}$ we define a distribution function

$$F_n(x) := \frac{1}{n}\#\{m\le n\colon f(m)\le x\}.$$

In this paper we prove a conjecture of Erdös, which asserts that in order for the sequence F_n to be (weakly) convergent, it is sufficient that there exist two numbers $a<b$ such that $\lim_{n\to\infty}(F_n(b)-F_n(a))$ exists and is positive.

The proof is based upon the use of the Stone–Čech compactification $\beta\mathbb{N}$ of $\mathbb{N}$ to mimic the behaviour of an additive function as a sum of independent random variables.

Анотація. Для дійсної адитивної функції $f\colon \mathbb{N}\to\mathbb{R}$ при всіх $n\in\mathbb{N}$ ми визначаємо функцію розподілу

$$F_n(x) := \frac{1}{n}\#\{m\le n\colon f(m)\le x\}.$$

У статті ми доводимо гіпотезу Ердеша, яка стверджує, що для (слабкої) збіжності послідовності F_n достатньою умовою є існування двох чисел $a<b$ таких, що границя $\lim_{n\to\infty}(F_n(b)-F_n(a))$ існує і додатна.

Доведення базується на використанні компактифікації Стоуна–Чеха (Stone–Čech) $\beta\mathbb{N}$ для $\mathbb{N}$, що дає змогу дослідити поведінку адитивної функції, трактуючи її як суму незалежних випадкових величин.

Аннотация. Для вещественной адитивной функции $f\colon \mathbb{N}\to\mathbb{R}$ при всех $n\in\mathbb{N}$ мы определяем функцию распределения

$$F_n(x) := \frac{1}{n}\#\{m\le n\colon f(m)\le x\}.$$

В статье мы доказываем гипотезу Эрдеша, в которой утверждается, что для (слабой) сходимости последовательности F_n достаточным условием является существование двух чисел $a<b$ таких, что предел $\lim_{n\to\infty}(F_n(b)-F_n(a))$ существует и положителен.

Доказательство основано на использовании компактификации Стоуна–Чеха (Stone–Čech) $\beta\mathbb{N}$ для $\mathbb{N}$, что дает возможность исследовать поведение адитивной функции, трактуя ее как сумму независимых случайных величин.

1. Introduction

A function $f\colon \mathbb{N}\to\mathbb{R}$ is called *additive* if $f(mn)=f(m)+f(n)$ for any coprime integers m and n. Then f is defined by its values $f(p^k)$ on prime powers p^k (p prime, $k\in\mathbb{N}$) and $f(1)=0$.

Given a real-valued additive function f, one can define, for each $n\in\mathbb{N}$, a distribution function

$$F_n(x) := \frac{1}{n}\#\{m\le n\colon f(m)\le x\}. \tag{1.1}$$

An old conjecture of Erdös in 1947 (see Erdös [4]) asserts that in order for the sequence F_n to be (weakly) convergent (in this case we say that the additive function f possesses a *limit distribution*), it is sufficient that there exist two numbers $a<b$ such

2000 *Mathematics Subject Classification.* Primary 11N37, 11N60, 11K65.

Key words and phrases. Probabilistic number theory, additive functions.

that

$$\lim_{n\to\infty}(F_n(b)-F_n(a)) \quad \text{exists and is positive.} \tag{1.2}$$

In 1992 A. Hildebrand [6] could show that the conclusion of Erdös' conjecture is valid, provided (1.2) is strengthened to

$$L_a := \lim_{n\to\infty} F_n(a) \quad \text{and} \quad L_b := \lim_{n\to\infty} F_n(b) \tag{1.3}$$

both exists and $L_a \neq L_b$. Some further discussions are contained in [10] and [11].

In this paper we show that the above conjecture of Erdös holds.

Theorem. *Let $f\colon \mathbb{N} \to \mathbb{R}$ be an additive function. In order for the distributions* (1.1) *to converge, it is sufficient that* (1.2) *holds for some $a < b$.*

The proof is based upon a method, introduced in [7, 8] using the Stone–Čech compactification $\beta\mathbb{N}$ of $\mathbb{N}$ to mimic the behaviour of an additive function as a sum of independent random variables.

2. Finitely distributed additive functions

An additive function f is said to be *finitely distributed* if there are positive constants c_1 and c_2, and an unbounded sequence $n_1 < n_2 < \dots$ so that for every i there exists a sequence

$$a_1^{(i)} < a_2^{(i)} < \cdots < a_{t_i}^{(i)} < n_i$$

satisfying

$$\left| f\left(a_r^{(i)}\right) - f\left(a_s^{(i)}\right)\right| < c_1, \qquad t_i > c_2 n_i,\ 1 \le r, s \le t_i.$$

The necessary and sufficient condition that f should be finitely distributed is that there should exist a constant c and an additive function h so that

$$f(n) = c\log n + h(n) \tag{2.1}$$

where both the series

$$\sum_{|h(p)|>1} \frac{1}{p}, \qquad \sum_{\substack{p\\ |h(p)|\le 1}} \frac{h^2(p)}{p} \tag{2.2}$$

converge (Erdös [3], 1946). Further characterizations of finitely distributed additive functions can be found in Ch. 7 of Elliott's book [2]. For our purpose we shall apply the following ([2, p. 259]).

Proposition. *If the additive function has a representation* (2.1) *with convergent series* (2.2)*, then, if we define*

$$\alpha(n) = c\log n + \sum_{\substack{p\le n\\ |h(p)|\le 1}} \frac{h(p)}{p}, \tag{2.3}$$

the distribution functions

$$G_n(x) := \frac{1}{n}\#\{m \le n\colon f(m) - \alpha(n) \le x\} \tag{2.4}$$

weakly converge to some distribution function $G(x)$.

If (1.2) holds then f is finitely distributed. Now, assume that (2.1) holds and $\alpha(n)$ is unbounded. Then, if $\alpha(n'_k) \to \infty$, $k \to \infty$, for some subsequence (n'_k), by (1.2),

$$\lim_{k\to\infty}\left\{G_{n'_k}(b-\alpha(n'_k)) - G_{n'_k}(a-\alpha(n'_k))\right\} = \lim_{k\to\infty}\left(F_{n'_k}(b) - F_{n'_k}(a)\right) > 0 \tag{2.5}$$

whereas the left side in (2.5) tends to zero since G_n converge weakly to some distribution function. Then, since $\alpha(n) = c\log n + O(\log\log n)$, we conclude $c = 0$, i.e. $f = h$, and

$$A(n) := \sum_{\substack{p\le n\\ |f(p)|\le 1}} \frac{f(p)}{p} = O(1) \quad \text{for all } n \in \mathbb{N}. \tag{2.6}$$

In the following we assume that

$$\sum_{\substack{p\\ |f(p)|\le 1}} \frac{f(p)}{p} \quad \text{diverges,} \tag{2.7}$$

which implies (see [3], Theorem II) that $G(x)$ is continuous and strictly increasing for all $x \in \mathbb{R}$.

For each $n \in \mathbb{N}$ define the additive function f_n by

$$f_n(p^k) = \begin{cases} f\left(p^k\right) & \text{if } p \le n, \\ 0, & \text{otherwise.} \end{cases}$$

and put, for $A \subset \mathbb{N}$,

$$\delta_n(A) := \frac{1}{n}\#\{m \le n \colon m \in A\}.$$

If the limit

$$\delta(A) := \lim_{n\to\infty} \delta_n(A) \tag{2.8}$$

exists we say that A *possesses the asymptotic density* $\delta(A)$.

If some sequence $\{n'_k\}$ is given we write

$$\delta'(A) := \lim_{k\to\infty} \delta_{n'_k}(A) \tag{2.9}$$

in the case the limit (2.9) exists.

With these notations we show

Lemma 1. *Assume that* (1.2) *holds. Then*

$$\lim_{n\to\infty} \delta(\{m \colon f_n(m) \in (a,b]\}) = \delta(\{m \colon f(m) \in (a,b]\}) =: c_0 > 0. \tag{2.10}$$

Proof. Observe that $\delta(\{m \colon f_n(m) \in (a,b]\})$ always exists. Assume that (2.10) does not hold. Then there exists a sequence $\{n_k\}$ of natural numbers such that

$$\lim_{k\to\infty} \delta(\{m : f_{n_k}(m) \in (a,b]\}) = c' \ne c_0.$$

Since $A(n_k) = O(1)$ there exists some subsequence $\{n'_k\}$ of $\{n_k\}$ so that

$$\lim_{k\to\infty} A(n'_k) =: A'$$

exists. Choose k_1 such that for every $k_0 \ge k_1$

$$\begin{aligned} &\left|\delta\left(\left\{m \colon f_{n'_{k_0}}(m) \in (a,b]\right\}\right) - c_0\right| \\ &\quad = \left|\delta'\left(\left\{m \colon f_{n'_{k_0}}(m) \in (a,b]\}\right) - \delta'(\{m \colon f(m) \in (a,b]\right\}\right)\right| \\ &\quad \ge \frac{|c_0 - c'|}{2}. \end{aligned} \tag{2.11}$$

On the other hand we shall show that

$$\lim_{k_0\to\infty} \overline{\lim_{k\to\infty}}\, \delta_{n'_k}\left(\left\{m \colon \left|f(m) - f_{n'_{k_0}}(m)\right| > \varepsilon\right\}\right) = 0 \tag{2.12}$$

for every $\varepsilon > 0$ which contradicts (2.11).

For the proof of (2.12) put

$$\mathcal{P}_0 := \{p\colon |f(p)| > 1\} \cup \left\{p^k\colon k \geq 2\right\}.$$

Define the functions

$$h'(m) := \sum_{\substack{p^k \| m \\ p^k \in \mathcal{P}_0 \\ p > n'_{k_0}}} f\left(p^k\right)$$

and

$$j(m) := \sum_{\substack{p \| m \\ |f(p)| \leq 1 \\ p > n'_{k_0}}} f(p).$$

From our definitions of these functions

$$f(m) - f_{n'_{k_0}}(m) = j(m) - \left(A(n'_k) - A\left(n'_{k_0}\right)\right) + h'(m) + \left(A(n'_k) - A\left(n'_{k_0}\right)\right).$$

We shall prove that for every $\varepsilon > 0$ each of the three expressions

$$L_1(k_0) = \overline{\lim_{k\to\infty}}\, \delta_{n'_k}\left(\left\{m\colon \left|j(m) - \left(A(n'_k) - A\left(n'_{k_0}\right)\right)\right| > \varepsilon\right\}\right),$$

$$L_2(k_0) = \overline{\lim_{k\to\infty}}\, \delta_{n'_k}(\{m\colon |h'(m)| > \varepsilon\})$$

and

$$L_3(k_0) = \overline{\lim_{k\to\infty}}\, \delta_{n'_k}\left(\left\{m\colon \left|A\left(n'_k\right) - A\left(n'_{k_0}\right)\right| > \varepsilon\right\}\right)$$

converge to zero as $k_0 \to \infty$. We may readily estimate the first of these three expressions by appealing to the Turan–Kubilius inequality. In our present circumstances it becomes

$$\frac{1}{n'_k} \sum_{m=1}^{n'_k} \left|j(m) - \left(A\left(n'_k\right) - A\left(n'_{k_0}\right)\right)\right|^2 \ll \sum_{\substack{n'_{k_0} < p \leq n'_k \\ |f(p)| \leq 1}} \frac{|f(p)|^2}{p}.$$

Appealing to the convergence of the second sum in (2.2) we see that

$$L_1(k_0) \ll \frac{1}{\varepsilon^2} \sum_{\substack{n'_{k_0} < p \\ |f(p)| \leq 1}} \frac{|f(p)|^2}{p} = o(1) \quad \text{as } k_0 \to \infty.$$

The estimate $L_3(k_0) = o(1)$ as $k_0 \to \infty$ is obvious.

If an integer m is counted in the expression $L_2(k_0)$ it must satisfy one of two divisibility criteria.

First, it may be divisible by the square of a prime $p > n'_{k_0}$. The frequency of these integers is at most

$$\delta_{n'_k}\left(\left\{m\colon p^2 | m, p > n'_{k_0}\right\}\right) \leq \sum_{n'_{k_0} < p} \frac{1}{p^2} = o(1) \quad \text{as } k_0 \to \infty.$$

Next, it may be exactly divisible by a prime in the range $n'_{k_0} < p$ for which $|f(p)| > 1$. From the hypothesis (2.2) we deduce that the frequencies of such integers is at most

$$\sum_{\substack{n'_{k_0} < p \\ |f(p)| > 1}} \frac{1}{p} = o(1) \quad \text{as } k_0 \to \infty$$

and thus $L_2(k_0) = o(1)$ as $k_0 \to \infty$. We have now shown that (2.12) holds and completed the proof of Lemma 1. □

In the next step we identify the additive function f with a sum $\sum_{p \text{ prime}} X_p$ of independent random variables.

3. Additive functions as a sum of independent random variables

For the sake of simplicity we restrict ourselves to strongly additive functions. Then f can be written in the form

$$f = \sum_p f(p)\varepsilon_p$$

where

$$\varepsilon_p(n) = \begin{cases} 1 & \text{if } p|n, \\ 0, & \text{otherwise.} \end{cases}$$

If $\mathcal{A}$ denotes the algebra generated by the sets

$$A_p := \{n \in \mathbb{N}\colon p|n\}, \quad p \text{ prime},$$

then obviously each $A \in \mathcal{A}$ possesses an asymptotic density $\delta(A)$ and $\delta(A_p) = \frac{1}{p}$ (p prime). Thus δ defines a content on $\mathcal{A}$. Now the construction runs as follows. (For details see [7, 8].) We embed $\mathbb{N}$, endowed with the discrete topology, in the Stone–Čech compactification $\beta\mathbb{N}$,

$$\mathbb{N} \hookrightarrow \beta\mathbb{N}$$

and, if for any $A \subset \mathbb{N}$, the closure of A in $\beta\mathbb{N}$ is denoted by $\bar{A}$, then

$$\bar{\mathcal{A}} := \{\bar{A} \subset \beta\mathbb{N}\colon A \in \mathcal{A}\}$$

is an algebra, too. The extension $\bar{\delta}$ of δ

$$\bar{\delta}(\bar{A}) := \delta(A), \quad \bar{A} \in \bar{\mathcal{A}},$$

defines a premeasure on $\bar{\mathcal{A}}$ and leads to a measure P, induced by

$$\delta^*(A) := \overline{\lim_{n\to\infty}}\, \delta_n(A) \quad \text{for all } A \subset \mathbb{N},$$

and to a probability space $(\Omega, \sigma(\bar{\mathcal{A}}), \mathsf{P})$ with $\Omega = \beta\mathbb{N}$ and with $\mathsf{P}(\bar{A}_p) = 1/p$, p prime.

There is a unique extension of ε_p to a function $\bar{\varepsilon}_p$ on Ω, and putting $X_p = f(p)\bar{\varepsilon}_p$

$$f = \sum_p f(p)\varepsilon_p \to X = \sum_p f(p)\bar{\varepsilon}_p = \sum_p X_p$$

$$f_n \to S_n := \sum_{p\le n} X_p$$

with

$$\mathsf{P}(X_p = f(p)) = \frac{1}{p}$$

and

$$\mathsf{P}(X_p = 0) = 1 - \frac{1}{p}.$$

The $\bar{\varepsilon}_p$ are independent, i.e. $X = \sum_p X_p$ is a sum of independent random variables. If (1.2) holds then, by Lemma 1,

$$\lim_{n\to\infty} \mathsf{P}(S_n \in (a,b]) = c_0 > 0$$

and, by Proposition, $\sum_p X_p$ is essentially convergent (for the definition see [13, p. 262]).

Putting

$$a_p = \mathsf{E}(X_p^c), \qquad Y_p = X_p - a_p, \qquad T_n := \sum_{p\le n} Y_p$$

then $\lim_{n\to\infty} T_n$ holds a.s.. (Here X_p^c denotes the truncation of X_p at (a positive) c, i.e. we replace X_p by $X_p = X$ or 0 according as $|X_p| < c$ or $|X_p| \geq c$.) Denote $Y := \lim T_n$ a.s.

It is well-known that the a.s. convergence of $Y = \sum_p Y_p$ is equivalent to the weak convergence of the distributions of the partial sums of that series. Moreover, by Kolmogorov's three series theorem, $Y = \sum_p Y_p$ converges a.s. if and only if the series

$$\sum_p \mathsf{E}\left(Y_p^c\right), \qquad \sum_p \mathsf{P}(|Y_p| > c), \qquad \sum_p \mathrm{Var}\left(Y_p^c\right) \tag{3.1}$$

converge.

We choose $c = 1$, i.e. $a_p = \mathsf{E}(X_p^1)$ and put (see (2.6))

$$A(n) = \sum_{p\leq n} a_p.$$

Then $A(n) = O(1)$ and, the divergence of the sequence $A(n)$ implies (see [3, Theorem 2]).

Lemma 2. *Let $Y = \sum_p Y_p$ with $Y_p = X_p - a_p$ as above, where the partial sums $\sum_{p\leq N} a_p$ are bounded and divergent. Then the distribution function $G(x) = \mathsf{P}(Y \leq x)$ is continuous and strictly monotone for all $x \in \mathbb{R}$.*

Remark. The divergence of the sequence $A(n)$ implies

$$\begin{aligned} \sum_p a_p^- &= -\infty, \\ \sum_p a_p^+ &= +\infty \end{aligned} \tag{3.2}$$

where $a_p^+ = \max(a_p, 0)$ and $a_p^- = \max(-a_p, 0)$. Then the strict monotonicity of the distribution function $G(x)$ in Lemma 2 can be directly proved by a result of A. Hildebrand [6].

For this we define, following the notation of Hildebrand in [6], p. 1206, the range of a random variable X as the set

$$R(X) = \{x \in \mathbb{R}\colon\ \mathsf{P}(|X - x| \leq \varepsilon) > 0 \text{ for every } \varepsilon > 0\},$$

that is, it is equal to the set of points of increase of the distribution function $F(x) = \mathsf{P}(X \leq x)$. The form of this set was described by A. Hildebrand in Lemma 2 of [6] when X is given as an a.s. convergent series of independent random variables. A special version of this result is contained in the following lemma.

Lemma 3. *Let $\sum_{n=0}^{\infty} X_n$ be an a.s. convergent series of independent random variables and let X denote its sum. Suppose that for every $\varepsilon > 0$ and $n \geq n_0 = n_0(\varepsilon)$ there exist numbers $c_n^- = c_n^-(\varepsilon)$, $c_n^+ = c_n^+(\varepsilon) \in R(X_n)$ with $|c_n^-| \leq \varepsilon$ and $|c_n^+| \leq \varepsilon$ such that*

$$\lim_{N\to\infty} \sum_{n=n_0}^{N} c_n^- = -\infty$$

and

$$\lim_{N\to\infty} \sum_{n=n_0}^{N} c_n^+ = +\infty.$$

Then $R(X) = \mathbb{R}$.

Now it is easy to prove the assertions of Lemma 2. Put

$$c_p^- = \begin{cases} f(p) - a_p & \text{if } -\frac{\varepsilon}{2} \leq f(p) < 0, \\ 0, & \text{otherwise.} \end{cases}$$

and

$$c_p^+ = \begin{cases} f(p) - a_p & \text{if } 0 < f(p) \le \frac{\varepsilon}{2}, \\ 0, & \text{otherwise.} \end{cases}$$

Then obviously, c_p^-, $c_p^+ \in R(Y_p)$, $|c_p^-| \le \frac{\varepsilon}{2} + |a_p| \le \varepsilon$ and $|c_p^+| \le \frac{\varepsilon}{2} + |a_p| \le \varepsilon$ for $p > n_0 = n_0(\varepsilon)$ since $|a_p| \le 1/p$. Further,

$$\begin{aligned}\sum_{n_0 \le p \le N} c_p^- &= \sum_{\substack{n_0 \le p \le N \\ -\frac{\varepsilon}{2} < f(p) < 0}} f(p) - \sum_{n_0 \le p \le N} a_p \\ &< \sum_{\substack{n_0 \le p \le N \\ -\frac{\varepsilon}{2} < f(p) < 0}} \frac{f(p)}{p} + O(1) \\ &< \sum_{\substack{n_0 \le p \le N \\ -1 < f(p) < 0}} \frac{f(p)}{p} + O(1) \\ &= \sum_{n_0 \le p \le N} a_p^- + O(1) \to -\infty \quad \text{as } N \to \infty.\end{aligned}$$

Here the last inequality holds because of the convergence of the second series in (3.1). Similarly,

$$\lim_{N \to \infty} \sum_{n_0 \le p \le N} c_p^+ = +\infty.$$

We use Lemma 3 and recall that the divergence of the series (2.7) implies, by Levy's theorem, the continuity of $G(x)$ to end the proof of Lemma 2.

This ends the remark.

For every subsequence $n' = (n'_k)$ of the natural numbers we defined

$$\delta'(A) = \lim_{k \to \infty} \delta_{n'_k}(A)$$

if the limit exists. This leads to a content δ' on $\mathcal{A}$ and a measure P' on $\beta\mathbb{N}$ induced by

$$\delta'^*(A) = \overline{\lim_{k \to \infty}}\, \delta_{n'_k}(A) \quad \text{for all } A \subset \mathbb{N}.$$

Obviously, if $\Omega_0 \subset \beta\mathbb{N}$ is P-measurable it is P'-measurable and $\mathsf{P}(\Omega_0) = P'(\Omega_0)$.

Since every bounded real-valued function g on $\mathbb{N}$ extends uniquely to a (continuous) function $\bar{g}$ on $\beta\mathbb{N}$ (for details see R. Walker [14, p. 8 et seq.]), we conclude

$$\Omega_0 := \overline{\{m\colon f(m) \in (a, b]\}} = \{\omega\colon \bar{f}(\omega) \in [a, b]\},$$

where $\bar{f}$ is the unique extension of the (bounded) function $f_{(a,b]}$, defined by

$$f_{(a,b]}(m) = \begin{cases} f(m), & \text{if } f(m) \in (a, b], \\ |a| + |b| + 1, & \text{if } f(m) \notin (a, b]. \end{cases}$$

If (1.2) holds then

$$\mathsf{P}(\Omega_0) = c_0 > 0.$$

4. Proof of the conjecture of Erdös

We suppose that $A(n)$ is not convergent so that

$$\underline{A} := \liminf_{n \to \infty} A(n) < \limsup_{n \to \infty} A(n) =: \bar{A}, \tag{4.1}$$

and we shall show that this leads to a contradiction.

We fix two increasing sequences $n' = \{n'_k\}$ and $n'' = \{n''_k\}$ of positive integers so that

$$\underline{A} = \lim_{k\to\infty} A(n''_k) \quad \text{and} \quad \bar{A} = \lim_{k\to\infty} A(n'_k).$$

We put

$$g_n = \sum_{p\le n} (f(p)\varepsilon_p - a_p)$$

and define

$$g' = g_{n'_1} + \sum_{k=1}^{\infty} \left(g_{n'_{k+1}} - g_{n'_k} \right). \tag{4.2}$$

Then

$$\left\{m\colon g'(m) \in (a - \overline{A}, b - \overline{A}]\right\} = \{m\colon f(m) \in (a, b]\}$$

since $g'(m) = f(m) - \overline{A}$ for every $m \in \mathbb{N}$. Further

$$\delta'\left(\left\{m\colon g'(m) \in (a - \overline{A}, b - \overline{A}]\right\}\right) = \lim_{k\to\infty} \delta'\left(\left\{m\colon g_{n'_k}(m) \in (a - \overline{A}, b - \overline{A}]\right\}\right) = c_0.$$

In the same way we define

$$g'' = g_{n''_1} + \sum_{k=1}^{\infty} \left(g_{n''_{k+1}} - g_{n''_k} \right)$$

with $g''(m) = f(m) - \underline{A}$, $m \in \mathbb{N}$, and obtain

$$\delta''\left(\{m\colon g''(m) \in (a - \underline{A}, b - \underline{A}]\}\right) = \lim_{k\to\infty} \delta''\left(\left\{m\colon g_{n''_k}(m) \in (a - \underline{A}, b - \underline{A}]\right\}\right) = c_0.$$

Defining the corresponding extensions $\overline{g'}$ and $\overline{g''}$ and P' and P'', respectively, we arrive at

$$\Omega_0 = \left\{\omega\colon \overline{g'}(\omega) \in [a - \overline{A}, b - \overline{A}]\right\} = \left\{\omega\colon \overline{g''}(\omega) \in [a - \underline{A}, b - \underline{A}]\right\}$$

together with

$$P'\left(\left\{\omega\colon \overline{g'}(\omega) \in [a - \overline{A}, b - \overline{A}]\right\}\right) = P''\left(\left\{\omega\colon \overline{g''}(\omega) \in [a - \underline{A}, b - \underline{A}]\right\}\right) = c_0.$$

Since

$$g' \quad \text{corresponds to} \quad Y' = \lim_{k\to\infty} T_{n'_k}$$

$$g'' \quad \text{corresponds to} \quad Y'' = \lim_{k\to\infty} T_{n''_k}$$

and since

$$Y = \sum_p Y_p = \lim_{n\to\infty} T_n$$

converges a.s. with respect to P and possesses an everywhere continuous distribution function we conclude

(i) $\{\omega\colon Y'(\omega) \in [a - \overline{A}, b - \overline{A}]\} = \Omega'_0$ with $P'(\Omega_0 \triangle \Omega'_0) = 0$,
(ii) $P'(\{\omega\colon Y'(\omega) \in [a - \overline{A}, a - \underline{A}]\}) \le P'(\{\omega\colon Y'(\omega) \ne Y''(\omega)\}) = 0$ and
(iii) $P'(\{\omega\colon Y'(\omega) \in [a - \underline{A}, b - \overline{A}]\}) = c_0$.

Observe, that (iii) implies that

$$a - \underline{A} < b - \overline{A}.$$

Since $\mathsf{P}(\{\omega\colon Y(\omega) \in [a - \overline{A}, a - \underline{A}]\})$ exists it must be zero by (ii), i.e.

$$\mathsf{P}\left(\left\{\omega\colon Y(\omega) \in [a - \overline{A}, a - \underline{A}]\right\}\right) = 0. \tag{4.3}$$

In the same way we show

$$\mathsf{P}\left(\left\{\omega : Y(\omega) \in [b - \overline{A}, b - \underline{A}]\right\}\right) = 0. \tag{4.4}$$

(4.3) and (4.4) contradict the monotonicity of $G(x)$, and thus the assertion of Theorem 1 holds.

Acknowledgement. I would like to acknowledge and express my gratitude to my colleagues and friends Imre Kátai, Oleg Klesov and Robert Wagner for fruitful discussions and helpful comments.

References

1. P. D. T. A. Elliott, *A conjecture of Kátai*, Acta Arith. **XXVI** (1974), 11–20.
2. P. D. T. A. Elliott, *Probabilistic Number Theory I/II*, Springer, New York, 1978/1980.
3. P. Erdös, *On the distribution function of additive functions*, Annals of Math. **47** (1946), 1–20.
4. P. Erdös, *Some remarks and corrections to one of my papers*, Bull. Amer. Math. Soc. **53** (1947), 761–763.
5. A. Hildebrand, *Additive and multiplicative functions on shifted primes*, Proc. London Math. Soc. **59(2)** (1989), 209–232.
6. A. Hildebrand, *A note on the convergence of sums of independent random variables*, The Annals of Prob. **20** (1992), 1204–1212.
7. K.-H. Indlekofer, *A new method in probabilistic number theory*, Probability Theory and Applications, Math. Appl. **80** (1992), 299–308.
8. K.-H. Indlekofer, *New approach to probabilistic number theory — compactifications and integration*, Advanced Studies in Pure Mathematics **49** (2005), 133–170.
9. I. Kátai, *On the distribution of arithmetical functions on the set prime plus one*, Composito Math. **19** (1968), 278–289.
10. F. Kong and Q. Tang, *A theorem on the convergence of sums of independent random variables*, Math. Acta Sci. **21 B** (2001), 331–338.
11. F. Kong and Q. Tang, *Notes on Erdös' conjecture*, Math. Acta Sci. **20(4)** (2000), 533–541.
12. P. Levy, *Sur les séries dont les termes sont des variables éventuellement indépendents*, Studia Math. **3** (1931), 119–155.
13. M. Loève, *Probability theory I*, Springer, Heidelberg–New York, 1977.
14. R. Walker, *The Stone–Čech Compactification*, Springer, Heidelberg–New York, 1974.

Department of Mathematics, University of Paderborn, Warburger Strasse 100, 33098 Paderborn, Germany

E-mail address: k-heinz@math.upb.de

Received 29/01/2013

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 30–39

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 30–39

CONSISTENCY AND ASYMPTOTIC NORMALITY OF PERIODOGRAM ESTIMATOR OF HARMONIC OSCILLATION PARAMETERS

UDC 519.21

A. V. IVANOV AND B. M. ZHURAKOVSKYI

Abstract. The problem of detection of hidden periodicities is considered in the paper. In the capacity of useful signal model the harmonic oscillation observed on the background of random noise being a local functional of Gaussian strongly dependent stationary process is taken. For estimation of unknown angular frequency and amplitude of harmonic oscillation periodogram estimator is chosen, for which sufficient conditions of asymptotic normality are obtained and limit normal distribution is found.

Анотація. У статті розглядається задача виявлення прихованої періодичності. В якості моделі корисного сигналу взято гармонічне коливання, що спостерігається на фоні випадкового шуму, який є локальним функціоналом від гаусівського сильно залежного стаціонарного процесу. Для оцінки невідомих кутової частоти та амплітуди гармонічного коливання була обрана періодограмна оцінка, для якої отримано достатні умови асимптотичної нормальності та знайдено граничний нормальний розподіл.

Аннотация. В статье рассматривается задача выявления скрытой периодичности. В качестве модели полезного сигнала взято гармоническое колебание, которое наблюдается на фоне случайного шума, являющимся локальным функционалом от гауссовского сильно зависимого стационарного процесса. Для оценки неизвестных угловой частоты и амплитуды гармонического колебания была выбрана периодограмная оценка, для которой получены достаточные условия асимптотической нормальности и найдено предельное нормальное распределение.

1. Introduction

Detection of hidden periodicities is a problem that has a long history started by Lagrange in XVIII century [1].

In statistical setting the detection of hidden periodicities is the estimation of unknown amplitudes and angular frequencies, generally speaking, of the sum of harmonic oscillations by observation of this sum on the background of a random noise masking these oscillations.

There are many publications on the subject. Among them first of all we have to mention the works by Whittle [2], Walker [3], Hannan [4], Dorogovtsev, Grechka [5], Ivanov [6], Knopov [7], Quinn and Hannan [8], etc. A good survey of the topic one can find in [9].

In the paper the problem of detecting hidden periodicities is considered in the case when we observe the only harmonic oscillation on the background of random noise being a local functional of Gaussian stationary process with strong dependence. For estimation of unknown parameters the periodogram estimator is chosen.

In the proofs we use approach of the paper [4] where the case of weakly dependent Gaussian stationary noise has been considered.

2000 *Mathematics Subject Classification.* Primary 62J02; Secondary 62J99.

Key words and phrases. Hidden periodicities, periodogram estimator, harmonic oscilation.

2. The main result

Suppose the observed random process is of the form

$$X(t) = A_0 \cos \varphi_0 t + \varepsilon(t), \tag{1}$$

where $A_0 > 0$, $\varphi_0 \in (\underline{\varphi}, \overline{\varphi})$, $0 < \underline{\varphi} < \overline{\varphi} < \infty$, and the random noise $\varepsilon(t)$ satisfies the following conditions:

A1. $\varepsilon(t)$, $t \in \mathbb{R}^1$, is a local functional of a Gaussian stationary process $\xi(t)$, that is $\varepsilon(t) = G(\xi(t))$, $G(x)$, $x \in \mathbb{R}^1$, is a Borel function such that $\mathsf{E}\,\varepsilon(0) = 0$, $\mathsf{E}\,\varepsilon^2(0) < \infty$.

A2. $\xi(t)$, $t \in \mathbb{R}^1$, is a real mean square continuous measurable Gaussian stationary process defined on the probability space (Ω, F, P), $\mathsf{E}\,\xi(0) = 0$.

Assume also that one of the next conditions is fulfilled:

A3. Covariance function (c.f.) of the process $\xi(t)$ is $\mathsf{E}\,\xi(t)\xi(0) = B(t) = L(|t|)|t|^{-\alpha}$, $\alpha \in (0,1)$, where $L(t)$, $t \geq 0$, is a nondecreasing slowly varying at infinity function, $\mathsf{E}\,\xi^2(0) = B(0) = 1$.

A4. C.f. of the process $\xi(t)$ is $B(t) = \cos \psi t (1 + t^2)^{-\alpha/2}$, $\alpha \in (0,1)$, $\psi > 0$ is some number, $\varphi_0 \neq \psi$.

Suppose that for a function $G(x) \in L_2\left(\mathbb{R}^1, \varphi(x)\,dx\right)$,

$$\varphi(x) = (2\pi)^{-1/2} e^{-x^2/2},$$

$C_1(G) \neq 0$ or $C_1(G) = \cdots = C_{m-1}(G) = 0$, $C_m(G) \neq 0$, where

$$C_k(G) = \int_{-\infty}^{+\infty} G(t) H_k(t) \varphi(t)\,dx, \qquad k \geq 0,$$

and $H_k(t)$ are Hermite polynomials. Then the number $m \geq 1$ is said to be Hermite rank of G.

We also assume that function $G(\cdot)$ from condition A1 satisfies assumption

B1. $m\alpha > 1$, where α is a parameter of c.f. B.

We need in a result proved in [10].

Lemma 1. *If conditions **A1**, **A2**, and **A3** or **A4** are satisfied, then*

$$\mathsf{E}\left(\sup_{\lambda \in \mathbb{R}^1} \frac{1}{T} \left|\int_0^T e^{-i\lambda t} \varepsilon(t)\,dt\right|\right)^2 \to 0, \qquad T \to \infty.$$

Consider the functional

$$Q_T(\varphi) = \left|\frac{2}{T} \int_0^T X(t) e^{i\varphi t}\,dt\right|^2. \tag{2}$$

The periodogram estimator of the frequency φ_0 is said to be any random variable (r.v.) $\varphi_T \in [\underline{\varphi}, \overline{\varphi}]$ such that $Q_T(\varphi_T) = \max_{\varphi \in [\underline{\varphi}, \overline{\varphi}]} Q_T(\varphi)$.

Theorem 1. *If conditions of Lemma 1 are satisfied, then $\varphi_T \xrightarrow{\mathsf{P}} \varphi_0$, $T \to \infty$.*

Proof. For any fixed φ consider a behavior, as $T \to \infty$, of the value

$$\begin{aligned} Q_T(\varphi) = {} & \frac{4}{T^2}\left(A_0^2 \left|\int_0^T \cos\varphi_0 t e^{i\varphi t}\,dt\right|^2 + \left|\int_0^T \varepsilon(t) e^{i\varphi t}\,dt\right|^2\right) \\ & + \frac{4}{T^2}\left(2A_0 \operatorname{Re}\left[\int_0^T \cos\varphi_0 t e^{i\varphi t}\,dt \int_0^T \varepsilon(t) e^{-i\varphi t}\,dt\right]\right). \end{aligned} \tag{3}$$

As

$$T^{-1}\left|\int_0^T \cos\varphi_0 t e^{i\varphi t}\,dt\right| \le 1,$$

then due to lemma 1, the 2nd and the 3rd summands in the right-hand side of (3) tend to 0, as $T\to\infty$, in probability. Next we have for $\varphi\in[\underline{\varphi},\overline{\varphi}]$

$$\frac{2}{T}\left|\int_0^T \cos\varphi_0 t e^{i\varphi t}\,dt\right| = \begin{cases} \frac{e^{i(\varphi-\varphi_0)T}-1}{i(\varphi-\varphi_0)T}+\frac{e^{i(\varphi+\varphi_0)T}-1}{i(\varphi+\varphi_0)T}, & \varphi\neq\varphi_0,\\ \frac{e^{i\varphi_0 T}-1}{2i\varphi_0 T}+1, & \varphi=\varphi_0.\end{cases} \tag{4}$$

From (3) and (4) it follows that

$$Q_T(\varphi_0)\xrightarrow{\mathsf{P}} A_0^2, \qquad T\to\infty, \tag{5}$$

$$Q_T(\varphi)\xrightarrow{\mathsf{P}} 0, \qquad T\to\infty, \tag{6}$$

uniformly on any set

$$\Phi_\delta=\left\{\varphi\in\left[\underline{\varphi},\overline{\varphi}\right]:\ |\varphi-\varphi_0|\ge\delta\right\}, \qquad \delta>0.$$

By definition of φ_T

$$\begin{aligned}\mathsf{P}\left(|\varphi_T-\varphi_0|\ge\delta\right)&=\mathsf{P}\left(|\varphi_T-\varphi_0|\ge\delta, Q_T(\varphi_T)\ge Q_T(\varphi_0)\right)\\ &\le\mathsf{P}\left(\sup_{\varphi\in\Phi_\delta}Q_T(\varphi)\ge Q_T(\varphi_0)\right)\to 0, \qquad T\to\infty,\end{aligned}$$

according to (5) and (6). □

We define the periodogram estimator of amplitude A_0 as $A_T=Q_T^{1/2}(\varphi_T)$.

Lemma 2. *If conditions of Lemma 1 are satisfied, then*

$$Q_T(\varphi_T)\xrightarrow{\mathsf{P}} A_0^2, \qquad T\to\infty.$$

Proof. Using (3), one can write

$$\begin{aligned}0&\le Q_T(\varphi_T)-\ Q_T(\varphi_0)\\ &=\frac{4A_0^2}{T^2}\left|\int_0^T\cos\varphi_0 t e^{i\varphi_T t}\,dt\right|^2-\frac{4A_0^2}{T^2}\left|\int_0^T\cos\varphi_0 t e^{i\varphi_0 t}\,dt\right|^2+\eta_T,\end{aligned} \tag{7}$$

$$\eta_T\xrightarrow{\mathsf{P}} 0, \qquad T\to\infty.$$

As from (4) we have

$$\sup_{\varphi\in[\underline{\varphi},\overline{\varphi}]}\frac{1}{T}\left|\int_0^T\cos\varphi_0 t e^{i\varphi t}\,dt\right|\le 1 \quad\text{and}\quad \lim_{T\to\infty}\frac{1}{T}\left|\int_0^T\cos\varphi_0 t e^{i\varphi_0 t}\,dt\right|=1,$$

then

$$\varlimsup_{T\to\infty}\sup_{\varphi\in[\underline{\varphi},\overline{\varphi}]}\left\{\frac{4A_0^2}{T^2}\left|\int_0^T\cos\varphi_0 t e^{i\varphi t}\,dt\right|^2-\frac{4A_0^2}{T^2}\left|\int_0^T\cos\varphi_0 t e^{i\varphi_0 t}\,dt\right|^2\right\}\le 0. \tag{8}$$

Taking into account relations (7), (8), we get

$$Q_T(\varphi_T)-\ Q_T(\varphi_0)\xrightarrow{\mathsf{P}} 0, \qquad T\to\infty.$$

According to (5) $Q_T(\varphi_0)\xrightarrow{\mathsf{P}} A_0^2$, $T\to\infty$, so

$$Q_T(\varphi_T)\xrightarrow{\mathsf{P}} A_0^2 \quad \text{as } T\to\infty.$$

□

Theorem 2. *If conditions of Lemma 1 are satisfied, then*

$$T\left(\varphi_T-\varphi_0\right) \xrightarrow{\mathsf{P}} 0, \qquad T\to\infty.$$

Proof. From lemma 2 and (7) it follows

$$\frac{2}{T^2}\left|\int_0^T \cos\varphi_0 t e^{i\varphi_T t}\,dt\right|^2-\frac{2}{T^2}\left|\int_0^T \cos\varphi_0 t e^{i\varphi_0 t}\,dt\right|^2 \xrightarrow{\mathsf{P}} 0, \qquad T\to\infty. \tag{9}$$

In order to satisfy (9), it is necessary and sufficient that (see. (4))

$$\left|\frac{e^{i(\varphi_T-\varphi_0)T}-1}{i\left(\varphi_T-\varphi_0\right)T}+\frac{e^{i(\varphi_T+\varphi_0)T}-1}{i\left(\varphi_T+\varphi_0\right)T}\right|^2-\left|\frac{e^{i\varphi_0 T}-1}{2i\varphi_0 T}+1\right|^2 \xrightarrow{\mathsf{P}} 0, \tag{10}$$

or

$$\frac{\sin\frac{1}{2}(\varphi_T-\varphi_0)T}{\frac{1}{2}(\varphi_T-\varphi_0)T}\xrightarrow{\mathsf{P}} 1 \quad \text{as } T\to\infty.$$

But the latter is possible if and only if

$$T(\varphi_T-\varphi_0)\xrightarrow{\mathsf{P}} 0, \qquad T\to\infty. \qquad \square$$

Consider a vector function

$$a(t)=(a_1(t),a_2(t),\ldots,a_q(t))', \qquad t\ge 0, \tag{11}$$

and family of matrix measures $\mu_T\left(d\lambda\right)=\left(\mu_T^{jl}(d\lambda)\right)_{j,l=1}^{q}$,

$$\mu_T^{jl}\left(d\lambda\right)=\left(a_T^j(\lambda)\overline{a_T^i(\lambda)}\right)\left(\int_{-\infty}^{+\infty}\left|a_T^j(\lambda)\right|^2 d\lambda\right)^{-1/2}\left(\int_{-\infty}^{+\infty}\left|a_T^l(\lambda)\right|^2 d\lambda\right)^{-1/2} d\lambda,$$

$$a_T^j(\lambda)=\int_0^T e^{i\lambda t}a_j(t)\,dt, \qquad j,l=1,\ldots,q.$$

Assume that $\mu_T(d\lambda)$ weakly converges, as $T\to\infty$, to a matrix measure $\mu(d\lambda)$, that is for any continuous bounded function $b(\lambda)$, $\lambda\in\mathbb{R}^1$,

$$\int_{-\infty}^{+\infty} b(\lambda)\,\mu_T(d\lambda)\to\int_{-\infty}^{+\infty} b(\lambda)\,\mu(d\lambda), \qquad T\to\infty.$$

Then the measure $\mu(d\lambda)$ is said to be spectral measure of vector function (11).

To determine the spectral measure of vector (11) one can use the relations [11]

$$\lim_{T\to\infty} d_{iT}^{-1}d_{jT}^{-1}\int_0^T a_i(t+s)a_j(t)\,dt=\int_{-\infty}^{+\infty} e^{i\lambda s}\,\mu_{ij}\left(d\lambda\right), \qquad i,j=1,\ldots,q,$$

with

$$d_{iT}^2=\int_0^T a_i^2(t)\,dt, \qquad i=1,\ldots,q.$$

Let for $j\ge m$

$$f^{*j}(\lambda)=\int_{\mathbb{R}^{j-1}} f(\lambda-\lambda_2-\cdots-\lambda_j)\prod_{i=2}^{j} f\left(\lambda_i\right)\,d\lambda_2\,\ldots\,d\lambda_j$$

be the j-th convolution of the spectral density $f(\lambda)$ of the random process ξ. Note that $B^k(\cdot)\in L_1\left(\mathbb{R}^1\right)$, $k\ge m$, so all the $f^{*j}(\lambda)$, $k\ge m$, are continuous bounded functions.

Further we formulate the general theorem on asymptotic normality of certain vector integrals [12] and will use in the paper partial cases of this result.

Theorem 3. *Suppose assumptions **A1**, **A2**, **B1**, and **A3** or **A4** are fulfilled. In addition the vector function* (11) *possesses spectral measure* $\mu(d\lambda)$ *and*

$$\sup_{t\in[0,T]} d_{iT}^{-1} |a_i(t)| \le k_i \cdot T^{-1/2}, \qquad i = 1, \dots, q, \tag{12}$$

$k_i < +\infty$, $i = 1, \dots, q$, *are some constants.*

Then the vector

$$b_T = d_T^{-1} \int_0^T G(\xi(t))\, a(t)\, dt, \qquad d_T = \operatorname{diag}(d_{iT})_{i=1}^q.$$

is asymptotically, as $T \to \infty$, *normal* $N(0, K)$ *where*

$$K = 2\pi \sum_{k=m}^{\infty} \frac{C_k^2}{k!} \int_{-\infty}^{\infty} f^{*k}(\lambda)\, \mu(d\lambda, \theta). \tag{13}$$

Corollary. *If the conditions of Lemma 1 and **B1** are satisfied, then the random vector*

$$\left(d_{1T}^{-1} \int_0^T \varepsilon(t) \sin\varphi_0 t\, dt, d_{2T}^{-1} \int_0^T \varepsilon(t) t \sin\varphi_0 t\, dt \right)'$$

is asymptotically, as $T \to \infty$, *normal* $N(0, K_1)$ *with*

$$K_1 = 2\pi \sum_{j=m}^{\infty} \frac{C_j^2}{j!} f^{*j}(\varphi_0) \begin{pmatrix} 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 1 \end{pmatrix}.$$

It follows from this fact that the vector

$$\left(T^{-1/2} \int_0^T \varepsilon(t) \sin\varphi_0 t\, dt, T^{-3/2} \int_0^T \varepsilon(t) t \sin\varphi_0 t\, dt \right)' \tag{14}$$

is asymptotically, as $T \to \infty$, normal $N(0, K_2)$ with

$$K_2 = 2\pi \sum_{j=m}^{\infty} \frac{C_j^2}{j!} f^{*j}(\varphi_0) \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{6} \end{pmatrix}.$$

Similarly, one can obtain asymptotic normality of the vector

$$\left(T^{-1/2} \int_0^T \varepsilon(t) \cos\varphi_0 t\, dt, T^{-3/2} \int_0^T \varepsilon(t) t \cos\varphi_0 t\, dt \right)' \tag{15}$$

with the same covariance matrix K_2.

Lemma 3. *If the conditions of Lemma 1 and **B1** are satisfied, then* $T^{-1/2} Q_T'(\varphi_0)$ *is asymptotically, as* $T \to \infty$, *normal* $N(0, K_3)$ *with*

$$K_3 = \frac{4}{3} \pi A_0^2 \sum_{j=m}^{\infty} \frac{C_j^2}{j!} f^{*j}(\varphi_0).$$

Proof. Obviously

$$Q_T(\varphi) = \frac{4}{T^2} \left[\int_0^T X(t) \cos\varphi t\, dt \right]^2 + \frac{4}{T^2} \left[\int_0^T X(t) \sin\varphi t\, dt \right]^2.$$

Then

$$
\begin{aligned}
T^{-1/2}Q'_T(\varphi_0) &= -\frac{8}{T^{5/2}}\int_0^T X(t)\cos\varphi_0 t\,dt\int_0^T X(t)t\sin\varphi_0 t\,dt \\
&\quad+\frac{8}{T^{5/2}}\int_0^T X(t)\sin\varphi_0 t\,dt\int_0^T X(t)t\cos\varphi_0 t\,dt \\
&= -\frac{8}{T^{5/2}}\left[\frac{1}{2}\int_0^T A_0\cos^2\varphi_0 t\,dt\int_0^T A_0 t\sin 2\varphi_0 t\,dt\right. \\
&\qquad+\int_0^T A_0\cos^2\varphi_0 t\,dt\int_0^T \varepsilon(t)t\sin\varphi_0 t\,dt \\
&\qquad+\frac{1}{2}\int_0^T A_0 t\sin 2\varphi_0 t\,dt\int_0^T \varepsilon(t)\cos\varphi_0 t\,dt \\
&\qquad\left.+\int_0^T \varepsilon(t)\cos\varphi_0 t\,dt\int_0^T \varepsilon(t)t\sin\varphi_0 t\,dt\right] \\
&\quad+\frac{8}{T^{5/2}}\left[\frac{1}{2}\int_0^T A_0\sin 2\varphi_0 t\,dt\int_0^T A_0 t\cos^2\varphi_0 t\,dt\right. \\
&\qquad+\frac{1}{2}\int_0^T A_0\sin 2\varphi_0 t\,dt\int_0^T \varepsilon(t)t\cos\varphi_0 t\,dt \\
&\qquad+\int_0^T A_0 t\cos^2\varphi_0 t\,dt\int_0^T \varepsilon(t)\sin\varphi_0 t\,dt \\
&\qquad\left.+\int_0^T \varepsilon(t)\sin\varphi_0 t\,dt\int_0^T \varepsilon(t)t\cos\varphi_0 t\,dt\right] \\
&= \sum_{i=1}^{8} S_i.
\end{aligned}
$$

Evidently $S_1, S_5 \to 0$, $T\to\infty$. Using lemma 1, we have $S_3 \xrightarrow{\mathsf{P}} 0$, $T\to\infty$. Convergence to 0 in probability of summands S_4, S_6 and S_8 arises from asymptotic normality of integrals

$$T^{-3/2}\int_0^T \varepsilon(t)t\sin\varphi_0 t\,dt \quad\text{and}\quad T^{-3/2}\int_0^T \varepsilon(t)t\cos\varphi_0 t\,dt.$$

So,

$$T^{-1/2}Q'_T(\varphi_0) = S_2+S_7+\eta_T^{(2)}, \qquad \eta_T^{(2)} \xrightarrow{\mathsf{P}} 0, \qquad T\to\infty,$$

and

$$
\begin{aligned}
T^{-\frac{1}{2}}Q'_T(\varphi_0) &= 2A_0T^{-\frac{1}{2}}\int_0^T \varepsilon(t)\sin\varphi_0 t\,dt - 4A_0T^{-\frac{3}{2}}\int_0^T \varepsilon(t)t\sin\varphi_0 t\,dt+\eta_T^{(3)} \\
&= 2A_0(b_{1T}-2b_{2T})+\eta_T^{(3)}, \qquad \eta_T^{(3)} \xrightarrow{\mathsf{P}} 0, \qquad T\to\infty.
\end{aligned}
\tag{16}
$$

Using asymptotic normality of vector (14), find the variance $4A_0^2 D\,(b_{1T}-2b_{2T})$. Let $\lambda=(\lambda_1,\lambda_2)$, $b_T=(b_{1T},b_{2T})$. It is easily seen that

$$\mathsf{E}\,e^{i\langle\lambda,b_T\rangle} \to \exp\left\{-\frac{1}{2}\langle K_2\lambda,\lambda\rangle\right\}, \qquad T\to\infty.$$

Taking $\lambda=(\tau,-2\tau)$ we obtain

$$\mathsf{E}\,e^{i\tau(b_{1T}-2b_{2T})} \to \exp\left\{-\frac{\tau^2}{2}\left(K_2^{11}-4K_2^{12}+4K_2^{22}\right)\right\},$$

that is r.v. $b_{1T}-2b_{2T}$ is asymptotically normal $N\left(0,\frac{1}{6}\right)$. So, $T^{-1/2}Q'_T\left(\varphi_0\right)$ is asymptotically, as $T\to\infty$, normal $N(0,K_3)$. □

Lemma 4. *If conditions of Lemma 1 and* ***B1*** *are fulfilled, then for any r.v.* $\widetilde{\varphi}_T$ *satisfying inequality* $|\widetilde{\varphi}_T-\varphi_0|\le|\varphi_T-\varphi_0|$ *with probability 1, for all* $T>0$,

$$\frac{1}{T^2}Q''_T\left(\widetilde{\varphi}_T\right)\xrightarrow{\mathsf{P}}-\frac{1}{6}A_0^2,\qquad T\to\infty.$$

Proof. Write

$$\begin{aligned}\frac{1}{T^2}Q''_T\left(\widetilde{\varphi}_T\right)=&\,8\left[\int_0^T t\sin\widetilde{\varphi}_T tX(t)\,dt\right]^2-\frac{8}{T}\int_0^T\cos\widetilde{\varphi}_T tX(t)\,dt\frac{1}{T^3}\int_0^T t^2\cos\widetilde{\varphi}_T tX(t)\,dt\\&+8\left[\frac{1}{T^2}\int_0^T t\cos\widetilde{\varphi}_T tX(t)\,dt\right]^2\\&-\frac{8}{T}\int_0^T\sin\widetilde{\varphi}_T tX(t)\,dt\frac{1}{T^3}\int_0^T t^2\sin\widetilde{\varphi}_T tX(t)\,dt\\=&\sum_{i=1}^4 Q_i.\end{aligned}$$

Then the integral

$$\begin{aligned}\frac{1}{T}\int_0^T\cos\widetilde{\varphi}_T tX(t)\,dt&=\frac{1}{T}\int_0^T\cos\widetilde{\varphi}_T t\left[A_0\cos\varphi_0 t+\varepsilon(t)\right]\,dt\\&=\frac{A_0}{T}\int_0^T\cos\widetilde{\varphi}_T t\cos\varphi_0 t\,dt+\frac{1}{T}\int_0^T\cos\widetilde{\varphi}_T t\varepsilon(t)\,dt\\&=\frac{1}{2T}\int_0^T A_0\left(\cos\left(\widetilde{\varphi}_T-\varphi_0\right)t+\cos\left(\widetilde{\varphi}_T+\varphi_0\right)t\right)\,dt+\eta_T^{(4)}\\&=\frac{A_0}{2T}\int_0^T\cos\left(\widetilde{\varphi}_T-\varphi_0\right)t\,dt+\eta_T^{(5)},\\&\quad\eta_T^{(4)},\eta_T^{(5)}\xrightarrow{\mathsf{P}}0,\qquad T\to\infty.\end{aligned}$$

Using the lemma conditions and the result of the theorem 2 we obtain

$$\frac{1}{T}\int_0^T\cos\widetilde{\varphi}_T tX(t)\,dt\xrightarrow{\mathsf{P}}\frac{A_0}{2},\qquad T\to\infty.$$

Using similar calculations we get

$$\frac{1}{T^3}\int_0^T t^2\cos\widetilde{\varphi}_T tX(t)\,dt\xrightarrow{\mathsf{P}}\frac{A_0}{6},\qquad T\to\infty,$$

so, $Q_2\xrightarrow{\mathsf{P}}-\frac{2}{3}A_0^2$. Similarly, $Q_3\xrightarrow{\mathsf{P}}A_0^2/2$, and $Q_1,Q_4\xrightarrow{\mathsf{P}}0$, $T\to\infty$. Then

$$\frac{1}{T^2}Q''_T\left(\widetilde{\varphi}_T\right)\xrightarrow{\mathsf{P}}-\frac{2A_0^2}{3}+\frac{A_0^2}{2}=-\frac{A_0^2}{6},\qquad T\to\infty.$$ □

Theorem 4. *If the conditions of Lemma 1 and* ***B1*** *are satisfied, then* $T^{3/2}\left(\varphi_T-\varphi_0\right)$ *is asymptotically, as* $T\to\infty$, *normal with zero mean and variance*

$$48\pi\sum_{j=m}^{\infty}\frac{c_j^2}{j!}f^{*j}\left(\varphi_0\right).$$

Proof. As $Q'_T(\varphi_T) = 0$, then

$$Q'_T(\varphi_0) + Q''_T(\widetilde{\varphi}_T)(\varphi_T - \varphi_0) = 0 \tag{17}$$

with some r.v. $\widetilde{\varphi}_T$, satisfying

$$|\widetilde{\varphi}_T - \varphi_0| \le |\varphi_T - \varphi_0|, \qquad T \to \infty.$$

From (17)

$$T^{3/2}(\varphi_T - \varphi_0) = -\frac{T^{-1/2}Q'_T(\varphi_0)}{T^{-2}Q''_T(\widetilde{\varphi}_T)}.$$

The theorem follows now from lemmas 3 and 4. □

Theorem 5. *If the conditions of Lemma 1 and* ***B1*** *are satisfied, then the normed estimator $T^{1/2}(A_T - A_0)$ is asymptotically, as $T \to \infty$, normal with zero mean and variance*

$$4\pi \sum_{j=m}^{\infty} \frac{c_j^2}{j!} f^{*j}(\varphi_0).$$

Proof. Write

$$T^{1/2}(A_T - A_0) = T^{1/2}\left[Q_T^{1/2}(\varphi_T) - A_0\right] = T^{1/2}\left[Q_T(\varphi_T) - A_0^2\right]\left[Q_T^{1/2}(\varphi_T) + A_0\right]^{-1}.$$

From Lemma 2

$$Q_T^{1/2}(\varphi_T) + A_0 \xrightarrow{\mathsf{P}} 2A_0, \qquad T \to \infty. \tag{18}$$

We have

$$T^{1/2}\left[Q_T(\varphi_T) - Q_T(\varphi_0)\right] = T^{1/2}Q'_T(\varphi_0)(\varphi_T - \varphi_0) + \frac{1}{2}T^{1/2}Q''_T(\widetilde{\varphi}_T)(\varphi_T - \varphi_0)^2$$

with some $\widetilde{\varphi}_T$ such that $|\widetilde{\varphi}_T - \varphi_0| \le |\varphi_T - \varphi_0|$. The value

$$T^{1/2}Q'_T(\varphi_0)(\varphi_T - \varphi_0) = T^{-1/2}Q'_T(\varphi_0)\,T(\varphi_T - \varphi_0) \tag{19}$$

tends to 0 in probability, according to theorem 2 and lemma 3. The expression

$$\frac{T^{1/2}}{2}Q''_T(\widetilde{\varphi}_T)(\varphi_T - \varphi_0)^2 = \frac{1}{2T^2}Q''_T(\widetilde{\varphi}_T)\,T^{3/2}(\varphi_T - \varphi_0)\,T(\varphi_T - \varphi_0)$$

tends to 0 in probability as it follows from lemma 4 and theorems 2 and 4. Using (18) and (19) we can conclude that asymptotic distribution of $T^{1/2}(A_T - A_0)$ is the same as asymptotic distribution of

$$\frac{T^{1/2}}{2A_0}\left[Q_T(\varphi_0) - A_0^2\right]. \tag{20}$$

From (3) and (4) it is seen that $Q_T(\varphi_0) - A_0^2$ behaves at infinity as

$$Z_T(\varphi_0) = \frac{4}{T^2}\left|\int_0^T \varepsilon(t)e^{i\varphi_0 t}\,dt\right|^2 + \frac{8A_0}{T^2}\,\mathrm{Re}\left\{\int_0^T \cos\varphi_0 t e^{i\varphi_0 t}\,dt \int_0^T \varepsilon(t)e^{-i\varphi_0 t}\,dt\right\}.$$

Consider

$$\frac{1}{T^{3/2}}\left|\int_0^T \varepsilon(t)e^{i\varphi_0 t}\,dt\right|^2 = \frac{1}{T^{1/2}}\int_0^T \varepsilon(t)e^{i\varphi_0 t}\,dt\frac{1}{T}\int_0^T \varepsilon(t)e^{-i\varphi_0 t}\,dt \xrightarrow{\mathsf{P}} 0, \qquad T \to \infty,$$

because the 1st integral is asymptotically normal and the 2nd tends to 0 in probability.

So, it remains to analyze the behavior of

$$
\begin{aligned}
&T^{-3/2}\operatorname{Re}\left\{\int_0^T \cos\varphi_0 t e^{i\varphi_0 t}\,dt \int_0^T \varepsilon(t)e^{-i\varphi_0 t}\,dt\right\}\\
&\quad= T^{-3/2}\int_0^T \cos^2\varphi_0 t\,dt \int_0^T \varepsilon(t)\cos\varphi_0 t\,dt\\
&\qquad+ T^{-3/2}\int_0^T \cos\varphi_0 t\sin\varphi_0 t\,dt \int_0^T \varepsilon(t)\sin\varphi_0 t\,dt\\
&\quad= \frac{1}{2T^{1/2}}\int_0^T \varepsilon(t)\cos\varphi_0 t\,dt + \eta_T^{(6)}, \qquad \eta_T^{(6)} \xrightarrow{\mathsf{P}} 0, \qquad T\to\infty.
\end{aligned}
$$

As it was shown earlier, $T^{-1/2}\int_0^T \varepsilon(t)\cos\varphi_0 t\,dt$ is asymptotically normal with parameters 0 and $\pi\sum_{j=m}^{\infty}\frac{c_j^2}{j!}f^{*j}(\varphi_0)$. Using this fact we obtain that $T^{1/2}\left(Q_T(\varphi_0)-A_0^2\right)$ is asymptotically normal with parameters 0 and $16\pi A_0^2\sum_{j=m}^{\infty}\frac{c_j^2}{j!}f^{*j}(\varphi_0)$, so $T^{1/2}(A_T-A_0)$ is asymptotically normal with zero mean and variance $4\pi\sum_{j=m}^{\infty}\frac{c_j^2}{j!}f^{*j}(\varphi_0)$. □

Theorem 6. *If conditions **A1**, **A2**, **B1**, and **A3** or **A4** are satisfied, then the random vector*

$$\left(T^{1/2}(A_T-A_0),\ T^{3/2}(\varphi_T-\varphi_0)\right)'$$

is asymptotically normal, as $T\to\infty$, with zero mean and covariance matrix

$$2\pi\sum_{j=m}^{\infty}\frac{C_j^2}{j!}f^{*j}(\varphi_0)\begin{pmatrix} 2 & 0\\ 0 & 24A_0^{-2}\end{pmatrix}.$$

Proof. In the proofs of lemma 3, theorems 4 and 5 it was shown that

$$
\begin{aligned}
T^{\frac{3}{2}}(\varphi_T-\varphi_0) &= 12A_0^{-1}T^{-\frac{1}{2}}\int_0^T \varepsilon(t)\sin\varphi_0 t\,dt\\
&\qquad - 24A_0^{-1}T^{-3/2}\int_0^T \varepsilon(t)t\sin\varphi_0 t\,dt + \eta_T^{(7)},
\end{aligned}
\tag{21}
$$

$$T^{1/2}(A_T-A_0) = 2T^{-1/2}\int_0^T \varepsilon(t)\cos\varphi_0 t dt + \eta_T^{(6)}, \qquad T\to\infty; \tag{22}$$

$\eta_T^{(6)}\xrightarrow{\mathsf{P}}0$, $\eta_T^{(7)}\xrightarrow{\mathsf{P}}0$, $T\to\infty$.

We have, for any u_1, u_2,

$$
\begin{aligned}
&u_1T^{1/2}(A_T-A_0) + u_2T^{3/2}(\varphi_T-\varphi_0)\\
&\quad= u_1 2T^{-1/2}\int_0^T \varepsilon(t)\cos\varphi_0 t\,dt + u_2 12A_0^{-1}T^{-1/2}\int_0^T \varepsilon(t)\sin\varphi_0 t\,dt\\
&\qquad - u_2 24A_0^{-1}T^{-3/2}\int_0^T \varepsilon(t)t\sin\varphi_0 t\,dt + \eta_T^{(8)}\\
&\quad= v_1\xi_{1T} + v_2\xi_{2T} + v_3\xi_{3T} + \eta_T^{(8)},
\end{aligned}
\tag{23}
$$

where $v_1 = u_1\frac{2}{\sqrt{2}}$, $v_2 = u_2\frac{12}{\sqrt{2}}A_0^{-1}$, $v_3 = -u_2\frac{24}{\sqrt{6}}A_0^{-1}$, $\xi_{1T} = \sqrt{2}T^{-1/2}\int_0^T \varepsilon(t)\cos\varphi_0 t\,dt$, $\xi_{2T} = \sqrt{2}T^{-1/2}\int_0^T \varepsilon(t)\sin\varphi_0 t\,dt$, $\xi_{3T} = \sqrt{6}T^{-3/2}\int_0^T \varepsilon(t)t\sin\varphi_0 t\,dt$, $\eta_T^{(8)}\xrightarrow{\mathsf{P}}0$, $T\to\infty$.

Note that the spectral measure $\mu(d\lambda)$ of the vector $(\cos\varphi_0 t, \sin\varphi_0 t, t\sin\varphi_0 t)$ is

$$\mu(d\lambda) = \begin{pmatrix} \alpha & -i\beta & -i\beta\\ -i\beta & \alpha & \frac{\sqrt{3}}{2}\alpha\\ -i\beta & \frac{\sqrt{3}}{2}\alpha & \alpha\end{pmatrix}, \tag{24}$$

where α is a measure concentrated at $\pm\varphi_0$, and $\alpha(\{\pm\varphi_0\}) = \frac{1}{2}$, β is a signed measure concentrated at $\pm\varphi_0$ and $\beta(\{\pm\varphi_0\}) = \pm\frac{1}{2}$.

Using result of the Theorem 3, we obtain

$$\mathsf{E}\exp\left\{i\left(v_1\xi_{1T} + v_2\xi_{2T} + v_3\xi_{3T}\right)\right\} \to \exp\left\{-\frac{1}{2}\left\langle Kv, v\right\rangle\right\},$$

where, from (13) and (24) it follows

$$\begin{aligned}\langle Kv, v\rangle &= 2\pi\sum_{k=m}^{\infty}\frac{C_k^2}{k!}f^{*k}(\varphi_0)\left(v_1^2 + v_2^2 + v_3^2 + \sqrt{3}v_2v_3\right)\\ &= 2\pi\sum_{k=m}^{\infty}\frac{C_k^2}{k!}f^{*k}(\varphi_0)\left(2u_1^2 + 24A_0^{-2}u_2^2\right). \qquad \square\end{aligned}$$

References

1. M. G. Serebrennikov and A. A. Pervozvanskyi, *The Detection of Hidden Periodicities*, "Nauka", Moscow, 1965. (Russian)
2. P. Whittle, *The simultaneous estimation of a time series harmonic components and covariance structure*, Trabajos Estadistica **3** (1952), 43–57.
3. A. M. Walker, *On the estimation of a harmonic component in a time series with stationary dependent residuals*, Advances in Appl. Probability **5** (1973), 217–241.
4. E. J. Hannan, *The estimation of frequency*, J. Appl. Probability **10** (1973), 510–519.
5. G. P. Hrechka and A. Ya. Dorogovtsev, *On asymptotical properties of periodogram estimator of harmonic oscillation frequency and amplitude*, Numerical and Applied Math. **28** (1976), 18–31.
6. A. V. Ivanov, *A solution of the problem of detecting hidden periodicities*, Theor. Probability and Math. Statist. **20** (1980), 51–68.
7. P. S. Knopov, *Optimal Estimators of Parameters of Stochastic Systems*, "Naukova Dumka", Kiev, 1981. (Russian)
8. B. G. Quinn and E. J. Hannan, *The Estimation and Tracking of Frequency*, Cambridge University Press, New York, 2001.
9. M. Artis, M. Hoffmann, D. Nachane, and J. Toro, *The Detection of Hidden Periodicities: a Comparison of Alternative Methods*, EUI Working Paper, No. ECO 2004/10, Badia Fiesolana, San Domenico (FI).
10. A. V. Ivanov and B. M. Zhurakovskyi, *The estimator consistency of least squares parameters of a sum of harmonic oscillations in the models with strongly dependent noise*, Naukovi visti **4** (2010), 60–66.
11. I. A. Ibragimov and Y. A. Rozanov, *Gaussian Random Processes*, Applications of Mathematics, vol. 9, Springer-Verlag, Berlin–Heidelberg–New York, 1978.
12. A. V. Ivanov and B. M. Zhurakovskyi, *Detection of hidden periodicities in the model with long range dependent noise*, International Conference Modern Stochastic: Theory and Applications II, Kiev, 2010, pp. 99–100.
13. A. V. Ivanov, *Consistency of the least squares estimator of the amplitudes and angular frequencies of a sum of harmonic oscillations in models with long-range dependence*, Theor. Probab. Math. Statist. **80** (2010), 61–69.

Department of Mathematical Analysis and Probability Theory, Faculty of Physics and Mathematics, National Technical University of Ukraine, "Kiev Politechnic Institute", Peremohy ave., 37, Kyiv 03056, Ukraine

E-mail address: alexntuu@gmail.com

Department of Mathematical Analysis and Probability Theory, Faculty of Physics and Mathematics, National Technical University of Ukraine, "Kiev Politechnic Institute", Peremohy ave., 37, Kyiv 03056, Ukraine

E-mail address: zhurak@gmail.com

Received 22/12/2012

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 40–50

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 40–50

QUANTITATIVE AND QUALITATIVE LIMITS FOR EXPONENTIAL ASYMPTOTICS OF HITTING TIMES FOR BIRTH-AND-DEATH CHAINS IN A SCHEME OF SERIES

UDC 519.21

N. V. KARTASHOV

Abstract. We consider time-homogeneous discrete birth-and-death Markov chain (X_t) and investigate the asymptotics of the hitting time $\tau_n = \inf(t \geq 1 \colon X_t \geq n)$ as well as the chain position before this time in the scheme of series as $n \to \infty$. In our case one-step probabilities of the chain vary simultaneously with n. The proofs are based on the explicit two-side inequalities with numerical bounds for the survival probability $\mathsf{P}(\tau_n > t)$. These inequalities can be used also for the pre-limit finite-time schemes. We have applied the results obtained for construction the uniform asymptotic representations of the corresponding risk function.

Анотація. Ми розглядаємо однорідний за часом дискретний ланцюг народження та загибелі (X_t) та вивчаємо асимптотику моменту досягнення $\tau_n = \inf(t \geq 1 \colon X_t \geq n)$ і стану ланцюга до цього моменту у схемі серій, де $n \to \infty$ та одночасно змінюються перехідні імовірності ланцюга за один крок. Доведення спираються на відповідні двобічні явні нерівності для ймовірності виживання $\mathsf{P}(\tau_n > t)$ з числовими границями. Останні можна використати і у дограничних схемах. Наведено застосування у вигляді асимптотичних розвинень для відповідної функції ризику.

Аннотация. Мы рассматриваем однородную дискретную цепь рождения и гибели (X_t) и изучаем асимптотику момента достижения $\tau_n = \inf(t \geq 1 \colon X_t \geq n)$ и положения цепи до этого момента, в схеме серий, где $n \to \infty$ и одновременно изменяются переходные вероятности цепи за один шаг. Доказательства основаны на соответствующих двусторонних неравенствах для вероятности выживания $\mathsf{P}(\tau_n > t)$ с явными числовыми ограничениями. Последние можно использовать и в допредельных схемах. Приведены применения в виде асимптотичных представлений для соответствующих функций риска.

1. Introduction

The task of investigation of the distribution stability for general Markov chains under the broad assumptions about the nature of jumps is expounded in details in the author's monograph [2]. Some applications of the theory are included there as well. The proofs are based on the analytical operator methods. The book includes some new inequalities for the renewal process asymptotics and the solutions of the renewal equation.

Foundations of the stability theory for stochastic models are set in the monograph by Zolotarev [5]. Important achievements in the stability theory are included in the book by Mayn and Tweedie [4].

This paper is based on the author's results placed in [2, Ch.7]. These results were obtained earlier but they have not been published. The comparison with paper [3] can be useful. The similar but not identical results were obtained earlier in [6].

2. Main results

Let us consider the time-homogeneous birth-and-death Markov chain

$$X = (X_t, t = 0, 1, \dots)$$

2000 *Mathematics Subject Classification.* Primary 60J45; Secondary 60A05, 60K05.

Key words and phrases. Birth and death Markov chains, hitting times, exponential asymptotics.

with values in a discrete space $E = \mathbb{Z}_+$. A matrix of one-step transition probabilities $P = (p_{ij}, i, j \in E)$ has entries $p_{i,i-1} = q_i$, $p_{i,i+1} = p_i$, $p_{ii} = r_i = 1 - p_i - q_i$ when $i \geq 1$, and $p_{01} = p_0$, $p_{00} = 1 - p_0 = q_0$. We assume that the chain is not reducible: p_i, $q_i > 0$. The symbols $\mathsf{P}_i(\cdot)$ and $\mathsf{E}_i(\cdot)$ will be used to denote the conditional probability and expectation given $\{X_0 = i\}$.

Let us define the hitting moment of the "distant" level as

$$\tau_n = \inf(t \geq 1\colon X_t \geq n). \tag{1}$$

We investigate the asymptotics of the time τ_n in a scheme of series where $n \to \infty$. In our case the one-step transition probabilities $(p_i,\, q_i)$ could change. For instance, they could depend on n.

Let's introduce the following notation for $t \geq 0$

$$\theta_t = \prod_{1 \leq i \leq t} (q_i/p_i), \qquad \theta_0 = 1, \qquad \sigma_t = \sum_{0 \leq i < t} \theta_i, \qquad \varkappa_t = 1/(p_t\theta_t), \qquad t \geq 0. \tag{2}$$

Consider the aggregate parameters

$$\begin{aligned} \lambda_n &= \left(1 + \sum_{i \leq j \in E_n} \varkappa_i\theta_j\right)^{-1} = \left(1 + \sum_{i \in E_n} \varkappa_i(\sigma_n - \sigma_i)\right)^{-1}, \\ \omega_n &= \lambda_n - \lambda_n^2 + \lambda_n^2 \sum_{i \leq j < k \leq l \in E_n} \varkappa_i\theta_j\varkappa_k\theta_l \\ &= \lambda_n - \lambda_n^2 + \lambda_n^2 \sum_{i < k \in E_n} \varkappa_i(\sigma_k - \sigma_i)\varkappa_k(\sigma_n - \sigma_k). \end{aligned} \tag{3}$$

Hereafter we will use the summation sign without upper and lower indexes assuming summation on the hole index set $E_n \equiv \{0, 1, \ldots, n-1\}$. It worth to mention that the process continuity implies the entire determination for the distribution of the time τ_n by $(p_i, q_i, i \in E_n)$.

The following estimation can be applicable to any scheme of series as well as for the fixed n.

Theorem 2.1. *The following inequality holds true*

$$\sup_{t \geq 0} \left|\mathsf{P}_0(\tau_n > t) - \left(1 - m_n^{-1}\right)^t\right| \leq 2\omega_n(1 + \lambda_n)p_0/\lambda_n\sigma_n(1 - \omega_n), \tag{4}$$

where

$$m_n^{-1} = \lambda_n/(1 + \omega_n).$$

Remark 2.1. It follows from the definitions (3) that $0 < \omega_n \leq 1/2$ in (4).

Corollary 2.1. *Let $n \to \infty$ in a scheme of series in such a way that $\lambda_n \to 0$ and*

$$\omega_n p_0 = o(\lambda_n\sigma_n), \qquad n \to \infty. \tag{5}$$

Then

$$\sup_{x \geq 0} |\mathsf{P}_0(\tau_n/m_n > x) - \exp(-x)| \to 0, \qquad n \to \infty.$$

This convergence is uniform in the scheme of series if the relation (5) *is uniform too.*

Corollary 2.2. *Let the chain X be unchangeable for the scheme of series, irreducible and ergodic, and $n \to \infty$. Then $\lambda_n \to 0$, $\omega_n \to 0$, and the following representation is true*

$$\sup_{x \geq 0} |\mathsf{P}_0(\lambda_n\tau_n > x) - \exp(-x)| = O(\omega_n), \qquad n \to \infty. \tag{6}$$

Corollary 2.3. *Let n and the distribution of the chain X be fixed excepting $p_0 \to 0$. Then the following representation holds true*

$$\sup_{x\geq 0} |\mathsf{P}_0(p_0\tau_n > x) - \exp(-x)| = O(p_0), \qquad p_0 \to 0. \tag{7}$$

We can obtain from (4) the limit results for the specially structured schemes of series at one time. Here are some examples.

Corollary 2.4. *Let the transition probabilities in a scheme of series for the birth-and-death chain satisfy the relationship*

$$p_i = \varepsilon_n v_i + o(\varepsilon_n), \qquad q_i = \varepsilon_n u_i + o(\varepsilon_n), \qquad i \geq 1,\ n \to \infty, \tag{8}$$

for some $\varepsilon_n \to 0$, and $v_i, u_i > 0$. Let us use the denotations

$$\theta_t = \prod_{i=1}^{t} (u_i/v_i), \qquad \sigma_t = \sum_{s=1}^{t-1} \theta_s, \qquad \chi_t = 1/(v_t\theta_t), \qquad t \geq 1. \tag{9}$$

We assume that in a scheme of series

$$\sigma_n \to \infty, \qquad \sum_{t\geq 1} \chi_t \equiv \chi = O(1), \qquad n \to \infty. \tag{10}$$

Then, subject to

$$\overline{\omega}_n \equiv \sigma_n^{-1} \sum_{1\leq i<k<n} \chi_i(\sigma_k - \sigma_i)\chi_k = o(1), \qquad n \to \infty,$$

the uniform convergence is true

$$\sup_{t\geq 0} \left| \mathsf{P}_0(\tau_n > t) - \left(1 - m_n^{-1}\right)^t \right| = O\left(\overline{\omega}_n + \sigma_n^{-1}\right) = o(1), \qquad n \to \infty.$$

Remark 2.2. If the coefficients v_i, u_i are bounded and separated from zero, then the conditions (10) are equivalent to the ergodicity of the the birth-and-death chain with jump probabilities $(u_i/(u_i + v_i), v_i/(u_i + v_i))$, $i \geq 1$.

To analyze the asymptotics of joint distribution of the time τ_n and the chain value X till this time (the comparison can be made with [6]) we additionally assume that there is a systematic shift to zero

$$q_i > p_i, \qquad i \geq 1, \tag{11}$$

and state 0 in a scheme of series is asymptotically positive and attainable:

$$\varliminf_{n\to\infty} \lambda_n\sigma_n > 0, \qquad 0 < \varliminf_{n\to\infty} p_0 \leq \varlimsup_{n\to\infty} p_0 < 1. \tag{12}$$

It was established in the proof of the Corollary 2.2 (see the limit relation (38)) that the conditions (12) hold for any fixed irreducible ergodic birth-and-death chain.

We define the speed-of-mixing indicator as

$$\delta_n = \min_{1\leq i<n} (q_i - p_i) > 0. \tag{13}$$

Theorem 2.2. *Let the conditions* (11), (12) *and $\lambda_n \to 0$, $\omega_n \to 0$ hold true in a scheme of series as $n \to \infty$ so that*

$$\lambda_n \ln \lambda_n^{-1} = o(\delta_n^4), \qquad n \to \infty. \tag{14}$$

Then, for every $s_0 > 0$ the uniform representation holds true

$$\begin{aligned} &\sup_{s\geq s_0, B\subset E} \left| \mathsf{P}_0(\lambda_n\tau_n > s,\ X_{s/\lambda_n} \in B) - \pi^n(B)\exp(-s) \right| \\ &\quad = O(\omega_n + \lambda_n\delta_n^{-4}\ln(1/\lambda_n\delta_n)) = o(1), \qquad n \to \infty, \end{aligned} \tag{15}$$

where the discrete distribution $\pi^n = (\pi_i^n, i \in E^n)$ *can be defined through* (2), (3) *by equalities*

$$\pi_n^n = \lambda_n, \qquad \pi_i^n = \lambda_n \varkappa_i(\sigma_n - \sigma_i), \qquad i < n, \qquad \pi^n(B) = \sum_{i \in B} \pi_i^n. \tag{16}$$

Corollary 2.5. *Let the chain* X *do not change in a scheme of series and be irreducible and ergodic. Then, the sufficient condition for the convergence to zero of the left-hand part of* (15) *is*

$$\varliminf_{n\to\infty} n\delta_n / \ln n > 3/2. \tag{17}$$

Remark 2.3. For comparison with (17) we remark that

(a) in a class of birth-and-death chains satisfying the conditions (13) and $\delta_n \to 0$, $n \to \infty$ the sufficient condition of ergodicity is

$$\varliminf_{n\to\infty} n\delta_n / \ln n > 1,$$

(b) under additional assumptions $r_i \equiv 0$ and $q_i - p_i \downarrow 0$, $i \to \infty$, it follows from the condition

$$\varliminf_{n\to\infty} n\delta_n < 1/2$$

that the chain is not ergodic.

3. Proofs

The proofs in this section are based on the Corollary 7.5 [2, Ch.VII].

In order to use it we consider the auxiliary finite chain $X^n = (X_t^n, t \ge 0)$ with the set of states $E^n \equiv \{0, 1, \ldots, n\} = E_n \cup \{n\}$ and the transition probabilities

$$P_n = (p_{ij}(n), i, j \in E^n),$$

where $p_{ij}(n) = p_{ij}$ as $i \in E_n$ and

$$p_{n0}(n) = 1.$$

It is evident that the distributions for the time τ_n in (1) for chains (X_t) and (X_t^n) are equal.

Lemma 3.1. *A chain* X^n *has the unique invariant probability* $\pi^n = (\pi_i^n, i \in E^n)$ *where*

$$\pi_n^n = \lambda_n, \qquad \pi_j^n = \lambda_n \varkappa_j(\sigma_n - \sigma_j) = \lambda_n \sum_{i<j\in E_n} \theta_i \varkappa_j, \qquad j < n. \tag{18}$$

Proof. The system of equations for $x_i \equiv \pi_i^n$ has a form

$$\begin{gathered} x_0 q_0 + x_1 q_1 + x_n = x_0, \\ x_{i-1}p_{i-1} + x_i r_i + x_{i+1} q_{i+1} = x_i, \qquad 1 \le i < n-1, \\ x_{n-2}p_{n-2} + x_{n-1} r_{n-1} = x_{n-1}, \\ x_{n-1} p_{n-1} = x_n. \end{gathered} \tag{19}$$

We obtain the following equations from the first and the second rows

$$x_{i-1}p_{i-1} - x_i q_i = x_i p_i - x_{i+1} q_{i+1} = x_n, \qquad 1 \le i < n-1. \tag{20}$$

And finally, using the third, the forth rows of (19) and from (20) we recurrently calculate when $0 \le k < n$

$$\begin{aligned} x_k &= x_n q_k^{-1} \theta_{k-1}^{-1} \left[\sum_{i=k}^{n-3} \theta_i + \theta_{n-3} q_{n-2} (p_{n-1} + q_{n-1}) / p_{n-2} p_{n-1} \right] \\ &= x_n q_k^{-1} \theta_{k-1}^{-1} \left[\sum_{i=k}^{n-3} \theta_i + \theta_{n-2} (1 + q_{n-1}/p_{n-1}) \right] = x_n \sum_{i=k}^{n-1} \varkappa_k \theta_i. \end{aligned} \tag{21}$$

The condition of normalization $\sum_{k=0}^{n} x_k = 1$ implies (18). $\square$

In order to prove the Theorem 2.1 we shall use the Corollary 7.5 [2, Ch.VII] for the chain X^n with a set E^n and invariant probability π^n. Let us mention that every transition kernel $Q(x, A)$ and the corresponding linear operator in the descrete space can be defined by the matrix $Q(x, A) = \sum_{y \in A} Q_{xy}$, $Q_{xy} = Q(x, \{y\})$. Operation of multiplication kernels by measures, functions and kernels corresponds to the multiplication of the matrices by rows, columns and matrices.

In particular, the system for the kernel $\overline{R} = (\overline{R}_{xy}, x, y \in E^n)$ in the formulation of the Corollary 7.5 [2, Ch.VII] is as following

$$\begin{gathered}\overline{R}_{xy} = \sum_{k \in E^n} P_{xk}(n)\overline{R}_{ky} + P_{xy}(n) - \pi_y^n,\\ \sum_{k \in E^n} \pi_k^n \overline{R}_{ky} = 0, \qquad x, y \in E^n,\end{gathered} \tag{22}$$

where the last equation arises from the Lemma 3.1 since π^n is the eigenvector for the matrix P_n.

Moreover, it follows from the defining $\overline{R}$ as a sum of series of powers of P_n (Corollary 7.5 [2, Ch.VII]) that the operators $\overline{R}$ and P_n commutate so the equations (22) are equivalent to the system

$$\begin{gathered}\overline{R}_{xy} = \sum_{k \in E^n} \overline{R}_{xk} P_{ky}(n) + P_{xy}(n) - \pi_y^n,\\ \sum_{k \in E^n} \overline{R}_{xk} = 0, \qquad x, y \in E^n.\end{gathered} \tag{23}$$

Lemma 3.2. *The solutions of systems* (22), (23) *for* $x = n$ *or* $y = n$ *are as following*

$$\overline{R}_{nn} = -\omega_n, \tag{24}$$

$$\overline{R}_{kn} = \lambda_n \sum_{i \le j < k} \varkappa_i \theta_j + \lambda_n - \omega_n, \qquad k < n, \tag{25}$$

$$\overline{R}_{nk} = \varkappa_k(\sigma_n - \sigma_k)(\lambda_n - \omega_n) + \lambda_n \sum_{k < i < n} \varkappa_k(\sigma_n - \sigma_i)\varkappa_i(\sigma_i - \sigma_k), \qquad k < n. \tag{26}$$

Proof. Denote $x_k = \overline{R}_{kn}$. Taking into account (18) we put $y = n$ into (22) and obtain the system

$$\begin{gathered}x_0 = q_0 x_0 + p_0 x_1 - \lambda_n,\\ x_i = q_i x_{i-1} + r_i x_i + p_i x_{i+1} - \lambda_n, \qquad 0 < i < n-1,\\ x_{n-1} = q_{n-1} x_{n-2} + r_{n-1} x_{n-1} + p_{n-1} x_n + p_{n-1} - \lambda_n,\\ x_n = x_0 - \lambda_n.\end{gathered} \tag{27}$$

The following equalities are deduced from the first and the second rows

$$\begin{gathered}p_0(x_1 - x_0) = \lambda_n,\\ (x_{i+1} - x_i)/\theta_i = \lambda_n \varkappa_i + (x_i - x_{i-1})/\theta_{i-1}, \qquad 1 \le i < n-1.\end{gathered}$$

By recurrent calculation we get

$$x_{k+1} - x_k = \lambda_n \sum_{i=0}^{k} \varkappa_i \theta_k, \qquad 0 \le k < n-1,$$

$$x_k = x_0 + \lambda_n \sum_{j<k} \sum_{i=0}^{j} \varkappa_i \theta_j, \qquad 0 \le k < n. \tag{28}$$

Putting the equalities (28) and (27) into the second equation (22) we obtain that

$$\begin{aligned}
0 &= \sum_{k\in E^n} \pi_k^n x_k = x_0 - \lambda_n^2 + \lambda_n \sum_{k\in E_n} \sum_{j<k} \sum_{i=0}^{j} \varkappa_i \theta_j \\
&= x_0 - \lambda_n^2 + \lambda_n^2 \sum_{i\le j<k\in E_n} \varkappa_i \theta_j \varkappa_k (\sigma_n - \sigma_k) = x_0 - \lambda_n^2 + \omega_n - \lambda_n + \lambda_n^2 \\
&= x_0 + \omega_n - \lambda_n,
\end{aligned}$$

and deduce from (27) the identities in (24), (25).

For proving (26) we use the denote $x_k \equiv \overline{R}_{nk}$ and use (23) when $x = n$

$$\begin{aligned}
x_0 &= q_0 x_0 + q_1 x_1 + x_n + 1 - \pi_0^n, \\
x_k &= p_{k-1} x_{k-1} + r_k x_k + q_{k+1} x_{k+1} - \pi_k^n, \qquad 1 \le k < n-1, \\
x_{n-1} &= p_{n-2} x_{n-2} + r_{n-1} x_{n-1} - \pi_{n-1}^n, \\
x_n &= p_{n-1} x_{n-1} - \pi_n^n,
\end{aligned} \tag{29}$$

where the probabilities π_k^n are defined in the Lemma 3.1.

Using the first two equations (29) and the recurrent calculations we deduce that

$$p_k x_k - q_{k+1} x_{k+1} = -\sum_{i=0}^{k} \pi_i^n + 1 + x_n = \sum_{i=k+1}^{n} \pi_i^n + x_n, \qquad 0 \le k < n-2. \tag{30}$$

Multiplying (30) by θ_k and summing over $k = 0, \ldots, n-3$ we obtain

$$x_k \theta_k p_k = x_{n-2} \theta_{n-2} p_{n-2} + \sum_{j=k}^{n-3} \theta_j \left(x_n + \sum_{i=j+1}^{n} \pi_i^n \right).$$

Taking into account the last two equations in (29) and the identity

$$\theta_{n-2}(p_{n-1} + q_{n-1})/p_{n-1} = \theta_{n-2} + \theta_{n-1}$$

we deduce that

$$x_k = x_n(\sigma_n - \sigma_k) + \sum_{j=k}^{n-1} \varkappa_k \theta_j \sum_{i=j+1}^{n} \pi_i^n, \qquad 0 \le k < n. \tag{31}$$

And finally, putting there the values π_i^n from (16) and $x_n = \overline{R}_{nn}$ from (24) concludes the proof of the Lemma 3.2. □

Proof of Theorem 2.1. Let us utilize the inequality (7.40) from the Corollary 7.5 [2, Ch.VII] to the chain $X = X^n$ on $E = E^n$ with time $\tau_H = \tau_n$ and the set $H = \{n\}$. An invariant measure and the chain potential are calculated in the Lemmas 3.1 and 3.2.

In the notations of (7.39) [2, Ch.VII]

$$r_{HH} = \sup_{x\in H} \left|\overline{R}\right|(x, H) = \left|\overline{R}_{nn}\right| = \omega_n, \tag{32}$$

under (24) since $\omega_n > 0$. It follows from (25), (18), (3)

$$\begin{aligned} r_{\pi H} &= \int \pi(dx) \left|\overline{R}\right|(x,H) = \sum_{k=0}^{n} \pi_k^n \left|\overline{R}_{kn}\right| \\ &\le \lambda_n \left|\overline{R}_{nn}\right| + \sum_{k=0}^{n-1} \lambda_n \varkappa_k (\sigma_n - \sigma_k) \max \left(\omega_n, \lambda_n + \lambda_n \sum_{i \le j < k} \varkappa_i \theta_j \right) \\ &= \lambda_n \omega_n + \max \left[\lambda_n \omega_n \sum_{k=0}^{n-1} \varkappa_k (\sigma_n - \sigma_k), \right. \\ &\qquad \left. \lambda_n^2 \sum_{k=0}^{n-1} \varkappa_k (\sigma_n - \sigma_k) \left(1 + \sum_{i \le j < k} \varkappa_i \theta_j \right) \right] \\ &\le \lambda_n \omega_n + \max \left[\lambda_n \omega_n (\lambda_n^{-1} - 1), \lambda_n^2 (\lambda_n^{-1} - 1) \right] + \omega_n - \lambda_n + \lambda_n^2 \\ &= \lambda_n \omega_n + \max[\omega_n (1 - \lambda_n), \omega_n] = \omega_n (1 + \lambda_n). \end{aligned} \tag{33}$$

Furthermore, according to the equality (7.41) [2, Ch. VII]

$$m_H^{-1} = (E_\pi \tau_H)^{-1} = \pi_n^n \sum_{t \ge 0} (-1)^t (\overline{R}_{nn})^t = \lambda_n (1 + \omega_n)^{-1} = m_n^{-1}. \tag{34}$$

And finally, the constant a in the Corollary 7.5 [2, Ch.VII] is the upper limit for the density of the initial distribution of α (it is concentrated in 0) regarding the measure π^n

$$a = 1/\pi_0^n = 1/\lambda_n \varkappa_0 \sigma_n = p_0/\lambda_n \sigma_n. \tag{35}$$

Putting the relations (32), (33), (34) and (35) into the inequality (7.40) of the Corollary 7.5 [2, Ch.VII] we have proved the estimate (4) in the Theorem 1. □

Proof of Remark 1. The positiveness of $\omega_n > 0$ follows from condition $\lambda_n < 1$ in definition (3). Let us denote as

$$s_n = \sum_{i \le j \in E_n} \varkappa_i \theta_j > 0$$

the sum included in (3). Using the last definition

$$\lambda_n = (1 + s_n)^{-1},$$

$$\begin{aligned} \omega_n &= \left(s_n + \sum_{i \le j < k \le l \in E_n} \varkappa_i \theta_j \varkappa_k \theta_l \right) / (1 + s_n)^2 \\ &\le \left(s_n + 1/2 \sum_{i \le j \in E_n} \varkappa_i \theta_j \sum_{k \le l \in E_n} \varkappa_k \theta_l \right) / (1 + s_n)^2 \\ &= (s_n + s_n^2/2)/(1 + s_n)^2 \le 1/2 < 1. \end{aligned} \tag{36}$$

□

Proof of Corollary 2.1. The proof can be concluded from the inequality (4) since the right-hand part of (4) equals to $O(p_0 \omega_n / \lambda_n \sigma_n)$ in view of (36). From the other side, the relation in the left-hand part after the substitution $t = [x m_n]$ is equivalent to

$$(1 - m_n^{-1})^{[x m_n]} \to \exp(-x), \qquad n \to \infty, \tag{37}$$

uniformly on $x \ge 0$ since $m_n^{-1} \le \lambda_n \to 0$. □

Proof of Corollary 2.2. The well-known recurrence and positivity criteria for the birth-and-death chain [1] correspond to the divergence of $\sigma_n \to \infty$ and convergence of $\varkappa = \sum_{i\geq 0} \varkappa_i < \infty$.

Let us calculate

$$(\lambda_n \sigma_n)^{-1} = \sigma_n^{-1}\left(1 + \sum_{i \leq j \in E_n} \varkappa_i \theta_j\right) = \sigma_n^{-1} + \sum_{i\geq 0} \varkappa_i \left(1 - \sigma_i \sigma_n^{-1}\right) \mathbb{1}_{i<n}$$
$$\to \sum_{i\geq 0} \varkappa_i = \varkappa \in (0,\infty), \qquad n \to \infty, \tag{38}$$

using the Lebesgue theorem on majorized convergence.

So, $\lambda_n \sim 1/\varkappa\sigma_n \to 0$, $n \to \infty$.

Similarly, it follows from the representation

$$\omega_n = \lambda_n - \lambda_n^2 + \lambda_n^2 \sum_{0\leq i<k<n} \varkappa_i(\sigma_k - \sigma_i)\varkappa_k(\sigma_n - \sigma_k)$$
$$\leq \lambda_n + (\lambda_n\sigma_n)^2 \sum_{i,k\in E_n} \varkappa_i\varkappa_k(\sigma_k - \sigma_i)^+ \sigma_n^{-1} \mathbb{1}_{i<k<n},$$

and the monotonicity of σ_n, using the Lebesgue theorem on majorized convergence, that $\overline{\lim}_{n\to\infty} \omega_n = 0$.

Taking into account (38) and the Remark 1 we can conclude that the right-hand part of (4) $2\omega_n(1+\lambda_n)p_0/\lambda_n\sigma_n(1-\omega_n)$ is equal to $O(\omega_n)$.

Utilization of approximation (37) in its left-hand part, convergence of $\omega_n \to 0$ and the estimate $|\exp(-x - x\varepsilon) - \exp(-x)| \leq \varepsilon$, $x, \varepsilon \geq 0$ result in (6) $\square$

Proof of Corollary 2.3. Let us use the representations (4) of the Theorem 2.1, where n is fixed. Since p_0 is included into (3) only as a part of $\varkappa_0$, then $\lambda_n = 1/(1 + L/p_0) \sim p_0/L$, $p_0 \to 0$, $\omega_n = p_0/L + o(p_0)$, $p_0 \to 0$, $\sigma_n^2 = C$ for some constants $L, C > 0$. Thus, (7) follows from (4). $\square$

Proof of Corollary 2.4. It follows from the definitions (3), (8), (9) that

$$\lambda_n^{-1} = 1 + \sigma_n/p_0 + \varepsilon_n^{-1} \sum_{1\leq i<n} \chi_i(\sigma_n - \sigma_i) \sim \sigma_n\chi/\varepsilon_n \to \infty, \qquad n \to \infty.$$

Simultaneously,

$$\omega_n = \lambda_n - \lambda_n^2 + \lambda_n^2 \sum_{1\leq i<k<n} \varepsilon_n^{-2}\chi_i(\sigma_k - \sigma_i)\chi_k(\sigma_n - \sigma_k) \sim \lambda_n + \lambda_n^2\sigma_n^2\overline{\omega}_n/\varepsilon_n^2$$
$$\sim \lambda_n + \chi^{-2}\overline{\omega}_n = o(1), \qquad n \to \infty. \qquad \square$$

Proof of Theorem 2.2. Let us use the inequality (7.43) of the Corollary 7.5 [2, Ch.VII] to the chain $X = X^n$ on $E = E^n$ with time $\tau_H = \tau_n$ and the set $H = \{n\}$. An invariant probability and the potential of the chain X^n are calculated in the Lemmas 3.1 and 3.2.

The estimate for new variation of the potential in (7.42) can be deduced from the equalities (24), (26) since

$$r_H = 1 + \sup_{x\in H} \left|\overline{R}\right|(x,E) = 1 + \sum_{k=0}^{n} \left|\overline{R}_{nk}\right|$$
$$\leq 1 + \left|\overline{R}_{nn}\right| + (\lambda_n + \omega_n)\sum_{k<n} \varkappa_k(\sigma_n - \sigma_k) + \lambda_n \sum_{k<i<n} \varkappa_k(\sigma_i - \sigma_k)\varkappa_i(\sigma_n - \sigma_k) \tag{39}$$
$$= 1 + \omega_n + (\lambda_n + \omega_n)\left(\lambda_n^{-1} - 1\right) + \left(\omega_n - \lambda_n + \lambda_n^2\right)\lambda_n^{-1} = 1 + 2\omega_n\lambda_n^{-1}.$$

The relations (3) are also used in the expressions above.

Since $\pi(H) = \pi_n^n = \lambda_n$ then utilizing (32) to the firs term in the right-hand part of (7.43) [2, Ch. VII] we obtain the inequality

$$\pi(H)2r_H/(1-r_{HH}) \leq \lambda_n 2\left(1+2\omega_n\lambda_n^{-1}\right)/(1-\omega_n) = O(\omega_n), \qquad n\to\infty, \tag{40}$$

where Remark 2.1 is taken into account as well. The representation $\lambda_n = O(\omega_n)$ is the evident conclusion from (3).

Thus, in order to apply (7.43) we need to find such constants $r_\alpha \in (0,1)$ and $b<\infty$ that

$$|\mathsf{P}_0(X_t^n \in B) - \pi^n(B)| \leq b(1-r_\alpha)^t \tag{41}$$

for all $t>0$, $B\subset E$.

Let us use the Theorem 3.6 [2, Ch.III].

We define the following norm on the space of measures $\mu = (\mu_i, i\in E^n)$

$$\|\mu\| = \sum_{i\in E^n} v^i\left|\mu_i\right|, \tag{42}$$

where the constant $v>1$ will be choose later. The form of the corresponding operator norm on $L(E^n)$ is placed in [2, p. 1.1]. Let us mention that since $v>1$:

$$\begin{aligned} |\mathsf{P}_0(X_t^n \in B) - \pi^n(B)| &\leq \left|\alpha P_n^t - \pi^n\right|(E^n) \leq \left\|\alpha P_n^t - \pi^n\right\| = \left\|\alpha(P_n^t - \Pi_n)\right\| \\ &\leq \|\alpha\|\left\|P_n^t - \Pi_n\right\| = \left\|P_n^t - \Pi_n\right\|, \end{aligned} \tag{43}$$

where $P_n^t \equiv (P_n)^t$ and $\alpha_i = \delta_{i0}$ is the initial distribution of the matrix Π_n that has equal rows of the type π^n.

Let us transform the matrix P_n as $P_n = T_n + h\circ\beta$, where the function

$$h = (\delta_{i0}, i\in E^n),$$

the measure $\beta = (p_0, 1-p_0, 0,\ldots,0) = (P_{0j}(n), j\in E^n)$, and the matrix

$$T_n = (P_{ij}(n)1_{i>0}, i,j\in E^n).$$

So, the condition (C) [2, p.3.3] is true when $n=1$ (in denotations of [2]).

Let us calculate the operator norm $\rho_n \equiv \|T_n\|$:

$$\begin{aligned} \rho_n &= \max_{i\in E^n} v^{-i}\sum_{j\in E^n} v^j P_{ij}(n)1_{i>0} \\ &= \max\left\{\max_{1\leq i<n} v^{-i}\left(q_i v^{i-1} + r_i v^i + p_i v^{i+1}\right), v^{-n}\right\} \\ &= \max_{1\leq i<n}\left(1-(v-1)\left(q_i v^{-1} - p_i\right)\right) = 1-(v-1)v^{-1}\min_{1\leq i<n}(q_i - p_i - (v-1)p_i) \\ &\leq 1-(v-1)v^{-1}(\delta_n - (v-1)/2), \end{aligned} \tag{44}$$

taking into account (13) and the condition (11) under which $\delta_n>0$ and $p_i<1/2$.

Let us put $v = 1+\delta_n$. Then (44) implies the following inequalities

$$\rho_n \leq 1-\delta_n^2/2(1+\delta_n) < 1-\delta_n^2/4.$$

The condition (T) from [2, p. 3.3] is fulfilled when $m=1$ (in denotations of [2]) and the following representation holds true uniformly in a scheme of series

$$(1-\rho_n)^{-1} = O\left(\delta_n^{-2}\right), \qquad n\to\infty. \tag{45}$$

Thus, all the conditions of the Theorem 3.6 [2, Ch. III] are true and in the denotations of the Theorem: $n=m=1$, $\alpha=\beta$, $h=h$, $P=P_n$, $\pi=\pi_n$, $\Pi=\Pi_n$, $T=T_n$, $\rho=\rho_n$, and the norm $\|\cdot\|$ is defined in (42). In particular, for the parameter σ in (3.31) [2] we get the estimate

$$\sigma \leq k\|\alpha\|/(1-\rho) = O\left(\delta_n^{-2}\right), \qquad n\to\infty. \tag{46}$$

In order to applied (3.30) we choose (3.29) according to (3.32)

$$\omega \le \omega_1 = 2\exp\bigl(-(1-\pi h)\ln(\alpha h)/\pi h(1-(\alpha h))\bigr) - 1 = O(1), \qquad n \to \infty, \tag{47}$$

where we used the equalities $\pi h = \pi_0^n = \lambda_n \sigma_n / p_0$, $\alpha h = p_0$ and the condition of distancing from zero (12).

From (45), (46), (47) we can calculate the asymptotics for the parameter θ_0 in (3.29) [2]:

$$(1-\theta_0)^{-1} = O\left((1-\rho_n)^{-1}\right) O(\sigma\omega) = O\left(\delta_n^{-4}\right), \qquad n \to \infty. \tag{48}$$

Let us choose in (3.30) [2, p. 3.3] the parameter $\theta = (1+\theta_0)/2$.

Since $\theta - \theta_0 = (1-\theta_0)/2$, $1-\theta = (1-\theta_0)/2$ so from (46), (48) and from (3.30) [2, p. 3.3] we deduce the inequality (41) in the form

$$\left\| P_n^t - \Pi_n \right\| \le b_n (1-r_n)^t, \tag{49}$$

where

$$r_n^{-1} = \max\left((1-\rho_n)^{-1}, (1-\theta)^{-1}\right) = O\left(\delta_n^{-4}\right), \qquad n \to \infty, \tag{50}$$

$$b_n = (1+\sigma)/(\theta-\theta_0) = O\left(\delta_n^{-6}\right), \qquad n \to \infty. \tag{51}$$

Finally, we deduce the following inequality for the second term in the right-hand part of (7.43) [2, Ch. VII]

$$\begin{aligned}\pi(H) a r_\alpha^{-1} \ln(1 + be/a\pi(H)) &= \lambda_n a r_n^{-1} \ln(1 + b_n e/a\lambda_n) \\ &\le \lambda_n O(1) O\left(\delta_n^{-4}\right) \ln\left(O\left(\delta_n^{-6}\right)\lambda_n^{-1}\right) \\ &= O\left(\lambda_n \delta_n^{-4} \ln(1/\delta_n\lambda_n)\right), \qquad n \to \infty,\end{aligned} \tag{52}$$

taking into account the identity (35) under the boundedness conditions (12) and the estimates (50), (51).

The relation (14) $\lambda_n \ln \lambda_n^{-1} = o(\delta_n^4)$, $n \to \infty$ is equivalent to the convergence to zero of the last term in (15): $\lambda_n \delta_n^{-4} \ln(\lambda_n^{-1}\delta_n^{-1}) \to 0$, $n \to \infty$. Really, (14) follows from (15) since $\lambda_n \to 0$, $\delta_n \to 0$. From the other hand, from (14) we deduce that $\delta_n^{-4} = o(1/\lambda_n \ln \lambda_n^{-1})$ implying

$$\lambda_n \delta_n^{-4} \ln \delta_n^{-1} = \lambda_n o\left(\left(\lambda_n \ln \lambda_n^{-1}\right)^{-1} \ln(\lambda_n \ln \lambda_n^{-1})^{-1}\right) = o(1), \qquad n \to \infty,$$

which concludes (15).

Since after putting in (7.43) [2, Ch. VII] $t = \lambda_n^{-1} s$ it follows from the inequality $s > s_0 > 0$ that

$$\lambda_n^{-1} s_0 \ge t_0 \equiv r_\alpha^{-1} \ln^+(b/a\pi(H)) = O\left(\delta_n^{-4} \ln(1/\delta_n\lambda_n)\right), \qquad n \to \infty,$$

as the consequence from the convergence to zero of the right-hand part of (15) and therefore $t \ge t_0$ in the Corollary 7.5.

This substitution and taking into account (40) and (52) prove (15). □

Proof of Corollary 2.5. The convergence $\lambda_n \to 0$, $\omega_n \to 0$ was proved in the Corollary 2.2. Correctness of (12) follows from (38).

If $\lim \delta_n > 0$ then the uniform ergodicity holds true and the Corollary statement will be evident since under the condition $\lambda_n + \omega_n \to 0$, $n \to \infty$.

So, we can assume that $\delta_n \to 0$.

From the relation (38) $\lambda_n \sim \varkappa \sigma_n^{-1}$, $0 < \varkappa < \infty$ we find

$$\lambda_n \ln \lambda_n^{-1} \sim \sigma_n^{-1} \ln \sigma_n, \qquad n \to \infty. \tag{53}$$

Furthermore, it follows from the definition (13) that

$$\theta_t = \prod_{i=1}^{t}(1 + (q_i - p_i)/p_i) \geq (1 + 2\delta_n)^t,$$
$$\sigma_n = \sum_{t<n} \theta_t \geq (2\delta_n)^{-1}((1 + 2\delta_n)^n - 1),$$

These relations and (17) imply that

$$(1 + 2\delta_n)^n \geq \exp((2 - \varepsilon)n\delta_n) \geq \exp((3 + \varepsilon)\ln n) = n^{3+\varepsilon},$$
$$\sigma_n^{-1} = O(\delta_n n^{-3-\varepsilon}),$$

for some $\varepsilon > 0$ starting from some number.

So, in the consequence of (53) the condition (14) hold true:

$$\lambda_n \ln \lambda_n^{-1} = O\left(\delta_n n^{-3-\varepsilon} \ln(\delta_n^{-1} n^{3+\varepsilon})\right) = o\left(\delta_n^4\right), \qquad n \to \infty,$$

since $\delta_n^{1-\alpha} n^{1+\varepsilon/3} \to \infty$ for all sufficiently small $\alpha, \varepsilon > 0$ given (17). Hereof,

$$\left(\delta_n n^{-3-\varepsilon} \ln \delta_n^{-1}\right)/\delta_n^4 = \left(\delta_n n^{1+\varepsilon/3}\left(\ln \delta_n^{-1}\right)^{-1/3}\right)^{-3} \to 0, \qquad n \to \infty. \qquad \square$$

Proof of Remark 2.3. According to [1] the ergodicity of the chain is equivalent to the convergence of the series $\sum \varkappa_t$, which corresponds to the convergence of the series $\sum_{n\geq 1}\prod_{i=1}^{n}(p_i/q_i)$. By the definition (13) the convergence of the last series follows from the convergence of the following series

$$\sum_{n\geq 1}\prod_{i=1}^{n}(1 - \delta_n/q_i) \leq \sum_{n\geq 1}(1 - \delta_n)^n < \infty.$$

In the conditions (b) the equality $\delta_{i+1} = q_i - p_i$ hold true, so $q_i > 1/2$ and

$$\sum_{n\geq 1}\prod_{i=1}^{n}(p_i/q_i) \geq \sum_{n\geq 1}\prod_{i=1}^{n}(1 - 2\delta_{i+1}) = \infty. \qquad \square$$

We are grateful to the anonymous referee for the insightful comments that have significantly improved the paper.

References

1. K. L. Chung, *Markov Chains with Stationary Transition Probabilities*, Springer Verlag, 1960.
2. N. V. Kartashov, *Strong Stable Markov Chains*, VSP, Utrecht, The Netherlands, 1996.
3. N. V. Kartashov, *Determination of the spectral index of ergodicity of a birth-and-death process*, Ukrainian Mathematical Journal **52** (2000), no. 7, 1018–1028.
4. S. P. Mayn and R. L. Tweedie, *Markov chains and Stochastic Stability*, Springer-Verlag, 1993.
5. V. M. Zolotarev, *Modern Theory of Summation of Random Variables*, De Gruyter, 1997.
6. A. M. Zubkov, *Inequalities for transition probabilities with taboos and their applications*, Math. USSR Sb. **37** (1980), no. 4, 451–488.

Department of Probability Theory, Statistics and Actuarial Mathematics, Faculty of Mechanics and Mathematics, Taras Shevchenko National University of Kyiv, 64/13, Volodymyrska Street, City of Kyiv, Ukraine, 01601

E-mail address: nkartashov@skif.com.ua

Received 04/09/2013

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 51–63

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 51–63

COMPOUND KERNEL ESTIMATES FOR THE TRANSITION PROBABILITY DENSITY OF A LÉVY PROCESS IN $\mathbb{R}^n$

UDC 519.21

VICTORIA KNOPOVA

Abstract. We construct in the small-time setting the upper and lower estimates for the transition probability density of a Lévy process in $\mathbb{R}^n$. Our approach relies on the complex analysis technique and the asymptotic analysis of the inverse Fourier transform of the characteristic function of the respective process.

Анотація. Побудовано верхню та нижню оцінки для перехідної щільності процесу Леві в $\mathbb{R}^n$ при малих значеннях часового параметру. Підхід, який був використаний у статті, базується на застосуванні техніки комплексного аналізу, та асимптотичного аналізу оберненого перетворення Фур'є характеристичної функції, що відповідає процесу.

Аннотация. Построены верхняя и нижняя оценки для переходной плотности процесса Леви в $\mathbb{R}^n$ при малых значениях временного параметра. Используемый подход основан на применении техники комплексного анализа, и асимптотического анализа обратного преобразования Фурье характеристической функции, соответствующей процессу.

1. Introduction

Let Z_t be a real-valued Lévy process in $\mathbb{R}^n$ with characteristic exponent ψ, i.e.

$$\mathsf{E}\, e^{i\xi\cdot Z_t} = e^{-t\psi(\xi)}, \qquad \xi \in \mathbb{R}^n.$$

It is known that the characteristic exponent ψ admits the Lévy-Khinchin representation

$$\psi(\xi) = ia\cdot\xi - \frac{1}{2}\xi\cdot Q\xi + \int_{\mathbb{R}^n}\left(1 - e^{i\xi\cdot u} + i\xi\cdot u\mathbb{1}_{\|u\|<1}\right)\,\mu(du), \tag{1.1}$$

where $a \in \mathbb{R}^n$, Q is a positive semi-definite $n\times n$ matrix, and μ is a Lévy measure, i.e. $\int_{\mathbb{R}^n}(1\wedge\|u\|^2)\,\mu(du) < \infty$. In what follows we assume that $Q \equiv 0$, and

$$\mu(\mathbb{R}^n) = \infty. \tag{1.2}$$

Clearly, (1.2) is necessary for Z_t to possess a distribution density.

In the past decades such questions as the existence and properties of the transition probability density of Lévy and, more generally, Markov processes, attracted a lot of attention. Although some progress is already achieved, this problem is highly nontrivial. One can prove the existence of the transition probability density of a symmetric Markov process and study its properties by applying the Dirichlet form technique, see [2, 8, 4, 3, 5, 6, 7]. The other approach relies on versions of the Malliavin calculus for jump processes, see [20], [9]–[10], [23]–[25], and provides the pointwise small-time asymptotic of the transition probability density of a Markov process which is a solution to a Lévy-driven SDE. Under certain assumptions on the Lévy measure estimates on the transition probability density are obtained in [11, 12], see also the references therein for earlier results. In [16], which is the one-dimensional predecessor of the current paper, we investigated the transition probability density $p_t(x)$ of a Lévy process, and proposed a

2010 *Mathematics Subject Classification.* Primary 60G51; Secondary 60J75, 41A60.

Key words and phrases. Transition probability density, transition density estimates, Lévy processes, Laplace method.

specific form of estimates, which we call the *compound kernel estimates*, see Definition 1 below. The approach described in [16] relies on the asymptotic analysis of the inverse Fourier transform of the respective characteristic function. The analysis made in [16] shows that under rather general assumptions the *bell-like estimate*

$$p_t(x) \le \sigma_t g(\|x\|\sigma_t) \tag{1.3}$$

where $g \in L_1(\mathbb{R}^n)$, and σ_t is some "scaling function", is not possible. We also point out, that in the case of a Lévy process the results obtained in [23]–[25] and [10] fit in our observation. At the same time, the upper and lower compound kernel estimates give an adequate picture of behaviour of the transition probability density. In [18, 19] we investigate possible applications of the compound kernel estimates for the construction of the transition probability density of some class of Markov processes.

In this paper we investigate the transition probability density of a Lévy process in the multi-dimensional setting. In Section 2 we set the notation and formulate our main result Theorem 1. Section 3 is devoted to the proof of Theorem 1. In Section 4, Theorems 2 and 3, we treat the particular cases in which it is possible to construct a bell-like estimate (1.3). In Section 5 we illustrate our results by examples. As already mentioned, even if one can construct an estimate of the form (1.3), it may prove to be not informative. In particular, in Example 2 we consider the discretized analogue of an α-stable Lévy measure, and show that in the multi-dimensional setting the bell-like estimate for the respective transition probability density, which is given by Theorem 2, is not integrable in x. At the same time, the compound kernel estimate provided by Theorem 1 gives an adequate answer.

2. Settings and the main result

Notation: We denote by $\mathbb{S}^n$ a unit sphere in $\mathbb{R}^n$; $\xi \cdot \eta$ and $\|\xi\|$ denote, respectively, the scalar product of $\xi, \eta \in \mathbb{R}^n$ and the Euclidean norm of ξ in $\mathbb{R}^n$. We write $f \asymp g$ if there exist constants $c_1, c_2 > 0$ such that $c_1 f(x) \le g(x) \le c_2 f(x)$ for all $x \in \mathbb{R}$; $a \wedge b := \min(a, b)$.

To formulate the regularity assumption on the characteristic exponent ψ we introduce some auxiliary functions. For $x \in \mathbb{R}$ put

$$L(x) := x^2 \mathbb{1}_{\{|x|<1\}}, \qquad U(x) := x^2 \wedge 1, \tag{2.1}$$

and define for $\xi \in \mathbb{R}^n$ the functions

$$\begin{aligned} \psi^L(\xi) &:= \int_{\mathbb{R}^n} L(\xi \cdot u)\, \mu(du) = \int_{|(\xi \cdot u)| \le 1} (\xi \cdot u)^2\, \mu(du), \\ \psi^U(\xi) &:= \int_{\mathbb{R}^n} U(\xi \cdot u)\, \mu(du) = \int_{\mathbb{R}^n} \left((\xi \cdot u)^2 \wedge 1\right)\, \mu(du). \end{aligned} \tag{2.2}$$

Observe that we always have

$$(1 - \cos 1)\psi^L(\xi) \le \operatorname{Re} \psi(\xi) \le 2\psi^U(\xi). \tag{2.3}$$

In addition, we assume that functions ψ^L and ψ^U are comparable, i.e. the assumption below holds true.

A. There exists $\beta > 1$ such that $\sup_{l \in \mathbb{S}^n} \psi^U(rl) \le \beta \inf_{l \in \mathbb{S}^n} \psi^L(rl)$ for all r large enough.

In particular, assumption **A** implies the existence of the transition probability density of Z_t, see Lemma 1 in Section 3.

Define

$$\psi^*(r) := \sup_{l \in \mathbb{S}^n} \psi^U(rl),$$

and

$$\rho_t := \inf\{r\colon \psi^*(r) = 1/t\}. \tag{2.4}$$

We decompose Z_t into a sum

$$Z_t = \bar{Z}_t + \hat{Z}_t - a_t, \tag{2.5}$$

where

• $a_t \in \mathbb{R}^n$ is a vector with coordinates

$$(a_t)_i = t\left(a_i + \int_{1/\rho_t<\|u\|<1} u_i\,\mu(du)\right), \tag{2.6}$$

where the vector $a \in \mathbb{R}^n$ is that from representation (1.1), and ρ_t is defined in (2.4);

• for each $t > 0$ the random variables $\bar{Z}_t$ and $\hat{Z}_t$ are independent; the variable $\bar{Z}_t$ is infinitely divisible for each $t > 0$, with respective characteristic exponent

$$\psi_t(\xi) := t\int_{\rho_t\|u\|\le 1} (1 - e^{i\xi\cdot u} + i\xi\cdot u)\,\mu(du), \tag{2.7}$$

and $\hat{Z}_t$ admits for each $t > 0$ the compound Poisson distribution with the intensity measure

$$\Lambda_t(du) := t\,\mu(du)\mathbb{1}_{\{\rho_t\|u\|>1\}}. \tag{2.8}$$

If condition **A** is satisfied, then $\bar{Z}_t$ possesses a distribution density (see Lemma 2 below), which we denote by $\bar{p}_t(x)$. Therefore, we can represent $p_t(x)$ as

$$p_t(x) = (\bar{p}_t * P_t * \delta_{-a_t})(x), \tag{2.9}$$

where

$$P_t(dy) := e^{-\Lambda_t(\mathbb{R}^n)}\sum_{m=0}^{\infty}\frac{1}{m!}\Lambda_t^{*m}(dy), \tag{2.10}$$

and Λ_t^{*m} denotes the m-fold convolution of the measure Λ_t; by Λ_t^{*0} we understand the δ-measure at 0.

We are looking for a specific form of the estimate for $p_t(x)$, called the *compound kernel estimate*, see the definition below.

Definition 1. Let $\sigma, \zeta\colon (0,\infty) \to \mathbb{R}$, $h\colon \mathbb{R}^n \to \mathbb{R}$ be some functions, and $(Q_t)_{t\ge 0}$ be a family of finite measures on the Borel σ-algebra in $\mathbb{R}^n$. We say that a real-valued function g defined on a set $A \subset (0,\infty)\times\mathbb{R}^n$ satisfies the *upper compound kernel estimate* with parameters $(\sigma_t, h, \zeta_t, Q_t)$, if

$$g_t(x) \le \sum_{m=0}\frac{1}{m!}\int_{\mathbb{R}^n}\sigma_t h((x-y)\zeta_t)\,Q_t^{*m}(dy), \qquad (t,x)\in A. \tag{2.11}$$

If the analogue of (2.11) holds true with the sign $\ge$ instead of $\le$, then we say that the function g satisfies the *lower compound kernel estimate* with parameters $(\sigma_t, h, \zeta_t, Q_t)$.

Let us put a lexicographical order on $\mathbb{R}^n$; namely, we say that $x \le y$, $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, if there exists $1 \le m \le n$, such that for all $i < m$ either $x_i = y_i$, or $x_i < y_i$. Introducing such an order, we can define in the lexicographical sense the first argument of maximum x_t of the function $\bar{p}_t(x)$. Below we show that x_t indeed exists, and for every $t_0 > 0$ there exists $L = L(t_0)$ such that

$$\|x_t\| \le L/\rho_t, \qquad t \in (0, t_0].$$

Below we present our main result on the behaviour of the transition probability density of a Lévy process in $\mathbb{R}^n$.

Theorem 1. *Suppose that condition **A** is satisfied. Then for every $t_0 > 0$ there exist constants $b_i > 0$, $i = 1, \dots, 4$, such that the statements below hold true.*

I. *The function*
$$p_t(x+a_t), \qquad (t,x)\in(0,t_0]\times\mathbb{R}^n,$$
satisfies the upper compound kernel estimate with parameters $(\rho_t^n, f_{\mathrm{upper}}, \rho_t, \Lambda_t)$, *where*
$$f_{\mathrm{upper}}(x) = b_1 e^{-b_2\|x\|}. \tag{2.12}$$

II. *The function*
$$p_t(x+a_t-x_t), \qquad (t,x)\in(0,t_0]\times\mathbb{R}^n,$$
satisfies the lower compound kernel estimate with parameters $(\rho_t^n, f_{\mathrm{lower}}, \rho_t, \Lambda_t)$, *where*
$$f_{\mathrm{lower}}(x) = b_3 \mathbb{1}_{\|x\|\le b_4}. \tag{2.13}$$

One can obtain in the same fashion as in the statement I of the preceding theorem that $p_t(\cdot)\in C_b^\infty(\mathbb{R}^n)$, and construct the upper estimates for derivatives.

Proposition 1. *Suppose that condition* $\boldsymbol{A}$ *is satisfied. Then there exist constants* $b_1>0$ *and* $b_2>0$ *such that for any* $N\ge 1$, $k_i\ge 0$, $i=1,\dots,n$, *such that* $k_1+\dots+k_n=N$, *the function*
$$\left|\frac{\partial^N}{\partial x_1^{k_1}\dots\partial x_n^{k_n}} p_t(x+a_t)\right|, \qquad (t,x)\in(0,t_0]\times\mathbb{R}^n,$$
satisfies the upper compound kernel estimate with parameters $(\rho_t^{n+N}, f_{\mathrm{upper}}, \rho_t, \Lambda_t)$.

Clearly, in the case of a symmetric Lévy measure and a zero drift the statement of Theorem 1 holds true with $a_t = x_t = 0$. Moreover, one can get the sharper upper estimate for $p_t(x)$ and its derivatives.

Proposition 2. *Suppose that the process* Z_t *is symmetric, and condition* $\boldsymbol{A}$ *holds true. Then the first statement of Theorem 1 and Proposition 1 hold true with* a_t *replaced by zero, and* f_{upper} *replaced by*
$$f_{\mathrm{upper}}(x) = b_1 e^{-b_2\|x\|\ln(\|x\|+1)}. \tag{2.14}$$

3. Proofs

We start with the proof of the auxiliary lemma on the growth of ψ^U.

Lemma 1. *Under condition* $\boldsymbol{A}$ *we have for* $\|\xi\|$ *large enough*
$$\psi^U(\xi)\ge c\|\xi\|^{2/\beta}, \tag{3.1}$$
where $c>0$ *is some constant.*

Proof. For $l\in\mathbb{S}^n$ and $r>0$ let
$$\theta^U(rl) := \psi^U(e^r l), \qquad \theta^L(rl) := \psi^L(e^r l). \tag{3.2}$$
Note that the functions L and U satisfy
$$U(x_2)-U(x_1) = \int_{x_1}^{x_2} \frac{2}{x} L(x)\,dx, \qquad x_1<x_2.$$
Then, taking two parallel vectors ξ_1 and ξ_2, and applying the above relation with $x_1 = \xi_1\cdot u$, $x_2=\xi_2\cdot u$, where $u\in\mathbb{R}^n$ and $\|\xi_1\|\le\|\xi_2\|$, we derive by the Fubini theorem
$$\begin{aligned}\psi^U(\xi_2)-\psi^U(\xi_1) &= \int_{\mathbb{R}^n}\big[U((\xi_2,u)) - U((\xi_1,u))\big]\,\mu(du)\\ &= \int_{\mathbb{R}^n}\int_{\|\xi_1\|}^{\|\xi_2\|}\frac{2}{r}L(r(l\cdot u))\,dr\,\mu(du)\\ &= \int_{\|\xi_1\|}^{\|\xi_2\|}\frac{2}{r}\psi^L(lr)\,dr,\end{aligned} \tag{3.3}$$

where $l := \xi_1/\|\xi_1\|$. Thus, by (3.3) and condition **A** we have

$$\theta^U(\xi_2) - \theta^U(\xi_1) \geq \frac{2}{\beta}\int_{\|\xi_1\|}^{\|\xi_2\|} \theta^U(vl)\, dv, \tag{3.4}$$

implying that $\exp\{-\frac{2}{\beta}\|\xi_2\|\}\theta^U(\xi_2) \geq \exp\{-\frac{2}{\beta}\|\xi_1\|\}\theta^U(\xi_1)$. Thus,

$$\psi^U\left(e^{\|\xi_2\|}l\right) = \theta^U(\xi_2) \geq c_1 e^{\frac{2}{\beta}\|\xi_2\|},$$

where $c_1 := e^{-\frac{2}{\beta}\|\xi_1\|}\inf_{l\in\mathbb{S}^n}\theta^U(\xi_1) > 0$. Taking $\inf_{l\in\mathbb{S}^n}$ in the left-hand side of the preceding inequality, we arrive at (3.1). □

The proof of Theorem 1 and Proposition 1 rely on the following lemma.

Lemma 2. *For each $t > 0$ the variable $\overline{Z}_t$ possesses the density $\bar{p}_t(x)$, which satisfies*

$$\left|\frac{\partial^N}{\partial x_1^{k_1}\dots\partial x_n^{k_N}}\overline{p}_t(x)\right| \leq b_1\rho_t^{N+n}e^{-b_2\rho_t\|x\|}, \qquad x\in\mathbb{R}^n,\ t\in(0,t_0], \tag{3.5}$$

for any $N\geq 0$, $k_i\geq 0$, $i=1,\dots,n$, such that $k_1+\dots+k_n=N$.

Proof. For $n=1$ we have

$$t\mu\{u\colon \rho_t\|u\|\geq 1\}\leq t\psi^*(\rho_t)=1.$$

For $n\geq 2$ the situation is similar, but a bit more complicated: since

$$\begin{aligned}
\mu\{u\colon \|u\|\geq r\} &\leq \sum_{i=1}^n \mu\{u\colon |u_i|\geq r\} + \mu\{u\colon \|u\|\geq r, |u_i|<r, i=1,\dots,n\}\\
&\leq \sum_{i=1}^n \mu\{u\colon |u_i|\geq r\} + \mu\{u\colon r/2\leq |u_i|<r, i=1,\dots,n\}\\
&= \sum_{i=1}^n \mu\{u\colon |u_i|\geq r\}\\
&\quad + \mu\{u\colon |u_i|\geq r,\ 1\leq i\leq n\} - \mu\{u\colon |u_i|\geq r/2, 1\leq i\leq n\}\\
&\leq \sum_{i=1}^n \mu\{u\colon |u_i|\geq r\} + \mu\{u\colon \exists i\colon |u_i|\leq r\}\\
&\leq (n+1)\psi^*(1/r),
\end{aligned}\tag{3.6}$$

we arrive at $t\mu\{u\colon \rho_t\|u\|\geq 1\}\leq n+1$. Therefore,

$$\begin{aligned}
\operatorname{Re}\psi_t(\xi) &= t\operatorname{Re}\psi(\xi) - t\int_{\rho_t\|u\|\geq 1}(1-\cos(\xi\cdot u))\,\mu(du)\\
&\geq t\operatorname{Re}\psi(\xi) - 2t\mu\{u\colon \rho_t\|u\|\geq 1\}\\
&= t\operatorname{Re}\psi(\xi) - 2(n+1) \geq t\left(\frac{1-\cos 1}{\beta}\right)\psi^U(\xi) - 2(n+1)\\
&\geq c_1 t\|\xi\|^{2/\beta} - 2(n+1).
\end{aligned}\tag{3.7}$$

where in the last line we used (3.1). Thus, by Lemma 1 the variable $\overline{Z}_t$ possesses a distribution density $\overline{p}_t\in C_b^\infty(\mathbb{R}^n)$, and for any $N\geq 0$, $k_1+\ldots+k_n=N$, we have

$$\frac{\partial^N}{\partial x_1^{k_1}\dots\partial x_n^{k_n}}\overline{p}_t(x) = (2\pi)^{-n}\int_{\mathbb{R}^n}(-ix_1)^{k_1}\dots(-ix_n)^{k_n}e^{-ix\cdot\xi-\psi_t(\xi)}\,d\xi. \tag{3.8}$$

Put $H(t,x,z) := -iz\cdot x - \psi_t(z)$. Note that by the structure of ψ_t the function $H(t,x,z)$ can be extended analytically (with respect to z) to $\mathbb{C}^n$. Applying the Cauchy theorem, we derive

$$\begin{aligned}\frac{\partial^N}{\partial x_1^{k_1}\dots\partial x_n^{k_n}}\overline{p}_t(x) &= (2\pi)^{-n}\int_{\mathbb{R}^n}(-iz_1)^{k_1}\dots(-iz_n)^{k_n}e^{H(t,x,z)}\,dz\\ &= (2\pi)^{-n}\int_{\mathbb{R}^n}\prod_{j=1}^N(-iy_j+\eta_j)^{k_j}e^{x\cdot\eta-ix\cdot y-\psi_t(y+i\eta)}\,dy.\end{aligned}$$

for any $\eta\in\mathbb{R}^n$ satisfying $\|\eta\|\le\rho_t$. Since the proof of the above equality repeats line by line the proof of [16, Lemma 3.4], see also [14] and [15] for the n-dimensional case, we omit the details.

For $\|\eta\|\le\rho_t$ we have

$$\begin{aligned}\operatorname{Re} H(t,x,y+i\eta) &= x\cdot\eta - t\int_{\rho_t\|u\|\le 1}\left(1-\eta\cdot u-e^{-u\cdot\eta}\right)\mu(du)\\ &\quad - t\int_{\rho_t\|u\|\le 1}e^{-\eta\cdot u}(1-\cos(y\cdot u))\,\mu(du)\\ &\le x\cdot\eta-\psi_t(i\eta)-e^{-1}\operatorname{Re}\psi_t(y),\end{aligned}$$

which implies the upper bound

$$\left|\frac{\partial^N}{\partial x_1^{k_1}\dots\partial x_n^{k_n}}\overline{p}_t(x)\right| \le c_2 e^{\eta\cdot x-\psi_t(i\eta)}\int_{\mathbb{R}^n}(\|\eta\|+\|y\|)^N e^{-e^{-1}\operatorname{Re}\psi_t(y)}\,dy. \tag{3.9}$$

Put

$$c := \sup_{|s|\le 1}\left|\frac{1-s-e^{-s}}{s^2}\right|, \qquad s\in\mathbb{R}.$$

Using again the inequality $\|\eta\|\le\rho_t$ and that $\{u\colon \rho_t\|u\|\le 1\}\subset\{u\colon |\eta\cdot u|\le 1\}$, we derive

$$-\psi_t(i\eta)\le ct\int_{\rho_t\|u\|\le 1}|\eta\cdot u|^2\mu(du)\le ct\psi^*(\rho_t)=c.$$

Thus, taking in (3.9) the vector η with coordinates $\eta_i=-\rho_t\operatorname{sign}x_i$, $i=1,\dots,n$, we get

$$\left|\frac{\partial^N}{\partial x_1^{k_1}\dots\partial x_n^{k_n}}\overline{p}_t(x)\right| \le c_3 e^{-\rho_t\|x\|}\int_{\mathbb{R}^n}\left(\rho_t^N+\|y\|^N\right)e^{-e^{-1}\operatorname{Re}\psi_t(y)}\,dy, \tag{3.10}$$

where $c_3\equiv c_3(n,N)>0$ is some constant. Recall that in (3.7) we proved that

$$\operatorname{Re}\psi_t(y)\ge tc_4\psi^U(y)-2,$$

where $c_4 := \frac{1-\cos 1}{\beta}$. Therefore,

$$\left|\frac{\partial^N}{\partial x_1^{k_1}\dots\partial x_n^{k_n}}\overline{p}_t(x)\right| \le c_5 e^{-\rho_t\|x\|}\sup_{l\in\mathbb{S}^n}\left(\rho_t^N I_{n-1}(t,c_6,l)+I_{N+n-1}(t,c_6,l)\right),$$

where $c_6 := e^{-1}c_4$, and

$$I_k(t,\lambda,l) := \int_0^\infty e^{-\lambda t\theta^U(vl)+(k+1)v}\,dv, \qquad k\ge 0. \tag{3.11}$$

To finish the proof we need to show that

$$\sup_{l\in\mathbb{S}^n} I_k(t,\lambda,l)\le c_7\rho_t^{k+1}. \tag{3.12}$$

We get

$$\begin{aligned}\sup_{l\in\mathbb{S}^n} I_k(t,\lambda,l) &= \rho_t^{k+1}\sup_{l\in\mathbb{S}^n}\int_0^\infty e^{-\lambda t[\theta^U(vl)-\theta^U(v_tl)]+(k+1)(v-v_t)-\lambda t\theta^U(v_tl)}\,dv\\ &\le \rho_t^{k+1}\int_0^\infty e^{-\lambda t\inf\limits_{l\in\mathbb{S}^n}[\theta^U(vl)-\theta^U(v_tl)]+(k+1)(v-v_t)-\lambda t\inf\limits_{l\in\mathbb{S}^n}\theta^U(v_tl)}\,dv\\ &\le \rho_t^{k+1}\left[\int_0^{v_t}+\int_{v_t}^\infty\right] e^{-\lambda t\inf\limits_{l\in\mathbb{S}^n}[\theta^U(vl)-\theta^U(v_tl)]+(k+1)(v-v_t)}\,dv,\end{aligned}$$

where $v_t:=\ln\rho_t$, and in the last line we used that θ^U is non-negative. To estimate the first integral observe that

$$\int_0^{v_t} e^{-\lambda t[\theta^U(vl)-\theta^U(v_tl)]+(k+1)(v-v_t)}\,dv \le e^{\lambda t\psi^U(l\rho_t)}\int_0^{v_t} e^{(k+1)(v-v_t)}\,dv\le \frac{e^\lambda}{k+1}. \quad (3.13)$$

Using condition **A** and (3.4) we derive

$$\begin{aligned}\left[\theta^U(vl)-\theta^U(v_tl)\right] &= 2\int_{v_t}^v \theta^L(rl)\,dr \ge \frac{2}{\beta}\int_{v_t}^v\theta^U(rl)\,dr\\ &= \frac{2}{\beta}\theta^U(v_tl)(v-v_t)+\frac{4}{\beta}\int_{v_t}^v\int_{v_t}^r\theta^L(sl)\,ds\,dr\\ &\ge \frac{2}{\beta}\theta^U(v_tl)(v-v_t)+\frac{4}{\beta^2}\int_{v_t}^v\int_{v_t}^r\theta^U(sl)\,ds\,dr\\ &\ge \frac{2}{\beta}\theta^U(v_tl)(v-v_t)+\frac{4}{\beta^2}\theta^U(v_tl)(v-v_t)^2.\end{aligned}$$

Further, by (2.3) and condition **A** we have

$$t\inf_{l\in\mathbb{S}^n}\theta^U(v_tl)\ge\frac{t(1-\cos 1)}{2\beta}\sup_{l\in\mathbb{S}^n}\psi^U(\rho_tl)=\frac{t(1-\cos 1)}{2\beta}\sup_{l\in\mathbb{S}^n}\psi^*(\rho_t)=\frac{1-\cos 1}{2\beta}, \quad (3.14)$$

implying

$$t\inf_{l\in\mathbb{S}^n}\left[\theta^U(vl)-\theta^U(v_tl)\right]\ge b(v-v_t)+2b\beta^{-1}(v-v_t)^2,$$

where $b=(1-\cos 1)/\beta^2$. Thus,

$$\int_{v_t}^\infty e^{-t\lambda\inf_{l\in\mathbb{S}^n}[\theta^U(vl)-\theta^U(v_tl)]+(k+1)(v-v_t)}\,dv\le\int_0^\infty e^{(k+1)w-b\lambda w-\frac{2b\lambda}{\beta}w^2}\,dw<\infty. \quad (3.15)$$

Combining (3.13) and (3.15) we get (3.12), which finishes the proof. □

If the Lévy measure μ is symmetric, one can refine the upper estimate in (3.5).

Lemma 3. *Let condition **A** hold true, and suppose in addition that the Lévy measure μ is symmetric. Then for any $N\ge 0$, and any $k_i\ge 0$, $i=1,\dots,n$, $k_1+\ldots+k_n=N$, we have*

$$\left|\frac{\partial^N}{\partial x_1^{k_1}\dots\partial x_n^{k_N}}\overline{p}_t(x)\right|\le b_1\rho_t^{N+n}e^{-b_2\rho_t\|x\|\ln(\rho_t\|x\|+1)},\qquad x\in\mathbb{R}^n,\ t\in(0,t_0]. \quad (3.16)$$

Proof. By the same argument as in [16, Lemma 3.6] we have for *any* $\eta\in\mathbb{R}^n$

$$\left|\frac{\partial^N}{\partial x_1^{k_1}\dots\partial x_n^{k_N}}\overline{p}_t(x)\right|\le(2\pi)^{-n}e^{\eta\cdot x-\psi_t(i\eta)}\int_{\mathbb{R}^n}(\|y\|+\|\eta\|)^N e^{-\mathrm{Re}\,\psi_t(y)}dy. \quad (3.17)$$

By Lemma 2, the integral in (3.17) is estimated from above by $c_1(\|\eta\|^N\rho_t^n+\rho_t^{N+n})$, where $c_1>0$ is some constant. For $\psi_t(i\eta)$ we have

$$\begin{aligned}
-\psi_t(i\eta) &= t\int_{\rho_t\|u\|\le 1}[\cosh(\eta\cdot u)-1]\,\mu(du) = t\theta\big(\|\eta\|/\rho_t\big)\int_{\rho_t\|u\|\le 1}(\eta\cdot u)^2\,\mu(du)\\
&\le t\theta\big(\|\eta\|/\rho_t\big)\big(\|\eta\|/\rho_t\big)^2\sup_{l\in\mathbb{S}^n}\int_{\rho_t\|u\|\le 1}\rho_t^2(l\cdot u)^2\,\mu(du)\\
&\le (\cosh(\|\eta\|/\rho_t)-1)t\psi^*(\rho_t)\\
&= \cosh(\|\eta\|/\rho_t)-1,
\end{aligned}$$

where $\theta(s):=s^{-2}(\cosh s-1)$, $s\ge 0$, is increasing. Since sofar η was arbitrary, take η with coordinates satisfying $\operatorname{sign}\eta_i=-\operatorname{sign}x_i$, $i=1,\dots,n$. Then

$$\left|\frac{\partial^N}{\partial x_1^{k_1}\dots\partial x_n^{k_N}}\overline{p}_t(x)\right|\le c_2\rho_t^{N+n}e^{-\|x\|\|\eta\|+\cosh(\|\eta\|/\rho_t)}. \tag{3.18}$$

Minimizing the expression under the exponent in (3.18) in $\|\eta\|$, we arrive at (3.16). □

Proof of Theorem 1. Upper bound. The proof of the upper bound follows from Lemmas 1 and 2, and representation (2.9).

Lower bound. From Lemma (2) we know that the function $\overline{p}_t(x)$ is continuous in x, and bounded from above by $b_1\rho_t^n$. Without loss of generality we may assume that $\int_{\rho_t\|x\|\le 1}\overline{p}_t(x)dx\ge 1/2$. Then

$$1/2\le\int_{\rho_t\|x\|\le L}\overline{p}_t(x)\,dx\le\frac{w_nL^n}{\rho_t^n}\max_{x\in\mathbb{R}^n}\overline{p}_t(x),$$

where w_n is the volume of a unit ball in $\mathbb{R}^n$. Let x_t be the "smallest" in the lexicographical sense point in which the maximum of $\overline{p}_t(x)$ is achieved. For the off-diagonal lower bound we get using the Taylor formula:

$$\begin{aligned}
\overline{p}_t(x) &\ge \overline{p}_t(x_t)-\left|\sum_{i=1}^n(x-x_t)_i\int_0^1\frac{\partial}{\partial x_i}\overline{p}_t(x_t+r(x-x_t))\,dr\right|\\
&\ge \overline{p}_t(x_t)-\left(\sum_{i=1}^n\int_0^1\left|\frac{\partial}{\partial x_i}\overline{p}_t(x_t+r(x-x_t))\right|^2dr\right)^{1/2}\|x-x_t\| \qquad (3.19)\\
&\ge \frac{1}{2w_nL^n}\rho_t^n-c_1(n)\rho_t^{n+1}\|x-x_t\|\\
&= c_2(n)\rho_t^n(1-c_3(n)\rho_t\|x-x_t\|),
\end{aligned}$$

where in the second line form below we used the on-diagonal estimate

$$\left|\frac{\partial}{\partial y_i}\overline{p}_t(y)\right|\le c(n)\rho_t^{n+1}.$$ □

4. Bell-like estimates

In this section we discuss some particular cases in which we pose more restrictive assumptions on the regularity of the tail of the Lévy measure. We show that under certain assumptions it is possible to write more explicit upper and lower estimates for $p_t(x)$. At the same time, we emphasize that although such estimates can be more explicit, they suppress the vital information about the transition probability density, given by the compound kernel estimates. Moreover, as we will see below, a bell-like estimate may heavily depend on the space dimension.

We begin with some notions on *sub-exponential distributions* in the multi-dimensional setting, see [22] and [21] for more details. We keep the notation of Theorem 1.

Definition 2. [22] We say that G is a *sub-exponential* distribution on $\mathbb{R}^n$ (and write $G \in \mathcal{L}(\mathbb{R}^n)$) if for all $x \in \mathbb{R}^n$ such that $\min_i x_i < \infty$, we have

$$\lim_{t\to\infty} \frac{1-G^{*2}(tx)}{1-G(tx)} = 1. \tag{4.1}$$

Theorem below generalizes the one-dimensional result, proved in [16].

Theorem 2. *Let condition* $\boldsymbol{A}$ *hold true, and suppose that there exist a distribution function* $G \in \mathcal{L}(\mathbb{R}^n)$*, such that*

$$t\mu\left(\{u\colon \|\rho_t u\| > \|v\|\}\right) \le C(1-G(v)), \qquad \|v\| \ge 1,\ t \in (0,t_0], \tag{4.2}$$

where $C > 0$ *is some constant, independent of* t*. Then for every* $t_0 > 0$ *there exist some constant* $C_1 > 0$*, such that*

$$p_t(x+a_t) \le C_1 \rho_t^n \left(f_{\text{upper}}(\rho_t x) + 1 - G(x\rho_t)\right), \qquad x \in \mathbb{R}^n,\ t \in (0,t_0], \tag{4.3}$$

where f_{upper} *is defined by* (2.12)*. If the inequality* (4.2) *holds true with the sign* $\ge$*, then*

$$p_t(x+a_t-x_t) \ge C_2 \rho_t^n (f_{\text{lower}}(\rho_t x) + 1 - G(\rho_t x)), \qquad x \in \mathbb{R}^n,\ t \in (0,t_0], \tag{4.4}$$

where $C_2 > 0$ *is some constant, and* f_{lower} *is defined in* (2.13).

In [16] we proved a version of Theorem 2 in the case when the measure μ is absolutely continuous, and the density is sub-exponential in the sense of [13]. Up to our knowledge sub-exponential *densities* are not studied in the multi-dimensional case, see, however, [22] for a brief comment. We strongly believe that the result analogous to those proved in [16] also can be proved in the multi-dimensional setting, after establishing the necessary properties of sub-exponential densities analogous to those presented in [13]. However, it is possible to prove a version of Theorem 2 under the assumption of a power decay of the Lévy density.

Theorem 3. *Let condition* $\boldsymbol{A}$ *hold true. Suppose that* $\mu(du) = m(u)du$*, and for* $\|u\| \ge 1$ *we have the estimate*

$$t\rho_t^{-n} m\left(u\rho_t^{-1}\right) \le \|u\|^{-n-b}, \qquad t \in (0,t_0], \tag{4.5}$$

where $b > 0$*. Then*

$$p_t(x+a_t) \le c_1 \frac{\rho_t^n}{(1+\rho_t\|x\|)^{n+b}}, \qquad x \in \mathbb{R}^n,\ t \in (0,t_0]. \tag{4.6}$$

If the inequality (4.5) *holds true with the sign* $\ge$*, then*

$$p_t(x+a_t-x_t) \ge c_2 \frac{\rho_t^n}{(1+\rho_t\|x\|)^{n+b}}, \qquad x \in \mathbb{R}^n,\ t \in (0,t_0]. \tag{4.7}$$

The proof of Theorem 2 relies on the results obtained in [22]. In order to make the presentation self-contained, we quote these results below.

It is shown in [22, Theorem 7, Corollary 11] that for a distribution function G the conditions

G1. For $\forall a$, $x \in \mathbb{R}^n$, $a \ge 0$, $x \ge 0$, such that $\min_i x_i < \infty$, $\lim\limits_{t\to\infty} \frac{1-G(tx-a)}{1-G(tx)} = 1$;

G2. All marginals G_i of G are sub-exponential (i.e., $G_i \in \mathcal{L}(\mathbb{R})$),

are equivalent to $G \in \mathcal{L}(\mathbb{R}^n)$, and imply that for $x \ge 0$, $\min x_i < \infty$, and $a \in \mathbb{R}^n$, $a \ge 0$, one has

$$\lim_{t\to\infty} \frac{1-H(tx-a)}{1-G(tx)} = \lambda, \tag{4.8}$$

where

$$H(x) = \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} G^{*k}(x), \quad \lambda \in (0,\infty). \tag{4.9}$$

We also need [22, Theorem 10], which states that if the distribution function G satisfies **G1** and **G2**, and the distribution functions R and F are such that

$$\lim_{t\to\infty} \frac{1-F(tx-a)}{1-G(tx)} = \alpha, \tag{4.10}$$

$$\lim_{t\to\infty} \frac{1-R(tx-a)}{1-G(tx)} = \beta, \tag{4.11}$$

for some $\alpha, \beta \in \mathbb{R}$, and any $a, x \in \mathbb{R}^n$, $a, x \geq 0$, $\min_i x_i < \infty$, then

$$\lim_{t\to\infty} \frac{1-R*F(tx-a)}{1-G(tx)} = \alpha+\beta. \tag{4.12}$$

Proof of Theorem 2. By (4.9) we have

$$p_t(x) \leq \rho_t^n f_{\text{upper}}(x\rho_t) + c_1\rho_t^n \int_{\|v\|\geq 1} f_{\text{upper}}(x\rho_t - v)\, G(dv). \tag{4.13}$$

Note that for any $c>0$ the tail of a sub-exponential distribution in $\mathbb{R}$ decays slower than $e^{-c|y|}$ as $|y|\to\infty$, (see [13], also the comment in [16]), which implies that for any $c>0$ the tail of a sub-exponential distribution in $\mathbb{R}^n$ decays slower than $e^{-c\|x\|}$ as $\|x\|\to\infty$. Hence, for $R(x)=1-f_{\text{upper}}(x)$ we have (4.11) with $\beta=0$. Thus, by sub-exponentiality of G we have the relation (4.12) with $\alpha=1$, $\beta=0$, i.e.

$$\lim_{s\to\infty} \frac{\int_{\|v\|\geq 1} f(xs-v)\, dG(v)}{1-G(sx)} = 1.$$

Since $\rho_t\to\infty$ as $t\to 0$, we finally derive (4.3) for t small enough.

Similar argument works for the lower bound: in this case we take

$$R(x) = 1 - f_{\text{lower}}(x). \qquad \square$$

Proof of Theorem 3. Let $q(v) := (1+\|v\|)^{-n-b}$, and put

$$Q(v) := \sum_{k=1}^{\infty} q^{*k}(v)/k!, \qquad v\in\mathbb{R}^n.$$

By Theorem 1 and (4.5) we get

$$p_t(x) \leq c\rho_t^n \left(f_{\text{upper}}(x\rho_t) + \int_{\mathbb{R}^n} f_{\text{upper}}(x\rho_t - v) Q(v)\, dv \right). \tag{4.14}$$

Let us estimate $Q(v)$. We have:

$$\begin{aligned} q^{*2}(w) &= \int_{\mathbb{R}^n} \frac{1}{(1+\|v\|)^{n+b}(1+\|w-v\|)^{n+b}}\, dv \\ &= \left[\int_{\{\|w-v\|\leq 2^{-1}\|w\|\}} + \int_{\{\|w-v\|\geq 2^{-1}\|w\|\}} \right] \frac{1}{(1+\|v\|)^{n+b}(1+\|w-v\|)^{n+b}}\, dv \\ &= I_1 + I_2. \end{aligned}$$

To estimate I_1 observe that if $\|w-v\|\leq 2^{-1}\|w\|$, then $\|w\|\leq\|v\|\leq \frac{3}{2}\|w\|$, or $\frac{1}{2}\|w\|\leq \|v\|\leq\|w\|$, implying

$$\frac{1}{1+\|v\|} \leq \frac{2}{2+\|w\|}.$$

Therefore,

$$I_1 \leq \left(\frac{2}{2+\|w\|}\right)^{n+b} \int_{\mathbb{R}} \frac{1}{(1+\|v\|)^{n+b}}\, dv \leq c\left(\frac{2}{1+\|w\|}\right)^{n+b}.$$

Analogously, if $\|w-v\| \geq 2^{-1}\|w\|$, then

$$\frac{2}{2+\|w\|} \geq \frac{1}{1+\|w-v\|},$$

implying

$$I_2 \leq \left(\frac{2}{2+\|w\|}\right)^{n+b} \int_{\mathbb{R}} \frac{1}{(1+\|v\|)^{n+b}}\, dv \leq c\left(\frac{2}{1+\|w\|}\right)^{n+b}.$$

Thus, there exists a constant $C>0$ such that $q^{*2}(v) \leq Cq(v)$. By induction, $q^{*k}(v) \leq C^{k-1}q(v)$, implying $Q(v) \leq c_1 q(v)$, $v \in \mathbb{R}$. Finally, observe that

$$\begin{aligned}\int_{\mathbb{R}} f_{\text{upper}}(x-v)Q(v)\,dv &= \left[\int_{\|x-v\|\geq 2^{-1}\|x\|} + \int_{\|x-v\|\leq 2^{-1}\|x\|}\right] f_{\text{upper}}(x-v)Q(v)\,dv \\ &\leq c_2 f_{\text{upper}}(x/2) + c_3 Q(x) \leq c_4 Q(x).\end{aligned}$$

Thus, we arrive at

$$p_t(x) \leq c_5 \frac{\rho_t^n}{(1+\rho_t\|x\|)^{n+b}},$$

which proves the first part of the theorem. The same argument applies for the lower bound. □

5. Examples

Example 1. Let Z_t be an α-stable process, $\alpha \in (0,2)$, with the Lévy measure $\mu(du) = c_\alpha \|u\|^{-n-\alpha}du$, and the drift vector $b \in \mathbb{R}^n$. One can easily verify that condition **A** is satisfied, and $\rho_t = t^{-1/\alpha}$. Applying Theorem 3, we arrive at

$$p_t(x+bt) \asymp t^{-n/\alpha} \wedge \frac{t}{\|x\|^{1+\alpha}} \asymp t^{-n/\alpha} f\left(t^{-1/\alpha}\|x\|\right), \qquad x \in \mathbb{R}^n,\ t \in (0,t_0],$$

where

$$f(z) = 1 \wedge z^{-\alpha-n}, \qquad z>0, \tag{5.1}$$

and for the lower bound we used that due to the symmetry of the Lévy measure we have $x_t = 0$. Note that by the structure of μ the above estimates hold true for all $t>0$, $x \in \mathbb{R}^n$, and coincides in the case $b=0$ with the well-known estimate for the transition probability density of a symmetric α-stable process.

Observe that for $1<\alpha<2$ we have

$$t^{-1/\alpha}\|x-tb\| \geq t^{-1/\alpha} - t^{1-1/\alpha}\|b\| \geq t^{-1/\alpha}\|x\| - c\|b\|, \qquad t \in (0,t_0].$$

Thus, for such α we arrived at

$$p_t(x) \asymp t^{-n/\alpha} f(t^{-1/\alpha}\|x\|), \qquad t \in (0,t_0],\ x \in \mathbb{R}^n.$$

Example 2. Consider a "discretized version" of an α-stable Lévy measure in $\mathbb{R}^n$. Let $m_{k,\upsilon}(dy)$ be a uniform distribution on a sphere $\mathbb{S}_{k,\upsilon}$ centered at 0 with radius $2^{-k\upsilon}$, $\upsilon>0$, $k \in \mathbb{Z}$. Consider a Lévy process with characteristic exponent of the form (1.1), where

$$\mu(dy) = \sum_{k=-\infty}^{\infty} 2^{k\gamma} m_{k,\upsilon}(dy), \qquad 0<\gamma<2\upsilon,$$

and some drift coefficient $a \in \mathbb{R}^n$. Let us check that in this case $\psi^U(\xi) \asymp \psi^L(\xi) \asymp \|\xi\|^\alpha$, where $\alpha = \gamma/\upsilon$.

Let $k_0 := \upsilon^{-1}\log_2\|\xi\|$. We have

$$\begin{aligned}\psi^U(\xi) &\leq \int_{\mathbb{R}^n} \left(\|\xi\|^2\|y\|^2 \wedge 1\right)\, \mu(dy) \\ &= \|\xi\|^2 \int_{\|y\|\leq/\|\xi\|} \|y\|^2\, \mu(dy) + \int_{\|y\|>1/\|\xi\|} \mu(dy)\end{aligned}$$

$$
\begin{aligned}
&= \|\xi\|^2 \sum_{k\ge k_0} 2^{\gamma k - 2kv} + c_1 \sum_{k\le k_0} 2^{\gamma k} \\
&\le \|\xi\|^2 2^{k_0(\gamma-2v)} \sum_{k\ge k_0} 2^{-(k-k_0)(2v-\gamma)} + c_1 + 2^{\gamma k_0}\frac{1-2^{-\gamma k_0}}{1-2^{-\gamma}} \\
&\le \frac{2^{2v-\gamma}}{2^{2v-\gamma}-1}\|\xi\|^2 2^{\frac{2v-\gamma}{v}\log_2\|\xi\|} + c_2 2^{\frac{\gamma}{v}\log_2\|\xi\|} \le c_3\|\xi\|^\alpha.
\end{aligned}
$$

The above calculations and the inequality $(1-\cos 1)\psi^L(\xi) \le \int_{\mathbb{R}^n}(1-\cos(\xi\cdot y))\,\mu(dy)$ imply that

$$\psi^L(\xi) \le c_4\psi^U(\xi) \le c_5\|\xi\|^\alpha.$$

For the lower bound we have

$$
\begin{aligned}
\psi^L(\xi) &\ge \int_{\|y\|\ge 1/\|\xi\|} |\xi\cdot y|^2\,\mu(dy) \ge m_{k_0,v}\{l\in\mathbb{S}_{k_0,v}\colon |\cos(l_\xi\cdot l)|>\varepsilon\}\|\xi\|^2 2^{k_0(\gamma-2v)} \\
&= c_6\|\xi\|^\alpha,
\end{aligned}
$$

where $l_\xi := \xi/\|\xi\|$, implying

$$\inf_{\|l\|=1}\psi^L(\|\xi\|l) \ge c\|\xi\|^\alpha.$$

Thus, condition **A** is satisfied, and $\psi^L(\xi)\asymp\psi^U(\xi)\asymp\|\xi\|^\alpha$, which in turn gives $\rho_t\asymp t^{-1/\alpha}$.

Note that for $\|x\|>1$ we have

$$t\mu\left(\{u\colon \rho_t\|u\|>\|x\|\}\right) = t\sum_{n\le n(t,x)} 2^{\gamma n} \le Ct2^{\frac{\gamma}{v}\log_2(\rho_t/\|x\|)} = C\|x\|^{-\gamma/v} = C\|x\|^{-\alpha},$$

where $n(t,x) := \frac{1}{v}\log_2(\rho_t/\|x\|)$. Therefore, condition (4.2) of Theorem 2 holds true with $1-G(x)=\|x\|^{-\alpha}$, $\|x\|\ge 1$. By this theorem we have the following estimate for the respective transition probability density:

$$p_t(x+at) \le c_1 t^{-n/\alpha} f(t^{-1/\alpha}\|x\|) \tag{5.2}$$

where

$$f(z) = 1\wedge z^{-\alpha}, \qquad z>0. \tag{5.3}$$

However, as one may notice, such upper estimate is informative only in the case $n=1$ and $1<\alpha<2$, see [16] for the detailed analysis. In the other cases the upper bound is not integrable! On the other hand, Theorem 1 together with Proposition 2 provides that the transition probability density satisfies the upper compound kernel estimates with parameters $(t^{-1/\alpha}, f_{\text{upper}}, t^{-1/\alpha}, \Lambda_t)$, with

$$f_{\text{upper}}(x) = b_1 e^{-b_2\|x\|\log(1+\|x)}, \quad \text{and} \quad \Lambda_t(du) = t\mathbb{1}_{\{\|u\|\ge t^{1/\alpha}\}}\mu(du).$$

In this case the obtained upper bound is integrable.

Remark 1. The above example illustrates that even if the (re-scaled) Lévy measure can be dominated by a reasonably good function, the explicit upper estimate obtained in Theorem 2 can be extremely inexact. Heuristically, the condition (4.2) is imposed on the tail of the re-scaled *measure*, which suppresses its intrinsic behaviour. See, however, [12] for another approach in a similar situation. On the other hand, the condition on the behaviour of the *density* can lead to adequate results, as we saw in Example 1. Possibly, one can modify the assumption Theorem 2 and get more reasonable explicit estimates, but in fact it is not needed, since the compound kernel estimates obtained in Theorem 1 already contain the information, sufficient for many applications, see [18] and [19].

Acknowledgement. The author is very grateful to Alexei Kulik for fruitful discussions. The DFG Grant Schi 419/8-1, and the Scholarship of the President of Ukraine for young scientists (2011-2013) are gratefully acknowledged.

References

1. E. T. Copson, *Asymptotic Expansions*, Cambridge Uni. Press, Cambridge, 1965.
2. M. T. Barlow, R. B. Bass, Z.-Q. Chen, and M. Kassmann, *Non-local Dirichlet forms and symmetric jump processes*, Trans. Amer. Math. Soc. **361** (2009), 1963–1999.
3. M. T. Barlow, A. Grigoryan, and T. Kumagai, *Heat kernel upper bounds for jump processes and the first exit time*, J. Reine Angew. Math. **626** (2009), 135–157.
4. Z.-Q. Chen, *Symmetric jump processes and their heat kernel estimates*, Sci. China Ser. A. **52** (2009), 1423–1445
5. Z.-Q. Chen and T. Kumagai, *Heat kernel estimates for stable-like processes on d-sets*, Stoch. Proc. Appl. **108** (2003), 27–62.
6. Z.-Q. Chen and T. Kumagai, *Heat kernel estimates for jump processes of mixed types on metric measure spaces*, Probab. Th. Rel. Fields. **140(1–2)** (2008), 277–317.
7. Z.-Q. Chen, P. Kim, and T. Kumagai, *On Heat kernel estimates and parabolic Harnack inequality for jump processes on metric measure spaces*, Acta Math. Sin. (Engl. Ser.) **25** (2009), 1067–1086.
8. Z.-Q. Chen, P. Kim, and T. Kumagai, *Global heat kernel estimates for symmetric jump processes*, Trans. Amer. Math. Soc. **363(9)** (2011), 5021–5055.
9. Y. Ishikawa, *Asymptotic behavior of the transitions density for jump type processes in small time*, Tôhoku Math. J. **46** (1994), 443–456.
10. Y. Ishikawa, *Density estimate in small time for jump processes with singular Lévy measures*, Tôhoku Math. J. **53(2)** (2001), 183–202.
11. K. Kaleta and P. Sztonyk, *Upper estimates of transition densities for stable-dominated semigroups*, J. Evol. Equ. **13** (2013), 633-–650
12. K. Kaleta and P. Sztonyk, *Estimates of Transition Densities and their Derivatives for Jump Lévy Processes*, Preprint, 2013.
13. C. Klüppelberg, *Subexponential distributions and characterizations of related classes*, Probab. Th. Rel. Fields. **82** (1989), 259–269.
14. V. Knopova and R. L. Schilling, *Transition density estimates for a class of Lévy and Lévy-type processes*, J. Theor. Probab. **25** (2012), 144–170.
15. V. Knopova and A. Kulik, *Exact asymptotic for distribution densities of Lévy functionals*, Electronic J. Probab. **16** (2011), 1394–1433.
16. V. Knopova and A. Kulik, *Intrinsic small time estimates for distribution densities of Lévy processes*, Rand. Op. Stoch. Eq. (to appear)
17. V. Knopova, *Asymptotic behaviour of the distribution density some Lévy fnctionals in* $\mathbb{R}^n$, Th. Stoch. Proc. **17(33)** (2011), no. 2, 35–54.
18. V. Knopova and A. Kulik, *Parametrix Construction for Certain Lévy-type Processes and Applications*, Preprint, 2013.
19. V. Knopova and A. Kulik, *Intrinsic Compound Kernel Estimates for the Transition Probability Density of a Lévy Type Processes and their Applications*, Preprint, 2013.
20. R. Léandre, *Densité en temps petit d'un processus de sant*, Seminaire de Probabilitts XXI (J. Azéma, P. A. Meyer, and M. Yor, eds.), Lecture Notes in Math., vol. 1247, Springer, Berlin, 1987, pp. 81–99.
21. E. Omey, F. Mallor, and J. Santos, *Multivariate subexponential distributions and random sums of random vectors*, Adv. in Appl. Probab. **38(4)** (2006), 1028–1046.
22. E. Omey, *Subexponential distribution functions in* $\mathbb{R}^d$, J. Math. Sci. **138(1)** (2006), 5434–5449.
23. J. Picard, *On the existence of smooth densities for jump processes*, Prob. Th. Rel. Fields **105** (1996), 481–511.
24. J. Picard, *Density in small time for Lévy processes*, ESAIM Probab. Statist. **1** (1997), 358–389.
25. J. Picard, *Density in small time at accessible points for jump processes*, Stoch. Proc. Appl. **67** (1997), 251–279.

V. M. Glushkov Institute of Cybernetics, NAS of Ukraine, 40, Acad. Glushkov Ave., 03187, Kiev, Ukraine

E-mail address: vic_knopova@gmx.de

Received 01/06/2013

STORAGE IMPULSIVE PROCESSES ON INCREASING TIME INTERVALS

UDC 519.21

V. S. KOROLIUK, R. MANCA, AND G. D'AMICO

Abstract. The Storage Impulsive Process (SIP) $S(t)$ is a sum of (jointly independent) random variables defined on the embedded Markov chain of a homogeneous Markov process.

The SIP is considered in the series scheme on increasing time intervals t/ε, with a small parameter $\varepsilon \to 0$, $\varepsilon > 0$. The SIP is investigated in the average and diffusion approximation scheme. The large deviation problem is considered under corresponding scaling with an asymptotically small diffusion.

Анотація. Імпульсні процеси накопичення (ІПН) задаються сумами (незалежними в сукупності) випадкових величин, визначених на вкладеному ланцюгу Маркова однорідного марковського процесу.

ІПН розглядаються у схемі серій на зростаючих інтервалах часу t/ε, з малим параметром серії $\varepsilon \to 0$, $\varepsilon > 0$. ІПН досліджуються у схемах усереднення та дифузійної апроксимації. Проблема великих відхилень розглядається при відповідному нормуванні з асимптотично малою дифузією.

Аннотация. Импульсные процессы накопления (ИПН) задаются суммами (независимыми в совокупности) случайных величин, определенных на вложенной цепи Маркова однородного марковского процесса.

ИПН рассматриваются в схеме серий на возрастающих интервалах времени t/ε, с малым параметром серии $\varepsilon \to 0$, $\varepsilon > 0$. ИПН исследуются в схемах укрупнения и диффузионной аппроксимации. Проблема больших уклонений рассматривается при соответствующей нормировке с асимптотически малой диффузией.

1. Introduction

The Storage Impulsive Process (SIP) $S(t)$ is a sum of (jointly independent) random variables defined on the embedded Markov chain of a homogeneous Markov process

$$S(t) = u + \sum_{n=1}^{\nu(t)} \alpha_n(x_n), \qquad t \geq 0,\ u \in \mathbb{R}^d. \tag{1}$$

The time homogeneous Markov process $x(t)$, $t \geq 0$, is defined on a standard phase space $(E, \mathcal{E})$ by the generator

$$Q\varphi(x) = q(x) \int_E P(x, dy)[\varphi(y) - \varphi(x)], \qquad x \in E,$$

for a real valued test function $\varphi(x)$, $x \in E$, with a bounded sup-norm:

$$\|\varphi(x)\| := \sup_{x \in E} |\varphi(x)|.$$

The embedded Markov chain x_n, $n \geq 0$, is defined by

$$x_n := x(\tau_n), \qquad n \geq 0,$$

2000 *Mathematics Subject Classification.* Primary 60J45; Secondary 60K05.

Key words and phrases. Storage impulsive process; average, diffusion approximation; large deviation problem.

where the renewal moments of jumps are given by

$$\tau_{n+1} = \tau_n + \theta_{n+1}, \qquad n \geq 0,\ \tau_0 = 0,$$

and the sojourn times θ_{n+1}, $n \geq 0$, are such that

$$\mathsf{P}(\theta_{n+1} \geq t \mid x_n = x) = e^{-q(x)t} =: \mathsf{P}(\theta_x \geq t).$$

The stochastic kernel $P(x, B)$, $x \in E$, $B \in \mathcal{E}$, defines the transition probabilities of the embedded Markov chain

$$P(x, B) = \mathcal{P}\{x_{n+1} \in B \mid x_n = x\}.$$

The counting process is defined by

$$\nu(t) := \max\{n > 0 \colon \tau_n \leq t\}, \qquad t \geq 0.$$

The random variables in (1) have the distribution functions

$$\Phi_x(dv) = P\{\alpha_n(x) \in dv\} := \mathsf{P}\{\alpha_n(x_n) \in dv \mid x_n = x\}, \qquad x \in E.$$

The SIP may be considered as a random evolution process [1, Ch.2]. The switching Markov process $x(t)$, $t \geq 0$, describes a random environment.

A1: The main assumption is the uniform ergodicity of the Markov process $x(t)$, $t \geq 0$, with the stationary distribution $\pi(B)$, $B \in \mathcal{E}$, satisfying the equation:

$$\pi(dx)q(x) = q\rho(dx), \qquad q = \int_E \pi(dx)q(x).$$

The stationary distribution $\rho(B)$, $B \in \mathcal{E}$, of the embedded Markov chain x_n, $n \geq 0$, satisfies the equation

$$\rho(B) = \int_E \rho(dx)P(x, B), \qquad B \in \mathcal{E},\ \rho(E) = 1.$$

Provided that the main assumption **A1** takes place the potential operator R_0 may be given by a solution of the equation [1, Ch. 2]

$$QR_0 = R_0 Q = \Pi - I, \qquad \Pi\varphi(x) := \int_E \pi(dx)\varphi(x).$$

2. SIP on increasing time intervals in average scheme.

The SIP on increasing time intervals in average scheme is considered in the series scheme with the small parameter $\varepsilon \to 0$, $\varepsilon > 0$, in the following scaling:

$$S^\varepsilon(t) = u + \varepsilon \sum_{n=1}^{\nu(t/\varepsilon)} \alpha_n(x_n), \qquad t \geq 0,\ \varepsilon > 0,\ u \in \mathbb{R}^d. \tag{2}$$

The random evolution approach [1, Ch. 3, 5] is an effective method of asymptotic analysis (2) when $\varepsilon \to 0$.

Proposition 2.1. *The SIP* (2) *in the average scheme convergences weakly*

$$S^\varepsilon(t) \Rightarrow S^0(t) = u + \widehat{a}_0 t, \qquad \varepsilon \to 0, \tag{3}$$

where the average velocity is such that

$$\widehat{a}_0 = q\widehat{a}, \qquad \widehat{a} = \int_E \rho(dx)a(x), \qquad a(x) = \int_{\mathbb{R}^d} v\, \Phi_x(dv). \tag{4}$$

Proof of Proposition 2.1 is based on the random evolution approach [1, Ch. 3] by using a solution of the singular perturbation problem [1, Ch. 5].

Remark 2.1. For simplcity without loss of generality the proof is realized for the SIP given on real line $\mathbb{R}$, $d = 1$.

According to the definition of a random evolution [1, Ch. 2] we consider the two component Markov process

$$S^\varepsilon(t),\ x^\varepsilon(t) := x(t/\varepsilon), \qquad t \geq 0. \tag{5}$$

Lemma 2.1. *The Markov process* (5) *is characterized by the generator*

$$L^\varepsilon \varphi(u,x) = \varepsilon^{-1} q(x) \int_E P(x,dy) \int_{\mathbb{R}^d} \Phi_y(dv)[\varphi(u+\varepsilon v, y) - \varphi(u,x)]. \tag{6}$$

The proof of Lemma 2.1 is a direct consequence of the definition of the generator [1, Ch. 3].

Remark 2.2. The generator (6) may be rewritten as follows

$$L^\varepsilon \varphi(u,x) = \varepsilon^{-1} \left[Q + Q_0 \Phi_x^\varepsilon\right] \varphi(u,x), \tag{7}$$

where, by definition,

$$Q_0 \varphi(x) := q(x) \int_E P(x,dy)\varphi(y),$$

$$\Phi_x^\varepsilon \varphi(u) := \int_{\mathbb{R}^d} \Phi_x(dv)[\varphi(u+\varepsilon v) - \varphi(u)].$$

On a test function $\varphi(u)$ being smooth enough,

$$\Phi_x^\varepsilon \varphi(u) = \varepsilon[a(x)\varphi'(u) + \delta^\varepsilon(x)\varphi(u)]$$

with the negligible term:

$$\|\delta^\varepsilon(x)\varphi(u)\| \to 0, \qquad \varepsilon \to 0,\ \varphi(u) \in C^2(\mathbb{R}).$$

Lemma 2.2. *The generator* (7) *admits the following asymptotic expansion:*

$$L^\varepsilon \varphi(u,x) = \left[\varepsilon^{-1} Q + Q_0 \mathbb{A}(x) + \delta^\varepsilon(x)\right] \varphi(u,x),$$

where

$$\mathbb{A}(x)\varphi(u) := a(x)\varphi'(u),$$

and the negligible term is such that

$$\sup_{x\in E} \|\delta^\varepsilon(x)\varphi(u,x)\| \to 0, \qquad \varepsilon \to 0,\ \varphi(u,\cdot) \in C^2(\mathbb{R}).$$

Then a solution of the singular perturbation problem [1, Ch. 5] *may be used for the truncated operator*

$$L_0^\varepsilon \varphi(u,x) := \left[\varepsilon^{-1} Q + Q_0 \mathbb{A}(x)\right] \varphi(u,x). \tag{8}$$

Lemma 2.3. *The truncated operator* (8) *on a perturbed test function*

$$\varphi^\varepsilon(u,x) = \varphi(u) + \varepsilon\varphi_1(u,x),$$

admits the following asymptotic representation [1, Proposition 5.1]*:*

$$L_0^\varepsilon \varphi^\varepsilon(u,x) = \widehat{a}_0 \varphi'(u) + \delta^\varepsilon(x)\varphi(u).$$

The negligible term may be written in explicit form:

$$\delta^\varepsilon(x)\varphi(u) = \varepsilon Q_0 \mathbb{A}(x) R_0 \widehat{\mathbb{A}}(x)\varphi(u).$$

$$\widehat{\mathbb{A}}(x) := \widehat{\mathbb{A}}_0 - Q_0\mathbb{A}(x), \qquad \widehat{\mathbb{A}}_0 := \Pi Q_0 \mathbb{A}(x)\Pi.$$

Conclusion 2.1. *The generator* (6) *of the random evolution* (5) *admits the asymptotic representation*

$$L^\varepsilon \varphi^\varepsilon(u,x) = \widehat{a}_0 \varphi'(u) + \delta^\varepsilon(x)\varphi(u) \tag{9}$$

with the negligible term $\delta^\varepsilon(x)\varphi(u)$.

The representation (9) *implies the weak convergence* (3)–(4) [1, Ch. 6] *because the limit operator*

$$L^0\varphi(u) := \widehat{a}_0\varphi'(u), \qquad \varphi(u) \in C^1(\mathbb{R}), \tag{10}$$

defines the evolution

$$S^0(t) = u + \widehat{a}_0 t, \qquad t \geq 0,\ S^0(0) = u.$$

Remark 2.3. The limit operator (10) in the Euclidean space $\mathbb{R}^d$ has the following representation:

$$\widehat{a}_0\varphi'(u) := \sum_{k=1}^d \widehat{a}_k^0 \varphi_k'(u), \qquad \varphi_k'(u) := \partial\varphi(u)/\partial u_k,$$

$$\widehat{a}_k^0 = q\widehat{a}_k, \qquad \widehat{a}_k = \int_E \rho(dx) a_k(x), \qquad a_k(x) = \int_{\mathbb{R}} v_k \Phi_x(dv).$$

3. SIP in diffusion approximation scheme.

It is well known that the diffusion approximation of stochastic systems may be realized under some additional *Balance Condition (BC)*.

We consider two different BC for SIP, namely the *total* and *local* ones.

3.1. **SIP under total balance condition.** The SIP in the series scheme with the parameter $\varepsilon \to 0$, $\varepsilon > 0$, in the diffusion approximation scheme under the *Total Balance Condition (TBC)*:

$$a(x) = \int_{R^d} v\,\Phi_x(dv) \equiv 0, \tag{11}$$

is considered in the following scaling:

$$S^\varepsilon(t) = u + \varepsilon \sum_{n=1}^{\nu(t/\varepsilon^2)} \alpha_n(x_n), \qquad t \geq 0,\ \varepsilon > 0.$$

Proposition 3.1. *Under the TBC* (11), *the weak convergence*

$$S^\varepsilon(t) \Rightarrow W_\sigma(t), \qquad \varepsilon \to 0,$$

takes place.

The limit Brownian motion process $W_\sigma(t)$, $t \geq 0$, *is defined by the variance matrix*

$$\widehat{C} = \sigma^*\sigma = q\widehat{B},$$

$$\widehat{B} = \int_E \rho(dx) B(x), \qquad B(x) = \int_{\mathbb{R}^d} v^* v\,\Phi_x(dv).$$

Proof of Proposition 3.1. As in Section 2, we start by characterizing the coupled Markov process.

Lemma 3.1. *The Markov process*

$$S^\varepsilon(t),\ x^\varepsilon(t) := x\left(t/\varepsilon^2\right), \qquad t \geq 0,$$

is characterized by the generator

$$L^\varepsilon\varphi(u,x) = \varepsilon^{-2} q(x) \int_E P(x,dy) \int_{\mathbb{R}^d} \Phi_x(dv)[\varphi(u+\varepsilon v, y) - \varphi(u,x)]. \tag{12}$$

The generator (12) *may be rewritten as follows*

$$L^\varepsilon \varphi(u,x) = \varepsilon^{-2}[Q + Q_0 \varPhi_x^\varepsilon]\varphi(u,x), \tag{13}$$

where

$$\varPhi_x^\varepsilon \varphi(u) := \int_{\mathbb{R}^d} \Phi_x(dv)[\varphi(u+\varepsilon v) - \varphi(u)] = \varepsilon^2 \left[\frac{1}{2} B(x)\varphi''(u) + \delta^\varepsilon(x)\varphi(u)\right], \tag{14}$$

with the negligible term $\delta^\varepsilon(x)\varphi(u)$.

Lemma 3.2. *The generator* (13)–(14) *admits the asymptotic expansion*

$$L^\varepsilon \varphi(u,x) = \left[\varepsilon^{-2} Q + Q_0 \mathbb{B}(x)\right]\varphi(u,x) + \delta^\varepsilon(x)\varphi(u)$$

with negligible term $\delta^\varepsilon(x)\varphi(u)$. *Here by definition*

$$\mathbb{B}(x)\varphi(u) = \frac{1}{2} B(x)\varphi''(u). \tag{15}$$

Then the solution of singular perturbation problem [1, Ch. 5] *can be used for the truncated operator*

$$\mathbb{L}_0^\varepsilon \varphi(u,x) = \left[\varepsilon^{-2} Q + Q_0 \mathbb{B}(x)\right]\varphi(u,x). \tag{16}$$

Lemma 3.3. *The truncated operator (16) on a perturbed test function*

$$\varphi^\varepsilon(u,x) = \varphi(u) + \varepsilon^2 \varphi_2(u,x), \tag{17}$$

admits the asymptotic representation

$$L_0^\varepsilon \varphi^\varepsilon(u,x) = \frac{1}{2}\widehat{C}\varphi''(u) + \delta^\varepsilon(x)\varphi(u).$$

Proof. Considering (16) and (17),

$$\begin{aligned} L_0^\varepsilon \varphi^\varepsilon &= \left[\varepsilon^{-2} Q + Q_0 \mathbb{B}(x)\right]\left[\varphi(u) + \varepsilon^2 \varphi_2(u,x)\right] \\ &= \varepsilon^{-2} Q\varphi(u) + [Q\varphi_2(u,x) + Q_0 \mathbb{B}(x)\varphi(u)] + \delta^\varepsilon(x)\varphi(u). \end{aligned}$$

It is obvious

$$Q\varphi(u) = 0.$$

The equation

$$Q\varphi_2(u,x) + Q_0 \mathbb{B}(x)\varphi(u) = \widehat{L}_0 \varphi(u)$$

can be solved under the solvability condition [1, Ch.5]:

$$\widehat{L}_0 \Pi = \Pi Q_0 \mathbb{B}(x)\Pi.$$

Transforming (15) gives us

$$\widehat{L}_0 \varphi(u) = \frac{1}{2}\widehat{C}\varphi''(u).$$

Indeed

$$\begin{aligned} \widehat{L}_0 \varphi(u) &= \int_E \pi(dx) q(x) \int_E P(x,dy) \frac{1}{2} B(y)\varphi''(u) \\ &= \frac{1}{2} q \int_E \rho(dx) B(x)\varphi''(u) = \frac{1}{2} q \widehat{B}\varphi''(u). \end{aligned}$$

□

Remark 3.1. The limit generator $\widehat{L}_0$ in the Euclidean space $\mathbb{R}^d$ is represented as follows:

$$\widehat{L}_0 \varphi(u) = \frac{q}{2} \sum_{k,r=1}^{d} B_{kr} \varphi''_{kr}(u),$$

$$\widehat{B} = [B_{kr}; 1 \le k, r \le d], \qquad \varphi''_{kr}(u) := \partial^2 \varphi(u)/\partial u_k \partial u_r,$$

$$B_{kr} = \int_E \rho(dx) B_{kr}(x), \qquad B_{kr}(x) = \int_{\mathbb{R}} v_k v_r \, \Phi_x(dv).$$

The proof of Proposition 3.1 is finished by using the asymptotic representation

$$L^\varepsilon \varphi^\varepsilon(u,x) = \widehat{L}_0\varphi(u) + \delta^\varepsilon(x)\varphi(u), \tag{18}$$

and convergence Theorem 6.3 [1, Ch.6]. The negligible term in (18) may be written in the explicit form. □

3.2. **SIP under the Local Balance Condition (LBC).** The LBC means that the average value of jumps is such that

$$\widehat{a} := \int_E \rho(dx)a(x) \neq 0. \tag{19}$$

The SIP in the series scheme under the LBC (19) with the parameter $\varepsilon \to 0$, $\varepsilon > 0$, is considered in the following scaling:

$$S^\varepsilon(t) = u + \varepsilon \sum_{n=1}^{\nu(t/\varepsilon^2)} \alpha_n(x_n) - q\widehat{a}t/\varepsilon, \qquad t \geq 0. \tag{20}$$

Proposition 3.2. *Under the LBC* (19)*, the weak convergence*

$$S^\varepsilon(t) \Rightarrow W_\sigma(t), \qquad \varepsilon \to 0,$$

takes place.

The limit Brownian motion $W_\sigma(t)$, $t \geq 0$, is defined by the variance matrix

$$\begin{gathered}
\widehat{C} = \sigma^*\sigma = q\widehat{B}, \qquad \widehat{B} = \widehat{B}_0 + \widehat{B}_1, \\
\widehat{B}_0 = \int_E \rho(dx)B_0(x), \qquad B_0(x) = \int_{\mathbb{R}^d} v^*v\Phi_x(dv), \\
\widehat{B}_1 = \int_E \rho(dx)B_1(x), \qquad B_1(x) = 2\widehat{a}^*(x)R_0\widehat{a}(x), \\
\widehat{a}(x) := a_0(x) - q\widehat{a}, \\
a_0(x) := q(x)\int_E P(x,dy)a(y).
\end{gathered} \tag{21}$$

Here the potential operator R_0 is defined as the solution of the equation

$$QR_0 = R_0Q = \Pi - I$$

[1, Ch. 3].

Proof of Proposition 3.2. As in the previous section we start using the generator of the two component Markov process.

Lemma 3.4. *The two component Markov process $S^\varepsilon(t)$, $x^\varepsilon(t) := x(t/\varepsilon^2)$, $t \geq 0$, is characterized by the generator*

$$L^\varepsilon\varphi(u,x) = \varepsilon^{-2}q(x)\int_E P(x,dy)\int_{\mathbb{R}^d}\Phi(dv)[\varphi(u+\varepsilon v, y) - \varphi(u,x)] - \varepsilon^{-1}\widehat{a}_0\varphi'(u,x). \tag{22}$$

This generator can be written as follows

$$L^\varepsilon\varphi(u,x) = \left[\varepsilon^{-2}[Q + Q_0\Phi_x^\varepsilon] - \varepsilon^{-1}\widehat{\mathbb{A}}_0\right]\varphi(u,x) \tag{23}$$

with $\widehat{\mathbb{A}}_0\varphi(u) := \widehat{a}_0\varphi'(u)$

$$\begin{aligned}
\Phi_x^\varepsilon\varphi(u) &= \int_{\mathbb{R}^d}\Phi_x(dv)[\varphi(u+\varepsilon v) - \varphi(u)] \\
&= \varepsilon a(x)\varphi'(u) + \varepsilon^2\frac{1}{2}B(x)\varphi''(u) + \varepsilon^2\delta^\varepsilon(x)\varphi(u).
\end{aligned}$$

Lemma 3.5. *The generator* (22) *admits the asymptotic expansion*

$$L^\varepsilon \varphi(u,x) = \left[\varepsilon^{-2} Q + \varepsilon^{-1} \widehat{\mathbb{A}}(x) + Q_0 \mathbb{B}(x)\right] \varphi(u,x) + \delta^\varepsilon(x)\varphi(u,x).$$

Here

$$\begin{aligned} \widehat{\mathbb{A}}(x)\varphi(u) &= \widehat{a}(x)\varphi'(u), \\ \widehat{a}(x) &:= a_0(x) - \widehat{a}_0, \\ a_0(x) &:= q(x)\int_E P(x,dy)a(y). \end{aligned} \tag{24}$$

Note that the following balance condition

$$\Pi \widehat{a}(x) = 0 \tag{25}$$

takes place.

Now a solution of singular perturbation problem [1, Ch.5] can be used for the truncated operator

$$L_0^\varepsilon \varphi(u,x) = \left[\varepsilon^{-2} Q + \varepsilon^{-1} \widehat{\mathbb{A}}(x) + Q_0 \mathbb{B}(x)\right] \varphi(u,x). \tag{26}$$

Lemma 3.6. *The truncated operator* (26) *on a perturbed test function*

$$\varphi^\varepsilon(u,x) = \varphi(u) + \varepsilon\varphi_1(u,x) + \varepsilon^2 \varphi_2(u,x)$$

admits the asymptotic representation

$$L_0^\varepsilon \varphi^\varepsilon(u,x) = \frac{1}{2}\widehat{C}\varphi''(u) + \delta^\varepsilon(x)\varphi(u).$$

Proof. Let us consider

$$\begin{aligned} L_0^\varepsilon \varphi^\varepsilon(u,x) &= [\varepsilon^{-2} Q + \varepsilon^{-1}\widehat{\mathbb{A}}(x) + Q_0\mathbb{B}(x)][\varphi(u) + \varepsilon\varphi_1(u,x) + \varepsilon^2\varphi_2(u,x)] \\ &= \varepsilon^{-2} Q\varphi(u) + \varepsilon^{-1}[Q\varphi_1 + \widehat{\mathbb{A}}(x)\varphi] + [Q\varphi_2 + \widehat{\mathbb{A}}(x)\varphi_1 + Q_0\mathbb{B}(x)\varphi] \\ &\quad + \delta^\varepsilon(x)\varphi(u). \end{aligned}$$

We get the equations

$$\begin{aligned} Q\varphi(u) &= 0, \\ Q\varphi_1(u,x) + \widehat{\mathbb{A}}(x)\varphi(u) &= 0, \\ Q\varphi_2(u,x) + \widehat{\mathbb{A}}(x)\varphi_1(u,x) + Q_0\mathbb{B}(x)\varphi(u) &= \widehat{L}_0\varphi(u). \end{aligned}$$

The first equation is obvious. The second equation satisfies the solvability condition (25). Hence

$$\varphi_1(u,x) = R_0 \widehat{\mathbb{A}}(x)\varphi(u).$$

Now the third equation is

$$Q\varphi_2 + \left[\widehat{\mathbb{A}}_0(x) + Q_0\mathbb{B}(x)\right]\varphi(u) = \widehat{L}_0\varphi(u), \tag{27}$$

where

$$\widehat{\mathbb{A}}_0(x)\varphi(u) := \widehat{\mathbb{A}}(x) R_0 \widehat{\mathbb{A}}(x)\varphi(u). \tag{28}$$

The solvability condition for (27) gives

$$\widehat{L}_0\Pi = \Pi\left[\widehat{\mathbb{A}}_0(x) + Q_0\mathbb{B}(x)\right]\Pi.$$

Using (28), (24), and (15) we calculate the limit generator

$$\widehat{L}_0\varphi(u) = \frac{1}{2}\widehat{C}\varphi''(u),$$

where the variance matrix $\widehat{C}$ is represented in (21).

Note that (see (24))

$$\widehat{\mathbb{A}}_0(x)\varphi(u) = \widehat{\mathbb{A}}(x)R_0\widehat{\mathbb{A}}(x)\varphi(u) = \widehat{\mathbb{A}}(x)R_0\widehat{a}(x)\varphi'(u) = \widehat{a}(x)R_0\widehat{a}(x)\varphi''(u) = \frac{1}{2}B_1(x)\varphi''(u).$$

Here

$$\widehat{a}(x) = a_0(x) - \widehat{a}_0. \qquad \square \quad \square$$

4. Large deviation in the scheme of asymptotically small diffusion

The SIP in the scheme of asymptotically small diffusion is considered under two different balance conditions, namely total and local ones.

4.1. **The SIP under the total balance condition.** The total balance condition means that the mean values of jumps of SIP equal totaly zero:

$$a(x) = \int_{\mathbb{R}^d} v\,\Phi_x(dv) \equiv 0. \tag{29}$$

The SIP in the scheme of asymptotically small diffusion is considered in the following scaling [3]:

$$S^\varepsilon(t) = u + \varepsilon^2 \sum_{n=1}^{\nu(t/\varepsilon^3)} \alpha_n(x_n), \qquad t \ge 0,\ \varepsilon > 0,\ u \in \mathbb{R}^d. \tag{30}$$

The coupled Markov process

$$S^\varepsilon(t),\ x^\varepsilon(t) := x\left(t/\varepsilon^3\right), \qquad t \ge 0,$$

is defined by the generator

$$L^\varepsilon\varphi(u,x) = \varepsilon^{-3}q(x)\int_E P(x,dy)\int_{\mathbb{R}^d}\Phi_y(dv)\left[\varphi(u+\varepsilon^2 v, y) - \varphi(u,x)\right],$$

which can be rewritten as follows

$$L^\varepsilon\varphi(u,x) = \varepsilon^{-3}[Q + Q_0\Phi_x^\varepsilon]\varphi(u,x), \tag{31}$$

where, by definition,

$$\Phi_x^\varepsilon\varphi(u) := \int_{\mathbb{R}^d}\Phi_x(dv)\left[\varphi\left(u+\varepsilon^2 v\right) - \varphi(u)\right] = \varepsilon^4[\mathbb{B}(x)\varphi(u) + \delta^\varepsilon(x)\varphi(u)].$$

Here

$$\mathbb{B}(x)\varphi(u) := \frac{1}{2}B(x)\varphi''(u).$$

Hence the generator (31) admits the asymptotic expansion

$$\begin{aligned} L^\varepsilon\varphi(u,x) &= L_0^\varepsilon\varphi(u,x) + \delta^\varepsilon(x)\varphi(u,x), \\ L_0^\varepsilon\varphi(u,x) &= \left[\varepsilon^{-3}Q + \varepsilon Q_0\mathbb{B}(x)\right]\varphi(u,x). \end{aligned} \tag{32}$$

The truncated operator (32) on a perturbed test function

$$\varphi^\varepsilon(u,x) = \varphi(u) + \varepsilon^4\varphi_1(u,x),$$

admits the asymptotic representation

$$L_0^\varepsilon\varphi^\varepsilon(u,x) = \varepsilon[Q\varphi_1 + Q_0\mathbb{B}(x)\varphi(u)] + \delta^\varepsilon(x)\varphi(u). \tag{33}$$

The representations (32) and (33) give

$$L^\varepsilon\varphi^\varepsilon(u,x) = \varepsilon\left[\widehat{\mathbb{C}}\varphi(u) + \delta^\varepsilon(x)\varphi(u,x)\right],$$

where the main part

$$\varepsilon\widehat{\mathbb{C}}\varphi(u) = \varepsilon\frac{1}{2}\widehat{C}\varphi''(u)$$

is the generator of a small diffusion.

4.2. **Large deviation for SIP under the total balance condition.** We investigate the large deviation problem for SIP by using the asymptotic analysis of the exponential generator of large deviation

$$H^\varepsilon \varphi(u,x) = e^{-\varphi/\varepsilon} \varepsilon L^\varepsilon e^{\varphi/\varepsilon} \tag{34}$$

[2, Part I].

Proposition 4.1. *The large deviation for SIP* (30) *under the total balance condition* (29) *is realized by the exponential generator of small diffusion*

$$H\varphi(u) = \frac{1}{2}\widehat{C}[\varphi'(u)]^2, \tag{35}$$

$$\widehat{C} = q\int_E \rho(dx)B(x), \qquad B(x) = \int_{\mathbb{R}^d} v^* v\, \Phi_x(dv).$$

Proof of Proposition 4.1.

Lemma 4.1. *The exponential generator* (34) *on a perturbed test function*

$$\varphi^\varepsilon(u,x) = \varphi(u) + \varepsilon \ln\left[1 + \varepsilon^2 \varphi_1(u,x)\right]$$

admits the asymptotic representation

$$H^\varepsilon \varphi^\varepsilon(u,x) = Q\varphi_1 + \frac{1}{2} Q_0 B(x)[\varphi'(u)]^2 + h^\varepsilon(x)\varphi(u)$$

with the negligible term

$$\|h^\varepsilon(x)\varphi(u)\| \to 0, \qquad \varepsilon \to 0,\ \varphi(u) \in C^3(\mathbb{R}).$$

Proof of Lemma 4.1. Let us calculate

$$\begin{aligned}
H^\varepsilon \varphi^\varepsilon &= e^{-\varphi/\varepsilon}\left[1+\varepsilon^2\varphi_1\right]^{-1} \varepsilon L^\varepsilon [1+\varepsilon^2\varphi_1] e^{\varphi/\varepsilon} \\
&= e^{-\varphi/\varepsilon}\left[1-\varepsilon^2\varphi_1\right] \varepsilon L_0^\varepsilon [1+\varepsilon^2\varphi_1] e^{\varphi/\varepsilon} + h^\varepsilon(x)\varphi(u) \\
&= e^{-\varphi/\varepsilon}\left[1-\varepsilon^2\varphi_1\right] \varepsilon^{-2} Q \left[1+\varepsilon^2\varphi_1\right] e^{\varphi/\varepsilon} + e^{-\varphi/\varepsilon}\varepsilon^{-2} Q_0 \Phi_x^\varepsilon e^{\varphi/\varepsilon} + h^\varepsilon(x)\varphi(u) \\
&= Q\varphi_1 + \frac{1}{2} Q_0 B(x)[\varphi'(u)]^2 + h^\varepsilon(x)\varphi(u).
\end{aligned}$$

□

Now the solution of the singular perturbation problem [1, Ch.5] gives

$$H^\varepsilon \varphi^\varepsilon(u,x) = H\varphi(u) + h^\varepsilon(x)\varphi(u). \tag{36}$$

The asymptotic representation (36) completes the proof of Proposition 4.1. □

Remark 4.1. The exponential generator of small diffusion (35) in the Euclidean space $\mathbb{R}^d$, $d \geq 2$, is represented as follows:

$$H\varphi(u) = \frac{1}{2}\varphi'^*(u)\widehat{C}\varphi'(u),$$

where $\varphi'^*(u) = (\varphi_k'(u), 1 \leq k \leq d)$ is a vector-row, $\varphi'(u) = (\varphi_k'(u), 1 \leq k \leq d)$ is a vector-column, $\widehat{C} = [\widehat{C}_{kr;1\leq k,r\leq d}]$ is the variance matrix.

4.3. **Large deviation for SIP under the local balance condition.** *The Local Balance Condition (LBC)* means that the average value of jumps is not equal to zero:

$$\widehat{a} := \int_E \rho(dx)a(x) \neq 0. \tag{37}$$

The SIP under LBC (37) is considered in the following scaling:

$$S^\varepsilon(t) = u + \varepsilon^2 \sum_{n=1}^{\nu(t/\varepsilon^3)} \alpha_n(x_n) - \widehat{a}_0 t/\varepsilon. \tag{38}$$

Lemma 4.2. *The coupled Markov process $S^\varepsilon(t)$, $x^\varepsilon(t) := x(t/\varepsilon^3)$, $t \geq 0$, is determined by the generator (compare* (22)*)*

$$L^\varepsilon\varphi(u,x) = \varepsilon^{-3}q(x)\int_E P(x,dy)\int_{\mathbb{R}^d}\Phi_y(dv)\left[\varphi(u+\varepsilon^2 v,y)-\varphi(u,x)\right] - \varepsilon^{-1}\widehat{a}_0\varphi'_u(u,x).$$

Or, in a different form,

$$L^\varepsilon\varphi(u,x) = \left[\varepsilon^{-3}[Q + Q_0\Phi^\varepsilon_x] - \varepsilon^{-1}\widehat{\mathbb{A}}_0\right]\varphi(u,x),$$

$$\Phi^\varepsilon_x\varphi(u) = \int_{\mathbb{R}^d}\Phi_x(dv)\left[\varphi(u+\varepsilon^2 v)-\varphi(u)\right].$$

Proposition 4.2. *The large deviation for SIP* (38) *under the LBC* (37) *is realized by the exponential generator of small diffusion*

$$H\varphi(u) = \frac{1}{2}\widehat{C}[\varphi'(u)]^2, \tag{39}$$

$$\widehat{C} = q[\widehat{B}_1 + \widehat{B}_2],$$

$$\widehat{B}_k = \int_E \rho(dx)B_k(x), \qquad k = 1,2, \tag{40}$$

$$B_1(x) = \int_{\mathbb{R}^d} v^* v\Phi_x(dv), \qquad B_2(x) = 2\widehat{a}(x)R_0\widehat{a}(x),$$

$$\widehat{a}(x) = a_0(x) - \widehat{a}_0, \qquad a_0(x) := q(x)\int_E P(x,dy)a(x).$$

The exponential generator of large deviation (39)–(40) *contains two components. One of them is the variance matrix of the second moment of jumps. The second component $\widehat{B}_2$ is defined by the fluctuation of the first moment of jumps.*

Proof of Proposition 4.2. To prove the proposition we need the following lemma:

Lemma 4.3. *The exponential generator* (34) *under the local balance condition* (37) *on the perturbed test function*

$$\varphi^\varepsilon(u,x) = \varphi(u) + \varepsilon\ln\left[1 + \varepsilon\varphi_1(u,x) + \varepsilon^2\varphi_2(u,x)\right]$$

admits the asymptotic representation

$$\begin{aligned} H^\varepsilon\varphi^\varepsilon(u,x) &= \varepsilon^{-1}\left[Q\varphi_1 + \widetilde{\mathbb{A}}(x)\varphi(u)\right] \\ &\quad + \left[Q\varphi_2 - \varphi_1 Q\varphi_1 + \frac{1}{2}Q_0 B(x)[\varphi'(u)]^2\right] + h^\varepsilon(x)\varphi(u) \end{aligned} \tag{41}$$

with the negligible term

$$\|h^\varepsilon(x)\varphi(u)\| \to 0, \qquad \varepsilon \to 0,\ \varphi(u) \in C^3(\mathbb{R}).$$

Proof. Proof of Lemma 4.3 is based on the following asymptotic representations:

$$H^\varepsilon_Q\varphi^\varepsilon(u,x) := e^{-\varphi^\varepsilon/\varepsilon}\varepsilon^{-2}Qe^{\varphi^\varepsilon/\varepsilon} = \varepsilon^{-1}Q\varphi_1 + [Q\varphi_2 - \varphi_1 Q\varphi_1] + h^\varepsilon_q(x)\varphi(u),$$
$$H^\varepsilon_\varphi\varphi^\varepsilon(u,x) := e^{-\varphi^\varepsilon/\varepsilon}\varepsilon^{-2}Q_0\Phi^\varepsilon_x e^{\varphi^\varepsilon/\varepsilon} = \varepsilon^{-1}Q_0\mathbb{A}(x)\varphi(u) + Q_0\mathbb{A}(x)\varphi_1(u,x) + h^\varepsilon_\varphi(x)\varphi(u),$$
$$H^\varepsilon_a\varphi^\varepsilon(u,x) := e^{-\varphi^\varepsilon/\varepsilon}\widehat{\mathbb{A}}_0 e^{\varphi^\varepsilon/\varepsilon} = \varepsilon^{-1}\widehat{a}_0\varphi'(u) + h^\varepsilon_a(x)\varphi(u).$$

Thus, the relation

$$H^\varepsilon\varphi^\varepsilon(u,x) = [H^\varepsilon_Q + H^\varepsilon_\varphi - H^\varepsilon_a]\varphi^\varepsilon(u,x)$$

gives (41) with (see (24)–(25))

$$\widetilde{\mathbb{A}}(x)\varphi(u) := \widetilde{a}(x)\varphi'(u),$$
$$\widetilde{a}(x) := Q_0 a(x) - \widehat{a}_0. \qquad \square$$

Now the solution of the singular perturbation problem [1, Ch. 5] may be used for the equations

$$\begin{aligned} Q\varphi_1 + \widetilde{A}(x)\varphi(u) = 0, \Pi\widetilde{A}(x) = 0;\\ Q\varphi_2 - \varphi_1 Q\varphi_1 + \frac{1}{2}B_1(x)[\varphi'(u)]^2 = \widehat{H}\varphi(u). \end{aligned} \tag{42}$$

The first equation in (42) has the solution

$$\varphi_1(u,x) = R_0\widetilde{a}(x)\varphi'(u), \qquad Q\varphi_1 = \widetilde{a}(x)\varphi'(u).$$

Hence, the second equation in (42) may be rewritten as follows

$$Q\varphi_2 + \frac{1}{2}[B_1(x) + B_2(x)][\varphi'(u)]^2 = \widehat{H}\varphi(u)$$

with $B_2(x)$ given in (40).

The solvability condition [1, Ch. 5] for the last equation gives Proposition 4.2. $\square$

References

1. V. S. Koroliuk and N. Limnios, *Stochastic Systems in Marging Phase Space*, VSP, 2005.
2. J. Feng and T. G. Kurtz, *Large Deviations for Stochastic Processes*, Mathematical Surveys and Monographs, vol. 131, AMS, RI, 2006.
3. A. A. Mogulskii, *Large deviations for processes with independent increments*, Ann. Prob. **21** (1993), 202–215.
4. M. J. Freidlin and A. D. Wentzel, *Random Perturbation of Dynamical Systems*, Springer–Verlag, 1998.

Current address: IM, Kyiv, Ukraine
E-mail address: vskorol@imath.kiev.ua

University of Rome "La Sapienza", Italy
E-mail address: raimondo.manca@uniroma1.it

Universita'di Chieti, Chieti, Italy
E-mail address: g.damico@unich.it

Received 20/10/2012

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 75–90

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 75–90

LIFT ZONOID ORDER AND FUNCTIONAL INEQUALITIES

UDC 519.21

ALEXEI M. KULIK AND TARAS D. TYMOSHKEVYCH

Abstract. We introduce the notion of a weighted lift zonoid and show that the ordering condition on a measure μ, formulated in terms of the weighted lift zonoids of this measure, leads to certain functional inequalities for this measure, such as non-linear extensions of Bobkov's shift inequality and weighted inverse log-Sobolev inequality. The choice of the weight K, involved in our version of the inverse log-Sobolev inequality, differs substantially from those available in the literature, and requires the weight v, involved into the definition of the weighted lift zonoid, to equal the divergence of the weight K w.r.t. initial measure μ. We observe that such a choice may be useful for proving direct log-Sobolev inequality, as well.

Анотація. Введено поняття зваженого ліфт зоноїда та показано, що умова порядку на міру μ, накладена у термінах зважених ліфт зоноїдів цієї міри, приводить до таких функціональних нерівностей на цю міру, як нелінійне узагальнення нерівності зсуву Бобкова та зваженої оберненої логарифмічної нерівності Соболєва. Вибір ваги K у нашій версії оберненої логарифмічної нерівності Соболєва істотно відрізняється від наявних у літературі, та вимагає, щоб вага v з означення зваженого ліфт зоноїда дорівнювала дивергенції ваги K відносно вихідної міри μ. Ми показуємо, що такий вибір також може бути корисним при доведенні прямої логарифмічної нерівності Соболєва.

Аннотация. Введено понятие взвешенного лифт зоноида и показано, что условие порядка на меру μ, наложенное в терминах взвешенных лифт зоноидов этой меры, приводит к таким функциональным неравенствам для этой меры, как нелинейное обобщение неравенства сдвига Бобкова и взвешенного обратного логарифмического неравенства Соболева. Выбор веса K в нашей версии обратного логарифмического неравенства Соболева существенно отличается от имеющихся в литературе, и требует, чтобы вес v из определения взвешенного лифт зоноида был равен дивергенции веса K относительно исходной меры μ. Мы показываем, что такой выбор также может быть полезным при доказательстве прямого логарифмического неравенства Соболева.

1. Introduction

The notions of *zonoid* and *lift zonoid*, introduced in [9], have a diverse field of applications. Because the lift zonoid determines the underlying measure uniquely, this concept can be used in multivariate statistics for measuring the variability of laws of random vectors, and for ordering these laws, see [10]. The concept of *zonoid equivalence* appears to be both naturally motivated by financial applications, and useful for proving extensions of the ergodic theorem for zonoid stationary and zonoid swap-invariant random sequences, see [12, 13]. Lift zonoids lead naturally to definitions of associated α-trimming and data depth, see [9] and [7], and to barycentric representation of the points of a space with a given measure, see [9] and [11].

In this paper, we explore a new field, where the notion of lift zonoid can be applied naturally. As a straightforward extension of the definition of lift zonoid, we introduce a *weighted lift zonoid* $\hat{Z}^v(\mu)$ with a vector-valued weight function v. We show that, for properly chosen weights v, the ordering condition on a measure μ, formulated in terms

2010 *Mathematics Subject Classification.* Primary 26D10, 39B62, 47D07, 60E15, 60J60.

Key words and phrases. Lift zonoid, weight, shift inequality, log-Sobolev inequality.

The second author was partially supported by the Leonard Euler program, DAAD project # 55518603.

of the weighted lift zonoid of this measure, leads to certain functional inequalities for this measure, such as non-linear extensions of Bobkov's *shift inequality* [3] and weighted *inverse log-Sobolev inequality*. Weighted versions of the classical functional inequalities (Poincaré, log-Sobolev, etc) have been studied recently in various contexts. The choice of the weight K, involved in our version of inverse log-Sobolev inequality, is specific and differs substantially from those available in the literature. This choice is strongly motivated by (an extension of) the functional form of Bobkov's shift inequality, and requires the weight v, involved into the definition of the weighted lift zonoid, to equal the *divergence* of the weight K w.r.t. initial measure μ. We observe that such a choice may be useful for proving (weighted) direct log-Sobolev inequality, as well. In the case of a bounded weight, this may lead to new sufficient conditions for the log-Sobolev inequality. We illustrate the range of applications of these conditions in two examples in Section 4.

2. Weighted lift zonoids, non-linear shift inequalities, and weighted inverse log-Sobolev inequalities

Let μ be a probability measure on the Borel σ-algebra in $\mathbb{R}^d$, and $v\colon \mathbb{R}^d \to \mathbb{R}^d$ be a measurable function such that

$$\int_{\mathbb{R}^d} \|v(x)\|\,\mu(dx) < \infty;$$

here and below we denote by $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^d$. We define the *weighted zonoid* $Z^v(\mu)$ with the weight v as the set of all the points in $\mathbb{R}^d$ of the form

$$\int_{\mathbb{R}^d} g(x)v(x)\,\mu(dx) \tag{1}$$

with arbitrary Borel measurable $g\colon \mathbb{R}^d \to [0,1]$. The *weighted lift zonoid* $\hat{Z}^v(\mu)$ is defined as the weighted zonoid of the measure $\delta_1 \times \mu$ in $\mathbb{R}^{d+1}$. Equivalently, the weighted zonoid $Z^v(\mu)$ and the weighted lift zonoid $\hat{Z}^v(\mu)$ are the sets of the points of the form

$$\mathsf{E}\,g(X)v(X) \in \mathbb{R}^d \quad \text{and} \quad (\mathsf{E}\,g(X), \mathsf{E}\,g(X)v(X)) \in \mathbb{R}^{d+1} \tag{2}$$

respectively, where X is a random vector with the distribution μ. This definition is a straightforward generalization of the definitions of the zonoid and the lift zonoid (see [10], Definition 2.1), where the function v has the form $v(x) = x$.

The lift zonoid $\hat{Z}(\mu)$ is a convex compact set in $\mathbb{R}^{d+1}$, symmetric w.r.t. the point $(\frac{1}{2}, \frac{1}{2}\,\mathsf{E}\,X)$, which identifies the underlying measure μ uniquely; see [10]. On the other hand, it can be seen easily that the definition of the weighted lift zonoid $\hat{Z}^v(\mu)$ would not change if one restricts the class of Borel measurable functions g within it to the class of the functions of the form

$$g(x) = G(v(x)), \qquad G\colon \mathbb{R}^d \to [0,1] \text{ is Borel measurable.}$$

This observation leads immediately to the identity $\hat{Z}^v(\mu) = \hat{Z}(\mu \circ v^{-1})$; that is, the weighted lift zonoid $\hat{Z}^v(\mu)$ equals the (usual) lift zonoid of the image of the measure μ under the mapping v. As a corollary, we get that the weighted lift zonoid $\hat{Z}^v(\mu)$ is a convex compact set in $\mathbb{R}^{d+1}$, symmetric w.r.t. the point $((1/2), (1/2)Ev(X))$, and identifies the image measure $\mu \circ v^{-1}$ uniquely.

The following theorem motivates the above definition of the weighted lift zonoid. To formulate it, we need to introduce some notation. Denote by γ_c the centered Gaussian measure in $\mathbb{R}^d$ with the covariance matrix $c^2 I_{\mathbb{R}^d}$. Let

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \qquad \Phi(x) = \int_{-\infty}^{x} \varphi(y)\,dy, \quad x \in \mathbb{R},$$

be the standard Gaussian distribution density function and the standard Gaussian cumulative distribution function, respectively, and let

$$I(p) = \varphi(\Phi^{-1}(p)), \qquad p \in (0,1), \quad I(0) = I(1) = 0, \tag{3}$$

be the *Gaussian isoperimetric function.*

For any measurable f on $\mathbb{R}^d$, we write $\mathbf{E}_\mu f$ for the integral of f w.r.t. μ; function f may be vector-valued, then the integral is understood in the component-wise sense. For a function f taking values in $\mathbb{R}^+$, its μ-entropy is defined by

$$\mathbf{Ent}_\mu f = \mathbf{E}_\mu(f \log f) - (\mathbf{E}_\mu f) \log(\mathbf{E}_\mu f),$$

with the convention $0 \log 0 = 0$.

In what follows, we assume that the measure μ has *the logarithmic gradient* v_μ; that is, a function $v_\mu \colon \mathbb{R}^d \to \mathbb{R}$, integrable w.r.t. μ and such that for every smooth $f \colon \mathbb{R}^d \to \mathbb{R}$ with a compact support

$$\mathbf{E}_\mu \nabla f = - \mathbf{E}_\mu (v_\mu f). \tag{4}$$

This assumption is equivalent to the following, see Proposition 3.4.3 in [6]: there exists the density p_μ of the measure μ w.r.t. the Lebesgue measure, which belongs to the Sobolev class $W_{1,1}(\mathbb{R}^d)$; in this case

$$[v_\mu]_i = \frac{\partial_{x_i} p_\mu}{p_\mu}, \qquad i = 1, \ldots, d.$$

Theorem 1. *I. The following three statements are equivalent.*

A. $\widehat{Z}^{v_\mu}(\mu) \subset \widehat{Z}(\gamma_c)$.

B. *For any smooth function* $f \colon \mathbb{R}^d \to [0,1]$ *with a compact support, one has*

$$\| \mathbf{E}_\mu \nabla f \| \le c I(\mathbf{E}_\mu f). \tag{5}$$

C. *For any* $h \in \mathbb{R}^d, A \in \mathcal{B}(\mathbb{R}^d)$

$$\Phi\left(\Phi^{-1}(\mu(A)) - c\|h\|\right) \le \mu(A+h) \le \Phi\left(\Phi^{-1}(\mu(A)) + c\|h\|\right). \tag{6}$$

*II. Under the conditions **A**–**C** above, the following inverse log-Sobolev inequality holds true: for any smooth function* $f \colon \mathbb{R}^d \to [0, \infty)$ *with a compact support,*

$$\| \mathbf{E}_\mu \nabla f\|^2 \le 2c \, \mathbf{Ent}_\mu f \, \mathbf{E}_\mu f. \tag{7}$$

Remark 1. By the definition (see Definition 5.1 in [10]), two measures μ_1 and μ_2 are related by the *lift zonoid order* (notation: $\mu_1 \preccurlyeq_{LZ} \mu_2$), if

$$\hat{Z}(\mu_1) \subset \hat{Z}(\mu_2).$$

Recall that $\widehat{Z}^{v_\mu}(\mu)$ equals the lift zonoid of $\nu_\mu := \mu \circ v_\mu^{-1}$; that is, of the distribution of the logarithmic gradient of the measure μ. Hence statement **A** can be equivalently formulated as follows: the distribution ν_μ of the logarithmic gradient of the measure μ is dominated in the sense of the lift zonoid order by the canonical Gaussian measure in $\mathbb{R}^d$.

Theorem 1 is not a genuinely new one. The equivalence of the relations **B** and **C** is used by S. Bobkov in [3] as a key ingredient in the proof of the *shift inequality* (6) (in [3], the measure μ is supposed to be a product-measure, but the proof of the equivalence of (5) and (6) in fact does not rely on this assumption). The outline of the proof of (7) under (5) and (6) is given in [2]. What we would like to emphasize is that condition **B**, usually called the *functional version of the shift inequality*, is equivalent to the relation **A**, which according to Remark 1 can be written as the lift zonoid order relation

$$\nu_\mu \preccurlyeq_{LZ} \gamma_c. \tag{8}$$

It is instructive to compare (8) with the following necessary and sufficient condition for the functional version of the shift inequality to hold, given in [3] in the case where the measure μ is a product-measure with equal marginals μ_1. This condition states that there exists $c > 0$ such that (5) holds true, if and only if there exists $\varepsilon > 0$ such that

$$\int_{\mathbb{R}} e^{\varepsilon x^2} \nu_{\mu_1}(dx) \leq 2; \tag{9}$$

in addition, the optimal constant c in (5) and ε in (9) are connected by the relation

$$\frac{1}{\sqrt{6\varepsilon}} \leq c \leq \frac{4}{\sqrt{\varepsilon}}. \tag{10}$$

For the product measure $\mu(dx) = \prod_{i=1}^d \mu_1(dx_i)$, respective distribution of the logarithmic gradient is again a product measure

$$\nu_\mu(dx) = \prod_{i=1}^d \nu_{\mu_1}(dx_i),$$

and in this case, due to Corollary 5.3 in [10], (8) is equivalent to

$$\nu_{\mu_1} \preccurlyeq_{LZ} \gamma_c^1, \tag{11}$$

where γ_c^1 is the $\mathcal{N}(0, c^2)$-Gaussian measure on $\mathbb{R}$. Both (9) and (11) are conditions on the tails of the distribution of the logarithmic gradient of μ_1, but (11) is more precise because it involves the same c with (5).

The main result of this section, Theorem 2 below, is a generalization of Theorem 1 and is motivated by an observation that in Theorem 1 the equivalence of the relations **A** and **B** follows in a very straightforward way from the integration-by-parts formula (4); see the proof of Theorem 2 below. With this observation in mind, we introduce a wide class of weights which admit an analogue of the integration-by-parts formula (4). To do that, we recall that the μ-divergence of a function $g\colon \mathbb{R}^d \to \mathbb{R}^d$, if exists, is defined as the function $\delta_\mu(g) \in L_1(\mathbb{R}^d, \mu)$ such that for every smooth $f\colon \mathbb{R}^d \to \mathbb{R}$ with a compact support

$$\mathbf{E}_\mu(\nabla f, g)_{\mathbb{R}^d} = \mathbf{E}_\mu\, f\delta_\mu(g).$$

The μ-divergence is well defined, for instance, for any $g \in C^1$ bounded together with its partial derivatives; in this case,

$$\delta_\mu(g) = -\sum_{i=1}^d [v_\mu]_i g_i - \sum_{i=1}^d \partial_{x_i} g_i.$$

This follows directly from (4); see [6], Chapter 6 for more information on this subject. Let function $v\colon \mathbb{R}^d \to \mathbb{R}$ be such that, for some function K taking values in $d\times d$-matrices,

$$v_i = \delta_\mu(K_i), \qquad i = 1, \dots, d, \tag{12}$$

where K_i denotes the i-th row of the matrix K. Then for every smooth f with a compact support

$$\mathbf{E}_\mu(K\nabla f) = \mathbf{E}_\mu\, fv; \tag{13}$$

here and below we treat elements of $\mathbb{R}^d$ as vectors-columns. Formula (13) is a straightforward extension of the integration-by-parts formula (4), where the gradient ∇ is replaced by the "weighted gradient" $K\nabla$ with the matrix-valued weight K, and the logarithmic gradient v_μ is replaced by the μ-divergence of K. Furthermore, if K satisfies some extra regularity condition, e.g.

$$K\colon \mathbb{R}^d \to \mathbb{R}^{d\times d} \quad \text{is Lipschitz}, \tag{14}$$

then for every $h \in \mathbb{R}^d$ there exists a flow of solutions $\{\Psi_t^{K,h}(x), t \in \mathbb{R}, x \in \mathbb{R}^d\}$ of the Cauchy problem

$$d\Psi_t(x) = (K^*h)(\Psi_t(x))\, dt, \qquad \Psi_0(x) = x. \tag{15}$$

Theorem 2. *I. Let $v = (v_i)_{i=1}^d$ satisfy* (12). *Then the following two statements are equivalent.*

A1. $\widehat{Z}^v(\mu) \subset \widehat{Z}(\gamma_c)$.

B1. *For any smooth function $f\colon \mathbb{R}^d \to [0,1]$ with a compact support, one has*

$$\| \mathbf{E}_\mu K\nabla f\| \le cI(\mathbf{E}_\mu f). \tag{16}$$

If, in addition, the matrix-valued function K satisfies (14)*, then* ***A1*** *and* ***B1*** *are equivalent to the following.*

C1. *For any $h \in \mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$*

$$\Phi\left(\Phi^{-1}(\mu(A)) - c\|h\|\right) \le \mu\left(\left[\Psi_1^{K,h}\right]^{-1}(A)\right) \le \Phi\left(\Phi^{-1}(\mu(A)) + c\|h\|\right). \tag{17}$$

II. Under the condition ***A1****, equivalently* ***B1****, the following weighted inverse log-Sobolev inequality holds true: for any smooth function $f\colon \mathbb{R}^d \to [0,\infty)$ with a compact support,*

$$\| \mathbf{E}_\mu K\nabla f\|^2 \le 2c^2\, \mathbf{Ent}_\mu f\, \mathbf{E}_\mu f. \tag{18}$$

Note that condition **A1** is just the lift zoonoid order relation for the image measure of μ under v:

$$\mu\circ v^{-1} \preccurlyeq_{LZ} \gamma_c. \tag{19}$$

Before giving the proof of Theorem 2, let us summarize: a lift zonoid order condition (8) is a *criterion* for the shift inequality, written either in its explicit form (6), or in its functional form (5). This equivalence is rather flexible in the following sense: if the logarithmic gradient v_μ in (8) is replaced by another weight v of the form

$$v = \delta_\mu(K) \tag{20}$$

(see (12)), then respective lift zonoid order condition (19) is still equivalent to the weighted version (16) of the functional form of a (generalized) shift inequality. The explicit form of the (generalised) shift inequality in that case is available as well, and concerns, instead of linear shifts, the transformations of the initial measure μ by the flows of solutions to (15).

Proof of Theorem 2: statement I. The lift zonoid $\hat{Z}(\gamma)$ of a standard Gaussian measure γ in $\mathbb{R}^d$ can be identified in the following way: for a given $\alpha \in (0,1)$, the section of $\hat{Z}(\gamma)$ by the hyper-plane $\{\alpha\}\times\mathbb{R}^d$ has the projection on the last d coordinates equal to the ball centered at 0 and having the radius $I(\alpha)$; see [9], Section 6.3 or [11], Proposition 3.4. It is easy to see from the definition of the lift zonoid that

$$\hat{Z}(\gamma_c) = c\hat{Z}(\gamma).$$

Hence condition **A1** can be equivalently written as follows: for every Borel measurable $g\colon \mathbb{R}^d \to [0,1]$ such that $\mathbf{E}_\mu\, g = \alpha$,

$$\|\mathbf{E}_\mu(gv)\| \le cI(\alpha) = cI(\mathbf{E}_\mu\, g).$$

By the standard approximation argument, the above condition is equivalent to a similar one with Borel measurable g's replaced by smooth and compactly supported f's. Because for such f by (13)

$$\|\mathbf{E}_\mu(fv)\| = \|\mathbf{E}_\mu(K\nabla f)\|,$$

conditions **A1** and **B1** are equivalent.

The proof of the equivalence of **B1** and **C1** follows the same lines with the S.Bobkov's proof from [3] for the case of product measures and linear shifts; to make the exposition self-sufficient here we expose the key steps of this proof.

Denote $R_r(p) = \Phi(\Phi^{-1}(p)+r)$, $r\ge 0$, $p\in(0,1)$. Then the following properties hold true:

- for every $r \geq 0$ the function R_r is concave;
- the family $\{R_r, r \geq 0\}$ is a semigroup w.r.t. the composition of the functions, i.e.

$$R_{r_1} \circ R_{r_2} = R_{r_1 + r_2};$$

- the function R_0 is an identity, and the "generator" of the semigroup $\{R_r, r \geq 0\}$ equals the Gaussian isoperimetric function I in the sense that

$$\frac{R_r(p) - p}{r} \to I(p), \qquad r \to 0+.$$

Similarly, the family of functions $S_r(p) = \Phi(\Phi^{-1}(p) - r)$, $r \geq 0$, $p \in (0,1)$ has the following properties:

- for every $r \geq 0$ the function S_r is convex;
- the family $\{S_r, r \geq 0\}$ is a semigroup w.r.t. the composition of the functions;
- the function S_0 is an identity, and the "generator" of the semigroup $\{S_r, r \geq 0\}$ equals $(-I)$.

Observe that **C1** is equivalent to the following.

C2. For any $h \in \mathbb{R}^d$ and Borel measurable $f\colon \mathbb{R}^d \to [0,1]$

$$S_{c\|h\|}(\mathbf{E}_\mu f) \leq \mathbf{E}_\mu \left(f \circ \Psi_1^{K,h} \right) \leq R_{c\|h\|}(\mathbf{E}_\mu f). \tag{21}$$

Indeed, taking $f = \mathbb{1}_A$ we get **C1** from **C2**. Inversely, under **C1** by the concavity of R_r and Jensen's inequality we have

$$\mathbf{E}_\mu \left(f \circ \Psi_1^{K,h} \right) = \int_0^\infty \mu \left(\left\{ x\colon f\left(\Psi_1^{K,h}(x)\right) \geq t \right\} \right) dt \leq \int_0^\infty R_{c\|h\|} \left(\mu(\{x : f(x) \geq t\}) \right) dt$$
$$\leq R_{c\|h\|} \left(\int_0^\infty \mu(\{x\colon f(x) \geq t\})\, dt \right) = R_{c\|h\|}(\mathbf{E}_\mu f).$$

The proof of the left hand side inequality in (21) is similar and omitted. Hence **C1** and **C2** are equivalent.

To get **B1** from **C2**, take th instead of h and differentiate the right hand side inequality in (21) w.r.t. t at the point $t = 0$. In more details, denote $f_t(x) = f\big(\Psi_1^{K,th}(x)\big)$, then

$$f_t(x) = f\left(\Psi_t^{K,h}(x)\right),$$

and therefore there exits a continuous derifative

$$\partial_t f_t(x) = \left((\nabla f)\left(\Psi_t^{K,h}(x)\right), (K^* h)\left(\Psi_t^{K,h}(x)\right) \right)_{\mathbb{R}^d}.$$

Because f is smooth and compactly supported and K satisfies (14), this derivative is bounded as a function of $(t,x) \in [0,T] \times \mathbb{R}^d$ for every fixed T. Therefore by the dominated convergence theorem

$$\frac{1}{t}\left(\mathbf{E}_\mu f_t - \mathbf{E}_\mu f\right) \to \mathbf{E}_\mu \left(\nabla f, K^* h\right)_{\mathbb{R}^d} = \left(\mathbf{E}_\mu K \nabla f, h\right)_{\mathbb{R}^d}, \qquad t \to 0+.$$

Because

$$\frac{1}{t}(R_{ct\|h\|}(\mathbf{E}_\mu f) - \mathbf{E}_\mu f) = c\|h\| I(\mathbf{E}_\mu f),$$

we get from (21)

$$(\mathbf{E}_\mu K \nabla f, h)_{\mathbb{R}^d} \leq c\|h\| I(\mathbf{E}_\mu f), \qquad h \in \mathbb{R}^d.$$

Taking sup over all h with $\|h\| = 1$, we get (16).

To get **C2** from **B1**, consider first the case where f is smooth and compactly supported and such that $0 < \mathbf{E}_\mu f < 1$. By (16), for a given $h \in \mathbb{R}^d$ we have that

$$\left.\frac{d}{dt}\right|_{t=0} \mathbf{E}_\mu f_t = (\mathbf{E}_\mu K \nabla f, h)_{\mathbb{R}^d} \leq c\|h\| I(\mathbf{E}_\mu f).$$

Recall that

$$\frac{d}{dt}\bigg|_{t=0} R_{c\|h\|t}(\mathbf{E}_\mu f) = c\|h\| I(\mathbf{E}_\mu f).$$

Therefore for every $\varrho > 1$ there exists $\delta = \delta(f) > 0$ such that for every $t \in (0, \delta)$:

$$\mathbf{E}_\mu f_t \le R_{\varrho c\|h\|t}(\mathbf{E}_\mu f). \tag{22}$$

Note that if $t_1 \in (0, \delta(f))$ and $t_2 \in (0, \delta(f_{t_1}))$, then

$$\mathbf{E}_\mu f_{t_1+t_2} = \mathbf{E}_\mu \left(f_{t_1} \circ \Psi_{t_2}^{K,h} \right) \le R_{\varrho c\|h\|t_1}(\mathbf{E}_\mu f_{t_1}) \le R_{\varrho c\|h\|(t_1+t_2)}(\mathbf{E}_\mu f); \tag{23}$$

here we have used the flow property of $\{\Psi_t^{K,h}, t \in \mathbb{R}\}$, the semigroup property of $\{R_r, r \ge 0\}$, and monotonicity of R_r. Because the derivative $\partial_t f_t$ is uniformly continuous w.r.t. $(t, x) \in [0, T] \times \mathbb{R}^d$ for every fixed T, it can be shown that

$$\delta_T = \inf_{t \in [0,T]} \delta(f_t) > 0.$$

Then, applying (23) at most T/δ_T times, we get that (22) holds true for every $t \in [0, T]$. Consequently, (22) holds true for every $t \in \mathbb{R}^+$ and ϱ therein can be replaced by 1. This gives the right hand side inequality in (21) for smooth and compactly supported f such that $0 < \mathbf{E}_\mu f < 1$. By an approximation argument, this can be extended to any measurable $f : \mathbb{R}^d \to [0, 1]$. The proof of the left hand side inequality in (21) is completely analogous and omitted. □

Proof of Theorem 2: statement II. The following lemma is a straightforward extension of a part of Proposition 2 in [2] (the one which states the equivalence of $P_1(c)$ and $P_2(c)$ in the notation of [2]).

Lemma 1. *Statement **B1** is equivalent to the following.*

B2.: *For any smooth function $f : \mathbb{R}^d \to [0, 1]$ with a compact support, one has*

$$\sqrt{(\mathbf{E}_\mu I(f))^2 + \frac{1}{c^2} \| \mathbf{E}_\mu K \nabla f \|^2} \le I(\mathbf{E}_\mu f). \tag{24}$$

The proof is completely analogous to the one from [2], therefore we just sketch it. The implication **B2** $\Rightarrow$ **B1** is trivial. To get the inverse implication, recall first that the standard Gaussian measure γ^d on $\mathbb{R}^d$ satisfies **B2** with $c = 1$ and identity matrix K; see [2], Section 2. Consider a smooth function $f : \mathbb{R}^d \to [0, 1]$ with a compact support, and let $F(r) = \mu(\{x : f(x) \le r\})$ be its distribution function w.r.t. μ. Assume that F is absolutely continuous w.r.t. Lebesque measure on $\mathbb{R}$, and take $r \in \mathbb{R}$, $\varepsilon > 0$. Define $\psi_\varepsilon(x) = \mathbb{I}_{[0,r]}(x) + (1 - \frac{x-r}{\varepsilon})\mathbb{I}_{[r,r+\varepsilon]}(x)$. Applying **B1** to the function $g = \psi_\varepsilon(f)$ and tending $\varepsilon \to 0$, we get

$$F'(r)\|\theta(r)\| \le cI(F(r)) \quad \text{for} \quad \mu \circ f^{-1}\text{-a.a. } r \in \mathbb{R}, \tag{25}$$

where $\theta(r) = E_\mu(K\nabla f | f = r)$. Denote $k = F^{-1} \circ \Phi$, then k transforms the standard Gaussian measure γ^1 on $\mathbb{R}$ to $\mu \circ f^{-1}$. Taking the derivative in the identity $F(k) = \Phi$, we get $k' F'(k) = \varphi$. Then from (25) with $r = k(x)$ we get inequality

$$\frac{1}{c} \|\theta(k(x))\| \le k'(x) \tag{26}$$

valid γ^1-a.s. We have already mentioned that a standard Gaussian measure satisfies **B2** with $c = 1$ and identity K; for the case $d = 1$ this can be written as

$$\sqrt{\left(\int_{\mathbb{R}} I(g) d\gamma^1 \right)^2 + \left(\int_{\mathbb{R}} g' d\gamma^1 \right)^2} \le I \left(\int_{\mathbb{R}} g d\gamma^1 \right).$$

Applying this inequality to $g = k$ and using (26) we get

$$\sqrt{\left(\int_0^1 I(r)dF(r)\right)^2 + \left(\int_0^1 \frac{1}{c}\|\theta(r)\|dF(r)\right)^2} \le I\left(\int_0^1 rdF(r)\right); \tag{27}$$

here we have took into account that the image of γ^1 under k is $\mu \circ f^{-1}$, and $\mu \circ f^{-1}$ is supported in $[0,1]$. Using the inequality

$$\int_0^1 \|\theta(r)\|dF(r) \ge \left\|\int_0^1 \theta(r)dF(r)\right\| = \|\,\mathbf{E}_\mu K\nabla f\|,$$

we complete the proof of the required statement. The additional assumption of $\mu \circ f^{-1}$ to be absolutely continuous can be removed by an approximation argument. □

According to Lemma 1, to prove statement II of Theorem 2 it is enough to show that **B2** implies (18) for any non-negative smooth compactly supported f. Take ε small, then εf takes values in $[0,1]$ and one can apply **B2**. After trivial transformations, we get

$$\frac{1}{c^2}\|\,\mathbf{E}_\mu K\nabla f\|^2 \le \frac{I^2(\varepsilon\,\mathbf{E}_\mu f) - (\mathbf{E}_\mu I(\varepsilon f))^2}{\varepsilon^2}.$$

Hence the required statement would follow from the relation

$$\lim_{\varepsilon\to 0+} \frac{I^2(\varepsilon\,\mathbf{E}_\mu f) - (\mathbf{E}_\mu I(\varepsilon f))^2}{\varepsilon^2} = 2\,\mathbf{Ent}_\mu f\,\mathbf{E}_\mu f. \tag{28}$$

This relation can be proved straightforwardly using the following asymptotic expansion:

$$I(\varepsilon) = \varepsilon\sqrt{2\log\frac{1}{\varepsilon}} - \frac{\varepsilon\log(2\log\frac{1}{\varepsilon})}{2\sqrt{2\log\frac{1}{\varepsilon}}} + \frac{\varepsilon}{\sqrt{2\log\frac{1}{\varepsilon}}} + \frac{\varepsilon\kappa(\varepsilon)}{\sqrt{2\log\frac{1}{\varepsilon}}}, \tag{29}$$

where $\kappa(\varepsilon) \to 0$, $\varepsilon \to 0+$; the detailed exposition is straightforward but cumbersome and therefore is omitted. The asymptotic expansion (29) follows from the standard expansion

$$\Phi(t) = -\frac{1}{t}\varphi(t) + \frac{1}{t^3}\varphi(t) + O\left(t^{-5}\varphi(t)\right), \qquad t \to -\infty,$$

which holds true e.g. by the integration-by-parts formula. □

Remark 2. The above proof of statement II follows, in main lines, the one sketched in [2] (the proof of the implication $P_3(c) \Rightarrow P_6(c\sqrt{2})$ in Proposition 2), where the authors referred to Beckner's lectures at the Institut Henri Poincaré. However, instead of using the equivalence

$$I(\varepsilon) \sim \varepsilon\sqrt{2\log\frac{1}{\varepsilon}}, \qquad \varepsilon \to 0,$$

which apparently is not sufficient to provide (28), we use stronger asymptotic expansion (29).

Let us mention that a more explicit condition, sufficient for the lift zonoid relation (19) tohold true, can be given in a way similar to (9).

Proposition 1. *There exists* $c > 0$ *such that* (19) *holds true, if and only if, there exists* $\varepsilon > 0$ *such that*

$$\mathbf{E}_\mu\, e^{\varepsilon(v,h)^2_{\mathbb{R}^d}} \le 2, \qquad \|h\| \le 1. \tag{30}$$

The optimal constant c *in* (19) *and* ε *in* (30) *are connected by the relation* (10).

Because the lift zonoid order relation is equivalent to the same relation for all one-dimensional projections (see Section 5 in [10]), statement of Proposition 1 follow immediately from the one-dimensional statement given below.

Lemma 2. *For a measure ν on $\mathbb{R}$ there exists $c > 0$ such that*

$$\nu \preccurlyeq_{LZ} \gamma_c$$

with $\gamma_c \sim \mathcal{N}(0, c)$ if, and only if, there exists $\varepsilon > 0$ such that

$$\int_{\mathbb{R}} e^{\varepsilon x^2}\, dx \le 2;$$

in that case, the optimal constants c, ε are connected by the relation (10).

The proof of Lemma 2 is contained, in fact, in the proof of Lemma 4.1 in [3], hence we omit it here.

At the end of this section, let us indicate one further research possibility related to the above results. In [4], an approach is proposed, making it possible to give explicit bounds for ergodic rates of solutions to Lévy driven SDE's, which has a wide range of further applications e.g. to limit theorems for functionals of such processes, see [14]–[16]. The key ingredient of this approach is a stochastic control based on perturbations of time coordinates of jumps of the Lévy noise. A natural question is whether such an approach remains practical when perturbations of jump amplitudes are used instead, which is typical in the stochastic calculus of variations for processes with jumps. In this context, it would be helpful to bound from below the size of the absolutely continuous part of the image of the Lévy measure of the noise under a non-linear mapping which corresponds to the perturbation of the noise. The above results seemingly can be useful here, because shift inequalities yield upper bounds for the size of singular component of the image of a measure: respective result was obtained in [3] in the context of linear shift inequalities (6), and can be extended easily to non-linear shift inequalities (17).

3. Weighted log-Sobolev inequalities in $\mathbb{R}$

Theorem 2 above gives a sufficient condition for a weighted inverse log-Sobolev inequality, based on a pair of functions v and K related by (20). The main result of this section, Theorem 3 below, shows that the use of the same pair may lead to sufficient conditions for the (direct) log-Sobolev inequality, either in a weighted or in a classical form. What is surprising is that, even in the simplest one-dimensional case, Theorem 3 leads to new sufficient conditions for the log-Sobolev inequality, when compared with those available in a literature; see below Proposition 2, Proposition 3, and two examples in Section 4. We believe that the reason for that is a proper choice of the *pair* of the weight functions v and K, involved in (31) and connected by (20).

Theorem 3. *Let $d = 1$ and functions v and K be related by* (20). *Assume that for some $\alpha > 0$*

$$Kv' \ge \alpha. \tag{31}$$

Assume in addition that the functions K and

$$a := 2KK' + K^2 v_\mu \tag{32}$$

belong to C^∞, have at most linear growth at ∞, and all their derivatives have at most polynomial growth at ∞.

Then for every smooth f with a compact support

$$\mathbf{Ent}_\mu f^2 \le \frac{2}{\alpha}\, \mathbf{E}_\mu (Kf')^2. \tag{33}$$

As a corollary, if K is bounded then μ satisfies the (classical) log-Sobolev inequality: for every absolutely continuous f such that both f and f' are square integrable w.r.t. μ,

$$\mathbf{Ent}_\mu f^2 \le \frac{2}{\alpha} \left(\sup_x K^2(x) \right) \mathbf{E}_\mu (f')^2. \tag{34}$$

Remark 3. The proof of Theorem 3 is based on the classic Bakry–Emery criterion; see below. We strongly believe that similar technique is applicable in the multidimensional case as well, but because of possible non-commutativity of matrix-valued weights which appear therein, now we can not give a multidimensional version of Theorem 3; this is a subject for a further research.

Remark 4. The additional assumptions on the functions K and a to be smooth and to satisfy certain growth bounds, in particular cases, can be removed by an approximation procedure; see e.g. Propositions 2 and 3 below.

Proof of Theorem 3. Consider a Markov process X defined as the strong solution to the SDE

$$dX_t = a(X_t)\,dt + \sqrt{2}K(X_t)\,dW_t;$$

see (32) for the formula for the coefficient a. Then on the Schwartz space $\mathcal{S}(\mathbb{R})$ of C^∞ functions s.t. all their derivatives decay at ∞ faster than any polynomial, the generator L of the process X has the form

$$Lf = af' + bf'' = v_\mu f' + (bf')', \qquad b := K^2.$$

By the construction, the measure μ is a symmetric measure for the semigroup $\{T_t\}$ generated by the process X:

$$\mathbf{E}_\mu\, fT_tg = \mathbf{E}_\mu\, gT_tf, \qquad t \geq 0;$$

in particular,

$$\mathbf{E}_\mu\, T_tf = \mathbf{E}_\mu\, f, \qquad t \geq 0,$$

i.e. μ is an invariant measure for X. The class $\mathcal{G} = \mathcal{S}(\mathbb{R})$ is an algebra, invariant w.r.t. superpositions with C^∞-functions and dense in every $L_p(\mu)$, $p \geq 1$. In addition, thanks to the smoothness conditions and growth bounds imposed on coefficients a and K, the class $\mathcal{G}$ is invariant w.r.t. the semigroup T_t and the generator L. Define for $f, g \in \mathcal{G}$

$$\Gamma(f,g) = \frac{1}{2}(L(fg) - fLg - gLf), \qquad \Gamma_2(f,g) = \frac{1}{2}(L\Gamma(f,g) - \Gamma(Lf,g) - \Gamma(f,Lg)).$$

We will prove that

$$\Gamma_2(f,f) \geq \alpha\Gamma(f,f), \qquad f \in \mathcal{G}, \tag{35}$$

then the required statement would follow from the Bakry–Emery criterion [1].

Straightforward calculations give

$$\Gamma(f,g) = bf'g',$$

$$\begin{aligned} 2\Gamma_2(f,f) &= (ab' + bb'' - 2a'b)(f')^2 - 2bb'f'f'' + 2b^2(f'')^2 \\ &= \left(ab' + bb'' - 2a'b - \frac{(b')^2}{2}\right)(f')^2 + \left(\frac{b'f'}{\sqrt{2}} - bf''\sqrt{2}\right)^2 \\ &\geq \left(ab' + bb'' - 2a'b - \frac{(b')^2}{2}\right)(f')^2. \end{aligned}$$

Hence to prove (35) it is enough to show that

$$2ab' + 2bb'' - 4a'b - (b')^2 \geq 4\alpha b. \tag{36}$$

Recall that

$$v = \delta_\mu(K) = -Kv_\mu - K',$$

hence we can express the coefficients a and b through the functions K and v:

$$a = KK' - Kv, \qquad b = K^2.$$

Substituting these expressions into (36), after some transformations, which are straightforward but cumbersome and therefore omitted, we re-write (36) to the following form:

$$K^3 v' \geq \alpha K^2.$$

The last inequality clearly holds true under (31). Hence, applying the Bakry–Emery criterion, we get (33) for every $f \in \mathcal{S}(\mathbb{R})$.

If K is bounded, then for every $f \in \mathcal{S}(\mathbb{R})$ (34) holds true as a corollary of (33). It is a standard procedure to approximate a given absolutely continuous f such that $f, f' \in L_2(\mu)$ by a sequence of smooth compactly supported f_n in such a way that $f_n \to f$ and $f_n' \to f'$ in $L_2(\mu)$; see e.g. the proof of Corollary 2.6.10 in [6]. Passing to the limit in (34) for f_n, $n \geq 1$, we complete the proof. □

There is a wide choice for the pair of functions v and K related by (20). Below we give two versions of Theorem 3 which correspond to particular choices of this pair. The first one arise when one just takes $v(x) = x - \langle \mu \rangle$,

$$\langle \mu \rangle = \int_{\mathbb{R}} y \, \mu(dy).$$

Proposition 2. *Let measure μ on $\mathbb{R}$ have the first absolute moment and have a positive continuous distribution density p_μ. Denote*

$$\bar{K}_\mu(x) = \frac{1}{p_\mu(x)} \int_x^\infty (y - \langle \mu \rangle) p_\mu(y) \, dy, \qquad x \in \mathbb{R}.$$

The following statements hold true.

I. *If* $\inf_x \bar{K}_\mu(x) = \alpha > 0$, *then for every smooth f with a compact support*

$$\mathbf{Ent}_\mu f^2 \leq \frac{2}{\alpha} \mathbf{E}_\mu \left(\bar{K}_\mu f' \right)^2.$$

II. *If, in addition,* $\sup_x \bar{K}_\mu(x) = \beta < \infty$, *then for every absolutely continuous f such that both f and f' are square integrable w.r.t. μ,*

$$\mathbf{Ent}_\mu f^2 \leq 2 \bar{c}_\mu \, \mathbf{E}_\mu (f')^2$$

with

$$\bar{c}_\mu = \frac{\beta^2}{\alpha}.$$

In the second version of Theorem 3, we choose K in a more intrinsic way, namely, we take K such that $\delta_\mu(K) = v$ with

$$v = \Phi^{-1}(F_\mu), \qquad F_\mu(x) = \mu((-\infty, x]), \tag{37}$$

then $\mu \circ v^{-1} = \gamma$, $\gamma \sim \mathcal{N}(0,1)$. Such a choice of the weight v is motivated by our intent to have

$$\hat{Z}^v(\mu) = \hat{Z}(\gamma);$$

that is, to make the order condition (19) with $c = 1$ as precise as it is possible, i.e. to replace an inequality by an identity. Because $\hat{Z}^v(\mu) = Z(\mu \circ v^{-1})$ identifies the law of $\mu \circ v^{-1}$ uniquely, such an intent naturally leads to the formula (37).

Proposition 3. *Let measure μ on $\mathbb{R}$ have a positive continuous distribution density p_μ. Denote*

$$\hat{K}_\mu(x) = \frac{I(F_\mu(x))}{p_\mu(x)}.$$

The following statements hold true.

I. *For every smooth f with a compact support,*

$$\mathbf{Ent}_\mu f^2 \leq 2 \, \mathbf{E}_\mu (\hat{K}_\mu f')^2. \tag{38}$$

II. *If, in addition, $\hat{K}_\mu$ is bounded, then for every absolutely continuous f such that both f and f' are square integrable w.r.t. μ,*

$$\mathbf{Ent}_\mu f^2 \le 2\hat{c}_\mu \, \mathbf{E}_\mu (f')^2$$

with

$$\hat{c}_\mu = \sup_x (\hat{K}_\mu(x))^2.$$

Remark 5. Define the *isoperimetric function* of the measure μ by

$$I_\mu(p) = p_\mu\left(F_\mu^{-1}(p)\right), \qquad p \in (0,1), \qquad I_\mu(0) = I_\mu(1) = 1.$$

Then, clearly, the function I defined by (3) equals I_γ, $\gamma \sim \mathcal{N}(0,1)$. The function $\hat{K}_\mu(x)$ above can be expressed as the ratio

$$\left.\frac{I_\gamma(p)}{I_\mu(p)}\right|_{p=F_\mu(x)},$$

and under the conditions of Proposition 3 the function F_μ gives a one-to-one correspondence between $(-\infty,\infty)$ and $(0,1)$. Hence the constant $\hat{c}_\mu$ above can be alternatively expressed as

$$\hat{c}_\mu = \left(\sup_{p\in(0,1)} \frac{I_\gamma(p)}{I_\mu(p)}\right)^2.$$

Proofs of Proposition 2 and Proposition 3. If $v(x) = x - \langle\mu\rangle$, we have $\bar{K}_\mu v' = \bar{K}_\mu$, and therefore the assumption $\inf \bar{K}_\mu = \alpha > 0$ made in Proposition 2 implies the principal condition (31). For the function v defined by (37) and the function $\hat{K}_\mu$, this condition takes even a more simple form because straightforward calculation shows that

$$\hat{K}_\mu v' = 1.$$

Hence one can expect that statements of Proposition 2 and Proposition 3 would follow from the version of the Bakry–Emery criterion given in Theorem 3. However, we can not apply this theorem here directly, because of extra smoothness and growth conditions on functions K and a, imposed therein. The strategy of the proof will be the following: first, we consider a family of measures, which approximate μ properly and satisfy both (31) for the respective pair of K and v, and extra smoothness and growth conditions on respective functions K and a. Then, by passing to a limit, we get respective weighted log-Sobolev inequality, i.e. prove statements I in Propositions 2, 3. Finally, using the same approximation procedure as in the proof of Theorem 3 above, we extend the class of f in the case where the weight K is bounded.

To shorten the exposition, we explain in details the way this strategy is implemented for the proof of Proposition 3, only. The detailed proof of Proposition 2 is similar and omitted. We also does not repeat the approximation arguments from the proof of Theorem 3 above, and concentrate on the proof of (38) for smooth compactly supported f.

Consider first the following auxiliary case: $p_\mu \in C^\infty$, and for some $R > 0$

$$p_\mu(x) = \varphi(x), \qquad |x| \ge R. \tag{39}$$

Then v_μ (which, let us recall, equals p'_μ/p_μ) and $\hat{K}_\mu$ belong to C^∞ and

$$v_\mu(x) = -x, \qquad \hat{K}_\mu(x) = 1, \qquad |x| \ge R.$$

Then the functions $K = \hat{K}_\mu$ and a defined by (32) satisfy the assumptions of Theorem 3. Hence, applying Theorem 3, we get (38).

Next, consider the general case. Fix some function $\chi \in C^\infty$ taking values in $[0,1]$, such that $\chi(0) = 0$, $\chi(x) = 1$, $x \ge 1$, and define

$$\varphi_{r,\delta}(x) = \varphi(x)(\delta + (1-\delta)\chi(|x| + r)), \qquad x \in \mathbb{R};$$

then every $\varphi_{r,\delta}$, $r>0$, $\delta\geq 0$ belongs to C^∞. Denote

$$M(r)=\int_{\mathbb{R}}\varphi_{r,0}(x)\,dx,$$

then M is a strictly decreasing function on $[0,\infty)$ and $M(0)<1$. For a given $Q>0$, consider the restriction p_μ^Q of p_μ to the segment $[-Q,Q]$, and assume that Q is large enough for

$$\int_{|x|>Q}p_\mu(x)\,dx<M(0).$$

Then for every δ small enough there exists unique $r=r(Q,\delta)>0$ such that

$$\int_{\mathbb{R}}\left(p_\mu^Q(x)+\varphi_{r,\delta}(x)\right)\,dx=1.$$

Take some non-negative $\psi\in C^\infty$, supported in $[-1,1]$ and such that $\int_{\mathbb{R}}\psi(x)\,dx=1$, and consider the probability measure $\mu_{Q,\varepsilon}$ with the density

$$p_{\mu_{Q,\delta}}(x)=\frac{1}{\delta}\int_{[-\delta,\delta]}p_\mu^Q(y)\psi\left(\frac{x-y}{\delta}\right)\,dy+\varphi_{r,\delta}(x).$$

By the construction, every $\mu_{Q,\delta}$ has positive C^∞ density and satisfy (39) for some large R. Therefore, (38) holds true with $\mu_{Q,\delta}$ instead of μ. It can be seen easily that

$$p_{\mu_{Q,\delta}}\to p_\mu,\qquad K_{\mu_{Q,\delta}}\to K_\mu,\qquad \delta\to 0,\ Q\to\infty,$$

uniformly on every finite segment. Passing to the limit, we obtain (38) for the initial measure μ and arbitrary smooth and compactly supported f. □

4. Examples

Example 1. Let μ on $\mathbb{R}$ have a positive C^1-density p_μ, such that for some $a,R>0$

$$v_\mu(x)x\geq -ax^2,\qquad |x|>R \tag{40}$$

Let us show that then condition $\inf\bar{K}_\mu>0$ from Proposition 2 holds true. Changing the variables $x\mapsto x-\langle\mu\rangle$, we can restrict ourselves to the case of $\langle\mu\rangle=0$. Then we have for $x>R$

$$\begin{aligned}\bar{K}_\mu(x)&=\int_x^\infty y\exp(\log p_\mu(y)-\log p_\mu(x))\,dy\\&=\int_x^\infty y\exp\left(\int_x^y v_\mu(z)\,dz\right)\,dy\geq\int_x^\infty y\exp\left(-a\int_x^y z\,dz\right)\,dy\\&=e^{ax^2/2}\int_x^\infty ye^{-ay^2/2}\,dy=1/a.\end{aligned}$$

Similar relation holds true for $x<-R$; to see this, one should note that

$$\bar{K}_\mu=-\frac{1}{p_\mu(x)}\int_{-\infty}^x yp_\mu(y)\,dy$$

because μ is centered. Finally, because $p_\mu\in C^1$ is positive, $\bar{K}_\mu$ has positive infimum over $[-R,R]$, which completes the proof.

Similarly, if in addition for some $b>0$

$$v_\mu(x)x\leq -bx^2,\qquad |x|>R, \tag{41}$$

then $\sup\bar{K}_\mu<\infty$. Hence, by statement II of Proposition 2, for a measure μ satisfying (40) and (41) the log-Sobolev inequality holds true.

Note that (41) is just the well known drift condition, sufficient for the Poincaré inequality, e.g. Theorem 3.1 and Remark 3.2 in [8]. However, various sufficient conditions for the log-Sobolev inequality, available in the literature, typically require additional assumptions on the *curvature*, which in the current context equals $-v'_\mu$. Namely, the famous Bakry–Emery condition ([1])) requires $-v'_\mu \geq \delta > 0$; conditions by Wang ([17]) and Cattiaux–Guillin ([8], Theorem 5.1) are more flexible, but still contain a requirement that the curvature is bounded from below, i.e. in our case

$$-v'_\mu \geq \delta \tag{42}$$

with some $\delta \in \mathbb{R}$. The above condition (40) can be understood as an "integral" version of (42), and it is easy to give an example of measure μ satisfying (40) and (41) such that (42) fails.

Example 2. Let γ^3 be a standard Gaussian measure on $\mathbb{R}^3$, and B_R be a ball of radius R, touching the origing and with the center located at the first basis vector e_1; that is, $B_R = B(Re_1, R)$. Denote by $\gamma^{3,R}$ the measure γ^3 conditioned outside the ball B_R:

$$\gamma^{3,R}(A) = \frac{\gamma^3(A \setminus B_R)}{\gamma^3(\mathbb{R}^3 \setminus B_R)}.$$

Consider a measure μ_R on $\mathbb{R}$ which is a projection of $\gamma^{3,R}$ on the first coordinate. We will show that there exists some constant $\hat{c}$ such that uniformly by $R \geq 0$ the constants $\hat{c}_\mu$ for the measures $\mu = \mu_R$ from Proposition 3 are dominated by $\hat{c}$. This would yield that for the family μ_R, $R \geq 0$ the log-Sobolev inequality holds true with uniformly bounded constants.

For a given $x \in [0, 2R]$, the section of the ball B_R by the hyperplane

$$\{y = (y_1, y_2, y_3) \colon y_1 = x\},$$

projected on the last two coordinates, is the ball in $\mathbb{R}^2$, centered at the origin and having the radius

$$r_R(x) = \sqrt{2Rx - x^2}.$$

Define $r(x) = 0$ for $x \notin [0, 2R]$. Then we have for $\mu = \mu_R$

$$p_\mu(x) = C_R \varphi(x) \psi_2(r_R(x)),$$

where

$$C_R = \left(\gamma^3(\mathbb{R}^3 \setminus B_R)\right)^{-1},$$

$$\psi_2(r) = \int_{\|y\| \geq r} \frac{1}{2\pi} e^{-(y_1^2+y_2^2)/2}\, dy_1\, dy_2 = \frac{1}{2\pi} \int_0^{2\pi} \int_r^\infty e^{-\rho^2/2} \rho\, d\rho\, d\theta = e^{-r^2/2}.$$

Consequently,

$$p_\mu(x) = \frac{C_R}{\sqrt{2\pi}} \begin{cases} e^{-Rx}, & x \in [0, 2R], \\ e^{-x^2/2}, & \text{otherwise}. \end{cases} \tag{43}$$

To bound $\hat{K}_\mu(x)$ consider separately three cases.

I. $x < 0$. Recall that $I'(p) = -\Phi^{-1}(p)$. Then for any $c > 1$ we have

$$[I(c\Phi(x))]' = -\Phi^{-1}(c\Phi(x)) c\varphi(x) \leq (-x) c\varphi(x) = c\varphi'(x)$$

because Φ^{-1} is an increasing function. Clearly, both $I(c\Phi(x))$ and $\varphi(x)$ vanish as $x \to -\infty$, hence

$$I(c\Phi(x)) = \int_{-\infty}^x [I(c\Phi(y))]'\, dy \leq c \int_{-\infty}^x \varphi'(y)\, dy = c\varphi(x), \qquad x \leq \Phi^{-1}(1/c). \tag{44}$$

Note that for $x < 0$

$$F_\mu(x) = C_R \Phi(x), \qquad p_\mu(x) = C_R \varphi(x),$$

and $C_R > 1$. In addition, the half-space $\{y = (y_1, y_2, y_3) \colon y_1 \leq x\}$ is contained in $\mathbb{R}^3 \setminus B_R$, hence

$$\Phi(x) = \gamma^3(\{y = (y_1, y_2, y_3) : y_1 \leq x\}) \leq \frac{1}{C_R} \Leftrightarrow x \leq \Phi^{-1}\left(\frac{1}{C_R}\right),$$

and we can apply (44) to get

$$\hat{K}_\mu(x) = \frac{I(C_R\Phi(x))}{C_R\varphi(x)} \le 1, \qquad x < 0.$$

II. $x > 2R$. In this case $1 - F_\mu(x) = C_R(1 - \Phi(x))$. Recall that $I(p) = I(1-p)$ and $\Phi^{-1}(1 - \Phi(x)) = -x$, hence we can use the same argument as in the case **I** to show that $\hat{K}_\mu(x) \le 1$, because for any $c > 1$

$$I(c(1 - \Phi(x)) = -c\int_x^\infty \Phi^{-1}\big(c(1 - \Phi(y))\big)\, dy \le c\int_x^\infty y\varphi'(y)\, dy = c\varphi(x).$$

III. $x \in [0, 2R]$. Recall that there exists a constant c_* such that

$$I(p) \le c_* p\sqrt{\log\frac{1}{p}}, \qquad p \in \left(0, \frac{1}{2}\right).$$

One has

$$C_R\gamma^3(\{y = (y_1, y_2, y_3)\colon y_1 > R\}) \le 1 - F_\mu(x) \le C_R\gamma^3(\{y = (y_1, y_2, y_3)\colon y_1 > 0\}) < \frac{1}{2},$$

hence we can write, using the identity $I(p) = I(1-p)$,

$$\hat{K}_\mu(x) = \frac{I(1 - F_\mu(x))}{p_\mu(x)} \le c_* \frac{1 - F_\mu(x)}{p_\mu(x)}\sqrt{\log\frac{1}{1 - F_\mu(x)}}.$$

Because $C_R > 1$, we have

$$\log\frac{1}{1 - F_\mu(x)} \le \log\frac{1}{1 - F_\mu(2R)} = \log\frac{1}{C_R(1 - \Phi(2R))} \le \log\frac{1}{1 - \Phi(2R)} \le c^*(1 + R)^2$$

with some $c^* > 2$. By (43), we have

$$\frac{1 - F_\mu(x)}{p_\mu(x)} = e^{Rx}\left(\int_x^{2R} e^{-Ry}\, dy + \int_{2R} e^{-y^2/2}\, dy\right),$$

and the right hand side term can be estimated either by

$$e^{Rx}\int_x^\infty e^{-Ry}\, dy = \frac{1}{R},$$

(when R is large), or by

$$e^{2R^2}\int_0 e^{-y^2/2}\, dy = \sqrt{\frac{\pi}{2}}e^{2R^2}$$

(when R is small). Then for any $R > 0$ for $\mu = \mu_R$

$$\hat{c}_\mu = \sup_x \hat{K}_\mu \le \hat{c} := c_* c^* \sup_{Q>0} \min\left(\frac{1+Q}{Q}, \sqrt{\frac{\pi}{2}}(1+Q)e^{2Q^2}\right);$$

for $R = 0$ the measure μ just equals γ and therefore $\hat{c}_\mu = 1$.

This example is motivated by the manuscript [5], where the problem of estimating of the Poincaré constant for a Gaussian measure conditioned outside a ball is considered. One approach proposed therein is based on the decomposition of variance, and requires an estimate for the Poincaré constant of one-dimensional projection of the "punctured" Gaussian measure on the line which contains the center of the ball. Such an estimate depend on the position and the size of the ball, see Lemma 4.7 in [5], and the case of a large ball touching the origin relates the case (4) of that lemma. Our estimate for the log-Sobolev constant implies that the Poincaré constant for μ is uniformly bounded by $\hat{c}$, which drastically improves the bound ce^{R^2} from Lemma 4.7 [5], statement (4). Heuristically, the reason for this is the following. The measure μ contain "cavities", which appear due to the "puncturing" procedure, and if the ball is "large" and is located not so "far from the origin", then these "cavities" make the bounds for the Poincaré inequality obtained via classic sufficient conditions to be very inaccurate. On the other hand, the form of the weight $\hat{K}_\mu$ in Proposition 3 is highly adjusted to these "cavities", which makes respective bounds more precise. We believe that similar calculations can be made in a general setting, i.e. for arbitrary $d \geq 2$ and arbitrary position and size of the ball; this is a subject of a further research.

Acknowledgements. The paper was prepared partially during a visit of authors to the University of Potsdam; the authors express their deep gratitude to the University of Potsdam and personally to Sylvie Roelly for the hospitality. The authors are also grateful to Emmanuel Boissard for fruitful discussions concerning the manuscript [5].

References

1. D. Bakry and M. Emery, *Hypercontractivity for diffusion semi-groupes*, Comptes rendus des seances de l'Academie des sciences. Serie 1 **299** (1984), no. 15, 775–778.
2. F. Barthe, D. Cordero-Erausquin, and M. Fradelizi, *Gaussian shift inequalities and norms of barycenters*, Studia Math. **146** (2001), 245–259.
3. S. G. Bobkov, *The size of singular component and shift inequalities*, The Annals of Probability **27** (1999), no. 1, 416–431.
4. S. V. Bodnarchuk and A. M. Kulik, *Stochastic control based on time-change transformations for stochastic processes with Lévy noise*, Prob. Theory and Math. Stat. **86** (2012), 11–27.
5. E. Boissard, P. Cattiaux, A. Guillin, and L. Miclo, *Ornstein–Uhlenbeck Pinball: I. Poincaré Inequaliies in a Punctured Domain*, Manuscript.
6. V. I. Bogachev, *Differentiable Measures and the Malliavin Calculus*, AMS, Providence, 2010.
7. I. Cascos, *Data depth: multivariate statistics and geometry*, New Perspectives in Stochastic Geometry (W. S. Kendall and I. Molchanov, eds.), Oxford Univ. Press, Oxford, 2010, pp. 398–426.
8. P. Cattiaux and A. Guillin, *Functional inequalities via Lyapunov conditions*, arXiv:1001.1822.
9. G. A. Koshevoy and K. Mosler, *Zonoid trimming for multivaiate distributions*, Ann. Stat. **25:5** (1997), 1998–2017.
10. G. A. Koshevoy and K. Mosler, *Lift zonoids, random convex hulls and the variability of random vectors*, Bernoulli **4** (1998), 377–399.
11. A. M. Kulik and T. D. Tymoshkevych, *Lift zonoid and barycentric representation on a Banach space with a cylinder measure*, Lithuanian Mathematical Journal **53** (2013), no. 2, 181–195.
12. I. Molchanov and M. Schmutz, *Exchangeability type properties of asset prices*, Adv. Appl. Probab. **43** (2011), 666–687.
13. I. Molchanov, M. Schmutz, and K. Stucki, *Invariance properties of random vectors and stochastic processes based on the zonoid concept*, arXiv:1203.6085.
14. A. Yu. Veretennikov and A. M. Kulik, *On extended Poisson equation for weakly ergodic Markov processes*, Prob. Theory and Math. Stat. **85** (2011), 22–38.
15. A. Yu. Veretennikov and A. M. Kulik, *Diffusion approximation for systems with weakly ergodic Markov perturbations I*, Prob. Theory and Math. Stat. **87** (2012), 1–16.
16. A. Yu. Veretennikov and A. M. Kulik, *Diffusion approximation for systems with weakly ergodic Markov perturbations II*, Prob. Theory and Math. Stat. **88** (2013), 1–16.
17. F. Y. Wang, *Functional inequalities for empty essential spectrum*, J. Funct. Anal. **170** (2000), no. 1, 219–245.

Institute of Mathematics Nat. Acad. Sci. of Ukraine, Tereshchenkivska str. 3, 01601, Kyiv, Ukraine
E-mail address: kulik.alex.m@gmail.com

National Taras Shevchenko University of Kyiv, Academician Glushkov pr. 4-e, 03127, Kyiv, Ukraine
E-mail address: tymoshkevych@gmail.com

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 91–103
Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 91–103

ASYMPTOTIC BEHAVIOR OF THE INTEGRAL FUNCTIONALS FOR UNSTABLE SOLUTIONS OF ONE-DIMENSIONAL ITÔ STOCHASTIC DIFFERENTIAL EQUATIONS

UDC 519.21

G. L. KULINICH, S. V. KUSHNIRENKO, AND Y. S. MISHURA

Abstract. We consider the stochastic one-dimensional differential equations with homogeneous drift and unit diffusion. The drift satisfies conditions supplying the unstable property of the unique strong solution. The explicit form of normalizing factor for certain integral functionals of unstable solution is established to provide the weak convergence to the limiting process. As a result we get the new class of limiting processes that are the functionals of Bessel diffusion processes.

Анотація. Розглядається одновимірне стохастичне диференціальне рівняння з однорідним коефіцієнтом переносу та одиничною дифузією. Коефіцієнт переносу задовольняє умови, при яких єдиний сильний розв'язок даного рівняння є нестійким. Знайдено явний вигляд нормування для певних функціоналів інтегрального типу від нестійкого розв'язку, що забезпечує слабку збіжність до граничного процесу. Отримано новий клас граничних процесів, які є певними функціоналами від бесселівських дифузійних процесів.

Аннотация. Рассматривается одномерное стохастическое дифференциальное уравнение с однородным коэффициентом сноса и единичной диффузией. Коэффициент сноса удовлетворяет условиям, при которых единственное сильное решение данного уравнения является неустойчивым. Найден явный вид нормировки для определенных функционалов интегрального типа от неустойчивого решения, что обеспечивает слабую сходимость к предельному процессу. Получен новый класс предельных процессов, которые являются определенными функционалами от бесселевских диффузионных процессов.

1. Introduction

Let $(\Omega, \Im, \mathsf{P})$ be the complete probability space and $W = \{W(t), t \geq 0\}$ be one-dimensional Wiener process on this space. Let the function $a = a(x) \colon \mathbb{R} \to \mathbb{R}$ be measurable and bounded. It is well-known (see, e.g. [15] and [14], Theorem 4) that the stochastic differential equation with the homogeneous drift and the unit diffusion

$$d\xi(t) = a(\xi(t))\,dt + dW(t), \qquad t \geq 0, \tag{1}$$

has the unique strong solution $\xi = \{\xi(t), t \geq 0\}$ for any initial condition $\xi(0) = x_0 \in \mathbb{R}$.

Definition 1.1. Solution $\xi = \{\xi(t), t \geq 0\}$ of equation (1) is called unstable if for any constant $N > 0$

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \mathsf{P}\{|\xi(s)| < N\}\, ds = 0.$$

Definition 1.2. Solution $\xi = \{\xi(t), t \geq 0\}$ of equation (1) has ergodic distribution $G(x)$ if for all $x \in \mathbb{R}$

$$\lim_{t\to\infty} \mathsf{P}\{\xi(t) < x\} = G(x).$$

2000 *Mathematics Subject Classification.* Primary 60H10.
Key words and phrases. Itô stochastic differential equation, unstable solution, asymptotic behavior of integral functionals.

Definition 1.3. The family $\{\zeta_T(t), t \geq 0\}$ of stochastic processes is said to converge weakly as $T \to \infty$ to the process $\{\zeta(t), t \geq 0\}$ if for any $L > 0$ measures $\mu_T[0, L]$ that correspond to the processes $\zeta_T(\cdot)$ on the interval $[0,\, L]$ converge weakly to the measure $\mu[0,\, L]$ that corresponds to the process $\zeta(\cdot)$.

Throughout the paper we suppose that the drift coefficient a satisfies assumption

(A_1) there exists such $C > 0$ that for any $x \in \mathbb{R}$

$$|xa(x)| \leq C.$$

In this case we can say that the class of equations (1) is located on the border between the equations whose solutions have ergodic distribution, and the equations with unstable solutions. To illustrate this observation, consider the drift coefficient of the form $a(x) = \frac{ax}{1+x^2}$ and introduce the function

$$f(x) = \exp\left\{-2\int_0^x a(v)\,dv\right\}. \tag{2}$$

Note that in our case $f(x) = (1+x^2)^{-a}$. In the paper [11] two cases were considered, namely, $a < -\frac{1}{2}$, $a > -\frac{1}{2}$. It was proved that in the case $a < -\frac{1}{2}$ the solution ξ of equation (1) has ergodic distribution, is transient and moreover

$$\lim_{t\to\infty} \mathsf{P}\{\xi(t) < x\} = \left[\int_{\mathbb{R}} \frac{dv}{f(v)}\right]^{-1} \left[\int_{-\infty}^x \frac{dv}{f(v)}\right] = \left[\int_{\mathbb{R}} (1+v^2)^a dv\right]^{-1} \left[\int_{-\infty}^x (1+v^2)^a dv\right]. \tag{3}$$

At the same time in the case $a > -\frac{1}{2}$ the solution ξ of equation (1) is unstable and recurrent and furthermore the process $r_T(t) = \frac{|\xi(tT)|}{\sqrt{T}}$ with normalizing factor $\frac{1}{\sqrt{T}}$ weakly converges as $T \to \infty$ to the Bessel process $r(t)$ that is the solution of the Itô's equation

$$dr^2(t) = (2a+1)\,dt + 2r(t)\,d\widehat{W}(t) \tag{4}$$

with some Wiener process $\{\widehat{W} = \widehat{W}(t), t \geq 0\}$. Here the weak convergence is considered in the uniform topology on the space of continuous functions. The case $a = -\frac{1}{2}$ is critical in the sense that for $a = -\frac{1}{2}$ the process is recurrent, $\mathsf{P}\{\overline{\lim}_{t\to\infty} \xi(t) = +\infty\} = \mathsf{P}\{\underline{\lim}_{t\to\infty} \xi(t) = -\infty\} = 1$, however, we do not know the normalizing factor that supplies the weak convergence.

The assertion that value $a = -\frac{1}{2}$ is critical can be illustrated by the following examples:

1) If $a(x) = -\frac{1}{2}\frac{x}{1+x^2} - 2\,\frac{x}{(1+x^2)\ln(1+x^2)}$ then the solution ξ of equation (1) has the ergodic distribution and moreover, we have in equality (3) $f(x) = \sqrt{1+x^2}\,\left[\ln(1+x^2)\right]^2$.

2) If $a(x) = -\frac{1}{2}\frac{x}{1+x^2} + \frac{x}{(1+x^2)\ln(1+x^2)}$ then the solution ξ of equation (1) is unstable, and stochastic process $\frac{\xi(tT)}{\sqrt{T}}$ converges to degenerate process $r(t) \equiv 0$ as $T \to \infty$.

The present paper is devoted to the asymptotic behavior of the integral functionals $\beta(t) = \int_0^t g(\xi(s))\,ds$ as $t \to \infty$. We suppose that $g = g(x)\colon \mathbb{R} \to \mathbb{R}$ is locally integrable function, ξ is the solution of equation (1). Also, introduce some additional notations. Denote Ψ the class of functions $\psi = \psi(r) > 0$, $r \geq 0$, that are non-decreasing and regularly varying (at infinity) with index $\alpha > 0$, i.e., $\lim_{T\to\infty} \frac{\psi(rT)}{\psi(T)} = r^\alpha$ for all $r > 0$.

Now, take function f that is defined via the relation (2), some constant $b \in \mathbb{R}$ and denote

$$q(x) = \frac{f(x)}{\psi(|x|)} \int_0^x \frac{g(u)}{f(u)}\,du - \bar{b}(x), \qquad \bar{b}(x) = b\,\mathrm{sign}\,x. \tag{5}$$

Suppose additionally that the drift coefficient a and function g satisfy assumption (A_2) (i) with one of the additional restrictions (ii), (iii) or (iv) and also one of the assumptions (A_3) and (A_4):

(A_2) (i) There exist the constants c_i, $i = 1, 2$ such that

$$\lim_{|x|\to\infty} \left[\frac{1}{x} \int_0^x va(v)\, dv - \bar{c}(x) \right] = 0, \tag{6}$$

where

$$\bar{c}(x) = \begin{cases} c_1, & x > 0, \\ c_2, & x < 0, \end{cases}$$

and moreover, one of the following restrictions on the coefficients hold:

(ii) $c_1 = c_2 = c_0 > -\frac{1}{2}$;
(iii) $c_1 > \frac{1}{2}$, $c_2 < \frac{1}{2}$;
(iv) $c_1 < \frac{1}{2}$, $c_2 > \frac{1}{2}$.

(A_3) (i) there exists a constant $C > 0$ such that $f(x) \le C$ for any $x \in \mathbb{R}$ and
(ii) there exist such $b \in \mathbb{R}$ and function $\psi \in \Psi$ that

$$\lim_{|x|\to\infty} \frac{1}{x} \int_0^x \frac{q^2(u)}{f(u)}\, du = 0; \tag{7}$$

(A_4) (i) there exists a constant $\delta > 0$ such that $0 < \delta \le f(x)$ for any $x \in \mathbb{R}$ and
(ii) there exist such $b \in \mathbb{R}$ and function $\psi \in \Psi$ that

$$\lim_{|x|\to\infty} \frac{f(x)}{x} \int_0^x q^2(u)\, du = 0. \tag{8}$$

In the present paper in order to proof that under the conditions (A_1), (A_2) and one of the conditions (A_3) and (A_4) random variable $\frac{\beta(t)}{\sqrt{t}\psi(\sqrt{t})}$ with normalizing factor $\frac{1}{\sqrt{t}\psi(\sqrt{t})}$ has the limit distribution as $t \to \infty$, we study the limit behavior as $T \to \infty$ of the process

$$\beta_T(t) = \frac{1}{\sqrt{T}\psi(\sqrt{T})} \int_0^{tT} g(\xi(s))\, ds,$$

with parameter $T > 0$. Theorems 2.1 and 2.2 describe the limit behavior mentioned above.

Remark 1.1. It is very easy to see that any of conditions (A_3) and (A_4) supply the convergence

$$\lim_{|x|\to\infty} \frac{1}{x} \int_0^x q^2(u)\, du = 0. \tag{9}$$

If condition (A_3) holds then

$$\frac{1}{x} \int_0^x q^2(u)\, du \le C \frac{1}{x} \int_0^x \frac{q^2(u)}{f(u)}\, du \to 0 \quad \text{as } |x| \to \infty.$$

If condition (A_4) holds then

$$\frac{1}{x} \int_0^x q^2(u)\, du \le \frac{1}{\delta} \frac{f(x)}{x} \int_0^x q^2(u)\, du \to 0 \quad \text{as } |x| \to \infty.$$

Moreover, if $0 < \delta \le f(x) \le C$ then (9) is equivalent both to (A_3), (ii) and (A_4), (ii). However, neither (A_3), (ii) and (A_4), (ii) nor (9) do not supply convergence $q(x) \to 0$ as $|x| \to \infty$. In other words, under any of these conditions function q can admit "explosions".

Remark 1.2. The function $q(x)$ (see Example 2.1) satisfies the condition (9). Obviously, $q(x) \nrightarrow 0$ as $|x| \to \infty$.

As to previous results in this direction, it was proved in [11] that under the condition (A_2) solution ξ of equation (1) is unstable. Moreover, in the case when (A_2), (ii) holds then $\frac{|\xi(tT)|}{\sqrt{T}}$ weakly converges as $T \to \infty$ to process r that is the solution of equation (4) with $a = c_0$. In the case when (A_2), (iii) holds then $\frac{\xi(tT)}{\sqrt{T}}$ weakly converges to process r with $a = c_1$, and in the case when (A_2), (iv) holds then $\frac{-\xi(tT)}{\sqrt{T}}$ weakly converges to process r with $a = c_2$. Asymptotic behavior of the process $\beta_T(t)$ in the case when conditions (A_2), (i) and (ii) hold and additionally $q(x) \to 0$ as $|x| \to \infty$ were considered in the papers [5] and [12]. The results of the paper [5] are generalized in the present paper to the case of the functions $q = q(x)$ with possible "explosions" (conditions (A_3) and (A_4)) and are extended to the cases when (A_2), (i) and (iii) or (A_2), (i) and (iv) hold. Moreover, the proofs from [5] are essentially simplified in the present paper due to the representation (12). The paper [12] contains similar result for the functional $\beta_T(t)$ of the solution ξ of equation (1) on the half-axis $(0, +\infty)$ with the instant reflection of the solution at zero point, and in this case it was supposed that $\psi(|x|) = |x|^\alpha$, $\alpha \geq 0$, $q(x) \to 0$ as $x \to \infty$.

The most complete results concerning the asymptotic behavior of the functionals $\beta_T(t)$ are proved for the equations (1) with more restrictive assumption on the drift coefficient, namely, $\left|\int_0^x a(u)\,du\right| \leq C$ (see [8] – [10]). The paper [8] contains the weak convergence of distributions of $\beta_T(t)$ in the case when $q(x) \to 0$ as $|x| \to \infty$. In the paper [9] the weak convergence of distributions of $\beta_T(t)$ was obtained under assumption (9) on function $q = q(x)$. In the paper [10] the necessary and sufficient conditions of weak convergence were obtained that are connected, in some sense, to (9).

The asymptotic behavior of the integral functionals of the form $\int_0^t g_T(\xi_T(s))\,d\mu_T(s)$, where $\xi_T(t)$ are the solutions of stochastic differential equations and $\mu_T(t)$ is the family of martingales that converge in probability, was considered in the paper [3, §5, Chapter IX] under the assumption of locally uniform convergence of the coefficients of the equation.

The paper is organized as follows: principal results are proved in Section 2 while an auxiliary lemma is relegated to Section 3. Section 4 concludes.

2. The main results

In what follows we denote C or C with some subscripts constants whose values are not so important and can change from line to line.

Theorem 2.1. *Let ξ be the solution of equation* (1) *with the drift coefficient a satisfying assumptions (A_1), (A_2), (i) and one of the assumptions (A_2), (ii), (iii) or (iv).*

Then the stochastic process

$$\beta_T(t) = \frac{1}{\sqrt{T}\psi(\sqrt{T})} \int_0^{tT} g(\xi(s))\,ds$$

converges as $T \to \infty$ weakly in the unform topology of the space of continuous functions to the process

$$\beta(t) = 2b\left[\frac{r^{\alpha+1}(t)}{\alpha+1} - \int_0^t r^\alpha(s)\,d\widehat{W}(s)\right],$$

where $r(t) \geq 0$ is the solution of stochastic differential equation

$$dr^2(t) = (2a+1)\,dt + 2r(t)\,d\widehat{W}(t).$$

Here $a = c_0$ in the case when assumption (A_2), (ii) is satisfied, $a = c_1$ in the case when assumption (A_2), (iii) is satisfied and $a = c_2$ in the case when assumption (A_2), (iv) is satisfied.

Proof. Introduce parameter $T > 0$ and set

$$r_T(t) = \frac{|\xi(tT)|}{\sqrt{T}}, \qquad W_T(t) = \frac{W(tT)}{\sqrt{T}}, \qquad \widehat{W}_T(t) = \int_0^t \operatorname{sign} \xi(sT)\, dW_T(s),$$
$$P_N = \mathsf{P}\left\{ \sup_{0 \le t \le L} r_T(t) > N \right\}, \qquad \alpha_T(t) = \frac{1}{T} \int_0^{tT} [\xi(s)a(\xi(s)) - \bar{c}(\xi(s))]\, ds,$$

where L and N are arbitrary positive constants. Evidently, for any fixed $T > 0$ process $W_T = \{W_T(t), t \ge 0\}$ is a Wiener process. Furthermore, it follows, e.g., from [2, Chapter 6, §3, Lemma 5] that

$$\int_0^t \mathsf{P}\{\xi(s) = 0\}\, ds = 0$$

for any $t > 0$. Therefore $\widehat{W}_T = \{\widehat{W}_T(t), t \ge 0\}$ for any $T > 0$ is continuous with probability 1 square integrable martingale with the quadratic characteristics $\langle \widehat{W}_T \rangle(t) = t$. It immediately follows from the Doob's theorem that $\widehat{W}_T$ is a Wiener process for any $T > 0$. Applying Itô's formula to the process r_T^2, we get

$$r_T^2(t) = \frac{x_0^2}{T} + \int_0^t [2\bar{c}(\xi(sT)) + 1]\, ds + 2\int_0^t r_T(s)\, d\widehat{W}_T(s) + 2\alpha_T(t).$$

Consider the function

$$F(x) = 2\int_0^x f(u) \left(\int_0^u \frac{g(v)}{f(v)}\, dv \right) du.$$

Obviously, function F has a continuous derivative F' and a.e. w.r.t. to the Lebesgue measure on $\mathbb{R}$ has a second derivative F'' that is locally integrable. Therefore we can apply an Itô's formula from [6, Chapter 2, §10] to $F(\xi(t))$ and get the equality

$$F(\xi(t)) - F(x_0) = \int_0^t \left[F'(\xi(s))a(\xi(s)) + \frac{1}{2} F''(\xi(s)) \right] ds + \int_0^t F'(\xi(s))\, dW(s) \tag{10}$$

with probability 1 for any $t \ge 0$. It is easy to see that a.e. w.r.t. to the Lebesgue measure on $\mathbb{R}$ the following equality holds

$$F'(x)a(x) + \frac{1}{2} F''(x) = g(x). \tag{11}$$

Applying (11) to (10) we get that

$$F(\xi(t)) - F(x_0) = \int_0^t g(\xi(s))\, ds + \int_0^t F'(\xi(s))\, dW(s)$$

with probability 1 for any $t \ge 0$. After some evident transformations we get from the last equality that

$$\beta_T(t) = \frac{1}{\sqrt{T}\psi(\sqrt{T})} \left[F(\xi(tT)) - F(x_0) - \int_0^{tT} F'(\xi(s))\, dW(s) \right].$$

Let us consider the first term

$$
\begin{aligned}
\frac{F(\xi(tT))}{\sqrt{T}\psi(\sqrt{T})} &= \frac{2}{\sqrt{T}\psi(\sqrt{T})}\int_0^{\xi(tT)} f(u)\left(\int_0^u \frac{g(v)}{f(v)}\,dv\right)du \\
&= \frac{2}{\sqrt{T}\psi(\sqrt{T})}\int_0^{\xi(tT)}\left(\frac{f(u)}{\psi(|u|)}\int_0^u \frac{g(v)}{f(v)}\,dv \pm \bar{b}(u)\right)\psi(|u|)\,du \\
&= \frac{2}{\sqrt{T}\psi(\sqrt{T})}\left(\int_0^{\xi(tT)} \bar{b}(u)\psi(|u|)\,du + \int_0^{\xi(tT)} q(u)\psi(|u|)\,du\right) \\
&= 2b\int_0^{\frac{|\xi(tT)|}{\sqrt{T}}} \frac{\psi\left(|u|\sqrt{T}\right)}{\psi(\sqrt{T})}\,du + \frac{2}{\sqrt{T}\psi(\sqrt{T})}\int_0^{\xi(tT)} q(u)\psi\left(|u|\right)\,du \\
&= 2b\int_0^{\frac{|\xi(tT)|}{\sqrt{T}}} |u|^\alpha\,du + 2b\int_0^{\frac{|\xi(tT)|}{\sqrt{T}}}\left(\frac{\psi\left(|u|\sqrt{T}\right)}{\psi(\sqrt{T})} - |u|^\alpha\right)du \\
&\quad + \frac{2}{\sqrt{T}\psi(\sqrt{T})}\int_0^{\xi(tT)} q(u)\psi\left(|u|\right)\,du,
\end{aligned}
$$

and transform the last term

$$
\begin{aligned}
\frac{1}{\sqrt{T}\psi(\sqrt{T})}&\int_0^{tT} F'(\xi(s))\,dW(s) = \frac{2}{\sqrt{T}\psi(\sqrt{T})}\int_0^{tT} f(\xi(s))\left(\int_0^{\xi(s)}\frac{g(u)}{f(u)}\,du\right)dW(s) \\
&= \frac{2}{\sqrt{T}\psi(\sqrt{T})}\left[\int_0^{tT}\bar{b}(\xi(s))\psi(|\xi(s)|)\,dW(s) + \int_0^{tT} q(\xi(s))\psi(|\xi(s)|)\,dW(s)\right] \\
&= \frac{2}{\psi(\sqrt{T})}\int_0^t \bar{b}(\xi(sT))\psi(|\xi(sT)|)\,\frac{dW(sT)}{\sqrt{T}} + 2\int_0^t q(\xi(sT))\frac{\psi(|\xi(sT)|)}{\psi(\sqrt{T})}\,dW_T(s) \\
&= 2b\int_0^t \frac{\psi(|\xi(sT)|)}{\psi(\sqrt{T})}\,d\widehat{W}_T(s) + 2\int_0^t q(\xi(sT))\frac{\psi(|\xi(sT)|)}{\psi(\sqrt{T})}\,dW_T(s) \\
&= 2b\int_0^t r_T^\alpha(s)\,d\widehat{W}_T(s) + 2b\int_0^t\left[\frac{\psi(r_T(s)\sqrt{T})}{\psi(\sqrt{T})} - r_T^\alpha(s)\right]d\widehat{W}_T(s) \\
&\quad + 2\int_0^t q(\xi(sT))\frac{\psi(r_T(s)\sqrt{T})}{\psi(\sqrt{T})}\,dW_T(s).
\end{aligned}
$$

Therefore

$$
\beta_T(t) = -\frac{F(x_0)}{\sqrt{T}\psi(\sqrt{T})} + 2b\int_0^{r_T(t)} u^\alpha du - 2b\int_0^t r_T^\alpha(s)\,d\widehat{W}_T(s) + 2\sum_{k=1}^4 S_T^{(k)}(t), \tag{12}
$$

where

$$
\begin{aligned}
S_T^{(1)}(t) &= b\int_0^{r_T(t)}\left[\frac{\psi(u\sqrt{T})}{\psi(\sqrt{T})} - u^\alpha\right]du, \\
S_T^{(2)}(t) &= \frac{1}{\sqrt{T}\psi(\sqrt{T})}\int_0^{\xi(tT)} q(u)\psi(|u|)\,du, \\
S_T^{(3)}(t) &= -b\int_0^t\left[\frac{\psi(r_T(s)\sqrt{T})}{\psi(\sqrt{T})} - r_T^\alpha(s)\right]d\widehat{W}_T(s), \\
S_T^{(4)}(t) &= -\int_0^t q(\xi(sT))\frac{\psi(r_T(s)\sqrt{T})}{\psi(\sqrt{T})}\,dW_T(s).
\end{aligned}
$$

It is known from [11] that under condition (A_1) the process $\{r_T(t), t \geq 0\}$ converges weakly as $T \to \infty$ to the process $\{r(t), t \geq 0\}$ that is the solution of equation (4) with $a = c_0$ in the case (A_2),(ii), with $a = c_1$ in the case (A_2), (iii) and with $a = c_2$ in the case (A_2), (iv). Furthermore, for any $L > 0$ and $\varepsilon > 0$ we have that

$$\lim_{N\to\infty} \overline{\lim_{T\to\infty}} P_N = 0,$$

$$\lim_{h\to 0} \overline{\lim_{T\to\infty}} \sup_{|t_1-t_2|\leq h;\, t_i\leq L} \mathsf{P}\left\{|r_T(t_2) - r_T(t_1)| > \varepsilon\right\} = 0. \tag{13}$$

Now we are in position to establish that $S_T^{(k)}$, $k = 1, \ldots, 4$, uniformly converge to zero in probability. In particular, it means that they satisfy equalities (13) as well. To start with, note that it follows from Lemma 3.1, evident inequalities

$$\mathsf{P}\left\{|\eta + \zeta| > \varepsilon\right\} \leq \mathsf{P}\left\{|\eta| > \frac{\varepsilon}{2}\right\} + \mathsf{P}\left\{|\zeta| > \frac{\varepsilon}{2}\right\}, \qquad \mathsf{P}\left\{|\eta| > \varepsilon\right\} \leq \frac{\mathsf{E}\, h(|\eta|)}{h(\varepsilon)}$$

with $h = x$ and $h = x^2$ and the properties of Itô's integrals that for any $\varepsilon > 0$, $L > 0$ and $T \geq T_N$, where T_N are introduced in Lemma 3.1, the following inequalities hold true:

$$\begin{aligned} \mathsf{P}\left\{\sup_{0\leq t\leq L} |S_T^{(1)}(t)| > \varepsilon\right\} &\leq P_N + \frac{2}{\varepsilon}\,\mathsf{E} \sup_{0\leq t\leq L} |S_T^{(1)}(t)| \chi_{\{r_T(t)\leq N\}} \\ &\leq P_N + \frac{2}{\varepsilon}\,|b| \int_0^N \left|\frac{\psi(u\sqrt{T})}{\psi(\sqrt{T})} - u^\alpha\right| du, \end{aligned} \tag{14}$$

$$\begin{aligned} &\mathsf{P}\left\{\sup_{0\leq t\leq L} \left|S_T^{(2)}(t)\right| > \varepsilon\right\} \leq \mathsf{P}\left\{\sup_{0\leq t\leq L} \left|\int_0^{\frac{\xi(tT)}{\sqrt{T}}} q(u\sqrt{T}) \frac{\psi(|u|\sqrt{T})}{\psi(\sqrt{T})}\, du\right| > \varepsilon\right\} \\ &\quad\leq P_N + \frac{2}{\varepsilon}\,\mathsf{E} \sup_{0\leq t\leq L} \left|\int_0^{\frac{\xi(tT)}{\sqrt{T}}} q(u\sqrt{T}) \frac{\psi(|u|\sqrt{T})}{\psi(\sqrt{T})}\, du\right| \chi_{\{r_T(t)\leq N\}} \\ &\quad\leq P_N + \frac{2}{\varepsilon}\, C_N \int_{-N}^{N} |q(u\sqrt{T})|\, du \leq P_N + \frac{2}{\varepsilon}\, C_N (2N)^{\frac{1}{2}} \left(\frac{1}{\sqrt{T}} \int_{-N\sqrt{T}}^{N\sqrt{T}} q^2(u)\, du\right)^{\frac{1}{2}}, \end{aligned} \tag{15}$$

$$\mathsf{P}\left\{\sup_{0\leq t\leq L} |S_T^{(3)}(t)| > \varepsilon\right\} \leq P_N + 4\left(\frac{2}{\varepsilon}\right)^2 b^2\, \mathsf{E} \int_0^L \left|\frac{\psi(r_T(s)\sqrt{T})}{\psi(\sqrt{T})} - r_T^\alpha(s)\right|^2 \chi_{\{r_T(s)\leq N\}}\, ds, \tag{16}$$

$$\begin{aligned} \mathsf{P}\left\{\sup_{0\leq t\leq L} |S_T^{(4)}(t)| > \varepsilon\right\} &\leq P_N + 4\left(\frac{2}{\varepsilon}\right)^2 \mathsf{E} \int_0^L q^2(\xi(sT)) \left[\frac{\psi(r_T(s)\sqrt{T})}{\psi(\sqrt{T})}\right]^2 \chi_{\{r_T(s)\leq N\}}\, ds \\ &\leq P_N + 4\left(\frac{2}{\varepsilon}\right)^2 C_N^2\, \mathsf{E} \int_0^L q^2(\xi(sT)) \chi_{\{r_T(s)\leq N\}}\, ds. \end{aligned} \tag{17}$$

Taking into account the convergence $\frac{\psi(|x|\sqrt{T})}{\psi(\sqrt{T})} - |x|^\alpha \to 0$ as $T \to \infty$, boundedness on the interval $|x| \leq N$ and relation (9), we let in inequalities (14) and (15) $T \to \infty$ and after that $N \to \infty$ and get

$$\sup_{0\leq t\leq L} \left|S_T^{(k)}(t)\right| \xrightarrow{\mathsf{P}} 0 \tag{18}$$

as $T \to \infty$ and for $k = 1, 2$.

Now we shall establish similar convergence for $k = 3, 4$. It is known from [4] that for any $0 < \delta < N < \infty$ the following convergence holds:

$$\sup_{0<\delta\le|x|\le N}\left|\frac{\psi(|x|\sqrt{T})}{\psi(\sqrt{T})} - |x|^\alpha\right| \to 0$$

as $T \to \infty$. Therefore, taking into account monotonicity of function $\psi(r)$, $r \ge 0$, we get the following convergence for any $0 < \delta < N$:

$$\begin{aligned}
&\mathsf{E}\int_0^L \left[\frac{\psi(r_T(s)\sqrt{T})}{\psi(\sqrt{T})} - r_T(s)\right]^2 \chi_{\{r_T(s)\le N\}}\,ds \\
&\quad\le L\sup_{\delta\le|x|\le N}\left|\frac{\psi(|x|\sqrt{T})}{\psi(\sqrt{T})} - |x|^\alpha\right|^2 + 2\int_0^L\left(\left[\frac{\psi(\delta\sqrt{T})}{\psi(\sqrt{T})}\right]^2 + \delta^2\right)ds \to 0,
\end{aligned}$$

if to tend at first $T \to \infty$ and after that $\delta \to 0$.

So, taking into account inequality (16) we get that convergence (18) holds for $S_T^{(3)}(t)$ as well. At last, in order to prove convergence (18) for $S_T^{(4)}(t)$, we apply Itô formula and get

$$\mathsf{E}\int_0^L q^2(\xi(sT))\chi_{\{|\xi(sT)|\le N\sqrt{T}\}}\,ds = \mathsf{E}\left[\Phi_T(\xi(LT)) - \Phi_T(x_0)\right],$$

where

$$\Phi_T(x) = \frac{1}{T}\int_0^x f(u)\left(\int_0^u \frac{q^2(v)}{f(v)}\chi_{\{|v|\le N\sqrt{T}\}}\,dv\right)du.$$

Now we consider separately conditions (A_3) and (A_4). It is easy to see that under condition (A_3) we have the following relations

$$\begin{aligned}
\frac{1}{x^2}\left|\int_0^x f(u)\left(\int_0^u \frac{q^2(v)}{f(v)}\,dv\right)du\right| &\le \frac{C}{x^2}\left|\int_0^x\left(\int_0^u \frac{q^2(v)}{f(v)}\,dv\right)du\right| \\
&= \frac{C}{x^2}\left|\int_0^x u\left(\frac{1}{u}\int_0^u \frac{q^2(v)}{f(v)}\,dv\right)du\right| \to 0 \quad \text{as } |x| \to \infty.
\end{aligned}$$

In turn, under condition (A_4) we have the following relations

$$\begin{aligned}
\frac{1}{x^2}\left|\int_0^x f(u)\left(\int_0^u \frac{q^2(v)}{f(v)}\,dv\right)du\right| &\le \frac{1}{\delta}\frac{1}{x^2}\left|\int_0^x f(u)\left(\int_0^u q^2(v)\,dv\right)du\right| \\
&= \frac{1}{\delta}\frac{1}{x^2}\left|\int_0^x u\left(\frac{f(u)}{u}\int_0^u q^2(v)\,dv\right)du\right| \to 0 \quad \text{as } |x| \to \infty.
\end{aligned}$$

Therefore, any of conditions (A_3) and (A_4) supply the following convergence

$$\frac{1}{x^2}\int_0^x f(u)\left(\int_0^u \frac{q^2(v)}{f(v)}\,dv\right)du \to 0$$

as $|x| \to \infty$. Therefore for any $\varepsilon > 0$ there exists such L_ε that for $|x| > L_\varepsilon$ we have inequality

$$\frac{1}{x^2}\left|\int_0^x f(u)\left(\int_0^u \frac{q^2(v)}{f(v)}\,dv\right)du\right| < \varepsilon. \tag{19}$$

Furthermore, since function $\frac{1}{x^2}\int_0^x f(u)\left(\int_0^u \frac{q^2(v)}{f(v)}\,dv\right)du$ is bounded at zero, there exists such $C_\varepsilon > 0$ that

$$\sup_{|x|\le L_\varepsilon}\frac{1}{x^2}\left|\int_0^x f(u)\left(\int_0^u \frac{q^2(v)}{f(v)}\,dv\right)du\right| \le C_\varepsilon. \tag{20}$$

Besides this,

$$\mathsf{E}\,\frac{|\xi(tT)|^2}{T} \le C + C_1 t. \tag{21}$$

Relations (19) and (20) together with (21) provide that

$$\begin{aligned}\mathsf{E}\,|\Phi_T(\xi(LT))| &\le \mathsf{E}\,\frac{|\xi(tT)|^2}{T}\cdot\frac{1}{|\xi(LT)|^2}\left|\int_0^{\xi(LT)} f(u)\left(\int_0^u \frac{q^2(v)}{f(v)}\,dv\right)du\right| \\ &\le \frac{C_\varepsilon}{T} + \varepsilon(C + C_1 t),\end{aligned}$$

whence $\mathsf{E}\,|\Phi_T(\xi(LT))| \to 0$ as $T\to\infty$. Evidently, $|\Phi_T(x_0)| \le \frac{C}{T}$. Therefore,

$$\mathsf{E}\int_0^L q^2(\xi(sT))\chi_{\{|\xi(sT)|\le N\sqrt{T}\}}\,ds \to 0$$

as $T\to\infty$. Together with (17) it means that the convergence (18) holds for $S_T^{(4)}(t)$ as well. Evidently, relation (13) holds for processes $\widehat{W}_T(t)$.

It means that we can apply Skorokhod representation theorem [13] and for any sequence $T_n\to\infty$ to choose the subsequence $T_n'\to\infty$, probability space $(\tilde{\Omega},\tilde{\Im},\tilde{\mathsf{P}})$ and processes $(\tilde{r}_{T_n'}(t),\tilde{W}_{T_n'}(t),\tilde{S}_{T_n'}^{(i)}(t), i=1,\dots,4)$ on this space so that the couple of processes will be stochastically equivalent to the process $(r_{T_n'}(t),\widehat{W}_{T_n'}(t),S_{T_n'}^{(i)}(t), i=1,\dots,4)$ and moreover,

$$\tilde{r}_{T_n'}(t)\xrightarrow{\tilde{\mathsf{P}}}\tilde{r}(t),\qquad \tilde{W}_{T_n'}(t)\xrightarrow{\tilde{\mathsf{P}}}\tilde{W}(t),\qquad \tilde{S}_{T_n'}^{(i)}(t)\xrightarrow{\tilde{\mathsf{P}}}\tilde{S}^{(i)}(t),\qquad i=1,\dots,4,$$

as $T_n'\to\infty$. In our case, according to (18), $\tilde{S}^{(i)}(t)=0$, $i=1,\dots,4$, and the processes $\tilde{r}(t)$, $\tilde{W}(t)$ satisfy equations (4) with $a=c_0$ in the case (A_2), (ii), $a=c_1$ in the case (A_2), (iii) and $a=c_2$ in the case (A_2), (iv), see [11].

According to equality (12) we have that the functional $\beta_{T_n'}(t)$ is stochastically equivalent to the functional $\tilde{\beta}_{T_n'}(t)$ for which we have similar equality

$$\tilde{\beta}_{T_n'}(t) = -\frac{F(x_0)}{\sqrt{T_n'}\psi(\sqrt{T_n'})} + 2b\int_0^{\tilde{r}_{T_n'}(t)} u^\alpha\,du - 2b\int_0^t \tilde{r}_{T_n'}^\alpha(s)\,d\tilde{W}_{T_n'}(s) + 2\sum_{i=1}^4 \tilde{S}_{T_n'}^{(i)}(t). \tag{22}$$

It is possible to get the limit as $T_n'\to\infty$ [13] in this equality and get that $\tilde{\beta}_{T_n'}(t)\xrightarrow{\tilde{\mathsf{P}}}\tilde{\beta}(t)$, where

$$\tilde{\beta}(t) = 2b\left[\int_0^{\tilde{r}(t)} u^\alpha du - \int_0^t \tilde{r}^\alpha(s)\,d\tilde{W}(s)\right]. \tag{23}$$

It follows from the strong uniqueness of the solution of equation (4) (see, e.g., [7]) that the distributions of the limit process $\tilde{\beta}(t)$ are unique as well. Therefore, it follows from arbitrary choice of $T_n\to\infty$ that the finite-dimensional distributions of the processes $\beta_T(t)$ tend as $T\to\infty$ to the corresponding distributions of the process $\tilde{\beta}(t)$ that is defined by equality (23). In order to establish the weak convergence of the processes $\beta_T(t)$ to the process $\tilde{\beta}(t)$, it is sufficient to prove tightness, i.e., to prove that for any $L>0$

$$\lim_{h\to 0}\varlimsup_{T\to\infty}\mathsf{P}\left\{\sup_{|t_1-t_2|\le h;\,t_i\le L}|\beta_T(t_2)-\beta_T(t_1)|>\varepsilon\right\}=0. \tag{24}$$

Tightness of the processes $r_T(t)$ was established in [11] and it was mentioned that tightness of $S_T^{(i)}(t)=0$, $i=1,\dots,4$, follows from (18). Furthermore, taking into account the properties of stochastic Itô integrals, we get the following bounds for any $\varepsilon > 0$, $L > 0$ and $N > 0$:

$$\begin{aligned}
&\mathsf{P}\left\{\sup_{|t_1-t_2|\le h; t_i\le L}\left|\int_0^{r_T(t_2)} u^\alpha du - \int_0^{r_T(t_1)} u^\alpha du\right| > \varepsilon\right\}\\
&\qquad \le P_N + \mathsf{P}\left\{N^\alpha \sup_{|t_1-t_2|\le h;\, t_i\le L} |r_T(t_2) - r_T(t_1)| > \frac{\varepsilon}{2}\right\}
\end{aligned} \tag{25}$$

and

$$\begin{aligned}
&\mathsf{P}\left\{\sup_{|t_1-t_2|\le h; t_i\le L}\left|\int_{t_1}^{t_2} r_T^\alpha(s)\, d\widehat{W}_T(s)\right| > \varepsilon\right\}\\
&\quad \le P_N + \mathsf{P}\left\{4\sup_{kh\le L}\sup_{kh\le t\le (k+1)h}\left|\int_{kh}^{t} r_T^\alpha(s)\chi_{\{r_T(s)\le N\}}\, d\widehat{W}_T(s)\right| > \frac{\varepsilon}{2}\right\}\\
&\quad \le P_N + \sum_{kh<L}\mathsf{P}\left\{\sup_{kh\le t\le (k+1)h}\left|\int_{kh}^{t} r_T^\alpha(s)\chi_{\{r_T(s)\le N\}}\, d\widehat{W}_T(s)\right| > \frac{\varepsilon}{8}\right\}\\
&\quad \le P_N + \sum_{kh<L}\left(\frac{8}{\varepsilon}\right)^4 \mathsf{E}\sup_{kh\le t\le (k+1)h}\left[\int_{kh}^{t} r_T^\alpha(s)\chi_{\{r_T(s)\le N\}}\, d\widehat{W}_T(s)\right]^4\\
&\quad \le P_N + \sum_{kh\le L}\left(\frac{8}{\varepsilon}\right)^4\left(\frac{4}{3}\right)^4 \mathsf{E}\left[\int_{kh}^{(k+1)h} r_T^\alpha(s)\chi_{\{r_T(s)\le N\}}\, d\widehat{W}_T(s)\right]^4\\
&\quad \le P_N + \left(\frac{8}{\varepsilon}\right)^4\left(\frac{4}{3}\right)^4\cdot 36N^{4\alpha}\sum_{kh\le L} h^2\\
&\quad \le P_N + \frac{ChN^{4\alpha}}{\varepsilon^4}.
\end{aligned} \tag{26}$$

In the last inequality the following upper bound for the fourth moment of the Itô's integral w.r.t. the Wiener process from [1] or [13] was used:

$$\mathsf{E}\left(\int_a^b f(t)\, dW(t)\right)^4 \le 36(b-a)\int_a^b \mathsf{E}\,|f(t)|^4\, dt.$$

It follows from (25) and (26) that the right-hand side of (12) is tight, i.e., satisfies (24). So, we have tightness (24) and consequently $\beta_T(t)$ weakly converges as $T \to \infty$ to the process $\beta(t)$ whence the proof follows. □

Example 2.1. Consider equation (1) with the drift coefficient of the form $a(x) = \frac{x}{1+x^2}$. In this case $f(x) = (1+x^2)^{-1}$ and the function $q(x)$ from (5) can be rewritten as

$$q(x) = \frac{1}{\psi(|x|)(1+x^2)}\int_0^x g(u)\left(1+u^2\right)\, du - b\,\mathrm{sign}\, x.$$

Let $\psi(|x|) = |x|$ is slowly varying (at infinity) function ($\alpha = 1$), then

$$\int_0^x g(u)\left(1+u^2\right)\, du = bx\left(1+x^2\right) + q(x)|x|\left(1+x^2\right) = x\left(1+x^2\right)\left[b + q(x)\,\mathrm{sign}\, x\right],$$

whence

$$g(x) = \frac{1}{1+x^2}\left[x\left(1+x^2\right)\left(b + q(x)\,\mathrm{sign}\, x\right)\right]'$$

a.e. w.r.t. to the Lebesgue measure on $\mathbb{R}$.

Consider the continuous function with "explosions"

$$q(x) = \begin{cases} q_1(x), & x \in \Delta_n, \\ 0, & x \notin \Delta_n, \end{cases}$$

where $q_1(x) > 0$, $\max_{x\in\Delta_n} q_1(x) = 1$, $\Delta_n = \big(n; n + \frac{1}{n^3}\big)$, $n \in \mathbb{N}$. Continuing $q(x)$ in a symmetric way to $(-\infty, 0)$, we obtain that the function $q(x)$, $x \in \mathbb{R}$, satisfies the condition (7) with the function $f(x) = (1 + x^2)^{-1} \le C$.

If we put $q(x)$ in the last allocated equality we get $g(x)$ such that the stochastic process

$$\beta_T(t) = \frac{1}{\sqrt{T}\psi(\sqrt{T})} \int_0^{tT} g(\xi(s))\, ds = \frac{1}{T} \int_0^{tT} g(\xi(s))\, ds$$

converges as $T \to \infty$ weakly to the process

$$\beta(t) = 2b \left[\frac{r^2(t)}{2} - \int_0^t r(s)\, d\widehat{W}(s) \right]$$

where $r(t) \ge 0$ is the solution of stochastic differential equation

$$dr^2(t) = 3\, dt + 2r(t)\, d\widehat{W}(t).$$

In this case $\beta(t) = 3bt$.

Remark 2.1. Analyzing the proof of Theorem 2.1 it is easy to see that it is true even in the case when we establish just the weak convergence of the processes $r_T(t)$ to the process $r(t)$ and the representation (12) in which $\sup_{0\le t\le L} \big|S_T^{(k)}(t)\big| \xrightarrow{\mathsf{P}} 0$, $k = 1, \ldots, 4$ as $T \to \infty$ for any $L > 0$.

In this connection, we can deduce the following statement as a corollary of Theorem 2.1.

Theorem 2.2. *Let ξ be a solution of equation* (1) *and let convergence relation* (6) *holds. Also, let locally integrable real-valued function g is such that there exists non-decreasing function $\psi(r)$, $r \ge 0$ that is regularly varying at infinity of order $\alpha > 0$ and $q(x) \to 0$ as $|x| \to \infty$. Here q is defined in* (5)*. Then Theorem 2.1 holds.*

Proof. Indeed, apply the representation (12). Similarly to proof of Theorem 2.1 we get that $\sup_{0\le t\le L} \big|S_T^{(k)}(t)\big| \xrightarrow{\mathsf{P}} 0$ as $T \to \infty$, $k = 1, 2, 3$. Convergence $\sup_{0\le t\le L} \big|S_T^{(4)}(t)\big| \xrightarrow{\mathsf{P}} 0$ as $T \to \infty$ follows directly from inequality (17) and convergence $q(x) \to 0$ as $|x| \to \infty$. In order to finish the proof of the present theorem, it is sufficient to apply Remark 2.1. □

Example 2.2. Consider the class of equations (1) with the drift coefficient of the form

$$a(x) = \frac{x\bar{c}(x)}{1 + x^2},$$

where

$$\bar{c}(x) = \begin{cases} c_1, & x > 0, \\ c_2, & x < 0, \end{cases} \qquad c_1 = c_2 = c_0, \qquad 2c_0 + 1 > 0.$$

1) Let $c_0 = 1$. In this case $f(x) = (1 + x^2)^{-1}$ and in order to satisfy the assumptions of Theorem 2.2 the function $q(x)$ can be rewritten as

$$q(x) = \frac{1}{|x|(1 + x^2)} \int_0^x g_0(1 + u^2)\, du - b \operatorname{sign} x.$$

If $b = \frac{g_0}{3}$, $g(x) = g_0$, $\psi(|x|) = |x|$ is slowly varying (at infinity) function ($\alpha = 1$), then $q(x) \to 0$ as $|x| \to \infty$ and the stochastic process $\beta_T(t) = \frac{1}{\sqrt{T}\psi(\sqrt{T})} \int_0^{tT} g_0\, ds = \frac{1}{T} g_0 \int_0^{tT} ds$ converges as $T \to \infty$ weakly to the process

$$\beta(t) = 2b \left[\frac{r^{\alpha+1}(t)}{\alpha+1} - \int_0^t r^\alpha(s)\, d\widehat{W}(s) \right] = \frac{2}{3} g_0 \left[\frac{r^2(t)}{2} - \int_0^t r(s)\, d\widehat{W}(s) \right]$$

where $r(t) \ge 0$ is the solution of stochastic differential equation

$$dr^2(t) = 3\, dt + 2r(t)\, d\widehat{W}(t).$$

In this case $\beta(t) = g_0 t$.

2) Let $c_0 = \frac{1}{2}$, so $f(x) = (1 + x^2)^{\frac{-1}{2}}$. If $g(x) = g_0$, $\psi(|x|) = |x|$, $b = \frac{g_0}{2}$ and

$$q(x) = \frac{1}{|x|\sqrt{1+x^2}} g_0 \int_0^x \sqrt{1+u^2}\, du - \frac{g_0}{2} \operatorname{sign} x \to 0 \quad \text{as } |x| \to \infty,$$

then the stochastic process $\beta_T(t)$ converges as $T \to \infty$ weakly to the process $\beta(t) = g_0 t$.

3) If $c_0 = 1$, $g(x) = \sin^2 x$, $\psi(|x|) = |x|$, $b = \frac{1}{6}$, then

$$q(x) = \frac{1}{|x|(1+x^2)} \int_0^x \left(1+u^2\right) \sin^2 u\, du - \frac{1}{6} \operatorname{sign} x \to 0 \quad \text{as } |x| \to \infty.$$

The stochastic process $\beta_T(t) = \frac{1}{\sqrt{T}\psi(\sqrt{T})} \int_0^{tT} \sin^2(\xi(s))\, ds = \frac{1}{T} \int_0^{tT} \sin^2(\xi(s))\, ds$ converges as $T \to \infty$ weakly to the process

$$\beta(t) = \frac{1}{3} \left[\frac{r^2(t)}{2} - \int_0^t r(s)\, d\widehat{W}(s) \right] = \frac{t}{2}.$$

3. Auxiliary result

Now we prove an auxiliary result concerning regularly varying functions $\psi(r)$, $r \ge 0$, that was applied in the proof of Theorem 2.1.

Lemma 3.1. *Let the function $\psi(r)$, $r \ge 0$ be positive, non-decreasing and regularly varying (at infinity) with index $\alpha \ge 0$. Then for an arbitrary $N > 0$ there exist constants $C_N < \infty$, $0 < T_N < \infty$ such that uniformly on $T \ge T_N$*

$$\sup_{0 \le r \le N} \frac{\psi(r\sqrt{T})}{\psi(\sqrt{T})} \le C_N.$$

Proof. It is clear that

$$\sup_{0 \le r \le N} \frac{\psi(r\sqrt{T})}{\psi(\sqrt{T})} \le \frac{\psi(N\sqrt{T})}{\psi(\sqrt{T})}.$$

Since for regularly varying function $\psi(r)$ we have convergence

$$\frac{\psi(N\sqrt{T})}{\psi(\sqrt{T})} \to N^\alpha,$$

as $T \to \infty$, then for $\varepsilon = 1$ there exists a constant $T_N < \infty$ such that for all $T \ge T_N$ the following inequality holds true

$$\frac{\psi(N\sqrt{T})}{\psi(\sqrt{T})} \le N^\alpha + 1.$$

Hence the statement of Lemma 3.1 is proved for $C_N = N^\alpha + 1$. □

References

1. I. I. Gikhman and A. V. Skorokhod, *Introduction to the Theory of Random Processes*, W. B. Saunders Co., Philadelphia, Pa.–London–Toronto, Ont., 1969.
2. I. I. Gikhman and A. V. Skorokhod, *Stochastic Differential Equations and their Applications*, "Naukova dumka", Kiev, 1982. (Russian)
3. J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, 2-nd ed., Springer, Berlin, 2003.
4. J. Karamata, *Sur un mode de croissance reguliere des fonctions*, Mathematica (Cluj) **4** (1930), 38–53.
5. E. Kas'kun, *On the asymptotic behaviour of functionals of integral type on diffusion processes*, Visn., Mat. Mekh., Kyiv. Univ. Im. Tarasa Shevchenka **1** (1998), 16–24. (Ukrainian)
6. N. V. Krylov, *Controlled Diffusion Processes*, Springer, 2008.
7. G. L. Kulinich, *On asymptotic behaviour of unstable solutions to a system of stochastic diffusion equations*, Proceedings of the School-Seminar on the Theory of Random Processes, Druskininkai, 25–30 November, 1974, Institute of Physics and Mathematics of LitSSR, Vilnius, 1975, pp. 169–201.
8. G. L. Kulinich, *Limit distributions for functionals of integral type of unstable diffusion processes*, Theor. Prob. and Math. Statist. **11** (1976), 82–86.
9. G. L. Kulinich, *On the asymptotic behavior of the solution of one dimensional stochastic diffusion equation*, Lecture Notes in Control and Information Sciencs, vol. 25, Springer-Verlag 1980, pp. 334–343.
10. G. L. Kulinich, *On necessary and sufficient conditions for convergence of homogeneous additive functionals of diffusion processes*, Proceedings of the second Ukrainian-Hungarian conference: New trends in probability and mathematical statistics (M. Arato and M. Yadrenko, eds.), vol. 2, "TViMS", Kyiv, 1995, pp. 381–390.
11. G. L. Kulinich and E. P. Kas'kun, *On the asymptotic behavior of solutions of a certain class of one-dimensional Ito stochastic differential equations*, Theory Probab. Math. Stat. **56** (1998), 97–105.
12. G. U. Minbaeva, *Limit theorems for integral functionals of a process with instantaneous reflection*, Theory Probab. Math. Stat. **47** (1993), 101–105.
13. A. V. Skorokhod, *Studies in the Theory of Random Processes*, Addison-Wesley Publ. Co., Inc., Reading. Mass., 1965.
14. A. Yu. Veretennikov, *On the strong solutions of stochastic differential equations*, Theory Probab. Appl. **24** (1979), no. 2, 354–366.
15. A. K. Zvonkin, *A transformation of the phase space of a diffusion process that removes the drift*, Mathematics of the USSR, Sbornik **93** (1974), no. 1, 129–149.

Department of General Mathematics, Faculty of Mechanics and Mathematics, Taras Shevchenko National University of Kyiv, 64/13, Volodymyrska Street, Kyiv, Ukraine, 01601

E-mail address: zag_mat@univ.kiev.ua

Department of General Mathematics, Faculty of Mechanics and Mathematics, Taras Shevchenko National University of Kyiv, 64/13, Volodymyrska Street, Kyiv, Ukraine, 01601

E-mail address: bksv@univ.kiev.ua

Department of Probability Theory, Statistics and Actuarial Mathematics, Faculty of Mechanics and Mathematics, Taras Shevchenko National University of Kyiv, 64/13 Volodymyrska Street, Kyiv, Ukraine, 01601

E-mail address: yumishura@gmail.com

Received 12/12/2012

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 104–114

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 104–114

POISSON APPROXIMATION OF PROCESSES WITH LOCALLY INDEPENDENT INCREMENTS WITH MARKOV SWITCHING

UDC 519.21

N. LIMNIOS AND I. V. SAMOILENKO

Abstract. In this paper, the weak convergence of additive functionals of processes with locally independent increments and with Markov switching in the scheme of Poisson approximation is investigated. Singular perturbation problem for the generator of Markov process is used to prove the relative compactness.

Анотація. В роботі досліджено слабку збіжність адитивних функціоналів від процесів з локально незалежними приростами та марковським перемиканням в схемі пуассонової апроксимації. Для доведення відносної компактності процесу використано задачу сингулярного збурення для генератора марковського процесу.

Аннотация. В работе исследована слабая сходимость аддитивных функционалов от процессов с локально независимыми приращениями и марковским переключением в схеме пуассоновской аппроксимации. При доказательстве относительной компактности процесса используется задача сингулярного возмущения для генератора марковского процесса.

1. Introduction

Let us consider the following stochastic additive functional

$$\xi(t) = \xi(0) + \int_0^t \eta(ds; x(s)), \qquad t \ge 0,$$

where $x(t)$, $t \ge 0$, is a jump Markov process with the state space $(E, \mathcal{E})$ and $\eta(t, x)$ is a family of processes with independent increments, $x \in E$, $t \ge 0$ with state space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. This is an important process since we have as particular cases the following well-known stochastic systems:

• The integral functional

$$\alpha(t) = \int_0^t a(x(s))\, ds, \qquad t \ge 0$$

where a is a deterministic measurable function defined on $(E, \mathcal{E})$.

• The dynamical system

$$\dot{u}(t) = C(u(t), x(t)), \qquad t \ge 0,$$

where C is a deterministic $\mathbb{R}^d$-function defined on $\mathbb{R}^d \times E$.

• The compound Poisson process

$$\zeta(t) = \sum_{k=1}^{\nu(t)} a(x_k),$$

where x_k is the embedded Markov chain of the jump Markov process $x(t)$.

2000 *Mathematics Subject Classification.* Primary 60J55, 60B10, 60F17, 60K10; Secondary 60G46, 60G60.

Key words and phrases. Poisson approximation, semimartingale, Markov process, independent increments process, piecewise deterministic Markov process, weak convergence, singular perturbation.

In this paper we establish weak convergence results in a semimartingale framework. In fact, we prove that for time scaled switching Markov process $x(t/\varepsilon)$, the additive semimartingale $\xi^\varepsilon(t)$, $t \geq 0$, $\varepsilon > 0$, weakly converges to a Poisson process with drift. The main difference from the results obtained in [8] is infinity of the measure of jumps corresponding to the processes with locally independent increments (see definition (2) below). The large deviations problem for the processes of this type was studied in [12].

The proof is given in two steps. In the first one we obtain relative compactness of the semimartingales representation of the family ξ^ε, $\varepsilon > 0$, by proving the following two facts [4]:

$$\lim_{c\to\infty} \sup_{\varepsilon \leq \varepsilon_0} \mathsf{P}\left\{\sup_{t\leq T} |\xi^\varepsilon(t)| > c\right\} = 0,$$

known as the compact containment condition (CCC), and

$$\mathsf{E}\,|\xi^\varepsilon(t) - \xi^\varepsilon(s)|^2 \leq k|t-s|,$$

for some positive constant $k > 0$. But due to infinity of the measure of jumps of the process we should check additional conditions (see Theorem A in Appendix).

In the second step we prove convergence of predictable characteristics of the semimartingales, which are integral functionals of the form:

$$\int_0^t a(\xi^\varepsilon(s), x^\varepsilon(s))ds,$$

by using singular perturbation technique as presented in [7].

Finally, we apply Theorem IX.3.27 from Jacod and Shiryayev [6] in order to prove the weak convergence of semimartingale.

The original part of this work is the use of relative compactness proof scheme given for averaging approximation (Bogolubov) to obtain a Poisson approximation result. Moreover, this kind of additive functionals are very useful in practice since they include the well-known stochastic systems.

The paper is organized as follows. In Section 2 we present the process with locally independent increments and the switching Markov process. In the same section we present the main results of Poisson approximation. In Section 3 we present the proof of the theorem. Two theorems we refer to are presented in the Appendix.

2. Main results

Let us consider the set of real numbers $\mathbb{R}$, and $(E, \mathcal{E})$, a *standard state space*, (i.e., E is a Polish space and $\mathcal{E}$ its Borel σ-algebra). Let $C_3(\mathbb{R})$ be a measure-determining class of real-valued bounded functions, such that $g(u)/u^2 \to 0$, as $|u| \to 0$ for $g \in C_3(\mathbb{R})$ and $C_2(\mathbb{R})$ be a measure-determining class of all continuous bounded functions which are 0 around 0 (see [6, 7]). We note that $C_2(\mathbb{R}) \subset C_3(\mathbb{R})$.

The additive functional $\xi^\varepsilon(t)$, $t \geq 0$, $\varepsilon > 0$, on $\mathbb{R}$ in the series scheme with small series parameter $\varepsilon \to 0$, $\varepsilon > 0$, is defined by the stochastic additive functional ([7, Section 3.3.1])

$$\xi^\varepsilon(t) = \xi_0^\varepsilon + \int_0^t \eta^\varepsilon(ds; x(s/\varepsilon)). \tag{1}$$

The family of processes with *locally independent increments* $\eta^\varepsilon(t;x)$, $t \geq 0$, $x \in E$, on $\mathbb{R}$, is defined by the generators (see [1, Section I.2], [7, Section 1.2.4])

$$\mathbf{\Gamma}^\varepsilon(x)\varphi(u) = b_\varepsilon(u;x)\varphi'(u) + \int_{\mathbb{R}} \left[\varphi(u+v) - \varphi(u) - v\varphi'(u)\mathbb{1}_{(|v|\leq 1)}\right] \Gamma^\varepsilon(u, dv; x), \tag{2}$$

where $\varphi(u)$ is real-valued twice differentiable function on $\mathbb{R}$ vanishing at infinity, with the sup-norm $\|\varphi\| = \sup_{u\in\mathbb{R}} |\varphi(u)|$, $\varphi(u) \in C_0^2(\mathbb{R})$, $b_\varepsilon(u;x) = \int_{\mathbb{R}} v\,\Gamma^\varepsilon(u, dv; x)$, and

$\Gamma^\varepsilon(u, dv; x)$ is the intensity kernel that satisfies the condition

$$\Gamma^\varepsilon(u, \{0\}; x) = 0.$$

Let $\mathbf{B}$ be the Banach space, that is a complete linear normed space, of all bounded real-valued measurable functions on E, with the sup-norm $\|\varphi\| = \sup_{x\in E} |\varphi(x)|$, $\varphi(x) \in \mathbf{B}$. The switching Markov process $x(t)$, $t \geq 0$, on the standard phase space $(E, \mathcal{E})$, is defined by the generator

$$\mathbf{Q}\varphi(x) = q(x) \int_E P(x, dy)[\varphi(y) - \varphi(x)], \tag{3}$$

where $q(x)$, $x \in E$, is the intensity of jumps function of $x(t)$, $t \geq 0$, and $P(x, dy)$ is the transition kernel of the embedded Markov chain x_n, $n \geq 0$, defined by $x_n = x(\tau_n)$, $n \geq 0$, with $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_n \leq \dots$ the jump times of $x(t)$, $t \geq 0$.

It is worth noticing that the coupled process $\xi^\varepsilon(t)$, $x(t/\varepsilon)$, $t \geq 0$, is a Markov additive process (see, e.g., [7, Section 2.5]).

Let Π be a projector onto null-subspace of reducible-invertible operator Q (see in details [7, Section 1.2]), defined in (3):

$$\Pi\varphi(x) = \int_E \pi(dx)\varphi(x).$$

The following relation is true

$$Q\Pi = \Pi Q = 0.$$

The Poisson approximation of Markov additive process (2) is considered under the following conditions.

C1: The Markov process $x(t)$, $t \geq 0$, is uniformly ergodic with $\pi(B)$, $B \in \mathcal{E}$, its stationary distribution.

C2: *Poisson approximation.* The family of processes with locally independent increments $\eta^\varepsilon(t; x)$, $t \geq 0$, $x \in E$, satisfies the Poisson approximation conditions [7, Section 7.2.3]:

PA1: Approximation of the mean values:

$$b_\varepsilon(u; x) = \int_{\mathbb{R}} v\, \Gamma^\varepsilon(u, dv; x) = \varepsilon[b(u; x) + \theta_b^\varepsilon(u; x)],$$

and

$$c_\varepsilon(u; x) = \int_{\mathbb{R}} v^2 \Gamma^\varepsilon(u, dv; x) = \varepsilon[c(u; x) + \theta_c^\varepsilon(u; x)].$$

PA2: Poisson approximation condition for intensity kernel

$$\Gamma_g^\varepsilon(u; x) = \int_{\mathbb{R}} g(v)\, \Gamma^\varepsilon(u, dv; x) = \varepsilon[\Gamma_g(u; x) + \theta_g^\varepsilon(u; x)]$$

for all $g \in C_3(\mathbb{R})$, and the function $\Gamma_g(u; x)$ is bounded for each $g \in C_3(\mathbb{R})$, that is,

$$|\Gamma_g(u; x)| \leq C_g \quad \text{(a constant depending on } g\text{)}.$$

The kernel $\Gamma(u, dv; x)$ is defined on the class $C_3(\mathbb{R})$ by the relation

$$\Gamma_g(u; x) = \int_{\mathbb{R}} g(v)\, \Gamma(u, dv; x), \qquad g \in C_3(\mathbb{R}).$$

The above negligible terms θ_g^ε, θ_b^ε, θ_c^ε satisfy the condition

$$\sup_{x\in E} |\theta_\cdot^\varepsilon(u; x)| \to 0, \qquad \varepsilon \to 0.$$

In addition the following conditions are used:

C3: *Uniform square-integrability*:

$$\lim_{c\to\infty} \sup_{x\in E} \int_{|v|>c} v^2\, \Gamma(u, dv; x) = 0.$$

C4: *Growth condition*: there exists a positive constant L such that

$$|b(u;x)| \le L(1+|u|) \quad \text{and} \quad |c(u;x)| \le L(1+|u|^2),$$

C5: *Linear growth of kernel*: we assume that $\Gamma(u,B;x)$ is absolutely continuous with respect to Lebesgue measure dv in $\mathbb{R}$, that is,

$$\Gamma(u,dv;x) = \Lambda(u,v;x)\,dv,$$

thus $\Lambda(u,v;x)$ is the Radon–Nikodym derivative of $\Gamma(u,B;x)$ and the following inequality holds:

$$|\Lambda(u,v;x)| \le Lf(v)(1+|u|)$$

for any real-valued non-negative function $f(v)$, $v\in\mathbb{R}$, such that

$$\int_{\mathbb{R}\setminus\{0\}} (1+f(v))v^2\,dv < \infty.$$

The main result of our work is the following.

Theorem 1. *Under conditions* **C1**–**C5** *the weak convergence*

$$\xi^\varepsilon(t) \Rightarrow \xi^0(t), \qquad \varepsilon\to 0$$

takes place.

The limit process $\xi^0(t)$, $t\ge 0$, is defined by the generator

$$\overline{\mathbf{\Gamma}}\varphi(u) = \widehat{b}(u)\varphi'(u) + \int_{\mathbb{R}} \left[\varphi(u+v)-\varphi(u)-v\varphi'(u)\mathbb{1}_{(|v|\le 1)}\right]\widehat{\Gamma}(u,dv), \tag{4}$$

where the average deterministic drift is defined by

$$\widehat{b}(u) = \Pi b(u;x) = \int_E \pi(dx)b(u;x),$$

and the average intensity kernel is defined by

$$\widehat{\Gamma}(u,dv) = \Pi\,\Gamma(u,dv;x) = \int_E \pi(dx)\Gamma(u,dv;x).$$

REMARK 1. The limit generator in the Euclidean space $\mathbb{R}^d$, $d>1$, is represented in the following view:

$$\overline{\mathbf{\Gamma}}\varphi(u) = \sum_{k=1}^d \widehat{b}_k(u)\varphi'_k(u) + \int_{\mathbb{R}^d}\left[\varphi(u+v)-\varphi(u)-\sum_{k=1}^d v_k\varphi'_k(u)\mathbb{1}_{(|v|\le 1)}\right]\widehat{\Gamma}(u,dv),$$

$$\varphi'_k(u) := \partial\varphi(u)/\partial u_k, \qquad 1\le k\le d.$$

3. Proof of Theorem 1

The proof of Theorem 1 is based on the semimartingale representation of the additive functional process (1). The method used here to prove the weak convergence is quite different from the method proposed by other authors ([4]–[6], [9]–[18]): the main point is to prove convergence of predictable characteristics of semimartingales which are integral functionals of some switching Markov processes.

According to Theorems 6.27 and 7.16 [2] the predictable characteristics of the semimartingale (2) have the following representations:

- $B^\varepsilon(t) = \varepsilon^{-1}\int_0^t b_\varepsilon(\xi^\varepsilon(s);x_s^\varepsilon)\,ds = \int_0^t b(\xi^\varepsilon(s);x_s^\varepsilon)\,ds + t\theta_b^\varepsilon$ — the first predictable characteristic;
- $C^\varepsilon(t) = \varepsilon^{-1}\int_0^t c_\varepsilon(\xi^\varepsilon(s);x_s^\varepsilon)\,ds = \int_0^t c(\xi^\varepsilon(s);x_s^\varepsilon)\,ds + t\theta_c^\varepsilon$ — the second modified characteristic;
- $\Gamma^\varepsilon(t) = \varepsilon^{-1}\int_0^t\int_{\mathbb{R}} h(v)\,\Gamma^\varepsilon(\xi^\varepsilon(s),dv;x_s^\varepsilon)\,ds = \int_0^t\int_{\mathbb{R}} h(v)\,\Gamma(\xi^\varepsilon(s),dv;x_s^\varepsilon)\,ds + t\theta_h^\varepsilon$, where $x_t^\varepsilon := x(t/\varepsilon)$, $t\ge 0$, and $\sup_{x\in E}|\theta_\cdot^\varepsilon|\to 0$, $\varepsilon\to 0$, $h(v)$ is the truncated function.

The jump martingale part of the semimartingale (2) is represented as follows

$$\mu^\varepsilon(t) = \int_0^t \int_{|v|\le 1} v \left[\mu^\varepsilon\left(\xi^\varepsilon(s), ds, dv; x_s^\varepsilon\right) - \Gamma^\varepsilon\left(\xi^\varepsilon(s), dv; x_s^\varepsilon\right)\, ds\right].$$

Here $\mu^\varepsilon(u, ds, dv; x)$, $x \in E$, is the family of counting measures of jumps of the process, namely

$$\mathsf{E}\,\mu^\varepsilon(u, ds, dv; x) = \Gamma^\varepsilon(u, dv; x)\, ds.$$

We can see now that predictable characteristics depend on the process $\xi^\varepsilon(s)$. Thus, to prove convergence of $\xi^\varepsilon(s)$ we should prove convergence of predictable characteristics dependent on $\xi^\varepsilon(s)$. To avoid this difficulty, we combine two methods. The one based on semimartingales theory, is combined with a solution of singular perturbation problem instead of ergodic theorem.

We split the proof of Theorem 1 in the following two steps.

3.1. **Relative compactness.** At this step we establish the relative compactness of the family of processes $\xi^\varepsilon(t)$, $t \ge 0$, $\varepsilon > 0$, by using the approach developed in [10]. Let us remind that the space of all probability measures defined on the standard space $(E, \mathcal{E})$ is also a Polish space; so the relative compactness and tightness are equivalent.

Proposition 1. *Under assumption* **C4,C5,PA1**, *the following compact containment condition (CCC) holds:*

$$\lim_{c\to\infty} \lim_{\varepsilon\to 0} \mathsf{P}\left\{\sup_{t\le T} |\xi^\varepsilon(t)| > c\right\} = 0. \tag{5}$$

Proof. The proof of this corollary follows from Kolmogorov's inequality by using the estimation of Lemma 1. □

Lemma 1. *Under assumption* **C4**, **C5**, **PA1** *there exists a constant* $k_T > 0$, *independent of* ε *and dependent on* T, *such that*

$$\mathsf{E} \sup_{t\le T} |\xi^\varepsilon(t)|^2 \le k_T.$$

Proof. (Following [10]). For a process $y(t)$, $t \ge 0$, let us define the process

$$y_t^\dagger = \sup_{s\le t} |y(s)|.$$

It follows from **PA1** and **C4** that for any fixed $t > 0$

$$\begin{aligned} \int_0^t \int_{\mathbb{R}} v^2 \Gamma^\varepsilon(\xi^\varepsilon(s), dv; x_s^\varepsilon)\, ds &= \varepsilon \int_0^t c(\xi^\varepsilon(s); x_s^\varepsilon) ds + \varepsilon t \theta_c^\varepsilon(u; x) \\ &\le \varepsilon t \left[L\left(1 + \left((\xi_t^\varepsilon)^\dagger\right)^2\right) + \theta_c^\varepsilon(u; x)\right] < \infty \quad \mathsf{P}\text{-a.s.} \end{aligned}$$

The increasing process $\int_0^t \int_{\mathbb{R}\setminus\{0\}} v^2 \Gamma^\varepsilon(\xi^\varepsilon(s), dv; x_s^\varepsilon) ds$ is continuous in t so that by Theorem 28 from [3, Ch.5] (see also Theorem 1.6.3 [11]) it is the compensator of

$$\int_0^t \int_{\mathbb{R}\setminus\{0\}} v^2 \mu^\varepsilon\left(\xi^\varepsilon(s), d(s/\varepsilon), dv; x_s^\varepsilon\right).$$

Therefore, (1) is the special semimartingale with the decomposition

$$\xi^\varepsilon(t) = u + A_t^\varepsilon + M_t^\varepsilon, \tag{6}$$

where $u = \xi^\varepsilon(0)$; A_t^ε is the predictable drift (see [4]):

$$A_t^\varepsilon = \int_0^t b\left(\xi^\varepsilon(s), x_s^\varepsilon\right)\, ds + \int_0^t \int_{|v|>1} v\Gamma\left(\xi^\varepsilon(s), dv; x_s^\varepsilon\right)\, ds + \theta_A^\varepsilon(t),$$

and M_t^ε is the locally square integrable martingale

$$M_t^\varepsilon = \int_0^t \int_{\mathbb{R}\setminus\{0\}} v[\mu(\xi^\varepsilon(s), ds, dv; x_s^\varepsilon) - \Gamma(\xi^\varepsilon(s), dv; x_s^\varepsilon)ds] + \theta_M^\varepsilon(t),$$

and for every finite $T > 0$

$$\sup_{0\le t\le T} |\theta_\cdot^\varepsilon(t)| \to 0, \qquad \varepsilon \to 0.$$

From (6) we have

$$\left((\xi_t^\varepsilon)^\dagger\right)^2 \le 3\left[u^2 + \left((A_t^\varepsilon)^\dagger\right)^2 + \left((M_t^\varepsilon)^\dagger\right)^2\right]. \tag{7}$$

Conditions **C4**–**C5** imply that

$$\begin{aligned}(A_t^\varepsilon)^\dagger &\le L\int_0^t \left(1 + (\xi_s^\varepsilon)^\dagger\right) ds + L\int_0^t \int_{|v|>1} |v| f(v) \left(1 + (\xi_s^\varepsilon)^\dagger\right) dv\, ds \\ &\le L(1 + r_1)\int_0^t \left(1 + (\xi_s^\varepsilon)^\dagger\right) ds,\end{aligned} \tag{8}$$

where $r_1 = \int_{\mathbb{R}\setminus\{0\}} v^2 f(v)\, dv$.

Now, by Doob's inequality (see, e.g., [11, Theorem 1.9.2]),

$$\mathsf{E}\left((M_t^\varepsilon)^\dagger\right)^2 \le 4\left|\mathsf{E}\langle M^\varepsilon\rangle_t\right|,$$

where by condition **C5** we obtain

$$|\langle M^\varepsilon\rangle_t| = \left|\int_0^t\int_{\mathbb{R}\setminus\{0\}} v^2\Gamma\left(\xi^\varepsilon(s), dv; x_s^\varepsilon\right) ds\right| \le Lr_1\int_0^t \left[1 + \left((\xi_s^\varepsilon)^\dagger\right)^2\right] ds. \tag{9}$$

Inequalities (7)–(9) and Cauchy–Bunyakovsky–Schwarz inequality,

$$\left[\int_0^t \varphi(s)\, ds\right]^2 \le t\int_0^t \varphi^2(s)\, ds$$

imply

$$\mathsf{E}\left((\xi_t^\varepsilon)^\dagger\right)^2 \le k_1 + k_2\int_0^t \mathsf{E}\left((\xi_s^\varepsilon)^\dagger\right)^2 ds,$$

where k_1 and k_2 are positive constants independent of ε.

By Gronwall inequality (see, e.g., [4, p. 498]), we obtain

$$\mathsf{E}\left((\xi_t^\varepsilon)^\dagger\right)^2 \le k_1 \exp(k_2 t).$$

Hence the lemma is proved. □

Lemma 2. *Under assumption* **C4, C5, PA1** *there exists a constant $k > 0$, independent of ε such that*

$$\mathsf{E}\,|\xi^\varepsilon(t) - \xi^\varepsilon(s)|^2 \le k|t - s|.$$

Proof. In the same manner with (7), we may write

$$|\xi^\varepsilon(t) - \xi^\varepsilon(s)|^2 \le 2|A_t^\varepsilon - A_s^\varepsilon|^2 + 2|M_t^\varepsilon - M_s^\varepsilon|^2.$$

By using Doob's inequality, we obtain

$$\mathsf{E}\,|\xi^\varepsilon(t) - \xi^\varepsilon(s)|^2 \le 2\,\mathsf{E}\{|A_t^\varepsilon - A_s^\varepsilon|^2 + 8\,|\langle M^\varepsilon\rangle_t - \langle M^\varepsilon\rangle_s|\}.$$

Now (8), (9), and assumption **C5** imply

$$|A_t^\varepsilon - A_s^\varepsilon|^2 + 8\,|\langle M^\varepsilon\rangle_t - \langle M^\varepsilon\rangle_s| \le k_3 \left[1 + \left((\xi_T^\varepsilon)^\dagger\right)^2\right] |t-s|,$$

where k_3 is a positive constant independent of ε.

From the last inequality and Lemma 1 the desired conclusion is obtained. □

Finally, we have to use the Theorem 8.2.1 from [11] that states the relative compactness of semimartingales (see Appendix).

Lemma 3. *Under conditions* ***C1–C5*** *the family of processes* $\xi^\varepsilon(t)$ *is relatively compact.*

Proof. To verify the relative compactness we should check the conditions **LP1–LP5** of Theorem A (see Appendix).

We easily see that **LP1** follows from the conditions **C4–C5**.

Show, that under (5) **LP2** and second part of **LP5** hold. Really, by the condition **C5** and the definition of $\Gamma(u,v;x)$ on the set $\{\sup_{t\le T}|\xi^\varepsilon(t)\le c|\}$ we have for the function $g(v) = \begin{cases} 0, & |v|\le l \\ 1, & |v|>l \end{cases}$:

$$\begin{aligned}\int_0^T \int_{|v|>l} \Gamma^\varepsilon\left(\xi^\varepsilon(s), dv; x_s^\varepsilon\right)\, ds &= \int_0^T \int_{\mathbb{R}} g(v)\Gamma^\varepsilon\left(\xi^\varepsilon(s), dv; x_s^\varepsilon\right)\, ds \\ &= \varepsilon \int_0^T \int_{\mathbb{R}} g(v)\Gamma(\xi^\varepsilon(s), dv; x_s^\varepsilon) ds + \varepsilon T\theta_g^\varepsilon \\ &\le \varepsilon TL(1+c)\int_{|v|>l} f(v)dv + \varepsilon T\theta_g^\varepsilon \\ &\le \varepsilon \frac{TL(1+c)}{l}\int_{\mathbb{R}} v^2 f(v)dv + \varepsilon T\theta_g^\varepsilon \to 0, \\ l\to\infty, \qquad \varepsilon\to 0.\end{aligned}$$

Using of condition **PA2** here is stipulated by the fact that $g(v)\in C_2(\mathbb{R}^d)\subset C_3(\mathbb{R}^d)$.

By the same way we get

$$\begin{aligned}\int_0^T \int_{|v|\le\delta} v^2\widehat{\Gamma}\left(\xi^\varepsilon(s), dv\right)\, ds &= \int_0^T \int_{|v|\le\delta}\int_E v^2\Gamma\left(\xi^\varepsilon(s), dv; x\right)\,\pi(dx)\,ds \\ &\le TL(1+c)\int_{|v|\le\delta} v^2 f(v)\,dv \to 0, \qquad \delta\to 0,\end{aligned}$$

and by the conditions $\Gamma^\varepsilon(u,\{0\};x)=0$, **PA1** and **C5**

$$\begin{aligned}\int_0^T \int_{|v|\le\delta} v^2\Gamma^\varepsilon\left(\xi^\varepsilon(s), dv; x_s^\varepsilon\right)\, ds &\le \int_0^T \int_{\mathbb{R}} v^2\Gamma^\varepsilon\left(\xi^\varepsilon(s), dv; x_s^\varepsilon\right)\, ds \\ = \varepsilon\int_0^T c(\xi^\varepsilon(s); x_s^\varepsilon)ds + \varepsilon T\theta_c^\varepsilon(u;x) &\le \varepsilon TL\left(1+c^2\right) + \varepsilon T\theta_c^\varepsilon(u;x)\to 0, \\ \varepsilon\to 0, \qquad \delta\to 0.\end{aligned}$$

It is clear, that **LP3**, **LP4** and the first part of **LP5** follows from the weak convergence of predictable characteristics. Thus, the final step in proof of this Lemma will be made in the next subsection by the verifying of Lemma 4. □

3.2. **Convergence of predictable characteristics.** The next step of proof concerns the convergence of the predictable characteristics. To do that, we apply the results of Sections 3.2–3.3 in [7] and Theorem 6.3 from [7] (see Appendix).

Lemma 4. *Let's point $A^\varepsilon(t)$ any of three predictable characteristics of the process $\xi^\varepsilon(t)$. The following weak convergence takes place*

$$A^\varepsilon(t) \Rightarrow A^0(t),$$

where

$$A^0(t) := \int_0^t \widehat{a}\left(\xi^0(s)\right)\, ds,$$

here

$$\widehat{a}(u) := \int_E \pi(dx) a(u;x).$$

Proof. We consider the three component Markov process $A^\varepsilon(t)$, $\xi^\varepsilon(t)$, x_t^ε, $t \geq 0$, which can be characterized by the martingale

$$\mu_t^\varepsilon = \varphi(A^\varepsilon(t), \xi^\varepsilon(t), x_t^\varepsilon) - \int_0^t \mathbf{L}^\varepsilon \varphi(A^\varepsilon(s), \xi^\varepsilon(s), x_t^\varepsilon)\, ds.$$

The generator $\mathbf{L}^\varepsilon$ of the martingale has the following representation [7]

$$\mathbf{L}^\varepsilon = \varepsilon^{-1}\mathbf{Q} + \mathbf{\Gamma}^\varepsilon + \mathbf{A}^\varepsilon, \tag{10}$$

with $\mathbf{\Gamma}^\varepsilon$ given by (3), $\mathbf{Q}$ given by (4), and $\mathbf{A}^\varepsilon(u;x)\varphi(v) = \mathbf{A}\varphi(v) + \widetilde{\theta}_a^\varepsilon$, where $\mathbf{A}\varphi(v) := a(u;x)\varphi'(v)$, and $\widetilde{\theta}_a^\varepsilon \to 0$, $\varepsilon \to 0$.

In order to prove the convergence of predictable characteristics, it is sufficient to study the action of the generator $\mathbf{L}^\varepsilon$ on test functions of two variables $\varphi(v,x)$.

Thus, it has the representation

$$\mathbf{L}^\varepsilon \varphi(v,x) = [\varepsilon^{-1}\mathbf{Q} + \mathbf{A}]\varphi(v,x) + \widetilde{\theta}_a^\varepsilon \varphi(v,x). \tag{11}$$

The solution of the singular perturbation problem at the test functions $\varphi^\varepsilon(v,x) = \varphi(v) + \varepsilon\varphi_1(v,x)$ in the form $\mathbf{L}^\varepsilon\varphi^\varepsilon = \widehat{\mathbf{L}}\varphi + \theta^\varepsilon\varphi$ can be found in the following manner. We have:

$$\begin{aligned}\mathbf{L}^\varepsilon\varphi^\varepsilon(v,x) &= [\varepsilon^{-1}\mathbf{Q} + \mathbf{A}][\varphi(v) + \varepsilon\varphi_1(v,x)]\\ &= \varepsilon^{-1}\mathbf{Q}\varphi(v) + [\mathbf{Q}\varphi_1(v,x) + \mathbf{A}\varphi(v)] + \varepsilon\mathbf{A}\varphi_1(u,x) + \widetilde{\theta}_a^\varepsilon\varphi(v,x).\end{aligned}$$

We may write down the following equalities:

$$\mathbf{Q}\varphi(v) = 0,$$
$$\mathbf{Q}\varphi_1(v,x) + \mathbf{A}\varphi(v) = \widehat{\mathbf{L}}\varphi.$$

From the first equality we see that the function $\varphi(v)$ belongs to the null-space of operator Q and thus does not depend on x. So, using the solvability condition, we have from the second equality

$$\widehat{\mathbf{L}}\varphi(v) = \Pi\mathbf{A}\Pi\varphi(v) + \Pi\mathbf{Q}\Pi\varphi_1(v,x) = \Pi\mathbf{A}\Pi\varphi(v).$$

That is

$$\widehat{\mathbf{L}} = \widehat{\mathbf{A}}, \tag{12}$$

where $\widehat{\mathbf{A}}\varphi(v) = \int_E \pi(dx)a(u;x)\varphi'(v)$.

Now Theorem B may be applied (see Appendix).

We see from (11) and (12) that the solution of singular perturbation problem for $\mathbf{L}^\varepsilon\varphi^\varepsilon(u,v;x)$ satisfies the conditions **CD1, CD2**. Condition **CD3** of this theorem implies that the quadratic characteristics of the martingale, corresponding to a coupled Markov process, is relatively compact. The same result follows from the CCC (see Corollary 1 and Lemma 2) by [6]. Thus, the condition **CD3** follows from the Corollary 1 and Lemma 2.

As soon as $A^\varepsilon(0) = A^0(0)$, $\xi^\varepsilon(0) = \xi^0(0)$ we see that the condition **CD4** is also satisfied. Thus, all the conditions of above Theorem 2 are satisfied, so the weak convergence $A^\varepsilon(t) \Rightarrow A^0(t)$ takes place.

Lemma is proved. □

Thus, by the weak convergence of predictable characteristics, we obtain **LP3**, **LP4** and the first part of **LP5**. As a result, by the Theorem 8.2.1 from [11] the process $\xi^\varepsilon(t)$ is relatively compact and Lemma 3 is proved.

The final step of the proof of Theorem 1 is achieved now by using Theorem IX.3.27 in [6]. Indeed all the conditions of this theorem are fulfilled.

As we have mentioned, the square integrability condition 3.24 follows from CCC (see [6]). The strong dominating hypothesis is true with the majoration functions are presented in the Conditions **C4**–**C5**. Condition **C5** implies the condition of big jumps for the last predictable measure of Theorem IX.3.27 in [6]. Conditions iv and v of Theorem IX.3.27 [6] are obviously fulfilled.

The weak convergence of predictable characteristics is proved by solving the singularly perturbation problem for the generator (10).

The last condition (3.29) of Theorem IX.3.27 is also fulfilled due to CCC proved in Proposition 1 and Lemma 2. Thus, the weak convergence is true.

We can see now that the limit Markov process is characterized by the following predictable characteristics

$$B^0(t) = \int_0^t \widehat{b}\left(\xi^0(s)\right)\, ds, \qquad C^0(t) = \int_0^t \widehat{c}\left(\xi^0(s)\right)\, ds, \qquad \Gamma_g^0(t) = \int_0^t \widehat{\Gamma}_g\left(\xi^0(s)\right)\, ds.$$

Here $C^0(t)$ is the second modified characteristic of the limit process. So, according to [2] the limit Markov process $\xi^0(t)$ can be expressed by the generator (4).

Theorem 1 is proved.

4. Appendix

Theorem A ([11, Theorem 8.2.1])**.** *Let Q^ε be the distribution of probabilities for $\mathbb{P}^\varepsilon$-semimartingale $\xi^\varepsilon = (\xi^\varepsilon(t), \mathcal{F}_t^\varepsilon)$ with the triplet of predictable characteristics $\mathcal{T}^\varepsilon = (B^\varepsilon, C^\varepsilon, \Gamma^\varepsilon)$ and Q is the distribution of semimartingale $\xi^0 = (\xi^0(t), \mathcal{D}_t^Q)$ with triplet $\mathcal{T}^0 = (B^0, C^0, \Gamma^0)$.*

If for the triplet $\mathcal{T}$ the following condition is true:

LP1:

$$\left|\int_E b(\xi(t), x)\, \pi(dx)\right| \le L(1 + \xi^\dagger(t)),$$

$$\left|\int_E c(\xi(t), x)\, \pi(dx)\right| \le L\left(1 + \left(\xi^\dagger(t)\right)^2\right),$$

and for any nonnegative measurable function $f(v) \le v^2 \wedge 1$

$$\int_E \int_{\mathbb{R}} f(v)\Gamma(\xi(t), dv; x)\, \pi(dx) \le L\left(1 + \xi^\dagger(t)\right).$$

And for the triplets $\mathcal{T}^\varepsilon$ for any fixed $T > 0$:

LP2:

$$\lim_{l\to\infty} \lim_{\varepsilon\to 0} \sup \int_0^T \int_{|v|>l} \Gamma^\varepsilon(\xi^\varepsilon(s), dv; x_s^\varepsilon)\, ds = 0,$$

$$\lim_{\delta\to 0} \lim_{\varepsilon\to 0} \sup \int_0^T \int_{|v|\le\delta} v^2 \widehat{\Gamma}(\xi^\varepsilon(s), dv)\, ds = 0.$$

LP3: *For every bounded measurable function* $f(v)$

$$\lim_{\varepsilon\to 0}\sup_{t\le T}\left|\int_0^t\int_{|v|>\delta} f(v)\left[\Gamma^\varepsilon\left(\xi^\varepsilon(s),dv;x_s^\varepsilon\right)-\widehat{\Gamma}\left(\xi^\varepsilon(s),dv\right)\right]\,ds\right|=0.$$

LP4:

$$\lim_{\varepsilon\to 0}\sup_{t\le T}\left|\int_0^t\left[b^\varepsilon\left(\xi^\varepsilon(s),x_s^\varepsilon\right)-\widehat{b}\left(\xi^\varepsilon(s)\right)\right]\,ds\right|=0.$$

LP5:

$$\lim_{\varepsilon\to 0}\sup_{t\le T}\left|\int_0^t\left[c^\varepsilon\left(\xi^\varepsilon(s),x_s^\varepsilon\right)-\widehat{c}\left(\xi^\varepsilon(s)\right)\right]ds\right|=0,$$

$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0}\sup\int_0^T\int_{|v|\le\delta} v^2\Gamma^\varepsilon\left(\xi^\varepsilon(s),dv;x_s^\varepsilon\right)\,ds=0,$$

then under compact containment condition

$$\lim_{c\to\infty}\overline{\lim_{\varepsilon\to 0}}\,\mathsf{P}\{|\xi^\varepsilon(0)|\ge c\}=0$$

the family Q^ε, $\varepsilon>0$, *is relatively compact.*

Theorem B ([7, Theorem 6.3])**.** *Put* $C_0^2(\mathbb{R}\times E)$ *be the space of real-valued twice continuously differentiable by the first argument functions, defined on* $\mathbb{R}\times E$ *and vanishing at infinity, and* $C(\mathbb{R}\times E)$ *is the space of real-valued continuous bounded functions defined on* $\mathbb{R}\times E$.

Let the following conditions hold for a family of coupled Markov processes $\xi^\varepsilon(t)$, $x^\varepsilon(t)$, $t\ge 0$, $\varepsilon>0$:

CD1: *There exists a family of test functions* $\varphi^\varepsilon(u,x)$ *in* $C_0^2(\mathbb{R}\times E)$, *such that*

$$\lim_{\varepsilon\to 0}\varphi^\varepsilon(u,x)=\varphi(u),$$

uniformly on u, x.

CD2: *The following convergence holds*

$$\lim_{\varepsilon\to 0}\mathbf{L}^\varepsilon\varphi^\varepsilon(u,x)=\mathbf{L}\varphi(u),$$

uniformly on u,x. *The family of functions* $\mathbf{L}^\varepsilon\varphi^\varepsilon$, $\varepsilon>0$, *is uniformly bounded, and* $\mathbf{L}\varphi(u)$ *and* $\mathbf{L}^\varepsilon\varphi^\varepsilon$ *belong to* $C(\mathbb{R}\times E)$.

CD3: *The quadratic characteristics of the martingales that characterize a coupled Markov process* $\xi^\varepsilon(t)$, $x^\varepsilon(t)$, $t\ge 0$, $\varepsilon>0$, *have the representation*

$$\langle\mu^\varepsilon\rangle_t=\int_0^t\zeta^\varepsilon(s)\,ds,$$

where the random functions ζ^ε, $\varepsilon>0$, *satisfy the condition*

$$\sup_{0\le s\le T}\mathsf{E}\,|\zeta^\varepsilon(s)|\le c<+\infty.$$

CD4: *The convergence of the initial values holds and*

$$\sup_{\varepsilon>0}\mathsf{E}\,|\zeta^\varepsilon(0)|\le C<+\infty.$$

Then the weak convergence

$$\xi^\varepsilon(t)\Rightarrow\xi(t),\qquad\varepsilon\to 0,$$

takes place.

References

1. J. Bertoin, *Lévy Processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996.
2. E. Çinlar, J. Jacod, P. Protter, and M. J. Sharpe, *Semimartingale and Markov processes*, Z. Wahrschein. verw. Gebiete **54** (1980), 161–219.
3. C. Dellacherie, *Capacités et processus stochastiques*, Springer-Verlag, Berlin–Heidelberg–New York, 1972.
4. S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, J. Wiley, New York, 1986.
5. I. I. Gihman and A. V. Skorohod, *Theory of Stochastic Processes*, vol. 1–3, Springer, Berlin, 1974.
6. J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer-Verlang, Berlin, 1987.
7. V. S. Koroliuk and N. Limnios, *Stochastic Systems in Merging Phase Space*, World Scientific Publishers, Singapore, 2005.
8. V. S. Koroliuk, N. Limnios, and I. V. Samoilenko, *Poisson approximation of processes with locally independent increments with Markov switching*, Theory of Stoch. Proc. **15(31)** (2009), no. 1, 40–48; corrections in *Letter from the publishers*, Theory of Stoch. Proc. **15(31)** (2009), no. 2, 164.
9. H. J. Kushner, *Weak Convergence Methods and Singular Perturbed Stochastic Control and Filtering Problems*, Birkhäuser, Boston, 1990.
10. R. Sh. Liptser, *The Bogolubov averaging principle for semimartingales*, Proceedings of the Steklov Institute of Mathematics (1994), no. 4.
11. R. Sh. Liptser and A. N. Shiryayev, *Theory of Martingales*, Kluwer Academic Publishers, Dordrecht, the Netherlands, 1989.
12. I. V. Samoilenko, *Large deviations for random evolutions with independent increments in the scheme of the Poisson approximation*, Theor. Probability and Math. Statist. **85** (2012), 107–114.
13. K.-I. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 1999.
14. D. W. Stroock and S. R. S. Varadhan, *Multidimensional Diffusion Processes*, Springer-Verlag, Berlin, 1979.
15. A. V. Skorokhod, *Asymptotic Methods in the Theory of Stochastic Differential Equations*, AMS, Providence, 1989.
16. A. V. Skorokhod, F. C. Hoppensteadt, and H. Salehi, *Random Perturbation Method with Application in Science and Engineering*, Springer, 2002.
17. D. S. Silvestrov, *Limit Theorems for Randomly Stopped Stochastic Processes*, Probability and its Applications, Springer, 2004.
18. M. N. Sviridenko, *Martingale approach to limit theorems for semi-Markov processes*, Theor. Probab. Appl. (1986), 540–545.

Laboratoire de Mathématiques Appliquées, Université de Technologie de Compiègne, France

E-mail address: nlimnios@dma.utc.fr

Institute of Mathematics, Ukrainian National Academy of Science, Kyiv, Ukraine

E-mail address: isamoil@imath.kiev.ua

Received 18/10/2012

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 115–129

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 115–129

MINIMAX-ROBUST FILTERING PROBLEM FOR STOCHASTIC SEQUENCE WITH STATIONARY INCREMENTS

UDC 519.21

M. M. LUZ AND M. P. MOKLYACHUK

Abstract. The problem of optimal estimation of the linear functional $A\xi=\sum_{k=0}^{\infty}a(k)\xi(-k)$ which depends on unknown values of a stochastic sequence $\xi(k)$ with stationary nth increments from observations of the sequence $\xi(k)+\eta(k)$ at points of time $k=0,-1,-2,\dots$ is considered. Formulas for calculation the mean-square error and spectral characteristic of the optimal linear estimate of the functional are derived under the condition of spectral certainty, where spectral densities of the sequences $\xi(k)$ and $\eta(k)$ are exactly known. The minimax (robust) method of estimation is applied in the case where spectral densities are not known exactly, but sets of admissible spectral densities are given. Formulas that determine the least favorable spectral densities and the minimax spectral characteristics are proposed for some special sets of admissible spectral densities.

Анотація. Досліджується задача оптимального оцінювання функціонала $A\xi=\sum_{k=0}^{\infty}a(k)\xi(-k)$ від невідомих значень стохастичної послідовності $\xi(k)$ зі стаціонарними n-ми приростами за спостереженнями послідовності $\xi(k)+\eta(k)$ у моменти часу $k=0,-1,-2,\dots$. Знайдені формули для обчислення середньоквадратичної похибки та спектральної характеристики оптимальної оцінки функціонала за умови спектральної визначеності, тобто коли спектральні щільності послідовностей $\xi(m)$ та $\eta(m)$ відомі. У тому випадку, коли спектральні щільності невідомі, а задані лише множини допустимих спектральних щільностей, застосовано мінімаксний метод оцінювання. Для заданих множин допустимих спектральних щільностей визначені найменш сприятливі спектральні щільності та мінімаксні спектральні характеристики оптимальної лінійної оцінки функціонала.

Аннотация. Исследуется задача оптимального оценивания функционала $A\xi=\sum_{k=0}^{\infty}a(k)\xi(-k)$ от неизвестных значений стохастической последовательности $\xi(k)$ со стационарными n-ми приращениями по наблюдениям последовательности $\xi(k)+\eta(k)$ в моменты времени $k=0,-1,-2,\dots$. Найдены формулы для вычисления среднеквадратической ошибки и спектральной характеристики оптима-льной оценки функционала в том случае когда спектральные плотности последовательностей $\xi(m)$ и $\eta(m)$ точно известны. В том случае, когда спектральные плотности неизвестны, а заданы лишь множества допустимых спектральных плотностей, используется минимаксный метод оценивания. Для заданных множеств допустимых спектральных плотностей определены наименее благоприятные спектральные плотности и минимаксные спектральные характеристики оптимальной линейной оценки функционала.

1. Introduction

Traditional methods of solution of extrapolation, interpolation and filtering problems for stationary stochastic processes and sequences were developed by A. N. Kolmogorov [11], N. Wiener [26], A. M. Yaglom [28] under the condition of spectral certainty where spectral densities of the considered stochastic processes are exactly known. In the case where spectral densities are not exactly known, but a set of admissible spectral densities is given, we can apply the minimax method for solving extrapolation, interpolation and filtering problems, which allows us to determine estimates that minimize the value of the mean-square error for all densities from a given class.

2010 *Mathematics Subject Classification.* Primary 60G10, 60G25, 60G35; Secondary 62M20, 93E10, 93E11.

Key words and phrases. Sequences with stationary increments, robust estimate, mean-square error, least favorable spectral density, minimax spectral characteristic.

A survey of results in minimax (robust) methods of data processing is proposed by S. A. Kassam and H. V. Poor [10]. The paper by Ulf Grenander [7] should be marked as the first one where the minimax approach to extrapolation problem for stationary processes was developed. J. Franke [8], J. Franke and H. V. Poor [9] investigated the minimax extrapolation and filtering problems for stationary sequences with the help of convex optimization methods. In the works by M. P. Moklyachuk [14]–[17] problems of extrapolation, interpolation and filtering for stationary processes and sequences were studied. Methods of solution the minimax-robust estimation problems for vector-valued stationary sequences and processes were developed by M. P. Moklyachuk and O. Yu. Masyutka [19]–[23]. Methods of solution the minimax-robust estimation problems (extrapolation, interpolation and filtering) for linear functionals which depend on unknown values of periodically correlated stochastic processes were proposed by I. I. Dubovets'ka and M. P. Moklyachuk [2]–[6]. M. M. Luz and M. P. Moklyachuk [12]–[13] investigated the minimax interpolation problem for stochastic sequences $\xi(m)$ with stationary n-th increments from observations of the sequence with an additive noise and from observations without noise.

In this paper we investigate the problem of optimal linear filtering of a functional $A\xi=\sum_{k=0}^{\infty}a(k)\xi(-k)$ which depends on unobserved values of a stochastic sequence $\xi(m)$ with nth stationary increments based on observations of the sequence $\xi(k)+\eta(k)$ at points $k=0,-1,-2,\dots,$ where $\eta(k)$ is a stochastic sequence with stationary nth increments which is uncorrelated with the sequence $\xi(k)$. This filtering problem is solved in the case of spectral certainty where spectral densities of sequences $\xi(m)$ and $\eta(m)$ are exactly known as well as in the case of spectral uncertainty where spectral densities of sequences are not exactly known, but a set of admissible spectral densities is given. Formulas that determine the least favorable spectral densities and minimax (robust) spectral characteristics of the optimal linear estimate of the functional are proposed in the case of spectral uncertainty for concrete classes of admissible spectral densities.

2. Stochastic stationary increment sequence. Spectral representation

Stochastic processes with stationary n-th increments were introduced and investigated by A. M. Yaglom [27], M. S. Pinsker [25], A. M. Yaglom and M. S. Pinsker [24].

Definition 2.1. For a given stochastic sequence $\{\xi(m), m\in\mathbb{Z}\}$ a sequence

$$\xi^{(n)}(m,\mu)=(1-B_{\mu})^n\xi(m)=\sum_{l=0}^{n}(-1)^lC_n^l\xi(m-l\mu), \tag{1}$$

where B_{μ} is a backward shift operator with step $\mu\in\mathbb{Z}$, such that $B_{\mu}\xi(m)=\xi(m-\mu)$, is called stochastic nth increment sequence with step $\mu\in\mathbb{Z}$.

For the stochastic nth increment sequence $\xi^{(n)}(m,\mu)$ the following relations hold true:

$$\xi^{(n)}(m,-\mu)=(-1)^n\xi^{(n)}(m+n\mu,\mu), \tag{2}$$

$$\xi^{(n)}(m,k\mu)=\sum\nolimits_{l=0}^{(k-1)n}A_l\xi^{(n)}(m-l\mu,\mu),\qquad k\in\mathbb{N}, \tag{3}$$

where coefficients $\{A_l, l=0,1,2,\dots,(k-1)n\}$ are determined by the representation

$$(1+x+\dots+x^{k-1})^n=\sum_{l=0}^{(k-1)n}A_lx^l.$$

Definition 2.2. The stochastic nth increment sequence $\xi^{(n)}(m,\mu)$ generated by stochastic sequence $\{\xi(m), m\in\mathbb{Z}\}$ is wide sense stationary if the mathematical expectations

$$\mathsf{E}\xi^{(n)}(m_0,\mu)=c^{(n)}(\mu),$$

$$\mathsf{E}\xi^{(n)}(m_0+m,\mu_1)\xi^{(n)}(m_0,\mu_2)=D^{(n)}(m,\mu_1,\mu_2)$$

exist for all m_0, μ, m, μ_1, μ_2 and do not depend on m_0. The function $c^{(n)}(\mu)$ is called the mean value of the nth increment sequence and the function $D^{(n)}(m,\mu_1,\mu_2)$ is called the structural function of the stationary nth increment sequence (or the structural function of nth order of the stochastic sequence $\{\xi(m), m\in\mathbb{Z}\}$).

The stochastic sequence $\{\xi(m), m\in\mathbb{Z}\}$ which determines the stationary nth increment sequence $\xi^{(n)}(m,\mu)$ by formula (1) is called sequence with stationary nth increments.

Theorem 2.1. *The mean value $c^{(n)}(\mu)$ and the structural function $D^{(n)}(m,\mu_1,\mu_2)$ of the stochastic stationary nth increment sequence $\xi^{(n)}(m,\mu)$ can be represented in the following forms:*

$$c^{(n)}(\mu) = c\mu^n, \tag{4}$$

$$D^{(n)}(m,\mu_1,\mu_2) = \int_{-\pi}^{\pi} e^{i\lambda m}\left(1-e^{-i\mu_1\lambda}\right)^n\left(1-e^{i\mu_2\lambda}\right)^n\frac{1}{\lambda^{2n}}\,dF(\lambda), \tag{5}$$

where c is a constant, $F(\lambda)$ is a left-continuous nondecreasing bounded function with $F(-\pi)=0$. The constant c and the function $F(\lambda)$ are determined uniquely by the increment sequence $\xi^{(n)}(m,\mu)$.

From the other hand, a function $c^{(n)}(\mu)$ which has the form (4) *with a constant c and a function $D^{(n)}(m,\mu_1,\mu_2)$ which has the form* (5) *with a function $F(\lambda)$ which satisfies the indicated conditions are the mean value and the structural function of some stationary nth increment sequence $\xi^{(n)}(m,\mu)$.*

Using representation (5) of the structural function of a stationary nth increment sequence $\xi^{(n)}(m,\mu)$ and the Karhunen theorem [1], we obtain the following spectral representation of the stationary nth increment sequence $\xi^{(n)}(m,\mu)$:

$$\xi^{(n)}(m,\mu) = \int_{-\pi}^{\pi} e^{im\lambda}\left(1-e^{-i\mu\lambda}\right)^n\frac{1}{(i\lambda)^n}\,dZ(\lambda), \tag{6}$$

where $Z(\lambda)$ is an orthogonal stochastic measure on $[-\pi,\pi)$ connected with the spectral function $F(\lambda)$ by the relation

$$\mathsf{E}Z(A_1)\overline{Z(A_2)} = F(A_1\cap A_2) < \infty. \tag{7}$$

Example 2.1. Consider an ARIMA(0,1,1) sequence defined by the equation

$$\xi_m = \xi_{m-1} + \varepsilon_m + a\varepsilon_{m-1},$$

where ε_m is a sequence of uncorrelated identically distributed random variables with mean value 0 and variance σ^2. If we take $\eta_m = \xi_m - \xi_{m-1}$ we obtain a moving average sequence $\eta_m = \varepsilon_m + a\varepsilon_{m-1}$. Thus, ξ_m is a stochastic sequence with stationary increments of the 1st order. The spectral function $F(\lambda)$ of the sequence ξ_m can be calculated as follows

$$F(\lambda) = \frac{\sigma^2}{4\pi}\int_{-\pi}^{\lambda}\frac{u^2}{1-\cos u}\left(1+2a\cos u + a^2\right)du.$$

Here are some values of the structural function;

$$D^{(1)}(0,1,1) = \sigma^2\left(1+a^2\right), \qquad D^{(1)}(0,1,2) = \sigma^2\left(1+a+a^2\right),$$

$$D^{(1)}(0,2,2) = 2\sigma^2(1+a+a^2),$$

$$D^{(1)}(m,1,1) = \begin{cases}\sigma^2(1+a^2), & m=0,\\ \sigma^2 a, & m=-1,1,\\ 0, & \text{otherwise},\end{cases}$$

$$D^{(1)}(m,1,2) = \begin{cases} \sigma^2(1+a+a^2), & m=-1,0, \\ \sigma^2 a^2, & m=-2,1, \\ 0, & \text{otherwise,} \end{cases}$$

$$D^{(1)}(m,2,2) = \begin{cases} 2\sigma^2(1+a+a^2), & m=0, \\ \sigma^2(1+2a+a^2), & m=-1,1, \\ \sigma^2 a^2, & m=-2,2, \\ 0, & \text{otherwise.} \end{cases}$$

3. Filtering problem for the functional $A\xi$

Let a stochastic sequence $\{\xi(m), m \in \mathbb{Z}\}$ define a stationary nth increment $\xi^{(n)}(m,\mu)$ with an absolutely continuous spectral function $F(\lambda)$ which has spectral density $f(\lambda)$. Let $\{\eta(m), m \in \mathbb{Z}\}$ be a stochastic sequence, uncorrelated with the sequence $\xi(m)$, which determines a stationary nth increment $\eta^{(n)}(m,\mu)$ with an absolutely continuous spectral function $G(\lambda)$ whith has spectral density $g(\lambda)$. Without loss of generality we will assume that the mean values of the increment sequences $\xi^{(n)}(m,\mu)$ and $\eta^{(n)}(m,\mu)$ equal to 0. Let us suppose that we know values of the sequence $\xi(m)+\eta(m)$ at points $m=0,-1,-2,\dots$. Consider the problem of mean-square optimal linear estimation of the functional

$$A\xi = \sum_{k=0}^{\infty} a(k)\xi(-k)$$

of unknown values of the sequence $\xi(m)$ from observation of the sequence $\xi(m)+\eta(m)$ at points $m=0,-1,-2,\dots$. We will consider the case where the step $\mu>0$.

From (1) we can obtain the formal equation

$$\xi(-k) = \frac{1}{(1-B_\mu)^n}\xi^{(n)}(-k,\mu) = \sum_{i=k}^{\infty} d_\mu(i-k)\xi^{(n)}(-i,\mu), \tag{8}$$

where $\{d_\mu(i)\colon i \geq 0\}$ are coefficients from decomposition $\sum_{i=0}^{\infty} d_\mu(i)x^i = \left(\sum_{l=0}^{\infty} x^{\mu l}\right)^n$. From equation (8) one can find the following relations:

$$\sum_{k=0}^{\infty} a(k)\xi(-k) = \sum_{i=0}^{\infty} \xi^{(n)}(-i,\mu) \sum_{k=0}^{i} a(k) d_\mu(i-k),$$

$$\sum_{k=0}^{\infty} b_\mu(k)\xi^{(n)}(-k,\mu) = \sum_{i=0}^{\infty} \xi(-i) \sum_{l=0}^{\min\{n,[\frac{i}{\mu}]\}} (-1)^l C_n^l b_\mu(i-l\mu).$$

From the last two relations we obtain the following representation of the functional $A\xi$:

$$A\xi = \sum_{k=0}^{\infty} a(k)\xi(-k) = \sum_{k=0}^{\infty} b_\mu(k)\xi^{(n)}(-k,\mu) = B\xi,$$

$$b_\mu(k) = \sum_{m=0}^{k} a(m) d_\mu(k-m) = (\mathbf{D}^\mu \mathbf{a})_k, \qquad k \geq 0, \tag{9}$$

where $\mathbf{D}^\mu$ is a linear operator with elements $\mathbf{D}^\mu_{k,j} = d_\mu(k-j)$ if $0 \leq j \leq k$ and $\mathbf{D}^\mu_{k,j} = 0$ if $j>k$; $\mathbf{a} = (a(0),a(1),a(2),\dots)$. Let $\widehat{A}\xi$ denote the mean-square optimal linear estimate of the functional $A\xi$ from observations of stochastic sequence $\xi(m)+\eta(m)$ at points $m = 0,-1,-2,\dots$ and let $\widehat{B}\xi$ denote the mean-square optimal linear estimate of the functional $B\xi$ from observations of the stochastic nth increment sequence $\xi^{(n)}(m,\mu)+\eta^{(n)}(m,\mu)$ at points $m=0,-1,-2,\dots$.

Let $\Delta(f,g,\widehat{A}\xi) = \mathsf{E}|A\xi - \widehat{A}\xi|^2$ be the mean-square error of the estimate $\widehat{A}\xi$ of the functional $A\xi$ and let $\Delta(f,g,\widehat{B}\xi) = \mathsf{E}|B\xi - \widehat{B}\xi|^2$ be the mean-square error of the estimate $\widehat{B}\xi$ of the functional $B\xi$. Since $A\xi = B\xi$, the following equality holds true:

$$\widehat{A}\xi = \widehat{B}\xi. \tag{10}$$

Therefore, the following relations hold true

$$\Delta(f,g,\widehat{A}\xi) = \mathsf{E}|A\xi - \widehat{A}\xi|^2 = \mathsf{E}|B\xi - \widehat{B}\xi|^2 = \Delta(f,g,\widehat{B}\xi).$$

To find the mean-square optimal estimate of the functional $B\xi$ we use the Hilbert space orthogonal projection method proposed by A. M. Kolmogorov [11]. Suppose that conditions

$$\sum_{k=0}^{\infty}|b_\mu(k)|<\infty, \qquad \sum_{k=0}^{\infty}(k+1)|b_\mu(k)|^2<\infty, \tag{11}$$

$$\sum_{k=0}^{\infty}|(\mathbf{D}^\mu\mathbf{a})_k|<\infty, \qquad \sum_{k=0}^{\infty}(k+1)|(\mathbf{D}^\mu\mathbf{a})_k|^2<\infty \tag{12}$$

are satisfied.

Let $H^0\big(\xi_\mu^{(n)}+\eta_\mu^{(n)}\big)$ be the closed linear subspace of the Hilbert space $H=L_2(\Omega,\mathfrak{F},\mathsf{P})$ of the second order random variables generated by values $\big\{\xi^{(n)}(k,\mu)+\eta^{(n)}(k,\mu)\colon k\leq 0\big\}$, $\mu>0$. Consider also a closed linear subspace $L_2^0(f+g)$ of the Hilbert space $L_2(f+g)$ generated by functions

$$\left\{e^{i\lambda k}\left(1-e^{-i\lambda\mu}\right)^n\frac{1}{(i\lambda)^n}\colon k\leq 0\right\}.$$

From the formula

$$\xi^{(n)}(k,\mu)+\eta^{(n)}(k,\mu)=\int_{-\pi}^{\pi}e^{i\lambda k}\left(1-e^{-i\lambda\mu}\right)^n\frac{1}{(i\lambda)^n}\,dZ_{\xi^{(n)}+\eta^{(n)}}(\lambda)$$

one can verify the existence of one to one correspondence between element

$$e^{i\lambda k}(1-e^{-i\lambda\mu})^n/(i\lambda)^n$$

from the space $L_2^0(f+g)$ and element $\xi^{(n)}(k,\mu)+\eta^{(n)}(k,\mu)$ from the space $H^0\big(\xi_\mu^{(n)}+\eta_\mu^{(n)}\big)$. Every linear estimate $\widehat{B}\xi$ of the functional $B\xi$ admits representation

$$\widehat{B}\xi=\int_{-\pi}^{\pi}h_\mu(\lambda)\,dZ_{\xi^{(n)}+\eta^{(n)}}(\lambda), \tag{13}$$

where $h_\mu(\lambda)$ is the spectral characteristic of the estimate $\widehat{B}\xi$. The optimal estimate $\widehat{B}\xi$ is a projection of the element $B\xi$ on the subspace $H^0(\xi_\mu^{(n)}+\eta_\mu^{(n)})$. This estimate $\widehat{B}\xi$ is determined by the following conditions:

1) $\widehat{B}\xi\in H^0\big(\xi_\mu^{(n)}+\eta_\mu^{(n)}\big)$;
2) $(B\xi-\widehat{B}\xi)\perp H^0\big(\xi_\mu^{(n)}+\eta_\mu^{(n)}\big)$.

It follows from condition 2) that for all $k\leq 0$ the function $h_\mu(\lambda)$ satisfies the relation

$$\begin{aligned}
&\mathsf{E}(B\xi-\widehat{B}\xi)\overline{(\xi^{(n)}(k,\mu)+\eta^{(n)}(k,\mu))}\\
&\quad=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(B_\mu\left(e^{i\lambda}\right)\left(1-e^{-i\lambda\mu}\right)^n\frac{1}{(i\lambda)^n}-h_\mu(\lambda)\right)e^{-i\lambda k}\left(1-e^{i\lambda\mu}\right)^n\frac{1}{(-i\lambda)^n}f(\lambda)\,d\lambda\\
&\quad-\frac{1}{2\pi}\int_{-\pi}^{\pi}h_\mu(\lambda)e^{-i\lambda k}\left(1-e^{i\lambda\mu}\right)^n\frac{1}{(-i\lambda)^n}g(\lambda)\,d\lambda\\
&\quad=0.
\end{aligned}$$

From the previous relation we derive the following relations

$$\int_{-\pi}^{\pi}\left(B_\mu\left(e^{i\lambda}\right)\left(1-e^{-i\lambda\mu}\right)^n\frac{f(\lambda)}{(i\lambda)^n}-h_\mu(\lambda)(f(\lambda)+g(\lambda))\right)\frac{\left(1-e^{i\lambda\mu}\right)^n}{(-i\lambda)^n}e^{-i\lambda k}\,d\lambda=0,$$
$$k\leq 0,$$

which yields

$$h_\mu(\lambda)=B_\mu\left(e^{i\lambda}\right)\left(1-e^{-i\lambda\mu}\right)^n\frac{1}{(i\lambda)^n}\frac{f(\lambda)}{f(\lambda)+g(\lambda)}-\frac{(-i\lambda)^nC_\mu\left(e^{i\lambda}\right)}{\left(1-e^{i\lambda\mu}\right)^n(f(\lambda)+g(\lambda))},$$
$$B_\mu(e^{i\lambda})=\sum_{k=0}^{\infty}b_\mu(k)e^{-i\lambda k},\qquad C_\mu(e^{i\lambda})=\sum_{k=1}^{\infty}c_\mu(k)e^{i\lambda k}.$$

It follows from condition 1) we conclude that the spectral characteristic $h_\mu(\lambda)$ admits the representation

$$h_\mu(\lambda)=h(\lambda)\left(1-e^{-i\lambda\mu}\right)^n\frac{1}{(i\lambda)^n},\qquad h(\lambda)=\sum_{k=-\infty}^{0}s(k)e^{i\lambda k},$$

where

$$\int_{-\pi}^{\pi}|h(\lambda)|^2\left|1-e^{i\lambda\mu}\right|^{2n}\frac{f(\lambda)+g(\lambda)}{\lambda^{2n}}\,d\lambda<\infty,$$
$$\frac{(i\lambda)^nh_\mu(\lambda)}{\left(1-e^{-i\lambda\mu}\right)^n}\in L_2^0,$$
$$\int_{-\pi}^{\pi}\left(B_\mu(e^{i\lambda})\frac{f(\lambda)}{f(\lambda)+g(\lambda)}-\frac{\lambda^{2n}C_\mu\left(e^{i\lambda}\right)}{\left(1-e^{-i\lambda\mu}\right)^n\left(1-e^{i\lambda\mu}\right)^n(f(\lambda)+g(\lambda))}\right)e^{-i\lambda l}\,d\lambda=0,$$
$$l\geq 1. \tag{14}$$

Let the following conditions holds true:

$$\int_{-\pi}^{\pi}\frac{f(\lambda)}{f(\lambda)+g(\lambda)}\,d\lambda<\infty,\qquad \int_{-\pi}^{\pi}\frac{\lambda^{2n}}{\left|1-e^{i\lambda\mu}\right|^{2n}(f(\lambda)+g(\lambda))}\,d\lambda<\infty. \tag{15}$$

Set

$$R_{k,j}=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-i\lambda(j+k)}\frac{f(\lambda)}{f(\lambda)+g(\lambda)}\,d\lambda,$$
$$P^\mu_{k,j}=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{i\lambda(j-k)}\frac{\lambda^{2n}}{\left|1-e^{i\lambda\mu}\right|^{2n}(f(\lambda)+g(\lambda))}\,d\lambda,$$
$$Q^\mu_{k,j}=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{i\lambda(j-k)}\frac{\left|1-e^{i\lambda\mu}\right|^{2n}f(\lambda)g(\lambda)}{\lambda^{2n}(f(\lambda)+g(\lambda))}\,d\lambda.$$

Then (14) is equivalent to the following linear system:

$$\sum_{m=0}^{\infty}R_{l,m}b_\mu(m)=\sum_{k=1}^{\infty}P^\mu_{l,k}c_\mu(k),\qquad l\geq 1.$$

These system can be rewritten as

$$\mathbf{R}\mathbf{b}_\mu=\mathbf{P}_\mu\mathbf{c}_\mu, \tag{16}$$

where $\mathbf{c}_\mu=(c_\mu(1),c_\mu(2),c_\mu(3),\dots)$, $\mathbf{b}_\mu=(b_\mu(0),b_\mu(1),b_\mu(2),\dots)$, $\mathbf{P}_\mu$, $\mathbf{R}$ are linear operators in the space ℓ_2 defined by $(\mathbf{P}_\mu)_{l,k}=P^\mu_{l,k}$, $l,k\geq 1$, $(\mathbf{R})_{l,m}=R_{l,m}$, $l\geq 1$, $m\geq 0$. A solution $\mathbf{c}_\mu$ of the last equation defines the linear estimate $\widehat{B}\xi$ which is a projection of the element $B\xi$ from the Hilbert space H on the subspace $H^0\big(\xi_\mu^{(n)}+\eta_\mu^{(n)}\big)$. Since the space $H^0\big(\xi_\mu^{(n)}+\eta_\mu^{(n)}\big)$ is closed and convex, the projection $B\xi$ is uniquely determined for arbitrary sequence $b_\mu(0), b_\mu(1), b_\mu(2), \dots$ satisfying conditions (11). Thus equation (16) has a unique solution for an arbitrary $\mathbf{b}_\mu \neq 0$ and the linear operator $\mathbf{P}_\mu\colon \ell_2 \to X$, $X = \{\mathbf{x}_\mu \in \ell_2\colon \mathbf{x}_\mu = \mathbf{R}\mathbf{b}_\mu, \text{ where } \mathbf{b}_\mu \text{ satisfies (11)}\}$, has the inverse $(\mathbf{P}_\mu)^{-1}$.

Consequently, the unknown coefficients can be calculated by the formula

$$c_\mu(k) = \left(\mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{b}_\mu\right)_k,$$

where $(\mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{b}_\mu)_k$ is the kth element of the vector $\mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{b}_\mu$. Thus, spectral characteristics $h_\mu(\lambda)$ of the optimal estimate $\widehat{B}\xi$ of the functional $B\xi$ is calculated by the formula

$$h_\mu(\lambda) = B_\mu\left(e^{i\lambda}\right)\left(1 - e^{-i\lambda\mu}\right)^n \frac{1}{(i\lambda)^n}\frac{f(\lambda)}{f(\lambda)+g(\lambda)} - \frac{(-i\lambda)^n \sum_{k=1}^{\infty}\left(\mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{b}_\mu\right)_k e^{i\lambda k}}{\left(1-e^{i\lambda\mu}\right)^n (f(\lambda)+g(\lambda))}. \quad (17)$$

The mean-square error of the estimate is calculated by the formula

$$\begin{aligned}
\Delta(f,g;\widehat{B}\xi) &= \mathsf{E}|B\xi - \widehat{B}\xi|^2 \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|B_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n} g(\lambda) + \lambda^{2n}\sum_{k=1}^{\infty}\left(\mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{b}_\mu\right)_k e^{i\lambda k}\right|^2}{\lambda^{2n}\left|1-e^{i\lambda\mu}\right|^{2n}(f(\lambda)+g(\lambda))^2} f(\lambda)\,d\lambda \\
&\quad + \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|B_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n} f(\lambda) - \lambda^{2n}\sum_{k=1}^{\infty}\left(\mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{b}_\mu\right)_k e^{i\lambda k}\right|^2}{\lambda^{2n}\left|1-e^{i\lambda\mu}\right|^{2n}(f(\lambda)+g(\lambda))^2} g(\lambda)\,d\lambda \\
&= \left\langle \mathbf{R}\mathbf{b}_\mu, \mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{b}_\mu\right\rangle + \langle \mathbf{Q}_\mu\mathbf{b}_\mu, \mathbf{b}_\mu\rangle,
\end{aligned} \quad (18)$$

where $\mathbf{Q}_\mu$ is a linear operator in the space ℓ_2 defined by elements $(\mathbf{Q}_\mu)_{l,k} = Q^\mu_{l,k}$, $l,k \geq 0$.

Let us summarize our reasoning and present the results in the form of theorem.

Theorem 3.1. *Let stochastic sequences $\{\xi(m), m \in \mathbb{Z}\}$ and $\{\eta(m), m \in \mathbb{Z}\}$ determine stationary nth increment sequences $\xi^{(n)}(m,\mu)$ and $\eta^{(n)}(m,\mu)$ with absolutely continuous spectral functions $F(\lambda)$ and $G(\lambda)$ which have spectral densities $f(\lambda)$ and $g(\lambda)$ satisfying conditions* (15). *Let coefficients $\{b_\mu(k)\colon k \geq 0\}$ satisfy conditions* (11). *The optimal linear estimate $\widehat{B}\xi$ of the functional $B\xi$ of known elements $\xi^{(n)}(m,\mu)$, $m \leq 0$, $\mu > 0$ from observations of the sequence $\xi^{(n)}(m,\mu) + \eta^{(n)}(m,\mu)$ at points $m = 0, -1, -2, \dots$ is calculated by formula* (13). *The spectral characteristic $h_\mu(\lambda)$ of the optimal estimate $\widehat{B}\xi$ is calculated by formula* (17). *The value of the mean-square error $\Delta(f,g;\widehat{B}\xi)$ is calculated by formula* (18).

As a corollary from theorem 3.1 we can obtain the optimal estimate of the unknown value of the increment $\xi^{(n)}(m,\mu)$, $m \leq 0$, from observations of the sequence $\xi(k)+\eta(k)$ at points $k = 0, -1, -2, \dots$. Let us take a vector b_μ with element 1 at the $(-m)$th position and elements 0 at the remaining positions in (17). Then the spectral characteristic $\varphi_m(\lambda,\mu)$ of the estimate

$$\widehat{\xi}^{(n)}(m,\mu) = \int_{-\pi}^{\pi} \varphi_m(\lambda,\mu)\, dZ_{\xi^{(n)}+\eta^{(n)}}(\lambda) \quad (19)$$

is calculated by the formula

$$\varphi_m(\lambda,\mu) = e^{i\lambda m}\left(1 - e^{-i\lambda\mu}\right)^n \frac{1}{(i\lambda)^n}\frac{f(\lambda)}{f(\lambda)+g(\lambda)} - \frac{(-i\lambda)^n \sum_{k=1}^{\infty}\left(\mathbf{P}_\mu^{-1}\mathbf{r}_m\right)_k e^{i\lambda k}}{\left(1-e^{i\lambda\mu}\right)^n (f(\lambda)+g(\lambda))}, \quad (20)$$

where $\mathbf{r}_m = (R_{1,-m}, R_{2,-m}, \dots)$. The mean-square error of the estimate is calculated by the formula

$$\Delta\left(f,g;\widehat{\xi}^{(n)}(m,\mu)\right)$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|e^{i\lambda m}\left|1-e^{i\lambda\mu}\right|^{2n} g(\lambda) + \lambda^{2n}\sum_{k=1}^{\infty}\left(\mathbf{P}_\mu^{-1}\mathbf{r}_m\right)_k e^{i\lambda k}\right|^2}{\left|1-e^{i\lambda\mu}\right|^{2n}(f(\lambda)+g(\lambda))^2} f(\lambda)\,d\lambda$$
$$+ \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|e^{i\lambda m}\left|1-e^{i\lambda\mu}\right|^{2n} f(\lambda) - \lambda^{2n}\sum_{k=1}^{\infty}\left(\mathbf{P}_\mu^{-1}\mathbf{r}_m\right)_k e^{i\lambda k}\right|^2}{\left|1-e^{i\lambda\mu}\right|^{2n}(f(\lambda)+g(\lambda))^2} g(\lambda)\,d\lambda. \tag{21}$$

Thus, we have the following statement.

Corollary 3.1. *The optimal linear estimate $\widehat{\xi}^{(n)}(m,\mu)$ of the unknown value of the stochastic increment sequence $\xi^{(n)}(m,\mu)$, $m \le 0$, $\mu > 0$, from observations of the sequence $\xi(k)+\eta(k)$ at points $k = 0,-1,-2,\dots$ can be calculated by formula* (19). *The spectral characteristic $\varphi_m(\lambda,\mu)$ of the optimal estimate $\widehat{\xi}^{(n)}(m,\mu)$ is calculated by formula* (20). *The value of mean-square error $\Delta(f,g;\widehat{\xi}^{(n)}(m,\mu))$ is calculated formula* (21).

Consider now the smoothing problem for the stationary nth increment sequence $\xi^{(n)}(m,\mu)$ which consists of finding the mean-square optimal linear estimate $\widehat{\xi}^{(n)}(0,\mu)$ of the unknown value of the increment $\xi^{(n)}(0,\mu)$, $\mu > 0$, from observations of the stochastic sequence $\xi(k)+\eta(k)$ at points $k = 0,-1,-2,\dots$.

Let $r(k) = R_{k,0}$, $k \in \mathbb{Z}$. Then $\{r(k)\colon k \in \mathbb{Z}\}$ are the Fourier coefficients of the function $\frac{f(\lambda)}{f(\lambda)+g(\lambda)}$ which have the property $r(k) = \overline{r}(-k)$, $k \in \mathbb{Z}$, where $\overline{r}(k)$ denotes a conjugate element to $r(k)$. Let $\{V^{\mu}_{k,j}\colon k,j \ge 1\}$ be the coefficients which determine a linear operator $\mathbf{V}_\mu = (\mathbf{P}_\mu)^{-1}$. Then we have relations

$$\sum_{l\ge 1} V^{\mu}_{l,j} P_{k,l} = \delta_{k,j}, \qquad k,j \ge 1, \tag{22}$$

where $\delta_{k,j}$ is the Kronecker symbol. Using formulas (20) and (22) we obtain the spectral characteristic of the optimal estimate $\widehat{\xi}^{(n)}(0,\mu)$ of the unknown value of the increment $\xi^{(n)}(0,\mu)$:

$$\varphi(\lambda,\mu) = \frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}\sum_{k=0}^{\infty}\overline{r}(k)e^{-i\lambda k}.$$

The optimal estimate of the increment $\xi^{(n)}(0,\mu)$ is calculated by the formula

$$\widehat{\xi}^{(n)}(0,\mu) = \sum_{k=0}^{\infty}\overline{r}(k)\xi^{(n)}(-k,\mu) = \sum_{j=0}^{\infty}(\xi(-j)+\eta(-j))\sum_{l=0}^{\min\{n,[\frac{j}{\mu}]\}}(-1)^l C_n^l\,\overline{r}(j-l\mu). \tag{23}$$

The mean-square error of the estimate $\widehat{\xi}^{(n)}(0,\mu)$ is calculated by the formula

$$\Delta\left(f,g;\widehat{\xi}^{(n)}(0,\mu)\right) = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\overline{V}^{\mu}_{k,j}\overline{r}(j)r(k) + \sum_{l\in\mathbb{Z}} r(l)g_\mu(-l), \tag{24}$$

where $\{g_\mu(k)\colon k \in \mathbb{Z}\}$ are the Fourier coefficients of the function $|1-e^{i\lambda\mu}|^{2n}g(\lambda)\lambda^{-2n}$.

Corollary 3.2. *The optimal estimate $\widehat{\xi}^{(n)}(0,\mu)$ of the unknown value $\xi^{(n)}(0,\mu)$ of the stationary nth increment sequence $\xi^{(n)}(m,\mu)$, $\mu > 0$, from observations of the sequence $\xi(k)+\eta(k)$ at points $k = 0,-1,-2,\dots$ is calculated by formula* (23). *The value of the mean-square error $\Delta(f,g;\widehat{\xi}^{(n)}(0,\mu))$ of the estimate $\widehat{\xi}^{(n)}(0,\mu)$ is calculated by formula* (24).

Theorem 3.1 and corollaries 3.1, 3.2 determine solutions of the filtering problems for the nth increment sequence $\widehat{\xi}^{(n)}(m,\mu)$ and the linear functional $B\xi$ which are based on the Fourier coefficients of functions

$$\frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+g(\lambda))}, \qquad \frac{f(\lambda)}{f(\lambda)+g(\lambda)}, \qquad \frac{|1-e^{i\lambda\mu}|^{2n}f(\lambda)g(\lambda)}{\lambda^{2n}(f(\lambda)+g(\lambda))}.$$

However, the problem of finding the inverse operator $(\mathbf{P}_\mu)^{-1}$ to the operator $\mathbf{P}_\mu$ determined by the Fourier coefficients of the function $\frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+g(\lambda))}$ is a complicated problem in most cases. Therefore, we propose a method of finding the operator $(\mathbf{P}_\mu)^{-1}$ under the condition that the functions

$$\frac{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+g(\lambda))}{\lambda^{2n}}, \qquad \frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+g(\lambda))} \tag{25}$$

admit the canonical factorizations

$$\frac{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+g(\lambda))}{\lambda^{2n}} = \left|\sum_{k=0}^{\infty}\varphi_\mu(k)e^{-i\lambda k}\right|^2, \tag{26}$$

$$\frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+g(\lambda))} = \left|\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right|^2. \tag{27}$$

Using the coefficients $\varphi_\mu(k)$, $\psi_\mu(k)$, $k\geq 0$, from factorizations (26), (27), we define linear operators Φ_μ and Ψ_μ in the space ℓ_2. Let $(\Phi_\mu)_{k,j}=\varphi_\mu(k-j)$ and $(\Psi_\mu)_{k,j}=\psi_\mu(k-j)$ if $1\leq j\leq k$, $(\Phi_\mu)_{k,j}=0$ and $(\Psi_\mu)_{k,j}=0$ if $j>k$ and $k,j\geq 1$. The defined operators admit the following relation: $\Psi_\mu\Phi_\mu=\Phi_\mu\Psi_\mu=I$, where I is the identity operator. Moreover, the operator $\mathbf{P}_\mu$ allows the factorization $\mathbf{P}_\mu=\overline{\Psi}'_\mu\Psi_\mu$. Thus, $(\mathbf{P}_\mu)^{-1}=\Phi_\mu\overline{\Phi}'_\mu$ and the coefficients of the operator $\mathbf{V}_\mu=(\mathbf{P}_\mu)^{-1}$ are calculated by the formula

$$V^\mu_{k,j}=\sum_{p=1}^{\min(k,j)}\varphi_\mu(k-p)\overline{\varphi}_\mu(j-p), \qquad k,j\geq 1.$$

These observations can be summarized in the form of the following theorem.

Theorem 3.2. *Let functions* (25) *admit the canonical factorizations* (26) *and* (27) *respectively. Then the inverse operator $\mathbf{P}_\mu^{-1}$ to the operator $\mathbf{P}_\mu$ is calculated by the formula $\mathbf{P}_\mu^{-1}=\Phi_\mu\overline{\Phi}'_\mu$, where the linear operator Φ_μ in ℓ_2 space is determined by the coefficients $(\Phi_\mu)_{k,j}=\varphi_\mu(k-j)$ if $1\leq j\leq k$ and $(\Phi_\mu)_{k,j}=0$ if $j<k$, $k,j\geq 1$.*

Using theorem 3.1 we can find the optimal estimate

$$\widehat{A}\xi=\int_{-\pi}^{\pi}h^{(a)}_\mu(\lambda)\,dZ_{\xi^{(n)}+\eta^{(n)}}(\lambda) \tag{28}$$

of the functional $A\xi$. The spectral characteristic of the estimate $\widehat{A}\xi$ is calculated by the formula

$$h^{(a)}_\mu(\lambda)=A_\mu\left(e^{i\lambda}\right)\left(1-e^{-i\lambda\mu}\right)^n\frac{1}{(i\lambda)^n}\frac{f(\lambda)}{f(\lambda)+g(\lambda)}-\frac{(-i\lambda)^n\sum_{k=1}^{\infty}\left(\mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}}{\left(1-e^{i\lambda\mu}\right)^n(f(\lambda)+g(\lambda))}, \tag{29}$$

where $A_\mu(e^{i\lambda})=\sum_{k=0}^{\infty}(\mathbf{D}^\mu\mathbf{a})_k e^{-i\lambda k}$. The mean-square error can be calculated by formula

$$\begin{aligned}
&\Delta\left(f,g;\widehat{A}\xi\right)\\
&\quad=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n}g(\lambda)+\lambda^{2n}\sum_{k=1}^{\infty}\left(\mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right|^2}{\left|1-e^{i\lambda\mu}\right|^{2n}(f(\lambda)+g(\lambda))^2}f(\lambda)\,d\lambda\\
&\quad\quad+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n}f(\lambda)-\lambda^{2n}\sum_{k=1}^{\infty}\left(\mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right|^2}{\left|1-e^{i\lambda\mu}\right|^{2n}(f(\lambda)+g(\lambda))^2}g(\lambda)\,d\lambda\\
&\quad=\left\langle\mathbf{R}\mathbf{D}^\mu\mathbf{a},\mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{D}^\mu\mathbf{a}\right\rangle+\langle\mathbf{Q}_\mu\mathbf{D}^\mu\mathbf{a},\mathbf{D}^\mu\mathbf{a}\rangle.
\end{aligned}\tag{30}$$

Theorem 3.3. *Let uncorrelated stochastic sequences $\{\xi(m),m\in\mathbb{Z}\}$ and $\{\eta(m),m\in\mathbb{Z}\}$ define stationary nth increment sequences $\xi^{(n)}(m,\mu)$ and $\eta^{(n)}(m,\mu)$ with absolutely continuous spectral functions $F(\lambda)$ and $G(\lambda)$ which have spectral densities $f(\lambda)$ and $g(\lambda)$ satisfying conditions* (15)*. Let conditions* (12) *be satisfied. The optimal linear estimate $\widehat{A}\xi$ of the functional $A\xi$ of unknown elements $\xi(m)$, $m\leq 0$, from observations of the sequence $\xi(m)+\eta(m)$ at points $m=0,-1,-2\dots$ is calculated by formula* (28)*. The spectral characteristic $h_\mu^{(a)}(\lambda)$ of the optimal estimate $\widehat{A}\xi$ is calculated by formula* (29)*. The value of the mean-square error $\Delta(f,g;\widehat{A}\xi)$ is calculated by formula* (30)*. If the function $|1-e^{i\lambda\mu}|^{2n}\lambda^{-2n}(f(\lambda)+g(\lambda))$ admits the canonical factorization* (26)*, the operator $\mathbf{P}_\mu^{-1}$ from formulas* (29) *and* (30) *can be represented as $\mathbf{P}_\mu^{-1}=\Phi_\mu\overline{\Phi}_\mu'$.*

Example 3.1. Consider an ARIMA(0,1,2) sequence $\{\xi(m),m\in\mathbb{Z}\}$. The first order increments of the sequence $\xi(m)$ are stationary and the increments with step $\mu=1$ form a one-sided moving average sequence of order 2. Let the sequence $\xi(m)$ have the spectral density

$$f(\lambda)=\frac{\lambda^2\left|1-\phi e^{-i\lambda}\right|^2\left|1-\psi e^{-i\lambda}\right|^2}{|1-e^{-i\lambda}|^2}.$$

Consider an other stochastic sequence $\{\eta(m),m\in\mathbb{Z}\}$ with stationary increments of order 1 uncorrelated with $\xi(m)$ such that increments of the sequence $\{\xi(m)+\eta(m),m\in\mathbb{Z}\}$ with step 1 form a moving average sequence of order 1 and the spectral density has the form

$$f(\lambda)+g(\lambda)=\frac{\lambda^2\left|1-\phi e^{-i\lambda}\right|^2}{|1-e^{-i\lambda}|^2}.$$

Consider a real number sequence $\{a(k)\colon k\geq 0\}$ which is defined as follows: $a(0)=1$, $a(k)=-2^{-k}$ for $k\geq 1$. This sequence satisfies conditions (12). The problem is to find the optimal mean-square linear estimate $\widehat{A}\xi$ of the functional $A\xi=\sum_{k=0}^{\infty}a(k)\xi(-k)$ of unknown values $\xi(k)$, $k\leq 0$, of the sequence $\xi(m)$ from observations $\xi(k)+\eta(k)$, $k=0,-1,-2,\dots$. To calculate the spectral characteristic of the optimal estimate $\widehat{A}\xi$ of the functional $A\xi$ we use formula (29). The operator $\mathbf{P}_\mu=\mathbf{P}$ is determined by coefficients $(\mathbf{P})_{l,k}=\frac{\psi^p}{1-\psi^2}$, $|k-l|=p$, $l,k\geq 1$. The inverse operator $\mathbf{V}=\mathbf{P}^{-1}$ is defined by coefficients $(\mathbf{V})_{1,1}=1$, $(\mathbf{V})_{l,l}=1+\phi^2$ if $l\geq 2$, $(\mathbf{V})_{l,k}=-\phi$ if $|l-k|=1$, $l,k\geq 1$, and $(\mathbf{V})_{l,k}=0$ otherwise. The operator $\mathbf{R}$ is defined by coefficients $(\mathbf{R})_{1,0}=1$ and $(\mathbf{R})_{l,k}=0$ if $l\geq 1$, $k\geq 0$, $(l,k)\neq(1,0)$. The operator $\mathbf{D}^\mu=\mathbf{D}$ is defined by coefficients $d_\mu(k)=1$, $k\geq 0$. The spectral characteristic $h_1(\lambda)$ of the estimate $\widehat{A}\xi$ is calculated by the formula $h_1(\lambda)=\sum_{k=0}^{\infty}s(k)e^{-i\lambda k}\frac{1-e^{-i\lambda}}{i\lambda}$, where $s(0)=1-\frac{1}{2}\psi+\psi^2+\phi\psi\frac{2-\phi^2}{1-\phi^2}$, $s(k)=2^{-k-1}(2-5\psi+2\psi^2)+\phi^{k+1}\psi$, $k\geq 1$.

Denote $A(j) = \min\{n, [j/\mu]\}$, $j \geq 0$. Then the estimate $\widehat{A}\xi$ of the functional $A\xi$ is calculated by the formula

$$\widehat{A}\xi = \sum_{k=0}^{\infty} s(k)\left(\xi^{(n)}(-k,\mu) + \eta^{(n)}(-k,\mu)\right) = \sum_{j=0}^{\infty}(\xi(-j) + \eta(-j))\sum_{l=0}^{A(j)}(-1)^l C_n^l s(j - l\mu).$$

4. Minimax-robust method of filtering

The value of the mean-square error $\Delta(h_\mu^{(a)}(f,g); f,g) := \Delta(f,g;\widehat{A}\xi)$ and the spectral characteristic $h_\mu^{(a)}(f,g)$ of the optimal linear estimate $\widehat{A}\xi$ of the functional $A\xi$ of unknown values $\xi(m)$ based on observations of the stochastic sequence $\xi(k)+\eta(k)$ are determined by formulas (29) and (30) under the condition that spectral densities $f(\lambda)$ and $g(\lambda)$ of stochastic sequences $\xi(m)$ and $\eta(m)$ are known. In the case where spectral densities are not exactly known, but a set $\mathcal{D} = \mathcal{D}_f \times \mathcal{D}_g$ of admissible spectral densities is given, the minimax (robust) approach to estimation of functionals of the unknown values of stochastic sequence with stationary increments is reasonable. In other words we are interesting in finding an estimate that minimizes the maximum of the mean-square error for all spectral densities from a given class $\mathcal{D}$ of admissible spectral densities simultaneously.

Definition 4.1. For a given class of spectral densities $\mathcal{D} = \mathcal{D}_f \times \mathcal{D}_g$ spectral densities $f_0(\lambda) \in \mathcal{D}_f$, $g_0(\lambda) \in \mathcal{D}_g$ are called least favorable in the class $\mathcal{D}$ for the optimal linear filtering of the functional $A\xi$ if

$$\Delta(f_0, g_0) = \Delta(h(f_0, g_0); f_0, g_0) = \max_{(f,g)\in\mathcal{D}_f\times\mathcal{D}_g} \Delta(h(f,g); f,g).$$

Definition 4.2. For a given class of spectral densities $\mathcal{D} = \mathcal{D}_f \times \mathcal{D}_g$ a spectral characteristic $h^0(e^{i\lambda})$ of the optimal linear estimate of the functional $A\xi$ is called minimax-robust if there are satisfied conditions

$$h^0\left(e^{i\lambda}\right) \in H_{\mathcal{D}} = \bigcap_{(f,g)\in\mathcal{D}_f\times\mathcal{D}_g} L_2^0(f+g),$$

$$\min_{h\in H_{\mathcal{D}}} \max_{(f,g)\in\mathcal{D}_f\times\mathcal{D}_g} \Delta(h; f,g) = \max_{(f,g)\in\mathcal{D}_f\times\mathcal{D}_g} \Delta\left(h^0; f,g\right).$$

Using the derived formulas and the introduced definitions we can conclude that the following statement holds true.

Lemma 4.1. *Spectral densities $f_\mu^0 \in \mathcal{D}_f(\lambda)$, $g_\mu^0 \in \mathcal{D}_g(\lambda)$ which satisfy conditions* (15) *are least favorable in the class $\mathcal{D} = \mathcal{D}_f \times \mathcal{D}_g$ for the optimal linear filtering of the functional $A\xi$ if operators $\mathbf{P}_\mu^0$, $\mathbf{R}^0$, $\mathbf{Q}_\mu^0$ constructed with the help of the Fourier coefficients of the functions*

$$\frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f_\mu^0(\lambda)+g_\mu^0(\lambda))}, \qquad \frac{f_\mu^0(\lambda)}{f_\mu^0(\lambda)+g_\mu^0(\lambda)}, \qquad \frac{\left|1-e^{i\lambda\mu}\right|^{2n} f_\mu^0(\lambda) g_\mu^0(\lambda)}{\lambda^{2n}(f_\mu^0(\lambda)+g_\mu^0(\lambda))}$$

determine a solution of the conditional extremum problem

$$\begin{aligned}&\max_{f\in\mathcal{D}}\left(\langle \mathbf{R}\mathbf{D}^\mu\mathbf{a}, \mathbf{P}_\mu^{-1}\mathbf{R}\mathbf{D}^\mu\mathbf{a}\rangle + \langle \mathbf{Q}_\mu\mathbf{D}^\mu\mathbf{a}, \mathbf{D}^\mu\mathbf{a}\rangle\right)\\ &\qquad = \left\langle \mathbf{R}^0\mathbf{D}^\mu\mathbf{a}, \left(\mathbf{P}_\mu^0\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right\rangle + \left\langle \mathbf{Q}_\mu^0\mathbf{D}^\mu\mathbf{a}, \mathbf{D}^\mu\mathbf{a}\right\rangle.\end{aligned} \tag{31}$$

The minimax spectral characteristic is determined as $h^0 = h_\mu(f_\mu^0, g_\mu^0)$ if $h_\mu(f_\mu^0, g_\mu^0) \in H_{\mathcal{D}}$.

The function h^0 and the pair (f_μ^0, g_μ^0) form a saddle point of the function $\Delta(h; f, g)$ on the set $H_{\mathcal{D}} \times \mathcal{D}$. The saddle point inequalities

$$\Delta\left(h; f_\mu^0, g_\mu^0\right) \geq \Delta\left(h^0; f_\mu^0, g_\mu^0\right) \geq \Delta\left(h^0; f, g\right) \quad \forall f \in \mathcal{D}_f, \forall g \in \mathcal{D}_g, \forall h \in H_{\mathcal{D}}$$

hold true if $h^0 = h_\mu(f_\mu^0, g_\mu^0)$ and $h_\mu(f_\mu^0, g_\mu^0) \in H_{\mathcal{D}}$, where (f_μ^0, g_μ^0) is a solution of the following conditional extremum problem

$$\widetilde{\Delta}(f,g) = -\Delta(h_\mu(f_\mu^0, g_\mu^0); f, g) \to \inf, \qquad (f,g) \in \mathcal{D},$$

$$\Delta\left(h_\mu\left(f_\mu^0, g_\mu^0\right); f, g\right)$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n} g_\mu^0(\lambda) + \lambda^{2n}\sum_{k=1}^{\infty}\left(\left(\mathbf{P}_\mu^0\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right|^2}{\lambda^{2n}\left|1-e^{i\lambda\mu}\right|^{2n}\left(f_\mu^0(\lambda)+g_\mu^0(\lambda)\right)^2} f(\lambda)\,d\lambda$$
$$+ \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n} f_\mu^0(\lambda) - \lambda^{2n}\sum_{k=1}^{\infty}\left(\left(\mathbf{P}_\mu^0\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right|^2}{\lambda^{2n}\left|1-e^{i\lambda\mu}\right|^{2n}\left(f_\mu^0(\lambda)+g_\mu^0(\lambda)\right)^2} g(\lambda)\,d\lambda.$$

This conditional extremum problem is equivalent to the unconditional extremum problem

$$\Delta_{\mathcal{D}}(f,g) = \widetilde{\Delta}(f,g) + \delta(f,g \mid \mathcal{D}_f \times D_g) \to \inf,$$

$\delta(f,g|\mathcal{D}_f \times D_g)$ is the indicator function of the set $\mathcal{D}_f \times \mathcal{D}_g$. Solution (f_μ^0, g_μ^0) to this unconditional extremum problem is characterized by the condition $0 \in \partial\Delta_{\mathcal{D}}(f_\mu^0, g_\mu^0)$ [18].

5. Least favorable spectral densities in the class $\mathcal{D}_f \times \mathcal{D}_g$

Consider the problem of optimal linear filtering of the functional $A\xi$ for the set of spectral densities $\mathcal{D} = \mathcal{D}_f \times \mathcal{D}_g$, where

$$\mathcal{D}_f^0 = \left\{ f(\lambda) \;\middle|\; \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\lambda)d\lambda \le P_1 \right\}, \qquad \mathcal{D}_g^0 = \left\{ g(\lambda) \;\middle|\; \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\lambda)d\lambda \le P_2 \right\}.$$

Let us assume that densities $f_\mu^0 \in \mathcal{D}_f$, $g_\mu^0 \in \mathcal{D}_g$ and functions

$$h_{\mu,f}\left(f_\mu^0, g_\mu^0\right) = \frac{\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n} g_\mu^0(\lambda) + \lambda^{2n}\sum_{k=1}^{\infty}\left(\left(\mathbf{P}_\mu^0\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right|}{|\lambda|^n\left|1-e^{i\lambda\mu}\right|^{n}\left(f_\mu^0(\lambda)+g_\mu^0(\lambda)\right)}, \tag{32}$$

$$h_{\mu,g}\left(f_\mu^0, g_\mu^0\right) = \frac{\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n} f_\mu^0(\lambda) - \lambda^{2n}\sum_{k=1}^{\infty}\left(\left(\mathbf{P}_\mu^0\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right|}{|\lambda|^n\left|1-e^{i\lambda\mu}\right|^{n}\left(f_\mu^0(\lambda)+g_\mu^0(\lambda)\right)} \tag{33}$$

are bounded. In this case the functional $\Delta(h_\mu(f_\mu^0, g_\mu^0); f, g)$ is continuous and bounded in $\mathcal{L}_1 \times \mathcal{L}_1$ space. It comes from the condition $0 \in \partial\Delta_{\mathcal{D}}(f_\mu^0, g_\mu^0)$ that the least favorable densities $f_\mu^0(\lambda) \in \mathcal{D}_f$, $g_\mu^0(\lambda) \in \mathcal{D}_g$ satisfy the equations

$$\begin{aligned}&\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n} g_\mu^0(\lambda) + \lambda^{2n}\sum_{k=1}^{\infty}\left(\left(\mathbf{P}_\mu^0\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right| \\ &\qquad = \alpha_1|\lambda|^n\left|1-e^{i\lambda\mu}\right|^n\left(f_\mu^0(\lambda)+g_\mu^0(\lambda)\right),\end{aligned} \tag{34}$$

$$\begin{aligned}&\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n} f_\mu^0(\lambda) - \lambda^{2n}\sum_{k=1}^{\infty}\left(\left(\mathbf{P}_\mu^0\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right| \\ &\qquad = \alpha_2|\lambda|^n\left|1-e^{i\lambda\mu}\right|^n\left(f_\mu^0(\lambda)+g_\mu^0(\lambda)\right),\end{aligned} \tag{35}$$

where $\alpha_1 \ge 0$ and $\alpha_2 \ge 0$ are constants such that $\alpha_1 \ne 0$ if $\frac{1}{2\pi}\int_{-\pi}^{\pi} f_\mu^0(\lambda)\,d\lambda = P_1$ and $\alpha_2 \ne 0$ if $\frac{1}{2\pi}\int_{-\pi}^{\pi} g_\mu^0(\lambda)\,d\lambda = P_2$. Thus, the following statements hold true.

Theorem 5.1. *Let spectral densities $f_\mu^0(\lambda) \in \mathcal{D}_f$ and $g_\mu^0(\lambda) \in \mathcal{D}_g$ satisfy conditions* (15) *and let functions $h_{\mu,f}(f_\mu^0, g_\mu^0)$, $h_{\mu,g}(f_\mu^0, g_\mu^0)$ determined by equations* (32), (33) *be bounded. The spectral densities $f_\mu^0(\lambda)$ and $g_\mu^0(\lambda)$ determined by relations* (34), (35) *are least favorable in the class $\mathcal{D} = \mathcal{D}_f \times \mathcal{D}_g$ for the optimal linear filtering problem for the functional*

$A\xi$ if they determine a solution of the extremum problem (31). *The function $h_\mu(f_\mu^0, g_\mu^0)$ determined by* (29) *is the minimax spectral characteristic of the optimal estimate of the functional $A\xi$.*

Theorem 5.2. *Let the spectral density $f(\lambda)$ be known, the spectral density $g_\mu^0(\lambda) \in \mathcal{D}_g$ and let conditions* (15) *be satisfied. Let the function $h_{\mu,g}(f, g_\mu^0)$ be bounded. The spectral density $g_\mu^0(\lambda)$ is least favorable in the class $\mathcal{D}_g$ for the optimal linear filtering of the functional $A\xi$ if it is of the form*

$$g_\mu^0(\lambda) = \max\left\{0, \frac{\left|A_\mu(e^{i\lambda})\left|1-e^{i\lambda\mu}\right|^{2n} f(\lambda) - \lambda^{2n}\sum_{k=1}^{\infty}\left(\left(\mathbf{P}_\mu^0\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right|}{\alpha_2|\lambda|^n|1-e^{i\lambda\mu}|^n} - f(\lambda)\right\}$$

and the pair (f, g_μ^0) determines a solution of the extremum problem (31). *The function $h_\mu(f, g_\mu^0)$ determined by* (29) *is the minimax spectral characteristic of the optimal estimate of the functional $A\xi$.*

6. Least favorable spectral densities in the class $\mathcal{D} = \mathcal{D}_u^v \times \mathcal{D}_\varepsilon$

Consider the problem of the optimal linear filtering of the functional $A\xi$ for the set of spectral densities $\mathcal{D} = \mathcal{D}_u^v \times \mathcal{D}_\varepsilon$, where

$$\mathcal{D}_u^v = \left\{ f(\lambda) \,\middle|\, v(\lambda) \leq f(\lambda) \leq u(\lambda), \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\lambda)d\lambda \leq P_1 \right\},$$

$$\mathcal{D}_\varepsilon = \left\{ g(\lambda) \,\middle|\, g(\lambda) = (1-\varepsilon)g_1(\lambda) + \varepsilon w(\lambda), \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\lambda)d\lambda \leq P_2 \right\}.$$

Here spectral densities $u(\lambda)$, $v(\lambda)$, $g_1(\lambda)$ are known and fixed, and spectral densities $u(\lambda)$, $v(\lambda)$ are bounded.

Let $f_\mu^0(\lambda) \in \mathcal{D}_u^v$, $g_\mu^0(\lambda) \in \mathcal{D}_\varepsilon$ be spectral densities such that functions $h_{\mu,f}(f_\mu^0, g_\mu^0)$, $h_{\mu,g}(f_\mu^0, g_\mu^0)$ determined by (32), (33) are bounded. From the condition $0 \in \partial\Delta_{\mathcal{D}}(f_\mu^0, g_\mu^0)$ we find the following equations that determine the least favorable densities

$$\begin{gathered}\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n} g_\mu^0(\lambda) + \lambda^{2n}\sum_{k=1}^{\infty}\left(\left(\mathbf{P}_\mu^0\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right| \\ = \alpha_1|\lambda|^n\left|1-e^{i\lambda\mu}\right|^n\left(f_\mu^0(\lambda)+g_\mu^0(\lambda)\right)\left(\gamma_1(\lambda)+\gamma_2(\lambda)+\alpha_1^{-1}\right),\end{gathered} \tag{36}$$

$$\begin{gathered}\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n} f_\mu^0(\lambda) - \lambda^{2n}\sum_{k=1}^{\infty}\left(\left(\mathbf{P}_\mu^0\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right| \\ = \alpha_2|\lambda|^n\left|1-e^{i\lambda\mu}\right|^n\left(f_\mu^0(\lambda)+g_\mu^0(\lambda)\right)\left(\varphi(\lambda)+\alpha_2^{-1}\right),\end{gathered} \tag{37}$$

where $\gamma_1 \leq 0$ and $\gamma_1 = 0$ if $f_\mu^0(\lambda) \geq v(\lambda)$; $\gamma_2(\lambda) \geq 0$ and $\gamma_2 = 0$ if $f_\mu^0(\lambda) \leq u(\lambda)$; $\varphi(\lambda) \leq 0$ and $\varphi(\lambda) = 0$ when $g_\mu^0(\lambda) \geq (1-\varepsilon)g_1(\lambda)$. The following statements hold true.

Theorem 6.1. *Let spectral densities $f_\mu^0(\lambda) \in \mathcal{D}_u^v$, $g_\mu^0(\lambda) \in \mathcal{D}_\varepsilon$ satisfy conditions* (15). *Let functions $h_{\mu,f}(f_\mu^0, g_\mu^0)$ and $h_{\mu,g}(f_\mu^0, g_\mu^0)$ determined by* (32), (33) *be bounded. Spectral densities $f_\mu^0(\lambda)$ and $g_\mu^0(\lambda)$ determined by equations* (36), (37) *are least favorable in the class $\mathcal{D} = \mathcal{D}_u^v \times \mathcal{D}_\varepsilon$ for the optimal linear filtering of the functional $A\xi$ if they determine a solution of extremum problem* (31). *The minimax spectral characteristic $h_\mu(f_\mu^0, g_\mu^0)$ of the optimal estimate of the functional $A\xi$ is determined by* (29).

Theorem 6.2. *Let the spectral density $f(\lambda)$ be known, the spectral density $g_\mu^0(\lambda) \in \mathcal{D}_\varepsilon$ and let conditions* (15) *be satisfied. Let the function $h_{\mu,g}(f, g_\mu^0)$ determined by* (29) *be* *bounded. The spectral density $g^0_\mu(\lambda)$ is least favorable in the class $\mathcal{D}_\varepsilon$ for the optimal linear filtering of the functional $A\xi$ if it is of the form*

$$g^0_\mu(\lambda)=\max\left\{(1-\varepsilon)g_1(\lambda),f_1(\lambda)\right\},$$

$$f_1(\lambda)=\frac{\alpha_2\left|A_\mu\left(e^{i\lambda}\right)\left|1-e^{i\lambda\mu}\right|^{2n}f(\lambda)-\lambda^{2n}\sum_{k=1}^{\infty}\left(\left(\mathbf{P}^0_\mu\right)^{-1}\mathbf{R}^0\mathbf{D}^\mu\mathbf{a}\right)_k e^{i\lambda k}\right|}{|\lambda|^n|1-e^{i\lambda\mu}|^n}-f(\lambda),$$

and the pair (f,g^0_μ) determines a solution of the extremum problem (31). The function $h_\mu(f,g^0_\mu)$ determined by (29) is minimax spectral characteristic of the optimal estimate of the functional $A\xi$.

7. Conclusions

In this article we found a solution of the filtering problem for linear functionals $A\xi=\sum_{k=0}^{\infty}a(k)\xi(-k)$ which depend on unobserved values of a stochastic sequence $\xi(m)$ with stationary nth increments at points $m=0,-1,-2,\dots$. Estimate is based on observations of a sequence $\xi(m)+\eta(m)$ at points $m=0,-1,-2,\dots$, where $\eta(m)$ is an uncorrelated with $\xi(m)$ sequence with stationary nth increments. We derived formulas for computing the value of the mean-square error and the spectral characteristic of the optimal linear estimate of the functional in the case where spectral densities of sequences are exactly known. In the case of spectral uncertainty, where spectral densities are not exactly known, but a set of admissible spectral densities is specified, the minimax-robust method is applied. Formulas that determine the least favorable spectral densities and minimax (robust) spectral characteristics are derived for some special sets of admissible spectral densities.

References

1. I. I. Gikhman and A. V. Skorokhod, *The Theory of Stochastic Processes. I*, Springer, Berlin, 2004.
2. I. I. Dubovets'ka, O. Yu. Masyutka, and M. P. Moklyachuk, *Interpolation of periodically correlated stochastic sequences*, Theory Probab. Math. Stat. **84** (2012), 43–56.
3. I. I. Dubovets'ka and M. P. Moklyachuk, *Filtration of linear functionals of periodically correlated sequences*, Theory Probab. Math. Stat. **86** (2012), 43–55.
4. I. I. Dubovets'ka and M. P. Moklyachuk, *Extrapolation of of periodically correlated processes processes from observations with noise*, Theory Probab. Math. Stat. **88** (2013), 60–75.
5. I. I. Dubovets'kaand M. P. Moklyachuk, *Filtering of periodically correlated processes*, Prykl. Stat. Aktuarna Finans. Mat. (2012), no. 2, 149–158,
6. I. I. Dubovets'ka and M. P. Moklyachuk, *Minimax estimation problem for periodically correlated stochastic processes*, Journal of Mathematics and System Science **3** (2013), no. 1, 26–30.
7. U. Grenander, *A prediction problem in game theory*, Ark. Mat. **3** (1957), 371–379.
8. J. Franke, *Minimax robust prediction of discrete time series*, Z. Wahrsch. Verw. Gebiete **68** (1985), 337–364.
9. J. Franke and H. V. Poor, *Minimax-robust filtering and finite-length robust predictors*, Robust and Nonlinear Time Series Analysis, Lecture Notes in Statistics, vol. 26, Springer-Verlag, 1984, pp. 87–126.
10. S. A. Kassam and H. V. Poor, *Robust techniques for signal processing: A survey*, Proc. IEEE. **73** (1985), 433–481.
11. A. N. Kolmogorov, *Selected Works by A. N. Kolmogorov*, Vol. II: Probability theory and mathematical statistics (A. N. Shiryayev, ed.), Mathematics and Its Applications. Soviet Series, vol. 26, Kluwer Academic Publishers, Dordrecht, 1992.
12. M. M. Luz and M. P. Moklyachuk, *Interpolation of functionals of stochactic sequenses with stationary increments*, Theory Probab. Math. Stat. **87** (2012), 94–108.
13. M. M. Luz and M. P. Moklyachuk, *Interpolation of functionals of stochastic sequences with stationary increments from observations with noise*, Prykl. Stat. Aktuarna Finans. Mat. (2012), no. 2, 131–148.
14. M. P. Moklyachuk, *Minimax filtration of linear transformations of stationary sequences*, Ukr. Math. J. **43** (1991), no. 1, 75–81.

15. M. P. Moklyachuk, *Robust procedures in time series analysis*, Theory Stoch. Process. **6** (2000), no. 3–4, 127–147.
16. M. P. Moklyachuk, *Game theory and convex optimization methods in robust estimation problems*, Theory Stoch. Process. **7** (2001), no. 1–2, 253–264.
17. M. P. Moklyachuk, *Robust Estimations of Functionals of Stochastic Processes*, "Kyivskyi Universytet", Kyiv, 2008.
18. M. P. Moklyachuk, *Nonsmooth Analysis and Optimization*, "Kyivskyi Universytet", Kyiv, 2008.
19. M. P. Moklyachuk and O. Yu. Masyutka, *Interpolation of multidimensional stationary sequences*, Theory Probab. Math. Stat. **73** (2006), 125–133.
20. M. P. Moklyachuk and O. Yu. Masyutka, *On the problem of filtration of vector stationary sequences*, Theory Probab. Math. Stat. **75** (2007), 109–119.
21. M. P. Moklyachuk and O. Yu. Masyutka, *Robust filtering of stochastic processes*, Theory Stoch. Process. **13** (2007), no. 1–2, 166–181.
22. M. P. Moklyachuk and O. Yu. Masyutka, *Minimax prediction problem for multidimensional stationary stochastic processes*, Commun. Stat. Theory Methods. **40** (2011), no. 19–20, 3700–3710.
23. M. Moklyachuk and O. Masyutka, *Minimax-robust estimation technique for stationary stochastic processes*, LAP LAMBERT Academic Publishing, 2012.
24. M. S. Pinsker and A. M. Yaglom, *On linear extrapolaion of random processes with nth stationary incremens*, Dokl. Akad. Nauk SSSR **94** (1954), no. 3, 385–388.
25. M. S. Pinsker, *The theory of curves with nth stationary incremens in Hilbert spaces*, Izv. Akad. Nauk SSSR. Ser. Mat. **19** (1955), no. 3, 319–344.
26. N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series. With Engineering Applications*, M.I.T. Press, Cambridge, Mass., 1966.
27. A. M. Yaglom, *Correlation theory of stationary and related random processes with stationary nth increments*, Mat. Sbornik **37(79)** (1955), no. 1, 141–196.
28. A. M. Yaglom, *Correlation theory of stationary and related random functions. Vol. 1: Basic results; Vol. 2: Suplementary notes and references*, Springer Series in Statistics, Springer-Verlag, New York, 1987.

Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, Kyiv 01601, Ukraine
E-mail address: maksim_luz@ukr.net

Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, Kyiv 01601, Ukraine
E-mail address: mmp@univ.kiev.ua

Received 27/11/2012

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 130–139

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 130–139

PROPERTIES OF INTEGRALS WITH RESPECT TO FRACTIONAL POISSON PROCESS WITH THE COMPACT KERNEL

UDC 519.21

Y. MISHURA AND V. ZUBCHENKO

Abstract. We study the properties of the fractional Poisson process with the Molchan–Golosov kernel. The kernel can be characterized as a compact since it is non-zero on compact interval. The integral of nonrandom function with respect to the centered and non-centered fractional Poisson processes with the Molchan–Golosov kernel is defined. The second moments of these integrals in terms of the norm of the integrand in $L_{1/H}([0,T])$ space are obtained. Moment estimates for the higher moments of these integrals are established via the Bichteler–Jacod inequality.

Анотація. Вивчено властивості дробово-пуассонівських процесів з ядром Молчана–Голосова, яке можна охарактеризувати як компактне, тому що воно ненульове лише на компактному інтервалі. Визначено інтеграли від невипадкових функцій за центрованим та нецентрованим дробово-пуассонівськими процесами з ядром Молчана–Голосова. Оцінено другі моменти цих інтегралів в термінах норми підінтегральної функції в просторі $L_{1/H}([0,T])$ та одержано моментні оцінки за допомогою нерівності Біхтелера–Жакода.

Аннотация. Изучены свойства дробно-пуассоновских процессов с ядром Молчана–Голосова, которое можно охарактеризовать как компактное, поскольку оно ненулевое на компактном интервале. Определены интегралы от неслучайной функции по центрированному и нецентрированному дробно-пуассоновским процессам с ядром Молчана–Голосова. Оценены вторые моменты этих интегралов в терминах нормы подинтегральной функции в пространстве $L_{1/H}([0,T])$ и получены оценки для моментов высшего порядка с помощью неравенства Бихтелера–Жакода.

1. Introduction

Models based on a fractional Brownian motion are an important tool for the study of many theoretical and applied problems. Due to the structure of its covariance function, the fractional Brownian motion that is the process parametrized by its Hurst index, allows to model the dependence on the past history of the process. It is known that for Hurst parameter $H > 1/2$ the fractional Brownian motion has so-called long-range dependence property, for $H \in (0, 1/2)$ it is a process with short memory, and for $H = 1/2$ we have the standard Brownian motion.

At the same time, many natural, technical and economic phenomena are characterized by the instantaneous change in the dynamics of the studied characteristics that cannot be described with the help of the fractional Brownian motion. In particular, such dynamics is typically seen in "jumps" of interest rates, exchange rates, financial indices. Models with jumps can be described with the help of stochastic differential equations that include Poisson measure (see, e.g., [17] and references therein). However, current dynamics of these processes depends essentially on their past history. So construction of models which are able to reflect effectively such features of the process is relevant. Particularly, it is significant for estimation and forecasting of future dynamics of complex financial instruments based on interest rates and financial indices. That's why we are interested in

2000 *Mathematics Subject Classification.* Primary 60G22; Secondary 60G51.

Key words and phrases. Fractional Poisson process; integral representation of the fractional Poisson process; Mandelbrot – van Ness kernel; Molchan–Golosov kernel; integral with respect to the fractional Poisson process; Bichteler–Jacod inequality.

study of the processes that can combine dependence on the past with an instant change or "jumping" change of the dynamics.

Combining randomness with independence on the past history of the process, the property of long memory and "jumping" change of the dynamics of characteristics under investigation can be expressed mathematically correspondingly by the standard Brownian motion, the fractional Brownian motion and by the Lévy processes.

Mathematical model combining dependence on the past and possibility of instantaneous change of characteristics can be expressed, in particular, by the fractional Poisson process.

There are different approaches to the definition of the fractional Poisson process. In the paper [1] several methods of the fractional Poisson process construction are proposed. One of them is the following: it is assumed that for the fractional Poisson process $N_\nu(t)$, $t > 0$, its distribution $p_k = \mathsf{P}\{N_\nu(t) = k\}$, $k \geq 0$, solves the following equation

$$\frac{d^\nu p_k}{dt^\nu} = -\lambda p_k + \lambda p_{k-1}, k \geq 0,$$

where $p_{-1}(t) = 0$ and $p_k(0) = \mathbb{1}_{\{k=0\}}$ and for $m \in \mathbb{N}$

$$\frac{d^\nu u(t)}{dt^\nu} = \begin{cases} \frac{1}{\Gamma(m-\nu)} \int_0^t \frac{1}{(t-s)^{1+\nu-m}} \frac{d^m}{ds^m} u(s)\, ds, & \text{for } m-1 < \nu < m, \\ \frac{d^m}{dt^m} u(t), \quad s \in [0,T], & \text{for } \nu = m, \end{cases}$$

is the fractional derivative in the sense of Dzherbashyan–Caputo.

Another method is to replace the factorial functions in the distribution of the Poisson process by the Gamma functions. In works [4, 5, 14] the so-called "renewal" approach is used. In contrast to classical characterization of the usual Poisson process as a renewal process, which is constructed as the sum of non-negative independent random variables with exponential distribution, it is assumed that these random variables have Mittag-Leffler distribution. One more approach to the fractional Poisson process construction is the use of so-called "inverse subordinator" method [8].

In order to introduce our approach, we perform certain analogy with a fractional Brownian motion, see, e.g. [10]. Besides the definition of the latter as a Gaussian process with some covariance structure, the fractional Brownian motion can be represented as the integral of a nonrandom kernel with respect to the standard Brownian motion. Examples of kernels used for such representation are the Mandelbrot – van Ness with infinite support and the compactly supported Molchan–Golosov kernel.

Using such representation, it is natural to define the fractional Poisson process as the integral of one of such kernels with respect to the Poisson process (Lévy process). The fractional Lévy processes was first defined using Mandelbrot – van Ness kernel in the work [2], the theory was developed in the paper [7]. The general definition of the fractional Lévy process by using the Molchan–Golosov kernel is given in the work [16].

In this paper we conduct further research of the fractional Poisson processes with the Molchan–Golosov kernel. The integral of a nonrandom function with respect to the centered and non-centered fractional Poisson processes with the Molchan–Golosov kernel is defined. We estimate second moments of such integrals in terms of the norm of the integrand in $L_{1/H}([0,T])$ space. Moment estimates for the higher moments of these integrals via the Bichteler–Jacod inequality are established.

2. Main definitions

The fractional Brownian motion $B^H = \{B_t^H, t \in \mathbb{R}\}$ with Hurst index $H \in (0,1)$ is a Gaussian process with zero mean and the covariance

$$\mathsf{E}\, B_t^H B_s^H = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right).$$

In what follows we consider $H \in (1/2, 1)$. In order to represent a fractional Brownian motion via a Brownian motion we can use both the Mandelbrot – van Ness and the Molchan–Golosov kernel.

The Mandelbrot – van Ness kernel $f_H(t,s)$ is given by

$$f_H(t,s) = c_H\left((t-s)_+^{H-1/2} - (-s)_+^{H-1/2}\right), \qquad s,t \in \mathbb{R},$$

where

$$c_H = \left(\int_0^\infty \left((1+s)^{H-1/2} - s^{H-1/2}\right)^2 ds + \frac{1}{2H}\right)^{-1/2} = \frac{(2H\sin\pi H\Gamma(2H))^{1/2}}{\Gamma(H+1/2)}.$$

The Molchan–Golosov kernel $z_H(t,s)$ is given by

$$z_H(t,s) = \frac{C_H}{\Gamma(H-1/2)} s^{1/2-H} \int_s^t u^{H-1/2}(u-s)^{H-3/2}\, du, \qquad 0 < s < t.$$

In the work [16] it is proved that actually $c_H = C_H$.

The dynamics of $z_H(t,\cdot)$ is equivalent to the dynamics of $\cdot^{1/2-H}$ in the neighborhood of zero and to the dynamics $(t-\cdot)^{H-1/2}$ in the neighborhood of t, see, e.g. [3]. In particular, $z_H(t,\cdot)$ is locally square integrable on $(0,t)$ for every $t \in (0,\infty)$. Also, for $H > 1/2$ the kernel $z_H(t,\cdot)$ is continuous when $s \neq 0$ and has a continuous derivative on $(0,t)$.

The fractional Brownian motion can be represented by integration of the nonrandom kernel with respect to a Brownian motion, in particular:

– by integration over an infinite interval of the Mandelbrot – van Ness kernel:

$$\left(B_t^H\right)_{t\in\mathbb{R}} = \left(\int_{-\infty}^t f_H(t,s)\, dW_s\right)_{t\in\mathbb{R}}$$

– or by integration over a compact interval of the Molchan–Golosov kernel:

$$\left(B_t^H\right)_{t\geq 0} = \left(\int_0^t z_H(t,s)\, dW_s\right)_{t\geq 0}. \tag{1}$$

The right-sided Riemann–Liouville fractional integral operator $I_{T-}^\alpha f$ of order α on $[0,T]$ is defined as

$$(I_{T-}^\alpha f)(s) := \begin{cases} \frac{1}{\Gamma(\alpha)}\int_s^T f(u)(u-s)^{\alpha-1}\, du, & s\in[0,T], \text{ for } \alpha > 0, \\ f(s), & s \in [0,T], \text{ for } \alpha = 0, \end{cases}$$

$$I_{T-}^{-\alpha} f := D_{T-}^\alpha f, \qquad \alpha \in (0,1),$$

where $D_{T-}^\alpha f$ is *the right-sided Riemann–Liouville fractional derivative operator* of order α on $[0,T]$, which is defined as

$$(D_{T-}^\alpha f)(s) := \begin{cases} -\frac{d}{ds}(I_{T-}^{1-\alpha} f)(s), & s \in (0,T), \text{ for } \alpha \in (0,1), \\ -\frac{d}{ds} f(s), & s \in (0,T), \text{ for } \alpha = 1, \\ f(s), & s \in (0,T), \text{ for } \alpha = 0. \end{cases}$$

The right-sided Riemann–Liouville fractional integral operator of order α on $\mathbb{R}$ is defined as

$$(I_-^\alpha f)(s) := \begin{cases} \frac{1}{\Gamma(\alpha)}\int_s^\infty f(u)(u-s)^{\alpha-1} du, & s \in \mathbb{R}, \text{ for } \alpha > 0, \\ f(s), \quad s \in \mathbb{R}, \text{ for } \alpha = 0. \end{cases}$$

$$I_-^{-\alpha} f := D_-^\alpha f, \quad \alpha \in (0,1),$$

where $D^{\alpha}_{-}f$ is *the right-sided Riemann–Liouville fractional derivative operator* of order α on $\mathbb{R}$:

$$(D^{\alpha}_{-}f)(s) := \begin{cases} -\frac{d}{ds}\left(I^{1-\alpha}_{-}f\right)(s) = -\frac{1}{\Gamma(1-\alpha)}\frac{d}{ds}\int_s^{\infty} f(u)(u-s)^{-\alpha}\,du, & \\ & s\in\mathbb{R},\ \text{for } \alpha\in(0,1), \\ -\frac{d}{ds}f(s), & s\in\mathbb{R},\ \text{for } \alpha=1, \\ f(s), & s\in\mathbb{R},\ \text{for } \alpha=0. \end{cases}$$

The centered fractional Poisson process with the Mandelbrot – van Ness kernel. Investigation of a fractional Poisson process with the Mandelbrot – van Ness kernel and the integral with respect to this process is carried out in the work [7]. Below we give an overview of the main results.

Definition 2.1. Two-sided centered Poisson process $(\tilde{\lambda}_t)_{t\in\mathbb{R}}$ is defined as follows: $\tilde{\lambda}_t = \tilde{\lambda}^{(1)}_t$, if $t\geq 0$ and $\tilde{\lambda}_t = -\tilde{\lambda}^{(2)}_{(-t)-} := -\lim_{\varepsilon\to 0+}\tilde{\lambda}^{(2)}_{(-t-\varepsilon)}$, if $t<0$, where $\tilde{\lambda}^{(1)}$ and $\tilde{\lambda}^{(2)}$ are independent and identically distributed centered Poisson processes.

Definition 2.2. Let $(\tilde{\lambda}_t)_{t\in\mathbb{R}}$ be a two-sided Poisson process on $\mathbb{R}$, $f_H(t,s)$ is the Mandelbrot – van Ness kernel. For $H\in(1/2,1)$ a stochastic process

$$X_t = \int_{-\infty}^{t} f_H(t,s)\,d\tilde{\lambda}_s,$$

is called a fractional Poisson process with the Mandelbrot – van Ness kernel. This integral exists in L^2-sense (as the limit in L^2 of integrals of a sequence of approximating $f_H(t,s)$ step functions; the limit does not depend on the choice of the sequence of approximating functions).

The fractional Poisson process X_t can be represented as follows:

$$X_t = \int_{\mathbb{R}} \left(I^{H-1/2}_{-}\mathbb{1}_{(0,t)}\right)(s)\,d\tilde{\lambda}_s,$$

where I_{-} is the right-sided Riemann-Liouville fractional integral operator on $\mathbb{R}$.

Define the space $\mathcal{H}$ as the completion of $L^1(\mathbb{R})\cap L^2(\mathbb{R})$ with respect to the norm

$$\|g\|_{\mathcal{H}} := \left(\lambda\int_{\mathbb{R}}\left(I^{H-1/2}_{-}f\right)^2(s)\,ds\right)^{1/2}.$$

It is known from [7] that for the functions $f\in L^1(\mathbb{R})\cap L^2(\mathbb{R})$

$$\int_{\mathbb{R}}\left(I^{H-1/2}_{-}f\right)^2(s)\,ds<\infty.$$

Let $\phi\colon\mathbb{R}\to\mathbb{R}$ be a simple function:

$$\phi(s) = \sum_{i=1}^{n-1} a_i\mathbb{1}_{[s_i,s_{i+1})}(s),$$

where $a_i\in\mathbb{R}$, $i=1,\ldots,n$ and $-\infty<s_1<s_2<\ldots<s_n<\infty$. Notice that simple functions belong to the space $\mathcal{H}$.

The integral with respect to the fractional Poisson process with the Mandelbrot – van Ness kernel is defined for simple functions at first. Let ϕ be a simple function. Then

$$\int_{\mathbb{R}}\phi(s)\,dX_s = \int_{\mathbb{R}}\left(I^{H-1/2}_{-}\phi\right)\,d\tilde{\lambda}_s.$$

Also from [7] we have the following L^2-isometry:

$$\mathsf{E}\left(\int_{\mathbb{R}}\phi(s)\,dX_s\right)^2 = \mathsf{E}\left(\int_{\mathbb{R}}\left(I^{H-1/2}_{-}\phi\right)\,d\tilde{\lambda}_s\right)^2 = \lambda\int_{\mathbb{R}}\left(I^{H-1/2}_{-}\phi\right)^2(s)\,ds = \|\phi\|^2_{\mathcal{H}}.$$

We can extend the definition of the integral with respect to the fractional Poisson process for the class of functions $f \in \mathcal{H}$. Namely,

$$\int_{\mathbb{R}} f(s)\, dX_s = \int_{\mathbb{R}} \left(I_-^{H-1/2} f \right)(s)\, d\tilde{\lambda}_s$$

with equality in L^2-sense.

The noncentered fractional Poisson process Y_t with the Molchan–Golosov kernel is defined as follows:

$$Y_t = \int_0^t z_H(t,s)\, d\lambda_s,$$

where λ_s is the simple Poisson process with intensity λ, $z_H(t,s)$ is the Molchan–Golosov kernel, and the integral exists in the pathwise sense due to step structure of the Poisson process and smooth properties of $z_H(t,s)$, mentioned above.

The centered fractional Poisson process $\tilde{Y}_t$ with the Molchan–Golosov kernel is defined as follows:

$$\tilde{Y}_t = \int_0^t z_H(t,s)\, d\tilde{\lambda}_s,$$

where $\tilde{\lambda}_s = \lambda_s - \lambda s$ is the centered Poisson process. $\tilde{Y}_t$ is defined as the integral with respect to the square integrable martingale. So the centered fractional Poisson process exists as the integral in L^2 sense.

Later on we shall consider both integrals with respect to the non-centered and centered fractional Poisson process.

3. Distribution characteristics of the fractional Poisson process with the Molchan–Golosov kernel

Using well-known formulas for the integrals with respect to the Poisson process, we obtain the following first and second noncentral moments for the noncentered fractional Poisson process with the Molchan–Golosov kernel:

$$\begin{aligned} m_1 &= \lambda \int_0^t z_H(t,s)\, ds = \lambda C_H \int_0^t u^{H-1/2} \int_0^u s^{1/2-H} (u-s)^{H-3/2}\, ds\, du \\ &= \lambda C_H \frac{\pi}{\sin(\pi(3/2-H))} \int_0^t u^{H-1/2}\, du = \lambda C_H \frac{\pi}{\sin(\pi(3/2-H))} \frac{t^{H+1/2}}{H+1/2}, \end{aligned} \tag{2}$$

$$\begin{aligned} m_2 &= \lambda^2 \left(\int_0^t z_H(t,s)\, ds \right)^2 + \lambda \int_0^t z_H^2(t,s)\, ds \\ &= \lambda^2 \left(C_H \frac{\pi}{\sin(\pi(3/2-H))} \frac{1}{H+1/2} \right)^2 t^{2H+1} + \lambda t^{2H}. \end{aligned}$$

Here we have used the equality

$$\int_0^t z_H^2(t,s)\, ds = t^{2H}$$

that follows from the representation (1) of the fractional Brownian motion and the form of its covariance function.

We know that the fractional Brownian motion has stationary increments. Now we investigate whether the property of stationarity of increments holds for the fractional Poisson process with the Molchan–Golosov kernel.

Lemma 3.1. *Both for centered and noncentered fractional Poisson process with the Molchan–Golosov kernel the property of stationarity of increments in general does not hold.*

Proof. Consider the noncentered process, for the centered one the proof is similar. We investigate whether the characteristic function of the fractional Poisson process

$$\mathsf{E}\exp\{iuY_t\}, \qquad u \in \mathbb{R},\ 0 < t < \infty,$$

and that of its increment

$$\mathsf{E}\exp\{iu(Y_{t+t_1} - Y_{t_1})\}, \qquad u \in \mathbb{R},\ 0 < t, t_1 < \infty,$$

are equal.

We use propositions 2.4, 2.6 [13] and results by [7]. Thus if some process Z allows the representation $Z_t = \int_{\mathbb{R}} f(t,s)\,dL_s$, where L is Lévy process with characteristic triplet $(0,0,\nu)$ without Gaussian component, such that $\mathsf{E}\,L_1 = 0$, $\mathsf{E}\,L_1^2 < \infty$, then

$$\mathsf{E}(\exp(iuZ_t)) = \exp\left(\int_{\mathbb{R}}\int_{\mathbb{R}} \left(e^{if(t,s)ux} - 1 - if(t,s)ux\right)\nu(dx)\,ds\right).$$

Therefore for fractional Poisson process Y_t with the Molchan–Golosov kernel we obtain the following characteristic function:

$$\mathsf{E}\exp[iuY_t] = \exp\left\{\int_{\mathbb{R}} \lambda\left(\exp\{iuz_H(t,s)\mathbb{1}_{[0,t]}(s)\} - 1\right)\,ds\right\}. \tag{3}$$

Further,

$$\begin{aligned} Y_{t+t_1} - Y_{t_1} &= \int_0^{t+t_1} z_H(t+t_1,s)\,d\lambda_s - \int_0^{t_1} z_H(t_1,s)\,d\lambda_s \\ &= \int_0^{t+t_1} \left(z_H(t+t_1,s) - z_H(t_1,s)\right)\,d\lambda_s, \end{aligned}$$

where in the last equality we use that according to definition we have $z_H(t,s) = 0$ if condition $0 < s < t$ does not hold. So

$$\begin{aligned} &\mathsf{E}\exp\{iu(Y_{t+t_1} - Y_{t_1})\} \\ &\quad = \exp\left\{\int_{\mathbb{R}} \lambda\big(\exp\{iu(z_H(t+t_1,s) - z_H(t_1,s))\mathbb{1}_{[0,t+t_1]}(s)\} - 1\big)\,ds\right\}. \end{aligned} \tag{4}$$

We compare (3) and (4). It is sufficient to compare

$$z_H(t,s)\cdot\mathbb{1}_{[0,t]}(s) = c_H s^{1/2-H}\int_s^t u^{H-1/2}(u-s)^{H-3/2}\,du\cdot\mathbb{1}_{[0,t]}(s), \tag{5}$$

and

$$\begin{aligned} &(z_H(t+t_1,s) - z_H(t_1,s))\cdot\mathbb{1}_{[0,t+t_1]}(s) \\ &\quad = c_H s^{1/2-H}\int_{t_1}^{t+t_1} u^{H-1/2}(u-s)^{H-3/2}\,du\cdot\mathbb{1}_{[0,t+t_1]}(s). \end{aligned} \tag{6}$$

As (5) and (6) are not equal, for the noncentered fractional Poisson process with the Molchan–Golosov kernel the property of stationarity of increments in general does not hold. □

4. Integral with respect to the fractional Poisson process with the Molchan–Golosov kernel and estimate of its second moment in terms of the norm of the integrand in $L_{1/H}([0,T])$ space

Consider the noncentered fractional Poisson process Y_t with the Molchan–Golosov kernel. Let a function f be defined on $[0,T]$, $H \in (\frac{1}{2},1)$. Define the following operator:

$$\left(K_T^H f\right)(s) = C_H s^{1/2-H}\left(I_{T-}^{H-1/2}(\cdot)^{H-1/2} f\right)(s), \qquad s \in (0,T),$$

where $I_{T-}^{H-1/2}$ is the right-sided Riemann–Liouville fractional operator defined in Section 2.

Introduce the spaces

$$L^2_{H,Pois}([0,T]) = \{f\colon [0,T] \to \mathbb{R} \mid K^H_T f \in L^2([0,T])\}$$

with the norm

$$\|f\|_{L^2_{H,Pois}([0,T])} = \left\|K^H_T f\right\|_{L^2([0,T])},$$

and

$$\tilde{L}^2_{H,Pois}([0,T]) = \left\{ f \in L^2_{H,Pois}([0,T]) \text{ and } (\cdot)^{H-1/2} f(\cdot) \in L_p([0,T]) \text{ for some } p > \frac{1}{H-1/2} \right\}$$

with the same norm.

Define for $f \in L^2_{H,Pois}([0,T])$ the integral with respect to the fractional Poisson processes in the following way:

$$\int_0^T f(s)\,dY_s = \int_0^T \left(K^H_T f\right)(s)\,d\lambda_s \tag{7}$$

and

$$\int_0^T f(s)\,d\tilde{Y}_s = \int_0^T \left(K^H_T f\right)(s)\,d\tilde{\lambda}_s. \tag{8}$$

Thus, we have the analogy with the construction of the integral with respect to the fractional Brownian motion. Note that from (2)

$$\tilde{Y}_t = \int_0^t z_H(t,s)\,d\tilde{\lambda}_s = Y_t - \lambda \int_0^t z_H(t,s)\,ds = Y_t - EY_t,$$

and

$$\int_0^T f(s)\,d\tilde{Y}_s := \int_0^T f(s)\,dY_s - \int_0^T f(s)\,d(EY_s),$$

where both integrals exist in L^2-sense.

Lemma 4.1. 1. *For $f \in L^2_{H,Pois}([0,T])$ both integrals* (7) *and* (8) *exist in L^2 sense.*
2. *For $f \in \tilde{L}^2_{H,Pois}([0,T])$ integral* (7) *exists in the pathwise sense.*

Proof. 1. Consider the noncentered case, the centered one is considered similarly. It holds [11] that $(K^H_T \mathbb{1}_{[0,t)})(s) = z_H(t,s)$. Using properties of integrals with respect to the Poisson process for step functions we have:

$$\mathsf{E}\left(\int_0^T \left(K^H_T f\right)(s)\,d\lambda_s\right)^2 = \lambda^2 \left(\int_0^T (K^H_T f)(s)\,ds\right)^2 + \lambda \int_0^T \left(K^H_T f\right)^2(s)\,ds$$
$$\le (\lambda^2 T + \lambda)\|f\|^2_{L^2_{H,Pois}([0,T])}, \tag{9}$$

$$\mathsf{E}\left(\int_0^T (K^H_T f)(s)d\tilde{\lambda}_s\right)^2 = \lambda \int_0^T \left(K^H_T f\right)^2(s)\,ds = \lambda\|f\|^2_{L^2_{H,Pois}([0,T])},$$

where λ is the intensity of the Poisson process. Note that according to [12] step functions are dense in $L^2_{H,Pois}([0,T])$. Therefore, we can approximate the function $f \in L^2_{H,Pois}([0,T])$ by step functions f_n in $L^2_{H,Pois}([0,T])$ and to define the integral of the function f with respect to the fractional Poisson process using as follows:

$$\int_0^T f(s)\,dY_s = \lim_{n\to\infty} \int_0^T f_n(s)\,dY_s \quad \text{— convergence in } L^2(\mathsf{P}).$$

2. Consider the integrand of the right side of equality (7):

$$\left(K^H_T f\right)(s) = C_H s^{1/2-H} \left(I^{H-1/2}_{T-} \cdot^{H-1/2} f\right)(s).$$

Functions belonging to the space $\tilde{L}^2_{H,Pois}([0,T])$ satisfies the conditions of the Theorem 3.6 [15]. According to it the function $\left(I_{T-}^{H-1/2} \cdot^{H-1/2} f\right)(s)$ is Hölder of order $H-\frac{1}{2}-\frac{1}{p}$ on $(0,T)$.

In the right side of equality (7) the integration is with respect to the Poisson process, which is a process of bounded variation on $[0,T]$. Also, according to the properties of the Poisson process a.s. there exists such $\varepsilon(\omega)>0$, that $\lambda_s=0$ for all $s\in[0,\varepsilon(\omega)]$. Thus, on $(\varepsilon(\omega),T)$ a.s. we have the continuous function, that can be integrated with respect to the process of bounded variation. Therefore, integral (7) exists in the pathwise sense. □

Remark 4.1. To estimate the second moment of the integral with respect to the fractional Poisson process with the Molchan–Golosov kernel we need to estimate $\int_0^T (K_T^H f)^2(s)\,ds$. It can be done similarly to the fractional Brownian motion case [9] with the help of (9). Denote $\alpha = H-\frac{1}{2}$. Then

$$\begin{aligned}
\mathsf{E}\left(\int_0^T f(s)\,dY_s\right)^2 &= \mathsf{E}\left(\int_0^T \left(K_T^H f\right)(s)\,d\lambda_s\right)^2 \le \left(\lambda^2 T+\lambda\right)\int_0^T \left(K_T^H f\right)^2(s)\,ds \\
&= C\int_0^T s^{-2\alpha}\left(\int_s^T f(u)u^{\alpha}(u-s)^{\alpha-1}\,du\right)^2 ds \\
&\le CB(1-2\alpha,\alpha)\int_0^T\int_0^T f(u)f(v)|u-v|^{2\alpha-1}\,du\,dv \\
&\le C\|f\|^2_{L_{1/H}([0,T])}.
\end{aligned}$$

5. Estimate of higher moments of integral with respect to the fractional Poisson process with the Molchan–Golosov kernel

Let $f\in L^2_{H,Pois}([0,T])$. Recall that

$$\int_0^T f(s)\,dY_s = \int_0^T \left(K_T^H f\right)(s)\,d\lambda_s = \int_0^T \left(K_T^H f\right)(s)\,d\tilde{\lambda}_s + \int_0^T \left(K_T^H f\right)(s)\lambda\,ds, \tag{10}$$

and the first integral in the right-hand side of (10) exists as the integral with respect to the square-integrable martingale $\tilde{\lambda}_s=\lambda_s-\lambda s$.

Now we are in the position to establish moment inequalities for integral with respect to the noncentered fractional Poisson process with the Molchan–Golosov kernel. For the centered process the similar bounds hold with obvious modification.

Theorem 5.1. *Let $f\in\tilde{L}^2_{H,Pois}([0,T])$, $H\in(\frac{1}{2},1)$. Then for any k such that $0<k<\frac{1}{2H-1}$ there exists the constant $C_k=C(H,k)$, such that for any $T>0$*

$$\mathsf{E}\left|\int_0^T f(s)\,dY_s\right|^{2k} \le C_k\left\|K_T^H f\right\|^{2k}_{L^{2k}_{[0,T]}} + C_k\lambda^{2k}\left(\int_0^T u^{H-1/2}|f(u)|\,du\right)^{2k}.$$

Proof. We consider moments of the order $2k$:

$$\begin{aligned}
\mathsf{E}\left|\int_0^T f(s)\,dY_s\right|^{2k} &= \mathsf{E}\left(\left|\int_0^T \left(K_T^H f\right)(s)\,d\tilde{\lambda}_s + \int_0^T \left(K_T^H f\right)(s)\lambda\,ds\right|\right)^{2k} \\
&\le 2^{2k}\,\mathsf{E}\left(\left|\int_0^T \left(K_T^H f\right)(s)\,d\tilde{\lambda}_s\right|\right)^{2k} + 2^{2k}\left(\int_0^T |(K_T^H f)(s)|\lambda\,ds\right)^{2k} \\
&:= I_1+I_2.
\end{aligned}$$

To bound the first integral, we use the Bichteler–Jacod inequality (see, e.g., [6]):

$$I_1 \le C \int_0^T \left(\left(K_T^H f\right)^{2k} \lambda + \left(\left(K_T^H f\right)^2 (s)\lambda \right)^k \right) ds \le C_k \int_0^T \left(K_T^H f\right)^{2k} (s)\, ds$$
$$= C_k \|K_T^H f\|^{2k}_{L^{2k}_{[0,T]}}.$$

Establish whether the last integral exists:

$$\tilde I_1 := \int_0^T \left(K_T^H f\right)^{2k} (s)\, ds = C_H^{2k} \int_0^T s^{(1/2-H)2k} \left(I_{T-}^{H-1/2} \cdot^{H-1/2} f\right)^{2k} (s)\, ds.$$

Remind that according to the definition of the space $\tilde L^2_{H,Pois}([0,T])$ there exists some $p > \frac{1}{H-1/2}$: $(\cdot)^{H-1/2} f(\cdot) \in L_p([0,T]$. So the same way as in the proof of the Lemma 4.1 we can establish that the function $\left(I_{T-}^{H-1/2} \cdot^{H-1/2} f\right)(s)$ is Hölder of order $H - \frac12 - \frac1p$ on $(0,T)$. Thus $\tilde I_1$ is finite if and only if

$$\int_0^T s^{(1/2-H)2k}\, ds < \infty,$$

and due to the condition $k < \frac{1}{2H-1}$ the integral $\tilde I_1$ is finite.

For estimation of I_2 we use the equality

$$\begin{aligned}\int_0^T \left(K_T^H f\right)(s)\, ds &= C_H \int_0^T s^{1/2-H} \int_s^T u^{H-1/2} f(u) (u-s)^{H-3/2}\, du\, ds \\ &= C_H \int_0^T u^{H-1/2} f(u) \int_0^u s^{1/2-H} (u-s)^{H-3/2}\, ds\, du \\ &= C_H \frac{\pi}{\sin(\pi(3/2-H))} \int_0^T u^{H-1/2} f(u)\, du.\end{aligned}$$

Therefore

$$\int_0^T \left| \left(K_T^H f\right)(s) \right| ds \le C_H \int_0^T u^{H-1/2} |f(u)|\, du,$$

and

$$I_2 \le C_k \lambda^{2k} \left(\int_0^T u^{H-1/2} |f(u)|\, du \right)^{2k}. \qquad \square$$

References

1. L. Beghin and E. Orsingher, *Fractional Poisson process and related planar random motions*, Bernoulli **12** (2006), no. 6, 1099–1126.
2. A. Benassi, S. Cohen, and J. Istas, *Identification and properties of real harmonizable fractional Lévy motions*, Bernoulli **8** (2002), 97–115.
3. C. Jost, *Transformation formulas for fractional Brownian motion*, Stochastic Processes and their Applications **116** (2006), 1341–1357.
4. F. Mainardi, R. Gorenflo, and E. Scalas, *A fractional generalization of the Poisson processes*, Vietnam J. Math. **32** (2004), 53–64.
5. F. Mainardi, R. Gorenflo, and A. Vivoli, *Renewal processes of Mittag–Leffler and Wright type*, Fract. Calc. Appl. Anal. **8** (2005), no. 1, 7–38.
6. C. Marinelli, C. Prévôt, and M. Röckner, *Regular dependence on initial data for stochastic evolution equations with multiplicative Poisson noise*, Journal of Functional Analysis **258** (2010), no. 2, 616–649.
7. T. Marquardt, *Fractional Lévy process with an application to long memory moving average processes*, Bernoulli **12** (2006), no. 6, 1099–1126.
8. M. M. Meerschaert, E. Nane, and P. Vellaisamy, *The fractional Poisson process and the inverse stable subordinator*, Electron. J. Probab. **16** (2011), no. 59, 1600–1620.
9. J. Memin, Yu. Mishura, and E. Valkeila, *Inequalities for the moments of Wiener integrals with respect to fractional Brownian motions*, Stat. Prob. Lett. **51** (2001), 197–206.

10. Yu. Mishura, *Stochastic calculus for fractional Brownian motion and related processes*, Springer, Berlin, 2008.
11. G. Molchan and J. Golosov, *Gaussian stationary processes with asymptotic power spectrum*, Soviet Mathematics Doklady **10** (1969), no. 1, 134–137.
12. V. Pipiras and M. Taqqu, *Are classes of deterministic integrands for fractional Brownian motion on an interval complete?*, Bernoulli **7** (2001), no. 6, 873–897.
13. B. S. Rajput and J. Rosinski, *Spectral representaion of infinitely divisible processes*, Probab. Th. Rel. Fields **82** (1989), 451–487.
14. O. N. Repin and A. I. Saichev, *Fractional Poisson law*, Radiophysics and Quantum Electronics **43** (2000), no. 9, 738–741.
15. S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional integrals and derivatives. Theory and applications*, Gordon and Breach Science Publishers, 1993.
16. H. Tikanmaki and Yu. Mishura, *Fractional Lévy process as a result of compact interval intergal transformation*, Stochastic Analysis and Applications **29** (2011), 1081–1101.
17. V. P. Zubchenko, *Properties of solutions of stochastic differential equations with random coefficients, non-Lipschitzian diffusion, and Poisson measures*, Theor. Probability and Math. Statist. **82** (2011), 11–26.

Department of probability theory, statistics and actuarial mathematics, Mechanics and mathematics faculty, Kyiv National Taras Shevchenko University, Volodymyrska st., 64, Kyiv, 01601, Ukraine

E-mail address: myus@univ.kiev.ua

Department of probability theory, statistics and actuarial mathematics, Mechanics and mathematics faculty, Kyiv National Taras Shevchenko University, Volodymyrska st., 64, Kyiv, 01601, Ukraine

E-mail address: v_zubchenko@ukr.net

Received 26/02/2013

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 140–155

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 140–155

QUASI-STATIONARY DISTRIBUTIONS FOR PERTURBED DISCRETE TIME REGENERATIVE PROCESSES

UDC 519.21

MIKAEL PETERSSON

Abstract. Non-linearly perturbed discrete time regenerative processes with regenerative stopping times are considered. We define the quasi-stationary distributions for such processes and present conditions for their convergence. Under some additional assumptions, the quasi-stationary distributions can be expanded in asymptotic power series with respect to the perturbation parameter. We give an explicit recurrence algorithm for calculating the coefficients in these asymptotic expansions. Applications to perturbed alternating regenerative processes with absorption and perturbed risk processes are presented.

Анотація. У статті розглядаються процеси відновлення з дискретным часом із нелінійними збуреннями. Визначено квазі-стаціонарні розподіли для таких процесів та представлено умови для їх збіжності. При деяких додаткових умовах для квазі-стаціонарних розподілів можна виписати асимптотичні розклади у степеневі ряди відносно параметру збурення. Представлено точний рекуррентний алгоритм для обчислення коефіцієнтів цих асимптотичних розкладів. Представлено застосування результатів для процесів відновлення із збуреннями з поглинанням та для процесів ризику із збуреннями.

Аннотация. В статье рассматриваются процессы восстановления с дискретным временем с нелинейными возмущениями. Определены квази-стационарные распределения для таких процессов и представлены условия их сходимости. При некоторых дополнительных условиях, для квази-стационарных распределений могут быть выписаны асимптотические разложения в степенные ряды относительно параметра возмущения. Представлен точный рекуррентный алгоритм для вычисления коэффициентов этих асимптотических разложений. Представлены приложения результатов для процессов восстановления с возмущениями с поглощением и для процессов риска с возмущениями.

1. Introduction

Many stochastic systems has a random lifetime, the process is terminated due to some rare event. This means that the stationary distribution of such process will be degenerated. However, before the lifetime of the system goes to an end, one can often observe something that resembles a stationary distribution. It is often of interest to describe such behaviour, so-called quasi-stationary phenomena.

In this paper we study such phenomena for discrete time regenerative processes with regenerative stopping time. Roughly speaking, such a process $\xi(n)$, $n = 0, 1, \dots$, regenerates at random times $\tau_1, \tau_2, \dots$, and has random lifetime μ which regenerates jointly with the process.

In particular, such processes includes discrete time semi-Markov processes with absorption. For example, $\xi(n)$ can be a Markov chain, $\tau_1, \tau_2, \dots$, the return times to some fixed state and μ, the first hitting time of some fixed state.

As a special case, when $\mu = \infty$ almost surely, this class of processes includes regenerative processes without stopping time.

2000 *Mathematics Subject Classification.* Primary 60K05, 34E10; Secondary 60K25.

Key words and phrases. Regenerative process, Renewal equation, Non-linear perturbation, Quasi-stationary distribution, Asymptotic expansion, Risk process.

Under some conditions, it can be shown that for such processes there exists a probability distribution $\pi(A)$ such that

$$\mathsf{P}\{\xi(n) \in A/\mu > n\} \to \pi(A) \quad \text{as } n \to \infty.$$

We call this distribution the *quasi-stationary* distribution and use it to describe the quasi-stationary phenomena of the process. In the case $\mu = \infty$ almost surely, $\pi(A)$ is the usual stationary distribution.

Quasi-stationary distributions have been studied intensively since the 1960's. Some of the important early works are Vere-Jones (1962), Kingman (1963), Darroch and Seneta (1965, 1967) and Seneta and Vere-Jones (1966).

In this paper, we consider the case when $\xi(n)$ is perturbed and that the perturbation is described by a small parameter ε. Furthermore, it is assumed that some continuity conditions hold at $\varepsilon = 0$ for certain characteristics of the process $\xi^{(\varepsilon)}(n)$, regarded as a function of ε. This allows us to interpret $\xi^{(\varepsilon)}(n)$ as a perturbed version of the process $\xi^{(0)}(n)$.

We want the quasi-stationary distribution $\pi^{(\varepsilon)}(A)$ of the process $\xi^{(\varepsilon)}(n)$ to be an approximation of the quasi-stationary distribution $\pi^{(0)}(A)$ of the process $\xi^{(0)}(n)$, that is $\pi^{(\varepsilon)}(A) \to \pi^{(0)}(A)$ as $\varepsilon \to 0$.

We give conditions such that the quasi-stationary distribution can be expanded as

$$\pi^{(\varepsilon)}(A) = \pi^{(0)}(A) + f_1(A)\varepsilon + \cdots + f_k(A)\varepsilon^k + o(\varepsilon^k),$$

where the coefficients $f_1(A), \ldots, f_k(A)$ can be calculated from an explicit recurrence algorithm.

Theoretical results are illustrated by example to the model of an alternating regenerative process with absorption. Under perturbation conditions on distributions of sojourn times and absorption probabilities, we give explicit the asymptotic expansion for the quasi-stationary distribution for such a process.

It is also shown how the results can be used in order to obtain approximations of the ruin probability for a discrete time risk process. We describe how an asymptotic expansion of the ruin probability can be obtained under perturbation conditions on claim probabilities and claim distributions.

The results in the present paper continue the line of research studies of the perturbed renewal equation in discrete time in Gyllenberg and Silvestrov (1994), Englund and Silvestrov (1997), Silvestrov (2000) and Petersson and Silvestrov (2012, 2013).

Corresponding results for perturbed regenerative processes in continuous time can be found in the book Gyllenberg and Silvestrov (2008) where one can also find an extended bibliography of works in the area.

Some works related to asymptotic expansions for perturbed Markov chains are Kartashov (1988, 1996), Latouche (1988), Hassin and Haviv (1992), Khasminskii, Yin and Zhang (1996), Yin and Zhang (2003), Altman, Avrachenkov and Núñes-Queija (2004), Koroliuk and Limnios (2005) and Yin and Nguyen (2009).

2. Quasi-Stationary Distributions for Regenerative Processes

For every $\varepsilon \geq 0$, let $\xi^{(\varepsilon)}(n)$ be a regenerative process in discrete time with a measurable phase space (X, Γ) and regeneration times $0 = \tau_0^{(\varepsilon)} < \tau_1^{(\varepsilon)} < \ldots$, and let $\mu^{(\varepsilon)}$ be a random variable defined on the same probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and taking values in the set $\{0, 1, \ldots, \infty\}$.

We call $\mu^{(\varepsilon)}$ a regenerative stopping time for the regenerative process $\xi^{(\varepsilon)}(n)$ if for any $A \in \Gamma$, the probabilities $P^{(\varepsilon)}(n, A) = \mathsf{P}\big\{\xi^{(\varepsilon)}(n) \in A, \mu^{(\varepsilon)} > n\big\}$ satisfies the renewal equation,

$$P^{(\varepsilon)}(n, A) = q^{(\varepsilon)}(n, A) + \sum_{k=0}^{n} P^{(\varepsilon)}(n-k, A) f^{(\varepsilon)}(k), \qquad n = 0, 1, \ldots, \tag{1}$$

where

$$q^{(\varepsilon)}(n, A) = \mathsf{P}\left\{\xi^{(\varepsilon)}(n) \in A, \mu^{(\varepsilon)} > n, \tau_1^{(\varepsilon)} > n\right\}$$

and

$$f^{(\varepsilon)}(n) = \mathsf{P}\left\{\tau_1^{(\varepsilon)} = n, \mu^{(\varepsilon)} > \tau_1^{(\varepsilon)}\right\}.$$

Note that the defect $f^{(\varepsilon)}$ of the distribution $f^{(\varepsilon)}(n)$ is given by the stopping probability in one regeneration period for the process $\xi^{(\varepsilon)}(n)$, that is,

$$f^{(\varepsilon)} = 1 - \sum_{n=0}^{\infty} f^{(\varepsilon)}(n) = \mathsf{P}\left\{\mu^{(\varepsilon)} \le \tau_1^{(\varepsilon)}\right\}.$$

We consider the case where the stopping probability in one regeneration period for the limiting process may be positive, i.e., $f^{(0)} \in [0, 1)$.

Assume that the distributions $f^{(\varepsilon)}(n)$ satisfy the following conditions:

A: (**a**) $f^{(\varepsilon)}(n) \to f^{(0)}(n)$ as $\varepsilon \to 0$, $n = 0, 1, \ldots$, where the limiting distribution is non-periodic and not concentrated in zero.

(**b**) $f^{(\varepsilon)} \to f^{(0)} \in [0, 1)$ as $\varepsilon \to 0$.

B: There exists $\delta > 0$ such that

(**a**) $\overline{\lim}_{0 \le \varepsilon \to 0} \sum_{n=0}^{\infty} e^{\delta n} f^{(\varepsilon)}(n) < \infty$.

(**b**) $\sum_{n=0}^{\infty} e^{\delta n} f^{(0)}(n) > 1$.

Let us consider the characteristic equation

$$\sum_{n=0}^{\infty} e^{\rho n} f^{(\varepsilon)}(n) = 1. \tag{2}$$

The following result from Petersson and Silvestrov (2012, 2013) gives some basic properties of $\rho^{(\varepsilon)}$ that will be used in what follows.

Lemma 2.1. *Assume that* **A** *and* **B** *hold. Then there exists a unique non-negative solution $\rho^{(\varepsilon)}$ of the characteristic equation* (2) *for ε small enough and $\rho^{(\varepsilon)} \to \rho^{(0)} < \delta$ as $\varepsilon \to 0$.*

For the rest of the paper, assume that **A** and **B** hold so that $\rho^{(\varepsilon)}$ is well defined for ε small enough. Also, to avoid repetition, we assume that ε always is small enough to satisfy the statements of Lemma 2.1. If both sides in (1) are multiplied by $e^{\rho^{(\varepsilon)} n}$, we see that the transformed probabilities $\tilde{P}(n, A) = e^{\rho^{(\varepsilon)} n} P(n, A)$ satisfy

$$\tilde{P}^{(\varepsilon)}(n, A) = \tilde{q}^{(\varepsilon)}(n, A) + \sum_{k=0}^{n} \tilde{P}^{(\varepsilon)}(n-k, A) \tilde{f}^{(\varepsilon)}(k), \qquad A \in \Gamma, \tag{3}$$

where

$$\tilde{q}^{(\varepsilon)}(n, A) = e^{\rho^{(\varepsilon)} n} q^{(\varepsilon)}(n, A), \qquad \tilde{f}^{(\varepsilon)}(n) = e^{\rho^{(\varepsilon)} n} f^{(\varepsilon)}(n).$$

It follows from the definition of $\rho^{(\varepsilon)}$, that (3) is a proper renewal equation. In order to apply the classical discrete time renewal theorem, the following condition is imposed on the tail probabilities of $\tau_1^{(\varepsilon)} \wedge \mu^{(\varepsilon)}$.

C: There exists $\gamma > 0$ such that

$$\overline{\lim_{\varepsilon \to 0}} \sum_{n=0}^{\infty} e^{(\rho^{(0)} + \gamma) n} q^{(\varepsilon)}(n, X) < \infty.$$

For any $\varepsilon \geq 0$, we define the *quasi-stationary distribution* of $\xi^{(\varepsilon)}(n)$ by

$$\pi^{(\varepsilon)}(A) = \frac{\sum_{n=0}^{\infty} e^{\rho^{(\varepsilon)} n} q^{(\varepsilon)}(n, A)}{\sum_{n=0}^{\infty} e^{\rho^{(\varepsilon)} n} q^{(\varepsilon)}(n, X)}, \qquad A \in \Gamma. \tag{4}$$

Under conditions **A**, **B** and **C** the quasi-stationary distribution is well defined for sufficiently small ε.

Let us denote

$$\Gamma_0 = \left\{ A \in \Gamma \colon q^{(\varepsilon)}(n, A) \to q^{(0)}(n, A) \text{ as } \varepsilon \to 0, n = 0, 1, \dots \right\}$$

We assume the following:

D: $X \in \Gamma_0$.

Note that Γ_0 is an algebra but does not necessarily coincide with Γ.

The first part of the following result motivates why it is natural to call $\pi^{(\varepsilon)}(A)$ quasi-stationary distributions. The second part gives conditions for convergence of $\pi^{(\varepsilon)}(A)$ for sets $A \in \Gamma_0$.

Theorem 2.2. *Assume that* **A**, **B** *and* **C** *hold.*

(**i**) *Then there exists* $\varepsilon_0 > 0$ *such that for every* $\varepsilon \leq \varepsilon_0$,

$$\mathsf{P}\left\{ \xi^{(\varepsilon)}(n) \in A / \mu^{(\varepsilon)} > n \right\} \to \pi^{(\varepsilon)}(A) \quad \text{as } n \to \infty, \ A \in \Gamma.$$

(**ii**) *If, in addition, condition* **D** *holds, then*

$$\pi^{(\varepsilon)}(A) \to \pi^{(0)}(A) \quad \text{as } \varepsilon \to 0, \ A \in \Gamma_0.$$

Proof. First note that if the limiting distribution $f^{(0)}(n)$ is non-periodic, then there exists a finite positive integer N such that

$$\gcd\left\{ 1 \leq n \leq N \colon f^{(0)}(n) > 0 \right\} = 1.$$

It follows from condition **A** that the distributions $\tilde{f}^{(\varepsilon)}(n)$ are non-periodic for ε sufficiently small, say $\varepsilon \leq \varepsilon_1$. Let $\tilde{m}_1^{(\varepsilon)}$ denote the expectation of $\tilde{f}^{(\varepsilon)}(n)$. Since $\rho^{(0)} < \delta$ we can choose $\delta_0 > 0$ such that $\rho^{(0)} < \delta - \delta_0$. Let $C = \sup_{n \geq 0} n e^{-\delta_0 n}$. Since $\rho^{(\varepsilon)} \to \rho^{(0)}$ and condition **B** holds it follows that

$$\begin{aligned}
\varlimsup_{\varepsilon \to 0} \tilde{m}_1^{(\varepsilon)} &= \varlimsup_{\varepsilon \to 0} \sum_{n=0}^{\infty} n e^{\rho^{(\varepsilon)} n} f^{(\varepsilon)}(n) \leq \varlimsup_{\varepsilon \to 0} \sum_{n=0}^{\infty} n e^{(\delta - \delta_0) n} f^{(\varepsilon)}(n) \\
&\leq C \varlimsup_{\varepsilon \to 0} \sum_{n=0}^{\infty} e^{\delta n} f^{(\varepsilon)}(n) < \infty.
\end{aligned}$$

It follows that $\tilde{m}_1^{(\varepsilon)}$ is finite for all ε small enough, say $\varepsilon \leq \varepsilon_2$. Condition **C** implies that for any $A \in \Gamma$

$$\begin{aligned}
\varlimsup_{\varepsilon \to 0} \sum_{n=0}^{\infty} \tilde{q}^{(\varepsilon)}(n, A) &= \varlimsup_{\varepsilon \to 0} \sum_{n=0}^{\infty} e^{\rho^{(\varepsilon)} n} \mathsf{P}\left\{ \xi^{(\varepsilon)}(n) \in A, \mu^{(\varepsilon)} > n, \tau_1^{(\varepsilon)} > n \right\} \\
&\leq \varlimsup_{\varepsilon \to 0} \sum_{n=0}^{\infty} e^{(\rho^{(0)} + \gamma) n} \mathsf{P}\left\{ \tau_1^{(\varepsilon)} \wedge \mu^{(\varepsilon)} > n \right\} < \infty,
\end{aligned}$$

so there exists $\varepsilon_3 > 0$ such that $\sum_{n=0}^{\infty} \tilde{q}^{(\varepsilon)}(n, A) < \infty$ for all $\varepsilon \leq \varepsilon_3$.

Define $\varepsilon_0 = \min\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$. It follows from the classical discrete time renewal theorem that for any $\varepsilon \leq \varepsilon_0$,

$$\tilde{P}^{(\varepsilon)}(n, A) \to \frac{1}{\tilde{m}_1^{(\varepsilon)}} \sum_{n=0}^{\infty} \tilde{q}^{(\varepsilon)}(n, A) \quad \text{as } n \to \infty, \ A \in \Gamma. \tag{5}$$

Part (**i**) follows from relation (5) and the following equality,

$$\mathsf{P}\left\{\xi^{(\varepsilon)}(n) \in A/\mu^{(\varepsilon)} > n\right\} = \tilde{P}^{(\varepsilon)}(n, A) \,/\, \tilde{P}^{(\varepsilon)}(n, X).$$

Lemma 2.1, condition **C** and the definition of Γ_0 implies that

$$\lim_{\varepsilon\to 0} \sum_{n=0}^{\infty} e^{\rho^{(\varepsilon)} n} q^{(\varepsilon)}(n, A) = \sum_{n=0}^{\infty} e^{\rho^{(0)} n} q^{(0)}(n, A) < \infty, \qquad A \in \Gamma_0. \tag{6}$$

Part (**ii**) now follows from relations (4) and (6), and condition **D**. □

3. Asymptotic Expansions of Quasi-Stationary Distributions

A problem with $\pi^{(\varepsilon)}(A)$ is that the expression defining it is rather complicated. Both numerator and denominator are represented as infinite sums and involves $\rho^{(\varepsilon)}$, which is only given as the solution to the nonlinear equation (2). However, under some perturbation conditions, $\pi^{(\varepsilon)}(A)$ can be expanded in an asymptotic power series with respect to ε.

In order to do this, we first need to expand $\rho^{(\varepsilon)}$. This can be done under some perturbation conditions on the following mixed power-exponential moments of the distributions $f^{(\varepsilon)}(n)$,

$$\phi^{(\varepsilon)}(\rho, r) = \sum_{n=0}^{\infty} n^r e^{\rho n} f^{(\varepsilon)}(n), \qquad \rho \geq 0,\ r = 0, 1, \dots$$

To expand the quasi-stationary distribution, some perturbation conditions on the following mixed power-exponential moment type functionals of $q^{(\varepsilon)}(n, A)$ are also needed,

$$\omega^{(\varepsilon)}(\rho, r, A) = \sum_{n=0}^{\infty} n^r e^{\rho n} q^{(\varepsilon)}(n, A), \qquad \rho \geq 0,\ r = 0, 1, \dots,\ A \in \Gamma.$$

The perturbation conditions are the following:

$\mathbf{P_1^{(k)}}$: $\phi^{(\varepsilon)}(\rho^{(0)}, r) = \phi^{(0)}(\rho^{(0)}, r) + a_{1,r}\varepsilon + \cdots + a_{k-r,r}\varepsilon^{k-r} + o(\varepsilon^{k-r})$, for $r = 0, \dots, k$, where $|a_{n,r}| < \infty$, $n = 1, \dots, k-r$, $r = 0, \dots, k$.

$\mathbf{P_2^{(k)}}$: $\omega^{(\varepsilon)}(\rho^{(0)}, r, A) = \omega^{(0)}(\rho^{(0)}, r, A) + b_{1,r}(A)\varepsilon + \cdots + b_{k-r,r}(A)\varepsilon^{k-r} + o(\varepsilon^{k-r})$, for $r = 0, \dots, k$, where $A \in \Gamma_0$ and $|b_{n,r}(A)| < \infty$, $n = 1, \dots, k-r$, $r = 0, \dots, k$.

For convenience, we define $a_{0,r} = \phi^{(0)}(\rho^{(0)}, r)$ and $b_{0,r} = \omega^{(0)}(\rho^{(0)}, r, A)$ for $r = 0, \dots, k$ and $A \in \Gamma_0$.

Now we are ready to give the expansion of $\pi^{(\varepsilon)}(A)$. The details are presented in the following theorem.

Theorem 3.1. *Suppose that* **A**, **B** *and* $\mathbf{P_1^{(k)}}$ *hold.*

(**i**) *Then the root $\rho^{(\varepsilon)}$ of the characteristic equation* (2) *has the asymptotic expansion*

$$\rho^{(\varepsilon)} = \rho^{(0)} + c_1\varepsilon + \cdots + c_k\varepsilon^k + o(\varepsilon^k).$$

The coefficients $c_1, \dots, c_k$ are given by the recurrence formulas

$$\begin{aligned} c_1 &= -a_{1,0}/a_{0,1}, \\ c_n &= -\frac{1}{a_{0,1}}\left(a_{n,0} + \sum_{q=1}^{n-1} a_{n-q,1}c_q \right. \\ &\qquad \left. + \sum_{m=2}^{n}\sum_{q=m}^{n} a_{n-q,m} \cdot \sum_{n_1,\dots,n_{q-1} \in D_{m,q}} \prod_{p=1}^{q-1} \frac{c_p^{n_p}}{n_p!}\right), \qquad n = 2, \dots, k, \end{aligned} \tag{7}$$

where $D_{m,q}$ is the set of all nonnegative integer solutions to the system

$$n_1 + \cdots + n_{q-1} = m,\ n_1 + \cdots + (q-1)n_{q-1} = q.$$

(**ii**) *If, in addition,* $\mathbf{C}$*,* $\mathbf{D}$ *and* $\mathbf{P_2^{(k)}}$ *hold, then for any* $A \in \Gamma_0$ *the following asymptotic expansion holds,*

$$\pi^{(\varepsilon)}(A) = \pi^{(0)}(A) + f_1(A)\varepsilon + \cdots + f_k(A)\varepsilon^k + o(\varepsilon^k).$$

The coefficients $f_1(A), \ldots, f_n(A)$ *are given by*

$$f_n(A) = \frac{1}{d_0(X)} \left(d_n(A) - \sum_{q=0}^{n-1} d_{n-q}(X) f_q(A) \right), \tag{8}$$

where

$$d_0(A) = \omega^{(0)}(\rho^{(0)}, 0, A)$$

and $f_0(A) = \pi^{(0)}(A)$*. The coefficients* $d_1(A), \ldots, d_k(A)$ *are given by*

$$\begin{aligned} d_1(A) &= b_{1,0}(A) + b_{0,1}(A)c_1, \\ d_n(A) &= b_{n,0}(A) + \sum_{q=1}^{n} b_{n-q,1}(A)c_q \\ &\quad + \sum_{m=2}^{n}\sum_{q=m}^{n} b_{n-q,m}(A) \cdot \sum_{n_1,\ldots,n_{q-1} \in D_{m,q}} \prod_{p=1}^{q-1} \frac{c_p^{n_p}}{n_p!}, \qquad n = 2, \ldots, k. \end{aligned} \tag{9}$$

Proof. For the proof of part (**i**), see Petersson and Silvestrov (2012, 2013). Here we give the proof of part (**ii**).

Let $\Delta^{(\varepsilon)} = \rho^{(\varepsilon)} - \rho^{(0)}$. Using the Taylor expansion of the exponential function, we obtain for any $n = 0, 1, \ldots,$

$$e^{\rho^{(\varepsilon)} n} = e^{\rho^{(0)} n} \left(\sum_{r=0}^{k} \frac{n^r (\Delta^{(\varepsilon)})^r}{r!} + \frac{n^{k+1}(\Delta^{(\varepsilon)})^{k+1}}{(k+1)!} e^{|\Delta^{(\varepsilon)}| n} \theta_{k+1}^{(\varepsilon)}(n) \right),$$

where $0 \leq \theta_{k+1}^{(\varepsilon)}(n) \leq 1$. Since $\rho^{(\varepsilon)} \to \rho^{(0)}$, there exists $\beta < \rho^{(0)} + \gamma$ and $\varepsilon_1 = \varepsilon_1(\beta)$ such that

$$\rho^{(0)} + |\Delta^{(\varepsilon)}| < \beta, \ \varepsilon \leq \varepsilon_1.$$

Let $\tilde{C}_r = \sup_{n \geq 0} n^r e^{(\rho^{(0)} + \gamma - \beta) n}$. From condition $\mathbf{C}$ it follows that there exists $\varepsilon_2 > 0$ and a constant C_r such that

$$\begin{aligned} \omega^{(\varepsilon)}(\beta, r, A) &= \sum_{n=0}^{\infty} n^r e^{\beta n} q^{(\varepsilon)}(n, A) \\ &\leq \tilde{C}_r \sum_{n=0}^{\infty} e^{(\rho^{(0)} + \gamma) n} \mathsf{P}\left\{ \tau_1^{(\varepsilon)} \wedge \mu^{(\varepsilon)} > n \right\} \leq C_r, \qquad \varepsilon \leq \varepsilon_2. \end{aligned}$$

Define $\varepsilon_0 = \varepsilon_0(\beta) := \min\{\varepsilon_1(\beta), \varepsilon_2\}$. Substituting the Taylor expansion of $e^{\rho^{(\varepsilon)} n}$ into the definition of $\omega^{(\varepsilon)}(\rho^{(\varepsilon)}, 0, A)$ yields

$$\begin{aligned} \omega^{(\varepsilon)}(\rho^{(\varepsilon)}, 0, A) &= \omega^{(\varepsilon)}(\rho^{(0)}, 0, A) + \omega^{(\varepsilon)}(\rho^{(0)}, 1, A)\Delta^{(\varepsilon)} + \cdots \\ &\quad + \omega^{(\varepsilon)}(\rho^{(0)}, k, A)(\Delta^{(\varepsilon)})^k / k! + r_{k+1}^{(\varepsilon)} (\Delta^{(\varepsilon)})^{k+1}, \end{aligned} \tag{10}$$

where

$$r_{k+1}^{(\varepsilon)} = \frac{1}{(k+1)!} \sum_{n=0}^{\infty} n^{k+1} e^{(\rho^{(0)} + |\Delta^{(\varepsilon)}|) n} \theta_{k+1}^{(\varepsilon)}(n) q^{(\varepsilon)}(n, A).$$

If $\varepsilon \leq \varepsilon_0$, the right hand side of (10) is finite and

$$r_{k+1}^{(\varepsilon)} \leq \frac{1}{(k+1)!} \omega^{(\varepsilon)}(\beta, k+1, A) \leq \frac{C_{k+1}}{(k+1)!}.$$

It follows that there exists a finite constant M_{k+1} and numbers $0 \leq \theta_{k+1}^{(\varepsilon)} \leq 1$ such that

$$r_{k+1}^{(\varepsilon)} = M_{k+1}\theta_{k+1}^{(\varepsilon)}, \qquad \varepsilon \leq \varepsilon_0. \tag{11}$$

Since **A**, **B** and $\mathbf{P_1^{(k)}}$ hold, it follows from part (**i**) that

$$\Delta^{(\varepsilon)} = c_1\varepsilon + \cdots + c_k\varepsilon^k + o(\varepsilon^k). \tag{12}$$

Substituting (11), (12) and condition $\mathbf{P_2^{(k)}}$ into the right hand side of (10) when $k = 0$ we see that $\omega^{(\varepsilon)}(\rho^{(\varepsilon)}, 0, A) \to \omega^{(0)}(\rho^{(0)}, 0, A)$ as $\varepsilon \to 0$, which means that we have the representation

$$\omega^{(\varepsilon)}(\rho^{(\varepsilon)}, 0, A) = \omega^{(0)}(\rho^{(0)}, 0, A) + \omega_0^{(\varepsilon)}(A), \tag{13}$$

where $\omega_0^{(\varepsilon)}(A) \to 0$ as $\varepsilon \to 0$.

Now assume that $k = 1$. If we substitute (11), (12), (13) and condition $\mathbf{P_2^{(k)}}$ into the right hand side of (10), divide by ε and let $\varepsilon \to 0$, it is found that

$$\frac{\omega_0^{(\varepsilon)}(A)}{\varepsilon} \to b_{1,0}(A) + b_{0,1}(A)c_1 \quad \text{as } \varepsilon \to 0. \tag{14}$$

Using (13) and (14) we obtain the asymptotic representation

$$\omega^{(\varepsilon)}(\rho^{(\varepsilon)}, 0, A) = \omega^{(0)}(\rho^{(0)}, 0, A) + d_1(A)\varepsilon + \omega_1^{(\varepsilon)}(A),$$

where $d_1(A) = b_{1,0}(A) + b_{0,1}(A)c_1$ and $\omega_1^{(\varepsilon)}(A)$ is of order $o(\varepsilon)$.

If $k \geq 2$, we can continue in this way and build an asymptotic expansion of order k for $\omega^{(\varepsilon)}(\rho^{(\varepsilon)}, 0, A)$. Once the existence of the expansion is proved, the coefficients can be found by collecting the coefficients of equal powers of ε in the expansion of the following expression,

$$\begin{aligned}
&\big(b_{0,0}(A) + \cdots + b_{k,0}(A)\varepsilon^k + o(\varepsilon^k)\big) \\
&\quad + \big(b_{0,1}(A) + \cdots + b_{k-1,1}(A)\varepsilon^{k-1} + o(\varepsilon^{k-1})\big) \\
&\qquad \times \big(c_1\varepsilon + \cdots + c_k\varepsilon^k + o(\varepsilon^k)\big) + \cdots \\
&\quad + (b_{0,k}(A) + o(1))\,\big(c_1\varepsilon + \cdots + c_k\varepsilon^k + o(\varepsilon^k)\big)^k/k! + o(\varepsilon^k).
\end{aligned}$$

This yields the expansion

$$\omega^{(\varepsilon)}(\rho^{(\varepsilon)}, 0, A) = \omega^{(0)}(\rho^{(0)}, 0, A) + d_1(A)\varepsilon + \cdots + d_k(A)\varepsilon^k + o(\varepsilon^k), \tag{15}$$

where the coefficients $d_1(A), \ldots, d_k(A)$ are given according to (9).

The quasi-stationary distribution can be written as

$$\pi^{(\varepsilon)}(A) = \frac{\omega^{(\varepsilon)}(\rho^{(\varepsilon)}, 0, A)}{\omega^{(\varepsilon)}(\rho^{(\varepsilon)}, 0, X)}, \qquad A \in \Gamma.$$

For sets $A \in \Gamma_0$, the numerator can be expanded as in equation (15). By condition **D**, we always have $X \in \Gamma_0$ so the denominator can also be expanded. Thus, for any $A \in \Gamma_0$,

$$\pi^{(\varepsilon)}(A) = \frac{\omega^{(0)}(\rho^{(0)}, 0, A) + d_1(A)\varepsilon + \cdots + d_k(A)\varepsilon^k + o(\varepsilon^k)}{\omega^{(0)}(\rho^{(0)}, 0, X) + d_1(X)\varepsilon + \cdots + d_k(X)\varepsilon^k + o(\varepsilon^k)}. \tag{16}$$

Using (16), we can build the expansion of $\pi^{(\varepsilon)}(A)$ similarly to how we built the expansion of $\omega^{(\varepsilon)}(\rho^{(\varepsilon)}, 0, A)$. To do this, first note that with $k = 0$ in (16) it immediately follows that $\pi^{(\varepsilon)}(A) \to \pi^{(0)}(A)$, which means that we have the representation

$$\pi^{(\varepsilon)}(A) = \pi^{(0)}(A) + \pi_0^{(\varepsilon)}(A), \tag{17}$$

where $\pi_0^{(\varepsilon)}(A) \to 0$ as $\varepsilon \to 0$.

Now put $k = 1$ in (16). Since $\omega^{(0)}(\rho^{(0)}, 0, X) > 0$, it follows that the denominator of (16) is positive for ε small enough. Substituting (17) into the left hand side of (16), rearranging and using the identity $\pi^{(0)}(A)d_0(X) = d_0(A)$ gives the following for sufficiently small ε,

$$\pi_0^{(\varepsilon)}(A)d_0(X) + d_1(X)f_0(A) + o(\varepsilon) = d_1(A)\varepsilon + o(\varepsilon).$$

Dividing both sides by ε and letting $\varepsilon \to 0$, we conclude that

$$\frac{\pi_0^{(\varepsilon)}(A)}{\varepsilon} \to \frac{1}{d_0(X)}\left(d_1(A) - d_1(X)f_0(A)\right) \quad \text{as } \varepsilon \to 0.$$

Using this and (17), the following asymptotic representation is obtained,

$$\pi^{(\varepsilon)}(A) = \pi^{(0)}(A) + f_1(A)\varepsilon + \pi_1^{(\varepsilon)}(A),$$

where $f_1(A) = (d_1(A) - d_1(X)f_0(A))/d_0(X)$ and $\pi_1^{(\varepsilon)}(A)$ is of order $o(\varepsilon)$.

This proves part (**ii**) when $k = 1$.

If $k \geq 2$ we can continuing in this way and prove that the asymptotic expansion of $\pi^{(\varepsilon)}(A)$ exists. When we know that the expansion exists, the coefficients can be found in the following way. Consider the equation

$$\begin{aligned}&\left(f_0(A) + f_1(A)\varepsilon + \cdots + f_k(A)\varepsilon^k + o(\varepsilon^k)\right)\\ &\qquad \times \left(d_0(X) + d_1(X)\varepsilon + \cdots + d_k(X)\varepsilon^k + o(\varepsilon^k)\right)\\ &\quad = \left(d_0(A) + d_1(A)\varepsilon + \cdots + d_k(A)\varepsilon^k + o(\varepsilon^k)\right).\end{aligned}$$

The coefficients $f_k(A)$ are obtained by equating the coefficients of ε^k in both sides of this equation. This yields

$$\pi^{(\varepsilon)}(A) = \pi^{(0)}(A) + f_1(A)\varepsilon + \cdots + f_k(A)\varepsilon^k + o(\varepsilon^k),$$

where $f_1(A), \ldots, f_k(A)$ are given according to the recurrent relation in equation (8). $\square$

4. Perturbed Alternating Regenerative Processes

In this section, we consider a perturbed alternating regenerative process with absorption. We assume that the process $\eta^{(\varepsilon)}(n)$ starts in state 1 and stays there for a time with distribution $g_1^{(\varepsilon)}(n)$ before it jumps down to state 0. Then the process remains in state 0 for a time with distribution $g_0^{(\varepsilon)}(n)$. Now, with some small probability $p^{(\varepsilon)}$ the process is absorbed in state -1 or with probability $1 - p^{(\varepsilon)}$ the process starts over in state 1.

Such a process can be interpreted as the state of a machine which is successively repaired after break-downs. The states 0 and 1 then represents that the machine is broken or working while -1 is the absorption state of fatal non-repairable failure.

Respectively, $g_1^{(\varepsilon)}(n)$ is the distribution of the time between repair and failure and $g_0^{(\varepsilon)}(n)$ is the distribution of the time to locate the error after a break-down. The absorption probability $p^{(\varepsilon)}$ corresponds to a fatal error such that the machine can not be repaired. The first hitting time to the state -1 is the lifetime of the system.

We assume the following condition, preventing instant jumps:

E: $g_0^{(\varepsilon)}(0) = g_1^{(\varepsilon)}(0) = 0$ for all $\varepsilon \geq 0$.

Mathematically, this is described by a discrete time semi-Markov process.

Let $\left(\eta_k^{(\varepsilon)}, \kappa_k^{(\varepsilon)}\right)$ be a Markov renewal chain with phase space $X \times \{1, 2, \ldots\}$, where $X = \{-1, 0, 1\}$, and with transition probabilities

$$q_{ij}^{(\varepsilon)}(n) = \mathsf{P}\left\{\eta_{k+1}^{(\varepsilon)} = j, \kappa_{k+1}^{(\varepsilon)} = n/\eta_k^{(\varepsilon)} = i\right\}, \qquad i, j \in X,\ n = 1, 2, \ldots,$$

given by

$$q_{ij}^{(\varepsilon)}(n) = \begin{cases} g_1^{(\varepsilon)}(n) & i=1, j=0, \\ (1-p^{(\varepsilon)})g_0^{(\varepsilon)}(n) & i=0, j=1, \\ p^{(\varepsilon)}g_0^{(\varepsilon)}(n) & i=0, j=-1, \\ \chi(n=1) & i=j=-1, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\nu^{(\varepsilon)}(n) = \max\{k\colon \gamma^{(\varepsilon)}(k) \le n\}$, where $\gamma^{(\varepsilon)}(0) = 0$ and $\gamma^{(\varepsilon)}(k) = \kappa_1^{(\varepsilon)} + \cdots + \kappa_k^{(\varepsilon)}$ for $k \ge 1$.

The discrete time semi-Markov process $\eta^{(\varepsilon)}(n)$ can be defined by the following relation,

$$\eta^{(\varepsilon)}(n) = \eta^{(\varepsilon)}_{\nu^{(\varepsilon)}(n)}, \qquad n = 0, 1, \ldots .$$

Let

$$\nu_j^{(\varepsilon)} = \min\left\{k \ge 1\colon \eta_k^{(\varepsilon)} = j\right\}.$$

Then the absorption time is given by $\mu^{(\varepsilon)} = \gamma^{(\varepsilon)}\big(\nu_{-1}^{(\varepsilon)}\big)$ and the first regeneration time is given by $\tau_1^{(\varepsilon)} = \gamma^{(\varepsilon)}\big(\nu_1^{(\varepsilon)}\big)$.

The process described above is illustrated in Figure 1.

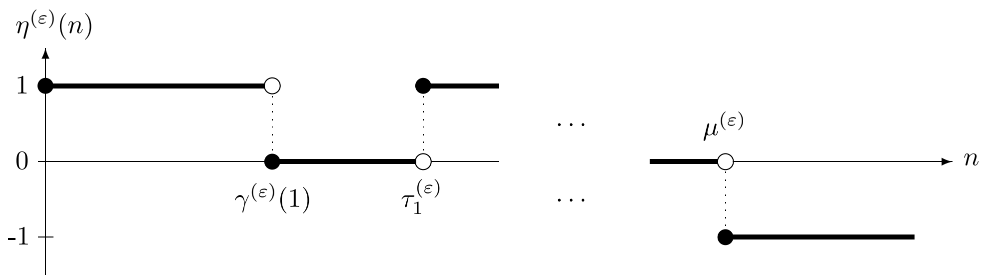


FIGURE 1. Realization of the process $\eta^{(\varepsilon)}(n)$.

In the definition of a regenerative process with regenerating stopping time it is assumed that the regeneration times are proper random variables. In the process defined above this is not the case. However, the transition probabilities from the absorbing state can be modified in such a way that the return times to state 1 are proper random variables, and that the probabilities $\mathsf{P}\big\{\eta^{(\varepsilon)}(n) = i, \mu^{(\varepsilon)} > n\big\}$ are the same for the modified process. We can then apply the results from Sections 2 and 3 to the modified process and interpret the results for the original process.

The weak continuity conditions at $\varepsilon = 0$ are formulated in terms of the local characteristics of the alternating regenerative process as follows.

F: **(a)** $g_i^{(\varepsilon)}(n) \to g_i^{(0)}(n)$ as $\varepsilon \to 0$, $n = 0, 1, \ldots$, $i = 0, 1$.
(b) $p^{(\varepsilon)} \to p^{(0)} \in [0, 1)$ as $\varepsilon \to 0$.

We also need the following non-periodicity condition.

G: At least one of the distributions $g_0^{(0)}(n)$ and $g_1^{(0)}(n)$ is non-periodic.

We introduce the following mixed power-exponential moment generating functions for distributions of sojourn times,

$$\psi_i^{(\varepsilon)}(\rho, r) = \sum_{n=0}^{\infty} n^r e^{\rho n} g_i^{(\varepsilon)}(n), \qquad \rho \ge 0,\ r = 0, 1, \ldots,\ i = 0, 1. \tag{18}$$

Also, consider the following mixed power-exponential moment generating functions,

$$\phi^{(\varepsilon)}(\rho, r) = \sum_{n=0}^{\infty} n^r e^{\rho n} f^{(\varepsilon)}(n), \qquad \rho \geq 0,\ r = 0, 1, \ldots, \tag{19}$$

where

$$f^{(\varepsilon)}(n) = \mathsf{P}\left\{\tau_1^{(\varepsilon)} = n, \mu^{(\varepsilon)} > \tau_1^{(\varepsilon)}\right\}, \qquad n = 0, 1, \ldots$$

For the exponential moment generating functions, the following relation is obtained,

$$\begin{aligned} \phi^{(\varepsilon)}(\rho, 0) &= \left(1 - p^{(\varepsilon)}\right) \sum_{n=0}^{\infty} e^{\rho n}\, \mathsf{P}\left\{\kappa_1^{(\varepsilon)} + \kappa_2^{(\varepsilon)} = n\right\} \\ &= \left(1 - p^{(\varepsilon)}\right) \psi_0^{(\varepsilon)}(\rho, 0) \psi_1^{(\varepsilon)}(\rho, 0), \qquad \rho \geq 0. \end{aligned} \tag{20}$$

From this it follows that existence of (18) and (19) for ε small enough is guaranteed by the following Cramér type condition:

H: There exists $\delta > 0$ such that
- **(a)** $\overline{\lim}_{0 \leq \varepsilon \to 0} \psi_i^{(\varepsilon)}(\delta, 0) < \infty$, $i = 0, 1$.
- **(b)** $(1 - p^{(0)}) \psi_0^{(0)}(\delta, 0) \psi_1^{(0)}(\delta, 0) > 1$.

We will also use the following mixed power-exponential moment generating functions,

$$\omega_i^{(\varepsilon)}(\rho, r) = \sum_{n=0}^{\infty} n^r e^{\rho n} q_i^{(\varepsilon)}(n), \qquad \rho \geq 0,\ r = 0, 1, \ldots,\ i = 0, 1, \tag{21}$$

where

$$q_i^{(\varepsilon)}(n) = \mathsf{P}\left\{\eta^{(\varepsilon)}(n) = i, \tau_1^{(\varepsilon)} \wedge \mu^{(\varepsilon)} > n\right\}, \qquad n = 0, 1, \ldots,\ i = 0, 1.$$

If condition **E**–**H** hold, then condition $\mathbf{A} - \mathbf{D}$ hold, so the results in Section 2 can be applied. Lemma 2.1 implies that for ε small enough there exists a unique root $\rho^{(\varepsilon)}$ of the characteristic equation

$$\phi^{(\varepsilon)}(\rho, 0) = 1. \tag{22}$$

It is worth noticing that the solution to equation (22) satisfies $\rho^{(\varepsilon)} = 0$ if and only if $p^{(\varepsilon)} = 0$, and $\rho^{(\varepsilon)} > 0$ if and only if $p^{(\varepsilon)} > 0$.

It follows from Theorem 2.2 that that for ε sufficiently small,

$$\lim_{n \to \infty} \mathsf{P}\left\{\eta^{(\varepsilon)}(n) = j / \mu^{(\varepsilon)} > n\right\} = \pi_j^{(\varepsilon)}\left(\rho^{(\varepsilon)}\right), \qquad j = 0, 1,$$

where

$$\pi_j^{(\varepsilon)}\left(\rho^{(\varepsilon)}\right) = \frac{\omega_j^{(\varepsilon)}(\rho^{(\varepsilon)}, 0)}{\omega_0^{(\varepsilon)}(\rho^{(\varepsilon)}, 0) + \omega_1^{(\varepsilon)}\left(\rho^{(\varepsilon)}, 0\right)}, \qquad j = 0, 1. \tag{23}$$

If conditions $\mathbf{P_1^{(k)}}$ and $\mathbf{P_2^{(k)}}$ hold for the generating functions (19) and (21), it follows from Theorem 3.1 that we can build an asymptotic expansion for the quasi-stationary distribution (23). However, it is more convenient to use perturbation conditions for local characteristics of the process $\eta^{(\varepsilon)}(n)$. Therefore, we formulate perturbation conditions on the generating functions of the distributions of sojourn times and the absorption probabilities, and then show how these conditions are related to $\mathbf{P_1^{(k)}}$ and $\mathbf{P_2^{(k)}}$.

We assume the following:

$\mathbf{P_3^{(k)}}$: $p^{(\varepsilon)} = p^{(0)} + p[1]\varepsilon + \cdots + p[k]\varepsilon^k + o(\varepsilon^k)$, where $|p[n]| < \infty$, $n = 1, \ldots, k$.

$\mathbf{P_4^{(k)}}$: $\psi_i^{(\varepsilon)}(\rho^{(0)}, r) = \psi_i^{(0)}(\rho^{(0)}, r) + \psi_i[1, r]\varepsilon + \cdots + \psi_i[k - r, r]\varepsilon^{k-r} + o(\varepsilon^{k-r})$, for $r = 0, \ldots, k$, $i = 0, 1$, where $|\psi_i[n, r]| < \infty$, $n = 1, \ldots, k - r$, $r = 0, \ldots, k$, $i = 0, 1$.

Observe that for $n = 0, 1, \dots,$

$$\mathsf{P}\{\eta^{(\varepsilon)}(n) = i, \tau_1^{(\varepsilon)} \wedge \mu^{(\varepsilon)} > n\} = \begin{cases} \mathsf{P}\{\kappa_1^{(\varepsilon)} \leq n, \kappa_1^{(\varepsilon)} + \kappa_2^{(\varepsilon)} > n\} & i = 0, \\ \mathsf{P}\{\kappa_1^{(\varepsilon)} > n\} & i = 1. \end{cases}$$

Using this relation, we obtain for $\rho \geq 0$,

$$\omega_i^{(\varepsilon)}(\rho, 0) = \begin{cases} \psi_1^{(\varepsilon)}(\rho, 0)\varphi_0^{(\varepsilon)}(\rho, 0) & i = 0, \\ \varphi_1^{(\varepsilon)}(\rho, 0) & i = 1, \end{cases} \tag{24}$$

where, for $i = 0, 1$,

$$\varphi_i^{(\varepsilon)}(\rho, 0) = \begin{cases} (\psi_i^{(\varepsilon)}(\rho, 0) - 1)/(e^\rho - 1) & \rho > 0, \\ \psi_i^{(\varepsilon)}(0, 1) & \rho = 0. \end{cases} \tag{25}$$

Under condition $\mathbf{H}$, the derivative of any order of the function $\varphi_i^{(\varepsilon)}(\rho, 0)$ exists for $0 \leq \rho \leq \beta < \delta$ and sufficiently small ε. Denote the derivative of order r of this function by $\varphi_i^{(\varepsilon)}(\rho, r)$. It follows directly from (25) that

$$\psi_i^{(\varepsilon)}(\rho, 0) = \varphi_i^{(\varepsilon)}(\rho, 0)(e^\rho - 1) + 1, \qquad \rho \geq 0. \tag{26}$$

By differentiating equation (26) r times and rearranging, it follows that the derivative of order $r = 1, 2, \dots,$ of the function $\varphi_i^{(\varepsilon)}(\rho, 0)$ is given by the recursive relation

$$\varphi_i^{(\varepsilon)}(\rho, r) = \begin{cases} (\psi_i^{(\varepsilon)}(\rho, r) - e^\rho \sum_{j=0}^{r-1} \binom{r}{j} \varphi_i^{(\varepsilon)}(\rho, j))/(e^\rho - 1) & \rho > 0, \\ (\psi_i^{(\varepsilon)}(0, r+1) - \sum_{j=0}^{r-1} \binom{r+1}{j} \varphi_i^{(\varepsilon)}(0, j))/(r+1) & \rho = 0. \end{cases}$$

In the following, suppose that condition $\mathbf{P_3^{(k)}}$ holds, together with condition $\mathbf{P_4^{(k)}}$ if $\rho^{(0)} > 0$, or together with condition $\mathbf{P_4^{(k+1)}}$ if $\rho^{(0)} = 0$. Then the following asymptotic expansion holds,

$$\varphi_i^{(\varepsilon)}\left(\rho^{(0)}, r\right) = \varphi_i^{(0)}\left(\rho^{(0)}, r\right) + \varphi_i[1, r]\varepsilon + \cdots + \varphi_i[k-r, r]\varepsilon^{k-r} + o(\varepsilon^{k-r}). \tag{27}$$

Denote $\varphi_i[0, r] = \varphi_i^{(0)}(\rho^{(0)}, r)$.

In the case $\rho^{(0)} > 0$, the coefficients for $i = 0, 1$, are given by

$$\begin{aligned} &\varphi_i[n, r]\left(e^{\rho^{(0)}} - 1\right) \\ &\quad = \begin{cases} \psi_i[n, 0] - \delta(n, 0) & n = 0, \dots, k,\ r = 0, \\ \psi_i[n, r] - e^{\rho^{(0)}} \sum_{j=0}^{r-1} \binom{r}{j} \varphi_i[n, j] & n = 0, \dots, k-r,\ r = 1, \dots, k. \end{cases} \end{aligned} \tag{28}$$

In the case $\rho^{(0)} = 0$, the coefficients for $i = 0, 1$, are given by

$$\begin{aligned} &\varphi_i[n, r](r+1) \\ &\quad = \begin{cases} \psi_i[n, 1] & n = 0, \dots, k,\ r = 0, \\ \psi_i[n, r+1] - \sum_{j=0}^{r-1} \binom{r+1}{j} \varphi_i[n, j] & n = 0, \dots, k-r,\ r = 1, \dots, k. \end{cases} \end{aligned} \tag{29}$$

Differentiating equation (20) and (24) r times with respect to ρ and evaluating at $\rho = \rho^{(0)}$ yields for any $r = 0, 1, \dots,$

$$\phi^{(\varepsilon)}\left(\rho^{(0)}, r\right) = \left(1 - p^{(\varepsilon)}\right) \sum_{j=0}^{r} \binom{r}{j} \psi_0^{(\varepsilon)}\left(\rho^{(0)}, j\right) \psi_1^{(\varepsilon)}\left(\rho^{(0)}, r-j\right), \tag{30}$$

$$\omega_0^{(\varepsilon)}\left(\rho^{(0)}, r\right) = \sum_{j=0}^{r} \binom{r}{j} \psi_1^{(\varepsilon)}\left(\rho^{(0)}, j\right) \varphi_0^{(\varepsilon)}\left(\rho^{(0)}, r-j\right), \tag{31}$$

$$\omega_1^{(\varepsilon)}\left(\rho^{(0)}, r\right) = \varphi_1^{(\varepsilon)}\left(\rho^{(0)}, r\right). \tag{32}$$

It follows from equations (27)–(32) that conditions $\mathbf{P_1^{(k)}}$ and $\mathbf{P_2^{(k)}}$ are implied by conditions $\mathbf{P_3^{(k)}}$ and $\mathbf{P_4^{(k)}}$ in the case $\rho^{(0)} > 0$, and by conditions $\mathbf{P_3^{(k)}}$ and $\mathbf{P_4^{(k+1)}}$ in the case $\rho^{(0)} = 0$. We can find the relations between the coefficients by using arithmetic rules of asymptotic expansions.

The coefficients in condition $\mathbf{P_1^{(k)}}$ are for any $n = 0, \dots, k-r$ and $r = 0, \dots, k$ given by

$$\begin{gathered} a_{0,r} = \left(1 - p^{(0)}\right) h_{0,r}, \qquad a_{n,r} = \left(1 - p^{(0)}\right) h_{n,r} - \sum_{i=1}^{n} p[i] h_{n-i,r}, \\ h_{n,r} = \sum_{i=0}^{n} \sum_{j=0}^{r} \binom{r}{j} \psi_0[i,j] \psi_1[n-i, r-j]. \end{gathered} \tag{33}$$

The coefficients in condition $\mathbf{P_2^{(k)}}$ are for any $n = 0, \dots, k-r$ and $r = 0, \dots, k$ given by

$$\begin{gathered} b_{n,r}(\{0\}) = \sum_{i=0}^{n} \sum_{j=0}^{r} \binom{r}{j} \psi_1[i,j] \varphi_0[n-i, r-j], \\ b_{n,r}(\{1\}) = \varphi_1[n,r], \qquad b_{n,r}(X) = b_{n,r}(\{0\}) + b_{n,r}(\{1\}). \end{gathered} \tag{34}$$

It follows from Theorem 3.1 that an asymptotic expansion of order k exists for the quasi-stationary distribution (23). We can build the expansion using equations (7), (8), (9), (28), (29), (33) and (34).

5. Perturbed Risk Processes

This section shows how the results of the present paper can be used to obtain approximations for the ruin probability in a perturbed discrete time risk process.

For each $\varepsilon \geq 0$, let $X_1^{(\varepsilon)}, X_2^{(\varepsilon)}, \dots$ be a sequence of non-negative, independent and identically distributed random variables and set

$$Z_u^{(\varepsilon)}(n) = u + n - \sum_{k=1}^{n} X_k^{(\varepsilon)}, \qquad n = 0, 1, \dots,$$

where u is a non-negative integer.

We can interpret $Z_u^{(\varepsilon)}(n)$ as the capital of an insurance company (in units equivalent to expected premium per time unit) and $X_n^{(\varepsilon)}$ as the claims at moment n.

Let us denote $p^{(\varepsilon)} = \mathsf{P}\{X_1^{(\varepsilon)} > 0\}$ and $\mu^{(\varepsilon)} = \sum_{u=0}^{\infty} u g^{(\varepsilon)}(u)$ where

$$g^{(\varepsilon)}(u) = \mathsf{P}\left\{X_1^{(\varepsilon)} = u / X_1^{(\varepsilon)} > 0\right\}, \qquad u = 0, 1, \dots$$

An object of interest is the infinite time horizon ruin probability which is defined as

$$\Psi^{(\varepsilon)}(u) = \mathsf{P}\left\{\min_{n \geq 0} Z_u^{(\varepsilon)}(n) < 0\right\}, \qquad u = 0, 1, \dots$$

Define $\alpha^{(\varepsilon)} := \mathsf{E} X_1^{(\varepsilon)} = p^{(\varepsilon)} \mu^{(\varepsilon)}$. It can be shown that if $\alpha^{(\varepsilon)} \geq 1$, then $\Psi^{(\varepsilon)}(u) = 1$ for all $u \geq 0$. In the case $\alpha^{(\varepsilon)} \leq 1$, the ruin probability $\Psi^{(\varepsilon)}(u)$ satisfies the following discrete time renewal equation,

$$\Psi^{(\varepsilon)}(u) = q^{(\varepsilon)}(u) + \sum_{k=0}^{u} \Psi^{(\varepsilon)}(u-k) f^{(\varepsilon)}(k), \qquad u = 0, 1, \dots, \tag{35}$$

where, for $u = 0, 1, \dots,$

$$G^{(\varepsilon)}(u) = \sum_{k=0}^{u} g^{(\varepsilon)}(k), \qquad f^{(\varepsilon)}(u) = \alpha^{(\varepsilon)} \frac{1 - G^{(\varepsilon)}(u)}{\mu^{(\varepsilon)}}, \qquad q^{(\varepsilon)}(u) = \sum_{k=u+1}^{\infty} f^{(\varepsilon)}(k).$$

A derivation of this equation can be found, for example, in Petersson and Silvestrov (2012, 2013). It is similar with the well-known technique for deriving the corresponding renewal equation for a continuous time risk process, given, for example, in Feller (1966) and Grandell (1991).

We now introduce the following mixed power-exponential moment generating functions,

$$\varphi^{(\varepsilon)}(\rho, r) = \sum_{u=0}^{\infty} u^r e^{\rho u} \left(1 - G^{(\varepsilon)}(u)\right), \qquad \rho \geq 0,\ r = 0, 1, \ldots$$

Let us assume the following conditions:

I: **(a)** $g^{(\varepsilon)}(u) \to g^{(0)}(u)$ as $\varepsilon \to 0$, $u = 0, 1, \ldots$
(b) $p^{(\varepsilon)} \to p^{(0)}$ as $\varepsilon \to 0$.

J: There exists $\delta > 0$ such that
(a) $\overline{\lim}_{0 \leq \varepsilon \to 0} \varphi^{(\varepsilon)}(\delta, 0) < \infty$.
(b) $(\alpha^{(0)}/\mu^{(0)})\varphi^{(0)}(\delta, 0) > 1$.

Under conditions **I** and **J** there exists a unique non-negative root $\rho^{(\varepsilon)}$ for sufficiently small ε to the characteristic equation

$$\sum_{u=0}^{\infty} e^{\rho u} f^{(\varepsilon)}(u) = 1. \tag{36}$$

Using this, we can transform the renewal equation (35) into the following form,

$$\widetilde{\Psi}^{(\varepsilon)}(u) = \tilde{q}^{(\varepsilon)}(u) + \sum_{k=0}^{u} \widetilde{\Psi}^{(\varepsilon)}(u-k)\tilde{f}^{(\varepsilon)}(k), \qquad u = 0, 1, \ldots, \tag{37}$$

where

$$\widetilde{\Psi}^{(\varepsilon)}(u) = e^{\rho^{(\varepsilon)} u} \Psi^{(\varepsilon)}(u), \qquad \tilde{q}^{(\varepsilon)}(u) = e^{\rho^{(\varepsilon)} u} q^{(\varepsilon)}(u), \qquad \tilde{f}^{(\varepsilon)}(u) = e^{\rho^{(\varepsilon)} u} f^{(\varepsilon)}(u).$$

There is a close connection between renewal equations and regenerative processes. In fact, the solution $x(n)$ of a discrete time renewal equation where the distribution $f(n)$ and the forcing function $q(n)$ satisfies $0 \leq q(n) \leq 1 - \sum_{k=0}^{n} f(k)$ for all $n \geq 0$, can be related to the one-dimensional distributions of some discrete time regenerative process.

In our case, there exists a discrete time regenerative process $\xi^{(\varepsilon)}(n)$, $n = 0, 1, \ldots$, with regeneration times $0 = \tau_0^{(\varepsilon)} < \tau_1^{(\varepsilon)} < \ldots$, and phase space $\{0, 1\}$ such that for $u = 0, 1, \ldots$, we have

$$\mathsf{P}\left\{\xi^{(\varepsilon)}(u) = 1\right\} = \widetilde{\Psi}^{(\varepsilon)}(u) \tag{38}$$

and

$$\mathsf{P}\left\{\xi^{(\varepsilon)}(u) = 1, \tau_1^{(\varepsilon)} > u\right\} = \tilde{q}^{(\varepsilon)}(u), \qquad \mathsf{P}\left\{\tau_1^{(\varepsilon)} = u\right\} = \tilde{f}^{(\varepsilon)}(u). \tag{39}$$

We next show how this process can be constructed. It is similar with the construction in the corresponding continuous time model, given in Ekheden and Silvestrov (2011).

Let $\kappa_1^{(\varepsilon)}, \kappa_2^{(\varepsilon)}, \ldots$, be a sequence of independent random variables, each with distribution $\tilde{f}^{(\varepsilon)}(n)$, and let $U_1^{(\varepsilon)}, U_2^{(\varepsilon)}, \ldots$, be a sequence of independent random variables uniformly distributed on the interval $[0, 1]$. Furthermore, we assume that the two sequences are independent.

For $k = 0, 1, \ldots$, let

$$v^{(\varepsilon)}(k) = \begin{cases} \tilde{q}^{(\varepsilon)}(k)/(1 - \tilde{F}^{(\varepsilon)}(k)) & \text{if } 1 - \tilde{F}^{(\varepsilon)}(k) > 0, \\ 0 & \text{if } 1 - \tilde{F}^{(\varepsilon)}(k) = 0, \end{cases}$$

where $\tilde{F}^{(\varepsilon)}(k) = \sum_{u=0}^{k} \tilde{f}^{(\varepsilon)}(u)$, $k = 0, 1, \ldots$.

Let us now, for every $n = 1, 2, \ldots$, define a random process by

$$\eta_n^{(\varepsilon)}(k) = \chi\left(U_n^{(\varepsilon)} \leq v^{(\varepsilon)}(k)\right), \qquad k = 0, 1, \ldots .$$

Using this, we can define a regenerative process $\xi^{(\varepsilon)}(n)$ with regeneration times $\tau_k^{(\varepsilon)} = \kappa_1^{(\varepsilon)} + \cdots + \kappa_k^{(\varepsilon)}$, $k = 1, 2, \ldots$, by

$$\xi^{(\varepsilon)}(n) = \eta^{(\varepsilon)}_{\nu^{(\varepsilon)}(n)+1}\left(\zeta^{(\varepsilon)}(n)\right), \qquad n = 0, 1, \ldots,$$

where $\nu^{(\varepsilon)}(n) = \max\{k \colon \tau_k^{(\varepsilon)} \leq n\}$ is the number of regenerations up to and including time n, and $\zeta^{(\varepsilon)}(n) = n - \tau^{(\varepsilon)}_{\nu^{(\varepsilon)}(n)}$ is the time since the last regeneration.

By definition, $\xi^{(\varepsilon)}(n)$ is a regenerative process with phase space $\{0, 1\}$ and regeneration times $0 = \tau_0^{(\varepsilon)} < \tau_1^{(\varepsilon)} < \ldots$. It can be checked that for this process, relations (38) and (39) hold.

If conditions $\mathbf{I}-\mathbf{J}$ hold, then conditions $\mathbf{A}$–$\mathbf{D}$ hold for the functions $\tilde{f}^{(\varepsilon)}(u)$ and $\tilde{q}^{(\varepsilon)}(u)$. It follows from Theorem 2.2 that

$$\mathsf{P}\left\{\xi^{(\varepsilon)}(u) = 1\right\} \to \pi^{(\varepsilon)}, \qquad u \to \infty, \tag{40}$$

where

$$\pi^{(\varepsilon)} = \frac{\sum_{u=0}^{\infty} \mathsf{P}\left\{\xi^{(\varepsilon)}(u) = 1, \tau_1^{(\varepsilon)} > u\right\}}{\sum_{u=0}^{\infty} \mathsf{P}\left\{\tau_1^{(\varepsilon)} > u\right\}}. \tag{41}$$

Rewriting in terms of the claim distributions, relations (40) and (41) yield

$$e^{\rho^{(\varepsilon)} u}\Psi^{(\varepsilon)}(u) \to \frac{\sum_{n=0}^{\infty} e^{\rho^{(\varepsilon)} n} \sum_{k=n+1}^{\infty}(1 - G^{(\varepsilon)}(k))}{\sum_{n=0}^{\infty} n e^{\rho^{(\varepsilon)} n}(1 - G^{(\varepsilon)}(n))} \quad \text{as } u \to \infty. \tag{42}$$

Relation (42) can be seen as a discrete time analogue of the classical Cramér–Lundberg approximation.

Let us introduce the following mixed power-exponential moment generating functions,

$$\psi^{(\varepsilon)}(\rho, r) = \sum_{u=0}^{\infty} u^r e^{\rho u} g^{(\varepsilon)}(u), \qquad \rho \geq 0,\ r = 0, 1, \ldots,$$

$$\omega^{(\varepsilon)}(\rho, r) = \sum_{u=0}^{\infty} u^r e^{\rho u} \sum_{k=u+1}^{\infty} (1 - G^{(\varepsilon)}(k)), \qquad \rho \geq 0,\ r = 0, 1, \ldots .$$

In order to build an asymptotic expansion for the stationary distribution $\pi^{(\varepsilon)}$, we impose the following perturbation conditions:

$\mathbf{P_5^{(k)}}$: $p^{(\varepsilon)} = p^{(0)} + p[1]\varepsilon + \cdots + p[k]\varepsilon^k + o(\varepsilon^k)$, where $|p[n]| < \infty$, $n = 1, \ldots, k$.

$\mathbf{P_6^{(k)}}$: $\psi^{(\varepsilon)}(\rho^{(0)}, r) = \psi^{(0)}(\rho^{(0)}, r) + \psi[1, r]\varepsilon + \cdots + \psi[k - r, r]\varepsilon^{k-r} + o(\varepsilon^{k-r})$, for $r = 0, \ldots, k$, where $|\psi[n, r]| < \infty$, $n = 1, \ldots, k - r$, $r = 0, \ldots, k$.

The moment generating functions $\varphi^{(\varepsilon)}(\rho, 0)$, $\psi^{(\varepsilon)}(\rho, 0)$ and $\omega^{(\varepsilon)}(\rho, 0)$ are linked by the relations

$$\omega^{(\varepsilon)}(\rho, 0) = \begin{cases} (\varphi^{(\varepsilon)}(\rho, 0) - \varphi^{(\varepsilon)}(0, 0))/(e^\rho - 1) & \rho > 0, \\ \varphi^{(\varepsilon)}(0, 1) & \rho = 0, \end{cases} \tag{43}$$

and

$$\varphi^{(\varepsilon)}(\rho, 0) = \begin{cases} (\psi^{(\varepsilon)}(\rho, 0) - 1)/(e^\rho - 1) & \rho > 0, \\ \psi^{(\varepsilon)}(0, 1) & \rho = 0. \end{cases} \tag{44}$$

Using relations (43) and (44) we can build asymptotic expansions for $\varphi^{(\varepsilon)}(\rho^{(0)}, r)$ and $\omega^{(\varepsilon)}(\rho^{(0)}, r)$ using the same techniques as in Section 4. From this one can continue and obtain asymptotic expansions for the stationary distribution $\pi^{(\varepsilon)}$ and the root $\rho^{(\varepsilon)}$ of the characteristic equation (36) as follows,

$$\begin{aligned} \pi_l^{(\varepsilon)} &= \pi^{(0)} + \pi_1\varepsilon + \cdots + \pi_l\varepsilon^l + o(\varepsilon^l), \\ \rho_r^{(\varepsilon)} &= \rho^{(0)} + a_1\varepsilon + \cdots + a_r\varepsilon^r + o(\varepsilon^r). \end{aligned} \tag{45}$$

Using (42) and (45) we obtain approximations of the ruin probability of the form

$$\widehat{\Psi}_{r,l}^{(\varepsilon)}(u) = e^{-\rho_r^{(\varepsilon)}u}\pi_l^{(\varepsilon)}. \tag{46}$$

By different choices of the parameters r and l one can control the highest order of moments of claim distributions involved in the approximation.

For any non-negative integer-valued function $u^{(\varepsilon)} \to \infty$ in such a way that $\varepsilon^r u^{(\varepsilon)} \to \lambda_r \in [0,\infty)$, as $\varepsilon \to \infty$, this approximation has asymptotic relative error zero, meaning that

$$\frac{\Psi^{(\varepsilon)}(u^{(\varepsilon)})}{\widehat{\Psi}_{r,l}^{(\varepsilon)}(u^{(\varepsilon)})} \to 1 \quad \text{as } \varepsilon \to 0.$$

In the case $\rho^{(0)} > 0$, the approximation in equation (46) generalises the Cramér–Lundberg approximation for discrete time risk processes while the case $\rho^{(0)} = 0$ corresponds to a generalisation of the diffusion approximation.

References

1. E. Altman, K. E. Avrachenkov, and R. Núñes-Queija, *Perturbation analysis for denumerable Markov chains with application to queueing models*, Adv. in Appl. Probab. **36** (2004), no. 3, 839–853.
2. J. Darroch and E. Seneta, *On quasi-stationary distributions in absorbing discrete-time finite Markov chains*, J. Appl. Probab. **2** (1965), 88–100.
3. J. Darroch and E. Seneta, *On quasi-stationary distributions in absorbing continuous-time finite Markov chains*, J. Appl. Probab. **4** (1967), 192–196.
4. E. Ekheden and D. S. Silvestrov, *Coupling and explicit rate of convergence in Cramér-Lundberg approximation for reinsurance risk processes*, Comm. Statist. Theory Methods **40** (2011), no. 19–20, 3524–3539.
5. E. Englund and D. S. Silvestrov, *Mixed large deviation and ergodic theorems for regenerative processes with discrete time*, Theory Stoch. Process. **3(19)** (1997), no. 1–2, 164–176.
6. W. Feller *An Introduction to Probability Theory and Its Applications, Vol. II*, Wiley Series in Probability and Statistics, Wiley, New York, 1966, 1971.
7. J. Grandell, *Aspects of Risk Theory*, Probability and Its Applications, Springer, New York, 1991.
8. M. Gyllenberg and D. S. Silvestrov, *Quasi-stationary distributions of a stochastic metapopulation model*, J. Math. Biol. **33** (1994), 35–70.
9. M. Gyllenberg and D. S. Silvestrov, *Quasi-Stationary Phenomena in Nonlinearly Perturbed Stochastic Systems*, De Gruyter Expositions in Mathematics, vol. 44, Walter de Gruyter, Berlin, 2008.
10. R. Hassin and M. Haviv, *Mean passage times and nearly uncoupled Markov chain*, SIAM J. Disc. Math. **5** (1992), 386–397.
11. N. V. Kartashov, *Asymptotic expansions and inequalities in stability theorems for general Markov chains with relatively bounded perturbations. Stability problems of stochastic models*, J. Soviet Math. **40** (1988), no. 4, 509–518.
12. N. V. Kartashov, *Strong Stable Markov Chains*, VSP/TViMS, Utrecht/Kiev, 1996.
13. R. Z. Khasminskii, G. Yin, and Q. Zhang, *Singularly perturbed Markov chains: quasi-stationary distribution and asymptotic expansion*, Proceedings of Dynamic Systems and Applications, vol. 2, Atlanta, GA, 1995, Dynamic, Atlanta, GA, 1996, pp. 301–308.
14. J. F. C. Kingman, *The exponential decay of Markovian transition probabilities*, Proc. London Math. Soc. **13** (1963), 337–358.
15. V. S. Koroliuk and N. Limnios, *Stochastic Systems in Merging Phase Space*, World Scientific, Singapore, 2005.
16. G. Latouche, *Perturbation analysis of a phase-type queue with weakly correlated arrivals*, Adv. Appl. Probab. **20** (1988), 896–912.

17. M. Petersson and D. S. Silvestrov, *Asymptotic expansions for perturbed discrete time renewal equations and regenerative processes*, Report 2012:12, Mathematical Statistics, Stockholm University, 2012.
18. E. Seneta and D. Vere-Jones, *On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states*, J. Appl. Probab. **3** (1966), 403–434.
19. D. S. Silvestrov, *Nonlinearly perturbed Markov chains and large deviations for lifetime functionals*, Recent Advances in Reliability Theory: Methodology, Practice and Inference (N. Limnios and M. Nikulin, eds.), Birkhäuser, Boston, 2000, pp. 135–144.
20. D. S. Silvestrov and M. Petersson, *Exponential expansions for perturbed discrete time renewal equations*, Applied Reliability Engineering and Risk Analysis: Probabilistic Models and Statistical Inference (I. Frenkel, A. Karagrigoriou, A. Kleyner, and A. Lisnianski, eds.), Wiley, Chichester, 2013, pp. 349–362.
21. D. Vere-Jones, *Geometric ergodicity in denumerable Markov chains*, Quart. J. Math. **13** (1962), 7–28.
22. G. Yin and D. T. Nguyen, *Asymptotic expansions of backward equations for two-time-scale Markov chains in continuous time*, Acta Math. Appl. Sin. Engl. Ser. **25** (2009), no. 3, 457–476.
23. G. Yin and Q. Zhang, *Discrete-time singularly perturbed Markov chains*, Stochastic Modelling and Optimization, Springer, New York, 2003, pp. 1–42.

Department of Mathematics, Stockholm University, SE-106 91 Stockholm, SWEDEN
E-mail address: mikpe@math.su.se

Received 11/11/2012

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 156–166

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 156–166

ASYMPTOTIC PROPERTIES OF CORRECTED SCORE ESTIMATOR IN AUTOREGRESSIVE MODEL WITH MEASUREMENT ERRORS

UDC 519.21

D. S. PUPASHENKO, S. V. SHKLYAR, AND A. G. KUKUSH

Abstract. The autoregressive model with errors in variables with normally distributed control sequence is considered. For the main sequence, two cases are dealt with: (a) main sequence has stationary distribution, and (b) initial distribution is arbitrary, independent of the control sequence and has finite fourth moment. Here the elements of the main sequence are not observed directly, but surrogate data that include a normally distributed additive error are observed. Errors and main sequence are assumed to be mutually independent.

We estimate unknown parameter using the Corrected Score method and in both cases prove strict consistency and asymptotic normality of the estimator. To prove asymptotic normality we apply the theory of strong mixing sequences. Finally, we compare the efficiency of the Least Squares (naive) estimator and the Corrected Score estimator in the forecasting problem and conclude that the naive estimator gives better forecast.

Анотація. Розглядається модель аторегресії з похибками у змінних і нормально розподіленою керуючою послідовністю. Для головної послідовності моделі розглянуто два випадки: а) головна послідовність має стаціонарний розподіл; б) початковий розподіл є довільним, не залежить від керуючої послідовності і має четвертий момент. Елементи головної послідовності не спостерігаються безпосередньо, натомість спостерігаються сурогатні дані, що включають нормально розподілену адитивну похибку. Похибки і головна послідовність є незалежними в сукупності.

Коефіцієнт авторегресії оцінюється методом виправленої оціночної функції. В обох випадках доведено строгу консистентність і асимптотичну нормальність оцінки. Доведення асимптотичної нормальності спирається на властивості коефіцієнта сильного перемішування. В задачі прогнозу порівнюється ефективність (наївної) оцінки найменших квадратів і виправленої оцінки і робиться висновок, що наївна оцінка забезпечує кращий прогноз.

Аннотация. Рассматривается модель аторегрессии с ошибками в переменных и нормально распределенной управляющей последовательностью. Для главной последовательности рассмотрены два случая: а) главная последовательность имеет стационарное распределение; б) начальное распределение является произвольным, не зависит от управляющей последовательности и имеет четвертый момент. Элементы главной последовательности не наблюдаются непосредственно, а вместо них наблюдаются суррогатные данные, включающие нормально распределенную аддитивную ошибку. Ошибки и главная последовательность независимы в совокупности.

Коэффициент авторегрессии оценивается методом исправленной оценочной функции. В обоих случаях доказаны строгая состоятельность и асимптотическая нормальность оценки. Доказательство асимптотической нормальности опирается на свойства коэффициента сильного перемешивания. В задаче прогноза сравнивается эффективность (наивной) оценки наименьших квадратов и исправленной оценки и делается вывод, что наивная оценка обеспечивает лучший прогноз.

1. Introduction

Introduce an autoregressive (AR) sequence

$$X_n - \mu = a(X_{n-1} - \mu) + b\varepsilon_n, \quad n \geq 1, \qquad X_0 \sim N\left(\mu, \sigma^2\right), \tag{1}$$

where

2000 *Mathematics Subject Classification.* Primary 62F12; Secondary 62F10.

Key words and phrases. Autoregressive model, measurement errors, stationary process, strong mixing sequences, least squares estimation, corrected score estimation, efficiency comparison.

- coefficients a, b and mean μ are unknown parameters, such that $|a| < 1$ and $b > 0$,
- $\{X_0, \varepsilon_n, n \geq 1\}$ are independent random variables, $\varepsilon_n \sim N(0,1)$, $n \geq 1$.

Properties and applications of such models were studied, e.g., in McQuarrie and Tsai [10].

We are interested in estimators of the parameters a and μ. In case where there is no errors in variables, estimators of these parameters can be constructed by the Least Squares (LS) method with elementary criterion function

$$q_{LS}(X_k, X_{k-1}; a, \mu) = ((X_k - \mu) - a(X_{k-1} - \mu))^2.$$

Here we consider a situation where elements of the main sequence are not observed directly, but surrogate data that include additive errors are observed. Control sequence of the model is normally distributed and main sequence is stationary distributed, or as a different case, initial distribution is arbitrary, independent of the control sequence and has finite fourth moment.

Estimation of the parameters in autoregressive model with measurement error was considered in Dedecker et al. [7]. They proposed an estimation procedure based on modified least square criterion involving a suitably chosen weight function.

Other consistent estimators exist in this model. Letting $q \to \infty$ as the sample size is increasing, Chanda [6] applies Yule–Walker ARMA(p,q) estimator for errors-in-variables AR(p) model. The estimator does not use the error variance. Moreover, the errors are allowed to be slightly autocorrelated. Under some conditions, Chanda's estimator is consistent and asymptotically normal, but it is not $\sqrt{n}$-consistent.

In present paper we apply Corrected Score (CS) method (see Carroll et al. [4, Ch. 4]). We observe $W_k = X_k + V_k$, $k \geq 0$, where $V_k \sim N(0, \sigma_V^2)$ and $\{X_0, V_k, \varepsilon_k, k \geq 0\}$ are mutually independent. Consider the elementary score function of LS estimator

$$\psi_{0,LS}(X_k, X_{k-1}; a, \mu) = \frac{1}{2}\frac{\partial}{\partial a} q_{LS}(X_k, X_{k-1}; a, \mu).$$

We construct a new score $q_{CS}(W_k, W_{k-1}; a, \mu)$ as a solution to the deconvolution equation

$$\mathsf{E}_{a_0,\mu_0}(\psi_{0,CS}(W_k, W_{k-1}; a, \mu) \mid X_k, X_{k-1}) = \psi_{0,LS}(X_k, X_{k-1}; a, \mu) \quad \text{a.s.},$$

for all $a, \mu \in \mathbb{R}$. Then the CS estimator $(\hat{a}_n, \hat{\mu}_n)^T$ is defined as a solution to equation

$$\sum_{k=1}^{n} \psi_{0,CS}(W_k, W_{k-1}; a, \mu) = 0, \qquad a, \mu \in \mathbb{R}.$$

The true parameter a satisfies $|a| < 1$, and it will be shown below that $|\hat{a}_n| < 1$, for all $n \geq n_0(w)$ a.s.

In this paper we construct the CS estimator explicitly and study its asymptotic properties as $n \to \infty$.

The paper is organized as follows. The CS is given explicitly in Section 2. The strict consistency and asymptotic normality of the estimator are presented in Section 3, and Section 4 concludes. Proofs of the main results are given in Appendix.

We use the following notations. z^T is transposed vector z, E stands for expectation of a random variable, $\xrightarrow{P1}$ and $\xrightarrow{d}$ denote the convergence a.s. and in distribution respectively, $a_n \overset{P1}{\approx} b_n$ means that $a_n - b_n \xrightarrow{P1} 0$, as $n \to \infty$.

2. Construction of corrected score estimator

Rewrite model (1) in a more convenient way.

Lemma 2.1. *For the model* (1) *it holds*

$$X_n - \mu = b\sum_{i=1}^{n} a^{n-i}\varepsilon_i + a^n(X_0 - \mu), \qquad n \geq 1. \tag{2}$$

Proof. This statement is straightforward and can be proved by induction. □

From now on we suppose that $\{W_k, k = 0, \ldots, n\}$ are observed instead of

$$\{X_k, k = 0, \ldots, n\},$$

where the additive error $V_k \sim N(0, \sigma_V^2)$ and $\{V_k, X_k, k \geq 0\}$ are mutually independent.

First, for the unknown AR coefficient a and mean μ we construct the LS estimators (LSEs). To do that we introduce the objective function:

$$Q_{LS}(W_0, \ldots, W_n; a, \mu) = \frac{1}{n}\sum_{k=1}^{n}((W_k - \mu) - a(W_{k-1} - \mu))^2,$$

and minimize it with respect to a and μ. Necessary and sufficient conditions for minimizing are:

$$\begin{cases} \frac{\partial Q_{LS}}{\partial a} = \frac{2}{n}\sum_{k=1}^{n}(a(W_{k-1} - \mu) - (W_k - \mu))(W_{k-1} - \mu) = 0, \\ \frac{\partial Q_{LS}}{\partial \mu} = \frac{2}{n}\sum_{k=1}^{n}(a(W_{k-1} - \mu) - (W_k - \mu))(1 - a) = 0. \end{cases}$$

Solving this system of equations, we get the LSE of the mean μ

$$\hat{\mu}_n = \frac{\sum_{k=1}^{n} W_k W_{k-1} \sum_{k=1}^{n} W_{k-1} - \sum_{k=1}^{n} W_{k-1}^2 \sum_{k=1}^{n} W_k}{n(\sum_{k=1}^{n} W_k W_{k-1} - \sum_{k=1}^{n} W_{k-1}^2) + (\sum_{k=1}^{n} W_{k-1})^2 - \sum_{k=1}^{n} W_{k-1} \sum_{k=1}^{n} W_k},$$

provided the denominator is nonzero, and the LSE of the parameter a is

$$\hat{a}_n^{\mathrm{LS}} = \frac{\sum_{k=1}^{n}(W_k - \hat{\mu}_n)(W_{k-1} - \hat{\mu}_n)}{\sum_{k=1}^{n}(W_{k-1} - \hat{\mu}_n)^2}. \tag{3}$$

Because the LSE $\hat{\mu}$ is too complicated to be investigated, we use the sample mean that provides a strict consistent estimator of the mean μ,

$$\hat{\mu}_n = \frac{1}{n}\sum_{k=0}^{n-1} W_k \xrightarrow{P1} \mu, \quad \text{as } n \to \infty.$$

We prove that the $\hat{\mu}_n$ is asymptotically normal using the Central Limit Theorem (CLT) (see Billingsley [1, Th 27.4]) and results of Bosq and Blanke [3, p. 47–48] in order to ensure that we deal with a geometrically strong mixing sequence.

Next we construct an estimator of the regression coefficient a by the CS method. We introduce a function $\psi_{LS}(X_0, \ldots, X_n; a, \mu)$ as

$$\psi_{LS}(X_0, \ldots, X_n; a, \mu) = \frac{1}{2}\frac{\partial Q_{LS}}{\partial a} = \frac{a}{n}\sum_{k=1}^{n}(X_{k-1} - \mu)^2 - \frac{1}{n}\sum_{k=1}^{n}(X_k - \mu)(X_{k-1} - \mu).$$

We search for a function $\psi_{CS}(W_0, \ldots, W_n; a, \mu)$ that satisfies the deconvolution equation

$$\mathsf{E}(\psi_{CS}(W_0, \ldots, W_n; a, \mu) \mid X_0, \ldots, X_n) = \psi_{LS}(X_0, \ldots, X_n; a, \mu) \quad a.s. \tag{4}$$

To do that we obtain polinomial functions $h(W_k; \mu)$ and $g(W_{k-1}, W_k; \mu)$ that solve equations

$$\mathsf{E}(h(W_k; \mu) \mid X_k) = (X_k - \mu)^2 \quad \text{a.s.}, \tag{5}$$

$$\mathsf{E}(g(W_{k-1}, W_k; \mu) \mid X_{k-1}, X_k) = (X_k - \mu)(X_{k-1} - \mu) \quad \text{a.s.} \tag{6}$$

of the following form

$$h(W_k; \mu) = (W_k - \mu)^2 - \sigma_V^2,$$
$$g(W_{k-1}, W_k; \mu) = (W_{k-1} - \mu)(W_k - \mu).$$

Hence we get a polynomial solution to (4)

$$\psi_{CS}(W_0,\dots,W_n;a,\mu)=\frac{a}{n}\sum_{k=1}^{n}\left((W_{k-1}-\mu)^2-\sigma_V^2\right)-\frac{1}{n}\sum_{k=1}^{n}(W_{k-1}-\mu)(W_k-\mu).$$

Plugging-in the sample mean $\hat{\mu}_n$ and equating $\psi_{CS}(W_0,\dots,W_n;a,\hat{\mu}_n)$ to zero we get the CS estimator of a,

$$\hat{a}_n=\frac{\sum_{k=1}^{n}(W_k-\hat{\mu}_n)(W_{k-1}-\hat{\mu}_n)}{\sum_{k=1}^{n}(W_{k-1}-\hat{\mu}_n)^2-n\sigma_V^2}. \tag{7}$$

Remark 2.1. *The denominator of* (7) *is nonzero starting from certain random number, i.e., for all $n\ge n_0(\omega)$ a.s.*

Proof of Remark 2.1 is a part of proof of Theorem 3.2, see Appendix.

3. Main results

Asymptotic properties of CS estimator. We state the consistency and asymptotic normality of the CS estimator (7) as $n\to\infty$.

Theorem 3.1. *In model* (1), *let $\{X_k, k\ge 1\}$ be a stationary process. Assume that variables $\{X_0,\varepsilon_k,V_{k-1},k\ge 1\}$ are mutually independent, then the CS estimator* (7) *is strictly consistent.*

For Theorems 3.2 and 3.4, do not assume that X_0 has a stationary distribution of underlying AR sequence. In particular, assume (1) without requirement that $X_0\sim N(\mu,\sigma^2)$.

Theorem 3.2. *Assume that $\{X_k,k\ge 1\}$ in AR* (1) *has an arbitrary initial distribution with finite fourth moment and variables $\{X_0,\varepsilon_k,V_{k-1},k\ge 1\}$ are mutually independent. Then CS estimator* (7) *is strictly consistent.*

Theorem 3.3. *In AR* (1) *let $\{X_k,k\ge 1\}$ be a stationary process. Assume that variables $\{X_0,\varepsilon_k,V_{k-1},k\ge 1\}$ are mutually independent. Then the CS estimator* (7) *is asymptotically normal with positive asymptotic variance*

$$\sigma_\infty^2=1-a^2+2\left(1-a^2\right)\frac{\sigma_V^2}{\sigma^2}+\left(2a^2+1\right)\frac{\sigma_V^4}{\sigma^4}. \tag{8}$$

Theorem 3.4. *Assume that $\{X_k,k\ge 1\}$ in AR* (1) *has arbitrary initial distribution with finite fourth moment and variables $\{X_0,\varepsilon_k,V_{k-1},k\ge 1\}$ are mutually independent. Then the CS the estimator* (7) *is asymptotically normal with positive asymptotic variance*

$$\sigma_\infty^2=1-a^2+2\left(1-a^2\right)^2\frac{\sigma_V^2}{b^2}+\left(2a^2+1\right)\left(1-a^2\right)^2\frac{\sigma_V^4}{b^4}. \tag{9}$$

Remark 3.1. *In case of known parameter μ, the CS estimator of a is defined by* (7) *setting $\hat{\mu}_n=\mu$. Then the estimator remains strictly consistent and asymptotically normal with unchanged asymptotic variance* (8).

Proofs of Theorems 3.1 to 3.4 can be found in Appendix.

Comparison of the LS and CS estimators. We compare the efficiency of the LS estimator (3) and CS estimator (7) in the forecasting problem.

As two forecasts of the forthcoming observation W_{n+1} we take the values

$$W_{n+1}^{\mathrm{LS}}=\hat{\mu}_n+\hat{a}_n^{\mathrm{LS}}(W_n-\hat{\mu}_n),\qquad W_{n+1}^{\mathrm{CS}}=\hat{\mu}_n+\hat{a}_n^{\mathrm{CS}}(W_n-\hat{\mu}_n).$$

To find an optimal forecast $\mathsf{E}(W_{n+1}|W_n)$ first we calculate the correlation coefficient between W_n and W_{n+1},

$$\rho = \frac{a\sigma^2}{\sigma^2+\sigma_V^2}.$$

Then we use a theorem from Kartashov [8] which states that for jointly Gaussian random variables $(\xi_1,\xi_2) \sim N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$, the conditional expectation can be calculated as

$$\mathsf{E}(\xi_1 \mid \xi_2 = y) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2).$$

Thus, the optimal forecast is

$$\mathsf{E}(W_{n+1} \mid W_n) = \mu + \rho(W_n-\mu) = \mu + \frac{a\sigma^2}{\sigma^2+\sigma_V^2}(W_n-\mu).$$

But the parameters of the model are unknown, and instead one can use two forecasts constructed above. Because the CS estimator is strictly consistent, i.e. $\hat{a}_n^{\mathrm{CS}} \xrightarrow{P1} a$, as $n\to\infty$, and

$$\hat{a}_n^{\mathrm{LS}} \xrightarrow{P1} a\frac{\sigma^2}{\sigma^2+\sigma_V^2} \quad \text{as } n\to\infty,$$

we have:

$$W_{n+1}^{\mathrm{CS}} - \mu = (a+o(1))(W_n-\mu) \quad \text{a.s.}$$

and for the LS forecast,

$$W_{n+1}^{\mathrm{LS}} - \mu = \hat{a}_n^{\mathrm{LS}}(W_n-\mu) = \left(a\frac{\sigma^2}{\sigma^2+\sigma_V^2} + o(1)\right)(W_n-\mu) \quad \text{a.s.,}$$

where $o(1)$ is a sequence of random variables that converges to 0 a.s.

Hence like in the example from Cheng and Van Ness [5, p. 70], we conclude that the naive LS estimator yields better forecast.

4. Conclusion

In this paper we considered the autoregressive model with measurement error. We proved the strict consistency and asymptotic normality of the CS estimator. Also we compared the efficiency of the LS (naive) estimator and CS estimator in the forecasting problem and showed that the naive estimator gives better forecast, though the naive estimator is inconsistent as $n\to\infty$.

Appendix

Proof of Theorem 3.1. We suppose that the main sequence of AR (1) has stationary distribution. Initial distribution is $X_0 \sim N(\mu,\sigma^2)$, therefore using stationarity of the process we get that $\sigma^2 = \frac{b^2}{1-a^2}$.

To show the strict consistency rewrite the estimator (7):

$$\begin{aligned}\hat{a}_n = {} & \frac{\frac{1}{n}\sum_{k=1}^n V_kV_{k-1} + \frac{1}{n}\sum_{k=1}^n (X_k-\hat{\mu}_n)(X_{k-1}-\hat{\mu}_n)}{\frac{1}{n}\sum_{k=1}^n (X_{k-1}-\hat{\mu}_n)^2 + \frac{1}{n}\sum_{k=1}^n V_{k-1}^2 + \frac{2}{n}\sum_{k=1}^n (X_{k-1}-\hat{\mu}_n)V_{k-1} - \sigma_V^2} \\ & + \frac{\frac{1}{n}\sum_{k=1}^n (X_k-\hat{\mu}_n)V_{k-1} + \frac{1}{n}\sum_{k=1}^n V_k(X_{k-1}-\hat{\mu}_n)}{\frac{1}{n}\sum_{k=1}^n (X_{k-1}-\hat{\mu}_n)^2 + \frac{1}{n}\sum_{k=1}^n V_{k-1}^2 + \frac{2}{n}\sum_{k=1}^n (X_{k-1}-\hat{\mu}_n)V_{k-1} - \sigma_V^2}.\end{aligned} \tag{10}$$

We find the limits as $n\to\infty$ for all terms in (10) separately.

First consider the sequence

$$\frac{1}{n}\sum_{k=1}^n (X_{k-1}-\hat{\mu}_n)^2.$$

Rewriting it as

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\hat{\mu}_n)^2 = \frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\mu)^2 + (\mu-\hat{\mu}_n)\frac{2}{n}\sum_{k=1}^{n}(X_{k-1}-\mu) + (\mu-\hat{\mu}_n)^2$$

and using strict consistency of sample mean $\hat{\mu}_n$, we get that the last two terms are vanishing as $n\to\infty$, hence

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\hat{\mu}_n)^2 \overset{P1}{\approx} \frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\mu)^2.$$

To get the limit of the sequence

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\mu)^2,$$

we use the ergodic theorem for stationary processes (see Korolyuk et al. [9]). Conditions of the ergodic theorem can be verified, and we get a limit of the first term in denominator of (10),

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\hat{\mu}_n)^2 \overset{P1}{\approx} \frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\mu)^2 \overset{P1}{\to} \mathsf{E}(X_0-\mu)^2=\sigma^2 \quad \text{as } n\to\infty. \tag{11}$$

To get a limit for the second term we use the strong law of large numbers (SLLN):

$$\frac{1}{n}\sum_{k=1}^{n}V_{k-1}^2 \overset{P1}{\to} \mathsf{E}\,V_0^2=\sigma_V^2 \quad \text{as } n\to\infty. \tag{12}$$

By similar technique we get limits of all terms in (10) as $n\to\infty$:

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\hat{\mu}_n)V_{k-1} \overset{P1}{\to} 0, \tag{13}$$

$$\frac{1}{n}\sum_{k=1}^{n}V_kV_{k-1} \overset{P1}{\to} 0, \tag{14}$$

$$\frac{1}{n}\sum_{k=1}^{n}(X_k-\hat{\mu}_n)(X_{k-1}-\hat{\mu}_n) \overset{P1}{\to} a\sigma^2, \tag{15}$$

$$\frac{1}{n}\sum_{k=1}^{n}(X_k-\hat{\mu}_n)V_{k-1} \overset{P1}{\to} 0, \tag{16}$$

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\hat{\mu}_n)V_k \overset{P1}{\to} 0. \tag{17}$$

Plugging limits (11)–(17) in expression (10), we get that $\hat{a}_n \overset{P1}{\to} a$ as $n\to\infty$. □

Proofs of Remark 2.1 and Theorem 3.2. We denote stationary distributed random variables satisfying (1) as $\{X_k^{\mathrm{st}}, k\ge 1\}$, with initial distribution $X_0^{\mathrm{st}}\sim N(\mu,\sigma^2)$. We assume that $\{X_0, X_0^{\mathrm{st}}, \varepsilon_k, V_{k-1}, k\ge 1\}$ are mutually independent.

Equality (2) implies that

$$X_n-\mu = (X_n^{\mathrm{st}}-\mu) + a^n\left(X_0 - X_0^{\mathrm{st}}\right), \tag{18}$$

hence $X_n - X_n^{\mathrm{st}} \xrightarrow{P1} 0$ as $n\to\infty$.

Now we have to find a limit of (10) as $n\to\infty$.

First consider

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\hat{\mu}_n)^2 \overset{P1}{\approx} \frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\mu)^2.$$

We plug expression (18) in the latter sequence, hence

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\mu)^2 &= \frac{1}{n}\sum_{k=1}^{n}\left((X_{k-1}^{\mathrm{st}}-\mu)+a^{k-1}(X_0-X_0^{\mathrm{st}})\right)^2\\ &= \frac{1}{n}\sum_{k=1}^{n}\left(X_{k-1}^{\mathrm{st}}-\mu\right)^2+\frac{2}{n}\left(X_0-X_0^{\mathrm{st}}\right)\sum_{k=1}^{n}a^{k-1}\left(X_{k-1}^{\mathrm{st}}-\mu\right)\\ &\quad+\frac{1}{n}\left(X_0-X_0^{\mathrm{st}}\right)^2\sum_{k=1}^{n}a^{2(k-1)}.\end{aligned} \tag{19}$$

In the proof of Theorem 3.1, we have shown convergence of the first term in expression (19):

$$\frac{1}{n}\sum_{k=1}^{n}\left(X_{k-1}^{\mathrm{st}}-\mu\right)^2 \xrightarrow{P1} \sigma^2 \quad \text{as } n\to\infty.$$

Since $|a|<1$, we get that

$$\frac{1}{n}\left(X_0-X_0^{\mathrm{st}}\right)^2\sum_{k=1}^{n}a^{2(k-1)} \xrightarrow{P1} 0 \quad \text{as } n\to\infty.$$

For the second term of (19) we proceed as follows. Denote corresponding random sequence as

$$Y_n=\frac{2}{n}\sum_{k=1}^{n}a^{k-1}\left(X_{k-1}^{st}-\mu\right):$$

- First using Chebyshev's inequality

$$\sum_{n=1}^{\infty}\mathsf{P}(|Y_n|>C)\le\sum_{n=1}^{\infty}\frac{\mathsf{E}\,|Y_n|^2}{C^2}<\infty$$

  we show that for each $C>0$, it holds $\sum_{n=1}^{\infty}\mathsf{P}(|Y_n|>C)<\infty$.
- Then Borel–Cantelli lemma implies that $\forall C>0\ \exists n_0\ \forall n\ge n_0\colon |Y_n|\le C$ a.s.

Therefore, to prove that $Y_n \xrightarrow{P1} 0$ as $n\to\infty$, it is enough to show $\sum_{n=1}^{\infty}\mathsf{E}\,|Y_n|^2<\infty$.

After quite cumbersome calculations, we can show that

$$\sum_{n=1}^{\infty}\mathsf{E}\left(\frac{2}{n}\sum_{k=1}^{n}a^{k-1}(X_{k-1}^{\mathrm{st}}-\mu)\right)^2$$

converges.

Hence

$$\frac{2}{n}\left(X_0-X_0^{\mathrm{st}}\right)\sum_{k=1}^{n}a^{k-1}\left(X_{k-1}^{\mathrm{st}}-\mu\right)\xrightarrow{P1}0 \quad \text{as } n\to\infty.$$

Therefore, plugging all limits found above in (19), we obtain

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\hat{\mu}_n)^2 \xrightarrow{P1} \sigma^2 \quad \text{as } n\to\infty. \tag{20}$$

Similarly we get limits of all terms in (10) as $n \to \infty$:

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\hat{\mu}_n)V_{k-1} \xrightarrow{P1} 0, \tag{21}$$

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\hat{\mu}_n)(X_k-\hat{\mu}_n) \xrightarrow{P1} a\sigma^2, \tag{22}$$

$$\frac{1}{n}\sum_{k=1}^{n}(X_k-\hat{\mu}_n)V_{k-1} \xrightarrow{P1} 0, \tag{23}$$

$$\frac{1}{n}\sum_{k=1}^{n}(X_{k-1}-\hat{\mu}_n)V_k \xrightarrow{P1} 0. \tag{24}$$

We plug (12), (14), (20)–(24) in expression (10) and get that $\hat{a}_n \xrightarrow{P1} a$, as $n \to \infty$, hence the estimator (7) is strictly consistent. A limit of the denominator in (10) is nonzero, therefore, the statement of Remark 2.1 holds true. □

Now, we state lemmas for mixing coefficients and mixing sequences.

First note that for two σ-algebras $\mathcal{G}$ and $\mathcal{H}$ on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, the strong mixing coefficient is defined as follows:

$$\alpha(\mathcal{G},\mathcal{H}) = \sup_{A\in\mathcal{G},B\in\mathcal{H}} |\,\mathsf{P}(A\cap B) - \mathsf{P}(A)\,\mathsf{P}(B)|.$$

For a random sequence $\{X_k, k \ge 0\}$, denote

$$\alpha^X(m) = \sup_{k\ge 0} \alpha(\sigma(X_0,\ldots,X_k), \sigma(X_{k+m}, X_{k+m+1},\ldots)).$$

The sequence $\{X_k, k \ge 0\}$ is called *a strong mixing process* if $\lim_{m\to\infty} \alpha^X(m) = 0$. It is called *a geometrically strong mixing* (GSM) process if

$$\alpha^X(m) \le br^m, \qquad m \ge 0,$$

for some $0 < r < 1$ and $b > 0$.

Now, we state a helpful lemma which is a direct consequence of the definition of strong mixing sequences (see Billingsley [2]).

Lemma 4.1. *Let $\{X_n, n \ge 0\}$ be a random sequence and $Z_n = (X_{n-l},\ldots,X_n)^T$, $n \ge l$. Then for α-mixing coefficients associated to sequences $\{X_n, n \ge 0\}$ and $\{Z_n, n \ge l\}$, the following relation holds true:*

$$\alpha^X(m) = \alpha^Z(m+l), \qquad m \ge 0$$

Corollary 4.1. *Let $\{X_n, n \ge 0\}$ be a random sequence and $Z_n = (X_{n-l},\ldots,X_n)^T$, $n \ge l$. For a Borel measurable vector function f, consider a sequence*

$$f(Z) = \{f(Z_n), n \ge l\}.$$

Then

$$\alpha^X(m) \ge \alpha^{f(Z)}(m+l), \qquad m \ge 0.$$

If $\{X_n, n \ge 0\}$ is a strong mixing sequence then $\{f(Z_n), n \ge l\}$ is a strong mixing sequence as well. If $\{X_n, n \ge 0\}$ is a GSM sequence then so is $\{f(Z_n), n \ge l\}$.

Lemma 4.2. *Let $(\Omega_1, \mathcal{F}_1, \mathsf{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathsf{P}_2)$ be two probability spaces. Let $\mathcal{G}_1$ and $\mathcal{H}_1$ be two sub-σ-algebras of $\mathcal{F}_1$ and let $\mathcal{G}_2$ and $\mathcal{H}_2$ be two independent sub-σ-algebras of $\mathcal{F}_2$. Then*

$$\alpha(\sigma(\mathcal{G}_1\times\mathcal{G}_2), \sigma(\mathcal{H}_1\times\mathcal{H}_2)) = \alpha_1(\mathcal{G}_1,\mathcal{H}_1),$$

Here for the calculating mixing coefficient α_1 we use measure P_1; and for α product measure $\mathsf{P} = \mathsf{P}_1 \times \mathsf{P}_2$ is used.

Proof. Denote $\mathcal{G} = \sigma(\mathcal{G}_1 \times \mathcal{G}_2)$, $\mathcal{H} = \sigma(\mathcal{H}_1 \times \mathcal{H}_2)$. Expectation in $(\Omega_2, \mathcal{F}_2, \mathsf{P}_2)$ is denoted as E_2. For $A \in \sigma(\mathcal{F}_1 \times \mathcal{F}_2)$, denote the section $A_{\omega_2} := \{\omega_1 \in \Omega_1 \mid (\omega_1, \omega_2) \in A\} \in \mathcal{F}_1$.

Let $A \in \mathcal{G}$ and $B \in \mathcal{H}$. Then $\mathsf{P}_1(A_{\omega_2})$ and $\mathsf{P}_1(B_{\omega_2})$ are independent random variables. Hence

$$\mathsf{P}(A)\,\mathsf{P}\,(B) = \mathsf{E}_2(\mathsf{P}_1(A_{\omega_2}))\ \mathsf{E}_2(\mathsf{P}_1(A_{\omega_2})) = \mathsf{E}_2(\mathsf{P}_1(B_{\omega_2})\ \mathsf{P}_1(B_{\omega_2})).$$

We have

$$|\,\mathsf{P}_1(A_{\omega_2} \cap B_{\omega_2}) - \mathsf{P}_1(A_{\omega_2})\,\mathsf{P}_1(A_{\omega_2})| \le \alpha_1(\mathcal{G}_1, \mathcal{H}_1) \quad \mathsf{P}_2\text{-a.s.},$$

$$\begin{aligned}|\,\mathsf{P}(A \cap B) - \mathsf{P}(A)\,\mathsf{P}(B)| &= |\,\mathsf{E}_2(\mathsf{P}_1((A \cap B)_{\omega_2})) - \mathsf{E}_2(\mathsf{P}_1(A_{\omega_2})\ \mathsf{P}_1(B_{\omega_2}))| \\ &= |\,\mathsf{E}_2(\mathsf{P}_1(A_{\omega_2} \cap B_{\omega_2}) - \mathsf{P}_1(A_{\omega_2})\ \mathsf{P}_1(B_{\omega_2}))| \le \alpha_1(\mathcal{G}_1, \mathcal{H}_1).\end{aligned}$$

Varying A and B, we get

$$\alpha(\mathcal{G}, \mathcal{H}) \le \alpha_1(\mathcal{G}_1, \mathcal{H}_1). \tag{25}$$

From the other hand

$$\begin{aligned}\alpha(\mathcal{G}, \mathcal{H}) &= \sup_{A \in \mathcal{G},\, B \in \mathcal{H}} |\,\mathsf{P}(A \cap B) - \mathsf{P}(A)\,\mathsf{P}(B)| \\ &\ge \sup_{A_1 \in \mathcal{G}_1,\, B_1 \in \mathcal{H}_1} |\,\mathsf{P}((A_1 \times \Omega_2) \cap (B_1 \times \Omega_2)) - \mathsf{P}(A_1 \times \Omega_2)\,\mathsf{P}(B_1 \times \Omega_2)| \\ &= \sup_{A_1 \in \mathcal{G}_1,\, B_1 \in \mathcal{H}_1} |\,\mathsf{P}_1(A_1 \cap B_1) - \mathsf{P}_1(A_1)\,\mathsf{P}_1(B_1)| = \alpha_1(\mathcal{G}_1, \mathcal{H}_1).\end{aligned} \tag{26}$$

Inequalities (25) and (26) imply the statement of Lemma. □

Under conditions of Lemma 4.2, a similar relation holds true for ϕ-mixing coefficients:

$$\phi(\sigma(\mathcal{G}_1 \times \mathcal{G}_2), \sigma(\mathcal{H}_1 \times \mathcal{H}_2)) = \phi_1(\mathcal{G}_1, \mathcal{H}_1),$$

where

$$\phi(\mathcal{G}, \mathcal{H}) = \sup_{A \in \mathcal{G},\, B \in \mathcal{H},\, \mathsf{P}(B) \neq 0} \bigl|\mathsf{P}(A \mid B) - \mathsf{P}(A)\bigr|.$$

Proof of Theorem 3.3. Now, the process $\{X_k, k \ge 0\}$ is stationary, $X_0 \sim N(\mu, \sigma^2)$ and $\sigma^2 = \frac{b^2}{1-a^2}$.

From expression (7) for estimator $\hat{\alpha}_n$ we get

$$\sqrt{n}(\hat{a}_n - a) = \frac{\frac{1}{\sqrt{n}} \sum_{k=1}^n (W_{k-1} - \hat{\mu}_n)(W_k - \hat{\mu}_n - a(W_{k-1} - \hat{\mu}_n)) + \sqrt{n} a \sigma_V^2}{\frac{1}{n} \sum_{k=1} n \sum_{k=1}^n (W_{k-1} - \hat{\mu})^2 - \sigma_V^2} =: \frac{A_n}{B_n}.$$

From the proof of Theorem 3.1 we get a limit of the denominator:

$$B_n \xrightarrow{P1} \sigma^2. \tag{27}$$

Now, rewrite the numerator. Because $\sum_{k=1}^n (W_{k-1} - \hat{\mu}_n) = 0$, we have

$$\begin{aligned}A_n &= \frac{1}{\sqrt{n}} \sum_{k=1}^n (W_{k-1} - \hat{\mu}_n)(W_k - a W_{k-1}) + \sqrt{n} a \sigma_V^2 \\ &= \frac{1}{\sqrt{n}} \sum_{k=1}^n (W_{k-1} - \mu)(W_k - \mu - a(W_{k-1} - \mu)) + \sqrt{n} a \sigma_V^2 \\ &\quad - \frac{\hat{\mu}_n - \mu}{\sqrt{n}} \sum_{k=1}^n (W_k - \mu - a(W_{k-1} - \mu)).\end{aligned}$$

By the classical CLT,

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n (W_k - \mu - a(W_{k-1} - \mu)) = \frac{1}{\sqrt{n}} \sum_{k=1}^n (V_k - a V_{k-1} + b \varepsilon_k)$$

converges in distribution. Remember that $\hat{\mu}_n$ is a consistent estimator of μ. Then by Slutsky lemma,

$$\frac{\hat{\mu}_n - \mu}{\sqrt{n}} \sum_{k=1}^{n} (W_k - \mu - a(W_{k-1} - \mu)) \xrightarrow{\mathsf{P}} 0 \quad \text{as } n \to \infty.$$

Denote

$$\begin{aligned} Z_k &= (W_{k-1} - \mu)(W_k - \mu - a(W_{k-1} - \mu)) + a\sigma_V^2 \\ &= (W_{k-1} - \mu)(V_k - aV_{k-1} + b\varepsilon_k) + a\sigma_V^2. \end{aligned}$$

With this notation, $A_n \overset{P}{\approx} \widetilde{A} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} Z_k$.

The AR process $\{X_k - \mu, k \geq 0\}$ is a GSM sequence, see Bosq, Blanke [3, Ex. 1.5, p. 47]. By Lemma 4.2 $\{(X_k - \mu, V_k)^T, k \geq 0\}$ is a GSM sequence too. Then by Corollary 4.1 $\{Z_k, k \geq 1\}$ is a GSM sequence. Also $\{Z_k, k \geq 1\}$ is a strictly stationary process with $\mathsf{E}\, Z_k = 0$ and $\mathsf{E}\, Z_k^{12} < \infty$. Applying CLT, we get

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Z_k \xrightarrow{d} N\left(0, \sigma_A^2\right)$$

with $\sigma_A^2 = \mathsf{E}\, Z_1^2 + 2 \sum_{k=2}^{\infty} \mathsf{E}\, Z_1 Z_k$. After some calculations we have

$$\begin{aligned} \mathsf{E}\, Z_1^2 &= \left(1 - a^2\right) \sigma^4 + 2\sigma^2 \sigma_V^2 + \left(2a^2 + 1\right) \sigma_V^4, \\ \mathsf{E}\, Z_1 Z_2 &= -a^2 \sigma^2 \sigma_V^2, \\ \mathsf{E}\, Z_1 Z_k &= 0, \qquad k \geq 3. \end{aligned}$$

Therefore

$$\sigma_A^2 = \left(1 - a^2\right) \sigma^4 + 2\left(1 - a^2\right) \sigma^2 \sigma_V^2 + \left(2a^2 + 1\right) \sigma_V^4.$$

Finally,

$$A_n \xrightarrow{d} N\left(0, \sigma_A^2\right), \tag{28}$$

$$\sqrt{n}(\hat{a} - a) = \frac{A_n}{B_n} \xrightarrow{d} N(0, \sigma_\infty^2)$$

with

$$\sigma_\infty^2 = \frac{\sigma_A^2}{\sigma^4} = 1 - a^2 + 2\left(1 - a^2\right) \frac{\sigma_V^2}{\sigma^2} + \left(2a^2 + 1\right) \frac{\sigma_V^4}{\sigma^4}.$$

Obviously $\sigma_\infty^2 > 0$. Thus, $\hat{\alpha}_n$ is an asymptotically normal estimator of a. □

Proof of Theorem 3.4. Proof of this theorem differs from the proof of Theorem 3.3 only when we deal with numerator $\widetilde{A}_n$. For the case of stationary initial distribution we denote it as $\widetilde{A}_n^{\mathrm{st}}$. Then using relation (18) we rewrite $\widetilde{A}_n$ for an arbitrary distribution as follows:

$$\begin{aligned} \widetilde{A}_n &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} V_k V_{k-1} + \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - \mu)(X_{k-1} - \mu) + \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - \mu) V_{k-1} \\ &\quad + \frac{1}{\sqrt{n}} \sum_{k=1}^{n} V_k (X_{k-1} - \mu) - \frac{a}{\sqrt{n}} \sum_{k=1}^{n} (X_{k-1} - \mu)^2 - \frac{a}{\sqrt{n}} \sum_{k=1}^{n} V_{k-1}^2 \\ &\quad - \frac{2a}{\sqrt{n}} \sum_{k=1}^{n} (X_{k-1} - \mu) V_{k-1} + a\sqrt{n}\sigma_V^2 \\ &= \widetilde{A}_n^{\mathrm{st}} + \frac{1}{\sqrt{n}} (X_0 - X_0^{\mathrm{st}}) \sum_{k=1}^{n} a^{k-1} (V_k - aV_{k-1} + b\varepsilon_k). \end{aligned}$$

Because $\widetilde{A}_n^{\mathrm{st}}$ converges in distribution, it remains to prove only that the last term

$$\frac{1}{\sqrt{n}}(X_0 - X_0^{\mathrm{st}}) \sum_{k=1}^{n} a^{k-1}(V_k - aV_{k-1} + b\varepsilon_k)$$

converges to 0 in probability. We have

$$\begin{aligned}\mathsf{E}\left|\frac{1}{\sqrt{n}} \sum_{k=1}^{n} a^{k-1}(V_k - aV_{k-1} + b\varepsilon_k)\right| &\le \frac{1}{\sqrt{n}}\,\mathsf{E} \sum_{k=1}^{n} |a^{k-1}|(|V_k| + |aV_{k-1}| + |b\varepsilon_k|) \\ &\le \frac{1}{\sqrt{n}} \sum_{k=1}^{n} |a^{k-1}|(\mathsf{E}\,|V_k| + \mathsf{E}\,|aV_{k-1}| + \mathsf{E}\,|b\varepsilon_k|).\end{aligned}$$

Because the sum $(\mathsf{E}\,|V_k| + \mathsf{E}\,|aV_{k-1}| + \mathsf{E}\,|b\varepsilon_k|)$ can be bounded by some constant c and $|a| < 1$, we have:

$$\mathsf{E}\left|\frac{1}{\sqrt{n}} \sum_{k=1}^{n} a^{k-1}(V_k - aV_{k-1} + b\varepsilon_k)\right| \le \frac{c}{\sqrt{n}} \frac{1 - |a|^n}{1 - |a|} \to 0 \quad \text{as } n \to \infty.$$

Hence we obtain that $A_n \overset{\mathsf{P}}{\approx} \widetilde{A}_n^{\mathrm{st}}$ and from (27), (28) with Slutsky lemma for all $|a| < 1$ we get:

$$\sqrt{n}(\hat{a}_n - a) = \frac{A_n}{B_n} \xrightarrow{d} \zeta \sim N\left(0, \sigma_\infty^2\right) \quad \text{as } n \to \infty. \qquad \square$$

References

1. P. Billingsley, *Probability and Measure*, John Wiley & Sons Inc., 1995.
2. P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons Inc., 1999.
3. D. Bosq and D. Blanke, *Inference and Prediction in Large Dimensions*, John Wiley & Sons Inc., 2007.
4. R. J. Carroll, D. Ruppert, and L. A. Stefanski, *Measurement Error in Nonlinear Models*, Chapman & Hall., London, 1995.
5. C. L. Cheng and J. Van Ness, *Statistical Regression with Measurement Error*, Chapman & Hall, London, 1999.
6. K. C. Chanda, *Asymptotic properties of estimators for autoregressive models with errors in variables*, The Annals of Statistics **24** (1996), 423–430.
7. J. Dedecker, A. Samson, and M.-L. Taupin, *Estimation in Autoregressive Model with Measurement Error*, Chapman & Hall, London, 2011.
8. M. V. Kartashov, *Probability, Processes, Statistics*, Kyiv University, Kyiv, 2008.
9. V. S. Korolyuk, N. I. Portenko, A. V. Skorokhod, and A. F. Turbin, *A Manual on Probability Theory and Mathematical Statistics*, “Nauka”, Moscow, 1985.
10. A. D. R. McQuarrie and C.-L. Tsai, *Regression and Time Series Model Selection*, Singapore, World Scientific, 1998.

Taras Shevchenko National University of Kyiv, Volodymyrska st. 64, 01601, Kyiv, Ukraine
E-mail address: daria.pupashenko@gmail.com

Taras Shevchenko National University of Kyiv, Volodymyrska st. 64, 01601, Kyiv, Ukraine
E-mail address: shklyar@mail.univ.kiev.ua

Taras Shevchenko National University of Kyiv, Volodymyrska st. 64, 01601, Kyiv, Ukraine
E-mail address: alexander_kukush@univ.kiev.ua

Received 01/04/2013

Теорія ймовір. та матем. статист.
Вип. 89, 2013, стор. 167–180

Teor. Ĭmovir. ta Matem. Statyst.
No. 89, 2013, pp. 167–180

MIXED STOCHASTIC DELAY DIFFERENTIAL EQUATIONS

UDC 519.21

G. SHEVCHENKO

Abstract. We consider a stochastic delay differential equation driven by a Hölder continuous process Z and a Wiener process. Under fairly general assumptions on coefficients of the equation, we prove that it has a unique solution. We also give a sufficient condition for finiteness of moments of the solution and prove that the solution depends on the driver Z continuously.

Анотація. Розглядається стохастичне диференціальне рівняння із затримкою, кероване процесом Z, що задовольняє умову Гельдера, та вінерівським процесом. За достатньо загальних припущень на коефіцієнти доведено, що рівняння має єдиний розв'язок. Також наведено достатню умову для скінченності моментів розв'язку та показано, що розв'язок неперервно залежить від процесу Z.

Аннотация. Рассматривается стохастическое дифференциальное уравнение, движимое процессом Z, удовлетворяющим условию Гельдера, и винеровским процессом. При достаточно общих предположениях на коэффициенты доказано, что уравнение имеет единственное решение. Также приведено достаточное условие конечности моментов решения и показано, что решение непрерывно зависит от процесса Z.

1. Introduction

This paper is devoted to a stochastic differential equation of the form

$$X(t) = X(0) + \int_0^t a(s, X)\,ds + \int_0^t b(s, X)\,dW(s) + \int_0^t c(s, X)\,dZ(s),$$

where W is a Wiener process, Z is a Hölder continuous process with Hölder exponent greater than $1/2$, the coefficients a, b, c depend on the past of the process X. The integral with respect to W is understood in the usual Itô sense, while the one with respect to Z is understood in the pathwise sense. (A precise definition of all objects is given in Section 2.) We will call this equation a *mixed stochastic delay differential equation*; the word *mixed* refers to the mixed nature of noise, while the word *delay* is due to dependence of the coefficients on the past.

In the pure Wiener case, where $c = 0$, this equation was considered by many authors, often by the name "stochastic functional differential equation". For overview of their results we refer a reader to [9, 12], where also the importance of such equations is explained, and several particular results arising in applications are given.

In the pure "fractional" case, where $b = 0$, there are only few results devoted to such equations, considering usually the case where $Z = B^H$ is a fractional Brownian motion (for us, it is also the most important example of the driver Z). In [4, 5], the existence of a solution is shown for the coefficients of the form $a(t, X) = a(X(t))$, $b(t, X) = b(X(t - r))$, and $H > 1/2$. It is also proved that the solution has a smooth density, and the convergence of solutions is established for a vanishing delay. A similar equation

2010 *Mathematics Subject Classification.* 60H10, 34K50, 60G22.

Key words and phrases. Fractional Brownian motion; Wiener process; stochastic delay differential equation; mixed stochastic differential equation.

The research was partially supported by President's grant for young scientists, project GP/F49/002.

constrained to stay non-negative is considered in [1]. Existence and uniqueness of solution for an equation with general coefficients, also in the case $H > 1/2$, are established in [2, 8]. For such equation, it is proved in [8] that the solution possesses infinitely differentiable density, and in [3], that the solution generates a continuous random dynamical system. In [13], the unique solvability is established for an equation with $H > 1/3$ and coefficients of the form $f(X(t), X(t-r_1), X(t-r_2), \dots)$.

Concerning mixed stochastic delay differential equations, there are no results known to author. There are some literature devoted to mixed equations without delay. The existence and uniqueness were proved, under different conditions, in [6, 7, 10, 11, 16]. Integrability and convergence results for mixed equations were established in [11, 15, 16, 17], and Malliavin regularity was proved in [17].

In this paper we show that a mixed stochastic delay differential equation has a unique solution under rather general assumptions about coefficients. We also provide a condition for the solution to have finite moments of all orders, and a result on the continuity of the solution with respect to the driver Z. The latter result allows, in particular, to approximate the solution to a mixed stochastic delay differential equation by solutions to usual stochastic delay differential equations having a random drift.

2. Preliminaries

Let $\big(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t, t \ge 0\}, \mathsf{P}\big)$ be a complete filtered probability space satisfying the usual assumptions.

First we fix some notation: throughout the article, $|\cdot|$ will denote the absolute value of a real number, the Euclidean norm of a vector, or the operator norm of a matrix. The symbol C will denote a generic constant, whose value may change from one line to another. To emphasize its dependence on some parameters, we will put them into subscripts.

We need some notation in order to introduce the main object. For a fixed $r > 0$, let $\mathcal{C} = C([-r, 0]; \mathbb{R}^d)$ be the Banach space of continuous $\mathbb{R}^d$-valued functions defined on the interval $[-r, 0]$ endowed with the supremum norm $\|\cdot\|_{\mathcal{C}}$. For a stochastic process $\xi = \{\xi(t), t \in [-r, T]\}$ and $t \in [0, T]$ define a *segment* $\xi_t \in \mathcal{C}$ by $\xi_t(s) = \xi(t + s)$, $s \in [-r, 0]$. Let $a\colon [0, T] \times \mathcal{C} \to \mathbb{R}^d$, $b_i\colon [0, T] \times \mathcal{C} \to \mathbb{R}^d$, $i = 1, \dots, m$, $c_j\colon [0, T] \times \mathcal{C} \to \mathbb{R}^d$, $j = 1, \dots, l$, be measurable functions, $Z = \{Z(t), t \in [0, T]\}$ be an $\mathbb{F}$-adapted process in $\mathbb{R}^l$ such that its trajectories are almost surely Hölder continuous of order $\gamma > 1/2$. Let also $\eta\colon [-r, 0] \to \mathbb{R}^d$ be a θ-Hölder continuous function with $\theta > 1 - \gamma$.

Our main object is the following stochastic delay differential equation in $\mathbb{R}^d$:

$$\begin{aligned} X(t) = X(0) + \int_0^t a(s, X_s)\, ds + \sum_{i=1}^m \int_0^t b_i(s, X_s)\, dW_i(s) \\ + \sum_{j=1}^l \int_0^t c_j(s, X_s)\, dZ_j(s), \qquad t \in [0, T], \end{aligned} \tag{2.1}$$

with the "initial condition" $X(s) = \eta(s)$, $s \in [-r, 0]$. In the rest of the paper a shorter notation will be used for equation (2.1) and its ingredients:

$$X(t) = X(0) + \int_0^t a(s, X_s)\, ds + \int_0^t b(s, X_s)\, dW(s) + \int_0^t c(s, X_s)\, dZ(s). \tag{2.2}$$

We remark that it is possible to consider an equation with coefficients depending on the whole past of the process X. This can be achieved by just taking $r = T$

The integral with respect to W in (2.2) will be understood in the Itô sense. The integral with respect to Z is a generalized Lebesgue–Stieltjes integral, defined as follows [18]. For

$\alpha \in (0,1)$, define the fractional derivatives

$$\big(D_{a+}^{\alpha} f\big)(x) = \frac{1}{\Gamma(1-\alpha)}\left(\frac{f(x)}{(x-a)^{\alpha}} + \alpha \int_a^x \frac{f(x)-f(u)}{(x-u)^{1+\alpha}}\,du\right),$$
$$\big(D_{b-}^{1-\alpha} g\big)(x) = \frac{e^{-i\pi\alpha}}{\Gamma(\alpha)}\left(\frac{g(x)}{(b-x)^{1-\alpha}} + (1-\alpha) \int_x^b \frac{g(x)-g(u)}{(u-x)^{2-\alpha}}\,du\right).$$

Assuming that $D_{a+}^{\alpha} f \in L_1[a,b]$, $D_{b-}^{1-\alpha} g_{b-} \in L_\infty[a,b]$, where $g_{b-}(x) = g(x) - g(b)$, the generalized (fractional) Lebesgue–Stieltjes integral $\int_a^b f(x)\,dg(x)$ is defined as

$$\int_a^b f(x)\,dg(x) = e^{i\pi\alpha} \int_a^b \big(D_{a+}^{\alpha} f\big)(x)\big(D_{b-}^{1-\alpha} g_{b-}\big)(x)\,dx.$$

Moreover, we have the estimate

$$\left|\int_a^b f(x)\,dg(x)\right| \le C\,\|g\|_{0,\alpha;[a,b]} \int_a^b \left(\frac{|f(s)|}{(s-a)^{\alpha}} + \int_a^s \frac{|f(s)-f(u)|}{(s-u)^{1+\alpha}}\,du\right) ds, \tag{2.3}$$

where

$$\|g\|_{\alpha;[a,b]} = \sup_{a\le u<v\le b}\left(\frac{|g(v)-g(u)|}{(v-u)^{1-\alpha}} + \int_u^v \frac{|g(u)-g(z)|}{(z-u)^{2-\alpha}}\,dz\right).$$

In what follows we fix some $\alpha \in (1-\gamma, \theta\wedge 1/2)$ and put $h(t,s) = (t-s)^{-1-\alpha}$. Define $\|X\|_{\infty,t} = \sup_{s\in[-r,t]} |X(s)|$, $\|X\|_{1,t} = \int_0^t \|X_{\cdot+t-s} - X_{\cdot}\|_{\infty,s}\, h(t,s)\,ds$, $\|X\|_t = \|X\|_{\infty,t} + \|X\|_{1,t}$. It is clear that both $\|X\|_{\infty,t}$ and $\|X\|_{1,t}$ are non-decreasing in t.

By a solution to equation (2.2), we will understand a pathwise continuous $\mathbb{F}$-adapted process X such that $\|X\|_T < \infty$ a.s., and (2.2) holds almost surely for all $t \in [0,T]$.

The following assumptions on the coefficients of (2.2) will be assumed throughout the article:

H1. Linear growth: for all $\psi \in \mathcal{C}$, $t \in [0,T]$,
$$|a(t,\psi)| + |b(t,\psi)| + |c(t,\psi)| \le C(1 + \|\psi\|_{\mathcal{C}}).$$

H2. For all $t \in [0,T]$, $\psi \in \mathcal{C}$, c has a Fréchet derivative $\partial_\psi c(t,\psi) \in L(\mathcal{C},\mathbb{R}^d)$, bounded uniformly in $t \in [0,T]$, $\psi \in \mathcal{C}$:
$$\|\partial_\psi c(t,\psi)\|_{L(\mathcal{C},\mathbb{R}^d)} \le C.$$

H3. The functions a, b and $\partial_\psi c$ are locally Lipschitz continuous in ψ: for any $R > 1$, $t \in [0,T]$, and all $\psi_1,\psi_2 \in \mathcal{C}$ with $\|\psi_1\|_{\mathcal{C}} \le R$, $\|\psi_2\|_{\mathcal{C}} \le R$,
$$\begin{aligned}&|a(t,\psi_1) - a(t,\psi_2)| + |b(t,\psi_1) - b(t,\psi_2)| + \|\partial_\psi c(t,\psi_1) - \partial_\psi c(t,\psi_2)\|_{L(\mathcal{C},\mathbb{R}^d)}\\ &\quad \le C_R\,\|\psi_1 - \psi_2\|_{\mathcal{C}}.\end{aligned}$$

H4. The functions c and $\partial_\psi c$ are Hölder continuous in t: for some $\beta \in (1-\gamma,1)$ and for all $s,t \in [0,T]$, $\psi \in \mathcal{C}$
$$|c(s,\psi) - c(t,\psi)| \le C|s-t|^{\beta}(1 + \|\psi\|_{\mathcal{C}}), \quad \|\partial_\psi c(s,\psi) - \partial_\psi c(t,\psi)\|_{L(\mathcal{C},\mathbb{R}^d)} \le C|s-t|^{\beta}.$$

The condition H4 allows, for instance, to consider an important particular case, namely, a linear equation.

3. Auxiliary results

First we establish some a priori estimates for the solution of (2.2).

Lemma 3.1. *Let X be a solution of* (2.2), *and $p \ge 1$, $N \ge 1$. Let also $A_{N,t} = \big\{\|Z\|_{\alpha;[0,t]} \le N\big\}$ for $t \in [0,T]$. Then*
$$\mathsf{E}\left[\|X\|_T^p\, \mathbb{1}_{A_{N,T}}\right] \le C_{N,p}.$$

Proof. Assume without loss of generality that $p > 4/(1-2\alpha)$.

For $R > 0$ define $B_{R,t} = \left\{\|X\|_{\infty,t} + \|X\|_{1,t} \le R\right\}$ and $\mathbb{1}_t = \mathbb{1}_{A_{N,t}\cap B_{R,t}}$.

Let $\omega \in A_{N,t}$. Write for $t \in [0,T]$

$$|X(t)| \le |X(0)| + |I^a(t)| + \left|I^b(t)\right| + |I^c(t)|,$$

where $I^a(t) = \int_0^t a(s,X_s)\,ds$, $I^b(t) = \int_0^t b(s,X_s)\,dW(s)$, $I^c(t) = \int_0^t c(s,X_s)\,dZ(s)$. Estimate, using (2.3),

$$|I^a(t)| \le \int_0^t |a(s,X_s)|\,ds \le C\int_0^t (1+\|X_s\|_{\mathcal{C}})\,ds \le C\left(1+\int_0^t \|X\|_{\infty,s}\,ds\right);$$

$$\begin{aligned}
|I^c(t)| &\le CN\int_0^t \left(|c(s,X_s)|\,s^{-\alpha} + \int_0^s |c(s,X_s)-c(u,X_u)|\,h(s,u)\,du\right) ds\\
&\le CN\int_0^t \Bigg((1+\|X_s\|_{\mathcal{C}})\,s^{-\alpha}\\
&\qquad + \int_0^s \left(|s-u|^{\beta}(1+\|X_s\|_{\mathcal{C}}) + \|X_s-X_u\|_{\mathcal{C}}\right) h(s,u)\,du\Bigg) ds\\
&\le CN\left(1+\int_0^t \left(\|X\|_{\infty,s}\,s^{-\alpha} + \|X\|_{1,s}\right) ds\right).
\end{aligned}$$

Therefore, we have

$$|X(t)| \le CN\left(1+\int_0^t \left(\|X\|_{\infty,s}\,s^{-\alpha} + \|X\|_{1,s}\right) ds\right) + \left|I^b(t)\right|,$$

whence

$$\|X\|_{\infty,t} \le CN\left(1+\int_0^t \left(\|X\|_{\infty,s}\,s^{-\alpha} + \|X\|_{1,s}\right) ds\right) + \left\|I^b\right\|_{\infty;[0,t]}. \tag{3.1}$$

Further, let $0 \le s \le t$. Then for $u \le s-t$,

$$|X(u+t-s) - X(u)| = |\eta(u+t-s) - \eta(u)| \le H_\eta (t-s)^{\theta},$$

where $H_\eta = \sup_{-r\le x<y\le 0} \frac{|\eta(y)-\eta(x)|}{(y-x)^{\theta}}$ is the θ-Hölder seminorm of η. Similarly, for $u \in (s-t,0]$,

$$\begin{aligned}
|X(u+t-s) - X(u)| &\le |X(u+t-s) - X(0)| + |\eta(0)-\eta(u)|\\
&\le |X(u+t-s) - X(0)| + H_\eta (t-s)^{\theta}.
\end{aligned}$$

Consequently, we can write

$$\|X\|_{1,t} \le H_\eta \int_0^t (t-s)^{\theta+\alpha-1} ds + J^a(t) + J^b(t) + J^c(t) \le C + J^a(t) + J^b(t) + J^c(t),$$

where $J^b(t) = \int_0^t \sup_{u\in[s-t,s]} \left|\int_{u\vee 0}^{u+t-s} b(v,X_v)\,dW(v)\right| h(t,s)\,ds$,

$$\begin{aligned}
J^a(t) &= \int_0^t \sup_{u\in[s-t,s]} \left|\int_{u\vee 0}^{u+t-s} a(v,X_v)\,dv\right| h(t,s)\,ds\\
&\le C\int_0^t \max_{u\in[s-t,s]} \int_{u\vee 0}^{u+t-s} \left(1+\|X_v\|_{\mathcal{C}}\right) dv\, h(t,s)\,ds\\
&\le C\left(1+\int_0^t\int_s^t \|X\|_{\infty,z}\,dz\,h(t,s)\,ds\right) \le C\left(1+\int_0^t \|X\|_{\infty,z}\,(t-z)^{-\alpha}\,dz\right);
\end{aligned}$$

$$J^c(t) = \int_0^t \sup_{u\in[s-t,s]} \left|\int_{u\vee 0}^{u+t-s} c(v,X_v)\,dZ(v)\right| h(t,s)\,ds \le CN\left(J_1^c(t) + J_2^c(t)\right)$$

with

$$\begin{aligned}
J_1^c(t) &= \int_0^t \max_{u\in[s-t,s]} \int_{u\vee 0}^{u+t-s} |c(v,X_v)|\,(v-u\vee 0)^{-\alpha}\,dv\,h(t,s)\,ds\\
&\le C\int_0^t \max_{u\in[-r,s]} \int_{u\vee 0}^{u+t-s} \big(1+\|X_v\|_{\mathcal{C}}\big)(v-u\vee 0)^{-\alpha}\,dv\,h(t,s)\,ds\\
&\le C\left(1+\int_0^t\int_s^t \|X\|_{\infty,z}\,(z-s)^{-\alpha}\,dz\,h(t,s)\,ds\right)\\
&\le C\left(1+\int_0^t \|X\|_{\infty,z}\,(t-z)^{-2\alpha}\,dz\right);
\end{aligned}$$

$$\begin{aligned}
J_2^c(t) &= \int_0^t \max_{u\in[-r,s]} \int_{u\vee 0}^{u+t-s}\int_{u\vee 0}^{v} |c(v,X_v)-c(z,X_z)|\,h(v,z)\,dz\,dv\,h(t,s)\,ds\\
&\le C\int_0^t \max_{u\in[-r,s]} \int_{u\vee 0}^{u+t-s}\int_{u\vee 0}^{v} \Big(|v-z|^{\beta}+\|X_v-X_z\|_{\mathcal{C}}\Big)\,h(v,z)\,dz\,dv\,h(t,s)\,ds\\
&\le C\int_0^t \max_{u\in[-r,s]} \int_{u\vee 0}^{u+t-s} \Big(|v-u\vee 0|^{\beta-\alpha}+\|X\|_{1,v}\Big)\,dv\,h(t,s)\,ds\\
&\le C\int_0^t \left((t-s)^{\beta-2\alpha}+\int_s^t \|X\|_{1,v}\,dv\,h(t,s)\right)ds\\
&\le C\left(1+\int_0^t \|X\|_{1,v}\,(t-v)^{-\alpha}\,dv\right).
\end{aligned}$$

To estimate J_1^c, we have used the computation

$$\begin{aligned}
\int_0^z (z-s)^{-\alpha}(t-s)^{-1-\alpha}\,ds &= \Big|s=z-(t-z)v\Big| = (t-z)^{-2\alpha}\int_0^{\frac{z}{t-z}} v^{-\alpha}(1+v)^{-1-\alpha}\,dv\\
&\le (t-z)^{-2\alpha}\int_0^\infty v^{-\alpha}(1+v)^{-1-\alpha}\,dv = \mathrm{B}(1-\alpha,2\alpha)(t-z)^{-2\alpha}.
\end{aligned}$$

Summing the estimates for $\|X\|_{1,t}$, we get

$$\|X\|_{1,t} \le CN\left(1+\int_0^t \Big(\|X\|_{\infty,s}\,(t-s)^{-2\alpha}+\|X\|_{1,s}\,(t-s)^{-\alpha}\Big)\,ds\right)+J^b(t). \tag{3.2}$$

Combining this with (3.1), we obtain

$$\|X\|_t \le CN\int_0^t \|X\|_s\,g(t,s)\,ds + \big\|I^b\big\|_{\infty;[0,t]} + J^b(t)$$

for $\omega\in A_{N,t}$, where $g(t,s)=s^{-\alpha}+(t-s)^{-2\alpha}$.

Using the Hölder inequality, we can estimate

$$\|X\|_t^p \le C_pN^p\int_0^t \|X\|_s^p\,g(t,s)\,ds\left(\int_0^t g(t,s)\,ds\right)^{p/q} + C_p\left(\big\|I^b\big\|_{\infty;[0,t]}^p + \big(J^b(t)\big)^p\right),$$

whence

$$\mathsf{E}\left[\|X\|_t^p\,\mathbb{1}_t\right] \le C_{N,p}\left(\int_0^t \mathsf{E}\left[\|X\|_s^p\,\mathbb{1}_s\right]g(t,s)\,ds + \mathsf{E}\left[\big\|I^b\big\|_{\infty;[0,t]}^p\,\mathbb{1}_t\right] + \mathsf{E}\left[\big(J^b(t)\big)^p\,\mathbb{1}_t\right]\right). \tag{3.3}$$

We now proceed to the estimation of the last two expressions. It is obvious that for any $0\le u\le s\le t$,

$$\left|\int_u^s b(v,X_v)\,dW(v)\right|\mathbb{1}_t \le \left|\int_u^s b(v,X_v)\mathbb{1}_v\,dW(v)\right|.$$

Therefore, by the Burkholder inequality,

$$
\begin{aligned}
\mathsf{E}\left[\left\|I^b\right\|^p_{\infty;[0,t]} \mathbb{1}_t\right] &= \mathsf{E}\left[\sup_{s\in[0,t]}\left|\int_0^s b(v,X_v)\,dW(v)\right|^p \mathbb{1}_t\right] \\
&\le \mathsf{E}\left[\sup_{s\in[0,t]}\left|\int_0^s b(v,X_v)\mathbb{1}_v\,dW(v)\right|^p\right] \\
&\le C_p\mathsf{E}\left[\left(\int_0^t |b(s,X_s)|^2\,\mathbb{1}_s\,ds\right)^{p/2}\right] \le C_p\int_0^t \mathsf{E}\left[(1+\|X_s\|_{\mathcal{C}}\,\mathbb{1}_s)^p\right]ds \\
&\le C_p\int_0^t \left(1+\mathsf{E}\left[\|X_s\|^p_{\mathcal{C}}\,\mathbb{1}_s\right]\right)ds \le C_p\left(1+\int_0^t \mathsf{E}\left[\|X\|^p_{\infty,s}\,\mathbb{1}_s\right]ds\right).
\end{aligned}
$$

Further, we have

$$
\mathsf{E}\left[\left(J^b(t)\right)^p\mathbb{1}_t\right] \le C_p\mathsf{E}\left[\left(\int_0^t \sup_{u\in[s-t,s]}\left|\int_{u\vee 0}^{u+t-s} b(v,X_v)\mathbb{1}_v\,dW(v)\right| h(t,s)\,ds\right)^p\right]. \quad (3.4)
$$

It follows from the Garsia–Rodemich–Rumsey inequality that for any $r,z\in[0,t]$

$$
\left|\int_r^z b(v,X_v)\mathbb{1}_v\,dW(v)\right| \le C_p\xi(t)\,|r-z|^{1/2-2/p},
$$

where

$$
\xi(t) = \left(\int_0^t\int_0^y \frac{\left|\int_x^y b(v,X_v)\mathbb{1}_v\,dW(v)\right|^p}{|x-y|^{p/2}}\,dx\,dy\right)^{1/p}.
$$

We can estimate

$$
\begin{aligned}
\mathsf{E}\left[\xi(t)^p\right] &= \int_0^t\int_0^y \frac{\mathsf{E}\left[\left|\int_x^y b(v,X_v)\mathbb{1}_v\,dW(v)\right|^p\right]}{|x-y|^{p/2}}\,dx\,dy \\
&\le C_p\int_0^t\int_0^y \frac{\mathsf{E}\left[\left(\int_x^y (1+\|X_v\|^2_{\mathcal{C}})\mathbb{1}_v\,dv\right)^{p/2}\right]}{(y-x)^{p/2}}\,dx\,dy \\
&\le C_p\int_0^t\int_0^y \frac{(y-x)^{p/2-1}\mathsf{E}\left[\int_x^y \left(1+\|X\|^p_{\infty,v}\,\mathbb{1}_v\right)dv\right]}{(y-x)^{p/2}}\,dx\,dy \\
&\le C_p\left(1+\int_0^t\int_0^y \mathsf{E}\left[\|X\|^p_{\infty,v}\,\mathbb{1}_v\right]\int_0^v (y-x)^{-1}\,dx\,dv\,dy\right) \\
&= C_p\left(1+\int_0^t \mathsf{E}\left[\|X\|^p_{\infty,v}\,\mathbb{1}_v\right]\int_v^t \log\frac{y}{y-v}\,dy\,dv\right) \\
&\le C_p\left(1+\int_0^t \mathsf{E}\left[\|X\|^p_{\infty,v}\,\mathbb{1}_v\right]dv\right).
\end{aligned}
$$

Therefore, taking into account that $p>4/(1-2\alpha)$, i.e. $2/p+1/\alpha-1/2<0$, we get from (3.4)

$$
\begin{aligned}
\mathsf{E}\left[J^b(t)^p\mathbb{1}_t\right] &\le C_p\mathsf{E}\left[\xi(t)^p\right]\left(\int_0^t (t-s)^{-2/p-1/2-\alpha}\,ds\right)^p \\
&\le C_p\left(1+\int_0^t \mathsf{E}\left[\|X\|^p_{\infty,v}\,\mathbb{1}_v\right]dv\right).
\end{aligned}
$$

Plugging the estimates of I^b and J^b into (3.3), we get

$$
\mathsf{E}\left[\|X\|^p_t\,\mathbb{1}_t\right] \le C_{N,p}\left(1+\int_0^t \mathsf{E}\left[\|X\|^p_s\,\mathbb{1}_s\right]g(t,s)\,ds\right).
$$

Since $g(t,s) \le (T^\alpha+1)t^{2\alpha}s^{-2\alpha}(t-s)^{-2\alpha}$, we can apply the generalized Gronwall lemma [14, Lemma 7.6] and obtain $\mathsf{E}\left[\|X\|_T^p \mathbb{1}_T\right] \le C_{N,p}$. By letting $R\to\infty$ and using the Fatou lemma, we arrive at the required statement. □

The following lemma establishes estimates for the distance between solutions of mixed stochastic delay differential equations with different drivers. To formulate it, assume that $\overline{Z}$ is another γ-Hölder $\mathbb{F}$-adapted process, and consider the equation

$$\overline{X}(t) = X(0) + \int_0^t a(s,\overline{X}_s)\,ds + \int_0^t b(s,\overline{X}_s)\,dW(s) + \int_0^t c(s,\overline{X}_s)\,d\overline{Z}(s) \tag{3.5}$$

with the same initial condition $\overline{X}(s)=\eta(s)$, $s\in[-r,0]$.

Lemma 3.2. *Let X and $\overline{X}$ be the solutions of* (2.2) *and* (3.5) *respectively, $p\ge 4/(1-2\alpha)$, $N\ge 1$, $R\ge 1$. Assume also that $\|Z\|_{\alpha;[0,T]}\le N$ and $\left\|\overline{Z}\right\|_{\alpha;[0,T]}\le N$. Then*

$$\mathsf{E}\left[\left\|X-\overline{X}\right\|_{\infty,T}^p \mathbb{1}_{B_{R,T}}\right] \le C_{N,R,p}\mathsf{E}\left[\left\|Z-\overline{Z}\right\|_{\alpha;[0,T]}^p\right],$$

where $B_{R,t}=\left\{\|X\|_t\le R, \left\|\overline{X}\right\|_t\le R\right\}$ for $t\in[0,T]$.

Proof. The proof will be similar to that of Lemma 3.1, so we will omit some details. Put $\Delta(t)=\left\|X-\overline{X}\right\|_t$, $\Delta_d(t)=d(s,X_s)-d(s,\overline{X}_s)$ for $d\in\{a,b,c\}$, and $\Delta_Z(t)=Z(t)-\overline{Z}(t)$. By assumption H3, $\Delta_d(t)\le C_R\left\|X_t-\overline{X}_t\right\|_{\mathcal{C}}\le C_R\Delta(t)$.

Let $\omega\in B_{R,t}$. Write for $t\in[0,T]$

$$\left|X(t)-\overline{X}(t)\right| \le |I^a(t)| + \left|I^b(t)\right| + |I^c(t)| + \left|I^Z(t)\right|,$$

where $I^a(t)=\int_0^t \Delta_a(s)\,ds$, $I^b(t)=\int_0^t\Delta_b(s)\,dW(s)$, $I^c(t)=\int_0^t\Delta_c(s)\,dZ(s)$, $I^Z(t)=\int_0^t c(s,\overline{X}_s)\,d\Delta_Z(t)$. We estimate the terms one by one, starting with I^a:

$$|I^a(t)| \le \int_0^t |\Delta_a(s)|\;ds \le C_R\int_0^t \Delta(s)\,ds.$$

Similarly to $I^c(t)$ in the proof of Lemma 3.1,

$$\left|I^Z(t)\right| \le C\left\|\Delta_Z\right\|_{\alpha;[0,t]}\int_0^t\left(\left\|\overline{X}\right\|_{\infty,s}s^{-\alpha}+\left\|\overline{X}\right\|_{1,s}\right)ds \le CR\left\|\Delta_Z\right\|_{\alpha;[0,t]}.$$

Further,

$$\begin{aligned}|I^c(t)| &\le CN\int_0^t\left(|\Delta_c(s)|\,s^{-\alpha}+\int_0^s|\Delta_c(s)-\Delta_c(u)|\,h(s,u)\,du\right)ds\\ &\le C_RN\int_0^t\left(\Delta(s)s^{-\alpha}+\int_0^s|\Delta_c(s)-\Delta_c(u)|\,h(s,u)\,du\right)ds.\end{aligned}$$

Similarly to [14, Lemma 7.1], it can be shown that assumptions H3 and H4 imply that for any $s,u\in[0,T]$ and $\psi_1,\dots,\psi_4\in\mathcal{C}$ with $\|\psi_i\|\le R$, $i=1,\dots,4$,

$$\begin{aligned}&|c(s,\psi_1)-c(u,\psi_2)-c(s,\psi_3)+c(u,\psi_4)|\\ &\quad\le C_R\Big(\|\psi_1-\psi_2-\psi_3+\psi_4\|_{\mathcal{C}}\\ &\qquad\qquad+\|\psi_1-\psi_3\|_{\mathcal{C}}\left(|s-u|^\beta+\|\psi_1-\psi_2\|_{\mathcal{C}}+\|\psi_3-\psi_4\|_{\mathcal{C}}\right)\Big).\end{aligned} \tag{3.6}$$

Therefore, we can estimate $|I^c(t)| \le C_R N \sum_{k=1}^d I_k^c(t)$, where

$$I_1^c(t) = \int_0^t \Delta(s) s^{-\alpha}\, ds;$$

$$\begin{aligned} I_2^c(t) &= \int_0^t \int_0^s \left\| X_s - \overline{X}_s - \overline{X}_u + \overline{X}_s \right\|_{\mathcal{C}} h(s,u)\, du\, ds \le \int_0^t \left\| X - \overline{X} \right\|_{1,s} ds \\ &\le \int_0^t \Delta(s)\, ds; \end{aligned}$$

$$I_3^c(t) = \int_0^t \int_0^s \left\| X_s - \overline{X}_s \right\|_{\mathcal{C}} (s-u)^{\beta-\alpha-1}\, du\, ds \le C \int_0^t \left\| X - \overline{X} \right\|_{\infty,s} ds \le C \int_0^t \Delta(s)\, ds;$$

$$\begin{aligned} I_4^c(t) &= \int_0^t \int_0^s \left\| X_s - \overline{X}_s \right\|_{\mathcal{C}} \left(\|X_s - X_u\|_{\mathcal{C}} + \left\| \overline{X}_s - \overline{X}_u \right\|_{\mathcal{C}} \right) h(s,u)\, du \\ &\le \int_0^t \left\| X_s - \overline{X}_s \right\|_{\infty,s} \left(\|X\|_{1,s} + \left\| \overline{X} \right\|_{1,s} \right) \le 2R \int_0^t \Delta(s)\, ds. \end{aligned}$$

Therefore, we have

$$\left\| X - \overline{X} \right\|_{\infty,t} \le C_{N,R} \left(\|\Delta_Z\|_{\alpha;[0,t]} + \int_0^t \Delta(s) s^{-\alpha}\, ds \right) + \left\| I^b \right\|_{\infty;[0,t]}. \qquad (3.7)$$

Further, let $0 \le s \le t$. Then for $u \le s - t$,

$$\left| X(u+t-s) - \overline{X}(u+t-s) - X(u) + \overline{X}(u) \right| = 0;$$

for $u \in (s-t, 0]$

$$\left| X(u+t-s) - \overline{X}(u+t-s) - X(u) + \overline{X}(u) \right| = \left| X(u+t-s) - \overline{X}(u+t-s) \right|.$$

Consequently, we can write

$$\left\| X - \overline{X} \right\|_{1,t} \le J^a(t) + J^b(t) + J^c(t) + J^Z(t),$$

where

$$J^a(t) = \int_0^t \sup_{u \in [s-t,s]} \left| \int_{u \vee 0}^{u+t-s} \Delta_a(v)\, dv \right|,$$

$$J^b(t) = \int_0^t \sup_{u \in [s-t,s]} \left| \int_{u \vee 0}^{u+t-s} \Delta_b(v)\, dW(v) \right| ds,$$

$$J^c(t) = \int_0^t \sup_{u \in [s-t,s]} \left| \int_{u \vee 0}^{u+t-s} \Delta_c(v)\, dZ(v) \right|,$$

$$J^Z(t) = \int_0^t \sup_{u \in [s-t,s]} \left| \int_{u \vee 0}^{u+t-s} c(\overline{X}_v, v)\, d\Delta_Z(v) \right|.$$

Estimate

$$J^a(t) \le C_R \int_0^t \max_{u \in [s-t,s]} \int_{u \vee 0}^{u+t-s} \Delta(v)\, dv\, h(t,s)\, ds \le C \int_0^t \Delta(z)(t-z)^{-\alpha}\, dz.$$

Similarly to $J^c(t)$ in the proof of Lemma 3.1,

$$\begin{aligned} J^Z(t) &= C \|\Delta_Z\|_{\alpha;[0,t]} \left(1 + \int_0^t \left(\left\| \overline{X} \right\|_{\infty,s} (t-s)^{-2\alpha} + \left\| \overline{X} \right\|_{1,s} (t-s)^{-\alpha} \right) ds \right) \\ &\le C_R \|\Delta_Z\|_{\alpha;[0,t]}. \end{aligned}$$

Further, using (3.6), we can estimate, analogously to $I^c(t)$ above,

$$J^c(t) \le C_{N,R} \sum_{k=1}^4 J_k^c(t),$$

where

$$\begin{aligned}
J_1^c(t) &= \int_0^t \max_{u\in[s-t,s]} \int_{u\vee 0}^{u+t-s} |\Delta_c(v)|\,(v-u\vee 0)^{-\alpha}\,dv\,h(t,s)\,ds\\
&\le C_R \int_0^t\int_s^t \Delta(z)(z-s)^{-\alpha}\,dz\,h(t,s)\,ds \le C\int_0^t \Delta(z)(t-z)^{-2\alpha}\,dz;\\
J_2^c(t) &= \int_0^t \max_{u\in[-r,s]} \int_{u\vee 0}^{u+t-s}\int_{u\vee 0}^{v} \left\|X-\overline{X}-X+\overline{X}\right\|_{\infty,z} h(v,z)\,dz\,dv\,h(t,s)\,ds\\
&\le C\int_0^t \max_{u\in[-r,s]} \int_{u\vee 0}^{u+t-s} \left\|X-\overline{X}\right\|_{1,v}\,dv\,h(t,s)\,ds\\
&\le C\int_0^t\int_s^t \left\|X-\overline{X}\right\|_{1,v} dv\,h(t,s)\,ds \le C\int_0^t \Delta(v)(t-v)^{-\alpha}\,dv;\\
J_3^c(t) &= \int_0^t \max_{u\in[-r,s]} \int_{u\vee 0}^{u+t-s}\int_{u\vee 0}^{v} \left\|X-\overline{X}\right\|_{\infty,v}(v-z)^{\beta-\alpha-1}\,dz\,dv\,h(t,s)\,ds\\
&\le C\int_0^t \max_{u\in[-r,s]} \int_{u\vee 0}^{u+t-s} \left\|X-\overline{X}\right\|_{\infty,v}\,dv\,h(t,s)\,ds\\
&\le C\int_0^t\int_s^t \left\|X-\overline{X}\right\|_{\infty,v}\,dv\,h(t,s)\,ds \le C\int_0^t \Delta(v)(t-v)^{-\alpha}\,dv;\\
J_4^c(t) &= \int_0^t \max_{u\in[-r,s]} \int_{u\vee 0}^{u+t-s}\int_{u\vee 0}^{v} \left\|X-\overline{X}\right\|_{\infty,v}\left(\|X_v-X_z\|_{\mathcal{C}}+\left\|\overline{X}_v-\overline{X}_z\right\|_{\mathcal{C}}\right)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times h(v,z)\,dz\,dv\,h(t,s)\,ds\\
&\le C\int_0^t \max_{u\in[-r,s]} \int_{u\vee 0}^{u+t-s} \left\|X-\overline{X}\right\|_{\infty,v}\left(\|X\|_{1,v}+\left\|\overline{X}\right\|_{1,v}\right)dv\,h(t,s)\,ds\\
&\le CR\int_0^t\int_s^t \left\|X-\overline{X}\right\|_{\infty,v}\,dv\,h(t,s)\,ds \le CR\int_0^t \Delta(v)(t-v)^{-\alpha}\,dv.
\end{aligned}$$

Summing the estimates for $\left\|X-\overline{X}\right\|_{1,t}$, we get

$$\left\|X-\overline{X}\right\|_{1,t} \le C_{N,R}\left(\|\Delta_Z\|_{\alpha;[0,t]} + \int_0^t \Delta(s)(t-s)^{-2\alpha}\,ds\right) + J^b(t).$$

Combining this with the estimate(3.7), we obtain

$$\left\|X-\overline{X}\right\|_t \le C_{N,R}\left(\|\Delta_Z\|_{\alpha;[0,t]} + \int_0^t \|X\|_s\, g(t,s)\,ds\right) + \left\|I^b\right\|_{\infty;[0,t]} + J^b(t)$$

for $\omega\in B_{R,t}$, where $g(t,s)=s^{-\alpha}+(t-s)^{-2\alpha}$. The rest of the proof goes exactly as in the Lemma 3.1. Namely, denoting $\mathbb{1}_t=\mathbb{1}_{B_{R,t}}$, we obtain

$$\mathsf{E}\left[\|X\|_t^p\,\mathbb{1}_t\right] \le C_{N,p}\left(\mathsf{E}\left[\|\Delta_Z(t)\|_{\alpha;[0,t]}^p\right] + \int_0^t \mathsf{E}\left[\|X\|_s^p\,\mathbb{1}_s\right] g(t,s)\,ds\right),$$

which implies the required statement with the help of the generalized Gronwall lemma. □

4. Existence and uniqueness of solution

Now we have everything to establish the unique solvability of (2.2).

Theorem 4.1. *Equation* (2.2) *has a unique solution.*

Proof. For convenience, the proof will be divided into several logical steps.

Step 1. Approximations by usual stochastic delay differential equations. Fix some $N \geq 1$ and define $\tau_N = \inf\left\{t > 0\colon \|Z\|_{\alpha;[0,t]} \geq N\right\}$, $Z^N(t) = Z(t \wedge \tau_N)$, $t \geq 0$. For each integer $n \geq 1$ define a smooth approximation of Z^N by

$$Z^{N,n}(t) = n \int_{(t-1/n)\vee 0}^{t} Z^N(s)\, ds$$

and consider the equation

$$X^{N,n}(s) = X(0) + \int_0^t a\left(s, X_s^{N,n}\right)\, dt + \int_0^t b\left(s, X_s^{N,n}\right)\, dW(s) + \int_0^t c\left(s, X_s^{N,n}\right)\, dZ^{N,n}(s)$$

with the same initial condition $X^{N,n}(s) = \eta(s)$, $s \in [-r, 0]$. Since $Z^{N,n}$ is absolutely continuous, this is a usual stochastic delay differential equation (or, in the terminology of [12], stochastic functional differential equation)

$$X^{N,n}(s) = X(0) + \int_0^t d^{N,n}\left(s, X_s^{N,n}\right)\, dt + \int_0^t b\left(s, X_s^{N,n}\right)\, dW(s) \tag{4.1}$$

with a random drift $d^{N,n}(s,\psi) = a(s,\psi) + c(s,\psi)\frac{d}{ds}Z^{N,n}(s)$. Clearly, $\left|\frac{d}{ds}Z^{N,n}(s)\right| \leq nN$. Therefore, the coefficients of (4.1) satisfy the linear growth condition: for all $s \in [0,T]$, $\psi \in \mathcal{C}$,

$$\left|d^{N,n}(s,\psi)\right| + |b(s,\psi)| \leq C_{N,n}\left(1 + \|\psi\|_{\mathcal{C}}\right), \tag{4.2}$$

and the local Lipschitz condition: for any $R > 0$ and all $s \in [0,T]$, $\psi_1, \psi_2 \in \mathcal{C}$ with $\|\psi_1\|_{\mathcal{C}} \leq R$, $\|\psi_2\|_{\mathcal{C}} \leq R$,

$$\left|d^{N,n}(s,\psi_1) - d^{N,n}(s,\psi_2)\right| + |b(s,\psi_1) - b(s,\psi_2)| \leq C_{N,n,R}\,\|\psi_1 - \psi_2\|_{\mathcal{C}}. \tag{4.3}$$

In [12, Theorem I.2] and in [9, Chapter 5, Theorem 2.5], the unique solvability of (4.1) was formulated for non-random coefficients satisfying conditions (4.2) and (4.3). However, the arguments given there are easily seen to extend to adapted coefficients satisfying (4.2) and (4.3) with a non-random constant, which is the case here. Thus, (4.1) has a unique solution.

Step 2. Convergence of approximations. First we show that, for a fixed $N \geq 1$, the sequence $\left\{X^{N,n}, n \geq 1\right\}$ is fundamental in probability in the norm $\|\cdot\|_T$. Indeed, it is easy to show (see e.g. [11, Lemma 2.1]) that $\left\|Z^{N,n} - Z^N\right\|_{\alpha;[0,T]} \to 0$, $n \to \infty$, a.s. Then, in view of the boundedness, $\mathsf{E}\left[\left\|Z^{N,n} - Z^N\right\|_{\alpha;[0,T]}^p\right] \to 0$ for any $p \geq 1$. Therefore, Lemma 3.2 and the Markov inequality imply that

$$\mathsf{P}\left(\left\|X^{N,n} - X^{N,m}\right\|_T > \varepsilon, \left\|X^{N,n}\right\|_T \leq R, \left\|X^{N,m}\right\|_T \leq R\right) \to 0, \qquad n,m \to \infty, \tag{4.4}$$

for any $\varepsilon > 0$, $R \geq 1$. Hence,

$$\limsup_{n,m\to\infty} \mathsf{P}\left(\left\|X^{N,n} - X^{N,m}\right\|_T > \varepsilon\right) \leq 2\sup_{n\geq 1} \mathsf{P}\left(\left\|X^{N,n}\right\|_T > R\right)$$

for any $\varepsilon > 0$, $R \geq 1$. The convergence $\mathsf{E}\left[\left\|Z^{N,n} - Z^N\right\|_{\alpha;[0,T]}^p\right] \to 0$, $n \to \infty$ implies that $\sup_{n\geq 1} \mathsf{E}\left[\left\|Z^{N,n}\right\|_{\alpha;[0,T]}^p\right] < \infty$. Then, due to Lemma 3.1 and the Markov inequality,

$$\sup_{n\geq 1} \mathsf{P}\left(\left\|X^{N,n}\right\|_T > R\right) \to 0, \qquad n \to \infty,$$

whence, letting $R \to \infty$ in (4.4), we deduce $\mathsf{P}\left(\left\|X^{N,n} - X^{N,m}\right\|_T > \varepsilon\right) \to 0$, $n,m \to \infty$, as required. Therefore, there exists some random process X^N such that $\left\|X^{N,n} - X^N\right\|_T \to 0$, $n \to \infty$, in probability. There is an almost surely convergent subsequence, and without loss of generality we can assume that $\left\|X^{N,n} - X^N\right\|_T \to 0$, $n \to \infty$, a.s.

Step 3. The limit provides a solution. In order to prove that X^N solves equation (2.2) with Z replaced by Z^N, we need to show that the integrals in (4.1) converge to the correspondent integrals for X^N. Since the convergence $\left\|X^{N,n} - X^N\right\|_T \to 0$, $n \to \infty$, implies the uniform convergence on $[0,T]$, we easily obtain

$$\int_0^t a(s, X_s^{N,n})\, ds \to \int_0^t a(s, X_s^N)\, ds \quad \text{a.s.,} \qquad n \to \infty,$$

Similarly to $I^c(t)$ and $I^Z(t)$ in the proof of Lemma 3.2, we have

$$\begin{aligned}&\left|\int_0^t c\left(s, X_s^N\right)\, dZ^N(s) - \int_0^t c\left(s, X_s^{N,n}\right)\, dZ^{N,n}(s)\right| \\ &\qquad \le C_N\left(\left\|X^N\right\|_t + \left\|X^{N,n}\right\|_t\right)\left(\left\|Z^N - Z^{N,n}\right\|_{\alpha;[0,t]} + \int_0^t \left\|X^N - X^{N,n}\right\|_t\, ds\right) \to 0\end{aligned}$$

as $n \to \infty$ a.s. Finally, denoting $\mathbb{1}_t = \mathbb{1}_{\|X^N\|_t < R, \|X^{N,n}\|_t < R}$, we have

$$\begin{aligned}&\mathsf{E}\left[\left(\int_0^t b\left(s, X_s^N\right)\, dW(s) - \int_0^t b\left(s, X_s^{N,n}\right)\, dW(s)\right)^2 \mathbb{1}_t\right] \\ &\qquad \le \int_0^t \mathsf{E}\left[\left(b\left(s, X_s^N\right) - b\left(s, X_s^{N,n}\right)\right)^2 \mathbb{1}_s\right]\, ds \\ &\qquad \le \int_0^t \mathsf{E}\left[\left\|X^N - X^{N,n}\right\|_s^2 \mathbb{1}_s\right]\, ds \to 0, \qquad n \to \infty.\end{aligned}$$

So we have that

$$\left(\int_0^t b\left(s, X_s^N\right)\, dW(s) - \int_0^t b\left(s, X_s^{N,n}\right)\, dW(s)\right) \mathbb{1}_t \to 0, \qquad n \to \infty$$

in probability. Thanks to the convergence $\left\|X^{N,n} - X^N\right\|_T \to 0$, $n \to \infty$, the event $\left\{\left\|X^N\right\|_t < R\right\}$ implies $\left\{\left\|X^{N,n}\right\|_t < R\right\}$ for n large enough, therefore we have the convergence of the integrals in probability on $\left\{\left\|X^N\right\|_t < R\right\}$ and arbitrary $R \ge 1$, therefore on Ω. Thus, we have that X^N is a solution to

$$X^N(s) = X(0) + \int_0^t a\left(s, X_s^N\right)\, dt + \int_0^t b\left(s, X_s^N\right)\, dW(s) + \int_0^t c\left(s, X_s^N\right)\, dZ^N(s)$$

with $X^N(s) = \eta(s)$, $s \in [-r, 0]$.

From Lemma 3.2, it is obvious that the processes X^N and X^M with $M \ge N$ coincide a.s. on the set $A_{N,R} = \left\{\|Z\|_{\alpha;[0,T]} \le N\right\}$. Therefore, there exists a process X such that for each $N \ge 1$, $X^N = X$ a.s. on $A_{N,T}$. Consequently, X solves (2.2) on each of the sets $A_{N,T}$, $N \ge 1$, hence, almost surely.

Finally, the uniqueness follows from Lemma 3.2: each solution to (2.2) must coincide with X on each of the sets $A_{N,T}$, hence, almost surely. □

5. Integrability and convergence of solutions

Now we investigate the question when the moments of X are finite. Naturally, we need to require certain integrability of the driver Z. It is quite involved to prove the integrability under assumptions H1–H4 (for equations without delay, the corresponding result is proved in [15]). So we prove the integrability under an additional assumption that b is bounded.

Theorem 5.1. *Assume, that, in addition to H1–H4, $|b(t,\psi)| \le C$ for any $t \in [0,T]$, $\psi \in \mathcal{C}$, and $\mathsf{E}\left[\exp\left\{c\,\|Z\|_{\alpha;[0,T]}^{1/(1-\alpha)}\right\}\right] < \infty$ for all $c > 0$. Then the solution of (2.2) satisfies $\mathsf{E}\left[\|X\|_T^p\right] < \infty$ for all $p \ge 1$, in particular, all moments of the solution are finite.*

Proof. The proof follows the scheme of [16, Lemma 4.1]. We will use the notation of Lemma 3.1.

Define for $\lambda > 0$, $t \in [0,T]$, $a \in \{1,\infty\}$ $\|X\|_{\lambda;a} = \sup_{s\in[0,T]} e^{-\lambda s} \|X\|_{a,s}$. Denote also $\zeta = \|I^b\|_{\infty;[0,T]} + J^b(T)$. Then from (3.2) we get for $\omega \in A_{N,t}$

$$\begin{aligned}
\|X\|_{\lambda;\infty} &\le CN\left(1 + \sup_{s\le T} e^{-\lambda s}\int_0^s \left(\|X\|_{\infty,u} u^{-\alpha} + \|X\|_{1,u}\right) du\right) + \zeta \\
&\le CN\left(1 + \sup_{s\le T}\int_0^s e^{\lambda(u-s)}\left(e^{-\lambda u}\|X\|_{\infty,u} u^{-\alpha} + e^{-\lambda u}\|X\|_{1,u}\right) du\right) + \zeta \\
&\le CN\left(1 + \sup_{s\le T}\int_0^s e^{\lambda(u-s)}\left(u^{-\alpha}\|X\|_{\lambda;\infty} + \|X\|_{\lambda;1}\right) du\right) + \zeta \\
&\le CN\left(1 + \lambda^{\alpha-1}\|X\|_{\lambda;\infty} + \lambda^{-1}\|X\|_{\lambda;1}\right) + \zeta,
\end{aligned}$$

where we have used the estimate

$$\begin{aligned}
\sup_{s\le T}\int_0^s e^{\lambda(u-s)} u^{-\alpha}\, du &= \sup_{s\le T}\lambda^{-1}\int_0^{\lambda s} e^{-z}(s - z/\lambda)^{-\alpha}\, dz \\
&= \sup_{s\le T}\lambda^{\alpha-1}\int_0^{\lambda s} e^{-z}(\lambda s - z)^{-\alpha}\, dz \le \lambda^{\alpha-1}\sup_{a>0}\int_0^a e^{-z}(a-z)^{-\alpha}\, dz = C\lambda^{\alpha-1}.
\end{aligned}$$

Similarly, from (3.1),

$$\begin{aligned}
\|X\|_{\lambda;1} &\le N\left(1 + \sup_{s\le T} e^{-\lambda s}\int_0^s \left(\|X\|_{\infty,u}(s-u)^{-2\alpha} + \|X\|_{1,u}(s-u)^{-\alpha}\right) du\right) + \zeta \\
&\le N\Bigg(1 + \sup_{s\le T}\int_0^s e^{\lambda(u-s)}\Big(e^{-\lambda u}\|X\|_{\infty,u}(s-u)^{-2\alpha} \\
&\qquad\qquad + e^{-\lambda u}\|X\|_{1,u}(s-u)^{-\alpha}\Big) du\Bigg) \\
&\quad + \zeta \\
&\le N\left(1 + \sup_{s\le T}\int_0^s e^{\lambda(u-s)}\left(\|X\|_{\lambda;\infty}(s-u)^{-2\alpha} + \|X\|_{\lambda;1}(s-u)^{-\alpha}\right) du\right) + \zeta \\
&\le CN\left(1 + \lambda^{2\alpha-1}\|X\|_{\lambda;\infty} + \lambda^{\alpha-1}\|X\|_{\lambda;1}\right) + \zeta.
\end{aligned}$$

Therefore, we have arrived at the system of inequalities

$$\begin{aligned}
\|X\|_{\lambda;\infty} &\le KN\left(1 + \lambda^{\alpha-1}\|X\|_{\lambda;\infty} + \lambda^{-1}\|X\|_{\lambda;1}\right) + \zeta, \\
\|X\|_{\lambda;t} &\le KN\left(1 + \lambda^{2\alpha-1}\|X\|_{\lambda;\infty} + \lambda^{\alpha-1}\|X\|_{\lambda;1}\right) + \zeta.
\end{aligned}$$

Setting $\lambda = 4KN^{1/(1-\alpha)}$, it is easy to deduce from this system that

$$\|X\|_{\lambda;\infty} + \|X\|_{\lambda;1} \le CN^{1/(1-\alpha)}(1+\zeta),$$

whence

$$\|X\|_T \le e^{\lambda T}\left(\|X\|_{\lambda;\infty} + \|X\|_{\lambda;1}\right) \le C\exp\left\{CN^{1/(1-\alpha)}\right\}(1+\zeta)$$

for $\omega \in A_{N,t}$. Thus, in order to prove the required result, it remains to show that all moments of ζ are finite. The argument is similar to the estimation of I^b and J^b in Lemma 3.1, so we omit some details.

Take arbitrary $p > 4/(1-2\alpha)$. By the Burkholder inequality,

$$\mathsf{E}\left[\left\|I^b\right\|^p_{\infty;[0,T]}\right] \le C_p \mathsf{E}\left[\left(\int_0^T |b(s,X_s)|^2\, ds\right)^{p/2}\right] < \infty. \tag{5.1}$$

Further, we have

$$\mathsf{E}\left[J^b(T)^p\right] \le C_p \mathsf{E}\left[\left(\int_0^T \sup_{u\in[s-T,s]}\left|\int_{u\vee 0}^{u+T-s} b(v,X_v)dW(v)\right| h(T,s)ds\right)^p\right]. \tag{5.2}$$

By the Garsia–Rodemich–Rumsey inequality, for any $r,z\in[0,T]$

$$\left|\int_r^z b(v,X_v)\mathbb{1}_v\, dW(v)\right| \le C_p\xi(T)\left|r-z\right|^{1/2-2/p},$$

where

$$\xi(T) = \left(\int_0^t\int_0^y \frac{\left|\int_x^y b(v,X_v)\, dW(v)\right|^p}{|x-y|^{p/2}}\, dx\, dy\right)^{1/p}.$$

From the estimate

$$\begin{aligned}\mathsf{E}\left[\xi(T)^p\right] &= \int_0^T\int_0^y \frac{\mathsf{E}\left[\left|\int_x^y b(v,X_v)\, dW(v)\right|^p\right]}{|x-y|^{p/2}}\, dx\, dy\\ &\le C_p\int_0^T\int_0^y \frac{\mathsf{E}\left[\left(\int_x^y |b(v,X_v)|^2\, dv\right)^{p/2}\right]}{(y-x)^{p/2}}\, dx\, dy \le C_p\int_0^T\int_0^y 1\, dx\, dy < \infty\end{aligned}$$

we obtain, as in Lemma 3.1, $\mathsf{E}\left[J^b(T)^p\right] < \infty$. Taking into account (5.1), we get that $\mathsf{E}\left[\zeta^p\right] < \infty$, thus finishing the proof. □

Remark 5.2. The assumption on Z from Theorem 5.1 is fulfilled e.g. for a fractional Brownian motion with Hurst parameter $H > 1/2$ (with any $\alpha > 1-H$), see e.g. [16, Theorem 4].

Finally, we state a result on stability of solutions to (2.2) with respect to the driver Z. Its proof virtually repeats Step 3 of the proof of Theorem 4.1 and therefore is omitted. Let for $n\ge 1$

$$Z^n = \{Z^n(t), n\ge 1\}$$

be an $\mathbb{F}$-adapted γ-Hölder continuous process, and X^n be a solution to

$$X^n(s) = X(0) + \int_0^t a(s,X^n_s)\, dt + \int_0^t b(s,X^n_s)\, dW(s) + \int_0^t c(s,X^n_s)\, dZ^n(s) \tag{5.3}$$

with the initial condition $X^n(s) = \eta(s)$, $s\in[-r,0]$.

Proposition 5.3. *Let X and X^n be solutions of* (2.2) *and* (5.3) *respectively, and $\left\|Z-Z^n\right\|_{\alpha;[0,T]} \to 0$, $n\to\infty$, in probability. Then $\left\|X-X^n\right\|_T \to 0$, $n\to\infty$, in probability.*

References

1. M. Besalú and C. Rovira, *Stochastic delay equations with non-negativity constraints driven by fractional Brownian motion*, Bernoulli **18** (2012), no. 1, 24–45.
2. B. Boufoussi and S. Hajji, *Functional differential equations driven by a fractional Brownian motion*, Comput. Math. Appl. **62** (2011), no. 2, 746–754.
3. L. H. Duc, B. Schmalfuss, and S. Siegmund, *Generation of random dynamical systems from fractional stochastic delay differential equations*, arXiv:math.DS/1309.6478 (2013).
4. M. Ferrante and C. Rovira, *Stochastic delay differential equations driven by fractional Brownian motion with Hurst parameter $H > \frac{1}{2}$*, Bernoulli **12** (2006), no. 1, 85–100.

5. M. Ferrante and C. Rovira, *Convergence of delay differential equations driven by fractional Brownian motion*, J. Evol. Equ. **10** (2010), no. 4, 761–783.
6. J. Guerra and D. Nualart, *Stochastic differential equations driven by fractional Brownian motion and standard Brownian motion*, Stoch. Anal. Appl. **26** (2008), no. 5, 1053–1075.
7. K. Kubilius, *The existence and uniqueness of the solution of an integral equation driven by a p-semimartingale of special type*, Stochastic Process. Appl. **98** (2002), no. 2, 289–315.
8. J. A. León and S Tindel, *Malliavin calculus for fractional delay equations*, J. Theoret. Probab. **25** (2012), no. 3, 854–889.
9. X. Mao, *Stochastic Differential Equations and Applications*, 2nd ed., Horwood Publishing, Chichester, 2007.
10. Yu. S. Mishura and G. M. Shevchenko, *Stochastic differential equation involving Wiener process and fractional Brownian motion with Hurst index* $H > \frac{1}{2}$, Comm. Statist. Theory Methods **40** (2011), no. 19–20, 3492–3508.
11. Yu. S. Mishura and G. M. Shevchenko, *Mixed stochastic differential equations with long-range dependence: Existence, uniqueness and convergence of solutions*, Comput. Math. Appl. **64** (2012), no. 10, 3217–3227.
12. S.-E. A. Mohammed, *Stochastic differential systems with memory: theory, examples and applications*, Stochastic Analysis and Related Topics, VI (Geilo, 1996), Progr. Probab., vol. 42, Birkhäuser Boston, Boston, MA, 1998, pp. 1–77.
13. A. Neuenkirch, I. Nourdin, and S. Tindel, *Delay equations driven by rough paths*, Electron. J. Probab. **13** (2008), 2031–2068.
14. D. Nualart and A. Răşcanu, Differential equations driven by fractional Brownian motion, Collect. Math. **53** (2002), no. 1, 55–81.
15. G. M. Shevchenko, *Integrability of solutions to mixed stochastic differential equations*, Ukr. Math. Bulletin **10** (2013), no. 4, 559–574.
16. G. M. Shevchenko, *Mixed fractional stochastic differential equations with jumps* Stochastics (2013). (to appear)
17. G. M. Shevchenko and T. O. Shalaiko, *Malliavin regularity of solutions to mixed stochastic differential equations*, Stat. Probab. Letters **83** (2013), no. 12, 2638–2646.
18. M. Zähle, *Integration with respect to fractal functions and stochastic calculus. I*, Probab. Theory Relat. Fields **111** (1998), no. 3, 333–374.

Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, 64 Volodymyrska, 01601 Kyiv, Ukraine

E-mail address: zhora@univ.kiev.ua

Received 26/02/2013