DIFFUSION APPROXIMATION OF SYSTEMS WITH WEAKLY ERGODIC MARKOV PERTURBATIONS. II

UDC 519.21

A. YU. VERETENNIKOV AND A. M. KULIK

ABSTRACT. This paper is a continuation of the paper [A. Yu. Veretennikov and A. M. Kulik, Diffusion approximation for systems with weakly ergodic Markov perturbations. I, Theory Probab. Math. Statist. 87 (2012), 13–29]. Some corollaries of the general results are given in several particular cases being of their own interest. An example of a process being a solution of a stochastic differential equation with a Lévy noise is considered; we show that the assumptions imposed on the process can effectively be verified.

1. INTRODUCTION

This is a continuation of the paper [1]. The continuation is devoted to studies of an asymptotic behavior as \( \varepsilon \to 0 \) of the family \( Y^\varepsilon, \varepsilon > 0, \) of solutions of stochastic differential equations whose coefficients depend on the “fast” random component

\[
X^\varepsilon(t) = X(t^{\varepsilon^{-1}}), \quad t \in \mathbb{R}^+,
\]

where \( X \) is a homogeneous Markov process for which the convergence of transition probabilities to the invariant distribution is non-uniform with respect to the initial values and holds, generally speaking, in a weaker mode of convergence than the convergence in variation.

In the paper [1], we introduced the main objects of investigation; also we stated and proved the main result. In the current paper, we provide some corollaries of this result for two particular cases being of their own interest, namely for the case where the process \( X \) admits a bound of the rate of convergence to the invariant distribution in the full variation norm being non-uniform with respect to the initial state of the process and for the case where the process \( X \) admits a bound of the rate of convergence in the Kantorovich–Rubinstein distance.

The central limit theorem for additive functionals of a process \( X \) that admits a bound of the rate of convergence in the Kantorovich–Rubinstein distance is a special case of our results; we compare this central limit theorem with similar results of the paper [2]. As an example, we consider a process \( X \) being a solution of a stochastic differential equation with a Lévy noise and show that the assumptions imposed on the process \( X \) in the framework of our approach can effectively be verified in this case.

We also provide the proofs of auxiliary results of Section 4.2 of the paper [1] (those proofs were omitted in [1] because of lack of place there). The notation used throughout the paper is the same as in [1].
2. Sojourn probabilities for compact sets and estimates for increments of the process $Y^\varepsilon$

In this section, we provide the proofs of the auxiliary results stated in Section 4.2 of the paper [1]. For convenience, we repeat the statements of these assertions.

Lemma 2.1. Let assumptions (1)--(5) of Theorem 3.1 in [1] hold. We also assume that relation (4.4) in [1] holds for an arbitrary function $f \in C^3(\mathbb{R}^m)$ satisfying condition (4.1) in [1].

Then, for an arbitrary $T \in \mathbb{R}^+$,

(2.1) $\sup_{\varepsilon \in (0, 1)} \sup_{t \in [0, T]} \mathbb{P}(|Y^\varepsilon(t)| > R) \rightarrow 0$, $R \rightarrow +\infty$.

Moreover, if $X$ satisfies condition $S(\psi_1, Q_1)$, then

(2.2) $\sup_{\varepsilon \in (0, 1)} \mathbb{P}\left(\sup_{t \in [0, T]} |Y^\varepsilon(t)| > R\right) \rightarrow 0$, $R \rightarrow +\infty$.

Proof. Note that the trajectories of $Y^\varepsilon$ are continuous. Thus relation (2.2) is equivalent to

(2.3) $\sup_{\varepsilon \in (0, 1)} \sup_{\tau \in S^*(T)} \mathbb{P}(|Y^\varepsilon(\tau)| > R) \rightarrow 0$, $R \rightarrow +\infty$.

We prove (2.3); the proof of (2.1) is similar, hence we mention only the differences between the two proofs.

Consider the function $f(y) = \ln(1 + |y|^2)$. It is clear that $f(y) \rightarrow +\infty$ as $|y| \rightarrow +\infty$, and thus it is sufficient to show that

(2.4) $\sup_{\tau \in S^*(T)} \mathbb{P}(f(Y^\varepsilon(\tau)) > R) \rightarrow 0$, $R \rightarrow +\infty$,

for all $T \in \mathbb{R}^+$.

Since (4.1) in [1] holds with the function $f(y) = \ln(1 + |y|^2)$, the assumptions of the lemma imply (4.4) in [1]. This means that (2.4) follows if the corresponding terms in (4.4) of [1] are such that

(2.5) $\sup_{\varepsilon \in (0, 1)} \sup_{\tau \in S^*(T)} \mathbb{P}\left(\varepsilon^{1/2}|u_f(X^\varepsilon(\tau), Y^\varepsilon(\tau))| > R\right) \rightarrow 0$, $R \rightarrow \infty$,

(2.6) $\sup_{\varepsilon \in (0, 1)} \sup_{\tau \in S^*(T)} \mathbb{P}\left(|\int_0^\tau K_f^\varepsilon(X^\varepsilon(s), Y^\varepsilon(s)) \, ds| > R\right) \rightarrow 0$, $R \rightarrow \infty$,

(2.7) $\sup_{\varepsilon \in (0, 1)} \sup_{\tau \in S^*(T)} \mathbb{P}(|M_f^\varepsilon(\tau)| > R) \rightarrow 0$, $R \rightarrow \infty$.

Using conditions $E(d, r_1, \psi_1)$ and $M_{\psi_1}(\phi_1, \psi_1)$ together with assumption 4 of Theorem 3.1 in [1] and bound (3.2) in [1] for the extended potential, we get

$$|\mathcal{R}A_i(x, y)| \leq 2||A_i||_{\phi_1, p_1} \left(\int_0^\infty r_1^{1/p_1}(t) \, dt\right) \left(\int_X \psi_1 \, d\pi\right)^{1/q_1} \psi_1(x)(1 + |y|),$$

where $i = 1, \ldots, m$.

According to equality in (4.7) [1] defining the corrector $u_f$, the latter result implies that

(2.8) $|u_f(x, y)| \leq C\psi_1(x)$, $x \in \mathbb{X}$, $y \in \mathbb{R}^m$,

in view of the bound imposed on the first derivative of $f$ in (4.1) of [1]. Combining this with condition $S(\psi_1, Q_1)$ we get
\[(2.9) \quad \sup_{\varepsilon \in (0, 1)} \sup_{\tau \in \mathcal{S}^\varepsilon(T)} \varepsilon^{1/2} \mathbb{E} \left| u_f(X^\varepsilon(\tau), Y^\varepsilon(\tau)) \right| < +\infty, \]

whence (2.5) follows.

If a weaker condition \( \mathbf{W}(\psi_1, Q_1) \) is used instead of \( \mathbf{S}(\psi_1, Q_1) \), we nevertheless follow the same reasoning but obtain a weaker version of relation (2.9), namely
\[(2.10) \quad \sup_{\varepsilon \in (0, 1)} \sup_{\tau \in [0, T]} \varepsilon^{1/2} \mathbb{E} \left| u_f(X^\varepsilon(t), Y^\varepsilon(t)) \right| < +\infty. \]

It is clear that (2.12) implies (2.11), and this completes the proof of (2.6).

Turning to the proof of (2.5), we first show that
\[(2.14) \quad |M^\varepsilon_f(T)| \leq f(Y^\varepsilon(T)) + \varepsilon^{1/2} |u_f(X^\varepsilon(T), Y^\varepsilon(T))| + \int_0^T |K^\varepsilon_f(X^\varepsilon(s), Y^\varepsilon(s))| \, ds. \]

Moreover, \( M^\varepsilon_f \) is a martingale, and thus
\[(2.15) \quad \mathbb{E} f(Y^\varepsilon(T)) = f(y^0) - \varepsilon^{1/2} \mathbb{E} u_f(X^\varepsilon(T), Y^\varepsilon(T)) + \varepsilon^{1/2} \mathbb{E} u_f(X^\varepsilon(0), Y^\varepsilon(0)) + \mathbb{E} \int_0^T K^\varepsilon_f(X^\varepsilon(t), Y^\varepsilon(t)) \, ds \leq f(y^0) + \varepsilon^{1/2} \mathbb{E} |u_f(X^\varepsilon(T), Y^\varepsilon(T))| + \varepsilon^{1/2} \mathbb{E} |u_f(X^\varepsilon(0), Y^\varepsilon(0))| + \mathbb{E} \int_0^T |K^\varepsilon_f(X^\varepsilon(s), Y^\varepsilon(s))| \, ds. \]

Inequalities (2.14), (2.13) and bounds (2.10), (2.12) imply (2.13).

By the Doob optional stopping time theorem applied to the submartingale \( |M^\varepsilon_f| \) we conclude that
\[ \mathbb{E} \left| M^\varepsilon_f(\tau) \right| \leq \mathbb{E} |M^\varepsilon_f(T)| \]
for all \( \tau \in \mathcal{S}^\varepsilon(T) \). Recall that \( \tau \) is a discrete stopping time, and thus we do not need to impose any additional restriction on the filtration \( \mathbb{F}^\varepsilon \) or on the trajectory \( M^\varepsilon_f \) in order to apply the Doob theorem. Therefore (2.13) implies that
\[ \sup_{\varepsilon \in (0, 1)} \sup_{\tau \in \mathcal{S}^\varepsilon(T)} \mathbb{E} \left| M^\varepsilon_f(\tau) \right| < +\infty, \]
whence (2.14) follows. The proof of (2.4) is complete.

Note that we did not use condition $S(\psi_1, Q_1)$ in the proof of relations (2.6) and (2.7), and thus relations (2.6), (2.7), and (2.10) hold even without condition $S(\psi_1, Q_1)$. This observation completes the proof of (2.1).

□

Lemma 2.2. If all the assumptions of Lemma 2.1 hold, then, given arbitrary $r > 0$ and $R > 0$, there exists a constant $C_{r,R} \in \mathbb{R}^+$ such that the inequality

$$
P \left[ |Y^\varepsilon(\tau + u) - Y^\varepsilon(\tau)| > r \mid F^\varepsilon_\tau \right] \leq C_{r,R} \left( u + \varepsilon \psi_2(X^\varepsilon(\tau)) \right)
$$

holds almost surely in the set $\{ |Y^\varepsilon(\tau)| \leq R \}$ for all $u \in (0, 1], \varepsilon \in (0, 1]$, and

$$
\tau \in \bigcup_{T \in \mathbb{R}^+} S^\varepsilon(T).
$$

Proof. We cover the ball $\bar{B}(0, R) = \{ |y| \leq R \}$ by a finite collection of open balls $B(z_i, r/3)$, $i = 1, \ldots, I$, where $r > 0$ is fixed. Consider the corresponding collection of functions $f_i(y) = \theta((y - z_i)/r)$, $i = 1, \ldots, I$, where $\theta \in C^2(\mathbb{R}^d)$ is a fixed non-negative function that assumes two values, namely 0 in the set $\{ |y| \leq 1/3 \}$ and 1 in the set $\{ |y| \geq 2/3 \}$. Then there exists a constant $C$ depending only on $r$ and $\theta$ such that

$$
\|\nabla^j f_i(y)\| \leq C(1 + |y|)^{-j}, \quad j = 1, 2, 3, \quad i = 1, \ldots, I.
$$

The derivatives of the function $f_i$ equal zero in the ball $B(z_i, r/3)$ and outside the ball $B(z_i, 2r/3)$. Now by formula (4.7) in [1] that defines the corrector $u_{f_i}$, the equality $u_{f_i}(x, y) = 0$ holds if $|y - z_i| < r/3$ or if $|y - z_i| > 2r/3$. Thus, given an arbitrary $\tau \in S^\varepsilon$, the equality

$$
f_i(Y^\varepsilon(\tau)) + \varepsilon^{1/2} u_{f_i}(X^\varepsilon(\tau), Y^\varepsilon(\tau)) = 0
$$

holds in the set $\{ Y^\varepsilon(\tau) \in B(z_i, r/3) \}$. On the other hand, for all $u \in \mathbb{R}^+$, the inequality

$$
|Y^\varepsilon(\tau + u) - Y^\varepsilon(\tau)| > r,
$$

being true in the same set, implies the relation $|Y^\varepsilon(\tau + u) - z_i| > 2r/3$. As a result, the equality

$$
f_i(Y^\varepsilon(\tau + u)) + \varepsilon^{1/2} u_{f_i}(X^\varepsilon(\tau + u), Y^\varepsilon(\tau + u)) = 1
$$

also holds. This consideration together with representation (4.4) in [1] and Doob’s optional stopping time theorem yields

$$
P \left[ |Y^\varepsilon(\tau + u) - Y^\varepsilon(\tau)| > r, Y^\varepsilon(\tau) \in B(z_i, r/3) \right] \leq E \left[ f_i(Y^\varepsilon(\tau + u)) + \varepsilon^{1/2} u_{f_i}(X^\varepsilon(\tau + u), Y^\varepsilon(\tau + u)) \right]
$$

$$
- f_i(Y^\varepsilon(\tau)) - \varepsilon^{1/2} u_{f_i}(X^\varepsilon(\tau), Y^\varepsilon(\tau)) \left| F^\varepsilon_\tau \right| 1_{Y^\varepsilon(\tau) \in B(z_i, r/3)}
$$

$$
= E \left[ \int_0^u K_{f_i}^\varepsilon(X^\varepsilon(\tau + s), Y^\varepsilon(\tau + s)) \, ds \left| F^\varepsilon_\tau \right| 1_{Y^\varepsilon(\tau) \in B(z_i, r/3)} \right].
$$

Then we get

$$
\|K_{f_i}^\varepsilon\|_{d_2, \psi_2, p_2, \pi_2} \leq C, \quad \varepsilon \in (0, 1],
$$

by (2.17) and relations (4.5) and (4.6) in [1].

A minor change of the proof of Proposition A.1 in [1] with $f = K_{f_i}^\varepsilon$, $d = d_2$, and $\psi = \psi_2$ allows us to prove

$$
E \left| K_{f_i}^\varepsilon(X^\varepsilon(\tau + s), Y^\varepsilon(\tau + s)) \right| \chi \leq C E \left( 1 + r_2^{1/p_2} (s) \psi_2(X^\varepsilon(\tau)) \right) \chi, \quad s \in \mathbb{R}^+,
$$
for an arbitrary $F^ε_τ$-measurable nonnegative bounded random variable $χ$. This is equivalent to

$$E \left[ K^ε_f(X^ε(τ + s), Y^ε(τ + s)) \left| F^ε_τ \right. \right] \leq C \left( 1 + r^1_2 / p^2(s)ψ_2(X^ε(τ)) \right), \quad s \in \mathbb{R}^+,$$

almost surely.

Therefore relation (2.19) implies that

$$P \left[ |Y^ε(τ + u) - Y^ε(τ)| > r | F^ε_τ \right]$$

$$\leq E \left[ \int_0^{τ+u} K^ε_f(X^ε(s), Y^ε(s)) \left| F^ε_τ \right. \right]$$

$$\leq C \left( 1 + r^1_2 / p^2((s - τ)/ε)ψ_2(X^ε(τ)) \right)$$

$$= C \left( u + ε \left( \int_0^u r^1_2 / p^2(s) ds \right) ψ_2(X^ε(τ)) \right)$$

$$\leq C \left( u + εQ(ε^{-1}) ψ_2(X^ε(τ)) \right)$$

in the set $\{Y^ε(τ) \in B(z_i, r/3)\}$. We have used the assumption $u \leq 1$ in the latter inequality. Since the set $\{Y^ε(τ) | R\}$ is covered by a finite collection of sets

$$\{Y^ε(τ) \in B(z_i, r/3)\}, \quad i = 1, \ldots, I,$$

inequalities (2.19), $i = 1, \ldots, I$ imply the result required. □

**Lemma 2.3.** Let all the assumptions of Lemma 2.1 hold. Then

$$\limsup_{ε \to 0} \sup_{τ \in [0, T]} \sup_{u \leq ε \leq δ} E \left[ |Y^ε(t + u) - Y^ε(t)| \wedge 1 \right] \to 0, \quad δ \to 0,$$

for all $T \in \mathbb{R}^+$. Moreover, if $X$ satisfies condition $S(ψ_i, Q_i)$, $i = 1, 2$, then

$$\limsup_{ε \to 0} \sup_{τ \in S^ε(T)} \sup_{u \leq δ} E \left[ |Y^ε(τ + u) - Y^ε(τ)| \wedge 1 \right] \to 0, \quad δ \to 0.$$

**Proof.** Lemma 2.2 implies that

$$E \left[ |Y^ε(τ + u) - Y^ε(τ)| \wedge 1 \right]$$

$$\leq E \left( E \left[ |Y^ε(τ + u) - Y^ε(τ)| \wedge 1 \right| F^ε_τ \right. \right) + P \left( |Y^ε(τ)| > R \right)$$

$$\leq r + E \left( P \left[ |Y^ε(τ + u) - Y^ε(τ)| > r \left| F^ε_τ \right. \right. \right)$$

$$+ P \left( \sup_{τ \in [0, T]} |Y^ε(t)| > R \right)$$

$$\leq r + C_{τ, R} \left( u + εQ_2(ε^{-1}) E ψ_2(X^ε(τ)) \right) + P \left( \sup_{τ \in [0, T]} |Y^ε(t)| > R \right.$$}

for all $τ \in S^ε(T)$ and positive $u$, $r$, and $R$. Thus condition $S(ψ_2, Q_2)$ implies that

$$\limsup_{δ \to 0} \limsup_{ε \to 0} \sup_{τ \in S^ε(T)} \sup_{0 \leq u \leq δ} E \left[ |Y^ε(τ + u) - Y^ε(τ)| \wedge 1 \right]$$

$$\leq r + \sup_{ε \in (0, 1]} P \left( \sup_{τ \in [0, T]} |Y^ε(t)| > R \right).$$

Applying the second statement of Lemma 2.1 (here we use condition $S(ψ_1, Q_1)$) we conclude that both terms on the right hand side of the latter inequality approach zero as $r \to 0, R \to ∞$. This proves (2.21).

The proof of (2.20) is similar, and thus we omit it. □
3. A PARTICULAR CASE: DISTANCE IN VARIATION

In this section, we consider an important particular case of a process $X$ that admits a non-uniform with respect to the initial value bound for the rate of convergence to the invariant distribution with respect to the distance in variation. This bound is written as follows.

$$TV(r,\psi). A \ process \ X \ has \ a \ unique \ invariant \ distribution \ \pi \ and$$

$$\|P_t(x,\cdot) - \pi\|_{\text{var}} \leq r(t)\psi(x), \quad x \in X, \ t \geq 0,$$

where the function $\psi$ assumes values in $[1, +\infty)$ and where the function $r: \mathbb{R}^+ \to \mathbb{R}^+$ is bounded and such that $r(t) \to 0$ as $t \to \infty$.

There are a number of papers where condition (3.1) is effectively checked for various classes of Markov processes; see the survey of the literature in [3]. The distance in variation equals the minimal metric generated by the discrete metric $d(x,y) = 1_{x \neq y}$; see [5]. Thus condition $TV(r,\psi)$ is a particular case of our general condition $E(d,r,\psi)$.

The description of the “weight Hölder spaces” $H_{\phi,d,p}, H_{\phi,p,\kappa}(d,\pi)$ and $\hat{H}_{\phi,p,\kappa}(d,\pi)$ explained in [1] is especially simple in the case of the discrete metric $d$, namely the norms in these spaces are majorized by the corresponding “weight sup-norms”. This allows us to simplify essentially the conditions imposed in Theorems 3.1 and 3.2 of [1] on the coefficients of the initial equation.

**Proposition 3.1.** Let all the assumptions of Theorem 3.1 in [1] hold, where $d_1 = d_2 = d$ is a discrete metric. Then conditions (3.4)–(3.7) in [1] follow from the following inequalities:

$$\sum_i |A_i(x,y)| + \sum_{i,j} |\partial_y A_j(x,y)| + \sum_{i,j,k} |\partial_{y,y} A_k(x,y)| \leq C\phi_1(x)(1 + |y|);$$

$$\sum_i |a_i(x,y)| + \sum_{i,j} (|A_i(x,y)\partial_y \mathcal{R}A_j(x,y)| + |a_i(x,y)\partial_y \mathcal{R}A_j(x,y)|)$$

$$\sum_{i,j} |b_{ij}(x,y)\partial_{y,y} \mathcal{R}A_k(x,y)| \leq C\phi_2(x)(1 + |y|);$$

$$\sum_{i,j} (|b_{ij}(x,y)| + |A_i(x,y)\mathcal{R}A_j(x,y)| + |a_i(x,y)\mathcal{R}A_j(x,y)|)$$

$$\sum_{i,j,k} |b_{ij}(x,y)\mathcal{R}A_k(x,y)| \leq C\phi_2(x)(1 + |y|^2);$$

$$\sum_{i,j,k} |b_{ij}(x,y)\mathcal{R}A_k(x,y)| \leq C\phi_2(x)(1 + |y|^3).$$

**Proof.** We drop the subscripts and write $\phi$, $p$ instead of $\phi_1$, $p_1$, or $\phi_2$, $p_2$, respectively. It is clear that

$$\|f\|_{\phi,d,p} = \sup_{x_1 \neq x_2} \frac{|f(x_1) - f(x_2)|}{d^{1/p}(x_1,x_2)(\phi(x_1) + \phi(x_2))} \leq 2\sup_x \frac{|f(x)|}{\phi(x)} = 2\|f\|_{\phi}.$$ 

Thus the function $f = f(x,y)$ belongs to $H_{\phi,p,\kappa}(d,\pi)$ if

$$\sup_{x \in X, y \in \mathbb{R}^n} \frac{|f(x,y)|}{\phi(x)(1 + |y|)\kappa} + \int_X \sup_{y \in \mathbb{R}^n} \frac{|f(x,y)|}{(1 + |y|)\kappa} \pi(dx) < \infty.$$ 

The expression under the sign of integral in the second term of (3.6) is majorized by the first term multiplied by $\phi(x)$. Assumption (2) of Theorem 3.1 in [1] implies that
\( \phi \in L_1(\pi) \). Therefore \( f \in \mathcal{H}_{\phi,p,\kappa}(d, \pi) \) if
\[
(3.7) \quad \|f\|_{\phi,\kappa} := \sup_{x \in \mathcal{X}, y \in \mathbb{R}^m} \frac{|f(x, y)|}{\phi(x)(1 + |y|)^\kappa} < \infty.
\]
Moreover, the same condition implies that \( f \) belongs to \( \hat{\mathcal{H}}_{\phi,p,\kappa}(d, \pi) \), since
\[
\int_{\mathcal{X}} \sup_{y \in \mathbb{R}^m} \mathbb{E}_N \left( \frac{f(x, y)}{(1 + |y|)^\kappa} \right) \pi(dx) \leq \int_{\mathcal{X}} \sup_{y \in \mathbb{R}^m} \mathbb{E}_N (\|f\|_{\phi,\kappa}\phi(x)) \pi(dx) \to 0,
\]
\( N \to \infty \). \( \square \)

\textbf{Remark 3.1.} Conditions (3.2)–(3.5) correspond to the conditions imposed on the coefficients of the initial equation in [4] (\( \phi, \psi \), and \( r \) grow slower than polynomials in [4]). The results of Theorems 3.1 and 3.2 in [1] also correspond to results of Theorem 3 in [4]. This indicates that our approach does not introduce additional restrictions on the model as compared to [4], but it extends the area of applications essentially (recall that \( X \) is a diffusion process in [4]; in contrast, we do not assume any structure of the process \( X \) in the current paper as well as in [1]).

\textbf{Remark 3.2.} The extended potential \( \mathcal{R}f \) of a function \( f \) admits a bound for its growth; see Theorem 3.1 in [3]. This allows one to introduce other sufficient conditions for (3.3)–(3.5) that do not use extended potentials \( \mathcal{R}A_i \) and their derivatives. For example, let (3.2) hold and
\[
(3.8) \quad \sum_{i,j} |b_{ij}(x, y)| \leq C\phi_3(x)(1 + |y|^2),
\]
where the function \( \phi_3 \) is such that \( \phi_3\psi_1 \leq \phi_2 \). Then (3.5) also holds; see bound (3.2) in [3]. Similar sufficient conditions can also be given for (3.3) and (3.4).

\textbf{Remark 3.3.} The assumption introduced in Section 2.1 of [1] that \( \mathcal{X} \) possesses the structure of a metric space is too restrictive. Recall that the distance in variation is the minimal metric generated by a discrete metric in the general case, that is, this property holds for any family of probability measures in an arbitrary measurable space; see for example [4]. One can check explicitly that if \( d_1 = d_2 = d \) is a discrete metric, then the proof of Theorem 3.1 in [1] is extended without changes to the case where \( (\mathcal{X}, \mathcal{X}) \) is a Borel measurable space and \( X \) is a measurable process.

\textbf{Remark 3.4.} The following stronger form of property (3.1) can often be found in the literature:
\[
(3.9) \quad \|P_t(x, \cdot) - \pi\|_{\phi,\text{var}} \leq r(t)\psi(x), \quad x \in \mathcal{X}, \ t \geq 0,
\]
where \( \|\nu\|_{\phi,\text{var}} \) is the weighted full variation of the charge \( \nu \) defined as
\[
\|\nu\|_{\phi,\text{var}} = \int_{\mathcal{X}} \phi(d\nu_+ + d\nu_-).
\]
Here \( \nu_+ \) and \( \nu_- \) are the components in the Haar decomposition \( \nu = \nu_+ - \nu_- \). The class of \( V \)-ergodic processes with \( \phi = \psi = V \) is of special interest. The properties of the semigroup and potential of a \( V \)-ergodic process defined in the space \( \mathcal{B}_V(\mathcal{X}) \) of measurable functions with the weighted sup-norm \( \|f\|_V = \sup_x |f(x)|/\phi(x) \) are similar to properties of usual semigroups and potential of a uniformly ergodic Markov process. Thus one may expect that the classical semigroup proof of theorems on the diffusion approximation ([7] [8]) can be extended to the case of the class of \( V \)-ergodic perturbation processes \( X \). Note, however, that usually the property of being \( V \)-ergodic is related to some condition that a process is “exponentially recursive”. One of the versions of this condition reduces
to the Lyapunov type condition with the extended generator \( A \) (see Definition 2.1 in [3]) of the process \( X \), namely
\[
AV \leq -aV + b
\]
for some \( a, b > 0 \). Combined with the corresponding mixing condition, the latter inequality implies bound (3.9) with \( \phi = \psi = V \) and \( r(t) = Ce^{-ct} \); see for example [9]. If the process \( X \) satisfies a weaker form of (3.10) with a sublinear right hand side, then the process (not necessarily being \( V \)-ergodic) may satisfy condition (3.9) with different \( \phi \) and \( \psi \) and with a subexponential or polynomial function \( 1/r \); see for example [10]–[12] as well as Remark 5.1 and Example 5.1 below. Therefore, the class of processes \( X \) satisfying condition (3.1) and for which Theorems 3.1 and 3.2 in [1] hold with \( d_1 = d_2 = d \) is wider than the class of \( V \)-ergodic processes.


Let \( \rho \) be the initial metric in the space \( X \) and let \( d = \rho^p, \ p \in [1, +\infty) \). Then the corresponding minimal probability distance is given by
\[
d_p(\mu, \nu) = [W_p(\mu, \nu)]^p,
\]
where \( W_p \) is the probability Kantorovich–Rubinstein metric. Note that \( W_p \) is often called the Wasserstein \( p \)-metric (see Section 1.1.3 in [13] and the references therein concerning the history of this notion). Similarly to what has been demonstrated in the preceding section for the distance in variation, general assumptions of Theorems 3.1 and 3.2 in [1] can be rewritten in a shorter and more explicit form in this case.

In what follows we assume that
\[
W_1(P_t(x, \cdot), P_t(x', \cdot)) \leq v_1(t)\rho(x, x'), \quad x, x' \in X, \ t \geq 0,
\]
where a bounded function \( v_1 \) is such that \( v_1(t) \to 0 \) as \( t \to \infty \). It is not hard to show that there exists a unique invariant measure \( \pi \) of the process \( X \) and
\[
W_1(\delta_x, \pi) < \infty, \quad x \in X.
\]

**Theorem 4.1.** Let inequality (4.1) hold and, additionally,
\[
W_2(P_t(x, \cdot), \pi) \leq v_2(t)W_2(\delta_x, \pi) < \infty, \quad x \in X, \ t \geq 0,
\]
where a bounded function \( v_2 \) is such that \( v_2(t) \to 0 \) as \( t \to \infty \).

Assume that any function \( A_i(\cdot, y), y \in \mathbb{R}^m, \ i = 1, \ldots, m \), is centered and
\[
\sum_i |A_i(x, y) - A_i(x', y)| + \sum_i |a_i(x, y) - a_i(x', y)| + \sum_{ij} |b_{ij}(x, y) - b_{ij}(x', y)|
\]
\[
\leq C\rho^\alpha(x, x') (1 + \rho(x, x_*) + \rho(x', x_*))^{1-\alpha} (1 + |y|);
\]
\[
\sum_{ij} |\partial_{y_i} A_j(x, y) - \partial_{y_i} A_j(x', y)| + \sum_{ijk} |\partial_{y_i y_j}^2 A_j(x, y) - \partial_{y_i y_j}^2 A_j(x', y)|
\]
\[
\leq C\rho^\alpha(x, x') (1 + \rho(x, x_*) + \rho(x', x_*))^{1-\alpha};
\]
\[
\sum_i |A_i(x, y)| + \sum_i |a_i(x, y)| + \sum_{ij} |b_{ij}(x, y)| \leq C(1 + \rho(x, x_*)).
\]

Let
\[
\int_0^\infty (v_1^2(t) + v_2^2(t)) \ dt < \infty
\]
for some \( \alpha \in (0, 1] \) and some (as a consequence, for all) \( x_* \in X \).
Then statements I–III of Theorem 3.1 in [1] hold. Moreover, if the process \( X \) is such that
\begin{equation}
\frac{1}{\sqrt{T}} \mathbb{E} 1_{\rho(X(0), x_*) \leq R} \sup_{t \leq T} \rho^2(x_*, X(t)) \to 0, \quad T \to \infty, 
\end{equation}
for an arbitrary number \( R > 0 \), then the result of Theorem 3.2 in [1] is valid.

**Proof.** Without loss of generality, we assume that, for some \( R > 0 \), the initial value \( X(0) \) belongs to the ball \( B(0, R) \) with probability one. The general case can be reduced to this particular one with the help of a standard localization procedure with respect to the initial value.

Put \( p_1 = 1/\alpha, \ q_1 = 1/(1 - \alpha), \ r_1 = v_1, \) and
\[
d_1(x, y) = \rho(x, y), \quad \phi_1(x) = (1 + \rho(x, x_*))^{1-\alpha}, \quad \psi_1(x) = 1 + \rho(x, x_*).
\]
By (4.1) and triangle inequality for the metrics \( W_1 \) and \( \rho \),
\[
W_1 (\rho_1(x, \cdot), \pi) = \rho_1 \left( \int_X \rho(x', x_*) \delta_{x'}(dx'), \int_X \rho(x', \cdot) \pi(dx') \right) \leq r_1(t) W_1 (\delta_x, \pi)
\]
\[
= r_1(t) \int_X \rho(x, x') \pi(dx') \leq r_1(t) \left[ \int_X \rho(x_*, x') \pi(dx') + \rho(x, x_*) \right],
\]
whence condition \( \mathbf{E}(d_1, r_1, \psi_1) \) follows.

If \( \alpha = 1 \), then \( \phi_1 \equiv 1 \) and condition \( \mathbf{M}_{q_1}(\phi_1, \psi_1) \) obviously holds. If \( \alpha \in (0, 1) \), then this condition is equivalent to
\begin{equation}
\int_X \rho(x', x_*) \rho_1(x, dx') \leq 1 + \rho(x, x_*) \quad t \geq 0.
\end{equation}
Using the triangle inequality for \( d_1 \) once more, we get
\begin{equation}
\int_X \rho(x', x_*) \rho_1(x, dx') = d_1 (\delta_{x_*}, \rho_1(x, \cdot)) \leq d_1 (\delta_{x_*}, \pi) + d_1 (\pi, \rho_1(x, \cdot)).
\end{equation}
Therefore, (4.8) follows from \( \mathbf{E}(d_1, r_1, \psi_1) \).

Further, we put \( p_2 = 2/\alpha, \ q_2 = 2/(2 - \alpha), \ r_2 = v_2^2, \) and
\[
d_2(x, y) = \rho^2(x, y), \quad \phi_2(x) = (1 + \rho(x, x_*)^{2-\alpha}, \quad \psi_2(x) = 2 \left( 1 + \rho^2(x, x_*) \right).
\]
Then (4.2) coincides with condition \( \mathbf{E}(d_2, r_2, \psi_2/2) \) which obviously implies condition \( \mathbf{E}(d_2, r_2, \psi_2) \). A reasoning similar to the above one proves that condition \( \mathbf{M}_{q_2}(\phi_2, \psi_2) \) holds. A minor difference in the proof is that, instead of the triangle inequality for \( d_2 \), the following variant of it holds:
\[
d_2(\theta, \vartheta) \leq 2 \left( d_2(\theta, \kappa) + d_2(\vartheta, \kappa) \right).
\]
Since
\begin{equation}
E_\kappa \psi_1(X(t)) = 1 + \int_X \rho(x', x_*) \rho_1(x, dx'),
\end{equation}
condition (4.9) implies \( \mathbf{P}(\psi_1) \) in view of \( \mathbf{E}(d_1, r_1, \psi_1) \). Therefore, assumption (1) of Theorem 3.1 in [1] holds.

As in the case of (4.10), we use (4.9) and \( \mathbf{E}(d_1, r_1, \psi_1) \):
\[
\mathbf{E} \psi_1(X(t)) = 1 + \mathbb{E} \int_X \rho(x', x_*) \rho_1(X(0), dx') \leq 1 + d_1 (\delta_{x_*}, \pi) + \mathbb{E} d_1 (\pi, \rho_1(X(0), \cdot))
\]
\[
\leq 1 + d_1 (\delta_{x_*}, \pi) + r_1(t) \mathbb{E} d_1 (\pi, \rho_1(X(0))) \leq 1 + (1 + r_1(t)) d_1 (\delta_{x_*}, \pi) + \mathbb{E} \rho(X(0), x_*).
\]
Therefore, according to the additional assumption that $X(0) \in B(0, R)$ almost surely and since the function $r_1$ is bounded,

$$\sup_{t \geq 0} E \psi_1(X(t)) < \infty,$$

whence condition $W(\psi_1, Q_1)$ follows.

Condition $W(\psi_2, Q_2)$ can be checked in a similar way. Doing so, we take into account that the function $r_2^{1/p_2}$ is integrable in $\mathbb{R}^+$ in view of (4.6), and thus the function $Q_2$ is bounded. Therefore, assumption (2) of Theorem 3.1 in [1] holds.

The same reasoning proves assumption (6) of Theorem 3.1 in [1] with

$$\varrho(x) = C(1 + \rho(x, x_s)).$$

Under an additional assumption (4.7), one can also show that condition $S(\psi_i, Q_i)$, $i = 1, 2$, holds.

Assumption (3) of Theorem 3.1 in the paper [1] follows from (4.6), while assumption (4) of the same theorem follows from the restrictions imposed on the coefficient $A$ in Theorem 4.1.

It remains to check assumption (5), that is, in relations (3.5)–(3.7), of Theorem 3.1 in [1]. We are going to check only the second relation in (3.6) of [1], namely we check that

$$A_i \mathcal{R} A_j \in \bar{H}_{\phi_2, p_2, 2}(d_2, \pi).$$

The proof of other relations is similar and sometimes is even simpler.

Using the elementary inequality $(1 + a + b)^{1-\alpha} \leq (1 + a)^{1-\alpha} + (1 + b)^{1-\alpha}$ we derive from (4.3) that

$$\|A_j(\cdot, y)\|_{\phi_1, d_1, p_1} \leq C(1 + |y|)$$

for $y \in \mathbb{R}^m$. Applying Theorem 3.1 in [3] with $f = A_j(\cdot, y)$, $d = d_1$, $p = p_1$, $\phi = \phi_1$, and $\psi = \psi_1$, we obtain

$$|\mathcal{R} A_j(x, y)| \leq C(1 + \rho(x, x_s)) (1 + |y|).$$

Further, condition (4.11) is a stronger version of condition $\hat{E}(d_1, r_1, \psi_1)$ of Section 2 in [3] where the term $(\psi_1(x) + \psi_1(x'))$ is changed by 1. Then a stronger version of bound (4.2) in [3] holds, where the right hand side contains the factor $(\psi_1(x) + \psi_1(x'))^{1/p_2}$ instead of $(\psi_1(x) + \psi_1(x'))$. Integrating this estimate with respect to $t \in \mathbb{R}^+$ and reasoning similarly to the proof of statement 2 of Theorem 3.1 in [3] we get

$$|\mathcal{R} A_j(x, y) - \mathcal{R} A_j(x', y)| \leq C \rho_2^2(x, x') (1 + \rho(x, x_s) + \rho(x', x_s))^{1-\alpha} (1 + |y|).$$

By assumption of the theorem, relations similar to (4.12) and (4.13) hold for the function $A_i$. Hence

$$|A_i(x, y) \mathcal{R} A_j(x, y) - A_i(x', y) \mathcal{R} A_j(x', y)|$$

$$\leq |A_i(x, y) - A_i(x', y)| \cdot |\mathcal{R} A_j(x, y)| + |A_i(x', y)| \cdot |\mathcal{R} A_j(x, y) - \mathcal{R} A_j(x', y)|$$

$$\leq C \rho_2^2(x, x') (1 + \rho(x, x_s) + \rho(x', x_s))^{2-\alpha} (1 + |y|)^2$$

$$\leq C (d_2(x, y))^{p_2} (\phi_2(x) + \phi_2(x')) (1 + |y|)^2,$$

that is

$$\|A_i \mathcal{R} A_j\|_{d_2, \phi_2, p_2, 2} < +\infty.$$

On the other hand, relation (4.12) and a similar relation for $A_i$ imply

$$|A_i(x, y) \mathcal{R} A_j(x, y)| \leq C (1 + \rho^2(x, x_s)) (1 + |y|)^2 = C \psi_2(x)(1 + |y|)^2.$$

The function $\psi_2$ is integrable with respect to $\pi$, and thus $\psi_2$ is uniformly integrable. The proof of (4.11) is complete.
Therefore, all the assumptions of Theorem 3.1 [1] are valid for \( d_1, d_2, p_1, p_2, \phi_1, \phi_2, \psi_1, \) and \( \psi_2 \) chosen above. If additionally condition (4.7) holds, then the same is true for the assumptions of Theorem 3.2 in [1]. Using these theorems we complete the proof. □

Below is a particular case being of its own interest. This result can be called the central limit theorem for additive functionals of the initial Markov process \( X \).

**Corollary 4.1.** Let conditions (4.1) and (4.2) hold with the functions \( v_1 \) and \( v_2 \) satisfying (4.6) for some \( \alpha \in (0, 1] \). Let \( f : X \to \mathbb{R}^m \) be a centered function such that

\[
|f(x) - f(x')| \leq C \rho^{\alpha}(x, x') \left( 1 + \rho(x, x) + \rho(x', x) \right)^{1-\alpha}, \quad x, x' \in X,
\]

for some (as a consequence, for all) \( x_* \in X \).

Then the following statements are true:

I. (Individual central limit theorem). The family of random vectors

\[
\frac{1}{\sqrt{T}} \int_0^T f(X(s)) \, ds
\]

weakly converges as \( T \to \infty \) to a Gaussian random vector in \( \mathbb{R}^m \) with zero mean and covariance matrix

\[
B = (B_{ij})_{i,j=1}^m,
\]

\[
B_{ij} = \int_X (f_i R f_j + f_j R f_i) \, d\pi = \int_{-\infty}^{\infty} \mathbb{E} f_i \left( X^{st}(0) \right) f_j \left( X^{st}(t) \right) \, dt,
\]

where \( X^{st} \) denotes a stationary version of the process \( X \) defined in the whole line \( \mathbb{R} \).

II. (Functional central limit theorem). Under an additional assumption (4.7), the family of stochastic processes

\[
\frac{1}{\sqrt{T}} \int_0^{Tt} f(X(s)) \, ds, \quad t \in \mathbb{R}^+,
\]

weakly converges as \( T \to \infty \) in the space \( C(\mathbb{R}^+, \mathbb{R}^m) \) to a Brownian motion with the covariance matrix (4.15).

Corollary 4.1 follows from Theorem 4.1 with \( a \equiv 0, b \equiv 0, \) and \( A(x, y) = f(x) \).

It is worthwhile to compare the first statement of Corollary 4.1 with the individual central limit theorem given in [2]. Assumption H1) in [2] coincides with our condition (4.1) if \( v_1 \) is an exponential function. Further, there is no counterpart of our condition (4.2) in [2]; on the other hand, assumptions H2) and H3) in [2] impose restrictions on moments of order \( 2 + \delta \) for the function \( \rho(\cdot, x_*) \). The differences mentioned above are minor, since if (4.1) holds, then, typically, all additional conditions can quite easily be checked; see Section 6 in [2] and Section 5 below.

We impose the “weighted H"older condition” (4.14) on the function \( f \), which is weaker than the Lipschitz condition imposed on \( f \) in [2] (note that these two conditions coincide for \( \alpha = 1 \)). This difference can be significant in some particular but still interesting cases. The kernels of additive functionals arising in these cases satisfy minimal regularity assumptions; see, for example, a remark in Section 6 of [2].

Finally, we should like to mention that the assumptions for the individual central limit theorem of Corollary 4.1 are direct counterparts of those in a more general Theorem 4.1 on the diffusion approximation. The latter assumptions are less restrictive (and more general, sometimes) than the conditions for the individual central limit theorem in [2]. This indicates once more that the method applied to the diffusion approximation in [1] as well as in the current paper is rather effective; also see Remark 3.2.
5. Example: A solution of a stochastic differential equation with Lévy noise

Consider a Markov process $X$ defined as a solution of a stochastic differential equation with Lévy noise in $\mathbb{R}^m$:

\begin{equation}
    dX(t) = a(X(t)) \, dt + \int_{|u| \leq 1} c(X(t), u) \, \nu(dt, du) + \int_{|u| > 1} c(X(t), u) \, \nu(dt, du).
\end{equation}

Here $\nu$ is a Poisson point measure in $\mathbb{R}^+ \times \mathbb{R}^k$ with intensity measure $dt \, \mu(du)$, $\mu$ is the Lévy measure of $\nu$, and $\bar{\nu}(dt, du) = \nu(dt, du) - dt \, \mu(dt)$ is the corresponding compensated point measure. The coefficients $a$ and $c$ are assumed to satisfy the usual conditions being sufficient for the existence of a strong solution of (5.1) (namely, we assume the local Lipschitz property and linear growth). We check the assumptions of Theorems 3.1 and 3.2 in [1] for such a process.

For simplicity, we assume that

$$
\int_{\mathbb{R}^k} |c(x, u)| \, \mu(du) < \infty, \quad x \in \mathbb{R}^m.
$$

Put

$$
\tilde{a}(x) = a(x) + \int_{|u| > 1} c(x, u) \, \mu(du), \quad \tilde{a}(x) = a(x) - \int_{|u| \leq 1} c(x, u) \, \mu(du).
$$

We start with a sufficient condition for the ergodicity of the process $X$ with respect to the distance in variation. We collect some related results from the paper [14] in the following assertion.

**Proposition 5.1.** Assume that the following assumptions hold:

1. For some $R \geq 0$ and $\alpha > 0$,

   $$(a(x), x) \leq -\alpha |x|^2, \quad |x| \geq R.$$

2. Given an arbitrary vector $w \in \mathbb{R}^k \setminus \{0\}$, there exists a number $\varrho \in (0, 1)$ such that

   $$
   \mu\left(\{u \in \mathbb{R}^k: (u, w) \geq \varrho |u||w|, |u| \leq \delta\}\right) > 0
   $$

   for an arbitrary $\delta > 0$; moreover,

   $$
   \int_{|u| > 1} |u|^2 \Pi(du) < +\infty.
   $$

3. The function $c$ admits the representation $c = c_1 + c_2$, where

   $$
   |c_1(x, u)| \leq \vartheta(x)|u| \quad \text{and} \quad \frac{\vartheta(x)}{|x|} \to 0, \quad |x| \to \infty;
   $$

   $$
   |x + c_2(x, u)| \leq |x|, \quad x \in \mathbb{R}^m, \quad |u| > 1, \quad c_2(\cdot, u) \equiv 0, \quad |u| \leq 1.
   $$

4. For some point $x_*$, there exists a neighborhood $O_{x_*}$ such that

   $$
   c(x, u) = \chi(x)u + \delta(x, u), \quad x \in O_{x_*};
   $$

   moreover,

   $$
   |\delta(x_*, u)| + |\nabla_x \delta(x_*, u)| = o(|u|), \quad |u| \to 0,
   $$

   and

   $$
   \text{rank} [\nabla \tilde{a}(x_*) \chi(x_*) - \nabla \chi(x_*) \tilde{a}(x_*)] = m.
   $$

Then the stochastic process $X$ satisfies condition $\text{TV}(r, \psi)$ with

$$
    r(t) = e^{-ct} \quad \text{and} \quad \psi(x) = C \left(1 + |x|^2\right).
$$
Remark 5.1. Assumptions (2) and (4) above imply the local integral condition that a process is non-recursive; see condition LD in [14]. Assumptions (1) and (3) imply a Lyapunov type condition of the form (3.10). If such a condition does not hold, then, nevertheless, condition TV(r, ψ) may hold for the process X with an appropriate function r and some non-exponential function ψ, and coefficients of equation (5.1) are non-degenerate. The following sufficient condition (5.6) (1 + Ux,γ ≤ |x|γ) may hold for the process X with an appropriate function r. For example, if condition LD in [14] holds and bounds (5.10) and (5.12) given below are satisfied for an arbitrary γ > 2, then (see the proof of Theorem 2 in [10]), given an arbitrary number δ > 0, there exists γ > 2 such that the process X satisfies condition TV(r, ψ) with
\[ r(t) = (1 + t)^{-δ} \quad \text{and} \quad ψ(x) = C(1 + |x|^γ). \]

The above sufficient condition for the ergodicity with respect to the distance in variation involves some (not always very restrictive) assumptions that the Lévy measure μ and coefficients of equation (5.1) are non-degenerate. The following sufficient condition for the ergodicity with respect to the Kantorovich–Rubinstein metric does not involve such assumptions.

**Proposition 5.2.** Let
\[
(5.2) \quad (\tilde{a}(x) - \tilde{a}(x'), x - x') \leq -α|x - x'|^2, \quad x, x' \in \mathbb{R}^m,
\]
\[
\int_{\mathbb{R}^k} |c(x, u) - c(x', u)|^2 \, μ(du) \leq β|x - x'|^2, \quad x, x' \in \mathbb{R}^m,
\]
for some α and β such that α > β/2.

Then
\[
(5.3) \quad W_2(P_t(x, ·), P_t(x', ·)) \leq e^{-t(α - β/2)}|x - x'|, \quad x, x' \in \mathbb{R}, \ t \geq 0.
\]

Note that (5.2) is a condition where the coefficient \( \tilde{a} \) is dissipative. Bounds of the type (5.3) are quite typical for dissipative systems; see, for example, [15], Section 11.5, or [16], Section 16.2. Thus we omit the corresponding proof. Note that (5.3) implies conditions (4.1) and (4.2).

Therefore, we can effectively check condition E(d, r, ψ) for a solution of equation (5.1) by using either condition TV(r, ψ) or the set of conditions (4.1) and (4.2).

Then we show that the rest of assumptions of Theorems 3.1 and 3.2 in [1] also hold in the case under consideration.

For example, let
\[
(5.4) \quad (a(x), x) \leq -B_1|x|^α + B_2, \quad x \in \mathbb{R}^m,
\]
for some α ∈ (0, 2) and B1, B2 > 0 and, for an arbitrary γ > 2,
\[
(5.5) \quad \sup_{x \in \mathbb{R}^m} \int_{\mathbb{R}^k} (|c(x, u)| + |c(x, u)|^γ) \, μ(du) < \infty.
\]
Let U ∈ C^2(\mathbb{R}^m) be a non-negative function such that U(x) ≤ |x|, x ∈ \mathbb{R}^m, and let U(x) = |x|, |x| ≥ 1. By Itô’s formula (Theorem 5.1 of Chapter II in [17]),
\[
(5.6) \quad (1 + U(X(t)) + t)^γ = \int_0^t Q_{U,γ}(X(s), s) \, ds + M_{U,γ}(t),
\]
where M_{U,γ} is a local martingale and where
\[
Q_{U,γ}(x, s) = γ(1 + U(x) + s)^{γ-1}(\tilde{a}(x), U'(x))
+ \int_{\mathbb{R}^k} ((1 + U(x + c(x, u)) + s)^γ - (1 + U(x) + s)^γ) \, μ(du).
\]
By construction, the function $U$ possesses the Lipschitz property. Applying the mean value theorem we derive from here that
\[
(1 + U(x + c(x, u)) + s)^\gamma - (1 + U(x) + s)^\gamma \leq C \left((1 + U(x) + s)^{\gamma - 1} |c(x, u)| + |c(x, u)|^\gamma\right)
\]
for some constant $C$. Taking into account (5.4) and (5.5), we conclude that
\[
Q_{U,\gamma}(x, s) \leq 0 \quad \text{and} \quad Q_{U,2\gamma}(x, s) \leq 0, \quad |x| \geq R, \quad s \geq 0,
\]
for some positive $R$. Thus if $E[X(0)]^\gamma < \infty$, then the process $M_{U,\gamma}$ is a martingale. The proof of this fact is quite standard and thus omitted here (see, for example, the proof of the first part of Proposition 3.1 in [18]). Therefore,
\[
E(1 + U(X(t)) + t)^{\gamma} \mathbb{1}_{\tau_R > t} \leq E(1 + U(X(0)))^\gamma,
\]
\[
\tau_R := \inf\{t: |X(t)| \leq R\}
\]
by (5.7). In particular,
\[
E(U(X(t)))^\gamma \mathbb{1}_{\tau_R > t} \leq E(1 + U(X(0)))^\gamma,
\]
\[
P(\tau_R > t) \leq t^{-\gamma} E(1 + U(X(0)))^\gamma \leq t^{-\gamma} E(1 + |X(0)|)^\gamma.
\]
Applying analogous inequalities with $2\gamma$ instead of $\gamma$ together with the Cauchy inequality, we obtain
\[
E(U(X(t)))^\gamma \mathbb{1}_{\tau_R > t} \leq t^{-\gamma} E(1 + U(X(0)))^{2\gamma}.
\]
Repeating the proof of Lemma A.3 in [19] (with minor changes), we deduce from (5.11) that
\[
\sup_{|x| \leq R, t \geq 0} E(U(X(t)))^\gamma < \infty.
\]
In combination with (5.9), the latter relation leads to the inequality
\[
E|X(t)|^\gamma \leq C E(1 + |X(0)|)^\gamma
\]
for some constant $C$. The proof of (5.12) is based on the strong Markov property of the process $X$ and on the construction of the function $U$.

The necessary properties of the process $X$ easily follow from bound (5.12); see Section 2.2 in [1]. For example, if one puts $\phi(x) = 1 + |x|^{\gamma/p}$, then (5.8) implies condition $M_p(\phi, \psi)$ with $\psi(x) = C(1 + |x|^\gamma)$. The same bound for the same function $\psi$ yields condition $P(\psi)$ and condition $W(\psi, Q)$ if $Q$ is such that
\[
\varepsilon Q(\varepsilon^{-1}) \to 0, \quad \varepsilon \to 0.
\]
To check condition $S(\psi, Q)$, we follow the lines of the proof in [4] given for diffusion processes. Similarly to (5.6), we write
\[
|X(t)|^\gamma = \int_0^t \vartheta_\gamma(X(s)) \, ds + M_\gamma(t),
\]
\[
M_\gamma(t) = \int_0^t \int_{R^k} (|X(s-) + c(X(s-), u)|^\gamma - |X(s-)|^\gamma) \, \nu(ds, du).
\]
Applying the Taylor formula we establish the following bound:
\[
(|x + c(x, u)|^\gamma - |x|^\gamma)^2 \leq C \left(|x|^{2\gamma - 2} |c(x, u)|^2 + |c(x, u)|^{2\gamma}\right).
\]
Thus inequality (5.12) with $2\gamma - 2$ instead of $\gamma$ implies that
\[
E M^2_\gamma(t) \leq C t E(1 + |X(0)|)^{2\gamma - 2}.
\]
Then
\[ |X(t)|^\gamma \leq |X(0)|^\gamma + \left( \int_0^t \vartheta_\gamma(X(s)) \, ds \right)_+ + |M_\gamma(t)|, \]
whence
\[ \sup_{t \leq T} |X(t)|^{2\gamma} \leq 3 \left( |X(0)|^{2\gamma} + T \int_0^T \vartheta_\gamma(X(s))^2 \, ds + \sup_{t \leq T} M^2_\gamma(t) \right) \]
by the Cauchy and Jensen inequalities applied to the convex function \( x \mapsto x^{2\gamma} \).

Similarly to (5.7), one can show that the function \( \vartheta_\gamma + \) is bounded. Now the latter inequality, Doob’s maximal moment inequality for the martingale \( M_\gamma \), and (5.14) imply that
\[
E \sup_{t \leq T} |X(t)|^{2\gamma} \leq C \left( 1 + T^2 \right) \left( 1 + |X(0)|^{2\gamma} \right).
\]
Let \( \psi(x) = 1 + |x|^\gamma \). Using the Hölder inequality with some \( h > 1 \) and the preceding inequality with \( h \gamma \) instead of \( 2\gamma \), we get
\[ \sup_{\tau \in \mathcal{S}(\tau)} E \psi(X^\tau(\tau)) \leq E \sup_{t \leq \varepsilon^{-2} T^2} \psi(X(t)) \leq 1 + C^{1/h} \left( 1 + \varepsilon^{-2} T^2 \right)^{1/h} \left( E(1 + |X(0)|^{h \gamma}) \right)^{1/h}. \]
Therefore, condition \( S(\psi, Q) \) holds with \( \psi(x) = 1 + |x|^\gamma \) if
\[ E|X(0)|^{h \gamma} < \infty \]
for some \( h > 1 \) and if \( Q \) is such that \( \varepsilon^{1-2/h} Q(\varepsilon^{-1}) \to 0 \) as \( \varepsilon \to 0 \). In particular, the above reasoning for \( h > 4 \) implies that condition \( S(\psi, Q) \) holds for \( Q_1(t) = \sqrt{t} \).

In conclusion, we briefly discuss two examples.

Example 5.1. Let \( X \) be a solution of the following one dimensional stochastic differential equation:
\[
dX(t) = a(X(t)) \, dt + dZ(t),
\]
where \( a(x) = -\arctg x \) and \( Z \) is a one dimensional Lévy process. Let the Lévy measure \( \mu \) of this process be non-degenerate and such that
\[
\int_{\mathbb{R}} (|u| + |u|^\gamma) \, \mu(du) < +\infty, \quad \gamma > 1.
\]
As shown above, bounds (5.10) and (5.12) hold for an arbitrary \( \gamma > 2 \). Thus the process \( X \) satisfies condition \( TV(r, \psi) \) with \( r(t) = (1 + t)^{-\delta} \) and \( \psi(x) = C(1 + |x|^\gamma) \); see Remark 5.1 (the proof of condition LD in [14] is the same as that of Proposition 0.1 in [13] and uses Proposition 4.8 in [14]).

As we have seen above, the other restrictions imposed on the process \( X \) in Theorems 3.1 and 3.2 of [1] can easily be checked for such functions \( \psi \) in view of the bound (5.12). Note that if
\[ \int_{\mathbb{R}} (e^{c|u|} - 1) \, \mu(du) = \infty, \quad c > 0, \]
then one can explicitly check that condition \( TV(r, \psi) \) does not hold for the process \( X \) for any function \( \psi \) and an exponential function \( r \).

Example 5.2. Let \( X \) be a solution in \( \mathbb{R}^2 \) of a stochastic differential equation (5.16) with \( a(x) = -x \) and \( Z = (Z_1, 0) \), where \( Z_1 \) is a square integrable one dimensional Lévy process. Then the distributions of solutions starting from points \( (x_1, x_2) \) and \( (x'_1, x'_2) \) are mutually singular if \( x_2 \neq x'_2 \). This means that the convergence in variation of transition probabilities does not hold for the process \( X \).
On the other hand, applying Proposition 5.2 we get bound (5.3) for the rate of convergence of transition probabilities in the Kantorovich–Rubinstein metric. In turn, this bound allows one to apply Theorem 4.1 and Corollary 4.1. In this case, condition (4.7) follows from bound (5.15) if the Lévy measure of the process $Z_1$ satisfies condition (5.17).

Note that several recent publications ([16], [20]–[24]; also see a survey in the introduction to the paper [3]), contain a wide range of models involving stochastic processes for which, similarly to the processes considered above, the transition probabilities converge to the invariant distribution in a weaker mode of convergence than the convergence in variation.

Bibliography


School of Mathematics, University of Leeds, United Kingdom; Institute for Information Transmission Problems, Moscow, Russia

E-mail address: A.Veretennikov@leeds.ac.uk

Institute of Mathematics, National Academy of Science, Tereschenkivs’ka Street, 3, 01601, Kyiv, Ukraine

E-mail address: kulik@imath.kiev.ua

Received 15/MAY/2012

Translated by N. SEMENOV