Zvonkin’s transform and the regularity of solutions to double divergence form elliptic equations

Vladimir I. Bogachev\textsuperscript{a,b}, Michael Röckner\textsuperscript{c}, and Stanislav V. Shaposhnikov\textsuperscript{a,b}

\textsuperscript{a}Department of Mechanics and Mathematics, Moscow State University, Moscow, Russia; \textsuperscript{b}National Research University Higher School of Economics, Moscow, Russia; \textsuperscript{c}Fakultät für Mathematik, Universität Bielefeld, Bielefeld, Germany

\textbf{ABSTRACT}

We study qualitative properties of solutions to double divergence form elliptic equations (or stationary Kolmogorov equations) on $\mathbb{R}^d$. It is shown that the Harnack inequality holds for nonnegative solutions if the diffusion matrix $A$ is nondegenerate and satisfies the Dini mean oscillation condition and the drift coefficient $b$ is locally integrable to some power $p > d$. We establish new estimates for the $L^p$-norms of solutions and obtain a generalization of the known theorem of Hasminskii on the existence of a probability solution to the stationary Kolmogorov equation to the case where the matrix $A$ satisfies Dini’s condition or belongs to the class VMO. These results are based on a new analytic version of Zvonkin’s transform of the drift coefficient.

\textbf{1. Introduction}

We study qualitative properties of solutions to the double divergence form elliptic equation (or the stationary Kolmogorov equation)

$$\partial_x \partial_y (a^{ij} \varphi) - \partial_x (b^i \varphi) = 0$$

on an open set $\Omega \subset \mathbb{R}^d$. The matrix $A = (a^{ij})$ is supposed to be symmetric and positive definite, $a^{ij}$ and $b^i$ are Borel functions. Set

$$L \varphi = a^{ij} \partial_x \partial_y \varphi + b^i \partial_x \varphi, \quad L^* \varphi = \partial_x \partial_y (a^{ij} \varphi) - \partial_x (b^i \varphi).$$

Then Equation (1.1) can be written in a shorter form

$$L^* \varphi = 0.$$

A function $\varphi$ in the class $L^1_{\text{loc}}(\Omega)$ of locally integrable functions on $\Omega$ is called a solution to Equation (1.1) if

$$a^{ij} \varphi, b^i \varphi \in L^1_{\text{loc}}(\Omega)$$

and for every function $\varphi \in C_0^\infty(\Omega)$ the equality

$$L^* \varphi = 0$$

holds.
\[
\int_{\Omega} L\varphi(x)\varrho(x) \, dx = 0
\]
is fulfilled. A nonnegative solution \( \varrho \) to the Kolmogorov equation (1.1) satisfying the condition
\[
\int_{\Omega} \varrho(x) \, dx = 1
\]
is called a probability solution.

An important example of a double divergence form elliptic equation is delivered by the stationary Kolmogorov equation for invariant measures of a diffusion process. Various properties of solutions to such equations were studied by many authors. The principal questions are

1. the existence of solutions, especially, of probability solutions,
2. the existence of solution densities and their properties such as local boundedness, continuity and Sobolev differentiability,
3. local separation of densities from zero, that is, certain forms of the Harnack inequality.

In case of locally Lipschitz coefficients the existence of a probability solution is given by the classical theorem of Hasminskii [1] under the existence of a Lyapunov function. This theorem was generalized in [2] (see also further generalizations in [3, Chapter 2], [4,5]), where either the diffusion coefficient is nondegenerate and locally Sobolev with the order of integrability higher than dimension along with the same local integrability of the drift coefficient or both the diffusion and drift coefficients are continuous. It was shown in [6] that the solution density is locally Sobolev in the first case and its continuous version is locally separated from zero. It was proved in [7] and [8] that in the case where the matrix \( A = (a^{ij}) \) is nondegenerate and satisfies Dini’s condition and the coefficients \( b^i \) are bounded, the solution has a continuous version, and when the coefficients \( a^{ii} \) are Hölder continuous, then the solution has a Hölder continuous version. These results have been generalized in [9] to the case of integrable \( b^i \). Analogous results have been obtained in [10] and [11] under the assumption that the matrix \( A \) satisfies the Dini mean oscillation condition, which is weaker than the classical Dini condition. Note also the paper [12], where some additional regularity of solutions along level sets has been established. Apart from the study of stationary distributions of diffusion processes, an important motivation for investigation of double divergence form elliptic equations comes from consideration of the properties of Green’s functions for elliptic operators with irregular coefficients (see, e.g., [13,14], and [15]). In the papers [13,14] some interesting counter-examples were constructed and the so-called renormalized solutions were studied, in particular, an example was constructed of a positive definite and continuous diffusion matrix \( A \) for which the equation \( \partial_{x_k} \partial_{x_l} (a^{il} \varrho) = 0 \) has a locally unbounded solution. The Harnack inequality for double divergence form equations with the matrix \( A \) belonging to the Sobolev class with a sufficiently high order of integrability is a corollary of the Harnack inequality for divergence form elliptic equations (see [3, Chapter 3]). The paper [16] gives a survey of the qualitative theory of divergence and
nondivergence elliptic equations including the maximum principle and the Harnack inequality. However, in case of a merely Hölder continuous matrix $A$ the double divergence form equation cannot be reduced to a divergence form equation, moreover, the classical results about the regularity of solutions to divergence form elliptic equations are not true for solutions to double divergence form equations. In the case where the matrix $A$ satisfies Dini’s condition, the Harnack inequality was obtained in [17] for $b = 0$, and for any bounded drift $b$ it was established in [9]. The proof consisted in obtaining certain inequalities for solutions generalizing classical mean value theorems and heavily used the boundedness of $b$. Another way of proving the Harnack inequality for $b = 0$ was suggested in [11], where the reasoning employs estimates for the modulus of continuity of the solution and some properties of renormalized solutions from the paper [18] considering double divergence form equations without first order terms. In [6] and [9] (see also [3, Chapter 1]) the integrability of solutions was investigated in some cases when the diffusion matrix does not satisfy Dini’s condition. In particular, it was shown that if $A$ belongs to the class $VMO$ and the coefficient $b$ is locally integrable to some power $p > d$, then the solution belongs to all $L^p_{loc}(\Omega)$. In spite of a considerable number of papers devoted to double divergence form elliptic equations, the answers to the following questions, certain specifications of general problems 1–3 mentioned above, have remained open so far:

- What are optimal conditions for the Harnack inequality for nonnegative solutions to double divergence form equations? In particular, does the Harnack inequality hold for $A$ satisfying Dini’s condition and an unbounded locally Lebesgue integrable drift $b$?
- What are optimal conditions for a high local integrability of solutions? In particular, the dependence of the integrability of the solution on the modulus of continuity of the matrix $A$ has not been studied.
- What are optimal conditions for the existence and uniqueness of a probability solution to the stationary Kolmogorov equation?

Here we obtain new results related to these questions.

i. We prove that the Harnack inequality holds for nonnegative solutions on $\mathbb{R}^d$ if the diffusion matrix $A$ is nondegenerate and satisfies the Dini mean oscillation condition and the drift coefficient $b$ is locally integrable to some power $p > d$.

ii. We establish new estimates for the $L^p$-norms of solutions and obtain sufficient conditions for the local exponential integrability. Note that it was asserted in [9] that in the case of a locally bounded coefficient $b$ and a nondegenerate matrix $A$ of class VMO the solution is locally exponentially integrable. However, the justification given there contains a gap, namely, a wrong dependence on $p$ of the constant in a priori $L^p$-estimates of second derivatives of solutions to the equation $\text{tr}(AD^2u) = f$. In the general case, the dependence of the constant on $p$ is influenced by the modulus of continuity of $A$. In the present paper we derive an estimate that takes the modulus of continuity of $A$ into account.
iii. Finally, an important new result of our paper is a generalization of the known theorem of Hasminskii on the existence of a probability solution to the stationary Kolmogorov equation to the case where the matrix $A$ belongs to the class $\text{VMO}$. Results on existence of positive or probability solutions to the stationary Kolmogorov equation in case of irregular coefficients are useful for constructing diffusion processes (see [5]). We also discuss uniqueness of probability solutions and their probabilistic interpretation.

These results are obtained with the aid of a new approach to the study of regularity of solutions to double divergence form equations based on Zvonkin’s transform, well known in the theory of diffusion processes, which applies for smoothing the drift coefficient (more precisely, we deal with its elliptic version, in the original paper [19] this transform was used for parabolic equations). In recent years Zvonkin’s transform has been applied for the study of diffusion processes with generalized coefficients (see [20–22]). In this paper we apply Zvonkin’s transform not to random processes, but to solutions of the Kolmogorov equation, moreover, we do not assume any connection of solutions with diffusion processes. It is shown below that with the aid of a suitable change of coordinates an integrable drift can be transformed into a continuously differentiable drift such that the new diffusion matrix enables us to apply known results about regularity of solutions. This leads to substantial generalizations of some results and simplification of proofs of other results on regularity of solutions to double divergence form elliptic equations. Note that change of coordinates has proved to be also useful in the study of uniqueness of solutions (see [23]).

Finally, we note that double divergence form equations play a major role in the study of various nonlinear problems, for example, nonlinear Fokker–Planck–Kolmogorov equations, (see [24]), systems of equations in mean field games (see [25,26]), in homogenization of differential equations (see [27]).

In §2 we construct Zvonkin’s transform in the analytic setting, in §3 we study the regularity of solutions, and in §4 we apply our results for proving existence and uniqueness of probability solutions.

### 2. Zvonkin’s transform

We first illustrate our approach by example of a smooth change of coordinates.

Let $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ be a diffeomorphism of class $\mathcal{C}^2$ and $\Psi = \Phi^{-1}$. Set

$$q^{km}(y) = a_{ij}(\Psi(y)) \partial_{x_i} \Phi^k(\Psi(y)) \partial_{x_j} \Phi^m(\Psi(y)),$$

$$h^k(y) = a_{ij}(\Psi(y)) \partial_{x_i} \partial_{x_j} \Phi^k(\Psi(y)) + b^i(\Psi(y)) \partial_{x_i} \Phi^i(\Psi(y)),$$

$$r(y) = q(\Psi(y)) |\det \nabla \Psi(y)|.$$

Then on the domain $\Omega' = \Phi(\Omega)$ the function $\sigma$ satisfies the equation $L^* \sigma = 0$ with the operator

$$L\varphi(y) = q^{km}(y) \partial_{y_k} \partial_{y_m} \varphi(y) + h^k(y) \partial_{y_k} \varphi(y).$$

Observe that the double divergence structure of the equation does not change. One can construct a mapping $\Phi$ of the form $\Phi(x) = x + u(x)$, where $u = (u^1, ..., u^d)$, such that $u^k$ is a solution to the elliptic equation
\[ Lu^k - b^k. \]

For \( \lambda \) sufficiently large, this equation possesses a solution for which \( \Phi \) is a diffeomorphism. Using \( \Phi \) to change variables we obtain a new drift coefficient \( h(y) = \lambda u(\Psi(y)) \).

It turns out that under fairly general assumptions about the coefficients (see below) the vector field \( h \) is continuously differentiable and the regularity of the matrix \( (q_{km}) \) is not worse than that of the original matrix \( (a_{ij}) \). This circumstance enables us after the change of coordinates to apply the known results on the regularity of solutions and obtain the desired properties for the function \( r \), hence also for original solution \( u \). The main difficulty consists in constructing the mapping \( u \). Here we employ some recent results of N.V. Krylov on the solvability of elliptic equations in Sobolev spaces (see [28] and [29]).

Under broad assumptions, we construct a diffeomorphism, which will be called Zvonkin’s transform.

Throughout this section we assume that the following conditions are fulfilled.

\( H_a \) The coefficients \( a_{ij} \) are defined on all of \( \mathbb{R}^d \) and for some constant \( \nu > 0 \) and all \( x \in \mathbb{R}^d \) the following inequalities hold:

\[
\nu \cdot I \leq A(x) \leq \frac{1}{\nu} \cdot I,
\]

where \( I \) is the unit matrix.

\( VMO \) The coefficients \( a_{ij} \) belong to the class \( VMO \), that is, there exists a continuous increasing function \( \omega \) on \([0, +\infty)\) such that \( \omega(0) = 0 \) and

\[
\sup_{z \in \mathbb{R}^d} r^{-2d} \int_{B(z, r)} \int_{B(z, r)} |a_{ij}(x) - a_{ij}(y)| \, dx \, dy \leq \omega(r), \quad r > 0.
\]

\( H_b \) \( b \in L_{loc}^{d+} \), which means that for every ball \( B \) there is a number \( p = p(B) > d \) such that the restriction of \( |b| \) to \( B \) belongs to \( L^p(B) \).

The assumption that the diffusion matrix satisfies the aforementioned conditions on the whole space \( \mathbb{R}^d \) does not restrict the generality of our considerations, although the equation will be considered on a domain. Moreover, in many problems it is useful to have global changes of variables rather than local. For our purposes of proving local Harnack inequalities or the continuity of solutions it suffices to extend the coefficients on the whole space with preservation of the required conditions. The drift coefficient can be extended by zero outside a fixed ball \( B \) and the diffusion coefficient can be extended by the formula \( \psi A + (1 - \psi)I \) with a smooth function \( \psi \) that equals 1 on \( B \) and 0 outside a larger ball. Of course, it is important here that the equation holds on a ball, but not on the whole space.

The standard inner product and norm in \( \mathbb{R}^d \) are denoted by \( \langle x, y \rangle \) and \( |x| \), the standard Lebesgue measure of a set \( E \subset \mathbb{R}^d \) is denoted by \( |E| \). Let \( B(x_0, R) \) denote the open ball of radius \( R \) centered at \( x_0 \).

Let \( B(x_0, 4R) \subset \Omega \). Set \( \beta(x) = b(x) \) if \( x \in B(x_0, 4R) \) and \( \beta(x) = 0 \) if \( x \not\in B(x_0, 4R) \). Then \( \beta \in L^p(\mathbb{R}^d) \) and
\[ \|\beta\|_{L^p(\mathbb{R}^d)} = \|b\|_{L^p(B(x_0,4R))}. \]

Let \( 1 \leq k \leq d \). Let us consider on \( \mathbb{R}^d \) the elliptic equation
\[ \text{tr}(AD^2u) + \langle \beta, \nabla u \rangle - \lambda u = -\beta^k, \quad \lambda > 0. \] (2.1)

**Proposition 2.1.** For every \( \delta > 0 \) there exists \( \lambda > 0 \) such that for every \( k \leq d \) Equation (2.1) has a solution \( u \in C^1(\mathbb{R}^d) \cap W^{p,2}(\mathbb{R}^d) \) for which
\[ \sup_{x \in \mathbb{R}^d} |\nabla u(x)| \leq \delta, \quad \|u\|_{W^{p,2}(\mathbb{R}^d)} \leq M, \]
where the constant \( M \) depends only on \( d, \nu, \omega \) and \( \|b\|_{L^p(B(x_0,4R))} \).

**Proof.** According to [29, Chapter 6, Section 4, Theorem 1], there exist numbers \( \lambda_0 > 0 \) and \( N_0 \) such that for all \( \lambda > \lambda_0 \) and every function \( v \in W^{p,2}(\mathbb{R}^d) \) we have the inequality
\[ \lambda\|v\|_{L^p(\mathbb{R}^d)} + \|\nabla v\|_{W^{p,2}(\mathbb{R}^d)} \leq N_0(\text{tr}(AD^2v) - \lambda v\|_{L^p(\mathbb{R}^d)}). \]

For every function \( f \in L^p(\mathbb{R}^d) \) there exists a unique solution \( v \in W^{p,2}(\mathbb{R}^d) \) of the equation
\[ \text{tr}(AD^2v) - \lambda v = f. \]

Since \( p > d \), by the embedding theorem for every function \( v \in W^{p,2}(\mathbb{R}^d) \) one has the estimate
\[ \|\nabla v\|_{L^\infty(\mathbb{R}^d)} \leq N_1\|v\|_{W^{p,2}(\mathbb{R}^d)}. \]

By [29, Chapter 1, Section 5, Corollary 2] there exists a constant \( N_2 \) such that
\[ \|\nabla v\|_{L^p(\mathbb{R}^d)} \leq N_2\|D^2v\|_{L^p(\mathbb{R}^d)} + N_2\|v\|_{L^p(\mathbb{R}^d)}. \]

Therefore,
\[ \|\nabla v\|_{L^\infty(\mathbb{R}^d)} \leq N_3\|D^2v\|_{L^p(\mathbb{R}^d)} + N_3\|v\|_{L^p(\mathbb{R}^d)}. \]

Let \( \varepsilon > 0 \). By a standard reasoning, replacing \( x \) with \( \varepsilon x \), we obtain the inequality
\[ \|\nabla v\|_{L^\infty(\mathbb{R}^d)} \leq \varepsilon^{1-d/p}N_3\|D^2v\|_{L^p(\mathbb{R}^d)} + \varepsilon^{-1-d/p}N_3\|v\|_{L^p(\mathbb{R}^d)}. \]

Thus, we can assume that for every \( \varepsilon > 0 \) there exists a constant \( N_4 = N_4(\varepsilon) \) for which
\[ \|\nabla v\|_{L^\infty(\mathbb{R}^d)} \leq \varepsilon\|v\|_{W^{p,2}(\mathbb{R}^d)} + N_4\|v\|_{L^p(\mathbb{R}^d)}. \]

Let us now estimate the expression
\[ \|\langle \beta, \nabla v \rangle\|_{L^p(\mathbb{R}^d)}. \]

We have
\[ \|\langle \beta, \nabla v \rangle\|_{L^p(\mathbb{R}^d)} \leq N_4\|\beta\|_{L^p(\mathbb{R}^d)}\|v\|_{L^p(\mathbb{R}^d)} + \varepsilon\|\beta\|_{L^p(\mathbb{R}^d)}\|v\|_{W^{p,2}(\mathbb{R}^d)}. \]

Take \( \varepsilon_0 > 0 \) and \( \lambda_1 \geq \lambda_0 \) such that
\[ \varepsilon_0N_0\|\beta\|_{L^p(\mathbb{R}^d)} < \frac{1}{2}, \quad N_4N_0\|\beta\|_{L^p(\mathbb{R}^d)} < \frac{\lambda_1}{2}. \]
Then for every $k$ and $v \in W^{p,2}(\mathbb{R}^d)$ we have

$$\lambda \|v\|_{L^p(\mathbb{R}^d)} + \|v\|_{W^{p,2}(\mathbb{R}^d)} \leq 2N_0 \|\beta\|_{L^p(\mathbb{R}^d)}.$$

This estimate remains unchanged if we replace $\beta$ by $t\beta$, where $t \in [0,1]$. Set

$$L_t v = t(\text{tr}(AD^2 v) + \langle \beta, \nabla v \rangle - \lambda v) + (1-t)(\text{tr}(AD^2 v) - \lambda v), \quad t \in [0,1].$$

The continuous operators $L_t$ from $W^{p,2}(\mathbb{R}^d)$ to $L^p(\mathbb{R}^d)$ satisfy the condition

$$\|L_t v\|_{L^p(\mathbb{R}^d)} \geq \frac{1}{2N_0} \|v\|_{W^{p,2}(\mathbb{R}^d)}.$$

Hence the standard method of continuation with respect to a parameter ensures the existence of a solution $u$ to Equation (2.1). By the embedding theorem $u \in C^1(\mathbb{R}^d)$. In addition,

$$\|u\|_{L^p(\mathbb{R}^d)} \leq \frac{2}{\lambda} N_0 \|\beta\|_{L^p(\mathbb{R}^d)}, \quad \|u\|_{W^{p,2}(\mathbb{R}^d)} \leq 2N_0 \|\beta\|_{L^p(\mathbb{R}^d)}.$$

It has been shown above that for every $\varepsilon > 0$ we have

$$\|\nabla u\|_{L^\infty(\mathbb{R}^d)} \leq \varepsilon \|u\|_{W^{p,2}(\mathbb{R}^d)} + N_4 \|u\|_{L^p(\mathbb{R}^d)},$$

where the right-hand side is estimated from above by

$$2\varepsilon N_0 \|\beta\|_{L^p(\mathbb{R}^d)} + \frac{2}{\lambda} N_0 N_4 \|\beta\|_{L^p(\mathbb{R}^d)}.$$

Therefore, taking $\varepsilon$ sufficiently small and $\lambda$ sufficiently large we can obtain the desired estimate $\|\nabla u\|_{L^\infty(\mathbb{R}^d)} \leq \delta$. \hfill \(\square\)

Remark 2.2. Since $u \in W^{p,2}(\mathbb{R}^d)$ with $p > d$, by the embedding theorem $u \in C^{1+(1-d/p)}(\mathbb{R}^d)$, hence the function $u$ is bounded and the derivatives $\partial_i u$ satisfy the Hölder condition of order $1 - d/p$. Note also that our construction of $u$ does not use a special form of $\beta$: only the condition $\beta \in L^p(\mathbb{R}^d)$ is needed.

Let $u = (u^1, \ldots, u^d)$, where each $u^k$ is a solution to Equation (2.1) from Proposition 2.1. Below $u'$ and $\Phi'$ denote the Jacobi matrices of the mappings $u$ and $\Phi$. Let us take a number $\delta$ from the hypotheses of Proposition 2.1 such that for all $x \in \mathbb{R}^d$ the inequality

$$\|u'(x)\| \leq \frac{1}{2} \quad \text{and} \quad \frac{1}{2} \leq \det(I + u'(x)) \leq 2$$

is fulfilled. Set

$$\Phi(x) = x + u(x).$$

We now establish some properties of the mapping $\Phi$.

**Proposition 2.3.**

i. The mapping $\Phi$ is a diffeomorphism of $\mathbb{R}^d$ of class $C^1$, moreover, the functions $\partial_i \Phi^k$ are locally Hölder continuous.
ii. The inequalities
\[ \frac{1}{2} |x - y| \leq |\Phi(x) - \Phi(y)| \leq 2|x - y| \]
hold.

**Proof.** Since \( u \) is a contracting mapping with the Lipschitz constant 1/2, the mapping \( \Phi \) is a homeomorphism and the stated inequalities hold. The inclusion \( u \in C^1(\mathbb{R}^d) \) yields that \( \Phi \in C^1(\mathbb{R}^d) \) and the Jacobi matrix has the form \( \Phi' = I + u' \). By the estimate \( \|u'(x)\| \leq 1/2 \) the matrix \( \Phi' \) is invertible. Therefore, \( \Phi \) is a diffeomorphism. \( \square \)

We recall that \( B(x_0, 4R) \subset \Omega \). Let \( \Psi = \Phi^{-1} \) and \( y_0 = \Phi(x_0) \) and consider the ball \( B(y_0, 2R) \). According to the inequalities in (ii) of **Proposition 2.3**, we have the inclusions
\[ B(x_0, R) \subset \Psi(B(y_0, 2R)) \subset B(x_0, 4R). \]

**Proposition 2.4.** Let \( \varrho \in L^1_{\text{loc}}(\Omega) \) be a solution to **Equation (1.1)**. Then the function
\[ \sigma(y) = |\det \Psi'(y)| \varrho(\Psi(y)) \]
on \( B(y_0, 2R) \) satisfies the equation \( \mathcal{L}^* \sigma = 0 \), where
\[ \mathcal{L}f(y) = q^{km}(y)\partial_{y_k} \partial_{y_m} f(y) + h^k(y) \partial_{y_k} f(y) \]
and the coefficients have the form
\[ q^{km}(y) = a^{ij}(\Psi(y)) \partial_{x_i} \Phi^k(\Psi(y)) \partial_{x_j} \Phi^m(\Psi(y)), \quad h^k(y) = \lambda u^k(\Psi(y)). \]

**Proof.** The function \( \varrho \) belongs to \( L^r(B(x_0, 4R)) \) for all \( r \in [1, d/(d - 1)] \) (see the comments at the beginning of the proof of [9, Theorem 2.1]). In particular, for \( p > d \) the value \( p/(p - 1) \) is less than \( d/(d - 1) \) and the inclusion \( \varrho \in L^{p/(p - 1)}(B(x_0, 4R)) \) holds. Therefore, in the integral equality determining the solution, in place of test functions of class \( C^\infty_0(\Omega) \) we can substitute test functions of class \( W^{p, 2}(\Omega) \) with compact support in \( B(x_0, 4R) \). For every function \( \varphi \in C^\infty_0(B(y_0, 2R)) \) (outside the ball \( B(y_0, 2R) \) we always extend \( \varphi \) by zero), the function \( \varphi \circ \Phi \) belongs to \( W^{p, 2}(B(x_0, 4R)) \), has compact support in \( B(x_0, 4R) \) and satisfies the equality
\[
\int_{B(x_0, 4R)} \left( \left[ a^{ij}(x) \partial_{x_i} \Phi^k(x) \partial_{x_j} \Phi^m(x) \right] \partial_{y_k} \partial_{y_m} \varphi(\Phi(x)) + \left[ a^{ij}(x) \partial_{x_i} \partial_{x_j} \Phi^k(x) + b^i(x) \partial_{x_i} \Phi^k(x) \right] \partial_{y_k} \varphi(\Phi(x)) \right) \varrho(x) \, dx = 0.
\]
Since \( b = \beta \) on \( B(x_0, 4R) \) and \( \Phi^k = x_k + u^k(x) \), we have
\[ a^{ij}(x) \partial_{x_i} \partial_{x_j} \Phi^k(x) + b^i(x) \partial_{x_i} \Phi^k(x) = \lambda u^k(x). \]
Then
\[
\int_{B(x_0, 4R)} \left( q^{km}(\Phi(x)) \partial_{y_k} \partial_{y_m} \varphi(\Phi(x)) + \lambda u^k(x) \partial_{y_k} \varphi(\Phi(x)) \right) \varrho(x) \, dx = 0.
\]
Using the change of variable \( y = \Phi(x) \) and taking into account that the support of \( \varphi \) belongs to \( B(y_0, 2R) \), we obtain

\[
\int_{B(y_0, 2R)} (q^{km}(y)\partial_{xk}\partial_{y\mu}\varphi(y) + \lambda u^k(\Psi(y))\partial_{xk}\varphi(y))q(\Psi(y))|\det\Psi'(y)|\ dy = 0.
\]

Since \( \varphi \) was arbitrary, we conclude that \( \sigma(y) = q(\Psi(y))|\det\Psi'(y)| \) is a solution to the equation \( \mathcal{L}^*\sigma = 0 \). \( \square \)

Observe that the vector field \( h(y) = \lambda u(\Psi(y)) \) is continuously differentiable on the ball \( B(y_0, 2R) \). In addition, the derivatives of \( \Phi \) also satisfy the Hölder condition. Therefore, the function \( \sigma \) on \( B(y_0, 2R) \) satisfies the equation \( \mathcal{L}^*\sigma = 0 \), in which the coefficients \( q^{mk} \) of the second order terms form a nondegenerate matrix and belong to the class \( \text{VMO} \) and the coefficients \( h^k \) are continuous on \( B(y_0, 2R) \). This enables us to apply the results from the papers [7, 9, 11], and [8] to the function \( \sigma \) and then to transfer these results to \( q \). Let us give an example demonstrating a simple derivation of the known result of [9, Theorem 3.1]) from the case of a nice drift.

We recall that a mapping satisfies Dini’s condition if for its modulus of continuity \( \omega \) we have

\[
\int_0^1 \frac{\omega(t)}{t} \ dt < \infty.
\]

Example 2.5. If conditions \( H_a \) and \( H_b \) are fulfilled and the matrix \( A \) satisfies Dini’s condition, then every solution \( q \in L^1_{\text{loc}}(\Omega) \) to Equation (1.1) has a continuous version.

Proof. Let us fix a ball \( B(x_0, R/2) \subset B(x_0, 4R) \subset \Omega \) and verify the existence of a continuous version of \( q \) on the ball \( B(x_0, R/2) \). Let \( \Phi \) be the diffeomorphism constructed above. By Proposition 2.4 the function \( \sigma(y) = q(\Psi(y))|\det\Psi'(y)| \) satisfies on \( B(y_0, 2R) \) an equation with some coefficients for which the hypotheses of [8, Theorem 1] are fulfilled, that is, the matrix \( (q^{mk}) \) is nondegenerate and the functions \( q^{mk}, h^k \) satisfy Dini’s condition. Hence \( \sigma \) has a continuous version on \( B(y_0, R) \). Since \( \Phi \) is a diffeomorphism of class \( C^1 \), the mappings \( \Phi \) and \( \Psi \) take measure zero sets to measure zero sets and a modification of the function \( \sigma \) on a measure zero set yields a change of \( q \) on a measure zero set. Therefore, the function \( q \) has a continuous version on \( B(x_0, R/2) \). \( \square \)

Note that on the ball \( B(x_0, R/2) \subset B(x_0, 4R) \subset \Omega \) the modulus of continuity of the solution \( q \) depends only on \( d, p, R, \omega, \nu, \) and \( \|b\|_{L/(B(x_0, 4R))} \).

In a similar way, by using [7, Theorem 2] one can derive the Hölder continuity of the solution, provided that the functions \( a^j \) are Hölder continuous. However, unlike [9, Theorem 3.1], this method does not ensure the Hölder order of the solution to be equal to the Hölder order of the matrix \( A \), because the expression for the coefficients \( q^{mk} \) involves the derivatives of the mapping \( \Phi \), but their Hölder order depends on \( d \) and \( p \).

3. Regularity of solutions

In this section we apply Zvonkin’s transform for establishing the regularity of solutions. We first discuss the case where the matrix \( A \) satisfies the classical Dini condition, then
consider the Dini mean oscillation condition, and finally study the integrability of solutions without the assumption about Dini’s condition.

The next assertion generalizes the Harnack inequality to the case where the diffusion matrix satisfies Dini’s condition and the drift coefficient is locally unbounded (and is merely integrable to some power larger than the dimension). In the known results, the drift coefficient is either zero or locally bounded, which has been substantially used in the proofs.

**Theorem 3.1.** If \( a^j \) and \( b^i \) satisfy conditions \( H_a \) and \( H_b \) and the matrix \( A \) satisfies Dini’s condition, then the continuous version of every nonnegative solution \( \varrho \in L^1_{\text{loc}}(\Omega) \) to Equation (1.1) satisfies the Harnack inequality, that is, for every ball \( B(x_0, R/2) \subset B(x_0, 4R) \subset \Omega \) there exists a number \( C \) such that

\[
\sup_{x \in B(x_0, R/2)} \varrho(x) \leq C \inf_{x \in B(x_0, R/2)} \varrho(x),
\]

where \( C \) depends on \( R, \omega, d, \nu, p \), and \( \|b\|_{L^p(B(x_0, 4R))} \), but does not depend on the solution \( \varrho \).

**Proof.** Let \( B(x_0, R/2) \subset B(x_0, 4R) \subset \Omega \) and let \( \Phi \) be the diffeomorphism constructed above. As above, according to Proposition 2.4 the function

\[
\sigma(y) = \varrho(\Psi(y))|\det \Psi'(y)|
\]

on \( B(y_0, 2R) \) satisfies the equation with coefficients for which the hypotheses of [9, Corollary 3.6] are fulfilled, i.e., the matrix \( (q_{mk}) \) is nondegenerate, the functions \( q_{mk} \) satisfy Dini’s condition and the functions \( h^k \) are bounded. Therefore, there exists a number \( C \) depending on the objects listed above such that

\[
\sup_{y \in B(y_0, R)} \sigma(y) \leq C \inf_{y \in B(y_0, R)} \sigma(y).
\]

Since \( 2^{-1} \leq |\det \Psi'(y)| \leq 2 \), we have

\[
\sup_{y \in B(y_0, R)} \varrho(\Psi(y)) \leq 4C \inf_{y \in B(y_0, R)} \varrho(\Psi(y)).
\]

By the inclusion \( B(x_0, R/2) \subset \Psi(B(y_0, R)) \) we have

\[
\sup_{x \in B(x_0, R/2)} \varrho(x) \leq \sup_{y \in B(y_0, R)} \varrho(\Psi(y)), \quad \inf_{x \in B(x_0, R/2)} \varrho(x) \leq \inf_{y \in B(y_0, R)} \varrho(\Psi(y)).
\]

Therefore, \( \sup_{x \in B(x_0, R/2)} \varrho(x) \leq 4C \inf_{x \in B(x_0, R/2)} \varrho(x) \).

**Remark 3.2.** Using the method suggested in [30] and [31], increasing the dimension, it is possible to add a potential term to the drift coefficient. Let \( \varrho \) be a solution to the equation

\[
\partial_{x_i} \partial_{x_j} (a^i j \varrho) - \partial_{x_i} (b^i \varrho) + c \varrho = 0
\]

on \( \mathbb{R}^d \). Then on \( \mathbb{R}^{d+1} = \mathbb{R}^d \times \mathbb{R}^1 \) the function \( \varrho \) satisfies the equation

\[
\partial_{x_i} \partial_{x_j} (a^i j \varrho) + \partial^2_{x_i} \varrho - \partial_{x_i} (b^i \varrho) - \partial_y (cy \varrho) = 0.
\]
Using this approach, one can apply the result obtained above to the equation with the potential term $c \varrho$, but in this case it is necessary to assume a higher integrability of the coefficients: $b', c \in L^p_{\text{loc}}(\Omega)$ with $p > d + 1$.

Note also that the equation with a nonzero drift coefficient $b$ can be transformed in a similar way into an equation without the drift. Let $\varrho$ be a solution to Equation (1.1) on $\mathbb{R}^d$. Then on $\mathbb{R}^d \times (0, 1) \subset \mathbb{R}^{d+1}$ the function $\varrho$ satisfies the equation

$$\partial_x \partial_y (a^y \varrho) + M \partial_y^2 \varrho - \partial_y \partial_x \left( \left( 2 + \frac{y}{2} \right) b^y \varrho \right) - \partial_x \partial_y \left( 2 + \frac{y}{2} \right) b^y \varrho = 0.$$ 

The new matrix $\tilde{A}$ has the form

$$\begin{pmatrix}
a^{11} & a^{12} & \cdots & -\left( 2 + \frac{y}{2} \right) b^1 \\
a^{21} & a^{22} & \cdots & -\left( 2 + \frac{y}{2} \right) b^2 \\
\vdots & \vdots & \ddots & \vdots \\
-\left( 2 + \frac{y}{2} \right) b^1 & -\left( 2 + \frac{y}{2} \right) b^2 & \cdots & M
\end{pmatrix}.$$ 

For every $\tilde{\xi} \in \mathbb{R}^{d+1}$ we have the equality

$$\langle \tilde{A} \tilde{\xi}, \tilde{\xi} \rangle = \sum_{i,j=1}^d a^{ij} \tilde{\xi}_{i} \tilde{\xi}_{j} - \sum_{j=1}^d (4 + y)b^j \tilde{\xi}_j \tilde{\xi}_{d+1} + M \tilde{\xi}_{d+1}^2.$$ 

If the functions $b^i$ are bounded and $A \geq \nu \cdot I$, then for $M$ sufficiently large the matrix $\tilde{A}$ is positive definite. A certain drawback of such transformations is the necessity to impose on $b$ the same restrictions as on $a^y$, for example, to require the continuity and Dini’s condition. In the next remark we show how one can accomplish smoothing of coefficients $b$ and $c$ with the aid of renormalization of solutions and Zvonkin’s transform.

**Remark 3.3.** Let $\varrho \in L^1_{\text{loc}}(\Omega)$ be a solution to Equation (3.1), where $c \in L^p_{\text{loc}}(\Omega)$ for some $p > d$, the coefficients $a^y$, $b^i$ satisfy conditions $H_a$ and $H_b$ and the matrix $A$ is continuous. Let $B(y_0, 4R) \subset \Omega$ and let $\Phi$ be the diffeomorphism constructed before Proposition 2.3. Set $y_0 = \Phi(x_0)$ and $\Psi = \Phi^{-1}$. Similarly, to Proposition 2.4, one verifies that the function

$$\sigma(y) = |\det \Psi'(y)| \varrho(\Psi(y))$$

satisfies the equation $L^* \sigma + g \sigma = 0$ on $B(y_0, 2R)$, where

$$Lf(y) = q^{km}(y) \partial_{y_k} \partial_{y_m} f(y) + h^k(y) \partial_{y_k} f(y)$$

and the coefficients have the form

$$q^{km}(y) = a^{ij}(\Psi(y)) \partial_{y_i} \Phi^k(\Psi(y)) \partial_{y_j} \Phi^m(\Psi(y)), \quad h^k(y) = \lambda u^k(\Psi(y)), \quad g(y) = c(\Psi(y)).$$

Let us observe that $h^k$ is a continuously differentiable function on $B(y_0, 2R)$, $g \in L^p(B(y_0, 2R))$, $Q = (q^{km})$ satisfies condition $H_a$, and the function $q^{km}$ is continuous. Let $\gamma > 0$. 

Let us consider the Dirichlet problem
\[ \mathcal{L}u + (g - \gamma)u = 0 \quad \text{on} \quad B(y_0, R), \quad u = 1 \quad \text{on} \quad \partial B(y_0, R). \] (3.2)

Since we do not assume that the coefficient \( g \) is bounded from above, for completeness we give a short justification of the existence of a positive solution under the condition that the number \( \gamma \) is sufficiently large. It is clear that it suffices to consider the Dirichlet problem with zero boundary condition and some right-hand side. Set \( B = B(y_0, R) \). By [32, Theorem 9.14 and Theorem 9.15] (see also [29, Chapter 8]) there exists \( \gamma_0 > 0 \) such that for every \( \gamma > \gamma_0 \) the Dirichlet problem \( \mathcal{L}v - \gamma v = f \) on \( B \) with \( v = 0 \) on \( \partial B \) has a solution \( v \in W^{p,2}(B) \cap W^{p,1}_0(B) \) for every function \( f \in L^p(B) \). In addition, for all \( \gamma > \gamma_0 \) and \( v \in W^{p,2}(B) \cap W^{p,1}_0(B) \) we have the estimate
\[ \gamma \| v \|_{L^p(B)} + \sqrt{\gamma} \| \nabla v \|_{L^p(B)} + \| D^2v \|_{L^p(B)} \leq N_1 \| \mathcal{L}v - \gamma v \|_{L^p(B)}, \]

By the embedding theorem \( \| v \|_{L^{\infty}(B)} \leq N_2 \| Dv \|_{L^p(B)} \) and
\[ \| g v \|_{L^p(B)} \leq N_2 \| g \|_{L^p(B)} \| Dv \|_{L^p(B)} \]

Therefore, for \( \gamma \) sufficiently large and every \( v \in W^{p,2}(B) \cap W^{p,1}_0(B) \) we have the estimate
\[ \| v \|_{W^{p,2}(B)} \leq N_3 \| \mathcal{L}v + (g - \gamma)v \|_{L^p(B)}. \]

The method of continuation with respect to a parameter gives a solution \( v \in W^{p,2}(B) \cap W^{p,1}_0(B) \) to the equation \( \mathcal{L}v + (g - \gamma)v = f \) for every \( f \in L^p(B) \), hence there is a solution \( u \) to the Dirichlet problem (3.2). We show that \( u \) is positive. For a function \( \eta \) let \( \eta^- \) and \( \eta^+ \) be the negative and positive parts of \( \eta \). Let \( w = -u \). If \( w \leq 0 \), then the proof is complete. If \( w \) is positive somewhere, then \( \sup_B w = \sup_B w^+ \). Since
\[ \mathcal{L}w - (g - \gamma)^-w = -(g - \gamma)^+w \geq -(g - \gamma)^+w^+, \]
by the maximum principle (see [32, Theorem 9.1]) we have
\[ \sup_B w^+ \leq N\|(g - \gamma)^+w^+\|_{L^1(B)} \leq N\|(g - \gamma)^+\|_{L^1(B)} \sup_B w^+. \]

Take \( \gamma \) so large that \( N\|(g - \gamma)^+\|_{L^1(B)} < 1 \). Then \( \sup_B w^+ \leq 0 \). Therefore, \( u \geq 0 \). The strict positivity follows from the Harnack inequality (see, e.g., [33]). Finally, we observe that by the Sobolev embedding theorem the function \( u \) has a continuously differentiable version, moreover, \( \partial_x u \) belongs to some Hölder space. We shall work with this version.

Taking for \( \varphi \) in the integral equality
\[ \int [\mathcal{L}\varphi + g\varphi] \sigma \ dy = 0 \]
the function \( u\psi \), where \( \psi \in C_0^\infty(B(y_0, R)) \), we arrive at the equality
\[ \int \left[ \mathcal{L}\psi + 2\left( \frac{1}{\mu} Q\nabla u, \nabla \psi \right) + \gamma \psi \right] u \sigma \ dy = 0. \]

Therefore, the function \( u \sigma \) is a solution to the equation \( \tilde{\mathcal{L}}^*(u \sigma) + \gamma(u \sigma) = 0 \) with the operator
By construction, the function \( u \) is Hölder continuous and strictly positive. Hence the function \( 1/u \) is also Hölder continuous. Thus, after these transformations we obtain the equation in which the coefficient \( c \) is constant and the coefficient \( b \) is continuous and even belongs to some Hölder class. Next, we apply the transformation from Remark 3.2 and arrive at the equation with zero coefficients \( b \) and \( c \). Moreover, if the original matrix \( A \) satisfies Dini's condition, then after these transformations it also satisfies this condition, in particular, Theorem 3.1 extends to equations with the zero order term \( c_0 \), provided that \( c \in L^p_{loc}(\Omega) \).

As already noted above, in [11, Lemma 4.2] the Harnack inequality was established under the assumption that \( b = 0 \) and \( A \) satisfies the Dini mean oscillation condition. Using Zvonkin’s transform and the methods of killing first and zero order terms explained above, we can generalize this assertion and obtain an analog of Theorem 3.1 for the matrix \( A \) that satisfies the Dini mean oscillation condition.

Following [10] and [11], we shall say that a measurable function \( f \) on an open set \( \Omega \subset \mathbb{R}^d \) satisfies the Dini mean oscillation condition if for some \( t_0 > 0 \)

\[
\int_0^{t_0} \frac{w(r)}{r} \, dr < \infty,
\]

where

\[
w(r) = \sup_{x \in \Omega} \frac{1}{|\Omega(x,r)|} \int_{\Omega(x,r)} |f(y) - f_\Omega(x,r)| \, dy, \quad \Omega(x,r) = \Omega \cap B(x,r)
\]

\[
f_\Omega(x,r) = \frac{1}{|\Omega(x,r)|} \int_{\Omega(x,r)} f(y) \, dy.
\]

Clearly, the classical Dini condition implies the Dini mean oscillation condition. However, the latter is weaker. Let us consider the following example from [10]:

\[
a^\gamma(0) = \delta^\gamma, \quad a^\gamma(x) = (1 + |\ln |x||^{-\gamma})\delta^\gamma, \quad 0 < \gamma < 1, 0 < |x| < \frac{1}{2}.
\]

It is clear that \( A = (a^\gamma) \) does not satisfy the classical Dini condition. By the Poincaré inequality \( w(r) \leq C|\ln r|^{-\gamma-1} \) with some constant \( C > 0 \), so \( A \) satisfies the Dini mean oscillation condition.

If there exists a number \( N \) such that \( |\Omega(x,r)| \geq Nr^d \) for all \( x \in \Omega \) and \( r \leq \operatorname{diam} \Omega \), then, according to [34, Lemma A1], any measurable function \( f \) satisfying the Dini mean oscillation condition, has a version uniformly continuous on \( \Omega \), moreover, its modulus of continuity is estimated by \( \int_0^{t_0} \frac{w(r)}{r} \, dr \). We shall work with this continuous version. In addition, according to [11, Lemma 2.1], the product \( fg \) satisfies the Dini mean oscillation condition if \( g \) satisfies the classical Dini condition and \( f \) satisfies the Dini mean oscillation condition. Note that there are discontinuous functions in VMO, while the continuity implies VMO locally. Thus, the Dini mean oscillation condition is stronger than VMO on balls.
Suppose that a function $f$ is defined on $\mathbb{R}^d$ and $|\Omega(x, r)| \geq Nr^d$ for all points $x \in \Omega$ and $r \leq \text{diam } \Omega$. Then for all $x \in \Omega$ and some constant $C(d, N) > 0$ the estimate

$$\frac{1}{|\Omega(x, r)|} \int_{\Omega(x, r)} |f(y) - f_\Omega(x, r)| \, dy \leq C(d, N) \frac{1}{|B(x, r)|} \int_{B(x, r)} |f(y) - f_B(x, r)| \, dy$$

holds, where

$$f_B(x, r) = \frac{1}{|B(x, r)|} \int_{B(x, r)} f(y) \, dy.$$

Therefore, in order to verify the Dini mean oscillation, condition it suffices to show that

$$\int_0^\infty \frac{\tilde{w}(r)}{r} \, dr < \infty, \quad \text{where} \quad \tilde{w}(r) = \sup_{x \in \Omega} \frac{1}{|B(x, r)|} \int_{B(x, r)} |f(y) - f_B(x, r)| \, dy.$$

We need the following observations.

**Lemma 3.4.**

i. Suppose that a function $f$ on $B(z, 4R) \subset \mathbb{R}^d$ satisfies the Dini mean oscillation condition, $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ is a diffeomorphism of class $C^1$, the functions $\partial_\alpha \Phi^j$ are Hölder continuous and $2^{-1}|x - y| \leq |\Phi(x) - \Phi(y)| \leq 2|x - y|$. Then the function $f \circ \Phi$ satisfies the Dini mean oscillation condition on $B(z', R)$, where $z' = \Phi^{-1}(z)$.

ii. Suppose that a function $f$ on $B(z, 4R) \subset \mathbb{R}^d$ satisfies the Dini mean oscillation condition. Then the function

$$F(x_1, \ldots, x_d, x_{d+1}) = f(x_1, x_2, \ldots, x_d)$$

satisfies the Dini mean oscillation condition on $B((z, z_{d+1}), R)$ for every $z_{d+1}$.

**Proof.** Let $0 < r < 2R$ and

$$w(r) = \sup_{x \in B(z, 2R)} \frac{1}{|B(x, r)|} \int_{B(x, r)} |f(y) - f_B(x, r)| \, dy.$$

Let us prove (i). By assumption the functions $\partial_\alpha \Phi^j$ are Hölder continuous of some order $\gamma \in (0, 1)$. It suffices to verify that the Dini mean oscillation condition is fulfilled for the function $g(x) = f(\Phi(x))|\det \Phi'(x)|$. Let $x \in B(z', 2R)$ and $0 < r < R$. We have

$$\frac{1}{|B(x, r)|} \int_{B(x, r)} |g(y) - g_B(x, r)| \, dv \leq \frac{1}{|B(x, r)|} \int_{B(\Phi(x), 2r)} |f(y) - \frac{g_B(x, r)}{|\det \Phi'(\Phi^{-1}(y))|}| \, dy.$$

Let us estimate the difference

$$\frac{g_B(x, r)}{|\det \Phi'(\Phi^{-1}(y))|} - f_B(\Phi(x), 2r).$$

If we replace $|\det \Phi'(\Phi^{-1}(y))|^{-1}$ by $|\det \Phi'(x)|^{-1}$, then the difference is estimated by $N_1 r^\gamma$ with some constant $N_1 > 0$. Observe that...
In addition, for some constant \( C > 0 \) we have the estimate

\[
1 - \frac{\Phi(B(x), r)}{|\det \Phi'(x)| |B(x, r)|} \leq C r^\gamma.
\]

Therefore, we arrive at the inequality

\[
\frac{1}{|B(x, r)|} \int_{B(x, r)} |g(v) - g_B(x, r)| \, dv \leq N_3(w(2r) + r^\gamma),
\]

which shows that \( g \) satisfies the Dini mean oscillation condition.

Let us prove assertion (ii). Let \( c_d \) denote the volume of the unit ball in \( \mathbb{R}^d \).

Let \( (x, x_{d+1}) \in B((z, z_{d+1}), R) \) and \( 0 < r < R \). We observe that by Fubini’s theorem

\[
F_B((x, x_{d+1}), r) = \frac{2}{c_{d+1}} \int_{B(0, 1)} f(x_1 + ry_1, \ldots, x_d + ry_d) \sqrt{1 - y_1^2 - \cdots - y_d^2} \, dy_1 \cdots dy_d
\]

\[
= \frac{2}{c_{d+1}} \int_{B(0, 1)} (f(x_1 + ry_1, \ldots, x_d + ry_d) - f_B(x, r)) \sqrt{1 - y_1^2 - \cdots - y_d^2} \, dy_1 \cdots dy_d + f_B(x, r),
\]

where the first term is estimated in absolute value by \( Nw(r) \) with some constant \( N > 0 \).

We have

\[
\frac{1}{|B((x, x_{d+1}), r)|} \int_{B((x, x_{d+1}), r)} |F(v) - F_B((x, x_{d+1}), r)| \, dv
\]

\[
\leq Nw(r) + \frac{1}{|B((x, x_{d+1}), r)|} \int_{B((x, x_{d+1}), r)} |F(v) - f_B(x, r)| \, dv.
\]

Applying Fubini’s theorem to the second term, we estimate it by

\[
\frac{2r}{|B((x, x_{d+1}), r)|} \int_{B(x, r)} |f(y) - f_B(x, r)| \, dy.
\]

Thus, for some constant \( C(d, N) > 0 \) we obtain
\[
\frac{1}{|B((x, x_{d+1}), r)|} \int_{B((x, x_{d+1}), r)} |F(y) - F_B((x, x_{d+1}), r)| \, dy \leq C(d, N)w(r),
\]

which implies our claim.

Our next result generalizes Example 2.5 and Theorem 3.1 to the case where the diffusion matrix satisfies the Dini mean oscillation condition.

**Theorem 3.5.** Suppose that condition \( \mathbf{H}_a \) is fulfilled, on every ball the matrix \( A \) satisfies the Dini mean oscillation condition with some function \( \omega \), and \( b^i, c \in L^{d+}_{{\text{loc}}} (\Omega) \). Suppose also that \( 2L_1 \) is a solution to the equation

\[
\frac{\partial}{\partial x_j} (a^{ij} \frac{\partial r}{\partial x_i}) - \frac{\partial}{\partial x_i} (b^i r) + cr = 0.
\]

Then the function \( r \) has a continuous version. Moreover, if \( r \geq 0 \), then the continuous version of \( r \) satisfies the Harnack inequality, i.e., for every ball \( B(x_0, R/2) \subset B(x_0, 4R) \subset \Omega \) there exists a number \( C \) such that

\[
\sup_{x \in B(x_0, R/2)} r(x) \leq C \inf_{x \in B(x_0, R/2)} r(x),
\]

where \( C \) depends on \( R, \omega, d, \nu, \eta \), and \( \|c\|_{L^p(B(x_0, 4R))} \) and \( \|b\|_{L^p(B(x_0, 4R))} \), and does not depend on \( r \). The modulus of continuity of \( r \) on \( B(x_0, R/2) \) depends on the same objects.

**Proof.** Justification is similar to the reasoning in Example 2.5 and Theorem 3.1, but in place of results from [8] and [9] we apply [10, Theorem 1.10] and [11, Lemma 4.2], because Zvonkin's transform combined with Remark 3.3 and the transformation from Remark 3.2 enable us to reduce the proof to the case \( b = 0 \) and \( c = 0 \), moreover, the elements of the new matrix \( A \) satisfy the Dini mean oscillation condition by Lemma 3.4.

Let us carefully describe the reduction procedure of our general equation to an equation with vanishing coefficients \( b \) and \( c \). Using Zvonkin's transform \( \Phi \) with \( \Psi = \Phi^{-1} \), we obtain the equation \( \tilde{\mathcal{L}}^* \sigma + g \sigma = 0 \), where

\[
\sigma(y) = |\det \Psi'(y)| \rho(\Psi(y))|
\]

\[
\mathcal{L}f(y) = q^{km}(y)\frac{\partial}{\partial x_k} f(y) + h^k(y) \frac{\partial}{\partial x_k} f(y),
\]

and the coefficients have the form

\[
q^{km}(y) = a^{ij}(\Psi(y))\frac{\partial}{\partial x_i} \Phi^k(\Psi(y))\frac{\partial}{\partial x_j} \Phi^m(\Psi(y)), \quad h^k(y) = \lambda u^k(\Psi(y)), \quad g(y) = c(\Psi(y)).
\]

Applying the arguments from Remark 3.3 we construct a positive continuous differentiable function \( \eta \) with Hölder continuous derivatives such that

\[
\tilde{\mathcal{L}} \eta + (g - \gamma) \eta = 0, \quad \gamma > 0.
\]

The function \( \theta = \sigma \eta \) satisfies the equation

\[
\tilde{\mathcal{L}}^* \theta + \gamma \theta = 0,
\]

where

\[
\tilde{\mathcal{L}}f = \mathcal{L}f + 2\left\langle \frac{1}{\eta} \nabla \eta, \nabla f \right\rangle.
\]

Observe that by Lemma 3.4 the matrix \( Q \) satisfies the Dini mean oscillation condition.
with a modulus $\tilde{\omega}$ expressed through $\omega$. The drift coefficient
\[
\lambda u(\Psi(y)) + \frac{1}{\eta}Q\nabla\eta
\]
also satisfies the Dini mean oscillation condition, since $u^k(\Psi(y))$, $\eta(y)$, $\partial_y\eta(y)$ are H"older continuous and $\eta$ is strictly positive. The H"older norm of $u^k(\Psi(y))$ is estimated by a constant depending on $\nu$, $\omega$, and $\|b\|_{L^p}$. The H"older norms of $\eta$ and $\partial_y\eta$ are estimated by a constant depending on $\nu$, $\tilde{\omega}$, $\lambda\sup_y|u^k(\Psi(y))|$, and $\|g\|_{L^p}$. Note also that $\lambda$ and $\gamma$ depend on $\nu$, $\omega$, $\|b\|_{L^p}$, and $\|g\|_{L^p}$. Applying the transformations from Remark 3.2 we obtain an equation with zero drift and potential terms.

In the theory of divergence form and non-divergence form elliptic equations the Harnack inequality is used for estimating the modulus of continuity of the solution and showing that it is H"older continuous. A simple one-dimensional example $(A_Q)^{n} = 0$ demonstrates that a solution $g(x) = (c_1x + c_2)/A(x)$ need not be more regular than the coefficient $A$. For example, if $A$ is not H"older continuous, but satisfies Dini’s condition, then $g(x)$ is the same. However, also for double divergence form equations the Harnack inequality can be used for estimating the modulus of continuity, see [9, Remark 3.9].

We now discuss the integrability of solutions without the assumption about Dini’s condition. Let us recall that, according to [9, Theorem 2.1], if $a^{ij}$ and $b^i$ satisfy conditions $H_a$ and $H_b$ and $a^{ij} \in VMO$, then every solution $g \in L^1_{\text{loc}}(\Omega)$ is locally integrable to every power $p \geq 1$. If the functions $a^{ij}$ satisfy the Dini mean oscillation condition, then the solution is locally bounded and even continuous. It is of interest to study integrability when $A$ is slightly better than $VMO$, but does not satisfy Dini’s condition.

Suppose that the coefficients $a^{ij}$ and $b^i$ satisfy conditions $H_a$ and $H_b$ and that
\[
\|A(x) - A(y)\| \leq \omega(|x - y|),
\]
where $\omega$ is an increasing continuous function on $[0, +\infty)$ and $\omega(0) = 0$. We also assume that for some $C_\omega > 0$ and all $t \geq 0$ the inequality
\[
\omega(t) \geq C_\omega t^{1-d/p}
\]
holds. Set
\[
\Lambda(t) = -\int_{1/t}^{1} \ln(\omega^{-1}(s)) \, ds,
\]
where $\omega^{-1}$ is the inverse function. Integrating by parts, it is readily verified that
\[
\Lambda\left(\frac{1}{t}\right) = \int_{\omega^{-1}(t)}^{\omega^{-1}(1)} \frac{\omega(s)}{s} \, ds - \ln\omega^{-1}(1) + t\ln\omega^{-1}(t).
\]
We observe that if Dini’s condition is fulfilled, then the function $\Lambda(t)$ is bounded from above (and bounded on $[1, +\infty)$).

**Theorem 3.6.** Let $g \in L^1_{\text{loc}}(\Omega)$ satisfy Equation (1.1). Then, for every closed ball $B \subset \Omega$ and every $\delta > 0$, there exists a constant $c = c(B, \delta) > 0$ such that for all $q \geq 1$ one has
\[ \ln \|q\|_{L^q(B)} \leq c + c\Lambda(cq) + \delta \ln q. \]

The proof is given below.

We first consider the case of a bounded drift coefficient \( b \) and by Zvonkin’s transform extend the result to the case of integrable \( b \).

The next auxiliary assertion is standard, but in order to control the dependence of constants on parameters we include a justification.

**Lemma 3.7.** Let \( \alpha = (\alpha^i) \) be a constant symmetric positive definite matrix such that
\[ \gamma \cdot 1 \leq \alpha \leq \frac{1}{\gamma} \cdot 1. \]

Let \( f \in C\infty_{0}(B(0,1)) \) and
\[ 1 < r < d/(d-1), \quad 1 < t < d/(d-1). \]

Then there exists a function \( u \in C\infty(B(0,1)) \) such that \( \text{tr}(\alpha D^2 u) = f \) and
\[
\|u\|_{L^r(B(0,1))} + \|\nabla u\|_{L^t(B(0,1))} \leq C(r,d,\gamma)\|f\|_{L^t(B(0,1))},
\]
\[
\|D^2 u\|_{L^r(B(0,1))} \leq C(d,t,\gamma)\|f\|_{L^t(B(0,1))},
\]

where the constants \( C(d,t,\gamma) \) and \( C(r,d,\gamma) \) do not depend on \( t \).

**Proof.** Changing coordinates, we reduce the problem to the case of the unit matrix \( \alpha \).

Let \( u \) be the solution to the Dirichlet problem \( \Delta u = f \) on \( B(0,2), \ u = 0 \) on \( \partial B(0,2) \). We estimate \( \nabla u \).

Let \( g^i \in C\infty_{0}(B(0,2)) \) and let \( v \) be the solution to the Dirichlet problem \( \Delta v = \text{div} \ g \) on \( B(0,2), \ v = 0 \) on \( \partial B(0,2) \). By [3, Corollary 1.7.6] and the embedding theorem one has
\[ \sup_{B(0,2)} |v(x)| \leq C_1(r,d)\|g\|_{L^{r'}(B(0,2))}. \]

Then
\[ \int \langle \nabla u, g \rangle \ dx = - \int v\Delta u \ dx \leq C_1(r,d)\|g\|_{L^{r'}(B(0,2))}\|f\|_{L^t(B(0,2))}. \]

Since \( g \) was arbitrary, we obtain the desired estimate of the norm \( \|\nabla u\|_{L^t(B(0,2))} \).

Similarly the norm \( \|u\|_{L^r(B(0,2))} \) is estimated.

Now let \( \zeta \in C\infty_{0}(B(0,2)) \) and \( \zeta = 1 \) on \( B(0,1) \). Applying [32, Theorem 9.8 and Theorem 9.9] to \( \zeta u \), we obtain the inequality
\[ \|D^2(\zeta u)\|_{L^r(B(0,2))} \leq C_1(d,t)\|\Delta(\zeta u)\|_{L^t(B(0,2))}, \]

where the right-hand side is estimated by
\[ C_2(d,t)\|f\|_{L^t(B(0,2))} + \|\nabla u\|_{L^t(B(0,2))} + \|u\|_{L^t(B(0,2))}. \]

It remains to apply the estimates for \( u \) and \( \nabla u \) obtained above.

\[ \square \]

**Lemma 3.8.** Let \( q \) be a solution to Equation (1.1) with a bounded drift coefficient. Then, for every ball \( B \) and every \( \delta > 0 \), there exists \( c = c(B,\delta) > 0 \) such that for all \( q \geq 1 \) we
have
\[
\ln \|q\|_{L^q(B)} \leq c + c\Lambda(q) + \delta \ln q.
\]

**Proof.** It suffices to obtain our estimate for large \( q \). Without loss of generality we can assume that \( B = B(0, 1) \) and \( B(0, 4) \subset \Omega \). Let \( x_0 \in B(0, 1) \) and \( 0 < \lambda < 1 \) be fixed. The function \( \sigma(y) = q(x_0 + \lambda y) \) on the ball \( B(0, 2) \) satisfies the equation
\[
\partial_y(y) (q(y)\sigma) - \partial_y(h(y)\sigma) = 0,
\]
where \( q(y) = a(y) = \lambda b(y) \) and \( h(y) = \lambda b(y) \). Let
\[
Q_0 = (q_0^i, q_0^j, x_0^i, x_0^j), \quad \zeta \in C^\infty_0(B(0, 1)), \quad 0 \leq \zeta \leq 1, \quad \zeta(y) = 1 \text{ if } y \in B(0, 1/2).
\]
By Lemma 3.7, for every function \( f \in C^\infty_0(B(0, 1)) \), there exists a smooth solution \( u \) to the equation \( \text{tr}(Q_0 D^2 u) = f \) satisfying estimates (3.3) and (3.4). Then
\[
\int f \zeta \sigma \ dy = \int \text{tr}((Q_0 - Q) D^2 u) \zeta \sigma \ dy - \int [\text{tr}(Q D^2 \zeta) + 2 \langle Q \nabla u, \nabla \zeta \rangle + \langle h, \nabla(\zeta u) \rangle] \sigma \ dy.
\]
Let \( 1 < t < \tau < d/(d - 1) \) and \( 1 < r < d/(d - 1) \). Observe that
\[
\int \text{tr}((Q_0 - Q) D^2 u) \zeta \sigma \ dy \leq \omega(\lambda) \| D^2 u \|_{L^r(B(0, 1))} \| \zeta \|_{L^r(B(0, 1))}.
\]
By Lemma 3.7 we have \( \| D^2 u \|_{L^r(B(0, 1))} \leq C_1 t |f|_{L^r(B(0, 1))} \), where \( C_1 \) does not depend on \( t \). Therefore,
\[
\int \text{tr}((Q_0 - Q) D^2 u) \zeta \sigma \ dy \leq C_1 t \omega(\lambda) |f|_{L^r(B(0, 1))} \| \zeta \|_{L^r(B(0, 1))}.
\]
The expression
\[
- \int [\text{tr}(Q D^2 \zeta) + 2 \langle Q \nabla u, \nabla \zeta \rangle + \langle h, \nabla(\zeta u) \rangle] \sigma \ dy \tag{3.5}
\]
is estimated from above by
\[
C(\zeta) \left[ \sup_y \| Q(y) \| + \sup_y |h(y)| \right] \int_{B(0, 1)} (|u(y)| + |\nabla u(y)|) \| \sigma(y) \| \ dy.
\]
By Hölder’s inequality we obtain
\[
\int_{B(0, 1)} (|u(y)| + |\nabla u(y)|) \| \sigma(y) \| \ dy \leq \left( \| u \|_{L^t(B(0, 1))} + \| \nabla u \|_{L^t(B(0, 1))} \right) \| \sigma \|_{L^r(B(0, 1))}.
\]
Applying again Lemma 3.7, we can estimate (3.5) by
\[
C_2 |f|_{L^t(B(0, 1))} \| \sigma \|_{L^r(B(0, 1))},
\]
where \( C_2 \) does not depend on \( t \). Thus, we arrive at the estimate
\[
\int f \zeta \sigma \ dy \leq C_1 t \omega(\lambda) |f|_{L^r} \| \zeta \|_{L^r} + C_2 |f|_{L^t} \| \sigma \|_{L^r(B(0, 1))}.
\]
Set now \( t' = q \). Let \( C_1 q \omega(\lambda) = 1/2 \), i.e., \( \lambda = \omega^{-1} \left( \frac{1}{2C_1 q} \right) \), where the number \( q \) is so large
that $\lambda < 1$. Recall that $\zeta = 1$ on $B(0, 1/2)$. Therefore, we have
\[ \|\sigma\|_{L^s(B(0, 1/2))} \leq 2C_2\|\sigma\|_{L^r(B(0, 1))}. \]
Let $r' < s < q$. Applying Hölder’s inequality, we arrive at the estimate
\[ \|\sigma\|_{L^s(B(0, 1/2))} \leq C_3\|\sigma\|_{L^r(B(0, 1))}, \]
where $C_3$ does not depend on $q$, $\lambda$ and $s$. Returning to the original coordinates, we obtain
\[ \|q\|_{L^s(B(x_0, \lambda/2))} \leq \lambda^{d/q-d/s}C_3\|q\|_{L^r(B(x_0, \lambda))}. \]
Let $1 < R < 2$. Covering the ball $B(0, R)$ by finitely many balls $B(x_i, \lambda/2)$, we arrive at the estimate
\[ \|q\|_{L^s(B(0, R))} \leq C_4\lambda^{-d/s}\|q\|_{L^r(B(0, R+\lambda))}, \]
where $C_4$ does not depend on $s$, $R$ and $\lambda$. Let
\[ \beta > 1, \quad \frac{\ln C_4}{\ln \beta} < \delta. \]
Set $q_m = 2C_1\beta^m$, $s = 2C_1\beta^{m-1}$ and
\[ \lambda_m = \omega^{-1}\left(\frac{1}{2C_1\beta^m}\right), \quad R = r_m = 2 - \sum_{k=0}^{m} \lambda_k. \]
Since $\omega^{-1}(s) \leq (s/C)^{p/(p-d)}$, the series $\sum_k \lambda_k$ converges. We can assume that $k_0$ is so large that $r_m > 1$ for all $m > k_0$ and $\beta^{-k_0+1} < 1$. Thus, for all $m > k_0$ the inequality
\[ \|q\|_{L^m(B(0, r_m))} \leq e^{\nu_m}\|q\|_{L^{m-1}(B(0, r_{m-1}))}, \quad \nu_m = \delta \ln \beta - d(2C_1)^{-1}\beta^{-m+1}\ln \lambda_m, \]
holds. By iterations from $k$ to $k_0$ we obtain the inequality
\[ \ln \|q\|_{L^m(B(0, r_k))} \leq C_5 + \sum_{m=k_0}^{k} \nu_m, \]
where $C_5$ does not depend on $k$. Observe that
\[ \sum_{m=k_0}^{k} \nu_m \leq \delta \ln q_k - \frac{d}{2C_1} \sum_{m=k_0}^{k} \beta^{-m+1}\ln \lambda_m. \]
We have the estimate
\[ -\sum_{m=k_0}^{k} \beta^{-m+1}\ln \omega^{-1}\left(\frac{1}{2C_1\beta^m}\right) \leq -\frac{\beta}{\beta - 1}\int_{\beta^{-k_0+1}}^{\beta^{-k}} \ln \omega^{-1}\left(\frac{t}{2C_1}\right) \, dt. \]
Thus,
\[ \ln \|q\|_{L^m(B(0, 1))} \leq C_5 + \delta \ln q_k + C_6\varLambda(q_k). \]
If now $\beta^{-1}q_k \leq q \leq q_k$ and $c$ is greater than $C_5$, $C_6$ and $\beta$, we obtain the resulting estimate $\ln \|q\|_{L^m(B(0, 1))} \leq c + c\varLambda(eq) + \delta \ln q$. \qed
We now prove Theorem 3.6.

**Proof of Theorem 3.6.** It suffices to show that given
\[ B(x_0, R/2) \subset B(x_0, 4R) \subset \Omega \]
and \( \delta > 0 \), there exists a constant \( c > 0 \) for which
\[ \ln \|q\|_{L^\infty(B(x_0,R/2))} \leq c + c\Lambda(cq) + \delta \ln q. \]
Let \( \Phi \) be the diffeomorphism constructed before Proposition 2.3. According to Proposition 2.4, the function
\[ \sigma(y) = q(\Psi(y))|\det\Psi'(y)| \]
on the ball \( B(y_0, 2R) \) centered at \( y_0 = \Phi(x_0) \) is a solution to an equation whose coefficients satisfy the hypotheses of Lemma 3.8, that is, the matrix \( (q^{mk}) \) satisfies condition \( H_a \) and the drift coefficient \( h \) is bounded. In addition, \( \Psi \) is Lipschitz, the functions \( \partial_s \Phi \) are Hölder continuous of order \( 1 - d/p \) and \( \omega(t) \geq Ct^{1-d/p} \). Hence there exists a number \( N > 0 \) such that
\[ \|Q(y) - Q(z)\| \leq N\omega(N|y - z|). \]
Observe that
\[ \Lambda_N \left( \frac{1}{t} \right) = -\int_t^1 \ln \left( \frac{1}{N} \omega^{-1} \left( \frac{s}{N} \right) \right) \, ds \leq N\Lambda \left( \frac{N}{t} \right) + \ln N. \]
By Lemma 3.8, there exists a number \( c = c(y_0, R, \delta) > 0 \) such that
\[ \ln \|\sigma\|_{L^\infty(B(y_0,R))} \leq c + c\ln N + cN\Lambda(cNq) + \delta \ln q. \]
Since \( \Phi \) is a \( C^1 \)-diffeomorphism, an analogous estimate holds for the original function \( q \). \( \square \)

**Corollary 3.9.** Under the hypotheses of Theorem 3.6, for every closed ball \( B \subset \Omega \) the following assertions are true.

i. If \( \lim_{t \to 0_+} \omega(t)|\ln t| = 0 \), then
\[ \exp (\gamma_1 |q|^{1/2}) \in L^1(B) \quad \text{for all} \quad \gamma_1, \gamma_2 > 0. \]

ii. If the function \( \omega(t)|\ln t| \) is bounded on \((0,1]\), then exist numbers \( \gamma_1, \gamma_2 > 0 \) such that
\[ \exp (\gamma_1 |q|^{1/2}) \in L^1(B). \]

iii. If for some \( \beta \in (0,1) \) the function \( \omega(t)|\ln t|^\beta \) is bounded on \((0,1]\), then for some \( \gamma > 0 \) we have
\[ \exp (\gamma |\ln (|q| + 1)|^{1/(1 - \beta)}) \in L^1(B). \]

**Proof.** Let us consider case (i), when \( \omega(t) = o(|\ln t|^{-1}) \). Let \( \varepsilon > 0 \). There is \( s_0 \in (0,1) \) such that for all \( s \in (0,s_0) \) one has \( |\ln \omega^{-1}(s)| \leq \varepsilon s^{-1} \), which follows from the estimate \( \omega(t) \leq \varepsilon |\ln t|^{-1} \). Therefore,
\[ \Lambda(q) \leq C_1(\varepsilon) + \varepsilon \ln q, \]

so that

\[ \int_B |q|^q \, dx \leq C_2 q^\varepsilon. \]

Let \( q = \gamma_2 k \) and \( \varepsilon \gamma_2 = 1/2 \). We obtain

\[ \int_B \left( \frac{|q|^{\gamma_2}}{k!} \right)^k \, dx \leq \frac{C_2^k k^{k/2}}{k!}. \]

By Stirling’s formula this yields convergence of the series

\[ \int_B \sum_k \left( \frac{\gamma_1 |q|^{\gamma_2}}{k!} \right)^k \, dx, \]

which implies the integrability of \( \exp(\gamma_1 |q|^{\gamma_2}) \). Assertion (ii) is proved similarly.

Let us prove assertion (iii). The boundedness of the function \( \omega(t) \ln t^{-\beta} \) implies the estimate \( |\ln \omega^{-1}(s)| \leq M_1 s^{1-\beta} \) with some constant \( M_1 > 0 \). By Theorem 3.6, for all \( q \geq 1 \) and some \( M_2 > 0 \) we obtain the estimate

\[ \ln \|q\|_{L^\beta(B)} \leq M_2 q^{(1-\beta)/\beta}. \]

Chebyshev’s inequality gives

\[ \ln |\{x \in B : |q(x)| > t\}| \leq -q \ln t + M_2 q^{1/\beta}. \]

Let \( q = C(\beta, M_2) \ln t^{\beta/(1-\beta)} \), where \( C(\beta, M_2) = (\beta/M_2)^{\beta/(1-\beta)} \). There is a number \( t_0 \) such that for all \( t \geq t_0 \) one has \( q \geq 1 \) and for some constant \( M_3 > 0 \) the inequality

\[ \ln |\{x \in B : |q(x)| > t\}| \leq -M_3 \ln t^{1/(1-\beta)} \]

holds. For any continuously differentiable function \( f \) with \( f' > 0 \) we have

\[ \int_B f(|q(x)|) \, dx \leq \int_0^{t_0} |\{x \in B : f(|q(x)|) > t\}| \, dt + \int_{f(t_0)}^{+\infty} f'(s) \exp \left( -M_3 |\ln s|^{1/(1-\beta)} \right) \, ds. \]

It is readily verified that for \( f(t) = \exp(\gamma |\ln (t + 1)|^{1/(1-\beta)}) \) with \( \gamma < M_3 \) the corresponding integral in the right-hand side is finite, which gives the desired integrability.

## 4. Existence and uniqueness of probability solutions

In this section we apply the results obtained above for constructing positive and probability solutions to the Kolmogorov equation.

The next result is a modification of [3, Theorem 1.5.6] for unbounded coefficients.

**Lemma 4.1.** Let \( \Omega \subset \mathbb{R}^d \) and \( B(x_0, 4R) \subset \Omega \). Suppose that the coefficients \( a^{ij} \) and \( b^i \) satisfy conditions \( H_a \) and \( H_b \) and the matrix \( A \) satisfies condition \( VMO \) on \( \Omega \). Then, for any numbers \( 0 < r_1 < r_2 < R \), there exists a constant \( C > 0 \) such that
\[ \int_{B(x_0, r_2)} q(x) \, dx \leq C \int_{B(x_0, r_1)} q(x) \, dx \]

for every nonnegative solution \( q \) to Equation (1.1) on \( \Omega \).

**Proof.** Let \( \Phi \) be the diffeomorphism constructed above and \( \Psi = \Phi^{-1} \). Set \( y_0 = \Phi(x_0) \).

Note that

\[ \Psi(B(y_0, r_1/2)) \subset B(x_0, r_1), \quad B(x_0, r_2) \subset \Psi(B(y_0, 2r_2)). \]

Let \( q \) be a nonnegative solution to Equation (1.1). By Proposition 2.4 the function

\[ \sigma(y) = q(\Psi(y)) |\det \Psi'(y)| \]

on \( B(y_0, 2R) \) satisfies the equation with bounded coefficients. According to [3, Theorem 1.5.6] there exists a constant \( C > 0 \) such that

\[ \int_{B(y_0, 2r_2)} \sigma(y) \, dy \leq C \int_{B(y_0, r_1/2)} \sigma(y) \, dy \]

and \( C \) depends only on the coefficients and the numbers \( r_1, r_2 \). Then

\[ \int_{B(x_0, r_2)} q(x) \, dx \leq \int_{\Psi(B(y_0, 2r_2))} q(x) \, dx \leq C \int_{\Psi(B(y_0, r_1/2))} q(x) \, dy \leq C \int_{B(x_0, r_1)} q(x) \, dx, \]

which completes the proof. \( \square \)

Our next theorem generalizes assertion (i) in [3, Theorem 2.4.1] to the case where the coefficients \( a_{ij} \) satisfy condition VMO. It was assumed in the cited theorem that the functions \( a_{ij} \) belong locally to the Sobolev class \( W^{p, 1} \) with some \( p > d \).

**Theorem 4.2.** Suppose that the coefficients \( a_{ij} \) and \( b^i \) are defined on all of \( \mathbb{R}^d \), \( b^i \in L^d_{\text{loc}}, \) and for every ball \( B \) we can find a number \( \nu_B > 0 \) and a continuous nonnegative increasing function \( w_B \) on \( [0, +\infty) \) such that \( w_B(0) = 0 \) and

\[ \nu_B \cdot I \leq A(x) \leq \frac{1}{\nu_B} \cdot I, \quad \sup_{x \in B} r^{-2d} \int_{B(x,r)} |a_{ij}(y) - a_{ij}(z)| \, dy \, dz \leq w_B(r) \quad \forall \ r > 0. \]

Then there exists a nonnegative solution \( q \) to Equation (1.1) on \( \mathbb{R}^d \) such that

\[ \int_{B(0,1)} q(x) \, dx = 1. \]

**Proof.** Let \( \zeta \in C_0^\infty(\mathbb{R}^d) \) have support in \( B(0, 1) \), \( \zeta \geq 0 \) and \( \|\zeta\|_{L^1(\mathbb{R}^d)} = 1 \). For every \( k \in \mathbb{N} \), let \( \zeta_k \) denote the function \( k^d \zeta(kx) \) and let

\[ a_{ij}^k = a_{ij}^* \zeta_k, \quad b^i_k = b^i \ast \zeta_k. \]

The functions \( a_{ij}^k \) and \( b^i_k \) are infinitely differentiable, and on every ball \( B_R = B(0, R) \) we have

\[ \nu_{B_{R+1}} \cdot I \leq A_k \leq \frac{1}{\nu_{B_{R+1}}} \cdot I, \]

\[ \|b_k\|_{L^p(B_R)} \leq \|b\|_{L^p(B_{R+1})}, \quad p_R = p(B_{R+1}). \]
Moreover,
\[
\sup_{x \in B_r} r^{-2d} \int_{B(x,r)} \int_{B(x,r)} |a_{ij}^k(y) - a_{ij}^k(z)| \, dy \, dz \leq w_{B_{r+1}}(r).
\]
According to [3, Theorem 2.4.1], there exists a positive smooth solution \( \varrho_k \) to the equation
\[
\partial_{x_i} \partial_{x_j} (a_{ij}^k \varrho_k) - \partial_{x_i} (b_i^j \varrho_k) = 0.
\]
Multiplying by a constant, we can assume that
\[
\int_{B_1} \varrho_k(x) \, dx = 1.
\]
Applying Lemma 4.1, for every ball we obtain the inequality
\[
\int_{B_R} \varrho_k(x) \, dx \leq C_R \int_{B_1} \varrho_k(x) \, dx = C_R,
\]
where \( C_R \) does not depend on \( k \). Let \( b \in L^{p_R}(B_R) \) and \( p_R > d \). Note that
\[
q_R = \frac{p_R}{p_R - 1} < \frac{d}{d - 1}.
\]
According to [9, Theorem 2.1], the norms \( \| \varrho_k \|_{L^{p_R}(B_R)} \) are uniformly bounded in \( k \) for each \( R > 0 \). Thus, the sequence \( \{ \varrho_k \} \) has a subsequence such that for every natural number \( N \) it converges weakly in \( L^{p_N}(B_N) \) to some nonnegative function \( \varrho \). In addition, on every ball \( B_N \) the sequence \( \{a_{ij}^k\} \) converges uniformly to \( a^{ij} \) and the sequence \( \{b_i^j\} \) converges to \( b^i \) in \( L^{p_N}(B_N) \). Passing to the limit in the integral identities determining the solutions \( \varrho_k \), we obtain an analogous integral identity for \( \varrho \). Therefore, the function \( \varrho \) is a solution to Equation (1.1). It is clear that \( \varrho \geq 0 \) and \( \int_{B_1} \varrho(x) \, dx = 1 \). \( \square \)

**Remark 4.3.** If in addition to the hypotheses of Theorem 4.2 the matrix \( A \) satisfies the Dini mean oscillation condition on every ball, then there exists a continuous positive solution \( \varrho \) to Equation (1.1).

As a corollary we obtain a generalization of the Hasminskii theorem, which follows from Theorem 4.2 and [3, Corollary 2.3.3]. Let \( W^{d,2}_{loc}(\mathbb{R}^d) \) be the local Sobolev class of all functions belonging to \( L^d_{loc}(\mathbb{R}^d) \) along with their first and second Sobolev derivatives.

**Corollary 4.4.** If in addition to the hypotheses of Theorem 4.2 there is a function \( V \) of class \( W^{d,2}_{loc}(\mathbb{R}^d) \) along with numbers \( C > 0 \) and \( R > 0 \) such that
\[
\lim_{|x| \to +\infty} V(x) = +\infty, \quad LV(x) \leq -C \quad \text{if} \quad |x| > R,
\]
then there exists a probability solution \( \varrho \) to Equation (1.1) on \( \mathbb{R}^d \). Moreover, if \( A \) satisfies the Dini mean oscillation condition on every ball, then \( \varrho \) has a continuous positive version.

Note that assertion (ii) in [3, Theorem 2.4.1] gives a nonzero nonnegative solution to Equation (1.1) under the assumption that the matrix \( A \) is locally positive definite and
bounded and the drift coefficient $b$ is locally bounded. If, in addition, there is a Lyapunov function, then there exists a probability solution. In the results obtained above the condition on $A$ is stronger and the condition on $b$ is weaker.

Zvonkin’s transform enables us to obtain the following modification of the Hasminskii theorem generalizing [22, Theorem 4.10(iii)] to the matrix $A$ of class VMO.

**Corollary 4.5.** Suppose that the coefficients $a^i$ and $b^i$ are defined on all of $\mathbb{R}^d$ and we can find a number $\nu > 0$ and an increasing continuous function $\omega$ on $[0, +\infty)$ such that $\omega(0) = 0$ and

$$\nu \cdot I \leq A(x) \leq \frac{1}{\nu} \cdot I,$$

$$\sup_{z \in \mathbb{R}^d} \int_{B(z, r)} r^{-2d} \int_{B(z, r)} |a^{ij}(x) - a^{ij}(y)| \, dx \, dy \leq \omega(r).$$

Let also $b^i = b_1^i + b_2^i$, where $b_2 \in L^p(\mathbb{R}^d)$ with $p > d$ and

$$\lim_{|x| \to \infty} \langle b_1(x), x \rangle \leq C_1 - C_2 |x|^{\kappa + 1}, \quad |b_1(x)| \leq C_3 (1 + |x|^\kappa),$$

for some numbers $C_1, C_2, C_3 > 0$ and $\kappa > 0$. Then there exists a probability solution $\phi$ to Equation (1.1) on $\mathbb{R}^d$.

**Proof.** Applying Proposition 2.1, for every $\delta > 0$ we construct a mapping $u = (u^1, \ldots, u^d)$ such that $u^k$ is a solution to the equation

$$\text{tr}(A \nabla u^k) + \langle b_2, \nabla u^k \rangle - \lambda u^k = -b_2^k,$$

$u^k \in L^\infty(\mathbb{R}^d)$ and $\sup_x |\nabla u^k(x)| < \delta$. According to Proposition 2.3, for $\delta > 0$ sufficiently small the mapping $\Phi(x) = x + u(x)$ is a diffeomorphism of $\mathbb{R}^d$. Let $\Psi = \Phi^{-1}$. It is readily seen that $|\Psi(y)| \to \infty$ as $|y| \to \infty$. Let us consider the equation

$$\partial_y \partial_{y^m} (q^{km} \sigma) - \partial_y (h^k \sigma) = 0,$$  \hspace{1cm} (4.1)

where

$$q^{km}(y) = a^{ij}(\Psi(y)) \partial_{y^j} \Phi^k(\Psi(y)) \partial_{y^m} \Phi^m(\Psi(y)),$$

$$h(y) = \lambda u(\Psi(y)) + b_1^i(\Psi(y)) \partial_{y^i} \Phi(\Psi(y)).$$

Since $\Phi(x) = x + u(x)$, we have

$$\langle h(y), y \rangle(\Phi(x)) = \lambda \langle u(x), x + u(x) \rangle + \langle b_1(x), x + u(x) \rangle + \langle b_1^i(x) \partial_{y^i} u(x), x + u(x) \rangle.$$ 

Let $|u(x)| \leq M$ and $|\nabla u(x)| \leq \delta$. Then

$$\langle h(y), y \rangle(\Phi(x)) \leq \lambda M^2 + \lambda M |x| + M |b_1(x)| + \delta |b_1(x)| |x| + \delta M |b_1(x)| + \langle b_1(x), x \rangle.$$ 

Using our condition on $b_1$, we can estimate the right side by

$$C_1 + \lambda M^2 + \lambda M |x| + (1 + \delta)MC_3(1 + |x|^\kappa) + \delta C_3(1 + |x|^{\kappa + 1}) - C_2 |x|^{\kappa + 1}.$$ 

Let $\delta C_3 < C_2$. Then $\langle h(y), y \rangle \to -\infty$ as $|y| \to \infty$. Therefore, the hypotheses of Corollary 4.4 are fulfilled and there exists a probability solution $\sigma$ of Equation (4.1). Finally, it is clear that $\phi(x) = \sigma(\Phi(x))|\det \Phi(x)|$ is a probability solution to Equation (1.1). 

In connection with the existence theorems proved above we discuss a probabilistic interpretation of probability solutions of the Kolmogorov equation. Let $a^{ij}$ and $b^i$ be
Borel functions on $\mathbb{R}^d$ and let $\mu$ be a Borel probability measure satisfying the stationary Kolmogorov equation

$$\partial_x \partial_y (a^{ij} \mu) - \partial_y (b^j \mu) = 0,$$

understood in the sense of the integral equality

$$\int_{\mathbb{R}^d} \left[ a^{ij} \partial_x \partial_y \phi + b^j \partial_x \phi \right] d\mu = 0 \quad \forall \phi \in C_0^\infty(\mathbb{R}^d).$$

**Theorem 4.6.** Suppose that $a^{ij}, b^j \in L^1_{loc}(\mu)$ and

$$\int_{\mathbb{R}^d} \frac{\|A(x)\| + |(b(x), x)|}{1 + |x|^2} \mu(dx) < \infty.$$

Then, for every $T > 0$, one can find a probability space $(Q, \mathcal{F}, P)$, a filtration $\mathcal{F}_t$, and a continuous random process $\xi_t$ and a Wiener process $w_t$ adapted to the filtration $\mathcal{F}_t$ such that

$$d\xi_t = b(\xi_t) \ dt + \sqrt{2A(\xi_t)}dw_t$$
on $[0, T]$ and the distribution of the random variable $\xi_t(\omega)$ equals $\mu$ for all $t$. Moreover, if it is known that for every probability measure $\sigma$ on $\mathbb{R}^d$ there exists a unique weak solution $\xi^\sigma_0$ of the indicated stochastic equation on $[0, +\infty)$ with initial distribution $\sigma$ and $P_\sigma$ is the distribution of $\xi^\sigma$, then the measure $\mu$ is invariant for the semigroup

$$T_t f(x) = \int_{C([0, +\infty), \mathbb{R}^d)} f(\xi_t) \ P_\delta (d\xi)$$
on the space of bounded Borel functions, that is, for every bounded continuous function $f$ and every $t \geq 0$ we have the identity

$$\int_{\mathbb{R}^d} T_t f(x) \mu(dx) = \int_{\mathbb{R}^d} f(x) \mu(dx).$$

**Proof.** By the Ambrosio–Figalli–Trevisan superposition principle (see [35] and [36]) there exists a Borel probability measure $P_\mu$ on $C([0, T], \mathbb{R}^d)$ satisfying the following conditions:

i. $P_\mu(\{\xi : \xi_t \in Q\}) = \mu(Q)$ for every Borel set $Q$ and every $t \in [0, T]$,

ii. for every function $f \in C_0^\infty(\mathbb{R}^d)$, the mapping

$$(\xi, t) \mapsto f(\xi_t) - f(\xi_0) - \int_0^t \left[ a^{ij}(\xi_s) \partial_{x_i} \partial_{x_j} f(\xi_s) + b^j(\xi_s) \partial_{x_j} f(\xi_s) \right] ds$$

is a martingale with respect to $P_\mu$ and the filtration $\sigma(\xi_s, s \leq t)$. Therefore, according to [37, Proposition 2.1, Ch. IV], one can find a probability space $(Q, \mathcal{F}, P)$, a filtration $\mathcal{F}_t$, a continuous random process $\xi_t$ and a Wiener process $w_t$ adapted to $\mathcal{F}_t$ such that

$$d\xi_t = b(\xi_t) \ dt + \sqrt{2A(\xi_t)}dw_t.$$
on $[0, T]$ and the distribution of $\xi_t$ coincides with $P_\mu$, in particular, the one-dimensional distribution of the process $\xi_t$ does not depend on $t$ and equals $\mu$.

Assume now that the given stochastic equation has a unique weak solution for every initial distribution and $P_{\delta_x}$ is the distribution of the solution with initial condition $\delta_x$. We observe that due to our assumption about the uniqueness of solutions the mapping $\sigma \mapsto P_\sigma$ is Borel measurable when the spaces of measures are equipped with their weak topologies (or with metrics generating them). Indeed, even without any assumptions about uniqueness, the set of all pairs $(P, \sigma)$ of probability measures for which $P$ is a measure on the space $C([0, +\infty), \mathbb{R}^d)$ of continuous paths and $\sigma$ is a measure on $\mathbb{R}^d$ such that $\sigma$ is the image $P$ under the mapping $x \mapsto x(0)$ and the process

$$f(x(t)) - f(x(0)) - \int_0^t Lf(x(s)) \, ds$$

is a martingale with respect to the measure $P$ for all functions $f \in C^2_b(\mathbb{R}^d)$, is a Borel set in the product $\mathcal{P}(C([0, +\infty), \mathbb{R}^d)) \times \mathcal{P}(\mathbb{R}^d)$. This is seen from the fact that this set can be described by a countable number of equalities of the form

$$\int \psi_i \, dP = 0$$

with some countable collection of bounded Borel functions $\psi_i$ on the path space along with the equality of the measure $\sigma$ to the image of $P$ with respect to the operator $x \mapsto x(0)$. A countable collection $\{\psi_i\}$ arises because for verification of the martingale property we can use a countable collection of functions $f$, moreover, the martingale property itself can be verified only for rational times, and the comparison of the corresponding conditional expectations also employs a countable collection of functions.

According to [37, Theorem 5.3, Ch. IV] the measures $P_{\delta_x}$ form a Markov family and $T_t f$ is a semigroup on the space of bounded continuous functions. Note that for every cylindrical set $C$ one has

$$P_\mu(C) = \int_{\mathbb{R}^d} P_{\delta_x}(C) \, \mu(dx),$$

which implies the equality

$$\int_{C([0, +\infty), \mathbb{R}^d)} f(\xi_t) \, P_\mu(d\xi) = \int_{\mathbb{R}^d} \int_{C([0, +\infty), \mathbb{R}^d)} f(\xi_t) \, P_x(d\xi) \, \mu(dx)$$

for every bounded continuous function $f$. It remains to observe that the left side of the last equality is the integral of $f$ against the measure $\mu$ and the right side is the integral of $T_t f$ against the measure $\mu$. \hfill \Box

In the general case the stationary Kolmogorov equation can have several different probability solutions (see [3, Chapter 4]). Sufficient conditions for uniqueness of probability solutions have been obtained in [38] in the case where the matrix $A$ satisfies the classical Dini condition, but by Theorem 3.5 and the results in [11] their justification extends without changes to the case of the matrix $A$ satisfying the Dini mean oscillation condition. Thus, the following assertion is true.
Theorem 4.7. Suppose that \(a^j\) and \(b^i\) are defined on all of \(\mathbb{R}^d\), \(b^i \in L^1_{\text{loc}}\), and for every ball \(B\) one can find a number \(\nu_B > 0\) and a continuous nonnegative increasing function \(w_B\) on \([0, 1]\) such that \(w_B(0) = 0\), the integral \(\int_0^1 \frac{w(t)}{t} \, dt\) converges, and

\[
\nu_B \cdot 1 \leq A(x) \leq \frac{1}{\nu_B} \cdot 1, \quad \sup_{x \in B} \frac{1}{|B(x,r)|} \int_{B(x,r)} |a^j(y) - a^j_B(x,r)| \, dy \leq w_B(r) \quad \forall \ r \in (0, 1].
\]

Let \(q\) be a probability solution to Equation (1.1) such that at least one of following conditions is fulfilled:

i. \((1 + |x|)^{-2}a^j, (1 + |x|)^{-1}b^i \in L^1(q \, dx),

ii. there exists a function \(V \in C^2(\mathbb{R}^d)\) with \(\lim_{|x| \to \infty} V(x) = +\infty\) and \(LV \leq C_1 + C_2 V\).

Then \(q\) is a unique probability solution.

Applying Zvonkin’s transform, we obtain the following sufficient condition for uniqueness that agrees with Corollary 4.5.

Corollary 4.8. Suppose that the coefficients \(a^j\) and \(b^i\) are defined on all of \(\mathbb{R}^d\) and we can find a number \(\nu > 0\) and an increasing continuous function \(\omega\) on \([0, +\infty)\) such that

\[
\nu \cdot 1 \leq A(x) \leq \frac{1}{\nu} \cdot 1, \quad \sup_{x \in \mathbb{R}^d} \frac{1}{|B(x,r)|} \int_{B(x,r)} |a^j(y) - a^j_B(x,r)| \, dy \leq \omega(r),
\]

\(\omega(0) = 0\) and the integral \(\int_0^1 \frac{\omega(r)}{r} \, dr\) converges.

Let also \(b^i = b^i_1 + b^i_2\), where \(b_2 \in L^p(\mathbb{R}^d)\) with \(p > d\) and

\[
\lim_{|x| \to \infty} \langle b_1(x), x \rangle \leq C_1 - C_2 |x|^{k+1}, \quad |b_1(x)| \leq C_3 (1 + |x|^k),
\]

for some \(C_1, C_2, C_3 > 0\) and \(k > 0\). Then there exists a unique probability solution \(q\) to Equation (1.1) on \(\mathbb{R}^d\).

Proof. The existence is proved in Corollary 4.5. After the change of coordinates as in the proof of Corollary 4.5 the coefficients of the new equation satisfy condition (ii) in Theorem 4.7, hence a probability solution is unique.

We do not know whether the continuity of the matrix \(A\) is sufficient for the uniqueness of a probability solution.

It is proved in [38] that if the matrix \(A\) satisfies the classical Dini condition, then the ratio \(\nu = \sigma/q\) of two positive solutions \(\sigma\) and \(q\) belongs to the Sobolev space \(W^{2,1}_{\text{loc}}(\mathbb{R}^d)\). The same is true if \(A\) satisfies the Dini mean oscillation condition. Note that in the general case a solution to Equation (1.1) (even with a Hölder continuous matrix \(A\)) does not have Sobolev derivatives.
Acknowledgments

We thank two anonymous referees for very useful corrections and suggestions. S.V. Shaposhnikov is a winner of the contest “Young Mathematics of Russia”, and thanks its jury and sponsors.

Disclosure statement

The authors report there are no competing interests to declare.

Funding

This research is supported by the Russian Science Foundation Grant 22-11-00015 (the results in Sections 2 and 3) and the CRC 1283 at Bielefeld University (the results in Section 4).

References

[1] Khasminskii, R. Z. (2012). Stochastic Stability of Differential Equations. 2nd ed. Heidelberg: Springer.
[2] Bogachev, V. I., Röckner, M. (2001). A generalization of Khasminskii’s theorem on the existence of invariant measures for locally integrable drifts. Theory Probab. Appl. 45(3): 363–378. DOI: 10.1137/S0040585X97978348.
[3] Bogachev, V. I., Krylov, N. V., Röckner, M., Shaposhnikov, S. V. (2015). Fokker–Planck–Kolmogorov Equations. Providence, RI: American Mathematical Society.
[4] Bogachev, V. I., Röckner, M., Shaposhnikov, S. V. (2012). On positive and probability solutions of the stationary Fokker–Planck–Kolmogorov equation. Dokl. Math. 85(3):350–354. DOI: 10.1134/S1064562412030143.
[5] Lee, H., Trutnau, G. (2021). Existence and regularity of infinitesimally invariant measures, transition functions and time-homogeneous Itô-SDEs. J. Evol. Equ. 21(1):601–623. DOI: 10.1007/s00028-020-00593-y.
[6] Bogachev, V. I., Krylov, N. V., Röckner, M. (2001). On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions. Commun. Partial Differ. Equ. 26(11–12):2037–2080. DOI: 10.1081/PDE-100107815.
[7] Sjögren, P. (1973). On the adjoint of an elliptic linear differential operator and its potential theory. Ark. Mat. 11(1–2):153–165. DOI: 10.1007/BF02388513.
[8] Sjögren, P. (1975). Harmonic spaces associated with adjoints of linear elliptic operators. Ann. Inst. Fourier. 25(3–4):509–518. DOI: 10.5802/aif.595.
[9] Bogachev, V. I., Shaposhnikov, S. V. (2017). Integrability and continuity of solutions to double divergence form equations. Annali di Matematica. 196(5):1609–1635. DOI: 10.1007/s10231-016-0631-2.
[10] Dong, H., Kim, S. (2017). On C1, C2, and weak type-(1, 1) estimates for linear elliptic operators. Comm. Partial Differ. Equ. 42(3):417–435. DOI: 10.1080/03605302.2017.1278773.
[11] Dong, H., Escauriaza, L., Kim, S. (2018). On C1, C2, and weak type-(1, 1) estimates for linear elliptic operators: Part II. Math. Ann. 370(1–2):447–489. DOI: 10.1007/s00208-017-1603-6.
[12] Leitao, R., Pimentel, E. A., Santos, M. S. (2020). Geometric regularity for elliptic equations in double-divergence form. Analysis and PDE. 13(4):1129–1144. DOI: 10.2140/apde.2020.13.1129.
[13] Bauman, P. (1984). Positive solutions of elliptic equations in nondivergent form and their adjoints. Ark. Mat. 22(1–2):153–173. DOI: 10.1007/BF02384378.
[14] Bauman, P. (1984). Equivalence of the Green’s functions for diffusion operators in $\mathbb{R}^n$: a counter example. Proc. Amer. Math. Soc. 91:64–68.
[15] Fabes, E. B., Stroock, D. W. (1984). The $L^p$-integrability of Green’s functions and fundamental solutions for elliptic and parabolic equations. *Duke Math. J.* 51(4):977–1016. DOI: 10.1215/S0012-7094-84-05145-7.

[16] Apushkinskaya, D. E., Nazarov, A. I. (2022). The normal derivative lemma and surrounding issues. *Uspekhi Mat. Nauk.* 77(2):3–68. (Russian Math. Surveys. 77(2):189–249).

[17] Mamedov, F. I. (1992). On the Harnack inequality for the equation formally adjoint to a linear elliptic differential equation. *Sib. Matem. Zh.* 33(5):100–106. (in Russian; English transl.: *SibMathJ.* 33(5): 835–841).

[18] Escudero, L. (2000). Bounds for the fundamental solution of elliptic and parabolic equations in nondivergence form. *Comm. Partial Differ. Equ.* 25(5–6):821–845. DOI: 10.1080/03605300008821533.

[19] Zvonkin, A. K. (1974). A transformation of the phase space of a diffusion process that removes the drift. 93. *Math. USSR Sb.* 22(1):129–149. (in Russian; English transl.: *Math. USSR-Sbornik* 22(1): 129–149). DOI: 10.1070/SM1974v022n01ABEH001689.

[20] Flandoli, F., Issoglio, E., Russo, F. (2016). Multidimensional stochastic differential equations with distributional drift. *Trans. Amer. Math. Soc.* 369(3):1665–1688. DOI: 10.1090/tran/6729.

[21] Röckner, M., Zhang, X. (2021). Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli.* 27(2):1131–1158. DOI: 10.3150/20-BEJ1268.

[22] Zhang, X., Zhao, G. (2018). Singular Brownian diffusion processes. *Commun. Math. Stat.* 6(4):533–581. DOI: 10.1007/s40304-018-0164-7.

[23] Krasovskii, T. I. (1979). Degenerate elliptic equations and nonuniqueness of solutions to the Kolmogorov equation. *Dokl. Math.* 100(1):354–357. DOI: 10.1134/S1064562419040112.

[24] Bogachev, V. I., Röckner, M., Shaposhnikov, S. V. (2019). Convergence in variation of solutions of nonlinear Fokker–Planck–Kolmogorov equations to stationary measures. *J. Func. Anal.* 276(12):3681–3713. DOI: 10.1016/j.jfa.2019.03.014.

[25] Cardaliaguet, P., Poirretta, A. (2020). An introduction to mean field game theory. In Cardaliaguet, P., Porretta, A. (eds.) *Mean Field Games. Lecture Notes Math.*, vol. 2281. Cham: Springer.

[26] Carmona, R., Delarue, F. (2018). *Probabilistic Theory of Mean Field Games with Applications. I. Mean Field FBSDEs, Control, and Games. II. Mean Field Games with Common Noise and Master Equations.* Springer, Cham.

[27] Blanc, X., Le Bris, C., Lions, P.-L. (2018). On correctors for linear elliptic homogenization in the presence of local defects. *Comm. Partial Differ. Equ.* 43(6):965–997. DOI: 10.1080/03605302.2018.1484764.

[28] Krylov, N. V. (2007). Parabolic and elliptic equations with VMO coefficients. *Commun. Partial Differ. Equ.* 32(3):453–475. DOI: 10.1080/03605300600781626.

[29] Krylov, N. V. (2008). *Lectures on Elliptic and Parabolic Equations in Sobolev Spaces.* Providence, RI: American Mathematical Society.

[30] Kruzhkov, S. N. (1964). A priori bounds and some properties of solutions of elliptic and parabolic equations. *Matem. Sb.* 65(4):522–570. (in Russian; English transl.: *Ten papers on differential equations and functional analysis*, Am. Math. Soc. Trans. Ser. 2. 68, pp. 169–220. Amer. Math. Soc. Providence, Rhode Island, 1968).

[31] Oleinik, O. A., Kruzhkov, S. N. (1961). Quasi-linear second-order parabolic equations with many independent variables. *Uspehi Mat. Nauk.* 16(5):105–146. (in Russian; English transl. Russ. Math. Surv. 16(5): 115–155). DOI: 10.1070/RM1961v016n05A04114.

[32] Gilbarg, D., Trudinger, N. S. (1977). *Elliptic Partial Differential Equations of Second Order.* Berlin, New York: Springer-Verlag.

[33] Trudinger, N. S. (1980). Local estimates for subsolutions and supersolutions of general second order elliptic quasilinear equations. *Invent Math.* 61(1):67–79. DOI: 10.1007/BF01389895.

[34] Hwang, S., Kim, S. (2020). Green’s function for second order elliptic equations in non-divergence form. *Potential Anal.* 52(1):27–39. DOI: 10.1007/s11118-018-9729-z.
[35] Trevisan, D. (2016). Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron. J. Probab.* 21(none):41. pp. DOI: 10.1214/16-EJP4453.

[36] Bogachev, V. I., Röckner, M., Shaposhnikov, S. V. (2021). On the Ambrosio–Figalli–Trevisan superposition principle for probability solutions to Fokker–Planck–Kolmogorov equations. *J Dyn Diff Equat.* 33(2):715–739. DOI: 10.1007/s10884-020-09828-5.

[37] Ikeda, N., Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes.* 2nd ed. Amsterdam: North-Holland.

[38] Bogachev, V. I., Shaposhnikov, S. V. (2021). Uniqueness of probability solutions to the Kolmogorov equation with a diffusion matrix satisfying Dini’s condition. *Dokl. Math.* 104(3):322–325. DOI: 10.1134/S1064562421060041.