Journal of Mathematical Sciences, Vol. 179, No. 1, November, 2011

ON SOBOLEV SOLUTIONS OF POISSON EQUATIONS IN $\mathbb{R}^d$ WITH A PARAMETER

A. Yu. Veretennikov

School of Mathematics, University of Leeds, LS2 9JT, UK

Institute of Information Transmission Problems, RAS, Moscow, Russia

a.veretennikov@leeds.ac.uk

To the 70th birthday of Professor N. V. Krylov

Smoothness with respect to a parameter is established under mild assumptions on the regularity of coefficients for Sobolev solutions of the Poisson equations in the whole $\mathbb{R}^d$ *in the "ergodic case." An assertion of this kind serves as one of the key tools in diffusion approximation and some other limit theorems. Bibliography*: 12 *titles.*

1 Introduction

Let us consider the Poisson equation in $\mathbb{R}^d$, $d \geqslant 1$,

$$L(x,y)u(x,y) = -f(x,y), \quad x \in \mathbb{R}^d, \tag{1.1}$$

where $y \in \mathbb{R}^\ell$ ($\ell \geqslant 1$) is a parameter and

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

with $a = \sigma\sigma^*/2$. Such equations *with a parameter and in the whole space* $\mathbb{R}^d$ – not in a bounded domain – are important for functional limit theorems in probability. To be well defined, this equation requires some sort of boundary condition; the role of the latter takes a behavior of solution at infinity which will be discussed shortly. Under such condition about growth, the solution turns out to be defined up to an additive constant, which is also quite natural due to $L1 \equiv 0$. To fix this constant, it is convenient to use the "centering" condition (cf., for example, [1])

$$\int\limits_{\mathbb{R}^d} u(x,y)\,\mu_\infty^y(dx) = 0. \tag{1.2}$$

Translated from *Problems in Mathematical Analysis* **61**, October 2011, pp. 43–68.

1072-3374/11/1791-0048

Here, $\mu^y_\infty(dx)$ is a (unique) invariant measure of the Markov process with a generator $L(\cdot, y)$ given y. In order to guarantee the *existence* of solution, a necessary condition is again "centering" for the right-hand side f,

$$\int_{\mathbb{R}^d} f(x,y)\, \mu^y_\infty(dx) = 0, \tag{1.3}$$

for each $y \in \mathbb{R}^\ell$ (cf., for example, [2]). The meaning of the condition (1.3) is analogous to the centering in the standard Central Limit Theorem.

The problem addressed in this paper is the smoothness of the solution u of the Poisson equation (1.1) with respect to the parameter y, which would suffice for application of the Itô formula with some diffusions plugged in at both variables, $u(X_t, Y_t)$, say, where Y_t is another diffusion. This is important in diffusion approximation and other limit theorems (cf. [3, 4, 1]). Note that the most general Poisson equations which arise in relation to diffusion approximation do not admit "potential term" of zero order, even though such equations may be also of a certain theoretical interest. Hence we would like to have a simple sufficient condition for two derivatives – either classical or, at least, Sobolev – with respect to y and, of course, in x as well. However, derivatives with respect to x are not a problem because we have them for free by virtue of [2] (cf. the reminder in Proposition 2 below). In the paper [1], Equation (1.1) was investigated in the Hölder classes of functions via fundamental solutions, and all derivatives with respect to the parameter and the state variable were classical. The approach suggested in that paper required a certain regularity (at least, C^2) of coefficients *with respect to the state variable* x, which may be really desirable to relax. In this paper, we present another idea which provides similar smoothness results in y (i.e., two derivatives in y) under quite different set of assumptions. Namely, instead of the regularity of coefficients a, b we assume here the regularity of the right-hand side f. This change allows us to tackle a much wider class of operators L, although the class of right-hand sides becomes more narrow. In this respect note that, in the discrete time theory, there is a series of results in Lipschitz classes, both for coefficients and right-hand sides (cf. [5]). Hence it seems as if the "total regularity" of coefficients and of the right-hand side in all known cases equals two, at least, informally (i.e., if we accept to assign regularity one to the Lipschitz condition). So, it may be said that in this paper we relax conditions on regularity with respect to x on the expense of regularity of the right-hand side f.

Note that under our assumptions derivatives of solution u with respect to the parameter turn out to be classical. Another remark is that the assumptions on the right-hand side ((H_f), (H'_f), and (H''_f) below) may be relaxed to similar growth of L_p norms of f and its derivatives and the latter may be understood also in the Sobolev sense.

Unlike [1], in the present paper we do not consider transition densities, nor fundamental solutions. However, the strategy implemented in [1] will be used essentially; just, now we will be working with measures rather than with densities. Some new technicalities and difficulties arise here, while some old ones become easier or even disappear, – such as, e.g., singularities at time zero. All assumptions will be formulated in the next section.

In the proofs below, there will be a repeating reasoning about series with a general term that consists of an exponentially decreasing value and another polynomially increasing (in x or in R) multiplier. This polynomially increasing term may look as an extension which complicates the calculus. However, in fact, we, apparently, cannot avoid this increasing term, broadly speaking, because solution of an elliptic partial differential equation with a bounded right-hand side is, generally speaking unbounded. In other words, those series below with power functions in R are natural and seem to be more or less unavoidable.

In order to tackle the problem (1.1), (1.2) in appropriate functional classes, we will need to deal with the corresponding auxiliary SDE,

$$dX_t^y = b(X_t^y, y)dt + \sigma(X_t^y, y)dB_t, \qquad t \geqslant 0, \qquad X_0^y = X_0, \tag{1.4}$$

with some (nonrandom) initial data X_0, where X_t^y takes values in $\mathbb{R}^d$, and $\{B_t;\, t \geqslant 0\}$ is a standard d–dimensional Brownian motion $\sigma(x,y) = \sqrt{2a(x,y)}$.

In the author's opinion, it would be desirable to simplify the calculus below, which could possibly allow to cover more general models. One of possible ways to that might be to use some ideas from another forthcoming paper [6] based on the notion of extended Poisson equation and weaker metrics. We postpone this task till further studies.

The paper consists of Introduction and three further sections: assumptions and the main result are contained in Section 2, auxiliaries in Section 3, and the main technical part about derivatives of the heat functions – solutions of some parabolic equations – in Section 4, the latter being split itself into several subsections related to various derivatives. We deal with functions, vector-valued functions and matrix-valued functions. However, for the notational simplicity all Euclidean norms are denoted by mean of the same modulus. The crucial point is the derivative representation (2.3) of Theorem 2: we *guess* this formula and then prove that the right-hand side there serves as a derivative $\partial_y p(x,g,y)$. It is possible to study further smoothness under further natural additional assumptions on the coefficients, but this is not the goal of this paper.

In the calculus below, all constants C and m may change from line to line.

2 Assumptions and the Main Result

Both coefficients a and b are assumed to be *bounded* (although this, of course, may be considerably relaxed); a is uniformly continuous with respect to x and nondegenerate uniformly with respect to (x,y), i.e. there exist two constants $0 < \lambda \leqslant \Lambda < \infty$ such that the following matrix inequalities hold (in the sense of nonnegative definiteness),

(H_a)
$$\lambda I \leqslant a(x,y) \leqslant \Lambda I,$$

The existence of an invariant probability measure μ_∞^y is ensured by the following recurrence condition:

(H_b)
$$\lim_{|x|\to\infty} \sup_y \langle b(x,y), x\rangle = -\infty$$

(cf., for example, [7]. The assumptions (H_a) and (H_b) also ensure uniqueness of the invariant measure, as well as uniqueness of solution of (1.4) in law and strong Markov property (cf. [8] about the latter). For the function f, we assume that there exist $C, m_0 > 0$ such that

(H_f)
$$\sup_y |f(x,y)| \leqslant C(1 + |x|^{m_0}), \qquad \text{in addition to (1.3).}$$

Throughout the paper, we assume that two derivatives of f with respect to x have moderate growth: with some m_0,

(H_f')
$$\sup_y |f_x(x,y)| \leqslant C(1 + |x|^{m_0}),$$

and

$$(H''_f) \qquad \sup_y |f_{xx}(x,y)| \leqslant C(1+|x|^{m_0}).$$

By (H_g), (H'_g), and (H''_g) we understand the same conditions for a generic function $(g(x), x \in \mathbb{R}^d)$, which does not depend on y (but *no centering*).

The family of regularity assumptions on all coefficients is as follows: we say that assumption $(H^{0,j})$ holds with some integer $j \geqslant 0$ if there exists m_0 such that

$$(H^{0,j}) \qquad \begin{aligned} &(a,b)(x,\cdot) \in C^j_b, \quad f(x,\cdot) \in C^j, \\ &\sup_y \sum_{|k|\leqslant j} |\partial^k_y f(x,y)| \leqslant C(1+|x|^{m_0}). \end{aligned}$$

Without the lower index b in C, no boundedness is assumed: instead, moderate (as polynomial) growth will be used in all such cases. $(a,b)(x,\cdot) \in C^0_b$ means bounded continuity in the variable y, which is uniform with respect to x, i.e., with some modulus of continuity in y only. In the sequel, only $(H^{0,1})$ and $(H^{0,2})$ will be used in Theorems 1–5.

Let us denote

$$p_t(x, g(\cdot); y) := \langle g(\cdot), \mu^y_{x,t} - \mu^y_\infty \rangle \equiv E_x g(X_t) - \int g(x')\, \mu^y_\infty(dx') \tag{2.1}$$

assuming that all the integrals here exist. The same notation will be also applied in case of *centered* functions depending on y, $(f(x,y),\ x \in \mathbb{R}^d, y \in \mathbb{R}^\ell)$,

$$p_t(x, f; y) := \langle f(\cdot, y), \mu^y_{x,t} - \mu^y_\infty \rangle = \langle f(\cdot, y), \mu^y_{x,t} \rangle. \tag{2.2}$$

Further, let $L_{i,2} := C^i_2\, \partial^{2-i} L / \partial y^{2-i}$. By induction introduce the vector-valued functions

$$p^{(1)}_t(x, g; y) \equiv q_t(x, g; y) := \int_0^t E_x \frac{\partial L}{\partial y}(X^y_s, y) p_{t-s}(X^y_s, g; y)\, ds, \tag{2.3}$$

$$q_t(x, f; y) := q_t(x, f(\cdot, y^*); y)|_{y^*=y} = \int_0^t E_x \frac{\partial L}{\partial y}(X^y_s, y) p_{t-s}(X^y_s, f; y)\, ds \tag{2.4}$$

and the matrix-valued functions

$$\begin{aligned} p^{(2)}_t(x, g; y) \equiv q^{(1)}_t(x, g; y) := &\int_0^t ds\, p_s\left(x, L_{0,2}(\cdot) p_{t-s}(\cdot, g)\right) \\ &+ \int_0^t ds \sum_{|i|=1} q_s\left(x, L_{i,2}(\cdot) p_{t-s}(\cdot, g)\right), \end{aligned} \tag{2.5}$$

assuming that all the integrals are well defined.

Let us denote (assuming all expressions exist)

$$p^{(1)}_t(x, f, y) := q_t(x, f(\cdot, y^*), y)|_{y^*=y} + \langle f_y(\cdot, y), \mu^y_{x,t} - \mu^y_\infty \rangle, \tag{2.6}$$

$$q^{(1)}_t(x, f(\cdot, y), y) := q^{(1)}_t(x, f(\cdot, y^*), y)|_{y^*=y}, \tag{2.7}$$

and

$$p_t^{(2)}(x,f,y) := q_t^{(1)}(x,f(\cdot,y),y) + 2q_t(x,f_y(\cdot,y),y) + \langle f_{yy}(\cdot,y), \mu^y_{x,t} - \mu^y_\infty\rangle. \tag{2.8}$$

Here, $q_t(x,f(\cdot,y),y)$ is treated so as if $f(\cdot,y^*)$ did not depend on y and then a substitution $y^* = y$ is performed and $q_t^{(1)}(x,f(\cdot,y),y)$ is understood similarly. The existence of $p_t^{(1)}(x,g,y)$ and $p_t^{(2)}(x,g,y)$ (and, hence, of $p_t^{(1)}(x,f,y)$ and $p_t^{(2)}(x,f,y)$) will be established below under the corresponding assumptions in Subsections 4.1 and 4.4 respectively. In the calculus, the parameter y may be occasionally dropped for the sake of simplicity. If g is a vector-valued function or a matrix-valued function, then any $p_t^{(i)}(x,g,y)$ will be understood respectively (with $p_t^{(0)}(x,g,y) := p_t(x,g,y)$), i.e., as a vector-valued function or a matrix-valued function of the corresponding dimension. We emphasize that fundamental solutions are *not* used in the paper and notation like $p_t(x,g,y)$ mean some integration with respect to the distribution of X_t.

Theorem 1. *Let assumptions (H_b), (H_a), (H_f), (H'_f), (H''_f), and $(H^{0,2})$ be satisfied. Then the unique centered solution u in the class $\bigcap_{p>1} W^2_{p,\mathrm{loc}} \bigcap C$ possesses two continuous derivatives with respect to y with the representations*

$$u_y(x,y) = \int_0^\infty p_t^{(1)}(x,f;y)\,ds, \tag{2.9}$$

and

$$u_{yy}(x,y) = \int_0^\infty p_t^{(2)}(x,f;y)\,ds. \tag{2.10}$$

All the expressions and integrals in (2.9) and (2.10) are well defined and u_y and u_{yy} are moderately growing in x functions uniformly with respect to y.

Proof. By Proposition 2 below, the unique centered solution u of Equation (1.1) in the class $\bigcap_{p>1} W^2_{p,\mathrm{loc}} \bigcap C$, has a representation,

$$u(x,y) = \int_0^\infty E_x f(X_s^y, y)\,ds \equiv \int_0^\infty p_t(x,f(\cdot,y);y)\,ds. \tag{2.11}$$

Note that convergence of all integrals in (2.9) and (2.10) and their continuity follow from the Proposition 1 and Theorems 2, 3 and 5 below, in particular, from the bounds (3.1), (4.46) and (2.5). The fact that both expressions for u_y and u_{yy}, indeed, represent the corresponding full derivatives of the function u, is a consequence of convergence rate bounds and from standard theorems of Analysis about differentiability under the integral. Polynomial growth of both derivatives in x is proved in Theorems 2 and 5 below. Theorem 1 is proved. □

Remark 1. For the first derivative u_y assumption $(H^{0,1})$ suffices instead of $(H^{0,2})$. More generally, assumption $(H^{0,j})$ $(j \geqslant 3)$ would imply j derivatives with respect to y, which can be proved by induction. Some of the conclusions might be interesting to extend under certain growth conditions on b and a (or σ). We do not pursue either of these goals here.

3 Auxiliaries

Before we proceed further, let us remind some previous results related to our problem. The first estimate is a special case of mixing bounds established in [7]. Proposition 1 is the basis for convergence of all integrals in (2.9), (2.10) and (2.11). Here, $\|\cdot\|_{TV}$ is the metric of total variation.

Proposition 1 ([7])**.** *Under assumptions* (H_a) *and* (H_b)*, the following convergence bounds hold for the process* X_t^y*,* $t \geqslant 0$*, uniformly with respect to* y*: for every* $k > 0$*, there exist* $C, m > 0$ *such that for any* $t \geqslant 0$*,*

$$\sup_y \|\mu_{x,t}^y - \mu_\infty^y\|_{TV} \leqslant C\frac{(1+|x|^m)}{(1+t^k)}. \tag{3.1}$$

Moreover, for every $k > 0$ *and* $m > 0$*, there exist* $C, m' > 0$ *such that*

$$\sup_y \sup_{0\leqslant t<+\infty} \int |x'|^m \mu_{x,t}^y(dx') \leqslant C\frac{1+|x|^{m'}}{1+t^k}, \tag{3.2}$$

and for every $m > 0$*,*

$$\sup_y \int |x'|^m \mu_\infty^y(dx') < \infty. \tag{3.3}$$

Moreover, for any $k > 0$ *and* $n > 0$*, and every function* g *satisfying* $|g(x)| \leqslant C(1+|x|^n)$*, there exist* $C, m > 0$ *such that*

$$\sup_y \|\langle g, \mu_{x,t}^y - \mu_\infty^y\rangle\|_B \leqslant C\frac{(1+|x|^m)}{1+t^k}. \tag{3.4}$$

The following result from [2] concerns the existence, uniqueness, and some estimates of solutions of Equation (1.1), with a minor adjustment to our present setting, which is a bit less general than in [2].

Proposition 2 ([2])**.** *Under assumptions* (H_a) *and* (H_b)*, the uniform continuity of the matrix* a *and growth of* $f(x,y)$ *in* x *not faster than polynomially for any* y*, there exists a solution of Equation* (1.1) *in the class of functions from the Sobolev space* $\bigcap_{p>1} W^2_{p,\mathrm{loc}} \bigcap C$ *which are locally bounded and grow at most polynomially in* $|x|$*, as* $|x| \to \infty$*, unique up to an additive constant which can be chosen so that for any* y *the centering equality (1.2) holds.*

Moreover, for this solution the representation (2.11) (*cf.* above) *holds with the following bounds*:

- *If for some* $\beta \geqslant 0$

  $$|f(x,y)| \leqslant C(y)(1+|x|^\beta),$$

  then for any $\beta' > \beta + 2$

  $$(|u(x,y)| + |\nabla_x u(x,y)|) \leqslant C_1(y)(1+|x|)^{\beta'} \tag{3.5}$$

  with some $C_1(y)$*.*

- *If for some* $\beta < 0$,
$$|f(x,y)| \leqslant C(y)(1+|x|)^{\beta-2},$$
then solution u *and its gradient* $\nabla_x u$ *are both bounded; moreover,*
$$|u(x,y)|(1+|x|)^{-\beta+2} + |\nabla_x u(x,y)| \leqslant C_1(y) \tag{3.6}$$
with some $C_1(y)$.

- *If for some* $\beta > 4$,
$$|f(x,y)| \leqslant C(y)(1+|x|^{\beta-2}),$$
then for some constant $C_1(y)$
$$|u(x,y)| + |\nabla_x u(x,y)| \leqslant C_1(y)(1+|x|^{\beta}). \tag{3.7}$$

Remark 2. The gradient $\nabla_x u$ is continuous due to the embedding theorem, see [9]. The boundedness provided by (3.6) in some cases is a very desirable property, even though often a moderate growth could be also sufficient. The representation (2.11) is a natural extension of many earlier results, in particular, we refer to [10] (for bounded domains) and [5] (for discrete equations in the whole space $\mathbb{R}^d$).

4 Derivatives of Heat Functions

4.1 First derivatives $\partial_y p_t(x,g;y)$ and $D_y p_t(x,f;y)$

Remind the notation:
$$p_t(x,g;y) := \langle g(\cdot), \mu^y_{x,t} - \mu^y_\infty \rangle$$
and
$$p_t(x,f;y) := \langle f(\cdot,y), \mu^y_{x,t} - \mu^y_\infty \rangle.$$
Under assumptions (H_a) and (H_b), this is a well defined operation, at least, for every g or f with a moderate growth with respect to x, due to Proposition 1. Occasionally, it will be convenient to use an equivalent semigroup notation for the same object,
$$p_t(x,g;y) \equiv \langle g, \mu^y_{x,t} \rangle - \langle g, \mu^y_\infty \rangle \equiv T_t(g - \langle g, \mu^y_\infty \rangle),$$
where T_t is a semigroup of linear operators with a generator L on the space of Borel functions g with a moderate growth, with a weighted norm
$$\|g\|_m = \operatorname*{ess\,sup}_x (1+|x|^m)^{-1}|g(x)| < \infty.$$
Note that the functions $p_t(x,g;y)$ and $p_t(x,g;y)$ are continuous in y, for example, due to the probabilistic representation via the SDE (1.4), continuity of solution of this SDE in y in mean, moment inequalities and the Lebesgue dominated convergence theorem. We now establish the differentiability of $p_t(x,g;y)$ and $p_t(x,f;y)$ with respect to y; $D_y p$, as usual, signifies the full derivative function (as f also depends on y), while ∂_y (or, occasionally, ∇_y) – the partial derivative with respect to the third variable. For the purpose of growth control, let us introduce one

more notation: we say that the function $\varphi(x)$, $x \in \mathbb{R}^d$, belongs to the space $L^{\mathrm{mod}}_{p,\mathrm{loc}}$ (for moderate growth) if and only if there exists $m \geqslant 0$ such that

$$\sup_x (1+|x|)^{-m} \Bigg(\int\limits_{v:\,|v-x|\leqslant 1} |\varphi(v)|^p \, dv \Bigg)^{1/p} < \infty. \tag{4.1}$$

Otherwise, this space may be defined as a family of functions for which there exists m such that

$$\sup_x (1+R)^{-m} \Bigg(\int\limits_{v:\,|v|\leqslant R} |\varphi(v)|^p \, dv \Bigg)^{1/p} < \infty. \tag{4.2}$$

Similarly, the notation $u \in W^{1,2,mod}_{p,\mathrm{loc}}$ means that $u, u_t, u_x, u_{xx} \in L^{\mathrm{mod}}_{p,\mathrm{loc}}$. Either expression (4.1) or (4.2) may be chosen as a norm of the space $L^{\mathrm{mod}}_{p,\mathrm{loc}}$; we will call the elements of this space *functions of L_p with a moderate growth.* Note that under (H_f), solution given in Proposition 2 belongs to $\bigcap_{p>1} W^{1,2,mod}_{p,\mathrm{loc}}$, by virtue of *a priori* bounds (cf. [9, Lemma 2.3.3]).

In the following calculus, we treat expressions like $\partial_y p$ or $D_y p$ as if they were functions, for simplicity of notation. For second derivatives (in the next section), likewise, $\partial^2_y p$ is a matrix-valued function, but again we will use simplified notation as it were a function with values in R^1. This simplification is not harmful; in particular, in a bit different manner, it could be understood as derivatives with respect to some specific variables, e.g., $\partial_{y_k} p$, etc. Notations (H_g) and (H'_g) below signifies assumptions (H_f) and (H'_f) applied to a function g, which does not depend on y, i.e., a moderate growth of g and ∇g respectively. It is methodically convenient to prove all estimates firstly for $p_t(x, g; y)$ and then expand them to more general $p_t(x, f; y)$. Respectively, all further theorems are split into two parts.

Theorem 2. *Suppose that (H_b), (H_a), (H_g), (H'_g), (H''_g), and $(H^{0,1})$ are satisfied and a function g has an L_p moderate growth in the sense of (4.2) (say) with some $p \geqslant d+1$. Then for each $t > 0$, $x, x' \in \mathbb{R}^d$, $y \in \mathbb{R}^\ell$, $p_t(x, g; y)$ is continuously differentiable in the variable y, and the gradient $\partial_y p_t(x, g; y)$ is given by formula (2.3),*

$$\begin{aligned} \partial_y p_t(x, g, y) &= q_t(x, g; y) \;\; (\equiv p^{(1)}_t(x, g; y)) \\ &:= \int_0^t E_x \frac{\partial L}{\partial y}(X^y_s, y) p_{t-s}(X^y_s, g; y)\, ds. \end{aligned}$$

The vector-valued function $q_t(\cdot, g; \cdot)$ is continuous in y for any t, x, and it has a moderate growth in x uniformly with respect to y and locally uniformly with respect to t.

Similarly, if (H_b), (H_a), (H_f), (H'_f), (H''_f) and $(H^{0,1})$ are satisfied, then the following representation holds:

$$\begin{aligned} D_y p_t(x, f, y) &= q_t(x, f(\cdot, y); y) + E_x f_y(X^y_t, y) \\ &\equiv q_t(x, f(\cdot, y); y) + p_t(x, f_y(\cdot, y); y) \\ &\equiv q_t(x, f(\cdot, y); y) + \langle f_y(\cdot, y), \mu^y_{x,t} - \mu^y_\infty \rangle, \end{aligned} \tag{4.3}$$

and the vector-valued function $q_t(\cdot, f; \cdot)$ is continuous in y for any t, x and has a moderate growth in x uniformly with respect to y and locally uniformly with respect to t.

Remark 3. In the proof of this theorem below, $t > 0$ is fixed and some constants in the calculus *may* depend on this value of t. Formula (4.3) itself is a consequence of an "educated guess" due to a formal differentiation with respect to y of the equation on p_t, i.e., $\partial_t p_t - L(y)p_t = 0,\ p_0 = g(x)$. This should formally result in

$$\partial_t(\partial_y p_t) - L(y)(\partial_y p_t) = (\partial_y L(y))p_t, \quad \partial_y p_0 = 0,$$

or

$$\partial_t(q_t) - L(y)(q_t) = (\partial_y L(y))p_t, \quad q_0 = 0. \tag{4.4}$$

It is known that solution of the latter equation is given by (2.3).

Proof. 1. In coordinates, Equation (2.3) reads, for any i,

$$\partial_{y^i} p_t(x, g, y) = \int\limits_0^t E_x \frac{\partial L}{\partial y^i}(X_s^y, y) p_{t-s}(X_s^y, g; y)\, ds.$$

All formulas in the following proof may be understood in this way in coordinates.

For any function $g(x)$, $x \in \mathbb{R}^d$, we introduce the notation denote

$$\begin{aligned} &\widehat{g}(x, y) := g(x) - \langle g, \mu_\infty^y \rangle, \\ &w^y(s, x) := E_x \widehat{g}(X_{t-s}^y, y) \equiv p_{t-s}(x, g; y). \end{aligned} \tag{4.5}$$

For given t and y and the function g we set

$$\begin{aligned} v(s, x; y) &:= \left(\frac{\partial L}{\partial y} p_{t-s}\right)(x, g; y) \\ &= \Big(\sum_{ij} a_y^{ij}(x, y)\partial^2_{x^i x^j} + \sum_i b_y^i(x, y)\partial_{x^i}\Big) E_x \widehat{g}(X_{t-s}^y, y). \end{aligned} \tag{4.6}$$

(Remind that here t is fixed; below we will use another notation for this function with any t, $f_{t-s}^{(1)}$.) By virtue of (H_f) and *a priori* bounds (cf. [9, Chapter 4, Section 10] and [11, Theorem 2])

$$\begin{aligned} \|E_x \widehat{g}(X_{t-s}^y, y)\|_{W_p^{1,2}([0,t]\times B_R)} &\equiv \|w^y\|_{W_p^{1,2}([0,t]\times B_R)} \\ &\leqslant C(\|w^y\|_{L_p([0,t]\times B_{R+1})} + \|\widehat{g}^y\|_{W_p^2(B_{R+1})}) \end{aligned}$$

and because

$$\|\widehat{g}^y\|_{W_p^2(B_{R+1})} \leqslant C(1+R)^m$$

(straightforward from the assumptions and integration if we take into account that the invariant measure μ_∞^y integrates any power function $|x|^k$; cf. Proposition 1) and

$$\|w^y\|_{L_p([0,t]\times B_{R+1})} \leqslant C(1+R)^m,$$

(straightforward), we conclude that with some $m > 0$,

$$\|E_x \widehat{g}(X_{t-s}^y, y)\|_{W_p^{1,2}([0,t]\times B_R)} \leqslant C(1+R)^m,$$

and, consequently,

$$p_{t-s}(x, g; y) = E_x \widehat{g}(X_{t-s}^y, y) \in W_{p,\mathrm{loc}}^{1,2,\mathrm{mod}} \qquad \forall\ p > 1, \tag{4.7}$$

as a function of the variables $(s,x)\in[0,t]\times\mathbb{R}^d$ (see also [11]). Since functions a_y and b_y are bounded, we conclude that for every y,

$$v(s,x;y)\in L^{\mathrm{mod}}_{p,\mathrm{loc}} \qquad \forall\ p>1, \tag{4.8}$$

i.e., more precisely,

$$\int\limits_0^t ds \int\limits_{B_R} |v(s,x;y)|^p\,dx \leqslant C(1+R)^m, \tag{4.9}$$

where C may depend on t. Therefore, using Krylov's estimate [12, Theorem 2.2.4] and Proposition 1, it can be shown that for any $t>0$, the function

$$\begin{aligned} q_t(x,g;y) &:= \int\limits_0^t E_x\left(\frac{\partial L}{\partial y}(X_s,y)p_{t-s}(X_s,g;y)\right)ds \\ &\equiv \int\limits_0^t E_x v(s,X_s;y)\,ds \end{aligned} \tag{4.10}$$

is well defined, bounded and continuous in y and has a moderate growth in x. For completeness of presentation, we provide the details.

We use two values

$$p=d+1 \quad \text{and} \quad p'=2(d+1)$$

(here any $p'>d+1$ could be used). Krylov's estimate ensures that

$$E_x \int\limits_0^{t\wedge\tau_R} v(s,X_s;y)\,ds<\infty \qquad \forall\ R>0,$$

where

$$\tau_R:=\inf(t\geqslant 0:\ |X_t|\geqslant R).$$

Moreover,

$$E_x \int\limits_0^{t\wedge\tau_R} v(s,X_s;y)\,ds\leqslant N\|v\|_{L_{d+1}([0,t]\times B_R)}.$$

By (4.8), we have

$$\|v\|_{L_{d+1}([0,t]\times B_R)}\leqslant C\Bigg(\int\limits_0^t ds\int\limits_{B_R}(1+|x|^m)^{d+1}\,dx\Bigg)^{\frac{1}{d+1}}\leqslant Ct^{\frac{1}{d+1}}(1+R)^{md} \tag{4.11}$$

and

$$\|v\|_{L_{2(d+1)}([0,t]\times B_R)}\leqslant C\Bigg(\int\limits_0^t ds\int\limits_{B_R}(1+|x|^m)^{2(d+1)}\,dx\Bigg)^{\frac{1}{2(d+1)}}\leqslant Ct^{\frac{1}{2(d+1)}}(1+R)^{md}. \tag{4.12}$$

Let $R \geqslant |x|$. Consider a sequence of hitting times,

$$\tau_0 = 0, \quad \tau^1 := \tau_R, \quad \tau^{n+1} := \inf(t \geqslant \tau^n : \; |X_t - X_{\tau^n}| \geqslant R), \quad \inf(\varnothing) = \infty.$$

Then, using the Cauchy–Bunyakowsky–Schwarz inequality, we estimate,

$$\begin{aligned}
E_x \int_0^t |v|(s, X_s; y)\, ds &= E_x \sum_{k=0}^{\infty} 1(\tau^k < t \leqslant \tau^{k+1}) \int_0^t |v|(s, X_s; y)\, ds \\
&\leqslant E_x \sum_{k=0}^{\infty} 1(\tau^k < t \leqslant \tau^{k+1}) \int_0^{\tau^{k+1}\wedge t} |v|(s, X_s; y)\, ds \\
&= \sum_{k=0}^{\infty} E_x\, 1(\tau^k < t \leqslant \tau^{k+1}) \int_0^{\tau^{k+1}\wedge t} |v|(s, X_s; y)\, ds \\
&\leqslant \sum_{k=0}^{\infty} \Big(E_x\, 1(\tau^k < t \leqslant \tau^{k+1})\Big)^{1/2} \left(E_x \left(\int_0^{\tau^{k+1}\wedge t} |v|(s, X_s; y)\, ds \right)^2 \right)^{1/2} \\
&\leqslant \sum_{k=0}^{\infty} \Big(E_x\, 1(\tau^k < t \leqslant \tau^{k+1})\Big)^{1/2} \left(E_x \left(\int_0^{\tau_{(k+1)R}\wedge t} |v|(s, X_s; y)\, ds \right)^2 \right)^{1/2}.
\end{aligned}$$

Now,

$$E_x\, 1(\tau^{k-1} < t \leqslant \tau^k) \leqslant P_x(\tau^{k-1} < t).$$

By the strong Markov property and induction,

$$\begin{aligned}
&\sup_x P_x(\tau^n < t) \leqslant q_t^n, \\
&q_t := \sup_x P_x(\sup_{0\leqslant s\leqslant t} |X_s - x| \geqslant R) < 1.
\end{aligned} \tag{4.13}$$

Next, by (4.12),

$$\begin{aligned}
E_x \left(\int_0^{t\wedge\tau_{(k+1)R}} |v|(s, X_s; y)\, ds \right)^2 &\leqslant E_x \left(t \int_0^{\tau_{(k+1)R}} |v|^2(s, X_s; y)\, ds \right) \\
&\leqslant C\, t\, \|v\|^2_{L_{2(d+1)}([0,t]\times B_{(k+1)R})} \leqslant C\, t^{1+\frac{1}{d+1}}\, (1+R)^{2md}(k+1)^{2md}.
\end{aligned}$$

Now we can complete the estimate above as follows:

$$\begin{aligned}
E_x \int_0^t |v|(s, X_s; y)\, ds &\leqslant \sum_{k=0}^{\infty} \Big(E_x\, 1(\tau^k < t \leqslant \tau^{k+1})\Big)^{1/2} \left(E_x \left(\int_0^{\tau_{(k+1)R}} |v|(s, X_s; y)\, ds \right)^2 \right)^{1/2} \\
&\leqslant \sum_{k=0}^{\infty} C q_t^{(k-1)/2} t^{\frac{d+2}{2(d+1)}} (1+R)^{md}(k+1)^{md} \leqslant C(1+R)^{md}.
\end{aligned}$$

This implies that, indeed, for any $t > 0$ the function $q_t(\cdot, g; y)$ in (4.10) is well defined and (locally uniformly in x and uniformly in y) bounded. Remind that R was chosen so as to satisfy the only condition $|x| \leqslant R$. Hence, indeed, the function $q_t(\cdot, g; y)$ defined in (4.10) has a moderate growth in x.

2. Let us show the continuity of $q_t(x, g; \cdot)$ for every x and t. A natural idea is again to use Krylov's estimate with some $p \geqslant d + 1$ (cf. [12, Theorem 2.3.4])

$$
\begin{aligned}
&|q_t(x, g; y) - q_t(x, g; y')| \\
&\leqslant \int_0^t \left| E_x(v(s, X_s^y, y) - v(s, X_s^{y'}, y') \right| ds \\
&\leqslant \int_0^t \left| E_x(v(s, X_s^y, y) - v(s, X_s^y, y')) \right| ds + \int_0^t \left| E_x(v(s, X_s^y, y') - v(s, X_s^{y'}, y')) \right| ds \\
&\leqslant C\|v(y) - v(y')\|_{L_p([0,t]\times\mathbb{R}^d])} + \int_0^t \left| E_x(v(s, X_s^y, y') - v(s, X_s^{y'}, y')) \right| ds. \qquad (4.14)
\end{aligned}
$$

Here, the first term looks small if $|y' - y|$ is small and the second term, apparently, tends to zero as $|y' - y| \to 0$ because $E_x(v(s, X_s^y, y')$ ought to be a solution of some parabolic equation for which continuity in y is more or less a standard result in the theory of partial differential equations. Note, by the way, that despite the notation used above $E_x(v(s, X_s^y, y') - v(s, X_s^{y'}, y'))$, there is no need to deal with strong solutions, as the latter expression may be considered just as a short notation for $(E_x v(s, X_s^y, y') - E_x v(s, X_s^{y'}, y'))$.

Nevertheless, in this simple form the second inequality above is not very helpful, since the right-hand side in (4.14) may just diverge (as $v(y)$ may not be in L_p in the whole space, but only locally). However, because of the polynomial growth condition in x, we may localize the right-hand side and to tackle this localized version similarly to the previous step. So, let us establish an upper bound

$$
\int_0^t \int_{B_R} |v(s, x, y) - v(s, x, y')|^p \, dx ds \leqslant C\rho(|y' - y|)(1 + R)^m \qquad (4.15)
$$

with some bounded ρ such that $\rho(0) = 0$ and $\rho(s) \to 0$, $s \to 0$, and then apply an appropriate version of Krylov's estimate, from which the desired continuity of $q_t^g(x, \cdot)$ would follow.

3. Note that the function v given by (4.6) is a linear combination of the first and second derivatives of solution of some parabolic partial differential equation (cf. Equation (4.19) below). Hence, we may use *a priori* local bounds for Sobolev norms of such solutions. Remind that the function $\widehat{g}$ was defined in (4.5). Slightly abusing notation, we have

$$
\begin{aligned}
&v(s, x, y) - v(s, x, y') \\
&= (a_y(x, y)\partial_x^2 + b_y(x, y)\partial_x)E_x\widehat{g}(X_{t-s}^y, y) - (a_y(x, y')\partial_x^2 + b_y(x, y')\partial_x)E_x\widehat{g}(X_{t-s}^{y'}, y') \\
&= (a_y(x, y) - a_y(x, y'))\partial_x^2 E_x\widehat{g}(X_{t-s}^y, y) + (b_y(x, y) - b_y(x, y'))\partial_x E_x\widehat{g}(X_{t-s}^y, y) \\
&\quad + a_y(x, y')\partial_x^2(E_x\widehat{g}(X_{t-s}^y, y) - E_x\widehat{g}(X_{t-s}^{y'}, y')) + b_y(x, y')\partial_x(E_x\widehat{g}(X_{t-s}^y, y) - E_x\widehat{g}(X_{t-s}^{y'}, y')).
\end{aligned}
$$

So, using the notation (eventually, the modulus of continuity in (4.15) may differ)

$$\rho(z) := \sup_{(y,y'\colon |y-y'|\leqslant z)} \sup_x \left(|a_y(x,y) - a_y(x,y')| + |b_y(x,y) - b_y(x,y')|\right), \tag{4.16}$$

we estimate,

$$\begin{aligned} |v(s,x,y) - v(s,x,y')| \leqslant \rho(y-y')(|\partial_x^2 E_x\widehat{g}(X^y_{t-s})| + |\partial_x E_x\widehat{g}(X^y_{t-s})| \\ + C|\partial_x^2(E_x\widehat{g}(X^y_{t-s},y) - E_x\widehat{g}(X^{y'}_{t-s},y'))| \\ + C|\partial_x(E_x\widehat{g}(X^y_{t-s},y) - E_x\widehat{g}(X^{y'}_{t-s},y'))|. \end{aligned} \tag{4.17}$$

Thus,

$$\begin{aligned} \left(\int_0^t\int_{B_R} |v(s,x,y) - v(s,x,y')|^p\, dxds\right)^{1/p} \leqslant C\,\rho(y-y')\,\|E_x\widehat{g}(X^y_{t-s},y)\|_{W^{1,2}_p([0,t]\times B_R)} \\ + C\left(\int_0^t\int_{B_R} dsdx\, |\partial_x^2(E_x\widehat{g}(X^y_{t-s},y) - E_x\widehat{g}(X^{y'}_{t-s},y'))|^p\right)^{1/p} \\ + C\left(\int_0^t\int_{B_R} dsdx\, |\partial_x(E_x\widehat{g}(X^y_{t-s},y) - E_x\widehat{g}(X^{y'}_{t-s},y'))|^p\right)^{1/p}. \end{aligned} \tag{4.18}$$

Now we will use the fact that the function $w^y(s,x) := E_x\widehat{g}(X^y_{t-s},y) \equiv p_{t-s}(x,g;y)$ is a solution of the Cauchy problem in $[0,t]\times\mathbb{R}^d$,

$$w^y_s(s,x) + \frac{1}{2}\sum_{i,j} a^{ij}(s,x)w^y_{x^ix^j}(s,x) + \sum_i b^i(s,x)w^y_{x^i}(s,x) = 0,\ w^y(t,x) = \widehat{g}(x,y). \tag{4.19}$$

Therefore, the first term in the right-hand side of (4.18) contains a multiplier $\rho(y-y')$ – which tends to zero as $y'\to y$ – and another multiplier $\|E_x\widehat{g}(X^y_{t-s},y)\|_{W^{1,2}_p([0,t]\times B_R)}$, which admits a bound (4.7), which is the first part leading to the estimate (4.15). Namely, we get,

$$\begin{aligned} &\int_0^t\int_{B_R} |(a_y(x,y) - a_y(x,y'))\partial_x^2 E_x\widehat{g}(X^y_{t-s},y)|^p dsdx \\ &+ \int_0^t\int_{B_R} |(b_y(x,y) - b_y(x,y'))\partial_x E_x\widehat{g}(X^y_{t-s},y)|^p dsdx \leqslant C\rho(|y'-y|)(1+R)^m. \end{aligned}$$

4. The second part of this consideration leading to (31) is to establish the bound

$$\left(\int_0^t\int_{B_R} dsdx\, |\partial_x^2(E_x\widehat{g}(X^y_{t-s},y) - E_x\widehat{g}(X^{y'}_{t-s},y'))|^p\right)^{1/p} \leqslant C\widetilde{\rho}(|y'-y|)(1+R)^m, \tag{4.20}$$

possibly with some new modulus of continuity $\widetilde{\rho}(|y'-y|)$. We have,

$$\partial_x^2(E_x\widehat{g}(X^y_{t-s},y) - E_x\widehat{g}(X^{y'}_{t-s},y')) = \partial_x^2(w^y(s,x) - w^{y'}(s,x)).$$

Denote

$$z(s,x,y,y') := w^y(s,x) - w^{y'}(s,x).$$

Then, $z(\cdot,y,y')$ is a Sobolev solution (i.e, in $\cap_{p>1}W^{1,2}_{p,\mathrm{loc}}$) of the equation,

$$\begin{aligned} & z_s + \frac{1}{2}\sigma\sigma^*(s,x,y)z_{xx} + b(s,x,y)z_x = -F(s,x,y,y'), \\ & z(t,x,y,y') = \langle g, \mu^{y'}_\infty - \mu^y_\infty\rangle, \end{aligned} \tag{4.21}$$

where

$$F(s,x,y,y') = \frac{1}{2}(\sigma\sigma^*(s,x,y') - \sigma\sigma^*(s,x,y))\, u_{xx}(s,x,y') + (b(s,x,y') - b(s,x,y))\, u_x(s,x,y').$$

Remind that

$$\|\sigma\sigma^*(s,x,y') - \sigma\sigma^*(s,x,y)\| + |b(s,x,y') - b(s,x,y)| \leqslant \rho(|y-y'|).$$

By virtue of [9],

$$\begin{aligned} \|z(\cdot,y,y')\|_{W^{1,2}_p([0,t]\times B_R)} \leqslant C(& \|z(\cdot,y,y')\|_{L_p([0,t]\times B_{R+1})} \\ & + \|z(t,\cdot,y,y')\|_{W^2_p(B_{R+1})} + \|F(\cdot,y,y')\|_{L_p([0,t]\times B_{R+1})}). \end{aligned}$$

(We apply a slightly weaker bound with a stronger norm $\|\cdot\|_{W^2}$ instead of $\|\cdot\ \|_{W^{2-2/p}}$.) Here, for example, due to Krylov's estimates,

$$\|z(\cdot,y,y')\|_{L_\infty(B_R)} \leqslant N\|F(\cdot,y,y')\|_{L_p([0,t]\times B_{R+1})}.$$

Hence also

$$\|z(\cdot,y,y')\|_{L_p(B_R)} \leqslant N(1+R)^m\|F(\cdot,y,y')\|_{L_p([0,t]\times B_{R+1})}.$$

and

$$\|z(\cdot,y,y')\|_{W^{1,2}_p([0,t]\times B_R)} \leqslant C(1+R)^m(\|F(\cdot,y,y')\|_{L_p([0,t]\times B_{R+1})} + \|z(t,\cdot,y,y')\|_{W^2_p(B_{R+1})}).$$

Note that the function $z(t,\cdot,y,y')$ at t does not depend on x, so that

$$\|z(t,\cdot,y,y')\|_{W^2_p(B_{R+1})} = \|z(t,\cdot,y,y')\|_{L_p(B_{R+1})} \leqslant C(1+R)^{m/p}|\langle g, \mu^{y'}_\infty - \mu^y_\infty\rangle|.$$

Here,

$$\langle g, \mu^{y'}_\infty - \mu^y_\infty\rangle \to 0 \quad \text{as } y'-y \to 0,$$

i.e., again with some m and some modulus of continuity $\widetilde{\rho}$,

$$|\langle g, \mu^{y'}_\infty - \mu^y_\infty\rangle| \leqslant C\widetilde{\rho}(|y'-y|)(1+R)^m \tag{4.22}$$

by (3.4) and the locally uniform continuity of $\langle g, \mu^y_{x,t}\rangle$ in y; the latter may be shown similarly to the calculus above, but we postpone it till next step. Further, since

$$\|F(\cdot,y,y')\|_{L_p([0,t]\times B_{R+1})} \leqslant C\rho(|y'-y|)\|u(\cdot,y,y')\|_{W^{1,2}_p([0,t]\times B_R)} \leqslant C\rho(|y'-y|)(1+(R+1))^m,$$

we get the estimate (4.20), as required. Note that similarly and a bit easier the estimate with the first derivatives follows, with some m and some modulus of continuity $\widetilde{\rho}$,

$$\left(\int_0^t\int_{B_R} dsdx\, |\partial_x(E_x\widehat{g}(X^y_{t-s},y) - E_x\widehat{g}(X^{y'}_{t-s},y'))|^p\right)^{1/p} \leqslant C\widetilde{\rho}(|y'-y|)(1+R)^m.$$

Thus, (4.15) is established.

5. Now let us show a more reasonable analogue of (4.14), starting from the preliminary estimate,

$$|q_t(x,g;y) - q_t(x,g;y')| \leqslant \int_0^t \left|E_x(v(s,X^y_s,y) - v(s,X^y_s,y'))\right| ds + \int_0^t \left|E_x(v(s,X^y_s,y') - v(s,X^{y'}_s,y'))\right| ds. \tag{4.23}$$

Consider the first term on the right-hand side. By (4.15) and (4.13), we have

$$\begin{aligned}
&\int_0^t \left|E_x(v(s,X^y_s,y) - v(s,X^y_s,y'))\right| ds \\
&\leqslant \sum_n E_x 1(\tau_n \leqslant t < \tau_{n+1}) \int_0^t \left|(v(s,X^y_s,y) - v(s,X^y_s,y'))\right| ds \\
&\leqslant \sum_n E_x 1(\tau_n \leqslant t)\, 1(\sup_{s\leqslant t}|X^y_s| \leqslant (n+1)R) \int_0^t \left|(v(s,X^y_s,y) - v(s,X^y_s,y'))\right| ds \\
&\leqslant \sum_n q_t^{n/2}\left(E_x \int_0^{t\wedge\tau_{n+1}} \left|(v(s,X^y_s,y) - v(s,X^y_s,y'))\right| ds\right)^{1/2} \\
&\leqslant \sum_n q_t^{n/2} N\|(v(s,\cdot,y) - v(s,\cdot,y'))\|^{1/2}_{L_{d+1}([0,t]\times B_{(n+1)(R+1)})} \\
&\leqslant \sum_n q_t^{n/2} C_t(1+(n+1)(1+R))^m\rho(|y'-y|) \\
&\leqslant C(t,m)\,(1+R)^m\,\rho^{1/2}(|y'-y|)
\end{aligned}$$

with some $m > 0$.

6. Consider the second term in (4.23),

$$\int_0^t \left| E_x(v(s, X_s^y, y') - v(s, X_s^{y'}, y')) \right| ds$$

Denote

$$\widetilde{z}(s, x, y, y') := E_x(v(s, X_s^y, y') - v(s, X_s^{y'}, y')), \quad 0 \leqslant s \leqslant t,$$
$$z^1(s, x, y, y') := E_x(v(s, X_s^{y'}, y').$$

Then $\widetilde{z}$ is a Sobolev solution (i.e., in $W^{1,2}_{p,\mathrm{loc}}$) of the equation

$$\begin{aligned} &\partial_s \widetilde{z}(s, x, y, y') + L^y \widetilde{z}(s, x, y, y') = (L^y - L^{y'}) z^1(s, x, y, y'), \\ &\widetilde{z}(0, x, y, y') = 0. \end{aligned} \tag{4.24}$$

Note that z^1 has a moderate growth in $W^{1,2}_{p,\mathrm{loc}}$, which implies that the right-hand side in (4.24) has a moderate growth in $L_{p,\mathrm{loc}}$. Hence, arguing similarly to the step 4 above, we can see that

$$\|\widetilde{z}(\cdot, y, y')\|_{W^{1,2}_{p,\mathrm{loc}}([0,t]\times B_R)} \leqslant C_t\, \rho(|y' - y|)\, (1 + R)^m \tag{4.25}$$

with some $m > 0$. In particular, due to embedding theorems (cf., for example, [9]),

$$\|\widetilde{z}(\cdot, y, y')\|_{L_\infty([0,t]\times B_R)} \leqslant C_t\, \rho(|y' - y|)\, (1 + R)^m. \tag{4.26}$$

So, integrating with respect to s, we obtain,

$$\int_0^t \left| E_x(v(s, X_s^y, y') - v(s, X_s^{y'}, y')) \right| ds \leqslant C_t\, \rho(|y' - y|)\, (1 + R)^m. \tag{4.27}$$

Here, we may take any $R \geqslant |x|$. Hence, we get a desired estimate,

$$|q_t(x, g; y) - q_t(x, g; y')| \leqslant C\widetilde{\rho}(|y' - y|)(1 + |x|)^m. \tag{4.28}$$

7. Let us show that

$$q_t(x, g; y) = \partial_y p_t(x, g; y),$$

i.e., the function $q_t(x, g; y)$ is, indeed, derivative (gradient) of $p_t(x, g; y)$ with respect to y. We apply the hint from [1], used there in a slightly different setup. For $1 \leqslant i \leqslant \ell$, let e_i denote the unit vector in the ith direction of $\mathbb{R}^\ell$, and let $h \neq 0$. We define

$$q_t^{h,g}(x; y) := \frac{p_t(x, g; y + he_i) - p_t(x, g; y)}{h},$$

omitting the index i for brevity. Note that $q_0^{h,g} = 0$ and

$$
\begin{aligned}
\frac{\partial q_t^{h,g}}{\partial t} &= \frac{\partial_t p_t(x,g;y+he_i) - \partial_t p_t(x,g;y)}{h} \\
&= \frac{L(y+he_i)p_t(x,g;y+he_i) - L(y)p_t(x,g;y)}{h} \\
&= L(y)\frac{p_t(x,g;y+he_i) - p_t(x,g;y)}{h} \\
&\quad + \frac{L(y+he_i)p_t(x,g;y+he_i) - L(y)p_t(x,g;y+he_i)}{h} \\
&= L(y)q_t^{h,g} + \frac{L(y+h)-L(y)}{h} p_t(x,g,y+h). \qquad (4.29)
\end{aligned}
$$

So, treating $((L(y+h)-L(y))/h)\,p_t(x,g,y+h)$ as a right-hand side of the equation on $q_t^{h,g}$, we get the representation

$$
q_t^{h,g}(x;y) = \int_0^t ds\, E_x \left(\frac{L(y+h)-L(y)}{h} p_{t-s}(X_s^y, g; y+h) \right), \qquad (4.30)
$$

where we can pass to the limit as $h \to 0$ to get the desired assertion, due to the Lebesgue dominated convergence theorem. Indeed,

$$
\begin{aligned}
&\int_0^t ds\, E_x \left(\frac{L(y+h)-L(y)}{h} p_{t-s}(X_s^y, g; y+h) \right) \\
&= \int_0^t ds\, E_x \left(\frac{b(X_s^y, y+h) - b(X_s^y, y)}{h} D_x p_{t-s}(X_s^y, g; y+h) \right) \\
&\int_0^t ds\, E_x \left(\frac{a(X_s^y, y+h) - a(X_s^y, y)}{h} D_{xx} p_{t-s}(X_s^y, g; y+h) \right),
\end{aligned}
$$

and the applicability of the Lebesgue theorem is due to the condition (4.7) and Proposition 1. Hence

$$
\partial_{y^i} p_t(x,g;y) = \int_0^t E_x \left(\frac{\partial L}{\partial y^i}(X_s, y) p_{t-s}(X_s, g; y) \right) ds.
$$

Since the right-hand side here is continuous in y, the desired result follows.

8. Now consider $p_t(x,f,y)$ and show formula (4.3) and continuity in y. Due to (2.6),

$$
p_t^{(1)}(x,f,y) := q_t(x, f(\cdot,y), y) + \langle f_y(\cdot,y), \mu_{x,t}^y - \mu_\infty^y \rangle
$$

So, in comparison to $p_t(x,g,y)$, the only difference is a new term:

$$
E_x \partial_y (f(X_t^{y'}, y))|_{y'=y} \equiv p_t(x, f_y(\cdot,y); y) \equiv \langle f_y(\cdot,y), \mu_{x,t}^y - \mu_\infty^y \rangle. \qquad (4.31)
$$

First of all, this expression is well defined, i.e., the integral converges. Since f has a moderate growth in x, this follows straightforward from Proposition 1.

The continuity in y will be proved if we show that the function given by (4.31) is continuous in this variable, because (locally uniform) continuity of the first term in (4.3) is tackled by the first part of the theorem (about g). Continuity of the term $\langle f_y(\cdot, y), \mu^y_{x,t}\rangle$ follows from the same calculus, too. Hence, it remains to show that $\langle f_y(\cdot, y), \mu^y_\infty\rangle$ is continuous in y. But the latter follows from the former, from assumption $(H^{0,1})$ and from the estimate (3.4) of Proposition 1. Moderate growth of (4.31) also follows from the same estimates. Finally, (4.3) is a direct consequence of the differentiation rule for the composite function along with the first part of the theorem about g. Theorem 2 is proved. □

4.2 Behavior of $\partial_y p_t(x, g; y)$ and $D_y p_t(x, f; y)$ at $t \to \infty$

Starting from this section, we will be concerned with large time behavior. Hence, we ought to take care about all constants in the calculus, which should not depend on time. The base for that will be the bounds from [9] and [11], Krylov's estimate from [12] and the inequality (3.4), in all of which the constants do not depend on time. Recall that

$$q_t(x, g; y) = \int_0^t E_x \frac{\partial L}{\partial y}(X_s, y) p_{t-s}(X_s, g; y)\, ds.$$

We use the notation

$$f^1_t(x, g; y) := \frac{\partial L}{\partial y} p_t(x, g; y). \tag{4.32}$$

Note that, in fact, this is the same function as v from (4.6), but for the latter the value t was fixed. Now it is convenient to work with reversed time and we change notation. Let us establish some simple auxiliary bounds with any $p \geqslant d + 1$. Denote

$$h_t(x, y) := \langle f_y(x, y), \mu^y_{x,t} - \mu^y_\infty\rangle.$$

Lemma 1. *Under the assumptions of Theorem 1 with $(H^{0,1})$ instead of $(H^{0,2})$ – including both for g and f – for any k there exist C and m such that for any $t > 0$ the function $u(s, x) := p_s(x, g; y)$ satisfies*

$$\|u\|_{W^{1,2}_p([0,t]\times B_R)} \leqslant C(1 + |R|^m). \tag{4.33}$$

In particular, uniformly with respect to y

$$\|f^1_\cdot(\cdot; g, y)\|_{L_p([0,t]\times B_R)} \leqslant C(1 + |R|^m). \tag{4.34}$$

Also, for any $t_1 > 0$ and k there exist $m, K > 0$ such that for any $t > t_1$

$$\|u\|_{W^{1,2}_p([t_1,t]\times B_R)} \leqslant C\frac{(1 + |R|^m)}{1 + t_1^k}, \tag{4.35}$$

and, in particular,

$$\|f^1\|_{L_p([t_1,t]\times B_R)} \leqslant C\frac{(1 + |R|^m)}{1 + t_1^k}, \tag{4.36}$$

In addition, for any k there exist $m, C > 0$ such that for any x, $t \geqslant 0$,

$$|h_t(x, y)| \leqslant C\,\frac{1 + |x|^m}{1 + t^k}. \tag{4.37}$$

Proof. By [11, Theorem 2] and [9, Chapter 4, Section 10], we have

$$\|u\|_{W^{1,2}_p([0,t]\times B_R)} \leqslant C(\|u\|_{L_p([0,t]\times B_{R+1})} + \|g\|_{W^2_p(B_{R+1})}). \tag{4.38}$$

By the assumption on g, there exist C, m such that

$$\|g\|_{W^2_p(B_{R+1})} \leqslant C(1+|R|^m).$$

Further,

$$\|u\|_{L_\infty([0,t]\times B_R)} \leqslant C\big(1+\sup_{0\leqslant s\leqslant t} E_x|g(X_s)|\big) \leqslant C(1+|R|^m+t^m).$$

Whence, a similar bound holds – with some new C, m – for the norm $\|u\|_{L_p([0,t]\times B_{R+1})}$, too. So, by (4.38), we obtain the provisional estimates

$$\|u\|_{W^{1,2}_p([0,t]\times B_R)} \leqslant C(1+|R|^m+t^m) \tag{4.39}$$

and

$$\|f^1_\cdot(\cdot; g, y)\|_{L_p([0,t]\times B_R)} \leqslant C(1+|R|^m+t^m). \tag{4.40}$$

Now let us establish (4.35). Remind the estimate (3.4); we need similar bounds for $\partial_x p_t$ and $\partial^2_x p_t$. Remind *a priory* bounds (of which we show a particular case) for $0 < T_0 < T$,

$$\begin{aligned}\|u\|_{W^{1,2}_p([T_0,T]\times B_R)} \leqslant C\big(&\|u\|_{L_p([0,T]\times B_{R+1})}\\ &+\|(\partial_s - L)u\|_{L_p([0,T]\times B_{R+1})} + \|u(T_0,\cdot)\|_{W^2_p(B_{R+1})}\big)\end{aligned} \tag{4.41}$$

and also, for $0 < \delta < T_0 < T$,

$$\|u\|_{W^{1,2}_p([T_0,T]\times B_R)} \leqslant C\big(\|u\|_{L_p([T_0-\delta,T]\times B_{R+1})} + \|(\partial_s - L)u\|_{L_p([T_0-\delta,T]\times B_{R+1})}\big) \tag{4.42}$$

(cf. [11, Theorem 2] and [9, Chapter 4, Section 10]). In our case, $(\partial_s - L)u = 0$. By (3.4),

$$\|u\|_{L_p([t_1,t]\times B_R)} \leqslant C\frac{(1+|R|^{m'})}{1+t_1^k}.$$

By (4.42),

$$\|u\|_{W^{1,2}_p([t_1,t]\times B_R)} \leqslant C\frac{(1+|R|^{m'})}{1+t_1^k}.$$

So, we obtain (4.35) which implies (4.36) as well. Finally, combining (4.35) and (4.36) with (4.39) and (4.40), we get (4.33) and (4.34). The last inequality holds due to (3.4). Lemma 1 is proved. □

Our next task is to show that the following integral converges,

$$\int_0^\infty dt\,[q_t(x,g;y) - q_\infty(g,y)]\,,$$

where

$$q_\infty(g,y) = \lim_{t\to\infty} q_t(x,g;y) = \partial_y p_\infty(g,y), \tag{4.43}$$

and that we have a representation

$$q_\infty(g,y) = \int\limits_0^\infty \left(\int f_s^1(x'',g;y)\,\mu_\infty^y(dx'') \right) ds. \tag{4.44}$$

This representation may be considered as an "educated guess," given (2.3). So, let us *define* q_∞ by (4.44); then, of course, (4.43) is to be proved. Since, clearly, $\langle f_y(x,y), \mu_t^y - \mu_\infty^y \rangle \to 0$, $t \to \infty$, let us also set

$$p_\infty^{(1)}(f,y) := q_\infty(f(\cdot,y^*),y)|_{y^*=y} = \int\limits_0^\infty \left(\int f_s^1(x'',g;y)\,\mu_\infty^y(dx'') \right) ds. \tag{4.45}$$

Theorem 3. *Assume that* (H_b), (H_a), (H_g), (H_g'), (H_g''), *and* $(H^{0,1})$ *hold. Then for each* k *there exist* $C, m > 0$ *such that for all* $y \in \mathbb{R}^\ell$, $t \geqslant 1$,

$$|q_t(x,g;y) - q_\infty(g;y)| \leqslant C\frac{(1+|x|^m)}{(1+t)^k}, \tag{4.46}$$

$$|q_\infty(g,y)| \leqslant C; \tag{4.47}$$

moreover,

$$q_\infty(g,y) = \partial_y p_\infty(g,y). \tag{4.48}$$

Furthermore, if (H_b), (H_a), (H_f), (H_f'), (H_f''), *and* $(H^{0,1})$ *hold and also for each* k *there exist* $C, m > 0$ *such that for all* $y \in \mathbb{R}^\ell$, $t \geqslant 1$,

$$|p_t^{(1)}(x,f;y) - p_\infty^{(1)}(f;y)| \leqslant C\frac{(1+|x|^m)}{(1+t)^k}, \tag{4.49}$$

$$|p_\infty^{(1)}(f,y)| \leqslant C; \tag{4.50}$$

moreover,

$$p_\infty^{(1)}(f,y) = \partial_y p_\infty(f,y). \tag{4.51}$$

Proof. 1. Assuming (4.46) is established, the statement (4.48) follows from taking the limit as $t \to \infty$ in the identity

$$p_t(x,g;y+he_i) - p_t(x,g;y) = \int\limits_0^h q_t^i(x,g;y+\alpha e_i)\,d\alpha$$

since the function

$$q_t^i(x,g,\cdot) := \partial_{y^i} p_t(x,g;y) = \int\limits_0^t E_x \frac{\partial L}{\partial y^i}(X_s,y) p_{t-s}(X_s,g;y)\,ds$$

is continuous, being a component of the vector function q_t. Existence of q_∞ and the bound (4.47) follow from the calculus in last step of this proof and due to Lemma 1.

2. Let us establish (4.46). We drop y for simplicity. We have

$$q_t(x,g) - q_\infty(x,g) = \int\limits_0^t ds \int f_s^1(x'',g)\,\mu_{x,t-s}^y(dx'') - \int\limits_0^\infty ds \int f_s^1(x'',g)\,\mu_\infty(dx'')$$
$$= \int\limits_0^{t/2} ds \int f_s^1(x'',g)\,(\mu_{x,t-s}^y(dx'') - \mu_\infty^y(dx''))$$
$$+ \int\limits_{t/2}^t ds \int f_s^1(x'',g)\,\mu_{x,t-s}^y(dx'') - \int\limits_{t/2}^\infty ds \int f_s^1(x'',g)\,\mu_\infty^y(dx''). \tag{4.52}$$

Now, the assertion (4.46) follows from the estimates of Lemma 1 – in particular, applied with $t_1 = t/2$ – and Proposition 1. Indeed, we have with any $a^{-1} + b^{-1} = 1$, by the Hölder inequality and having in mind (3.1), (3.2) and (3.3):

$$|q_t(x,g) - q_\infty(x,g)| \leqslant \int\limits_0^{t/2} ds \left(\int |f_s^1(x'',g)|^a\,|\mu_{x,t-s}^y(dx'') - \mu_\infty^y(dx'')|\right)^{1/a} ||\mu_{x,t-s}^y - \mu_\infty^y\|_{TV}^{1/b}$$
$$+ \int\limits_{t/2}^t ds \int |f_s^1(x'',g)|\,\mu_{x,t-s}^y(dx'') + \int\limits_{t/2}^\infty ds \int |f_s^1(x'',g)|\,\mu_\infty^y(dx'').$$

3. The term with $0 \leqslant s \leqslant t/2$ we may estimate similarly to the calculus in the proof of Theorem 2 (cf. step 1). We have

$$\left(\int |f_s^1|^a(x'',g)\,|\mu_{x,t-s}^y(dx'') - \mu_\infty^y(dx'')|\right)^{1/a}$$
$$\leqslant \left(E_x|f_s^1|^a(X_{t-s},g,y)\right)^{1/a} + \left(\int E_z|f_s^1|^a(X_{t-s},g,y)\mu_\infty^y(dz)\right)^{1/a}. \tag{4.53}$$

Let us consider the first expression here integrated with respect to s:

$$\int\limits_0^{t/2} \left(E_x|f_s^1|^a(X_{t-s},g,y)\right)^{1/a} ds = \frac{t}{2}\,\frac{2}{t} \int\limits_0^{t/2} \left(E_x|f_s^1|^a(X_{t-s},g,y)\right)^{1/a} ds$$

$$\leqslant \frac{t}{2} \left(\frac{2}{t} \int\limits_0^{t/2} E_x|f_s^1|^a(X_{t-s},g,y)\,ds\right)^{1/a} = \frac{t^{1-1/a}}{2^{1-1/a}} \left(\int\limits_0^{t/2} E_x|f_s^1|^a(X_{t-s},g,y)\,ds\right)^{1/a}.$$

According to Krylov's estimate, the latter integral admits the bound:

$$\int\limits_0^{t/2} E_x|f_s^1|^a(X_{t-s},g,y)\,ds \leqslant N\|f_s^1\|_{L_{a(d+1)}([0,t/2]\times\mathbb{R}^d)}.$$

So, if the last norm were finite, we would have achieved our goal to estimate this term. However, since the norm of f^1 is only locally bounded with a possible moderate growth in the appropriate L_p sense, the last bound is, generally speaking, useless and we ought to repeat our localization procedure. *Given* t, we modify stopping times as follows:

$$\tau^0 = 0, \quad \tau^{n+1} := \inf(s \geqslant \tau^n : \ |X_s - X_{\tau^n}| \geqslant Rt), \quad \inf(\varnothing) = \infty.$$

The value of R we assume here large enough, but independent on x, so that

$$\sup_x P_x(\tau^n < t/2) \leqslant q^n, \quad q := \sup_t \sup_x P_x(\sup_{0\leqslant s\leqslant t} |X_s - x| \geqslant Rt) < 1. \tag{4.54}$$

The second inequality here for R large enough (say, $R > (\|b\|_{L_\infty} \vee \|\sigma\|_{L_\infty})$ follows from the Chebyshev–Markov and Burkholder–Davis–Gundy inequalities,

$$\begin{aligned}
&P_x\left(\sup_{0\leqslant s\leqslant t/2} |X_s - x| \geqslant Rt\right) \\
&\leqslant P_x\left(\sup_{0\leqslant s\leqslant t/2}\left|\int_0^{t/2} b(X_r)dr\right| \geqslant Rt/2\right) + P_x\left(\sup_{0\leqslant s\leqslant t/2}\left|\int_0^{t/2} \sigma(X_r)dW_r\right| \geqslant Rt/2\right) \\
&\leqslant (Rt/2)^{-1} E_x \sup_{0\leqslant s\leqslant t/2}\left|\int_0^{s} \sigma(X_r)\,dW_r\right| \leqslant (Rt/2)^{-1}\sqrt{E_x\int_0^{t/2} \|\sigma(X_r)\|^2\,dr} \leqslant \frac{\|\sigma\|_{L_\infty}}{R} < 1.
\end{aligned}$$

Now,

$$\begin{aligned}
&E_x \int_0^{t/2} |f^1_s|^a(X_{t-s}; g, y)\,ds \\
&\quad = E_x \sum_{k=0}^{\infty} 1(\tau^k < t/2 \leqslant \tau^{k+1}) \int_0^{t/2} |f^1_s|^a(X_{t-s}; g, y)\,ds \\
&\quad \leqslant \sum_{k=0}^{\infty}\left(E_x\, 1(\tau^k < t/2 \leqslant \tau^{k+1})\right)^{1/2}\left(E_x\left(\int_0^{\tau^{k+1}\wedge(t/2)} |f^1_s|^a(X_{t-s}; g, y)\,ds\right)^2\right)^{1/2}.
\end{aligned}$$

Further,

$$\begin{aligned}
E_x\left(\int_0^{(t/2)\wedge\tau^{k+1}} |f^1_s|^a(X_{t-s}; g, y)\,ds\right)^2 &\leqslant E_x\left(t\int_0^{\tau^{k+1}} |f^1_s|^{2a}(X_{t-s}; g, y)\,ds\right) \\
&\leqslant C\,t\,\|f^1\|^2_{L_{2a(d+1)}([0,t]\times B_{|x|+(k+1)Rt})} \leqslant C\,t^{1+\frac{1}{d+1}}\,(1+|x|+Rt)^{2md}(k+1)^{2md}.
\end{aligned}$$

Hence, simplifying a little bit the formulas, we may write

$$\begin{aligned}
E_x \int_0^{t/2} |f^1_s|^a(X_{t-s}; g, y)\,ds &\leqslant C(1+t^{m'})(1+|x|+R)^{m'} \sum_{k=0}^{\infty} (q^{k/2}\,(k+1)^{2md})^{1/2} \\
&= C(1+t^{m'})(1+|x|+R)^{m'}.
\end{aligned} \tag{4.55}$$

The point is that the right-hand side here increases as a fixed power function (in time), while convergence in total variation to the stationary distribution is faster than any polynomial rate. But before turning to the details, we ought to consider the second term in (4.53).

4. We estimate, as above,

$$\int_0^{t/2}\left(\int E_z|f^1_s|^a(X_{t-s},g,y)\mu^y_\infty(dz)\,ds\right)^{1/a}$$

$$\leqslant \frac{t^{1-1/a}}{2^{1-1/a}}\left(\int\left(\int_0^{t/2} E_z|f^1_s|^a(X_{t-s},g,y)\,ds\right)\mu^y_\infty(dz)\right)^{1/a}$$

$$\leqslant C(1+t^{m'})\int(1+|z|^{m'}+R^{m'})\mu^y_\infty(dz)\leqslant C(1+t^{m'})(1+R^{m'}),$$

where we used (4.55).

5. Returning to (4.52), we conclude that

$$\int_0^{t/2} ds\int|f^1_s|(x'',g)\,|\mu^y_{x,t-s}(dx'')-\,\mu^y_\infty(dx'')|\leqslant C\frac{(1+|x|^{m''})(1+R)^{m'}(1+t^{m''})}{(1+t^{k'})},\qquad(4.56)$$

where R, m', and m'' are some fixed values while k' may be chosen as large as we like; the choice of k' here affects only the multiplier C.

6. The next integral in (4.52) may be estimated by virtue of the bounds (4.36), (3.2) and (3.3) as follows, by Krylov's estimate ($p=d+1$),

$$\int_{t/2}^{t} ds\int|f^1_s|(x'',g)\,\mu^y_{x,t-s}(dx'')=E_x\int_{t/2}^{t}|f^1_s|(X_{t-s})\,ds\leqslant N\|f^1\|_{L_p([t/2,t]\times\mathbb{R}^d)}.$$

As earlier in similar expressions, the right-hand side here may diverge, so, again a localization procedure is needed. Given t, let

$$\tau^0=t/2,\qquad \tau^{n+1}:=\inf(s\geqslant\tau^n:\ |X_s-X_{\tau^n}|\geqslant Rt),\quad \inf(\varnothing)=\infty.$$

The value of R we assume to be large enough, $R>\|b\|\vee\|\sigma\|$, so that

$$\sup_x P_x(\tau^n<t/2)\leqslant q^n,\quad q:=\sup_t\sup_x P_x(\sup_{t/2\leqslant s\leqslant t}|X_s-X_{t/2}|\geqslant Rt)<1\qquad(4.57)$$

and $R\geqslant|x|$. Then

$$E_x\int_{t/2}^{t}|f^1_s|(X_{t-s};g,y)\,ds=E_x\sum_{k=0}^{\infty}1(\tau^k<t\leqslant\tau^{k+1})\int_{t/2}^{t}|f^1_s|(X_{t-s};g,y)\,ds$$

$$\leqslant\sum_{k=0}^{\infty}\left(E_x\,1(\tau^k<t/2\leqslant\tau^{k+1})\right)^{1/2}\left(E_x\left(\int_{t/2}^{\tau^{k+1}\wedge t}|f^1_s|(X_{t-s};g,y)\,ds\right)^2\right)^{1/2}.$$

Further, due to (4.36),

$$E_x\left(\int\limits_{t/2}^{t\wedge\tau^{k+1}} |f_s^1|(X_{t-s};g,y)\,ds\right)^2 \leqslant E_x\left(\frac{t}{2}\int\limits_{t/2}^{\tau^{k+1}} |f_s^1|^2(X_{t-s};g,y)\,ds\right)$$

$$\leqslant C\,t\,\|f^1\|^2_{L_{2a(d+1)}([t/2,t]\times B_{(k+1)Rt})} \leqslant C\,\frac{t^{1+\frac{1}{d+1}}\,(1+Rt)^{2md}(k+1)^{2md}}{1+(t/2)^k},$$

where k may be chosen arbitrarily large. Hence

$$E_x\int\limits_{t/2}^{t} |f_s^1|(X_{t-s};g,y)\,ds \leqslant C\frac{(1+t^{m'})(1+R)^{m'}}{1+(t/2)^k}\sum_{k=0}^{\infty}\left(q^{k/2}\,(k+1)^{2md}\right)^{1/2}$$

$$= C\frac{(1+t^{m'})(1+R)^{m'}}{1+t^k} \leqslant C\frac{(1+R)^{m'}}{1+t^k}.$$

Here, $R > \|b\| \vee \|\sigma\| \vee |x|$, so we finally get,

$$E_x\int\limits_{t/2}^{t} |f_s^1|(X_{t-s};g,y)\,ds \leqslant C\frac{(1+|x|^{m'})}{1+t^k}. \tag{4.58}$$

7. Now consider the last integral in (4.52),

$$\int\limits_{t/2}^{\infty}\int |f_s^1|(x'',g)\,\mu_\infty^y(dx'')\,ds = \int\left(\int\limits_{t/2}^{\infty} E_z|f_s^1|(X_s,g)\,ds\right)\mu_\infty^y(dz)$$

$$= \int\left(\sum_{j=1}^{\infty}\int\limits_{jt/2}^{(j+1)t/2} E_z|f_s^1|(X_s,g)\,ds\right)\mu_\infty^y(dz).$$

Quite similarly to the previous step (cf. (4.58)), each term in the above sum admits the estimate

$$\int\limits_{jt/2}^{(j+1)t/2} E_z|f_s^1|(X_s,g)\,ds \leqslant C\frac{(1+|z|)^{m'}}{1+(jt)^k}.$$

Therefore, we get,

$$\int\limits_{t/2}^{\infty}\int |f_s^1|(x'',g)\,\mu_\infty^y(dx'')\,ds \leqslant \int\left(\sum_{j=1}^{\infty} C\frac{(1+|z|)^{m'}}{1+(jt)^k}\right)\mu_\infty^y(dz) \leqslant \frac{C}{1+t^k},$$

as required. All estimates for $q_t(x,g;y)$ are established.

8. The claims about $p_t^{(1)}(x,f;y)$ follow straightforward from the corresponding inequalities for $q_t(x,g;y)$ by virtue of (2.6), (4.37), (4.44) and (4.45). Theorem 3 is proved. □

4.3 Derivatives $\partial_x q_t(x,g;y)$ and $\partial_x p_t^{(1)}(x,f;y)$ at $t\to\infty$

We now study the derivative $\partial_x q_t(x,g;y)$. The assumptions of Theorem 1 with $(H^{0,2})$ relaxed to $(H^{0,1})$. are satisfied throughout this subsection about the coefficients of the SDE and the function g. Recall that

$$q_t(x,g;y) = \int_0^t E_x \frac{\partial L}{\partial y}(X_s,y) p_{t-s}(X_s,g;y)\,ds.$$

We have

$$\begin{aligned}\partial_x q_t(x,g;y) &= \partial_x \int_0^t E_x \frac{\partial L}{\partial y}(X_s,y) p_{t-s}(X_s,g;y)\,ds \\ &= \partial_x \int_0^{t/2} E_x \frac{\partial L}{\partial y}(X_s,y) p_{t-s}(X_s,g;y)\,ds + \partial_x \int_{t/2}^{t} E_x \frac{\partial L}{\partial y}(X_s,y) p_{t-s}(X_s,g;y)\,ds \\ &=: I_1 + I_2. \end{aligned} \tag{4.59}$$

To estimate both integrals $I_{1,2}$, let

$$v^1(s,x) := E_x \int_s^t \frac{\partial L}{\partial y} u(t-r,X_r)\,dr, \quad 0\leqslant s\leqslant t,$$

$$v^2(s,x) := E_x \int_s^{t/2} \frac{\partial L}{\partial y} u(t-r,X_r)\,dr, \quad 0\leqslant s\leqslant t/2.$$

Note that $I_1 = \partial_x v^2(t/2,x)$ and $I_2 = \partial_x v^1(0,x)$. The idea of estimation is that the value $v^1(t/2,x)$ is small in some $W_p^{1,2}$ space because $u_{xx}(t-s,\cdot)$ is small when $t-s$ is large. On the other hand, the value $v^2(0,x)$ is small in a similar Sobolev sense because of the integration with respect to μ_s^y over large values of s, the latter measure being close to μ_∞^y. Both functions are solutions of certain parabolic partial differential equations. So, estimates in Sobolev spaces will be based on *a priori* bounds from [9]. Let us now show the details.

Lemma 2. *Under assumptions (H_b), (H_a), (H_g), (H_g'), (H_g''), and $(H^{0,1})$, for any $k>0$ there exist $C,m>0$ such that for every $t\geqslant 0$,*

$$\|v^1\|_{W_p^{1,2}([0,t]\times B_R)} \leqslant C(1+R^m), \tag{4.60}$$

$$\|v^1\|_{W_p^{1,2}([0,t-t_1]\times B_R)} \leqslant \frac{C(1+R^m)}{1+t_1^k}. \tag{4.61}$$

Proof. The function $(v^1(s,x):\ 0\leqslant s\leqslant t,\ x\in\mathbb{R}^d)$ solves the equation

$$v_s^1(s,x) + Lv^1(s,x) = -\frac{\partial L}{\partial y} u(t-s,x), \quad v^1(t,x)=0. \tag{4.62}$$

Note that for any $t_1 \leqslant t_2$ and $D \in \mathscr{B}(\mathbb{R}^d)$,

$$\|\frac{\partial L}{\partial y} u(t-\cdot,\cdot)\|_{L_p([t_1,t_2]\times D)} \leqslant C\|u(t-\cdot,\cdot)\|_{W_p^{1,2}([t_1,t_2]\times D)}.$$

Due to *a priori* bounds (4.41) and/or (4.42) and taking into account the terminal condition $v^1(t,x) \equiv 0$, we have

$$\|v^1\|_{W_p^{1,2}([0,t]\times B_R)} \leqslant C\big(\|v^1\|_{L_p([0,t]\times B_{R+1})} + \|u\|_{W_p^{1,2}([0,t]\times B_{R+1})}\big),$$

and with $0 < t_1 < t$,

$$\|v^1\|_{W_p^{1,2}([0,t_1]\times B_R)} \leqslant C\big(\|v^1\|_{L_p([0,t_1+\delta]\times B_{R+1})} + \|u\|_{W_p^{1,2}([t-t_1-\delta,t]\times B_{R+1})}\big).$$

Here, as we know, for any $0 < t_1 \leqslant t$ with arbitrarily large k and some C, m

$$\|u\|_{W_p^{1,2}([0,t]\times B_R)} \leqslant C(1+R^m)$$

and

$$\|u\|_{W_p^{1,2}([t_1,t]\times B_R)} \leqslant C\frac{(1+R^m)}{1+t_1^k}.$$

On the other hand, due to Krylov's estimate (with some new m),

$$\begin{aligned}\|v^1\|_{L_p([0,t]\times B_{R+1})} &\leqslant CR^{d/p}\, t^{1/p}\, \|v^1\|_{L_\infty([0,t]\times B_{R+1})} \\ &\leqslant CR^{d/p}\, t^{1/p}\, \|u\|_{W_p^{1,2}([0,t]\times B_{R+1})} \leqslant Ct^{1/p}\,(1+R^m)\end{aligned} \tag{4.63}$$

and, similarly,

$$\begin{aligned}\|v^1\|_{L_p([0,t-t_1]\times B_{R+1})} &\leqslant CR^{d/p}\, t^{1/p}\, \|v^1\|_{L_\infty([0,t-t_1]\times B_{R+1})} \\ &\leqslant CR^{d/p}\, t^{1/p}\, \|u\|_{W_p^{1,2}([t_1,t]\times B_{R+1})} \leqslant C\frac{(1+R^m)}{1+t_1^k}.\end{aligned}$$

Note that the latter two inequalities imply the following improvement of (4.63):

$$\|v^1\|_{L_p([0,t]\times B_{R+1})} \leqslant C(1+R^m).$$

Hence we obtain (4.60) and (4.61). Lemma 2 is proved. □

Lemma 3. *Under assumptions* (H_b), (H_a), (H_g), (H_g'), (H_g''), *and* $(H^{0,1})$, *for any* $k > 0$ *there exist* $C, m > 0$ *such that for every* $t \geqslant 0$,

$$\|\partial_x v^2\|_{L_p([0,t/2]\times B_R)} + \|\partial_x^2 v^2\|_{L_p([0,t/2]\times B_R)} \leqslant \frac{C(1+R^m)}{1+t^k}. \tag{4.64}$$

Proof. Let us use the representation

$$v^2(s,x) := E_x \int\limits_s^{t/2} \frac{\partial L}{\partial y} u(t-r, X_r)\, dr$$

$$= \int\limits_s^{t/2} E_x \frac{\partial L}{\partial y} u(t-r, X_r)\, dr = \int\limits_s^{t/2} \Big\langle \frac{\partial L}{\partial y} u(t-r, x'), \mu_{x,r}^y(dx')\Big\rangle\, dr.$$

So,

$$\partial_x v^2(s,x) = \partial_x \int\limits_s^{t/2} \Big\langle \frac{\partial L}{\partial y} u(t-r,x'), \mu^y_{x,r}(dx') \Big\rangle\, dr$$

$$= \partial_x \int\limits_s^{t/2} \Big\langle \frac{\partial L}{\partial y} u(t-r,x'), (\mu^y_{x,r} - \mu^y_\infty)(dx') \Big\rangle\, dr.$$

Hence

$$\|\partial_x v^2\|_{L_p([0,t/2]\times B_R)} \leqslant \|\widetilde{v}^2\|_{W^{1,2}_p([0,t/2]\times B_R)},$$

where

$$\widetilde{v}^2(s,x) := \int\limits_s^{t/2} \Big\langle \frac{\partial L}{\partial y} u(t-r,x'), (\mu^y_{x,r} - \mu^y_\infty)(dx') \Big\rangle\, dr.$$

Repeating our localization arguments and using the estimate (3.4) as in the previous theorem, we get the bound

$$\|\widetilde{v}^2\|_{W^{1,2}_p([0,t/2]\times B_R)} \leqslant \frac{C(1+R^m)}{1+t^k}.$$

The same bound also guarantees a similar estimate for the second derivative,

$$\|\partial^2_x v^2\|_{L_p([0,t/2]\times B_R)} \leqslant \|\widetilde{v}^2\|_{W^{1,2}_p([0,t/2]\times B_R)}.$$

Lemma 3 is proved. □

Theorem 4. *Let assumptions* (H_b), (H_a), (H_g), (H'_g), (H''_g), *and* $(H^{0,1})$ *hold. Then for all* $k \geqslant 0$ *there exist real numbers* C *and* m *such that for all* $y \in \mathbb{R}^\ell$, $x, x' \in \mathbb{R}^d$ *and all* $t \geqslant 1$

$$|\partial_x q_t(x,g;y)| \leqslant C\frac{1+|x|^m}{(1+t)^k}. \tag{4.65}$$

Moreover, under assumptions (H_b), (H_a), (H_g), (H'_f), (H''_f), *and* $(H^{0,1})$, *for all* $k \geqslant 0$ *there exist real numbers* C *and* m *such that for all* $y \in \mathbb{R}^\ell$, $x, x' \in \mathbb{R}^d$ *and all* $t \geqslant 1$

$$\Big|\partial_x p^{(1)}_t(x,f;y)\Big| \leqslant C\frac{1+|x|^m}{(1+t)^k}. \tag{4.66}$$

Proof. 1. The proof of (4.65) follows straightforward from Lemmas 2 and 3 by virtue of the embedding theorems [9, Lemma 2.3.3].

2. Due to (2.6), the difference between $p^{(1)}_t(x,f,y)$ and $q_t(x,f(\cdot,y^*),y)|_{y^*=y}$ consists of the integral $\langle f_y(\cdot,y), \mu^y_{x,t} - \mu^y_\infty\rangle \equiv p_t(x,f_y;y)$ (cf. (2.2)). So, it suffices to show the estimate

$$|\partial_x p_t(x,f_y;y)| \leqslant C\frac{1+|x|^m}{(1+t)^k}. \tag{4.67}$$

This follows straight away from (4.35) and embedding theorems (cf. again [9, Lemma 2.3.3]). Theorem 4 is proved. □

4.4 Second derivatives $\partial_y^2 p_t(x,g;y)$ and $D_y^2 p_t(x,f;y)$

We now study the matrix-valued functions $q_t^{(1)}(x,g,y)$ and $p_t^{(2)}(x,g,y)$ defined in (2.5) and (2.8). Remind that

$$p_t^{(2)}(x,g;y) \equiv q_t^{(1)}(x,g;y)$$

$$:= \int_0^t ds\, p_s\left(x, L_{0,2}(\cdot,y)p_{t-s}(\cdot,g;y)\right) + \int_0^t ds \sum_{|i|=1} q_s\left(x, L_{i,2}(\cdot,y)p_{t-s}(\cdot,g;y)\right)$$

and

$$p_t^{(2)}(x,f,y) := q_t^{(1)}(x,f(\cdot,y),y) + 2q_t(x,f_y(\cdot,y),y) + \langle f_{yy}(\cdot,y), \mu_{x,t}^y - \mu_\infty^y \rangle.$$

Here, $(H^{0,2})$ is assumed instead of $(H^{0,1})$. The function $p^{(2)}(x,g;y)$ formally should satisfy the equation

$$\partial_t p_t^{(2)}(x,g;y) = L\, p_t^{(2)}(x,g;y) + \sum_{0\leqslant|i|\leqslant 1} L_{i,2} p_t^{(i)}(x,g;y), \quad p_0^{(2)}(x,g;y) = 0, \tag{4.68}$$

with

$$L_{i,2} := C_2^i \frac{\partial^{2-i} L}{\partial y^{2-i}}.$$

This follows from the (formal) differentiation with respect to y of Equation (4.4) on $p^{(1)} \equiv q$. This is a reason to *define* $p_t^{(2)}(x,g;y) \equiv q_t^{(1)}(x,g;y)$ by formula (2.5),

$$p_t^{(2)}(x,g;y) := \int_0^t ds\, p_s\left(x, L_{0,2}(\cdot)p_{t-s}(\cdot,g)\right) + \int_0^t ds \sum_{|i|=1} q_s\left(x, L_{i,2}(\cdot)p_{t-s}(\cdot,g)\right).$$

Then we have the following result.

Theorem 5. (I). *Suppose that (H_a), (H_b), (H_g), (H_g'), (H_g''), and $(H^{0,2})$ hold. Then the matrix-valued function $p_t^{(2)}(x,g;y) \equiv q_t^{(1)}(x,g;y)$ is well defined by (2.5), continuous with respect to y, has a moderate growth in x uniformly with respect to y and the equality holds for every x,y,t,*

$$\partial_y^2 p_t(x,g;y) = p_t^{(2)}(x,g;y) \equiv q_t^{(1)}(x,g;y). \tag{4.69}$$

Moreover, there exists a limit

$$\lim_{t\to\infty} q_t^{(1)}(x,g;y) =: q_\infty^{(1)}(g;y), \tag{4.70}$$

this limit admits a representation

$$q_\infty^{(1)}(g;y) = \int_0^\infty ds\, p_s\left(x, L_{0,2}(\cdot,y)p_\infty(\cdot,g;y)\right) + \int_0^\infty ds \sum_{|i|=1} q_s\left(x, L_{i,2}(\cdot)p_\infty(\cdot,g)\right), \tag{4.71}$$

and for any $k>0$ there exist $C,m>0$ such that (uniformly over y)

$$|q_t^{(1)}(x,g;y) - q_\infty^{(1)}(g;y)| \leqslant C\frac{1+|x|^m}{(1+t)^k}, \quad t \geqslant 0. \tag{4.72}$$

Moreover, for any $k > 0$ there exist $C, m' > 0$ such that

$$|\partial_x q_t^{(1)}(x, g; y)| \leqslant \frac{C(1 + |x|^{m'})}{(1 + t)^k}, \quad t \geqslant 0. \tag{4.73}$$

Moreover, for any $T > 0$, $(q_t^{(1)}(x, g),\ 0 \leqslant t \leqslant T)$ has a moderate growth in x in the sense of $W^{0,1}_{p,\mathrm{loc}}$ for any $p > 1$, i.e.,

$$\|q_t^{(1)}(x, g; y)\|_{W_p^{0,1}([0,t]\times B_R)} \leqslant C(1 + R^{m'}), \quad t \geqslant 0, \tag{4.74}$$

and, moreover, for any $k > 0$ there exist $C, m' > 0$ such that

$$\|\partial_x q_t^{(1)}(x, g; y)\|_{W_p^{0,1}([t_1,t]\times B_R)} \leqslant \frac{C(1 + R^{m'})}{(1 + t_1)^k}, \quad t \geqslant 0, \tag{4.75}$$

(II). *Let assumptions (H_a), (H_b), (H_f), (H'_f), (H''_f), and $(H^{0,2})$ hold. Then the matrix-valued function $p_t^{(2)}(x, f; y)$ is well defined by (2.8), is continuous with respect to y, has a moderate growth in x uniformly with respect to y and the equality holds for every x, y, t*

$$\partial_y^2 p_t(x, f; y) = p_t^{(2)}(x, f; y). \tag{4.76}$$

Moreover, there exists a limit

$$\lim_{t\to\infty} p_t^{(2)}(x, f; y) =: p_\infty^{(2)}(f; y), \tag{4.77}$$

this limit admits a representation (cf. above (4.71) and (4.45))

$$p_\infty^{(2)}(f; y) = q_\infty^{(1)}(f; y) + 2q_\infty(x, f_y(\cdot, y), y) \tag{4.78}$$

and for any $k > 0$ there exist $C, m > 0$ such that (uniformly over y)

$$|p_t^{(2)}(x, f; y) - p_\infty^{(2)}(f; y)| \leqslant C\frac{1 + |x|^m}{(1 + t)^k}, \quad t \geqslant 0. \tag{4.79}$$

Moreover, for any $k > 0$ there exist $C, m' > 0$ such that

$$|\partial_x p_t^{(2)}(x, f; y)| \leqslant \frac{C(1 + |x|^{m'})}{(1 + t)^k}, \quad t \geqslant 0. \tag{4.80}$$

Moreover, for any $T > 0$, $(p_t^{(2)}(x, f),\ 0 \leqslant t \leqslant T)$ has a moderate growth in x in the sense of $W^{0,1}_{p,\mathrm{loc}}$ for any $p > 1$, i.e.,

$$\|p_t^{(2)}(x, f; y)\|_{W_p^{0,1}([0,t]\times B_R)} \leqslant C(1 + R^{m'}), \quad t \geqslant 0, \tag{4.81}$$

and, moreover, for any $k > 0$ there exist $C, m' > 0$ such that

$$\|\partial_x p_t^{(2)}(x, f; y)\|_{W_p^{0,1}([t_1,t]\times B_R)} \leqslant \frac{C(1 + R^{m'})}{(1 + t_1)^k}, \quad t \geqslant 0, \tag{4.82}$$

Remark 4. Here, the main assertions are (4.69), (4.70), and (4.72), i.e., the existence of the second derivative $\partial_y^2 u$, its continuity and some properties at $t \to \infty$, which all, in particular, may help apply Itô's (in our case, more precisely, Itô–Krylov's) formula in the *corrector method* (cf., for example, [1]). However, (4.74) and (4.75) may be helpful in further studies such as establishing higher derivatives.

Proof. 1. If we denote

$$f_t^2(x;g;y) := \sum_{0\leqslant|i|\leqslant 1} L_{i,2}(x,y)p_t^{(i)}(x,g;y),$$

then formula (2.5) may be written in the form,

$$p_t^{(2)}(x,g;y) := \int_0^t E_x f_{t-s}^2(X_s,g;y)\,ds.$$

Now let us remind which properties of the functions v and f^1 (cf. (4.6) and (4.32)) have been used in the proofs of Theorems 2 and 3, which guarantee similar properties for $p_t^{(1)}(x,g;y)$. Clearly, if we show the same properties of f^2, then the desired results about $p_t^{(2)}(x,g;y)$ will follow. The crucial about v were inequalities (4.9) and (4.15), while about f^1 the estimates (4.34) and (4.36) sufficed; also notice that practically (4.9) follows from (4.34). So, all we need to know about f^2 now is the following four assertions:

$$\|p_\cdot(\cdot;g)\|_{W_p^{1,2}([0,t]\times B_R)} + \|\partial_x p_\cdot^{(1)}(\cdot;g)\|_{W_p^{0,1}([0,t]\times B_R)} \leqslant C(1+|R|^m), \tag{4.83}$$

$$\|p_\cdot(\cdot;g)\|_{W_p^{1,2}([t_1,t]\times B_R)} + \|\partial_x p_\cdot^{(1)}(\cdot;g)\|_{W_p^{0,1}([t_1,t]\times B_R)} \leqslant C\frac{(1+|R|^{m'})}{1+t_1^k}, \tag{4.84}$$

$$\int_0^t\int_{B_R} |p_s(x,g;y) - p_s(x,g;y')|^p\,dxds \leqslant C_t\rho(|y'-y|)(1+R)^m, \tag{4.85}$$

$$\int_0^t\int_{B_R} |q_s(x,g;y) - q_s(x,g;y')|^p\,dxds \leqslant C_t\widetilde{\rho}(|y'-y|)(1+R)^m, \tag{4.86}$$

the latter one with *some* modulus of continuity (even depending on R). Note that, in fact, (4.83) follows from the estimates (4.33) from Lemma 1, (4.60) from Lemma 2 and (4.64) from Lemma 3; (4.84) also holds due to (4.35) from Lemma 1, (4.61) from Lemma 2 and again (4.64) from Lemma 3; and (4.85) have been established above in (4.15). So, it only remains to check (4.86).

2. The function q satisfies the equation

$$\partial_t q_t - L(x,y)q_t = (\partial_y L(x,y))p_t, \quad q_0 = 0$$

(cf. (4.4)). Hence the difference

$$z_t(x,y,y') := p_t^{(1)}(x,g;y) - p_t^{(1)}(x,g;y')$$

solves

$$
\begin{aligned}
&\partial_t(p_t^{(1)}(x,g;y) - p_t^{(1)}(x,g;y')) - L(y)(p_t^{(1)}(x,g;y) - p_t^{(1)}(x,g;y')) \\
&= (L(y) - L(y'))p_t^{(1)}(x,g;y') + (\partial_y L(y))p_t(x,g;y) - (\partial_y L(y'))p_t(x,g;y'), \qquad (4.87) \\
&p_0^{(1)}(x,g;y) - p_0^{(1)}(x,g;y') = 0.
\end{aligned}
$$

Denote

$$
-F_t(x,y,y') := (L(y) - L(y'))p_t^{(1)}(x,g;y') + (\partial_y L(y))p_t(x,g;y) - (\partial_y L(y'))p_t(x,g;y').
$$

Then

$$
\partial_t z_t(x,y,y') - L(y)z_t(x,y,y') = -F_t(x,y,y'), \quad z_0(x,y,y') \equiv 0. \qquad (4.88)
$$

It is already known that the right-hand side in Equation (4.88) is continuous with respect to the parameter y in the sense of L_p locally in (x,t) (with a moderate growth in x). Hence, we may conclude similarly to the step 3 in the proof of Theorem 2 that, indeed, (4.86) holds.

3. The claims about $p_t^{(2)}(x,f;y)$ follow due to the first part of Theorem (about $q_t^{(1)}(x,f;y)$) and by Theorems 2–4, which guarantee all similar bounds for the second and third terms in the right-hand side of (2.8), i.e., for $2q_t(x, f_y(\cdot,y), y)$ and $\langle f_{yy}(\cdot,y), \mu_{x,t}^y - \mu_\infty^y\rangle$. Thus, Theorem 5 is proved. □

References

1. É. Pardoux and A. Yu. Veretennikov, "On Poisson equation and diffusion approximation 2," *Ann. Probab.* **31**, No. 3, 1166–1192 (2003).

2. É. Pardoux and A. Yu. Veretennikov, "On the Poisson equation and diffusion approximation. I," *Ann. Probab.* **29**, No. 3, 1061–1085 (2001).

3. S. N. Ethier and T. G. Kurtz, *Markov Processes. Characterization and Convergence.* John Wiley & Sons, New York etc. (1986)

4. G. C. Papanicolaou, D. W. Stroock, and S. R. S. Varadhan, "Martingale approach to some limit theorems," In: *Conference on Statistical Mechanics, Dynamical Systems and Turbulence*, M. Reed ed., Duke Univ. Math. Series 3 (1977).

5. A. Benveniste, M. Metivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*, Springer, Berlin etc. (1990)

6. A. M. Kulik and A. Yu. Veretennikov, "On extended Poisson equation for weakly ergodic Markov processes," [Submitted]

7. A. Yu. Veretennikov, "On polynomial mixing and convergence rate for stochastic difference and differential equations"' [in Russian], *Theory Prob. Appl.* **44**, No. 2, 312–327 (1999); English transl.: *Theory Probab. Appl.* **44**, No. 2, 361–374 (2000).

8. N. V. Krylov, "On the selection of a Markov process from a system of processes and the construction of quasi–diffusion processes" [in Russian], *Izv. AN SSSR, Ser. Mat.* **7**, No. 3, 691-708 (1973); English transl.: *Math. USSR Izv.* **7**, No. 3, 691-709 (1973).

9. O. Ladyzenskaja, V. Solonnikov, and N. Ural'ceva *Linear and Quasilinear Equations of Parabolic Type*, Am. Math. Soc., Providence, RI (1973).

10. E. B. Dynkin, *Markov Processes*, Vols. 1-2, Springer, Berlin etc. (1965)

11. A. Yu. Veretennikov, "Parabolic equations and Ito's stochastic equations with coefficients discontinuous in the time variable," *Mat. Zametki* **31**, 549–557 (1982) ; English transl.: *Math. Notes* **31**, 278–283.

12. N. V. Krylov, *Controlled Diffusion Processes*, Springer, Berlin etc. (1980).

Submitted on August 14, 2011

Revised version on September 16, 2011