ON STOCHASTIC EQUATIONS WITH DRIFT IN $L_d$

N.V. KRYLOV

Abstract. For the Itô stochastic equations in $\mathbb{R}^d$ with drift in $L_d$, several results are discussed such as the existence of weak solutions, the existence of the corresponding Markov process, the Aleksandrov type estimates of their Green’s functions, which yield their summability to the power of $d/(d - 1)$, the Fabes-Stroock type estimates which show that Green’s functions are summable to a higher degree, the Fanghua Lin type estimates, which are one of the main tools in the $W_2^p$-theory of fully nonlinear elliptic equations, the fact that Green’s functions are in the class $A_\infty$ of Muckenhoupt and a few other results.

1. Introduction

Let $\mathbb{R}^d$ be an Euclidean space of points $x = (x^1, ..., x^d)$. We assume that $d \geq 2$ and denote

$$B_R(x) = \{y \in \mathbb{R}^d : |y - x| < R\}, \quad B_R = B_R(0), \quad D_i = \frac{\partial}{\partial x^i}, \quad D_{ij} = D_i D_j.$$

First we address the issue of existence of solutions of Itô’s equations with drift term in $L_d$.

Example 1.1. Let $b(x) = -x|x|^{-2}(d/2)$. We have $b \in L_{d-\varepsilon}(B_1)$ for any $\varepsilon \in (0, 1)$ but not for $\varepsilon = 0$ and it turns out that there is no solutions of the equation $dx_t = dw_t + b(x_t) dt$ starting at zero, where $w_t$ is a $d$-dimensional Wiener process.

Indeed, if we assume the contrary, then for the equation to make sense

$$\int_0^t |b(x_t)| \, dt$$

should be finite. On the other hand, by Itô’s formula $|x_t|^2$ turns out to be at least a local martingale and, since it is nonnegative and starts from zero, it is zero. Then

$$0 = x_t = w_t + \int_0^t b(x_t) \, dt,$$

that is the sum of $w_t$ and a function of bounded variation is zero, which is impossible.

2010 Mathematics Subject Classification. 60H10, 60H20, 60H30, 60J60.

Key words and phrases. Itô equations, weak uniqueness, higher summability of Green’s functions.
Thus, in the general case we can only hope that the existence of solutions of stochastic equations holds if \( b \in L_d \). To formulate a result which contains the existence theorem we introduce some necessary objects.

Introduce \( S \) as the set of \( d \times d \) symmetric matrices, and for \( \delta \in (0, 1) \) let \( S_\delta \) be the subset of \( S \) consisting of matrices whose eigenvalues are between \( \delta \) and \( \delta^{-1} \).

Let \( b(x), b^{(k)}(x), k = 1, 2, \ldots, \) be \( \mathbb{R}^d \)-valued Borel functions on \( \mathbb{R}^d \) such that, for a constant \( \| b \| < \infty \),

\[
\|b\|_{L_d(\mathbb{R}^d)}, \|b^{(k)}\|_{L_d(\mathbb{R}^d)} \leq \|b\|, \quad k = 1, 2, \ldots,
\]

and \( b^{(k)} \to b \) as \( k \to \infty \) in \( L_d(\mathbb{R}^d) \). Let \( a(x), a^{(k)}(x), k = 1, 2, \ldots, \) be Borel functions on \( \mathbb{R}^d \) with values in \( S_\delta \) for some \( \delta \in (0, 1) \) such that \( a^{(k)} \to a \) as \( k \to \infty \) (a.e.).

**Theorem 1.1.** Take \( x \in \mathbb{R}^d \). (i) There exists a probability space \((\Omega, \mathcal{F}, P)\), a filtration of \( \sigma \)-fields \( \mathcal{F}_t \subset \mathcal{F}, t \geq 0 \), a process \( w_t, t \geq 0 \), which is a \( d \)-dimensional Wiener process relative to \( \{\mathcal{F}_t\} \), and an \( \mathcal{F}_t \)-adapted process \( x_t \) such that (a.s.) for all \( t \geq 0 \)

\[
x_t = x + \int_0^t \sqrt{a(x_s)} \, dw_s + \int_0^t b(x_s) \, ds. \quad (1.1)
\]

(ii) Furthermore, let \( x^{(k)} \in \mathbb{R}^d, k = 0, 1, 2, \ldots, \) and let \( x^{(k)} \to x \) as \( k \to \infty \). Assume that for each \( k = 1, 2, \ldots \) there exists a probability space \((\Omega^{(k)}, \mathcal{F}^{(k)}, P^{(k)})\), a filtration of \( \sigma \)-fields \( \mathcal{F}_t^{(k)} \subset \mathcal{F}^{(k)}, t \geq 0 \), a process \( w_t^{(k)}, t \geq 0 \), which is a \( d \)-dimensional Wiener process relative to \( \{\mathcal{F}_t^{(k)}\} \), and an \( \mathcal{F}_t^{(k)} \)-adapted process \( x_t^{(k)} \) such that (a.s.) for all \( t \geq 0 \)

\[
x_t^{(k)} = x^{(k)} + \int_0^t \sqrt{a^{(k)}(x_s^{(k)})} \, dw_s^{(k)} + \int_0^t b^{(k)}(x_s^{(k)}) \, ds. \quad (1.2)
\]

Then the set of distributions of \( x^k \) on \( C([0, \infty), \mathbb{R}^d) \) is tight and any weakly converging subsequence of distributions converges weakly to the distribution of one of solutions of (1.1) as described in (i).

This theorem is proved in Section 3 by using Skorokhod’s embedding method.

Once the solvability of (1.1) is established, the question of its weak uniqueness arises. A standard way (but there are other ways as well) of treating it is the following when we argue formally without caring about rigorousness at the moment.

Introduce

\[
L = (1/2)a^{ij} D_{ij} + b^i D_i.
\]

Take smooth \( f, \lambda > 0, R > |x|, \) and find a bounded sufficiently regular solution of

\[
\lambda u - Lu = f \quad (1.3)
\]
in $B_R$ ($B_\infty := \mathbb{R}^d$) with zero boundary data (no boundary data if $R = \infty$). For $n = 0, 1, \ldots$, let $g_n(y_0, \ldots, y_n), y_k \in \mathbb{R}^d, k = 0, 1, \ldots, n$, be smooth bounded functions. Let $0 = t_0 \leq t_1 < t_2 < \cdots < \infty$. Use Itô’s formula to get that on the set $\{t_n \leq \tau\},$ where $\tau = \tau_R$ is the first exit time of $x_t$ from $B_R$,

$$u(x_{t\wedge \tau})e^{-t\wedge \tau} = u(x_{t_n})e^{-\lambda t_n} - \int_{t_n}^{t\wedge \tau} e^{-\lambda s} f(x_s) \, ds$$

$$+ \int_{t_n}^{t\wedge \tau} e^{-s} Du(x_s) \sqrt{a(x_s)} \, dw_s.$$

Take expectations of both sides multiplied by the indicator of $\{t_n < \tau\}$ times $g_n(x_{t_0}, \ldots, x_{t_n})$. The expectation containing the stochastic integral, naturally, disappears and after letting $t \to \infty$ we obtain

$$Eg_n(x_{t_0}, \ldots, x_{t_n})u(x_{t_n})I_{t_n < \tau}e^{-\lambda t_n} = \int_{t_n}^{\infty} e^{-\lambda s} Eg_n(x_{t_0}, \ldots, x_{t_n})f(x_s)I_{s < \tau} \, ds.$$ 

(1.4)

If the left-hand side is uniquely defined (that is, independent of which solutions $x_t$ we take) for any $\lambda > 0$ and smooth $f$, then, for $s \geq t_n$ and smooth $f$,

$$Eg_n(x_{t_0}, \ldots, x_{t_n})f(x_s)I_{s < \tau_R} \quad \text{and hence} \quad Eg_{n+1}(x_{t_0}, \ldots, x_{t_{n+1}})I_{t_{n+1} < \tau_R}$$

are uniquely defined. For $n = 0$ the left-hand side of (1.4) is $Eg_0u(x_0) = g_0u(x)$ that is independent of which solution we take. By induction this allows us to conclude that all quantities in (1.5) are uniquely defined for all $n$. Letting $R \to \infty$ (if the above $R < \infty$) yields weak uniqueness.

There are the following obstacles to implement this scheme if $b \in L_d$. To apply Itô’s formula we, generally, need $u \in W^2_{d}$ and $Lu \in L_d$. Then $b^iD_iu$ should be in $L_d$, but $b$ is only in $L_d$ and then, apparently, $Du$ needs to be bounded. However, by embedding theorems the fact that $u \in W^2_{d}$ does not imply that $Du$ is bounded. On the other hand, if $u \in W^2_p$ with $p < d$ and $b \in L_d$ then $b^iD_iu \in L_p$ and the above mismatch does not occur, but we need to know that Itô’s formula is applicable to $u \in W^2_p$ for some $p < d$.

However, if $b$ is bounded, Itô’s formula is applicable to $u \in W^2_{d-\varepsilon}$ for some $\varepsilon > 0$ (a consequence of a Fabes-Stroock result from [6] if $b \equiv 0$, which was carried over to bounded $b$ in [3] and to $b \in L_{d+\varepsilon}$ in [7]). We would be in business if we knew that this also holds if $b \in L_d$. Then weak uniqueness would follow from the solvability of (1.3) in $W^2_{d-\varepsilon}$ for small $\varepsilon > 0$.

For $R \in (0, \infty)$ introduce $0^2(B_R) = W^2_p(B_R) \cap \{u : u|_{\partial B_R} = 0\}$,

$$0^2(B_\infty) = W^2_p(\mathbb{R}^d),$$

$$L_0 = (1/2)\delta^{ij}D_{ij}.$$
Theorem 1.2. There is a \( d_0 = d_0(d, \delta, \|b\|) \in (d/2, d) \) such that, if \( p \in [d_0, d) \) and for some \( R, \lambda > 0 \), and any \( t \in [0, 1] \) and \( u \in \dot{W}^2_p(B_R) \) we have

\[
\|u\|_{\dot{W}^2_p(B_R)} \leq N \|\lambda u - (t\Delta + (1-t)L_0)u\|_{L^p(B_R)}
\]

with \( N \) independent of \( u \) and \( t \), then for any smooth \( f \) equation (1.3) has a unique solution in \( \dot{W}^2_p(B_R) \).

To the best of the author's knowledge the most general conditions on the coefficients \( a^{ij} \) when (1.6) holds with \( R = \infty \) (for \( p > d \)) is given in [14], where the solvability in \( W^2_p \) spaces is proved for second-order elliptic equations with coefficients which are measurable in one direction and VMO in the orthogonal directions in each small ball with the direction depending on the ball. In a subsequent article the author intends to show that this result also holds for \( p \geq d_0 \). Of course, we know from [2] that (1.6) holds for all \( p > 1 \) if \( a^{ij} \) are continuous. In that case weak uniqueness with bounded \( b \) is known from Stroock-Varadhan [23]. If \( a^{ij} \in VMO \), (1.6) for all \( p > 1 \) and \( R < \infty \) is proved in [4].

Most likely the conclusion of Theorem 1.2 is false if \( p \geq d \) even if \( a^{ij} = \delta^{ij} \).

If \( p < d \), the proof of Theorem 1.2 uses the fact that

\[
\int_{\mathbb{R}^d} b^p|Du|^p \, dx \leq \left( \int_{\mathbb{R}^d} b^d \, dx \right)^{p/d} \left( \int_{\mathbb{R}^d} |Du|^{pd/(d-p)} \, dx \right)^{(d-p)/d} \leq N \int_{\mathbb{R}^d} |D^2u|^p \, dx,
\]

where the last inequality follows from the embedding theorem saying that \( Du \in L_q \) if \( D^2u \in L_p \) and

\[ 2 - \frac{d}{p} = 1 - \frac{d}{q}, \quad q = \frac{pd}{d-p}. \]

It is worth saying that Ladyzhenskaya-Ural’tseva in [20] studied the case of \( b \in L_{d+\varepsilon} \) with \( p = 2 \) in (1.6) and continuous \( a^{ij} \). Actually in this situation the assumption of Theorem 1.2 is satisfied for any \( p \in (1, \infty) \) and hence the above described method of proving weak uniqueness works. But the case \( b \in L_d \) is excluded. In the classical book Gilbarg-Trudinger [10] integrable drifts are not treated.

The last ingredient in the above scheme of how to prove weak uniqueness on the basis of Theorem 1.2 is Itô’s formula. To state it we need some notation and assumptions used throughout the paper.

Let \( d_1 \) be a positive integer, \((\Omega, \mathcal{F}, P)\) be a complete probability space, and let \((w_t, \mathcal{F}_t)\) be a \(d_1\)-dimensional Wiener process on this space with complete, relative to \( \mathcal{F}, P, \sigma\)-fields \( \mathcal{F}_t \). Let \( \sigma_t, t \geq 0 \), be a progressively measurable process with values in the set of \( d \times d_1\)-matrices and let \( b_t, t \geq 0 \), be an \( \mathbb{R}^{d_1} \)-valued progressively measurable process. Assume that for any \( T \in [0, \infty) \)
and \( \omega \)
\[
\int_0^T (|\sigma_t|^2 + |b_t|) \, dt < \infty. \tag{1.7}
\]

Under this condition the stochastic process
\[
x_t = \int_0^t \sigma_s \, dw_s + \int_0^t b_s \, ds \tag{1.8}
\]
is well defined. Fix a nonnegative Borel \( \mathcal{B} \) on \( \mathbb{R}^d \) and \( \delta \in (0,1) \).

**Assumption 1.1.** We have \( \|b\| := \|b\|_{L_d(\mathbb{R}^d)} < \infty \) and
\[
|b_t| \leq b(x_t), \quad a_t \in S_\delta
\]for all \((\omega, t)\), where \( a_t = \sigma_t \sigma_t^* \).

Introduce
\[
L_t u(x_t) = (1/2) a_t^{ij} D_{ij}(x_t) + b_t^i D_i(x_t).
\]

**Theorem 1.3.** Under Assumption 1.1 there is a \( d_0 = d_0(d, \delta, \|b\|) \in (d/2, d) \) such that if \( p \geq d_0 \) and \( u \in W^2_p(\mathbb{R}^d) \), then (a.s.) for all \( t \geq 0 \)
\[
u(x_t) = u(0) + \int_0^t L_s u(x_s) \, ds + \int_0^t D_i u(x_s) \sigma_s^{ik} \, dw_s^k \tag{1.10}
\]
and the last term is a square integrable martingale.

This theorem is proved in Section 3.

The above results and the discussion after Theorem 1.1 immediately yield the following.

**Theorem 1.4.** In the setting of Theorem 1.1 suppose that for a \( p \in [d_0, d) \) and any \( R, \lambda \in (0, \infty), t \in [0, 1], u \in W^2_p(B_R) \) we have
\[
\|u\|_{W^2_p(B_R)} \leq N \|\lambda u - (t \Delta + (1-t)L_0) u\|_{L^p(B_R)}
\]
with \( N \) independent of \( u \) and \( t \) (may depend on \( \lambda, R, \ldots \)). Then solutions of (1.1) are weakly unique.

In the heart of the above results lies the following estimate.

**Theorem 1.5.** Under Assumption 1.1 there is a \( d_0 \in (d/2, d) \), depending only on \( d, \delta, \|b\| \), such that, for any \( \lambda > 0, p \geq d_0 \), and nonnegative Borel \( f(x) \) given on \( \mathbb{R}^d \) we have
\[
E \int_0^\infty e^{-\lambda t} \Psi_\lambda(x_t) f(x_t) \, dt \leq N \lambda^{d/(2p)-1} \|f\|_{L^p(\mathbb{R}^d)}. \tag{1.11}
\]
where \( N \) depends only on \( d, \delta, \) and \( \|b\| \), \( \Psi_\lambda(x) = \exp(\sqrt{\lambda} \nu |x|) \), \( \nu = \mu/4 \), and \( \mu \) is taken from Theorem 2.1.
This theorem is proved in Corollary 2.4.

The above results allows one to construct Markov diffusion processes corresponding to $L$. To show how to do this we need the following, which would be a simple consequence of Theorem 4.5.1 of [23] were $b$ supposed to be bounded.

**Lemma 1.6.** Let $a$ and $b$ be the same as before Theorem 1.1. Suppose that we are given a continuous process $x_t$, $t \geq 0$, such that $x_0 = 0$, for any $T \in (0, \infty)$

$$
\int_0^T |b(x_t)| \, dt < \infty
$$

(a.s.), and for any twice continuously differentiable function $u(x)$ with compact support the process

$$
u(x_t) - \int_0^t Lu(x_s) \, ds
$$

(1.12) is a local martingale with respect to the filtration of $\sigma$-fields $\{\mathcal{N}_t = \sigma(x_s; s \leq t), t \geq 0\}$. Then there exists a $d$-dimensional Wiener process $(w_t, \bar{N}_t)$, $t \geq 0$, where $\bar{N}_t$ is the completion of $\mathcal{N}_t$, such that (1.1) is satisfied with $x = 0$.

Proof. First observe that by using cut-off functions one easily shows that (1.12) is a local martingale for any twice continuously differentiable function $u(x)$. Then, we claim that the following processes are local martingales

$$X_t := x_t - \int_0^t b(x_s) \, ds,$$

$$x_t^* - \int_0^t (a(x_s) + b(x_s) x_s^* + x_s b^*(x_s)) \, ds,$$

$$A_t := X_t X_t^* - \int_0^t a(x_s) \, ds.$$

Indeed, the first two processes are obtained from (1.12) for $u = x, xx^*$. Concerning the last one introduce $\gamma_R$ as the minimum of $\tau_R$ and

$$\inf\{t \geq 0 : \int_0^t |b(x_s)| \, ds \geq R\}.$$  

Also let

$$\Phi_t = \int_0^t b(x_s) I_{s<\gamma_R} \, ds.$$  

Observe that $X_{t \wedge \gamma_R}$ and $\Phi_t$ are bounded and simple manipulations show that

$$A_{t \wedge \gamma_R} = \int_0^t X_{s \wedge \gamma_R} \, d\Phi_s^* - X_{t \wedge \gamma_R} \Phi_t^* + \int_0^t (d\Phi_s) X_s^{\gamma_R} - \Phi_t X_t^{\gamma_R},$$

which by the Lemma from Appendix 2 of [12] shows that $A_{t \wedge \gamma_R}$ is a martingale.
By the above claim the quadratic variation process of the local martingale $X_t$ is

$$\int_0^t a(x_s) \, ds.$$ 

After that our assertion follows directly from Theorem III.10.8 of [13]. The lemma is proved. □

Remark 1.1. We used a result from [13], where the initial definition (see there Definition II.8.2) of a martingale is different from commonly used and, owing to Doob’s optional stopping theorem, seemingly admits wider class of processes than the martingales in the usual sense. However, just considering stopping times taking only two values, one easily sees that, actually, martingales from [13] are martingales in the classical sense.

Theorem 1.7. Let $a$ and $b$ be as in Lemma 1.6. Then there exists a continuous strong Markov process $X = (x_t, \infty, M_t, P_x)$ (the terminology taken from [5]) in $\mathbb{R}^d$ such that for any $x \in \mathbb{R}^d$ and $t \geq 0$

$$E_x \int_0^t |b(x_s)| \, ds < \infty \quad (1.13)$$

and for any twice continuously differentiable function $u$ with compact support the process (1.12) is a local martingale relative to $P_x$. Furthermore, $(x_t, \infty, M_{t+}, P_x)$ is a Markov process.

Proof. The proof of this theorem follows the proof of Theorem 3 of [11] and we only point out the necessary changes related to the fact that, unlike [11] where $b$ is Borel bounded, our $b \in L_d$.

As in [11] we set $\Omega = \mathcal{C}([0, \infty), \mathbb{R}^d)$ and for $\omega = \omega : \in \Omega$ define $x_t(\omega) = \omega_t$. Also set $\mathcal{M}_t = \mathcal{N}_t = \sigma(x_s; s \leq t)$ and by $\Pi_x$ denote the set of probability measures $P$ on $(\Omega, \mathcal{N}_{\infty})$ such that $P(x_0 = x) = 1$ and the process (1.12) is a local martingale for any twice continuously differentiable function $u$. According to Theorem 1.1 and Theorem 1.5, assuring that (1.13) holds for solutions of (1.1), $\Pi_x \neq \emptyset$.

Owing to Lemma 1.6, Corollary 1.2 of [17] and Theorem 1.5 are applicable, that is, for $P \in \Pi_x$ and any $n \geq 0$

$$E \max_{r \in [s,t]} |x_r - x_s|^{2n} \leq N(t - s)^n, \quad (1.14)$$

where $N = N(n, d, \delta, \|b\|)$ and (1.11) holds with $N = N(d, \delta, \|b\|)$. In particular, the assertions of Lemmas 5 and 6 of [11] are valid. After that the proof goes the same way as in [11] once more using Theorem 1.1, this time its second statement, while proving that $\{\Pi_x\}$ is a $B$-system in the terminology of [11]. The theorem is proved. □

Remark 1.2. In a subsequent article we will show that the process $(x_t, \infty, M_{t+}, P_x)$ is strong Markov with strong Feller semigroup.
Theorem 1.7 provides existence of a Markov diffusion process corresponding to the operator $L$. One knows that, generally, this process is not unique in any sense. In this connection we present some results such as Corollary 4.2 which are the main tools in proving the Harnack inequality and Hölder continuity property of harmonic functions for the corresponding diffusion processes with drift in $L_d$.

We also deal with some issues from the theory of partial differential equations. For instance, Corollary 3.1, in particular, provides the maximum principle for elliptic equations with measurable $a^{ij}$ and drift in $L_d$ for solutions in $W^2_p$ with $p < d$ (in case $p = d$ this is a classical Aleksandrov’s result and, if $b$ is bounded, $p < d$ is allowed according to the results in [6], [7]). Theorem 3.2 is indispensable in the Sobolev space theory of fully nonlinear elliptic equations while studying the possibility to pass to the limit in such equations.

The rest of the article is organized as follows. In Section 2 we prove Theorem 1.5. This allows us to prove Theorems 1.1 and 1.3 in Section 3. Section 4 is devoted to studying fine properties of our processes such as estimating the time spent in sets of small measure, the probability to reach such sets, Fanghua Lin estimates playing a major role in the Sobolev space theory of fully nonlinear elliptic equations, boundary behavior of solutions of the corresponding elliptic equations with first order coefficients in $L_d$, and the probability to pass through narrow tubes, which in the first draft of the paper was one of cornerstones of everything else. We also prove the doubling property of the corresponding Green’s measures and the fact that their densities are in the class $A_{\infty}$ of Muckenhoupt.

We finish the introduction with some notation and the stipulation about constants. In the proofs of various results we use the symbol $N$ to denote finite nonnegative constants which may change from one occurrence to another and we do not always specify on which data these constants depend. In these cases the reader should remember that, if in the statement of a result there are constants called $N$ which are claimed to depend only on certain parameters, then in the proof of the result the constants $N$ also depend only on the same parameters unless specifically stated otherwise. Of course, if we write $N = N(...)$, this means that $N$ depends only on what is inside the parentheses. Another point is that when we say that certain constants depend only on such and such parameters we mean, in particular, that the dependence is such that these constants stay bounded as the parameters vary in compact subsets of their ranges.

Introduce $|\Gamma|$ as the volume of $\Gamma \subset \mathbb{R}^d$,

$$
\int_B f(x) \, dx = \frac{1}{|B|} \int_B f(x) \, dx, \quad a_\pm = a^{\pm} = (1/2)(|a| \pm a).
$$

Use the notation $u^{(\varepsilon)} = u \ast \zeta_{\varepsilon}$, where $\zeta_{\varepsilon}(x) = \varepsilon^{-d}\zeta(x/\varepsilon)$, $\varepsilon > 0$, and $\zeta$ is a nonnegative $C^\infty$-function with support in $B_1$ whose integral is equal to one.
If $B$ is a ball and $\eta$ is a positive number, by $\eta B$ we denote a concentric ball whose radius is $\eta$ times that of $B$.

If it is not explicitly stated otherwise, by $x_t$ we always mean the process defined by (1.8) and let $\tau_R(x)$ be the first exit time of $x + x_t$ from $B_R$ (equal to infinity if $x + x_t$ never exits from $B_R$). Also let $\tau_R = \tau_R(0)$.

2. Green’s functions

We suppose throughout the article that Assumption 1.1 is satisfied. Recall that Theorem 2.17 of [17] implies that if $p \geq d$, then there exists constants $N$ and $\mu > 0$, depending only on $d, p, \delta$, and $\|b\|$, such that for any $\lambda > 0$ and Borel nonnegative $f$ given on $\mathbb{R}^d$ we have

$$E \int_0^\infty e^{-\lambda t} f(x_t) dt \leq N \lambda^{-d/(2p) - 1} \|\Psi^{-1}_\lambda f\|_{L_p(\mathbb{R}^d)},$$

(2.1)

where $\Psi_\lambda(x) = \exp(\sqrt{\lambda \mu} |x|)$.

Here is a straightforward consequence of (2.1).

**Theorem 2.1.** Let $p \geq d$. Then there exists constants $N$ and $\mu > 0$, depending only on $d, \delta, p,$ and $\|b\|$, and for any $\lambda > 0$ there exists a nonnegative Borel function $G_\lambda(x)$ (Green’s function of $x$) on $\mathbb{R}^d$ such that for any Borel nonnegative $f$ given on $\mathbb{R}^d$ we have

$$E \int_0^\infty e^{-\lambda t} f(x_t) dt = \int_{\mathbb{R}^d} f(x) G_\lambda(x) dx,$$

$$\|\Psi_\lambda G_\lambda\|_{L_p/(p-1)(\mathbb{R}^d)} \leq N \lambda^{-d/(2p) - 1}.$$

(2.2)

The highest power of summability of $G_\lambda$ guaranteed by this theorem occurs when $p = d$ and this is $d/(d - 1)$. It turns out that, actually, $G_\lambda$ is summable to a higher power. The proof of this is based on Gehring’s lemma from [9], Aleksandrov’s estimates, and the following, which is a particular case of Lemma 2.13 in [17]

**Lemma 2.2.** There is a constant $N = N(d, \delta, \|b\|)$ such that for any $R \in (0, \infty)$

$$NE \int_0^{\tau_R \wedge R^2} e^{-t} dt \geq R^2 \wedge 1.$$  

(2.3)

**Theorem 2.3.** There are constants $\varepsilon \in (0, 1)$ and $N$, depending only on $d, \delta,$ and $\|b\|$, such that for any ball $B$ of radius $R \leq 1/2$ and $p \geq d_0 := d - \varepsilon$, we have

$$\|G_1\|_{L_p/(p-1)(B)} \leq NR^{-d/p} \|G_1\|_{L_1(2B)},$$

(2.4)

which is equivalently rewritten as

$$\left( \int_B G_1^{p/(p-1)} dx \right)^{(p-1)/p} \leq N \int_{2B} G_1 dx.$$
Proof. We basically follow the arguments in [6]. Take $R \in (0, 1/2]$, a closed ball $B$ of radius $R$ and let $B'$ be the concentric open ball of radius $2R$. Define recursively

$$
\gamma_1 = \inf \{ t \geq 0 : x_t \in B \}, \quad \tau_1 = \inf \{ t \geq \gamma_1 : x_t \not\in B' \},
\gamma_{n+1} = \inf \{ t \geq \tau_n : x_t \in B \}, \quad \tau_{n+1} = \inf \{ t \geq \gamma_{n+1} : x_t \not\in B' \}.
$$

Then for any nonnegative Borel $f$ vanishing outside $B$ with $\|f\|_{L^d(B)} = 1$ we have

$$
\int_B f G_1(x) \, dx = E \int_0^\infty e^{-t} f(x_t) \, dt
= \sum_{n=1}^\infty E e^{-\gamma_n} E \left( \int_{\gamma_n}^{\tau_n} e^{-(t-\gamma_n)} f(x_t) \, dt \mid F_{\gamma_n} \right).
$$

Next we use the conditional version of the Aleksandrov estimate to see that the conditional expectation above is less than $NR \|f\|_{L^d(B')} = NR$. After that we use the conditional version of (2.3) to get that

$$
R^2 \leq NE \left( \int_{\gamma_n}^{\tau_n} e^{-(t-\gamma_n)} \, dt \mid F_{\gamma_n} \right).
$$

Then we obtain

$$
\int_B f G_1(x) \, dx \leq NR^{-1} \sum_{n=1}^\infty E e^{-\gamma_n} E \left( \int_{\gamma_n}^{\tau_n} e^{-(t-\gamma_n)} \, dt \mid F_{\gamma_n} \right)
= NR^{-1} \sum_{n=1}^\infty E \int_{\gamma_n}^{\tau_n} e^{-t} \, dt
\leq NR^{-1} E \int_0^\infty e^{-t} I_{B'}(x_t) \, dt = NR^{-1} \int_{B'} G_1(x) \, dx.
$$

The arbitrariness of $f$ implies that

$$
\left( \int_B G_1^{d/(d-1)}(x) \, dx \right)^{(d-1)/d} \leq N \int_{B'} G_1(x) \, dx.
$$

Now the assertion of the theorem for $p = d_0$ follows directly from the corrected version of the famous Gehring’s lemma proved as Proposition 5.1 in [8]. For larger $p$ it suffices to use Hölder’s inequality. The theorem is proved.

Remark 2.1. Take $\nu \in \mathbb{R}$. Then the assertion of Theorem 2.3 remains true if we replace (2.4) with

$$
\|\Psi G_1\|_{L^p((p-1)/p)(B)} \leq NR^{-d/p}\|\Psi G_1\|_{L^1(2B)}, \quad (2.5)
$$

where $\Psi(x) = \exp(\nu|x|)$, and allow $N$ to depend also on $\nu$. This follows from the fact that the supremum of $\Psi$ over $B$ is less than a constant independent of $R \leq 1/2$ times the infimum of $\Psi$ over $2B$. 

\end{proof}

\begin{remark}

\end{remark}
Here is a substantial improvement of Theorem 2.1. Below and many times in the future we use self-similarity transformations like \( x_t \rightarrow cx_t/c^2 \), where \( c > 0 \) is a constant. This transformation changes \( \sigma_t \) and \( b_t \) in a well known way, which will bring about a new function \( b \) (see (1.9)). A remarkable fact is that this new \( b \) has the same \( L_d(\mathbb{R}^d) \)-norm as the original one.

**Corollary 2.4.** For any \( \lambda > 0 \) and \( p \geq d_0 \) we have
\[
\|\Psi\lambda G_\lambda\|_{L_p/(p-1)(\mathbb{R}^d)} \leq N\lambda^{d/(2p)-1},
\] (2.6)
where \( N \) depends only on \( p, d, \delta \), and \( \|b\| \), \( \Psi\lambda(x) = \exp(\sqrt{\lambda} \nu|x|) \), \( \nu = \mu/4 \), and \( \mu \) is taken from Theorem 2.1. In other words, for any nonnegative Borel \( f(x) \) given on \( \mathbb{R}^d \) estimate (1.11) holds:
\[
E \int_0^\infty e^{-\lambda t}\Psi\lambda(x_t)f(x_t)\,dt \leq N\lambda^{d/(2p)-1}\|f\|_{L_p(\mathbb{R}^d)}.
\]

Indeed, the case of arbitrary \( \lambda > 0 \) reduces to the one where \( \lambda = 1 \) by using self-similarity and for \( \lambda = 1 \) it suffices to note that, for \( q = p/(p-1) \),
\[
\int_{\mathbb{R}^d} e^{q\nu|x|}G_1^q(x)\,dx = N \int_{\mathbb{R}^d} \left[ \int_{|y-x| \leq 1} e^{q\nu|y|}G_1^q(y)\,dy \right] dx
\]
\[
\leq N \int_{\mathbb{R}^d} e^{-q\nu|x|} \left[ \int_{|y-x| \leq 1} e^{2q\nu|y|}G_1^q(y)\,dy \right] dx
\]
and then, to estimate the interior integral, use (2.5) and the fact that owing to (2.2) and Hölder’s inequality
\[
\int_{\mathbb{R}^d} e^{2\nu|x|}G_1^1(x)\,dx \leq N.
\] (2.7)

**Corollary 2.5.** For any \( x \in \mathbb{R}^d, p \geq d_0, \) and Borel nonnegative \( f \) vanishing outside \( B_1(x) \) we have
\[
E \int_0^\infty e^{-t}\int f(x_t)\,dt \leq N e^{-\nu|x|}\|f\|_{L_p(B_1(x))},
\]
where \( N \) depends only on \( p, d, \delta \), and \( \|b\| \) and \( \nu \) is the same as in Corollary 2.4.

**Theorem 2.6.** There is a constant \( N = N(p, d, \delta, \|b\|) \) such that for any \( n = 1, 2, \ldots, \) nonnegative Borel \( f \) on \( \mathbb{R}^d, T \in (0, \infty), \) and \( p \geq d_0 \) we have
\[
E \left[ \int_0^T f(x_t)\,dt \right]^n \leq n!N^nT^n[1-d/(2p)]\|\Psi_{1/T}f\|^n_{L_p(\mathbb{R}^d)},
\] (2.8)
where \( \Psi_{1/T} \) is taken from Corollary 2.4.

Proof. We are going to use the induction on \( n \). If the assertion of the theorem holds for some \( n \geq 1 \), then by using its conditional version and the fact that
\[
E \left[ \int_0^T f(x_t)\,dt \right]^{n+1} = (n+1)E \int_0^T f(x_t)E \left[ \int_t^T f(x_s)\,ds \right]^n \mid \mathcal{F}_t \right] dt,
\]
we see that the left-hand side is less than
\[
(n + 1)N^n \| \Psi^{-1}_1 f \|^{n}_{L^p(\mathbb{R}^d)} E \int_0^T f(x_t)(T - t)^{n[1 - d/(2p)]} dt
\]
\[
\leq (n + 1)!N^n T^{n[1 - d/(2p)]} \| \Psi^{-1}_1 f \|^{n}_{L^p(\mathbb{R}^d)} E \int_0^T f(x_t) dt.
\]
It follows that we only need to prove (2.8) for \( n = 1 \).

In that case denote the left-hand side of (2.8) by \( F(T) \) and observe that for any \( \lambda > 0 \) owing to Corollary 2.4 we have
\[
e^{-\lambda T} F(T) \leq \lambda \int_0^\infty F(t) e^{-\lambda t} dt
\]
\[
= E \int_0^\infty e^{-\lambda t} f(x_t) dt \leq N \lambda^{d/(2p) - 1} \| f \|_{L^p(\mathbb{R}^d)}.
\]
For \( \lambda = 1/T \) we get (2.8) with \( n = 1 \). The theorem is proved. \( \square \)

Next theorem improves Theorem 1.1 of [17] in what concerns the range of \( p \) for uniformly nondegenerate processes.

**Theorem 2.7.** There exists a constant \( N = N(p, d, \delta, \| b \|) \) such that, for any \( R > 0, p \geq d \), and Borel nonnegative \( f \) given on \( B_R \), we have
\[
E \int_0^\tau f(x_t) dt \leq NR^{2 - d/p} \| f \|_{L^p(B_R)}.
\]

Proof. Hölder’s inequality allows us to only concentrate on \( p = d \). Scalings show that we may assume that \( R = 1 \). Also we may assume that \( f \) is bounded and is zero outside \( B_1 \). In that case denote by \( \mathfrak{M} \) the set of stopping times \( \gamma \leq \tau := \tau_1 \), and set
\[
u_\gamma = E \left[ \int_\gamma^\tau f(x_t) dt \mid \mathcal{F}_\gamma \right], \quad \bar{\nu} = \operatorname{esssup}_{\gamma \in \mathfrak{M}} \nu_\gamma.
\]
Observe that for any \( \omega \) and \( \lambda > 0 \) it holds that
\[
\int_\gamma^\tau f(x_t) dt = \int_\gamma^\tau e^{-\lambda(t-\gamma)} f(x_t) dt + \lambda \int_\gamma^\infty e^{-\lambda(t-\gamma)} I_{\gamma \leq t < \tau} \left[ \int_t^\tau f(x_s) ds \right] dt.
\]
By the conditional version of (1.11) (recall that \( p = d \)) (a.s.)
\[
E \left[ \int_\gamma^\tau e^{-\lambda(t-\gamma)} f(x_t) dt \mid \mathcal{F}_\gamma \right]
\]
\[
\leq E \left[ \int_\gamma^\infty e^{-\lambda(t-\gamma)} f(x_t) dt \mid \mathcal{F}_\gamma \right] \leq N \lambda^{d/(2p) - 1} \| f \|_{L^p(B_1)}.
\]
Hence,
\[
u_\gamma \leq N \lambda^{d/(2p) - 1} \| f \|_{L^p(B_1)}
\]
\[
+ \lambda E \left[ \int_\gamma^\infty e^{-\lambda(t-\gamma)} I_{\gamma \leq t < \tau} E \left[ \int_t^\tau f(x_s) ds \mid \mathcal{F}_t \right] dt \mid \mathcal{F}_\gamma \right],
\]
where the last term is dominated by
\[ \lambda \bar{u} E \left[ \int_0^\gamma e^{-\lambda(t-\gamma)} I_{t \leq t < \tau} \, dt \mid \mathcal{F}_\gamma \right] \leq \lambda \bar{u} E \left[ \int_0^\gamma dt \mid \mathcal{F}_\gamma \right] \leq N_1 \lambda \bar{u} \]
(a.s.), where the last inequality follows from the conditional version of Corollary 2.1 of [17]. Thus, (a.s.)
\[ u_\gamma \leq N \lambda^{d/(2p)-1} \|f\|_{L_p(B_1)} + N_1 \lambda \bar{u}. \]
Since \( \gamma \) is arbitrary within \( \mathcal{M} \),
\[ \bar{u} \leq N \lambda^{d/(2p)-1} \|f\|_{L_p(B_1)} + N_1 \lambda \bar{u} \]  
(2.10)
(a.s.), and since \( \bar{u} < \infty \) (\( f \) is bounded), by taking \( \lambda = 1/(2N_1) \), we arrive at
\[ \bar{u} \leq N \|f\|_{L_p(B_1)}. \]
The theorem is proved. \( \square \)

Remark 2.2. Equation (2.10) implies that \( (p =) d_0 \geq d/2 \). Of course, the example of the Wiener process with no drift shows more than that, namely, \( d_0 > d/2 \).

We finish the section with a result which will be used in a subsequent article.

**Theorem 2.8.** Let \( p \geq d_0 \). Then there exists constants \( N \) and \( \mu > 0 \), depending only on \( d, p, \) and \( \|b\| \), and there exists \( R_0 = R_0(d, \|b\|) \geq 2 \), such that for any \( \lambda > 0 \), \( R \in [0, \infty) \), and Borel nonnegative \( f \) given on \( \mathbb{R}^d \) we have
\[ E \int_0^\infty e^{-\lambda \phi_t(B_0^R)} f(x_t) \, dt \leq N (R \sqrt{\lambda} + R_0)^{2-d/p} \lambda^{d/(2p)-1} \|\Psi_{R, \lambda}^{-1} f\|_{L_p(\mathbb{R}^d)}, \]
(2.11)
where \( \Psi_{R, \lambda}(x) = \exp \left( \sqrt{\lambda} \mu \text{dist} \left( x, B_{R+R_0/\sqrt{\lambda}} \right) \right) \) and
\[ \phi_t(B_0^R) = \int_0^t I_{|x_s| \geq R} \, ds. \]

This theorem looks very much like Theorem 2.18 of [17] proved for possibly degenerate processes for \( p \geq d \) rather than \( p \geq d_0 \). Theorem 2.8 is proved in the same way as Theorem 2.18 of [17] on the basis of sharper estimates of Green’s functions in the special case of uniformly nondegenerate processes. We only need to use again Lemma 2.8 of [17] on the basis of sharper estimates of Green’s functions in the special case of uniformly nondegenerate processes.

3. **Itô’s formula and solvability of stochastic equations with drift in \( L_d \)**

Recall that Assumption 1.1 is supposed to be satisfied throughout the article. First we deal with Itô’s formula.

**Proof of Theorem 1.3.** Since \( p > d/2 \) (see Remark 2.2), by embedding theorems, \( u \) is bounded and continuous. Furthermore, since \( p \geq d_0 \), by
embedding theorems, the $L_q$-norms of $|Du|$ over any ball of radius one are bounded by the same constant, where $q = d_0d/(d - d_0)$. Next since $2d_0 \leq q$ ($d_0 > d/2$), for any $\lambda > 0$ and $\Psi_{\lambda}$ from Corollary 2.4, it holds that $\Psi_{\lambda}^{-1} |Du|^2 \in L_{d_0}(\mathbb{R}^d)$. Therefore, by Theorem 2.6

$$E \int_0^T |Du(x_t)|^2 \, dt < \infty$$

for any $T \in (0, \infty)$, which proves that the stochastic integral in (1.10) is indeed a square integrable martingale.

We prove (1.10) by passing to the limit from smooth functions $u_n$ which converge to $u$ in $W^2_0(\mathbb{R}^d)$. In light of what is said in the previous paragraph, $u_n \to u$ uniformly in $\mathbb{R}^d$ and there is no difficulty to pass to the limit in the stochastic term. In the deterministic term there could be only one expression of concern

$$E \int_0^T |b_t||D(u_n - u)|(x_t) \, dt$$

which owing to the condition $|b_t| \leq b(x_t)$ and Theorem 2.6 is less than a constant independent of $n$ times

$$\|\Psi_{1/T}^{-1} bD(u_n - u)\|_{L_{d_0}(\mathbb{R}^d)}.$$

The latter by H"older’s inequality is estimated by

$$\|b\| \|\Psi_{1/T}^{-1} D(u_n - u)\|_{L_q(\mathbb{R}^d)}$$

where the last term, by embedding theorem, is less than a constant independent of $n$ times the $W^2_0(\mathbb{R}^d)$-norm of $u_n - u$ that tends to zero as $n \to \infty$. This proves the theorem. □

Next we deal with stochastic equations with drift in $L_d$.

**Proof of Theorem 1.1.** Having in mind mollifiers we see that assertion (ii) implies (i). The proof of (ii) is achieved by repeating the proof of Theorem 2.6.1 of [12] with only a few changes which we point out below. By Corollary 1.2 of [17] for any $m = 1, 2, \ldots, 0 \leq s \leq t$

$$E \max_{r \in [s,t]} |x^{(k)}_r - x^{(k)}_x|^{2m} \leq N(t - s)^m,$$

where $N$ is independent of $k$. This yields the tightness of distributions.

Then in exactly the same way as in the proof of Theorem 2.6.1 of [12], for any weakly converging sequence $\{k'\}$ of distributions of $x^{(k')}$ by using Skorokhod’s embedding theorem, we find a probability space $(\Omega, \mathcal{F}, P)$, $d$-dimensional Wiener processes $(\tilde{w}^{(k')}_t, \tilde{\omega}^{(k')}_t)$ defined on this space, and $\tilde{\mathcal{F}}^{(k')}_t$ adapted continuous processes $\tilde{x}^{(k')}_t$ such that, for some $(\tilde{x}_t, \tilde{\omega}_t)$ we have $(\tilde{x}^{(k')}_t, \tilde{\omega}^{(k')}_t) \to (\tilde{x}_t, \tilde{\omega}_t)$ in probability for any $t \geq 0$ and for any $k'$ it holds that with probability one for all $t \geq 0$

$$\tilde{x}^{(k')}_t = x^{(k')}_t + \int_0^t \sqrt{a^{(k')}(\tilde{x}^{(k')}_s)} \, d\tilde{w}^{(k')}_s + \int_0^t b^{(k')}(\tilde{x}^{(k')}_s) \, ds. \quad (3.1)$$
Observe that, in light of Theorem 2.6, as for \( \bar{x}^{(k')} \), we have that for any, first, continuous and, hence, for all Borel nonnegative \( f \),

\[
E \int_0^T f(\bar{x}_t) \, dt \leq NT^{1/2} \| \Psi_{1/T}^{-1} f \|_{L_d(\mathbb{R}^d)}.
\]

Then one passes to the limit in the first term on the right in (3.1) by literally repeating the corresponding part of the proof of Theorem 2.6.1 of [12].

In what concerns the second term, it suffices to observe that for any \( k_0 \)

\[
\lim_{k' \to \infty} \int_0^T |b^{(k')}(\bar{x}^{(k')}_t) - b(\bar{x}_t)| \, dt \leq \lim_{k' \to \infty} \int_0^T |b^{(k')}(\bar{x}^{(k')}_t) - b^{(k_0)}(\bar{x}^{(k')}_t)| \, dt
\]

\[
+ E \int_0^T |b^{(k_0)}(\bar{x}_t) - b(\bar{x}_t)| \, dt \leq N \lim_{n' \to \infty} \| b^{(k')} - b^{(k_0)} \|_{L_d(\mathbb{R}^d)}
\]

\[
= N \| b - b^{(k_0)} \|_{L_d(\mathbb{R}^d)},
\]

where the constants \( N \) are independent of \( k_0 \), which after sending \( k_0 \to \infty \) shows that the first expression is zero and this, as in the proof of Theorem 2.6.1 of [12], allows us to finish the present proof. The theorem is proved. □

Introduce \( L(\delta, \| b \|) \) as the set of operators

\[
L = (1/2)a^{ij} D_{ij} + b^i D_i,
\]

where \( (a^{ij}) \) is an \( \mathbb{S}_\delta \)-valued Borel function on \( \mathbb{R}^d \) and \( b = (b^i) \) is an \( \mathbb{R}^d \)-valued Borel function on \( \mathbb{R}^d \) such that \( \| b \|_{L_d(\mathbb{R}^d)} \leq \| b \| \).

Here is a generalization for uniformly nondegenerate processes of the famous Lemma 8 of Aleksandrov [1] for functions in \( W_p^2 \) with \( p \) that could be \( < d \). This result for bounded \( b \) is found in [3] and for \( b \in L_{d+\varepsilon}(\Omega) \) in [7].

**Corollary 3.1.** Let \( p \in [d_0, \infty) \), \( D \) be a bounded domain in \( \mathbb{R}^d \), and \( u \in W_p^2 \cap C(\bar{D}) \). Let \( c \) be a nonnegative measurable function on \( D \) and let \( L \in L(\delta, \| b \|) \). Then in \( D \)

\[
u \leq N \| (Lu - cu)_- \|_{L_p(\partial D)} + \sup_{\partial D} u_+,
\]

where \( N \) depends only on \( p, d, \delta, \| b \| \), and the diameter of \( D \).

Proof. It suffices to prove (3.2) only in \( D' = D \cap \{ u > 0 \} \) assuming that this domain is not empty. Then (3.2) will become stronger if we replace \( D \) with \( D' \). Also observe that on \( D' \) we have \( (Lu)_- \leq (Lu - cu)_- \) since \( cu \geq 0 \).

It follows that it suffices to prove (3.2) in case that \( u \geq 0 \) and \( c \equiv 0 \) on \( D \). Having in mind obvious approximation of \( D \) from inside with smooth domains and extending \( u \) outside of approximating domains, we may also assume that \( D \) is smooth and \( u \in W_p^2(\mathbb{R}^d) \).

In that case, by Theorem 1.1, for any \( x \in D \) we can find a solution \( x_t \) of equation (1.1). Below in the proof by \( x_t \) we mean this solution. In light of
Theorem 1.3, for any $T \in (0, \infty)$,

$$u(x) = -E \int_0^{\tau \wedge T} Lu(x_t) \, dt + Eu(x_{\tau \wedge T}),$$

where $\tau$ is the first exit time of $x_t$ from $D$. Since $E\tau < \infty$ and (cf. (2.9))

$$E \int_0^{\tau} I_D(x_t) |Lu(x_t)| \, dt \leq N\|Lu\|_{L_0(D)} < \infty,$$

we can pass to the limit in (3.3) as $T \to \infty$ and obtain

$$u(x) = -E \int_0^{\tau} Lu(x_t) \, dt + Eu(x_{\tau}) \leq E \int_0^{\tau} (Lu(x_t)) - \sup_{\partial D} u.$$

After that it only remains to use (2.9) again. \qed

Here is another consequence of Itô’s formula and our previous results for elliptic equations. Such results play a crucial role in the theory of fully nonlinear elliptic equation providing a tool allowing to pass to the limit under the sign of a nonlinear operator when there is no convergence of the derivatives of the functions to which the operator is applied (see, for instance, Section 4.2 in [16]).

The following result for $p \geq d$ and $R = \infty$ is obtained in [15], however, with $N$ in (3.4) depending on how fast $\|(|b| - \mu) + \|b\|_{L_d(\mathbb{R}^d)} \to 0$ as $\mu \to \infty$.

**Theorem 3.2.** Let $p \geq d_0$, $R \in (0, \infty]$, and $L \in \mathcal{L}(\delta, \|b\|)$. Then there exists a constant $N = N(p, d, \delta, \|b\|)$ such that for any $\lambda > 0$ and $u \in W^{2,p}_{p,\text{loc}}(B_R) \cap C(\overline{B_R})$ ($B_\infty = \mathbb{R}^d$, $C(\mathbb{R}^d)$ is the set of bounded continuous functions on $\mathbb{R}^d$) we have

$$\lambda u_+ \|u\|_{L_p(B_{R/2})} \leq N\|\lambda u - Lu\|_{L_p(B_R)} + N\lambda R^{d/p}e^{-R\sqrt{\lambda}/N} \sup_{\partial B_R} u_+, \quad (3.4)$$

where the last term should be dropped if $R = \infty$.

For $p \geq d$ this theorem is proved in [17] (see Theorem 3.1 there). The proof from [17] carries over to our present situation almost word for word.

4. Estimates of the time spent in sets of small measure

Here is the first result of this section, which will be proved after some discussion.

**Theorem 4.1.** For any $\kappa \in (0, 1)$ there exist constants $\mu \geq 1$ and $N$, depending only on $d, \delta, \|b\|,$ and $\kappa$, such that, for any $R \in (0, \infty)$, $x \in B_{\kappa R}$, and Borel set $\Gamma \subset B_R$, the expected time that $x + x_t$ spends in $\Gamma$ before exiting from $B_R$ is greater than or equal to $N^{-1}R^2(|\Gamma|/|B_R|)^{\mu}$:

$$E \int_0^{\tau_R(x)} I_\Gamma(x + x_t) \, dt \geq N^{-1}R^2(|\Gamma|/|B_R|)^{\mu}. \quad (4.1)$$

The results of the kind which follows are commonly used while establishing the Hölder continuity of harmonic functions for diffusion processes or elliptic operators.
Corollary 4.2. For any $\kappa \in (0, 1)$ there exists a constant $N = N(d, \delta, \|b\|, \kappa)$ such that, for any $R \in (0, \infty)$, $x \in B_{\kappa R}$, and closed set $\Gamma \subset B_R$, the probability that $x + x_t$ reaches $\Gamma$ before exiting from $B_R$ is greater than or equal to $N^{-1}(|\Gamma|/|B_R|)^{\mu-1/d}$:

$$P(\tau(x) < \tau_R(x)) \geq N^{-1}(|\Gamma|/|B_R|)^{\mu-1/d},$$

where $\tau(x)$ is the first time $x + x_t$ hits $\Gamma$ and $\mu$ is taken from Theorem 4.1.

Indeed, set $\gamma = |\Gamma|/|B_R|$ and observe that by Theorem 4.1 and the conditional version of Theorem 1.1 of [17]

$$R^2\gamma^\mu \leq NEI_{\tau(x)} < \tau(x) E\left[ \int_{T_{\tau(x)}} I_{\Gamma}(x + x_t) \, dt \mid F_{\tau(x)} \right]$$

$$\leq NEI_{\tau(x)} R|\Gamma|^{1/d} = NR^2\gamma^{1/d}P(\tau(x) < \tau_R(x)).$$

One more consequence of Theorem 4.1 is the following which will be used in a subsequent article to show that diffusion processes with drift in $L_d$ are strong Markov with strong Feller semigroup.

Corollary 4.3. For $R, \kappa, x$, and $\Gamma$ as in the theorem set

$$\phi_t(\Gamma)(x) = \int_0^t I_{\Gamma}(x + x_s) \, ds.$$

Then there exists $\theta > 0$, depending only on $d, \delta, \|b\|$, and $\kappa$, such that

$$P(\phi_{\tau_R(x)}(\Gamma)(x) \geq \theta \gamma^\mu R^2) \geq N^{-1}_1 \gamma^{2\mu},$$

(4.3)

where $\gamma = |\Gamma|/|B_R|$, $N = N(d, \delta, \|b\|, \kappa)$, and $\mu$ is the same as in Theorem 4.1.

Indeed, (4.1) implies that for any $\theta > 0$ and $\phi := \phi_{\tau_R(x)}(\Gamma)(x),

$$N^{-1}_1 R^2 (|\Gamma|/|B_R|)^{\mu} \leq \theta \gamma^\mu R^2 P(\phi \leq \theta \gamma^\mu R^2) + E\phi I_{\phi \geq \theta \gamma^\mu R^2}$$

$$\leq \theta \gamma^\mu R^2 + P^{1/2}(\phi \geq \theta \gamma^\mu R^2) E^{1/2} \tau_R(x)$$

$$\leq \theta \gamma^\mu R^2 + N_2 R^2 P^{1/2}(\phi \geq \theta \gamma^\mu R^2),$$

where the last inequality is obtained by Lemma 2.6 of [17]. We get (4.3) for $\theta = (1/2)N_1^{-1}$.

Corollary 4.4. For any $R \in (0, \infty), \kappa \in (0, 1), x \in B_{\kappa R}$, and Borel non-negative $f$

$$\int_{B_R} f^{1/(2\mu)}(y) \, dy \leq NR^{-d-1/\mu} \left( E \int_0^{\tau_R(x)} f(x + x_t) \, dt \right)^{1/(2\mu)},$$

where $N = N(d, \delta, \|b\|, \kappa)$. 
Indeed, without losing generality assuming that \( f = 0 \) outside \( B_R \) and setting
\[
u := E \int_0^{\tau_R(x)} f(x + x_t) \, dt,
\]
we have that for any \( \lambda > 0 \)
\[
u \geq \lambda E \int_0^{\tau_R} I_{f(x + x_t) \geq \lambda} \, dt \geq \lambda N^{-1} R^2 (|\{ f \geq \lambda \}|/|B_R|)^\mu,
\]
\[
|\{ f \geq \lambda \}| \leq N R^{-2/\mu} \lambda^{-1/\mu} |B_R| u^{1/\mu}.
\]
It follows that for any \( c > 0 \)
\[
\int_{B_R} f^{1/(2\mu)}(y) I_{f \leq c} \, dy = \left(1/(2\mu)\right) \int_c^\infty \lambda^{1/(2\mu) - 1} \left|\{ f(y) > \lambda \}\right| \, d\lambda 
\]
\[
\leq N R^{-2/\mu} |B_R| u^{1/\mu} c^{-1/(2\mu)}.
\]
Also
\[
\int_{B_R} f^{1/(2\mu)}(y) I_{f > c} \, dy \leq c^{1/(2\mu)} |B_R|.
\]
For \( c = u R^{-2} \) we have
\[
R^{-2/\mu} |B_R| u^{1/\mu} c^{-1/(2\mu)} = c^{1/(2\mu)} |B_R|,
\]
\[
\int_{B_R} f^{1/(2\mu)}(y) \, dy \leq N u^{1/(2\mu)} R^{d-1/\mu}.
\]
This is what is claimed.

Another corollary is a generalization of the Fanghua Lin estimate for operators with summable drift which is one of the main tools in the Sobolev space theory of fully nonlinear elliptic equations (see, for instance, [16]).

**Theorem 4.5.** Let \( R \in (0, \infty), \; p \in [d_0, \infty), \; u \in W^{2,1}_{p, \text{loc}}(B_R) \cap C(\overline{B_R}), \; L \in \mathcal{L}(\delta, \| b \|), \; \) and \( c \in L_{d_0}(B_R), \; c \geq 0. \) Then
\[
\left(\int_{B_R} |D^2 u|^{1/(2\mu)} \, dx\right)^{2\mu} \leq N \left(\int_{B_R} |Lu - cu|^p \, dx\right)^{1/p} + NR^{-2} \sup_{\partial B_R} |u|, \tag{4.4}
\]
where \( \mu \) is taken from Theorem 4.1 with \( \kappa = 1/2 \) and \( N \) depends only on \( d, \delta, \| b \|, p, \) and \( R^{2-d/d_0} \| c \|_{L_{d_0}(B_R)}. \)

Proof. Hölder’s inequality allows us to concentrate on the case of \( p = d_0. \)

On the account of moving \( R, \) we may assume that \( u \in W^2_p(B_R) \) and then, by using scaling, that \( R = 1. \) After that we observe that
\[
\|Lu\|_{L_p(B_1)} \leq \|Lu - cu\|_{L_p(B_1)} + \|c\|_{L_p(B_1)} \sup_{B_1} |u|
\]
\[
\leq (1 + N \| c \|_{L_p(B_1)}) \|Lu - cu\|_{L_p(B_1)} + \|c\|_{L_p(B_1)} \sup_{\partial B_1} |u|.
\]
and reduce the case of general \( c \) to the one with \( c \equiv 0 \). In that case, it is easy to see that for sufficiently small \( \varepsilon = \varepsilon(d, \delta) > 0 \) there is an operator \( L' \in L(\delta/2, \|b\|) \) such that for our function \( u \) we have

\[
L'u = Lu + \varepsilon |D^2u|.
\]

Then, if \( x'_t \) is the process corresponding to \( L' \) and starting at the origin by Itô's formula we get (cf. the proof of Corollary 3.1)

\[
\varepsilon E \int_0^\tau |D^2u|(x'_t) \, dt = -u(0) - E \int_0^\tau Lu(x'_t) \, dt + Eu(x'_\tau),
\]

where \( \tau \) is the first exit time of \( x'_t \) from \( B_1 \). After that it only remains to use Corollary 4.4, Corollary 3.1, and Theorem 2.7. \( \square \)

**Remark 4.1.** It is standard that Corollary 4.4 implies not only (4.4) but also a similar estimate in half balls and similar estimates for \( |Du| \) (see, for instance, Section 9.4 in [16]).

To prove Theorem 4.1 we need three lemmas. In their proofs we, actually, translate into probability language the arguments from [19] reproduced, for instance, in [16]. In turn, the arguments in [19] have their origin in [18] written in probability language. We start with the following.

**Lemma 4.6.** Let \( \kappa \in (0, 1) \). Then there is a constant \( \xi = \xi(\kappa, d, \delta, \|b\|) \in (0, 1) \) such that for any \( R \in (0, \infty) \), Borel set \( \Gamma \subset B_R \) satisfying \( |\Gamma| \geq \xi |B_R| \), and \( x \in B_{\kappa R} \) we have

\[
E \int_{\tau_R(x)}^\tau I_\Gamma(x + x_t) \, dt \geq \nu E \tau_R(x),
\]

where \( \nu = \nu(d, \delta, \|b\|, \kappa) \in (0, 1) \).

**Proof.** Fix \( x \) with \( |x| \leq \kappa R \) and define \( \gamma \) as the first exit time of \( x + x_t \) from \( B_{2R}(x) \). By Corollary 2.12 of [17] and Theorem 2.7

\[
E \tau_R(x) - E \int_{0}^{\tau_R(x)} I_\Gamma(x + x_t) \, dt = E \int_{0}^{\tau_R(x)} I_{B_R \setminus \Gamma}(x + x_t) \, dt
\]

\[
\leq E \int_{0}^{\gamma} I_{B_R \setminus \Gamma}(x + x_t) \, dt \leq NR |B_R \setminus \Gamma|^{1/d}
\]

\[
= NR^2 \left( 1 - |\Gamma|/|B_R| \right)^{1/d} \leq NE \tau_R(x)(1 - \xi)^{1/d},
\]

where the constants \( N \) depend only on \( \kappa, d, \delta, \) and \( \|b\| \). We see how to choose \( \xi \) to satisfy (4.5) with a \( \nu = \nu(d, \delta, \|b\|, \kappa) \in (0, 1) \). The lemma is proved. \( \square \)

Next, we need a fact from the geometric measure theory. Take \( R \in (0, \infty) \), \( \zeta \in (0, 1) \), and a Borel set \( \Gamma \subset B_R \) such that \( |\Gamma| < \zeta |B_R| \). Define \( \mathfrak{M} \) as the collection of all open balls \( B \subset B_R \) such that

\[
|\Gamma \cap B| \geq \zeta |B|.
\]
Observe that this collection is nonempty and the union of its elements
\[ \Gamma' = \bigcup_{B \in \mathfrak{A}} B. \]
contains almost any point of \( \Gamma \) since almost any point of \( \Gamma \) is its density point. Also observe that \( \Gamma' \) is an open set, since all the \( B' \)’s are open.

Then for \( \varepsilon \in (0,1) \) denote by \( \mathfrak{A}_\varepsilon \) the set of \( B \in \mathfrak{A} \) such that \( |B| \geq \varepsilon \). Finally, recall that if \( B \) is an open ball and \( \kappa \in (0,1) \), we write \( \kappa B \) for the concentric open ball of radius \( \kappa \) times that of \( B \) and set
\[ \Gamma'_\kappa = \bigcup_{B \in \mathfrak{A}} \kappa B, \quad \Gamma'_{\kappa,\varepsilon} = \bigcup_{B \in \mathfrak{A}_\varepsilon} \kappa B. \]

**Lemma 4.7.** 1. We have \( |\Gamma \setminus \Gamma'| = 0 \) and
\[ |\Gamma'| \geq \left( 1 + \frac{1 - \zeta}{3d} \right)|\Gamma|. \]
2. There exists \( \kappa = \kappa(d, \zeta) \in (0,1) \) and \( \theta = \theta(d, \zeta) > 1 \) such that for all sufficiently small \( \varepsilon > 0 \) there exists a closed set \( \Gamma''_\varepsilon \subset \Gamma'_{\kappa,\varepsilon} \) such that
\[ |\Gamma''_\varepsilon| \geq \theta|\Gamma|. \]

**Proof.** The first assertion, a parabolic version of which is found in [19], is proved in Lemma 1.1 of [21]. To prove the second one it suffices to observe that, obviously, \( \Gamma'_{\kappa,\varepsilon} \uparrow \Gamma'_\kappa \) as \( \varepsilon \downarrow 0 \) and, similarly to Lemma 2.4 of [19], \( |\Gamma'_\kappa| \geq \kappa^d|\Gamma'|. \) The lemma is proved. \( \square \)

Since \( \Gamma''_\varepsilon \subset \Gamma'_{\kappa,\varepsilon} \), the closed set \( \Gamma''_\varepsilon \) is covered by the family \( \left\{ \kappa B : B \in \mathfrak{A}_\varepsilon \right\} \).

Then there is finitely many \( B(1), ..., B(n) \in \mathfrak{A}_\varepsilon \) such that
\[ \Gamma''_\varepsilon \subset \bigcup_{i=1}^n \kappa B(i) =: \Pi_\varepsilon. \]

Then for \( x \in \Pi_\varepsilon \) define \( i(x) \) as the first \( i \in \{1, ..., n\} \) for which \( x \in \kappa B(i) \).

Also set \( B(0) = B_R \) and \( i(x) = 0 \) if \( x \in \partial B_R \).

Now define recursively \( \gamma^0 = 0 \), \( \tau^1 \) as the first time after \( \gamma^0 \) when \( x_t \) exits from \( B_R \setminus \Gamma''_\varepsilon \), \( \gamma^1 \) as the first time after \( \tau^1 \) when \( x_t \) exits from \( B(i(x_{\tau^1})) \), and generally, for \( n = 2, 3, ... \) define \( \tau^n \) as the first time after \( \gamma^{n-1} \) when \( x_t \) exits from \( B_R \setminus \Gamma''_\varepsilon \), \( \gamma^n \) as the first time after \( \tau^n \) when \( x_t \) exits from \( B(i(x_{\tau^n})) \). It is easy to check that so defined \( \tau^n \) and \( \gamma^n \) are stopping times and, since \( |B(i)| \geq \varepsilon \) and the trajectories of \( x_t \) are continuous, \( \tau^n \uparrow \tau_R \) as \( n \to \infty \). Furthermore, since (a.s.) \( \tau_R \) is finite, (a.s.) all the \( \tau^n \)’s equal \( \tau_R \) for all large \( n \).

Now in the above general constructions we set \( \zeta = \xi \), where \( \xi \) is taken from Lemma 4.6.

**Lemma 4.8.** Set \( \zeta = \xi \), where \( \xi \) is taken from Lemma 4.6. Then for \( x \) such that \( |x| \leq \kappa R \) we have
\[ E \int_0^{\tau_R(x)} I_{\Gamma}(x + x_t) \, dt \geq \nu E \int_0^{\tau_R(x)} I_{\Gamma''_\varepsilon}(x + x_t) \, dt, \quad (4.7) \]
where \( \nu \) is taken from Lemma 4.6.
Proof. By the conditional version of Lemma 4.6 (a.s.)
\[
E\left[\int_0^{\gamma k} I_{\gamma y}(x + x_t) dt \mid \mathcal{F}_{\tau k}\right] \leq E\left[\int_0^{\gamma k} I_{B(i(x,k))}(x + x_t) dt \mid \mathcal{F}_{\tau k}\right]
\]
\[
\leq \nu^{-1}E\left[\int_0^{\gamma k} I_{\Gamma \cap B(i(x,k))}(x + x_t) dt \mid \mathcal{F}_{\tau k}\right] \leq \nu^{-1}E\left[\int_0^{\gamma k} I_{\Gamma}(x + x_t) dt \mid \mathcal{F}_{\tau k}\right].
\]
Hence,
\[
E\int_0^{\tau_R(x)} I_{\gamma y}(x + x_t) dt = \sum_{k=1}^{\infty} E\int_0^{\gamma k} I_{\gamma y}(x + x_t) dt
\]
\[
\leq \nu^{-1}\sum_{k=1}^{\infty} E\int_0^{\gamma k} I_{\Gamma}(x + x_t) dt \leq \nu^{-1}E\int_0^{\tau_R(x)} I_{\Gamma}(x + x_t) dt.
\]
This proves the lemma.

Proof of Theorem 4.1. Take \( \xi \) from Lemma 4.6. The results of Lemma’s 4.7 and 4.8 can be summarized as follows: For \( x \in \kappa B_R \), if a measurable \( \Gamma \subset B_R \) is such that \(|\Gamma| < \xi|B_R|\), then there exists a closed set \( \Gamma_1 \subset B_R \) such that \(|\Gamma_1| \geq \theta|\Gamma|\), where \( \theta = \theta(d, \xi) > 1 \) is taken from Lemma 4.7, and
\[
E\int_0^{\tau_R(x)} I_{\Gamma}(x + x_t) dt \geq \nu E\int_0^{\tau_R(x)} I_{\Gamma_1}(x + x_t) dt
\]
with \( \nu \) from Lemma 4.8. Naturally, if \(|\Gamma_1| < \xi|B_R|\), which only happens if \(|\Gamma| \leq (\xi/\theta)|B_R|\), then there exists a closed set \( \Gamma_2 \subset B_R \) such that \(|\Gamma_2| \geq \theta|\Gamma_1| \geq \theta^2|\Gamma|\) and
\[
E\int_0^{\tau_R(x)} I_{\Gamma}(x + x_t) dt \geq \nu E\int_0^{\tau_R(x)} I_{\Gamma_1}(x + x_t) dt \geq \nu^2 E\int_0^{\tau_R(x)} I_{\Gamma_2}(x + x_t) dt.
\]
We continue in a natural way and see that, if \( n \) is such that \(|\Gamma_n| < \xi|B_R|\), which only happens if \(|\Gamma| \leq (\xi/\theta^n)|B_R|\), then there exists a closed set \( \Gamma_{n+1} \subset B_R \) such that \(|\Gamma_{n+1}| \geq \theta^{n+1}|\Gamma|\) and
\[
E\int_0^{\tau_R(x)} I_{\Gamma}(x + x_t) dt \geq \nu^{n+1} E\int_0^{\tau_R(x)} I_{\Gamma_{n+1}}(x + x_t) dt.
\]
Let \( n_0 \) be the largest \( n \) for which the construction of \( \Gamma_{n+1} \) is still possible, that is \(|\Gamma_{n_0}| < \xi|B_R|\) and \(|\Gamma_{n_0+1}| \geq \xi|B_R|\). Since \(|\Gamma_n| \geq \theta^n|\Gamma|\), we have
\[
n_0 \leq \left\lfloor \left( \ln \left( |B_R|/|\Gamma| \right) / \ln \theta \right) \right\rfloor.
\]
Since by Lemma 4.6
\[
E\int_0^{\tau_R(x)} I_{\Gamma_{n_0+1}}(x + x_t) dt \geq \nu E\tau_R(x),
\]
we have
\[
E\int_0^{\tau_R(x)} I_{\Gamma}(x + x_t) dt \geq \nu^{n_0+2} E\tau_R(x).
\]
We take into account that by Corollary 2.12 of [17] \( E \tau_R(x) \geq N_1^{-1} R^2 \) and come to (4.1) with

\[
\mu = -((\ln \nu)/\ln \theta, \quad N = N_1 \nu^{-3}.
\]

This takes care of the case in which \(|\Gamma| < \xi |B_R|\). To include the case \(|\Gamma| \geq \xi |B_R|\) it suffices to increase the above \( N \) in an obvious way. The theorem is proved. \( \square \)

**Corollary 4.9.** Let \( R \in (0, \infty), \gamma \in (0, 1), \) and assume that a closed set \( \Gamma \subseteq B_R \) is such that, for any \( r \in (0, R), \) \(|B_r \cap \Gamma| \geq \gamma |B_r|\). Then there exist constants \( \alpha \in (0, 1) \) and \( N, \) depending only on \( d, \delta, \|b\|, \) and \( \gamma, \) such that, for any \( x \in B_{R/2}, \)

\[
P(\tau_R(x) < \tau^*_R(x)) \leq N(x/R)^\alpha. \tag{4.8}
\]

Indeed, let \( R^n = R 2^{-n}, n = 0, 1, \ldots, \) Then by Corollary 4.2, for \(|x| \leq R_{n+1}, \) \( P(A_n) \leq q = q(d, \delta, \|b\|, \gamma) < 1. \) The conditional version of this says that on the set \( A_{n+1} \) we have (a.s.) \( P(A_n \mid \mathcal{F}_{R_{n+1}}) \leq q. \) Since, for \(|x| \leq R/2, \) \( A_0 = \bigcap_{n=0}^{n(x)} A_n, \) where \( n(x) = \lfloor \ln_2(R/x) \rfloor - 1 \) \((\geq 0), \) we have

\[
P(A_0) \leq q^{n(x)+1} \leq q^{-1}(x/R)^{\ln_2(1/q)},
\]

which is just a different form of (4.8).

The following result will be used in a subsequent article on fully nonlinear elliptic equations with singular lower order terms.

**Theorem 4.10.** Let \( D \) be a bounded domain in \( \mathbb{R}^d, \) \( 0 \in \partial D, \) and assume that for some constants \( \rho, \gamma > 0 \) and any \( r \in (0, \rho) \) we have \(|B_r \cap D^c| \geq \gamma |B_r|\). Then there exists \( \beta = \beta(d, \delta, \|b\|, \gamma) > 0 \) such that, for any nonnegative \( f \in L_{d_0}(D) \) and \( x \in D, \)

\[
u(x) := E \int_0^{\tau(x)} f(x + x_t) dt \leq N|x|^{\beta \|f\|_{L_{d_0}(D)}}, \tag{4.9}
\]

where \( \tau(x) \) is the first exit time of \( x + x_t \) from \( D \) and \( N \) depends only on \( d, \delta, \|b\|, \gamma, \rho, \) and the diameter of \( D. \)

Proof. In light of Theorem 2.7 we may concentrate on \( x \in B_{\rho/2} \) with \(|x| \leq 1. \) The conditional version of this theorem allows us to write that, for \( 2|x| \leq r < \rho \) and \( \tau^r(x) \) being the first exit time of \( x + x_t \) from \( B_r \cap D, \)

\[
u(x) = E \int_0^{\tau^r(x)} f(x + x_t) dt + E I_{\tau^r(x) < \tau(x)} E \left( \int_0^{\tau(x)} f(x + x_t) dt \mid \mathcal{F}_{r_t(x)} \right)
\]

\[
\leq N r^{2-d/(2p)} \|f\|_{L_{d_0}(D)} + N \|f\|_{L_{d_0}(D)} P(\tau^r(x) < \tau(x)).
\]

Observe that \( \{\tau^r(x) < \tau(x)\} \subseteq \{\tau^x < \tau^r(x)\}, \) where \( \Gamma_r = \bar{B}_r \cap D^c, \) and by Corollary 4.9 we have \( P(\tau^r(x) < \tau(x)) \leq N(x/r)^\alpha. \) Thus, for any \( r \in \{2|x|, \rho\} \)

\[
u(x) \leq N \|f\|_{L_{p}(D)} (r^{2-d/(2d_0)} + (x/r)^\alpha).
\]
By choosing \( r = \sqrt{2|x|/\beta} \), we get the result with \( \beta = \alpha/2 \) since \( 1 - d/(4d_0) > 1/2 > \alpha/2 \). The theorem is proved.

**Remark 4.2.** If \( b \) and \( f \) are bounded and a part of \( \partial D \) near the origin is flat, then one can take \( \beta = 1 \) in (4.9). However, even in the case of flat boundary and bounded \( f \), if \( b \in L_d \), then in the general case certainly \( \beta < 1 \) (see Example 4.1 in [22]) and most likely \( \beta \to 0 \) as \( \delta \to 0 \).

**Corollary 4.11.** Under the assumptions of Theorem 4.10 suppose that we are given a function \( u \in W^{2}_{d,\text{loc}}(D) \cap C(\bar{D}) \). Let \( w(r) \) be a concave continuous function on \([0, \infty)\) such that \( w(0) = 0 \) and \( |u(x) - u(0)| \leq w(|x|) \) for all \( x \in \partial D \). Then for \( x \in D \) we have

\[
|u(x) - u(0)| \leq N|x|^{\beta}||Lu||_{L_{d_0}(D)} + \omega(N|x|^{\beta/2}),
\]

(4.10)

where \( \beta \) is the same as in Theorem 4.10 and \( N \) depend on the data in the same way as in Theorem 4.10.

Indeed, define \( f = -Lu \). Clearly, we may assume that \( f \in L_{d_0}(D) \). Then take a sequence of domains \( D_n \subset \bar{D}_n \subset \bar{D} \) such that \( D_n \uparrow \bar{D} \), denote by \( \tau_n(x) \) the first exit time of \( x + x_{t} \) from \( D_n \), and use Itô’s formula (cf. the proof of Corollary 3.1) to conclude that

\[
u(x) = E \int_{0}^{\tau_n(x)} f(x + x_t) dt + E u(x + x_{\tau_n(x)}).
\]

In light of (2.9) we can pass to the limit as \( n \to \infty \) and owing to (4.9) to conclude

\[
u(x) - u(0) = E \int_{0}^{\tau(x)} f(x + x_t) dt + E [u(x + x_{\tau(x)}) - u(0)]
\]

\[
\leq N|x|^{\beta}||f||_{L_{d_0}(D)} + E w(|x + x_{\tau(x)}|).
\]

Here, by Jensen’s inequality, the last term is less than \( w \) evaluated at the square root of

\[
E|x + x_{\tau(x)}|^2 = |x|^2 + 2E \int_{0}^{\tau(x)} [t_r a_s + (x + x_s)b_s] ds
\]

\[
\leq |x|^2 + N|x|^{\beta}(1 + \text{diam}(D)||b||).
\]

This yields an estimate for \( \nu(x) - u(0) \) from above. Similarly one estimates it from below.

Next, we study the probability to pass through narrow tubes. We represent the points in \( \mathbb{R}^d \) as \( x = (x^1, x') \), where \( x^1 \in \mathbb{R} \) and \( x' \in \mathbb{R}^{d-1} \).

**Theorem 4.12.** Let \( \kappa \in [1/2, 1) \). Then there exist \( T_1 > 1 > T_0 > 0, p_0 > 0 \), depending only on \( \kappa, d, ||b||, \) and \( \delta \), such that, if \( R \in (0, \infty), n \in \{2, 3, \ldots\} \), and an open round cylinder \( C \) in \( \mathbb{R}^d \) with base being a ball in \( \mathbb{R}^{d-1} \) of radius \( R \) and length \( nR \) is defined by \( C = (0, nR) \times \{ x' : |x'| < R \} \), then, for any \( x \in (R, x') \) with \( |x'| \leq \kappa R \), the probability that \( x + x_t \) will first exit from the
cylinder through \(\{nR\} \times \{x': |x'| < (1-\kappa)R\}\) and this will happen in the time interval \([(n-1)T_0 R^2, (n-1)T_1 R^2]\) is greater than or equal to \(p_0^{-1}\).

Proof. As usual we may concentrate on \(R = 1\). Let us call the sections of \(C\) by hyperplanes \(x^1 = k\) disks. We contract them to their centers with the coefficient of contraction, say \(c\) and call the results \(c\)-subdisks. In this terminology we need to estimate from below the probability of the event \(A\) that our process starting from a point on the \(\kappa\)-subdisk lying at the distance 1 from the base on which \(x^1 = 0\) will first exit from \(C\) during the time interval \([nT_0, nT_1]\) through the (smaller) \((1-\kappa)\)-subdisk on the other base.

Let \(C_k = C \cap \{k < x^1 < k+2\}\), \(k = 0, 1, ..., n-2\). Then for \(A\) to happen it suffices for the process to consecutively exit from each \(C_k\), \(k = 0, 1, ..., n-2\), through the \((1-\kappa)\)-subdisk where \(x^1 = k + 2\) during the time interval \([(k+1)T_0, (k+1)T_1]\). By using conditional expectations we easily see that \(P(A) \geq p_0^{-1}\), where \(p_0\) is the estimate from below in terms of only \(d, \|b\|\), and \(\delta\) of the probability of the event \(A_0\) that our process will first exit from \(C_0\) through the \((1-\kappa)\)-subdisk on which \(x^1 = 2\) during the time interval \([T_0, T_1]\).

Let \(B\) be the open unit ball centered at \(x_0 = (2-\kappa, 0)\), which is slightly off the center of \(C_0\). Since \(|B \cap \{x^1 \geq 2\}| = N(\kappa) > 0\) and \(|(1, x') - x_0| < (1-\kappa)^2 + \kappa^2 < 1\) if \(|x'| \leq \kappa\), by Corollary 4.2 we have \(P(A_0') \geq 3p_0 = 3p_0(\kappa, d, \|b\|) > 0\), where \(A_0'\) is the event that \(x + x_t\) reaches \(\{x^1 \geq 2\}\) before exiting from \(B\), that is \(\{x^1 = 2\} \cap \{x': |x'| \leq 1-\kappa\}\) before exiting from \(B\). We also know that \(E\tau_B \leq N\), where \(\tau_B\) is the first exit time of \(x + x_t\) from \(B\). Hence, there is \(T_1 > 1\) such that \(P(\tau_B > T_1) \leq p_0\). We also know (see Theorem 2.10 in [17]) that there exists \(T_0 \in (0, 1)\) such that \(P(\tau_B \leq T_0) \leq p_0\). Hence, \(P(A_0', \tau_B \in [T_0, T_1]) \geq p_0\). Since \(B \cap \{x^1 < 2\} \subset C_0\), obviously, \(A_0' \cap \{\tau_B \in [T_0, T_1]\} \subset A_0\) and the theorem is proved. \(\square\)

Corollary 4.13. Let \(R \in (0, \infty)\) and \(|x| \leq R\). Then there is a constant \(N = N(d, \delta, \|b\|)\) such that the expected time spent by \(x_t\) in \(B_R(x)\) before exiting from \(B_{2R}\) is greater than \(N^{-1} R^2\):

\[
R^2 \leq NE \int_0^{\tau_2} I_{B_R(x)}(x_t) \, dt.
\]

Indeed, as always we may assume that \(R = 1\) and then by Theorem 4.12 with probability \(p = p(d, \delta, \|b\|) > 0\) the process \(x_t\) reaches \(\bar{B}_{1/2}(x)\) before exiting from \(B_2\). After that happens the expected time spent in \(B_1(x)\) before \(\tau_2\) is greater than the expected time spent in \(B_1(x)\) before exiting from it. Then it only remains to use Corollary 2.12 of [17], according to which the expected exit time from \(B_1(x)\) starting from a point in \(\bar{B}_{1/2}(x)\) is greater than \(N^{-1}\).

This corollary easily implies the so-called doubling property of the Green’s measure of \(x_t\). Let \(D \subset \mathbb{R}^d\) be a bounded domain containing the origin.
Then the Green’s measure of \( x_t \) in \( D \) is defined by

\[
G(\Gamma) = E \int_0^\tau I_\Gamma(x_t) \, dt,
\]

where \( \tau \) is the first exit time of \( x_t \) from \( D \). As we know from the above, \( G \) has a density summable to the power of \( d_0/(d_0 - 1) \).

**Theorem 4.14** (doubling property). Let a ball \( B \subset D \) be such that \( 2B \subset D \). Then \( G(B) \leq NG((1/2)B) \), where \( N = N(d, \delta, b) \).

Proof. We may assume that \( D \) is connected and the radius of \( B \) is one. Then define \( \tau_D \) as the first exit time of \( x_t \) from \( D \) and introduce recurrently, for \( n = 0, 1, 2, \ldots \),

\[
\gamma_n = 0, \quad \tau_n = \inf \{ t \geq \gamma_n : x_t \in \bar{B} \}, \quad \gamma^{n+1} = \inf \{ t \geq \tau_n : x_t \notin (3/2)B \}.
\]

By the conditional versions of, first, Corollary 2.1 of [17] and then Corollary 4.13 we have

\[
G(B) = \sum_{n=1}^\infty EI_{\tau_n < \tau_D} E \left( \int_{\tau_n}^{\gamma^{n+1}} I_B(x_t) \, dt \mid \mathcal{F}_{\tau_n} \right)
\]

\[
\leq N \sum_{n=1}^\infty EI_{\tau_n < \tau_D} \leq N \sum_{n=1}^\infty EI_{\tau_n < \tau_D} E \left( \int_{\tau_n}^{\gamma^{n+1}} I_{(1/2)B}(x_t) \, dt \mid \mathcal{F}_{\tau_n} \right)
\]

\[
= NG((1/2)B),
\]

which proves the theorem.

**Corollary 4.15** \((A_\infty\text{-property of } G)\). There are constants \( \mu \geq 1 \) and \( N \), depending only on \( d, \delta, \|b\| \), such that for any ball \( B \) satisfying \( 2B \subset D \) and Borel \( \Gamma \subset B \) we have

\[
N \frac{G(\Gamma)}{G(B)} \geq \left( \frac{|\Gamma|}{|B|} \right)^\mu.
\]

Proof. Take the same \( \gamma^n, \tau^n \) as in the proof of Theorem 4.14 and observe that by the conditional version of Theorem 4.1 on the set \( \{ \tau_n < \tau_D \} \) (a.s.)

\[
E \left( \int_{\tau_n}^{\gamma^{n+1}} I_{\Gamma}(x_t) \, dt \mid \mathcal{F}_{\tau_n} \right) \geq N^{-1} R^2 (|\Gamma|/|B|)^\mu,
\]

where \( R \) is the radius of \( B \). Furthermore,

\[
NR^2 \geq E \left( \int_{\tau_n}^{\gamma^{n+1}} I_B(x_t) \, dt \mid \mathcal{F}_{\tau_n} \right).
\]

After that it only remains to mimic (4.11).

Corollary 4.15 is almost identical to Corollary 2.3 in [6]. However, there are no lower order terms in [6] and the comparable situations would be only when \( x_t \) were a solution of (1.1). We refer the reader to the proof of Corollary 2.3 in [6] concerning \( A_\infty \)-weights only pointing out that Corollary 4.15 is not sufficient for proving even Theorem 4.1 because not arbitrary subsets of \( B \) could be considered. On the other hand, \( N \) and \( \mu \) in (4.12)
are independent of how close \( \partial D \) to the origin is in contrast with Theorem 4.1, where the starting point of the process is at a distance at least \((1 - \kappa)R\)
from the boundary.

**Remark 4.3.** Once we know that \( G \) is an \( A_\infty \)-weight, it is also an \( A_p \)-weight for certain large \( p \). In particular, on any closed \( \Gamma \subset D \), \( G^{-\alpha} \) is summable for some \( \alpha > 0 \).

Theorem 4.12 allows us to prove a few more properties of \( x_t \). By Corollary 2.7 of [17] there are constants \( N, \nu > 0 \) depending only on \( d, \delta \), and \( \|b\| \) such that for any \( R, t > 0 \),

\[
P(\tau_R \geq T) \leq Ne^{-\nu T/R^2}.
\]

This turns out to be very close to an optimal result.

**Lemma 4.16.** There are constants \( N, \nu > 0 \), depending only on \( d, \delta \), and \( \|b\| \), such that for any \( R, T > 0 \),

\[
NP(\tau_R > T) \geq e^{-\nu T/R^2}.
\]  

**Proof.** We may assume that \( R = 3 \). Then the cylinder \( C = (-1, 2) \times \{x' : \ |x'| < 1\} \subset B_3 \) and \( \tau_3 > \tau \), where \( \tau \) is the first exit time of \( x_t \) from \( C \).
Introduce, times of meandering: \( \tau_0 = 0 \) and, for \( n = 0, 1, 2, ... \), let \( \tau_{2n+1} \) be the first exit time of \( x_t \) from \((-1, 1) \times \{x' : \ |x'| \leq 1\} \) after \( \tau_{2n} \), \( \tau_{2n+2} \) be the first exit time of \( x_t \) after \( \tau_{2n+1} \) from \((0, 2) \times \{x' : \ |x'| \leq 1\} \). Also let \( \kappa = 1/2 \), take \( T_0, T_1 \) from Theorem 4.12, introduce

\[
A_{2n+1} = \{x_{\tau_{2n+1}} \in \{1\} \times \{x' : \ |x'| \leq 1/2\}, \tau_{2n+1} - \tau_{2n} \in [T_0, T_1]\},
\]

\[
A_{2n+2} = \{x_{\tau_{2n+2}} \in \{0\} \times \{x' : \ |x'| \leq 1/2\}, \tau_{2n+2} - \tau_{2n+1} \in [T_0, T_1]\},
\]

and define \( n_0 \) as the least integer such that \( n_0 T_0 \geq T \). Observe that on the set

\[
\bigcap_{i=1}^{n_0} A_i
\]

we have

\[
T \leq n_0 T_0 \leq \tau_{n_0} < \tau
\]

and hence

\[
P(\tau_3 > T) \geq P(\tau > T) \geq P\left(\bigcap_{k=1}^{n_0} A_k\right) \geq p_0^{n_0},
\]

where the last inequality follows from the conditional version of Theorem 4.12. This obviously proves the lemma.

The following result will be used in a subsequent paper for establishing Harnack’s inequality for caloric functions related to diffusion processes with drift in \( L_d \).
Theorem 4.17. Let \( R \in (0, \infty), \kappa, \eta \in (0, 1) \), \( x, y \in B_{\kappa R} \), and \( \eta^{-1}R^2 \geq t \geq \eta R^2 \). Then there exist \( N, \nu > 0 \), depending only on \( \kappa, \eta, d, \delta, \) and \( \|b\| \), such that, for any \( \rho \in (0, 1) \),

\[
NP(x + x_t \in B_{\rho R}(y), \tau_R(x) > t) \geq \rho'.
\] (4.14)

We prove this theorem after appropriate preparations.

Lemma 4.18. If \( \rho_0 \in (0, 1), \xi \in (0, \infty) \), and \( \kappa \in [1/2, 1) \), then there exists \( \mu = \mu(d, \delta, \|b\|, \kappa, \rho_0, \xi) > 0 \) such that

\[
P(x + x_{\xi R^2} \in B_{\rho_0 \kappa R}(y), \tau_R(x) > \xi R^2) \geq \mu,
\] (4.15)

whenever \( R \in (0, \infty), x, y \in B_{\kappa R} \).

Proof. While proving (4.15) we may assume that \( R = 1 \). Then observe that (4.15) becomes stronger if \( \rho_0 \) becomes smaller. Therefore we may assume that

\[
\rho_0 \leq \min (\kappa^{-1} - 1, \xi/T_1).
\] (4.16)

Then also assume, as the first case, that \( 2|x - y| \geq \kappa \rho_0 \) and connect \( x \) and \( y \) by a round cylinder of length \( nr \), where

\[
n = \left\lceil \frac{9|x - y|}{\kappa \rho_0} \right\rceil + 1, \quad r = \frac{|y - x|}{n-1}.
\]

More precisely our cylinder is given by

\[
C = \{ x + t(y - x)/|y - x| + re : t \in (-r, (n - 1)r), e \in \mathbb{R}^d, |e| < 1, e \perp x \}.
\]

It is not hard to check that, owing to \( \rho_0 \leq \kappa^{-1} - 1 \), we have \( C \subseteq B_1 \). Also as is easy to see

\[
\kappa \rho_0/7 \geq r \geq \kappa \rho_0/9.
\]

Define

\[
\tau = \inf\{ s : x + x_s \in B_r(y) \}.
\]

By Theorem 4.12 we obtain that with probability not less than \( p_0^{n-1} \) we have \( \tau \leq (n - 1)T_1 r^2 \) and \( \tau < \tau_1(x) \). Furthermore,

\[
(n - 1)T_1 r^2 = |y - x|T_1 r \leq 2T_1 r \leq T_1 \rho_0 \leq \xi.
\]

By Lemma 4.16, given that \( \tau \leq \xi \wedge \tau_1(x) \), the probability that the process \( x_t \) does not exit from \( B_r(x_t) \) before time \( \xi \), assuring that \( x + x_\xi \in B_{\kappa \rho_0}(y) \) and \( \tau_1(x) > \xi \), is bigger than \( N^{-1}e^{-\nu \xi /\rho_0^2} \). Hence,

\[
P(x + x_\eta \in B_{\kappa \rho_0}(y), \tau_1(x) > \xi) \geq p_0^{n-1}N^{-1}e^{-\nu \xi /\rho_0^2} \geq N^{-1}e^{-\nu /\rho_0^2} =: \mu,
\]

where the last \( \nu \) is perhaps different from the previous one. This proves (4.15) if \( 2|x - y| \geq \kappa \rho_0 \). If \( 2|x - y| < \kappa \rho_0 \) one does not need the first part of the proof. The lemma is proved. \( \square \)

Lemma 4.19. Let \( \kappa, \eta \in (0, 1) \). Then there are constants \( N, \nu > 0 \), depending only on \( \kappa, \eta, d, \delta, \) and \( \|b\| \), such that, for any \( R \in (0, \infty), \rho \in (0, 1) \), and \( x \in B_{\kappa R} \),

\[
NP(\tau_R(x) > \eta R^2, x + x_{\eta R^2} \in B_{\rho R}) \geq \rho'.
\] (4.17)
 Furthermore, obviously, for \( \xi = \eta, y = 0, \) and \( \rho_0 \) equal to the right-hand side of (4.16), means that

\[
P(x + x_{\eta R^2} \in B_{\kappa \rho_0 R}, \sup_{s \leq \eta R^2} |x + x_s| < R) \geq \mu, \tag{4.18}
\]

whenever \( R \in (0, \infty) \) and \( x \in B_{\kappa R} \). For \( n = 1, 2, \ldots \) introduce \((t_0 := 0)\)

\[
R_n = \rho_0^{-1} = \rho_0 R_{n-1}, \quad s_n = \eta R_n^2 = \eta \rho_0^{2(n-1)}, \quad t_n = \sum_{k=1}^n s_k,
\]

\[
A_n = \left\{ \sup_{s \leq s_n} |x + x_{s+t_{n-1}}| < R_n \right\}, \quad \Pi_n = \bigcap_{k=1}^n A_k
\]

and observe that by the conditional version of (4.18) on the set \( \{ y := x + x_{t_{n-1}} \in B_{\kappa R_n} \} \) we have (a.s.)

\[
P\left(y + (x_{t_n} - x_{t_{n-1}}) \in B_{\kappa R_{n+1}}, \sup_{s \leq s_n} |y + (x_{t_{n-1} + s} - x_{t_{n-1}})| < R_n \mid \mathcal{F}_{t_{n-1}} \right) \geq \mu.
\tag{4.19}
\]

Furthermore, obviously, for \( n \geq 2, \)

\[
P^n := P\left(x + x_{t_n} \in B_{\kappa R_{n+1}}, \Pi_n \right) \quad \geq \quad P\left(x + x_{t_{n-1}} \in B_{\kappa R_{n-1}}, \Pi_{n-1} \right),
\]

\[
x + x_{t_{n-1}} + (x_{t_n} - x_{t_{n-1}}) \in B_{\kappa R_{n+1}}, \sup_{s \leq s_n} |x + x_{t_{n-1}} + (x_{t_{n-1} + s} - x_{t_{n-1}})| < R_n,
\]

which in light of (4.19) yields \( P^n \geq \mu P^{n-1} \) and since for \( |x| < \kappa \) we have \( P^1 \geq \mu \) by (4.18), it holds that for \( |x| < \kappa \) and all \( n \geq 0 \)

\[
P\left(x + x_{t_n} \in B_{\kappa R_{n+1}}, \sup_{s \leq t_n} |x + x_s| < 1 \right) \geq \mu^{n+1}.
\tag{4.20}
\]

Now it is convenient to consider \( \eta \) as a variable and \( \kappa, d, \delta, \|b\| \) as fixed parameters and not to include them in the arguments of some functions which appear below. Observe that

\[
t_n = \eta \sum_{k=0}^{n-1} \rho_0^{2k}(\eta)
\]

is a strictly increasing function of \( \eta \) and \( t_n \geq \eta \). Therefore, for fixed \( \eta' \in (0, 1) \) and each \( n = 1, 2, \ldots \) there is \( \eta = \eta(n) = \eta(n, \eta') \in (0, 1) \) such that

\[
\eta' = \eta \sum_{k=0}^{n-1} \rho_0^{2k}(\eta) \quad (= t_n).
\]

Clearly, the sequence \( \eta(n) \) is decreasing and its limit \( \bar{\eta} \) is a function of \( \eta' \), which is strictly positive.

Then take \( \rho \in (0, \kappa \rho_0(\bar{\eta})) \) and define \( n(\rho) = n(\rho, \eta') \) as the biggest \( n \geq 1 \) satisfying

\[
\kappa \rho_0^2(\bar{\eta}) \geq \rho \tag{4.21}
\]
that is
\[ n(\rho) = \left\lfloor \frac{\ln(\rho/\kappa)}{\ln \rho_0(\bar{\eta})} \right\rfloor. \]

With so defined \( n = n(\rho) \) in light of (4.21) we have \( \kappa \rho_0^\eta(\eta(n)) \geq \rho \) and (4.20) yields
\[ P(\{ x + x_{\eta'} \in B_\rho, \sup_{s \leq \eta'} | x + x_s | < 1 \}) \geq \mu^{n(\rho)+1} \]
if \( \rho \in (0, \kappa \rho_0(\bar{\eta})) \) and \( | x | < \kappa \). Here
\[ \mu^{n(\rho)+1} \geq \mu \exp \left( \frac{\ln(\rho/\kappa)}{\ln \rho_0(\bar{\eta})} \ln \mu \right) = N \rho^\nu, \]
where \( N \) and \( \nu \) are defined by the above equality. It follows that (4.17) holds with \( R = 1 \) if \( \rho \in (0, \kappa \rho_0(\bar{\eta})) \). Then it automatically holds for larger \( \rho \) perhaps with a different \( N \). Arbitrary \( R \) are treated by self-similarity. The lemma is proved. □

**Proof of Theorem 4.17.** Let \( R_1 = R - | y | (\geq (1 - \kappa)R) \) and \( \xi = t/R_2 - \eta \).

By the conditional version of Lemma 4.19 on the set \( \{ z := x + x_{\xi R_2^2} \in B_{\kappa R_1}(y) \} \) we have (a.s.)
\[ NP \left( \sup_{s \in [\xi R_2^2, \xi R_2^2 + \eta R_2^2]} | z + x_s - x_{\xi R_2^2} - y | < R_1, \right. \]
\[ x + x_{\xi R_2^2 + \eta R_2^2} \in B_{\rho R_1}(y) \left| \mathcal{F}_\xi \right) \geq \rho^\nu. \]

By Lemma 4.18, where we take \( \rho_0 = R_1/R \) and \( \xi = (t - R_2^2 \eta)/R_2 \),
\[ P \left( \sup_{s \leq \xi R_2^2} | x + x_s | < R, x + x_{\xi R_2^2} \in B_{\kappa R_1}(y), \right) \geq \mu. \]

By combining these two facts and using that \( \xi R_2^2 + \eta R_2^2 = t \), we obviously come to (4.14). The theorem is proved. □

**References**

[1] A. D. Aleksandrov, *Uniqueness conditions and estimates for the solution of the Dirichlet problem*, Vestnik Leningrad. Univ., Vol. 18 (1963), No. 3, 5-29 in Russian; English translation in Amer. Mat. Soc. Transl., Vol. 68 (1968), No. 2, 89-119.

[2] S. Agmon, A. Douglas, and L. Nirenberg, *Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. I*, Comm. Pure Appl. Math., Vol. 12 (1959), No. 4, 623–727.

[3] X. Cabré, *On the Alexandroff-Bakelman-Pucci estimate and the reversed Hölder inequality for solutions of elliptic and parabolic equations*, Comm. Pure Appl. Math., Vol. 48 (1995), 539–570.

[4] F. Chiarenza, M. Frasca, and P. Longo, *W^{2,p}-solvability of the Dirichlet problem for nondivergence elliptic equations with VMO coefficients*, Trans. Amer. Math. Soc., Vol. 336 (1993), No. 2, 841–853.

[5] E. B. Dynkin, “Markov processes”, Fizmatgiz, Moscow, 1963 in Russian: English translation in Grundlehren Math. Wiss., Vols. 121, 122, Springer-Verlag, Berlin, 1965.
[6] E.B. Fabes and D.W. Stroock, *The $L^p$-integrability of Green’s functions and fundamental solutions for elliptic and parabolic equations*, Duke Math. J., Vol. 51 (1984), No. 4, 997–1016.

[7] K. Fok, *A nonlinear Fabes-Stroock result*, Comm. PDEs, Vol 23 (1998), No. 5-6, 967–983.

[8] M. Giaquinta and G. Modica, *Regularity results for some classes of higher order non linear elliptic systems*, J. Reine Angew. Math., 311(312), (1979), 145-169.

[9] F.W. Gehring, *The $L^p$-integrability of the partial derivatives of a quasiconformal mapping*, Acta Math., Vol. 130 (1973), 265–277.

[10] D. Gilbarg and N. S. Trudinger, “Elliptic partial differential equations of second order”, Reprint of the 2nd ed. Berlin Heidelberg New York 1983. Corr. 3rd printing 1998, 2001.

[11] N.V. Krylov, *On the selection of a Markov process from a system of processes and the construction of quasi-diffusion processes*, Izvestiya Akademii Nauk SSSR, seriya matematicheskaya, Vol. 37, No. 3 (1973), 691–708 in Russian; English translation in Math. USSR Izvestija, Vol. 7, No. 3 (1973), 691–709.

[12] N.V. Krylov, “Controlled diffusion processes”, Nauka, Moscow, 1977 in Russian; English transl. Springer, 1980.

[13] N.V. Krylov, “Introduction to the theory of diffusion processes”, Amer. Math. Soc., Providence, RI, 1995.

[14] N.V. Krylov, *Second-order elliptic equations with variably partially VMO coefficients*, J. Funct. Analysis, Vol. 257 (2009), 1695–1712.

[15] N.V. Krylov, *Some $L_p$-estimates for elliptic and parabolic operators with measurable coefficients*, Discrete and Continuous Dynamical Systems, Series B, Vol. 17 (2012), No. 6, 2073–2090.

[16] N.V. Krylov, “Sobolev and viscosity solutions for fully nonlinear elliptic and parabolic equations”, Mathematical Surveys and Monographs, 233, Amer. Math. Soc., Providence, RI, 2018.

[17] N.V. Krylov, *On stochastic Itô processes with drift in $L_d$*, http://arxiv.org/abs/2001.03660

[18] N.V. Krylov and M.V. Safonov, *An estimate of the probability that a diffusion process hits a set of positive measure*, Doklady Academyi Nauk SSSR, Vol. 245 (1979), No.1, 18–20 in Russian; English translation in Soviet Math. Dokl., Vol. 20 (1979), No. 2, 253–255.

[19] N.V. Krylov and M.V. Safonov, *A certain property of solutions of parabolic equations with measurable coefficients*, Izvestiya Academyi Nauk SSSR, seriya matematicheskaya, Vol. 44 (1980), No. 1, 161–175 in Russian; English translation in Math. USSR Izvestija, Vol. 16 (1981), No. 1, 151–164.

[20] O.A. Ladyzhenskaya and N.N. Ural’tseva, “Linear and quasilinear equations of elliptic type” Second edition, revised, “Nauka”, Moscow, 1973, 576 p. in Russian.

[21] M. V. Safonov, *Harnack inequalities for elliptic equations and Hölder continuity of their solutions*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov (LOMI), Vol. 96 (1980), 272–287 in Russian; English transl. in Journal of Soviet Mathematics, Vol. 21 (March 1983), No. 3, 851–863.

[22] M.V. Safonov, *Non-divergence elliptic equations of second order with unbounded drift*, Nonlinear partial differential equations and related topics, 211–232, Amer. Math. Soc. Transl. Ser. 2, 229, Adv. Math. Sci., 64, Amer. Math. Soc., Providence, RI, 2010.

[23] D. W. Stroock and S. R. S. Varadhan “Multidimensional diffusion processes”, Grundlehren Math. Wiss., Vol. 233, Springer-Verlag, Berlin and New York, 1979. 

E-mail address: nkrylov@umn.edu

127 Vincent Hall, University of Minnesota, Minneapolis, MN, 55455