HÖLDER CONTINUITY FOR A DRIFT-DIFFUSION EQUATION WITH PRESSURE

LUIS SILVESTRE AND VLAD VICOL

ABSTRACT. We address the persistence of Hölder continuity for weak solutions of the linear drift-diffusion equation with nonlocal pressure

$$u_t + b \cdot \nabla u - \triangle u = \nabla p, \qquad \nabla \cdot u = 0$$

on $[0, \infty) \times \mathbb{R}^n$, with $n \geq 2$. The drift velocity b is assumed to be at the critical regularity level, with respect to the natural scaling of the equations. The proof draws on Campanato's characterization of Hölder spaces, and uses a maximum-principle-type argument by which we control the growth in time of certain local averages of u. We provide an estimate that does not depend on any local smallness condition on the vector field b, but only on scale invariant quantities.

Annales de l'Institut Henri Poincaré (C) Analyse Non Linéaire. Volume 29, Number 4 (2012), 637–652.

1. INTRODUCTION

A classical problem in partial differential equations is to address the regularity of solutions to parabolic problems involving advection by a vector field b and diffusion

$$u_t + b \cdot \nabla u - \triangle u = 0. \tag{1.1}$$

If the vector field b is sufficiently regular, the solution u is expected to be regular as well. Naturally, this is expressed as a result of the type: if b is bounded with respect to some norm, then u is smooth in some sense. The appropriate norms for such statement are the ones that are either critical or subcritical with respect to the inherent scaling of the equation. More precisely, if u is a solution of (1.1), then for any $r > 0$ the function $u_r(x,t) = ru(rx, r^2t)$ solves an equation of the same form but with drift velocity given by $b_r(x,t) = rb(rx, r^2t)$. This change of variables acts as a *zoom in* that focuses on the local behavior of the solution u. An assumption on b is critical with respect to the scaling of the equation if the norm of b_r coincides with the norm of b, for any $r > 0$. The assumption would be subcritical if b_r has smaller norm than b for all small enough values of r, and supercritical otherwise.

As a rule of thumb, with current methods it seems impossible to obtain a regularity result for (1.1) with a supercritical assumption on b, since the transport part of the equation would be stronger than diffusion in the small scales. With subcritical assumptions on b, it is generally possible to treat equation (1.1) as a perturbation of the heat equation, and strong regularity results in this direction are available. The Kato class condition for b is probably the largest class that falls into this category. For results in the subcritical case see for example [1, 4, 25].

Obtaining regularity estimates for (1.1) depending only on *scale invariant* norms of b requires the use of non-perturbative techniques, since the drift term does not become negligible at any scale. To the best of our knowledge, the only results available are variations of the De Giorgi-Nash-Moser Harnack inequality [18], which states that weak solutions to (1.1) are C^α for positive time, for some small $\alpha > 0$. Results in this direction include a variety of critical assumptions on the vector field b. For $b \in L^p_t L^q_x$ with $n/q + 2/p = 1$, we refer to [16, Chapter 3] or [19]. For divergence-free drift $b \in L^\infty_t BMO^{-1}_x$, the Hölder regularity of weak solutions was proved only recently in [9] and [22]. For b in a space-time Morrey space, this result was obtained recently in [19]. See [17, 21, 26] for other conditions on b yielding Hölder regularity, such as the form boundedness condition.

Date: March 19, 2011.

The equation (1.1) is essentially a scalar equation, since even if u is a vector field, each component would satisfy the same equation. In contrast, the De Giorgi-Nash-Moser theory is hard to apply to actual systems. In this article we consider a Stokes system with drift, i.e. we add a pressure term as is common for the equations of fluid dynamics, and we look for a solution u which is divergence free. Given a divergence free vector field $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$, we consider the following evolution equation

$$u_t + b \cdot \nabla u - \triangle u = \nabla p \tag{1.2}$$

for a solution $u : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ which satisfies

$$\nabla \cdot u = 0. \tag{1.3}$$

The term pressure gradient may be computed from (1.2)–(1.3) by the formula

$$\nabla p = \nabla(-\triangle)^{-1} \operatorname{div}(b \cdot \nabla u). \tag{1.4}$$

In this paper we prove that if a scale invariant norm of b is bounded, then the C^α norm of u at time t is bounded in terms of its C^α norm at time zero. Our result is a propagation of regularity instead of a regularization result, in the sense that we require the initial data u_0 to be Hölder continuous, and $\alpha \in (0,1)$ is arbitrary.

The assumption on the divergence-free drift velocity b is that it is an L^p integrable function in time, with values in the L^1-based Morrey-Campanato space M^β, where $\beta \in [-1,1]$, and $p = 2/(1+\beta)$. We recall cf. [24, Definition 1.7.2] the definition of the L^1-based Morrey-Campanato spaces M^β. For any $\beta \in [-1,1]$, we say $f \in M^\beta$ if $f \in L^1_{loc}$ and

$$\sup_{x \in \mathbb{R}^n} \sup_{0<r<1} r^{-\beta} \inf_{P \in \mathcal{P}^\beta} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(z) - P(z)| \, \mathrm{d}z < +\infty, \tag{1.5}$$

where $\mathcal{P}^\beta = \{0\}$ if $-1 \le \beta < 0$, and $\mathcal{P}^\beta = \{\text{constant functions}\}$ if $0 \le \beta \le 1$. The condition $b(\cdot,t) \in M^\beta$ has a different flavor depending of the value of the exponent β, cf. [24, Theorem 5.3.1]: if $\beta = 1$, M^β coincides with the space of Lipschitz functions; if $\beta \in (0,1)$, it is exactly the Hölder class C^β; if $\beta = 0$, it corresponds to the class of functions of bounded mean oscillation BMO; while if $\beta \in [-1,0)$ it is the usual Morrey-Campanato space. In all these cases, the estimate in our main theorem depends *only* on the semi-norm $[b(\cdot,t)]_{M^\beta}$ associated to the space. In this paper we consider divergence-free drifts b such that

$$[b(\cdot,t)]_{M^\beta} := \sup_{x \in \mathbb{R}^3} \sup_{r>0} r^{-\beta} \mathcal{M}(x,t,r) \le g(t) \tag{1.6}$$

for some $g \in L^p([0,T])$, where we define

$$\mathcal{M}(x,t,r) = \frac{1}{r^n} \int_{B_r(x)} |b(z,t) - \bar{b}(x,r,t)| \, \mathrm{d}z = \int_{B_1(0)} |b(x+ry,t) - \bar{b}(x,r,t)| \, \mathrm{d}y \tag{1.7}$$

and $\bar{b}(x,r,t)$ is chosen to equal zero if $\beta \in [-1,0)$, the average of b over $B_r(x)$ if $\beta = 0$, respectively $b(x,t)$ if $\beta \in (0,1]$, which is equivalent to (1.5) (except for $\beta = 1$). We give further details on the precise assumptions on b in Section 3 below. Our main theorem in the case $\beta \in (-1,1]$ is given in Theorem 1.1 below, while the endpoint case $\beta = -1$ is addressed in Theorem 4.1 (see also Remark 1.3).

Theorem 1.1. *Assume $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ is a divergence-free vector field such that $b \in L^p([0,T]; M^\beta)$ with $\beta \in (-1,1]$ and $p = 2/(1+\beta)$. Assume also that $u_0 \in C^\alpha$ for some some $\alpha \in (0,1)$. Then there exists a weak solution $u : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ of the system*

$$u_t + b \cdot \nabla u - \triangle u = \nabla p \tag{1.8}$$

$$\nabla \cdot u = 0 \tag{1.9}$$

$$u(x,0) = u_0(x) \tag{1.10}$$

such that $u(x,t)$ is C^α in x for all positive time $t \in (0,T]$. Moreover, we have the estimate

$$[u(\cdot,t)]_{C^\alpha} \le C [u_0]_{C^\alpha},$$

for some positive universal constant $C = C(T, \alpha, \beta, [b]_{L^p_t M^\beta_x})$.

Remark 1.2. To the best of our knowledge, the result of Theorem 1.1 is new even if the pressure term was removed, and the scalar equation (1.1) was considered instead. Indeed, the De Giorgi-Nash-Moser iteration scheme provides a C^α estimate for the solution for some small α only, whereas our result provides a C^α estimate for any $\alpha \in (0, 1)$.

Remark 1.3. The assumption $b \in L^p_t M^\beta_x$ implies a local smallness condition in the sense that $||b||_{L^p([t-\tau,t],M^\beta)}$ becomes arbitrarily small as $\tau \to 0$, due to uniform integrability, but without any rate. This is not true for the endpoint case $p = \infty$. However, this is not the reason why we require $p < \infty$ in Theorem 1.1, and in fact this local smallness plays no role in our proof. Indeed, the constant C in the estimate of Theorem 1.1 depends only on the scaling-invariant norm of b, and not on any other feature of the vector field b, such as the modulus of continuity of the map $Q \mapsto \|b\|_{L^p_t M^\beta_x(Q)}$. Any argument that relies on the local smallness of b would make the constants in the estimates depend on the rate at which the local norm of b decays, and would hence be implicitly a subcritical result. The reasons why we exclude the case $p = +\infty$ in Theorem 1.1 are more subtle, but we remark that with our current method we could not prove the result in this case.

In the particular case when $b = u$, Theorem 1.1 becomes a no-blowup condition for solutions to Navier-Stokes equation. It says that if the norm of u remains bounded in $L^p_t M^\beta_x$, then u does not blow up on $\mathbb{R}^3 \times [0, T]$. This is a scale invariant condition that is slightly more general (but has the same scaling) than the classical Ladyzhenskaya-Foias-Prodi-Serrin condition $u \in L^p_t L^q_x$ with $2/p+n/q \leq 1$ (cf. [8, 15, 20, 23]). Note that the endpoint case $(q, p) = (3, \infty)$, when $n = 3$, was only treated recently in [10] (see also [11] and Theorem 4.1 below for a related statement). The result of Theorem 1.1 for the full range of $p \in [1, \infty)$ may be new for the Navier-Stokes equations as well, though it is comparable to other available regularity criteria in terms of scaling critical norms of u (cf. [3, 5, 6, 14] and references therein).

One important difficulty for proving Theorem 1.1 is to deal with the non-local pressure term on the right hand side of (1.2). There are very few results of this kind available for equations with pressure terms. In [26] the same equation (1.2) is considered and a Lipschitz estimate is shown under a sub-critical assumption on b (which includes $b \in L^p_t L^q_x$ with $2/p + n/q \leq 1 - \epsilon$, for any $\epsilon > 0$).

The idea of the proof is to write the Hölder regularity condition of $u(\cdot, t)$ in integral form using a classical theorem of Campanato [2]. Then we claim that these local integral estimates have a certain growth in time (in terms of integral estimates on b). In order to prove that these estimates hold for all time we argue by contradiction and look for the first point in which they would be invalidated. At that time we apply the equation and obtain a contradiction in a way that resembles maximum-principle-type arguments (see also [12, 13] for the SQG equations). The integral representation of the Hölder modulus of continuity allows us to take advantage of the divergence-free condition and the integral bound on b. The divergence free condition on u is used in the estimate for the gradient of pressure term. The general method of the proof introduced here seems to be new, as it may be applied to systems with pressure gradients, and we believe it may be applicable to other evolution equations in the future.

In section 4 we analyze the endpoint case $\beta = -1$. The method of this article is applicable in this case, but we need to impose an extra smallness condition on the vector field b (cf. (4.1)–(4.2) below).

We believe that the most important contribution of this article is the introduction of a new method to prove Hölder estimates for evolution equations. We believe it is particularly interesting that the method can be carried out even for systems that are coupled through the pressure term. We also provide an example of how Theorem 1.1 can help prove that a nonlinear equation is well posed. Let us consider the following *modified* energy critical Navier-Stokes equation in 3D (see [7] for a similar modified critically dissipative SQG equation)

$$\partial_t u + \left((-\Delta)^{-1/4} u \cdot \nabla\right) u - \Delta u = \nabla p \tag{1.11}$$

$$\operatorname{div} u = 0 \tag{1.12}$$

that is, $b = (-\Delta)^{-1/4} u$ in (1.8)–(1.9). It follows directly from Theorem 1.1 that the system (1.11)–(1.12) is well-posed in the classical sense. Indeed, the global existence of weak solutions $u \in L^\infty_t L^2_x \cap L^2_t \dot{H}^1_x$ is straightforward, as is the local existence of strong solutions. In particular, for any $T > 0$, we have that $(-\Delta)^{-1/4} u$ is a priori bounded in $L^2(0,T; H^{3/2}(\mathbb{R}^3))$, and by the Calderón-Zygmund theorem we have

$$\left\| (-\Delta)^{-1/4} u \right\|_{L^2(0,T;BMO)} < \infty \tag{1.13}$$

for any $T > 0$. Therefore, by applying Theorem 1.1 with $\beta = 0$, we obtain that $u(\cdot, t) \in C^\alpha$ on $(0, T]$, given that $u(\cdot, 0) = u_0 \in C^\alpha$, where $\alpha \in (0, 1)$ is arbitrary. From this estimate, it is easy to obtain higher regularity of u by a standard bootstrap argument.

Note that for the above system one could also use classical energy estimates at the level of vorticity, combined with Sobolev interpolation, to obtain the global well-posedness of the problem. On the other hand our method allows some extra flexibility in the relationship between b and u. As explained above, when $b = (-\Delta)^{-1/4} u$ the system is well posed. Following essentially the same idea we can obtain using Theorem 1.1 that the system is well posed for any of the following choices

- $b = a(x)(-\triangle)^{-1/4} u + \nabla q$ for any bounded function a in $\mathbb{R}^3$ and ∇q is the gradient of a scalar function that makes b divergence free. In this case we apply the a priori estimate $u \in L^4 L^3$ that is obtained by interpolation from the energy inequality, and gives $b \in L^4 L^6$.
- $b = \int k(x,y) u(y)\, \mathrm{d}y$ where $k(x,-) \in L^{2n/(n+2)}$ for any $x \in \mathbb{R}^n$ and $k(-,y)$ is divergence free for any $y \in \mathbb{R}^n$.
- $\triangle b = \operatorname{div}(u \otimes u) + \nabla q$.

We plan to explore other applications of this method in the future.

2. Preliminaries

In this section we state a few introductory remarks about the weak and classical solutions to (1.8)–(1.9), and recall a classical characterization of Hölder spaces in terms of local averages. Throughout the rest of the paper we will write $L^p L^q$ to denote $L^p_t L^q_x = L^p(0,T; L^q)$, and similarly $L^p M^\beta$ will be used instead of $L^p_t M^\beta_x$.

We first prove Theorem 1.1 assuming that the solution is classical (i.e. C^2 in space and C^1 in time). The important feature is that the a priori estimate (1.6) depends only on the assumptions of Theorem 1.1 and not on any further smoothness assumptions on b or u. Then we approximate any weak solution with classical solutions by using a mollification of b, and pass to the limit to obtain the result of Theorem 1.1 in full generality.

Definition 2.1 (Weak Solutions). *If $b \in L^1_{loc}([0,T] \times \mathbb{R}^n)$ is divergence-free, a function $u \in L^\infty([0,T] \times \mathbb{R}^n)$ is a weak solution of (1.2), if it is weakly divergence-free, and for all smooth, divergence-free, compactly supported test functions φ we have:*

$$\int_{\mathbb{R}^n} \varphi(x,T) u(x,T)\, \mathrm{d}x + \int_{[0,T]\times\mathbb{R}^n} u\, \left(-\varphi_t + b\nabla\varphi - \triangle\varphi\right)\, \mathrm{d}x\, \mathrm{d}t = \int_{\mathbb{R}^n} \varphi(x,0) u(x,0)\, \mathrm{d}x.$$

The following proposition is standard.

Proposition 2.2. *Let b^ε and u^ε be a sequence of smooth divergence-free vector fields. Assume that u^ε is a weak solution of (1.2) with drift velocity b^ε. Assume also that $b^\varepsilon \to b$ strongly in $L^1_{loc} L^1_{uloc}$. Then, up to a subsequence, u^ε converges weakly to a weak solution of (1.2).*

Using proposition 2.2, we immediately observe the following.

Proposition 2.3. *It is enough to prove Theorem 1.1 assuming that b is smooth and u is a classical solution.*

Proof of Proposition 2.3. The assumption $b \in L^p M^\beta$ implies in particular that $b \in L^1_{loc} L^1_{uloc}$. Using a mollification argument, we consider a sequence of smooth vector fields b^ε converging strongly to b in $L^1_{loc} L^1_{uloc}$.

Moreover, we choose b^ε such that $||b^\varepsilon||_{L^p M^\beta}$ is bounded uniformly with respect to ε (for example mollifying with a smooth function with fixed L^1 norm). For each of these vector fields, we solve the equation (1.2), for instance using the mild formulation and Picard iteration, to obtain a smooth solution u^ε. If the result of Theorem 1.1 is known for classical solutions, we would have that u^ε satisfies the estimate (1.6) uniformly in ε. Note that in particular we also obtain $u^\varepsilon \in L^\infty$. By Proposition 2.2, up to a subsequence, u^ε converges weakly to a weak solution u of (1.2), and therefore this solution u satisfies (1.6) as well. □

In order to prove the main theorem, we use a local integral characterization of Hölder spaces. For this purpose, let φ be a nonnegative, radially symmetric, smooth function supported in $B_1(0)$. Unless otherwise specified, the center of the unit ball B_1 in $\mathbb{R}^n$ shall be 0. Let us also assume that $\int \varphi(y)\ \mathrm{d}y = 1$. The following theorem (or a small variation of it) is proved in [2].

Theorem 2.4 (Campanato's characterization of Hölder spaces). *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be an L^2 function such that for all $r > 0$ and $x \in \mathbb{R}^n$, there exists a constant $\bar{f}$ such that*

$$\int_{B_1} |f(x + ry) - \bar{f}|^2 \varphi(y)\ \mathrm{d}y \leq A^2 r^{2\alpha} \tag{2.1}$$

for some positive constant A, and $\alpha \in (0, 1)$. Then the function f has a Hölder continuous representative such that

$$|f(x) - f(y)| \leq BA|x - y|^\alpha$$

where the constant B depends on dimension and α only.

The most natural choice of the constant $\bar{f}$ in the above theorem, for which the converse also holds. is to choose the average of f in the ball.

$$\bar{f} = \int_{B_1} f(x + ry)\varphi(y)\ \mathrm{d}y.$$

This is optimal in the sense that it minimizes the left hand side in (2.1) (see also (1.5)).

The theorem of Campanato is interesting because it provides a non-obvious equivalence between a Hölder modulus of continuity, which is a priori a pointwise property, and averages of differences of the function, which is an integral property. This relation will allow us to exploit the divergence free nature of the vector fields b and u when estimating the evolution of a Hölder modulus of continuity.

3. Evolution of a modulus of continuity

We will prove that the solutions of (1.8)–(1.10) do not lose regularity by showing that they always satisfy a time dependent Hölder modulus of continuity. This modulus of continuity will evolve and deteriorate with time, but it will stay bounded. In order to take advantage appropriately of the divergence-free character of the vector field u, we use the integral characterization of the modulus of continuity. Let φ be a radially symmetric weight supported in B_1 with mass one as in section 2. We denote the weighted mean of u on $B_r(x)$ by

$$\bar{u}(x, t, r) = \int_{B_1} u(x + ry, t)\varphi(y)\ \mathrm{d}y. \tag{3.1}$$

The integral version of the modulus of the continuity of u is then

$$\mathcal{I}(x, t, r) = \int_{B_1} |u(x + ry, t) - \bar{u}(x, t, r)|^2 \varphi(y)\ \mathrm{d}y. \tag{3.2}$$

Due to Theorem 2.4, if we knew that

$$\mathcal{I}(x, t, r) \leq f(t)^2 r^{2\alpha}, \tag{3.3}$$

for some function $f(t) > 0$, and all $r > 0$, then $[u(\cdot, t)]_{C^\alpha} \leq Cf(t)$ for some universal constant C. Our goal is to prove that estimate (3.3) holds for all $t > 0$, if it holds at $t = 0$, for some function $f(t)$ to be chosen appropriately.

As discussed in the introduction, our assumptions on b will be in terms of quantities similar to $\mathcal{I}$, which are distinguished by the parameter $\beta \in [-1,1]$ as follows.

(i) The Morrey-Campanaoto case. For $\beta \in [-1,0)$, let

$$\mathcal{M}(x,t,r) = \int_{B_1} |b(x+ry,t)|\,\mathrm{d}y = \frac{1}{r^n}\int_{B_r(x)} |b(z,t)|\,\mathrm{d}z. \tag{3.4}$$

We assume that there exists a positive function $g \in L_t^{2/(1+\beta)}$ such that

$$\sup_{x\in\mathbb{R}^n}\sup_{r>0} r^{-\beta}\mathcal{M}(x,r,t) \leq g(t) \Leftrightarrow \|b(\cdot,t)\|_{M^\beta} \leq g(t) \tag{3.5}$$

for all $t \geq 0$, where $\|\cdot\|_{m^s}$ denotes the usual Morrey norm (cf. [24]).

(ii) The BMO case. For $\beta = 0$, we let

$$\mathcal{M}(x,t,r) = \int_{B_1} |b(x+ry,t) - \bar{b}(x,r,t)|\,\mathrm{d}y = \frac{1}{r^n}\int_{B_r(x)} |b(z,t) - \bar{b}(x,r,t)|\,\mathrm{d}z, \tag{3.6}$$

where

$$\bar{b}(x,r,t) = \frac{1}{r^n}\int_{B_r(x)} b(z,t)\,\mathrm{d}z \tag{3.7}$$

is the usual mean of b on $B_r(x)$. We assume that there exists a positive function $g \in L_t^2$ such that

$$\sup_{x\in\mathbb{R}^n}\sup_{r>0} \mathcal{M}(0,x,r,t) \leq g(t) \Leftrightarrow \|b(\cdot,t)\|_{BMO} \leq g(t) \tag{3.8}$$

for all $t \geq 0$, where $\|\cdot\|_{BMO}$ denotes the norm on the space of functions with bounded mean oscillation.

(iii) The Hölder and Lipschitz cases. For $\beta \in (0,1]$, we consider

$$\mathcal{M}(x,t,r) = \int_{B_1} |b(x+ry,t) - b(x,t)|\,\mathrm{d}y = \frac{1}{r^n}\int_{B_r(x)} |b(z,t) - b(x,t)|\,\mathrm{d}z. \tag{3.9}$$

We assume that there exists a positive function $g \in L_t^{2/(1+\beta)}$ such that

$$\sup_{x\in\mathbb{R}^n}\sup_{r>0} r^{-\beta}\mathcal{M}(x,t,r) \leq g(t) \Leftrightarrow [b(\cdot,t)]_{C^\beta} \leq g(t) \tag{3.10}$$

for all $t \geq 0$, where $[\cdot]_{C^\beta}$ denotes the Hölder semi-norm. Note that $2/(1+\beta) \in [1,2)$ when $\beta \in (0,1]$.

We shall prove that if (3.5), (3.8), or respectively (3.10) holds, then we have $\mathcal{I}(x,t,r) < f(t)^2 r^{2\alpha}$ for all $t > 0$. The proof is in the flavor of a maximum principle. We show that if the inequality is satisfied at $t = 0$, it will be satisfied for all positive t. Kiselev, Nazarov and Volberg [13] used an argument is the same spirit for proving the well posedness of the critical dissipative SQG equation. But that as opposed to (1.8)–(1.9), it is a scalar equation that has an L^∞ maximum principle and the modulus of continuity can be studied pointwise in their case.

In order to prove Theorem 1.1, assume there is a first time t and some value of x where the *strict* modulus is invalidated, i.e.

$$\mathcal{I}(x,t,r) = f(t)^2 r^{2\alpha}. \tag{3.11}$$

By Proposition 2.3, we can assume that u is a smooth function vanishing at infinity. Therefore the equality in (3.11) of the modulus must be achieved at some $r > 0$ and $x \in \mathbb{R}^n$.

If we fix t and r, the function I achieves its maximum at x, and we obtain

$$0 = \nabla_x \mathcal{I} = \int_{B_1} (u(x+ry,t) - \bar{u}(x,t,r)) \cdot (\nabla_x u(x+ry,t) - \nabla_x \bar{u}(x,t,r))\varphi(y)\,\mathrm{d}y. \tag{3.12}$$

Due to the definition of $\bar{u}$ (3.1), and the fact that $\nabla_x \bar{u}(x,t,r)$ does not depend on y, we also have

$$0 = \int_{B_1} (u(x+ry,t) - \bar{u}(x,t,r)) \cdot \nabla_x u(x+ry,t)\varphi(y)\, \mathrm{d}y. \tag{3.13}$$

Since $\mathcal{I} < f(t)^2 r^{2\alpha}$ for all times prior to t, and since $(u-\bar{u})\varphi$ has zero mean, we thus conclude

$$2f'(t)f(t)r^{2\alpha} \leq \partial_t \mathcal{I} = \int_{B_1} (u(x+ry,t) - \bar{u}(x,t,r)) \cdot (\partial_t u(x+ry,t) - \partial_t \bar{u}(x,t,r))\varphi(y)\, \mathrm{d}y \tag{3.14}$$

$$= \int_{B_1} (u(x+ry,t) - \bar{u}(x,t,r)) \cdot \partial_t u(x+ry,t)\varphi(y)\, \mathrm{d}y. \tag{3.15}$$

The key to prove Theorem 1.1 is to find an appropriate upper bound for the right hand side of (3.15) in terms of $f(t)$ and $\mathcal{M}(r,t)$. Inserting the equation (1.2) in the right hand side of (3.15), we obtain

$$\begin{aligned} 2f'(t)f(t)r^{2\alpha} &\leq \int_{B_1} (u(x+ry,t) - \bar{u}(x,t,r)) \\ &\qquad \cdot \Big(- b(x+ry,t)\cdot \nabla_x u(x+ry,t) + \triangle_x u(x+ry,t) - \nabla p(x+ry,t)\Big)\varphi(y)\, \mathrm{d}y \\ &= \mathcal{A} + \mathcal{D} + \mathcal{P}. \end{aligned} \tag{3.16}$$

The following three lemmas give bounds to the three terms on the right side of (3.16). The advection term $\mathcal{A}$ is the simplest one to estimate. Observe that

$$\nabla_x u(x+ry) = \frac{1}{r}\nabla_y u(x+ry). \tag{3.17}$$

This identity, together with the assumption $\operatorname{div} b = 0$, allows us to integrate by parts the gradient into the weight φ and obtain a precise estimate for $\mathcal{A}$.

The dissipative term turns out to be negative, but we must analyze it with care in order to obtain a precise lower bound on its absolute value. In fact, note that if u is linear in $B_r(x)$ then $\mathcal{D} = 0$. We will obtain an estimate of $\mathcal{D}$ that measures how much u is forced to separate from a linear function, just from the values of $\mathcal{I}$ and $\partial_r \mathcal{I}$ at the point where the equality (3.11) holds.

Lastly, we obtain an upper bound for the pressure term $\mathcal{P}$, comparable to the advection term $\mathcal{A}$. This is to be expected since ∇p is obtained from $b\cdot\nabla u$ though an operator of order zero. However, the pressure estimate is more involved since the formula for the pressure is non-local and in order to obtain the desired estimate we need to take advantage of some cancellations that occur after integration by parts of Riesz kernels using that both b and u are divergence free.

We now carry out the estimates for the three terms on the right of (3.16) in the three lemmas below.

Lemma 3.1 (The advection term). *Let u and b be as in the statement of Theorem 1.1. Then we have*

$$\mathcal{A} = \int_{B_1} (u(x+ry,t) - \bar{u}(x,t,r)) \cdot (-b(x+ry,t)\cdot\nabla_x u(x+ry))\varphi(y)\, \mathrm{d}y \leq C r^{2\alpha-1} f(t)^2 \mathcal{M}(x,t,r) \tag{3.18}$$

for all $\beta \in [-1,1]$, where C is a positive constant depending only on α, φ, and n.

Proof of Lemma 3.1. Using the identities (3.13), (3.17), and integrating by parts, we obtain from (3.18) that

$$\mathcal{A} = \frac{1}{r}\int_{B_1} |u(x+ry,t) - \bar{u}(x,t,r)|^2 (b(x+ry,t) - \bar{b}(x,r,t))\cdot\nabla\varphi(y)\, \mathrm{d}y \tag{3.19}$$

where $\bar{b} = \bar{b}(x,r,t)$ is a constant with respect to y, to be chosen suitably in the three cases for $\beta \in [-1,1]$, as discussed above. From identity (3.19), the Hölder inequality, and Theorem 2.4, we directly obtain

$$\mathcal{A} \leq C r^{2\alpha-1} f(t)^2 \mathcal{M}(x,t,r), \tag{3.20}$$

for all $\beta \in [-1,1]$, where C is a positive constant depending on α, n, and φ through $\sup_{B_1} |\nabla\varphi|$. □

The second term corresponds to the viscosity and it is strictly negative, as we will show below.

Lemma 3.2 (The dissipative term). *Let u and b be as in the statement of Theorem 1.1. Then we have*

$$\mathcal{D} = \int_{B_1} \triangle_x u(x+ry,t) \cdot (u(x+ry,t) - \bar{u}(x,t,r))\varphi(y)\, \mathrm{d}y \leq -cf(t)^2 r^{2\alpha-2} \tag{3.21}$$

for all $r > 0$, where c is a sufficiently small positive constant, depending only on n, α, and φ.

Remark 3.3. Note that the constant c in Lemma 3.2 goes to zero as $\alpha \to 1$. This is the reason why Theorem 1.1 works for $\alpha < 1$ only.

In order to prove Lemma 3.2, we need the following technical result, which relates the quantities $\mathcal{D}$ and $\mathcal{I}$.

Lemma 3.4. *For a fixed x, we have*

$$\mathcal{D}(r) \geq c\frac{(2\mathcal{I}(r) - r\mathcal{I}'(r))^2}{r^2\mathcal{I}(r)} \tag{3.22}$$

for some sufficiently small positive constant c.

For simplicity we omited the x dependence in (3.22). Note that if f is a linear function, then $\mathcal{D} = 0$ and $\mathcal{I} = Cr^2$ for some constant C. Moreover, in this case $r\mathcal{I}' = 2\mathcal{I}$. We see that if $(r\mathcal{I}' - 2\mathcal{I})$ is non zero, then the function u cannot be linear. The inequality (3.22) gives a precise quantitative version of this fact.

We note that the identity

$$\int_{B_1} |f(y) - c|^2 \varphi(y)\, \mathrm{d}y = \int_{B_1}\int_{B_1} (f(y) - f(z))^2 \varphi(y)\varphi(z)\, \mathrm{d}y\, \mathrm{d}z \tag{3.23}$$

holds for any constant c, in particular for $c = \bar{f}$, giving equivalent definitions for $\mathcal{I}$ and $\mathcal{D}$ in terms of double-integrals. The proof of identity (3.23) is straightforward, and hence omitted.

With formula (3.23) in mind, we prove the following lemma, which is exactly the case $r = 1$ for (3.22). The proof of Lemma 3.4 follows for all other values of r from Lemma 3.5 by scaling.

Lemma 3.5. *Let $f : B_1 \to \mathbb{R}^n$ be any H^1 function. There is a constant C depending only on n and φ such that*

$$\begin{aligned}\iint_{B_1\times B_1} (f(y) - f(z))^2 \varphi(y)\varphi(z)\, \mathrm{d}y\, \mathrm{d}z - \iint_{B_1\times B_1} (y\cdot\nabla f(y) - z\cdot\nabla f(z))\cdot(f(y) - f(z))\varphi(y)\varphi(z)\, \mathrm{d}y\, \mathrm{d}z \\ \leq C\left(\iint_{B_1\times B_1} (\nabla f(y) - \nabla f(z))^2 \varphi(y)\varphi(z)\, \mathrm{d}y\, \mathrm{d}z\right)^{1/2} \left(\iint_{B_1\times B_1} (f(y) - f(z))^2 \varphi(y)\varphi(z)\, \mathrm{d}y\, \mathrm{d}z\right)^{1/2}.\end{aligned}$$

Proof of Lemma 3.5. We start by writing $f(y) - f(z)$ as an integral of ∇f along the segment between z and y. Thus, we have

$$\begin{aligned}&\iint_{B_1\times B_1} (f(y) - f(z))^2 \varphi(y)\varphi(z)\, \mathrm{d}y\, \mathrm{d}z \\ &\qquad = \iint_{B_1\times B_1}\int_0^1 (\nabla f(sy + (1-s)z)\cdot(y-z))\cdot(f(y) - f(z))\varphi(y)\varphi(z)\, \mathrm{d}s\, \mathrm{d}y\, \mathrm{d}z.\end{aligned}$$

Substituting in the first term of the left hand side, we obtain

$$\begin{aligned}
&\iint_{B_1\times B_1} (f(y)-f(z))^2\varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z \\
&\qquad\qquad - 2\iint_{B_1\times B_1} (y\cdot\nabla f(y) - z\cdot\nabla f(z))\cdot(f(y)-f(z))\varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z \\
&\quad = \iint_{B_1\times B_1}\int_0^1 \Big((\nabla f(sy+(1-s)z)-\nabla f(y))\cdot y - (\nabla f(sy+(1-s)z)-\nabla f(y))\cdot z\Big) \\
&\qquad\qquad \times (f(y)-f(z))\varphi(y)\varphi(z)\,\mathrm{d}s\,\mathrm{d}y\,\mathrm{d}z \\
&\quad \le \mathcal{I}^{1/2}\mathcal{E}^{1/2},
\end{aligned}$$

where $\mathcal{I}$ is as defined in (3.2), and $\mathcal{E}$ is defined as

$$\mathcal{E} = \iint_{B_1\times B_1}\int_0^1 \Big((\nabla f(sy+(1-s)z)-\nabla f(y))\cdot y - (\nabla f(sy+(1-s)z)-\nabla f(y))\cdot z\Big)^2\,\mathrm{d}s\,\mathrm{d}y\,\mathrm{d}z.$$

The quantity $\mathcal{E}$ is a bounded quadratic functional with respect to the vector field ∇f that vanishes whenever ∇f is constant. Therefore, there is a constant C such that

$$\mathcal{E} \le C\iint_{B_1\times B_1} (\nabla f(y)-\nabla f(z))^2\varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z = C\,\mathcal{D},$$

which concludes the proof. □

Proof of Lemma 3.2. We recall that at the first point of equality in (3.11), the function $\mathcal{I}$ achieves its maximum as a function of x. Then we have $\triangle_x\mathcal{I}\le 0$, and hence

$$\begin{aligned}
0 \ge &\int_{B_1} (\triangle_x u(x+ry,t) - \triangle_x\bar{u}(x,t,r))\cdot(u(x+ry,t)-\bar{u}(x,t,r))\varphi(y)\,\mathrm{d}y \\
&+ \int_{B_1} |\nabla u(x+ry,t) - \nabla_x\bar{u}(x,t,r)|^2\varphi(y)\,\mathrm{d}y.
\end{aligned}$$

Therefore, since $(u-\bar{u})\varphi$ has zero mean and $\triangle_x\bar{u}$ does not depend on y, we obtain the inequality

$$\mathcal{D} \le -\int_{B_1} |\nabla u(x+ry,t) - \nabla_x\bar{u}(x,t,r)|^2\varphi(y)\,\mathrm{d}y. \tag{3.24}$$

According to (3.24), $\mathcal{D}$ is negative, but we must estimate how negative it is. It is convenient to rewrite the formula for $\mathcal{I}(x,t,r)$ using (3.23) as

$$\begin{aligned}
\mathcal{I}(x,t,r) &= \int_{B_1} |u(x+ry,t)-\bar{u}(x,t,r)|^2\varphi(y)\,\mathrm{d}y \\
&= \iint_{B_1\times B_1} |u(x+ry,t)-u(x+rz,t)|^2\varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z.
\end{aligned} \tag{3.25}$$

At the point where $f(t)^2r^{2\alpha} = \mathcal{I}(x,t,r)$ for the first time, we have $\nabla_x\mathcal{I}=0$ and $\triangle_x\mathcal{I}\le 0$, so that

$$\begin{aligned}
0 &\ge \triangle_x\mathcal{I}(x,t,r) \\
&= 2\iint_{B_1\times B_1} |\nabla_x u(x+ry,t) - \nabla_x u(x+rz,t)|^2 \\
&\qquad + (\triangle_x u(x+ry,t) - \triangle_x u(x+rz,t))\cdot(u(x+ry,t)-u(x+rz,t))\varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z \\
&= 2\mathcal{D} + 2\iint_{B_1\times B_1} |\nabla_x u(x+ry,t) - \nabla_x u(x+rz,t)|^2\varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z,
\end{aligned}$$

where $\mathcal{D}$ is given by (3.24). We thus have proven the inequality

$$\begin{aligned}\mathcal{D} &\leq -\iint_{B_1\times B_1} |\nabla_x u(x+ry,t) - \nabla_x u(x+rz,t)|^2 \varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z \\ &= -\frac{1}{r^2}\iint_{B_1\times B_1} |\nabla_y u(x+ry,t) - \nabla_z u(x+rz,t)|^2 \varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z. \end{aligned} \tag{3.26}$$

The right hand side is clearly negative unless u is an affine function. We now need to estimate how negative it is, in terms of $\mathcal{I}$ and $\partial_r\mathcal{I}$. Note that at the first point of equality $\mathcal{I} = f(t)^2 r^{2\alpha}$, we must also have $\partial_r\mathcal{I} = 2\alpha f(t)^2 r^{2\alpha-1}$. We compute $\partial_r\mathcal{I}$ as

$$\begin{aligned}\partial_r\mathcal{I} &= 2\int_{B_1} (u(x+ry,t) - \bar{u}(x,t,r))\; y\cdot\nabla_x u(x+ry,t)\varphi(y)\,\mathrm{d}y \\ &= 2\iint_{B_1\times B_1} (u(x+ry,t) - u(x+rz,t))\cdot(y\cdot\nabla_x u(x+ry,t) - z\cdot\nabla_x u(x+rz,t))\varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z \\ &= \frac{2}{r}\iint_{B_1\times B_1} (u(x+ry,t) - u(x+rz,t))\cdot(y\cdot\nabla_y u(x+ry,t) - z\cdot\nabla_z u(x+rz,t))\varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z. \end{aligned} \tag{3.27}$$

From the expression (3.25) and (3.27), we can apply Lemma 3.5 to obtain

$$(2\mathcal{I} - r\partial_r\mathcal{I}) \leq C\mathcal{I}^{1/2}\left(\iint_{B_1\times B_1} |\nabla_y u(x+ry,t) - \nabla_z u(x+rz,t)|^2\varphi(y)\varphi(z)\,\mathrm{d}y\,\mathrm{d}z\right)^{1/2}.$$

Recalling the inequality (3.26), we obtain

$$\mathcal{D} \leq -\frac{c}{r^2}\frac{(2\mathcal{I} - r\partial_r\mathcal{I})^2}{\mathcal{I}} = -cf(t)^2(1-\alpha)^2 r^{2\alpha-2} \tag{3.28}$$

for some positive constant c, depending only on α, n, and φ. □

Lastly, we bound the pressure term $\mathcal{P}$ arising on the right side of (3.16).

Lemma 3.6 (The pressure term). *Let u and b be as in the statement of Theorem 1.1. Then we have*

$$\mathcal{P} = \int_{B_1} (u(x+ry,t) - \bar{u}(x,r,t))\cdot\nabla p(x+ry,t)\varphi(y)\,\mathrm{d}y \leq Cf(t)^2 g(t) r^{2\alpha+\beta-1} \tag{3.29}$$

for all $r > 0$ and $\beta \in [-1,1]$, where C is a positive constant, depending on α, β, n, and φ.

Proof of Lemma 3.6. Since all estimates in this section hold for a fixed time $t > 0$, we omit the time dependence of all functions. Recall that the function $\bar{b} = \bar{b}(x,r)$ (which is constant respect to y) is chosen to be $\bar{b} = b(x)$ in the Hölder case, the average of b over $B_r(x)$ in the BMO case, or $\bar{b} = 0$ in the Morrey-Campanato case.

In order to estimate the third term in (3.16), let us analyze the identity (1.4). Since u is divergence free, we have

$$\begin{aligned}\nabla_x p(x+ry) &= \nabla_x(-\triangle)_x^{-1}\,\mathrm{div}_x(b(x+ry)\cdot\nabla_x u(x+ry)) \\ &= \frac{1}{r}\nabla_y(-\triangle_y)^{-1}\,\mathrm{div}_y\Big((b(x+ry) - \bar{b}(x,r))\cdot\nabla_y(u(x+ry) - \bar{u}(x,r))\Big). \end{aligned} \tag{3.30}$$

We have the pressure term $\mathcal{P}$ equal to

$$\begin{aligned}&\int (u(x+ry) - \bar{u}(x,r))\cdot\nabla_x p(x+ry)\varphi(y)\,\mathrm{d}y \\ &= \frac{1}{r}\int (u(x+ry) - \bar{u}(x,r))\cdot\nabla_y(-\triangle_y)^{-1}\,\mathrm{div}_y\Big((b(x+ry) - \bar{b}(x,r))\cdot\nabla_y(u(x+ry) - \bar{u}(x,r))\Big)\varphi(y)\,\mathrm{d}y. \end{aligned}$$

We integrate the ∇_y by parts and use that u is divergence free to obtain that $\mathcal{P}$ equals

$$\frac{1}{r}\int (-\triangle_y)^{-1}\operatorname{div}_y\Big((b(x+ry)-\bar b(x,r))\cdot\nabla_y(u(x+ry)-\bar u(x,r))\Big)(u(x+ry)-\bar u(x,r))\cdot\nabla\varphi(y)\,\mathrm{d}y.$$

Using that b and u are divergence free, we re-write the above identity as

$$\begin{aligned}\mathcal{P}&=\frac{1}{r}\int (-\triangle_y)^{-1}\partial_i\partial_j\Big((b(x+ry)-\bar b(x,r))_i(u(x+ry)-\bar u(x,r))_j\Big)(u(x+ry)-\bar u(x))\cdot\nabla\varphi(y)\,\mathrm{d}y\\&=\frac{1}{r}\int R_iR_j\Big((b(x+ry)-\bar b(x,r))_i(u(x+ry)-\bar u(x,r))_j\Big)(u(x+ry)-\bar u(x,r))\cdot\nabla\varphi_y(y)\,\mathrm{d}y\\&=\frac{1}{r}\int (b(x+ry)-\bar b(x,r))_i(u(x+ry)-\bar u(x,r))_jR_iR_j\Big((u(x+ry)-\bar u(x,r))\cdot\nabla\varphi(y)\Big)\,\mathrm{d}y,\end{aligned}\tag{3.31}$$

where the Riesz transforms are taken with respect to the y variable.

The third factor inside the integral is given by Riesz transforms in y of the function

$$\psi_{x,r}(y)=\psi(y)=(u(x+ry)-\bar u(x,r))\cdot\nabla\varphi(y)=\operatorname{div}_y\Big((u(x+ry)-\bar u(x,r))\varphi(y)\Big),\tag{3.32}$$

since u is divergence free. The function ψ is supported in B_1 and we have $\int_{B_1}\psi(y)\,\mathrm{d}y=0$. Moreover, since by assumption $I(x,r)\le Cf(t)r^\alpha$ for some $\alpha\in(0,1)$, we have by Theorem 2.4 that $\|u(x+ry)-\bar u(x,r)\|_{L^\infty_y(B_1)}\le Cf(t)r^\alpha$, and therefore $\|\psi\|_{L^\infty_y(B_1)}\le Cf(t)r^\alpha$, uniformly in x. In fact, ψ is also C^α, with C^α norm bounded by $Cf(t)r^\alpha$ (note the scaling in y).

Gathering these bounds together, we see that $R_iR_j(\psi)$ must be bounded in $L^\infty(\mathbb{R}^n)$ by $Cf(t)r^\alpha$, uniformly in x. Indeed, the Riesz transforms are bounded on C^α and on L^2 (since ψ is supported on B_1 and is bounded there, its L^2 norm is also finite), and therefore $R_iR_j(\psi)\in L^2(\mathbb{R}^n)\cap C^\alpha(\mathbb{R}^n)\supset L^\infty(\mathbb{R}^n)$.

The Riesz transforms of functions with compact support are not compactly supported. The decay of the Riesz transform of a compactly supported function is normally of order $-n$. However, in this case since the function ψ has integral zero, and since it is a derivative, we have that $R_iR_j(\psi)$ decays like $|y|^{-n-2}$ for $|y|$ large. To see this, let K_{ij} be the Kernel associated with the Riesz transform and, using that $\operatorname{div}u=0$, we compute

$$\begin{aligned}R_{ij}(\psi)(y)=\int\psi(z)K_{ij}(y-z)\,\mathrm{d}z&=\int\varphi(z)(u(x+rz)-\bar u(x,r))\cdot\nabla K_{ij}(y-z)\,\mathrm{d}z\\&=\int\varphi(z)(u(x+rz)-\bar u(x,r))\cdot\Big(\nabla K_{ij}(y-z)-\nabla K_{ij}(y)\Big)\,\mathrm{d}z\end{aligned}\tag{3.33}$$

in principal value sense, since $\int\varphi(u-\bar u)=0$. Letting y be such that $|y|>2$, we obtain that $|\nabla^2K_{ij}(\xi)|\le C|y|^{n+2}$, for all ξ that lies between y and $y-z$, where C is a sufficiently large dimensional constant, gives the desired decay in $|y|$.

To summarize, we have proved that

$$R_iR_j[(u(x+ry)-\bar u(x,r))\cdot\nabla\varphi(y)]\le\begin{cases}Cf(t)r^\alpha & \text{if } |y|<2\\ \frac{Cf(t)r^\alpha}{|y|^{n+2}} & \text{if } |y|\ge 2\end{cases}\tag{3.34}$$

where C is a universal constant that does not depend on x. Recall that by our assumption and Campanato's theorem we also have $(u(x+ry)-\bar u(x,r))\le Cf(t)(ry)^\alpha$. Therefore, we obtain from (3.31) and (3.34) the estimate

$$\mathcal{P}\le Cf(t)^2r^{2\alpha-1}\left(\int_{B_2}|b(x+ry)-\bar b(x,r)|\,\mathrm{d}y+\int_{\mathbb{R}^n\setminus B_2}|b(x+ry)-\bar b(x,r)|\frac{1}{|y|^{n+2-\alpha}}\,\mathrm{d}y\right).\tag{3.35}$$

It is here where it is necessary to make the distinction between the Hölder and Morrey-Camapanato cases for the a priori assumptions on b, by making the specific choices for $\bar b$.

The Hölder case. If $\bar{b}(x,r) = b(x)$, it is clear that

$$\int_{B_2} |b(x+ry) - \bar{b}(x)| \, \mathrm{d}y \leq C\mathcal{M}(x, r/2) \leq Cg(t)r^{\beta} \tag{3.36}$$

for $\beta \in (0,1]$. To bound the tail of the integral arising from the Riesz transforms, we first change variables $ry = z$ so that

$$\begin{aligned}
\int_{\mathbb{R}^n \setminus B_2} |b(x+ry) - \bar{b}(x)| \frac{1}{|y|^{n+2-\alpha}} \, \mathrm{d}y &= r^{2-\alpha} \int_{\mathbb{R}^n \setminus B_{2r}} |b(x+z) - \bar{b}(x)| \frac{1}{|z|^{n+2-\alpha}} \, \mathrm{d}z \\
&= r^{2-\alpha} \int_{2r}^{\infty} \int_{\partial B_\rho} |b(x+z) - \bar{b}(x)| \, \mathrm{d}\sigma(z) \frac{1}{\rho^{n+2-\alpha}} \, \mathrm{d}\rho \\
&= r^{2-\alpha} \int_{2r}^{\infty} \frac{\partial}{\partial \rho} \left(\int_{B_\rho} |b(x+z) - \bar{b}(x)| \, \mathrm{d}\sigma(z) \right) \frac{1}{\rho^{n+2-\alpha}} \, \mathrm{d}\rho \\
&\leq Cr^{2-\alpha} \int_{2r}^{\infty} (\rho^n \mathcal{M}(x,\rho)) \frac{1}{\rho^{n+3-\alpha}} \, \mathrm{d}\rho \\
&\leq Cr^{2-\alpha} g(t) \int_{2r}^{\infty} \frac{1}{\rho^{3-\alpha-\beta}} \leq Cg(t)r^{\beta}.
\end{aligned} \tag{3.37}$$

From (3.35), (3.36), and (3.37), we obtain

$$\mathcal{P} \leq Cf(t)^2 g(t) r^{2\alpha+\beta-1}. \tag{3.38}$$

The BMO case. In this case we have $\bar{b}(x,r) = (1/r^n) \int_{B_r(x)} b(z) \, \mathrm{d}z$. The difference with the Hölder case lies in the tail of the integral due to the Riesz transform. We use the following classical fact about BMO functions: the difference between the mean on B_r an $B_{\lambda r}$ is bounded by $2^n \ln(1+\lambda)$ times the BMO norm, for all $\lambda > 1$. We split the integral in dyadic cylinders to find a bound for each and sum.

$$\begin{aligned}
\int_{\mathbb{R}^n \setminus B_2} |b(x+ry) - \bar{b}(x,r)| \frac{1}{|y|^{n+2-\alpha}} \, \mathrm{d}y &= \sum_{k=1}^{\infty} \int_{B_{2^{k+1}} \setminus B_{2^k}} |b(x+ry) - \bar{b}(x,r)| \frac{1}{|y|^{n+2-\alpha}} \, \mathrm{d}y \\
&\leq \sum_{k=1}^{\infty} \int_{B_{2^{k+1}} \setminus B_{2^k}} \left(|b(x+ry) - \bar{b}(x, 2^{k+1}r)| + |\bar{b}(x, 2^{k+1}r) - \bar{b}(x,r)| \right) \frac{1}{2^{k(n+2-\alpha)}} \, \mathrm{d}y \\
&\leq \sum_{k=1}^{\infty} \frac{1}{2^{k(2-\alpha)}} ||b||_{BMO} + \frac{1}{2^{k(n+2-\alpha)}} \int_{B_{2^{k+1}r} \setminus B_{2^k r}} |\bar{b}(x, 2^{k+1}r) - \bar{b}(x,r)| \, \mathrm{d}y \\
&\leq \sum_{k=1}^{\infty} \frac{1}{2^{k(2-\alpha)}} ||b||_{BMO} + \frac{1}{2^{k(n+2-\alpha)}} C 2^{kn} \log(1+2^{k+1}) ||b||_{BMO} \leq C||b||_{BMO}.
\end{aligned}$$

Thus, the tail of the integral in the bound of $\mathcal{P}$ (3.35) is bounded by $C||b||_{BMO} \leq Cg(t)$, as well as the fist term in (3.35). Therefore, in this case ($\beta = 0$) we also obtain

$$\mathcal{P} \leq Cf(t)^2 g(t) r^{2\alpha+\beta-1}.$$

The Morrey-Campanato case. In this case we have $\bar{b}(x,r) = 0$. The same proof as in the Hölder case above, but with $\bar{b} = 0$, shows that

$$\mathcal{P} \leq Cf(t)^2 g(t) r^{2\alpha+\beta-1} \tag{3.39}$$

when $\sup_{r>0} r^{-\beta} \mathcal{M}(x,r) < \infty$, and $\beta \in [-1, 0)$. □

Once we have estimated the three terms on the right side of (3.16), the proof of the main theorem is concluded as follows.

Proof of Theorem 1.1. Since the vector drift b is a priori assumed to lie in M^β, for some $\beta \in (-1,1]$, with $[b(\cdot,t)]_{M^\beta} \le g(t)$, and $g \in L_t^{2/(1+\beta)}$, we obtain from (3.20), (3.28), and (3.38) the estimate

$$f'(t)f(t)r^{2\alpha} \le C_* f(t)^2 g(t) r^{2\alpha+\beta-1} - c_* f(t)^2 r^{2\alpha-2} \tag{3.40}$$

which holds for some sufficiently large constant C_*, and some sufficiently small positive constant c_*. The above estimate implies

$$\frac{f'(t)}{f(t)} \le C_* g(t) r^{\beta-1} - c_* r^{-2} \le \bar{C} g(t)^{2/(1+\beta)} \tag{3.41}$$

for some positive constant $\bar{C} = \bar{C}(C_*, c_*, \beta)$. The last inequality was obtained maximizing the expression with respect to r.

On the other hand, if $g \in L_t^{2/(1+\beta)}$, we can choose f to be the solution to the ODE

$$f'(t) = 2\bar{C} g(t)^{2/(1+\beta)} f(t),$$

which contradicts the above inequality and makes it impossible for the Hölder modulus of continuity to ever be invalidated. Note that the above ODE has the explicit solution

$$f(t) = e^{\left(\int_0^t C g(s)^{2/(1+\beta)}\, \mathrm{d}s\right)} f(0),$$

which stays bounded for all t. □

4. The endpoint case $\beta = -1$

For the three-dimensional Navier-Stokes equations, i.e. $b = u$, obtaining regularity of the solutions in the endpoint case $u \in L^\infty L^3$ is highly non-trivial, and this issue was only settled recently by Iskauriaza, Seregin, and Sverak in [10] (see also [11] for the case $L^\infty H^{1/2}$). The case $(p,q) = (\infty, n)$ on the Ladyzhenskaya-Foias-Prodi-Serrin scale is of particular importance as it is the scaling-critical space for the initial data, and it gives the borderline space (on the Lebesgue scale) for constructing solutions via the Picard iteration scheme (so-called mild solutions). Note that $L^3 \subset M^{-1}$.

It turns out that for the linear system (1.8)–(1.10), the proof given above fails in the case $b \in L^\infty M^{-1}$, as the constant $\bar{C}$ blows up as $\beta \to -1$. The corresponding result for the Navier-Stokes equations (as in [10] or [11]) relies essentially on the nonlinear structure of the equation. In order to obtain a result in this direction for the liner equation (1.2), with the method of this article, we need to impose an extra smallness condition.

Theorem 4.1. *Let $T > 0$ be arbitrary. Assume that b is a divergence-free vector field in $L^\infty([0,T]; M^{-1})$, and let $u_0 \in L^2 \cap C^\alpha$ for some $\alpha \in (0,1)$. There exists a positive constant $\epsilon > 0$ such that if for all $t \in (0,T]$ there exists $r_*(t) > 0$ with*

$$\sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^n} \sup_{0<r<r_*(t)} r \int_{B_1} |b(x+ry,t)|\, \mathrm{d}y \le \epsilon \tag{4.1}$$

and

$$\int_0^T \frac{1}{r_*(t)^2}\, \mathrm{d}t < \infty, \tag{4.2}$$

then $u(\cdot,t) \in C^\alpha(\mathbb{R}^n)$ for all $t \in (0,T]$.

Remark 4.2. If $\|b\|_{L^\infty M^{-1}} \le \epsilon$, the theorem holds trivially. Additionally, note that for any $\phi \in L^3(\mathbb{R}^3)$, or any ϕ in the closure of C_0^∞ in M^{-1}, we have

$$\lim_{r\to 0} r \int_{B_1} |\phi(x+ry)|\, \mathrm{d}y = 0,$$

for all x, and hence (4.1) holds for some $r_* > 0$. Therefore, if b is continuous in time with values in L^3, or piecewise continuous with arbitrarily large jumps, or if all jumps are of size smaller than $\varepsilon/2$, the conditions

(4.1)–(4.2) are automatically satisfied, with $r_*(t)$ being a sufficiently small constant, proving regularity of the solution (see also [5, Theorem 3.1] for a similar result in the critical Besov space). Theorem 4.1 states that if the drift velocity $b(\cdot,t)$ is nicely behaved at sufficiently small scales $r_*(t)$, and if these scales do not go to 0 *too fast* at any point $t \in [0,T]$, i.e. $r_*(s)$ cannot vanish at t with a rate faster than $\sqrt{t-s}$, then the solution is regular until T. Note that with respect to time regularity, the conditions of Theorem 4.1 above are stronger than merely $b \in L^\infty_t L^3_x$, but weaker than $C_t L^3_x$.

Proof of Theorem 4.1. From (3.20), (3.28), (3.36), and (3.37), we obtain

$$2f'(t)f(t)r^{2\alpha} \leq C_* r^{2\alpha-1} f(t)^2 \mathcal{M}(x,t,r) - c_* f(t)^2 r^{2\alpha-2} + C_* f(t)^2 r^{2\alpha-1} \left(\mathcal{M}(x,t,r/2) + r^{2-\alpha} \int_{2r}^{\infty} \mathcal{M}(x,t,\rho) \frac{1}{\rho^{3-\alpha}} \, \mathrm{d}\rho \right), \tag{4.3}$$

where $\mathcal{M}(x,t,r) = \int_{B_1} |b(x+ry,t)| \, \mathrm{d}y$, and by assumption of the theorem we have

$$\sup_{0<r<r_*(t)} r\mathcal{M}(x,t,r) \leq \epsilon \tag{4.4}$$

and

$$\sup_{r>0} r\mathcal{M}(x,t,r) \leq \|b\|_{L^\infty_t M^{-1}_1} = B < \infty \tag{4.5}$$

uniformly in x,t. Inserting the bounds (4.4)–(4.5) into (4.3), we have

$$2r^2 \frac{d}{dt} \log f(t) \leq C_* r\mathcal{M}(x,t,r) - c_* + C_* r\mathcal{M}(x,t,r/2) + C_* r^{3-\alpha} \left(\int_{2r}^{r_*} (\rho \mathcal{M}(x,t,\rho)) \frac{1}{\rho^{4-\alpha}} \, \mathrm{d}\rho + \int_{r_*}^{\infty} (\rho \mathcal{M}(x,t,\rho)) \frac{1}{\rho^{4-\alpha}} \, \mathrm{d}\rho \right) \tag{4.6}$$

for all $r > 0$. To bound the right side of (4.6), we distinguish the cases $r/r_* \leq \delta$, and $r/r_* > \delta$, where we let

$$\delta = \min \left\{ \left(\frac{\epsilon}{B} \right)^{1/(3-\alpha)}, \frac{1}{2} \right\}. \tag{4.7}$$

Indeed, if $r/r_* \leq \delta$, we have $2r < r_*$ and $B(r/r_*)^{3-\alpha} \leq B\delta^{3-\alpha} \leq \epsilon$, so that by (4.1) implies

$$\begin{aligned} 2r^2 \frac{d}{dt} \log f(t) &\leq C_*\epsilon - c_* + 2C_*\epsilon + C_* r^{3-\alpha} C_\alpha \left(\epsilon r^{\alpha-3} + B r_*^{\alpha-3} \right) \\ &\leq C_*(3+2C_\alpha)\epsilon - c_*, \end{aligned} \tag{4.8}$$

for some positive constant $C_\alpha > 0$, and for all $0 < r \leq \delta \, r_*$. Therefore, if we choose ϵ as

$$\epsilon = \frac{c_*}{C_*(3+2C_\alpha)}, \tag{4.9}$$

then we obtain from (4.8) that

$$\frac{d}{dt} \log f(t) \leq 0 \tag{4.10}$$

for all $t \in [0,T]$, and all $0 < r \leq \delta \, r_*(t)$. On the other hand, if $r > \delta \, r_*(t)$, we bound the right side of (4.6) by making use of (4.5), namely

$$\frac{d}{dt} \log f(t) \leq \frac{1}{2r^2} \left(3C_* B + C_* B r^{3-\alpha} \int_{2r}^{\infty} \frac{1}{\rho^{4-\alpha}} \right) \leq \frac{C_* B(3+C_\alpha)}{2\delta^2 r_*(t)^2} \tag{4.11}$$

for some $C_\alpha > 0$, for all $r > \delta \, r_*(t)$. The proof of the theorem is then concluded since we may choose f as

$$f(t) = f(0) \exp \left(\frac{C_* B(3+C_\alpha)}{\delta^2} \int_0^t \frac{1}{r_*(s)^2} \right) \tag{4.12}$$

which is finite for all $t \leq T$ by (4.2), and it contradicts (4.10)–(4.11). □

ACKNOWLEDGMENT

Luis Silvestre was partially supported by NSF grant DMS-1001629 and the Sloan Foundation. Vlad Vicol was partially supported by and AMS-Simmons travel grant.

REFERENCES

1. D. G. Aronson, *Non-negative solutions of linear parabolic equations*, Ann. Scuola Norm. Sup. Pisa (3) **22** (1968), 607–694. MR 0435594 (55 #8553)
2. S. Campanato, *Propietà di hölderianità di alcune classi di funzioni*, Ann. Scuola Norm. Sup. Pisa (3) **17** (1963), 175–188. MR 0156188 (27 #6119)
3. C. H. Chan and A. Vasseur, *Log improvement of the Prodi-Serrin criteria for Navier-Stokes equations*, Methods Appl. Anal. **14** (2007), no. 2, 197–212. MR 2437103 (2009h:35313)
4. Z. Q. Chen and Z. Zhao, *Diffusion processes and second order elliptic operators with singular coefficients for lower order terms*, Math. Ann. **302** (1995), no. 2, 323–357. MR 1336339 (96e:60136)
5. A. Cheskidov and R. Shvydkoy, *The regularity of weak solutions of the 3D Navier-Stokes equations in $B^{-1}_{\infty,\infty}$*, Arch. Ration. Mech. Anal. **195** (2010), no. 1, 159–169. MR 2564471
6. ———, *Regularity problem for the 3d Navier-Stokes equations: the use of Kolmogorov's dissipation range*, Arxiv preprint arXiv:1102.1944v1 (2011).
7. P. Constantin, G. Iyer, and J. Wu, *Global regularity for a modified critical dissipative quasi-geostrophic equation*, Indiana Univ. Math. J. **57** (2008), no. 6, 2681–2692. MR 2482996 (2009k:35120)
8. C. Foiaş and G. Prodi, *Sur le comportement global des solutions non-stationnaires des équations de Navier-Stokes en dimension 2*, Rend. Sem. Mat. Univ. Padova **39** (1967), 1–34. MR 0223716 (36 #6764)
9. S. Friedlander and V. Vicol, *Global well-posedness for an advection-diffusion equation arising in magneto-geostrophic dynamics*, Ann. Inst. H. Poincaré Anal. Non Linéaire **28** (2011), no. 2, 283–301. MR 2784072
10. L. Iskauriaza, G. A. Seregin, and V. Sverak, *$L_{3,\infty}$-solutions of Navier-Stokes equations and backward uniqueness*, Uspekhi Mat. Nauk **58** (2003), no. 2(350), 3–44. MR 1992563 (2004m:35204)
11. C. E. Kenig and G. S. Koch, *An alternative approach to regularity for the Navier-Stokes equations in critical spaces*, Ann. Inst. H. Poincaré Anal. Non Linéaire **28** (2011), no. 2, 159–187. MR 2784068
12. A. Kiselev, *Regularity and blow up for active scalars*, Math. Model. Nat. Phenom. **5** (2010), no. 4, 225–255. MR 2662457
13. A. Kiselev, F. Nazarov, and A. Volberg, *Global well-posedness for the critical 2D dissipative quasi-geostrophic equation*, Invent. Math. **167** (2007), no. 3, 445–453. MR 2276260 (2008f:35308)
14. I. Kukavica, *Regularity for the Navier-Stokes equations with a solution in a Morrey space*, Indiana Univ. Math. J. **57** (2008), no. 6, 2843–2860. MR 2483004 (2010f:35286)
15. O. A. Ladyženskaja, *Uniqueness and smoothness of generalized solutions of Navier-Stokes equations*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **5** (1967), 169–185. MR 0236541 (38 #4836)
16. O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'tseva, *Linear and quasilinear equations of parabolic type*, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23, American Mathematical Society, Providence, R.I., 1967. MR 0241822 (39 #3159b)
17. V. Liskevich and Q. S. Zhang, *Extra regularity for parabolic equations with drift terms*, Manuscripta Math. **113** (2004), no. 2, 191–209. MR 2128546 (2006a:35127)
18. J. Moser, *On Harnack's theorem for elliptic differential equations*, Comm. Pure Appl. Math. **14** (1961), 577–591. MR 0159138 (28 #2356)
19. A. I. Nazarov and N. N. Ural'tseva, *The Harnack inequality and related properties of solutions of elliptic and parabolic equations with divergence-free lower-order coefficients*, Algebra i Analiz **23** (2011), no. 1, 136–168. MR 2760150 (2012d:35026)
20. G. Prodi, *Un teorema di unicità per le equazioni di Navier-Stokes*, Ann. Mat. Pura Appl. (4) **48** (1959), 173–182. MR 0126088 (23 #A3384)
21. Y. A. Semenov, *Regularity theorems for parabolic equations*, J. Funct. Anal. **231** (2006), no. 2, 375–417. MR 2195337 (2007b:35149)
22. G. A. Seregin, L. Silvestre, V. Šverák, and A. Zlatoš, *On divergence-free drifts*, Journal of Differential Equations **252** (2012), no. 1, 505 – 540.
23. J. Serrin, *On the interior regularity of weak solutions of the Navier-Stokes equations*, Arch. Rational Mech. Anal. **9** (1962), 187–195. MR 0136885 (25 #346)
24. H. Triebel, *Theory of function spaces. II*, Monographs in Mathematics, vol. 84, Birkhäuser Verlag, Basel, 1992. MR 1163193 (93f:46029)
25. Q. S. Zhang, *Gaussian bounds for the fundamental solutions of $\nabla(A\nabla u) + B\nabla u - u_t = 0$*, Manuscripta Math. **93** (1997), no. 3, 381–390. MR 1457736 (98c:35067)
26. ———, *Local estimates on two linear parabolic equations with singular coefficients*, Pacific J. Math. **223** (2006), no. 2, 367–396. MR 2221033 (2007k:35378)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO, 5734 UNIVERSITY AVENUE, CHICAGO, IL 60637.
E-mail address: luis@math.uchicago.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO, 5734 UNIVERSITY AVENUE, CHICAGO, IL 60637.
E-mail address: vicol@math.uchicago.edu