ON DIVERGENCE-FREE DRIFTS

GREGORY SEREGIN[a], LUIS SILVESTRE[b], VLADIMÍR ŠVERÁK[c], AND ANDREJ ZLATOŠ[d]

Abstract. We investigate the validity and failure of Liouville theorems and Harnack inequalities for parabolic and elliptic operators with low regularity coefficients. We are particularly interested in operators of the form $\partial_t - \Delta + b\cdot\nabla$ resp. $-\Delta + b\cdot\nabla$ with a divergence-free drift b. We prove the Liouville theorem and Harnack inequality when $b \in L_\infty(BMO^{-1})$ resp. $b \in BMO^{-1}$ and provide a counterexample demonstrating sharpness of our conditions on the drift. Our results generalize to divergence-form operators with an elliptic symmetric part and a BMO skew-symmetric part. We also prove the existence of a modulus of continuity for solutions to the elliptic problem in two dimensions, depending on the non-scale-invariant norm $\|b\|_{L_1}$. In three dimensions, on the other hand, bounded solutions with L_1 drifts may be discontinuous.

1. Introduction

This paper is motivated by questions about the behavior of solutions of elliptic and parabolic equations with low regularity drift terms. A classical example is

$$\partial_t u + b\cdot\nabla u - \Delta u = 0 \tag{1.1}$$

considered in $\mathbb{R}^n\times]0,\infty[$, where b is a time-dependent vector field in $\mathbb{R}^n$. Of particular interest to us will be the case of divergence-free b (i.e., $\operatorname{div} b = 0$), which is relevant for applications to incompressible flows.

To describe the regularity conditions on the drift term, it is useful to recall some elementary dimensional analysis. Equation (1.1) is invariant under the following scaling transformations:

$$u(x,t) \to u^{(\lambda)}(x,t) = u(\lambda x, \lambda^2 t)\ , \tag{1.2}$$

$$b(x,t) \to b^{(\lambda)}(x,t) = \lambda b(\lambda x, \lambda^2 t)\ , \tag{1.3}$$

where $\lambda > 0$. Following the usual convention (see, e.g., [3]), we can say that u has dimension 0 and b has dimension -1. The classical theory (see, e.g., [18]) studies the question of under which conditions (1.1) can be considered as a perturbation of the heat equation. The required regularity on b is usually expressed as $b \in \mathcal{B}$, with $\mathcal{B}$ a suitable function space. Typically the borderline spaces for which one can still prove most of the deeper results[1] are scale-invariant under the scaling (1.3) of b, that is, $||b^{(\lambda)}||_{\mathcal{B}} = ||b||_{\mathcal{B}}$ (see, e.g., [18, 29]) [2]. The reason for this

[a]University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK.
[b]University of Chicago, 5374 University Ave., Chicago, IL 60637, USA.
[c]University of Minnesota, 206 Church St. SE, Minneapolis, MN 55455, USA.
[d]University of Wisconsin—Madison, 480 Lincoln Dr., Madison, WI 53706, USA.

[1]such as, for instance, the Harnack inequality for positive solutions

[2]For example, the Lebesgue spaces $L_{q,p} = L_t^p L_x^q$ are scale-invariant if and only if $2/p + q/n = 1$.

is as follows. The arguments in the proofs of the "deeper properties"[3] typically have to work on all (small) scales and we therefore need to control b on all scales, which naturally leads to the scale-invariant spaces.

Similar considerations can be made for elliptic equations of the form

$$-\Delta u + b\cdot\nabla u = 0\ , \tag{1.4}$$

with the elliptic scaling

$$u(x) \to u^{(\lambda)}(x) = u(\lambda x)\ , \tag{1.5}$$

$$b(x) \to b^{(\lambda)}(x) = \lambda b(\lambda x)\,. \tag{1.6}$$

Let us now consider the condition $\operatorname{div} b = 0$ and its consequences. (The relevant references include, for example, [36, 31] in the parabolic case and [22, 15] in the elliptic case.) Among the most important consequences are the following[4].

(i) The energy identity

$$\int_{\mathbb{R}^n} |u(x,t_2)|^2\,dx + \int_{t_1}^{t_2}\int_{\mathbb{R}^n} |\nabla u|^2\,dx\,dt = \int_{\mathbb{R}^n} |u(x,t_1)|^2\,dx \tag{1.7}$$

is exactly the same as for the heat equation.

(ii) The integral $\int_{\mathbb{R}^n} u(x,t)\,dx$ is conserved:

$$\int_{\mathbb{R}^n} u(x,t_2)\,dx = \int_{\mathbb{R}^n} u(x,t_1)\,dx\,. \tag{1.8}$$

J. Nash showed in his famous paper [28] (inequality (8) on page 936) that one can obtain from (i) and (ii) the point-wise upper bound

$$|G(x,t;y,s)| \le \frac{C}{(t-s)^{n/2}} \tag{1.9}$$

on the fundamental solutions $G(x,t;y,s)$, with C depending only on the dimension. Therefore, this bound also holds for solutions of (1.1) when $\operatorname{div} b = 0$, with practically no other assumptions on b. The heuristic behind this estimate is that in an incompressible fluid, mixing can enhance the decay of, say, a temperature field but it cannot slow it down. Nash's simple argument proving this heuristics is very elegant. There are many other results in this direction, see for example [36, 31]. Bound (1.9) can also be integrated in time to obtain (global) estimates of $\sup_x |u(x)|$ for the elliptic problem

$$-\Delta u + b\cdot\nabla u = f\,, \tag{1.10}$$

with $f \in L_{n/2+\delta}$, a divergence-free b, and practically no other assumptions.

Since the condition $\operatorname{div} b = 0$ has such strong consequences for the L_∞-bounds, it is natural to ask about its effects on other properties of the solutions. For instance, can the standard

[3]The definition of what is meant by a "deeper property" is of course somewhat ambiguous. We already mentioned the Harnack inequality as an example. On the other hand, the weak maximum principle would not be considered as such in this context.

[4]To derive these consequences, one needs to assume that the formal integration by parts used to obtain them is valid. We are ignoring this technical issue for the moment.

assumptions on the drift term b needed, say, for the Harnack inequality be relaxed when $\operatorname{div} b = 0$? Similar questions have been considered, for example, in [31, 36].

It turns out that the condition $\operatorname{div} b = 0$ can be used to relax the regularity assumptions on b under which one can prove the Harnack inequality and other results. However, the effects are not as dramatic as in the case of Nash's upper bound (1.9), for which not much is needed beyond $\operatorname{div} b = 0$. In particular, it seems that even with the condition $\operatorname{div} b = 0$ one cannot significantly "break the scaling". Indeed, to be able to prove the "deeper regularity properties" of the solutions (as discussed above), we still need to assume that b belongs, at least locally, to a scale-invariant space $\mathcal{B}$. The norm can be weaker than in the absence of the assumption $\operatorname{div} b = 0$, but it still has to be scale invariant or stronger on the small scales. For example, the results of [31] imply that the Harnack inequality, Hölder continuity of solutions, and the Aronson estimate for fundamental solutions[5] remain true when $b \in L_\infty(L_\infty^{-1})$, where L_∞^{-1} denotes distributions which are first derivatives of bounded measurable functions. This should be compared to the condition $b \in L_{n,\infty}$, which naturally comes up when the assumption $\operatorname{div} b = 0$ is dropped[6]. Note that both L_n and L_∞^{-1} are scale-invariant.

The assumption $\operatorname{div} b = 0$ can be used to reformulate equation (1.1) in the following way. When $\operatorname{div} b = 0$, we can write $b = \operatorname{div} d$ for an anti-symmetric tensor $d = (d_{ij})$[7]. Moreover, by introducing a suitable "gauge condition"[8], we can assume that the derivatives of d have similar regularity as b. Since b has dimension -1 with respect to the natural scaling of (1.1), the tensor d has dimension 0, that is, it scales as

$$d(x,t) \to d(\lambda x, \lambda^2 t) \tag{1.11}$$

when u is scaled by (1.2).

Replacing b by the potential d, equation (1.1) becomes

$$\partial_t u - \operatorname{div}(A\nabla u) = 0 \tag{1.12}$$

where $A = \mathbb{I} + d$. This is a divergence-form equation with a non-symmetric leading term. Such equations (including the versions with lower-order terms) have been studied in [31] under the assumptions that the coefficients a_{ij} are bounded measurable functions satisfying the ellipticity condition

$$(A\xi)\cdot\xi \ge \nu|\xi|^2\,. \tag{1.13}$$

The results of [31] show, roughly speaking, that most of the results which are valid for symmetric A are also true in the non-symmetric case. The transformation of (1.1) to (1.12) has been used in many other works (see, for instance, [8]).

[5] $c_1(t-s)^{-n/2}\exp[c_2|x-y|^2/(t-s)] \le G(x,t;y,s) \le c_3(t-s)^{-n/2}\exp[c_4|x-y|^2/(t-s)]$, see [1].

[6] Strictly speaking, as far as we are aware, when we do not assume $\operatorname{div} b = 0$, most of the regularity results above are proved for $b \in L_{q,p}$ with $2/p + n/q = 1$ and $p < \infty$ (see [29]), but not in the borderline case $p = \infty,\ q = n$.

[7] For $n = 3$, this corresponds to introducing the vector potential $\tilde d$ such that $b = \operatorname{curl}\tilde d$.

[8] such as $d_{kl,j} + d_{jk,l} + d_{lj,k} = 0$, which for $n = 3$ and $b = \operatorname{curl}\tilde d$ corresponds to $\operatorname{div}\tilde d = 0$

In the elliptic case, Mazja and Verbitsky [22] studied (among other things) the bi-linear form

$$(u,v) \ \to \int_{\mathbb{R}^n} (A\nabla u)\cdot \nabla v \, dx \, . \tag{1.14}$$

The form is obviously continuous in $\dot{H}^1$ when A is bounded, but it turns out that the boundedness of the coefficients is not a necessary condition for the boundedness of the form. The form is still continuous on $\dot{H}^1$ if the symmetric part of A is bounded and the anti-symmetric part of A is in the John-Nirenberg space BMO (bounded mean oscillation). This is a consequence of the following two facts:
(i) If A is anti-symmetric, the form (1.14) can be factored through the determinants $\frac{\partial(u,v)}{\partial(x_i,x_j)}$.
(ii) The determinants have "better than expected" regularity: when u, v are in $\dot{H}^1$, the determinants are not only in L_1, but they are in fact in the Hardy space $\mathcal{H}^1$, the dual space of BMO (see [5]).

It is natural to expect that much of the classical regularity results for elliptic and parabolic equations with measurable coefficients in divergence form will remain valid if the leading part A is of the form $A = a + d$, with a symmetric, bounded and satisfying the usual ellipticity condition (1.13), and d anti-symmetric and belonging to BMO in the elliptic case, and to $L_\infty(BMO)$ in the parabolic case.

Indeed, let $Q_- = \mathbb{R}^n \times \mathbb{R}_-$ (with $\mathbb{R}_- =]-\infty, 0[$) and assume that

$$A = a + d, \tag{1.15}$$

where $a \in L_\infty(Q_-; \mathbb{M}^{n\times n})$ is a symmetric matrix satisfying

$$\nu \mathbb{I} \le a \le \nu^{-1}\mathbb{I} \tag{1.16}$$

and $d \in L_\infty(\mathbb{R}_-; BMO(\mathbb{R}^n; \mathbb{M}^{n\times n}))$ is a skew symmetric matrix, that is,

$$d = -d^* \tag{1.17}$$

for all $(x,t) \in Q_-$. Here $\nu > 0$, $\mathbb{I}$ is the identity in the space $\mathbb{M}^{n\times n}$ of $n\times n$-matrices and d^* is the transpose of d. Let also $B(x,r)$ be the ball of radius r centered at $x \in \mathbb{R}^n$, and $Q(z,r) = B(x,r)\times]t-r^2, t[$ a parabolic ball in $\mathbb{R}^{n+1}$ centered at point $z = (x,t)$. Finally, let $B = B(0,1)$ and $Q = Q(0,1)$. We then prove the following parabolic Harnack inequality and Liouville theorem for suitable weak solutions (see Definition 2.1) to (1.12).

Theorem 1.1. *If the matrix A satisfies conditions (1.15)–(1.17), then there exists $C > 0$, depending only on n, ν, and $\|d\|_{L_\infty(-1,0;BMO(B))}$, such that for any nonnegative suitable weak solution u to (1.12) on Q we have*

$$\sup_{(y,s)\in Q(z_R,R/2)} u(y,s) \le C \inf_{(y,s)\in Q(z,R/2)} u(y,s), \tag{1.18}$$

whenever $Q(z,R) \subset Q$. Here, $z_R = (x, t - R^2/2)$.

Theorem 1.2. *If the matrix A satisfies conditions (1.15)–(1.17), then the only bounded ancient suitable weak solutions to (1.12) on Q_- are the constant functions.*

Remark. Of course, the corresponding elliptic results follow immediately by taking time-independent solutions. In addition, in Section 3 we provide a second — short and elementary — proof of the Liouville theorem for weak (sub)solutions (see Definition 3.1) to (1.4) in $\mathbb{R}^2$.

Recall that the norm in the space $BMO(\Omega;\mathbb{M}^{n\times n})$ is

$$\|d\|_{BMO(\Omega;\mathbb{M}^{n\times n})} = \sup\left\{ \frac{1}{|B(0,r)|} \int\limits_{B(x,r)} |d-[d]_{x,r}|\,dx \,:\, B(x,r)\Subset\Omega \right\},$$

with $[d]_{x,r}$ the average of d over $B(x,r)$.

We note that the space BMO is invariant under the scaling (1.11), and hence these results are again in line with the argument that to preserve the "deeper properties" of the solutions, one cannot "break the scaling". One of the goals of this paper is to present some evidence for this based on studying the failure of Liouville theorems under appropriate conditions.

Let us first look at (1.4) in $\mathbb{R}^n$. By the Liouville theorem for (1.4) we mean the usual statement that a bounded solution in $\mathbb{R}^n$ has to be constant. This is of course true for $b\equiv 0$. For the time being let us assume that the vector field b is locally smooth, hence the solutions u are also locally smooth and the only obstacles to the validity of the Liouville theorem are global.

The results of Stampacchia [34] imply the following:

(L) *If $b\in L_n(\mathbb{R}^n)$, then the Liouville theorem for (1.4) holds.*

This is easy for $n=1$, and for $n=2$, there is also a relatively simple proof based on the energy estimate. The proof for $n\ge 3$ can be accomplished by using the Hölder estimate or the Harnack inequality (see Sections 7 and 8 of [34]). If $n\ge 2$, then by Theorem 3.2 in [29], the condition on b can be weakened to $\liminf_{R\to\infty}\sup_{|x|=R}||b||_{L_n(B(x,R\delta))} < c_n$ for some $\delta>0$, where $c_n>0$ is a fixed dimension-dependent constant. This result implies in particular that (L) remains true for $n\ge 2$ when

$$|b(x)|\le \frac{C}{|x|} \qquad \text{for large } |x|. \tag{1.19}$$

In dimension $n=1$, condition (1.19) is sufficient when $C\le 1$, as one can check by direct integration. With $C>1$, however, (1.19) is no longer sufficient. This can be illustrated by the example

$$b(x)=\frac{2x}{1+x^2} \qquad \text{and} \qquad u(x)=\arctan(x)\,, \tag{1.20}$$

which was pointed out in this context to one of the authors in 1997 by Joel Spruck. The trivial extension of this example to higher dimensions is

$$b(x_1,\dots,x_n)=\left(\frac{2x_1}{1+x_1^2},0,\dots,0\right) \qquad \text{and} \qquad u(x_1,\dots,x_n)=\arctan(x_1)\,. \tag{1.21}$$

We note that the vector field b in (1.21) belongs to the space $(BMO)^{-1}(\mathbb{R}^n)$, since

$$\frac{2x}{1+x^2}=\frac{d}{dx}\log(1+x^2) \tag{1.22}$$

and $\log(1+x_1^2) \in BMO(\mathbb{R}^n)$.

This example and Theorem 1.2, which establishes the Liouville theorem for $b \in (BMO)^{-1}$ and $\operatorname{div} b = 0$, together show that the divergence-free condition can play an important role in Liouville theorems for equations with drift terms. On the other hand, we now provide a counter-example to the Liouville theorem with a divergence-free b on $\mathbb{R}^2$ which is in some sense not too far from $(BMO)^{-1}$. Recall that the *stream function* of a divergence-free vector field b on $\mathbb{R}^2$ is $H : \mathbb{R}^2 \to \mathbb{R}$ such that

$$b(x) = \nabla^\perp H(x) = (H_{x_2}(x), -H_{x_1}(x)). \tag{1.23}$$

We therefore have

$$-\Delta + b \cdot \nabla = -\operatorname{div}(A\,\nabla), \tag{1.24}$$

where $A(x) = \mathbb{I} + d(x)$ has skew-symmetric part

$$d(x) = \begin{pmatrix} 0 & H(x) \\ -H(x) & 0 \end{pmatrix}.$$

Theorem 1.3. *There exists a divergence-free vector field $b \in C^\infty(\mathbb{R}^2)$ with all derivatives bounded and a stream function satisfying $|H(x)| \le C \ln|x| \ln\ln|x|$ for some C and all large enough $|x|$ such that* (1.4) *has a non-constant bounded classical solution.*

This illustrates, to some degree, the important role of scale invariance of the assumptions in the Liouville result. In particular, it seems unlikely that one can significantly "break the scaling" even if we assume that $\operatorname{div} b = 0$.

We conjecture that similar negative conclusions can be arrived at when considering questions about Hölder continuity of solutions of (1.4) (as well as the Harnack inequality). For example, it seems unlikely that the condition $\operatorname{div} b = 0$ is sufficient to get a C^α-bound on solutions u in the unit ball $B = B(0,1)$ under the assumptions $|u| \le C$ and $||b||_{L_{n-\delta}} \le C$. (Here we assume that all the functions involved are smooth, but only the indicated quantities are controlled, and we are interested in an a-priori bound.)

Related to this are our last two main results, concerning distributional solutions u (see Proposition 4.1) of (1.4) in B with divergence-free $b \in L_1(B)$. The first establishes a logarithmic modulus of continuity of such solutions in two dimensions, depending only on $\|b\|_{L_1(B)}$ and $\|u\|_{L_\infty(B)}$. However, due to the low regularity assumed on the vector field b and u solving (1.4) only in the distributional sense, our result is restricted to those solutions which can be obtained as weak-star L_∞-limits of solutions with drifts in $L_2(B)$.

Theorem 1.4. *Let B be the unit ball in $\mathbb{R}^2$ and let $(b_m, u_m) \in L_2(B) \times L_\infty(B)$ be a sequence of divergence-free drifts b_m and distributional solutions u_m to (1.4) with $b = b_m$. Assume that u_m are uniformly bounded in B and*

$$b_m \to b \quad \text{in} \quad L_1(B),$$

$$u_m \overset{\star}{\rightharpoonup} u \quad \text{in} \quad L_\infty(B).$$

Then the function u is a distributional solution to (1.4). Moreover,

$$u \in H^1_{\mathrm{loc}}(B) \cap C_{\mathrm{loc}}(B)$$

and at the origin u has the modulus of continuity

$$\sup_{x\in B(0,r)} |u(x)-u(0)| \le \frac{C\left(1+||b||_{L_1(B)}\right)^{1/2}}{\sqrt{-\log r}}||u||_{L_\infty(B)} \tag{1.25}$$

with a universal $C>0$.

In three (and more) dimensions this result is false. Indeed, there exists no modulus of continuity of classical solutions depending only on $\|b\|_{L_1(B)}$ and $\|u\|_{L_\infty(B)}$, and distributional solutions $u\in L_\infty(B)\cap H^1(B)$ with divergence-free $b\in L_1(B)$ may be discontinuous.

Theorem 1.5. *Let B be the unit ball in $\mathbb{R}^3$.*

(i) There is $c>0$ such that for each $\varepsilon>0$ there is a smooth divergence-free drift b with $||b||_{L_1(B)}\le c$ and a smooth u with $||u||_{L_\infty(B)}\le 1$, solving (1.4) in B and satisfying

$$u(0,0,\varepsilon)-u(0,0,0)\ge c^{-1}.$$

(ii) There is a divergence-free drift $b\in L_1(B)$ and a distributional solution $u\in H^1(B)\cap L_\infty(B)$ of (1.4) in B which can be approximated by a smooth sequence (b_m,u_m) in the sense of Theorem 1.4, but u is discontinuous at the origin.

Our paper is organized as follows. In the next section, we develop local regularity theory for parabolic operators (1.12) under the assumption that the skew-symmetric part of A is in BMO, and prove Theorems 1.2 and 1.1. The important step of our approach is a higher integrability of suitable weak solutions. This allows us to adopt Moser's method for proving the Harnack inequality that implies Hölder continuity of suitable weak solutions and Liouville type theorems for ancient suitable weak solutions. All these results hold true for the heat equation with a drift $b\in L_\infty(BMO^{-1})$ as a particular case. In this connection, we would like to mention the recent paper [10], of which we learned while writing the present manuscript. In [10], among other questions, the Cauchy problem for the heat operator with the drift term from $L_\infty(BMO^{-1})$ has been considered and Hölder continuity of solutions has been proved. The authors of [10] follow the Caffarelli-Vaseur approach [4]. In Section 3, an elementary proof of a Liouville theorem in the two-dimensional elliptic case is provided and Theorem 1.3 is proved. Theorems 1.4 and 1.5 are proved in Section 4.

Acknowledgement. GS was partially supported by the RFFI grant 08-01-00372-a. The other authors were supported in part by NSF grants DMS-1001629 (LS), DMS-0800908 (VS), DMS-1113017 and DMS-1056327 (AZ). LS and AZ also acknowledge partial support by Alfred P. Sloan Research Fellowships.

2. Some results for parabolic equations

The main goal of this section is to prove Theorems 1.2 and 1.1. We consider (1.12) in $Q_-=\mathbb{R}^n\times\mathbb{R}_-$, with the matrix A satisfying (1.15)–(1.17). We will study the so-called suitable weak solutions to (1.12). In what follows we will use the abbreviated notation

$$B(r)=B(0,r),\quad B=B(1),\quad Q(r)=Q(0,r),\quad Q=Q(1),$$

as well as $z=(x,t)$.

Definition 2.1. Function u is said to be a *suitable weak solution* to equation (1.12) in the parabolic ball $Q(R)$ if it satisfies

$$u \in L_{2,\infty}(Q(R)) \cap W^{1,0}_2(Q(R)), \tag{2.1}$$

$$\int\limits_{Q(R)} u\, \partial_t \varphi dz = \int\limits_{Q(R)} (A\nabla u)\cdot \nabla\varphi dz \qquad \forall \varphi \in C^\infty_0(Q(R)), \tag{2.2}$$

and for a.e. $t_0 \in]-R^2, 0[$, the local energy inequality

$$\frac{1}{2}\int\limits_{B(R)} \varphi(x,t_0)|u(x,t_0)|^2 dx + \int\limits_{-R^2}^{t_0}\int\limits_{B(R)} \varphi\nabla u\cdot a\nabla u dz \le$$

$$\le \frac{1}{2}\int\limits_{-R^2}^{t_0}\int\limits_{B(R)} |u|^2\partial_t\varphi dz - \int\limits_{-R^2}^{t_0}\int\limits_{B(R)} (A\nabla u)\cdot\nabla\varphi u dz \tag{2.3}$$

holds for all non-negative test-functions $\varphi \in C^\infty_0(B(R)\times]-R^2, R^2[)$.

The function $u : Q_- \to \mathbb{R}$ is called an *ancient suitable weak solution* to (1.12), if it is a suitable weak solution to (1.12) in $Q(R)$ for any $R > 0$.

It is not clear whether one can show that any solution to (1.12), subject to assumptions (2.1) and (2.2), satisfies local energy inequality (2.3). In this respect the situation is similar to the Navier-Stokes equations: there is a certain cancelation due to the skew symmetric matrix d which works well in global setting, i.e., when initial-boundary value problems are under consideration. The corresponding procedure is relatively routine and leads to the existence of global solutions which satisfy the inequalities in Definition 2.1 at least locally.

We now outline the main points of our approach. The structure of equation (1.12) admits a modification of the technique developed by J. Moser in [25]–[27] and get Hölder continuity of suitable weak solutions. This property, together with scaling invariance, leads to the Liouville theorem. The main tool of proving Hölder continuity is the Harnack inequality. We prove the Harnack inequality for smooth solution by the method of J. Moser. Extension of the Harnack inequality to suitable weak solutions is provided by higher integrability of the the spatial gradient. Here, our arguments use an approach due to M. Gianquinta and M. Struwe, see [12].

2.1. **Local set-up and higher integrability.** Equation (1.12) is invariant with respect to translations and the following scaling

$$u^\lambda(x,t) = u(\lambda x, \lambda^2 t), \qquad A^\lambda(x,t) = A(\lambda x, \lambda^2 t) \tag{2.4}$$

for any positive λ. This allows us to reduce all considerations to some canonical domain, say, to $Q = Q(1)$.

So, we consider equation (1.12) in the unit parabolic cylinder. Matrix A is split into two parts as in (1.15) with matrices $a \in L_\infty(Q;\mathbb{M}^{n\times n})$ and $d \in L_\infty(-1,0;BMO(B;\mathbb{M}^{n\times n}))$ satisfying conditions (1.16) and (1.17)

In what follows, we shall denote by c positive constants depending only on n and ν. We let $\|d\|_{L_\infty(BMO)}=\|d\|_{L_\infty(-1,0;BMO(B))}$ and denote mean values by

$$[f]_{x,r}=\frac{1}{|B(r)|}\int\limits_{B(x,r)}f(y)dy,\qquad (u)_{z_0,r}=\frac{1}{|Q(r)|}\int\limits_{Q(z_0,r)}u(z)dz.$$

The main result of this subsection is the following theorem.

Theorem 2.2. *Assume that u is a suitable weak solution to (1.12) in Q and matrices A, a, and b satisfy conditions (1.15)-(1.17). Then there exist two positive constants $p>2$ and C depending only on n, ν, and $\|d\|_{L_\infty(BMO)}$ such that $u\in L_p(Q(R))$ for any $R\in]0,1[$. Moreover, the following estimate is valid:*

$$\Big(\frac{1}{|Q(R)|}\int\limits_{Q(z_0,R)}|\nabla u|^pdz\Big)^{\frac{1}{p}}\leq C\Big(\frac{1}{|Q(6R)|}\int\limits_{Q(z_0,6R)}|\nabla u|^2dz\Big)^{\frac{1}{2}}\tag{2.5}$$

for all $Q(z_0,6R)\subset Q$ with $6R<\operatorname{dist}(x_0,\partial B)$ and $t_0-(6R)^2>-1$.

This theorem is a consequence of the reverse Hölder inequality, see [12] for further references. To prove the reverse Hölder inequality, we need a Caccioppoli's type inequality. To formulate it, let us introduce additional notation. Fix a non-negative cut-off functions $\varphi\in C^\infty_0(B(2))$ and $\chi_0(t)$ with the following properties:

$$\varphi(x)=1\qquad x\in B,\qquad \chi(t)=0\quad t\leq -4,$$

$$\chi_0(t)=(t+4)/3\quad -4<t<-1,\qquad \chi_0(t)=1\quad t\geq -1.$$

Now, for a point $z_0=(x_0,t_0)$ and for $R>0$ such that $Q(z_0,2R)\in Q$, we let

$$\chi_{t_0,2R}(t)=\chi_0((t-t_0)/R^2),\qquad \varphi_{x_0,2R}(x)=\varphi((x-x_0)/R).$$

And then we can introduce a mean value of u as in [12]

$$u_{x_0,2R}(t)=\int\limits_{B(x_0,2R)}u(x,t)\varphi^2_{x_0,2R}(x)dx\Big(\int\limits_{B(x_0,2R)}\varphi^2_{x_0,2R}(x)dx\Big)^{-1}.$$

In our particular situation, we have

Lemma 2.3. *(Caccioppoli's type inequality) Under assumptions of Theorem 2.2, the following inequality is valid:*

$$\frac{1}{2}\int\limits_B|\widehat{u}(x,t_0)|^2\varphi^2_{x_0,2R}(x)dx+\nu\int\limits_{-1}^{t_0}\int\limits_B\chi^2_{t_0,2R}\varphi^2_{x_0,2R}|\nabla\widehat{u}|^2dz\leq$$

$$\leq\frac{1}{2}\int\limits_{-1}^{t_0}\int\limits_B|\widehat{u}|^2\varphi^2_{x_0,2R}\partial_t\chi^2_{t_0,2R}dz-\int\limits_{-1}^{t_0}\int\limits_B\chi^2_{t_0,2R}(a\nabla\widehat{u})\cdot\nabla\varphi^2_{x_0,2R}\widehat{u}dz\tag{2.6}$$

$$-\int\limits_{-1}^{t_0}\int\limits_B \chi^2_{t_0,2R}((d-[d]_{x_0,2R})\nabla\widehat{u})\cdot\nabla\varphi^2_{x_0,2R}\widehat{u}dz,$$

where

$$\widehat{u}(x,t)=u(x,t)-u_{x_0,2R}(t).$$

Inequality (2.6) holds for a.a. $t_0\in]-1,0[$, for all $x_0\in B$, and for all $R>0$ subject to the additional condition $Q(z_0,R)\subset Q$.

Proof. There are two important points to note. The first one is that for any skew-symmetric matrix d_0, depending on t only, we have

$$\int\limits_Q d_0\nabla u\cdot\nabla\varphi u dz=0 \tag{2.7}$$

whenever $\varphi\in C^\infty_0(Q)$. The proof is straightforward integration by part.

The second point is that, see [12],

$$\partial_t u_{x_0,2R}\in L_{\frac32}(-1,0). \tag{2.8}$$

To see this, we take as test function in (2.2) the function $\varphi^2_{x_0,2R}(x)\eta(t)$ and conclude

$$\partial_t u_{x_0,2R}(t)=-\int\limits_{B(x_0,2R)} A(z)\nabla u(z)\cdot\nabla\varphi^2_{x_0,2R}(x)dx\Big/\int\limits_{B(x_0,2R)}\varphi^2_{x_0,2R}(x)dx \tag{2.9}$$

Next, we replace $u(x,t)$ with $\widehat{u}(x,t)+u_{x_0,2R}(t)$ in local energy inequality (2.3) and take $\varphi=\chi^2\varphi^2_{x_0,2R}$ with χ from $C^1_0(-1,1)$. Then terms which do not contain spatial derivatives can be transformed as follows

$$\frac12\int\limits_{B(x_0,2R)}|\widehat{u}(x,t_0)+u_{x_0,2R}(t_0)|^2\varphi^2_{x_0,2R}(x)dx=$$

$$=\frac12\int\limits_{B(x_0,2R)}|\widehat{u}(x,t_0)|^2\varphi^2_{x_0,2R}(x)dx+\frac12|u_{x_0,2R}(t_0)|^2\int\limits_{B(x_0,2R)}\varphi^2_{x_0,2R}dx,$$

and

$$\frac12\int\limits_{-1}^{t_0}\int\limits_B\varphi^2_{x_0,2R}(x)|\widehat{u}(x,t)+u_{x_0,2R}(t)|^2\partial_t\chi^2(t)dx\,dt=$$

$$\frac12\int\limits_{-1}^{t_0}\int\limits_B\varphi^2_{x_0,2R}(x)|\widehat{u}(x,t)|^2\partial_t\chi^2(t)dx\,dt-$$

$$-\int\limits_{-1}^{t_0}\chi^2(t)u_{x_0,2R}(t)\partial_t u_{x_0,2R}(t)dt\int\limits_{B(x_0,2R)}\varphi^2_{x_0,2R}dx+$$

$$+\frac{1}{2}|u_{x_0,2R}(t_0)|^2\chi^2(t_0)\int\limits_{B(x_0,2R)}\varphi^2_{x_0,2R}dx.$$

Now, the local energy inequality, together with the last two identities, implies

$$\frac{1}{2}\int\limits_{B(x_0,2R)}\chi^2(t_0)|\widehat{u}(x,t_0)|^2dx+\nu\int\limits_{-1}^{t_0}\int\limits_{B}\chi^2(t)\varphi^2_{x_0,2R}(x)|\nabla\widehat{u}(x,t)|^2dx\,dt\leq$$

$$\leq\frac{1}{2}\int\limits_{-1}^{t_0}\int\limits_{B}\varphi^2_{x_0,2R}(x)|\widehat{u}(x,t)|^2\partial_t\chi^2(t)dx\,dt-\int\limits_{-1}^{t_0}\int\limits_{B}\chi^2A\nabla\widehat{u}\cdot\nabla\varphi^2_{x_0,2R}\widehat{u}dx\,dt-$$

$$-\int\limits_{-1}^{t_0}\chi^2(t)u_{x_0,2R}(t)\partial_tu_{x_0,2R}(t)dt\int\limits_{B(x_0,2R)}\varphi^2_{x_0,2R}dx-$$

$$-\int\limits_{-1}^{t_0}\int\limits_{B}\chi^2A\nabla\widehat{u}\cdot\nabla\varphi^2_{x_0,2R}u_{x_0,2R}dx\,dt.$$

By the (2.9), the sum of the last two terms is zero and from (2.7) it follows that

$$\frac{1}{2}\int\limits_{B(x_0,2R)}|\widehat{u}(x,t_0)|^2\chi^2(t_0)\varphi^2_{x_0,2R}(x)dx+$$

$$+\nu\int\limits_{-1}^{t_0}\int\limits_{B}\chi^2(t)\varphi^2_{x_0,2R}(x)|\nabla\widehat{u}(x,t)|^2dx\,dt\leq$$

$$\leq\frac{1}{2}\int\limits_{-1}^{t_0}\int\limits_{B}\varphi^2_{x_0,2R}(x)|\widehat{u}(x,t)|^2\partial_t\chi^2(t)dx\,dt-$$

$$-\int\limits_{-1}^{t_0}\int\limits_{B}\chi^2a\nabla\widehat{u}\cdot\nabla\varphi^2_{x_0,2R}\widehat{u}dx\,dt-$$

$$-\int\limits_{-1}^{t_0}\int\limits_{B}\chi^2(d-[d]_{x_0,2R})\nabla\widehat{u}\cdot\nabla\varphi^2_{x_0,2R}\widehat{u}dx\,dt.$$

Here,

$$[d]_{x_0,2R}(t)=\frac{1}{|B(2R)|}\int\limits_{B(x_0,2R)}d(x,t)dx.$$

So, inequality (2.6) follows if we choose the cut-off function χ in an appropriate way. □

Proof of Theorem 2.2. Using known simple arguments, we can derive from (2.6) the following estimate

$$I \equiv \frac{1}{2}\int\limits_{B} |\widehat{u}(x,t_0)|^2\varphi^2_{x_0,2R}(x)dx + \int\limits_{-1}^{t_0}\int\limits_{B} \chi^2_{t_0,2R}\varphi^2_{x_0,2R}|\nabla\widehat{u}|^2dz \leq$$

$$\leq c\Big(\frac{1}{R^2}\int\limits_{Q(z_0,2R)} |\widehat{u}|^2dz + \frac{1}{R}\int\limits_{Q(z_0,2R)} (|\nabla\widehat{u}|\varphi_{x_0,2R}\chi_{t_0,2R})|\widehat{u}||d-[d]_{x_0,2R}|dz\Big).$$

We now fix an arbitrary number $s\in]1,2[$. Let us denote as usual $s'=s/(s-1)$. Then the right hand side of the latter inequality can be estimated with the help of Hölder inequality by

$$\frac{c}{R^2}\int\limits_{Q(z_0,2R)} |\widehat{u}|^2dz + \frac{c}{R}\int\limits_{t_0-(2R)^2}^{t_0}\Big(\int\limits_{B(x_0,2R)} |d-[d]_{x_0,2R}|^{s'}dx\Big)^{\frac{1}{s'}}\times$$

$$\times\Big(\int\limits_{B(x_0,2R)} (|\nabla\widehat{u}|\varphi_{x_0,2R}\chi_{t_0,2R})^s|\widehat{u}|^sdx\Big)^{\frac{1}{s}}.$$

Applying Hölder's inequality one more time, we find

$$I\leq \frac{c}{R^2}\int\limits_{Q(z_0,2R)} |\widehat{u}|^2dz+$$

$$+\frac{c}{R}R^{\frac{n}{s'}}\mathrm{ess}\sup_{t_0-(2R)^2<t<t_0}\ \sup_{B(x_0,2R)\subset B}\Big(\frac{1}{|B(2R)|}\int\limits_{B(x_0,2R)} |d-[d]_{x_0,2R}|^{s'}dx\Big)^{\frac{1}{s'}}\times$$

$$\times\int\limits_{t_0-(2R)^2}^{t_0}\Big(\int\limits_{B(x_0,2R)} |\nabla\widehat{u}|^2\varphi^2_{x_0,2R}\chi^2_{t_0,2R}dx\Big)^{\frac{1}{2}}\Big(\int\limits_{B(x_0,2R)} |\widehat{u}|^{\frac{2s}{2-s}}dx\Big)^{\frac{2-s}{2s}}\leq$$

$$\leq \frac{c}{R^2}\int\limits_{Q(z_0,2R)} |\widehat{u}|^2dz + \frac{c(s)}{R}R^{\frac{n}{s'}}\|d\|_{L_\infty(BMO)}\Big(\int\limits_{Q(z_0,2R)} |\nabla\widehat{u}|^2\varphi^2_{x_0,2R}\chi^2_{t_0,2R}dz\Big)^{\frac{1}{2}}\times$$

$$\times\Big(\int\limits_{t_0-(2R)^2}^{t_0}\Big(\int\limits_{B(x_0,2R)} |\widehat{u}|^{\frac{2s}{2-s}}dx\Big)^{\frac{2-s}{s}}dt\Big)^{\frac{1}{2}}.$$

Summarizing our efforts, we have

$$\frac{1}{2}\int\limits_{B} |\widehat{u}(x,t_0|^2\varphi^2_{x_0,2R}(x)dx + \int\limits_{-1}^{t_0}\int\limits_{B} \chi^2_{t_0,2R}\varphi^2_{x_0,2R}|\nabla\widehat{u}|^2dz \leq$$

$$\leq c(s)(1+\Gamma^2)R^{(\frac{n}{s'}-1)2}\int\limits_{t_0-(2R)^2}^{t_0}\Big(\int\limits_{B(x_0,2R)}|\widehat{u}|^{\frac{2s}{2-s}}dx\Big)^{\frac{2-s}{s}}dt, \tag{2.10}$$

where $\Gamma=\|d\|_{L_\infty(BMO)}$. Now, let us discuss simple consequences of (2.10) following [12]. By Poincare-Sobolev inequality, we have for

$$s\leq\frac{n}{n-1} \tag{2.11}$$

the following inequality

$$\Big(\int\limits_{B(x_0,2R)}|\widehat{u}|^{\frac{2s}{2-s}}dx\Big)^{\frac{2-s}{s}}\leq c(s)R^{n\frac{2-s}{s}+2-n}\int\limits_{B(x_0,2R)}|\nabla u|^2dx.$$

Combining (2.11) and (2.10), we find

$$\int\limits_{B}|\widehat{u}(x,t_0)|^2\varphi^2_{x_0,2R}(x)dx\leq c(s)(1+\Gamma^2)\int\limits_{Q(z_0,2R)}|\nabla u|^2dx.$$

Hence, assuming that $Q(z_0,3R)\subset Q$, we have the second estimate

$$\operatorname*{ess\,sup}_{t_0-R^2<t<t_0}\int\limits_{B(x_0,R)}|\widehat{u}|^2(x,t)dx\leq c(s)(1+\Gamma^2)\int\limits_{Q(z_0,3R)}|\nabla u|^2dx. \tag{2.12}$$

Now, our aim is going to be the so-called reverse Hölder inequality. We first assume that the number s satisfies the condition

$$1<s<\frac{2n}{2n-1},\qquad n=2,3,... \tag{2.13}$$

Obviously, (2.13) implies (2.11) and

$$\frac{2n}{2n-1}\leq\frac{4}{3}\leq\frac{4n}{3n-2}\leq 2,\qquad n=2,3,... \tag{2.14}$$

It is not difficult to show that under assumption (2.13) there exist numbers $0<\lambda<1$, $0<\mu<1$, and $1<r<2$ such that

$$\frac{2s}{2-s}=2\lambda+\frac{nr}{n-r}\mu$$
$$\lambda+\mu=1$$
$$\frac{nr}{n-r}\mu\frac{2-s}{s}=1.$$

Using these exponents, we derive from (2.10)

$$\int\limits_{Q(z_0,R)}|\nabla u|^2dz\leq c(s)(1+\Gamma^2)R^{(\frac{n}{s'}-1)2}\int\limits_{t_0-(2R)^2}^{t_0}\Big(\int\limits_{B(x_0,2R)}|\widehat{u}|^{2\lambda+\frac{nr}{n-r}\mu}dx\Big)^{\frac{2-s}{s}}dt\leq$$

$$\leq c(s)(1+\Gamma^2)R^{(\frac{n}{s'}-1)2}\int\limits_{t_0-(2R)^2}^{t_0}\Big(\int\limits_{B(x_0,2R)}|\widehat{u}|^2dx\Big)^{\frac{2-s}{s}\lambda}\Big(\int\limits_{B(x_0,2R)}|\widehat{u}|^{\frac{rn}{n-r}}dx\Big)^{\frac{2-s}{s}\mu}dt.$$

The last multiplier can be estimated with the help of Sobolev's inequality

$$\int\limits_{Q(z_0,R)}|\nabla u|^2dz\leq c(s)(1+\Gamma^2)R^{(\frac{n}{s'}-1)2}\mathrm{ess}\sup_{t_0-(2R)^2<t<t_0}\Big(\int\limits_{B(x_0,2R)}|\widehat{u}(x,t)|^2dx\Big)^{\frac12}\times$$

$$\times R^{\frac{2(r-1)}{r}}\Big(\int\limits_{Q(z_0,2R)}|\nabla u|^rdz\Big)^{\frac1r}.$$

To estimate the first multiplier on the right hand side of the last inequality, one can apply (2.12) in the following way

$$\int\limits_{B(x_0,2R)}|u(x,t)-u_{x_0,2R}(t)|^2dx\leq c\int\limits_{B(x_0,2R)}|u(x,t)-u_{x_0,4R}(t)|^2dx$$

$$\leq c(s)(1+\Gamma^2)\int\limits_{Q(z_0,6R)}|\nabla u|^2\,dz$$

for a.a. $t\in]t_0-(2R)^2,t_0[$. Combining the latter inequality, we arrive at the reverse Hölder inequality

$$\frac{1}{|Q(R)|}\int\limits_{Q(z_0,R)}|\nabla u|^2dz\leq c(s)(1+\Gamma^2)^2\Big(\frac{1}{|Q(6R)|}\int\limits_{Q(z_0,6R)}|\nabla u|^2dz\Big)^{\frac12}\times$$

$$\times\Big(\frac{1}{|Q(2R)|}\int\limits_{Q(z_0,2R)}|\nabla u|^rdz\Big)^{\frac1r}$$

which holds for some $r\in]1,2[$ and for any $Q(z_0,6R)\subset Q$. This leads to a higher integrability, see [12]. □

2.2. **Moser iteration.** To avoid some technical difficulties, we will assume that matrices a and b and solution u are sufficiently smooth in Q. Later we shall show how to remove this assumption. We also assume that our function u is strictly positive in the following sense

$$u(z)\geq\alpha_R>0\qquad\forall z\in\overline{Q}(R)\tag{2.15}$$

for any $0<R<1$. Sometimes assumption (2.15) is not necessary, but for simplicity we will assume it is satisfied. We fix the following notation

$$\varepsilon^2(m)=\Big|\frac{1}{2m}-1\Big|,\qquad p=\frac{2(n+2)}{n}$$

and, assuming that condition (2.13) holds, let

$$q=\frac{2s}{2-s},\qquad\gamma=\frac{p}{q}>1.$$

Lemma 2.4. *For any $m_1 \geq m_0 > 1/2$ and for any $0 < \varrho < r$ with $Q(z_0, r) \subset Q$, we have*

$$\sup_{z \in Q(z_0,\varrho)} u^{m_1}(z) \leq \frac{c_1(n, \nu, s, \Gamma, \varepsilon_0)}{(r-\varrho)^{\frac{n+2)}{q}}} \Big(\int\limits_{Q(z_0,r)} u^{m_1 q}(z) dz \Big)^{\frac{1}{q}}, \tag{2.16}$$

where $\varepsilon_0 = \varepsilon(m_0)$.

Proof. Set $w = u^m$. For any $m \neq 0$, we can derive from (1.12)

$$\frac{1}{2} \int\limits_{B(x_0,r)} \psi^2 \partial_t |w|^2 dx + \frac{2m-1}{m} \int\limits_{B(x_0,r)} \psi^2 a \nabla w \cdot \nabla w dx =$$

$$= - \Big(\int\limits_{B(x_0,r)} a \nabla w \cdot w \nabla \psi^2 dx + \int\limits_{B(x_0,r)} d \nabla w \cdot w \nabla \psi^2 dx \Big), \tag{2.17}$$

with a cut-off function ψ satisfying:

$$\psi(x,t) = \varphi(x)\chi(t),$$

$$\varphi(x) = 1 \quad x \in B(x_0, \varrho), \qquad \varphi(x) = 0 \quad x \notin B(x_0, r),$$

$$0 \leq \varphi \leq 1, \qquad |\nabla \varphi| \leq \frac{c}{r - \varrho},$$

$$\chi(t) = 0 \quad t < t_0 - r^2, \qquad \chi(t) = 1 \quad t > t_0 - \varrho^2,$$

$$\chi(t) = \frac{t - (t_0 - r^2)}{r^2 - \varrho^2} \quad t_0 - r^2 \leq t \leq t_0 - \varrho^2.$$

Next, we introduce the following sequence of exponents

$$l_0 = q, \qquad l_i = \gamma^i l_0, \qquad i = 0, 1, ..., \tag{2.18}$$

If we let

$$m_i = l_i m_1 / p, \qquad i = 1, 2, ...,$$

then we have

$$m_i q = l_{i-1} m_1, \qquad \varepsilon^2(m_i) = \frac{1}{2m_i} - 1 > \varepsilon_0^2, \qquad i = 1, 2, ... \tag{2.19}$$

Letting $m = m_i$ in (2.17) and taking into account (2.19), we find

$$\sup_{t_0 - \varrho^2 < t < t_0} \int\limits_{B(x_0,\varrho)} |w(x,t)|^2 dx + \varepsilon_0^2 \nu \int\limits_{Q(z_0,r)} \psi^2 |\nabla w|^2 dz \leq$$

$$\leq \frac{c}{(r-\varrho)^2} \int\limits_{Q(z_0,r)} \psi^2 |w|^2 dz + c\nu^{-1} \int\limits_{Q(z_0,r)} \psi |\nabla \psi| w |\nabla w| dz \tag{2.20}$$

$$+ c \int\limits_{Q(z_0,r)} |d - [d]_{x_0,r}| \psi |\nabla \psi| w |\nabla w| dz.$$

The same arguments as in Section 2 show that the latter inequality gives us:

$$|w|^2_{2,Q(z_0,\varrho)} \equiv \sup_{t_0-\varrho^2<t<t_0} \int\limits_{B(x_0,\rho)} |w(x,t)|^2dx + \int\limits_{Q(z_0,\varrho)} |\nabla w|^2dz \leq$$

$$\leq \frac{c(s,\varepsilon_0)}{(r-\varrho)^2}(1+\Gamma^2)r^{\frac{2(n+2)}{s'}}\Big(\int\limits_{Q(x_0,r)} |w|^qdz\Big)^{\frac{2}{q}}$$

with s satisfying condition (2.13). By the known embedding theorem, see [18], we have $\|w\|_{p,Q(z_0,\varrho)} \leq c|w|_{2,Q(z_0,\varrho)}$ with $p=\frac{2(n+2)}{n}$ and, hence,

$$\Big(\frac{1}{|Q(\varrho)|}\int\limits_{Q(z_0,\varrho)} |w|^pdz\Big)^{\frac{1}{p}} \leq c(s,\varepsilon_0)(1+\Gamma)\Big(\frac{r}{r-\varrho}\Big)\times$$

$$\times\Big(\frac{r}{\varrho}\Big)^{\frac{n}{2}}\Big(\frac{1}{|Q(r)|}\int\limits_{Q(z_0,r)} |w|^qdz\Big)^{\frac{1}{q}}. \tag{2.21}$$

It is worth noting that, under assumption (2.13) we have $p>q$.

Our further steps are routine. We let

$$\varrho = R_i = \frac{R}{2}+\frac{R}{2^{i+1}}, \qquad r=R_{i-1}, \qquad i=1,2,...,$$

in (2.21) and find

$$\Big(\frac{1}{|Q(R_i)|}\int\limits_{Q(z_0,R_i)} |u|^{m_1l_i}dz\Big)^{\frac{1}{l_i}} \leq$$

$$\leq (c(s,\varepsilon_0,\Gamma)2^i)^{\frac{1}{\gamma^{i-1}}}\Big(\frac{1}{|Q(R_{i-1})|}\int\limits_{Q(z_0,R_{i-1})} |u|^{m_1l_{i-1}}dz\Big)^{\frac{1}{l_{i-1}}}$$

for $i=1,2,....$ After iterations, we arrive at (2.17) with $\varrho=R/2$ and $r=R$. General case is deduced from this particular one with help of known arguments. $\square$

To see what happens if $0<m<1/2$, we have to introduce additional notation

$$Q^+(z_0,R)=B(x_0,R)\times]t_0,t_0+R^2[, \qquad Q^+(R)=Q^+(0,R),$$

$$\widetilde{Q}(z_0,R)=B(x_0,R)\times]t_0-R^2,t_0+R^2[, \qquad \widetilde{Q}(R)=\widetilde{Q}(0,R).$$

Lemma 2.5. *For any $0<m_1<1/2$ and for any $0<\varrho<r$ provided $\widetilde{Q}(z_0,r)\subset Q$, we have*

$$\sup_{z\in\widetilde{Q}(z_0,\varrho)} u^{m_1}(z) \leq \frac{c_2(n,\nu,s,\Gamma)}{(r-\varrho)^{\frac{n+2}{q}}}\Big(\int\limits_{\widetilde{Q}(z_0,r)} u^{m_1q}(z)dz\Big)^{\frac{1}{q}}. \tag{2.22}$$

Proof. We replace the function χ with the following one

$$\chi(t)=0 \quad t>t_0'+r^2, \qquad \chi(t)=1 \quad t<t_0'+\varrho^2,$$

$$\chi(t)=\frac{-t+(t_0'+r^2)}{r^2-\varrho^2} \quad t_0'+\varrho^2\leq t\leq t_0'+r^2, \qquad t_0'=t_0-\Big(\frac{3}{4}\Big)^2r^2.$$

Then from (2.17), we can derive (an analog of (2.20))

$$\sup_{t_0<t<t_0+\varrho^2}\int\limits_{B(x_0,\varrho)}|w(x,t)|^2dx+\varepsilon^2(m)\nu\int\limits_{Q^+(z_0',r)}\psi^2|\nabla w|^2dz\leq$$

$$\leq\frac{c}{(r-\varrho)^2}\int\limits_{Q^+(z_0',r)}\psi^2|w|^2dz+c\nu^{-1}\int\limits_{Q^+(z_0',r)}\psi|\nabla\psi|w|\nabla w|dz \tag{2.23}$$

$$+c\int\limits_{Q^+(z_0',r)}|d-[d]_{x_0,r}|\psi|\nabla\psi|\omega|\nabla\omega|dz, \qquad z_0'=(x_0,t_0').$$

Next, it is not so difficult to check that there exists a natural number k with the following property

$$\frac{1}{\gamma^2}m_1\leq m_1'=\gamma^{-(k-\frac12)}\frac12\leq m_1.$$

And then, for this number k, we have

$$m_1'<m_2'<...<m_k'<\frac12<m_{k+1}'<...,$$

where

$$m_i'=\frac{l_im_1'}{p}=\gamma^{i-1}m_1'=\gamma^{i-k-\frac12}\frac12, \qquad i=1,2,....$$

and numbers l_i is defined by (2.18). It is easy to check that

$$\varepsilon^2(m_i')\geq\gamma^{\frac12}-1, \qquad i=1,2,...,k,$$

and then repeating derivation of (2.21) for $w=u^{m_i'}$ with the same indices i, we find

$$\Big(\frac{1}{|Q(\varrho)|}\int\limits_{Q(z_0',\varrho)}|w|^pdz\Big)^{\frac1p}\leq c(s,\Gamma)\Big(\frac{r}{r-\varrho}\Big)\times$$

$$\times\Big(\frac{r}{\varrho}\Big)^{\frac n2}\Big(\frac{1}{|Q(r)|}\int\limits_{Q(z_0',r)}|w|^qdz\Big)^{\frac1q}. \tag{2.24}$$

Now, we consider (2.24) for

$$r=r_i, \qquad \varrho=r_{i-1}, \qquad r_i=\frac r4+\frac14\frac{r}{2^i}$$

and find

$$\Big(\frac{1}{|Q(r_i)|}\int\limits_{Q^+(z'_0,r_i)}|u|^{m'_1l_i}dz\Big)^{\frac{1}{l_i}}\leq$$

$$\leq(c(s,\Gamma)2^i)^{\frac{1}{\gamma^{i-1}}}\Big(\frac{1}{|Q(r_{i-1})|}\int\limits_{Q^+(z'_0,r_{i-1})}|u|^{m'_1l_{i-1}}dz\Big)^{\frac{1}{l_k}}$$

for $i=1,2,...,k$. After exactly k iterations, we have

$$\Big(\frac{1}{|Q(3r/4)|}\int\limits_{Q^+(z'_0,3r/4)}|u|^{m'_{k+1}q}dz\Big)^{\frac{1}{l_k}}=\Big(\frac{1}{|Q(3r/4)|}\int\limits_{Q(z_0,3r/4)}|u|^{m'_{k+1}q}dz\Big)^{\frac{1}{l_k}}\leq$$

$$\leq c(s,\Gamma)\Big(\frac{1}{|Q(r)|}\int\limits_{Q^+(z'_0,r)}|u|^{m'_1q}dz\Big)^{\frac{1}{q}}.$$

Since $m'_{k+1}>1/2$, we are in a position to apply Lemma 2.4 letting there $m_1=m_0=m'_{k+1}$ and conclude that

$$\sup_{z\in Q(z_0,r/2)}u^{m'_{k+1}}(z)\leq c(s,\Gamma)\Big(\frac{1}{r^{n+2}}\int\limits_{Q(z_0,3r/4)}u^{m'_{k+1}q}dz\Big)^{\frac{1}{q}}.$$

Taking into account definition (2.18) of l_k and combining the latter inequalities, we find

$$\sup_{z\in Q(z_0,r/2)}u^{m_1}(z)\leq\Big[c(s,\Gamma)\Big]^{(1+\gamma^{-k})\frac{m_1}{m'_1}}\Big(\frac{1}{r^{n+2}}\int\limits_{Q^+(z'_0,r)}|u|^{m'_1q}dz\Big)^{\frac{1}{q}\frac{m_1}{m'_1}}$$

and, by Hölder inequality, we have

$$\sup_{z\in Q(z_0,r/2)}u^{m_1}(z)\leq c(s,\Gamma)\Big(\frac{1}{r^{n+2}}\int\limits_{Q^+(z'_0,r)}|u|^{m'_1q}dz\Big)^{\frac{1}{q}}. \tag{2.25}$$

We may shift in time this estimate and show that

$$\sup_{z\in Q^+(z_0,r/2)}u^{m_1}(z)\leq c(s,\Gamma)\Big(\frac{1}{r^{n+2}}\int\limits_{Q^+(z''_0,r)}|u|^{m'_1q}dz\Big)^{\frac{1}{q}}, \tag{2.26}$$

where $z''_0=(x_0,t''_0)$ and $t''_0=t_0-\frac{5r^2}{16}$. From (2.25) and (2.26), estimate (2.22) with $\varrho=r/2$ follows. General case is deduced from this particular one with help of known arguments. □

Lemma 2.6. *For any $\varepsilon>0$ and for any $0<\varrho<r$ such that $Q(z_0,r)\subset Q$, we have*

$$\sup_{z\in Q(z_0,\varrho)}u^{-\varepsilon}(z)\leq\frac{c_3(n,\nu,s,\Gamma)}{(r-\varrho)^{\frac{n+2)}{q}}}\Big(\int\limits_{Q(z_0,r)}u^{-\varepsilon q}(z)dz\Big)^{\frac{1}{q}}. \tag{2.27}$$

Proof. We let $v = u^{-\varepsilon}$ and observe that by (1.12), the function v satisfies

$$\partial_t v - \mathrm{div} A\nabla v < 0.$$

We can repeat the proof of Lemma 2.4 with $m_1 = 1$ for v instead of u and then show (2.27). □

2.3. **Estimates of** $\ln u$.

Lemma 2.7. *Assume that u i a sufficiently smooth positive solution to equation (1.12) and $Q'(z_0,R) = B(x_0,2R)\times]t_0-R^2,t_0+R^2[\subset Q$. There exist two constants $c_4 = c_4(n,\nu)$ and a^R such that*

$$|\{z\in Q^+(z_0,R):\ -\ln u - a^R > s\}| \le \frac{c_4R^{n+2}}{s}, \tag{2.28}$$

$$|\{z\in Q(z_0,R):\ -\ln u - a^R < -s\}| \le \frac{c_4R^{n+2}}{s}. \tag{2.29}$$

Proof. To simplify notation, we shift and scale our variables in the following way

$$u^R(y,s) = u(x_0+Ry,t_0+R^2s), \qquad A^R(y,s) = A(x_0+Ry,t_0+R^2s)$$

for $(y,s)\in Q' = B(2)\times]-1,1[$. Since equation (1.12) is invariant with respect to this transformation, we may reduce our considerations to the cylinder Q' and, after proving our result for this particular case, get all the statements of the lemma with the help of inverse translation and dilatation. Without ambiguity, in what follows, we drop upper index R in the notation of functions u^R and A^R.

So, if we let $v = \ln u$, then by (1.12)

$$\partial_t v - \mathrm{div}(A\nabla v) + \nabla v\cdot a\nabla v = 0 \tag{2.30}$$

in Q'. Take and fix a smooth nonnegative cut-off function $\psi = \psi(x)$ so that $\psi = 1$ in B and $\psi = 0$ outside $B(2)$. Multiplying equations (2.30) by ψ^2 and integrating the product in x over $B(2)$ and in t over the interval $]t_1,t_2[$, we find

$$\int\limits_{B(2)} v\psi^2dx\Big|_{t_1}^{t_2} + \int\limits_{t_1}^{t_2}\int\limits_{B(2)} \nabla\psi^2\cdot A\nabla v dx\,dt + \int\limits_{t_1}^{t_2}\int\limits_{B(2)} \psi^2\nabla v\cdot a\nabla v dx\,dt = 0$$

and thus

$$\int\limits_{B(2)} v\psi^2dx\Big|_{t_1}^{t_2} + \int\limits_{t_1}^{t_2}\int\limits_{B(2)} \psi^2\nabla v\cdot a\nabla v dx\,dt \le$$

$$\le c\int\limits_{t_1}^{t_2}\int\limits_{B(2)} \psi|\nabla\psi||\nabla v|(|a| + |d-[d]_{0,2}|)dx\,dt.$$

After application of the Cauchy-Schwartz inequality, we have the following estimate

$$\int\limits_{B(2)} v\psi^2 dx\Big|_{t_1}^{t_2} + \frac{\nu}{2}\int\limits_{t_1}^{t_2}\int\limits_{B(2)} \psi^2|\nabla v|^2 dx\,dt \le c(\Gamma)(t_2-t_1). \tag{2.31}$$

From this point we essentially repeat arguments of J. Moser in [25], see Lemma 3 therein. We do this just for completeness. As it is pointed out in [25], we can choose our cut-off function ψ so that the following Poincarè-type inequality takes place

$$\int\limits_{B(2)} |v(x,t)-V(t)|^2\psi^2(x)dx \le c\int\limits_{B(2)} |\nabla v(x,t)|^2\psi^2(x)dx,$$

where

$$V(t) = \int\limits_{B(2)} v(x,t)\psi^2(x)dx\Big(\int\limits_{B(2)} \psi^2(x)dx\Big)^{-1}.$$

Making use of this inequality, we can derive from (2.31) the following relation

$$V(t_2)-V(t_1)+c_4^{-1}\int\limits_{t_1}^{t_2}\int\limits_{B} |v(x,t)-V(t)|^2 dx\,dt \le c_5(n,\nu,\Gamma)(t_2-t_1)$$

which can be reduced to the differential form

$$\frac{dV}{dt}(t)+c_4^{-1}\int\limits_{B}|v(x,t)-V(t)|^2dx \le c_5.$$

One may make this inequality homogeneous with help of the shift

$$w(x,t)=v(x,t)-V(0)-c_5t, \qquad W(t)=V(t)-V(0)-c_5t.$$

This give us the inequality

$$\frac{dW}{dt}(t)+c_4^{-1}\int\limits_{B}|w(x,t)-W(t)|^2dx \le 0 \tag{2.32}$$

and the initial condition

$$W(0)=0. \tag{2.33}$$

For $0<t<1$ and $s>0$, we introduce the family of sets

$$B_s^+(t)=\{x\in B: \ \ w(x,t)>s\}.$$

As it follows from (2.32) and (2.33), for those values of parameters t and s, we have $w(\cdot,t)-W(t)\ge s-W(t)>0$ on $B_s^+(t)$ and, hence,

$$\frac{dW}{dt}(t)+c_4^{-1}|B_s^+(t)|(s-W)^2\le 0$$

or

$$c_4(s-W)^{-2}\frac{d(s-W)}{dt}\ge |B_s^+(t)|$$

The latter identity can be integrated and, as a result, we find

$$\int_0^1 |B_s^+(t)|dt = \{(x,t) \in Q^+ : \ \ w(x,t) > s\} \leq \frac{c_4}{s}$$

which implies the first estimate (2.28). Other estimate (2.29) can be established in the same way. □

2.4. **Harnack inequality.**

Theorem 2.8. *Let u be a positive sufficiently smooth solution to (1.12) in Q. Then for any $Q(z_0,R) \subset Q$ we have the inequality*

$$\sup_{z\in Q(z_{0R},R/2)} u(z) \leq c_6(n,\nu,s,\Gamma) \inf_{z\in Q(z_0,R/2)} u(z), \tag{2.34}$$

where $z_{0R} = z_0 - (0, R^2/2)$.

Proof. Using translation and dilatation similar to those described in Section 4, one may consider our problem in a canonical domain, say, in Q'. Now our aim is to make use of estimates proved in Sections 3 and 4 plus some iteration technique in order to find a particular version of the Harnack inequality. It can be extended to the general case of Theorem 2.8 with the help of covering methods. So, in this section, we follow [27] with minor changes.

By Lemma 2.7, see (2.29), we know that

$$|\{z \in Q : \ \ -\ln u - a < -s\}| \leq \frac{c_4}{s} \tag{2.35}$$

for some constant a. As in [27], we introduce the following function

$$\varphi(r) = \sup_{z\in \widetilde{Q}(z_0,r)} \ln w(z)$$

with $z_0 = (0, 1/2)$, $r \in]1/2, 1/\sqrt{2}[$, and $w = e^a u$.

Now, our aim is to show that there exist a constant $\delta > 2$ depending only on n, ν, s, and $\vartheta \in [1/2, 1/\sqrt{2}[$ such that

$$\varphi(\vartheta) \leq \delta. \tag{2.36}$$

To this end, we derive from (2.35) the estimate

$$\int\limits_{\widetilde{Q}(z_0,r)} w^p dz \leq e^{p\varphi(r)} \frac{2c_4}{\varphi(r)} + e^{\frac{p}{2}\varphi(r)} \tag{2.37}$$

being true for any positive $m_1 = p/q$ and for any $r \geq \vartheta$. Let us choose m_1 so that terms on the right-hand side of (2.37) contribute to the sum equally. This suggests the following value for m_1

$$m_1 = \frac{1}{q\varphi(r)} \ln \frac{\varphi(r)}{2c_4}. \tag{2.38}$$

Obviously, there exists a constant $\delta_1 > 2$ depending only on s and c_4 such that if

$$\varphi(\vartheta) > \delta_1, \tag{2.39}$$

then m_1 defined above belongs to the interval $]0,1/2[$. If not, then $\delta=\delta_1$. If (2.39) holds, estimate (2.22) of Lemma 2.5 give us the relation

$$m_1\varphi(\varrho)\leq \ln\frac{c_2}{(r-\varrho)^{\frac{n+2}{q}}}+\frac{1}{q}\ln\Big(\int\limits_{\widetilde{Q}(z_0,r)} w^p dz\Big)$$

for any $\vartheta\leq\varrho<r$. Recalling the choice of m_1, we find from (2.37) that

$$\int\limits_{\widetilde{Q}(z_0,r)} w^p dz\leq 2e^{\frac{p}{2}\varphi(r)}$$

and thus

$$\varphi(\varrho)\leq\frac{1}{m_1}\ln\frac{c_2}{(r-\varrho)^{\frac{n+2}{q}}}+\frac{1}{2}\varphi(r).$$

The latter inequality can be rewritten with the help of (2.38) in the following way

$$\varphi(\varrho)\leq\frac{1}{2}\varphi(r)\Big[\frac{\ln(c_2(r-\varrho)^{-\frac{n+2}{q}})}{m_1\varphi(r)}+1\Big]=\frac{1}{2}\varphi(r)\Big[\frac{\ln c_2^q(r-\varrho)^{-(n+2)}}{\ln\frac{\varphi(r)}{2c_4}}+1\Big].$$

Then one can consider two cases. In the first case,

$$\frac{\ln c_2^q(r-\varrho)^{-(n+2)}}{\ln\frac{\varphi(r)}{2c_4}}\leq\frac{1}{2}$$

and thus

$$\varphi(\varrho)\leq\frac{3}{4}\varphi(r).$$

In the opposite case, we have

$$\varphi(\varrho)\leq\varphi(r)\leq\frac{\mu_1(n,\nu,s,\Gamma)}{(r-\varrho)^{2(n+2)}}.$$

Combining both cases, we find the following basic inequality

$$\varphi(\varrho)\leq\frac{3}{4}\varphi(r)+\frac{\mu_1}{(r-\varrho)^{2(n+2)}}$$

for any $1/2\leq\vartheta\leq\varrho<r\leq 1/\sqrt{2}$. It can be iterated in the known way, see [27], and the result of these iterations can be expresses in the form

$$\varphi(\vartheta)\leq\delta_2(n,\nu,s,\Gamma).$$

So, (2.36) is proved with $\delta=\max\{\delta_1,\delta_2\}$.

Next, let $z_*=(0,1)$ and $v=u^{-1}$. Then as it follows from Lemmata 2.6 and 2.7, see (2.27) and (2.28),

$$\sup_{z\in Q(z_0,\varrho)} v(z)\leq\frac{c_3}{(r-\varrho)^{\frac{n+2)}{q}}}\Big(\int\limits_{Q(z_0,r)} v^q(z)dz\Big)^{\frac{1}{q}}$$

and

$$|\{z\in Q(z_*,1):\ \ \ln(e^{-a}v)>s\}|\leq\frac{c_4}{s}.$$

The same arguments as above show that there exists a constant $\delta_3(n,\nu,s,\Gamma)$ such that

$$\sup_{z\in Q(z_*,\vartheta)}\ln(e^{-a}u^{-1}(z))\leq\delta_3.$$

This estimate, together with (2.36), implies a particular version of the Harnack inequality

$$\sup_{z\in\widetilde{Q}(z_0,\vartheta)}u(z)\leq e^{\delta\delta_3}\inf_{z\in Q(z_*,\vartheta)}u(z).$$

The general case can be obtained from the particular case with the help of covering technique, see [27], Lemma 4, and translation and dilatation. □

2.5. **Nonsmooth case.**

Proof of Theorem 1.1. Without loss of generality, we may assume that

$$\nabla u\in L_{p_0}(Q) \tag{2.40}$$

for some $p_0>2$. To provide (2.40), we can apply Theorem 2.2 and scaling. Obviously, one can construct smooth approximations of matrices a and d with the following properties:

$$a^{(j)}\to a\qquad\text{in}\quad L_p(Q)$$

$$d^{(j)}\to d\qquad\text{in}\quad L_p(Q)$$

for any $p>1$ and

$$a^{(j)}\to a\qquad\text{a.e. in}\quad Q$$

$$d^{(j)}\to d\qquad\text{a.e. in}\quad Q.$$

Moreover, matrices $A^{(j)}=a^{(j)}+d^{(j)}$, $a^{(j)}$ and $d^{(j)}$ satisfies conditions (1.15)–(1.17) with the same constants and $\|d^{(j)}\|_{L_\infty(BMO)}\leq\|d\|_{L_\infty(BMO)}$.

Then we consider the following initial boundary value problem

$$\partial_t w^{(j)}-\operatorname{div}(A^{(j)}\nabla w^{(j)})=f^{(j)},$$

$$w^{(j)}|_{\partial'Q}=0,$$

where $\partial'Q$ is a parabolic boundary of Q and

$$f^{(j)}\equiv\partial_t u-\operatorname{div}(A^{(j)}\nabla u).$$

Our claim is

$$f^{(j)}\to 0\qquad\text{in}\quad L_2(-1,0;H^{-1}),$$

where H^1 is the completion of smooth compactly supported in B functions with respect to the norm $\|u\|_{2,B}+\|\nabla u\|_{2,B}$. Indeed, it is not difficult to show that

$$\|f^{(j)}\|_{L_2(-1,0;H^{-1})}\leq\Big(\int\limits_Q(|a-a^{(j)}|^2+|d-d^{(j)}|^2)|\nabla u|^2dz\Big)^{\frac12}.$$

So, by (2.40), the right hand of the latter inequality goes to zero. On the other hand, for $w^{(j)}$, we have global energy estimate

$$|w^{(j)}|_{2,Q} \leq c\|f^{(j)}\|_{L_2(-1,0;H^{-1})}$$

which, in turn, means that

$$|w^{(j)}|_{2,Q} \to 0. \tag{2.41}$$

Now, we let $v^{(j)} = u - w^{(j)}$. Obviously, $v^{(j)}$ is a unique solution to the following initial boundary value problem

$$\partial_t v^{(j)} - \operatorname{div}(A^{(j)}\nabla v^{(j)}) = 0,$$

$$(v^{(j)} - u)|_{\partial' Q} = 0.$$

We know that $v^{(j)}$ possesses the following global properties

$$\partial_t v^{(j)} \in L_2(-1,0;H^{-1}), \qquad \nabla v^{(j)} \in L_2(Q)$$

and, moreover, it is nonnegative on the parabolic boundary of Q and smooth inside Q where the equation for $v^{(j)}$ can be reduced to the form

$$\partial_t v^{(j)} - (a^{(j)}_{kl} v^{(j)}_{,l})_{,k} - d^{(j)}_{kl,k} v^{(j)}_{,l} = 0.$$

Here, comma in lower indices stands for the differentiation with respect to the corresponding spatial variable and summation over repeated indices running from 1 to n is adopted. As it was shown in [18], see Chapter 3, Theorem 7.2, therein, for functions satisfying equation above, the maximum principle holds and thus $v^{(j)}$ remains to be nonnegative everywhere inside Q. Obviously function $v^{(j)} + \frac{1}{j}$ satisfies all the conditions of Theorem 2.8 and, hence,

$$\sup_{z\in Q(z_{0R},R/2)} v^{(j)}(z) \leq c_6 \inf_{Q(z_0,R/2)} v^{(j)}(z).$$

Passing to the limit as $j \to \infty$ and taking into account (2.41), we arrive at (1.18). □

2.6. **Liouville Theorem.** In this subsection, we assume that u is an ancient suitable weak solution to equation (1.12) which means that it is defined on $Q_- \equiv \mathbb{R}^n \times]-\infty, 0[$ and is a suitable weak solution in all parabolic balls $Q(z_0,1)$ with $z_0 = (x_0,t_0)$ for any $x_0 \in \mathbb{R}^n$ and for any $t_0 \leq 0$. Since our equation is invariant with respect to translation and usual parabolic dilatation, such a solutions will be suitable in all parabolic balls of the form $Q(a)$ for any positive a. Now, we shall show its Hölder continuity provided it is bounded.

Lemma 2.9. *Let u be an ancient suitable weak solution to equation (1.12). Then there are two constants c_7 and α which depend only on n, ν, s, satisfying condition (2.13), and $\Gamma = \|d\|_{L_\infty(BMO)}$ such that*

$$|u(z) - u(z_0)| \leq c_7 |z - z_0|^\alpha_{par} \sup_{z\in Q_-} |u(z)| \tag{2.42}$$

for any z and z_0 from Q_- with the parabolic distance $|z - z_0|_{par} = |x - x_0| + |t - t_0|^{\frac{1}{2}}$.

Proof. We let

$$M_R=\sup_{z\in Q(z_0,R)}u(z),\qquad M_{R/2}=\sup_{z\in Q(z_0,R/2)}u(z),$$
$$m_R=\inf_{z\in Q(z_0,R)}u(z),\qquad m_{R/2}=\inf_{z\in Q(z_0,R/2)}u(z).$$

If we let $v(z)=M_R-u(z)$, then will be a nonnegative suitable weak solution to equation (1.12) in parabolic ball $Q(z_0,R)$. By translation and parabolic dilatation, we can derive from Theorem 2.8 the following inequality for v

$$\inf_{z\in Q(z_0,R/2)}v(z)=M_R-M_{R/2}\geq\frac{1}{c_6}(M_R-u(z))\tag{2.43}$$

for all $z\in Q(z_{0R},R/2)$. On the other hand, for the same reason, we may apply Theorem 2.8 to function $w=u(z)-m_R$ and find

$$m_{R/2}-m_R\geq\frac{1}{c_6}(u(z)-m_R)\tag{2.44}$$

for all $z\in Q(z_{0R},R/2)$. Adding (2.43) and (2.44), we arrive at the inequality

$$\mathrm{osc}(z_0,R)-\mathrm{osc}(z_0,R/2)\geq\frac{1}{c_6}\mathrm{osc}(z_0,R),\tag{2.45}$$

where $\mathrm{osc}(z_0,R)=M_R-m_R$, or

$$\mathrm{osc}(z_0,R/2)\leq\vartheta\,\mathrm{osc}(z_0,R)$$

with $\vartheta=1-1/c_6$. After simple iterations, we have the series of the inequalities

$$\mathrm{osc}(z_0,R/2^k)\leq\vartheta^k\mathrm{osc}(z_0,R)$$

which can be reduced to the form

$$\mathrm{osc}(z_0,\varrho)\leq c_7\varrho^\alpha\vartheta\,\mathrm{osc}(z_0,R).$$

The latter is true for $z_0\in Q_-$ and all $0<\varrho<R<+\infty$ and certainly implies (2.42). □

Proof of Theorem 1.2. We let $M=\sup_{z\in Q_-}|u(z)|$. If we scaled our solution u and matrix A so that

$$u^R(y,s)=u(Ry,R^2s),\qquad A^R(y,s)=A(Ry,R^2y),$$

then as it is easy to see u^R and A^R satisfy the equation (1.12) in Q_- and

$$\nu=\nu^R,\qquad M=M^R=\sup_{z\in Q_-}|u^R(z)|,\qquad \Gamma=\Gamma^R=\|d^R\|_{L_\infty(BMO)}.$$

By Lemma 2.9, we have

$$|u^R(e)-u^R(0)|\leq c_7|e|^\alpha_{par}M^R$$

for any $e=(y,s)\in Q_-$. Making inverse scaling in the latter inequality, we find

$$|u(z)-u(0)|\leq c_7|z|^\alpha_{par}\frac{1}{R^\alpha}M$$

for any $z\in Q_-$ and for any $R>0$. By arbitrariness of R, we show that u must be a constant. □

3. An elementary proof of an elliptic Liouville theorem in 2D and a counterexample

In this section we explore the Liouville theorem for (1.4) in two dimensions. Assuming that the divergence-free vector field b is in the space $(BMO)^{-1}$, we provide an elementary, short, and self contained proof showing that bounded subsolutions (and supersolutions) are constant. Afterwards, we construct a counterexample to such a Liouville theorem for a divergence-free vector field whose stream function is bounded by $\ln|x|\ln\ln|x|$ ($\notin BMO$) for large $|x|$. This construction shows that the hypothesis $b \in (BMO)^{-1}$ is quite sharp.

If b is a smooth divergence-free vector field on $\mathbb{R}^2$, then it has a *stream function* $H : \mathbb{R}^2 \to \mathbb{R}$ as in (1.23) and we have (1.24). This relationship between b and A allows us to introduce the notion of a weak solution for very singular drifts.

Definition 3.1. Let b be a divergence-free drift from $BMO^{-1}(\mathbb{R}^2)$, that is, $H \in BMO(\mathbb{R}^2)$. We say that a function $u \in H^1_{\mathrm{loc}}(\mathbb{R}^2)$ is a weak subsolution to (1.4) in $\mathbb{R}^2$, that is, a weak solution to

$$-\Delta u + b \cdot \nabla u \le 0 \tag{3.1}$$

if for any nonnegative test function $v \in C_0^\infty(\mathbb{R}^2)$ we have

$$\int\limits_{\mathbb{R}^2} (A\nabla u) \cdot \nabla v \; dx \le 0. \tag{3.2}$$

Weak supersolutions are defined by reversing both inequalities.

We note that, as mentioned in the introduction, the bilinear form in (3.2) extends continuously to compactly supported $v \in \dot{H}^1$ and hence in (3.2) one can equivalently consider any nonnegative compactly supported $v \in H^1(\mathbb{R}^2)$.

Theorem 3.2. *Assume that $b \in BMO^{-1}(\mathbb{R}^2)$ is divergence-free, and let u be a weak subsolution to (1.4) in $\mathbb{R}^2$. If u is bounded then u is a constant.*

Remark. If the drift b is not too irregular, for example, $b \in L_{2,\mathrm{loc}}(\mathbb{R}^2)$, then distributional solutions to (3.1) can be defined for $u \in L_{2,\mathrm{loc}}(\mathbb{R}^2)$. In this case, bounded solutions are in $H^1_{\mathrm{loc}}(\mathbb{R}^2)$ and satisfy (3.2) automatically, as we will show in the next section.

The proof of Theorem 3.2 is an immediate consequence of the following lemma.

Lemma 3.3. *Let u be a bounded weak subsolution of*

$$-\mathrm{div}(A\nabla u) \le 0 \qquad \text{in } \mathbb{R}^2 \tag{3.3}$$

where $A(x) = a(x) + d(x)$ with a symmetric and d skew symmetric. Assume that there are $\lambda, \Lambda > 0$ such that for any $x, \xi \in \mathbb{R}^2$ we have

$$(a(x)\xi) \cdot \xi \ge \lambda |\xi|^2, \tag{3.4}$$

$$||a||_{L_\infty} \le \Lambda, \tag{3.5}$$

$$||d||_{BMO} \le \Lambda. \tag{3.6}$$

Then u is constant.

Remark. Note that we actually prove a Liouville theorem for bounded subsolutions. This is only possible in two dimensions. In higher dimensions one needs u to be a solution in order to show that it is constant even in the case of the Laplace equation.

Proof. Without loss of generality we can assume that u is nonnegative (otherwise we can add a constant). Let η be the test function

$$\eta(x) = \begin{cases} 1 & \text{if } |x| \leq 1 \\ 1 - \frac{\log|x|}{\log R} & \text{if } 1 \leq |x| \leq R \\ 0 & \text{if } |x| > R \end{cases}$$

We take $v = u\eta^2$ in (3.2) to obtain

$$\begin{aligned} 0 &\geq \int (A\nabla u) \cdot \nabla(u\eta^2) \, \mathrm{d}x \\ &\geq \int \eta^2 (a\nabla u) \cdot \nabla u \, \mathrm{d}x + 2\int u\eta(A\nabla u) \cdot \nabla\eta \, \mathrm{d}x \\ &\geq \lambda \int |\nabla u|^2 \eta^2 \, \mathrm{d}x + 2\int u\eta(A\nabla u) \cdot \nabla\eta \, \mathrm{d}x \end{aligned} \tag{3.7}$$

(all integrals are over $\mathbb{R}^2$ unless otherwise indicated). Therefore we have

$$\lambda \int |\nabla u|^2 \eta^2 \, \mathrm{d}x \leq 2 \left| \int u\eta(A\nabla u) \cdot \nabla\eta \, \mathrm{d}x \right| \tag{3.8}$$

We need to estimate the second term. For the symmetric part of A, we have

$$\begin{aligned} \left| \int u\eta(a\nabla u) \cdot \nabla\eta \, \mathrm{d}x \right| &\leq \Lambda ||u\nabla\eta||_{L_2} ||\eta\nabla u||_{L_2} \\ &\leq \frac{\lambda}{8} ||\eta\nabla u||_{L_2}^2 + C||\nabla\eta||_{L_2}^2 \\ &\leq \frac{\lambda}{8} ||\eta\nabla u||_{L_2}^2 + \frac{C}{\log R} \end{aligned} \tag{3.9}$$

for a constant C depending only on $||u||_{L_\infty}$, Λ, and λ.

Let $\bar{k}$ be the average of k in $B(R)$, the disk of radius R centered at the origin. It is easy to check that

$$\int u\eta(\bar{d}\nabla u) \cdot \nabla\eta \, \mathrm{d}x = 0$$

Now we estimate the contribution to the variable skew-symmetric part of the coefficients using Hölder inequality:

$$\left| \int u\eta(d\nabla u) \cdot \nabla\eta \, \mathrm{d}x \right| \leq C \left(\int_{B(R)} |d - \bar{d}|^4 \right)^{1/4} \left(\int u^4 |\nabla\eta|^4 \right)^{1/4} \left(\int \eta^2 |\nabla u|^2 \right)^{1/2}$$

$$\leq \frac{1}{4}\int \eta^2|\nabla u|^2 \,\mathrm{d}x + C\left(\int\limits_{B(R)} |d-\bar{d}|^4\right)^{1/2}\left(\int u^4|\nabla\eta|^4\right)^{1/2} \tag{3.10}$$

Since d is a BMO function, we have

$$\int\limits_{B(R)} |d-\bar{d}|^4 \,\mathrm{d}x \leq CR^2$$

and by direct computation using that u is bounded,

$$\int u^4|\nabla\eta|^4 \,\mathrm{d}x \leq \frac{C}{R^2(\log R)^4}$$

Therefore, (3.10) gives

$$\left|\int u\eta(d\nabla u)\cdot\nabla\eta \,\mathrm{d}x\right| \leq \frac{1}{4}\int \eta^2|\nabla u|^2 \,\mathrm{d}x + \frac{C}{(\log R)^2} \tag{3.11}$$

Adding (3.9) and (3.11), we estimate the right hand side of (3.8) to obtain

$$\int\limits_{B(1)} |\nabla u|^2 \,\mathrm{d}x \leq \int\limits_{\mathbb{R}^2} |\nabla u|^2\eta^2 \,\mathrm{d}x \leq \frac{C}{\log R} + \frac{C}{(\log R)^2}$$

for a constant C independent of R. We conclude the proof by taking $R \to \infty$. $\square$

Proof of Theorem 1.3. Let $h : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ be such that $h(s) = e^{1-e}s$ for $s \in [0, e^e]$ and $h(s) = \ln s \ln\ln s$ for $s \geq e^e$. For $x = (x_1, x_2)$ define $\hat{x} \equiv \min\{|x_1|, |x_2|\}$ and

$$\tilde{H}(x) \equiv C\,\mathrm{sgn}(x_1x_2)h(\hat{x}) \tag{3.12}$$

with C large. If now $H \equiv \eta * \tilde{H}$ for some radially symmetric smooth mollifier η supported on the unit disc and $b \equiv \nabla^\perp H$, then the hypotheses of the theorem are satisfied. Moreover, if $K^\pm \equiv \{x \mid \pm x_2 \geq |x_1| + 2\}$, then for some $c > 0$ independent of C we have

$$b_1(x) = 0 \qquad \text{and} \qquad -\mathrm{sgn}(x_2)b_2(x) \geq cCh'(\hat{x}) \qquad \text{if } x \in K^+ \cup K^-. \tag{3.13}$$

Let $B_t = B_t(\omega)$ with $\omega \in \Omega$ be the 2-dimensional Brownian motion with $B_0 = 0$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and for $x \in \mathbb{R}^2$ let $X_t^x = X_t^x(\omega)$ be the stochastic process with $X_0^x = x$ and satisfying the SDE

$$dX_t^x = -b(X_t^x)dt + \sqrt{2}dB_t. \tag{3.14}$$

It is then well known (see, e.g., [30, Lemma 7.8]) that if v solves

$$\partial_t v + b\cdot\nabla v - \Delta v = 0$$

on $\mathbb{R}^2$ with $v(0,x) = v_0(x)$, then its value equals the expectation

$$v(t,x) = \mathbb{E}(v_0(X_t^x)). \tag{3.15}$$

We let $v_0(x) = \mathrm{sgn}(x_2)$. Clearly $|v(t,x)| \le 1$ by (3.15), and the symmetry of the drift $(b_1(x_1,-x_2), b_2(x_1,-x_2)) = (b_1(x_1,x_2), -b_2(x_1,x_2))$ gives

$$v(t,x_1,0) = 0 \qquad \text{for all } (t,x_1). \tag{3.16}$$

So $\partial_t v(t,x_1,0) = 0$ and obviously $\mathrm{sgn}(x_2)\partial_t v(0,x_1,x_2) \le 0$ for $x_2 \ne 0$ because of $|v(t,x)| \le 1$ and the choice of v_0. This and the maximum principle for $\partial_t v$ give $\mathrm{sgn}(x_2)\partial_t v(t,x) \le 0$ for all (t,x) and so there exists $u(x) \equiv \lim_{t\to\infty} v(t,x)$. Parabolic regularity shows that u is a (bounded) solution of (1.4).

Let $A_{t,x} \equiv \{\omega : (X_s^x(\omega))_2 \ne 0 \text{ for all } s \in [0,t]\}$ and $A_x = \bigcup_{t>0} A_{t,x}$ (with $(Y)_2$ being the second coordinate of Y). Then (3.15), (3.16), and the strong Markov property for X_t^x imply

$$\begin{aligned} v(t,x) &= \mathrm{sgn}(x_2)\mathbb{P}(A_{t,x}), \\ u(x) &= \mathrm{sgn}(x_2)\mathbb{P}(A_x). \end{aligned}$$

We will now show that $\mathbb{P}(A_{(0,4)}) = \mathbb{P}(A_{(0,-4)}) > 0$ (the equality holds by symmetry), which implies $u(0,4) > 0 > u(0,-4)$.

The law of iterated logarithm (see, e.g., [30, Theorem 5.1.2]) implies that $\mathbb{P}(A) > 0$ for

$$A \equiv \{\omega : \ |B_s| < 1 \text{ for all } s \in [0, e^e] \text{ and } |B_s| < (3s \ln\ln s)^{1/2} \text{ for all } s \ge e^e\}.$$

For any $\omega \in A$, let $t > 0$ be the first time such that $X_t^{(0,4)}(\omega) \in \partial K^+$ (we assume such a time exists and will derive a contradiction). Then (3.14) and the definition of A give $t \ge e^e$, using $\mathrm{dist}((0,4), \partial K^+) > 1$ as well as that $b_1 = 0$ and $-b_2 \ge 0$ in K^+. So $X_s^{(0,4)} \in K^+$ for $s \in [0,t]$, and thus $b_1(X_s^{(0,4)}) = 0$ and $\hat{X}_s^{(0,4)} = |(X_s^{(0,4)})_1|$ for these s. This gives

$$|(X_s^{(0,4)})_1| = |\sqrt{2}(B_s)_1| < (6s\ln\ln s)^{1/2} \le (6t \ln\ln t)^{1/2} \tag{3.17}$$

and so (using (3.13))

$$-b_2(X_s^{(0,4)}) \ge cC \frac{\ln\ln(6t\ln\ln t)^{1/2} + 1}{(6t\ln\ln t)^{1/2}} \ge cC \left(\frac{\ln\ln t}{6t}\right)^{1/2}$$

for $s \in [0,t]$. This means

$$(X_t^{(0,4)})_2 \ge cC\left(\frac{\ln\ln t}{6t}\right)^{1/2} t + \sqrt{2}(B_t)_2 + 4 \ge \left(\frac{cC}{\sqrt{6}} - \sqrt{6}\right)(t\ln\ln t)^{1/2}.$$

If we choose $C \ge 18c^{-1}$, then this and (3.17) give

$$(X_t^{(0,4)})_2 \ge 2(6t\ln\ln t)^{1/2} \ge (6t\ln\ln t)^{1/2} + (6e^e)^{1/2} \ge |(X_t^{(0,4)})_1| + 9,$$

contradicting $X_t^{(0,4)} \in \partial K^+$. Therefore $X_t^{(0,4)}(\omega) \in K^+$ for all $t \ge 0$ and $\omega \in A$. This means that $A \subseteq A_{(0,4)}$ and so $0 < \mathbb{P}(A_{(0,4)}) = \mathbb{P}(A_{(0,-4)})$. Hence $u(0,4) > 0 > u(0,-4)$ and the result follows. $\square$

4. On a modulus of continuity in 2D and a counterexample in 3D

In this section we prove that distributional solutions of (1.4) with a divergence-free b in a two dimensional domain $\Omega \subset \mathbb{R}^2$ are continuous with a logarithmic modulus of continuity that we estimate explicitly. The modulus of continuity depends on a local bound for the H^1 norm of u and it does not essentially depend on any quantity associated with the vector field b. If $b \in L_{1,loc}$, then for suitable u (see Theorem 1.4) we can estimate the modulus of continuity in terms of the L_∞ norm of u instead of H^1 thanks to a local energy inequality. Because of the low regularity assumed for the vector field b, the a priori estimates are hard to extend to distributional solutions and this presents some technical difficulties that are explained below. The estimate is a version of the classical result that functions in the border-line Sobolev spaces which satisfy the maximum principle[9] are continuous, with logarithmic modulus of continuity.

We also show in this section that the same type of regularity result does not hold in three dimensions. Indeed, we construct an example of a function $u \in L_\infty(B) \cap H^1(B)$ (recall that $B = B(0,1)$ is the unit ball) and a vector field $b \in L_1(B)$, such that u solves (1.4) in the distributional sense and is discontinuous at the origin.

Proposition 4.1. *Assume that the drift $b \in L_{2,\mathrm{loc}}(\Omega)$ is divergence-free and $u \in L_\infty(\Omega)$ solves* (1.4) *in the distributional sense (that is,*

$$\int_\Omega b \cdot \nabla v dx = 0 = \int_\Omega (u\Delta v + bu \cdot \nabla v) dx \tag{4.1}$$

for all $v \in C_0^\infty(\Omega)$). Then $u \in H^1_{\mathrm{loc}}(\Omega)$ and if $B(x_0, r) \Subset \Omega$, there is a constant C depending only on r such that

$$||\nabla u||_{L_2(B(x_0,r/2))} \leq C \left(1 + ||b||_{L_1(B(x_0,r))}\right)^{1/2} ||u||_{L_\infty(B(x_0,r))}$$

Proof. The claim $u \in H^1_{\mathrm{loc}}(\Omega)$ is obvious from $bu \in L_{2,\mathrm{loc}}(\Omega)$.

This means that

$$\int_\Omega b \cdot \nabla v dx = 0 = \int_\Omega (\nabla u - bu) \cdot \nabla v dx \tag{4.2}$$

for $v \in H^1(\Omega)$ compactly supported in Ω. Let η be a smooth bump function such that

$$\begin{aligned} \eta &= 1 \qquad \text{in } B(x_0, r/2) \\ \eta &= 0 \qquad \text{in } \Omega \setminus B(x_0, r). \end{aligned}$$

We take $v = u\eta^2$ in (4.2) to obtain

$$0 = \int_{B(x_0,r)} |\nabla u|^2 \eta^2 + 2u\eta \nabla u \cdot \nabla \eta - b \cdot \nabla u u \eta^2 - 2b \cdot \nabla \eta u^2 \eta \, \mathrm{d}x$$

[9]Sometimes the terminology "monotone in the sense of Lebesgue" is used in this context, see e. g. [24].

$$= \int\limits_{B(x_0,r)} |\nabla u|^2\eta^2 + 2u\eta\nabla u\cdot\nabla\eta - b\cdot\nabla\eta u^2\eta \,\mathrm{d}x,$$

where we have used

$$0 = \int\limits_{B(x_0,r)} b\cdot\nabla(u^2\eta^2)dx = 2\int\limits_{B(x_0,r)} b\cdot\nabla\eta u^2\eta + b\cdot\nabla u u\eta^2 dx$$

with $u^2\eta^2 \in H^1(\Omega)$. Therefore

$$\begin{aligned}\int\limits_{B(x_0,r)} |\nabla u|^2\eta^2 \,\mathrm{d}x &= \int\limits_{B(x_0,r)} -2u\eta\nabla u\cdot\nabla\eta + b\cdot\nabla\eta\, u^2\eta \,\mathrm{d}x \\ &\le \frac{1}{2}\int\limits_{B(x_0,r)} |\nabla u|^2\eta^2 \,\mathrm{d}x + \int\limits_{B(x_0,r)} 2u^2|\nabla\eta|^2 + b\cdot\nabla\eta\, u^2\eta \,\mathrm{d}x\end{aligned}$$

and thus

$$\begin{aligned}\frac{1}{2}\int\limits_{B(x_0,r/2)} |\nabla u|^2 \,\mathrm{d}x &\le \int\limits_{B(x_0,r)} 2u^2|\nabla\eta|^2 + b\cdot\nabla\eta\, u^2\eta \,\mathrm{d}x \\ &\le C(1+||b||_{L_1(B(x_0,r))})||u||_{L_\infty}^2.\end{aligned}$$

□

It is worth noting that all statements of Proposition 4.1 hold true in higher dimensions.

We now find the modulus of continuity for functions satisfying the maximum principle and a bound in H^1. Note that in two dimensions, the space H^1 is borderline with respect to the Sobolev embeddings to spaces of continuous functions. The monotonicity of $\mathrm{osc}_{\partial B(r)} u$ is the extra assumption used in the theorem below to actually obtain an explicit modulus of continuity.

Theorem 4.2. *Let $u \in H^1(B)$ and assume that for any $r \in (0,1)$ the maximum principle holds in $B(r)$:*

$$\max_{B(r)} u = \max_{\partial B(r)} u,$$
$$\min_{B(r)} u = \min_{\partial B(r)} u.$$

Then u satisfies the following modulus of continuity estimate at the origin

$$\sup_{x\in B(r)} |u(x)-u(0)| \le \frac{C}{\sqrt{-\log r}}||\nabla u||_{L_2(B)}$$

for some constant C independent of u.

Proof. Let $r \in (0,1)$. We want to estimate $\mathrm{osc}_{B(r)} u = \max_{B(r)} u - \min_{B(r)} u$.

$$\int_{B\setminus B(r)} |\nabla u|^2 \,\mathrm{d}x = \int_r^1 \int_{\partial B(s)} |\nabla u|^2 \,\mathrm{d}\sigma\,\mathrm{d}s$$

since $|\nabla u|^2 = u_\sigma^2 + u_\nu^2$ where u_σ is the tangential derivative and u_ν is the normal one,

$$\geq \int_r^1 \int_{\partial B(s)} |u_\sigma|^2 \,\mathrm{d}\sigma\,\mathrm{d}s$$

Rewriting the integral using polar coordinates ($s\theta = \sigma$),

$$= \int_r^1 \frac{1}{s} \int_{\partial B} |u_\theta(s\theta)|^2 \,\mathrm{d}\theta\,\mathrm{d}s.$$

Since $H^1(\partial B) \subset C^\alpha(\partial B)$ from the one dimensional Sobolev imbedding,

$$\geq \int_r^1 \frac{C}{s} (\mathrm{osc}_{\partial B(s)} u)^2 \,\mathrm{d}s.$$

From the maximum principle, $\mathrm{osc}_{\partial B(s)} u$ is monotone in s, therefore,

$$\geq \int_r^1 \frac{C}{s} (\mathrm{osc}_{\partial B(r)} u)^2 \,\mathrm{d}s = (-C\log r)(\mathrm{osc}_{\partial B(r)} u)^2$$

Taking square roots of both sides we obtain

$$\mathrm{osc}_{B(r)} u = \mathrm{osc}_{\partial B(r)} u \leq \frac{C}{\sqrt{-\log r}} ||\nabla u||_{L_2(B)}$$

□

Consider now a drift $b \in L_1(B)$ and let $u \in L_\infty(B)$ be a distributional solution to (1.4). We are interested in whether u is still a continuous function and, if so, how to estimate its modulus of continuity. We do not know the answer to this question. However, it is in the affirmative if u is an appropriate limit of solutions with L_2 drifts, as in Theorem 1.4.

Proof of Theorem 1.4. The first claim is immediate from the definition of distributional solutions. Moreover, Proposition 4.1 and Theorem 4.2 show that u_m are locally uniformly bounded in H^1 as well as locally uniformly continuous with the modulus of continuity from (1.25), and the second claim follows. □

Finally, we show that that Theorem 1.4 does not hold in higher dimensions in general.

Proof of Theorem 1.5. As in the previous section, we will again consider vector fields with $b(Rx) = Rb(x)$, where $R(x_1, x_2, x_3) \equiv (x_1, x_2, -x_3)$. In addition, b will be axisymmetric with respect to the x_3-axis and with no angular component. Such divergence-free vector fields can again be obtained from a "stream function" $H : \mathbb{R}_0^+ \times \mathbb{R} \to \mathbb{R}$ with $H(0, z) = 0$ as

$$b_H(x) \equiv \nabla \times \left[\frac{H(\rho, z)}{2\rho^2}(-x_2, x_1, 0) \right] = \frac{1}{2\rho^2} \left(x_1 H_z(\rho, z), x_2 H_z(\rho, z), -\rho H_\rho(\rho, z) \right), \tag{4.3}$$

where $\rho \equiv \sqrt{x_1^2 + x_2^2}$ and $z \equiv x_3$. Notice that again we have $b_H \cdot \nabla H = 0 = b_H \cdot (x_2, -x_1, 0)$, so H is constant on the streamlines of b_H, and b_H has no angular component.

We now pick $\alpha \in (\frac{2}{3}, 1)$ and for $\rho^2 + z^2 < 1$ and $\rho \geq 0$ we let

$$\tilde{H}(\rho, z) \equiv \operatorname{sgn}(z) \begin{cases} \rho^2 z^{-2} & \rho^\alpha \leq |z|, \\ z^2 \rho^{-2} & |z|^\alpha \leq \rho, \\ (\rho|z|)^{2(1-\alpha)/(1+\alpha)} & |z| < \rho^\alpha \text{ and } \rho < |z|^\alpha. \end{cases} \tag{4.4}$$

Finally, for some large C we define $H_0 \equiv C\tilde{H}$ and $b_0 \equiv b_{H_0}$.

Notice that H_0 is continuous and vanishes on the axes, and $b_0(Rx) = Rb_0(x)$. We also have

$$-b_0(x) = C\frac{x}{|x_3|^3} \qquad \text{for } (x_1^2 + x_2^2)^{\alpha/2} \leq |x_3| \tag{4.5}$$

as well as $b_0 \in L_1(B)$ (because $|b_0(x)| \leq c|x|^{-2}$ for some $c > 0$ due to $\alpha \in (\frac{2}{3}, 1)$). Moreover, $\tilde{H}$ is smooth except on

$$P \equiv \left\{ \rho^2 + z^2 < 1 : \, \rho \geq 0 \text{ and } |z| \in \{\rho^\alpha, \rho^{1/\alpha}, 0\} \right\},$$

so b_0 is smooth except on

$$S \equiv \left\{ x \in B : \left(\sqrt{x_1^2 + x_2^2}, x_3 \right) \in P \right\}.$$

We therefore let H_ε be smooth such that $H_\varepsilon(\rho, -z) = -H_\varepsilon(\rho, z)$ and $H_\varepsilon = H_0$ outside the ε-neighborhood of P, the vector field $b_\varepsilon \equiv b_{H_\varepsilon}$ is also smooth and $|b_\varepsilon(x)| \leq c|x|^{-2}$ on B, as well as

$$\lim_{\varepsilon \to 0} \|b_\varepsilon - b_0\|_{L_1(B)} = 0. \tag{4.6}$$

Clearly $\nabla \cdot b_\varepsilon = 0$ for $\varepsilon > 0$, so (4.6) gives $\nabla \cdot b_0 = 0$ in the distributional sense.

For each $\varepsilon > 0$ we now construct (smooth) u_ε using b_ε in a way similar to our construction of u using b in the proof of Theorem 1.3. We are here on B so we set boundary conditions $u_\varepsilon(x) = \operatorname{sgn}(x_3)$ on ∂B and thus consider the stochastic process

$$dX_t^{x,\varepsilon} = -b_\varepsilon(X_t^{x,\varepsilon})dt + \sqrt{2}dB_t.$$

with $X_0^{x,\varepsilon} = x$ and stopping time

$$\tau_\varepsilon \equiv \inf \left\{ t \geq 0 \,\middle|\, X_t^{x,\varepsilon} \in \partial B \right\}.$$

We therefore obtain

$$-\Delta u_\varepsilon + b_\varepsilon \cdot \nabla u_\varepsilon = 0 \tag{4.7}$$

where

$$u_\varepsilon(x) = \mathbb{E}\left(\operatorname{sgn}\left(\left(X^{x,\varepsilon}_{\tau_\varepsilon}\right)_3\right)\right) = \operatorname{sgn}(x_3)\mathbb{P}\left((X^{x,\varepsilon}_t)_3 \neq 0 \text{ for all } t \in [0,\tau_\varepsilon]\right), \tag{4.8}$$

Each u_ε is a smooth solution of (4.7) and they are uniformly Hölder continuous away from the origin due to $|b_\varepsilon(x)| \le c|x|^{-2}$. Therefore there is a sequence $\varepsilon_k \to 0$ and u_0 such that $u_{\varepsilon_k} \to u_0$ locally uniformly on $B \setminus \{0\}$. (In fact,

$$\lim_{\varepsilon\to 0} u_\varepsilon(x) = u_0(x) = \operatorname{sgn}(x_3)\mathbb{P}\left(\left(X^{x,0}_t\right)_3 \neq 0 \text{ for all } t \in [0,\tau_0]\right)$$

for any $x \in B$, provided we set $u_0(0) \equiv 0$.) But this, $\|u_\varepsilon\|_{L_\infty} \le 1$, and (4.6) show that

$$\int_B (u_0 \Delta v + b_0 u_0 \cdot \nabla v)dx = \lim_{k\to\infty} \int_B (u_{\varepsilon_k}\Delta v + b_{\varepsilon_k} u_{\varepsilon_k} \cdot \nabla v)dx = 0$$

for any $v \in C_0^\infty(B)$. Thus

$$-\Delta u_0 + b_0 \cdot \nabla u_0 = 0 \tag{4.9}$$

in the distributional sense. The proof of Proposition 4.1 applies to each u_{ε_k}, implying that their weak limit $u_0 \in H^1_{\mathrm{loc}}(B)$. Since for $\varepsilon \ge 0$ we have $u_\varepsilon(Rx) = -u_\varepsilon(x)$ and $||b_\varepsilon||_{L_1(B)}$ is uniformly bounded, we only need to show

$$\lim_{z\downarrow 0}\lim_{\varepsilon\to 0} u_\varepsilon(0,0,z) > 0 \tag{4.10}$$

to conclude the proof of both (i) and (ii).

In fact, let us consider instead of $(0,0,z)$ with $z > 0$ any $y \in B$ with $y_3 > 0$ and $\sqrt{y_1^2 + y_2^2} \le \frac{1}{2} y_3^{1/\alpha}$. Let K_y be the cut-off cone

$$K_y \equiv \left\{ x \in \mathbb{R}^3 : \sqrt{x_1^2 + x_2^2} \le \frac{(2y_3)^{1/\alpha}}{4y_3} x_3 \right\} \subseteq \mathbb{R}^2 \times \mathbb{R}^+,$$

with upper and lower base consisting of discs D_y, E_y centered at $(0,0,2y_3)$, $(0,0,\frac{1}{2}y_3)$ and with radii $\frac{1}{2}(2y_3)^{1/\alpha}$, $\frac{1}{8}(2y_3)^{1/\alpha}$. Notice that its tip would be at the origin, were it not cut off. Then $\sqrt{x_1^2 + x_2^2} \le x_3^{1/\alpha}$ on K_y because $\alpha > \frac{2}{3}$, so (4.5) holds on $K_y \cap B$. Let σ be the exit time of $X^{y,\varepsilon}_t$ from $K_y \cap B$, which is the same for all $\varepsilon \lesssim y_3^{1/\alpha}$ because $b_\varepsilon \equiv b_0$ on $K_y \cap B$ in that case. We will show that

$$\mathbb{P}(X^{y,\varepsilon}_\sigma \in D_y \cup \partial B) \ge 1 - e^{-y_3^{-3+2/\alpha}}, \tag{4.11}$$

provided C from the definition of H_0 is large. This is sufficient since $X^{y,\varepsilon}_\sigma \in D_y \cup \partial B$ means either

$$(X^{y,\varepsilon}_t)_3 \neq 0 \qquad \text{for all } t \in [0,\tau_\varepsilon] \tag{4.12}$$

or $(X^{y,\varepsilon}_\sigma)_3 = 2y_3$. In the latter case we have $((X^{y,\varepsilon}_\sigma)_1^2 + (X^{y,\varepsilon}_\sigma)_2^2)^{1/2} \le \frac{1}{2}(X^{y,\varepsilon}_\sigma)_3^{1/\alpha}$, so we can bootstrap (4.11) and obtain (4.12) with probability at least

$$\prod_{k=1}^{\lfloor -\log_2 y_3 \rfloor} (1 - e^{-(2^k y_3)^{-3+2/\alpha}}) \ge \prod_{j=-\infty}^{0} (1 - e^{-(2^j)^{-3+2/\alpha}}) \equiv m > 0.$$

Thus (4.8) gives $u_\varepsilon(y) \geq m$ for any $y \in B$ such that $y_3 > 0$ and $\sqrt{y_1^2 + y_2^2} \leq \frac{1}{2} y_3^{1/\alpha}$ and any $\varepsilon \lesssim y_3^{1/\alpha}$. The claim of (4.10) now follows immediately.

It remains to prove (4.11) for some large C independent of y. The point here is that $X_t^{y,\varepsilon}$ starts well inside K_y (specifically, $\operatorname{dist}(y, \partial K_y) \geq d y_3^{1/\alpha}$ for some $d \in (0,1)$) and the (strong) drift $-b$ quickly pushes it towards D_y while the Brownian term will not affect this picture much during the short time needed to reach D_y, at least with probability close to 1. We have

$$X_\sigma^{y,\varepsilon} = y - \int_0^\sigma b_0(X_t^{y,\varepsilon}) dt + B_\sigma, \tag{4.13}$$

as well as

$$\mathbb{P}(|B_t| < d y_3^{1/\alpha} \text{ for all } t \in [0, 8C^{-1} y_3^3]) \geq 1 - e^{-y_3^{-3+2/\alpha}}, \tag{4.14}$$

provided C is large enough. Since the vector $-b_0(X_t^{y,\varepsilon})$ points 'inside' the mantle of K_y for $t \in [0, \sigma)$ (because of (4.5) and the fact that the cut-off tip of K_y is the origin) and has a positive third component, and $\operatorname{dist}(y, \partial K_y) \geq d y_3^{1/\alpha}$, this means that with probability at least $1 - e^{-y_3^{-3+2/\alpha}}$, the process $X_t^{y,\varepsilon}$ cannot exit K_y through the mantle or the bottom E_y before time $8C^{-1} y_3^3$. But we have $(-b_0(X_t^{y,\varepsilon}))_3 \geq C(2y_3)^{-2}$ for $t \in [0, \sigma]$, so (4.13), (4.14), and $y_3 + C(2y_3)^{-2} 8C^{-1} y_3^3 - c y_3^{1/\alpha} \geq 2y_3$ yield

$$\mathbb{P}(X_\sigma^{y,\varepsilon} \in D_y \cup \partial B \text{ and } \sigma \leq 8C^{-1} y_3^3) \geq 1 - e^{-y_3^{-3+2/\alpha}}.$$

This proves (4.11) and the result follows. $\square$

References

[1] Aronson, D. G., Non-negative solutions of linear parabolic equations, Ann. Scuola Norm. Sup. Pisa 22 (1968), 607–694.

[2] Aronson, D. G., Serrin, J., Local behavior of solutions of quasilinear parabolic equations, Arch. Rational Mech. Anal. 25 (1967), 81–122.

[3] Caffarelli, L., Kohn, R.-V., Nirenberg, L., Partial regularity of suitable weak solutions of the Navier-Stokes equations, Comm. Pure Appl. Math., Vol. XXXV (1982), pp. 771–831.

[4] Caffarelli, L., Vasseur, A., Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation, Annals of Mathematics, 171(2010), 1903–1930.

[5] Coifman, R., Lions, P.-L., Meyer, Y., Semmes, S., Compensated compactness and Hardy spaces, J. Math. Pures Appl. (9) 72 (1993), no. 3, 247–286.

[6] De Giorgi, E., Sulla dierenziabilita e l'analiticita delle estremali degli integrali multipli regolari, Mem. Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur. (3) 3 (1957), 25–43.

[7] Fabes, E. B.,and D. W. Stroock, A new proof of Moser's parabolic Harnack inequality using the old idea of Nash, Arch. Rational Mech. Anal. 96, no. 4 (1986), 327–338.

[8] Fannjiang, A., Papanicolaou, G., Diffusion in turbulence, Probability Theory and Related Fields, 105, (1996), 297–334.

[9] Ferretti, E., Safonov, M. V., Growth Theorems and Harnack Inequality for Parabolic Equations, Contemporary Mathematics, Volume 277, 2001, pp. 87-112.

[10] Friedlander, S., Vicol, V., Global well-posedness for an advective-diffusion equation arising in magneto-geostrophic dynamics, Annales de l'Institut Henri Poincare (C) Analyse Non Linéaire, 28 (2011), 283–301.

[11] Friedman, A., Partial Dierential Equations of Parabolic Type, Prentice Hall, 1964.

[12] Giaquinta, M., Stuwe, M., On the Partial Regularity of Weak Solutions of Nonlinear Parabolic Systems, Math. Z. 179, 437-451 (1982).

[13] Gilbarg, D., and Trudinger, N. S., Elliptic Partial Differential Equations of second Order,

[14] Ivanov, A. V., The Harnack inequality for generalized solutions of second-order quasilinear parabolic equations, Trudy Mat. Inst. Steklov 102 (1967), 51–84. 2nd ed., Springer-Verlag, Berlin,Heidelberg, New York,Tokyo, 1983.

[15] Kontovourkis, M., On elliptic equatons with low-regularity divergence-free drift terms and the steady-state Navier-Stokes eqautions in higher dimensions, Thesis, University of Minnesota, 2005.

[16] Koch, G., Nadirashvili, N., Seregin, G., and Sverak V., Liouville theorems for Navier-Stokes equations and applications, Acta Math. 203 (2009), no. 1, 83–105.

[17] Krylov, N. V., Nonlinear Elliptic and Parabolic Equations of Second Order, Nauka, Moscow, 1985 in Russian; English translation: Reidel, Dordrecht, 1987.

[18] Ladyzhenskaya, O. A., Solonnikov, V. A., Uralt'seva, N. N., Linear and quasi-linear equations of parabolic type, Moscow, 1967; English translation, American Math. Soc., Providence 1968.

[19] Ladyzhenskaya, O. A., Uraltseva,N. N., Linear and quasilinear equations of elliptic type, "Nauka", Moscow, 1973.

[20] Landis, E. M., Second Order Equations of Elliptic and Parabolic Type, Nauka, Moscow, 1971 in Russian; English transl.: Amer. Math. Soc., Providence, RI, 1997.

[21] Lieberman, G. M., Second Order Parabolic Dierential Equations, World Scientic, 1996.

[22] Mazja, V. G., Verbitsky, I.E., Form boundedness of the general second-order differential operator, Communications on Pure and Applied Mathematics, Vol. LIX, 1286-1329 (2006).

[23] Liskevich, V., Zhang, Qi S., Extra regularity for parabolic equations with drif terms, manuscripta math. 113(2004), pp. 191–209.

[24] Morrey, Ch. B., Multiple Integrals in the Calculus of Variations, Springer, 1966.

[25] Moser, J., A Harnack Inequality for Parabolic Differential Equations, Comm. Pure Appl. Math., Vol. XVII (1964), pp. 101–134.

[26] Moser, J., Correction to "A Harnack Inequality for Parabolic Differential Equations", Comm. Pure Appl. Math., Vol. XX (1967), pp. 231–236.

[27] Moser, J., On a Pointwise Estimate for Parabolic Differential Equations, Comm. Pure Appl. Math., Vol. XXIV (1971), pp. 727–740.

[28] Nash, J., Continuity of solutions of parabolic and elliptic equations, Amer. J. Math. 80 (1958), 931–954.

[29] Nazarov, A. I., Uraltseva, O.A., Quantitative properties of solutions to elliptic and parabolic equations with unbounded lower-order coefficients, SPb. Math. Society Preprint 2009-05, March 2009.

[30] B. Øksendal, *Stochastic Differential Equations,* Springer-Verlag, Berlin, 1995.

[31] Osada, H., Diffusion processes with generators of generalized divergence form, J. Math. Kyoto Univ., 27-4 (1087), 597–619.

[32] Safonov, M. V., Yuan, Yu, Doubling properties for second order parabolic equations, Annals of Mathematics 150 (1999), 313–327.

[33] Semenov, Yu.A., Regularity theorems for parabolic equations, Journal of Functional Analysis, 231(2005), pp. 375–417.

[34] Stampacchia, G., Le probl'eme de Dirichlet pour les .equations elliptiques du second ordre 'a coefficients discontinus, Ann. Inst. Fourier (Grenoble) 15 (1)(1965), pp. 189-258.

[35] Trudinger, N. S.. Pointwise estimates and quasilinear parabolic equations, Comm. Pure and Appl. Math. 21 (1968), 205–226.

[36] Zhang, Qi S., A strong regularity result for parabolic equations, Commun. Math. Phys., 244(2004), pp. 245–260.