

A Conditional Mean Square Estimate for the Solution of a SDE

Andriy Yurachkivsky¹

¹ Taras Shevchenko National University, Kyiv, Ukraine

Correspondence: Andriy Yurachkivsky, Taras Shevchenko National University, 64 Volodymyrska St., Kyiv 01601, Ukraine. E-mail: andriy.yurachkivsky@gmail.com

Received: July 2, 2014 Accepted: August 2, 2014 Online Published: August 14, 2014

doi:10.5539/ijsp.v3n4p1

URL: <http://dx.doi.org/10.5539/ijsp.v3n4p1>

Abstract

Let $\mathbb{F} = (\mathcal{F}(t), t \in \mathbb{R}_+)$ be a filtration on some probability space, and X be the strong solution of the equation $X(t) = \hat{X} + \int_0^t Q(s, X(s))d\iota(s) + \int_0^t \sigma(s, X(s-))dY(s)$, where \hat{X} is an $\mathcal{F}(0)$ -measurable \mathbb{R}^d -valued random variable, ι is a continuous increasing process with $\mathcal{F}(0)$ -measurable values at all times, Y is an \mathbb{R}^m -valued locally square integrable martingale with respect to \mathbb{F} subjected to some mild additional demands, Q and σ are continuous in $x \in \mathbb{R}^d$ random functions on $\mathbb{R}_+ \times \mathbb{R}^d$ (the former \mathbb{R}^d -valued and \mathbb{F} -progressive in $(\omega, t) \in \Omega \times \mathbb{R}_+$, the latter $(d \times m)$ -matrix-valued and \mathbb{F} -predictable). Suppose also that there exists an $\mathcal{F}(0) \otimes \mathcal{B}_+$ -measurable in (ω, t) nonnegative random process ψ such that, for all t, x , $x^\top Q(t, x) \leq -\psi(t)|x|^2$ and $\int_0^t \psi(s)d\iota(s) < \infty$. Under these assumptions, $E(|X(t)|^2 | \mathcal{F}(0))$ is evaluated from above.

Keywords: conditional expectation, stochastic equation, mean square stability

Mathematics Subject Classification: 60H10, 60G44

1. Introduction

The random processes under consideration are assumed given on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let \mathcal{F}^0 be a sub- σ -algebra of \mathcal{F} . We introduce the notation: $E^0 = E(\cdot | \mathcal{F}^0)$ (the definition of conditional expectation, in particular E^0 , adopted in this article does not demand finiteness of first absolute moment—see Section 2); \mathcal{V}_0^+ is the class of all increasing from zero numeral random processes whose values at all times are \mathcal{F}^0 -measurable random variables, $\mathcal{V}_0^{+,c}$ is its subclass of continuous processes. If, besides, a filtration $\mathbb{F} = (\mathcal{F}(t), t \in \mathbb{R}_+)$ is given, then we identify \mathcal{F}^0 with $\mathcal{F}(0)$. By $\ell\mathcal{M}_2$ we denote, following (Gikhman & Skorokhod, 1982, 2009), the class of all \mathbb{R}^m -valued (m will be determined by context, if matters) locally square integrable martingales w.r.t. \mathbb{F} .

Let X be the strong solution of a stochastic differential equation of the kind

$$X(t) = \hat{X} + \int_0^t Q(s, X(s))d\iota(s) + \int_0^t \sigma(s, X(s-))dY(s),$$

where $\iota \in \mathcal{V}_0^{+,c}$ and Y is chosen from some subclass of $\ell\mathcal{M}_2$ which is constructed and studied in Section 2 (and was introduced in (Yurachkivsky, 2013a, 2013b)). The goal of this article is to find an upper bound, much more exact than that provided by the Gronwall–Bellman lemma, for $E^0|X(t)|^2$. This is done in Section 3 containing the final result of the article together with its application to stability theory. In the case $\sigma = \text{const}$, an estimate for $E^0|X(t)|^2$ was found in (Yurachkivsky, 2013a), and it is the starting point for our present research. Sections 1 and 2 contain preparatory technical results of which the Fubini-type theorem for conditional expectations (Proposition 2.3) may be of interest on its own right.

Stability of order p of the solution of a SDE is usually studied with the aid of Lyapunov's functions (see, e.g., Khraminsky, 2012; Shen & Sun, 2011). But this approach is fruitful only when Q, σ and ι are nonrandom and Y has independent increments, so that X is a Markov process. Our Theorem 3.1 yields, as a byproduct, sufficient conditions for mean square stability without these restrictive assumptions. The only specific condition of that theorem—inequality (26)—is of quadratic nature, which explains why the theorem concerns only the case $p = 2$.

All vectors are thought of as columns; \int_a^b means $\int_{[a,b]}$. We use the Euclidean norm $|\cdot|$ of vectors and the operator norm $\|\cdot\|$ of matrices ($\|A\| = \sup_{|x| \leq 1} |Ax|$). The symbols $\text{Matr}_{d \times m}$, \mathfrak{S} and \mathfrak{S}_+ signify the class of all $d \times m$ matrices

with real entries, the class of all symmetric $m \times m$ matrices with real entries and its subclass of nonnegative (in the spectral sense) matrices, respectively. For $A, B \in \mathfrak{S}$, the inequality $A \leq B$ means that $B - A \in \mathfrak{S}_+$ (so that one may speak about increasing \mathfrak{S} -valued functions). The words “almost surely” are tacitly implied in relations between random variables, including the convergence relation unless it is explicitly written as the convergence in probability. Indicators are denoted by I with two possible modes of writing the set: I_B or $I\{\cdot \dots\}$.

The reference books for the notions and results of stochastic analysis used in this paper are (Elliott, 1982; Gikhman & Skorokhod, 1982, 2009; Jacod & Shiryaev, 1987; Liptser & Shiryaev, 1989).

2. Deterministic Preliminaries

Lemma 1.1 For any $S \in \mathfrak{S}$ and $B \in \text{Matr}_{d \times m}$, $\text{tr } BS B^\top \leq \|B\|^2 \text{tr } S$.

Proof. By a familiar property of symmetric matrices there exist real numbers $\lambda_1, \dots, \lambda_m$ and an orthonormal basis h_1, \dots, h_m in \mathbb{R}^m such that $S = \lambda_1 h_1 h_1^\top + \dots + \lambda_m h_m h_m^\top$. Let us take an arbitrary orthonormal basis e_1, \dots, e_d in \mathbb{R}^d . Then for any symmetric $d \times d$ matrix A one has $\text{tr } A = e_1^\top A e_1 + \dots + e_d^\top A e_d$. In particular,

$$\text{tr } BS B^\top = \sum_{i=1}^d e_i^\top BS B^\top e_i = \sum_{i=1}^d e_i^\top \sum_{j=1}^m \lambda_j B h_j h_j^\top B^\top e_i = \sum_{j=1}^m \lambda_j \sum_{i=1}^d e_i^\top (B h_j)(B h_j)^\top e_i,$$

i.e., $\text{tr } BS B^\top = \sum_{j=1}^m \lambda_j \text{tr} \left((B h_j)(B h_j)^\top \right)$. It remains to note that, firstly, for any $x \in \mathbb{R}^d$ $\text{tr } xx^\top = |x|^2$ and, secondly, $|B h_j| \leq \|B\|$ since $|h_j| = 1$. □

In the subsequent three statements, H is a Borel $\text{Matr}_{d \times m}$ -valued function and K is an increasing continuous \mathfrak{S} -valued function, both defined on $]a, b[\subset \mathbb{R}_+$. The (i, j) th entry of the the matrix $H(s)$ (respectively $K(s)$) will be denoted by $h_{ij}(s)$ (respectively $\kappa_{ij}(s)$). For an arbitrary natural n , we denote $\underline{n} = \{1, \dots, n\}$.

Lemma 1.2 Suppose that

$$\int_a^b \|H(s)\|^2 d \text{tr } K(s) < \infty. \tag{1}$$

Then for all $i \in \underline{d}$ and $\lambda, \mu, \nu \in \underline{m}$

$$\int_a^b h_{i\lambda}(s)^2 d\kappa_{\mu\nu}(s) < \infty. \tag{2}$$

Proof. Obviously, the module of each entry of any matrix does not exceed the operator norm of the latter. So to deduce (2) from (1) it suffices to show that for all $\mu, \nu \in \underline{m}$, $s \in]a, b[$ and $t \in]s, b[$

$$|\kappa_{\mu\nu}(t) - \kappa_{\mu\nu}(s)| \leq \text{tr } K(t) - \text{tr } K(s). \tag{3}$$

And this follows from the above-stated and the inequality $\|S\| \leq \text{tr } S$ for an arbitrary $S \in \mathfrak{S}_+$ (the norm of such a matrix equals to its greatest eigenvalue). □

Corollary 1.1 Under condition (1), the integrals $\int_a^b h_{i\mu}(s) d\kappa_{\mu\nu}(s) h_{j\nu}(s)$, $i, j \in \underline{d}, \mu, \nu \in \underline{m}$, are well-defined and hereon the integral $\int_a^b H(s) dK(s) H(s)^\top$ exists.

Lemma 1.3 Condition (1) implies that

$$\text{tr} \int_a^b H(s) dK(s) H(s)^\top \leq \int_a^b \|H(s)\|^2 d \text{tr } K(s). \tag{4}$$

Proof. For continuous H , the integrals on both sides of (4) are the limits of the Riemann–Stieltjes integral sums, so in this case the inequality is immediate from Lemma 1.1.

By the dominated convergence theorem the class of those functions H which satisfy (4) contains the limit of every pointwise convergent and uniformly bounded sequence of its elements. So it contains all bounded Borel functions, since each of them arises from continuous ones by virtue of at most countably many bounded pointwise passages to the limit. In case $\|H\|$ is unbounded we introduce the functions

$$H_n(s) = \frac{nH(s)}{n \vee \|H(s)\|}.$$

By construction $\|H_n(s)\| \leq n \wedge \|H(s)\|$, so by what was proved

$$\text{tr} \int_a^b H_n(s) dK(s) H_n(s)^\top \leq \int_a^b \|H(s)\|^2 d \text{tr} K(s).$$

It remains to show that

$$\int_a^b H_n(s) dK(s) H_n(s)^\top \rightarrow \int_a^b H(s) dK(s) H(s)^\top. \quad (5)$$

Denote the (k, p) th entry of $H_n(s)$ by $h_{nkp}(s)$. From (3) and the inequality $\|H_n(s)\| \leq \|H(s)\|$ we have

$$\left| \int_a^b h_{ni\mu}(s) h_{nj\nu}(s) d\kappa_{\mu\nu}(s) - \int_a^b h_{i\mu}(s) h_{j\nu}(s) d\kappa_{\mu\nu}(s) \right| \leq \int_a^b g_{nk}(s) d \text{tr} K(s), \quad (6)$$

where $\mathbf{k} = (i, j, \mu, \nu)$, $g_{nk} = |h_{ni\mu} h_{nj\nu} - h_{i\mu} h_{j\nu}|$. Obviously, $g_{nk} \leq 2\|H_n\|^2 + 2\|H\|^2$ and all the more $g_{nk} \leq 4\|H\|^2$. Herein $\lim_{n \rightarrow \infty} g_{nk}(s) = 0$ (since, evidently, $\lim_{n \rightarrow \infty} H_n(s) = H(s)$), which together with the last inequality and (1) yields by the dominated convergence theorem

$$\lim_{n \rightarrow \infty} \int_a^b g_{nk} d \text{tr} K(s) = 0.$$

And this jointly with (6) entails (5). \square

Lemma 1.4 Let f be a continuous at zero \mathbb{R}^d -valued function on \mathbb{R}^d such that $x^\top f(x) \leq c|x|^2$ for some $c \in \mathbb{R}$ and all $x \in \mathbb{R}^d$. Then $f(0) = 0$.

Proof. Denote $q = f(0)$. By the assumption $q^\top f(\varepsilon q) \leq c\varepsilon|q|^2$ for all $\varepsilon > 0$. Hence, letting $\varepsilon \rightarrow 0$ and taking to account continuity of f at zero, we get $|q|^2 \leq 0$. \square

2. Probabilistic Preliminaries

2.1 Extended Conditional Expectations

The definition of conditional expectation adopted in this article is due to Meyer (Shiryaev, 1996, Ch. II, § 7). It admits existence of the conditional expectation of a random variable with infinite first absolute moment. In this subsection, we recall some properties of thus generalized conditional expectation and prove several statements concerning this notion.

Denote $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{\infty\}$ and, for $a \in \mathbb{R}$, $a_+ = a \vee 0$, $a_- = -(a \wedge 0)$, so that $a = a_+ - a_-$. In what follows, “nonnegative” means “ $\overline{\mathbb{R}}_+$ -valued” (the value ∞ is not admitted).

Let \mathcal{G} be a sub- σ -algebra of \mathcal{F} . The conditional given \mathcal{G} expectation of an $\overline{\mathbb{R}}_+$ -valued random variable γ is defined, according to (Shiryaev, 1996), as the \mathcal{G} -measurable $\overline{\mathbb{R}}_+$ -valued random variable $E(\gamma|\mathcal{G})$ such that $E\gamma I_G = E(E(\gamma|\mathcal{G})I_G)$ for every $G \in \mathcal{G}$. For an \mathbb{R} -valued random variable γ such that $P\{E(\gamma_+|\mathcal{G}) = \infty = E(\gamma_-|\mathcal{G})\} = 0$ we set by definition $E(\gamma|\mathcal{G}) = E(\gamma_+|\mathcal{G}) - E(\gamma_-|\mathcal{G})$. Further $E(\gamma|\mathcal{G})$ is defined in the obvious way for \mathbb{R}^d -valued (and even \mathbb{C}^d -valued if one needs) γ . Thus defined (on some extension of $L_1(\Omega, \mathcal{F}, P)$) conditional expectation is called *extended*.

It is immediate from the above definition that $E(c\gamma|\mathcal{G}) = cE(\gamma|\mathcal{G})$ for every $c \in \mathbb{R}$ and \mathbb{R}^d -valued random variable γ such that $E(|\gamma| | \mathcal{G}) < \infty$. In particular, $E(-\gamma|\mathcal{G}) = -E(\gamma|\mathcal{G})$ for such γ . This together with Lemma 2.11 in (Yurachivsky, 2013a) leads to the following conclusion.

Proposition 2.1 Let α and β be \mathbb{R}^d -valued random variables such that $E(|\alpha| | \mathcal{G}) < \infty$, $E(|\beta| | \mathcal{G}) < \infty$. Then $E(\alpha \pm \beta|\mathcal{G}) = E(\alpha|\mathcal{G}) \pm E(\beta|\mathcal{G})$.

Lemma 2.1 (Yurachivsky, 2013a, Lemma 2.3) Let α and β be nonnegative random variables such that $\alpha \leq \beta$. Then $E(\alpha|\mathcal{G}) \leq E(\beta|\mathcal{G})$.

Corollary 2.1 Let $E(|\gamma| | \mathcal{G}) < \infty$. Then $|E(\gamma|\mathcal{G})| \leq E(|\gamma| | \mathcal{G})$.

Lemma 2.2 (Yurachivsky, 2013a, Lemma 2.13) Let α and β be nonnegative random variables, α \mathcal{G} -measurable. Then $E(\alpha\beta|\mathcal{G}) = \alpha E(\beta|\mathcal{G})$.

Lemma 2.3 (Yurachivsky, 2013a, Lemma 2.14; Yurachivsky, 2013b, Lemma 1.9) Let α and β be real-valued random variables such that $E(|\beta| | \mathcal{G}) < \infty$ and α is \mathcal{G} -measurable. Then $E(\alpha\beta|\mathcal{G}) = \alpha E(\beta|\mathcal{G})$.

Lemma 2.4 (Yurachkivsky, 2013a, Corollary 2.8) *For every sequence (γ_n) of nonnegative random variables one has $E(\lim \gamma | \mathcal{G}) \leq \lim E(\gamma | \mathcal{G})$.*

In this subsection, we consider \mathbb{R} -valued random processes. The total variation on $[a, b]$ of a function f will be denoted by $\text{var}_{[a,b]} f$.

The following statement is immediate from Lemma 2.1, Proposition 2.1, Corollary 2.1 and the definition of total variation.

Lemma 2.5 *Let F be a random process on $[a, b]$ such that $\text{var}_{[a,b]} F$ is a random variable (i. e., an \mathcal{F} -measurable function of $\omega \in \Omega$) and*

$$E^0 \left(|F(a)| + \text{var}_{[a,b]} F \right) < \infty. \tag{7}$$

Then: (i) $E^0 F(s)$ exists for all $s \in [a, b]$; (ii) the inequality $\text{var}_{[a,b]} E^0 F \leq E^0 \text{var}_{[a,b]} F$ holds.

Lemma 2.6 *Let ξ be a càdlàg random process on $[a, b]$. Assume that there exists a random variable Γ such that*

$$\sup_{a < s \leq b} |\xi(s)| \leq \Gamma \tag{8}$$

and $E^0 \Gamma < \infty$. Then $\lim_{s \rightarrow a+} E^0 \xi(s) = E^0 \xi(a)$.

Proof. Denote $\Xi(t) = \sup_{a < s \leq t} |\xi(s) - \xi(a)|$. Any càdlàg function is determined by its values on a dense subset of $[a, b]$, so the supremum may be taken over $s \in [a, t] \cap \mathbb{Q}$. Thus $\Xi(t)$ is an \mathcal{F} -measurable function of $\omega \in \Omega$, which entitles us to define $\rho(t, \varepsilon)$ by

$$\rho(t, \varepsilon) = E^0 I\{\Xi(t) > \varepsilon\}, \tag{9}$$

so that $P\{\Xi(t) > \varepsilon\} = E\rho(t, \varepsilon)$. Right-continuity of ξ implies that, for any $\varepsilon > 0$, the left-hand side of the last equality tends to zero as $t \rightarrow a+$. Hence, taking into account that $\rho(\cdot, \varepsilon)$ is an increasing process (since Ξ increases by construction and the operator E^0 is isotonic by Lemma 2.1), we get

$$\lim_{t \rightarrow a+} \rho(t, \varepsilon) = 0. \tag{10}$$

Denote $\Gamma_n = \Gamma I\{\Gamma > n\}$. Writing, for arbitrary $n \in \mathbb{N}$ and $\varepsilon \in]0, 2n]$, the identity

$$1 = I\{\Xi(t) > 2n\} + I\{\varepsilon < \Xi(t) \leq 2n\} + I\{\Xi(t) \leq \varepsilon\},$$

we obtain from (8), (9) and the definition of Ξ (with the use of Proposition 2.1 and Lemma 2.1, of course) $E^0 \Xi(t) \leq 2E^0 \Gamma_n + 2\rho(t, \varepsilon) + \varepsilon$, which together with (10) yields

$$\lim_{t \rightarrow a+} E^0 \Xi(t) \leq 2E^0 \Gamma_n + \varepsilon. \tag{11}$$

By construction the sequence (Γ_n) decreases to zero. By assumption $E^0 \Gamma < \infty$, whence by Lemma 2.1 $E^0 \Gamma_1 < \infty$. Then Lemma 2.9 in (Yurachkivsky, 2013a) asserts that $E^0 \Gamma_n \searrow 0$, which together with (11) where ε is arbitrary yields $E^0 \Xi(t) \rightarrow 0$ as $t \rightarrow a+$. It remains to note that $|E^0 \xi(s) - E^0 \xi(a)| \leq E^0 |\xi(s) - \xi(a)|$ by Corollary 2.1. \square

The proof of Lemma 2.6 will not change if we substitute a by an arbitrary inner point of $[a, b]$, so, under its assumptions,

$$\lim_{s \rightarrow t+} E^0 \xi(s) = E^0 \xi(t)$$

for any $t \in [a, b[$. If, moreover, ξ increases, then by Lemma 2.1 $E^0 \xi(t_1) \leq E^0 \xi(t_2)$ as $t_1 < t_2$. Hence, repeating, up to notation, the proof of Theorem II.7.4 (Shiryaev, 1996), we deduce the following statement (for classical conditional expectations, it is a particular case of that theorem).

Proposition 2.2 *Let F be a right-continuous increasing random process on $[a, b]$ such that for any $s \in [a, b]$, $E^0 |F(s)| < \infty$. Then there exists a right-continuous increasing random process G on $[a, b]$ such that*

$$P\{G(s) = E^0 F(s)\} = 1, \quad s \in [a, b]. \tag{12}$$

Corollary 2.2 *Let a càdlàg random process F on $[a, b]$ satisfy condition (7). Then there exists a càdlàg random process G on $[a, b]$ with property (12).*

Otherwise speaking, Corollary 2.2 asserts existence, under the above stated assumptions, of the càdlàg version of $E^0 F$. In what follows, we consider namely it.

We denote by \mathbb{F}^0 the filtration with $\mathcal{F}(t) = \mathcal{F}(0)$, $t \in \mathbb{R}_+$ (so that a random process is \mathbb{F}^0 -adapted iff its values at all times are $\mathcal{F}(0)$ -measurable random variables). Then \mathcal{V}_0^+ can be defined equivalently as the class of all starting from zero \mathbb{F}^0 -adapted increasing random processes.

Proposition 2.3 *Let F and φ be random processes on $[a, b]$, the former càdlàg and satisfying condition (7), the latter continuous and \mathbb{F}^0 -adapted. Then*

$$\int_a^b \varphi(s) dE^0 F(s) = E^0 \int_a^b \varphi(s) dF(s). \quad (13)$$

Note that the integral on the right-hand side of (13) is, due to continuity of φ , the limit of Riemann-Stieltjes sums and therefore an \mathcal{F} -measurable function of ω .

Proof. 1°. Denote

$$\zeta = \max_{a \leq s \leq b} |\varphi(s)|, \quad \chi = \int_a^b \varphi(s) dF(s);$$

by var we mean var. It follows from the assumptions on φ that, firstly, $\zeta(\omega) < \infty$ for all $\omega \in \Omega$ and, secondly, ζ is an \mathcal{F}^0 -measurable random variable. The evident inequality $|\chi| \leq \zeta \text{ var } F$ and condition (7) imply, by Lemmas 2.1 and 2.2, that $E^0 |\chi| \leq \zeta E^0 \text{ var } F$ and, consequently, $E^0 |\chi| < \infty$. Thus $E^0 \chi$ exists.

2°. Denote $G(s) = E^0 F(s)$. Condition (7) implies that $\text{var } G < \infty$ (by Lemma 2.5) and $\text{var } F < \infty$ (by Corollary 2.5 in (Yurachivsky, 2013a)). Let us impose the interim assumption $\text{var } \varphi < \infty$. Then the integration-by-parts formula yields

$$\chi = \varphi(b)F(b) - \varphi(a)F(a) - \int_a^b F(s) d\varphi(s). \quad (14)$$

\mathbb{F}^0 -adaptedness of φ implies by Lemma 2.2 that $E^0 |\varphi(s)F(s)| = |\varphi(s)|E^0 |F(s)|$. Hence and from finiteness of $E^0 |F(s)|$ established in the proof of Lemma 2.5 we get $E^0 |\varphi(s)F(s)| < \infty$, $s \in [a, b]$. Then it follows from (14) and item 1°, via Proposition 2.1 and Lemma 2.3, that

$$E^0 \chi = \varphi(b)G(b) - \varphi(a)G(a) - E^0 \int_a^b F(s) d\varphi(s). \quad (15)$$

The evident inequality $|F(s)| \leq |F(a)| + \text{var } F$ implies by Lemma 2.1 that $E^0 |F(s)| \leq E^0 (|F(a)| + \text{var } F)$, which together with (7) and the interim assumption $\text{var } \varphi < \infty$ shows that $\int_a^b (E^0 |F(s)|) |d\varphi(s)| < \infty$. Hence, recalling that φ is \mathbb{F}^0 -adapted, we get by Theorem 1.17 in (Yurachivsky, 2013b)

$$E^0 \int_a^b F(s) d\varphi(s) = \int_a^b G(s) d\varphi(s). \quad (16)$$

As was shown above, $\text{var } G < \infty$ and G may be, without loss of generality, considered càdlàg, so the integration-by-parts formula yields $\int_a^b \varphi(s) dG(s) = \varphi(b)G(b) - \varphi(a)G(a) - \int_a^b G(s) d\varphi(s)$, which together with (16) and (15) results in

$$\int_a^b \varphi(s) dG(s) = E^0 \chi. \quad (17)$$

And this is none other than equality (13).

3°. In this item, we impose the interim assumption $\varphi(a) = 0 = \varphi(b)$ instead of $\text{var } \varphi < \infty$. We extend φ to the whole real line, putting $\varphi(s) = 0$ as $s \notin [a, b]$, so that $\int |\varphi(t)| dt < \infty$ (\int means $\int_{\mathbb{R}}$) and φ is uniformly continuous on \mathbb{R} . Let us take an arbitrary nonnegative function $g \in C^1(\mathbb{R})$ such that $\int g(t) dt = 1$ and $g(t) = 0$ as $|t| > 1$. Set $g_n(t) = ng(nt)$,

$$\varphi_n(s) = \int \varphi(t) g_n(s-t) dt \equiv \int \varphi(s-t) g_n(t) dt.$$

By construction all trajectories of φ_n are continuously differentiable and therefore have finite variation on $[a, b]$. So, according to item 1°,

$$\int_a^b \varphi_n(s) dG(s) = E^0 \chi_n, \quad (18)$$

where $\chi_n = \int_a^b \varphi_n(s) dF(s)$ (existence of $E^0 \chi_n$ is justified in the same manner as was done for $E^0 \chi$ in item 1°). Obviously, $\varphi_n \rightrightarrows \varphi$, which together with $\text{var } G < \infty$ yields

$$\int_a^b \varphi_n(s) dG(s) \rightarrow \int_a^b \varphi(s) dG(s). \tag{19}$$

By construction the random process φ_n is \mathbb{F}^0 -adapted, since so is φ . Continuity of both φ and φ_n implies that $\|\varphi_n - \varphi\|_\infty = \sup_{s \in \mathbb{Q}} |\varphi_n(s) - \varphi(s)|$. Thus $\|\varphi_n - \varphi\|_\infty$ is an \mathcal{F}^0 -measurable random variable, whence with account of the evident inequality $|\chi_n - \chi| \leq \|\varphi_n - \varphi\|_\infty \text{var } F$ we get by Lemmas 2.1 and 2.2 $E^0 |\chi_n - \chi| \leq \|\varphi_n - \varphi\|_\infty E^0 \text{var } F$, which together with (7) and the relation $\varphi_n \rightrightarrows \varphi$ yields $E^0 |\chi_n - \chi| \rightarrow 0$. Herein $E^0 \chi_n - E^0 \chi = E^0 (\chi_n - \chi)$ by Proposition 2.1, and $|E^0 (\chi_n - \chi)| \leq E^0 |\chi_n - \chi|$ by Corollary 2.1. Consequently, $E^0 \chi_n \rightarrow E^0 \chi$, which together with (18) and (19) entails (17).

4°. In case $|\varphi(a)| + |\varphi(b)| > 0$ we take an arbitrary \mathbb{F}^0 -adapted continuous random process ϕ such that $\phi(a) = \varphi(a)$, $\phi(b) = \varphi(b)$,

$$\int_a^b \phi(s) dE^0 F(s) = E^0 \int_a^b \phi(s) dF(s) \tag{20}$$

(for example, $\phi(s) = ((b - s)\varphi(a) + (s - a)\varphi(b))/(b - a)$) and put $\varphi_0 = \varphi - \phi$, so that

$$E^0 \int_a^b \varphi(s) dF(s) = E^0 \int_a^b \varphi_0(s) dF(s) + E^0 \int_a^b \phi(s) dF(s). \tag{21}$$

According to item 3° $\int_a^b \varphi_0(s) dE^0 F(s) = E^0 \int_a^b \varphi_0(s) dF(s)$, which together with (20) and (21) entails (13). □

2.2 A Subclass of the Class of Locally Square Integrable Martingales

Let our probability space (Ω, \mathcal{F}, P) be endowed with a right-continuous flow of σ -algebras (or, in the terminology of (Jacod & Shiryaev, 1987; Liptser & Shiryaev, 1989), filtration) $\mathbb{F} = (\mathcal{F}(t), t \in \mathbb{R}_+)$. By \mathcal{K} we denote the class of all \mathbb{F} -adapted \mathbb{R}^m -valued (m will be determined by context, if matters) càdlàg random processes M satisfying the conditions:

M1. For all $t \quad E^0 |M(t)|^2 < \infty$.

M2. For all $t \geq s \geq 0 \quad E(M(t)|\mathcal{F}(s)) = M(s)$.

M3. The process $E^0 |M|^2$ is continuous.

Lemma 2.7 (Yurachkivsky, 2013a, Lemma 3.7) *Let M have properties **M1** and **M2**. Then the equality $E(M(\zeta)|\mathcal{F}(s)) = M(\zeta \wedge s)$ holds for every $s \in \mathbb{R}_+$ and bounded stopping time ζ .*

Proposition 2.4 (Yurachkivsky, 2013a, Theorem 3.9) $\mathcal{K} \subset \ell \mathcal{M}_2$.

For arbitrary random process ξ and nonnegative random variable τ , we denote $\xi^\tau(t) = \xi(t \wedge \tau)$.

Proposition 2.5 *For any $M \in \mathcal{K}$ and stopping time τ , $M^\tau \in \mathcal{K}$.*

Proof. By Proposition 2.4 $M \in \ell \mathcal{M}_2$, so $M^\tau \in \ell \mathcal{M}_2$, too. We consider, without loss of generality, that $M(0) = 0$. Then by Corollary 3.11 in (Yurachkivsky, 2013a)

$$E^0 |M^\tau(t)|^2 = E^0 \text{tr} \langle M^\tau \rangle (t). \tag{22}$$

Herein $\langle M^\tau \rangle = \langle M \rangle^\tau$ by Theorem I.1.19 in (Gikhman & Skorokhod, 2009) and therefore

$$E^0 \text{tr} \langle M^\tau \rangle (t) = E^0 \text{tr} \langle M \rangle (t \wedge \tau). \tag{23}$$

The process $\text{tr} \langle M \rangle$ increases, so by Lemma 2.1 $E^0 \text{tr} \langle M \rangle (t \wedge \tau) \leq E^0 \text{tr} \langle M \rangle (t)$ ($= E^0 |M(t)|^2$ by Corollary 3.11 in (Yurachkivsky, 2013a)). Thus $E^0 |M^\tau(t)|^2 < \infty$, which together with equalities (22) and (23) yields $E^0 |M^\tau(t)|^2 - E^0 |M^\tau(s)|^2 = E^0 \text{tr} \langle M \rangle (t \wedge \tau) - E^0 \text{tr} \langle M \rangle (s \wedge \tau)$ ($= E^0 (\text{tr} \langle M \rangle (t \wedge \tau) - \text{tr} \langle M \rangle (s \wedge \tau))$ by Proposition 2.1). Hence and from the evident inequalities $0 \leq f(t \wedge \tau) - f(s \wedge \tau) \leq f(t) - f(s)$ (for arbitrary $t > s \geq 0$ and increasing function f) we get by Lemma 2.1 and Proposition 2.1 $0 \leq E^0 |M^\tau(t)|^2 - E^0 |M^\tau(s)|^2 \leq E^0 \text{tr} \langle M \rangle (t) - E^0 \text{tr} \langle M \rangle (s)$ ($= E^0 |M(t)|^2 - E^0 |M(s)|^2$ by Corollary 3.11 in (Yurachkivsky, 2013a)). Thus the process $E^0 |M^\tau|^2$ is continuous, since by the choice of M so is $E^0 |M|^2$.

The equality $E^0(M^\tau(t)|\mathcal{F}(s)) = M^\tau(s)$ for $t \geq s$ follows from Lemma 2.7, if one sets $\zeta = t \wedge \tau$ (so that $\zeta \wedge s = \tau \wedge s$). \square

In (Yurachkivsky, 2013a), this statement was implicitly used in the proof of the main result.

3. The Main Result

In this section, we obtain, under appropriate assumptions, a conditional mean square estimate for the solution of the equation

$$X(t) = \dot{X} + \int_0^t Q(s, X(s))d\iota(s) + \int_0^t \sigma(s, X(s-))dZ(s). \quad (24)$$

Along with this equation in its general form we will consider its particular case

$$X(t) = \int_0^t Q(s, X(s))d\iota(s) + N(t). \quad (25)$$

First of all we impose the following assumptions:

S1. For every $Z \in \mathcal{K}$ Equation (24) has the unique strong solution on \mathbb{R}_+ .

S2. For every $N \in \mathcal{K}$ Equation (25) has the unique strong solution on \mathbb{R}_+ .

The Borel σ -algebra in \mathbb{R}_+ will be denoted by \mathcal{B}_+ .

The solution of (25) was evaluated in (Yurachkivsky, 2013a) by virtue of a special lemma cognate to the comparison theorems in (Ikeda & Watanabe, 1981, Ch. VI), without recourse to Lyapunov functions. That inequality underlies the derivation of our main result.

Theorem 3.1 *Let \dot{X} be an $\mathcal{F}(0)$ -measurable \mathbb{R}^d -valued random variable, ι be an \mathbb{F}^0 -adapted increasing continuous random process, Y be an \mathbb{R}^m -valued random process of class \mathcal{K} , Q and σ be continuous in $x \in \mathbb{R}^d$ random functions on $\mathbb{R}_+ \times \mathbb{R}^d$, the former \mathbb{R}^d -valued and \mathbb{F} -progressive in $(\omega, t) \in \Omega \times \mathbb{R}_+$, the latter $(d \times m)$ -matrix-valued and \mathbb{F} -predictable in (ω, t) . Assume the following: conditions **S1** and **S2** are satisfied; there exists an $\mathcal{F}(0) \otimes \mathcal{B}_+$ -measurable in (ω, t) nonnegative random process ψ such that*

$$x^\top Q(t, x) \leq -\psi(t)|x|^2 \quad (26)$$

and

$$\int_0^t \psi(s)d\iota(s) < \infty \quad (27)$$

for all $x \in \mathbb{R}^d$, $t > 0$; there exist random processes $L \in \mathcal{V}_0^+$ and $R \in \mathcal{V}_0^{+,c}$ such that

$$\|\sigma(t, x)\|^2 \leq L(t)|x|^2, \quad (28)$$

$$\text{tr}\langle Y \rangle(t) - \text{tr}\langle Y \rangle(s) \leq R(t) - R(s) \quad (29)$$

and

$$\int_0^t L(u)dR(u) < \infty \quad (30)$$

for all $x \in \mathbb{R}^d$, $t > s \geq 0$.

Then the strong solution of the equation

$$X(t) = \dot{X} + \int_0^t Q(s, X(s))d\iota(s) + \int_0^t \sigma(s, X(s-))dY(s) \quad (31)$$

satisfies, for all t , the inequality

$$E^0|X(t)|^2 \leq |\dot{X}|^2 \exp \left\{ \int_0^t L(s)dR(s) - 2 \int_0^t \psi(s)d\iota(s) \right\}. \quad (32)$$

Proof. 1°. Denote $\tau_n = \inf\{s: |X(s)| \geq n\}$, $Y_n(t) = Y(t \wedge \tau_n)$, so that τ_n is a stopping time,

$$|X(s-)| \leq n \quad \text{as} \quad s \leq \tau_n, \quad (33)$$

and Y_n belongs, by Proposition 2.5 and the choice of Y , to \mathcal{K} (and all the more to $\ell\mathcal{M}_2$ —see Proposition 2.4).

Let X_n be the strong solution of the equation

$$X_n(t) = \dot{X} + \int_0^t Q(s, X_n(s))dt(s) + \int_0^t \sigma(s, X_n(s-))dY_n(s) \quad (34)$$

(condition **S1** justifies this definition). Denote $\Xi_n(s) = \sigma(s, X_n(s-))$, $U_n(t) = \int_0^t \|\Xi_n(s)\|^2 d\text{tr}\langle Y_n \rangle(s)$, $T(t) = \int_0^t L(s)dR(s)$,

$$M_n(t) = \dot{X} + \int_0^t \Xi_n(s)dY_n(s), \quad (35)$$

so that

$$X_n(t) = \int_0^t Q(s, X_n(s))dt(s) + M_n(t). \quad (36)$$

Obviously, processes of class \mathcal{V}_0^+ (in particular L and R) are $\mathcal{F}(0) \otimes \mathcal{B}_+$ -measurable in (ω, t) . Then T possesses this property, too, and is therefore \mathbb{F}^0 -adapted.

By Theorem I.1.19 in (Gikhman & Skorokhod, 2009) $\langle Y_n \rangle(t) = \langle Y \rangle(t \wedge \tau_n)$. Consequently,

$$U_n(t) = \int_0^{t \wedge \tau_n} \|\Xi_n(s)\|^2 d\text{tr}\langle Y \rangle(s),$$

whence in view of (28), (29) and (33) we have for any $t_2 > t_1 \geq 0$

$$U_n(t_2) - U_n(t_1) \leq n^2(T(t_2) - T(t_1)). \quad (37)$$

Hence and from \mathbb{F}^0 -adaptedness of T we get by Lemmas 2.1 and 2.3

$$\mathbb{E}^0(U_n(t_2) - U_n(t_1)) \leq n^2(T(t_2) - T(t_1)). \quad (38)$$

In particular, $\mathbb{E}^0 U_n(t) < \infty$ for any t , whence by Proposition 2.1

$$\mathbb{E}^0(U_n(t_2) - U_n(t_1)) = \mathbb{E}^0 U_n(t_2) - \mathbb{E}^0 U_n(t_1). \quad (39)$$

The process R is, by assumption, continuous, so condition (30) implies continuity of T and therefore, in view of (39) and (38), continuity of $\mathbb{E}^0 U_n$. Then by Lemma 3.12 in (Yurachivsky, 2013a) (applied to H_n row-wise) $M_n \in \mathcal{K}$, hereon Theorem 4.3 in (Yurachivsky, 2013a) asserts that, under the above assumptions (including **S2**) on Q ,

$$\mathbb{E}^0 |X_n(t)|^2 \leq e^{-\Psi(t)} \left(|M_n(0)|^2 + \int_0^t e^{\Psi(s)} d\mathbb{E}^0 \text{tr}\langle M_n \rangle(s) \right), \quad (40)$$

where

$$\Psi(t) = 2 \int_0^t \psi(s)dt(s) \quad (41)$$

($< \infty$ by condition (27)). It remains to derive (32) from (40).

2°. As was found out, Y_n and M_n belong to $\ell\mathcal{M}_2$. Inequality (37) and condition (30) show that $U_n(t) < \infty$, or, minutely,

$$\int_0^t \|\Xi_n(s)\|^2 d\text{tr}\langle Y_n \rangle(s) < \infty.$$

Then from (35) we have by Theorem I.4.40 in (Jacod & Shiryaev, 1987)

$$\langle M_n \rangle(t) = \int_0^t \Xi_n(s)d\langle Y_n \rangle(s)\Xi_n(s)^\top.$$

Hereon Lemma 2.7 asserts that for any $b > a \geq 0$

$$\text{tr}\langle M_n \rangle(b) - \text{tr}\langle M_n \rangle(a) \leq U_n(b) - U_n(a). \quad (42)$$

By construction U_n is a nonnegative increasing random process, so

$$|U_n(a)| + \text{var}_{[a,b]} U_n = U_n(b) - U_n(a). \tag{43}$$

Denote $\Phi = e^\Psi$. The process ι , being \mathbb{F}^0 -adapted and continuous, is $\mathcal{F}(0) \otimes \mathcal{B}_+$ -measurable in (ω, t) ; for ψ this was assumed. Then it follows from (41) that Ψ is $\mathcal{F}(0) \otimes \mathcal{B}_+$ -measurable in (ω, t) and all the more \mathbb{F}^0 -adapted. Besides, it is continuous, since so is ι . These properties of Ψ together with relations (43), (38) and (30) imply by Proposition 2.3 that

$$\int_0^t \Phi(s) d\mathbb{E}^0 U_n(s) = \mathbb{E}^0 \int_0^t \Phi(s) dU_n(s). \tag{44}$$

Inequality (42) and finiteness of $\mathbb{E}^0 U_n(t)$ established above imply by Proposition 2.1 and Lemma 2.1 that

$$\mathbb{E}^0 \text{tr}\langle M_n \rangle(b) - \mathbb{E}^0 \text{tr}\langle M_n \rangle(a) \leq \mathbb{E}^0 U_n(b) - \mathbb{E}^0 U_n(a)$$

for all $b > a \geq 0$, which together with (44) yields

$$\int_0^t \Phi(s) d\mathbb{E}^0 \text{tr}\langle M_n \rangle(s) \leq \mathbb{E}^0 \int_0^t \Phi(s) dU_n(s).$$

Recalling the definition of U_n and the equality $\langle Y_n \rangle(t) = \langle Y \rangle(t \wedge \tau_n)$, we rewrite the last inequality in the form

$$\int_0^t \Phi(s) d\mathbb{E}^0 \text{tr}\langle M_n \rangle(s) \leq \mathbb{E}^0 \int_0^{t \wedge \tau_n} \Phi(s) \|\sigma(s, X_n(s-))\|^2 d\text{tr}\langle Y \rangle(s). \tag{45}$$

By condition (28) $\|\sigma(s, X_n(s-))\|^2 \leq L(s)|X_n(s-)|^2$, which together with condition (29) yields, for any nonnegative random variable ρ ,

$$\int_0^\rho \Phi(s) L(s) \|\sigma(s, X_n(s-))\|^2 d\text{tr}\langle Y \rangle(s) \leq \int_0^\rho \Phi(s) L(s) |X_n(s-)|^2 dR(s). \tag{46}$$

Continuity of R implies that

$$\int_0^\rho \Phi(s) L(s) |X_n(s-)|^2 dR(s) = \int_0^\rho \Phi(s) L(s) |X_n(s)|^2 dR(s),$$

which together with (45), (46) and Lemma 2.1 yields

$$\int_0^t \Phi(s) d\mathbb{E}^0 \text{tr}\langle M_n \rangle(s) \leq \mathbb{E}^0 \int_0^t \Phi(s) L(s) |X_n(s)|^2 dR(s). \tag{47}$$

Denoting $V_n = \Phi \mathbb{E}^0 |X_n|^2$ and noting that $M_n(0) = \dot{X}$ because of (35), we get from (40) and (47)

$$V_n(t) \leq |\dot{X}|^2 + \mathbb{E}^0 \int_0^t L(s) \Phi(s) |X_n(s)|^2 dR(s). \tag{48}$$

By assumption $R \in \mathcal{V}_0^+$, so Theorem 2.19 in (Yurachkivsky, 2013a) asserts that

$$\mathbb{E}^0 \int_0^t L(s) \Phi(s) |X_n(s)|^2 dR(s) = \int_0^t \mathbb{E}^0 (L(s) \Phi(s) |X_n(s)|^2) dR(s). \tag{49}$$

It was shown above that the process Ψ (and therefore Φ) is \mathbb{F}^0 -adapted; for L this was assumed. Thus by Lemma 2.2 $\mathbb{E}^0 (L(s) \Phi(s) |X_n(s)|^2) = L(s) V_n(s)$, which together with (48) and the definitions of V_n and T turns (48) into

$$V_n(t) \leq |\dot{X}|^2 + \int_0^t V_n(s) dT(s). \tag{50}$$

By construction T increases and is continuous (since R possesses these properties and L is nonnegative). So (50) yields by the Gronwall–Bellman lemma $V_n(t) \leq |\dot{X}|^2 e^{T(t)}$. Multiplying both sides of this inequality by $e^{-\Psi(t)}$, we get

$$\mathbb{E}^0 |X_n(t)|^2 \leq |\dot{X}|^2 e^{T(t) - \Psi(t)}. \tag{51}$$

Obviously, $\tau_n \rightarrow \infty$ and $X_n(t) = X(t)$ as $t < \tau_n$. So $X_n(t) \rightarrow X(t)$, whence by Lemma 2.4 $E^0|X(t)|^2 \leq \liminf E^0|X_n(t)|^2$. Comparing this with (51), we arrive at (32). \square

This theorem gives rise to some important conclusions concerning stochastic stability of the solution of Equation (31) (various kinds of stochastic stability are defined and studied in (Khasminsky, 2012)). Wishing to stress that this solution depends on the initial value \hat{X} , we shall denote it more minutely than earlier, namely $X(\cdot, \hat{X})$.

Corollary 3.1 *Let the conditions of Theorem 3.1 be fulfilled and let, besides,*

$$\lim_{N \rightarrow \infty} \sup_{t > 0} P \left\{ \int_0^t L(s) dR(s) - 2 \int_0^t \psi(s) dt(s) \geq N \right\} = 0.$$

Then for any $\varepsilon > 0$

$$\sup_{t \geq 0} P \left\{ E^0 \left| X(t, \hat{X}) \right|^2 > \varepsilon \right\} \rightarrow 0 \quad \text{as} \quad \hat{X} \xrightarrow{P} 0. \quad (52)$$

Corollary 3.2 *Let the conditions of Theorem 3.1 be fulfilled and let, besides,*

$$\sup_{t > 0} \left(\int_0^t L(s) dR(s) - 2 \int_0^t \psi(s) dt(s) \right) < \infty.$$

Then for any $\varepsilon > 0$ relation (52) holds.

Note that condition (26) implies, by Lemma 1.4, that $Q(t, 0) = 0$. Likewise $\sigma(t, 0) = 0$ because of (28). Thus $X(t, 0) = 0$. So relation (52) asserts a kind of stochastic stability (which may be called the *conditional mean square stability*) of the trivial solution of (31). The following straightforward consequence of Theorem 3.1 gives sufficient conditions for unconditional mean square stability.

Corollary 3.3 *Let the conditions of Corollary 3.2 be fulfilled with nonrandom ι, ψ, L and R . Then $\sup_{t > 0} E \left| X(t, \hat{X}) \right|^2 \rightarrow 0$ as $E \left| \hat{X} \right|^2 \rightarrow 0$.*

For the very special case when $d = m = 1, \iota(s) = s, Q(s, x) = ax, \sigma(s, x) = bx$, and Y is the standard Wiener process this result is well known (Khasminsky, 2012, Ch. VI, §3) (and almost trivial, since under these assumptions $E \left| X(t, \hat{X}) \right|^2$ can be calculated explicitly).

References

- Elliott, R. J. (1982). *Stochastic Calculus and Applications*. New York: Springer.
- Gikhman, I. I., & Skorokhod, A. V. (1982). *Stochastic Differential Equations and Their Applications* (in Russian). Kiev: Naukova Dumka.
- Gikhman, I. I., & Skorokhod, A. V. (2009). *Theory of Stochastic Processes* (Vol. 3, 2nd ed.). Berlin, Heidelberg: Springer.
- Ikeda, N., & Watanabe, Sh. (1981). *Stochastic Differential Equations and Diffusion Processes*. Amsterdam: North Holland Publishing Company.
- Jacod, J., & Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes*. Berlin: Springer.
- Khasminsky, R. Z. (2012). *Stochastic Stability of Differential Equations* (2nd ed.). Berlin, Heidelberg: Springer.
- Liptser R. Sh., & Shiryaev, A. N. (1989). *Theory of Martingales*. Dordrecht: Kluwer.
- Shen, L. J., & Sun, J. T. (2011). p -th moment exponential stability of stochastic differential equations with impulse effect, *Science China Information Sciences*, 54(8), 1702-1711. <http://dx.doi.org/10.1007/s11432-011-4250-7>
- Shiryaev, A. N. (1996). *Probability*. Berlin: Springer.
- Yurachivsky, A. (2013a). An upper bound for conditional second moment of the solution of a SDE. *Applied Mathematics*, 4(1), 135-143. <http://dx.doi.org/10.4236/am.2013.41023>
- Yurachivsky, A. (2013b). An ergodic-type theorem for generalized Ornstein-Uhlenbeck processes. *Random Operators and Stochastic Equations*, 21(3), 217-269. <http://dx.doi.org/10.1515/rose-2013-0011>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).