Ann. Inst. H. Poincaré
Anal. Non Linéaire 39 (2022), 793–818
DOI 10.4171/AIHPC/20


# Construction of maximal functions associated with skewed cylinders generated by incompressible flows and applications

Jincheng Yang

**Abstract.** We construct a maximal function associated with a family of skewed cylinders. These cylinders, which are defined as tubular neighborhoods of trajectories of a mollified flow, appear in the study of fluid equations such as the Navier–Stokes equations and the Euler equations. We define a maximal function subordinate to these cylinders and show it is of weak type $(1, 1)$ and strong type $(p, p)$ by a covering lemma. As an application, we give an alternative proof for the higher-derivatives estimate of smooth solutions to the three-dimensional Navier–Stokes equations.

## 1. Introduction

This paper is dedicated to the study of the maximal functions adapted to the Lagrangian description of a flow. When studying the motion of a fluid, there are two different but deeply connected descriptions to work with. The Eulerian formulation records physical quantities such as velocity, temperature, and pressure at fixed positions, while the Lagrangian formulation builds the frame of reference following each moving fluid parcel, and describes their motion and trajectories by a flow map. The transport phenomenon is easier to describe in the Lagrangian formulation, while the diffusion usually suits the Eulerian description better. Let us refer to the works of Constantin ([8]), Kukavica and Vicol ([9]) for the connection and distinction between these two descriptions in the context of Euler equations.

For both mathematical study and numerical simulation, sometimes it is necessary to switch between two descriptions. For instance, in computational fluid dynamics, the *vortex particle method* treats the fluid as a collection of vortex particles, moving along the trajectories generated by the velocity field, which is in turn recovered from vortex particles. Its early development was by Chorin on the study of the two-dimensional Navier–Stokes equations ([5]). The validity and convergence of this vortex method in three and two dimensions are confirmed by Beale and Majda in [1, 2]. We refer interested readers to the books of Raviart ([17]), Cottet and Koumoutsakos ([10]), and Majda and Bertozzi ([16])

*2020 Mathematics Subject Classification.* 42B25, 76D05, 35Q30.
*Keywords.* Maximal function, covering lemma, incompressible flows, Lagrangian/Eulerian representation, partial regularity.

for detailed bibliographies. Majda and Bertozzi also used the particle-trajectory method to show existence and uniqueness results for Euler equations. Even recently, hybrid numerical schemes are still a very active area ([14]). To avoid singularities in the computation, a mollification is applied to the velocity field. Therefore, particles are in fact moving along approximated trajectories of this mollified flow defined in Definition 1. Mollification is also needed for this Lagrangian formulation when the velocity field does not have enough regularity to define trajectories and flow maps, for instance, weak solutions to Navier–Stokes equations or Euler equations.

Before introducing our new maximal function, let us recall the classical one. For any real-valued or vector-valued function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ with $d \geq 1$, recall the *classical* maximal function $\mathcal{M} f$ is defined as

$$(\mathcal{M} f)(x) := \sup_{r>0} \fint_{B_r(x)} |f(y)| \, \mathrm{d}y = \sup_{r>0} \frac{1}{|B_r|} \int_{B_r(x)} |f(y)| \, \mathrm{d}y. \tag{1}$$

Here $B_r(x)$ is a $d$-dimensional ball with radius $r$ and center $x$, and $|B_r|$ stands for its $d$-dimensional Lebesgue measure $\mathcal{L}^d$. Throughout the article, we may use $|\cdot|$ to represent the spatial Lebesgue measure $\mathcal{L}^d$ or the space-time Lebesgue measure $\mathcal{L}^{d+1}$ depending on the context. The strength of the maximal function is that it captures the nonlocal information of a function, and in the meantime keeps the homogeneity: it commutes with rigid motion and scaling, as well as scalar multiplication. This maximal function $\mathcal{M}$ is a bounded operator on $L^p$ for $1 < p \leq \infty$; it is also bounded from $L^1$ to $L^{1,\infty}$, the weak $L^1$ space. However, if we include a time variable $t$ in an evolutionary problem – for instance, a transport equation – Euclidean balls in space-time are no longer the most natural objects to work with. Instead, we may consider using a space-time cylinder, or "skewed cylinder", transported in the space-time, to be more rigorously defined below. In this paper we will study such cylinders and construct a maximal function associated with them.

Consider a vector field $u: (S, T) \times \mathbb{R}^d \to \mathbb{R}^d$ satisfying

$$u \in L^1_{\mathrm{loc}}(S, T; \dot{W}^{1,p}(\mathbb{R}^d))$$

for some $1 \leq p \leq \infty$, where $d \geq 1$ and $-\infty \leq S < T \leq \infty$ are some finite or infinite initial and terminal times fixed throughout this article. Fix a spatial function $\varphi \in C_c^\infty(B_1)$ satisfying $\int \varphi \, \mathrm{d}x = 1$, $\varphi \geq 0$, where $B_1 \subset \mathbb{R}^d$ is a unit ball of dimension $d$. Define the usual mollifier function $\varphi_\varepsilon := \varepsilon^{-d} \varphi(\cdot/\varepsilon) \in C_c^\infty(B_\varepsilon)$. We denote a universal constant by $C$ if it depends only on $\varphi$ and $d$. Its value may change from line to line. We define the *spatially* mollified velocity $u_\varepsilon: (S, T) \times \mathbb{R}^d \to \mathbb{R}^d$ by

$$u_\varepsilon(t, x) := [u(t, \cdot) * \varphi_\varepsilon](x) = \int_{\mathbb{R}^d} u(t, x - y) \varphi_\varepsilon(y) \, \mathrm{d}y.$$

By convolution, we have $u_\varepsilon \in L^1_{\mathrm{loc}}(S, T; C^1(\mathbb{R}^d))$. Let us now define the mollified flow and the skewed cylinders.

**Definition 1** (Mollified flow, skewed cylinders). For some fixed $\varepsilon > 0$ and $(t,x) \in (S,T) \times \mathbb{R}^d$, define the *mollified flow* $X_\varepsilon(t,x;\cdot)$ to be the unique solution to the initial value problem

$$\begin{cases} \dot{X}_\varepsilon(t,x;s) = u_\varepsilon(s, X_\varepsilon(t,x;s)), \\ X_\varepsilon(t,x;t) = x, \end{cases} \qquad s \in (S,T),$$

where the dot means taking the derivative in the last argument $s$. Moreover, if $S + \varepsilon^2 < t < T - \varepsilon^2$, define the *skewed parabolic*[1] *cylinder* with center $(t,x)$ and radius $\varepsilon$ by

$$Q_\varepsilon(t,x) := \big\{(s,y) : |s-t| < \varepsilon^2,\ |y - X_\varepsilon(t,x;s)| < \varepsilon\big\}.$$

Heuristically speaking, skewed cylinders defined in Definition 1 are objects appearing in the Lagrangian formulation but written in Eulerian coordinates. Indeed, they are following the mollified flow and capturing particles that are close to the center trajectories. Similar to the difficulty of bridging these two formulations, the difficulty of working with these cylinders comes from the lack of control on the distortion. Without a uniform control on the velocity field, these skewed cylinders following different flows may include nonuniform geometric properties. Despite this technical challenge, the maximal function will provide with us a tool for overcoming this conceptual difficulty. Instead of taking the average in balls, now we construct a new maximal function that takes the average in the skewed cylinders that are “admissible”.

**Definition 2** (Admissibility, maximal function). Given $\varepsilon > 0$, $x \in \mathbb{R}^d$, $t \in (S + \varepsilon^2, T - \varepsilon^2)$, we define a skewed cylinder $Q_\varepsilon(t,x)$ by Definition 1. For $\eta > 0$, we say $Q_\varepsilon(t,x)$ is *$\eta$-admissible* if

$$\varepsilon^2 \fint_{Q_\varepsilon(t,x)} \mathcal{M}(\nabla u(s))(y)\, \mathrm{d}y\, \mathrm{d}s = \frac{1}{\varepsilon^d |Q_1|} \int_{Q_\varepsilon(t,x)} \mathcal{M}(\nabla u)\, \mathrm{d}y\, \mathrm{d}s < \eta. \tag{2}$$

Here $\mathcal{M}$ is the spatial-only maximal function defined in (1), and with a slight abuse of notation, we also use $|Q_1|$ to represent the $(d+1)$-dimensional space-time Lebesgue measure $\mathcal{L}^{d+1}$ of a cylinder with radius 1. For any locally integrable function $f \in L^1_{\mathrm{loc}}((S,T) \times \mathbb{R}^d)$, for every $(t,x) \in (S,T) \times \mathbb{R}^d$ we define a new maximal function $\mathcal{M}_{\mathcal{Q}}$ by

$$\mathcal{M}_{\mathcal{Q}}(f)(t,x) := \sup_{\varepsilon > 0} \big\{ \fint_{Q_\varepsilon(t,x)} |f(s,y)|\, \mathrm{d}y\, \mathrm{d}s : Q_\varepsilon(t,x) \text{ is } \eta\text{-admissible} \big\}.$$

Note that in the sup we actually need $\varepsilon^2 < \min\{t - S, T - t\}$ to define $Q_\varepsilon(t,x)$, and we will justify in Section 3 that admissible choices of $\varepsilon$ exist for almost every $(t,x)$ so that $\mathcal{M}_{\mathcal{Q}}$ is well defined.

The main result of this paper is the following.

[1]Parabolic scaling – $\varepsilon^2$ in time versus $\varepsilon$ in space – will not be indispensable in this paper. We only employ it because of its applications to the Navier–Stokes equations, but all the results can be generalized to other time-space scalings.

**Theorem 1.** *Let $\eta < \eta_0$ for some small universal constant $\eta_0 > 0$. If $u$ is divergence-free, and $\mathcal{M}(\nabla u) \in L^p((S,T) \times \mathbb{R}^d)$ for some $1 \le p \le \infty$,[2] then $\mathcal{M}_{\mathcal{Q}}$ associated with $\eta$-admissible cylinders generated by $u$ satisfies the following:*

(1) *$\mathcal{M}_{\mathcal{Q}}$ is of strong type $(\infty, \infty)$, i.e. for $f \in L^\infty((S,T) \times \mathbb{R}^d)$, it holds that*

$$\|\mathcal{M}_{\mathcal{Q}} f\|_{L^\infty((S,T)\times\mathbb{R}^d)} \le \|f\|_{L^\infty((S,T)\times\mathbb{R}^d)}.$$

(2) *$\mathcal{M}_{\mathcal{Q}}$ is of weak type $(1,1)$, i.e. for $f \in L^1((S,T) \times \mathbb{R}^d)$, $\lambda > 0$, the Lebesgue measure of the superlevel set satisfies*

$$\mathcal{L}^{d+1}\big(\{(t,x) \in (S,T) \times \mathbb{R}^d : (\mathcal{M}_{\mathcal{Q}} f)(t,x) > \lambda\}\big) \le \frac{C_1}{\lambda}\|f\|_{L^1((S,T)\times\mathbb{R}^d)}.$$

(3) *$\mathcal{M}_{\mathcal{Q}}$ is of strong type $(q,q)$ for any $1 < q < \infty$, i.e. for $f \in L^q((S,T) \times \mathbb{R}^d)$, it holds that*

$$\|\mathcal{M}_{\mathcal{Q}} f\|_{L^q((S,T)\times\mathbb{R}^d)} \le C_q\|f\|_{L^q((S,T)\times\mathbb{R}^d)}.$$

Let us now explain why we are interested in these skewed cylinders and the maximal function related to them. In many scaling-invariant partial differential equations, it is a common technique to zoom in near a point and conduct a local analysis in its neighborhood and use this obtained local information to deduce global results. This form of argument usually consists of two parts: one is a local theorem, which handles the rescaled problem near a point, and the second is a local-to-global step, which contributes to some global information. For instance, the three-dimensional Navier–Stokes equations

$$\partial_t u + u \cdot \nabla u + \nabla P = \Delta u, \quad \operatorname{div} u = 0 \tag{3}$$

are scaling invariant. In particular, $u_\lambda$ and $P_\lambda$ defined by

$$u_\lambda(t,x) = \lambda u(\lambda^2 t, \lambda x), \quad P_\lambda(t,x) = \lambda^2 P(\lambda^2 t, \lambda x)$$

are also solutions to (3). In [3], Caffarelli, Kohn, and Nirenberg investigated the partial regularity of suitable weak solutions to the Navier–Stokes equations by zooming in to a so-called *parabolic cylinder*, where *parabolic* refers to the fact that the spatial scale is $\lambda$ while the temporal scale is $\lambda^2$. They showed that if a suitable solution $u$ satisfies

$$\limsup_{r\to 0} \frac{1}{r}\int_{t-\frac{7}{8}r^2}^{t+\frac{1}{8}r^2}\int_{B_r(x)} |\nabla u(s,y)|^2 \,\mathrm{d}y\,\mathrm{d}s \le \eta$$

for some fixed small $\eta$, then $u$ is regular at $(t,x)$. From this local theorem, they used a covering argument to conclude a global result, that the parabolic measure $\mathcal{P}^1$ of the singular set is zero. This was an improvement from Scheffer's result ([18]) which stated

[2] In the case $1 < p \le \infty$, since $\mathcal{M}$ is a bounded operator on $L^p(\mathbb{R}^d)$, this condition is equivalent to $\nabla u \in L^p((S,T) \times \mathbb{R}^d)$.

the singular set has at most Hausdorff dimension $\frac{5}{3}$. The reason for this improvement is that $\iint |\nabla u|^2 \,\mathrm{d}x\,\mathrm{d}t$ has a stronger scaling than other quantities, which is $\iint |\nabla u_\lambda|^2 \,\mathrm{d}x\,\mathrm{d}t = \frac{1}{\lambda} \iint |\nabla u|^2 \,\mathrm{d}x\,\mathrm{d}t$.

Quantitative global results can also follow from this kind of scaling argument. Choi and Vasseur ([4, 21]) estimated higher derivatives, by locally controlling higher derivatives using the De Giorgi technique applied to quantities with the same strong scaling as $\iint |\nabla u|^2$. In particular, one must avoid using $\iint |u|^{\frac{10}{3}}$, which has a weaker scaling. However, without controlling the flux, the parabolic regularization cannot overcome the nonlinearity. A natural idea would be to utilize the Galilean invariance of Navier–Stokes equations and work in a neighborhood following the flow. Instead of working on parabolic cylinders, they worked on skewed parabolic cylinders as we defined above.

The advantage of using such skewed cylinders is that, by taking out the mean velocity, one can use the velocity gradient to control the velocity in the local study. The maximal function associated with these skewed cylinders then will help us better bridge the local study to global results.

Let us mention that a similar construction also appears in the recent development of convex integration for Euler equations by Isett ([11, 12]) and the subsequent work of Isett and Oh ([13]), where the authors call the mollified flow *coarse scale flow* and skewed cylinders *$u_\varepsilon$-adapted Eulerian cylinders*. The difference from the previous definition is that their apertures of mollification, radii of cylinder bases, and lengths of time spans are chosen differently from here. The purpose is however the same, which is to kill the mean velocity and to obtain dimensionally correct estimates.

Note that Theorem 1 has already been used in [22, Corollary 1] to show the following result.

**Theorem 2.** *Let $u$ be a suitable weak solution to the three-dimensional Navier–Stokes equations* (3) *with initial data $u|_{t=0} = u_0 \in L^2(\mathbb{R}^3)$. For any $q > \frac{4}{3}$, $K \subset\subset (0,T) \times \mathbb{R}^3$, there exists a constant $C_{q,K}$ depending on $q$ and $K$ such that the following holds,*

$$\|\nabla^2 u\|_{L^{\frac{4}{3},q}(K)} \le C_{q,K}\big(\|u_0\|_{L^2(\mathbb{R}^3)}^{\frac{3}{2}} + 1\big).$$

This is an improvement of [7] where the result was shown for $L^q$ with $q < \frac{4}{3}$, and of [15] where it was shown for $L^{\frac{4}{3},\infty}$.

In this paper we provide the first application of Theorem 1 to give an alternative proof for the results of Choi and Vasseur in [4], as an example of using the maximal function to obtain global results from local estimates.

**Theorem 3.** *Let $(u, P)$ be a smooth solution to* (3) *in $(0,T)$ with initial data $u_0 \in L^2$, let $d \ge 1$, $\alpha \in [0,2)$, denote $f = |(-\Delta)^{\frac{\alpha}{2}} \nabla^d u|$, $p = \frac{4}{d+1+\alpha}$. We have*

$$\|f \mathbf{1}_{\{f^p > C_{d,\alpha} t^{-2}\}}\|_{L^{p,\infty}((0,T)\times\mathbb{R}^3)}^p \le C \|u_0\|_{L^2(\mathbb{R}^3)}^2.$$

This paper is organized as follows. Bounds on the maximal function rely on a Vitali-type covering lemma, which is introduced in Section 2, where we define admissible cylinders and prove the covering lemma for them. We use this covering lemma to show some properties of the maximal function in Section 3. Finally, in Section 4 we use the maximal function to give an alternative proof for the higher-derivative estimates for the Navier–Stokes equations.

## 2. Covering lemma

In this section we derive some basic properties of the mollified flows and admissible cylinders, and then use them to prove the covering lemma.

### 2.1. Preliminaries

We first note the following easy pointwise estimate on the mollified velocity gradient.

**Lemma 4** (Pointwise estimate on $\nabla u_\varepsilon$). *For $(t,x) \in (S,T) \times \mathbb{R}^d$, $y \in \mathbb{R}^d$, and $\varepsilon, r > 0$, we have*

$$|\nabla u_\varepsilon(t,x)| \le C\varepsilon^{-d} \|\nabla u(t)\|_{L^1(B_\varepsilon(x))}, \tag{4}$$

$$|\nabla u_\varepsilon(t,x)| \le C\varepsilon^{-d} \Big(\frac{|y-x|}{\varepsilon} + 2\Big)^d \|\mathcal{M}(\nabla u(t))\|_{L^1(B_\varepsilon(y))}, \tag{5}$$

$$|\nabla u_\varepsilon(t,x)| \le C\varepsilon^{-d} \Big(\frac{|y-x| + r + \varepsilon}{r}\Big)^d \|\mathcal{M}(\nabla u(t))\|_{L^1(B_r(y))}. \tag{6}$$

*Proof.* The first estimate follows easily from the scaling that

$$\nabla u_\varepsilon(t,x) = \int_{\mathbb{R}^d} \nabla u(t, x-y)\varphi_\varepsilon(y)\,\mathrm{d}y \le \|\nabla u(t)\|_{L^1(B_\varepsilon(x))}\|\varphi_\varepsilon\|_{L^\infty}.$$

This indicates that by controlling the average of $\nabla u$ in a small ball $B_\varepsilon(x)$, we can control the size of the mollified gradient at the center $x$. To control the mollified gradient from elsewhere, we need a maximal function to gather nonlocal information. For any $x' \in B_\varepsilon(x)$, $y' \in B_r(y)$, we have $|y'-x'| \le |y-x| + r + \varepsilon =: Kr$, so $B_\varepsilon(x) \subset B_{Kr}(y')$ and the integral of $\nabla u$ can be bounded by

$$\begin{aligned}\int_{B_\varepsilon(x)} |\nabla u(t,z)|\,\mathrm{d}z \le \int_{B_{Kr}(y')} |\nabla u(t,z)|\,\mathrm{d}z &= |B_{Kr}(y')| \fint_{B_{Kr}(y')} |\nabla u(t,z)|\,\mathrm{d}z \\ &\le K^d |B_r| \mathcal{M}(\nabla u(t))(y').\end{aligned}$$

Since the above holds for any $y' \in B_r(y)$, by taking the average of the right-hand side in $B_r(y)$ we have

$$\begin{aligned}\|\nabla u(t)\|_{L^1(B_\varepsilon(x))} &\le K^d |B_r| \fint_{B_r(y)} \mathcal{M}(\nabla u(t))(y')\,\mathrm{d}y' \\ &= K^d \|\mathcal{M}(\nabla u(t))\|_{L^1(B_r(y))}.\end{aligned} \tag{7}$$

This bound and estimate (4) yield the third estimate, and the second estimate is a special case of the third when $r = \varepsilon$. ∎

As can be seen here, $\mathcal{M}(\nabla u)$ controls how mollified velocities alter in space. This observation motivates us to introduce the notion of admissibility in Definition 2. Let us provide a heuristic explanation for the choice of homogeneity in (2). Consider two skewed cylinders, both with radii of order $\varepsilon$, starting at the same time with distance also of order $\varepsilon$. If $\nabla u$ is of order $\varepsilon^{-2}\eta$, then their velocities roughly differ by $\varepsilon^{-1}\eta$, so in a time span of length $\varepsilon^2$, they at most diverge $\varepsilon\eta$ further away, so their distance will remain of order $\varepsilon$. This ensures cylinders do not deviate relatively too far away and will be crucial in the covering lemma.

**Remark 1.** For $1 < p < \infty$, (2) is weaker than the $L^p$ analogue

$$\varepsilon^2 \left( \fint_{Q_\varepsilon(t,x)} \mathcal{M}(|\nabla u|^p)\, \mathrm{d}y\, \mathrm{d}s \right)^{\frac{1}{p}} < \eta.$$

This is because Jensen's inequality implies that

$$\left( \fint_{Q_\varepsilon(t,x)} \mathcal{M}(\nabla u)\, \mathrm{d}y\, \mathrm{d}s \right)^p \le \fint_{Q_\varepsilon(t,x)} [\mathcal{M}(\nabla u)]^p\, \mathrm{d}y\, \mathrm{d}s$$

and

$$[\mathcal{M}(\nabla u)]^p(x) = \sup_{r>0} \left( \fint_{B_r(x)} |\nabla u|\, \mathrm{d}y \right)^p \le \sup_{r>0} \fint_{B_r(x)} |\nabla u|^p\, \mathrm{d}y = [\mathcal{M}(|\nabla u|^p)](x).$$

Next, let us discuss the trajectories of the mollified flow that passes through an admissible cylinder.

**Lemma 5.** *There exists a universal constant $\eta_1 > 0$ such that the following is true. Given $\varepsilon > 0$, $t_0 \in (S + \varepsilon^2, T - \varepsilon^2)$, and $x_0 \in \mathbb{R}^d$, suppose $Q_\varepsilon(t_0, x_0)$ is $\eta$-admissible as defined in Definition 2 with $\eta < \eta_1$. For any $(t_*, x_*) \in Q_\varepsilon(t_0, x_0)$, we have*

$$|X_\varepsilon(t_*, x_*; t) - X_\varepsilon(t_0, x_0; t)| \le 2\varepsilon \tag{8}$$

*at any given time $t \in (t_0 - \varepsilon^2, t_0 + \varepsilon^2)$.*

*Proof.* To ease the notation, we denote

$$X^*(t) := X_\varepsilon(t_*, x_*; t), \quad X^0(t) := X_\varepsilon(t_0, x_0; t), \quad \Delta X(t) := X^*(t) - X^0(t);$$

thus we need to show $|\Delta X(t)| \le 2\varepsilon$. We argue by contradiction and suppose $|\Delta X(s_*)| > 2\varepsilon$ at some $s_* \in (S^\alpha, T^\alpha)$. Without loss of generality, suppose $s_* > t_*$. Note that

$$|\Delta X(t_*)| = |X^*(t_*) - X^0(t_*)| = |x_* - X_\varepsilon(t_0, x_0; t_*)| < \varepsilon < 2\varepsilon$$

because $(t_*, x_*) \in Q_\varepsilon(t_0, x_0)$. Since $\Delta X$ is absolute continuous, there must exist an $r_* \in (t_*, s_*)$ such that

$$|\Delta X(t)| \le 2\varepsilon \text{ for any } t \in [t_*, r_*], \quad |\Delta X(r_*)| = 2\varepsilon. \tag{9}$$

For almost every $t \in [t_*, r_*]$, the growth rate of the difference $\Delta X$ can be bounded by

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Big|\Delta X(t)\Big| \le \Big|\frac{\mathrm{d}}{\mathrm{d}t}\Delta X(s)\Big| &= |\dot{X}^*(t) - \dot{X}^0(t)| \\ &= |u_\varepsilon(t, X^*(t)) - u_\varepsilon(t, X^0(t))| \\ &\le |\nabla u_\varepsilon(t, \xi_t)||\Delta X(t)|\end{aligned}$$

for some $\xi_t$ between $X^*(t)$ and $X^0(t)$. We can bound the gradient term by

$$\begin{aligned}|\nabla u_\varepsilon(t, \xi_t)| &\le C\varepsilon^{-d}\Big(\frac{|\xi_t - X^0(t)|}{\varepsilon} + 2\Big)^d \|\mathcal{M}(\nabla u(t))\|_{L^1(B_\varepsilon(X^0(t)))} \\ &\le C\varepsilon^{-d}\Big(\frac{|\Delta X(t)|}{\varepsilon} + 2\Big)^d \|\mathcal{M}(\nabla u(t))\|_{L^1(B_\varepsilon(X^0(t)))},\end{aligned}$$

using (5) for $x = \xi_t$ and $y = X^0(t)$. By (9), $|\Delta X(t)| \le 2\varepsilon$ for any $t \in [t_*, r_*]$, so in the above coefficient $C(\frac{|\Delta X(t)|}{\varepsilon} + 2)^d \le C(2+2)^d = C$, thus for almost every $t \in [t_*, r_*]$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}|\Delta X(t)| \le \frac{C}{\varepsilon^d}\|\mathcal{M}(\nabla u(t))\|_{L^1(B_\varepsilon(X^0(t)))}|\Delta X(t)|.$$

By Grönwall's inequality, we reach the conclusion that

$$\begin{aligned}|\Delta X(r_*)| &\le |\Delta X(t_*)| \exp\Big(\int_{t_*}^{r_*} \frac{C}{\varepsilon^d}\|\mathcal{M}(\nabla u(t))\|_{L^1(B_\varepsilon(X^0(t)))}\,\mathrm{d}t\Big) \\ &\le \varepsilon \exp\Big(\int_{t_0-\varepsilon^2}^{t_0+\varepsilon^2} \frac{C}{\varepsilon^d}\|\mathcal{M}(\nabla u(t))\|_{L^1(B_\varepsilon(X^0(t)))}\,\mathrm{d}t\Big) \\ &= \varepsilon \exp\Big(\frac{C}{\varepsilon^d}\|\mathcal{M}(\nabla u)\|_{L^1(Q_\varepsilon(t_0,x_0))}\Big) \\ &\le \varepsilon \exp(C\eta),\end{aligned}$$

which contradicts (9) when choosing $\eta < \eta_1 = \frac{1}{C}\log 2$. ■

To conclude this subsection, we discuss two streamlines with different $\varepsilon$ that start from the same location. Before that, we introduce some notation. Let $\alpha$ be an index. Given $\varepsilon_\alpha > 0$, $t^\alpha \in (S + {\varepsilon_\alpha}^2, T - {\varepsilon_\alpha}^2)$, $x^\alpha \in \mathbb{R}^d$, we abbreviate

$$\begin{aligned}X^\alpha(t) &:= X_{\varepsilon_\alpha}(t^\alpha, x^\alpha; t), \qquad B^\alpha(t) := B_{\varepsilon_\alpha}(X^\alpha(t)) \subset \mathbb{R}^d, \\ S^\alpha &:= t^\alpha - {\varepsilon_\alpha}^2, \qquad T^\alpha := t^\alpha + {\varepsilon_\alpha}^2, \\ Q^\alpha &:= Q_{\varepsilon_\alpha}(t^\alpha, x^\alpha) = \big\{(t,x) : S^\alpha < t < T^\alpha,\ x \in B^\alpha(t)\big\}.\end{aligned} \tag{10}$$

For $\lambda > 0$, we denote the spatial dilation of a cylinder $Q^\alpha$ by

$$\lambda Q^\alpha := \big\{(t,x) : S^\alpha < t < T^\alpha,\ x \in \lambda B^\alpha(t) = B_{\lambda\varepsilon_\alpha}(X^\alpha(t))\big\}. \tag{11}$$

Notice that different from upright cylinders or cubes, for $\varepsilon_1 < \varepsilon_2$, it is not known that $Q_{\varepsilon_1}(t,x) \subset Q_{\varepsilon_2}(t,x)$, because their center streamlines $X_{\varepsilon_{1,2}}$ solve different equations. As we will see later, this lack of monotonicity only poses a minor technical difficulty. For the same reason, note that $\lambda Q_\varepsilon(t,x) \neq Q_{\lambda\varepsilon}(t,x)$, and neither is necessarily contained in the other.

**Lemma 6.** *Recall that $\eta_1$ is a universal constant defined in Lemma 5. There exists a universal constant $\eta_0 < \eta_1$ such that the following is true. Given $\varepsilon_\alpha > \frac{1}{2}\varepsilon_\beta > 0$, $t^\alpha \in (S + {\varepsilon_\alpha}^2, T - {\varepsilon_\alpha}^2)$, $t^\beta \in (S + {\varepsilon_\beta}^2, T - {\varepsilon_\beta}^2)$, and $x^\alpha, x^\beta \in \mathbb{R}^d$, suppose $Q^\alpha = Q_{\varepsilon_\alpha}(t^\alpha, x^\alpha)$, $Q^\beta = Q_{\varepsilon_\beta}(t^\beta, x^\beta)$ are $\eta$-admissible as defined in Definition 2 with $\eta < \eta_0$. For any $(t_*, x_*) \in Q^\alpha \cap Q^\beta$, we have*

$$|X_{\varepsilon_\beta}(t_*, x_*; t) - X_{\varepsilon_\alpha}(t_*, x_*; t)| \le \varepsilon_\alpha \tag{12}$$

*at any given time $t \in (S^\alpha, T^\alpha) \cap (S^\beta, T^\beta)$.*

*Proof.* Denote

$$X^1(t) = X_{\varepsilon_\alpha}(t_*, x_*; t), \quad X^2(t) = X_{\varepsilon_\beta}(t_*, x_*; t), \quad \Delta X(t) = X^1(t) - X^2(t);$$

thus we need to show $|\Delta X(t)| \le \varepsilon_\alpha$. Note that

$$\Delta X(t_*) = X^2(t_*) - X^1(t_*) = X_{\varepsilon_\beta}(t_*, x_*; t_*) - X_{\varepsilon_\alpha}(t_*, x_*; t_*) = x_* - x_* = 0.$$

Similarly to the last lemma, we argue by contradiction and suppose there exists $r_* \in (t_*, \min\{T^\alpha, T^\beta\})$ such that

$$|\Delta X(t)| \le \varepsilon_\alpha \text{ for any } t \in [t_*, r_*], \quad |\Delta X(r_*)| = \varepsilon_\alpha. \tag{13}$$

For almost every $t \in [t_*, r_*]$, the time derivative of $\Delta X$ is calculated as

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Delta X(t) &= \dot{X}^2(t) - \dot{X}^1(t) \\
&= u_{\varepsilon_\beta}(t, X^2(t)) - u_{\varepsilon_\alpha}(s, X^1(t)) \\
&= u_{\varepsilon_\beta}(t, X^2(t)) - u_{\varepsilon_\alpha}(s, X^2(t)) + u_{\varepsilon_\alpha}(s, X^2(t)) - u_{\varepsilon_\alpha}(s, X^1(t)) \\
&= \int_{\varepsilon_\alpha}^{\varepsilon_\beta} \frac{\partial}{\partial \varepsilon} u_\varepsilon(t, X^2(t))\,\mathrm{d}\varepsilon + u_{\varepsilon_\alpha}(t, X^2(t)) - u_{\varepsilon_\alpha}(t, X^1(t)).
\end{aligned} \tag{14}$$

We will use $Q^\beta$ to control the first integral term and use $Q^\alpha$ to control the rest. Note that

$$\frac{\partial}{\partial \varepsilon} u_\varepsilon(t, x) = \frac{\partial}{\partial \varepsilon} \int_{\mathbb{R}^d} u(t, x - \varepsilon y)\varphi(y)\,\mathrm{d}y = \int_{\mathbb{R}^d} \nabla_x u(t, x - \varepsilon y) \cdot -y\varphi(y)\,\mathrm{d}y;$$

thus we can control its absolute value by

$$\begin{aligned}
\left|\frac{\partial}{\partial \varepsilon} u_\varepsilon(t, x)\right| &\le \varepsilon^{-d} \|\nabla u(t)\|_{L^1(B_\varepsilon(x))} \|y\varphi(y)\|_{L^\infty} \\
&= C\varepsilon^{-d} \|\nabla u(t)\|_{L^1(B_\varepsilon(x))} \\
&\le C\varepsilon^{-d} \left(\frac{|x - X^\beta(t)| + \varepsilon_\beta + \varepsilon}{\varepsilon_\beta}\right)^d \|\mathcal{M}(\nabla u(t))\|_{L^1(B_{\varepsilon_\beta}(X^\beta(t)))} \\
&= C{\varepsilon_\beta}^{-d} \left(\frac{|x - X^\beta(t)| + \varepsilon_\beta + \varepsilon}{\varepsilon}\right)^d \|\mathcal{M}(\nabla u(t))\|_{L^1(B^\beta(t))}.
\end{aligned} \tag{15}$$

Here, to control $\|\nabla u(t)\|_{L^1(B_\varepsilon(x))}$ we use (7) with $r = \varepsilon_\beta$ and $y = X^\beta(t)$ in the last inequality. Thanks to Lemma 5, $|X^2(t) - X^\beta(t)| \le 2\varepsilon_\beta$. Since $\varepsilon$ is between $\varepsilon_\beta$ and $\varepsilon_\alpha > \frac{1}{2}\varepsilon_\beta$, we have

$$\frac{|X^2(t) - X^\beta(t)| + \varepsilon_\beta + \varepsilon}{\varepsilon} \le \frac{3\varepsilon_\beta + \varepsilon}{\varepsilon} \le 7.$$

Hence, if we set $x = X^2(t)$ in (15), we would get

$$\left|\frac{\partial}{\partial\varepsilon} u_\varepsilon(t, X^2(t))\right| \le C{\varepsilon_\beta}^{-d} \|\mathcal{M}(\nabla u(t))\|_{L^1(B^\beta(t))},$$

thus we can bound the $\partial_\varepsilon$ term in (14) by

$$\begin{aligned}\left|\int_{\varepsilon_\alpha}^{\varepsilon_\beta} \frac{\partial}{\partial\varepsilon} u_\varepsilon(t, X^2(t))\,\mathrm{d}\varepsilon\right| &\le C|\varepsilon_\beta - \varepsilon_\alpha|{\varepsilon_\beta}^{-d} \|\mathcal{M}(\nabla u(t))\|_{L^1(B^\beta(t))} \\ &\le C\varepsilon_\alpha{\varepsilon_\beta}^{-d} \|\mathcal{M}(\nabla u(t))\|_{L^1(B^\beta(t))}.\end{aligned}$$

The remaining terms in (14) can be bounded similarly to Lemma 5 as

$$\begin{aligned}|u_{\varepsilon_\alpha}(t, X^2(t)) - u_{\varepsilon_\alpha}(t, X^1(t))| &\le |\nabla u_{\varepsilon_\alpha}(t, \xi_t)|\,|\Delta X(t)| \\ &\le C{\varepsilon_\alpha}^{-d} \|\mathcal{M}(\nabla u(t))\|_{L^1(B^\alpha(t))} |\Delta X(t)|.\end{aligned}$$

Combining these two bounds in (14), for almost every $t \in [t_*, r_*]$, the growth rate of $\Delta X$ is bounded by

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}|\Delta X(t)| \le{}& C\big({\varepsilon_\beta}^{-d} \|\mathcal{M}(\nabla u(t))\|_{L^1(B^\beta(t))} + {\varepsilon_\alpha}^{-d} \|\mathcal{M}(\nabla u(t))\|_{L^1(B^\alpha(t))}\big) \\ &\times (\varepsilon_\alpha + |\Delta X(t)|).\end{aligned}$$

By Grönwall's inequality, we would reach

$$\begin{aligned}\varepsilon_\alpha + |\Delta X(r_*)| \le{}& (\varepsilon_\alpha + |\Delta X(t_*)|) \exp\bigg(C\int_{t_*}^{r_*} {\varepsilon_\beta}^{-d} \|\mathcal{M}(\nabla u(t))\|_{L^1(B^\beta(t))} \\ &\qquad + {\varepsilon_\alpha}^{-d} \|\mathcal{M}(\nabla u(t))\|_{L^1(B^\alpha(t))}\,\mathrm{d}t\bigg) \\ \le{}& \varepsilon_\alpha \exp(2C\eta),\end{aligned} \tag{16}$$

which contradicts (13) when choosing $\eta < \eta_0 = \min\{\eta_1, \frac{1}{2C}\log 2\}$. ∎

### 2.2. Covering lemma for admissible cylinders

The goal of this section is to prove a Vitali-type covering lemma for $\eta$-admissible cylinders, provided $\eta < \eta_0$. The key ingredient is Proposition 7, which shows that if two cylinders intersect, then during their shared life span, they are uniformly close to each other. Based on this proposition, we conclude in Lemma 8 that for an $\eta$-admissible cylinder $Q^\alpha$, the union of all $\eta$-admissible cylinders with comparable or lower radius that intersect $Q^\alpha$ has a comparable total measure as $Q^\alpha$. The covering lemma will be a consequence of Lemma 8.

Throughout this subsection, we employ the notation introduced in (10).

**Proposition 7.** *For any pair of intersecting $\eta$-admissible cylinders $Q^\alpha$, $Q^\beta$ as in (10) with $\varepsilon_\beta < 2\varepsilon_\alpha$ and $\eta < \eta_0$ chosen in Lemma 6, at any $t \in (S^\alpha, T^\alpha) \cap (S^\beta, T^\beta)$, we have $B^\beta(t) \subset 9B^\alpha(t)$.*

That is, if $Q^\alpha$ intersects $Q^\beta$ with $\varepsilon_\beta < 2\varepsilon_\alpha$, then $Q^\beta \cap \{S^\alpha < t < T^\alpha\} \subset 9Q^\alpha$. Recall that $\lambda Q^\alpha$ is the spatial dilation defined in (11). The proof is based on Lemmas 5 and 6 which control the trajectories at the level of $Q_\varepsilon$. See Figure 1 for our strategy.

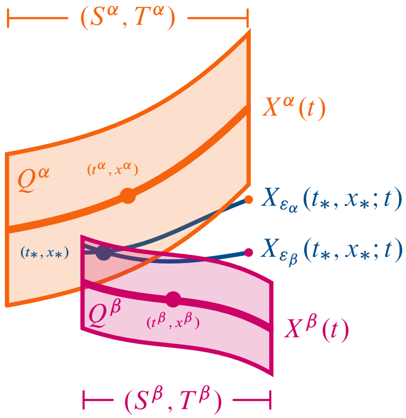


**Figure 1.** $Q^\alpha$ and $Q^\beta$ intersect.

*Proof of Proposition 7.* Let $\eta_0$ be chosen as in Lemma 6. Fix some $(t_*, x_*) \in Q^\alpha \cap Q^\beta$. For any $(t, x) \in Q^\beta$ with $S^\alpha < t < T^\alpha$, we apply the triangle inequality to estimate

$$\begin{aligned} |x - X^\alpha(t)| \le\ & |x - X^\beta(t)| \\ & + |X^\beta(t) - X_{\varepsilon_\beta}(t_*, x_*; t)| \\ & + |X_{\varepsilon_\beta}(t_*, x_*; t) - X_{\varepsilon_\alpha}(t_*, x_*; t)| \\ & + |X_{\varepsilon_\alpha}(t_*, x_*; t) - X^\alpha(t)| \\ \le\ & \varepsilon_\beta + 2\varepsilon_\beta + \varepsilon_\alpha + 2\varepsilon_\alpha. \end{aligned}$$

Here, the first term is because $x \in B^\beta(t)$, the second and the fourth are due to Lemma 5, and the third term is controlled by Lemma 6. Since $\varepsilon_\beta < 2\varepsilon_\alpha$, we have

$$|x - X^\alpha(t)| < 9\varepsilon_\alpha. \qquad \blacksquare$$

We remark here that if we take a sharper estimate in each step of Lemmas 5 and 6 (and require a smaller $\eta$), the factor 9 can be easily improved to $5 + \delta$ for any $\delta > 0$. The factor 5 is also the one that appeared in the original Vitali covering lemma for balls. Recall that in the proof of the Vitali lemma, an important reason why we get a comparable volume is because if two balls $B_{r_1}(x_1) \cap B_{r_2}(x_2) \neq \varnothing$ with $r_2 < 2r_1$, then $B_{r_2}(x_2) \subset 5B_{r_1}(x_1)$. Unfortunately, this geometric property cannot be realized in our case, because an admissible cylinder with (2) has no control over the past and future velocities. As a consequence, it is unlikely to cover $Q^\beta$ by dilation of $Q^\alpha$ in space-time. However, this requirement can be relaxed as the following. See [20, Section 1.1] for a more general setting.

**Lemma 8.** *Given a fixed $Q^\alpha$ and a family of $\{Q^\beta\}_{\beta\subset\Lambda}$ as in (10) such that for each $Q^\beta$, $Q^\alpha \cap Q^\beta \neq \varnothing$, $\varepsilon_\beta < 2\varepsilon_\alpha$, and they are $\eta$-admissible for $\eta < \eta_0$. Let $Q^\alpha_* = \bigcup_{\beta\in\Lambda} Q^\beta$ denote the union of this family. Then there exists a universal constant $C$ such that*

$$|Q^\alpha_*| \le C|Q^\alpha|.$$

*Proof.* Without loss of generality, we may assume that $\{Q^\beta\}_{\beta\subset\Lambda}$ is a finite collection. The general case can be proven using the finite case. Note that each $Q^\beta$ is an open set. For any compact subset $K \subset\subset Q^\alpha_*$, $K$ admits a finite open cover, thus $|K| \le C|Q^\alpha|$ using the finite case. Since the inequality holds for any compact subset $K$, it must also be true for $Q^\alpha_*$.

For each $Q^\beta$, we can break it into $Q^\beta = Q^\beta_+ \cup Q^\beta_- \cup Q^\beta_\circ$, where

$$\begin{aligned} Q^\beta_+ &= Q^\beta \cap \{t \ge T^\alpha\},\\ Q^\beta_- &= Q^\beta \cap \{t \le S^\alpha\},\\ Q^\beta_\circ &= Q^\beta \cap \{S^\alpha < t < T^\alpha\}. \end{aligned}$$

From Proposition 7, we can conclude that

$$\bigcup_{\beta\in\Lambda} Q^\beta_\circ \subset 9Q^\alpha \Rightarrow \left|\bigcup_{\beta\in\Lambda} Q^\beta_\circ\right| \le 9^d|Q^\alpha|. \tag{17}$$

As mentioned in the remark, we cannot bound the size of $\bigcup_{\beta\in\Lambda} Q^\beta_+$ or $\bigcup_{\beta\in\Lambda} Q^\beta_-$ directly by $Q^\alpha$, as their center streamlines can diverge away from $X^\alpha$ after $T^\alpha$. Fortunately, we do not need them to be close to $X^\alpha$, as long as we can show they remain a small distance from each other.

Let us measure $\bigcup_{\beta\in\Lambda} Q^\beta_+$. First, we group the cylinders by their radii. Denote

$$\Lambda_i = \left\{\beta \in \Lambda : 2^{-i}\varepsilon_\alpha \le \varepsilon_\beta < 2^{-i+1}\varepsilon_\alpha\right\}. \tag{18}$$

Because each $\varepsilon_\beta < 2\varepsilon_\alpha$, we have $\Lambda = \bigcup_{i\in\mathbb{N}} \Lambda_i$, hence we can write the union as

$$\bigcup_{\beta\in\Lambda} Q^\beta_+ = \bigcup_{i\in\mathbb{N}} \bigcup_{\beta\in\Lambda_i} Q^\beta_+. \tag{19}$$

Now we fix $i$ and estimate the size of $\bigcup_{\beta\in\Lambda_i} Q^\beta_+$. Clearly we can disregard the empty ones, and assume $T^\alpha < T^\beta$ for each $\beta \in \Lambda_i$. To begin with, set $\mathcal{Q}^{(0)}_i = \{Q^\beta_+\}_{\beta\in\Lambda_i}$. Then we repeat the following two steps: at the $j$th iteration ($j \ge 1$), (Step 1) select some $\beta_j$ such that $T^{\beta_j} = \max\{T^\beta : Q_+^\beta \in \mathcal{Q}_i^{(j-1)}\}$;

(Step 2) from $\mathcal{Q}_i^{(j-1)}$ we remove any $Q_+^\beta$ such that $B^\beta(T^\alpha) \cap B^{\beta_j}(T^\alpha) \neq \varnothing$, and denote the rest by $\mathcal{Q}_i^{(j)}$.

After finitely many iterations, $\mathcal{Q}_i^{(n+1)}$ will be empty, and we have a list of $Q_+^{\beta_1}, \dots, Q_+^{\beta_n}$. We claim that

$$\bigcup_{\beta \in \Lambda_i} Q_+^\beta \subset \bigcup_{j=1}^n 9Q_+^{\beta_j}. \tag{20}$$

To see why this is true, take any $Q_+^\beta \in \mathcal{Q}_i^{(0)}$. It must have been removed from $\mathcal{Q}_i^{(j-1)}$ at some step $j$ in the above process. This implies $B^\beta(T^\alpha) \cap B^{\beta_j}(T^\alpha) \neq \varnothing$, and $T^\beta \le T^{\beta_j}$. Also, we have $\varepsilon_\beta < 2\varepsilon_{\beta_j}$, which is actually true for any pair of cylinders by our selection of $\Lambda_i$ according to (18). Therefore, by Proposition 7 we have $B^\beta(t) \subset 9B^{\beta_j}(t)$ at any $t \in (S^\beta, T^\beta) \cap (S^{\beta_j}, T^{\beta_j})$. Because $S^\beta, S^{\beta_j} \le T^\alpha \le T^\beta \le T^{\beta_j}$, we have $Q_+^\beta \subset 9Q_+^{\beta_j}$ and this proves claim (20).

Note that by our construction, $\{B^{\beta_j}(T^\alpha)\}_{j=1}^n$ are pairwise disjoint, and they are all inside $9B^\alpha(T^\alpha)$ by Proposition 7. Therefore their total measure is

$$\begin{aligned}
\sum_{j=1}^n |Q_+^{\beta_j}| &\le \sum_{j=1}^n |B^{\beta_j}(T^\alpha)| \cdot 2(\varepsilon_{\beta_j})^2 \\
&\le \sum_{j=1}^n 2 \cdot |B^{\beta_j}(T^\alpha)| \cdot (2^{-i+1}\varepsilon_\alpha)^2 \\
&= 2 \cdot 4^{-i+1} {\varepsilon_\alpha}^2 \left| \bigcup_{j=1}^n B^{\beta_j}(T^\alpha) \right| \\
&\le 2 \cdot 4^{-i+1} {\varepsilon_\alpha}^2 |9B^\alpha(T^\alpha)| = 4^{-i+1} \cdot 9^d |Q^\alpha|.
\end{aligned}$$

Combining with claim (20), we have

$$\left| \bigcup_{\beta \in \Lambda_i} Q_+^\beta \right| \le \left| \bigcup_{j=1}^n 9Q_+^{\beta_j} \right| \le 9^d \sum_{j=1}^n |Q_+^{\beta_j}| \le 4^{-i+1} \cdot 9^{2d} |Q^\alpha|.$$

Finally, take the summation over $i$, and (19) yields

$$\left| \bigcup_{\beta \in \Lambda} Q_+^\beta \right| \le \sum_{i=0}^\infty \left| \bigcup_{\beta \in \Lambda_i} Q_+^\beta \right| \le \sum_{i=0}^\infty 4^{-i+1} \cdot 9^{2d} |Q^\alpha| = \frac{16}{3} \cdot 9^{2d} |Q^\alpha|.$$

The same proof also applies to $\bigcup_{\beta \in \Lambda} Q_-^\beta$. Therefore, together with estimate (17), we have proven that

$$\begin{aligned}
|Q_*^\alpha| = \left| \bigcup_{\beta \in \Lambda} Q^\beta \right| &\le \left| \bigcup_{\beta \in \Lambda} Q_+^\beta \right| + \left| \bigcup_{\beta \in \Lambda} Q_-^\beta \right| + \left| \bigcup_{\beta \in \Lambda} Q_\circ^\beta \right| \\
&\le \frac{16}{3} \cdot 9^{2d} |Q^\alpha| + \frac{16}{3} \cdot 9^{2d} |Q^\alpha| + 9^d |Q^\alpha| = C|Q^\alpha|. \qquad \blacksquare
\end{aligned}$$

We are finally ready to show the Vitali-type covering lemma.

**Proposition 9** (Covering lemma). *Let $\mathcal{A}$ be an index set and let*

$$\mathcal{Q} = \left\{Q^\alpha = Q_{\varepsilon_\alpha}(t^\alpha, x^\alpha) : \alpha \in \mathcal{A}\right\}$$

*be a collection of $\eta$-admissible cylinders, where $\eta < \eta_0$ is defined in Lemma 6 and $\varepsilon_\alpha$ are uniformly bounded. Then there is a pairwise disjoint subcollection $\mathcal{P} = \{Q^{\alpha_1}, Q^{\alpha_2}, \ldots, Q^{\alpha_n}, \ldots\}$ (finite or infinite) such that*

$$\sum_j |Q^{\alpha_j}| \geq \frac{1}{C}\left|\bigcup_{\alpha\in\mathcal{A}} Q^\alpha\right|,$$

*where $C$ is a universal constant.*

*Proof.* With the help of the previous lemma, the proof of the covering lemma is the same as the classical one in [19]. We select the subcollection $\mathcal{P}$ by the following procedure. To begin with, set $\mathcal{Q}^{(0)} = \mathcal{Q}$. Then repeat the following two steps: at the $j$th iteration ($j \geq 1$),

(Step 1) select some $\alpha_j$ such that $\varepsilon_{\alpha_j} > \frac{1}{2}\sup_{Q^\alpha\in\mathcal{Q}^{(j-1)}}\{\varepsilon_\alpha\}$;

(Step 2) from $\mathcal{Q}^{(j-1)}$ we remove any $Q^\alpha$ that intersects with $Q^{\alpha_j}$, and denote the rest by $\mathcal{Q}^{(j)}$.

This procedure may stop after a certain step if $\mathcal{Q}^{(n+1)} = \varnothing$, or it can continue indefinitely. We denote the chosen ones by $\mathcal{P} := \{Q^{\alpha_1}, \ldots, Q^{\alpha_n}, \ldots\}$ (finite or infinite). They are pairwise disjoint due to our strategy.

Suppose that $\sum_j |Q^{\alpha_j}| < \infty$; otherwise the conclusion is automatically true. Thus either $\mathcal{P}$ is a finite collection, or $\mathcal{P}$ is infinite and $\varepsilon_{\alpha_j} \to 0$ as $j \to \infty$. In either case, each $Q^\alpha$ must be removed from $\mathcal{Q}^{(j)}$ at some iteration. Otherwise, we would have $Q^\alpha \in \mathcal{Q}^{(j-1)}$ for all $j$, then Step 1 would imply that $\varepsilon_{\alpha_j} > \frac{1}{2}\varepsilon_\alpha$ for all $j$, thus the sequence $\varepsilon_{\alpha_j}$ cannot converge to zero. Now suppose $Q^\alpha \in \mathcal{Q}^{(j-1)} \setminus \mathcal{Q}^{(j)}$; then we have $Q^\alpha \cap Q^{\alpha_j} \neq \varnothing$, and $\varepsilon_\alpha < 2\varepsilon_{\alpha_j}$. This implies

$$Q^\alpha \subset Q_*^{\alpha_j} := \bigcup_{\alpha\in\mathcal{A}}\left\{Q^\alpha \in \mathcal{Q} : \varepsilon_\alpha < 2\varepsilon_{\alpha_j},\ Q^\alpha \cap Q^{\alpha_j} \neq \varnothing\right\}.$$

Thus $\bigcup_{\alpha\in\mathcal{A}} Q^\alpha \subset \bigcup_{j=1}^n Q_*^{\alpha_j}$, and finally we control the measure of the union by

$$\left|\bigcup_{\alpha\in\mathcal{A}} Q^\alpha\right| \leq \left|\bigcup_{j=1}^n Q_*^{\alpha_j}\right| \leq \sum_{j=1}^n |Q_*^{\alpha_j}| \leq C\sum_{j=1}^n |Q^{\alpha_j}|,$$

thanks to Lemma 8. ∎

## 3. Construction of the maximal function

In this section we use the covering lemma to generalize some results from classical harmonic analysis to our situation. First, we confirm the existence of $\eta$-admissible cylinders centering almost everywhere under some assumptions on $u$. Then we prove the main theorem for the maximal function on these skewed cylinders and show related results similar to the classical case.

### 3.1. Existence of admissible cylinders

To begin with, we need some assumptions to guarantee the existence of $\eta$-admissible cylinders centering almost everywhere, which are the following. For the entire Section 3 we make the following assumptions.

**Assumption 10.** *For some* $1 \leq p \leq \infty$,

(1) $\mathcal{M}(\nabla u) \in L^p((S,T) \times \mathbb{R}^d)$;

(2) $\operatorname{div} u = 0$.

**Proposition 11.** *Let* $\eta > 0$. *For almost every* $(t,x) \in (S,T) \times \mathbb{R}^d$, $Q_\varepsilon(t,x)$ *is* $\eta$*-admissible for sufficiently small* $\varepsilon$ *(depending on* $(t,x)$*). Moreover, we have*

$$\lim_{\varepsilon \to 0} \operatorname{diam}(Q_\varepsilon(t,x)) = 0,$$

*where* diam *refers to the* $(d+1)$*-dimensional diameter.*

Before showing the proof of Proposition 11, we first give a general lemma on the $L^1$ boundedness of the map $f \mapsto f_\varepsilon$ defined below. Given $f \in L^1_{\text{loc}}((S,T) \times \mathbb{R}^d)$, for $x \in \mathbb{R}^d$, $t \in (S,T)$, $\varepsilon > 0$, we define

$$f_\varepsilon(t,x) = \begin{cases} \displaystyle⨍_{Q_\varepsilon(t,x)} f(s,y)\,\mathrm{d}y\,\mathrm{d}s, & t \in (S+\varepsilon^2, T-\varepsilon^2), \\ 0, & t \in (S, S+\varepsilon^2] \cup [T-\varepsilon^2, T). \end{cases} \tag{21}$$

Then we have the following bound on $f_\varepsilon$.

**Lemma 12** ($L^1$ boundedness). *Given* $f \in L^1((S,T) \times \mathbb{R}^d)$ *we have*

$$\|f_\varepsilon\|_{L^1((S,T)\times\mathbb{R}^d)} \leq \|f\|_{L^1((S,T)\times\mathbb{R}^d)}.$$

*Proof.* A direct computation gives

$$\begin{aligned} &\int_{S+\varepsilon^2}^{T-\varepsilon^2} \int_{\mathbb{R}^d} |f_\varepsilon(t,x)|\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_{S+\varepsilon^2}^{T-\varepsilon^2} \int_{\mathbb{R}^d} \frac{1}{|Q_\varepsilon|} \left| \int_{Q_\varepsilon(t,x)} f(s,y)\,\mathrm{d}y\,\mathrm{d}s \right| \mathrm{d}x\,\mathrm{d}t \\ &\leq \frac{1}{|Q_\varepsilon|} \int_S^T \int_{\mathbb{R}^d} \int_{S+\varepsilon^2}^{T-\varepsilon^2} \int_{\mathbb{R}^d} |f(s,y)| \mathbf{1}_{\{(s,y)\in Q_\varepsilon(t,x)\}}\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}s \\ &= \frac{1}{|Q_\varepsilon|} \int_S^T \int_{\mathbb{R}^d} |f(s,y)| \mathcal{L}^{d+1}(\widetilde{Q}_\varepsilon(s,y))\,\mathrm{d}y\,\mathrm{d}s, \end{aligned} \tag{22}$$

where we define, for any fixed $(s, y) \in (S, T) \times \mathbb{R}^d$, the *dual Lagrangian cylinder* by (see [13] for a detailed discussion of these cylinders)

$$\widetilde{Q}_\varepsilon(s, y) := \left\{(t, x) \in (S + \varepsilon^2, T - \varepsilon^2) \times \mathbb{R}^d : (s, y) \in Q_\varepsilon(t, x)\right\}.$$

Then from the definition of $Q_\varepsilon(t, x)$, we can see that

$$\begin{aligned}\widetilde{Q}_\varepsilon(s, y) &\subset \left\{(t, x) : |t - s| < \varepsilon^2,\ |X_\varepsilon(t, x; s) - y| < \varepsilon\right\} \\ &= \left\{(t, x) : |t - s| < \varepsilon^2,\ x' := X_\varepsilon(t, x; s) \in B_\varepsilon(y)\right\} \\ &= \left\{(t, x) : |t - s| < \varepsilon^2,\ x' \in B_\varepsilon(y),\ x = X_\varepsilon(s, x'; t)\right\}.\end{aligned}$$

Because $u_\varepsilon$ is also divergence-free, the measure of a set is invariant under the flow, so we have

$$\mathcal{L}^d\left(\left\{X_\varepsilon(s, x'; t) : x' \in B_\varepsilon(y)\right\}\right) = \mathcal{L}^d(B_\varepsilon(y)).$$

Thus the measure of the dual cylinder is

$$\begin{aligned}\mathcal{L}^{d+1}(\widetilde{Q}_\varepsilon(s, y)) &\le \mathcal{L}^{d+1}\left(\left\{(t, x) : |t - s| < \varepsilon^2,\ x' \in B_\varepsilon(y),\ x = X_\varepsilon(s, x'; t)\right\}\right) \\ &= \int_{\max(S, s-\varepsilon^2)}^{\min(T, s+\varepsilon^2)} \mathcal{L}^d\left(\left\{X_\varepsilon(s, x'; t) : x' \in B_\varepsilon(y)\right\}\right) \mathrm{d}t \\ &\le 2\varepsilon^2 |B_\varepsilon| = |Q_\varepsilon|.\end{aligned}$$

Plugging into (22), we conclude that

$$\begin{aligned}\int_{S+\varepsilon^2}^{T-\varepsilon^2} \int_{\mathbb{R}^d} |f_\varepsilon(t, x)| \,\mathrm{d}x\,\mathrm{d}t &\le \frac{1}{|Q_\varepsilon|} \int_S^T \int_{\mathbb{R}^d} |f(s, y)| \mathcal{L}^{d+1}(\widetilde{Q}_\varepsilon(s, y)) \,\mathrm{d}y\,\mathrm{d}s \\ &\le \int_S^T \int_{\mathbb{R}^d} |f(t, x)| \,\mathrm{d}x\,\mathrm{d}t. \qquad \blacksquare\end{aligned}$$

*Proof of Proposition* 11. If $p = \infty$ in Assumption 10, the conclusions follow naturally from Definitions 1 and 2, as now both the velocity field and its gradient are locally bounded. We shall only focus on the case $p < \infty$ from now.

Without loss of generality, assume $\eta \le \eta_0$. For $S < t < T$, $x \in \mathbb{R}^d$, define

$$F(t, x) := [\mathcal{M}(\nabla u(t))(x)]^p \in L^1((S, T) \times \mathbb{R}^d),$$

and $F_\varepsilon(t, x)$ is defined the same as in (21). Lemma 12 shows that $\|F_\varepsilon\|_{L^1} \le \|F\|_{L^1}$. We want to show that for sufficiently small $\varepsilon$,

$$F_\varepsilon(t, x) \le \eta^p \varepsilon^{-2p}.$$

By Remark 1, this implies that $Q_\varepsilon(t, x)$ is $\eta$-admissible. Define the set of nonadmissible points by

$$\Omega_\varepsilon = \left\{(t, x) \in (S + \varepsilon^2, T - \varepsilon^2) \times \mathbb{R}^d : F_\varepsilon(t, x) > \eta^p \varepsilon^{-2p}\right\}.$$

By Chebyshev's inequality, its measure is bounded by

$$|\Omega_\varepsilon| \le \left|\left\{F_\varepsilon > \eta^p \varepsilon^{-2p}\right\}\right| \le \frac{\|F_\varepsilon\|_{L^1}}{\eta^p \varepsilon^{-2p}} \le \frac{\|F\|_{L^1}}{\eta^p}\varepsilon^{2p} \to 0$$

as $\varepsilon \to 0$. Therefore, $|\bigcap_{\varepsilon>0} \Omega_\varepsilon| = 0$, that is, the set of points at which no $\eta$-admissible cylinder centers has measure zero. In other words, for almost every point $(t, x)$, there exists $\varepsilon > 0$ such that $Q_\varepsilon(t, x)$ is $\eta$-admissible.

This is not enough to show the conclusion, because $\Omega_\varepsilon$ may not be monotone in $\varepsilon$. To see that $Q_\varepsilon(t, x)$ is $\eta$-admissible for all sufficiently small $\varepsilon$, let us define

$$\Omega'_\varepsilon = \left\{(t, x) \in (S + \varepsilon^2, T - \varepsilon^2) \times \mathbb{R}^d : F_\varepsilon(t, x) > \eta^p (2^{d+1}\varepsilon)^{-2p}\right\}.$$

Similarly to before, Chebyshev's inequality implies

$$|\Omega'_\varepsilon| \le \frac{\|F\|_{L^1}}{\eta^p}(2^{d+1}\varepsilon)^{2p}.$$

In particular, for each $i \ge 1$, we have a geometric decaying upper bound as

$$|\Omega'_{2^{-i}}| \le \frac{\|F\|_{L^1}}{\eta^p}(2^{d+1}2^{-i})^{2p}.$$

It is a summable geometric series in $i$; thus by the Borel–Cantelli lemma we have

$$\left|\limsup_{i\to\infty} \Omega'_{2^{-i}}\right| = \left|\bigcap_{I>0}\bigcup_{i>I} \Omega'_{2^{-i}}\right| = 0.$$

That is, for almost every $(t, x) \in (S, T) \times \mathbb{R}^d$, there exists $I > 0$ such that for all $i > I$, $(t, x) \notin \Omega'_{2^{-i}}$, i.e. for $\varepsilon_i = 2^{-i}$ we have

$$F_{\varepsilon_i}(t, x) = \fint_{Q_{\varepsilon_i}(t,x)} F(s, y)\,\mathrm{d}y\,\mathrm{d}s \le \eta^p (2^{d+1}\varepsilon_i)^{-2p}.$$

By Remark 1, Jensen's inequality implies

$$\varepsilon_i^2 \fint_{Q_{\varepsilon_i}(t,x)} \mathcal{M}(\nabla u)\,\mathrm{d}y\,\mathrm{d}s \le \varepsilon_i^2 \left(\fint_{Q_{\varepsilon_i}(t,x)} [\mathcal{M}(\nabla u)]^p\,\mathrm{d}y\,\mathrm{d}s\right)^{\frac{1}{p}} \le \frac{\eta}{4^{d+1}}.$$

That is, $Q_{\varepsilon_i}(t, x)$ is $(4^{-d-1}\eta)$-admissible.

We claim that if $Q_{\varepsilon_\alpha}(t_0, x_0)$ is $(4^{-d-1}\eta)$-admissible, then for every $\varepsilon_\beta$ within $\frac{\varepsilon_\alpha}{4} \le \varepsilon_\beta \le \frac{\varepsilon_\alpha}{2}$, $Q_{\varepsilon_\beta}(t_0, x_0) \subset \frac{3}{4}Q_{\varepsilon_\alpha}(t_0, x_0)$. This can be proven by the claim

$$|X_{\varepsilon_\beta}(t_0, x_0; t) - X_{\varepsilon_\alpha}(t_0, x_0; t)| \le \frac{\varepsilon_\alpha}{4} \quad \text{for all } t \in (t_0 - {\varepsilon_\beta}^2, t_0 + {\varepsilon_\beta}^2), \tag{23}$$

whose proof is a slight modification of Lemma 6. Define $Q^\alpha = Q_{\varepsilon_\alpha}(t_0, x_0)$ and $Q^\beta = Q_{\varepsilon_\beta}(t_0, x_0)$. If we proceed with the proof of Lemma 6, without knowing that $Q^\beta$ is $\eta$-admissible, the only difficulty will arise at the last step (16), when we want to bound the integral of ${\varepsilon_\beta}^{-d}\|\mathcal{M}(\nabla u(t))\|_{L^1(B^\beta(t))}$ in Grönwall's inequality. However, as long as (23) holds at time $t$, $B^\beta(t)$ is contained in $B^\alpha(t)$, thus

$$
{\varepsilon_\beta}^{-d}\|\mathcal{M}(\nabla u(t))\|_{L^1(B^\beta(t))} \le 4^d {\varepsilon_\alpha}^{-d}\|\mathcal{M}(\nabla u(t))\|_{L^1(B^\alpha(t))}
$$

while the integral of the latter is bounded by $4^d\eta$. Following the same continuity argument, we conclude (23) in the end.

By this claim, for every $\varepsilon$ between $\frac{\varepsilon_i}{4}$ and $\frac{\varepsilon_i}{2}$, we have

$$
Q_\varepsilon(t,x) \subset \frac{3}{4} Q_{\varepsilon_i}(t,x) \subset \left(\frac{3}{4}\right)^2 Q_{\varepsilon_{i-1}}(t,x) \subset \cdots
$$

which implies $\operatorname{diam}(Q_\varepsilon(t,x)) \to 0$ as $\varepsilon \to 0$. Although we do not have monotonicity for $Q_\varepsilon(t,x)$ in $\varepsilon$, we have this "monotonicity with gaps". Moreover, since $Q_\varepsilon(t,x) \subset Q_{\varepsilon_i}(t,x)$, $\varepsilon > \frac{\varepsilon_i}{4}$, we can bound $F_\varepsilon$ by

$$
\begin{aligned}
F_\varepsilon(t,x) = \frac{1}{|Q_\varepsilon|}\int_{Q_\varepsilon(t,x)} F \,\mathrm{d}y\,\mathrm{d}s \le \frac{|Q_{\varepsilon_i}|}{|Q_\varepsilon|} \fint_{Q_{\varepsilon_i}(t,x)} F \,\mathrm{d}y\,\mathrm{d}s &\le 4^{d+1}\eta^p (2^{d+1}\varepsilon_i)^{-2p} \\
&\le \eta^p \varepsilon_i^{-2p} \le \eta^p \varepsilon^{-2p}.
\end{aligned}
$$

Thus $(t,x) \notin \Omega_\varepsilon$ for every $\varepsilon \in [\frac{\varepsilon_i}{4}, \frac{\varepsilon_i}{2}]$ and for every $i > I$, that is, for every $\varepsilon \le 2^{-I-1}$. This means $Q_\varepsilon(t,x)$ is admissible for all $\varepsilon$ sufficiently small. ∎

Following this existence proposition, we furthermore have the following corollary on the $L^1$ convergence.

**Corollary 13** ($L^1$ convergence). *Let $f \in L^1((S,T)\times\mathbb{R}^d)$, and define $f_\varepsilon$ by (21); then*

$$
f_\varepsilon \to f \text{ in } L^1((S,T)\times\mathbb{R}^d) \quad \text{as } \varepsilon \to 0.
$$

*Proof.* For any $\delta > 0$, we can find $g \in C_c^\infty((S,T)\times\mathbb{R}^d)$ such that $\|f-g\|_{L^1} < \frac{\delta}{3}$. Denote $h = f - g$; then $\|h\|_{L^1} < \frac{\delta}{3}$ and, by Lemma 12, $\|h_\varepsilon\|_{L^1} < \frac{\delta}{3}$ also (we define $h_\varepsilon$ in the same way as (21)). Since $g$ is uniformly continuous, it is clear that as $\operatorname{diam}(Q_\varepsilon(t,x)) \to 0$,

$$
\|g - g_\varepsilon\|_{L^1} \le \int_{(S+\varepsilon^2, T-\varepsilon^2)\times\mathbb{R}^d} \fint_{Q_\varepsilon(t,x)} |g(t,x) - g(s,y)| \,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t < \frac{\delta}{3}
$$

for sufficiently small $\varepsilon$ such that $g(t,\cdot) = 0$ in $(S, S+\varepsilon^2) \cup (T-\varepsilon^2, T)$. Thus

$$
\|f - f_\varepsilon\|_{L^1} = \|g + h - g_\varepsilon - h_\varepsilon\|_{L^1} \le \|g - g_\varepsilon\|_{L^1} + \|h\|_{L^1} + \|h_\varepsilon\|_{L^1} < \delta
$$

provided $\varepsilon$ is small enough. ∎

### 3.2. Maximal function

The existence Proposition 11 ensures the maximal function is well defined almost everywhere. With the help of the covering lemma, we can prove the bounds for the maximal function. A lot of ideas are borrowed from [19]. We do not claim any originality for results in this section but only put them here for the sake of completeness.

*Proof of Theorem* 1. By the existence Proposition 11, for almost every $(t,x) \in (S,T) \times \mathbb{R}^d$, the set $\{\varepsilon > 0 : Q_\varepsilon(t,x) \text{ is } \eta\text{-admissible}\}$ is nonempty, so the maximal function $\mathcal{M}_{\mathcal{Q}}(f)$ is well defined almost everywhere.

(1) This is evident from the definition, since for any $(t,x)$ it holds that

$$\fint_{Q_\varepsilon(t,x)} |f(s,y)|\,\mathrm{d}y\,\mathrm{d}s \le \|f\|_{L^\infty}.$$

(2) For any $\lambda > 0$, let $E_\lambda = \{(t,x) : (\mathcal{M}_{\mathcal{Q}} f)(t,x) > \lambda\}$ be the superlevel set. Then by definition, there is an $\eta$-admissible $Q_\varepsilon$ centered at each point $(t,x) \in E_\lambda$, such that

$$|Q_\varepsilon| < \frac{1}{\lambda} \int_{Q_\varepsilon(t,x)} |f(s,y)|\,\mathrm{d}y\,\mathrm{d}s.$$

Their radii are thus uniformly bounded. Thanks to the covering lemma Proposition 9, we can choose a pairwise disjoint subcollection $\{Q_{\varepsilon_j}(t^j, x^j)\}$, such that

$$\sum_j |Q_{\varepsilon_j}| \ge \frac{1}{C}\Bigg|\bigcup_{(t,x)\in E_\lambda} Q_\varepsilon(t,x)\Bigg|.$$

Therefore the measure of the superlevel set can be bounded by

$$|E_\lambda| \le C \sum_j |Q_{\varepsilon_j}| \le \frac{C}{\lambda} \sum_j \int_{Q_{\varepsilon_j}(t^j,x^j)} |f|\,\mathrm{d}x\,\mathrm{d}t \le \frac{C}{\lambda} \int_{(S,T)\times\mathbb{R}^d} |f|\,\mathrm{d}x\,\mathrm{d}t.$$

(3) For the type $(q,q)$ part, we use Marcinkiewicz interpolation. Note that $\mathcal{M}_{\mathcal{Q}}$ is subadditive: $\mathcal{M}_{\mathcal{Q}}(f+g) \le \mathcal{M}_{\mathcal{Q}}(f) + \mathcal{M}_{\mathcal{Q}}(g)$. We can split $f = f_1 + f_2$ where $f_1 = f\chi_{|f|\le\frac{\lambda}{2}}$ and $f_2 = f\chi_{|f|>\frac{\lambda}{2}}$. First, the strong type $(\infty,\infty)$ estimate applied to $f_1$ yields

$$\|\mathcal{M}_{\mathcal{Q}}(f_1)\|_{L^\infty} \le \frac{\lambda}{2}.$$

Thus we have

$$\mathcal{M}_{\mathcal{Q}}(f) \le \mathcal{M}_{\mathcal{Q}}(f_1) + \mathcal{M}_{\mathcal{Q}}(f_2) \le \frac{\lambda}{2} + \mathcal{M}_{\mathcal{Q}}(f_2).$$

So $\mathcal{M}_{\mathcal{Q}}(f) > \lambda$ implies $\mathcal{M}_{\mathcal{Q}}(f_2) > \frac{\lambda}{2}$. Next, the weak type $(1,1)$ estimate applied to $f_2$ yields

$$\mu(E_\lambda) \le \mu\big(\{\mathcal{M}_{\mathcal{Q}}(f_2) > \tfrac{\lambda}{2}\}\big) \le \frac{2C}{\lambda}\|f_2\|_{L^1}.$$

By the layer cake representation, we have

$$\begin{aligned}\int_{(S,T)\times\mathbb{R}^d} [\mathcal{M}_{\mathcal{Q}}(f)]^q\,\mathrm{d}t\,\mathrm{d}x &= q\int_0^\infty \mu(E_\lambda)\lambda^{q-1}\,\mathrm{d}\lambda \\ &\le 2Cq\int_0^\infty \frac{1}{\lambda}\int_{(S,T)\times\mathbb{R}^d} |f|\chi_{|f|>\frac{\lambda}{2}}\lambda^{q-1}\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda\end{aligned}$$

$$
\begin{aligned}
&= 2Cq \int_{(S,T)\times\mathbb{R}^d} |f| \int_0^{2|f|} \lambda^{q-2}\,\mathrm{d}\lambda\,\mathrm{d}x\,\mathrm{d}t \\
&= \frac{2Cq\cdot 2^{q-1}}{q-1} \int_{(S,T)\times\mathbb{R}^d} |f|^q\,\mathrm{d}x\,\mathrm{d}t = C_q\|f\|_{L^q}^q .
\end{aligned}
$$

This finishes the proof of the theorem. ∎

This theorem together with the $L^1$ convergence will imply the almost everywhere convergence of $f_\varepsilon$.

**Corollary 14** (a.e. convergence). *Given $f \in L^1_{\mathrm{loc}}((S,T)\times\mathbb{R}^d)$, for almost every $(t,x) \in (S,T)\times\mathbb{R}^d$, we have $f_\varepsilon(t,x) \to f(t,x)$ as $\varepsilon \to 0$, where $f_\varepsilon$ is defined in* (21).

*Proof.* According to Proposition 11, $\mathrm{diam}(Q_\varepsilon(t,x)) \to 0$ for almost every $(t,x)$, so we can assume $f$ is compactly supported and thus integrable without loss of generality. By Corollary 13 ($L^1$ convergence) we can find a subsequence which converges to $f$ almost everywhere, hence it suffices to show the following oscillation function is zero almost everywhere: for $f \in L^1_{\mathrm{loc}}((S,T)\times\mathbb{R}^d)$, define the oscillation function by

$$
\Omega f(t,x) = \limsup_{\varepsilon\to 0} f_\varepsilon(t,x) - \liminf_{\varepsilon\to 0} f_\varepsilon(t,x).
$$

For a uniformly continuous function $g$, we have $\Omega g \equiv 0$ almost everywhere, again using the fact that $\mathrm{diam}(Q_\varepsilon(t,x)) \to 0$ by Proposition 11. Moreover, notice that as $\varepsilon \to 0$, $Q_\varepsilon(t,x)$ is $\eta$-admissible, so we have

$$
\begin{aligned}
\limsup_{\varepsilon\to 0} f_\varepsilon(t,x) &\le \limsup_{\varepsilon\to 0} |f_\varepsilon(t,x)| \le \mathcal{M}_{\mathcal{Q}}(f)(t,x), \\
-\liminf_{\varepsilon\to 0} f_\varepsilon(t,x) &\le \limsup_{\varepsilon\to 0} |f_\varepsilon(t,x)| \le \mathcal{M}_{\mathcal{Q}}(f)(t,x),
\end{aligned}
$$

so $\Omega f \le 2\mathcal{M}_{\mathcal{Q}}(f)$ almost everywhere. Now we fix $\lambda > 0$. For any given $\delta > 0$, we split $f = g + h$ with $g \in C_c^\infty((S,T)\times\mathbb{R}^d)$ and $\|h\|_{L^1} < \delta$; we have

$$
\Omega f \le \Omega h + \Omega g = \Omega h \le 2\mathcal{M}_{\mathcal{Q}}(h).
$$

By Theorem 1, the weak type $(1,1)$ estimate gives

$$
\mu\big(\{\Omega f > \lambda\}\big) \le \mu\big(\{\mathcal{M}_{\mathcal{Q}}(h) > \tfrac{\lambda}{2}\}\big) \le \frac{2C}{\lambda}\|h\|_{L^1} = \frac{2C}{\lambda}\delta.
$$

Set $\delta \to 0$; we obtain

$$
\mu\big(\{\Omega f > \lambda\}\big) = 0.
$$

This is true for any $\lambda > 0$; therefore we actually have

$$
\mu\big(\{\Omega f > 0\}\big) = 0.
$$

This means, for almost every $(t,x) \in (S,T)\times\mathbb{R}^d$, the oscillation is zero and

$$
\limsup_{\varepsilon\to 0} f_\varepsilon(t,x) = \liminf_{\varepsilon\to 0} f_\varepsilon(t,x) = \lim_{\varepsilon\to 0} f_\varepsilon(t,x) = f(t,x). \qquad ∎
$$

Using the definition of $\mathcal{M}_{\mathcal{Q}}$, it is easy to deduce the following.

**Corollary 15.** *For $f \in L^1_{\mathrm{loc}}((S,T)\times\mathbb{R}^d)$, $f \le \mathcal{M}_{\mathcal{Q}}(f)$ almost everywhere.*

To conclude this section, we present a slightly stronger result than the almost everywhere convergence.

**Theorem 16** ($\mathcal{Q}$-Lebesgue differentiation theorem). *Under the same assumption as Corollary 14, for almost every $(t,x)\in(S,T)\times\mathbb{R}^d$, we have*

$$\lim_{\varepsilon\to 0} ⨍_{Q_\varepsilon(t,x)} |f(s,y)-f(t,x)|\,\mathrm{d}y\,\mathrm{d}s = 0. \tag{24}$$

If (24) is true for $(t,x)$, we call it a *$\mathcal{Q}$-Lebesgue point of $f$*, and define the *$\mathcal{Q}$-Lebesgue set of $f$* to be the set of all $\mathcal{Q}$-Lebesgue points of $f$.

*Proof of Theorem 16.* Consider any rational number $q\in\mathbb{Q}$. Then $f-q\in L^1_{\mathrm{loc}}$; thus by Corollary 14 we have

$$|f-q|_\varepsilon(t,x) = ⨍_{Q_\varepsilon(t,x)} |f-q|(s,y)\,\mathrm{d}y\,\mathrm{d}s \to |f-q|(t,x) \quad \text{a.e. as } \varepsilon\to 0.$$

By taking a countable intersection over $q\in\mathbb{Q}$ of all the sets where the convergence $|f-q|_\varepsilon \to |f-q|$ happens, we have

$$|f-q|_\varepsilon(t,x)\to|f-q|(t,x) \quad \text{a.e. as } \varepsilon\to 0 \text{ for all } q\in\mathbb{Q}.$$

By the density of rational numbers, it holds that

$$|f-r|_\varepsilon(t,x)\to|f-r|(t,x) \quad \text{a.e. as } \varepsilon\to 0 \text{ for all } r\in\mathbb{R}.$$

In particular, letting $r=f(t,x)$ gives

$$|f-f(t,x)|_\varepsilon(t,x)\to|f(t,x)-f(t,x)|=0 \quad \text{a.e. as } \varepsilon\to 0.$$

This is equivalent to (24). ∎

## 4. Application to the Navier–Stokes equations

In this section we give an example of how to use the maximal function to bridge between the local study and global results. Here we provide an alternative proof for $L^p$-weak integrability for higher derivatives of three-dimensional Navier–Stokes equations. In [4, Proposition 2.2] (case $r=0$), the authors obtained the following local theorem. Recall that $B_r$ represents a ball of radius $r$ in $\mathbb{R}^3$.

**Proposition 17** ([4]). *Let $\varphi\in C_c^\infty(B_1)$ be radial, satisfying $0\le\varphi\le 1$, $\int\varphi\,\mathrm{d}x=1$, and $\varphi\equiv 1$ on $B_{\frac{1}{2}}$. There exists $\bar\eta>0$, such that if $v,p\in C^\infty((-4,0)\times\mathbb{R}^3)$ is a solution to*

$$\partial_t v + (v\cdot\nabla)v + \nabla p = \Delta v, \quad \operatorname{div} v = 0$$

*verifying both*

$$\int_{\mathbb{R}^3} \varphi(x) v(t,x)\, dx = 0 \quad \text{for almost every } t \in (-4,0),$$

$$\int_{-4}^{0}\int_{B_2}\Big(\big|\mathcal{M}(|\mathcal{M}(\nabla v)|^q)\big|^{\frac{2}{q}} + |\nabla^2 p| + \sum_{m=d}^{d+4}\sup_{\delta>0}|(\nabla^{m-1}h^\alpha)_\delta * \nabla^2 p|\Big)\, dx\, dt \le \bar\eta$$

*for some integer* $d \ge 1$, $\alpha \in [0,2)$, $q = \frac{12}{\alpha+6}$, *and* $(\nabla^m h^\alpha)_\delta$ *is defined by*

$$h^\alpha(x) := \frac{\varphi(\frac{x}{2}) - \varphi(x)}{|x|^{3+\alpha}}, \quad (\nabla^m h^\alpha)_\delta(x) := \frac{1}{\delta^3}(\nabla^m h^\alpha)\Big(\frac{x}{\delta}\Big),$$

*then*

$$|(-\Delta)^{\frac{\alpha}{2}}\nabla^d v| \le C_{d,\alpha} \quad \text{in } (-1/36, 0) \times B_{\frac{1}{6}}(0).$$

This local theorem aims to control the magnitude of the higher fractional derivatives using quantities involving $\nabla v$ and $\nabla^2 p$. Indeed, $\iint |\nabla v|^2\, dx\, dt$ and $\iint |\nabla^2 p|\, dx\, dt$ both have the best scaling of the equation, and it is not hard to see the integrand in Proposition 17 has the same scaling. The average zero condition ensures that the velocity $v$ is small as well, so that the quadratic flux term $v \cdot \nabla v$ is manageable in the parabolic regularization. Moreover, since the purpose is to control a nonlocal quantity $(-\Delta)^{\frac{\alpha}{2}}\nabla^d v$, we need to gather nonlocal information using the maximal function $\mathcal{M}$ and the $\nabla^m h^\alpha$-maximal function: $\sup_{\delta>0}|(\nabla^m h^\alpha)_\delta * \cdot|$.

Let $u$ be a smooth solution to the Navier–Stokes equations (3) in $(0,T)$. Since we need to center the cylinders at the terminal time for the local study, let us change our notation, and redefine

$$Q_\varepsilon(t,x) := \big\{(s,y) : t - \varepsilon^2 < s < t,\ |y - X_\varepsilon(t,x;s)| < \varepsilon\big\}$$

based on the velocity field $u$, which has $L^2$ gradient and divergence zero. Results for the covering lemma and the maximal function can all be applied to this family of skewed cylinders, as we are just recentering. By Galilean transform, the previous local proposition implies the following in the global coordinates.

**Corollary 18.** *There exists* $\bar\eta > 0$, *such that if* $u, P \in C^\infty((0,T)\times\mathbb{R}^3)$ *is a solution to* (3) *verifying for some* $(t,x) \in (0,T)\times\mathbb{R}^3$, $\varepsilon < \frac{1}{2}\sqrt{t}$,

$$\frac{1}{\varepsilon}\int_{Q_{2\varepsilon}(t,x)}\Big(\big|\mathcal{M}(|\mathcal{M}(\nabla u)|^q)\big|^{\frac{2}{q}} + |\nabla^2 P| + \sum_{m=d}^{d+4}\sup_{\delta>0}|(\nabla^{m-1}h^\alpha)_\delta * \nabla^2 P|\Big)\, dx\, dt \le \bar\eta,$$

*then*

$$|(-\Delta)^{\frac{\alpha}{2}}\nabla^d u|(t,x) \le \frac{C_{d,\alpha}}{\varepsilon^{d+\alpha+1}}.$$

*Proof.* For a fixed $(t,x)$, denote $r(s) = t + \varepsilon^2 s$ and $z(s) = X_\varepsilon(t,x;r)$; then

$$\dot r = \varepsilon^2, \quad \dot z = \varepsilon^2 u_\varepsilon(r,z).$$

We define the following change of coordinates:

$$\begin{aligned} v(s,y) &= \varepsilon u(r, z+\varepsilon y) - \varepsilon u_\varepsilon(r,z), \\ p(s,y) &= \varepsilon^2 P(r, z+\varepsilon y) + \varepsilon y \partial_s [u_\varepsilon(r,z)]. \end{aligned}$$

Then we have the following in the new variables:

$$\begin{aligned} \partial_s v &= \varepsilon \dot{r} \partial_t u + \varepsilon \dot{z} \cdot \nabla u - \varepsilon \dot{r} \partial_t u_\varepsilon - \varepsilon \dot{z} \cdot \nabla u_\varepsilon \\ &= \varepsilon^3 (\partial_t u + u_\varepsilon \cdot \nabla u - \partial_t u_\varepsilon - u_\varepsilon \cdot \nabla u_\varepsilon), \\ v \cdot \nabla_y v &= v \cdot \varepsilon^2 \nabla u = \varepsilon^3 (u \cdot \nabla u - u_\varepsilon \cdot \nabla u), \\ \Delta_y v &= \varepsilon^3 \Delta u. \end{aligned}$$

Combining these three, we obtain

$$\begin{aligned} \partial_s v + v \cdot \nabla v - \Delta v &= \varepsilon^3 (\partial_t u + u \cdot \nabla u - \Delta u - (\partial_t + u_\varepsilon \cdot \nabla) u_\varepsilon) \\ &= -\varepsilon^3 (\nabla P + (\partial_t + u_\varepsilon \cdot \nabla) u_\varepsilon). \end{aligned}$$

Moreover, since $\partial_s u_\varepsilon(r,z) = \dot{r} u_\varepsilon + \dot{z} \cdot \nabla u_\varepsilon = \varepsilon^2 (\partial_t + u_\varepsilon \cdot \nabla) u_\varepsilon$, we have

$$\nabla_y p = \varepsilon^2 \varepsilon \nabla P + \varepsilon \partial_s (u_\varepsilon(r,z)) = \varepsilon^3 (\nabla P + (\partial_t + u_\varepsilon \cdot \nabla) u_\varepsilon).$$

Therefore, $(v, p)$ is also a solution to the Navier–Stokes equations. Now we check that $v$, $p$ satisfy the assumptions of Proposition 17. First, since $t \in (4\varepsilon^2, T)$, we know that $(v, p)$ is a smooth solution for $s \in (-4, 0)$. Next we can verify that

$$\int_{\mathbb{R}^3} \varphi(y) v(s,y) \, \mathrm{d}y = \varepsilon \left( \int_{\mathbb{R}^3} \varphi(y) u(r, z+\varepsilon y) \, \mathrm{d}y - u_\varepsilon(r,z) \right) = 0.$$

For the last condition, the change of variable yields

$$\begin{aligned} \left| \mathcal{M}(|\mathcal{M}(\nabla v)|^q) \right|^{\frac{2}{q}} &= \varepsilon^4 \left| \mathcal{M}(|\mathcal{M}(\nabla u)|^q) \right|^{\frac{2}{q}}, \\ \nabla^2 p &= \varepsilon^2 \cdot \varepsilon^2 \nabla^2 P + 0 = \varepsilon^4 \nabla^2 P, \\ (\nabla^{m-1} h^\alpha)_\delta * \nabla^2 p &= \varepsilon^4 (\nabla^{m-1} h^\alpha)_{\varepsilon\delta} * \nabla^2 P. \end{aligned}$$

Since $Q_{2\varepsilon}(t,x) = \{(r, z+\varepsilon y) : (s,y) \in (-4,0) \times B_2\}$ has space-time dimension 5, the last condition of Proposition 17 is verified. As a consequence, we can bound

$$|(-\Delta)^{\frac{\alpha}{2}} \nabla^d v(s,y)| \le C_{d,\alpha} \quad \text{in } (-1/36, 0) \times B_{\frac{1}{6}}(0).$$

In particular, when $s = 0$, $y = 0$, we have

$$C_{d,\alpha} \ge |(-\Delta)^{\frac{\alpha}{2}} \nabla^d v(0,0)| = \varepsilon^{d+\alpha+1} |(-\Delta)^{\frac{\alpha}{2}} \nabla^d u(t,x)|. \qquad \blacksquare$$

Based on this, let us prove Theorem 3 using the maximal function $\mathcal{M}_{\mathcal{Q}}$.

*Proof of Theorem 3*. Define

$$F(t,x) = \left|\mathcal{M}(|\mathcal{M}(\nabla u)|^q)\right|^{\frac{2}{q}}(t,x) + |\nabla^2 P| + \sum_{m=d}^{d+4} \sup_{\delta>0} |(\nabla^{m-1}h^\alpha)_\delta * \nabla^2 P|(t,x).$$

It is well known that for the Navier–Stokes equations, smooth solutions satisfy the following energy inequality:

$$\|\nabla u\|^2_{L^2((0,T)\times\mathbb{R}^3)} \leq \frac{1}{2}\|u_0\|^2_{L^2(\mathbb{R}^3)}.$$

By the boundedness of the spatial maximal function in $L^2(\mathbb{R}^3)$ and in $L^{\frac{2}{q}}(\mathbb{R}^3)$, we have

$$\left\|\left|\mathcal{M}(|\mathcal{M}(\nabla u)|^q)\right|^{\frac{2}{q}}\right\|_{L^1} \leq \|\nabla u\|^2_{L^2}.$$

Moreover, using $-\Delta P = \operatorname{div}(u\cdot\nabla u) = \nabla u_i \cdot \partial_{x_i} u$, by the compensated compactness ([6]), we bound

$$\|\nabla^2 P\|_{L^1(0,T;\mathcal{H}^1(\mathbb{R}^3))} \leq \|\nabla u\|^2_{L^2((0,T)\times\mathbb{R}^3)}$$

where $\mathcal{H}^1$ is the Hardy space. It is continuously embedded in $L^1$, and we can use the Hardy norm to bound the $\nabla^m h^\alpha$-maximal function by

$$\left\|\sup_{\delta>0} |(\nabla^{m-1}h^\alpha)_\delta * \nabla^2 P(t)|\right\|_{L^1(\mathbb{R}^3)} \leq C_{m,\alpha}\|\nabla^2 P(t)\|_{\mathcal{H}^1(\mathbb{R}^3)}.$$

Combining the above estimates, we conclude that

$$\|F\|_{L^1((0,T)\times\mathbb{R}^3)} \leq C\|u_0\|^2_{L^2(\mathbb{R}^3)}. \tag{25}$$

Denote $\eta := \min\{\frac{\bar\eta}{|Q_1|}, (\eta_0)^2\}$, and for $(t,x)\in(0,T)\times\mathbb{R}^3$ we define

$$I(\varepsilon) := \varepsilon^4 \fint_{Q_\varepsilon(t,x)} F(s,y)\,\mathrm{d}y\,\mathrm{d}s = \frac{1}{\varepsilon|Q_1|}\int_{Q_\varepsilon(t,x)} F(s,y)\,\mathrm{d}y\,\mathrm{d}s.$$

For all the $\mathcal{Q}$-Lebesgue points $(t,x)$ of $F$, we claim that there exists a positive $\varepsilon = \varepsilon_{(t,x)}$ such that one of the following two cases is true:

Case 1. $\varepsilon_{(t,x)} < t^{\frac{1}{2}}$ and $I(\varepsilon_{(t,x)}) = \eta$.

Case 2. $\varepsilon_{(t,x)} = t^{\frac{1}{2}}$ and $I(\varepsilon_{(t,x)}) \leq \eta$.

The reason is that $\lim_{\varepsilon\to 0} I(\varepsilon) = 0^4 F(t,x) = 0$, and $I(\varepsilon)$ is clearly a continuous function of $\varepsilon$ when $\varepsilon > 0$. As $\varepsilon$ ranges from 0 to $t^{\frac{1}{2}}$, either $I(\varepsilon)$ reaches $\eta$ at some $\varepsilon_{(t,x)} < t^{\frac{1}{2}}$ (Case 1), or it remains smaller than $\eta$ until $\varepsilon_{(t,x)} = t^{\frac{1}{2}}$ (Case 2).

At this $\varepsilon = \varepsilon_{(t,x)}$ level, because

$$|\mathcal{M}(\nabla u)|^2 \leq \left|\mathcal{M}(|\mathcal{M}(\nabla u))|^q)\right|^{\frac{2}{q}} \leq F,$$

by Jensen we have

$$\varepsilon^4\left(\fint_{Q_\varepsilon(t,x)} |\mathcal{M}(\nabla u)|\,\mathrm{d}y\,\mathrm{d}s\right)^2 \leq \varepsilon^4 \fint_{Q_\varepsilon(t,x)} |\mathcal{M}(\nabla u)|^2\,\mathrm{d}y\,\mathrm{d}s \leq I(\varepsilon) \leq \eta,$$

which implies $Q_\varepsilon(t,x)$ is actually $\sqrt{\eta}$-admissible. So when in Case 1,

$$\eta = \varepsilon^4 \fint_{Q_\varepsilon(t,x)} F(s,y)\,\mathrm{d}y\,\mathrm{d}s \le \varepsilon^4 \mathcal{M}_{\mathcal{Q}} F(t,x).$$

Combining with Case 2, we conclude

$$\varepsilon_{(t,x)}^{-4} \le \max\Big\{t^{-2}, \tfrac{\mathcal{M}_{\mathcal{Q}} F(t,x)}{\eta}\Big\}.$$

Moreover, because $|Q_1| \cdot I(\varepsilon_{(t,x)}) \le \bar{\eta}$ in both cases, Corollary 18 claims that

$$\begin{aligned} |(-\Delta)^{\frac{\alpha}{2}} \nabla^d u|(t,x) &\le \frac{C_{d,\alpha}}{\varepsilon_{(t,x)}^{d+\alpha+1}} \\ \Rightarrow f^p(t,x) &\le C_{d,\alpha} \varepsilon_{(t,x)}^{-4} \le C_{d,\alpha} \max\Big\{t^{-2}, \tfrac{\mathcal{M}_{\mathcal{Q}} F(t,x)}{\eta}\Big\}. \end{aligned}$$

Finally, because $\mathcal{M}_{\mathcal{Q}}$ is of weak type $(1,1)$, $\|\mathcal{M}_{\mathcal{Q}} F\|_{L^{1,\infty}} \le C\|F\|_{L^1}$. Together with (25) we complete the proof of the theorem. ∎

**Funding.** The author was partially supported by the NSF grant DMS-RTG 1840314 (PI: Alexis Vasseur).

## References


[1] J. T. Beale and A. Majda, Vortex methods. I. Convergence in three dimensions. *Math. Comp.* **39** (1982), no. 159, 1–27 Zbl 0488.76024 MR 658212

[2] J. T. Beale and A. Majda, Vortex methods. II. Higher order accuracy in two and three dimensions. *Math. Comp.* **39** (1982), no. 159, 29–52 Zbl 0488.76025 MR 658213

[3] L. Caffarelli, R. Kohn, and L. Nirenberg, Partial regularity of suitable weak solutions of the Navier-Stokes equations. *Comm. Pure Appl. Math.* **35** (1982), no. 6, 771–831 Zbl 0509.35067 MR 673830

[4] K. Choi and A. F. Vasseur, Estimates on fractional higher derivatives of weak solutions for the Navier-Stokes equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **31** (2014), no. 5, 899–945 Zbl 1297.76047 MR 3258360

[5] A. J. Chorin, Numerical study of slightly viscous flow. *J. Fluid Mech.* **57** (1973), no. 4, 785–796 MR 395483

[6] R. Coifman, P.-L. Lions, Y. Meyer, and S. Semmes, Compensated compactness and Hardy spaces. *J. Math. Pures Appl. (9)* **72** (1993), no. 3, 247–286 Zbl 0864.42009 MR 1225511

[7] P. Constantin, Navier-Stokes equations and area of interfaces. *Comm. Math. Phys.* **129** (1990), no. 2, 241–266 Zbl 0725.35080 MR 1048693

[8] P. Constantin, An Eulerian-Lagrangian approach for incompressible fluids: local theory. *J. Amer. Math. Soc.* **14** (2001), no. 2, 263–278 Zbl 0997.76009 MR 1815212

[9] P. Constantin, I. Kukavica, and V. Vicol, Contrast between Lagrangian and Eulerian analytic regularity properties of Euler equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **33** (2016), no. 6, 1569–1588 Zbl 1353.35233 MR 3569243

[10] G.-H. Cottet and P. D. Koumoutsakos, *Vortex methods*. Cambridge University Press, Cambridge, 2000 Zbl 1140.76002 MR 1755095

[11] P. Isett, *Hölder continuous Euler flows in three dimensions with compact support in time*. Ann. of Math. Stud. 196, Princeton University Press, Princeton, NJ, 2017 Zbl 1367.35001 MR 3616258
[12] P. Isett, A proof of Onsager's conjecture. *Ann. of Math. (2)* **188** (2018), no. 3, 871–963 Zbl 1416.35194 MR 3866888
[13] P. Isett and S.-J. Oh, On nonperiodic Euler flows with Hölder regularity. *Arch. Ration. Mech. Anal.* **221** (2016), no. 2, 725–804 Zbl 1338.35344 MR 3488536
[14] E. Kaas, B. Sørensen, P. H. Lauritzen, and A. B. Hansen, A hybrid Eulerian–Lagrangian numerical scheme for solving prognostic equations in fluid dynamics. *Geosci. Model Dev.* **6** (2013), no. 6, 2023–2047
[15] P.-L. Lions, *Mathematical topics in fluid mechanics. Vol. 1: Incompressible models*. Oxford Lecture Ser. Math. Appl. 3, The Clarendon Press, Oxford University Press, New York, 1996 Zbl 0866.76002 MR 1422251
[16] A. J. Majda and A. L. Bertozzi, *Vorticity and incompressible flow*. Cambridge Texts Appl. Math. 27, Cambridge University Press, Cambridge, 2002 Zbl 0983.76001 MR 1867882
[17] P.-A. Raviart, An analysis of particle methods. In *Numerical methods in fluid dynamics (Como, 1983)*, pp. 243–324, Lecture Notes in Math. 1127, Springer, Berlin, 1985 Zbl 0598.76003 MR 802214
[18] V. Scheffer, The Navier-Stokes equations on a bounded domain. *Comm. Math. Phys.* **73** (1980), no. 1, 1–42 Zbl 0451.35048 MR 573611
[19] E. M. Stein, *Singular integrals and differentiability properties of functions*. Princeton Math. Ser. 30, Princeton University Press, Princeton, NJ, 1970 Zbl 0207.13501 MR 0290095
[20] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Princeton Math. Ser. 43, Princeton University Press, Princeton, NJ, 1993 Zbl 0821.42001 MR 1232192
[21] A. Vasseur, Higher derivatives estimate for the 3D Navier-Stokes equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27** (2010), no. 5, 1189–1204 Zbl 1396.35046 MR 2683756
[22] A. Vasseur and J. Yang, Second derivatives estimate of suitable solutions to the 3D Navier-Stokes equations. *Arch. Ration. Mech. Anal.* **241** (2021), no. 2, 683–727 Zbl 1473.35414 MR 4275744

Received 24 August 2020; revised 17 May 2021; accepted 17 August 2021.

**Jincheng Yang**
Department of Mathematics, The University of Texas at Austin, 2515 Speedway Stop C1200, Austin, TX 78712, USA; jcyang@math.utexas.edu