



ELSEVIER

On weak uniqueness for some diffusions with discontinuous coefficients[☆]

N.V. Krylov*

School of Mathematics, University of Minnesota, 127 Vincent Hall, Minneapolis, MN 55455, USA

Received 29 June 2003; received in revised form 15 March 2004; accepted 30 March 2004

Abstract

Several situations when one can prove weak uniqueness of solutions of Itô equations with discontinuous or/and degenerate coefficients are presented. In particular, the cases are considered in which the set of discontinuity is a cone, or a straight line, or else a discrete set of points.

© 2004 Elsevier B.V. All rights reserved.

MSC: 60H10; 35J15

Keywords: Weak uniqueness; Stochastic Itô equations; Second-order elliptic equations

1. Introduction

Let $d \in \{1, 2, \dots\}$ be an integer, let

$$\mathbb{R}^d = \{x = (x^1, \dots, x^d) : x^i \in \mathbb{R}, i = 1, \dots, d\}$$

be a d -dimensional Euclidean space, and let $a(x) = (a^{ij}(x))_{i,j=1}^d$ be a symmetric nonnegative matrix-valued and $b(x) = (b^i(x))_{i=1}^d$ an \mathbb{R}^d -valued Borel functions given on \mathbb{R}^d .

It may happen that, for an $x \in \mathbb{R}^d$, there exists a probability space, a d -dimensional continuous process $x_t, t \geq 0$, and a d -dimensional Wiener process $w_t, t \geq 0$, given on this space such that $w_{t+h} - w_t$ and $\sigma\{x_s, w_s : 0 \leq s \leq t\}$ are independent whenever $t, h \geq 0$ and (a.s.) for all $t \geq 0$

$$x_t = x + \int_0^t \sqrt{a(x_s)} dw_s + \int_0^t b(x_s) ds, \tag{1.1}$$

[☆] The work was partially supported by NSF Grant DMS-0140405.

* Tel.: 6126258338; fax: 6126262017.

E-mail address: krylov@math.umn.edu (N.V. Krylov).

where \sqrt{a} denotes the nonnegative definite square root of the matrix a . In that case the problem arises as to whether the distribution of x depends on the probability space and the particular Wiener process. If it does not, we say that weak uniqueness holds for (1.1) at point x . We say that weak uniqueness holds for (1.1) if it holds at any $x \in \mathbb{R}^d$.

The problem of weak uniqueness arises in very many situations. One of them is the problem of constructing diffusion processes, when concrete equations governing the trajectories are not important and what counts is the finite-dimensional *distributions* of trajectories. If we want our diffusion process to behave like a solution of (1.1) and possess the Markov property, then we either have to prove that weak uniqueness holds for (1.1) or else carefully choose “right” solutions among all possible solutions of (1.1). Stroock and Varadhan (1979) and Ikeda and Watanabe (1981) are good sources of information on the state of the problem in the last century and also good sources of further references related to constructing diffusion processes by using Itô stochastic equations.

One of strong motivations for this article was to present methods allowing us to treat the following example.

Example 1.1. In Khasminskii and Krylov (2001) the authors encountered the problem of weak uniqueness in a particular situation in which $d \geq 2$, $x = (x', x'')$, $x' \in \mathbb{R}$, $x'' \in \mathbb{R}^{d-1}$, and in Eq. (1.1)

1. $a(x) = I_{x' > 0} a_+(x'') + I_{x' \leq 0} a_-(x'')$, $b(x) = I_{x' > 0} b_+(x'') + I_{x' \leq 0} b_-(x'')$, $b^1 \equiv 0$;
2. $\delta \leq a_{\pm}^{11}(x'') \leq K$, where $\delta, K \in (0, \infty)$ are some constants;
3. the first and second order derivatives of $a_{\pm}(x'')$ and $b_{\pm}(x'')$ are bounded.

Equations with such coefficients arose in Khasminskii and Krylov (2001) as limit equations in an averaging problem. However, to be sure that not only subsequences of processes under consideration converge weakly but the whole sequence does, it is assumed in Khasminskii and Krylov (2001) that weak uniqueness holds in Example 1.1. As we will see later in Remark 3.1, this assumption holds automatically.

The matrix a in Example 1.1 is discontinuous and degenerate. Generally, even if a in (1.1) is uniformly nondegenerate but discontinuous, the problem of weak uniqueness is highly nontrivial. Nadirashvili (1997) and see also an excellent exposition of a generalization of that paper in Safonov (1999) gives an example in which $d \geq 3$, a is uniformly nondegenerate and bounded, $b \equiv 0$ and weak uniqueness just does not hold. Hence, we can only hope to find subclasses of equations for which weak uniqueness holds. Several of them are known for uniformly nondegenerate a . Weak uniqueness always holds if $d \leq 2$ (see, for instance, Krylov, 1969). The continuity of a even if it also depends on time is known to be sufficient since long ago as well (see Stroock and Varadhan, 1979). The result of Bramanti and Cerutti (1993) shows that it is enough for a to be in VMO rather than to be continuous. About 10 years ago the articles (Cerutti et al., 1993; Krylov, 1992; Safonov, 1994) appear showing that weak uniqueness holds if the set of discontinuity is discrete or more generally has zero α -Hausdorff measure with sufficiently small α . Other sufficient conditions are given when the set of

discontinuity has special geometric structure: (i) in Bass and Pardoux (1987) where a is piecewise constant on each polyhedron of a decomposition of \mathbb{R}^d into finite union of them, (ii) in Gao (1993) where the set of discontinuity is a hyperplane of co-dimension one.

However, none of the above references covers Example 1.1 just because a may degenerate. In a sense, our treatment of it is quite close to some arguments in Bass and Pardoux (1987), we only add random time change and rely on stronger and more general results. Some of these results deal with relating weak uniqueness problem for stochastic equation (1.1) to weak uniqueness for the Dirichlet problem for the operator L :

$$u \rightarrow Lu, \quad Lu(x) = (1/2)a^{ij}(x)u_{x^i x^j}(x) + b^i(x)u_{x^i}(x)$$

(we use here and throughout the paper the summation convention over repeated indices). This relation is important, in particular, because it allows us to use some results stated in PDE terms for the problem of weak uniqueness for uniformly elliptic operators. In connection with this it is worth emphasizing that the nondegeneracy assumption is not used in Theorems 2.18 and 2.19.

Our main results are stated in Section 2. A rather long discussion of them along with the discussion of Example 1.1 is given in Section 3. Subsequent Sections 4–6 contain the proofs of the main results. In the final Section 7 we discuss two general results one of which is a criterion for weak uniqueness and another one says that a slight “shaking” of any equation leads to an equation possessing weak uniqueness property.

We use the notation $C = C([0, \infty), \mathbb{R}^d)$ for the Polish space of continuous \mathbb{R}^d -valued functions given on $[0, \infty)$ with a concept of convergence which is equivalent to uniform convergence on finite intervals. By smooth functions or domains we mean infinitely differentiable functions and domains. If D is a domain in \mathbb{R}^d , by $C_0^\infty(D)$ we denote the space of infinitely differentiable functions on D with compact support, $L_p(D)$ stands for the Lebesgue space of functions f such that

$$\|f\|_{L_p(D)}^p = \int_D |f(x)|^p dx < \infty.$$

We also need the Sobolev space $W_p^2(D)$ of functions u with finite norm defined by

$$\|u\|_{W_p^2(D)} = \|u\|_{L_p(D)} + \sum_{i=1}^d \|u_{x^i}\|_{L_p(D)} + \sum_{i,j=1}^d \|u_{x^i x^j}\|_{L_p(D)},$$

where the derivatives $u_{x^i}, u_{x^i x^j}$ are understood in the sense of distributions. Everywhere in the article K, δ are fixed constants, $K, \delta \in (0, \infty)$.

The author is sincerely grateful to the referee for pointing out the article by Gao (1993).

2. Main results

Unless explicitly stated otherwise, we always assume that for all $x, \lambda \in \mathbb{R}^d$, we have

$$a(x) = a^*(x), \quad \|a(x)\| + |b(x)|^2 \leq K(1 + |x|^2), \quad a^{ij}(x)\lambda^i \lambda^j \geq \delta|\lambda|^2. \quad (2.1)$$

Remark 2.1. Under these conditions Eq. (1.1) has a solution on some probability space. To show this we take any sequence of smooth a_n, b_n satisfying (2.1) which are such that $a_n, b_n \rightarrow a, b$ as $n \rightarrow \infty$ (a.e.) in \mathbb{R}^d . Then, for each n , denote by x_t^n a solution of

$$x_t^n = x + \int_0^t \sqrt{a_n(x_s^n)} \, dw_s^n + \int_0^t b_n(x_s^n) \, ds \tag{2.2}$$

defined on a probability space, naturally, with the condition that w_t^n is a d -dimensional Wiener process and $w_{t+h}^n - w_t^n$ is independent of $w_s^n, x_s^n, s \leq t$ for any $t, h \geq 0$. Since, the functions a_n, b_n are smooth and satisfy (2.1), the processes x_t^n are well defined on any probability space carrying a Wiener process and the set of distributions of x_t^n on C is precompact. The latter and Skorokhod’s embedding theorem allows us to assume, if necessary upon passing to a subsequence, that, for all n, x_t^n and w_t^n are defined on the same probability space and $x_t^n \rightarrow x_t, w_t^n \rightarrow w_t$ (a.s.) for certain processes x_t, w_t . Then we pass to the limit in (2.2) and we prove that x_t satisfies (1.1). This line of argument is justified in Krylov (1977) if, additionally, a and b are bounded. In our case they are allowed to have linear growth as $|x| \rightarrow \infty$. The reader will easily make the necessary adjustments after noticing that, for any $T \in (0, \infty), E \sup_{t \leq T} |x_t^n|$ is bounded by a constant independent of n (see, for instance, Krylov, 1995).

It is convenient to localize the notion of weak uniqueness.

Definition 2.2. Let D be an open nonempty subset of \mathbb{R}^d and $x \in D$. For $y \in C$ introduce

$$\tau_D(y) = \inf \{t \geq 0 : y_t \notin D\} \quad (\inf \emptyset := \infty).$$

Let $x_t, t \geq 0$, be a solution of

$$x_t = x + \int_0^t \sqrt{a(x_s)} I_{s < \tau_D(x)} \, dw_s + \int_0^t b(x_s) I_{s < \tau_D(x)} \, ds \tag{2.3}$$

defined on a probability space carrying a Wiener process $w_t, t \geq 0$, and introduce \mathbb{P} as the distribution of x_t on C . We say that weak uniqueness holds for (2.3) at point x if \mathbb{P} , the law of x_t , is the same for all solutions of (2.3) perhaps given on different probability spaces. Sometimes the domain D under consideration needs to be emphasized. Then instead of saying that weak uniqueness holds for (2.3) at point x we say that weak uniqueness holds for (2.3, D) at point x . We say that weak uniqueness holds for (2.3) or for (2.3, D) if it holds at any point in D . In the latter case we also say that weak uniqueness holds for (1.1) until first exit time from D .

As in the case of Eq. (1.1) while speaking about solutions of (2.3) we always assume that $w_{t+h} - w_t$ and $\sigma\{x_s, w_s : 0 \leq s \leq t\}$ are independent whenever $t, h \geq 0$.

Remark 2.3. Observe that if x_t is a solution of (1.1), then obviously $x_{t \wedge \tau_D(x)}$ is a solution of (2.3). However, if we know that weak uniqueness holds for (2.3) with D replaced with a larger domain, say \mathbb{R}^d , formally, we cannot conclude that it holds for (2.3) as it is. Yet under the standing assumption (2.1) weak uniqueness of solutions defined on larger time intervals implies that on smaller ones. To be more precise, fix

$T \in (0, \infty)$ and assume that weak uniqueness holds for (2.3). Also let $x_t, t \geq 0$, be a continuous process which satisfies (2.3) only for $t \in [0, T]$ with a Wiener process w_t defined perhaps also only for $t \in [0, T]$. Then it turns out that the distribution of $x_{\cdot \wedge T}$ on C is uniquely determined by a, b, D , and T .

Indeed, according to Krylov (1973a,b), there exists a function $P_y(B)$ defined for $y \in \mathbb{R}^d$ and Borel $B \subset C$ such that $P_y(B)$ is Borel in y for every Borel B and, for every $y \in \mathbb{R}^d$, P_y is the distribution on C of a solution of (1.1) with y in place of x . By “piecing together” the distribution of $x_{\cdot \wedge \gamma}$ and P_{x_γ} , where $\gamma = \gamma(x) = T \wedge \tau_D(x)$, one gets a measure P on C corresponding to a process which “behaves” before γ as $x_{\cdot \wedge \gamma}$ and after γ as a solution of (1.1) starting at point x_γ . A rigorous way to do the “piecing” can be found in Krylov (1973a,b) or Stroock and Varadhan (1979) (see Theorem 6.1.2 in Stroock and Varadhan, 1979), where one also finds that P corresponds to a solution of (1.1), say \bar{x} . Then $\hat{x}_t := \bar{x}_{t \wedge \tau_D(\bar{x})}$ is a solution of (2.3). By the assumption the distribution of \hat{x} is uniquely determined by a, b , and D . By construction the distributions of $x_{\cdot \wedge \gamma(x)}$ and $\bar{x}_{\cdot \wedge \gamma(\bar{x})}$ coincide and it only remain to notice that $\bar{x}_{\cdot \wedge \gamma(\bar{x})} = \hat{x}_{\cdot \wedge T}$ whereas $x_{\cdot \wedge \gamma(x)} = x_{\cdot \wedge T}$ since $x_{t \wedge \tau_D(x)} = x_t$ for $t \leq T$.

Remark 2.4. Let $x \in D$. Then weak uniqueness holds for (2.3) at x if weak uniqueness holds for (2.3, Q) at x for any smooth bounded domain Q such that $x \in Q \subset \bar{Q} \subset D$.

Indeed, take any increasing sequence of smooth bounded domains D^n such that $x \in \bar{D}^n \subset D$ and $\bigcup_n D^n = D$. Also let x_t be a solution of (2.3). Then

$$x_{\cdot} = x_{\cdot \wedge \tau_D(x)} = \lim_{n \rightarrow \infty} x_{\cdot \wedge \tau_{D^n}(x)}$$

in the sense of convergence in C . Furthermore, $x_{t \wedge \tau_{D^n}(x)}$ solves (2.3) with D^n in place of D , so that if the distribution of $x_{t \wedge \tau_{D^n}(x)}$ is independent of the solution, the same holds for the distribution of x_{\cdot} .

Here are the main results of the article, the first of which is stated for any domain $D \subset \mathbb{R}^d$ perhaps coinciding with \mathbb{R}^d .

Theorem 2.5. *Assume that $0 \in D$. Then the following statements are equivalent:*

- (a) *weak uniqueness holds for (2.3);*
- (b) *for any bounded subdomain $Q \subset \bar{Q} \subset D \setminus \{0\}$ with boundary of class C^∞ , weak uniqueness holds for (2.3, Q);*
- (c) *weak uniqueness holds for (2.3) at 0;*
- (d) *for each point $y \in D$ there is an open ball $B_r(y)$ of radius $r > 0$ centered at y such that weak uniqueness holds for (2.3, $B_r(y)$).*

Remark 2.6. The implication (b) \Rightarrow (a) generalizes Theorem 5.5 of Bass and Pardoux, 1987 in which a is assumed to depend only on $x/|x|$.

Remark 2.7. The implication (d) \Rightarrow (a) would follow almost immediately from Theorem 6.6.1 of Stroock and Varadhan (1979), at least for bounded a and b if we knew

that one can continue the coefficients outside of each $B_r(y)$ so that new equations in the whole \mathbb{R}^d possess weak uniqueness. One could also prove this implication by “piecing together” solutions in each ball in the spirit of Remark 2.3. We prefer to give a rather short proof based on Theorems 2.14 and 2.11, which relate weak uniqueness for stochastic equations to that for elliptic operators.

Remark 2.8. It is well known that if a, b are locally Lipschitz continuous in D , then any solution is a measurable function of w , and hence weak uniqueness holds. In that case the third condition in (2.1) is not needed. Theorem 2.5 implies that, even if the local Lipschitz condition is satisfied only in $D \setminus \{0\}$, still weak uniqueness holds (but then the nondegeneracy condition is required).

Remark 2.9. There is another “classical” statement saying that weak uniqueness holds for (2.3) if and only if it holds for (2.3) with 0 in place of b .

This fact follows almost immediately from Girsanov’s theorem. We show how to derive the “only if” part leaving the “if” part to the reader. Let \hat{x}_t be a solution of (2.3) with 0 in place of b . Then by Girsanov’s theorem for any $T \in (0, \infty)$ and Borel function $f(y) \geq 0, y \in C$, depending only on $y_t, t \leq T$, we have

$$\begin{aligned}
 Ef(\hat{x}_t) = Ef(x_t) \exp \left(- \int_0^T I_{t < \tau_D(x_t)} b^* a^{-1}(x_t) dx_t \right. \\
 \left. + (1/2) \int_0^T I_{t < \tau_D(x_t)} |a^{-1/2} b(x_t)|^2 dt \right), \tag{2.4}
 \end{aligned}$$

where x_t is a solution of (2.3) for $t \in [0, T]$. By Remark 2.3 the distribution of $x_{\cdot \wedge T}$ is uniquely determined by a, b, D , and T , so that by (2.4) the same holds for $Ef(\hat{x}_t)$. Since this is true for any T , the distribution of \hat{x}_\cdot is uniquely determined by a, b, D , that is, weak uniqueness indeed holds for (2.3) with 0 in place of b .

Wider sufficient conditions for weak uniqueness for (2.3) to hold can be obtained on the basis of the theory of elliptic equations.

Let $c = c(x)$ be a nonnegative Borel bounded function on \mathbb{R}^d . It is well known that if a, b, c are smooth and D is a bounded C^∞ -domain, then there exists a unique (up to equivalence in y) Green’s function of the operator $L - c$ in D with zero boundary data, that is there exists a Borel nonnegative function $g(x, y), x, y \in \bar{D}$, such that for any $f \in C^\infty(\bar{D})$ the unique solution $u \in C^\infty(\bar{D})$ of the problem

$$Lu - cu = -f \quad \text{in } D, \quad u = 0 \quad \text{on } \partial D$$

is represented as

$$u(x) = \int_D g(x, y) f(y) dy, \quad x \in \bar{D}.$$

Definition 2.10. Let D be a bounded C^∞ -domain, $x \in \bar{D}$, and let g be a function of class $L_{d/(d-1)}(D)$. We call g a Green’s function of $L - c$ (in D with pole at x) and we

write $g \in G(D, L - c, x)$ if there exists a sequence of functions

$$a_n = (a_n^{ij})_{i,j=1}^d, \quad b_n = (b_n^i)_{i=1}^d, \quad c_n$$

given and infinitely differentiable in \bar{D} satisfying

$$a_n = a_n^*, \quad a_n^{ij} \lambda^i \lambda^j \geq \delta |\lambda|^2, \quad \|a_n\| + |b_n| + c_n \leq K, \quad c_n \geq 0 \tag{2.5}$$

in D for all $\lambda \in \mathbb{R}^d$, and such that $a_n, b_n, c_n \rightarrow a, b, c$ in D (a.e.) and $g_n(x, \cdot) \rightarrow g$ weakly in $L_{d/(d-1)}(D)$, where $g_n(x, y)$ is the Green’s function in D corresponding to the operator $L_n - c_n$, with $L_n u := a_n^{ij} u_{x^i x^j} + b_n^i u_{x^i}$.

We say that weak uniqueness at x holds for the operator $L - c$ in D if there is only one (up to equivalence) Green’s function of $L - c$ in D with pole at x . We say that weak uniqueness holds for the operator $L - c$ in D if it holds at any $x \in D$.

This definition can be found in Krylov (1992), few results of which will be also used in what follows quite often without direct citation. For instance, we know that, for each $x \in D$, the set $G(D, L - c, x)$ is a nonempty closed convex compact subset of $L_{d/(d-1)}(D)$. Furthermore, for $g \in G(D, L - c, x)$

$$\|g\|_{L_{d/(d-1)}(D)} \leq N, \tag{2.6}$$

where N depends only on δ, K , and D . Also, obviously $g \geq 0$ (a.e. D) if $g \in G(D, L - c, x)$.

Definition 2.10 is naturally extended to the case $D = \mathbb{R}^d$. However, in this situation we require that a and b be bounded and c be bounded away from zero in order to be sure that the set $G(\mathbb{R}^d, L - c, x)$ is a nonempty closed convex compact subset of $L_{d/(d-1)}(\mathbb{R}^d)$ and (2.6) holds.

Integrals of Green’s functions over subsets $\Gamma \subset D$ are known to be expected times which corresponding diffusion processes spend in Γ before exiting from D . The average of occupation measure for a process x . seems to be a rather weak characterization of its distribution in C . On that account the equivalence (a) \Leftrightarrow (d) in the following theorem may look rather surprising.

Theorem 2.11. *Let D be a bounded C^∞ -domain and let $0 \in D$. Then the following statements are equivalent:*

- (a) *weak uniqueness holds for equation (2.3);*
- (b) *weak uniqueness holds for the operator $L - 1$ in D ;*
- (c) *weak uniqueness holds for the operator $L - c$ in D ;*
- (d) *weak uniqueness at 0 holds for the operator $L - 1$ in D .*

If a and b are bounded, these statements are also equivalent when $D = \mathbb{R}^d$ and in (c) we only take c bounded away from zero.

Remark 2.12. It follows from Definition 2.10 that if $g \in G(D, L - c, x)$ and $u \in C^\infty(\bar{D})$ and $u = 0$ on ∂D , then

$$u(x) = - \int_D g(y)(Lu - cu)(y) \, dy. \tag{2.7}$$

In turn this fact and (2.6) imply that if the set $\mathbb{L} := \{Lu - cu : u \in C^\infty(\bar{D}), u|_{\partial D} = 0\}$ is dense in $L_d(D)$, then weak uniqueness holds for $L - c$ in D and, owing to Theorem 2.11, it also holds for the stochastic equation (2.3).

Remark 2.13. Assume that D is smooth and bounded or else $D = \mathbb{R}^d$ and then a and b are bounded. Assume that for any $f \in L_d(D)$ there exists a $u \in W_d^2(D)$ vanishing on the boundary (if $D \neq \mathbb{R}^d$) such that $Lu - u = f$. In that case obviously the set \mathbb{L} from Remark 2.12 is dense in $L_d(D)$ and therefore weak uniqueness holds for (2.3).

It is well known (see, for instance, Gilbarg and Trudinger, 1983) that if D is smooth and bounded and a is uniformly continuous in D , then the above solvability assumption is satisfied. Thus, weak uniqueness holds for (2.3) if a is uniformly continuous in D and D is smooth and bounded. Then Theorem 2.5 allows us to extend the result to unbounded D . We thus get a particular case of results of Stroock and Varadhan (1979), where time inhomogeneous equations are considered. According to Bramanti and Cerutti (1993) the continuity of a can be relaxed to the condition $a \in \text{VMO}$. Below we will see other examples when the solvability is known.

Our study of the relationship between weak uniqueness for (2.3) and weak uniqueness for the operator $L - c$ in D is based to a very large extent on the following two theorems.

Theorem 2.14. *Let Q be a domain in \mathbb{R}^d . Let D be a bounded C^∞ -domain, $x \in D \subset Q$, and let g be a nonnegative integrable function on D such that (2.7) holds for any $u \in C^\infty(\bar{D})$ with $u = 0$ on ∂D . Then there is a solution x_t of (2.3) with Q in place of D such that for any Borel bounded f*

$$\int_D g(y)f(y) \, dy = E \int_0^\tau f(x_t) \exp\left(- \int_0^t c(x_s) \, ds\right) \, dt, \tag{2.8}$$

where τ is the first exit time of x_t from D . In particular (when $Q = D$), (a) \Rightarrow (c) in Theorem 2.11.

Theorem 2.15. *Let D be a bounded C^∞ -domain, let x_t be a solution of (2.3) with $x \in D$, τ its first exit time from D , and let $\gamma \leq \tau$ be a stopping time with respect to the filtration of $\mathcal{F}_t = \sigma\{x_s : s \leq t\}$. Then there exists a function $g(y) = g(\omega, y)$ such that*

- (a) for almost any ω , $g \in G(D, L - c, x_\gamma)$;
- (b) for any bounded Borel function f given on D , we have almost surely

$$E \left\{ \int_\gamma^\tau f(x_t) \exp\left(- \int_\gamma^t c(x_r) \, dr\right) \, dt \middle| \mathcal{F}_\gamma \right\} = \int_D g(y)f(y) \, dy. \tag{2.9}$$

Eq. (2.9) is highly nontrivial even if $\gamma = 0$ since the set $G(D, L - c, x)$ is defined by using approximations by operators with smooth coefficients and x_t in (2.9) is just a solution of (2.3) obtained perhaps not as in Remark 2.1, but say by piecing together different solutions.

The following result concerns, basically, the case of piecewise continuous coefficients. We borrow it from Gao (1993) and provide it with a shorter proof.

Theorem 2.16. *There exists an $\varepsilon = \varepsilon(d, \delta, K) > 0$ such that if*

$$\|a(x) - a_{\pm}\|_{I_{\pm x^1} > 0} \leq \varepsilon \tag{2.10}$$

in D (a.e.), where a_{\pm} are some constant symmetric matrices, then weak uniqueness holds for (2.3).

Proof. According to Lorenzi (1972) for any $f \in L_d(\mathbb{R}^d)$ there exists a unique $u \in W_d^2(\mathbb{R}^d)$ such that $\tilde{L}u - u = f$, where $\tilde{L}u = (a_{+}^{ij} I_{x^1 > 0} + a_{-}^{ij} I_{x^1 < 0})u_{x^i x^j}$. Furthermore, there exists a constant $N = N(d, \delta, K)$ such that

$$\|u\|_{W_d^2(\mathbb{R}^d)} \leq N \|\tilde{L}u - u\|_{L_d(\mathbb{R}^d)}.$$

A trivial perturbation argument shows that in the above statements one can replace \tilde{L} with $\tilde{a}^{ij} \partial^2 / \partial x^i \partial x^j$ if \tilde{a} is a Borel symmetric matrix-valued function satisfying (2.10) in \mathbb{R}^d (a.e.) for small enough ε .

By Remark 2.13 weak uniqueness holds for (2.3, \mathbb{R}^d) with $\tilde{a}, 0$ in place of a, b , respectively. Theorem 2.5 allows us to go from \mathbb{R}^d to D and assert that weak uniqueness holds for (2.3, D) with $\tilde{a}, 0$ in place of a, b . By assumption it is possible to take \tilde{a} coinciding with a in D . Since Eq. (2.3) constructed from different sets of coefficients but coinciding in D , obviously have the same sets of solutions, we conclude that weak uniqueness holds for (2.3) with 0 in place of b . After that Remark 2.9 allows us to take any b . The theorem is proved.

In low dimensional cases no continuity of the coefficients is required.

Theorem 2.17. *If $d \leq 2$, then weak uniqueness holds for (2.3).*

This theorem (proved in Krylov, 1969) follows from Theorem 2.5, Remark 2.13, and the fact that in smooth bounded domains in one or two dimensions there is a solvability theory in the Sobolev spaces W_2^2 for elliptic equations with measurable coefficients (see, for instance, Ladyzhenskaya and Ural'tseva, 1964).

In the following two theorems the third condition in (2.1) is not needed. We will see that from their proofs. Recall the notion of pseudo-inverse matrix defined for any symmetric nonnegative matrix γ as

$$\gamma^{(-1)} := \lim_{\varepsilon \downarrow 0} \gamma(I\varepsilon + \gamma)^{-2},$$

where I is the unit matrix of the same dimensions as γ . The reader will understand the idea behind Theorem 2.18 easier if he assumes that $a_{12} = a_{21} = 0$. This theorem generalizes the main result of Cerutti et al. (1996) and also Theorem 4.3 of Bass and Pardoux (1987), where $a(x)$ is uniformly nondegenerate and is independent of x'' ,

so that conditions (2.12) are automatically satisfied. We use Theorem 2.18 to discuss Example 1.1 and therefore it is essential for us to admit degeneracy of a and its dependence on x'' .

Theorem 2.18. *Let $1 \leq d' < d$, $d'' = d - d'$, $D = D' \times \mathbb{R}^{d''}$ with a ball $D' \subset \mathbb{R}^{d'}$. We write $x = (x', x'')$, where $x' \in \mathbb{R}^{d'}$, $x'' \in \mathbb{R}^{d''}$. Represent $a = (a^{ij})_{i,j=1}^d$ as the block matrix*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

where

$$a_{11} = (a^{ij})_{i,j=1,\dots,d'}, \quad a_{21} = (a^{d'+i,j})_{i,j=1}^{d'',d'},$$

$$a_{12} = a_{21}^*, \quad a_{22} = (a^{d'+i,d'+j})_{i,j=1}^{d'',d''}.$$

Assume that $a_{11}(x) = a_{11}(x')$ and $b'(x) = b'(x')$, take an $x'_0 \in D'$, and assume that at x'_0 weak uniqueness holds for the equation

$$x'_t = x'_0 + \int_0^t I_{s < \tau_{D'}(x')} \sigma_{11}(x'_s) dw'_s + \int_0^t I_{s < \tau_{D'}(x')} b'(x'_s) ds, \tag{2.11}$$

where $\sigma_{11} = \sqrt{a_{11}}$. Finally, denote

$$\sigma_{21} = a_{21} \sigma_{11}^{(-1)}, \quad \sigma_{22} = (a_{22} - a_{21} a_{11}^{(-1)} a_{12})^{1/2}$$

and assume that for $j = 1, 2$ and $x' \in D'$, $y'', z'' \in \mathbb{R}^{d''}$ we have

$$\begin{aligned} \|\sigma_{2j}(x', y'') - \sigma_{2j}(x', z'')\| &\leq K|y'' - z''|, & \|\sigma_{2j}(x', y'')\| &\leq K(1 + |y''|), \\ |b''(x', y'') - b''(x', z'')| &\leq K|y'' - z''|, & |b''(x', y'')| &\leq K(1 + |y''|). \end{aligned} \tag{2.12}$$

Then weak uniqueness for (2.3) holds at (x'_0, x''_0) for any $x''_0 \in \mathbb{R}^{d''}$.

Various versions of the next result, which is quite useful, can be found in the literature (see, for instance, Theorem 6.5.2 of Stroock and Varadhan, 1979). We give it with a standard proof in order to make our discussion of Example 1.1 self contained.

Theorem 2.19. *Take an $x_0 \in D$ and let $h = h(x)$ be a real-valued Borel positive function on \mathbb{R}^d bounded along with $1/h$. Then weak uniqueness at x_0 holds for (2.3) if and only if it holds for the equation constructed from ha and hb in place of a and b , respectively.*

Proof. Obviously, it suffices to prove the “only if” part. For $y. \in C$ and $t, s \geq 0$ introduce

$$F(h, t, y.) = F(t, y.) = \int_0^t h(y_r) dr,$$

$$\Phi(h, t, y.) = \Phi(s, y.) = \inf\{t \geq 0 : F(t, y.) \geq s\}.$$

The function $F(t, y.)$ is Borel in $y.$ and continuous and strictly increasing in t . It follows that it is Borel in $(t, y.)$ as well as its time inverse $\Phi(t, y.)$. Below we also use that the function $(t, y.) \rightarrow y_t$ is Borel measurable.

Next let \bar{x}_t be a solution of (2.3) with (ha, hb) and x_0 in place of (a, b) and x , respectively. Then according to well-known results about random time change (see, for instance, Krylov, 1995) the process $x_t := \bar{x}_{\gamma(t)}$ with $\gamma(t) = \Phi(t, \bar{x})$ is a solution of (2.3) in its original form but with a new Wiener process and, of course, x_0 in place of x . It follows that the distribution of x in C is uniquely determined by (a, b, D) . Now observe that the formulas

$$\begin{aligned} \bar{x}_t &= x_{F(t, \bar{x})}, \quad F'(t, \bar{x}) = h(\bar{x}_t) = h(x_{F(t, \bar{x})}), \\ t &= \int_0^t h^{-1}(x_{F(s, \bar{x})})F'(s, \bar{x}) \, ds = \int_0^{F(t, \bar{x})} h^{-1}(x_s) \, ds \end{aligned}$$

imply that $t = F(h^{-1}, F(t, \bar{x}), x)$, $F(t, \bar{x}) = \Phi(h^{-1}, t, x)$, and

$$\bar{x}_t = x_{\Phi(h^{-1}, t, x)}.$$

Thus, \bar{x} is expressed as the result of application of a (deterministic) Borel function to x . This allows us to find the distribution of \bar{x} in terms of a, b, h , and D only and proves the theorem.

3. Discussion of the main results

Our discussion is split into several remarks. Although we believe that each of our main results is interesting in its own right the most interesting are applications of their various combinations. We will see that there are infinitely many combinations in which our main results play role of building blocks. It seems that one cannot formulate a general result encompassing all the combinations and we confine ourselves only to few examples.

Remark 3.1. Here we show that weak uniqueness holds in Example 1.1. We take $D = \mathbb{R}^d$ and use Theorem 2.19 with $h = 1/a^{11}$ to conclude that the whole issue of weak uniqueness in this situation is reduced to the case in which $a^{11} \equiv 1$. After that it suffices to use Theorem 2.18 with $d' = 1$ and notice that the square root of a twice differentiable nonnegative matrix-valued function is Lipschitz continuous (see Freidlin, 1968, Phillips and Sarason, 1968).

Remark 3.2. We comment on Theorem 2.5. Remark 2.13 and the implication (d) \Rightarrow (a) (in Theorem 2.5) show that weak uniqueness holds for (2.3) if a is only continuous rather than uniformly continuous in D . After that from the implication (b) \Rightarrow (a) we see that, actually, a need not be continuous at a fixed point $x_0 \in D$, since if a is continuous in $D \setminus \{x_0\}$, then weak uniqueness holds for any smooth subdomain of $D \setminus \{x_0\}$. Hence the discontinuity of a at only one point does not ruin weak uniqueness.

One can repeat the same argument for the case in which there are only two points of discontinuity, say x_0, y_0 . Then in any smooth subdomain of $Q \subset \bar{Q} \subset D \setminus \{y_0\}$ there can be only one point of discontinuity, by implication (b) \Rightarrow (a), weak uniqueness holds for (2.3, Q) and again by the same implication weak uniqueness holds for (2.3,

D). By induction we obviously can get that weak uniqueness holds if a has any finite number of discontinuity points.

Suppose now that the set of discontinuities of a is countable and has only one cluster point x_0 . Then in any smooth subdomain of $Q \subset \bar{Q} \subset D \setminus \{x_0\}$ there are only finitely many points of discontinuity, so that weak uniqueness holds for (2.3, Q) and hence weak uniqueness holds for (2.3, D).

By using the same argument we can go even further. Namely, if the set of discontinuities of a has only two cluster points, say x_0 and y_0 , then in any smooth subdomain of $Q \subset \bar{Q} \subset D \setminus \{y_0\}$ there can be only one cluster point, by the above, weak uniqueness holds for (2.3, Q) and hence, by the theorem, weak uniqueness holds for (2.3, D). Of course, the case of finitely many cluster points now presents no difficulty as well as the case of countably many cluster points having only one cluster points.

One can keep going in an obvious way considering cluster points of “higher degree”.

Remark 3.3. Safonov (1994) proves that if a is continuous apart from a closed set of α -Hausdorff measure zero and α is small enough, then weak uniqueness holds for L in smooth bounded domains. This result, certainly, covers the result of Remark 3.2 but it does not cover other situations considered below when weak uniqueness is known to hold not because of continuity of the coefficients but for some other reasons.

In this connection it could be interesting to know if the single point $\{0\}$ in part (b) of Theorem 2.5 can be replaced with a set of α -Hausdorff measure zero (without assuming any continuity).

Also notice that there is no control on α , so that we cannot say anything in the general case if the set of discontinuity is a straight line, say x^d -axis even if $d = 3$. However, if the diffusion coefficients of $x' := (x^1, x^2)$ depend only on x' and other coefficients are Lipschitz with respect to x^3 , then weak uniqueness holds as follows from Theorems 2.17 and 2.18.

Remark 3.4. Let Γ be a C^2 surface in \mathbb{R}^d of co-dimension 1, orientable or not. We assume that Γ is “open” in the sense that each of its points has a ball containing it which is split by Γ into two disconnected parts. Let S be a subset of $\Gamma \cap D$ such that S is open in the relative topology of Γ and $\bar{S} \subset \Gamma$. Introduce the distance between $x, y \in D \setminus \bar{S}$ as the infimum of lengths of smooth curves connecting x and y and lying in $D \setminus \bar{S}$. Assume that not only a is continuous in $D \setminus \bar{S}$ in the usual sense but also uniformly continuous with respect to the new distance. It turns out that in this case weak uniqueness holds for (2.3).

To show this we use Theorem 2.16 and the implication (d) \Rightarrow (a) in Theorem 2.5. Since a is continuous in the open set $D \setminus \bar{S}$, Remark 2.13 guarantees that for each $y \in D \setminus \bar{S}$ there is an $r > 0$ such that weak uniqueness holds for (2.3, $B_r(y)$).

If $y \in D$ happens to be on the relative boundary of S , then for small $r > 0$ the new distance between $x_1, x_2 \in B_r(y) \setminus \bar{S}$ is, obviously, smaller than $3r$. Therefore, for any $\varepsilon > 0$ there is a small ball $B_r(y)$ such that $\|a(x_1) - a(x_2)\| \leq \varepsilon$ for all $x_1, x_2 \in B_r(y) \setminus \bar{S}$, in particular, for almost all $x_1, x_2 \in B_r(y)$. Then Theorem 2.16 with $a_{\pm} = a(x_0)$, where x_0 is any fixed point in $B_r(y) \setminus \bar{S}$, yields weak uniqueness for (2.3, $B_r(y)$).

Finally, let $y \in D$ be in the relative interior of S . If for all small r the set $B_r(y) \cap S$ is flat, then weak uniqueness for (2.3, $B_r(y)$) with sufficiently small r follows immediately from Theorem 2.16 and the uniform continuity of a in each halfball of $B_r(y) \setminus S$. In the general case it suffices to do a C^2 smooth one-to-one change of coordinates which flattens S near y and observe that, although Eq. (2.3) will change, this transformation in no way affects weak uniqueness.

We see that for any $y \in D$ there is an $r > 0$ such that weak uniqueness holds for (2.3, $B_r(y)$). Now our assertion follows from Theorem 2.5.

Remark 3.5. Observe that if in the situation of Remark 3.4 we take a $y \in S$, $r > 0$ and replace S and D with $S' = S \setminus \bar{B}_r(y)$ and $D' = D \setminus \bar{B}_r(y)$, respectively, then a is uniformly continuous in $D' \setminus \bar{S}'$ with respect to the distance constructed from D' and S' . This is true since the latter distance is just bigger than the one constructed from D and S .

Remark 3.5 and Theorem 2.5 lead to the following generalization of Remark 3.4.

Remark 3.6. In the situation of Remark 3.4 allow Γ to lose C^2 smoothness at a point $x_0 \in S$ and require a to be uniformly continuous with respect to the new distance only in $(D \setminus \bar{S}) \setminus \bar{B}_r(x_0)$ for any $r > 0$. Then weak uniqueness holds for (2.3).

For instance, if a is uniformly continuous inside and outside of the cone $2|x^1| > |x|$, then weak uniqueness holds for (2.3, \mathbb{R}^d).

If Γ loses smoothness only at two points $x_0, y_0 \in S$ and a is uniformly continuous with respect to the new distance only in $(D \setminus \bar{S}) \setminus (\bar{B}_r(x_0) \cup \bar{B}_r(y_0))$, for each $r > 0$, then again weak uniqueness holds for (2.3). To see this it suffices to repeat the corresponding part of Remark 3.2. Then naturally come the case of infinitely many points of nonsmoothness of Γ having only one cluster point, the case of infinitely many such cluster points having only one cluster point of their own, and so on.

Remark 3.7. In Remark 3.6 the surface Γ could be composed of finitely many pieces whose closures are disjoint. In that case the argument just does not change. Theorem 2.5 allows us to extend the result to Γ consisting of infinitely many components having only one point of attraction, that is a point $x_0 \in D$ such that for any $r > 0$ only finitely many components of Γ are not contained in $B_r(x_0)$. Now one can repeat the same drill going to finitely many points of attraction of components, infinitely many points having only one cluster point and so on.

Remark 3.8. Theorem 2.18 allows us to increase the number of examples in which weak uniqueness holds by using “stratification”. For instance, consider the system consisting of (1.1) and the following equation in a Euclidean space \mathbb{R}^{d_1}

$$y_t = y + \int_0^t \sqrt{\tilde{a}(x_s, y_s)} dB_s + \int_0^t \tilde{b}(x_s, y_s) ds,$$

where \tilde{a} is a symmetric $d_1 \times d_1$ matrix valued function, \tilde{b} an \mathbb{R}^{d_1} -valued function, and B_t a d_1 -dimensional Wiener process independent of w . Assume that \tilde{a} and \tilde{b} are Borel

in (x, y) , Lipschitz continuous in y uniformly with respect to x , that they satisfy linear growth condition with respect to (x, y) , and \tilde{a} is uniformly nondegenerate. Finally, assume that weak uniqueness holds for (1.1) = (2.3, \mathbb{R}^d). Then weak uniqueness holds for our system as well.

Indeed, first observe that the system is written as one equation with block-diagonal diffusion matrix $\begin{pmatrix} a & 0 \\ 0 & \tilde{a} \end{pmatrix}$ and drift term $\begin{pmatrix} b \\ \tilde{b} \end{pmatrix}$. This corresponds to the case $a_{12} = a_{21} = 0$ in Theorem 2.18. Furthermore, $\sigma_{21} = 0$ and $\sigma_{22} = \sqrt{\tilde{a}}$ satisfies the Lipschitz condition in y uniformly with respect to x since the same is true by assumption for \tilde{a} and \tilde{a} is uniformly nondegenerate. After that it only remains to notice that weak uniqueness holds for (2.3, D) for any bounded domain D (Theorem 2.5), then weak uniqueness holds for our system until first exit time from $D \times \mathbb{R}^{d_1}$ by Theorem 2.18, and finally weak uniqueness holds for the system again by Theorem 2.5.

Now we can follow an already familiar pattern. Namely, let \tilde{a}, \tilde{b} be Lipschitz continuous in y only for (x, y) outside of any neighborhood of a closed countable set $\Gamma = \{(x_i, y_i), i = 1, 2, \dots\}$, which has only finitely many cluster points. Of course we allow the Lipschitz constant to depend on the neighborhood. Then, for any neighborhood, the coefficients can be easily modified throughout it so that new functions become uniformly Lipschitz in y in \mathbb{R}^{d_1} . Then weak uniqueness holds for the modified coefficients, by Theorem 2.5 weak uniqueness holds for the original system until first exit time from the complement of the closure of any neighborhood of Γ . After that Theorem 2.5 allows us to conclude that weak uniqueness for our system holds until first exit time from $(\mathbb{R}^d \times \mathbb{R}^{d_1}) \setminus (\Gamma \cup \{(x_1, y_1)\})$ if (x_1, y_1) is an isolated point of Γ . By following the argument of Remark 3.2 we can eliminate all of Γ and therefore weak uniqueness for our system holds again. Of course, Γ could also be a closed set having only countably many cluster points, which in turn have only finitely many cluster points, and so on.

Remark 3.9. The result of Remark 3.8 admits a localization (by Theorem 2.5) so that we get yet another sufficient condition for weak uniqueness until first exit time from domains. Also, it is worth noticing that the above stratification can be repeated few times each time increasing the dimension of the process and each time before and after stratifying one can also use time change (Theorem 2.19) with arbitrarily discontinuous functions h .

Remark 3.10. By having few sufficient conditions for weak uniqueness to hold until first exit time from domains we also have sufficient conditions for weak uniqueness to hold until first exit time from small balls. By relying on statement (d) of Theorem 2.5, and specifying the structure of the coefficients in each ball we can get numerous results on weak uniqueness, which in turn can be localized and fed back into the scheme of balls.

Now, what happens if for each point $y \in D$ but one there is an $r > 0$ such that we can prove that weak uniqueness holds for (2.3, $B_r(y)$)? Theorem 2.5 says that still weak uniqueness holds for (2.3, D). If there are only two “suspicious” points the

result is the same. One can keep going already familiar way and this argument seems to imply that no matter how general but concrete conditions we find on the coefficients guaranteeing that weak uniqueness holds, there always is a possibility to make them even more general.

In conclusion we discuss one of the main results of Bass and Pardoux (1987). This is the following.

Theorem 3.11. *Suppose \mathbb{R}^d can be divided up into finitely many polyhedra so that a is constant in the interior of each polyhedron. Then weak uniqueness holds for (2.3, \mathbb{R}^d).*

Proof. In Bass and Pardoux (1987) this theorem is proved by using stratification and the implication (b) \Rightarrow (a) in Theorem 2.5. Below we, basically, reproduce the argument from Bass and Pardoux (1987) with some shortcuts made possible by our other results.

First, one gets rid of b by using Girsanov’s theorem. Then we use induction on d . If $d=1$ the result follows either from Theorem 2.5, or Theorem 2.17, or else Theorem 2.19 with $h=1/a$. Actually, for $d=1$ the function a has bounded variation, so that the solution is even strong (see Nakao, 1972). Assuming that the result is true for $d-1$ we show that apart from finitely many point for each point $y \in \mathbb{R}^d$ there is an $r > 0$ such that weak uniqueness holds for (2.3, $B_r(y)$). This is enough owing to the above discussion of Theorem 2.5.

To describe the set of points to be excluded we denote by $\{\ell_i, i=1, \dots, k\}$ the finite set of hyperplanes each of which contains at least one of the $d-1$ dimensional faces of the polyhedra. Also let n_i be the unit normal vector to ℓ_i . Then we exclude all points y each of which belongs to the intersection of at least one subset $\{\ell_{i_m}, m=1, \dots, p\}$ of $\{\ell_i, i=1, \dots, k\}$ such that the corresponding set of $\{n_{i_m}, m=1, \dots, p\}$ generates \mathbb{R}^d . Obviously, we exclude only finitely many points (some of which can be even inside of some polyhedra).

Now, in a neighborhood of a point y , lying in the interior of a polyhedron, a is constant and weak (and even strong) uniqueness until first exit time from the neighborhood is trivial.

Let y belong to the boundary of a polyhedron and be not an exceptional point. Then the set $\{\ell_{i_m}, m=1, \dots, p\}$ of all planes containing y has the property that the set $\{n_{i_m}, m=1, \dots, p\}$ does not generate \mathbb{R}^d . Without losing generality we assume that $y=0$, $\{n_{i_m}, m=1, \dots, p\} \subset \mathbb{R}^{d-1} = \{x \in \mathbb{R}^d : x^d = 0\}$ and according to the notation from Theorem 2.18 we write $x = (x', x'')$, where $x'' = x^d$.

Now we claim that if x is close enough to $y=0$, then $a(x)$ depends only on x' . Indeed, for $a(x)$ to change, x should cross one of ℓ_i ’s. However, going in the direction of the x^d -axis, which is perpendicular to $\{n_{i_m}, m=1, \dots, p\}$ and therefore parallel to each of $\ell_{i_m}, m=1, \dots, p$, will certainly not lead to crossing either of them. Other hyperplanes will not be crossed either if we keep $|x'|$ and x'' small enough, since $y=0$ is at a positive distance from each of them.

After having proved our claim, we take $r > 0$ so small that $a(x)$ in $B_r(0)$ depends only on x' and for $|x'| < r$ we introduce $\tilde{a}(x') := a(x', 0)$. Observe that \tilde{a} keeps the same value unless x' with $|x'| < r$ crosses the trace on \mathbb{R}^{d-1} of at least one of $\ell_{i_m}, m=1, \dots, p$.

These traces naturally produce a partition of \mathbb{R}^{d-1} into finitely many (conic) polyhedra and \tilde{a} is constant in the intersection of each of them with $\{|x'| < r\}$. The latter property allows us to extend \tilde{a} to all of \mathbb{R}^{d-1} keeping it constant in each new conic polyhedron. We use the same notation for the extended \tilde{a} .

Next we take $D = B_r(0)$ and observe that (2.3) with this choice of D becomes

$$x_t = x + \int_0^t \sqrt{\tilde{a}(x'_s)} I_{s < \tau_D(x)} \, dw_s.$$

If here in place of D we take \mathbb{R}^d , then we get an equation for which weak uniqueness holds by Theorem 2.18 and the induction hypothesis. By localization statement in Theorem 2.5 weak uniqueness holds for $(2.3, D) = (2.3, B_r(y))$.

As is explained in the beginning of the proof now the assertion of our theorem follows from Theorem 2.5. The theorem is proved.

4. Proof of Theorem 2.18

We need several auxiliary facts which allow us to reduce Eq. (2.3) to a triangular form. This reduction is commonly done in filtering theory (see, for instance, Rozovskii, 1990).

The first assertion of the following lemma is a general property of symmetric non-negative matrices, and the second one is easily proved by passing to the diagonal form.

Lemma 4.1.

- (i) We have $a_{22} \geq a_{21} a_{11}^{(-1)} a_{12}$ in the matrix sense, so that σ_{22} is well defined.
- (ii) For $i = 1, 2$ define $\Pi_i = \sigma_{ii} \sigma_{ii}^{(-1)}$. Then $\Pi_i \sigma_{ii} = \sigma_{ii}$, $\sigma_{ii}^{(-1)} \sigma_{ii} = \Pi_i = \sigma_{ii}^{(-1)} \sigma_{ii}^2 \sigma_{ii}^{(-1)}$, $\sigma_{11}^{(-1)} \sigma_{11}^{(-1)} = a_{11}^{(-1)}$, $\sigma_{11}^{(-1)} a_{11} = \sigma_{11}$.

Next, let x_t be a solution of (2.3) with $x_0 = (x'_0, x''_0)$ in place of x . Take a d -dimensional Wiener process B_t independent of x, w and introduce

$$\begin{aligned} \bar{x}_t &= \int_0^t I_{s < \tau_{D'}(x')} \sqrt{a(x_s)} \, dw_s = x_t - x_0 - \int_0^t I_{s < \tau_{D'}(x')} b(x_s) \, ds, \\ \hat{w}'_t &= \int_0^t I_{s < \tau_{D'}(x')} [\sigma_{11}^{(-1)}(x'_s) \, dx'_s + (I_{d' \times d'} - \Pi_1(x'_s)) \, dB'_s] + \int_0^t I_{s \geq \tau_{D'}(x')} \, dB'_s, \\ \hat{w}''_t &= \int_0^t \sigma_{22}^{(-1)}(x_s) I_{s < \tau_{D'}(x')} [dx''_s - \sigma_{21}(x_s) \, d\hat{w}'_s] \\ &\quad + \int_0^t (I_{d'' \times d''} - \Pi_2(x_s)) I_{s < \tau_{D'}(x')} \, dB''_s + \int_0^t I_{s \geq \tau_{D'}(x')} \, dB''_s, \end{aligned}$$

where and below by $I_{k \times k}$ we mean the unit $k \times k$ matrix.

Lemma 4.2.

- (i) The process $\hat{w}_t = (\hat{w}'_t, \hat{w}''_t)$ is well defined.
- (ii) The processes \bar{x}_t and \hat{w}_t are \mathcal{F}_t -martingales, where \mathcal{F}_t is the completion of $\sigma\{x_s, w_s, B_s : s \leq t\}$.
- (iii) With respect to the filtration $\{\mathcal{F}_t\}$ we have $\langle \hat{w}, \hat{w} \rangle_t = tI_{d \times d}$, so that, by Lévy's theorem, $(\hat{w}_t, \mathcal{F}_t)$ is a Wiener process.

Proof. Assertion (i) means that the stochastic integrals defining \hat{w}_t exist, that is that the corresponding quadratic variation processes are finite. This is proved in the proof of assertion (iii) below, where we compute the quadratic variations. After that assertion (ii) follows from properties of stochastic integrals.

To prove assertion (iii) observe that

$$\begin{aligned} d\langle \hat{w}', \hat{w}' \rangle_t &= \sigma_{11}^{(-1)} I_{t < \tau_{D'}(x')} (d\langle \bar{x}', \bar{x}' \rangle_t) \sigma_{11}^{(-1)} + I_{t < \tau_{D'}(x')} (I_{d' \times d'} - \Pi_1) dt \\ &\quad + I_{t \geq \tau_{D'}(x')} I_{d' \times d'} dt \\ &= \sigma_{11}^{(-1)} I_{t < \tau_{D'}(x')} a_{11} \sigma_{11}^{(-1)} dt + I_{t < \tau_{D'}(x')} (I_{d' \times d'} - \Pi_1) dt \\ &\quad + I_{t \geq \tau_{D'}(x')} I_{d' \times d'} dt \\ &= I_{d'' \times d''} dt, \end{aligned}$$

where and below we drop the argument x_t . Also

$$\begin{aligned} d\langle \hat{w}'', \hat{w}'' \rangle_t &= \sigma_{22}^{(-1)} I_{t < \tau_{D'}(x')} (dV_t) \sigma_{22}^{(-1)} \\ &\quad + (I_{d'' \times d''} - \Pi_2) I_{t < \tau_{D'}(x')} dt + I_{t \geq \tau_{D'}(x')} I_{d'' \times d''} dt, \end{aligned} \tag{4.1}$$

where for $t < \tau_{D'}(x')$

$$\begin{aligned} dV_t &:= (d\bar{x}''_t - \sigma_{21} d\hat{w}'_t)(d\bar{x}''_t - \sigma_{21} d\hat{w}'_t)^* \\ &= d\langle \bar{x}'', \bar{x}'' \rangle_t - \sigma_{21} d\langle \hat{w}', \bar{x}'' \rangle_t - (d\langle \bar{x}'', \hat{w}' \rangle_t) \sigma_{21}^* + \sigma_{21} \sigma_{21}^* dt. \end{aligned}$$

Here, for $t < \tau_{D'}(x')$

$$\begin{aligned} d\langle \bar{x}'', \bar{x}'' \rangle_t &= a_{22} dt, \quad \sigma_{21} d\langle \hat{w}', \bar{x}'' \rangle_t = a_{21} \sigma_{11}^{(-1)} \sigma_{11}^{(-1)} d\langle \bar{x}', \bar{x}'' \rangle_t = a_{21} a_{11}^{(-1)} a_{12} dt, \\ \sigma_{21} \sigma_{21}^* &= a_{21} a_{11}^{(-1)} a_{12}. \end{aligned}$$

It follows that for $t < \tau_{D'}(x')$

$$\sigma_{22}^{(-1)} (dV_t) \sigma_{22}^{(-1)} = \sigma_{22}^{(-1)} (a_{22} - a_{21} a_{11}^{(-1)} a_{12}) \sigma_{22}^{(-1)} dt = \Pi_2 dt,$$

so that, owing to (4.1), $d\langle \hat{w}'', \hat{w}'' \rangle_t = I_{d'' \times d''} dt$. Finally,

$$\begin{aligned} d\langle \hat{w}', \hat{w}'' \rangle_t &= I_{t < \tau_{D'}(x')} \sigma_{11}^{(-1)} [d\langle \bar{x}', \bar{x}'' \rangle_t - (d\langle \bar{x}', \hat{w}' \rangle_t) \sigma_{21}^*] \sigma_{22}^{(-1)} \\ &= I_{t < \tau_{D'}(x')} \sigma_{11}^{(-1)} [a_{21}^* - a_{11} \sigma_{11}^{(-1)} \sigma_{21}^*] \sigma_{22}^{(-1)} dt \end{aligned}$$

and $a_{21}^* - \sigma_{11} \sigma_{21}^* = (I_{d' \times d'} - \Pi_1) a_{21}^*$, so that

$$\sigma_{11}^{(-1)} [a_{21}^* - a_{11} \sigma_{11}^{(-1)} \sigma_{21}^*] = \sigma_{11}^{(-1)} \Pi_1 [a_{21}^* - \sigma_{11} \sigma_{21}^*] = 0$$

and $d\langle \hat{w}', \hat{w}'' \rangle_t = 0$. The lemma is proved.

Lemma 4.3. *We have*

$$x'_t = x'_0 + \int_0^t I_{s < \tau_{D'}(x')} \sigma_{11}(x'_s) d\hat{w}'_s + \int_0^t I_{s < \tau_{D'}(x')} b'(x'_s) ds, \tag{4.2}$$

$$\begin{aligned} x''_t &= x''_0 + \int_0^t I_{s < \tau_{D'}(x')} \sigma_{21}(x_s) d\hat{w}'_s \\ &+ \int_0^t I_{s < \tau_{D'}(x')} \sigma_{22}(x_s) d\hat{w}''_s + \int_0^t I_{s < \tau_{D'}(x')} b''(x_s) ds. \end{aligned} \tag{4.3}$$

Proof. Denote by \hat{x}'_t the stochastic integral in (4.2) and by \hat{x}''_t the sum of the stochastic integrals in (4.3). Notice that

$$\begin{aligned} d\langle \hat{x}', \hat{x}' \rangle_t &= \sigma_{11} I_{t < \tau_{D'}(x')} d\langle \hat{w}', \hat{x}' \rangle_t \\ &= \sigma_{11} I_{t < \tau_{D'}(x')} \sigma_{11}^{(-1)} a_{11} dt = \Pi_1 a_{11} I_{t < \tau_{D'}(x')} dt = a_{11} I_{t < \tau_{D'}(x')} dt. \end{aligned}$$

Also by using some computations from the proof of Lemma 4.2 we obtain that

$$\begin{aligned} d\langle \hat{w}'', \hat{x}'' \rangle_t &= I_{t < \tau_{D'}(x')} \sigma_{22}^{(-1)} [d\langle \hat{x}'', \hat{x}'' \rangle_t - \sigma_{21} d\langle \hat{w}', \hat{x}'' \rangle_t] \\ &= I_{t < \tau_{D'}(x')} \sigma_{22}^{(-1)} (a_{22} - a_{21} a_{11}^{(-1)} a_{12}) dt = I_{t < \tau_{D'}(x')} \sigma_{22} dt, \end{aligned}$$

$$d\langle \hat{w}', \hat{x}'' \rangle_t = I_{t < \tau_{D'}(x')} \sigma_{21}^* dt.$$

This and the fact that x_t stays the same after $\tau_{D'}(x')$ allow us to find easily the quadratic variation of the martingales $\hat{x}'_t - \hat{x}'_t$ and $\hat{x}''_t - \hat{x}''_t$ and show that these are zeros. The lemma is proved.

The following statement and its proof are almost identical to the corresponding arguments in Bass and Pardoux (1987). We give the proof for the sake of completeness and because our situation is yet more general.

Lemma 4.4. *The processes x' and \hat{w}'' are independent.*

Proof. Let $t, h \geq 0, A \in \mathcal{F}_t$, where \mathcal{F}_t is introduced in Lemma 4.2, and let $B \in \mathcal{F}_{\geq t}^{\hat{w}''} := \sigma\{\hat{w}''_r - \hat{w}''_t : r \geq t\}$. Notice that A and $I_B(\hat{w}'_{t+h} - \hat{w}'_t)$ are independent since \hat{w} is a Wiener process relative to \mathcal{F}_t . Also the processes \hat{w}' and \hat{w}'' are independent as different components of the Wiener process \hat{w} . It follows that

$$EI_A I_B(\hat{w}'_{t+h} - \hat{w}'_t) = P(A) EI_B(\hat{w}'_{t+h} - \hat{w}'_t) = 0.$$

In particular, if $C \in \mathcal{F}_{\leq t}^{\hat{w}''} := \sigma\{\hat{w}''_r : r \leq t\} (\subset \mathcal{F}_t)$, then

$$EI_A I_{B \cap C}(\hat{w}'_{t+h} - \hat{w}'_t) = EI_{A \cap C} I_B(\hat{w}'_{t+h} - \hat{w}'_t) = 0.$$

A standard measure-theoretic argument allows us to conclude that for any $F \in \mathcal{F}_{\infty}^{\hat{w}''} := \sigma\{\hat{w}''_r : r \geq 0\}$ such that $P(F) > 0$, we have

$$E_F I_A(\hat{w}'_{t+h} - \hat{w}'_t) = 0,$$

where E_F is the expectation sign with respect to the conditional probability $P_F(\cdot) := P(\cdot | F)$. In other words, $(\hat{w}'_t, \mathcal{F}_t)$ is a martingale with respect to the new measure P_F .

The quadratic variation of \hat{w} . remains the same since it can be defined through the limit of sums of increments and P_F is absolutely continuous with respect to P . It follows that $(\hat{w}_t, \mathcal{F}_t)$ is a Wiener process with respect to the new measure P_F . Since we assumed that weak uniqueness holds for (2.11) and x'_t satisfies (4.2) we conclude that, for each Borel $H \subset C$

$$P_F(x' \in H) = P(x' \in H).$$

This is just another way to state the assertion of the lemma, which is thus proved.

Now we pass directly to the proof of Theorem 2.18. For $n = 0, 1, 2, \dots$ and $t \geq 0$ introduce $x''_t(0) = x'_0$,

$$x''_t(n+1) = x''_0 + \int_0^t I_{s < \tau_{D'}(x'_t)} \sigma_{21}(x'_s, x''_s(n)) d\hat{w}'_s + \int_0^t I_{s < \tau_{D'}(x'_t)} \sigma_{22}(x'_s, x''_s(n)) d\hat{w}''_s + \int_0^t I_{s < \tau_{D'}(x'_t)} b''(x'_s, x''_s(n)) ds,$$

$$n \geq 0.$$

One notices that

$$\int_0^t I_{s < \tau_{D'}(x'_t)} \sigma_{21}(x'_s, x''_s(n)) d\hat{w}'_s = \int_0^t I_{s < \tau_{D'}(x'_t)} a_{21} a_{11}^{(-1)}(x'_s, x''_s(n)) dx'_s$$

and by the induction on n one proves that $x''_t(n)$ is measurable with respect to the completion of $\sigma\{x'_s, \hat{w}''_s : s \leq t\}$, that is, for each t there exists a Borel function f such that $(x'_t, x''_t(n)) = f(x'_t, w''_t)$ (a.s.). Owing to Lemma 4.3 and the fact that the distribution of x'_t is uniquely determined by a_{11} , b' , and D' we conclude that the distribution of $(x'_t, x''_t(n))$ is uniquely determined by a , b , D , and n .

To prove the theorem now it only remains to show that $x''_t(n) \rightarrow x_t$ as $n \rightarrow \infty$ in probability. However, by the inequality $(a + b)^2 \leq 2a^2 + 2b^2$ and Hölder's inequality

$$\begin{aligned} E|x''_t(n+1) - x''_t|^2 &\leq 2E \int_0^t I_{s < \tau_{D'}(x'_t)} [|\sigma_{21}(x'_s, x''_s(n)) - \sigma_{21}(x_s)|^2 \\ &\quad + |\sigma_{22}(x'_s, x''_s(n)) - \sigma_{22}(x_s)|^2] ds \\ &\quad + 2tE \int_0^t I_{s < \tau_{D'}(x'_t)} |b''(x'_s, x''_s(n)) - b''(x_s)|^2 ds \\ &\leq N(1+t)E \int_0^t |x''_s(n) - x''_s|^2 ds, \end{aligned}$$

where N is independent of n and t . By induction this implies that $E|x''_t(n) - x''_t|^2 \leq N^n(1+t)^{2n}/n! \rightarrow 0$ as $n \rightarrow \infty$ and this brings the proof of the theorem to an end.

5. Proof of Theorems 2.11, 2.14, and 2.15

Proof of Theorem 2.14. As usual it suffices to prove (2.8) only for $f \in C^\infty(\bar{D})$. By Lemma 3.1 of Krylov (1992) the assumption of the theorem concerning g means

that g is a Green’s function of $L - c$ in D with pole at x . Next we proceed in three steps.

Step 1: First we prove (5.3), where x_t and τ^n are to be introduced below. Let a_n, b_n, c_n be as in Definition 2.10. Since Green’s functions in D of operators with smooth coefficients are independent of the values of the coefficients outside of D , we either redefine or continue a_n, b_n, c_n outside of D keeping their smoothness in such a way that they satisfy conditions (2.1) in all of \mathbb{R}^d and the convergence $a_n, b_n, c_n \rightarrow a, b, c$ holds in \mathbb{R}^d (a.e.). After that we take x_t^n, w_t^n, x_t, w_t from Remark 2.1 such that $x_t^n \rightarrow x$ and x_t satisfies (1.1).

It is known (see, for instance, Krylov, 1985) that there exists a constant N independent of n such that for any nonnegative Borel f

$$E \int_0^{\tau^n} f(x_t^n) dt \leq N \|f\|_{L_d(D)}, \tag{5.1}$$

where τ^n is the first exit time of x_t^n from D .

Next, let L_n be the elliptic operator corresponding to a_n, b_n and let $g^n(x, y)$ be the Green’s functions of $L_n - c_n$ in D with zero boundary data. Then for any $f \in C^\infty(\bar{D})$ the function

$$u_n(z) := \int_D g_n(z, y) f(y) dy, \quad z \in \bar{D} \tag{5.2}$$

is smooth in \bar{D} and satisfies $L_n u_n - c_n u_n = -f$ in D and $u_n = 0$ on ∂D . It follows by Itô’s formula that

$$u_n(x) = E \int_0^{\tau^n} f(x_t^n) \exp\left(-\int_0^t c_n(x_s^n) ds\right) dt.$$

We pass to the limit in this formula as $n \rightarrow \infty$. Observe that by virtue of (5.1) and the assumed bounded convergence: $c_n \rightarrow c$ (a.e.), the limit will not change if we replace c_n with c . Also use (5.2), the definition of g , and the fact that $x_t^n \rightarrow x_t$ for any t (a.s.). Then we get

$$\int_D g(y) f(y) dy = \lim_{n \rightarrow \infty} E \int_0^{\tau^n} f(x_t) \exp\left(-\int_0^t c(x_s) ds\right) dt. \tag{5.3}$$

Step 2: Now we prove that

$$\lim_{n \rightarrow \infty} E|\tau^n - \tau| = 0, \tag{5.4}$$

where τ is the first exit time of x_t from D . To do this we denote by v_n the smooth solutions of $L_n v_n = -1$ in D with zero boundary condition. By Itô’s formula $v_n(x) = E\tau^n$ and by (5.1) the functions v_n are uniformly bounded. Furthermore, as always, the convergence $x_t^n \rightarrow x_t$ implies that, for any $\varepsilon > 0$,

$$\overline{\lim}_{n \rightarrow \infty} \tau_\varepsilon^n \leq \tau \leq \underline{\lim}_{n \rightarrow \infty} \tau^n,$$

where τ_ε^n is the first exit time of x_t^n from $D(\varepsilon) := \{z \in D : \text{dist}(z, \partial D) > \varepsilon\}$. In particular, by Fatou’s lemma $E\tau < \infty$ and, since $(\tau - \tau^n)_+ \leq \tau$,

$$\overline{\lim}_{n \rightarrow \infty} E(\tau - \tau^n)_+ \leq E \overline{\lim}_{n \rightarrow \infty} (\tau - \tau^n)_+ = E(\tau - \underline{\lim}_{n \rightarrow \infty} \tau^n)_+ = 0.$$

It follows that to prove (5.4) it suffices to show that

$$\lim_{n \rightarrow \infty} E\tau^n = E\tau. \tag{5.5}$$

By Fatou’s lemma, for any constant $T, \varepsilon \in (0, \infty)$,

$$\overline{\lim}_{n \rightarrow \infty} E(T \wedge \tau_\varepsilon^n) \leq E\tau \leq \underline{\lim}_{n \rightarrow \infty} E\tau^n. \tag{5.6}$$

Itô’s formula yields

$$E\tau^n = v_n(x) = E\tau_\varepsilon^n + Ev_n(x_{\tau_\varepsilon^n}^n), \quad E(\tau^n - \tau_\varepsilon^n) \leq \sup\{v_n, \partial D(\varepsilon)\}.$$

By Lemma 5.1 below we conclude that $\sup_n E(\tau^n - \tau_\varepsilon^n) \rightarrow 0$ as $\varepsilon \downarrow 0$. Hence, from (5.6) we infer that

$$\overline{\lim}_{n \rightarrow \infty} E(T \wedge \tau^n) \leq E\tau \leq \underline{\lim}_{n \rightarrow \infty} E\tau^n. \tag{5.7}$$

Eq. (5.5) would now follow if we knew that τ^n are uniformly integrable or even more that $\sup_n E(\tau^n)^2 < \infty$.

However, denote by V_n the smooth solution of $L_n V_n = -v_n$ with zero boundary condition. Then v_n and V_n are uniformly bounded owing to (5.1). By Itô’s formula

$$Ev_n(x_t^n)I_{t < \tau^n} = -E[v_n(x_{\tau^n}^n) - v_n(x_t^n)]I_{t < \tau^n} = E(\tau^n - t)I_{t < \tau^n},$$

$$V_n(x) = E \int_0^{\tau^n} v_n(x_t^n) dt = E \int_0^{\tau^n} (\tau^n - t) dt = (1/2)E(\tau^n)^2$$

and hence $E(\tau^n)^2$ are bounded indeed. Therefore, by letting $T \rightarrow \infty$ in (5.7) we get (5.5) and this yields (5.4).

Step 3: Eqs. (5.4) and (5.3) lead to (2.8). It only remains to note that (2.8) will not change if we replace x_t with $y_t := x_{t \wedge \tau_Q(x)}$ and that y_t is a solution of (2.3) with Q in place of D . The theorem is proved.

Proof of Theorem 2.15. Denote by $P_\gamma = P_\gamma(\cdot) = P_\gamma(\omega, \cdot)$ a regular conditional distribution of x . on C given \mathcal{F}_γ . In other words,

- (i) $P_\gamma(\omega, B)$ is a probability measure on Borel subsets B of C for any ω ,
- (ii) $P_\gamma(\omega, B)$ is \mathcal{F}_γ -measurable in ω for any Borel $B \subset C$,
- (iii) for any nonnegative or bounded Borel function h given on $[0, \infty) \times C$, we have

$$E\{h(\gamma, x.) | \mathcal{F}_\gamma\} = \int_C h(\gamma, y.) P_\gamma(dy.) \quad (\text{a.s.}).$$

Convenient theorems guaranteeing existence of regular conditional probabilities can be found, for instance, in Stroock and Varadhan (1979) and Krylov (1973a,b).

It is proved in Krylov, 1977, that for any bounded Borel f given on D

$$E \left\{ \int_\gamma^\tau f(x_t) \exp\left(-\int_\gamma^t c(x_r) dr\right) dt | \mathcal{F}_\gamma \right\} \leq N \|f\|_{L_d(D)} \quad (\text{a.s.}),$$

where N is independent of f and ω . A standard measure-theoretic argument allows us to infer that there is a set Ω' having full probability such that for $\omega \in \Omega'$ and bounded Borel f given on D we have

$$\int_C \int_{\gamma(\omega)}^{\tau_D(y)} f(y_t) \exp\left(-\int_{\gamma(\omega)}^t c(y_r) dr\right) dt P_\gamma(\omega, dy) \leq N \|f\|_{L_d(D)}.$$

By the Riesz representation theorem it follows that for $\omega \in \Omega'$ there exists a nonnegative $g(y) = g(\omega, y)$ with $L_{d/(d-1)}(D)$ -norm bounded by the constant N and such that for any bounded Borel f given on D Eq. (2.9) holds. This proves assertion (b) of the theorem.

To prove (a) observe that by Itô’s formula for any smooth u vanishing on ∂D , we have (a.s.)

$$E \left\{ \int_\gamma^\tau (c - L)u(x_t) \exp\left(-\int_\gamma^t c(y_r) dr\right) dt \middle| \mathcal{F}_\gamma \right\} = u(x_\gamma).$$

Hence, owing to assertion (b) of the theorem, perhaps for somewhat reduced Ω' but still having full probability we get that for any $\omega \in \Omega'$ and any smooth u vanishing on ∂D

$$u(x_\gamma) = \int_D g(y)(c - L)u(y) dy.$$

By Lemma 3.1 of Krylov (1992) this means that for $\omega \in \Omega'$ the function $g(\omega, y)$ is a Green’s function of $L - c$ in D with pole at $x_\gamma(\omega)$. The theorem is proved.

Proof of Theorem 2.11. The last statement of the theorem can be proved by a rather easy modification of the proof for the case that D is a smooth bounded domain. As a matter of fact, for $D = \mathbb{R}^d$ the equivalence of (a) and (b) is proved as Theorem 3 in Krylov (1985) and the equivalence of (b), (c), and (d) can be proved as this is done for bounded domains in Krylov (1992).

The proof for smooth bounded D consists of three lemmas. First we recall that the implication (a) \Rightarrow (c) is proved in Theorem 2.14 and we state a standard and well-known result, which is usually proved by using simple barriers.

Lemma 5.1. *Let D be a bounded C^∞ -domain in \mathbb{R}^d and let a, b, c be functions of class $C^\infty(\bar{D})$ satisfying (2.5). Let $u \in C^\infty(\bar{D})$ be a solution of $Lu - cu = -f$ in D with zero boundary condition, where f is a smooth bounded function. Then $|u(x)| \leq N \text{dist}(x, \partial D) \sup |f|$, where the constant N depends only on δ, K , and D .*

The next result is just a restatement of Lemma 4.1 from Krylov (1992) and part of Theorem 2.1 from Krylov (1992).

Lemma 5.2. *Weak uniqueness holds for the operator $L - 1$ in D if and only if it holds for $L - c$ in D and if and only if it holds at 0 for $L - c$ in D . In particular, (b) \Leftrightarrow (c) \Leftrightarrow (d) in Theorem 2.11.*

We now see that to finish proving Theorem 2.11, it suffices to prove the following lemma.

Lemma 5.3. *Assume that for each constant $\lambda \geq 0$, weak uniqueness holds for the operator $L - \lambda$ in D . Then for Eq. (2.3) weak uniqueness holds in D .*

Proof. We closely follow the corresponding argument in Krylov (1985). Take an $x \in D$, a solution x_t of (2.3) and use the notation from Theorem 2.15. Also denote by $g_\lambda(z, y)$ the (unique) Green’s function of $L - \lambda$ in D with pole at z and introduce $G_\lambda f(z) = \int_D g_\lambda(z, y) f(y) dy$. Then by Theorem 2.15

$$\begin{aligned} & I_{t < \tau} E \left\{ \int_t^\tau e^{-\lambda(s-t)} f(x_s) ds \mid \mathcal{F}_t \right\} \\ &= E \left\{ \int_{t \wedge \tau}^\tau e^{-\lambda(s-t \wedge \tau)} f(x_s) ds \mid \mathcal{F}_{t \wedge \tau} \right\} = G_\lambda f(x_{t \wedge \tau}) = I_{t < \tau} G_\lambda f(x_t) \quad (\text{a.s.}) \end{aligned}$$

Also notice that $G_\lambda f(z)$ is Borel (actually, Hölder continuous, see, for instance, Krylov, 1992; Safonov, 1994).

In particular, for any $n \geq 1$, partition $0 = t_0 \leq t_1 \leq \dots \leq t_n = t$ and bounded continuous f_0, \dots, f_n, f given on D

$$\begin{aligned} & \int_0^\infty e^{-\lambda s} E f_0(x_{t_0}) \cdots \cdots f_n(x_{t_n}) f(x_{t+s}) I_{t+s < \tau} ds \\ &= E f_0(x_{t_0}) \cdots \cdots f_n(x_{t_n}) G_\lambda f(x_{t_n}) I_{t_n < \tau}. \end{aligned}$$

By using the induction on n and uniqueness of the Laplace transform of continuous functions we easily conclude that the expressions $E f_0(x_{t_0}) \cdots \cdots f_n(x_{t_n}) I_{t_n < \tau}$ will not change if we take a different solution of (2.3). The same is true for

$$\begin{aligned} & E f_0(x_{t_0}) \cdots \cdots f_n(x_{t_n}) I_{t_n < \tau \leq t_{n+1}} \\ &= E f_0(x_{t_0}) \cdots \cdots f_n(x_{t_n}) I_{t_n < \tau} - E f_0(x_{t_0}) \cdots \cdots f_n(x_{t_n}) I_{t_{n+1} < \tau}, \end{aligned}$$

where $t_{n+1} \geq t_n$, and for

$$\begin{aligned} & E f_0(x_{t_0}) \cdots \cdots f_n(x_{t_n}) I_{t_n < \tau} h(x_\tau) \\ &= \lim_{r \rightarrow \infty} \sum_{i=0}^\infty E f_0(x_{t_0}) \cdots \cdots f_n(x_{t_n}) I_{t_n+i/r < \tau \leq t_n+(i+1)/r} h(x_{t_n+i/r}) \end{aligned}$$

if h is bounded and continuous. Hence the formula

$$\begin{aligned} & E f_0(x_{t_0 \wedge \tau}) \cdots \cdots f_n(x_{t_n \wedge \tau}) \\ &= \sum_{i=0}^{n-1} E f_0(x_{t_0}) \cdots \cdots f_i(x_{t_i}) I_{t_i < \tau \leq t_{i+1}} h_i(x_\tau) + E f_0(x_{t_0}) \cdots \cdots f_n(x_{t_n}) I_{t_n < \tau}, \end{aligned}$$

where $h_i = f_{i+1} \cdots \cdots f_n$ for $i < n$, shows that the distribution of $x_\cdot = x_{\cdot \wedge \tau}$ is the same for all solutions of (2.3) and this finishes the proof of the lemma and Theorem 2.11. \square

6. Proof of Theorem 2.5

Theorem 2.1 of Krylov (1992) says that, if D is smooth and bounded, then the following statements are equivalent:

- (a') weak uniqueness holds for $L - 1$ in D ,
- (b') weak uniqueness holds for $L - 1$ in any smooth domain $Q \subset \bar{Q} \subset D \setminus \{0\}$,
- (c') weak uniqueness holds for $L - 1$ in D at 0.

Now, Theorem 2.14 shows that if (a) (of the theorem) is true, then weak uniqueness holds for $L - 1$ in Q , where Q is any bounded smooth subdomain of D . Then, the implication (a') \Rightarrow (b') entails that (a) \Rightarrow (b'), which along with Theorem 2.11 yields

$$(a) \Rightarrow (b).$$

Furthermore, by Theorem 2.11 (b) \Leftrightarrow (b'), which along with (b') \Rightarrow (a') proves that if (b) is true, then weak uniqueness holds for (2.3) with any smooth bounded domain in place of D . Remark 2.4 now allows us to conclude that

$$(b) \Rightarrow (a) \Rightarrow (c).$$

Next, owing to Theorem 2.14, (c) implies that weak uniqueness holds for $L - 1$ in Q at 0, where Q is any smooth bounded subdomain of D . The implication (c') \Rightarrow (a') shows that if (c) is true, then weak uniqueness holds for $L - 1$ in Q , which by Theorem 2.11 and Remark 2.4 leads to

$$(c) \Rightarrow (a).$$

Also notice that due to the above proved fact that (a) \Rightarrow (b), we obviously have

$$(a) \Rightarrow (d). \tag{6.1}$$

To finish with combining obvious implications, observe that by virtue of Theorem 2.11 and Remark 2.4 to prove the remaining implication (d) \Rightarrow (a) it suffices to prove that for bounded smooth D we have (d) \Rightarrow (a'). In other words, it only remains to check that weak uniqueness holds for $L - 1$ in a smooth bounded D under the additional assumption that

(A) for each $y \in D$ there is an open ball $B_r(y)$ such that weak uniqueness holds for (2.3, $B_r(y)$).

Below (A) is supposed to be fulfilled. Without losing generality (c.f. (6.1)) we may also assume that for the balls in (A) we have $\bar{B}_r(y) \subset D$.

For $f \in C^\infty(\bar{D})$ define

$$G(f, x) = \sup_{g \in G(D, L-1, x)} \int_D g(y) f(y) \, dy.$$

We need to prove that, for any $x \in D$, the set $G(D, L - 1, x)$ is a singleton, which obviously is equivalent to the fact that

$$G(f, x) = \inf_{g \in G(D, L-1, x)} \int_D g(y) f(y) \, dy = -G(-f, x) \tag{6.2}$$

for any $x \in D$ and $f \in C^\infty(\bar{D})$.

Fix an $f \in C^\infty(\bar{D})$. By Lemma 3.2 of Krylov (1992) there is a Borel function $\bar{g}(x, y)$ defined for $x, y \in \bar{D}$ such that, for any $x \in \bar{D}$, we have

$$\bar{g}(x, \cdot) \in G(D, L - 1, x), \quad \int_D \bar{g}(x, y) f(y) dy = G(f, x).$$

Furthermore, $G(f, x)$ is a (Hölder) continuous function in \bar{D} . Therefore, there exists an $x_0 \in \bar{D}$ such that

$$G(f, x_0) + G(-f, x_0) = \max_{x \in \bar{D}} [G(f, x) + G(-f, x)] =: m.$$

If $x_0 \in \partial D$, then $G(f, x_0) = G(-f, x_0) = 0$, $m = 0$.

In the second case $x_0 \in D$ and there exists $r_0 > 0$ such that $\bar{B}_{r_0}(x_0) \subset D$ and weak uniqueness holds for (2.3, $B_{r_0}(x_0)$). By Theorem 2.14

$$\begin{aligned} G(f, x_0) &= \int_D \bar{g}(x_0, y) f(y) dy = E \int_0^\tau f(x_t) e^{-t} dt \\ &= E \int_0^{\tau_0} f(x_t) e^{-t} dt + E e^{-\tau_0} E \left\{ \int_{\tau_0}^\tau f(x_t) e^{-(t-\tau_0)} dt \mid \mathcal{F}_{\tau_0} \right\}, \end{aligned}$$

where x_t is a solution of (2.3) with initial condition x_0 , τ and τ_0 are first exit times of x_t from D and $B_{r_0}(x_0)$, respectively. Invoking also Theorem 2.15 we get

$$G(f, x_0) \leq E \int_0^{\tau_0} f(x_t) e^{-t} dt + E e^{-\tau_0} G(f, x_{\tau_0}). \tag{6.3}$$

Observe that $x_{t \wedge \tau_0}$ is a solution of (2.3) with x_0 and $B_{r_0}(x_0)$ in place of x and D , respectively. Therefore, by assumption (A) and the choice of r_0 it is weakly unique. We conclude that, for Borel bounded functions ψ and ϕ the quantities

$$E \int_0^{\tau_0} \psi(x_t) e^{-t} dt \quad \text{and} \quad E e^{-\tau_0} \phi(x_{\tau_0})$$

are independent of what solution x_t we take. Therefore, by applying (6.3) to $-f$ in place of f , then adding up the results and recalling the definition of m we obtain

$$m \leq E e^{-\tau_0} [G(f, x_{\tau_0}) + G(-f, x_{\tau_0})] \leq m E e^{-\tau_0}.$$

Since $E e^{-\tau_0} < 1$, we get $m = 0$. Thus, $G(f, x) + G(-f, x) \leq 0$ on \bar{D} . Since obviously $G(f_1, x) + G(f_2, x) \geq G(f_1 + f_2, x)$, we conclude $0 \leq G(f, x) + G(-f, x) \leq 0$. This is equivalent to (6.2) and the theorem is proved. \square

7. Concluding remarks

Here we briefly discuss two facts known to the author for at least a quarter of a century which he could not apply in any single case. However, they still shed some additional light on the problem. The first fact is a necessary and sufficient condition for weak uniqueness and the second one says that if we slightly “shake” the coefficients, the solutions of any equation become weakly unique.

Remark 7.1. According to Krylov (1973a,b) there is at least one Markov process in \mathbb{R}^d corresponding to the operator L . If a is bounded and $b(x) = -x/|x|$ for large x , then as

easy to see any Markov process corresponding to L has a unique invariant probability measure. Due to Alexandrov’s estimate these measures are absolutely continuous. Let μ be one of them. Then for the corresponding semigroup T_t and any $u \in C_0^\infty(\mathbb{R}^d)$ we have $\int (T_t u - u) \mu(dx) = 0$ for $t \geq 0$. By dividing this equation by t and letting $t \downarrow 0$, we easily get that

$$\int_{\mathbb{R}^d} Lu(x) \mu(dx) = 0. \tag{7.1}$$

It turns out that, for our coefficients, weak uniqueness holds for (1.1) if and only if there is only one probability measure such that (7.1) holds for all $u \in C_0^\infty(\mathbb{R}^d)$.

Indeed, if μ satisfies (7.1), then according to Theorem 2 of Krylov (1985), μ is an invariant measure of a Markov process corresponding to L . This, obviously, proves the “only if” part.

On the other hand, assume that there is only one probability measure μ satisfying (7.1) for all $u \in C_0^\infty(\mathbb{R}^d)$. Fix $u \in C_0^\infty(\mathbb{R}^d)$ and $\lambda > 0$ and take Markov solutions of (1.1) for which $R_\lambda u(x) := E \int_0^\infty e^{-\lambda t} u(x_t) dt = \int_0^\infty e^{-\lambda t} T_t u(x) dt$ is maximal at all $x \in \mathbb{R}^d$ or minimal at all $x \in \mathbb{R}^d$ between all solutions of (1.1). That such Markov processes exist follows from Krylov (1973a,b) or Krylov (1986). Their invariant measures by assumption coincide with μ . Hence, $\int R_\lambda u \mu(dx)$, which equals $\lambda^{-1} \int u \mu(dx)$, is uniquely defined no matter $R_\lambda u$ is maximal or minimal. Since one of them is always bigger, we get that they coincide μ -(a.s.). Furthermore, $R_\lambda u$ are Hölder continuous (see Remark 1.1 in Krylov, 1986), and obviously, μ charges any ball. Hence $R_\lambda u(x)$ is the same at all points for any solution of (1.1). Now one concludes as in the proof of Lemma 5.3.

Remark 7.2. Consider the following case of equation with time dependent coefficients

$$x_t = \int_0^t \sqrt{a(s, x_s + f_s)} dw_s + \int_0^t b(s, x_s + f_s) ds \tag{7.2}$$

assuming that f is a deterministic function belonging to C and a and b are Borel functions satisfying (2.1), with $a(t, x)$ and $b(t, x)$ in place of $a(x)$ and $b(x)$, for all t, x, λ . It turns out that for any $T, \varepsilon > 0$ there exists an f such that $|f_t| \leq \varepsilon$ and weak uniqueness holds for solutions of (7.2) on the time interval $[0, T]$.

To explain this, take a Borel bounded real-valued function $h(t, x)$ equal to zero for large t and introduce

$$\pi^f \cdot h = \sup E \exp \int_0^\infty h(t, x_t + f_t) dt,$$

where the sup is taken over all solutions of (7.2) given on all possible probability spaces. By using the same arguments as in Krylov (1973a, b) we see that $\pi^f \cdot h$ is a Borel (upper semicontinuous) function of f and there exists a Borel function \mathbb{P}^f on C with values in the set of probability measures on C such that for each f , \mathbb{P}^f is the distribution of a solution of (7.2) and

$$\pi^f \cdot h = \int_C \left(\exp \int_0^\infty h(t, y_t + f_t) dt \right) \mathbb{P}^f(dy). \tag{7.3}$$

Now let B_t be a d -dimensional Wiener process given on a probability space (Ω, \mathcal{F}, P) and let $\nu > 0$ be a constant. At this point it is convenient to remember the martingale characterization of solutions of (7.2). Then it is easy to see that with respect to the measure $\mathbb{P}^{\nu B, (\omega)}(dy)P(d\omega)$ on $\Omega \times C$ the processes $x_t = x_t(\omega, y) := y_t$ and $z_t = z_t(\omega, y) := x_t + \nu B_t(\omega)$ satisfy

$$x_t = \int_0^t \sqrt{a(s, x_s + \nu B_s)} dw_s + \int_0^t b(s, x_s + \nu B_s) ds,$$

$$z_t = \int_0^t \sqrt{a(s, z_s)} dw_s + \nu B_t + \int_0^t b(s, z_s) ds,$$

where w is a Wiener process independent of B . Furthermore,

$$E\pi^{\nu B} h = E \exp \int_0^\infty h(t, z_t) dt. \tag{7.4}$$

However, the process z_t/ν has diffusion as close to the unit matrix as we wish if we just take ν large enough. We fix such a ν that according to Stroock and Varadhan (1979) the distribution of z is uniquely defined by a and b . Then the right-hand side of (7.4) is uniquely defined as well. Now if instead of sup we take inf in (7.3) we will come again to (7.4), implying that, for almost any trajectory f of νB ,

$$E \exp \int_0^\infty h(t, x_t + f_t) dt$$

is independent of what solution of (7.2) we take. This conclusion we have for each particular h . By using the separability of C and $C_0^\infty(\mathbb{R}^d)$, we obtain that for almost any trajectory f of νB , the distribution of x is independent of which solution of (7.2) we take. It only remains to observe that with positive probability $\nu|B_t| \leq \varepsilon$ for $t \in [0, T]$.

References

Bass, R., Pardoux, E., 1987. Uniqueness for diffusions with piecewise constant coefficients. *Probab. Theory Related Fields* 76, 557–572.

Bramanti, M., Cerutti, M.C., 1993. $W_p^{1,2}$ solvability for the Cauchy-Dirichlet problem for parabolic equations with VMO coefficients. *Comm. Partial Differential Equations* 18 (9–10), 1735–1763.

Cerutti, M.C., Escauriaza, L., Fabes, E.B., 1993. Uniqueness in the Dirichlet problem for some elliptic operators with discontinuous coefficients. *Ann. Mat. Pura Appl.* 163 (4), 161–180.

Cerutti, M.C., Fabes, E.B., Manselli, P., 1996. Uniqueness for elliptic equations with time-independent coefficients. *Progress in Elliptic and Parabolic Partial Differential Equations, Capri, 1994, Pitman Res. Notes Math. Ser.*, 350, Longman, Harlow, pp. 112–135.

Freidlin, M.I., 1968. The factorization of nonnegative definite matrices. *Teor. Veroyatnost. i Primenen.* 13 (2), 375–378 (in Russian).

Gao, P., 1993. The martingale problem for a differential operator with piecewise continuous coefficients. *Seminar on Stochastic Processes, Seattle, WA, 1992, Progress in Probability, Vol. 33, Birkhäuser, Boston, MA*, pp. 135–141.

Gilbarg, D., Trudinger, N.S., 1983. *Elliptic partial differential equations of second order*, 2nd Edition. *Grundlehren der Mathematischen Wissenschaften, Vol. 224, Springer, Berlin*.

Ikeda, N., Watanabe, S., 1981. *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam.

- Khasminskii, R., Krylov, N.V., 2001. On averaging principle for diffusion processes with null-recurrent fast component. *Stochastic Process Appl.* 93 (2), 229–240.
- Krylov, N.V., 1969. Diffusion in the plane with reflection. Construction of the process. *Sibirski Mat. J.* 10 (2), 343–354 (in Russian) (English translation in *Siberian Math. J.* 10 (2), 244–252).
- Krylov, N.V., 1973a. On the regularity of the conditional probability for stochastic processes. *Teoriya Veroyatnostei i eye Primeneniya* 18 (1), 151–155 (in Russian) (English translation: *Theor. Probab. Appl.* 18(1), 150–153).
- Krylov, N.V., 1973b. On the selection of a Markov process from a system of processes and the construction of quasi-diffusion processes. *Izv. Akad. Nauk SSSR Ser. Mat.* 37 (3), 691–708 (in Russian) (English translation in *Math. USSR Izvestija* 7(3), 691–709).
- Krylov, N.V., 1977. *Controlled Diffusion Processes*. Nauka, Moscow (in Russian) (English translation: Springer, Berlin, 1980).
- Krylov, N.V., 1985. Once more about the connection between elliptic operators and Itô's stochastic equations. In: *Statistics and Control of Stoch. Proceedings of the Steklov Seminar 1984*; Optimiz. Software Inc., Publications Division, New York, 1985, pp. 214–229.
- Krylov, N.V., 1986. An approach in the theory of controlled diffusion processes. *Teoriya Veroyatnostei i eye Primeneniya* 31 (4), 685–709 (in Russian) (English translation: *Theor. Probab. Appl.* 31(4), (1987) 604–626).
- Krylov, N.V., 1992. On one-point weak uniqueness for elliptic equations. *Comm. PDE* 17, 1759–1784.
- Krylov, N.V., 1995. *Introduction to the Theory of Diffusion Processes*. American Mathematical Society, Providence, RI.
- Ladyzhenskaya, O.A., Ural'tseva, N.N., 1964. *Linear and Quasi-Linear Elliptic Equations*. Nauka, Moscow (in Russian) (English translation: Academic Press, New York, 1968).
- Lorenzi, A., 1972. On elliptic equations with piecewise constant coefficients II. *Ann. Scuola Norm. Sup. Pisa* 26 (3), 839–870.
- Nakao, S., 1972. On the pathwise uniqueness of solutions of one-dimensional stochastic differential equations. *Osaka J. Math.* 9, 513–518.
- Nadirashvili, N.S., 1997. Nonuniqueness in the martingale problem and the Dirichlet problem for uniformly elliptic operators. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* 24 (4), 537–550.
- Phillips, R.S., Sarason, L., 1968. Elliptic-parabolic equations of the second order. *J. Math. Mech.* 17 (2), 891–917.
- Rozovskii, B.L., 1990. *Stochastic Evolution Systems*. Kluwer, Dordrecht.
- Safonov, M.V., 1994. On a weak uniqueness for some elliptic equations. *Comm. Partial Differential Equations* 19 (5–6), 943–957.
- Safonov, M.V., 1999. Nonuniqueness for second-order elliptic equations with measurable coefficients. *SIAM J. Math. Anal.* 30 (4), 879–895.
- Stroock, D.W., Varadhan, S.R.S., 1979. *Multidimensional diffusion processes*. Springer, New York.