Journal of Mathematical Sciences, Vol. 153, No. 3, 2008

SEMIMARTINGALE STOCHASTIC APPROXIMATION PROCEDURE AND RECURSIVE ESTIMATION

N. Lazrieva, T. Sharia, and T. Toronjadze UDC 519.2

ABSTRACT. The semimartingale stochastic approximation procedure, precisely, the Robbins–Monro type SDE, is introduced, which naturally includes both generalized stochastic approximation algorithms with martingale noises and recursive parameter estimation procedures for statistical models associated with semimartingales. General results concerning the asymptotic behavior of the solution are presented. In particular, the conditions ensuring the convergence, the rate of convergence, and the asymptotic expansion are established. The results concerning the Polyak weighted averaging procedure are also presented.

CONTENTS

0. Introduction

In 1951, in the famous paper of Robbins and Monro "A stochastic approximation method" [36], was created a method for dealing with the problem of location of roots of functions, which can be observed only with random errors. In fact, they carried a "stochastic" component in the classical Newton's method.

This method is known in probability theory as the Robbins–Monro (RM) stochastic approximation algorithm (procedure).

Since then, a considerable amount of works has been performed to relax the assumptions on the regression functions, as well as those on the structure of the measurement errors (see, e.g., [17, 23, 26–30, 41, 42]). In particular, in [28], the generalized stochastic approximation algorithms with deterministic regression functions and martingale noises (independent of the phase variable) as strong solutions of semimartingale SDEs were introduced.

Translated from Sovremennaya Matematika i Ee Prilozheniya (Contemporary Mathematics and Its Applications), Vol. 45, Martingale Theory and Its Application, 2007.

1072–3374/08/1533–0211

Starting from [1], a link between the RM stochastic approximation algorithm and the recursive parameter estimation procedures was intensively exploited. Later on recursive parameter estimation procedures for various special models (e.g., i.i.d. models, non-i.i.d. models in discrete time, diffusion models, etc.) was studied by a number of authors by using methods of stochastic approximation (see, e.g., [7, 17, 23, 26, 27, 38–40]). One should mention the fundamental book [32] of Nevelson and Khas'minski [32] among them.

In 1987, Lazrieva and Toronjadze proposed an heuristic algorithm for constructing the recursive parameter estimation procedures for statistical models associated with semimartingales (including both discrete and continuous time semimartingale statistical models) [18]. These procedures cannot be covered by the generalized stochastic approximation algorithm proposed by Melnikov, whereas in the i.i.d. case, the classical RM algorithm contains recursive estimation procedures.

To recover the link between the stochastic approximation and the recursive parameter estimation, in [19–21], Lazrieva, Sharia, and Toronjadze introduced the semimartingale stochastic differential equation, which naturally includes both generalized RM stochastic approximation algorithms with martingale noises and recursive parameter estimation procedures for semimartingale statistical models.

On the stochastic basis $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t\geq 0}, P)$ satisfying the usual conditions, let the following objects be given:

(a) a random field $H = \{H_t(u),\, t \geq 0,\, u \in \mathbb{R}^1\} = \{H_t(\omega, u),\, t \geq 0,\, \omega \in \Omega,\, u \in \mathbb{R}^1\}$ such that for each $u \in \mathbb{R}^1$, $H(u) = (H_t(u))_{t\geq 0} \in \mathcal{P}$ (i.e., the process is predictable);
(b) a random field $M = \{M(t, u),\, t \geq 0,\, u \in \mathbb{R}^1\} = \{M(\omega, t, u),\, \omega \in \Omega,\, t \geq 0,\, u \in \mathbb{R}^1\}$ such that for each $u \in \mathbb{R}^1$, $M(u) = (M(t, u))_{t\geq 0} \in \mathcal{M}^2_{\mathrm{loc}}(P)$;
(c) a predictable increasing process $K = (K_t)_{t\geq 0}$ (i.e., $K \in \mathcal{V}^+ \cap \mathcal{P}$).

In what follows, we restrict ourselves to consideration of the following particular cases:

1°. $M(u) \equiv m \in \mathcal{M}^2_{\mathrm{loc}}(P)$;
2°. for each $u \in \mathbb{R}^1$, $M(u) = f(u) \cdot m + g(u) \cdot n$, where $m \in \mathcal{M}^c_{\mathrm{loc}}(P)$, $n \in \mathcal{M}^{d,2}_{\mathrm{loc}}(P)$, the processes $f(u) = (f(t, u))_{t\geq 0}$ and $g(u) = (g(t, u))_{t\geq 0}$ are predictable, the corresponding stochastic integrals are well-defined, and $M(u) \in \mathcal{M}^2_{\mathrm{loc}}(P)$;
3°. for each $u \in \mathbb{R}^1$, $M(u) = \varphi(u) \cdot m + W(u) * (\mu - \nu)$, where $m \in \mathcal{M}^c_{\mathrm{loc}}(P)$, μ is an integer-valued random measure on $(R \times E, \mathcal{B}(R_+) \times \varepsilon)$, ν is its P-compensator, (E, ε) is the Blackwell space, and $W(u) = (W(t, x, u),\, t \geq 0,\, x \in E) \in \mathcal{P} \otimes \varepsilon$. Here, we also assume that all stochastic integrals are well defined.

In what follows, we denote by

$$\int_0^t M(ds, u_s),$$

where $u = (u_t)_{t\geq 0}$ is some predictable process, the following stochastic line integrals:

$$\int_0^t f(s, u_s)\, dm_s + \int_0^t g(s, u_s)\, dn_s$$

in case 2°;

$$\int_0^t \varphi(s, u_s)\, dm_s + \int_0^t \int_E W(s, x, u_s)(\mu - \nu)(ds, dx)$$

in case 3° provided they are well defined.

Consider the following semimartingale stochastic differential equation:

$$z_t = z_0 + \int_0^t H_s(z_{s-})\, dK_s + \int_0^t M(ds, z_{s-}), \quad z_0 \in \mathcal{F}_0. \tag{1}$$

The SDE (1) is called a Robbins–Monro (RM) type SDE if the drift coefficient $H_t(u)$, $t \geq 0$, $u \in \mathbb{R}^1$, satisfies the following conditions: for all $t \in [0, \infty)$ P-a.s.,

(A) $H_t(0) = 0$, $H_t(u)u < 0$ for all $u \neq 0$.

The problem of the strong solvability of the SDE (1) is well studied (see, e.g., [8, 9, 13]).

We assume that there exists a unique strong solution $z = (z_t)_{t \geq 0}$ of Eq. (1) defined on the whole interval of time $[0, \infty)$ such that $\widetilde{M} \in \mathcal{M}^2_{\text{loc}}(P)$, where

$$\widetilde{M}_t = \int_0^t M(ds, z_{s-}).$$

Some sufficient conditions for the latter can be found in [8, 9, 13].

The unique solution $z = (z_t)_{t \geq 0}$ of the RM type SDE (1) can be considered as a semimartingale stochastic approximation procedure.

In the present work, we consider the asymptotic behavior as $t \to \infty$ of the process $(z_t)_{t \geq 0}$ and also that of the averaged procedure $\overline{z} = \varepsilon^{-1}(z \circ \varepsilon)$ (see Sec. 3 for the definition of $\overline{z}$).

The work is organized as follows. In Sec. 1, we study the problem of convergence:

$$z_t \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.} \tag{2}$$

Our approach to this problem is based, first, on the description of the nonnegative semimartingale convergence sets given in Sec. 1.1 [19] (see also [19] for other references) and, second, on two representations ("standard" and "nonstandard") of the predictable process $A = (A_t)_{t \geq 0}$ in the canonical decomposition of the semimartingale $(z_t^2)_{t \geq 0}$, $z_t^2 = A_t + \text{mart}$ in the form of the difference of two predictable increasing processes A^1 and A^2. According to these representations, two groups of conditions **(I)** and **(II)** ensuring the convergence (2) are introduced.

In Sec. 1.2, the main theorem concerning (2) is formulated. Also, the relationship between groups **(I)** and **(II)** are investigated. In Sec. 1.3, some simple conditions for **(I)** and **(II)** are given.

In Sec. 1.4, a series of examples illustrating the efficience of all aspects of our approach are given. In particular, in Example 1.4.1 we introduced the recursive parameter estimation procedure for semimartingale statistical models and show how can it be reduced to the SDE (1). In Example 1.4.2, we show that the recursive parameter estimation procedure for discrete time general statistical models can also be embedded in the stochastic approximation procedure given by (1). This procedure was completely studied in [39]. In Example 1.4.3, we demonstrate that the generalized stochastic approximation algorithm proposed in [28] is covered by SDE (1).

In Sec. 2, we find the rate of convergence (see Sec. 2.2) and also show that under very mild conditions, the process $z = (z_t)_{t \geq 0}$ admits an asymptotic representation, in which the principal term is a normalized, locally square integrable martingale. In the context of the parametric statistical estimation, this implies the local asymptotic linearity of the corresponding recursive estimator. This result enables one to study the asymptotic behavior of process $z = (z_t)_{t \geq 0}$ using a suitable form of the central limit theorem for martingales (see [11, 12, 14, 25, 35]).

In Sec. 2.1, we introduce some notation and represent the normalized process $\chi^2 z^2$ in the form

$$\chi_t^2 z_t^2 = \frac{L_t}{\langle L \rangle_t^{1/2}} + R_t, \tag{3}$$

where $L = (L_t)_{t \geq 0} \in \mathcal{M}^2_{\text{loc}}(P)$ and $\langle L \rangle_t$ is the shifted square characteristic of L, i.e., $\langle L \rangle_t := 1 + \langle L \rangle_t^{F,P}$ (see also Sec. 2.1 for the definition of all objects in (3)).

In Sec. 2.2, assuming $z_t \to 0$ as $t \to \infty$ P-a.s., we give various sufficient conditions that ensure the convergence

$$\gamma_t^{\delta} z_t^2 \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.} \tag{4}$$

for all δ, $0 < \delta < \delta_0$, where $\gamma = (\gamma_t)_{t \geq 0}$ is a predictable increasing process and δ_0, $0 < \delta_0 \leq 1$, is some constant. In this section, we also give a series of examples illustrating these results.

In Sec. 2.3, assuming that Eq. (4) holds with the process asymptotically equivalent to χ^2, we study sufficient conditions that ensure the convergence

$$R_t \xrightarrow{P} 0 \quad \text{as} \quad t \to \infty, \tag{5}$$

which implies the local asymptotic linearity of recursive procedure $z = (z_t)_{t \geq 0}$. As an example illustrating the efficiency of the introduced conditions, we consider the RM stochastic approximation procedure with slowly varying gains (see [31]).

An important approach to stochastic approximation problems was proposed by Polyak [33] and Ruppert [38]. The main idea of this approach is the use of averaging iterates obtained from primary schemes. Since then, the averaging procedures were studied by a number of authors for various schemes of stochastic approximation (see [1–7, 31, 34]). The most important result of these studies is that the averaging procedures lead to asymptotically optimal estimates, and, in some cases, they converges to the limit more rapoidly than the initial algorithms.

In Sec. 3, the Polyak weighted averaging procedures of the initial process $z = (z_t)_{t \geq 0}$ are considered. They are defined as follows:

$$\overline{z}_t = \varepsilon_t^{-1}(g \circ K) \int_0^t z_s \, d\varepsilon_s(g \circ K), \tag{6}$$

where $g = (g_t)_{t \geq 0}$ is a predictable process, $g_t \geq 0$,

$$\int_0^t g_s dK_s < \infty, \quad \int_0^\infty g_t dK_t = \infty,$$

and $\varepsilon_t(X)$ as usual is the Dolean exponential.

Here, we state the conditions which guarantee the asymptotic normality of process $\overline{z} = (\overline{z}_t)_{t \geq 0}$ in the case of the continuous process under consideration.

The main result of this section is Theorem 3.3.1, in which, under the assumption that Eq. (4) holds with some increasing process $\gamma = (\gamma_t)_{t \geq 0}$ asymptotically equivalent to the process $(\Gamma_t^2 \langle L \rangle_t^{-1})_{t \geq 0}$, we give the conditions that ensure the convergence

$$\varepsilon_t^{1/2} \overline{z}_t \xrightarrow{d} \sqrt{2}\, \xi, \quad \xi \in N(0,1), \tag{7}$$

where

$$\varepsilon_t = 1 + \int_0^t \Gamma_s^2 \langle L \rangle_s^{-1} \beta_s dK_s.$$

As special cases, we obtain the results concerning averaging procedures for the standard RM stochastic approximation algorithms and those with slowly varying gains.

All the notations and facts concerning the martingale theory used in the presented work can be found in [12, 14, 25].

1. Convergence

1.1. The semimartingales convergence sets. Let $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t\geq 0}, P)$ be a stochastic basis satisfying the usual conditions, and let $X = (X_t)_{t\geq 0}$ be an F-adapted process with trajectories in the Skorokhod space D (denoted as $X = F \cap D$). Let $X_\infty = \lim_{t\to\infty} X_t$, and let $\{X \to\}$ denote the set where X_∞ exists and is a finite random variable (r.v.).

In this section, we study the structure of the set $\{X \to\}$ for a nonnegative special semimartingale X. Our approach is based on the multiplicative decomposition of positive semimartingales.

Denote by $\mathcal{V}^+$ ($\mathcal{V}$) the set of processes $A = (A_t)_{t\geq 0}$, $A_0 = 0$, $A \in F \cap D$ with nondecreasing (bounded variation on each interval $[0, t[$) trajectories. We write $X \in \mathcal{P}$ if X is a predictable process. Denote by S_P the class of special semimartingales, i.e., $X \in S_p$ if $X \in F \cap D$ and

$$X = X_0 + A + M,$$

where $A \in \mathcal{V} \cap \mathcal{P}$ and $M \in \mathcal{M}_{\text{loc}}$.

Let $X \in S_P$. Denote by $\varepsilon(X)$ the solution of the Dolean equation

$$Y = 1 + Y_- \cdot X,$$

where

$$Y_- \cdot X_t := \int_0^t Y_{s-} dX_s.$$

If $\Gamma_1, \Gamma_2 \in \mathcal{F}$, then $\Gamma_1 = \Gamma_2$ P-a.s. or $\Gamma_1 \subseteq \Gamma_2$ P-a.s. means that $P(\Gamma_1 \Delta \Gamma_2) = 0$ or $P(\Gamma_1 \cap (\Omega \setminus \Gamma_2)) = 0$, respectively, where Δ is the sign of the symmetric difference of sets.

Let $X \in S_P$. We set $A = A^1 - A^2$, where $A^1, A^2 \in \mathcal{V}^+ \cap \mathcal{P}$. Denote

$$\widehat{A} = (1 + X_- + A^2_-)^{-1} \circ A^2 \quad \left(:= \int_0^\cdot (1 + X_{s-} + A^2_{s-})^{-1} dA^1_s \right).$$

1.1.1. Theorem. *Let $X \in S_P$, $X \geq 0$. Then*

$$\{\widehat{A}_\infty < \infty\} \subseteq \{X \to\} \cap \{A^2_\infty < \infty\} \quad P\text{-a.s.}$$

Proof. Consider the process $Y = 1 + X + A^2$. Then

$$Y = Y_0 + A^1 + M, \quad Y_0 = 1 + X_0,$$

$Y \geq 1$, and $Y_-^{-1} \Delta A^1 \geq 0$. Thus, the processes $\widehat{A} = Y_-^{-1} \circ A^1$ and $\widehat{M} = (Y_- + \Delta A^1)^{-1} \cdot M$ are well defined, and, moreover, $\widehat{A} \in \mathcal{V}^+ \cap \mathcal{P}$, $\widehat{M} \in \mathcal{M}_{\text{loc}}$. Then using [25, Theorem 1, Sec. 5, Chap. 2], we obtain the following multiplicative decomposition:

$$Y = Y_0 \varepsilon(\widehat{A}) \varepsilon(\widehat{M}),$$

where $\varepsilon(\widehat{A}) \in \mathcal{V}^+ \cap \mathcal{P}$ and $\varepsilon(\widehat{M}) \in \mathcal{M}_{\text{loc}}$.

Note that $\Delta \widehat{M} > -1$. Indeed,

$$\Delta \widehat{M} = (Y_- + \Delta A^1)^{-1} \Delta M.$$

But

$$\Delta M = \Delta Y - \Delta A^1 = Y - (Y_- + \Delta A^1) > -(Y_- + \Delta A^1).$$

Therefore, $\varepsilon(\widehat{M}) > 0$ and $\{\varepsilon(\widehat{M}) \to\} = \Omega$ P-a.s.. On the other hand (see, e.g., [30, Lemma 2.5]),

$$\varepsilon_t(\widehat{A}) \uparrow \infty \Longleftrightarrow \widehat{A}_t \uparrow \infty \quad \text{as} \quad t \to \infty.$$

Hence

$$\{\widehat{A}_\infty < \infty\} \subseteq \{Y \to\} = \{X \to\} \cap \{A^2_\infty < \infty\},$$

since $A^2 < Y$ and $A^2 \in \mathcal{V}^+$. The theorem is proved. □

1.1.1. Corollary.

$$\{A^1_\infty < \infty\} = \{(1+X_-)^{-1} \circ A^1_\infty < \infty\} = \{\widehat{A}_\infty < \infty\} \quad P\text{-a.s.}$$

Proof. Obviously,

$$\{A^1_\infty < \infty\} \subseteq \{(1+X_-)^{-1} \circ A^1_\infty < \infty\} \subseteq \{\widehat{A}_\infty < \infty\} \subseteq \{X \to\} \cap \{A^2_\infty < \infty\} \quad P\text{-a.s.}$$

It remains to note that

$$\{A^1_\infty < \infty\} \cap \{X \to\} \cap \{A^2_\infty < \infty\} = \{\widehat{A}_\infty < \infty\} \cap \{X \to\} \cap \{A^2_\infty < \infty\} \quad P\text{-a.s.}$$

The corollary is proved. □

1.1.2. Corollary.

$$\{\widehat{A}_\infty < \infty\} \cap \{\varepsilon_\infty(\widehat{M}) > 0\} = \{X \to\} \cap \{A^2_\infty < \infty\} \cap \{\varepsilon_\infty(\widehat{M}) > 0\} \quad P\text{-a.s.},$$

as easily follows from the proof of Theorem 1.1.1.

1.1.1. Remark. The relation

$$\{A^1_\infty < \infty\} \subseteq \{X \to\} \cap \{A^2_\infty < \infty\} \quad P\text{-a.s.}$$

was proved in [25, Chap. 2, Sec. 6, Theorem 7] under the following additional assumptions:

(1) $EX_0 < \infty$;
(2) one of the following conditions (α) or (β) is satisfied:
  (α) there exists $\varepsilon > 0$ such that $A^1_{t+\varepsilon} \in \mathcal{F}_t$ for all $t > 0$;
  (β) for any predictable Markov moment σ,

$$E\Delta A^1_\sigma I_{\{\sigma<\infty\}} < \infty.$$

Let $A, B \in F \cap D$. We write $A \prec B$ if $B - A \in \mathcal{V}^+$.

1.1.3. Corollary. *Let $X \in S_P$, $X \geq 0$, $A \leq A^1 - A^2$, and $A \prec A^1$, where $A^1, A^2 \in \mathcal{V}^+ \cap \mathcal{P}$. Then*

$$\{A^1_\infty < \infty\} = \{(1+X_-)^{-1} \circ A^1_\infty < \infty\} \subseteq \{X \to\} \cap \{A^2_\infty < \infty\} \quad P\text{-a.s.}$$

Proof. Rewrite X in the form

$$X = X_0 + A^1 - \widetilde{A}^2 + M,$$

where $\widetilde{A}^2 = A^1 - A \in \mathcal{V}^1 \cap \mathcal{P}$. Then the required assertion follows from Theorem 1.1.1, Corollary 1.1.1, and the trivial inclusion $\{\widetilde{A}^2_\infty < \infty\} \subseteq \{A^2_\infty < \infty\}$. The corollary is proved. □

1.1.4. Corollary. *Let $X \in S_P$, $X \geq 0$, and let*

$$X = X_0 + X_- \circ B + A + M$$

with $B \in \mathcal{V}^+ \cap \mathcal{P}$, $A \in \mathcal{V} \cap \mathcal{P}$ and $M \in \mathcal{M}_{\text{loc}}$. Assume that for $A^1, A^2 \in \mathcal{V}^+ \cap \mathcal{P}$,

$$A \leq A^1 - A^2 \quad \textit{and} \quad A \prec A^1.$$

Then

$$\{A^1_\infty < \infty\} \cap \{B_\infty < \infty\} \subseteq \{X \to\} \cap \{A^2_\infty < \infty\} \quad P\text{-a.s.}$$

The proof is similar to the proof of Corollary 1.1.3 if we consider the process $X\varepsilon^{-1}(B)$.

1.1.2. Remark. Consider the discrete-time case.

Let $\mathcal{F}_0, \mathcal{F}_1, \dots$ be a nondecreasing sequence of σ-algebras, let X_n, β_n, ξ_n, $\zeta_n \in \mathcal{F}_n$, $n \geq 0$, be nonnegative r.v., and let

$$X_n = X_0 + \sum_{i=0}^{n} X_{i-1}\beta_{i-1} + A_n + M_n$$

(we mean that $X_{-1} = X_0$, $\mathcal{F}_{-1} = \mathcal{F}_0$ and $\beta_{-1} = \xi_{-1} = \zeta_{-1} = 0$), where $A_n \in \mathcal{F}_{n-1}$ with $A_0 = 0$ and $M \in \mathcal{M}_{\text{loc}}$. Note that X_n can be always represented in this form by taking

$$A_n = \sum_{i=0}^{n} (E(X_i|\mathcal{F}_{i-1}) - X_{i-1}) - \sum_{i=0}^{n} X_{i-1}\beta_{i-1}.$$

Denote

$$A_n^1 = \sum_{i=0}^{n} \xi_{i-1}, \quad A_n^2 = \sum_{i=0}^{n} \zeta_{i-1}.$$

It is clear that in this case,

$$A \prec A^1 \Longleftrightarrow \Delta A_n \le \xi_{n-1}$$

($\Delta A_n := A_n - A_{n-1}$, $n \ge 1$).

Therefore, in this case Corollary 1.1.4 can be formulated as follows: *For each n, let*

$$A_n \le \sum_{i=0}^{n} (\xi_{i-1} - \zeta_{i-1}), \quad \Delta A_n \le \xi_{n-1}.$$

Then

$$\left\{ \sum_{i=0}^{\infty} \xi_{i-1} < \infty \right\} \cap \left\{ \sum_{i=0}^{\infty} \beta_{i-1} < \infty \right\} \subseteq \{X \to\} \cap \left\{ \sum_{i=0}^{\infty} \zeta_{i-1} < \infty \right\} \quad P\text{-a.s.}$$

From this corollary, the result by Robbins and Siegmund (see [37]) follows. Indeed, the above inclusion holds if, in particular, $\Delta A_n \le \xi_{n-1} - \zeta_{n-1}$, $n \ge 1$, i.e., when

$$E(X_n \mid \mathcal{F}_{n-1}) \le X_{n-1}(1 + \beta_{n-1}) + \xi_{n-1} - \zeta_{n-1}, \quad n \ge 0.$$

In our terms, the previous inequality means that $A \prec A^1 - A^2$.

1.2. Main theorem. Consider the stochastic equation (RM procedure)

$$z_t = z_0 + \int_0^t H_s(z_{s-})\, dK_s + \int_0^t M(ds, z_{s-}), \quad t \ge 0, \quad z_0 \in \mathcal{F}_0, \tag{1.2.1}$$

or, in the differential form,

$$dz_t = H_t(z_{t-})dK_t + M(dt, z_{t-}), \quad z_0 \in \mathcal{F}_0.$$

Assume that there exists a unique strong solution $z = (z_t)_{t \ge 0}$ of (1.2.1) on the whole interval of time $[0, \infty)$, $\widetilde{M} \in \mathcal{M}^2_{\text{loc}}$, where

$$\widetilde{M}_t := \int_0^t M(ds, z_{s-}).$$

We study the problem of P-a.s. convergence $z_t \to 0$ as $t \to \infty$.

For this purpose, apply Theorem 1.1.1 to the semimartingale $X_t = z_t^2$, $t \ge 0$. Using the Itô formula, for the process $(z_t^2)_{t \ge 0}$, we obtain

$$dz_t^2 = dA_t + dN_t, \tag{1.2.2}$$

where

$$dA_t = V_t^-(z_{t-})dK_t + V_t^+(z_{t-})dK_t^d + d\langle \widetilde{M} \rangle_t,$$
$$dN_t = 2z_{t-}d\widetilde{M}_t + H_t(z_{t-})\Delta K_t\, d\widetilde{M}_t^d + d([\widetilde{M}]_t - \langle \widetilde{M} \rangle_t),$$

with

$$V_t^-(u) := 2H_t(u)u, \quad V_t^+(u) := H_t^2(u)\Delta K_t.$$

Note that $A = (A_t)_{t \ge 0} \in \mathcal{V} \cap \mathcal{P}$, $N \in \mathcal{M}_{\text{loc}}$.

Represent the process A in the form

$$A_t = A_t^1 - A_t^2 \tag{1.2.3}$$

with

$$(1)\ \begin{cases} dA_t^1 = V_t^+(z_{t-})dK_t^d + d\langle \widetilde{M}\rangle_t, \\ -dA_t^2 = V_t^-(z_{t-})dK_t, \end{cases}$$

$$(2)\ \begin{cases} dA_t^1 = [V_t^-(z_{t-})I_{\{\Delta K_t \neq 0\}} + V_t^+(z_{t-})]^+ dK_t^d + d\langle \widetilde{M}\rangle_t, \\ -dA_t^2 = \{V_t^-(z_{t-})I_{\{\Delta K_t = 0\}} - [V_t^-(z_{t-})I_{\{\Delta K_t \neq 0\}} + V_t^+(z_{t-})]^-\}dK_t, \end{cases}$$

where $[a]^+ = \max(0, a)$ and $[a]^- = -\min(0, a)$.

As follows from condition **(A)**, $\alpha_t(z_{t-}) \le 0$ for all $t \ge 0$, and, therefore, the representation (1.2.3)(1) directly corresponds to the usual (in the stochastic approximation procedures) standard form of the process A (in (1.2.2), $A = A^1 - A^2$ with A^1, A^2 from (1.2.3)(1)). Therefore, representation (1.2.3)(1) is said to be "standard," whereas representation (1.2.3)(2) is said to be "nonstandard."

Introduce the following set of conditions. For all $u \in \mathbb{R}^1$ and $t \in [0, \infty)$, we have

(A) for all $t \in [0, \infty)$, $H_t(0) = 0$ P-a.s. and $H_t(0)u < 0$ for all $u \neq 0$;

(B) (i) $\langle M(u)\rangle \ll K$,

(ii) $h_t(u) \le B_t(1 + u^2)$, $B_t \ge 0$, $B = (B_t)_{t\ge 0} \in \mathcal{P}$, $B \circ K_\infty < \infty$, where $h_t(u) = \dfrac{d\langle M(u)\rangle_t}{dK_t}$;

(I) (i$_1$) $I_{\{\Delta K_t \neq 0\}}|H_t(u)| \le C_t(1 + |u|)$, $C_t \ge 0$, $C = (C_t)_{t\ge 0} \in \mathcal{P}$, $C \circ K_t < \infty$,

(i$_2$) $C^2 \Delta K \circ K_\infty^d < \infty$,

(ii) for each $\varepsilon > 0$,

$$\inf_{\varepsilon \le |u| \le 1/\varepsilon} |V^-(u)| \circ K_\infty = \infty;$$

(II) (i) $[V_t^-(u)I_{\{\Delta K_t \neq 0\}} + V_t^+(u)]^+ \le D_t(1 + u^2)$, $D_t \ge 0$, $D = (D_t)_{t\ge 0} \in \mathcal{P}$, $D \circ K_\infty^d < \infty$,

(ii) for each $\varepsilon > 0$

$$\inf_{\varepsilon \le |u| \le 1/\varepsilon} \{|V^-(u)|I_{\{\Delta K_t = 0\}} + [V^-(u)I_{\{\Delta K_t \neq 0\}} + V^+(u)]^-\} \circ K_\infty = \infty.$$

1.2.1. Remark. When $M(u) \equiv m \in \mathcal{M}^2_{\text{loc}}$, we do not claim the condition $\langle m\rangle \ll K$ and replace the condition **(B)** by

(B′) $\langle m\rangle_\infty < \infty$.

1.2.2. Remark. Everywhere, we assume that all conditions are satisfied P-a.s.

1.2.3. Remark. It is obvious that **(I)**(ii)$\Longrightarrow C \circ K_\infty = \infty$.

1.2.1. Theorem. *Let conditions* **(A)**, **(B)**, **(I)** *or* **(A)**, **(B)**, **(II)** *be satisfied. Then*

$$z_t \to 0 \quad P\text{-a.s.} \quad \textit{as} \quad t \to \infty.$$

Proof. For example, assume that conditions **(A)**, **(B)** and **(I)** are satisfied. Then using Corollary 1.1.1 and (1.2.2) with the standard representation (1.2.3)(1) of process A, we obtain

$$\{(1 + z_-^2)^{-1} \circ A_\infty^1 < \infty\} \subseteq \{z^2 \to\} \cap \{A_\infty^2 < \infty\}. \tag{1.2.4}$$

But, from conditions **(B)** and **(I)**(i), we have

$$\{(1 + z_-^2)^{-1} \circ A_\infty^1 < \infty\} = \Omega \quad P\text{-a.s.},$$

and, therefore,

$$\{z^2 \to\} \cap \{A_\infty^2 < \infty\} = \Omega \quad P\text{-a.s.} \tag{1.2.5}$$

Denote $z_\infty^2 = \lim_{t\to\infty} z_t^2$ and $N = \{z_\infty^2 > 0\}$; assume that $P(N) > 0$. In this case, by simple arguments, from **(I)**(ii) we obtain

$$P(|V^-(z_-)| \circ K_\infty = \infty) > 0,$$

which contradicts (1.2.4). Hence $P(N) = 0$. The proof of the second case is quite similar. The theorem is proved. □

In the following propositions, the relationship between conditions **(I)** and **(II)** are given.

1.2.1. Proposition. (I)⇒(II).

Proof. From **(I)**(i_1), we have

$$[V_t^-(u)I_{\{\Delta K_t\neq 0\}} + V_t^+(u)]^+ \le V_t^+(u) \le C_t^2\Delta K_t(1+u^2),$$

and if we take $D_t = C_t^2\Delta K_t$, then **(II)**(i) follows from **(I)**(i_2).

Furthermore, from **(I)**(i_1), for each $\varepsilon > 0$ and u with $\varepsilon \le |u| \le 1/\varepsilon$, we have

$$|V_t^-(u)|I_{\{\Delta K_t=0\}} + [V_t^-(u)+V_t^+(u)]^- I_{\{\Delta K_t\neq 0\}} \ge |V_t^-(u)| - V_t^+(u) \ge |V_t^-(u)| - C_t^2\Delta K_t\Big(1+\frac{1}{\varepsilon^2}\Big).$$

Now **(II)**(ii) follows from **(I)**(i_2) and **(I)**(ii). The proposition is proved. □

1.2.2. Proposition. *Under* **(I)**(i), *we have* **(I)**(ii) ⇔ **(II)**(ii).

The proof follows from the previous proposition and the trivial implication **(II)**(ii)⇒**(I)**(ii).

1.3. Some simple sufficient conditions for (I) and (II). Introduce the following group of conditions: for each $u \in \mathbb{R}^1$ and $t \in [0,\infty)$,

(S.1) (i_1) $G_t|u| \le |H_t(u)| \le \widetilde{G}_t|u|$, $G_t \ge 0$, $G = (G_t)_{t\ge 0}$, $\widetilde{G} = (\widetilde{G}_t)_{t\ge 0} \in \mathcal{P}$, $\widetilde{G}\circ K_t < \infty$,
(i_2) $\widetilde{G}^2\Delta K \circ K_\infty^d < \infty$;
(ii) $G\circ K_\infty = \infty$;

(S.2) (i) $\widetilde{G}[-2+\widetilde{G}\Delta K]^+ \circ K_\infty^d < \infty$;
(ii) $G\{2I_{\{\Delta K=0\}} + [-2+\widetilde{G}\Delta K]^- I_{\{\Delta K\neq 0\}} \circ K_\infty = \infty$.

1.3.1. Proposition. (S.1) ⇒ (I), (S.1)(i_1), **(S.2) ⇒ (II).**

Proof. The first implication is obvious. For the second, note that

$$\begin{aligned} &V_t^-(u)I_{\{\Delta K_t\neq 0\}} + V_t^+(u) \\ &\qquad = -2|H_t(u)|\,|u|I_{\{\Delta K_t\neq 0\}} + H_t^2(u)\Delta K_t \le |H_t(u)|\,|u|\,[-2I_{\{\Delta K_t\neq 0\}} + \widetilde{G}_t\Delta K_t]. \end{aligned} \tag{1.3.1}$$

Therefore,

$$[V_t^-(u)I_{\{\Delta K_t\neq 0\}} + V_t^+(u)]^+ \le |H_t(u)|\,|u|\,[-2I_{\{\Delta K_t\neq 0\}} + \widetilde{G}_t\Delta K_t]^+ \le \widetilde{G}_t[-2I_{\{\Delta K_t\neq 0\}} + \widetilde{G}_t\Delta K_t]^+|u^2|,$$

and **(II)**(i) follows from **(S.2)**(i) if we take

$$D_t = \widetilde{G}_t[-2+\widetilde{G}_t\Delta K_t]^+ I_{\{\Delta K_t\neq 0\}}.$$

Furthermore, from (1.3.1), we have

$$|V_t^-(u)|I_{\{\Delta K_t=0\}} + [V_t^-(u)I_{\{\Delta K_t\neq 0\}} + V_t^+(u)]^- \ge u^2 G_t\{2I_{\{\Delta K_t=0\}} + [-2I_{\{\Delta K_t\neq 0\}} + \widetilde{G}_t\Delta K_t]^-\}$$

and **(II)**(ii) follows from (1.2.3). The proposition is proved. □

1.3.1. Remark.

(a) **(S.1)** ⇒ **(S.2)**;
(b) under **(S.1)**(i), we have **(S.1)**(ii) ⇔ **(S.2)**(ii);
(c) **(S.2)**(ii) ⇒ **(S.1)**(ii).

Summarizing the above, we arrive at the following conclusions: a) if the condition **(S.1)**(ii) is not satisfied, then **(S.2)**(ii) is also not satisfied; b) if **(S.1)**(i_1) and **(S.1)**(ii) are satisfied, but **(S.1)**(i_2) is violated, then nevertheless the conditions **(S.2)**(i) and **(S.2)**(ii) can be satisfied.

In this case, the nonstandard representations (1.2.3)(2) is useful.

1.3.2. Remark. Denote

$$\widetilde{G}_t \Delta K_t = 2 + \delta_t, \quad \delta_t \geq -2 \quad \text{for all} \quad t \in [0, \infty).$$

It is obvious that if $\delta_t \leq 0$ for all $t \in [0, \infty)$, then $[-2 + \widetilde{G}_t \Delta K_t]^+ = 0$. Therefore, **(S.2)**(i) is trivially satisfied and **(S.2)**(ii) takes the form

$$G\{2I_{\{\Delta K=0\}} + |\delta| I_{\{\Delta K \neq 0\}} \circ K_\infty = \infty. \tag{1.3.2}$$

Note that if $G \cdot \min(2, |\delta|) \circ K_\infty = \infty$, then (1.3.2) holds, and the simplest sufficient condition (1.3.2) is: for all $t \geq 0$,

$$G \circ K_\infty = \infty, \quad |\delta_t| \geq \text{const} > 0.$$

1.3.3. Remark. Let conditions **(A)**, **(B)**, and **(I)** be satisfied. Since we apply Theorem 1.1.1 and its corollaries on the semimartingales convergence sets given in Sec. 1.1, we get rid of many of the "usual" restrictions: "moment" restrictions, boundedness of regression function, etc.

1.4. Examples.

1.4.1. Example. Recursive parameter estimation procedures for statistical models associated with semimartingale.

1. *Basic model and regularity.* The object of consideration is a parametric filtered statistical model

$$\varepsilon = (\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \{P_\theta; \theta \in \mathbb{R}\})$$

associated with a one-dimensional $\mathbb{F}$-adapted RCLL process $X = (X_t)_{t \geq 0}$ in the following way: for each $\theta \in \mathbb{R}^1$, P_θ is a unique measure on $(\Omega, \mathcal{F})$ such that with this measure, X is a semimartingale with predictable characteristics $(B(\theta), C(\theta), \nu_\theta)$ (with respect to the standard truncation function $h(x) = xI_{\{|x| \leq 1\}}$). For simplicity, assume that all P_θ coincide on $\mathcal{F}_0$.

Assume that for each pair (θ, θ'), $P_\theta \overset{\text{loc}}{\sim} P_{\theta'}$. Fix $\theta = 0$ and denote $P = P_0$, $B = B(0)$, $C = C(0)$, and $\nu = \nu_0$.

Let $\rho(\theta) = (\rho_t(\theta))_{t \geq 0}$ be a local density process (likelihood ratio process):

$$\rho_t(\theta) = \frac{dP_{\theta,t}}{dP_t},$$

where, for each θ $P_{\theta,t} := P_\theta | \mathcal{F}_t$, $P_t := P | \mathcal{F}_t$ are restrictions of the measures P_θ and P to $\mathcal{F}_t$, respectively.

As is well known (see, e.g., [14, Chap. III, Sec. 3d, Theorem 3.24]), for each θ, there exist a $\widetilde{\mathcal{P}}$-measurable positive function

$$Y(\theta) = \{Y(\omega, t, x; \theta), \quad (\omega, t, x) \in \Omega \times R_+ \times R\},$$

and a predicable process $\beta(\theta) = (\beta_t(\theta))_{t \geq 0}$ with

$$|h(Y(\theta) - 1)| * \nu \in \mathcal{A}^+_{\text{loc}}(P), \quad \beta^2(\theta) \circ C \in \mathcal{A}^+_{\text{loc}}(P)$$

such that

$$\begin{aligned} &(1) \quad B(\theta) = B + \beta(\theta) \circ C + h(Y(\theta) - 1) * \nu, \\ &(2) \quad C(\theta) = C, \qquad (3) \quad \nu_\theta = Y(\theta) \cdot \nu. \end{aligned} \tag{1.4.1}$$

In addition, the function $Y(\theta)$ can be chosen in such a way that

$$a_t := \nu(\{t\}, R) = 1 \Longleftrightarrow a_t(\theta) := \nu_\theta(\{t\}, R) = \int Y(t, x; \theta) \nu(\{t\}) dx = 1.$$

We assume that the model is regular in the Jacod sense (see [15, Sec. 3, Definition 3.12]) at each point θ, i.e., the process $(\rho_{\theta'}/\rho_\theta)^{1/2}$ is locally differentiable with respect to θ' at θ with the derived process

$$L(\theta) = (L_t(\theta))_{t \geq 0} \in \mathcal{M}^2_{\text{loc}}(P_\theta).$$

In this case the Fisher information process is defined as

$$\widehat{I}_t(\theta) = \langle L(\theta), L(\theta)\rangle_t. \tag{1.4.2}$$

In [15, Sec. 2-c, Theorem 2.28], it was proved that the regularity of the model at point θ is equivalent to the differentiability of characteristics $\beta(\theta)$, $Y(\theta)$, and $a(\theta)$ in the following sense: there exist a predictable process $\dot{\beta}(\theta)$ and $\widetilde{\mathcal{P}}$-measurable function $W(\theta)$ with

$$\dot{\beta}^2(\theta) \circ C_t < \infty, \quad W^2(\theta) * \nu_{\theta,t} < \infty \quad \text{for all} \quad t \in \mathbb{R}_+$$

and such that for all $t \in \mathbb{R}_+$, we have as $\theta' \to \theta$

$$\begin{aligned} &(1) \quad (\beta(\theta') - \beta(\theta) - \dot{\beta}(\theta)(\theta' - \theta))^2 \circ C_t/(\theta' - \theta)^2 \overset{P_\theta}{\to} 0, \\ &(2) \quad \left(\left(\frac{Y(\theta')}{Y(\theta)}\right)^{1/2} - 1 - \frac{1}{2} W(\theta)(\theta' - \theta)\right)^2 * \nu_{\theta,t} \Big/ (\theta' - \theta)^2 \overset{P_\theta}{\to} 0, \\ &(3) \quad \sum_{\substack{s \le t \\ a_s(\theta) < 1}} \left[(1 - a_s(\theta'))^{1/2} - (1 - a_s(\theta))^{1/2} + \frac{1}{2} \frac{\widehat{W}_s^\theta(\theta)}{(1 - a_s(\theta))^{1/2}} (\theta' - \theta) \right]^2 \Big/ (\theta' - \theta)^2 \overset{P_\theta}{\to} 0, \end{aligned} \tag{1.4.3}$$

where

$$\widehat{W}_t^\theta(\theta) = \int W(t, x; \theta) \nu_\theta(\{t\}, dx).$$

In this case, $a_s(\theta) = 1 \Rightarrow \widehat{W}_s^\theta(\theta) = 0$ and the process $L(\theta)$ can be written as

$$L(\theta) = \dot{\beta}(\theta) \cdot (X^c - \beta(\theta) \circ C) + \left(\widehat{W}^\theta(\theta) + \frac{\widehat{W}^\theta(\theta)}{1 - a(\theta)} \right) * (\mu - \nu_\theta), \tag{1.4.4}$$

and

$$\widehat{I}(\theta) = \dot{\beta}^2(\theta) \circ C + (\widehat{W}^\theta(\theta))^2 * \nu_\theta + \sum_{s \le \cdot} \frac{(\widehat{W}_s^\theta(\theta))^2}{1 - a_s(\theta)}. \tag{1.4.5}$$

Denote

$$\Phi(\theta) = W(\theta) + \frac{\widehat{W}^\theta(\theta)}{1 - a(\theta)}\,.$$

One can consider another alternative definition of the regularity of the model (see, e.g., [35]) based on the following representation of the process $\rho(\theta)$:

$$\rho(\theta) = \varepsilon(M(\theta)),$$

where

$$M(\theta) = \beta(\theta) \cdot X^c + \left(Y(\theta) - 1 + \frac{\widehat{Y}(\theta) - a}{1 - a} I_{\{0 < a < 1\}} \right) * (\mu - \nu) \in \mathcal{M}_{\text{loc}}(P). \tag{1.4.6}$$

Here, X^c is a continuous martingale part of X under the measure P (see, e.g., [16, 28]).

We say that the model is regular if for almost all (ω, t, x), the functions $\beta : \theta \to \beta_t(\omega; \theta)$ and $Y : \theta \to Y(\omega, t, x; \theta)$ are differentiable (notation $\dot{\beta}(\theta) := \frac{\partial}{\partial\theta}\beta(\theta)$, $\dot{Y}(\theta) := \frac{\partial}{\partial\theta}Y(\theta)$) and differentiability under the sign of integral is possible. Then

$$\frac{\partial}{\partial\theta} \ln \rho(\theta) = L(\dot{M}(\theta), M(\theta)) := \widetilde{L}(\theta) \in \mathcal{M}_{\text{loc}}(P_\theta),$$

where $L(m, M)$ is the Girsanov transformation defined as follows: if $m, M \in \mathcal{M}_{\text{loc}}(P)$ and $Q \ll P$ with $\frac{dQ}{dP} = \varepsilon(M)$, then

$$L(m, M) := m - (1 + \Delta M)^{-1} \circ [m, M] \in \mathcal{M}_{\text{loc}}(Q).$$

It is not difficult to verify that

$$\widetilde{L}(\theta) = \dot{\beta}(\theta) \cdot (X^c - \beta(\theta) \circ C) + \widetilde{\Phi}(\theta) * (\mu - \nu_\theta), \tag{1.4.7}$$

where

$$\widetilde{\Phi}(\theta) = \frac{\dot{Y}(\theta)}{Y(\theta)} + \frac{\dot{a}(\theta)}{1 - a(\theta)}$$

with $I_{\{a(\theta)=1\}}\dot{a}(\theta) = 0$.

If we assume that for each $\theta \in \mathbb{R}^1$ $\widetilde{L}(\theta) \in \mathcal{M}^2_{\mathrm{loc}}(P_\theta)$, then the Fisher information process is

$$\widehat{I}_t(\theta) = \langle \widetilde{L}(\theta), \widetilde{L}(\theta) \rangle_t.$$

It should be noted that from the regularity of the model in the Jacod sense it follows that $L(\theta) \in \mathcal{M}^2_{\mathrm{loc}}(P_\theta)$, while under the latter regularity conditions $\widetilde{L}(\theta) \in \mathcal{M}^2_{\mathrm{loc}}(P_\theta)$ is an assumption in general.

In the sequel, we assume that the model is regular in both senses given above. Then

$$W(\theta) = \frac{\dot{Y}(\theta)}{Y(\theta)}, \quad \widehat{W}^\theta(\theta) = \dot{a}(\theta), \quad L(\theta) = \widetilde{L}(\theta).$$

2. *Recursive estimation procedure for MLE.* In [18], an heuristic algorithm was proposed for constructing recursive estimators of the unknown parameter θ asymptotically equivalent to the maximum likelihood estimator (MLE).

This algorithm was derived using the following reasoning.

Consider the MLE $\widehat{\theta} = (\widehat{\theta}_t)_{t \ge 0}$, where $\widehat{\theta}_t$ is a solution of estimation equation

$$L_t(\theta) = 0.$$

Assume that

(1) for each $\theta \in \mathbb{R}^1$, the process $(\widehat{I}_t(\theta))^{1/2}(\widehat{\theta}_t - \theta)$ is P_θ-stochastically bounded, and, in addition, the process $(\widehat{\theta}_t)_{t \ge 0}$ is a P_θ-semimartingale;
(2) for each pair (θ', θ), the process $L(\theta') \in \mathcal{M}^2_{\mathrm{loc}}(P_{\theta'})$ and is a P_θ-special semimartingale;
(3) the family $(L(\theta), \theta \in \mathbb{R}^1)$ is such that the Ito–Ventzel formula is applicable to the process $(L(t, \widehat{\theta}_t))_{t \ge 0}$ with respect to P_θ for each $\theta \in \mathbb{R}^1$;
(4) for each $\theta \in \mathbb{R}^1$, there exists a positive increasing predictable process $(\gamma_t(\theta))_{t \ge 0}$ asymptotically equivalent to $\widehat{I}_t^{-1}(\theta)$, i.e.,

$$\gamma_t(\theta)\widehat{I}_t(\theta) \xrightarrow{P_\theta} 1 \quad \text{as} \quad t \to \infty.$$

Under these assumptions, using the Itô–Ventzel formula for the process $(L(t, \widehat{\theta}_t))_{t \ge 0}$, we obtain an "implicit" stochastic equation for $\widehat{\theta} = (\widehat{\theta}_t)_{t \ge 0}$. Analyzing the orders of infinitesimality of terms of this equation and rejecting the high-order terms, we obtain the following SDE (recursive procedure):

$$d\theta_t = \gamma_t(\theta_{t-})L(dt, \theta_{t-}), \tag{1.4.8}$$

where $L(dt, u_t)$ is a stochastic line integral with respect to the family $\{L(t,u),\ u \in \mathbb{R}^1,\ t \in \mathbb{R}_+\}$ of P_θ-special semimartingales along the predictable curve $u = (u_t)_{t \ge 0}$.

To give an explicit form to the SDE (1.4.8) for the statistical model associated with the semimartingale X, assume for a moment that for each (u, θ) (including the case $u = \theta$),

$$|\Phi(u)| * \mu \in \mathcal{A}^+_{\mathrm{loc}}(P_\theta). \tag{1.4.9}$$

Then for each pair (u, θ), we have

$$\Phi(u) * (\mu - \nu_u) = \Phi(u) * (\mu - \nu_\theta) + \Phi(u)\left(1 - \frac{Y(u)}{Y(\theta)}\right) * \nu_\theta.$$

Based on this equality one can obtain the canonical decomposition of P_θ-special semimartingale $L(u)$ (with respect to measure P_θ):

$$L(u) = \dot\beta(u)\circ(X^c - \beta(\theta)\circ C) + \Phi(u) * (\mu - \nu_\theta)$$
$$+ \dot\beta(u)(\beta(\theta) - \beta(u))\circ C + \Phi(u)\left(1 - \frac{Y(u)}{Y(\theta)}\right) * \nu_\theta. \tag{1.4.10}$$

Now, in view of (1.4.10), the meaning of $L(dt, u_t)$ is

$$\int_0^t L(ds, u_{s-}) = \int_0^t \dot\beta_s(u_{s-})d(X^c - \beta(\theta)\circ C)_s + \int_0^t\int \Phi(s,x,u_{s-})(\mu - \nu_\theta)(ds,dx)$$
$$+ \int_0^t \dot\beta_s(u_s)(\beta_s(\theta) - \beta_s(u_s))dC_s + \int_0^t\int \Phi(s,x,u_{s-})\left(1 - \frac{Y(s,x,u_{s-})}{Y(s,x,\theta)}\right)\nu_\theta(ds,dx).$$

Finally, the recursive SDE (1.4.8) takes the form

$$\theta_t = \theta_0 + \int_0^t \gamma_s(\theta_{s-})\dot\beta_s(\theta_{s-})d(X^c - \beta(\theta)\circ C)_s + \int_0^t\int \gamma_s(\theta_{s-})\Phi(s,x,\theta_{s-})(\mu - \nu_\theta)(ds,dx)$$
$$+ \int_0^t \gamma_s(\theta)\dot\beta_s(\theta_s)(\beta_s(\theta) - \beta_s(\theta_s))dC_s + \int_0^t\int \gamma_s(\theta_{s-})\Phi(s,x,\theta_{s-})\left(1 - \frac{Y(s,x,\theta_{s-})}{Y(s,x,\theta)}\right)\nu_\theta(ds,dx). \tag{1.4.11}$$

1.4.1. Remark. One can give sufficient conditions more accurate than (1.4.9) (see, e.g., [12, 14, 25]) to ensure the fulfillment of decomposition (1.4.10).

Assume that there exists a unique strong solution $(\theta_t)_{t\geq 0}$ of the SDE (1.4.11).

To investigate the asymptotic properties of recursive estimators $(\theta_t)_{t\geq 0}$ as $t \to \infty$, precisely, the strong consistency, the rate of convergence, and the asymptotic expansion, we reduce the SDE (1.4.11) to the Robbins–Monro type SDE.

For this purpose, denote $z_t = \theta_t - \theta$. Then (1.4.11) can be rewritten as

$$z_t = z_0 + \int_0^t \gamma_s(\theta + z_{s-})\dot\beta(\theta + z_{s-})(\beta_s(\theta) - \beta_s(\theta + z_{s-})dC_s$$
$$+ \int_0^t\int \gamma_s(\theta + z_{s-})\Phi(s,x,\theta + z_{s-})\left(1 - \frac{Y(s,x,\theta + z_{s-})}{Y(s,x,\theta)}\right)\nu_\theta(ds,dx)$$
$$+ \int_0^t \gamma_s(\theta + z_s)\dot\beta_s(\theta + z_s)d(X^c - \beta(\theta)\circ C)_s$$
$$+ \int_0^t\int \gamma_s(\theta + z_{s-})\Phi(s,x,\theta + z_{s-})(\mu - \nu_\theta)(ds,dx). \tag{1.4.12}$$

For the definition of the objects K^θ, $\{H^\theta(u),\, u\in\mathbb{R}^1\}$, and $\{M^\theta(u),\, u\in\mathbb{R}^1\}$ we consider a version of characteristics (C, ν_θ) such that

$$C_t = C^\theta \circ A_t^\theta, \quad \nu_\theta(\omega, dt, dx) = dA_t^\theta B_{\omega,t}^\theta(dx),$$

where $A^\theta = (A_t^\theta)_{t\geq 0} \in \mathcal{A}^+_{\rm loc}(P_\theta)$, $C^\theta = (C_t^\theta)_{t\geq 0}$ is a nonnegative predictable process, $B^\theta_{\omega,t}(dx)$ is a transition kernel from $(\Omega\times\mathbb{R}_+,\mathcal{P})$ in $(R,\mathcal{B}(R))$ with $B^\theta_{\omega,t}(\{0\})=0$, and

$$\Delta A_t^\theta B^\theta_{\omega,t}(R)\leq 1$$

(see [14, Chap. 2, Sec. 2, Proposition 2.9]).

Put $K_t^\theta = A_t^\theta$,

$$\begin{aligned}H_t^\theta(u) &= \gamma_t(\theta+u)\bigg\{\dot\beta_t(\theta+u)(\beta_t(\theta)-\beta_t(\theta+u))C_t^\theta \\ &\quad + \int \phi(t,x,\theta+u)\bigg(1-\frac{Y(t,x,\theta+u)}{Y(t,x,\theta)}\bigg)B^\theta_{\omega,t}(dx)\bigg\}, \qquad (1.4.13)\\ M^\theta(t,u) &= \int_0^t \gamma_s(\theta+u)\dot\beta_s(\theta+u)d(X^c-\beta(\theta)\circ C)_s \\ &\quad + \int_0^t\int \gamma_s(\theta+u)\Phi(s,x,\theta+u)(\mu-\nu_\theta)(ds,dx). \qquad (1.4.14)\end{aligned}$$

Assume that for each u, $M^\theta(u)=(M^\theta(t,u))_{t\geq 0}\in\mathcal{M}^2_{\rm loc}(P_\theta)$. Then

$$\begin{aligned}\langle M^\theta(u)\rangle_t &= \int_0^t (\gamma_s(\theta+u)\dot\beta_s(\theta+u))^2 C_s^\theta dA_s^\theta + \int_0^t \gamma_s^2(\theta+u)\bigg(\int \Phi^2(s,x,\theta+u)B^\theta_{\omega,s}(dx)\bigg)dA_s^{\theta,c} \\ &\quad + \int_0^t \gamma_s^2(\theta+u)B^\theta_{\omega,t}(R)\bigg\{\int\Phi^2(s,x,\theta+u)q^\theta_{\omega,s}(dx) - a_s(\theta)\bigg(\int\Phi(s,x,\theta+u)q^\theta_{\omega,s}(dx)\bigg)^2\bigg\}dA_s^{\theta,d},\end{aligned}$$

where $a_s(\theta)=\Delta A_s^\theta B^\theta_{\omega,s}(R)$, $q^\theta_{\omega,s}(dx)I_{\{a_s(\theta)>0\}} = \frac{B^\theta_{\omega,s}(dx)}{B^\theta_{\omega,s}(R)}I_{\{a_s(\theta)>0\}}$.

Now we give a more detailed description of $\Phi(\theta)$, $\widehat{I}(\theta)$, $H^\theta(u)$, and $\langle M^\theta(u)\rangle$. Denote

$$\frac{d\nu_\theta^c}{d\nu^c} := F(\theta), \quad \frac{q^\theta_{\omega,t}(dx)}{q_{\omega,t}(dx)} := f_{\omega,t}(x,\theta) \quad (:= f_t(\theta)).$$

Then

$$Y(\theta) = F(\theta)I_{\{a=0\}} + \frac{a(\theta)}{a}f(\theta)I_{\{a>0\}}$$

and

$$\dot Y(\theta) = \dot F(\theta)I_{\{a=0\}} + \bigg(\frac{\dot a(\theta)}{a}f(\theta)+\frac{a(\theta)}{a}\dot f(\theta)\bigg)I_{\{a>0\}}.$$

Therefore,

$$\Phi(\theta) = \frac{\dot F(\theta)}{F(\theta)}I_{\{a=0\}} + \bigg\{\frac{\dot f(\theta)}{f(\theta)} + \frac{\dot a(\theta)}{a(\theta)(1-a(\theta))}\bigg\}I_{\{a>0\}} \qquad (1.4.15)$$

with

$$I_{\{a(\theta)>0\}}\int\frac{\dot f(\theta)}{f(\theta)}q^\theta(dx)=0.$$

Denote

$$\dot\beta(\theta)=\ell^c(\theta), \quad \frac{\dot F(\theta)}{F(\theta)} := \ell^\pi(\theta), \quad \frac{\dot f(\theta)}{f(\theta)} := \ell^\delta(\theta), \quad \frac{\dot a(\theta)}{a(\theta)(1-a(\theta))} := \ell^b(\theta).$$

The indices $i = c$, π, δ, and b have the following meaning: "c" corresponds to the continuous part, "π" to the Poisson type part, "δ" to the predictable moments of jumps (including a main special case, the discrete time case), and "b" to the binomial type part of the likelihood score $\ell(\theta) = (\ell^c(\theta), \ell^\pi(\theta), \ell^\delta(\theta), \ell^b(\theta))$.

In this notation, for the Fisher information process we have

$$\widehat{I}_t(\theta) = \int_0^t (\ell_s^c(\theta))^2 dC_s + \int_0^t \int (\ell_s^\pi(x;\theta))^2 B_{\omega,s}^\theta(dx) dA_s^{\theta,c} + \int_0^t B_{\omega,s}^\theta(R)\Big[\int (\ell_s^\delta(x;\theta))^2 q_{\omega,s}^\theta(dx)\Big] dA_s^{\theta,d} + \int_0^t (\ell_s^b(\theta))^2(1-a_s(\theta)) dA_s^{\theta,d}. \quad (1.4.16)$$

For the random field $H^\theta(u)$, we have

$$H_t^\theta(u) = \gamma_t(\theta+u)\Big\{\ell_t^c(\theta+u)(\beta_t(\theta)-\beta_t(\theta+u))C_t^\theta + \int \ell_t^\pi(x;\theta+u)\Big(1-\frac{F_t(x;\theta+u)}{F_t(x;\theta)}\Big) B_{\omega,t}^\theta(dx) I_{\{\Delta A_t^\theta=0\}} + \Big\{\int \ell_t^\delta(x;\theta+u) q_{\omega,t}^\theta(dx) + \ell_t^b(\theta+u)\,\frac{a_t(\theta)-a_t(\theta+u)}{a_t(\theta)}\Big\} B_{\omega,t}^\theta(R) I_{\{\Delta A_t^\theta>0\}}. \quad (1.4.17)$$

Finally, for $\langle M^\theta(u)\rangle$, we have

$$\langle M^\theta(u)\rangle_t = (\gamma(\theta+u)\ell^c(\theta+u))^2\, C^\theta \circ A_t^\theta + \int_0^t \gamma_s^2(\theta+u)\int (\ell_s^\pi(x;\theta+u))^2 B_{\omega,t}^\theta(dx) dA_s^{\theta,c} + \int_0^t \gamma_s^2(\theta+u) B_{\omega,s}^\theta(R)\Big\{\int (\ell_s^\delta(x;\theta+u)+\ell_s^b(\theta+u))^2 q_{\omega,s}^\theta(dx) - a_s(\theta)\Big(\int (\ell_s^\delta(x;\theta+u)+\ell_s^b(\theta+u)) q_{\omega,s}^\theta(dx)\Big)^2\Big\} dA_s^{\theta,d}. \quad (1.4.18)$$

Thus, we reduced SDE (1.4.12) to the Robbins–Monro type SDE with $K_t^\theta = A_t^\theta$, and $H^\theta(u)$ and $M^\theta(u)$ defined by (1.4.17) and (1.4.14), respectively.

As follows from (1.4.17)

$$H_t^\theta(0) = 0 \quad \text{for all} \quad t \ge 0, \quad P_\theta\text{-a.s.}$$

For condition **(A)** to be satisfied, it suffices to assume that for all $t \ge 0$, $u \ne 0$ P_θ-a.s.,

$$\dot\beta_t(\theta+u)(\beta_t(\theta)-\beta_t(\theta+u)) < 0,$$

$$\Big(\int \frac{\dot F(t,x,\theta+u)}{F(t,x,\theta+u)}\Big(1-\frac{F(t,x;\theta+u)}{F(t,x;\theta)}\Big) B_{\omega,t}^\theta(dx)\Big) I_{\{\Delta A_t^\theta=0\}} u < 0,$$

$$\Big(\int \frac{\dot f(t,x;\theta+u)}{f(t,x;\theta+u)}\, q_t^\theta(dx)\Big) I_{\{\Delta A_t^\theta>0\}} u < 0,$$

$$\dot a_t(\theta+u)(a_t(\theta)-a_t(\theta+u))u < 0,$$

and the simplest sufficient conditions for the latter is the monotonicity (P-a.s.) of the functions $\beta(\theta)$, $F(\theta)$, $f(\theta)$ and $a(\theta)$ with respect to θ.

1.4.2. Remark. In the case where the model is only regular in the Jacod sense, we save the same form of all the above-given objects (precisely of $\Phi(\theta)$) using the formal definitions:

$$\frac{\dot F(\theta)}{F(\theta)}\, I_{\{a(\theta)=0\}} := W(\theta) I_{\{a(\theta)=0\}},$$

$$\dot a(\theta) := \widehat{W}^\theta,$$

$$\frac{\dot f(\theta)}{f(\theta)} := W(\theta) I_{\{a(\theta)>0\}} - \frac{\widehat{W}^\theta(\theta)}{a(\theta)} I_{\{a(\theta)>0\}}.$$

1.4.2. Example. Discrete time.

(a) *Recursive MLE in parameter statistical models* Let $X_0, X_1, \dots, X_n, \dots$ be observations taking values in some measurable space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ such that the regular conditional densities of distributions (with respect to some measure μ) $f_i(x_i, \theta | x_{i-1}, \dots, x_0)$, $i \le n$, $n \ge 1$ exist, and $f_0(x_0, \theta) \equiv f_0(x_0)$, $\theta \in \mathbb{R}^1$, is the parameter to be estimated. Denote P_θ corresponding distribution on $(\Omega, \mathcal{F}) := (\mathcal{X}^\infty, \mathcal{B}(\mathcal{X}^\infty))$. Identify the process $X = (X_i)_{i\ge 0}$ with coordinate process and denote $\mathcal{F}_0 = \sigma(X_0)$ and $\mathcal{F}_n = \sigma\ (X_i,\ i \le n)$. If $\psi = \psi(X_i, X_{i-1}, \dots, X_0)$ is a r.v., then by $E_\theta(\psi | \mathcal{F}_{i-1})$ we mean the following version of conditional expectation:

$$E_\theta(\psi \mid \mathcal{F}_{i-1}) := \int \psi(z, X_{i-1}, \dots, X_0) f_i(z, \theta \mid X_{i-1}, \dots, X_0) \mu(dz)$$

if the last integral exists.

Assume that the usual regularity conditions are satisfied and denote

$$\frac{\partial}{\partial\theta} f_i(x_i, \theta \mid x_{i-1}, \dots, x_0) := \dot f_i(x_i, \theta \mid x_{i-1}, \dots, x_0),$$

by

$$l_i(\theta) := \frac{\dot f_i}{f_i} (X_i, \theta \mid X_{i-1}, \dots, X_0)$$

the maximum likelihood scores, and by

$$I_n(\theta) := \sum_{i=1}^n E_\theta(l_i^2(\theta) \mid \mathcal{F}_{i-1})$$

the empirical Fisher information. Also denote

$$b_n(\theta, u) := E_\theta(l_n(\theta + u) \mid \mathcal{F}_{n-1})$$

and indicate that for each $\theta \in \mathbb{R}^1$, $n \ge 1$

$$b_n(\theta, 0) = 0 \quad (P_\theta\text{-a.s.}). \tag{1.4.19}$$

Consider the following recursive procedure:

$$\theta_n = \theta_{n-1} + I_n^{-1}(\theta_{n-1}) l_n(\theta_{n-1}), \quad \theta_0 \in \mathcal{F}_0.$$

Fix θ, denote $z_n = \theta_n - \theta$, and rewrite the last equation in the form

$$\begin{aligned} z_n &= z_{n-1} + I_n^{-1}(\theta + z_{n-1}) b_n(\theta, z_{n-1}) + I_n^{-1}(\theta + z_{n-1}) \Delta m_n, \\ z_0 &= \theta - \theta, \end{aligned} \tag{1.4.20}$$

where $\Delta m_n = \Delta m(n, z_{n-1})$ with $\Delta m(n, u) = l_n(\theta + u) - E_\theta(l_n(\theta + u) | \mathcal{F}_{n-1})$.

Note that algorithm (1.4.20) is embedded in the stochastic approximation scheme (1.2.1) with

$$H_n(u) = I_n^{-1}(\theta + u) b_n(\theta, u) \in \mathcal{F}_{n-1}, \quad \Delta K_n = 1,$$
$$\Delta M(n, u) = I_n^{-1}(\theta + u) \Delta m(n, u).$$

This example clearly shows the necessity of considering the random fields $H_n(u)$ and $M(n, u)$.

In [39], the convergence $z_n \to 0$ P-a.s. as $n \to \infty$ was proved under conditions equivalent to **(A)**, **(B)**, and **(I)** connected with the standard representation (1.2.2)(1).

1.4.3. Remark. Let $\theta \in \Theta \subset \mathbb{R}^1$, where θ is an open proper subset of $\mathbb{R}^1$. It is possible that the objects $l_n(\theta)$ and $I_n(\theta)$ are defined only on the set Θ, but for each fixed $\theta \in \Theta$ the objects $H_n(u)$ and $M(n,u)$ are well-defined functions of variable u on all $\mathbb{R}^1$. Then under the conditions of Theorem 1.2.1, $\theta_n \to \theta$ P_θ-a.s. as $n \to \infty$ starting from an arbitrary θ_0. The example given below illustrates this situation. The same example also illustrates the efficiency of representation (1.2.3)(2).

(b) *Galton–Watson branching process with immigration.* Let the observable process be

$$X_i = \sum_{j=1}^{X_{i-1}} Y_{i,j} + 1, \quad i = 1, 2, \dots, n; \quad X_0 = 1,$$

$Y_{i,j}$ are i.i.d. random variables having the Poisson distribution with the parameter θ, $\theta > 0$, to be estimated. If $\mathcal{F}_i = \sigma(X_j,\, j \le i)$, then

$$P_\theta(X_i = m \mid \mathcal{F}_{i-1}) = \frac{(\theta X_{i-1})^{m-1}}{(m-1)!}\, e^{-\theta X_{i-1}}, \quad i = 1, 2, \dots; \quad m \ge 1.$$

From this, we have

$$l_i(\theta) = \frac{X_i - 1 - \theta X_{i-1}}{\theta}, \quad I_n(\theta) = \theta^{-1} \sum_{i=1}^n X_{i-1}.$$

The recursive procedure has the form

$$\theta_n = \theta_{n-1} + \frac{X_n - 1 - \theta_{n-1} X_{n-1}}{\sum_{i=1}^n X_{i-1}}, \quad \theta_0 \in \mathcal{F}_0, \tag{1.4.21}$$

and if, as usual $z_n = \theta_n - \theta$, then

$$z_n = z_{n-1} - \frac{z_{n-1} X_{n-1}}{\sum\limits_{i=1}^n X_{i-1}} + \frac{\varepsilon_n}{\sum\limits_{i=1}^n X_{i-1}}, \tag{1.4.22}$$

where $\varepsilon_n = X_n - 1 - \theta X_n$ is a P_θ-square integrable martingale-difference. In fact, $E_\theta(\varepsilon_n \,|\, \mathcal{F}_{n-1}) = 0$ and $E_\theta(\varepsilon_n^2 \,|\, \mathcal{F}_{n-1}) = \theta X_{n-1}$. In this case, $H_n(u) = -u X_{n-1} / \sum\limits_{i=1}^n X_{i-1}$, $\Delta M(n,u) = \Delta m_n = \varepsilon_n / \sum\limits_{i=1}^n X_{i-1}$, and $\Delta K = 1$ and, therefore, they are well defined on all $\mathbb{R}^1$.

Indicate that the solution of Eq. (1.4.21) coincides with MLE

$$\widehat{\theta}_n = \frac{\sum\limits_{i=1}^n (X_i - 1)}{\sum\limits_{i=1}^n X_{i-1}},$$

and it is easy to see that $(\widehat{\theta}_n)_{n \ge 1}$ is strongly consistent for all $\theta > 0$.

Indeed,

$$\widehat{\theta}_n - \theta = \frac{\sum\limits_{i=1}^n \varepsilon_i}{\sum\limits_{i=1}^n X_{i-1}},$$

and desirability follows from the strong law of large numbers for martingales and the well-known fact (see, e.g., [10]) that for all $\theta > 0$,

$$\sum_{i=1}^\infty X_{i-1} = \infty \quad (P_\theta\text{-a.s.}). \tag{1.4.23}$$

Derive this result as a consequence of Theorem 1.2.1.

First, note that for each $\theta > 0$, the conditions **(A)** and **(B′)** are satisfied. Indeed,

(A) $H_n(u)u = \dfrac{-u^2 X_{n-1}}{\sum\limits_{i=1}^{n} X_{i-1}} < 0$ for all $u \neq 0$ ($X_i > 0$, $i \geq 0$);

(B′) $\langle m\rangle_\infty = \theta \sum\limits_{n=1}^{\infty} \dfrac{X_{n-1}}{\left(\sum\limits_{i=1}^{n} X_{i-1}\right)^2} < \infty$ owing to (1.4.23).

Now to illustrate the efficiency of the group of conditions **(II)**, let us consider two cases: (1) $0 < \theta \leq 1$ and (2) θ is arbitrary, i.e., $\theta > 0$.

In case (1), conditions **(I)** are satisfied. In fact,

$$|H_n(u)| = \Big(X_{n-1}/\sum_{i=1}^{n} X_{i-1}\Big)|u|, \quad \sum_{n=1}^{\infty} X_{n-1}^2/\Big(\sum_{i=1}^{n} X_{i-1}\Big)^2 < \infty, \quad P_\theta\text{-a.s.}$$

But if $\theta > 1$, then the last series diverges and, therefore, the condition **(I)**(i) is not satisfied.

On the other hand, the proof of desirable convergence by verifying the conditions **(II)** is almost trivial. Indeed, use Remark 1.3.2 and take $\widetilde{G}_n = G_n = X_{n-1}/\sum\limits_{i=1}^{n} X_{i-1}$. Then $\sum\limits_{n=1}^{\infty} G_n = \infty$ P_θ-a.s. for all $\theta > 0$. Moreover, $\delta_n = -2 + \widetilde{G}_n < 0$, $|\delta_n| \geq 1$.

1.4.3. Example. RM algorithm with deterministic regression function.

Consider a particular case of algorithm (1.2.1) where $H_t(\omega, u) = \gamma_t(\omega)R(u)$ and the process $\gamma = (\gamma_t)_{t\geq 0} \in \mathcal{P}$, $\gamma_t > 0$ for all $t \geq 0$, $dM(t,u) = \gamma_t dm_t$, $m \in \mathcal{M}^2_{\text{loc}}$, i.e.,

$$dz_t = \gamma_t R(z_{t-})dK_t + \gamma_t dm_t, \quad z_0 \in \mathcal{F}_0.$$

(a) Let the following conditions be satisfied:

(A) $R(0) = 0$, $R(u)u < 0$ for all $u \neq 0$;

(B′) $\gamma^2 \circ \langle m\rangle_\infty < \infty$;

(1) $|R(u)| \leq C(1 + |u|)$, $C > 0$ is constant;

(2) for each $\varepsilon > 0$, $\inf\limits_{\varepsilon \leq u \leq \frac{1}{\varepsilon}} |R(u)| > 0$;

(3) $\gamma \circ K_t < \infty$, $\forall t \geq 0$, $\gamma \circ K_\infty = \infty$;

(4) $\gamma^2 \Delta K \circ K^d_\infty < \infty$.

Then $z_t \to 0$ P-a.s. as $t \to \infty$.

Indeed, it is easy to see that conditions **(A)**, **(B′)**, and (1)–(4) imply **(A)**, **(B)**, and **(I)** of Theorem 1.2.1.

In [28], this result was proved on the basis of the theorem on the semimartingale convergence sets noted in Remark 1.1.1. In the case where $K^d \neq 0$, this automatically leads to the "moment" restrictions and the additional assumption $|R(u)| \leq \text{const}$.

(b) As in case (a), let conditions **(A)** and **(B′)** be satisfied. Moreover, assume that for each $u \in \mathbb{R}^1$ and $t \in [0, \infty)$,

(1′) $V_t^-(u) + V_t^+(u) \leq 0$;

(2′) for all $\varepsilon > 0$,

$$I_\varepsilon := \inf_{\varepsilon \leq u \leq \frac{1}{\varepsilon}} \{-(V^-(u) + V^+(u))\} \circ K_\infty = \infty.$$

Then $z_t \to 0$ P-a.s. as $t \to \infty$.

Indeed, it is not difficult to verify that (1′), (2′)$\Rightarrow$ **(II)**.

The following question arises: is it possible that (1′) and (2′) can be satisfied? In addition, assume that

$$C_1|u| \leq |R(u)| \leq C_2|u|, \quad C_1, C_2 \quad \text{are constants,} \tag{1.4.24}$$

(3′) $2 - C_2\gamma_t\Delta K_t \geq 0$;

(4′) $\gamma(2 - C_2\gamma\Delta K) \circ K_\infty = \infty$.
Then (3′) ⇒ (1′) and (4′) ⇒ (2′).
Indeed,

$$V_t^-(u) + V_t^+(u) \le C_1\gamma_t|u|^2[-2 + C_2\gamma_t\Delta K_t] \le 0,$$
$$I_\varepsilon \ge C_1\varepsilon^2\{\gamma(2 - C_2\gamma\Delta K) \circ K_\infty\} = \infty.$$

1.4.4. Remark. (4′) ⇒ $\gamma \circ K_\infty = \infty$.

In [30], the convergence $z_t \to 0$ P-a.s. as $t \to \infty$ was proved under the following conditions:

(A) $R(0) = 0$, $R(u)u < 0$ for all $u \neq 0$;

(M) there exists a nonnegative predictable process $r = (r_t)_{t\ge0}$ integrable with respect to the process $K = (K_t)_{t\ge0}$ on any finite interval $[0, t]$ with the following properties:
- (a) $r \circ K_\infty = \infty$,
- (b) $A^1_\infty = \gamma^2\varepsilon^{-1}(-r \circ K) \circ \langle m\rangle_\infty < \infty$,
- (c) all jumps of process A^1 are bounded,
- (d) $r_tu^2 + \gamma_t^2\Delta K_tR^2(u) \le -2\gamma_tR(u)u$, for all $u \in \mathbb{R}^1$ and $t \in [0, \infty)$.

Show that **(M)**⇒**(B′)**, (1′), and (2′).

It is obvious that (b)⇒**(B′)**. Further, (d)⇒(1′), Finally, (2′) follows from (a) and (d) owing to the relation

$$I_\varepsilon := \inf_{\varepsilon\le|u|\le\frac{1}{\varepsilon}} -(V^-(u) + V^+(u)) \circ K_\infty \ge \varepsilon^2 r \circ K_\infty = \infty.$$

The implication is proved.

In the particular case where (1.4.24) holds and for all $t \ge 0$ $\gamma_t\Delta K_t \le q$, $q > 0$ is a constant and C_1 and C_2 in (1.4.24) are chosen such that $2C_1 - qC_2^2 > 0$, if we take $r_t = b\gamma_t$, $b > 0$, with $b < 2C_1 - qC_2^2$, then (a) and (d) are satisfied if $\gamma \circ K_\infty = \infty$.

But these conditions imply (3′) and (4′). In fact, on the one hand, $0 < 2C_1 - qC_2^2 \le C_1(2 - qC_2)$, and, therefore, (3′) follows, since $2 - C_2\gamma_t\Delta K_t \ge 2 - qC_2 > 0$. On the other hand, (4′) follows from $\gamma(2 - C_2\gamma\Delta K) \circ K_\infty \ge (2 - qC_2)\gamma \circ K_\infty = \infty$.

From the above, we may conclude that if conditions **(A)**, **(B′)**, (1.4.24), $\gamma_t\Delta K_t \le q$, $q > 0$, $2 - qC_2 > 0$, and $\gamma \circ K_\infty = \infty$ are satisfied, then the desirable convergence $z_t \to 0$ P-a.s. takes place, and, therefore, the choice of the process $r = (r_t)_{t\ge0}$ with properties (M) is not necessary (cf. Remark 1.2.3, Sec. 1.3, and [30]).

(c) *Linear model* (see, e.g., [28]). Consider the linear RM procedure

$$dz_t = b\gamma_tz_{t-}dK_t + \gamma_tdm_t, \quad z_0 \in \mathcal{F},$$

where $b \in B \subseteq (-\infty, 0)$, $m \in \mathcal{M}^2_{\text{loc}}$.
Assume that

$$\gamma^2 \circ \langle m\rangle_\infty < \infty, \tag{1.4.25}$$
$$\gamma \circ K_\infty = \infty, \tag{1.4.26}$$
$$\gamma^2\Delta K \circ K^d < \infty.$$

Then for each $b \in B$, conditions **(A)**, **(B′)**, and **(I)** are satisfied. Hence

$$z_t \to 0 \quad P\text{-a.s.} \quad \text{as} \quad t \to \infty. \tag{1.4.27}$$

Now let (1.4.25) and (1.4.26) be satisfied, but $P(\gamma^2\Delta K \circ K^d = \infty) > 0$.

At the same time, assume that $B = [b_1, b_2]$, $-\infty < b_1 \le b_2 < 0$, and for all $t > 0$, $\gamma_t\Delta K_t < |b_1|^{-1}$.

Then for each $b \in B$, (1.4.27) holds.

Indeed,

$$[V_t^-(u)I_{\{\Delta K_t\neq 0\}}+V_t^+(u)]^+=|b|\gamma_t u^2[-2+|b|\gamma_t\Delta K_t I_{\{\Delta K_t\neq 0\}}]^+ \\ \leq I_{\{\Delta K_t\neq 0\}}|b|\gamma_t u^2[-2+|b|\gamma_t\Delta K_t]^+=0,$$

and, therefore, **(II)**(i) is satisfied. On the other hand,

$$\inf_{\varepsilon\leq|u|\leq\frac{1}{\varepsilon}} u^2\{2\gamma|b|I_{\{\Delta K\neq 0\}}+b\gamma[2-|b|\gamma\Delta K]I_{\{\Delta K\neq 0\}}\}\circ K_\infty \\ \geq \varepsilon^2|b|\gamma[2-|b|\gamma\Delta K]\circ K_\infty\geq\varepsilon^2|b|\gamma\circ K_\infty=\infty.$$

Therefore, **(II)**(ii) also holds.

2. Rate of Convergence and Asymptotic Expansion

2.1. Notation and preliminaries. We consider the RM type stochastic differential equation (SDE)

$$z_t=z_0+\int_0^t H_s(z_{s-})dK_s+\int_0^t M(ds,z_{s-}). \tag{2.1.1}$$

As usual, we assume that there exists a unique strong solution $z=(z_t)_{t\geq 0}$ of Eq. (2.1.1) defined on the whole time interval $[0,\infty[$ and $\widetilde{M}=(\widetilde{M}_t)_{t\geq 0}\in\mathcal{M}^2_{\mathrm{loc}}(P)$, where $\widetilde{M}=\int_0^t M(ds,z_{s-})$ (see [8, 9, 13]).

Let us denote

$$\beta_t=-\lim_{u\to 0}\frac{H_t(u)}{u}$$

assuming that this limit exists and is finite for each $t\geq 0$ and define the random field

$$\beta_t(u)=\begin{cases}-\frac{H_t(u)}{u} & \text{if} \quad u\neq 0,\\ \beta_t & \text{if} \quad u=0.\end{cases}$$

It follows from **(A)** that for all $t\geq 0$ and $u\in\mathbb{R}^1$,

$$\beta_t\geq 0 \quad \text{and} \quad \beta_t(u)\geq 0 \quad P\text{-a.s.}$$

Further, rewrite Eq. (2.1.1) as

$$z_t=z_0-\int_0^t\beta_s z_{s-}I_{\{\beta_s\Delta K_s\neq 1\}}dK_s+\int_0^t M(ds,0)-\sum_{s\leq t}z_{s-}I_{\{\beta_s\Delta K_s=1\}} \\ +\int_0^t(\beta_s-\beta_s(z_{s-}))z_{s-}dK_s+\int_0^t(M(ds,z_{s-})-M(ds,0))$$

(we suppose that $M(\cdot,0)\not\equiv 0$).

Denote

$$\overline{\beta}_t=\beta_t I_{\{\beta_t\Delta K_t\neq 1\}},\quad \overline{R}_t^{(1)}=-\sum_{s\leq t}z_{s-}I_{\{\beta_s\Delta K_s=1\}},$$

$$\overline{R}_t^{(2)}=\int_0^t(\beta_s-\beta_s(z_{s-}))z_{s-}dK_s,\quad \overline{R}_t^{(3)}=\int_0^t(M(ds,z_{s-})-M(ds,0)).$$

In this notation,

$$z_t = z_0 - \int_0^t \overline{\beta}_s z_{s-} dK_s + \int_0^t M(ds,0) + \overline{R}_t,$$

where

$$\overline{R}_t = \overline{R}_t^{(1)} + \overline{R}_t^{(2)} + \overline{R}_t^{(3)}.$$

Solving this equation with respect to z, we have

$$z_t = \Gamma_t^{-1}\left(z_0 + \int_0^t \Gamma_s M(ds,0) + \int_0^t \Gamma_s d\overline{R}_s\right), \tag{2.1.2}$$

where

$$\Gamma_t = \varepsilon_t^{-1}(-\overline{\beta} \circ K).$$

Here, $\alpha \circ K_t = \int_0^t \alpha_s dK_s$ and $\varepsilon_t(A)$ is the Dolean exponential.

The process $\Gamma = (\Gamma_t)_{t \geq 0}$ is predictable (but not positive in general) and, therefore, the process $L = (L_t)_{t\geq 0}$ defined by

$$L_t = \int_0^t \Gamma_s M(ds,0)$$

belongs to the class $\mathcal{M}^2_{\mathrm{loc}}(P)$. It follows from Eq. (2.1.2) that

$$\chi_t z_t = \frac{L_t}{\langle L_t \rangle_t^{1/2}} + R_t,$$

where

$$\chi_t = \Gamma_t \langle L \rangle_t^{-1/2}, \quad R_t = \frac{z_0}{\langle L \rangle_t^{1/2}} + \frac{1}{\langle L \rangle_t^{1/2}} \int_0^t \Gamma_s d\overline{R}_s$$

and $\langle L \rangle$ is the shifted square characteristic of L, i.e., $\langle L \rangle_t := 1 + \langle L \rangle_t^{F,P}$.

This section is organized as follows. In Sec. 2.2 assuming that $z_t \to 0$ as $t \to \infty$ P-a.s., we give various sufficient conditions for the convergence

$$\gamma_t^{\delta} z_t^2 \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.} \tag{2.1.3}$$

for all $0 \leq \delta \leq \delta_0$, where $\gamma = (\gamma_t)_{t\geq 0}$ is a predictable increasing process and δ_0, $0 \leq \delta_0 \leq 1$, is some constant. There we also give a series of examples illustrating these results.

In Sec. 2.3 assuming that Eq. (2.1.3) holds with γ asymptotically equivalent to χ^2 (see the definition in Sec. 2.2), we study sufficient conditions for the convergence

$$R_t \xrightarrow{P} 0 \quad \text{as} \quad t \to \infty,$$

which implies the local asymptotic linearity of the solution.

We say that the process $\xi = (\xi_t)_{t\geq 0}$ *has some property eventually if for every* ω *in a set* Ω_0 *of* P *probability* 1, *the trajectory* $(\xi_t(\omega))_{t\geq 0}$ *of the process has this property on the set* $[t_0(\omega), \infty)$ *for some* $t_0(\omega) < \infty$.

Everywhere in this section, we assume that $z_t \to 0$ as $t \to \infty$ P-a.s.

2.2. Rate of convergence. Throughout this section we assume that $\gamma = (\gamma_t)_{t\geq 0}$ is a predictable increasing process such that P-a.s.

$$\gamma_0 = 1, \quad \gamma_\infty = \infty.$$

Also, assume that for each $u \in \mathbb{R}^1$, the processes $\langle M(u)\rangle$ and γ are locally absolutely continuous with respect to the process K and denote

$$h_t(u,v) = \frac{d\langle M(u), M(v)\rangle_t}{dK_t}, \quad g_t = \frac{d\gamma_t}{dK_t}$$

assuming that $g_t > 0$ for simplicity, and hence $I_{\{\Delta K_t \neq 0\}} = I_{\{\Delta\gamma_t \neq 0\}}$ P-a.s. for all $t > 0$.

In this section, we study the problem of the convergence

$$\gamma_t^\delta z_t \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.}$$

for all δ, $0 < \delta < \delta_0/2$, $0 < \delta_0 \leq 1$.

Note that consideration of the two control parameters δ and δ_0 substantially simplifies the application of the results and also clarifies their relation with the classical ones (see Examples 2.2.1 and 2.3.1).

We consider two approaches to this problem. The first approach is based on the results on the convergence sets of nonnegative semimartingales and on the so-called "nonstandard representations."

The second approach presented exploits the stochastic version of the Kronecker lemma. This approach is employed in [39] for the discrete time case under the assumption (2.2.23). A comparison of the results obtained in this section with those obtained before is also presented.

We also note that these two approaches give different sets of conditions in general. This fact is illustrated by the various examples.

Let us formulate some auxiliary results based on the convergence sets.

Suppose that $r = (r_t)_{t\geq 0}$ is a nonnegative predictable process such that

$$r_t \Delta K_t < 0, \quad r \circ K_t < \infty \quad P\text{-a.s.}$$

for each $t > 0$ and

$$r \circ K_\infty = \infty \quad P\text{-a.s.}$$

Denote by $\varepsilon_t = \varepsilon_t(-r \circ K)$ the Dolean exponential, i.e.,

$$\varepsilon_t = e^{-\int_0^t r_s dK_s^c} \prod_{s\leq t} (1 - r_s \Delta K_s).$$

Then, as is well known (see [25, 28]), the process $\varepsilon_t^{-1} = \{\varepsilon_t(-r\circ K)\}^{-1}$ is a solution of the linear SDE

$$\varepsilon_t^{-1} = \varepsilon_t^{-1} r_t dK_t, \quad \varepsilon_0^{-1} = 1,$$

and $\varepsilon_t^{-1} \to \infty$ as $t \to \infty$ (P-a.s.).

2.2.1. Proposition. *Suppose that*

$$\int_0^\infty \varepsilon_t^{-1}\varepsilon_{t-}[r_t - 2\beta_t(z_{t-}) + \beta_t^2(z_{t-})\Delta K_t]^+ dK_t < \infty \quad P\text{-a.s.}, \tag{2.2.1}$$

$$\int_0^\infty \varepsilon_t^{-1} h_t(z_{t-}, z_{t-}) dK_t < \infty \quad P\text{-a.s.}, \tag{2.2.2}$$

where $[x]^+$ *denotes the positive part of* x. *Then* $\varepsilon^{-1} z^2 \to$ P-a.s. (*the notation* $X \to$ *means that* $X = (X_t)_{t\geq 0}$ *has a finite limit as* $t \to \infty$).

Proof. Using the Ito formula, we have

$$d(\varepsilon_t^{-1}z_t^2) = z_{t-}^2 d\varepsilon_t^{-1} + \varepsilon_t^{-1}dz_t^2 = \varepsilon_t^{-1}z_{t-}^2(r_t - 2\beta_t(z_{t-}) + \beta_t^2(z_{t-})\Delta K_t)dK_t$$
$$+ \varepsilon_t^{-1}h_t(z_{t-}, z_{t-})dK_t + d(\mathrm{Mart}) = \varepsilon_t^{-1}z_{t-}^2 dB_t + dA_t^1 - dA_t^2 + d(\mathrm{Mart}),$$

where

$$dB_t = \varepsilon_t^{-1}\varepsilon_{t-}\left[r_t - 2\beta_t(z_{t-}) + \beta_t^2(z_{t-})\Delta K_t\right]^+ dK_t,$$
$$dA_t^1 = \varepsilon_t^{-1}h_t(z_{t-}, z_{t-})dK_t,$$
$$dA_t^2 = \varepsilon_t^{-1}\varepsilon_{t-}\left[r_t - 2\beta_t(z_{t-}) + \beta_t^2(z_{t-})\Delta K_t\right]^- dK_t.$$

Now, applying Corollary 1.1.4 to the nonnegative semimartingale $(\varepsilon_t^{-1}z_t^2)_{t\geq 0}$, we obtain

$$\{B_\infty < \infty\} \cap \{A_\infty^1 < \infty\} \subseteq \{\varepsilon^{-1}z^2 \to\} \cap \{A_\infty^2 < \infty\}$$

and the result follows from Eqs. (2.2.1) and (2.2.2). □

The following lemma is an immediate consequence of the Ito formula applied to the process $(\gamma_t^\delta)_{t\geq 0}$, $0 < \delta < 1$.

2.2.1. Lemma. *Suppose that* $0 < \delta < 1$. *Then*

$$\gamma_t^\delta = \varepsilon_t^{-1}(-r^\delta \circ K),$$

where

$$r_t^\delta = \overline{r}_t^\delta g_t/\gamma_t, \quad \overline{r}_t^\delta = \delta I_{\{\Delta\gamma_t = 0\}} + \frac{1 - (1 - \Delta\gamma_t/\gamma_t)^\delta}{\Delta\gamma_t/\gamma_t} I_{\{\Delta\gamma_t \neq 0\}}.$$

The following theorem is the main result based on the first approach.

2.2.1. Theorem. *Suppose that for each* δ, $0 < \delta < \delta_0$, $0 < \delta_0 \leq 1$,

$$\int_0^\infty \left(\frac{\gamma_{t-}}{\gamma_t}\right)^{-\delta} [r_t^\delta - 2\beta_t(z_{t-}) + \beta_t^2(z_{t-})\Delta K_t]^+ dK_t < \infty \quad P\text{-a.s.} \tag{2.2.3}$$

and

$$\int_0^\infty \gamma_t^\delta h_t(z_{t-}, z_{t-})dK_t < \infty \quad P\text{-a.s.} \tag{2.2.4}$$

Then $\gamma_t^\delta z_t^2 \to 0$ *as* $t \to \infty$ P-a.s. *for each* δ, $0 < \delta < \delta_0$, $0 < \delta_0 \leq 1$.

Proof. It follows from Proposition 2.2.1, Lemma 2.2.1 and the conditions (2.2.3) and (2.2.4) that

$$P\{\gamma^\delta z^2 \to\} = 1$$

for all δ, $0 < \delta < \delta_0$, $0 < \delta_0 \leq 1$. Now the result follows since

$$\{\gamma^\delta z^2 \to \quad \text{for all} \quad \delta,\ 0 < \delta < \delta_0\} \Rightarrow \{\gamma^\delta z^2 \to 0 \quad \text{for all} \quad \delta,\ 0 < \delta < \delta_0\}.$$

□

2.2.1. Remark. Note that if Eq. (2.2.3) holds for $\delta = \delta_0$, than it holds for all $\delta \leq \delta_0$.

Some simple conditions ensuring Eq. (2.2.3) are given in the following corollaries.

2.2.1. Corollary. *Suppose that the process*

$$\frac{\gamma}{\gamma_-} \tag{2.2.5}$$

is eventually bounded. Then for each δ, $0<\delta<\delta_0$, $0<\delta_0\le 1$,

$$\begin{aligned}
&\left\{\left[(\delta I_{\{\Delta\gamma=0\}} + I_{\{\Delta\gamma\neq 0\}}\,\frac{g}{\gamma} - 2\beta(z_-) + \beta^2(z_-)\Delta K\right]^+ \circ K_\infty < \infty\right\}\\
&\qquad\subseteq \left\{\left[\left(\delta + (1-\delta)\,\frac{\Delta\gamma}{\gamma}\right)\frac{g}{\gamma} - 2\beta(z_-) + \beta^2(z_-)\Delta K\right]^+ \circ K_\infty < \infty\right\}\\
&\qquad\qquad\subseteq \left\{\left(\frac{\gamma_-}{\gamma}\right)^{-\delta}[r^\delta - 2\beta(z_-) + \beta^2(z_-)\Delta K]^+ \circ K_\infty < \infty\right\}.
\end{aligned}$$

Proof. The proof immediately follows from the following simple inequalities

$$1-(1-x)^\delta \le \delta x + (1-\delta)x^2 \le x$$

if $0<x<1$ and $0<\delta<1$, which with $x=\Delta\gamma_t/\gamma_t$ yields

$$\overline{r}_t^\delta \le \left(\delta + (1-\delta)\,\frac{\Delta\gamma_t}{\gamma_t}\right) \le \big(\delta I_{\{\Delta\gamma_t=0\}} + I_{\{\Delta\gamma_t\neq 0\}}\big).$$

It remains to apply condition (2.2.5). □

In the next corollary, we need the following group of conditions: for δ, $0<\delta<\delta_0/2$,

$$\left[\delta\,\frac{g}{\gamma} - \beta(z)\right]^+ \circ K^c_\infty < \infty \quad P\text{-a.s.}; \tag{2.2.6}$$

$$\sum_{t\ge 0}\left[(1-\beta_t(z_{t-})\Delta K_t - \left(1-\frac{\Delta\gamma_t}{\gamma_t}\right)^\delta\right]^+ I_{\{\beta_t(z_{t-})\Delta K_t\le 1\}} < \infty \quad P\text{-a.s.}; \tag{2.2.7}$$

$$\sum_{t\ge 0}\left[(\beta_t(z_{t-})\Delta K_t - 1 - \left(1-\frac{\Delta\gamma_t}{\gamma_t}\right)^\delta\right]^+ I_{\{\beta_t(z_{t-})\Delta K_t\ge 1\}} < \infty \quad P\text{-a.s.} \tag{2.2.8}$$

2.2.2. Corollary. *Suppose that the process*

$$(\beta_t(z_{t-})\Delta K_t)_{t\ge 0} \tag{2.2.9}$$

is eventually bounded. If Eq. (2.2.5) *holds, then*

(1) $\{$(2.2.6), (2.2.7), *and* (2.2.8) *for all* δ, $0<\delta<\delta_0/2\} \Rightarrow \{$(2.2.3) *for all* δ, $0<\delta<\delta_0\}$;

(2) *if, in addition, the process* $\xi=(\xi_t)_{t\ge 0}$, *with* $\xi_t=\sup_{s\ge t}(\Delta\gamma_s/\gamma_s)$ *is eventually* <1, *then the reverse implication "$\Leftarrow$" holds in* (1);

(3) $\{$(2.2.6), (2.2.7), (2.2.8) *for* $\delta=\delta_0/2\} \Rightarrow \{$(2.2.6), (2.2.7), (2.2.8) *for all* δ, $0<\delta<\delta_0/2\}$ (*here,* δ_0 *is some fixed constant with* $0<\delta_0\le 1$).

Proof. By simple calculations, for all δ, $0<\delta<\delta_0$, $0<\delta_0\le 1$, we have

$$\begin{aligned}
&\int_0^\infty \left(\frac{\gamma_{t-}}{\gamma_t}\right)^{-\delta}\left[\left(\delta I_{\{\Delta\gamma_t=0\}} + \frac{1-(1-\Delta\gamma_t/\gamma_t)^\delta}{\Delta\gamma_t/\gamma_t}\,I_{\{\Delta\gamma_t\neq 0\}}\right)\frac{g_t}{\gamma_t} - 2\beta_t(z_{t-}) + \beta_t^2(z_{t-})\Delta K_t\right]^+ dK_t\\
&\qquad = \int_0^\infty \left[\delta\,\frac{g_t}{\gamma_t} - 2\beta_t(z_{t-})\right]^+ dK^c_t + \sum_{t\ge 0}\left(\frac{\gamma_{t-}}{\gamma_t}\right)^{-\delta}\left(1-\beta_t(z_{t-})\Delta K_t - \left(1-\frac{\Delta\gamma_t}{\gamma_t}\right)^{\delta/2}\right)\\
&\qquad\qquad \times\left[1-\beta_t(z_{t-})\Delta K_t + \left(1-\frac{\Delta\gamma_t}{\gamma_t}\right)^{\delta/2}\right]^+ I_{\{\beta_t(z_{t-})\Delta K_t\le 1\}}
\end{aligned}$$

$$+\sum_{t\ge 0}\left(\frac{\gamma_{t-}}{\gamma_t}\right)^{-\delta}\left(\beta_t(z_{t-})\Delta K_t - 1 + \left(1-\frac{\Delta\gamma_t}{\gamma_t}\right)^{\delta/2}\right)$$

$$\times\left[\beta_t(z_{t-})\Delta K_t - 1 - \left(1-\frac{\Delta\gamma_t}{\gamma_t}\right)^{\delta/2}\right]^{+} I_{\{\beta_t(z_{t-})\Delta K_t\ge 1\}}. \quad (2.2.10)$$

Now for the fulfillment of implications (1) and (2), it suffices to show that under conditions (2.2.5) and (2.2.9), the processes

$$\left(1-\beta(z_-)\Delta K + (1-\Delta\gamma/\gamma)^{\delta/2}\right) I_{\{\beta(z_-)\Delta K\le 1\}}$$

and

$$\left(\beta(z_-)\Delta K - 1 + (1-\Delta\gamma/\gamma)^{\delta/2}\right) I_{\{\beta(z_-)\Delta K\ge 1\}}$$

are eventually bounded, and, moreover, if $\xi < 1$ eventually, these processes are bounded from below by a strictly positive random constant. Indeed, for each $0<\delta<1$ and $t\ge 0$, if $\beta_t(z_{t-})\Delta K_t \le 1$, then

$$1-\sup_{s\ge t}\frac{\Delta\gamma_s}{\gamma_s} \le 1-\beta_t(z_{t-})\Delta K_t + (1-\Delta\gamma_t/\gamma_t)^{\delta/2} \le 2, \quad (2.2.11)$$

and if $\beta_t(z_{t-})\Delta K_t \ge 1$, then

$$1-\sup_{s\ge t}\frac{\Delta\gamma_s}{\gamma_s} \le \beta_t(z_{t-})\Delta K_t - 1 + (1-\Delta\gamma_t/\gamma_t)^{\delta/2} \le \beta_t(z_{t-})\Delta K_t. \quad (2.2.12)$$

The implication (3) simply follows from the inequality $(1-x)^{\delta} \le (1-x)^{1/2}$ if $0<x<1$ and $0<\delta<1/2$. □

The following result is an immediate consequence of Corollary 2.2.2.

2.2.3. Corollary. *Suppose that*

$$\sum_{t\ge 0} I_{\{\beta_t(z_{t-})\Delta K_t\ge 1\}} < \infty \quad \textit{and} \quad \sum_{t\ge 0}\left(\frac{\Delta\gamma_t}{\gamma_t}\right)^2 < \infty \quad P\text{-a.s.} \quad (2.2.13)$$

Then Eq. (2.2.7) *is equivalent to*

$$\int_0^{\infty}\left[\delta - \frac{\gamma_t\beta_t(z_{t-})}{\gamma_t}\right]^{+}\frac{d\gamma_t^d}{\gamma_t} < \infty \quad P\text{-a.s.} \quad (2.2.14)$$

and

$$\{(2.2.6),\ (2.2.14)\ \textit{for all}\ \delta,\ 0\le\delta\le\delta_0/2\} \Leftrightarrow \{(2.2.3)\ \textit{for all}\ \delta,\ 0<\delta<\delta_0\}.$$

Proof. Conditions (2.2.8) and (2.2.9) are automatically satisfied, and also $\xi<1$ eventually ($\xi=(\xi_t)_{t\ge 0}$ is the process with $\xi_t=\sup\limits_{s\ge t}(\Delta\gamma_s/\gamma_s)$). Therefore, it follows from Corollary 2.2.2 (2) that

$$\{(2.2.6),(2.2.7)\quad \text{for all}\quad \delta,\ 0<\delta<\delta_0/2\} \Rightarrow \{(2.2.3)\quad \text{for all}\quad \delta,\ 0<\delta<\delta_0\}.$$

It remains to prove that Eq. (2.2.7) is equivalent to Eq. (2.2.14). This immediately follows from the inequalities

$$[a+b]^+ \le [a]^+ + [b]^+, \quad \delta x \le 1-(1-x)^{\delta} \le \delta x + (1-\delta)x^2,$$
$$0<x<1, \quad 0<\delta<1,$$

applied to $x=(\Delta\gamma_s/\gamma_s)$ and to the expression

$$\left[1-\beta_t(z_{t-})\Delta K_t + (1-\Delta\gamma_t/\gamma_t)^{\delta}\right]^{+},$$

and from the condition $\sum_{t\ge 0}(\Delta\gamma_t/\gamma_t)^2 < \infty$ P-a.s. □

2.2.2. Remark. Condition (2.2.14) can be written as

$$\sum_{t\ge 0}\left[\delta\,\frac{\Delta\gamma_t}{\gamma_t} - \beta_t(z_{t-})\Delta K_t\right]^+ < \infty \quad P\text{-a.s.}$$

Below, using the stochastic version of the Kronecker lemma, we give an alternative group of conditions that ensure the convergence

$$\gamma_t^\delta z_t \to 0 \quad \text{as} \quad t\to\infty \quad P\text{-a.s.}$$

for all $0<\delta<\delta_0/2$, $0<\delta_0\le 1$.

Rewrite Eq. (2.1.1) in the form

$$z_t = z_0 + \int_0^t z_{s-}dB_s + G_t,$$

where

$$dB_t = -\overline{\beta}_t(z_{t-})dK_t, \quad \overline{\beta}_t(u) = \beta_t(u)I_{\{\beta_t(u)\Delta K_t\ne 1\}}$$

and

$$G_t = -\sum_{s\le t} z_{s-}I_{\{\beta_t(z_{t-})\Delta K_t = 1\}} + \int_0^t M(ds, z_{s-}). \tag{2.2.15}$$

Since $\Delta B_t = -\overline{\beta}_t(z_{t-})\Delta K_t \ne -1$, we can represent z as

$$z_t = \varepsilon_t(B)\left(z_0 + \int_0^t \varepsilon_s^{-1}(B)\,dG_s\right),$$

and multiplying this equation by γ_t^δ, we obtain

$$\gamma_t^\delta z_t = \operatorname{sign}\varepsilon_t(B)\Gamma_t^{(\delta)}\left(z_0 + \int_0^t \operatorname{sign}\varepsilon_s(B)\{\Gamma_s^{(\delta)}\}^{-1}\gamma_s^\delta\,dG_s\right), \tag{2.2.16}$$

where $\Gamma_t^{(\delta)} = \gamma_t^\delta|\varepsilon_t(B)|$.

2.2.1. Definition. We say that predictable processes $\xi = (\xi_t)_{t\ge 0}$ and $\eta = (\eta_t)_{t\ge 0}$ are equivalent as $t\to\infty$ and write $\xi \simeq \eta$ if there exists a process $\zeta = (\zeta_t)_{t\ge 0}$ such that

$$\xi_t = \zeta_t\eta_t,$$

and

$$0 < \zeta^1 < |\zeta| < \zeta^2 < \infty$$

eventually for some random constants ζ^1 and ζ^2.

The proof of the following result is based on the stochastic version of the Kronecker lemma.

2.2.2. Proposition. *Suppose that for all δ, $0<\delta<\delta_0/2$, $0<\delta_0\le 1$,*

(1) *there exists a positive decreasing predictable process $\overline{\Gamma}^{(\delta)} = (\overline{\Gamma}_t^{(\delta)})_{t\ge 0}$ such that*

$$\overline{\Gamma}_0^{(\delta)} = 1 \quad P\text{-a.s.}, \quad P\Big\{\lim_{t\to 0}\overline{\Gamma}_t^{(\delta)} = 0\Big\} = 1, \quad \Gamma^{(\delta)} \simeq \overline{\Gamma}^{(\delta)},$$

and

(2)

$$\sum_{t\geq 0} I_{\{\beta_t(z_{t-})\Delta K_t=1\}} < \infty \quad P\text{-a.s.,} \tag{2.2.17}$$

$$\int_0^\infty \gamma_t^{2\delta} h_t(z_{t-}, z_{t-})dK_t < \infty \quad P\text{-a.s.} \tag{2.2.18}$$

Then

$$\gamma_t^\delta z_t \to 0 \quad as \quad t \to \infty \quad P\text{-a.s.}$$

for all $0 < \delta < \delta_0/2$, $0 < \delta_0 \leq 1$.

Proof. Recall the stochastic version of the Kronecker lemma (see, e.g., [25, Chap. 2, Sec. 6]):

Kronecker lemma. *Suppose that* $X = (X_t)_{t\geq 0}$ *is s semimartingale and* $L = (L_t)_{t\geq 0}$ *is a predictable increasing process. Then*

$$\{L_\infty = \infty\} \cap \{Y \to\} \subseteq \left\{\frac{X}{L} \to 0\right\} \quad P\text{-a.s.,}$$

where $Y = (1 + L)^{-1} \cdot X$*.*

We set

$$(1 + L_t)^{-1} = \overline{\Gamma}_t^{(\delta)}, \quad X_t = \int_0^t (\Gamma_s^{(\delta)})^{-1} \operatorname{sign} \varepsilon_s(B)\gamma_s^\delta dG_s.$$

Then it follows from condition (1) that L is an increasing process with $L_\infty = \infty$ P-a.s. and

$$A = \{\overline{\Gamma}_\infty^{(\delta)} = 0\} \cap \left\{ \int_0^\cdot \overline{\Gamma}_s^{(\delta)}(\Gamma_s^{(\delta)})^{-1} \operatorname{sign} \varepsilon_s(B)\gamma_s^\delta dG_s \to \right\}$$

$$\subseteq \left\{ \frac{\overline{\Gamma}^{(\delta)}}{1 - \overline{\Gamma}^{(\delta)}} \int_0^\cdot (\Gamma_s^{(\delta)})^{-1} \operatorname{sign} \varepsilon_s(B)\gamma_s^\delta dG_s \to 0 \right\} \subseteq \{\gamma^\delta z \to 0\},$$

where the latter inequality follows from the relation $\overline{\Gamma}^{(\delta)} \simeq \Gamma^{(\delta)}$ and Eq. (2.2.16).

At the same time, from Eq. (2.2.15) and the well-known fact that if $M \in \mathcal{M}^2_{\mathrm{loc}}$, then $\{\langle M\rangle_\infty < \infty\} \subseteq \{M \to\}$ (see, e.g., [25]), we have

$$\{\overline{\Gamma}_\infty^{(\delta)} = 0\} \cap \left\{ \sum_{t\geq 0} I_{\{\beta_t(z_{t-})\Delta K_t=1\}} < \infty \right\} \cap \left\{ \int_0^\infty \gamma_t^{2\delta} h_t(z_{t-}, z_{t-})dK_t < \infty \right\} \subseteq A.$$

Now the result follows from Eqs. (2.2.17) and (2.2.18). □

Now we establish some simple results, which are useful for verifying condition (1) of Proposition 2.2.2.

By the definition of $\varepsilon_t(B)$,

$$\varepsilon_t(B) = e^{B_t^c} \prod_{s\leq t}(1 + \Delta B_s)$$

and since

$$\gamma_t^\delta = \exp\left(\delta \int_0^t \frac{d\gamma_s^c}{\gamma_s} - \sum_{s\leq t} \log\left(1 - \frac{\Delta\gamma_s}{\gamma_s}\right)^\delta\right),$$

we obtain

$$\Gamma_t^{(\delta)} = \exp\left(B_t^c + \delta\int\limits_0^t \frac{d\gamma_s^c}{\gamma_s} + \sum_{s\le t}\log\frac{|1+\Delta B_s|}{\left(1-\frac{\Delta\gamma_s}{\gamma_s}\right)^{\delta}}\right) = \exp\Big(-\int\limits_0^t D_s dC_s^{(\delta)}\Big), \tag{2.2.19}$$

where $D_t = 1/\gamma_t$ and

$$C_t^{(\delta)} = \int\limits_0^t\left(\left\{\frac{\beta_s(z_{s-})\gamma_s}{g_s} - \delta\right\} I_{\{\Delta\gamma_s=0\}} - \frac{\gamma_s}{\Delta\gamma_s}\log\frac{|1+\Delta B_s|}{\left(1-\frac{\Delta\gamma_s}{\gamma_s}\right)^{\delta}} I_{\{\Delta\gamma_s\neq 0\}}\right)d\gamma_s. \tag{2.2.20}$$

Integrating by parts

$$d(D_tC_t) = D_t dC_t + C_{t-}dD_t$$

and using the relation

$$d\left(\frac{1}{\gamma_t}\right) = -\frac{1}{\gamma_{t-}}\frac{d\gamma_t}{\gamma_t},$$

we obtain from Eq. (2.2.19) that

$$\Gamma_t^{(\delta)} = \exp\Big(-\frac{C_t^{(\delta)}}{\gamma_t} - \int\limits_0^t C_{s-}^{(\delta)}\frac{1}{\gamma_{s-}}\frac{d\gamma_s}{\gamma_s}\Big).$$

Therefore,

$$\Gamma_t^{(\delta)} = \zeta_t\overline{\Gamma}_t^{(\delta)}, \tag{2.2.21}$$

where

$$\overline{\Gamma}_t^{(\delta)} = \exp\Big(-\int\limits_0^t\left[\frac{C_{s-}^{(\delta)}}{\gamma_{s-}}\right]^+\frac{d\gamma_s}{\gamma_s}\Big), \quad \zeta_t = \exp\Big(-\frac{C_t^{(\delta)}}{\gamma_t} + \int\limits_0^t\left[\frac{C_{s-}^{(\delta)}}{\gamma_{s-}}\right]^+\frac{d\gamma_s}{\gamma_s}\Big).$$

The following proposition is an immediate consequence of Eq. (2.2.21).

2.2.3. Proposition. *Suppose that for each δ, $0<\delta<\delta_0/2$, $0<\delta_0\le 1$, the following conditions hold*:

(a) *There exist random constants $\underline{C}(\delta)$ and $\overline{C}(\delta)$ such that*

$$-\infty < \underline{C}(\delta) < \frac{C^{(\delta)}}{\gamma} < \overline{C}(\delta) < \infty$$

eventually, where $C^{(\delta)}/\gamma = (C_t^{(\delta)}/\gamma_t)_{t\ge 0}$.

(b) $$\int\limits_0^\infty\left[\frac{C_{t-}^{(\delta)}}{\gamma_{t-}}\right]^-\frac{d\gamma_t}{\gamma_t} < \infty \quad P\text{-a.s.}$$

(c) $$\int\limits_0^\infty\left[\frac{C_{t-}^{(\delta)}}{\gamma_{t-}}\right]^+\frac{d\gamma_t}{\gamma_t} = \infty \quad P\text{-a.s.}$$

Then $\Gamma^{(\delta)} \simeq \overline{\Gamma}^{(\delta)}$ for each δ, $0<\delta<\delta_0/2$.

2.2.4. Corollary. *Suppose that*

$$0 < \frac{C^{(\delta_0/2)}}{\gamma} < \frac{C^{(0)}}{\gamma} < \overline{C}(0) < \infty$$

eventually, where $\overline{C}(0)$ is some random constant and the processes $C^{(\delta_0/2)}$ and $C^{(0)}$ are defined in Eq. (2.2.20) for $\delta=\delta_0/2$ and $\delta=0$, respectively.

Then $\Gamma^{(\delta)} \simeq \overline{\Gamma}^{(\delta)}$ for each δ, $0<\delta<\delta_0/2$, $0<\delta_0\le 1$.

This result follows since, as is easy to verify,

$$C_t^{(\delta_0/2)} < C_t^{(\delta)} < C_t^{(0)} \quad \text{and} \quad C_t^{(\delta)} - C_t^{(\delta_0/2)} \geq \left(\frac{\delta_0}{2} - \delta\right)\gamma_t$$

for each δ, $0 < \delta < \delta_0/2$, which yields

$$\frac{\delta_0}{2} - \delta < \frac{C^{(\delta)}}{\gamma} < \overline{C}(0)$$

and

$$\left[\frac{C^{(\delta)}}{\gamma}\right]^+ > \frac{\delta_0}{2} - \delta \quad \text{and} \quad \left[\frac{C^{(\delta)}}{\gamma}\right]^- = 0$$

eventually.

We now formulate the main result of this approach, which is an immediate consequence of Propositions 2.2.2 and 2.2.3.

2.2.2. Theorem. *Suppose that conditions* (2.2.17), (2.2.18), *and the conditions of Proposition* 2.2.3 *hold for all* δ, $0 < \delta < \delta_0/2$, $0 < \delta_0 \leq 1$. *Then* P-a.s.,

$$\gamma_t^\delta z_t \to 0 \quad \textit{as} \quad t \to \infty$$

for all δ, $0 < \delta < \delta_0/2$, $0 < \delta_0 \leq 1$.

Consider the following two cases in more detail: (1) all the processes under consideration are continuous; (2) the discrete time case. In addition assume that $M(t, u) = M(t)$ for all $u \in \mathbb{R}^1$, $t \geq 0$.

In the case of continuous processes, conditions (2.2.7) and (2.2.8) are trivially satisfied, the condition (2.2.6) takes the form

$$\int_0^\infty \left[\delta - \frac{\gamma_t \beta_t(z_{t-})}{g_t}\right]^+ \frac{d\gamma_t}{\gamma_t} < \infty \quad P\text{-a.s.,} \tag{2.2.22}$$

and also

$$\{(2.2.22) \quad \text{for} \quad \delta = \delta_0/2\} \Rightarrow \{(2.2.22) \quad \text{for all} \quad \delta,\ 0 < \delta < \delta_0/2\}.$$

Further, since

$$\frac{C_t^{(\delta)}}{\gamma_t} = \frac{1}{\gamma_t} \int_0^t \frac{\beta_s(z_s)\gamma_s}{g_s}\, d\gamma_s - \delta \geq -\delta,$$

conditions (a)–(c) of Proposition 2.2.3 can be simplified to the following conditions:

(a′) The process

$$\left(\frac{1}{\gamma_t} \int_0^t \frac{\beta_s(z_s)\gamma_s}{g_s}\, d\gamma_s\right)_{t \geq 0}$$

is eventually bounded.

(b′) $\displaystyle\int_0^\infty \left[\frac{1}{\gamma_t} \int_0^t \frac{\beta_s(z_s)\gamma_s}{g_s}\, d\gamma_s - \delta\right]^- \frac{d\gamma_t}{\gamma_t} < \infty \quad P\text{-a.s.}$

(c′) $\displaystyle\int_0^\infty \left[\frac{1}{\gamma_t} \int_0^t \frac{\beta_s(z_s)\gamma_s}{g_s}\, d\gamma_s - \delta\right]^+ \frac{d\gamma_t}{\gamma_t} = \infty \quad P\text{-a.s.}$

Also, if (a′) holds and

(bc′) $\displaystyle\frac{C^{(\delta_0/2)}}{\gamma} = \left(\frac{1}{\gamma_t} \int_0^t \frac{\beta_s(z_s)\gamma_s}{g_s}\, d\gamma_s - \frac{\delta_0}{2}\right)_{t \geq 0} > 0$, eventually, then (b′) and (c′) hold for each δ, $0<\delta<\delta_0/2$.

In the discrete-time case, we additionally assume that

$$\sum_{t\geq 0}\left(\frac{\Delta\gamma_t}{\gamma_t}\right)^2<\infty \quad\text{and}\quad \sum_{t\geq 0}(\beta_t(z_{t-1}))^2<\infty \quad P\text{-a.s.} \tag{2.2.23}$$

Then the conditions of Corollary 2.2.3 are trivially satisfied. Hence conditions (2.2.3) and (2.2.14) are equivalent and can be written as

$$\sum_{t\geq 0}\left[\delta-\frac{\gamma_t\beta_t(z_{t-1})}{g_t}\right]^+\frac{\Delta\gamma_t}{\gamma_t}<\infty \quad P\text{-a.s.,} \tag{2.2.24}$$

and also

$$\{(2.2.24)\quad\text{for}\quad \delta=\delta_0/2\}\Rightarrow\{(2.2.24)\quad\text{for all}\quad \delta,\quad 0<\delta<\delta_0/2\}\,.$$

Note that the inverse implication "⇐" does not hold in general (see Example 2.2.3).

It is not difficult to verify that (a), (b), and (c) are equivalent to $(\widetilde{\text{a}})$, $(\widetilde{\text{b}})$, and $(\widetilde{\text{c}})$ defined as follows.

$(\widetilde{\text{a}})$ The process

$$\left(\frac{1}{\gamma_t}\sum_{s\leq t}\beta_s(z_{s-1})\gamma_s\right)_{t\geq 0}$$

is eventually bounded.

$(\widetilde{\text{b}})$ $$\sum_{t\geq 1}\left[\frac{1}{\gamma_{t-1}}\sum_{s<t}\beta_s(z_{s-1})\gamma_s-\delta\right]^-\frac{\Delta\gamma_t}{\gamma_t}<\infty \quad P\text{-a.s.}$$

$(\widetilde{\text{c}})$ $$\sum_{t\geq 1}\left[\frac{1}{\gamma_{t-1}}\sum_{s<t}\beta_s(z_{s-1})\gamma_s-\delta\right]^+\frac{\Delta\gamma_t}{\gamma_t}=\infty \quad P\text{-a.s.}$$

Also, if $(\widetilde{\text{a}})$ holds and

$(\widetilde{\text{bc}})$ $$\left(\frac{1}{\gamma_t}\sum_{s\leq t}\beta_s(z_{s-1})\gamma_s-\delta\right)_{t\geq 0}>\delta_0/2 \ \text{ eventually,}$$

then $(\widetilde{\text{b}})$ and $(\widetilde{\text{c}})$ hold for each δ, $0<\delta<\delta_0/2$.

Hence $\{(\widetilde{\text{a}}),(\widetilde{\text{bc}})\}\Rightarrow\{(\widetilde{\text{a}}),(\widetilde{\text{b}}),(\widetilde{\text{c}})$ for all δ, $0<\delta<\delta_0/2\}$. However, the inverse implication is not true (see Examples 2.2.3 and 2.2.4).

Note that the conditions imposed on the martingale part of Eq. (2.1.1) in Theorems 2.2.1 (see Eq. (2.2.4)) and 2.2.2 (see Eq. (2.2.18)) are identical. Therefore, assume that these conditions hold in all the examples given below.

2.2.1. Example. This example illustrates that Eq. (2.2.22) holds whereas (a′) is violated.

Let

$$K_t=\gamma_t=t+1 \quad\text{and}\quad \beta_t(u)\equiv(t+1)^{-(1/2+\alpha)},$$

where $0<\alpha<1/2$.

Substituting K_t, γ_t, β_t on the left-hand side of Eq. (2.2.22), we obtain

$$\int_0^\infty\left[\delta-(t+1)^{-(1/2-\alpha)}(t+1)\right]^+\frac{dt}{t+1}=\int_0^\infty\left[\delta-(t+1)^{1/2-\alpha}\right]^+\frac{dt}{t+1}\,.$$

Since $([\delta-(t+1)^{1/2-\alpha}]^+)_{t\geq 0}=0$ eventually, the condition (2.2.22) holds.

Conditions (a′) does not hold, since

$$\frac{1}{\gamma_t}\int_0^t\frac{\beta_s(z_s)\gamma_s}{g_s}\,d\gamma_s=\frac{1}{t+1}\int_0^t(s+1)^{1/2-\alpha}ds\propto(t+1)^{1/2-\alpha}\to\infty \quad\text{as}\quad t\to\infty.$$

Note that conditions (b′) and (c′) are satisfied.

Note that although Eq. (2.2.22) holds for all δ, $\delta > 0$, if, e.g.,

$$d\langle M\rangle_t = \frac{dt}{(t+1)^{3/2+\alpha}}\,,$$

conditions (2.2.4) only holds for δ satisfying $0 < \delta < \delta_0 = 1/2 + \alpha$.

2.2.2. Example. In this example, conditions ($\widetilde{\text{a}}$) and ($\widetilde{\text{bc}}$) hold for $\delta_0 = 1$, while Eq. (2.2.24) fails for some δ, $0 < \delta < 1/2 = \delta_0/2$.

Consider the discrete-time model with $K_t = \gamma_t = t$, $\beta_t(u) \equiv \beta_t$ and

$$\beta_t\gamma_t = \begin{cases} 1/2 + a & \text{if} \quad t \quad \text{is odd,} \\ 1/2 - b & \text{otherwise,} \end{cases}$$

where $0 < b < 1/2 \le a$. Then, since

$$\frac{1}{2} + a > \frac{1}{\gamma_t}\sum_{s\le t}\beta_s\gamma_s = \frac{1}{2} + \begin{cases} \frac{a-b}{2} & \text{if} \quad t = 2k, \quad k = 1,2,\dots \\ \frac{k(a-b)+a}{2k+1} & \text{if} \quad t = 2k+1, \quad k = 1,2,\dots \end{cases} > \frac{1}{2}\,,$$

conditions ($\widetilde{\text{a}}$) and ($\widetilde{\text{bc}}$) hold for $\delta_0 = 1$.

It is easy to verify that if $1/2 - b < \delta < 1/2$, then

$$\sum_{t\ge 1}[\delta - \beta_t\gamma_t]^+\frac{1}{t} = \sum_{t\ge 1}\left[\delta - \frac{1}{2} + b\right]^+\frac{1}{t}\, I_{\{t \text{ is even}\}} = \infty$$

which implies that Eq. (2.2.24) does not hold for all δ with $1/2 < b < \delta < 1/2$.

2.2.3. Example. In this discrete-time example, $\delta_0 = 1$ and

$$\{(2.2.24) \quad \text{for all} \quad \delta,\ 0 < \delta < 1/2\} \not\Rightarrow \{(2.2.24) \quad \text{for} \quad \delta = 1/2\}.$$

Suppose that $K_t = \gamma_t = t$, $\beta_t(u) \equiv \beta_t$ and

$$\beta_t\gamma_t = \left[\frac{1}{2} - \frac{1}{\log(t+1)}\right]^+.$$

Then for $0 < \delta < 1/2$ and large t,

$$[\delta - \beta_t\gamma_t]^+ = 0,$$

and it follows that

$$\sum_{t\ge 1}[\delta - \beta_t\gamma_t]^+\frac{1}{t} < \infty.$$

But, for $\delta = 1/2$,

$$\sum_{t\ge 1}\left[\frac{1}{2} - \beta_t\gamma_t\right]^+\frac{1}{t} \ge \sum_{t\ge 1}\frac{1}{t\log(t+1)}\, I_{\{\log(t+1)>1\}} = \infty.$$

Also, note that by the Toeplitz lemma,

$$\frac{1}{t}\sum_{s\le t}\beta_s\gamma_s = \frac{1}{t}\sum_{s\le t}\left[\frac{1}{2} - \frac{1}{\log(s+1)}\right]^+ \uparrow \frac{1}{2} \quad \text{as} \quad t \to \infty.$$

Therefore, for all δ, $0 < \delta < 1/2$, conditions ($\widetilde{\text{a}}$), ($\widetilde{\text{b}}$) and ($\widetilde{\text{c}}$) hold, whereas ($\widetilde{\text{bc}}$) does not.

2.2.4. Example. This is a discrete time example illustrating that Eq. (2.2.24) holds for $\delta = 1/2$ (hence for all $0 < \delta < 1/2$) and for all δ, $0 < \delta < 1/2$, conditions $(\widetilde{\mathrm{a}})$, $(\widetilde{\mathrm{b}})$ and $(\widetilde{\mathrm{c}})$, hold whereas $(\widetilde{\mathrm{bc}})$ does not.

Suppose that $K_t = \gamma_t = t$, $\beta_t(u) \equiv \beta_t$, and for $t > 0$,

$$\beta_t\gamma_t = \frac{1}{2} - \frac{1}{t}\,.$$

Then for $\delta = 1/2$, condition (2.2.24) follows, since

$$\sum_{t>2}\left[\frac{1}{2} - \beta_t\gamma_t\right]^+\frac{1}{t} = \sum_{t>2}\frac{1}{t^2} < \infty.$$

It remains to note that

$$\frac{1}{t}\sum_{s\le t}\beta_s\gamma_s \uparrow \frac{1}{2}$$

by the Toeplitz lemma.

2.2.5. Example. Here, we drop the "traditional" assumptions

$$\sum_{t>0}\left(\frac{\Delta\gamma_t}{\gamma_t}\right)^2 < \infty \quad \text{and} \quad \sum_{t\ge 0}(\beta_t(z_{t-1}))^2 < \infty \quad P\text{-a.s.}$$

and give an example where the conditions of Theorems 2.2.1 and 2.2.2 are satisfied.

Suppose that $K_t = t$ and the process γ and $\beta(u) = \beta$ are defined as follows: $\gamma_1 = 1$,

$$\gamma_t = \sum_{s=1}^{t} q^s = \frac{q}{1-q}\,(1-q^t), \quad \text{where} \quad q > 1,$$

and

$$\beta_t = \frac{\alpha}{\beta}\,\frac{\Delta\gamma_t}{\gamma_t}\,,$$

where $\alpha = q/(q-1)$ and β, $\beta > 1$, are some constants satisfying $(1-1/\alpha)^{1/2} > 1 - 1/\beta$. In this case,

$$\frac{\Delta\gamma_t}{\gamma_t} \to \frac{1}{\alpha} \quad \text{as} \quad t \to \infty$$

and

$$\beta_t\Delta K_t = \frac{\alpha}{\beta}\,\frac{\Delta\gamma_t}{\gamma_t} \to \frac{1}{\beta} < 1 \quad \text{as} \quad t \to \infty.$$

Therefore, the conditions of Corollary 2.2.3 hold, and it follows that conditions (2.2.3) and (2.2.14) are equivalent.

To verify Eq. (2.2.14), note that for all $0 < \delta < 1/2$,

$$\begin{aligned}
&\sum_{t>0}\left[1 - \beta_t(z_{t-})\Delta K_t - \left(1 - \frac{\Delta\gamma_t}{\gamma_t}\right)^{\delta}\right]^+ I_{\{\beta_t(z_{t-})\Delta K_t \le 1\}} \\
&\qquad \le \sum_{t>0}\left[1 - \beta_t(z_{t-})\Delta K_t - \left(1 - \frac{\Delta\gamma_t}{\gamma_t}\right)^{1/2}\right]^+ I_{\{\beta_t(z_{t-})\Delta K_t \le 1\}} \\
&\qquad\qquad \le \sum_{t>0}\left[1 - \frac{1}{\beta}\,\frac{q^t}{q^t-1} - \left(1 - \frac{1}{\alpha}\,\frac{q^t}{q^t-1}\right)^{1/2}\right]^+ I_{\{\beta_t(z_{t-})\Delta K_t \le 1\}}.
\end{aligned}$$

But since

$$1 - \frac{1}{\beta}\,\frac{q^t}{q^t-1} - \left(1 - \frac{1}{\alpha}\,\frac{q^t}{q^t-1}\right)^{1/2} \to 1 - \frac{1}{\beta} - \left(1 - \frac{1}{\alpha}\right)^{1/2} < 0,$$

we have

$$\left[1-\frac{1}{\beta}\,\frac{q^t}{q^t-1}-\left(1-\frac{1}{\alpha}\,\frac{q^t}{q^t-1}\right)^{1/2}\right]^+=0$$

for large t. Hence Eq. (2.2.14) holds.

To verify conditions (a), (b) and (c) of Theorem 2.2.2, note that by the Toeplitz lemma,

$$\frac{1}{\gamma_t}\,C_t^{(\delta)}=-\frac{1}{\gamma_t}\sum_{s\le t}\Delta\gamma_s\log\frac{|1-\beta_s(z_{s-1})|}{\left(1-\frac{\Delta\gamma_s}{\gamma_s}\right)^{\delta}}\,\frac{\gamma_s}{\Delta\gamma_s}\to a,$$

where

$$a=-\alpha\log\frac{1-1/\beta}{(1-1/\alpha)^{\delta}}>-\alpha\log\frac{1-1/\beta}{(1-1/\alpha)^{1/2}}>0,$$

which implies (a), (b), and (c).

2.3. Asymptotic expansion. In Sec. 2.1, we have derived the representation

$$\chi_t z_t=\frac{L_t}{\langle L\rangle_t^{1/2}}+R_t, \tag{2.3.1}$$

where all objects are defined there.

Throughout this section, we assume that

$$\langle L\rangle_\infty=\infty\quad P\text{-a.s.}$$

and there exists a predictable increasing process $\gamma=(\gamma_t)_{t\ge 0}$ such that $\gamma_0=1$, $\gamma_\infty=\infty$ P-a.s., the process γ/γ_- is eventually bounded, and

$$\gamma\simeq\Gamma^2\langle L\rangle^{-1}.$$

In this section, assuming that $\gamma_t^{\delta}z_t\to 0$ P-a.s. for all $0<\delta<\delta_0/2$ (for some $0<\delta_0\le 1$), we establish sufficient conditions for the convergence $R_t\xrightarrow{P}0$ as $t\to\infty$.

Consider the following conditions:

(d) There exists a nonrandom increasing process $(\langle\langle L\rangle\rangle_t)_{t\ge 0}$ such that

$$\frac{\langle L\rangle_t}{\langle\langle L\rangle\rangle_t}\xrightarrow{d}\zeta\quad\text{as}\quad t\to\infty,$$

where $\xrightarrow{d}$ denotes the convergence in distribution and $\zeta>0$ is some random variable.

(e) $$\sum_{t\ge 0}I_{\{\beta_t\Delta K_t=1\}}<\infty\quad P\text{-a.s.}$$

(f) There exists ε, $1/2-\delta_0/2<\varepsilon<1/2$, such that

$$\frac{1}{\langle L\rangle_t}\int_0^t|\beta_s-\beta_s(z_{s-})|\gamma_{s-}^{\varepsilon}\langle L\rangle_s dK_s\to 0\quad\text{as}\quad t\to\infty\quad P\text{-a.s.}$$

(g) $$\frac{1}{\langle L\rangle_t}\int_0^t\Gamma_s^2(h_s(z_{s-},z_{s-})-2h_s(z_{s-},0)+h_s(0,0))dK_s\xrightarrow{P}0\quad\text{as}\quad t\to\infty.$$

2.3.1. Theorem. *Suppose that* $\gamma_t^{\delta}z_t\to 0$ P-a.s. *for all* δ, $0<\delta<\delta_0/2$ $(0<\delta_0\le 1)$, *and conditions* (d)–(g) *are satisfied. Then*

$$R_t\xrightarrow{P}0\quad\textit{as}\quad t\to\infty.$$

Proof. Recall from Sec. 2.1 that

$$R_t = \frac{1}{\langle L\rangle_t^{1/2}}\, z_0 + R_t^{(1)} + R_t^{(2)} + R_t^{(3)},$$

where

$$R_t^{(1)} = -\frac{1}{\langle L\rangle_t^{1/2}} \sum_{s\le t} \Gamma_s z_{s-} I_{\{\beta_s \Delta K_s = 1\}},$$

$$R_t^{(2)} = \frac{1}{\langle L\rangle_t^{1/2}} \int_0^t \Gamma_s(\beta_s - \beta_s(z_{s-}))z_{s-} dK_s,$$

$$R_t^{(3)} = \frac{1}{\langle L\rangle_t^{1/2}} \int_0^t \Gamma_s(M(ds, z_{s-}) - M(ds, 0)).$$

Since $\langle L\rangle_t \to \infty$, we have $z_0/\langle L\rangle_t^{1/2} \to 0$ as $t\to\infty$. Further, it follows from (e) that the process $(I_{\{\beta_t \Delta K_t = 1\}})_{t\ge 0} = 0$ eventually, and, therefore, $R_t^{(1)} \to 0$ as $t \to \infty$.

Since the process γ/γ_- is bounded eventually and $\gamma_t^{1/2-\varepsilon} z_t \to 0$ as $t\to\infty$ P-a.s., we obtain that the process $\gamma^{1/2-\varepsilon} z_-$ is bounded eventually for each ε, $1/2 - \delta_0/2 < \varepsilon < 1/2$. Also, $|\Gamma|\langle L\rangle^{-1/2} \simeq \gamma^{1/2}$. Therefore, it follows that there exists an eventually bounded positive process $\eta = (\eta_t)_{t\ge 0}$ such that

$$|R_t^{(2)}| \le \frac{1}{\langle L\rangle_t^{1/2}} \int_0^t |\Gamma_s|\,|\beta_s - \beta_s(z_{s-})|\,|z_{s-}|\,dK_s$$
$$= \frac{1}{\langle L\rangle_t^{1/2}} \int_0^t |\beta_s - \beta_s(z_{s-})| \gamma_s^{\varepsilon} \langle L\rangle_s \eta_s \frac{dK_s}{\langle L\rangle_s^{1/2}} = \frac{1}{\langle L\rangle_t^{1/2}} \int_0^t D_s dC_s^{\varepsilon},$$

where

$$D_t = \frac{1}{\langle L\rangle_t^{1/2}}, \quad C_t^{\varepsilon} = \int_0^t |\beta_s - \beta_s(z_{s-})| \gamma_s^{\varepsilon} \langle L\rangle_s \eta_s dK_s.$$

Using the formulas $d(D_t C_t) = D_t dC_t + C_{t-} dD_t$ we obtain

$$|R_t^{(2)}| \le \left(\frac{1}{\langle L\rangle_t}\, C_t^{\varepsilon} - \frac{1}{\langle L\rangle_t^{1/2}} \int_0^t C_{s-}^{\varepsilon} d\langle L\rangle_s^{-1/2} \right).$$

It is easy to verify that

$$d(\langle L\rangle_t^{-1/2}) = -\frac{1}{\langle L\rangle_{t-}^{1/2}} \frac{d\langle L\rangle_t^{1/2}}{\langle L\rangle_t^{1/2}}$$

and

$$|R_t^{(2)}| \le \frac{1}{\langle L\rangle_t}\, C_t^{\varepsilon} + \frac{1}{\langle L\rangle_t^{1/2}} \int_0^t \frac{1}{\langle L\rangle_{s-}} C_{s-}^{\varepsilon} d\langle L\rangle_s^{1/2}.$$

Now, from the condition (f) and the Toeplitz lemma, $R_t^{(2)} \to 0$ P-a.s.

To prove the convergence $R^{(3)}_t \to 0$ note that by condition (d), it suffices to consider the case where $\langle L\rangle_t$ is nonrandom. Denote

$$N_t = \int_0^t \Gamma_s(M(ds, z_{s-}) - M(ds, 0)).$$

Using the Lenglart–Rebolledo inequality (see, e.g., [25, Chap. 1, Sec. 9] and [22]), we obtain

$$P\{\langle L\rangle_t^{-1/2} N_t > a\} = P\{\langle L\rangle_t^{-1} N_t^2 > a^2\} = P\{N_t^2 - \langle L\rangle_t \varepsilon > (a^2 - \varepsilon)\langle L\rangle_t\}$$
$$\le \frac{b}{(a^2-\varepsilon)\langle L\rangle_t} + P\{\langle N\rangle_t - \langle L\rangle_t \varepsilon > b\}$$

for any $a > 0$, $b > 0$ and $0 < \varepsilon < a^2$. Now, the result follows, since $\langle L\rangle_\infty = \infty$ P-a.s. and

$$\frac{1}{\langle L\rangle_t}\langle N\rangle_t = \frac{1}{\langle L\rangle_t}\int_0^t \Gamma_s^2(h_s(z_{s-}, z_{s-}) - 2h_s(z_{s-}, 0) + h_s(0,0))dK_s \xrightarrow{P} 0 \quad \text{as} \quad t \to \infty.$$

The theorem is proved. □

2.3.1. Remark. Suppose that P-a.s.,

$$\beta \circ K_\infty = \infty, \quad \inf_{t\ge 0} \beta_t I_{\{\Delta K_t \ne 0\}} > 0, \quad \sup_{t\ge 0} \beta_t \Delta K_t I_{\{\Delta K_t \ne 0\}} < 2.$$

Then, as is easy to see, $|\Gamma|$ is an increasing process with $|\Gamma_\infty| = \infty$ P-a.s..

2.3.2. Remark. (1) Condition (f) can be replaced by the following one: (f′) there exists $\varepsilon > (1-\delta_0)/\delta_0$ such that

$$\frac{1}{\langle L\rangle_t}\int_0^t |\beta_s - \beta_s(z_{s-})|\,|z_{s-}|^{-\varepsilon}\langle L\rangle_s dK_s \to 0 \quad \text{as} \quad t\to\infty \quad P\text{-a.s.}$$

(2) It follows from Eq. (2.3.1) that under the conditions of Theorem 2.3.1, the asymptotic behavior of the normalized process $(\chi_t z_t)_{t\ge 0}$ coincides with the asymptotic behavior of $(L_t/\langle L\rangle_t)_{t\ge 0}$ as $t\to\infty$.

(3) Assume that the first two conditions in Remark 2.3.1 hold, and, moreover,

$$\sup_{t\ge 0} \beta_t \Delta K_t I_{\{\Delta K_t \ne 0\}} < 1 \quad P\text{-a.s.}$$

In this case, $\overline{\beta}_t = \beta_t I_{\{\beta_t \Delta K_t \ne 1\}} = \beta_t$, $\Gamma = \varepsilon^{-1}(-\beta \circ K)$ is a positive increasing process, $\Gamma_t \uparrow \infty$ P-a.s. as $t\to\infty$, and if we suppose that $\Gamma \simeq \langle L\rangle$, then taking $\gamma = \langle L\rangle$, we obtain

$$\gamma \simeq \Gamma^2 \langle L\rangle^{-1} \simeq \Gamma$$

and under the conditions of Theorem 2.3.1,

$$\Gamma_t^{1/2} z_t = \frac{L_t}{\langle L\rangle_t^{1/2}} + R_t, \quad R_t \xrightarrow{P} 0 \quad \text{as} \quad t\to\infty.$$

Note that for the recursive parametric estimation procedures in the discrete time case, $\Gamma^2\langle L\rangle^{-1} = \Gamma$ (see [39]).

2.3.1. Example. The RM stochastic approximation procedure with slowly varying gains (see [31]). Consider the SDE

$$dz_t = -\frac{\alpha R(z_t)}{(1+K_t)^r}\,dK_t + \frac{\alpha}{(1+K_t)^r}\,dm_t.$$

Here, $K = (K_t)_{t\geq 0}$ is a continuous increasing nonrandom function with $K_\infty = \infty$, $1/2 < r < 1$, $0 < \alpha < 1$, $m = (m_t)_{t\geq 0} \in \mathcal{M}^2_{\text{loc}}(P)$, $d\langle m\rangle_t = \sigma_t^2 dK_t$, and $\sigma_t^2 \to \sigma^2 > 0$ as $t \to \infty$ and the nonrandom regression function R satisfies the following conditions:

$$R(0) = 0, \quad uR(u) > 0 \quad \text{if} \quad u \neq 0,$$

for each $\varepsilon > 0$, $\inf\limits_{\varepsilon<|u|<\frac{1}{\varepsilon}} uR(u) > 0$, and

$$R(u) = \beta u + v(u) \quad \text{with} \quad v(u) = O(u^2) \quad \text{as} \quad u \to 0.$$

In our notation,

$$\beta_t = \frac{\alpha\beta}{(1+K_t)^r} \quad \text{and} \quad \beta_t(u) = \frac{\alpha R(u)}{u(1+K_t)^r}\,.$$

It follows from Theorem 1.2.1 that $z_t \to 0$ P-a.s. as $t \to \infty$.

From Sec. 2.1, it follows that

$$\chi_t z_t = \frac{L_t}{\langle L\rangle_t^{1/2}} + R_t$$

with $\Gamma_t = \varepsilon_t^{-1}(-\beta \circ K)$,

$$L_t = \int_0^t \Gamma_s \frac{\alpha}{(1+K_s)^r}\, dm_s, \quad \chi_t^2 = \Gamma_t^2 \langle L\rangle_t^{-1},$$

and

$$R_t = \frac{1}{\langle L\rangle_t^{1/2}} \int_0^t \Gamma_s(\beta_s - \beta_s(z_{s-}))z_{s-}dK_s + \frac{z_0}{\langle L\rangle_t^{1/2}}\,.$$

On can verify that

$$(1+K_t)^{-r}\chi_t^2 \to \frac{2\beta}{\alpha\sigma^2}$$

as $t \to \infty$. Since

$$\frac{L_t}{\langle L\rangle_t^{1/2}} \xrightarrow{w} \mathcal{N}(0,1),$$

it follows that if the convergence $R_t \xrightarrow{P} 0$ holds, we have

$$(1+K_t)^{r/2} z_t \xrightarrow{w} \mathcal{N}\left(0, \frac{\alpha\sigma^2}{2\beta}\right). \tag{2.3.2}$$

It remains to prove that $R_t \xrightarrow{P} 0$ as $t \to \infty$. Let us first prove that if $1/2 < r < 1$, then P-a.s.,

$$(1+K_t)^{r\delta} z_t \to 0 \quad \text{for all} \quad \delta < 1 - \frac{1}{2r}\,. \tag{2.3.3}$$

It is easy to verify that

$$(1+K_t)^{2r\delta} = \varepsilon_t^{-1}\left(-\frac{2r\delta}{(1+K)} \circ K\right).$$

Therefore, conditions (2.2.3) and (2.2.4) of Theorem 2.2.1 can be rewritten as

$$\int_0^\infty \left[\frac{2r\delta}{(1+K_t)} - \frac{2\alpha\beta}{(1+K_t)^r} - \frac{2\alpha v(z_t)}{z_t(1+K_t)^r}\right]^+ dK_t < \infty \quad P\text{-a.s.} \tag{2.3.4}$$

and

$$\int_0^\infty (1+K_t)^{2r\delta} \frac{\alpha^2\sigma_t^2}{(1+K_t)^{2r}}\, dK_t < \infty \quad P\text{-a.s.} \tag{2.3.5}$$

Condition (2.3.4) holds since

$$\left[\frac{2r\delta}{(1+K_t)} - \frac{2\alpha\beta}{(1+K_t)^r} - \frac{2\alpha v(z_t)}{z_t(1+K_t)^r}\right]^+ = 0$$

eventually. Condition (2.3.5) is satisfied, since $2r-2r\delta > 1$ if $\delta < 1-1/(2r)$. Therefore, Theorem 2.2.1 yields Eq. (2.3.3). Conditions (d) and (e) of Theorem 2.3.1 are trivially fulfilled. To verify (f), note that from the Kronecker lemma, it suffices to verify that

$$\int_0^\infty |\beta_t - \beta_t(z_t)|\gamma_t^\varepsilon dK_t < \infty \quad P\text{-a.s.}$$

for some ε with $1/2 - \delta_0/2 < \varepsilon < 1/2$, $\delta_0 = 2 - 1/r$. For each δ, $0 < \delta < \delta_0/2 = 1 - 1/(2r)$, we have

$$\int_0^\infty |\beta_t - \beta_t(z_t)|\gamma_t^\varepsilon dK_t = \int_0^\infty \frac{|v(z_t)|}{|z_t|^2}\,|z_t|\gamma_t^\varepsilon(1+K_t)^{-r} dK_t \le \xi \int_0^\infty (1+K_t)^{-r(a+\delta-\varepsilon)} dK_t$$

for some random variables ξ. Therefore, it follows that if there exists a triple (r,δ,ε) satisfying the inequalities

$$\frac{1}{2} < r < 1, \quad 0 < \delta < \frac{1}{2}, \quad \varepsilon > 0, \quad r(1+\delta-\varepsilon) > 1,$$
$$\frac{1}{2r} - \frac{1}{2} < \varepsilon < \frac{1}{2}, \quad \delta < 1 - \frac{1}{2r},$$

then Eq. (2.3.2) holds. It is easy to verify that such a triple exists only for $r > 4/5$. Therefore, it follows that Eq. (2.3.2) holds for $r > 4/5$.

3. The Polyak Weighted Averaging Procedure

3.1. Preliminaries. Consider the RM type SDE

$$z_t = z_0 + \int_0^t H_s(z_s)\,dK_s + \int_0^t \ell_s(z_s)\,dm_s, \tag{3.1.1}$$

where

(1) $\{H_t(u),\ t \ge 0,\ u \in \mathbb{R}^1\}$ is a random field described in Sec. 0;
(2) $\{M(t,u),\ t \ge 0,\ u \in \mathbb{R}^1\}$ is a random field such that

$$M(u) = (M(t,u))_{t\ge 0} \in M^2_{\text{loc}}(P)$$

for each $u \in \mathbb{R}^1$ and $M(t,u) = \int_0^t \ell_s(u)\,dm_s$, where $m = (m_t)_{t\ge0} \in M^{2,c}_{\text{loc}}(P)$, $M(\cdot,0) \ne 0$; $\ell(u) = (\ell_t(u))_{t\ge0}$ is a predictable process for each $u \in \mathbb{R}^1$. Denote $\ell_s := \ell_s(0)$.
(3) $K = (K_t)_{t\ge0}$ is a continuous increasing process.

Suppose that this equation has a unique strong solution $z = (z_t)_{t\ge0}$ defined on the whole time interval $[0,\infty)$ such that

$$\big(M(t)\big)_{t\ge0} = \left(\int_0^t \ell_s(z_s)\,dm_s\right)_{t\ge0} \in M^{2,c}_{\text{loc}}(P).$$

In Sec. 1 we have established the conditions that guarantee the convergence

$$z_t \to 0, \quad \text{as} \quad t \to \infty \quad P\text{-a.s.} \tag{3.1.2}$$

In Sec. 2, assuming (3.1.2), we have stated the conditions under which the following property of $z = (z_t)_{t\geq 0}$ holds:

(a) for each δ, $0 < \delta < \delta_0$, $0 < \delta_0 \leq 1$,

$$\gamma_t^\delta z_t^2 \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.},$$

where $\gamma = (\gamma_t)_{t\geq 0}$ is a predictable increasing process with $\gamma_0 = 1$, $\gamma_\infty = \infty$ P-a.s.

Furthermore, assuming that $z = (z_t)_{t\geq 0}$ has property (a) with the process $\gamma = (\gamma_t)_{t\geq 0}$, equivalent to the process $\Gamma^2\langle L\rangle^{-1} = (\Gamma_t^2\langle L\rangle_t^{-1})_{t\geq 0}$ (i.e., $\lim\limits_{t\to\infty} \frac{\Gamma_t^2\langle L\rangle_t^{-1}}{\gamma_t} = \widetilde{\gamma}^{-1}$, $0 < \widetilde{\gamma} < \infty$), in Sec. 2, we have found the conditions under which the asymptotic expansion

$$\Gamma_t\langle L\rangle_t^{1/2} z_t = \frac{L_t}{\langle L\rangle_t^{1/2}} + R_t, \tag{3.1.3}$$

where $R_t \xrightarrow{P} 0$ as $t \to \infty$, holds.

Here, the objects γ_t, L_t, and $\langle L\rangle_t$ are defined as follows:

$$\Gamma_t = \varepsilon_t(\beta \circ K) := \exp\left(\int_0^t \beta_s \, dK_s\right),$$

where $\beta_t = -H_t'(0)$, $L_t = \int_0^t \Gamma_s \ell_s(0)\, dm_s$, and $\langle L\rangle$ is the shifted square characteristics of L, i.e., $\langle L\rangle_t = 1 + \langle L\rangle_t^{F,P}$, where $\langle L\rangle_t^{F,P} = \int_0^t \Gamma_s^2 \ell_s^2 dK_s$.

Now let us consider the following weighted averaging procedure:

$$\overline{z}_t = \frac{1}{\varepsilon_t(g\circ K)} \int_0^t z_s \, d\varepsilon_s(g \circ K), \tag{3.1.4}$$

where $g = (g_t)_{t\geq 0}$ is a predictable process, $g_t \geq 0$ for all $t \geq 0$, P-a.s.,

$$\varepsilon_t = \varepsilon_t(g\circ K) = \exp\left(\int_0^t g_s\, dK_s\right), \quad \int_0^t g_s\, dK_s < \infty, \quad t \geq 0, \quad \int_0^\infty g_s\, dK_s = \infty \quad P\text{-a.s.}$$

The aim of this section is to study the asymptotic properties of the process $\overline{z} = (\overline{z}_t)_{t\geq 0}$ as $t \to \infty$.

First it should be noted that if $z_t \to 0$ as $t \to \infty$ P-a.s., then by the Toeplitz lemma (see, e.g., [25]), it immediately follows that

$$\overline{z}_t \to 0 \quad \text{as} \quad t \to \infty \quad P - \text{a.s.}$$

In Sec. 3.2 we establish the asymptotic distribution of the process $\overline{z}$ in the "linear" case where $H_t(u) = -\beta_t u$, $M(t, u) \equiv M(t) = \int_0^t \ell_s \, dm_s$ with deterministic g, β, ℓ, and K, and $d\langle m\rangle_t = dK_t$.

The general case, i.e., when the process z in (3.1.4) is a strong solution of SDE (3.1.1), is considered in Sec. 3.3.

3.2. Asymptotic properties of $\overline{z}$. "Linear" case. In this section, we consider the "linear" case where the SDE (3.1.1) is of the form

$$dz_t = -\beta_t z_t \, dK_t + \ell_t \, dm_t, \quad z_0, \tag{3.2.1}$$

where $K = (K_t)_{t \geq 0}$ is a deterministic increasing function, $\beta = (\beta_t)_{t \geq 0}$, and $\ell = (\ell_t)_{t \geq 0}$ are deterministic functions, $\beta_t \geq 0$ for all $t \geq 0$,

$$\int_0^\infty \beta_s dK_s = \infty, \quad \int_0^t \beta_s dK_s < \infty \quad \forall t \geq 0, \quad \int_0^\infty \ell_s^2 dK_s < \infty.$$

Define the following objects:

$$\Gamma_t = \exp\Bigl(\int_0^t \beta_s \, dK_s\Bigr), \quad L_t = \int_0^t \Gamma_s \ell_s \, dm_s, \quad t \geq 0.$$

Under the above conditions, we have $\Gamma_\infty = \infty$, $\Gamma_\infty^2 \langle L \rangle_\infty^{-1} = \infty$. Indeed, the application of the Kronecker lemma (see, e.g., [25]) yields

$$\Gamma_t^{-2} \langle L \rangle_t = \frac{1}{\Gamma_t^2} \int_0^t \Gamma_s^2 \ell_s^2 \, dK_s \to 0 \quad \text{as} \quad t \to \infty,$$

since $\int_0^\infty \ell_s^2 \, dK_s < \infty$.

Solving Eq. (3.2.1), we obtain

$$z_t = \Gamma_t^{-1} \Bigl\{ z_0 + \int_0^t \Gamma_s \ell_s \, dm_s \Bigr\}, \quad t \geq 0. \tag{3.2.2}$$

From (3.2.2) and CLT for continuous martingales (see, e.g., [25]), it directly follows that

$$z_t \to 0 \quad \text{as} \quad t \to \infty, \tag{3.2.3}$$

$$\Gamma_t \langle L \rangle_t^{-1/2} z_t \xrightarrow{d} \xi \quad \text{as} \quad t \to \infty, \tag{3.2.4}$$

where "$\xrightarrow{d}$" denotes the convergence in distribution and ξ is a standard normal random variable ($\xi \in N(0,1)$).

Now let $\overline{z} = (\overline{z}_t)$ be an averaged process defined by (3.1.4) with the deterministic function $g = (g_t)_{t \geq 0}$,

$$\int_0^\infty g_t dK_t = \infty, \quad \int_0^t g_s dK_s < \infty \quad \forall t \geq 0.$$

Denote

$$B_t = \int_0^t \Gamma_s^{-1} d\varepsilon_s, \quad \widetilde{B}_t = \int_0^t (B_t - B_s)^2 d\langle L \rangle_s, \quad \varepsilon_t = \varepsilon_t(g \circ K).$$

3.2.1. Proposition. *Suppose that $\langle L \rangle_\infty = \infty$, $\langle L \rangle \circ B_\infty = \infty$, and $\widetilde{B}_\infty = \infty$. Then*

$$\varepsilon_t \widetilde{B}_t^{-1/2} \overline{z}_t \xrightarrow{d} \xi \quad \textit{as} \quad t \to \infty, \quad \xi \in N(0,1), \tag{3.2.5}$$

Proof. Substituting (3.2.2) in (3.1.4) and integrating by parts, we obtain

$$\overline{z}_t = \frac{z_0 B_t}{\varepsilon_t} + \varepsilon_t^{-1} \int_0^t (B_t - B_s) \, dL_s$$

Hence

$$\varepsilon_t \widetilde{B}_t^{-1/2} \overline{z}_t = z_0 B_t / (\widetilde{B}_t)^{1/2} + (\widetilde{B}_t)^{-1/2} \int_0^t (B_t - B_s)\, dL_s = I_t^1 + I_t^2. \tag{3.2.6}$$

First, we show that

$$I_t^1 \to 0 \quad \text{as} \quad t \to \infty.$$

It is easy to verify that

$$\widetilde{B}_t = \int_0^t (B_t - B_s)^2 d\langle L \rangle_s = 2 \int_0^t \left(\int_0^s \langle L \rangle_u \, dB_u \right) dB_s. \tag{3.2.7}$$

Rewrite $(I_t^1)^2$ in the form

$$(I_t^1)^2 = B_t^2 (\widetilde{B}_t)^{-1} = \frac{2 \int_0^t B_s \left(\int_0^s \langle L \rangle_u \, dB_u \right)^{-1} d\widetilde{B}_s}{\widetilde{B}_t}.$$

Since $\widetilde{B}_\infty = \infty$, applying the Toeplitz lemma, we obtain

$$\lim_{t\to\infty} (I_t^1)^2 = \lim_{t\to\infty} \frac{B_t}{\int_0^t \langle L \rangle_u \, dB_u}.$$

Furthermore, as $\int_0^\infty \langle L \rangle_u \, dB_u = \infty$, applying the Toeplitz lemma once again, we obtain

$$\lim_{t\to\infty} \frac{B_t}{\int_0^t \langle L \rangle_u \, dB_u} = \lim_{t\to\infty} \frac{\int_0^t \langle L \rangle_u^{-1} \langle L \rangle_u \, dB_u}{\int_0^t \langle L \rangle_u \, dB_u} = \lim_{t\to\infty} \frac{1}{\langle L \rangle_t} = 0.$$

It remains to show that

$$I_t^2 \xrightarrow{d} \xi \quad \text{as} \quad t \to \infty, \quad \xi \in N(0,1).$$

For any sequence $t_n \to \infty$ as $n \to \infty$, we define the sequence of martingales as follows:

$$M^n(u) = \frac{\int_0^{t_n u} (B_{t_n} - B_s)\, dL_s}{\left(\int_0^{t_n} (B_{t_n} - B_s)^2 d\langle L \rangle_s \right)^{1/2}}, \quad u \in [0,1].$$

Obviously, $\langle M^n \rangle_1 = 1$ for each $n \geq 1$, and from the CLT for continuous martingales, we have

$$M^n(1) = I_{t_n}^2 \xrightarrow{d} \xi \quad \text{as} \quad n \to \infty, \quad \xi \in N(0,1).$$

The proposition is proved. □

3.2.1. Remark. Note that $\varepsilon_\infty \widetilde{B}_\infty^{-1/2} = \infty$. Indeed, by the Toeplitz lemma,

$$\lim_{t\to\infty} \frac{\widetilde{B}_t}{\varepsilon_t^2} = \lim_{t\to\infty} \frac{\int\limits_0^t \left(\int\limits_0^s \langle L\rangle_u dB_u\right) \Gamma_s^{-1}\varepsilon_s^{-1} d\varepsilon_s^2}{\varepsilon_t^2} = \lim_{t\to\infty} \frac{1}{\Gamma_t\varepsilon_t} \int\limits_0^t \langle L\rangle_s \Gamma_s^{-1} d\varepsilon_s$$

$$= \lim_{t\to\infty} \frac{1}{\Gamma_t\varepsilon_t} \int\limits_0^t \langle L\rangle_s \Gamma_s^{-2}\Gamma_s \, d\varepsilon_s \le \lim_{t\to\infty} \frac{1}{\varepsilon_t} \int\limits_0^t \langle L_s\rangle \Gamma_s^{-2} d\varepsilon_s = 0,$$

since $\varepsilon_\infty = \infty$ and $\langle L\rangle_\infty \Gamma_\infty^{-2} = 0$.

Now let us define the process $\varepsilon_t^{(\alpha)} := \varepsilon_t(g^{(\alpha)} \circ K)$ as follows: let $(\alpha_t)_{t\ge 0}$ be a function, $\alpha_t \ge 0$ for all $t \ge 0$, and let $\lim\limits_{t\to\infty} \alpha_t = \alpha$, $0 < \alpha < \infty$. Define $\varepsilon^{(\alpha)}$ by the relation

$$\varepsilon_t^{(\alpha)} = 1 + \int\limits_0^t \alpha_s\beta_s \langle L\rangle_s^{-1} \Gamma_s^2 dK_s. \tag{3.2.8}$$

Note that

$$\langle L\rangle_t \Gamma_t^{-2} \varepsilon_t^{(\alpha)} g_t^{(\alpha)}/\beta_t = \alpha_t. \tag{3.2.9}$$

Indeed, it is easily seen that if

$$\varepsilon_t(\psi) = 1 + \int\limits_0^t \varphi_s dK_s,$$

then

$$\psi_t = \frac{\varphi_t}{\varepsilon_t(\psi \circ K)}.$$

Hence, if $\varepsilon_t(g^{(\alpha)} \circ K) = \varepsilon_t^{(\alpha)}$, then

$$g_t^{(\alpha)} = \alpha_t\beta_t \langle L\rangle_t^{-1} \Gamma_t^2 / \varepsilon_t^{(\alpha)},$$

and (3.2.9) follows.

It should be also noted that for each $(\alpha_t)_{t\ge 0}$ with $\lim\limits_{t\to\infty} \alpha_t = \alpha$,

$$\lim_{t\to\infty} \frac{\varepsilon_t^\alpha}{1} + \int\limits_0^t \alpha\beta_s \langle L\rangle_s^{-1} \Gamma_s^2 dK_s = 1.$$

3.2.2. Proposition. *Let $\overline{z}^{(\alpha)} = (\overline{z}_t^{(\alpha)})_{t\ge 0}$ be the averaged process corresponding to the averaging process $\varepsilon^{(\alpha)}$ (see (3.1.4)), i.e.,*

$$\overline{z}_t^{(\alpha)} = \frac{1}{\varepsilon_t^{(\alpha)}} \int\limits_0^t z_s d\varepsilon_s^{(\alpha)}, \quad t \ge 0.$$

Then

$$\left(1 + \int\limits_0^t \beta_s \langle L\rangle_s^{-1} \Gamma_s^2 dK_s\right)^{1/2} \overline{z}_t^{(\alpha)} \xrightarrow{d} \sqrt{2}\,\xi \quad \text{as} \quad t\to\infty, \quad \xi \in N(0,1).$$

Proof. By virtue of Proposition 3.2.1, it suffices to show that

$$\frac{\varepsilon_t^{(1)}}{(\varepsilon_t^{(\alpha)})^2(\widetilde{B}_t^{(\alpha)})^{-1}} \to 2 \quad \text{as} \quad t \to \infty, \tag{3.2.10}$$

where

$$B_t^{(\alpha)} = \int_0^t \Gamma_s^{-1}\, d\varepsilon_s^{(\alpha)}, \quad \widetilde{B}_t^{(\alpha)} = \int_0^t (B_t^{(\alpha)} - B_s^{(\alpha)})^2\, d\langle L\rangle_s.$$

We have

$$\lim_{t\to\infty} \frac{\varepsilon_t^{(1)}}{(\varepsilon_t^{(\alpha)})^2(\widetilde{B}_t^{(\alpha)})^{-1}} = \lim_{t\to\infty} \frac{\varepsilon_t^{(1)}}{\varepsilon_t^{(\alpha)}} \frac{\widetilde{B}_t^{(\alpha)}}{\varepsilon_t^{(\alpha)}} = \frac{1}{\alpha} \lim_{t\to\infty} \frac{\widetilde{B}_t^{(\alpha)}}{\varepsilon_t^{(\alpha)}}$$

$$= \frac{1}{\alpha} \lim_{t\to\infty} \frac{2\int_0^t \left(\int_0^s \langle L\rangle_u dB_u^{(\alpha)} \right) \Gamma_s^{-1} d\varepsilon_s^{(\alpha)}}{\varepsilon_t^{(\alpha)}} = \frac{2}{\alpha} \lim_{t\to\infty} \frac{1}{\Gamma_t} \int_0^t \langle L\rangle_s dB_s^{(\alpha)}.$$

Now, applying relation (3.2.9) and the Toeplitz lemma, we obtain

$$\lim_{t\to\infty} \frac{1}{\Gamma_t} \int_0^t \langle L\rangle_s dB_s^{(\alpha)} = \lim_{t\to\infty} \frac{1}{\Gamma_t} \int_0^t \langle L\rangle_s \Gamma_s^{-1} d\varepsilon_s^{(\alpha)}$$

$$= \lim_{t\to\infty} \frac{1}{\Gamma_t} \int_0^t \langle L\rangle_s \Gamma_s^{-2} \varepsilon_s^{(\alpha)} \frac{g_s^{(\alpha)}}{\beta_s} \Gamma_s \beta_s dK_s = \lim_{t\to\infty} \frac{1}{\Gamma_t} \int_0^t \alpha_s d\Gamma_s = \alpha.$$

□

3.2.1. Corollary. *Let $\gamma = (\gamma_t)_{t\geq 0}$ be an increasing process such that $\gamma_0 = 1$, $\gamma_\infty = \infty$, and let*

$$\lim_{t\to\infty} \frac{\langle L\rangle_t^{-1}\Gamma_t^2}{\gamma_t} = \widetilde{\gamma}^{-1} \quad \text{as} \quad t \to \infty,$$

where $\widetilde{\gamma}$ is a constant, $0 < \widetilde{\gamma} < \infty$. Then

(1) $\gamma_t^{1/2} z_t \xrightarrow{d} \widetilde{\gamma}^{1/2} \xi$ *as* $t \to \infty$;

(2) $\left(1 + \int_0^t \gamma_s \beta_s dK_s\right)^{1/2} \overline{z}_t^{(\alpha)} \xrightarrow{d} \sqrt{2\widetilde{\gamma}}\, \xi$ *as* $t \to \infty$;

(3) *if* $\gamma_s \beta_s = 1$ *eventually, then* $(1 + K_t)^{1/2} \overline{z}_t^{(\alpha)} \to \sqrt{2\widetilde{\gamma}}\, \xi$ *as* $t \to \infty$, $\xi \in N(0,1)$.

3.2.2. Remark. (1) Let

$$\gamma = (\gamma_t)_{t\geq 0} := \left(\frac{\beta_t}{\ell_t^2}\right)_{t\geq 0}$$

be an increasing process, $\gamma_0 = 1$, $\gamma_\infty = \infty$, $d\gamma \ll dK$. Then γ can be represented as a solution of the SDE $d\gamma_t = \gamma_t \lambda_t dK_t$, $\gamma_0 = 1$, with some $\lambda = (\lambda_t)_{t\geq 0}$.

Assume that $\lambda_t \to 0$ as $t \to \infty$ and $\lambda_t/\beta_t \to 0$ as $t \to \infty$. Then

$$\lim_{t\to\infty} \frac{\langle L\rangle_t^{-1}\Gamma_t^2}{\gamma_t} = 2.$$

Indeed,

$$\frac{\langle L\rangle_t^{-1}\Gamma_t}{\gamma_t} = \frac{\Gamma_t^2\gamma_t^{-1}}{\langle L\rangle_t}.$$

and integration by parts and the application of the Toeplitz lemma yield

$$\frac{\langle L\rangle_t^{-1}\Gamma_t^2}{\gamma_t} = \frac{\int_0^t 2\Gamma_s^2\beta_s\gamma_s^{-1}dK_s - \int_0^t \Gamma_s^2\gamma_s^{-2}\gamma_s\lambda_s dK_s}{\langle L\rangle_t}$$

$$= 2 - \frac{1}{\langle L\rangle_t}\int_0^t \frac{\lambda_s}{\gamma_s\ell_s^2}\,d\langle L\rangle_s = 2 - \frac{1}{\langle L\rangle_t}\int_0^t \frac{\lambda_s}{\beta_s}d\langle L\rangle_s \to 2 \quad \text{as} \quad t\to\infty.$$

Thus, if we set $\gamma_t = \frac{\beta_t}{\ell_t^2}$ in the above Corollary 3.2.1, then all the assertions hold with

$$\gamma_t = \frac{\beta_t}{\ell_t^2}, \quad \widetilde{\gamma} = \frac{1}{2};$$

(2) Let $\ell_t = \sigma\beta_t$, where β_t is a decreasing function, $\beta_t \to 0$ as $t\to\infty$, $d\beta_t = -\beta_t' dK_t$, and $\beta_t' > 0$. Then, if

$$\beta_t'/\beta_t^2 \to 0 \quad \text{as} \quad t\to\infty,$$

we have

$$\lim_{t\to\infty}\langle L\rangle_t^{-1}\Gamma_t^2\beta_t = 2\sigma^2.$$

From Proposition 3.2.2, it immediately follows that

$$(1+K_t)^{1/2}\overline{z}_t^{(\alpha)} \xrightarrow{d} \sqrt{2}\,\sigma\,\xi \quad \text{as} \quad t\to\infty.$$

3.2.3. Remark. Summarizing the above statements, we conclude that: as $t\to\infty$,

(a) $(\varepsilon_t^{(1)})^{1/2}\overline{z}_t^{(\alpha)} \xrightarrow{d} \sqrt{2}\,\xi$;
(b) $(\varepsilon_t^{(\alpha)})^{1/2}\overline{z}_t^{(\alpha)} \xrightarrow{d} \sqrt{\frac{2}{\alpha}}\,\xi$;
(c) $(\varepsilon_t^{(1)})^{1/2}\overline{z}_t^{(1)} \xrightarrow{d} \sqrt{2}\,\xi$;
(d) $\Gamma_t\langle L\rangle_t^{-1/2}z_t \xrightarrow{d} \xi$,

where $\xi \in N(0,1)$.

3.2.1. Example. Standard "linear" procedure.

Let $\beta_t = \alpha\beta(1+K_t)^{-1}$, $\ell_t = \alpha\sigma(1+K_t)^{-1}$, $\alpha\beta > 0$, and let $2\alpha\beta > 1$. Then $\Gamma_t^2\langle L\rangle_t^{-1} = \frac{2\alpha\beta-1}{\alpha^2\sigma^2}(1+K_t)$. Hence, from (3.2.4) it follows that

$$(1+K_t)^{1/2}z_t \xrightarrow{d} \frac{\alpha\sigma}{\sqrt{2\alpha\beta-1}}\,\xi \quad \text{as} \quad t\to\infty, \quad \xi\in N(0,1).$$

On the other hand,

$$\varepsilon_t^{(1)} = 1 + \int_0^t \beta_s\Gamma_s^2\langle L\rangle_s^{-1}\,dK_s = 1 + \frac{\alpha^2\sigma^2}{\beta(2\alpha\beta-1)}\,K_t,$$

and it follows from Proposition 3.2.2 that if we define

$$\overline{z}_t^{(1)} = \frac{1}{\varepsilon_t^{(1)}}\int_0^t z_s\,d\varepsilon_s^{(1)}, \quad \overline{z}_t = \frac{1}{1+K_t}\int_0^t z_s\,dK_s,$$

then

$$(1+K_t)^{1/2}\overline{z}_t^{(1)} \xrightarrow{d} \sigma\sqrt{\frac{2\alpha}{\beta(2\alpha\beta-1)}}\,\xi \quad \text{as} \quad t\to\infty,$$

$$(1+K_t)^{1/2}\overline{z}_t \xrightarrow{d} \sigma\sqrt{\frac{2\alpha}{\beta(2\alpha\beta-1)}}\,\xi \quad \text{as} \quad t\to\infty, \quad \xi\in N(0,1).$$

Hence the rate of convergence is the same, but the asymptotic variance of the averaged procedure $\overline{z}$ is smaller than that of the initial one.

3.2.2. Example. "Linear" procedure with slowly varying gains.

Let $\beta_t = \alpha\beta(1+K_t)^{-r}$, $\ell_t = \alpha\sigma(1+K_t)^{-r}$, $\alpha\beta > 0$, $\frac{1}{2} < r < 1$. Then the process $\gamma = (\gamma_t)_{t\geq 0}$ defined in Remark 3.2.2 is $\gamma_t = \frac{\beta}{\alpha\sigma^2}(1+K_t)^r$, $d\gamma_t = \frac{r\beta}{\alpha\sigma^2}(1+K_t)^r\frac{dt}{1+K_t}$. Hence $\lambda_t = \frac{r\beta}{\alpha\sigma^2}(1+K_t)^{-1}$ and $\lambda_t/\beta_t \to 0$ as $t\to\infty$. From Remark 3.2.2, it follows that

$$\lim_{t\to\infty}\frac{\Gamma_t^2\langle L\rangle_t^{-1}}{\gamma_t} = 2, \tag{3.2.11}$$

and from (3.2.4), we have

$$(1+K_t)^{r/2}z_t \xrightarrow{d} \sigma\sqrt{\frac{\alpha}{2\beta}}\,\xi \quad \text{as} \quad t\to\infty, \quad \xi\in N(0,1).$$

On the other hand,

$$\varepsilon_t^{(1)} = 1 + \int_0^t \beta_s\Gamma_s^2\langle L\rangle_s^{-1}dK_s = 1 + \int_0^t \beta_s\gamma_s\frac{\Gamma_s^2\langle L\rangle_s^{-1}}{\gamma_s}dK_s = 1 + \frac{\beta^2}{\sigma^2}\int_0^t \frac{\Gamma_s^2\langle L\rangle_s^{-1}}{\gamma_s}dK_s.$$

Hence, taking into the account (3.2.11), by the Toeplitz lemma, we have

$$\frac{\varepsilon_t^{(1)}}{1+K_t} \to 2\frac{\beta^2}{\sigma^2} \quad \text{as} \quad t\to\infty.$$

Therefore, from Remark 3.2.3 (c), we obtain

$$(1+K_t)^{1/2}\overline{z}_t^{(1)} \xrightarrow{d} \frac{\sigma}{\beta}\,\xi \quad \text{as} \quad t\to\infty, \quad \xi\in N(0,1).$$

and

$$(1+K_t)^{1/2}\overline{z}_t \xrightarrow{d} \frac{\sigma}{\beta}\,\xi \quad \text{as} \quad t\to\infty, \quad \xi\in N(0,1).$$

Note that if $\alpha\beta > 2$, then the asymptotic variance of $\overline{z}$ is smaller than that of the initial one.

3.2.3. Example. Let $\beta_t = (1+t)^{-(\frac{1}{2}+\alpha)}$, where α is a constant, $0 < \alpha < \frac{1}{2}$, $\ell_t^2 = (1+t)^{-(\frac{3}{2}+\alpha)}$. If we take $\gamma_t = \beta_t/\ell_t^2 = (1+t)^{-(\frac{1}{2}+\alpha)}(1+t)^{\frac{3}{2}+\alpha} = 1+t$, $d\gamma_t = \gamma_t\frac{1}{1+t}dt$, then $\lambda_t = (1+t)^{-1}$, $\frac{\lambda_t}{\beta_t} = (1+t)^{-1}(1+t)^{\frac{1}{2}+\alpha} = (1+t)^{\alpha-\frac{1}{2}} \to 0$ as $t\to\infty$. Therefore, from Remark 3.2.2 (1) it follows that

$$\lim_{t\to\infty}\frac{\Gamma_t^2\langle L\rangle_t^{-1}}{1+t} = 2,$$

and from Corollary 3.2.1 (1), we have

$$(1+t)^{1/2}z_t \xrightarrow{d} \sqrt{\frac{1}{2}}\,\xi \quad \text{as} \quad t\to\infty, \quad \xi\in N(0,1).$$

Now, if we define

$$\varepsilon_t^{(1)} = 1 + \int_0^t \beta_s \langle L\rangle_s^{-1}\Gamma_s^2 ds = 1 + \int_0^t \beta_s\gamma_s \frac{\Gamma_s^2\langle L\rangle_s^{-1}}{\gamma_s} ds = 1 + \int_0^t (1+s)^{\frac{1}{2}-\alpha} \frac{\Gamma_s^2\langle L\rangle_s^{-1}}{\gamma_s} ds,$$

then $\varepsilon_t^{(1)}/(1+t)^{3/2-\alpha} \to \frac{4}{3-2\alpha}$, and from Corollary 3.2.1 (2), we obtain

$$(1+t)^{3/2-\alpha}\overline{z}_t^{(1)} \to \sqrt{\frac{4}{3-2\alpha}}\,\xi \quad \text{as} \quad t\to\infty, \quad \xi \in N(0,1).$$

In the last two examples, the rate of convergence of the averaged procedure is higher than that of the initial one.

3.3. Asymptotic properties of $\overline{z}$. General case. In this section, we study the asymptotic properties of the averaged process $\overline{z} = (\overline{z})_{t\ge 0}$ defined by (3.1.4), where $z = (z_t)_{t\ge 0}$ is the strong solution of SDE (3.1.1).

In the sequel, we will need the following objects:

$$\beta_t = -H_t'(0), \quad \beta_t(u) = \begin{cases} -\frac{H_t(u)}{u} & \text{if} \quad u \neq 0, \\ \beta_t & \text{if} \quad u = 0, \end{cases}$$

$$\Gamma_t = \varepsilon_t(\beta\circ K) = \exp\left\{\int_0^t \beta_s dK_s\right\}, \quad L_t = \int_0^t \Gamma_s \ell_s dm_s, \quad \ell_t = \ell_t(0), \quad d\langle m\rangle_t = dK_t.$$

Assume that processes K, β, and ℓ are deterministic. Rewrite Eq. (3.1.1) in terms of these objects:

$$dz_t = -\beta_t z_t dK_t + \ell_t dm_t + (\beta_t - \beta_t(z_t))z_t dK_t + (\ell_t(z_t) - \ell_t)dm_t. \tag{3.3.1}$$

Furthermore, formally solving the last equation as the linear one with respect to z, we obtain

$$z_t = \Gamma_t^{-1}\left[z_0 + L_t + \int_0^t \Gamma_s d\overline{R}_1(s) + \int_0^t \Gamma_s d\overline{R}_2(s)\right], \tag{3.3.2}$$

where

$$\Gamma_t = \exp\left(\int_0^t \beta_s\, dS_s\right), \quad L_t = \int_0^t \Gamma_s \ell_s\, dm_s, \quad d\overline{R}_1(t) = \big(\beta_t - \beta_t(z_t)\big)z_t dK_t, \quad d\overline{R}_2(t) = \big(\ell_t(z_t) - \ell_t\big)dm_t.$$

Now we consider the following averaging procedure:

$$\overline{z}_t = \frac{1}{\varepsilon_t}\int_0^t z_s d\varepsilon_s, \tag{3.3.3}$$

where

$$\varepsilon_t := \varepsilon_t = 1 + \int_0^1 \Gamma_s^2 \langle L\rangle_s^{-1}\beta_s dK_s,$$

i.e., it is defined by relation (3.2.8) with $\alpha_t = 1$.

In the sequel, it will be assumed that the functions β, ℓ, K, and g satisfy all the conditions imposed on the corresponding functions in Propositions 3.2.1 and 3.2.2.

Let $\gamma = (\gamma)_{t\ge 0}$ be an increasing function such that $\gamma_0 = 1$, $\gamma_\infty = \infty$, and $\lim\limits_{t\to\infty} \dfrac{\Gamma_t^2\langle L\rangle_t^{-1}}{\gamma_t} = \widetilde{\gamma}^{-1}$.

3.3.1. Theorem. *Suppose that $\gamma_t^\delta z_t^2 \to 0$ as $t \to \infty$ for all δ, $0 < \delta < \delta_0$, $0 < \delta_0 \le 1$. Assume that the following conditions are satisfied*:

(i) *there exists δ, $0 < \delta < \delta_0/2$ such that*

$$\int_0^\infty \varepsilon_t^{-1/2} \gamma_t^{-\delta} \big|\beta_t(z_t) - \beta_t\big| dK_t < \infty \quad P\text{-a.s.};$$

(ii) $\dfrac{\langle N\rangle_t}{\langle L\rangle_t} \to 0$ *as $t \to \infty$, where*

$$N_t = \int_0^t \Gamma_s\big(\ell_s(z_s) - \ell_s\big) dm_s.$$

Then

$$\varepsilon_t^{1/2} \overline{z}_t \xrightarrow{d} \quad \sqrt{2}\,\xi \quad \textit{as} \quad t \to \infty, \quad \xi \in N(0,1).$$

Proof. Substituting (3.3.2) in (3.3.3), we obtain

$$\overline{z}_t = \frac{z_0 B_t}{\varepsilon_t} + \frac{1}{\varepsilon_t} \int_0^t L_s dB_s + R_t^1 + R_t^2, \tag{3.3.4}$$

where

$$R_t^i = \frac{1}{\varepsilon_t} \int_0^t \int_0^s \Big(L_u d\overline{R}_i(u)\Big) dB_s, \quad i = 1, 2,$$

and $dB_t \equiv \Gamma_t^{-1} d\varepsilon_t$.

Integration of the second term in (3.3.4) by parts results in

$$\overline{z}_t = \frac{z_0 B_t}{\varepsilon_t} + \frac{1}{\varepsilon_t} \int_0^t (B_t - B_s) dL_s + R_t^1 + R_t^2. \tag{3.3.5}$$

Denoting $\widetilde{B}_t = \int_0^t (B_t - B_s)^2 d\langle L\rangle_s$, we have

$$\varepsilon_t \widetilde{B}_t^{-1/2} \overline{z}_t = z_0 \frac{B_t}{(\widetilde{B}_t)^{1/2}} + \frac{\int_0^t (B_t - B_s) dL_s}{(\widetilde{B}_t^{1/2})} + \frac{R_t^1}{(\widetilde{B}_t)^{1/2}} + \frac{R_t^2}{(\widetilde{B}_t)^{1/2}}. \tag{3.3.6}$$

As is seen, the first two terms on the right-hand side of (3.3.6) coincide with those in (3.2.6), and since by our assumption, the conditions of Propositions 3.2.1 and 3.2.2 are satisfied, taking into the account (3.2.10) with $\alpha = 1$, one can conclude that it suffices to show that

$$\varepsilon_t^{1/2} R_t^i \xrightarrow{P} 0 \quad \text{as} \quad t \to \infty, \quad i = 1, 2. \tag{3.3.7}$$

Let us investigate the case $i = 1$:

$$\varepsilon_t^{1/2} R_t^1 = \frac{1}{\varepsilon_t^{1/2}} \int_0^t \Big(\int_0^s \Gamma_u d\overline{R}_1(u)\Big) dB_s = \frac{1}{\varepsilon_t^{1/2}} \int_0^t \Big(\int_0^s \Gamma_u d\overline{R}_1(u)\Big) \Gamma_s^{-1} d\,\varepsilon_s$$

$$= \frac{2}{\varepsilon_t^{1/2}} \int_0^t \Big(\int_0^s \Gamma_u d\overline{R}_1(u)\Big) \Gamma_s^{-1} \varepsilon_s^{1/2} d\varepsilon_s^{1/2}.$$

Since ε_t is an increasing process, $\varepsilon_\infty = \infty$, by virtue of the Toeplitz lemma, it suffices to show that

$$A_t = \frac{1}{\Gamma_t \varepsilon_t^{1/2}} \int_0^t \Gamma_s d\overline{R}_1(s) \to 0 \quad \text{as} \quad t \to \infty, \quad P\text{-a.s.}$$

For all δ, $0 < \delta < \delta_0/2$, since $\gamma_t^\delta |z_t| \to 0$ as $t \to \infty$, we have

$$|A_t| \le \frac{1}{\Gamma_t \varepsilon_t^{1/2}} \int_0^t \Gamma_s |\beta_s - \beta_s(z_s)||z_s| dK_s \le \text{const}(\omega) \frac{1}{\Gamma_t \varepsilon_t^{1/2}} \int_0^t \Gamma_s \gamma_s^{-\delta} |\beta_s - \beta_s(z_s)| dK_s$$

$$= \text{const}(\omega) \frac{1}{\Gamma_t \varepsilon_t^{1/2}} \int_0^t \Gamma_s \varepsilon_s^{1/2} \varepsilon_s^{-1/2} \gamma_s^{-\delta} |\beta_s - \beta_s(z_s)| dK_s.$$

Now the desired convergence $A_t \to 0$ as $t \to \infty$ follows from condition (i) and the Kronecker lemma applied to the last term of the previous inequalities.

Now let us consider the second term

$$\varepsilon_t^{1/2} R_t^2 = \frac{1}{\varepsilon_t^{1/2}} \int_0^t \left(\int_0^s \Gamma_u \big(\ell_u(z_u) - \ell_u\big) dm_u \right) \Gamma_s^{-1} d\varepsilon_s. \tag{3.3.8}$$

Denoting $N_t = \int_0^t \Gamma_s(\ell_s(z_s) - \ell_s) dm_s$ and integrating by parts, from (3.3.8), we obtain

$$\varepsilon_t^{1/2} R_t^2 = \frac{1}{\varepsilon_t^{1/2}} \int_0^t (B_t - B_s) dN_s.$$

Furthermore, for any sequence t_n, $t_n \to \infty$ as $n \to \infty$, let us consider the sequence of martingales Y_u^n, $u \in [0,1]$, defined as follows:

$$Y_u^n = \frac{1}{\varepsilon_{t_n}^{1/2}} \int_0^{t_n u} (B_{t_n} - B_s) dN_s, \quad \langle Y^n \rangle_1 = \frac{1}{\varepsilon_{t_n}} \int_0^{t_n} (B_{t_n} - B_s)^2 d\langle N \rangle_s.$$

Now, if we show that $\langle Y^n \rangle_1 \xrightarrow{P} 0$ as $n \to \infty$, then from the well-known fact that $\langle Y^n \rangle_1 \xrightarrow{P} 0 \Rightarrow Y_1^n \xrightarrow{P} 0$ (see, e.g., [25]), we obtain $\varepsilon_{t_n}^{1/2} R_{t_n}^2 \to 0$ as $n \to \infty$, and hence $\varepsilon_t^{1/2} R_t^2 \to 0$ as $t \to \infty$.

Thus, we need to show that

$$\frac{1}{\varepsilon_t} \int_0^t (B_t - B_s)^2 d\langle N \rangle_s \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.}$$

Using the relation

$$\int_0^t (B_t - B_s)^2 d\langle N \rangle_s = 2 \int_0^t \left(\int_0^s \langle N \rangle_u dB_u \right) dB_s,$$

we need to show that

$$\frac{1}{\varepsilon_t} \int_0^t (B_t - B_s)^2 d\langle N \rangle_s = 2 \frac{1}{\varepsilon_t} \int_0^t \left(\int_0^s \langle N \rangle_u dB_u \right) \Gamma_s^{-1} d\varepsilon_s \to 0 \quad \text{as} \quad t \to \infty \tag{3.3.9}$$

Applying the Toeplitz lemma to (3.3.9), it suffices to show that

$$\frac{1}{\Gamma_t}\int_0^t \langle N\rangle_s dB_s \to 0 \quad \text{as} \quad t\to\infty \quad P\text{-a.s.} \tag{3.3.10}$$

But

$$\frac{1}{\Gamma_t}\int_0^t \langle N\rangle_s dB_s = \frac{1}{\Gamma_t}\int_0^t \langle N\rangle_s \Gamma_s^{-1} d\varepsilon_s = \frac{1}{\Gamma_t}\int_0^t \langle N\rangle_s \langle L\rangle_s^{-1} d\Gamma_s \tag{3.3.11}$$

(recall that $d\varepsilon_s = \Gamma_s^2\langle L\rangle_s^{-1}\beta_s dK_s$).

Applying the Toeplitz lemma to (3.3.11) once again, we can see that (3.3.10) follows from condition (ii). □

3.3.1. Corollary. *Let $H_t(u) = -\beta_t u + v_t(u)$, where, for each $t\in[0,\infty)$, $\left|\frac{v_t(u)}{u^2} - v_t\right| \to 0$ as $u\to 0$ P-a.s. Assume that the following condition holds:*

(i′) *there exists δ, $0<\delta<\delta_0$, such that*

$$\int_0^\infty \varepsilon_t^{1/2}\gamma_t^{-2\delta}|v_t|dK_t < \infty.$$

Then condition (i) *of Theorem* 3.3.1 *holds.*

Proof. Since $|\beta_t(u)-\beta_t| = |\frac{v_t(u)}{u}|$, for δ, $0<\delta<\frac{\delta_0}{2}$, we have

$$\int_0^\infty \varepsilon_t^{1/2}\gamma^{-\delta}|\beta_s(z_t)-\beta_t|dK_t \le \int_0^\infty \varepsilon_t^{1/2}\gamma_t^{-\delta}\left|\frac{v_t(z_t)}{z_t^2}\right||z_t|dK_t$$

$$\le \text{const}(\omega)\int_0^\infty \varepsilon_t^{1/2}\gamma_t^{-2\delta}\left|\frac{v_t(z_t)}{z_t^2}\right|dK_t \le \text{const}(\omega)\int_0^\infty \varepsilon_t^{1/2}\gamma_t^{-2\delta}|v_t|dK_t < \infty.$$

The corollary is proved. □

3.3.2. Corollary. *Let $\ell_t(u)-\ell_t = \omega_t(u)$, where, for each $t\in[0,\infty)$,*

$$\left|\frac{\omega_t(u)}{u} - \omega_t\right| \to 0 \quad as \quad u\to 0 \quad P\text{-a.s.}$$

Assume that the following condition is satisfied:

(ii′) *there exists δ, $0<\delta<\delta_0$, such that*

$$\frac{1}{\langle L\rangle_t}\int_0^t \Gamma_s^2\gamma_s^{-\delta}|\omega_s|^2 ds \to 0 \quad as \quad t\to\infty \quad P\text{-a.s.}$$

Then condition (ii) *of Theorem* 3.3.1 *is satisfied.*

Proof. For all δ, $0<\delta<\delta_0$, we have

$$\langle N\rangle_t = \int_0^t \Gamma_s^2(\ell_s(z_s)-\ell_s)^2 dK_s = \int_0^t \Gamma_s^2\Big(\frac{\ell_s(z_s)-\ell_s}{z_s}\Big)^2 z_s^2 dK_s \le \text{const}(\omega)\int_0^t \Gamma_s^2\gamma_s^{-\delta}|\omega_s|^2 ds,$$

since $\gamma_t^\delta z_t^2 \to 0$ as $t \to \infty$ P-a.s., and

$$\Big|\frac{\ell_t(z_t) - \ell_t}{z_t} - \omega_t\Big| = \Big|\frac{\omega_t(z_t)}{(z_t)} - \omega_t\Big| \to 0 \quad \text{as} \quad t \to \infty.$$

Finally, we can conclude that the assertion of Theorem 3.3.1 holds if we replace conditions (i) and (ii) by (i′) and (ii′), respectively. □

3.3.1. Example. Averaging procedure for the RM stochastic approximation algorithm with slowly varying gain.

Let $H_t(u) = \frac{\alpha}{(1+K_t)^r} R(u)$, where $\frac{1}{2} < r < 1$, $R(u) = -\beta u + v(u)$, and $v(u) = 0(u^2)$ as $u \to 0$, $\ell_t = \frac{\sigma_t}{(1+K_t)^r}$, σ_t^2 is deterministic, $\sigma_t^2 \to \sigma^2$ as $t \to \infty$, and $K = (K_t)$ is a continuous increasing function with $K_\infty = \infty$. That is, we consider the following SDE:

$$z_t = z_0 + \int_0^t \frac{\alpha}{(1+K_s)^r} R(z_s)\, dK_s + \int_0^t \frac{\sigma_t}{(1+K_t)^r}\, dm_t$$

with $d\langle m\rangle_t = dK_t$.

If $r > \frac{4}{5}$, then, according to Example 2.3.1,

$$(1+K_t)^{r/2} z_t \xrightarrow{d} \sqrt{\frac{\alpha\sigma^2}{2\beta}}\,\xi \quad \text{as} \quad t \to \infty, \quad \xi \in N(0,1),$$

and, moreover, for all δ, $0 < \delta < \frac{\delta_0}{2}$, $\delta_0 = 2 - \frac{1}{r}$,

$$(1+K_t)^\delta z_t \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.},$$

Thus, for the convergence

$$(1+K_t)^{1/2} \overline{z}_t \xrightarrow{d} \sqrt{\frac{\sigma^2}{\beta^2}}\,\xi \quad \text{as} \quad t \to \infty, \quad \xi \in N(0,1),$$

it suffices to verify condition (i′) of Theorem 3.3.1, since condition (ii) is trivially satisfied.

In this example, the object $v_t(u)$ defined in Corollary 3.3.1 is

$$v_t(u) = \frac{\alpha v(u)}{(1+K_t)^r},$$

and for condition (i') of Corollary 3.3.1 to be satisfied, it sufficed to require the following: there exists δ, $0 < \delta < \delta_0$, $\delta_0 = 2 - \frac{1}{r}$, such that

$$\int_0^t (1+K_t)^{1/2}(1+K_t)^{-2\delta}(1+K_t)^{-r} dK_t < \infty,$$

or, equivalently, there exists δ, $0 < \delta < \delta_0$, $\delta_0 = r - \frac{1}{r}$ such that $r(1+\delta) - \frac{1}{2} > 1$.

It is not difficult to verify that if $r > \frac{5}{6}$, then such a δ exists.

REFERENCES

1. A. E. Albert and L. A. Gardner, Jr., “Stochastic approximations and nonlinear regression,” *M.I.T. Press Research Monograph*, **42**, The M.I.T. Press, Cambridge, (1967).
2. A. Le Breton, “About the averaging approach in Gaussian schemes for stochastic approximation,” *Math. Methods Statist.*, **2**, No. 4, 295–315 (1993).
3. A. Le Breton, “About Gaussian schemes in stochastic approximation,” *Stochastic Process. Appl.*, **50**, No. 1, 101–115 (1994).

4. A. Le Breton and A. A. Novikov, “Averaging for estimating covariances in stochastic approximation,” *Math. Methods Statist.*, **3**, No. 3, 244–266 (1994).
5. H. F. Chen, “Asymptotically efficient stochastic approximation,” *Stochastics Stochastics Rep.*, **45**, Nos. 1-2, 1–16 (1993).
6. B. Delyon and A. Juditsky, “Stochastic optimization with averaging of trajectories,” *Stochastics Stochastics Rep.*, **39**, Nos. 2-3, 107–118 (1992).
7. V. Fabian, “On asymptotically efficient recursive estimation,” *Ann. Statist.*, **6**, No. 4, 854–866 (1978).
8. L. I. Gal’chuk, “On the existence and uniqueness of solutions of stochastic equations with respect to semimartingales,” *Teor. Veroyatn. Primen.*, **23**, No. 4, 782–795 (1978).
9. I. I. Gikhman and A. V. Skorokhod, *Stochastic Differential Equations and Their Applications* [in Russian], Naukova Dumka, Kiev (1982).
10. A. A. Gushchin, “Asymptotic optimality of parameter estimators under the LAQ condition,” *Teor. Veroyatn. Primen.*, **40**, No. 2, 286–300 (1995). (1996).
11. P. Hall and C. C. Heyde, “Martingale limit theory and its application,” in: *Probability and Mathematical Statistics*, Academic Press, New York–London (1980).
12. J. Jacod, “Calcul stochastique et problèmes de martingales,” *Lect. Notes Math.*, **714** (1979).
13. J. Jacod and J. Mémin, “Weak and strong solutions of stochastic differential equations: existence and stability,” *Lect. Notes Math.*, **851** (1981).
14. J. Jacod and A. N. Shiryaev, “Limit theorems for stochastic processes,” *Grundlehren Math. Wiss.*, **288** (1987).
15. J. Jacod, “Regularity, partial regularity, partial information process for a filtered statistical model,” *Probab. Theory Related Fields*, **86**, No. 3, 305–335 (1990).
16. Yu. M. Kabanov, R. Sh. Litscer, and A. N. Shiryaev, “Absolute continuity and singularity of locally absolutely continuous probability distributions, I,” *Mat. Sb.*, **107**, No. 3, 364–415 (1978).
17. H. J. Kushner and G. G. Yin, “Stochastic approximation algorithms and applications,” *Appl. Math.*, **35** (1997).
18. N. L. Lazrieva and T. A. Toronjadze, “Ito–Ventzel’s formula for semimartingales, asymptotic properties of MLE and recursive estimation,” *Lect. Notes Control Inf. Sci.*, **96** (1987), pp. 346–355.
19. N. Lazrieva, T. Sharia, and T. Toronjadze, “The Robbins–Monro type stochastic differential equations, I. Convergence of solutions,” *Stochastics Stochastics Rep.*, **61**, Nos. 1-2, 67–87 (1997).
20. N. Lazrieva, T. Sharia, and T. Toronjadze, “The Robbins–Monro type stochastic differential equations, II. Asymptotic behavior of solutions,” *Stochastics Stochastics Rep.*, **75**, No. 3, 153–180 (2003).
21. N. Lazrieva and T. Toronjadze, “The Polyak weighted averaging procedure for Robbins–Monro type SDE,” *Proc. Razmadze Math. Inst.*, **124**, 115–130 (2000).
22. E. Lenglart, “Sur la convergence presque sure des martingales locales,” *C. R. Acad. Sci. Paris Sér. A-B*, **284**, No. 17, A1085–A1088 (1977).
23. D. Levanony, A. Shwartz, and O. Zeitouni, “Recursive identification in continuous-time stochastic processes,” *Stochastic Process. Appl.*, **49**, No. 2, 245–275 (1994).
24. R. Sh. Liptser, “A strong law of large numbers for local martingales,” *Stochastics*, **3**, No. 3, 217–228 (1980).
25. R. Sh. Liptser and A. N. Shiryaev, *Martingale Theory* [in Russian], Nauka, Moscow (1986).
26. L. Ljung, G. Pflug, and H. Walk, “Stochastic approximation and optimization of random systems,” in: *DMV Seminar,* **17**, Birkhäuser (1992).
27. L. Ljung, “Recursive least-squares and accelerated convergence in stochastic approximation schemes,” *Int. J. Adapt. Control Signal Process.*, **15**, No. 2, 169–178 (2001).
28. A. V. Melnikov, “Stochastic approximation procedures for semimartingales,” in: *Statistics and Control of Random Processes* [in Russian], Nauka, Moscow (1989), pp. 147–156.

29. A. V. Melnikov and A. E. Rodkina, "Martingale approach to the procedures of stochastic approximation," in: *Proc. Third Finnish-Soviet Symp. on Probability Theory and Mathematical Statistics, Turku, Finland, August 13-16, 1991*, VSP. Front. Pure Appl. Probab., **1**, Utrecht (1993), pp. 165–182.
30. A. V. Melnikov, A. E. Rodkina, and E. Valkeila, "On a general class of stochastic approximation algorithms," in: *Proc. Third Finnish-Soviet Symp. on Probability Theory and Mathematical Statistics, Turku, Finland, August 13-16, 1991*, VSP. Front. Pure Appl. Probab., **1**, Utrecht (1993), pp. 183–196.
31. A. V. Melnikov and E. Valkeila, "Martingale models of stochastic approximation and their convergence," *Teor. Veroyatn. Primen.*, **44**, No. 2, 278–311 (1999).
32. M. B. Nevelson and R. Z. Khas'minski, "Stochastic approximation and recurrent estimation," in: *Monogr. Probability Theory and Mathematical Statistics* [in Russian], Nauka, Moscow (1972).
33. B. T. Polyak, "A new method of stochastic approximation type," *Avtomat. Telemekh.*, **7**, 98–107 (1990).
34. B. T. Polyak and A. B. Juditsky, "Acceleration of stochastic approximation by averaging," *SIAM J. Control Optim.*, **30**, No. 4, 838–855 (1992).
35. B. L. S. Prakasa Rao, "Semimartingales and their statistical inference," in: *Monogr. on Statistics and Applied Probability*, **83** (1999).
36. H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statistics*, **22**, 400–407 (1951).
37. H. Robbins and D. Siegmund, "A convergence theorem for nonnegative almost supermartingales and some applications," in: *Optimizing Methods in Statistics. Proc. Sympos., Ohio State Univ., Columbus, Ohio, 1971*, Academic Press, New York (1971), pp. 233–257.
38. D. Ruppert, "Efficient estimations from a slowly convergent Robbins–Monro process," *Tech. Rep.* 781, *School of Oper. Res. and Indust. Eng.*, Cornell Univ. (1988).
39. T. Sharia, "On the recursive parameter estimation in the general discrete time statistical model," *Stochastic Process. Appl.*, **73**, No. 2, 151–172 (1998).
40. P. Spreij, "Recursive approximate maximum likelihood estimation for a class of counting process models," *J. Multivariate Anal.*, **39**, No. 2, 236–245 (1991).
41. G. Yin and I. Gupta, "On a continuous time stochastic approximation problem. Stochastic optimization," *Acta Appl. Math.*, **33**, No. 1, 3–20 (1993).
42. G. Yin, "Stochastic approximation: theory and applications," in: *Handbook of Stochastic Analysis and Applications. Statist. Textbooks Monogr.*, **163**, Dekker, New York (2002), pp. 577–624.

N. Lazrieva
Georgian–American University, Business School, Tbilisi, Georgia;
A. Razmadze Mathematical Institute, Tbilisi, Georgia
T. Sharia
Department of Mathematics, Royal Holloway, University of London
E-mail: t.sharia@shul.ac.uk
T. Toronjadze
Georgian–American University, Business School, Tbilisi, Georgia;
A. Razmadze Mathematical Institute, Tbilisi, Georgia
E-mail: toronj333@yahoo.com