

EXTENDED POISSON EQUATION FOR WEAKLY ERGODIC MARKOV PROCESSES

UDC 519.21

A. YU. VERETENNIKOV AND A. M. KULIK

ABSTRACT. Solvability conditions for a Poisson equation with an extended generator of a general Markov process are obtained. The predictable part in the Doob–Meyer decomposition is described for a process of the form $g(X(t), Y(t))$, where Y is a solution of a stochastic equation with the coefficients depending on X and where the function $g = g(x, y)$ is defined as a family of solutions of the Poisson equation.

1. INTRODUCTION

Let $X(t)$, $t \in \mathbb{R}^+$, be a homogeneous Markov process and let processes Y^ε , $\varepsilon > 0$, be defined as solutions of differential equations (possibly, stochastic differential equations) whose coefficients depend on a “fast” random component $X^\varepsilon(t) = X(t\varepsilon^{-1})$, $t \in \mathbb{R}^+$. Describing the limit behavior of Y^ε as $\varepsilon \rightarrow 0+$ is of significant interest, since there is a wide range of both theoretical and applied models, where “fast” and “slow” components are presented. A comprehensive survey of results of this type can be found, for example, in Sections 1 and 2 of the paper [1].

The current paper is motivated by the following natural question: how do the ergodic properties of the process X influence the results of the “diffusion approximation” type for the processes Y^ε ? The term “diffusion approximation” is usually used for a certain version of the functional central limit theorem.

The classical results on the diffusion approximation (see, for example [2], [3], [4], Theorem 12.2.4 in [5]; see also [6]) assume that the process X is *uniformly ergodic*. This means that the invariant distribution π is unique and that

$$(1) \quad \|P_t(x, dy) - \pi(dy)\|_{\text{var}} \leq r(t), \quad t \in \mathbb{R}^+,$$

uniformly with respect to points x of the space \mathbb{X} of states of the process X , where $P_t(x, dy)$ is the transition function of the process X , $\|\cdot\|_{\text{var}}$ the distance in variation, and r some function converging sufficiently quickly zero at $+\infty$. The assumption on the uniform ergodicity is quite natural for processes with a finite (or, more generally, with a compact) space of states. On the other hand, such an assumption is rather restrictive for more complicated systems. For example, if X is defined via a stochastic differential equation in \mathbb{R}^k whose coefficients grow at most linearly, then bound (1) does

2010 *Mathematics Subject Classification*. Primary 60H10, 60J10.

Key words and phrases. Markov process, extended generator, Poisson equation, Doob–Meyer decomposition.

This research was partially supported by the State Fund for Fundamental Researches, project $\Phi 40.1/023$.

This paper is based on the talk presented at the International Conference “Modern Stochastics: Theory and Applications II” held on September 7–11, 2010, at Kyiv National Taras Shevchenko University and dedicated to the anniversaries of the prominent Ukrainian scientists Anatoliï Skorokhod, Vladimir Korolyuk, and Igor Kovalenko.

not hold, since $P_t(x, K) \rightarrow 0$, $|x| \rightarrow \infty$, for an arbitrary compact set K and all fixed t . Estimates of type (1) are even less natural for Markov processes with infinite state spaces. For example, the distributions $P_t(x, \cdot)$ are singular for certain pairs of initial points for many Markov processes defined via partial stochastic differential equations. This, in particular, means that $P_t(x, dy)$ does not converge to the initial distribution in the sense of the convergence in variation.

A method of solving the problem of diffusion approximation was developed in [7] for the case where the “fast” random component is a diffusion process. This method is essentially based on the analysis of the following *Poisson equation* in \mathbb{R}^k :

$$(2) \quad Lu(x) = -f(x),$$

where

$$L = \sum_{i,j=1}^k a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^k b_i(x) \frac{\partial}{\partial x_i}$$

is the elliptic operator of the second order related to the “fast” diffusion component. As mentioned above, bound (1) does not hold, as a rule. Instead, bounds of the form

$$(3) \quad \|P_t(x, dy) - \pi(dy)\|_{\text{var}} \leq r(t)\psi(x), \quad t \in \mathbb{R}^+,$$

are satisfied under a rather wide class of natural assumptions imposed on the coefficients, where ψ is a “penalty” function. The function ψ as well as the function r are determined by the Lyapunov conditions for the coefficients of the process (see [8]–[10]). Bounds of type (3) are used in [7] to analyze equation (5).

Our aim is to generalize the method of paper [7] for the case where no specific structural assumptions are imposed on the “fast” random component. For example, we do not assume that X is a diffusion process, but we do assume that some bounds of the form (3) hold. More generally, we assume that

$$(4) \quad d(P_t(x, dy), \pi(dy)) \leq r(t)\psi(x), \quad t \in \mathbb{R}^+,$$

where d is a *probabilistic distance* (see Section 2 below). Therefore our main assumptions imply that the transition probabilities of the “fast” random component converge to the invariant distribution in a certain sense. This convergence is weaker than the convergence in variation, which explains the term “weakly ergodic process” used in the title of the current paper.

A variety of methods is described in the literature allowing one to check both conditions (3) and (4). Condition (3) is obtained in [11] with an exponential function r and in [12] with an subexponential function r . General conditions for (3) are given in [13] with an exponential function r in the case of solutions of stochastic differential equations with a Lévy noise. As shown in [13], the checking of irreducibility conditions used in the papers [11] and [12] and stated in terms of “petite sets” is hardly possible in the general case.

A bound of type (4) with an exponential function r and with an appropriate distance d appears in a natural way if a process is defined via a stochastic differential equation (possibly, via an infinite dimensional stochastic differential equation) possessing a certain *dissipative* property. Typical examples of such results are given in the books [14] (Section 11.5) and [15] (Section 16.2).

In some recent papers (see, for example, [16]–[19]), several methods are proposed that allow one to prove bounds similar to (4) under sharper assumptions than the dissipative assumption. We also mention the “general version of the Harris theorem” proposed in [20] (we follow the terminology of the paper [20]). This result derives bound (4) from an appropriate version of the Lyapunov condition. This Lyapunov condition involves a

set where the process X is irreducible in a certain sense related to the distance d (“ d -small set”).

The above discussion confirms that there are several examples of significant interest for which one can prove either explicit estimates for the rate of convergence to the invariant distribution that are nonuniform with respect to the initial point, or those for the convergence in variation as in (3), or those for a weaker probabilistic distance such as the one in (4). Our general aim is to prove results on the diffusion approximation for systems where the “fast” component is represented by a process of this kind.

This paper is the first, preparatory and technical, step in performing the program mentioned above. For a general Markov process X , we investigate an appropriate case of Poisson equation (2). In doing so, we answer two main questions arising in the context of the problem of diffusion approximation.

The first of these questions is how to define a solution of equation (2). In the paper [7], a *weak* solution of equation (2) is studied with an elliptic operator of the second order. It is also shown in [7] that this solution is, as a rule, unbounded in \mathbb{R}^k if one does not assume that the right hand side of (2) converges sufficiently fast to zero as $|x| \rightarrow \infty$. This shows that if one views the operator L in (2) as the generator of the semigroup generated by the initial process in the space $\mathbb{B}(\mathbb{R}^k)$ of measurable bounded functions, for example, then the class of functions f such that equation (2) has a solution becomes too small. We propose to use the *extended generator* \mathcal{A} defined in terms of an appropriate version of the Dynkin formula (see Definition 2.1) as an analogue of the operator L in (2) for a general Markov process. In other words, for a given measurable function $f: \mathbb{X} \rightarrow \mathbb{R}$, we search for a function u such that

$$(5) \quad \mathcal{A}u = -f.$$

Since equation (5) is given in terms of the extended generator \mathcal{A} , we call it the *extended Poisson equation*.

The second question studied in this paper is related to the structure of the predictable component in the Doob–Meyer decomposition of stochastic processes of the form

$$(6) \quad g(X(t), Y(t)), \quad t \in \mathbb{R}^+,$$

where X is the initial Markov process, Y is a process defined via the stochastic differential equation whose coefficients depend on X , and where $g = g(x, y)$ is a family of solutions of equations of the form (5) with the right hand side $f = f(x, y)$ depending on a parameter y . The processes described above appear in the context of the theory of diffusion approximation, since the standard method related to the *corrector* (see, for example, the proof of Theorem 3 in [7]) requires some properties of the semimartingale components of this corrector. These properties are proved in [7] and [21] with the help of an appropriate version of the Itô formula. Note, however, that this method depends essentially on a structure of the process X . In particular, some results on local properties of weak solutions of partial parabolic and elliptic equations of the second kind are used in [7] and [21]. On the other hand, this method is rather intractable in a quite simple case of a process X defined via a stochastic differential equation with a Lévy noise, since the analogs of the analytic results mentioned above do not exist in a needed generality for solutions of equations with pseudo-differential operators.

In the current paper, we propose another approach that does not require any specific assumption about the structure of a process and that allows one to describe the structure of the predictable component of stochastic processes represented in the form (6). In a forthcoming paper [22], these results will be applied to the problem of diffusion approximation.

2. NOTATION AND AUXILIARY RESULTS

Throughout the paper $X = \{X(t), t \in \mathbb{R}^+\}$ denotes a homogeneous Feller Markov process taking values in a separable metric space \mathbb{X} . We denote the transition probability of X by $P_t(x, dy)$ and assume that the process X has a càdlàg version for each distribution of the initial value $X(0)$. We denote by \mathbb{P}_x and \mathbb{E}_x , $x \in \mathbb{X}$, the distribution of this version in $\mathbb{D}(\mathbb{R}^+, \mathbb{X})$ with $X(0) = x$ and the corresponding expectation, respectively. The natural filtration of the process X is denoted by $\mathbb{F}^X = \{\mathcal{F}_t^X, t \in \mathbb{R}^+\}$.

Definition 2.1. We say that a measurable function $f: \mathbb{X} \rightarrow \mathbb{R}$ belongs to the domain of the extended generator \mathcal{A} of the process X if there exists a measurable function $g: \mathbb{X} \rightarrow \mathbb{R}$ such that the process

$$(7) \quad f(X(t)) - \int_0^t g(X(s)) ds, \quad t \in \mathbb{R}^+,$$

is well defined and is a local \mathbb{F}^X -martingale with respect to every measure \mathbb{P}_x , $x \in \mathbb{X}$. For such a pair (f, g) , we write $f \in \text{Dom}(\mathcal{A})$ and $\mathcal{A}f \stackrel{\text{df}}{=} g$.

The latter definition is motivated by the Dynkin formula. According to the Dynkin formula, if f belongs to the domain of a (usual) generator A , then $f \in \text{Dom}(\mathcal{A})$ and $\mathcal{A}f = Af$. Note, however, that $\text{Dom}(\mathcal{A})$ is usually wider than the domain of a (usual) generator A . For example, $\text{Dom}(\mathcal{A})$ does not necessarily consist of only bounded functions.

The notion of the extended generator traces back to the paper [23]. There are several technical modifications of the definition of the extended generator given in [23]. The definition above coincides with that introduced in [25] and is more general than that given in [24], Chapter VII, since we require that the process (7) is a *local* martingale and do not require that it is a martingale. Moreover, we do not require that the trajectories of this process are right continuous. Note that the function g in Definition 2.1 is not, generally speaking, uniquely defined; that is, \mathcal{A} is a multivalent mapping (see the discussion after Definition 1.8 in [24], Chapter VII).

Definition 2.2. We say that a measurable function $f: \mathbb{X} \rightarrow \mathbb{R}$ belongs to the domain of the *extended potential* \mathcal{R} of a process X if

$$\int_0^\infty |\mathbb{E}_x f(X(t))| dt < +\infty, \quad x \in \mathbb{X}.$$

The action of the extended potential at the function f is defined by

$$(8) \quad \mathcal{R}f(x) = \int_0^\infty \mathbb{E}_x f(X(t)) dt, \quad x \in \mathbb{X}.$$

Theorem 3.1 makes clear a relationship between the extended potential \mathcal{R} and equation (5). On one hand, $\mathcal{R}f$ is a solution of equation (5) under some additional assumptions. On the other hand, this equation cannot have other solutions in a certain class of functions.

Let d be a *premetric*, that is, a nonnegative, symmetric, lower semicontinuous function defined in $\mathbb{X} \times \mathbb{X}$ that vanishes at the “diagonal” $\{(x, x), x \in \mathbb{X}\} \subset \mathbb{X} \times \mathbb{X}$ and that takes positive values outside the “diagonal”. For arbitrary probability measures μ and ν in \mathbb{X} , put

$$d(\mu, \nu) = \inf_{(\xi, \eta) \in \mathcal{C}(\mu, \nu)} \mathbb{E} d(\xi, \eta),$$

where $\mathcal{C}(\mu, \nu)$ is a family of pairs (ξ, η) of random elements taking values in \mathbb{X} and such that $\xi \stackrel{d}{=} \mu$ and $\eta \stackrel{d}{=} \nu$. The latter definition is a version of the *minimal* (or *coupling*) distance between the probability distributions; see, for example, [26] or [27], Section 11.8.

The following is the main assumption on the ergodic properties of the stochastic process X , that is, an assumption on the limit behavior of its transition probabilities as $t \rightarrow +\infty$.

$\mathbf{E}(d, r, \psi)$. The process X has a unique invariant measure π and, moreover,

$$(9) \quad d(P_t(x, \cdot), \pi) \leq r(t)\psi(x), \quad x \in \mathbb{X}, t \geq 0,$$

where d is a premetric, the measurable function ψ takes values in $[1, +\infty)$, and the function $r: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is bounded, measurable, and such that $r(t) \rightarrow 0$ as $t \rightarrow \infty$.

It is sometimes convenient to accompany condition $\mathbf{E}(d, r, \psi)$ with the following auxiliary assumption.

$\hat{\mathbf{E}}(d, r, \psi)$. For given functions $r: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ and $\psi: \mathbb{X} \rightarrow [1, +\infty)$ and a premetric d ,

$$d(P_t(x, \cdot), P_t(y, \cdot)) \leq r(t)d(x, y) (\psi(x) + \psi(y)), \quad x, y \in \mathbb{X}, t \geq 0.$$

Below are two typical examples.

Example 2.1. Let $d(x, y) = \mathbb{1}_{x \neq y}$ be a discrete metric. Then

$$d(\mu, \nu) = \frac{1}{2} \|\mu - \nu\|_{\text{var}}$$

(see [29]), and thus condition $\mathbf{E}(d, r, \psi)$ coincides with condition (3). By the triangle inequality,

$$d(P_t(x, \cdot), P_t(y, \cdot)) \leq d(P_t(x, \cdot), \pi) + d(P_t(y, \cdot), \pi),$$

and we conclude that condition (3) also implies $\hat{\mathbf{E}}(d, r, \psi)$, since $d(x, y) = 1$ for $x \neq y$. As mentioned in Section 1, there are a number of papers devoted to checking condition (3) (see, for example, [8]–[13]; this list is not exhaustive).

Example 2.2. Let ρ be the initial metric in the space \mathbb{X} and let $d = \rho^p$, where $p \in [1, +\infty)$. Then

$$d(\mu, \nu) = [W_{\rho, p}(\mu, \nu)]^p,$$

where $W_{\rho, p}$ is the probabilistic *Kantorovich–Rubinstein metric*. In modern literature, $W_{\rho, p}$ is often called the p^{th} *Wasserstein metric*, although this name contradicts the background (see Section 1.1.3 in [26] and the references therein).

A simple sufficient condition for $\mathbf{E}(d, r, \psi)$ and $\hat{\mathbf{E}}(d, r, \psi)$ with an exponential function r is the so-called dissipative condition (see, for example, [14, Section 11.5] and [15, Section 16.2]). The cases $p = 2$ and $p = 1$ are treated in the papers [14] and [15], respectively. The same technique (with appropriate modifications) can also be used for other values of p .

Below is another moment assumption imposed on the transition probabilities.

$\mathbf{M}_p(\phi, \psi)$. For functions $\phi, \psi: \mathbb{X} \rightarrow [1, \infty)$ and a number $p \in [1, \infty]$,

$$(10) \quad \int_{\mathbb{X}} \phi^p(y) P_t(x, dy) \leq \psi(x), \quad x \in \mathbb{X}, t \geq 0.$$

Here the function ψ belongs to $L_1(\pi)$ (we agree that $1^\infty = 1$ and $a^\infty = \infty$ for $a > 1$).

If $\phi \equiv 1$, then condition $\mathbf{M}_p(\phi, \psi)$ obviously holds with an arbitrary number $p \in [1, \infty]$. This condition for an unbounded ϕ function follows from an appropriate version of the Lyapunov condition (see, for example, Proposition 1 in [7], inequalities (4) and (7) therein).

The function ψ involved in the right hand sides of inequalities (9) and (10) that describe the ergodic and moment properties of the stochastic process X , respectively, can be interpreted as a “penalty”. Note that the “penalty” depends on the initial state of this process in our case. The following condition controls, in a certain sense, the value of the “penalty”.

$\mathbf{P}(\psi)$. For an arbitrary $x \in \mathbb{X}$,

$$\mathbf{E}_x \int_0^\infty e^{-\lambda t} \psi(X(t)) dt < \infty, \quad \lambda > 0.$$

For $p \in [1, \infty)$ and a function $\phi: \mathbb{X} \rightarrow [1, +\infty)$, we denote by $H_{\phi,d,p}$ the class of functions $f: \mathbb{X} \rightarrow \mathbb{R}$ such that

$$(11) \quad \|f\|_{\phi,d,p} \stackrel{\text{df}}{=} \sup_{x \neq y} \frac{|f(x) - f(y)|}{d^{1/p}(x,y)(\phi(x) + \phi(y))} < +\infty.$$

It is natural to treat the class $H_{\phi,d,p}$ as a “weight Hölder space” of order $1/p$ and weight ϕ with respect to a premetric d . A “weight” function ϕ , being unbounded in the general case, is used in this definition in order to extend the corresponding classes of functions.

Apart from the process X , we consider a Wiener process W taking values in the space \mathbb{R}^l and two Poisson point measures ν_1 and ν_2 in $\mathbb{R}^+ \times \mathbb{U}_1$ and $\mathbb{R}^+ \times \mathbb{U}_2$, respectively. Here \mathbb{U}_1 and \mathbb{U}_2 are some measurable Borel spaces. The intensity measures ν_1 and ν_2 are assumed to be of the form $dt \times \mu_1(du)$ and $dt \times \mu_2(du)$, respectively. We further assume that the Wiener process W , point measures ν_1 and ν_2 , and process X are jointly independent. Put

$$\mathbb{F} = \{\mathcal{F}_t, t \in \mathbb{R}^+\}, \quad \mathcal{F}_t = \sigma(X(s), W(s), \nu_{1,2}([0, s] \times \cdot), s \leq t).$$

Consider the following stochastic differential equation in \mathbb{R}^m :

$$(12) \quad \begin{aligned} Y(t) = Y(0) &+ \int_0^t a(X(s), Y(s)) ds + \int_0^t b(X(s), Y(s)) dW_s \\ &+ \int_0^t \int_{\mathbb{U}_1} c_1(X(s), Y(s-), u) \nu_1(ds, du) \\ &+ \int_0^t \int_{\mathbb{U}_2} c_2(X(s), Y(s-), u) \tilde{\nu}_2(ds, du), \quad t \geq 0, \end{aligned}$$

where $\tilde{\nu}_2(dt, du) = \nu_2(dt, du) - dt \mu_2(du)$ is the compensated point measure corresponding to the point measure $\nu_2(dt, du)$.

Proposition 2.1. *Assume that the coefficients a , b , c_1 , and c_2 of stochastic differential equation (12) satisfy the following versions of the linear growth condition and local Lipschitz condition:*

- (a) *there exists a function $K: \mathbb{X} \rightarrow \mathbb{R}^+$ being bounded in every compact set of \mathbb{X} and such that*

$$\begin{aligned} &|a(x, y)| + |b(x, y)| + \int_{\mathbb{U}_1} |c_1(x, y, u)| \mu_1(du) + \left(\int_{\mathbb{U}_2} |c_2(x, y, u)|^2 \mu_2(du) \right)^{1/2} \\ &\leq K(x)(1 + |y|) \end{aligned}$$

for all $x \in \mathbb{X}$ and $y \in \mathbb{R}^m$;

- (b) *for an arbitrary $R > 0$, there exists a function $K_R: \mathbb{X} \rightarrow \mathbb{R}^+$ being bounded in every compact set of \mathbb{X} and such that*

$$\begin{aligned} &|a(x, y_1) - a(x, y_2)| + |b(x, y_1) - b(x, y_2)| + \int_{\mathbb{U}_1} |c_1(x, y_1, u) - c_1(x, y_2, u)| \mu_1(du) \\ &+ \left(\int_{\mathbb{U}_2} |c_2(x, y_1, u) - c_2(x, y_2, u)|^2 \mu_2(du) \right)^{1/2} \\ &\leq K_R(x) |y_1 - y_2|, \end{aligned}$$

for all $x \in \mathbb{X}$ and $y_1, y_2 \in \{y: |y| \leq R\}$.

Then, given the process X and an \mathcal{F}_0 -measurable random variable $Y(0)$, there exists a unique strong solution Y of stochastic differential equation (12) with respect to the filtration \mathbb{F} . The two component process (X, Y) defined in this way is a homogeneous Feller Markov process in $\mathbb{X} \times \mathbb{R}^m$.

The proof of this result is similar to the standard proofs of theorems for stochastic differential equations (see for example, [30], Chapter II.2). Thus we omit the proof of Proposition 2.1 here.

3. MAIN RESULTS

In this section, we state the main results of the paper. The proof of these results are given in Section 4. Throughout the rest of the paper π denotes a unique invariant measure for the process X . A function f is called centered if

$$\int_{\mathbb{X}} f \, d\pi = 0.$$

The first result gives sufficient conditions for the existence of the extended potential and describes its main properties.

Theorem 3.1. *1. Let conditions $\mathbf{E}(d, r, \psi)$ and $\mathbf{M}_q(\phi, \psi)$ hold for a process X . Moreover, we assume that the function r is such that*

$$(13) \quad \int_0^\infty r^{1/p}(t) \, dt < +\infty,$$

where $1/p + 1/q = 1$ and $p \in [1, \infty)$.

Then, given an arbitrary centered function $f \in H_{\phi, d, p}$, the generalized potential $\mathcal{R}f$ is well defined by equality (8), is centered, and admits the following bound:

$$(14) \quad |\mathcal{R}f(x)| \leq 2\|f\|_{\phi, d, p} \left(\int_0^\infty r^{1/p}(t) \, dt \right) \left(\int_{\mathbb{X}} \psi \, d\pi \right)^{1/q} \psi(x), \quad x \in \mathbb{X}.$$

If, in addition, condition $\mathbf{P}(\psi)$ holds, then the function $\mathcal{R}f$ is a solution of equation (5). Moreover, the process

$$(15) \quad \mathcal{R}f(X(t)) + \int_0^t f(X(s)) \, ds, \quad t \in \mathbb{R}^+,$$

is a (usual) \mathbb{F}^X -martingale with respect to each of the measures \mathbf{P}_x , $x \in \mathbb{X}$.

2. Let, in addition, condition $\hat{\mathbf{E}}(d, r, \psi)$ hold. Then, for every centered function

$$f \in H_{\phi, d, p},$$

the generalized potential $\mathcal{R}f$ belongs to the class $H_{\psi, d, p}$ and admits the following bound:

$$(16) \quad \|\mathcal{R}f\|_{\psi, d, p} \leq 2^{1/p} \|f\|_{\phi, d, p} \left(\int_0^\infty r^{1/p}(t) \, dt \right).$$

3. Let all the assumptions of statement 1 hold. In addition, assume that conditions $\mathbf{E}(d, r, \chi)$ and $\mathbf{M}_q(\psi, \chi)$ hold with a certain function χ . Let a function $f \in H_{\phi, d, p}$ be centered and let u be a centered solution of equation (5) that belongs to the class $H_{\psi, d, p}$. Then

$$u = \mathcal{R}f.$$

According to Theorem 3.1, the extended potential $\mathcal{R}f$ is a unique centered solution of equation (5) that belongs to the class $H_{\psi, d, p}$ if the assumptions of statements 1 and 3 of Theorem 3.1 hold.

The second of the main theorems describes the predictable component of the Doob–Meyer decomposition of a process of form (6) with the function $g = g(x, y)$ defined as

an extended potential $\mathcal{R}f$ of a certain function of two arguments $f = f(x, y)$ and with a process Y defined via the stochastic differential equation (12). In the definition of this potential, the variable y is fixed. In other words, we put

$$(17) \quad g(x, y) = \int_0^\infty \mathbf{E}_x f(X(t), y) dt, \quad x \in \mathbb{X}, y \in \mathbb{R}^m.$$

For a function of two variables $x \in \mathbb{X}$ and $y \in \mathbb{R}^m$, we write

$$(18) \quad \begin{aligned} \mathcal{A}_Y f(x, y) = & \sum_i a_i(x, y) \partial_{y_i} f(x, y) + \frac{1}{2} \sum_{i,j=1}^m \sum_{k=1}^d b_{ik}(x, y) b_{jk}(x, y) \partial_{y_i y_j}^2 f(x, y) \\ & + \int_{\mathbb{U}_1} (f(x, y + c_1(x, y, u)) - f(x, y)) \mu_1(du) \\ & + \int_{\mathbb{U}_2} \left(f(x, y + c_2(x, y, u)) - f(x, y) \right. \\ & \left. - \sum_{i=1}^m (c_2)_i(x, y, u) \partial_{y_i} f(x, y) \right) \mu_2(du), \end{aligned}$$

assuming that the right hand side is well defined. Denote by $\mathcal{S}(T)$ the family of all stopping times with respect to the filtration \mathbb{F} that are bounded by the number T . Also put $\mathcal{S} = \bigcup_T \mathcal{S}(T)$.

Theorem 3.2. *Let conditions $\mathbf{E}(d, r, \psi)$, $\mathbf{M}_q(\phi, \psi)$, and $\mathbf{P}(\psi)$ hold for a process X . Assume that a function r satisfies condition (13) with $1/p + 1/q = 1$, $p \in [1, \infty)$. Further, let the assumptions of Proposition 2.1 hold for the coefficients of stochastic differential equation (12). Assume further that a function $f: \mathbb{X} \times \mathbb{R}^m \rightarrow \mathbb{R}$ is twice continuously differentiable with respect to the variable y and that, for each fixed $y \in \mathbb{R}^m$, a function $f(\cdot, y)$ is centered. Finally assume that*

$$(19) \quad \sup_{y \in \mathbb{R}^m} \left(\|f(\cdot, y)\|_{\phi, d, p} + \sum_{i=1}^m \|\partial_{y_i} f(\cdot, y)\|_{\phi, d, p} + \sum_{i,j=1}^m \left\| \partial_{y_i y_j}^2 f(\cdot, y) \right\|_{\phi, d, p} \right) < \infty,$$

$$(20) \quad \sup_{y \in \mathbb{R}^m} \left(\left| \int_{\mathbb{X}} f(x, y) \pi(dx) \right| + \sum_{i=1}^m \left| \int_{\mathbb{X}} \partial_{y_i} f(x, y) \pi(dx) \right| + \sum_{i,j=1}^m \left| \int_{\mathbb{X}} \partial_{y_i y_j}^2 f(x, y) \pi(dx) \right| \right) < \infty.$$

Then the function g is well defined by equality (17) and is twice continuously differentiable with respect to the variable y . Moreover, the function $\mathcal{A}_Y g$ is well defined.

If the initial value $X(0)$ of the process X is such that

$$(21) \quad \mathbf{E} \int_0^\infty e^{-\lambda t} \psi(X(t)) dt < \infty, \quad \lambda > 0,$$

$$(22) \quad \mathbf{E} \psi(X(\tau)) < \infty, \quad \tau \in \mathcal{S},$$

then, for an arbitrary \mathcal{F}_0 -measurable initial value $Y(0)$ of Y , the stochastic process

$$(23) \quad g(X(t), Y(t)) + \int_0^t [f(X(s), Y(s)) - \mathcal{A}_Y g(X(s), Y(s))] ds, \quad t \in \mathbb{R}^+,$$

is a local \mathbb{F} -martingale.

Furthermore, if, for all $T \in \mathbb{R}^+$, the family of random variables $\psi(X(\tau))$, $\tau \in \mathcal{S}(T)$, is uniformly integrable and if there exists a process Z such that

$$(24) \quad |f(X(s), Y(s)) - \mathcal{A}_Y g(X(s), Y(s))| \leq Z(s), \quad s \in \mathbb{R}^+,$$

with probability one and

$$(25) \quad \mathbf{E} \int_0^t Z(s) ds < +\infty, \quad t \in \mathbb{R}^+,$$

then the process (23) is a (usual) \mathbb{F} -martingale.

Remark 3.1. Let all the assumptions of Theorem 3.2 hold. Assume also that, for an arbitrary $x \in \mathbb{X}$, the initial value $X(0) = x$ satisfies condition (22). Then the following result is of its own interest: the extended generator of the two dimensional Markov process (X, Y) acts as follows:

$$(26) \quad \mathcal{A}_{(X,Y)} = \mathcal{A}_X + \mathcal{A}_Y,$$

at the functions of the form (17), where \mathcal{A}_X denotes the action of the extended generator of the process X at a function of two variables (the action is considered with respect to the variable x if the variable y is fixed).

Below is an another version of the preceding result where we weaken the assumptions imposed on the process X . Instead, we strengthen the assumptions on the coefficients of stochastic differential equation (12).

Theorem 3.3. *Let all the assumptions of Theorem 3.2 hold except for (22). Instead, we assume that*

$$(27) \quad \mathbf{E} \psi(X(t)) < \infty, \quad t \in \mathbb{R}^+,$$

and that there exists a function $\varrho: \mathbb{X} \rightarrow \mathbb{R}^+$ such that

$$\mathbf{E} \int_0^\infty e^{-\lambda t} \psi(X(t)) \varrho(X(t)) dt < \infty, \quad \lambda > 0.$$

We finally assume that

$$|a(x, y)| + |b(x, y)|^2 + \int_{\mathbb{U}_1} |c_1(x, y, u)| \mu_1(du) + \int_{\mathbb{U}_2} |c_2(x, y, u)|^2 \mu_2(du) \leq \varrho(x)$$

for all $x \in \mathbb{X}$ and $y \in \mathbb{R}^m$.

Then the process given by (23) is a \mathbb{F} -martingale.

4. PROOFS

4.1. Auxiliary assertions.

Proposition 4.1. *Let $p \in [1, \infty)$. Assume that conditions $\mathbf{E}(d, r, \psi)$ and $\mathbf{M}_q(\phi, \psi)$ with $1/p + 1/q = 1$ hold for a process X . Then $H_{\phi, d, p} \subset L_1(\pi)$ and*

$$(28) \quad \left| \mathbf{E}_x f(X(t)) - \int_{\mathbb{X}} f d\pi \right| \leq 2 \|f\|_{\phi, d, p} \left(\int_{\mathbb{X}} \psi d\pi \right)^{1/q} r^{1/p}(t) \psi(x), \quad x \in \mathbb{X},$$

for all $t \in \mathbb{R}^+$ and $f \in H_{\phi, d, p}$.

If, additionally, condition $\hat{\mathbf{E}}(d, r, \psi)$ holds, then

$$(29) \quad \left| \mathbf{E}_x f(X(t)) - \mathbf{E}_{x'} f(X(t)) \right| \leq 2^{1/p} \|f\|_{\phi, d, p} (\psi(x) + \psi(x')) r^{1/p}(t) d^{1/p}(x, x'),$$

$$x, x' \in \mathbb{X},$$

for all $f \in H_{\phi, d, p}$ and $t \in \mathbb{R}^+$.

Proof. Let $x_* \in \mathbb{X}$ be a fixed point. Then

$$(30) \quad |f(x) - f(x_*)| \leq \|f\|_{\phi, d, p} d^{1/p}(x, x_*)(\phi(x) + \phi(x_*)), \quad x \in \mathbb{X},$$

for $f \in H_{\phi, d, p}$. According to condition $\mathbf{M}_q(\phi, \psi)$, the function ϕ belongs to the space $L_q(\pi)$, that is,

$$(31) \quad \int_{\mathbb{X}} \phi^q(x) \pi(dx) = \int_{\mathbb{X}} \int_{\mathbb{X}} \phi^q(y) P_t(x, dy) \pi(dx) \leq \int_{\mathbb{X}} \psi(x) \pi(dx).$$

On the other hand, inequality (9) applied with $x = x_*$ and $t = 0$ proves that the function $d^{1/p}(\cdot, x_*)$ belongs to $L_p(\pi)$. This implies that $f \in L_1(\pi)$.

To prove bound (28), we consider an arbitrary pair

$$(\xi, \eta) \in C(P_t(x, \cdot), \pi).$$

Then we write

$$\begin{aligned} \left| \mathbf{E}_x f(X(t)) - \int_{\mathbb{X}} f d\pi \right| &= |\mathbf{E}_x f(X(t)) - \mathbf{E}_\pi f(X(t))| \\ &= |\mathbf{E} f(\xi) - \mathbf{E} f(\eta)| \leq \mathbf{E} |f(\xi) - f(\eta)|. \end{aligned}$$

Now we substitute random variables ξ and η in (30) for x and x_* , respectively. Then take the expectation, and finally apply the Hölder inequality:

$$\mathbf{E} |f(\xi) - f(\eta)| \leq \|f\|_{\phi, d, p} (\mathbf{E} d(\xi, \eta))^{1/p} (\mathbf{E}(\phi(\xi) + \phi(\eta))^q)^{1/q}.$$

Hence

$$\mathbf{E} \phi^q(\eta) \leq \int_{\mathbb{X}} \psi d\pi$$

(see inequality (31)). Since the function $z \mapsto z^q$ is convex,

$$\mathbf{E}(\phi(\xi) + \phi(\eta))^q \leq 2^{q-1} (\mathbf{E} \phi^q(\xi) + \mathbf{E} \phi^q(\eta)) \leq 2^{q-1} \left(\psi(x) + \int_{\mathbb{X}} \psi d\pi \right)$$

in view of condition $\mathbf{M}_q(\phi, \psi)$. Taking into account the bound $\psi \geq 1$, we finally obtain

$$\mathbf{E}(\phi(\xi) + \phi(\eta))^q \leq 2^q \left(\int_{\mathbb{X}} \psi d\pi \right) \psi(x).$$

Therefore

$$\left| \mathbf{E}_x f(X(t)) - \int_{\mathbb{X}} f d\pi \right| \leq 2 \|f\|_{\phi, d, p} \left(\int_{\mathbb{X}} \psi d\pi \right)^{1/q} (\mathbf{E} d(\xi, \eta))^{1/p} \psi^{1/q}(x).$$

Since the pair $(\xi, \eta) \in C(P_t(x, \cdot), \pi)$ in this inequality is arbitrary, we conclude that

$$\left| \mathbf{E}_x f(X(t)) - \int_{\mathbb{X}} f d\pi \right| \leq 2 \|f\|_{\phi, d, p} \left(\int_{\mathbb{X}} \psi d\pi \right)^{1/q} (d(P_t(x, \cdot), \pi))^{1/p} \psi^{1/q}(x),$$

whence the desired result follows by condition $\mathbf{E}(d, r, \psi)$.

The proof of bound (29) is analogous: for an arbitrary pair

$$(\xi, \eta) \in C(P_t(x, \cdot), P_t(x', \cdot))$$

we get

$$\mathbf{E}(\phi(\xi) + \phi(\eta))^q \leq 2^{q-1} (\psi(x) + \psi(x')) = 2^{q/p} (\psi(x) + \psi(x'))$$

by inequality (10). Reasoning as above we prove that

$$|\mathbf{E}_x f(X(t)) - \mathbf{E}_{x'} f(X(t))| \leq 2^{1/p} \|f\|_{\phi, d, p} (\psi(x) + \psi(y))^{1/q} (d(P_t(x, \cdot), P_t(x', \cdot)))^{1/p}.$$

The latter inequality, together with condition $\hat{\mathbf{E}}(d, r, \psi)$, completes the proof. □

We mention a simple corollary that is important for our reasoning below. Putting $t = 0$ in inequality (28), we obtain the following bound:

$$(32) \quad |f(x)| \leq \left| \int_{\mathbb{X}} f \, d\pi \right| + L \|f\|_{\phi, d, p} \psi(x), \quad x \in \mathbb{X},$$

where

$$L = 2r^{1/p}(0) \left(\int_{\mathbb{X}} \psi \, d\pi \right)^{1/q}.$$

This implies that if the function $f = f(x, y)$ satisfies the conditions of Theorem 3.2, then the absolute value of the function f and the absolute values of its partial derivatives of the first order with respect to y and those of the second order are bounded by $C\psi(x)$. Here and in what follows, C denotes a constant whose precise value does not matter for us.

The following result follows from the above estimates, Fubini’s theorem, and the Lebesgue dominated convergence theorem. Hence we omit the details of the proof. For $\lambda \geq 0$, put

$$(33) \quad g_\lambda(x, y) = \int_0^\infty e^{-\lambda t} \mathbf{E}_x f(X(t), y) \, dt.$$

It is clear that the function g defined by (17) coincides with g_0 .

Proposition 4.2. *Let conditions $\mathbf{E}(d, r, \psi)$, $\mathbf{M}_p(\phi, \psi)$, and $\mathbf{P}(\psi)$ hold for the process X with some premetric d , functions ϕ, ψ , and r , and a number $p \in (1, \infty)$. We further assume that the function r satisfies condition (13) and that the function f satisfies the assumptions of Theorem 3.2.*

Then the function $g_\lambda(x, y)$ is well defined by relation (33). Moreover, this function is twice continuously differentiable with respect to y and its derivatives are equal to

$$(34) \quad \begin{aligned} \partial_{y_i} g_\lambda(x, y) &= \int_0^\infty e^{-\lambda t} \mathbf{E}_x \partial_{y_i} f(X(t), y) \, dt, \\ \partial_{y_i y_j}^2 g_\lambda(x, y) &= \int_0^\infty e^{-\lambda t} \mathbf{E}_x \partial_{y_i y_j}^2 f(X(t), y) \, dt. \end{aligned}$$

If x and y are fixed, then the function $g_\lambda(x, y)$ and its derivatives (34) are continuous with respect to λ .

Note that inequality (28) and definition (33) imply

$$(35) \quad |g_\lambda(x, y)| \leq C\psi(x), \quad |\partial_{y_i} g_\lambda(x, y)| \leq C\psi(x), \quad \left| \partial_{y_i y_j}^2 g_\lambda(x, y) \right| \leq C\psi(x).$$

Applying the finite difference formula and condition of the linear growth imposed on the coefficients of stochastic differential equation (12), we prove that $\mathcal{A}_Y g_\lambda$, $\lambda \geq 0$, is well defined by relation (18),

$$\mathcal{A}_Y g_\lambda(x, y) \rightarrow \mathcal{A}_Y g_0(x, y), \quad \lambda \rightarrow 0+, \quad x \in \mathbb{X}, \quad y \in \mathbb{R}^m,$$

and that, for an arbitrary compact set $K \subset \mathbb{X}$ and a number $R > 0$, there exists a constant $C_{K,R}$ such that

$$(36) \quad |\mathcal{A}_Y g_\lambda(x, y)| \leq C_{K,R} \psi(x), \quad x \in K, \quad |y| \leq R, \quad \lambda \geq 0.$$

4.2. Proof of the main results.

Proof of Theorem 3.2. As above, we do not restrict the generality if we assume that $m = 1$. Let τ be a Markov moment with respect to the filtration \mathbb{F} . The process X possesses the Markov property with respect to \mathbb{F} and, moreover, X is a Feller process and has càdlàg trajectories. Thus X is a strictly Markov process, whence

$$\begin{aligned}
 & e^{-\lambda(t \wedge \tau)} g_\lambda(X(t \wedge \tau), Y(t \wedge \tau)) \\
 &= e^{-\lambda(t \wedge \tau)} \int_0^\infty e^{-\lambda s} \mathbb{E}_x f(X(s), y) \Big|_{(x,y)=(X(t \wedge \tau), Y(t \wedge \tau))} ds \\
 (37) \quad &= e^{-\lambda(t \wedge \tau)} \int_0^\infty e^{-\lambda s} \mathbb{E} [f(X(s + t \wedge \tau), Y(t \wedge \tau)) \mid \mathcal{F}_{t \wedge \tau}] ds \\
 &= \int_{t \wedge \tau}^\infty e^{-\lambda v} \mathbb{E} [f(X(v), Y(t \wedge \tau)) \mid \mathcal{F}_{t \wedge \tau}] dv
 \end{aligned}$$

almost surely. We introduce the following auxiliary filtration:

$$\mathbb{H} = \{\mathcal{H}_t\}, \quad \mathcal{H}_t = \sigma(X(v), W(s), \nu_{1,2}([0, s] \times \cdot), s \leq t, v \in \mathbb{R}^+), \quad t \in \mathbb{R}^+.$$

It is clear that \mathbb{H} is an extension of \mathbb{F} . Nevertheless W and ν_1 and ν_2 are independent Wiener process and Poisson point measures with respect to the filtration \mathbb{H} , respectively. In particular, the process Y is a \mathbb{H} -semimartingale.

For almost all positive v , we have

$$\begin{aligned}
 f(X(v), Y(v)) &= f(X(v), Y(t)) + \int_t^v \partial_y f(X(v), Y(s)) a(X(s), Y(s)) ds \\
 &+ \int_t^v \partial_y f(X(v), Y(s)) b(X(s), Y(s)) dW(s) \\
 &+ \frac{1}{2} \int_t^v \partial_{yy}^2 f(X(v), Y(s)) b^2(X(s), Y(s)) ds \\
 &+ \int_t^v \int_{\mathbb{U}_1} [f(X(v), Y(s-) + c_1(X(s), Y(s-), u)) \\
 (38) \quad &\qquad\qquad\qquad - f(X(v), Y(s-))] \nu_1(ds, du) \\
 &+ \int_t^v \int_{\mathbb{U}_2} [f(X(v), Y(s-) + c_2(X(s), Y(s-), u)) \\
 &\qquad\qquad\qquad - f(X(v), Y(s-))] \tilde{\nu}_2(ds, du) \\
 &+ \int_t^v \int_{\mathbb{U}_2} [f(X(v), Y(s) + c_2(X(s), Y(s), u)) \\
 &\qquad\qquad\qquad - f(X(v), Y(s)) \\
 &\qquad\qquad\qquad - \partial_y f(X(v), Y(s)) c_2(X(s), Y(s), u)] \mu_2(du) ds, \\
 &\qquad\qquad\qquad t \in [0, v].
 \end{aligned}$$

Relation (38) follows from the Itô formula (see Theorem 5.1 of Chapter II in [28]) applied to the \mathbb{H} -semimartingale Y if the function

$$F = f(X(v), \cdot)$$

belongs to the class C^2 with respect to the variable y . The function F is nonrandom in the theorem cited above, and thus we cannot get (38) as a corollary of this theorem. Nevertheless, $F = f(X(v), \cdot)$ is \mathcal{H}_0 -measurable as a function of a random argument, and moreover the absolute value of this function and the absolute values of its derivatives of the first and second orders are bounded from above by the random variable $C\psi(X(v))$

being integrable for almost all $v > 0$ in view of (21). Reasoning similarly to the proof of Theorem 5.1 of Chapter II in [28], we obtain equality (38) (we omit the details).

For a compact set $K \subset \mathbb{X}$ and a number $R > 0$, we put

$$\tau_{K,R} = \inf\{t: X(t) \notin K \text{ or } |Y(t)| > R\}.$$

Since the trajectories of (X, Y) are càdlàg, there exists a sequence of compact sets $\{K_n\}$ such that

$$(39) \quad \tau_n := \tau_{K_n, n} \rightarrow \infty, \quad n \rightarrow \infty,$$

with probability one. On the other hand, the condition of the linear growth imposed on the coefficients of stochastic differential equation (12) implies that there are real constants C_n such that

$$(40) \quad \begin{aligned} & |a(X(s), Y(s))| + |b(X(s), Y(s))|^2 + \int_{\mathbb{U}_1} |c_1(X(s), Y(s), u)| \mu_1(du) \\ & + \int_{\mathbb{U}_2} |c_2(X(s), Y(s), u)|^2 \mu_2(du) \\ & \leq C_n, \quad s < \tau_n. \end{aligned}$$

Taking into account the bounds

$$(41) \quad |F(y+c) - F(y)| \leq |c| \sup_z |F'(z)|, \quad |F(y+c) - F(y) - F'(y)c| \leq \frac{c^2}{2} \sup_z |F''(z)|,$$

and that the function $F = f(X(v), \cdot)$ as well as its derivatives of the first and second orders are bounded by the random variable $C\psi(X(v))$, we conclude from equality (38) that

$$(42) \quad \begin{aligned} & f(X(v), Y(v)) - f(X(v), Y(t)) \\ & = M_v(v) - M_v(t) + \int_t^v \partial_y f(X(v), Y(s)) a(X(s), Y(s)) ds \\ & + \frac{1}{2} \int_t^v \partial_{yy}^2 f(X(v), Y(s)) b^2(X(s), Y(s)) ds \\ & + \int_t^v \int_{\mathbb{U}_1} [f(X(v), Y(s) + c_1(X(s), Y(s), u)) - f(X(v), Y(s))] \mu_1(du) ds \\ & + \int_t^v \int_{\mathbb{U}_2} [f(X(v), Y(s) + c_2(X(s), Y(s), u)) - f(X(v), Y(s)) \\ & \quad - \partial_y f(X(v), Y(s)) c_2(X(s), Y(s), u)] \mu_2(du) ds \\ & = M_v(v) - M_v(t) + \sum_{i=1}^4 (I_{i,v}(v) - I_{i,v}(t)), \quad t \in [0, v], \end{aligned}$$

for almost all $v > 0$, where $M_v(\cdot)$ is a local martingale with respect to \mathbb{H} with a localizing sequence $\{\tau_n\}$ defined by (39). Since $\mathcal{H}_t \supset \mathcal{F}_t$, we have

$$\mathbb{E} [M_v(v \wedge \tau_n) - M_v(t \wedge \tau_n) \mid \mathcal{F}_{t \wedge \tau_n}] = \mathbb{E} \left[\mathbb{E} [M_v(v \wedge \tau_n) - M_v(t \wedge \tau_n) \mid \mathcal{H}_t] \mid \mathcal{F}_{t \wedge \tau_n} \right] = 0$$

for $t \in [0, v]$. Now we deduce from (42) that equality

$$(43) \quad \begin{aligned} \mathbb{E} [f(X(v), Y(t \wedge \tau_n)) \mid \mathcal{F}_{t \wedge \tau_n}] & = \mathbb{E} [f(X(v), Y(v \wedge \tau_n)) \mid \mathcal{F}_{t \wedge \tau_n}] \\ & - \sum_{i=1}^4 \mathbb{E} [I_{i,v}(v \wedge \tau_n) - I_{i,v}(t \wedge \tau_n) \mid \mathcal{F}_{t \wedge \tau_n}] \end{aligned}$$

holds with probability one for almost all $v > 0$ and all $t \in [0, v]$. On the other hand, if $t > v$, then equality (43) holds in the set $\{v > \tau_n\}$ since $t \wedge \tau_n = v \wedge \tau_n = \tau_n$. Therefore equality (43) holds with probability one for almost all $v > t \wedge \tau_n$ if $t \geq 0$ is arbitrary.

Next we use (43) in equality (37) with $\tau = \tau_n$. The terms in (37) are considered separately. Since

$$|f(X(v), y)| \leq C\psi(X(v)),$$

we apply Fubini's theorem and derive from condition (21) that

$$(44) \quad \begin{aligned} & \int_{t \wedge \tau_n}^{\infty} e^{-\lambda v} \mathbf{E} [f(X(v), Y(v)) \mid \mathcal{F}_{t \wedge \tau_n}] dv \\ &= \mathbf{E} \left[\int_{t \wedge \tau_n}^{\infty} e^{-\lambda v} f(X(v), Y(v)) dv \mid \mathcal{F}_{t \wedge \tau_n} \right]. \end{aligned}$$

Using inequality (40) and an analogous bound for the derivative of the function $f(X(v), \cdot)$, we apply Fubini's theorem to the term in (43) corresponding to $i = 1$. Then

$$\begin{aligned} & \int_{t \wedge \tau_n}^{\infty} e^{-\lambda v} \mathbf{E} [I_{1,v}(v \wedge \tau_n) - I_{1,v}(t \wedge \tau_n) \mid \mathcal{F}_{t \wedge \tau_n}] dv \\ &= \int_{t \wedge \tau_n}^{\infty} e^{-\lambda v} \int_{t \wedge \tau_n}^{v \wedge \tau_n} \mathbf{E} [\partial_y f(X(v), Y(s)) a(X(s), Y(s)) \mid \mathcal{F}_{t \wedge \tau_n}] ds dv \\ &= \mathbf{E} \left[\int_{t \wedge \tau_n}^{\tau_n} \int_s^{\infty} e^{-\lambda v} \mathbf{E} [\partial_y f(X(v), Y(s)) a(X(s), Y(s)) \mid \mathcal{F}_s] dv ds \mid \mathcal{F}_{t \wedge \tau_n} \right]. \end{aligned}$$

Applying the Markov property of the process X and proceeding similarly to (37), we get

$$\begin{aligned} & \int_s^{\infty} e^{-\lambda v} \mathbf{E} [\partial_y f(X(v), Y(s)) a(X(s), Y(s)) \mid \mathcal{F}_s] dv \\ &= e^{-\lambda s} a(X(s), Y(s)) \int_0^{\infty} e^{-\lambda w} \mathbf{E} [\partial_y f(X(s+w), Y(s)) \mid \mathcal{F}_s] dw \\ &= e^{-\lambda s} a(X(s), Y(s)) \int_0^{\infty} e^{-\lambda w} [\mathbf{E}_x \partial_y f(X(w), y)]_{(x,y)=(X(s), Y(s))} dw \\ &= e^{-\lambda s} a(X(s), Y(s)) \partial_y g_{\lambda}(X(s), Y(s)). \end{aligned}$$

We refer to the first equality in (34) to justify the last line above. Therefore

$$(45) \quad \begin{aligned} & \int_{t \wedge \tau_n}^{\infty} e^{-\lambda v} \mathbf{E} [I_{1,v}(v \wedge \tau_n) - I_{1,v}(t \wedge \tau_n) \mid \mathcal{F}_{t \wedge \tau_n}] dv \\ &= \mathbf{E} \left[\int_{t \wedge \tau_n}^{\infty} e^{-\lambda s} a(X(s), Y(s)) \partial_y g_{\lambda}(X(s), Y(s)) ds \mid \mathcal{F}_{t \wedge \tau_n} \right]. \end{aligned}$$

Similar reasoning applies for the terms in (43) corresponding to $i = 2, 3, 4$, as well. A technical difference in the proof is that one needs to use a bound for the second derivative of the function $f(X(r), \cdot)$ for $i = 2, 4$ and estimates (41) in the case of $i = 3, 4$ to justify the application of Fubini's theorem. Omitting the details, we write the final relation in the proof:

$$\begin{aligned} & e^{-\lambda(t \wedge \tau_n)} g_{\lambda}(X(t \wedge \tau_n), Y(t \wedge \tau_n)) \\ &= \mathbf{E} \left[\int_{t \wedge \tau_n}^{\infty} e^{-\lambda s} (f(X(s), Y(s)) - \mathcal{A}_Y g_{\lambda}(X(s), Y(s))) ds \mid \mathcal{F}_{t \wedge \tau_n} \right]. \end{aligned}$$

Hence if $t_2 \geq t_1$, then

$$\begin{aligned} & \mathbb{E} \left[e^{-\lambda(t_2 \wedge \tau_n)} g_\lambda(X(t_2 \wedge \tau_n), Y(t_2 \wedge \tau_n)) \mid \mathcal{F}_{t_1 \wedge \tau_n} \right] - e^{-\lambda(t_1 \wedge \tau_n)} g_\lambda(X(t_1 \wedge \tau_n), Y(t_1 \wedge \tau_n)) \\ &= \mathbb{E} \left[\int_{t_1 \wedge \tau_n}^{t_2 \wedge \tau_n} e^{-\lambda s} (f(X(s), Y(s)) - \mathcal{A}_Y g_\lambda(X(s), Y(s))) ds \mid \mathcal{F}_{t_1 \wedge \tau_n} \right]. \end{aligned}$$

Now we pass to the limit $\lambda \rightarrow 0+$ in the latter equality. To justify this procedure, we refer to the Lebesgue dominated convergence theorem and use conditions (21) and (22). We also mention that (32), (35), and (36) imply

$$\begin{aligned} & |g_\lambda(X(t_2 \wedge \tau_n), Y(t_2 \wedge \tau_n))| \leq C\psi(X(t_2 \wedge \tau_n)), \\ & |f(X(s), Y(s)) - \mathcal{A}_Y g_\lambda(X(s), Y(s))| \mathbb{1}_{s \leq t_2 \wedge \tau_n} \leq C\psi(X(s)). \end{aligned}$$

For an arbitrary stopping time τ_n , $n \geq 1$, we have

$$(46) \quad \begin{aligned} & \mathbb{E} [g(X(t_2 \wedge \tau_n), Y(t_2 \wedge \tau_n)) \mid \mathcal{F}_{t_1 \wedge \tau_n}] - g(X(t_1 \wedge \tau_n), Y(t_1 \wedge \tau_n)) \\ &= \mathbb{E} \left[\int_{t_1 \wedge \tau_n}^{t_2 \wedge \tau_n} (f(X(s), Y(s)) - \mathcal{A}_Y g(X(s), Y(s))) ds \mid \mathcal{F}_{t_1 \wedge \tau_n} \right], \quad t_2 \geq t_1, \end{aligned}$$

whence we deduce that stochastic process (7) is a local \mathbb{F} -martingale. If, in addition, the family $\{\psi(X(\tau)), \tau \in \mathcal{S}(t_2)\}$ is uniformly integrable and if conditions (24) and (25) hold, then one can pass to the limit in (46) as $n \rightarrow \infty$ and prove that process (7) is a \mathbb{F} -martingale. \square

Proof of Theorem 3.3. The proof is similar to that presented above, and thus we do not discuss it in detail. From the beginning, we change all the Markov moments with $+\infty$. The Fubini theorem and Lebesgue dominated convergence theorem can nevertheless be applied, since, instead of inequalities (40) and (36), one can use the following bounds:

$$\begin{aligned} & |a(X(s), Y(s))| + |b(X(s), Y(s))|^2 + \int_{\mathbb{U}_1} |c_1(X(s), Y(s), u)| \mu_1(du) \\ & \quad + \int_{\mathbb{U}_2} |c_2(X(s), Y(s), u)|^2 \mu_2(du) \\ & \leq C\varrho(X(s)), \quad s \geq 0, \end{aligned}$$

$$|\mathcal{A}_Y g_\lambda(x, y)| \leq C\psi(x)\varrho(x), \quad x \in \mathbb{X}, y \in \mathbb{R}^m, \lambda \geq 0. \quad \square$$

Proof of Theorem 3.1. The existence of the generalized potential and its properties (14) and (16) follow directly from bounds (28) and (29). Further, let $a \equiv 1$, $b \equiv 0$, and $c_{1,2} \equiv 0$. Consider a function $f(x, y) = f(x)$ that does not depend on the “additional” variable y . Then all the assumptions of Theorem 3.3 hold with $\varrho \equiv 1$. Hence the generalized potential $\mathcal{R}f$ is a solution of equation (5) and process (15) is a \mathbb{F}^X -martingale with respect to every measure \mathbb{P}_x , $x \in \mathbb{X}$. This proves statements 1 and 2 of the theorem.

To prove statement 3, consider a certain solution u of equation (5). Then, by definition of the extended generator, the process

$$u(X(t)) + \int_0^t f(X(s)) ds, \quad t \geq 0,$$

is a martingale with respect to \mathbb{P}_x for all $x \in \mathbb{X}$. In particular,

$$\mathbb{E}_x u(X(t)) + \mathbb{E}_x \int_0^t f(X(s)) ds = u(x), \quad t \geq 0.$$

Thus $|f(x)| \leq C\psi(x)$, and condition $\mathbf{P}(\psi)$ justifies the application of Fubini's theorem, that is

$$u(x) = \int_0^t \mathbf{E}_x f(X(s)) ds + \mathbf{E}_x u(X(t)).$$

According to statement 1 proved above,

$$u(x) = \lim_{t \rightarrow +\infty} \left[\int_0^t \mathbf{E}_x f(X(s)) ds + \mathbf{E}_x u(X(t)) \right] = \mathcal{R}f(x) + \lim_{t \rightarrow +\infty} \mathbf{E}_x u(X(t)).$$

Now we apply bound (28) with u , χ , and ψ instead of f , ψ , and ϕ , respectively. As a result, we get

$$|\mathbf{E}_x u(X(t))| \leq 2\|u\|_{\psi, d, p} \left(\int_{\mathbb{X}} \chi d\pi \right)^{1/q} \chi(x)r^{1/p}(t), \quad t \geq 0.$$

Therefore $\mathbf{E}_x u(X(t)) \rightarrow 0$ as $t \rightarrow \infty$, and thus $u(x) = \mathcal{R}f(x)$. \square

BIBLIOGRAPHY

1. N. Abourashchi and A. Yu. Veretennikov, *On stochastic averaging and mixing*, *Theory Stoch. Process.* **16**(32) (2010), no. 1, 111–129. MR2779833 (2012a:60188)
2. R. Z. Khas'minskii, *A limit theorem for the solutions of differential equations with random right-hand sides*, **11** (1966), no. 3, 444–462; English transl. in *Theor. Probability Appl.* **11** (1966), no. 3, 390–406. MR0203789 (34:3637)
3. A. N. Borodin, *A limit theorem for the solutions of differential equations with a random right-hand side*, *Teor. Veroyatnost. i Primenen.* **22** (1977), no. 3, 498–512; English transl. in *Theor. Probability Appl.* **22** (1977), no. 3, 482–497. MR0517995 (58:24546)
4. V. V. Sarafyan and A. V. Skorokhod, *Dynamical systems with fast switchings*, *Teor. Veroyatnost. i Primenen.* **32** (1987), no. 4, 658–669; English transl. in *Theory Probab. Appl.* **32** (1987), no. 4, 595–607. MR927247 (89f:60064)
5. S. N. Ethier and T. G. Kurtz, *Markov Processes. Characterization and Convergence*, Wiley, New York, 1986. MR838085 (88a:60130)
6. V. S. Korolyuk and N. Limnios, *Stochastic Systems in Merging Phase Space*, World Scientific, New Jersey, 2005. MR2205562 (2007a:60004)
7. E. Pardoux and A. Yu. Veretennikov, *On Poisson equation and diffusion approximation 1*, *Ann. Probab.* **29** (2001), 1061–1085. MR1872736 (2002j:60120)
8. A. Yu. Veretennikov, *Bounds for the mixing rates in the theory of stochastic equations*, *Theory Probab. Appl.* **32** (1987), 273–281. MR902757 (89b:60144)
9. A. Yu. Veretennikov, *On polynomial mixing bounds for stochastic differential equations*, *Stoch. Process. Appl.* **70** (1997), 115–127. MR1472961 (99k:60158)
10. S. A. Klokov and A. Yu. Veretennikov, *Sub-exponential mixing rate for a class of Markov chains*, *Math. Comm.* **9** (2004), 9–26. MR2076227 (2005f:60148)
11. D. Down, S. P. Meyn, and R. L. Tweedie, *Exponential and uniform ergodicity of Markov processes*, *Ann. Probab.* **23** (1995) no. 4, 1671–1691. MR1379163 (97c:60181)
12. R. Douc, G. Fort, and A. Guillin, *Subgeometric rates of convergence of f -ergodic strong Markov processes*, *Stoch. Process. Appl.* **119** (2009), no. 3, 897–923. MR2499863 (2010j:60184)
13. A. M. Kulik, *Exponential ergodicity of the solutions to SDE's with a jump noise*, *Stoch. Proc. Appl.* **119** (2009), 602–632. MR2494006 (2010i:60176)
14. G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, 1992. MR1207136 (95g:60073)
15. S. Peszat and J. Zabczyk, *Stochastic Partial Differential Equations with Lévy Noise (an Evolution Equation Approach)*, Cambridge University Press, Cambridge, 2007. MR2356959 (2009b:60200)
16. M. Hairer and J. Mattingly, *Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing*, *Ann. Math.* **164** (2006), 993–1032. MR2259251 (2008a:37095)
17. M. Hairer and J. Mattingly, *Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations*, *Ann. Probab.* **36** (2008), 2050–2091. MR2478676 (2010i:35295)
18. T. Komorowski, Sz. Peszat, and T. Szarek, *On ergodicity of some Markov processes*, *Ann. Probab.* **38** (2010), 1401–1443. MR2663632 (2011f:60148)
19. T. Szarek, *The uniqueness of invariant measures for Markov operators*, *Studia Math.* **189** (2008), 225–233. MR2457488 (2009m:60166)

20. M. Hairer, J. C. Mattingly, and M. Scheutzow, *Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations*, Prob. Theory Rel. Fields (2009), DOI 10.1007/s00440-009-0250-6.
21. E. Pardoux and A. Yu. Veretennikov, *On Poisson equation and diffusion approximation 2*, Ann. Prob. **31** (2003), 1166–1192. MR1988467 (2004d:60156)
22. A. M. Kulik and A. Yu. Veretennikov, *On extended Poisson equation and diffusion approximation* (2011). (to appear) MR2933700 (2012m:60128)
23. H. Kunita, *Absolute continuity of Markov processes and generators*, Nagoya Math. J. **36** (1969), 1–26. MR0250387 (40:3626)
24. D. Revuz and M. and Yor, *Continuous Martingales and Brownian Motion*, third edition, Springer-Verlag, Berlin, 1999. MR1725357 (2000h:60050)
25. Z. Palmowski and T. Rolski, *A technique for exponential change of measure for Markov processes*, Bernoulli **8** (2002), no. 6, 767–785. MR1963661 (2004f:60152)
26. V. M. Zolotarev, *Modern theory of summation of random variables*, “Nauka”, Moscow, 1986; English transl., VSP, Utrecht, 1997. MR1640024 (99m:60002)
27. R. M. Dudley, *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2002. MR1932358 (2003h:60001)
28. N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland Publishing Co. and Kodansha, Ltd., Amsterdam–New York–Tokyo, 1981. MR637061 (84b:60080)
29. R. L. Dobrushin, *Prescribing a system of random variables by conditional distributions*, Teor. Veroyatnost. i Primenen. **15** (1970), no. 3, 469–497; English transl. in Theory Probab. Appl. **15** (1970), no. 3, 458–486. MR0298716 (45:7765)
30. I. I. Gihman and A. V. Skorohod, *Stochastic Differential Equations*, “Naukova Dumka”, Kiev, 1968; English transl., Springer-Verlag, Berlin–Heidelberg–New York, 1972. MR0346904 (49:11625)

SCHOOL OF MATHEMATICS, UNIVERSITY OF LEEDS, LS2 9JT, LEEDS, UNITED KINGDOM; INSTITUTE OF INFORMATION TRANSMISSION PROBLEMS, BOL'SHOI KARETNY STREET 19, 127994, MOSCOW, RUSSIA
E-mail address: A.Veretennikov@leeds.ac.uk

INSTITUTE OF MATHEMATICS, NATIONAL ACADEMY OF SCIENCE OF UKRAINE, TERESHCHENKIVS'KA STREET, 3, 01601 KYIV-4, UKRAINE
E-mail address: kulik@imath.kiev.ua

Received 06/SEP/2011
Translated by N. SEMENOV