Received: 22 November 2019 | Revised: 24 July 2020 | Accepted: 19 August 2020

DOI: 10.1002/mana.201900506

ORIGINAL PAPER

MATHEMATISCHE NACHRICHTEN

Maximal regularity for the Cauchy problem of the heat equation in *BMO*

Takayoshi Ogawa[1] | **Senjo Shimizu**[2]

[1]Mathematical Institute, Tohoku University, Sendai 980-8578, Japan

[2]Graduate School of Human and Environmental Studies, Kyoto University, Kyoto 606-8501, Japan

Correspondence
Senjo Shimizu, Graduate School of Human and Environmental Studies, Kyoto University, Kyoto 606-8501, Japan.
Email shimizu.senjo.5s@kyoto-u.ac.jp

Dedicated to Professor Kenji Yajima on the occasion of his seventieth birthday

Funding information
Japan Society for the Promotion of Science, Grant/Award Numbers: Challenging Research (Pioneering) JP20K20284, Grant in aid for Scientific Research (B) JP16H03945, (B) JP21H00992, Grant in aid for Scientific Research (S) JP19H05597; Fostering Joint International Research, Grant/Award Number: JP18KK0072

Abstract

We consider maximal regularity for the Cauchy problem of the heat equation in a class of bounded mean oscillations (BMO). Maximal regularity for non-reflexive Banach spaces is not obtained by the established abstract theory. Based on the symmetric characterization of BMO-expression, we obtain maximal regularity for the heat equation in BMO and its sharp trace estimate. Our result shows that the homogeneous initial estimate obtained by Stein [50] and Koch–Tataru [32] can be strengthened up to the inhomogeneous estimate for the external forces and the obtained estimates can be applicable to quasilinear problems. Our method is based on integration by parts and can also be applicable to other type of parabolic problems.

KEYWORDS

BMO, heat equations, maximal regularity, non-UMD Banach space, sharp trace estimate, VMO

MSC (2020)

Primary: 35K15, 35K45; Secondary: 42B37

1 | MAXIMAL REGULARITY FOR THE CAUCHY PROBLEM

We consider the initial value problem of the heat equation in the whole space:

$$\begin{cases} \partial_t u - \nu\Delta u = f, & t>0,\ x\in\mathbb{R}^n, \\ u(0,x) = u_0(x), & x\in\mathbb{R}^n, \end{cases} \tag{1.1}$$

where $\nu > 0$ is a constant, $u = u(t,x)\colon \mathbb{R}_+\times\mathbb{R}^n \to \mathbb{R}^n$ denotes the unknown function, and $f = f(t,x)\colon \mathbb{R}_+\times\mathbb{R}^n\to\mathbb{R}^n$ and $u_0 = u_0(x)\colon \mathbb{R}^n\to\mathbb{R}^n$ are given external and initial data.

Let X be a proper Banach space and let A be a closed linear operator in X with a dense domain $\mathcal{D}(A)$. Given $u_0\in X$ and $f\in L^\rho(0,T;X)$ $(1\le\rho\le\infty)$, we consider the abstract Cauchy problem

$$\begin{cases} \dfrac{d}{dt}u + Au = f, \quad t>0, \\ u(0) = u_0. \end{cases}$$

We say that A has maximal L^ρ-regularity if there exists a unique solution u such that $\frac{d}{dt}u,\ Au \in L^\rho(0,T;X)$ satisfy the estimate

$$\left\|\frac{d}{dt}u\right\|_{L^\rho(0,T;X)} + \|Au\|_{L^\rho(0,T;X)} \le C\Big(\|u_0\|_{(X,\mathcal{D}(A))_{1-1/\rho,\rho}} + \|f\|_{L^\rho(0,T;X)}\Big) \tag{1.2}$$

under the restriction $u_0 \in (X,\mathcal{D}(A))_{1-1/\rho,\rho}$, where $(X,\mathcal{D}(A))_{1-1/\rho,\rho}$ denotes the real interpolation space between X and $\mathcal{D}(A)$, and C is a positive constant independent of u_0 and f. Maximal regularity for parabolic equations was first developed by Ladyzhenskaya–Solonnikov–Ural'tseva [34]. Then the research of maximal regularity has progressed immensely in these last few decades such as [4, 12, 13, 14, 17, 18, 19, 20, 25, 26, 29, 48, 49]. In the general framework on Banach spaces X that satisfy *the unconditional martingale differences* (called as UMD), well established especially by Amann[1, 2], Denk–Hieber–Prüss [15, 16], Weis [53] and by [30, 33, 45].

On the other hand, maximal regularity for Banach spaces that are not UMD (e.g., non-reflexive Banach spaces such as L^1 or L^∞) requires a different treatment. For example, we have previously proven maximal regularity for homogenous Banach spaces in [36–38].

The class of *bounded mean oscillation* (*BMO*) introduced by John–Nirenberg [28] is one of such non-reflexive Banach spaces. A locally integrable function f is in the class of bounded mean oscillation $BMO = BMO(\mathbb{R}^n)$ if

$$\|f\|_{BMO} \equiv \sup_{x,R>0} \frac{1}{|B_R|}\int_{B_R(x_0)} \left|f(x) - \overline{f_{B_R}}\right| dx < \infty, \tag{1.3}$$

where $\overline{f_{B_R}}$ denotes the integral average of f over a ball $B_R(x_0)$;

$$\overline{f_{B_R}} \equiv \frac{1}{|B_R|}\int_{B_R(x_0)} f(x)\,dx. \tag{1.4}$$

Introducing the quotient space of all BMO-functions with a constant difference are equivalent, BMO is a Banach space with the norm $\|\cdot\|_{BMO}$. We denote $VMO = VMO(\mathbb{R}^n)$ as the BMO completion of $C_0(\mathbb{R}^n)$, where $C_0(\mathbb{R}^n)$ denotes the set of all continuous functions with compact supports. It is well known that the dual of VMO is the Hardy space $\mathcal{H}^1(\mathbb{R}^n)$ and the dual of the Hardy space is BMO (cf. Fefferman–Stein [21]).

Applying Stein's basic estimate [50, p. 158] to the formulation of a BMO by the Carleson measure, Koch–Tataru [32] showed the following estimate: For $\nu = 1$,

$$\left(\sup_{x_0,R>0}\int_0^{R^2}\left(\frac{1}{|B_R|}\int_{B_R(x_0)} \left|\nabla e^{t\Delta}u_0\right|^2 dx\right) dt\right)^{1/2} \le C_0\|u_0\|_{BMO}. \tag{1.5}$$

This is an expression of the BMO-semi-norm using the heat kernel. In view of the theory of evolution equations, the estimate (1.5) can be regarded as the homogeneous estimate of maximal regularity in (1.2). On the other hand, since BMO is a dual of the Hardy class $\mathcal{H}^1(\mathbb{R}^n)$ but not a pre-dual of itself, it is not reflexive. Further, since any UMD Banach space is necessarily reflexive (cf. Amann [1], Rubio de Francia [46]), the Banach space BMO is not UMD; therefore, we cannot apply the general theorem of maximal regularity by e.g. [1, 2, 4, 12, 13–20, 25, 26, 29, 30, 33, 45, 48, 49, 53] as we mentioned before.

In this paper, we derive maximal regularity for the Cauchy problem of the heat equation (1.1) with $\nabla u_0 \in BMO(\mathbb{R}^n)$ using the fact that the semi-norm of BMO is expressed by the L^2-framework.

It is worth mentioning that the well-posedness of the Cauchy problem for the incompressible Navier–Stokes equation is considered in various scaling invariant classes. Consider the initial value problem for the Navier–Stokes system in the whole space:

$$\begin{cases} \partial_t u - \Delta u + (u\cdot\nabla)u + \nabla p = 0, & t>0,\ x\in\mathbb{R}^n,\\ \operatorname{div} u = 0, & t>0,\ x\in\mathbb{R}^n,\\ u(0,x) = u_0(x), & x\in\mathbb{R}^n. \end{cases} \tag{1.6}$$

Equation (1.6) remains invariant under the scaling:

$$\begin{cases} u_\lambda(t,x) \equiv \lambda u\big(\lambda^2 t, \lambda x\big), \\ p_\lambda(t,x) \equiv \lambda^2 p\big(\lambda^2 t, \lambda x\big) \end{cases} \tag{1.7}$$

and the scaling invariant class is realized by the Bochner class such as the homogeneous Sobolev space $\dot H^s_p = \dot H^s_p(\mathbb{R}^n)$;

$$L^\theta\big(\mathbb{R}_+; \dot H^s_p(\mathbb{R}^n;\mathbb{R}^n)\big), \quad \frac{2}{\theta} + \frac{n}{p} = 1 + s,$$

where $1 \le \theta \le \infty$, $1 \le p < \infty$ and $-1 < s$. If we restrict ourselves to $\theta = \infty$, then

$$L^\infty\Big(\mathbb{R}_+; \dot H_p^{-1+n/p}(\mathbb{R}^n;\mathbb{R}^n)\Big)$$

is a typical invariant space. According to the well-known result by Fujita–Kato [23], the solvability of the scaling invariant problem in such a space is a basic and important problem for (1.6) (cf. [40, 44, 47]). Indeed, the global well-posedness of the solution for the incompressible Navier–Stokes equation is obtained in $L^\infty\big(0,T;\dot B^{-1+n/p}_{p,q}(\mathbb{R}^n)\big)$, $1 \le q < \infty$, or $BUC_w\big([0,\infty); BMO^{-1}\big(\mathbb{R}^3\big)\big) \cap C\,\big((0,\infty); L^\infty\big(\mathbb{R}^3\big)\big)$ (for the time local well-posedness in VMO^{-1} by Miura–Sawada [35]). A similar well-posedness was obtained by several authors [1, 9, 10, 24, 31, 41] and ill-posedness by [7, 52, 54].

Along with such a background, it is worth considering maximal regularity in BMO based on a Bochner or related function spaces. It may be possible to apply such estimates to other quasilinear parabolic problems and free boundary value problems. We shall discuss such a result elsewhere [39].

1.1 | The space of Koch–Tataru

Koch–Tataru [32] introduced the caloric extension for expressing the function of BMO^{-1} and obtained the global well-posedness for the incompressible Navier–Stokes equation. They introduced the following expression:

$$\|f\|_{BMO^{-1}} \equiv \sup_{x\in\mathbb{R}^n, R>0} \left(\frac{1}{|B_R(x)|} \int_0^{R^2} \int_{B_R(x)} \left|e^{t\Delta} f(x)\right|^2 dx\,dt \right)^{1/2} < \infty. \tag{1.8}$$

Then they constructed a global solution to (1.6) by the metric induced from

$$|||f||| \equiv \sup_{t<T} t^{1/2} \big\|e^{t\Delta} f\big\|_\infty + \|f\|_{BMO^{-1}}.$$

We should note that Frazier–Jawerth [22] (cf. Peetre [43]) introduced the equivalent norm of BMO in terms of Lizorkin–Triebel space (cf. Stein [50], Triebel [51]).

According to the celebrated result due to John–Nirenberg [28], it is reasonable to introduce an equivalent L^2-based semi-norm of BMO;

$$\|f\|_{BMO} \equiv \sup_{x_0\in\mathbb{R}^n, R>0} \left(\frac{1}{|B_R|} \int_{B_R(x_0)} \left| f(x) - \overline{f_{B_R}} \right|^2 dx \right)^{1/2},$$

where $\overline{f_{B_R}}$ is defined by (1.4). Then it is easy to see that the following expression gives an equivalent norm (cf. Brezis–Nirenberg [8]):

$$\|f\|_{BMO} \simeq \sup_{x_0\in\mathbb{R}^n, R>0} \left(\frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} |f(x) - f(y)|^2\,dx\,dy \right)^{1/2}.$$

Accordingly, we introduce a class of space-time functions that substitutes for the Bochner class $L^2(I;BMO)$.

Definition 1.1. Let $I=(0,T)$ for $T<\infty$ and let $1\le\theta\le 2$. A measurable function $f\in\widetilde{L^\theta(I;BMO(\mathbb{R}^n))}$ if

$$\|f\|_{\widetilde{L^\theta(I;BMO)}}\equiv\sup_{x_0\in\mathbb{R}^n,R>0}\left(\int_{I\cap(0,R^2)}\left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}|f(t,x)-f(t,y)|^2\,dx\,dy\right)^{\theta/2}dt\right)^{1/\theta}<\infty. \tag{1.9}$$

We also denote if $T=\infty$,

$$\|f\|_{\widetilde{L^\theta(\mathbb{R}_+;BMO)}}\equiv\sup_{x_0\in\mathbb{R}^n,R>0}\left(\int_0^{R^2}\left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}|f(t,x)-f(t,y)|^2\,dx\,dy\right)^{\theta/2}dt\right)^{1/\theta}<\infty. \tag{1.10}$$

These spaces are variants of spaces introduced by Chemin–Lerner [11] [1]. Let $I=(0,T)$ for some $0<T\le\infty$. Then a natural extension of the Chemin–Lerner type space for $\theta=2$ can be given by

$$\|f\|_{\widetilde{L^2(I_T;BMO)}}\equiv\sup_{x_0\in\mathbb{R}^n,R>0}\left(\int_0^T\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}|f(t,x)-f(t,y)|^2\,dx\,dy\,dt\right)^{1/2}<\infty.$$

The norm defined in (1.9) is a modification of this definition. In this sense we regard that the norm in (1.9) as one modification of the Chemin–Lerner type variation from the Bochner space. We also define that $f\in\widetilde{W^{1,2}(I;BMO(\mathbb{R}^n))}$ if both f and $\partial_t f\in\widetilde{L^2(I;BMO(\mathbb{R}^n))}$. We also define $\widetilde{L^2(I;VMO(\mathbb{R}^n))}$ and $\widetilde{W^{1,2}(I;VMO(\mathbb{R}^n))}$ in a manner similar to $VMO(\mathbb{R}^n)$. The homogeneous Sobolev spaces based on BMO and VMO are defined as follows. For any $s\in\mathbb{R}$,

$$B\dot{M}O^s=\{f\in\mathcal{S}^*;|\nabla|^s f\in BMO(\mathbb{R}^n)\},$$
$$V\dot{M}O^s=\{f\in\mathcal{S}^*;|\nabla|^s f\in VMO(\mathbb{R}^n)\},$$

where $\mathcal{S}^*$ denotes the tempered distribution and we regard them as Banach spaces by taking a quotient space using polynomial functions. The spaces $\widetilde{L^\theta\left(I;B\dot{M}O^s(\mathbb{R}^n)\right)}$ and $\widetilde{L^\theta\left(I;V\dot{M}O^s(\mathbb{R}^n)\right)}$ are analogously defined as above.

The above norm (semi-norm) (1.9) is equivalent to the norm below: For $1\le p\le 2$,

$$\|f\|_{\widetilde{L^2(I;BMO_p)}}\equiv\sup_{x_0\in\mathbb{R}^n,R>0}\left(\int_{I\cap(0,R^2)}\left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}|f(t,x)-f(t,y)|^p\,dx\,dy\right)^{2/p}dt\right)^{1/2}<\infty$$

(see Proposition 5.4 in the Appendix below). From the definition, it follows for $I=(0,T)$ with $0<T\le\infty$ that

$$L^2(I;BMO(\mathbb{R}^n))\subsetneq\widetilde{L^2(I;BMO(\mathbb{R}^n))}. \tag{1.11}$$

One can see that the above norm (semi-norm) (1.9) is equivalent to the following norm:

$$\|f\|_{BMO(\mathbb{R}^n;L^2(I))}\equiv\sup_{x_0\in\mathbb{R}^n,R>0}\left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}\left(\int_{I\cap(0,R^2)}|f(t,x)-f(t,y)|^2\,dt\right)dx\,dy\right)^{1/2}<\infty.$$

Besides

$$BMO\left(\mathbb{R}^n;L^2(I)\right)\simeq\widetilde{L^2(I;BMO(\mathbb{R}^n))}$$

when we restrict ourselves to a finite time interval $I = (0,T)$, it is a Bochner space of space-time variables [2]. From the definition, it generally follows that

$$L^2(\mathbb{R}_+; BMO(\mathbb{R}^n)) \subset BMO\big(\mathbb{R}^n; L^2(\mathbb{R}_+)\big) \subset \widetilde{L^2(\mathbb{R}_+; BMO(\mathbb{R}^n))}. \tag{1.12}$$

1.2 | Maximal regularity in *BMO* and the sharp trace estimate

We extend the estimate (1.5) for maximal regularity to the initial value problem for the heat equation (1.1) in BMO which is our main theorem.

Theorem 1.2 (Maximal regularity in BMO). *(1) For every $f \in \widetilde{L^2(I; VMO(\mathbb{R}^n))}$ and $\nabla u_0 \in VMO(\mathbb{R}^n)$, then the initial value problem (1.1) admits a unique solution $u \in \widetilde{W^{1,2}}(I; VMO(\mathbb{R}^n)) \cap \widetilde{L^2}\Big(\mathbb{R}_+; V\dot{M}O^2(\mathbb{R}^n)\Big)$ which satisfies the following estimate:*

$$\|\partial_t u\|_{\widetilde{L^2(I;VMO)}} + \nu\|\Delta u\|_{\widetilde{L^2(I;VMO)}} \le C_M\Big(\|\nabla u_0\|_{VMO} + \|f\|_{\widetilde{L^2(I;VMO)}}\Big), \tag{1.13}$$

where $C_M > 0$ is independent of $\nu > 0$ and $T > 0$.

(2) Besides if $f \in \widetilde{L^2(I;BMO(\mathbb{R}^n))}$ and $\nabla u_0 \in BMO(\mathbb{R}^n)$, then the initial value problem (1.1) admits a unique solution $u \in \widetilde{W^{1,2}}(I; BMO(\mathbb{R}^n)) \cap \widetilde{L^2}\Big(I; B\dot{M}O^2(\mathbb{R}^n)\Big)$ which satisfies the following estimate:

$$\|\partial_t u\|_{\widetilde{L^2(I;BMO)}} + \nu\|\Delta u\|_{\widetilde{L^2(I;BMO)}} \le C_M\Big(\|\nabla u_0\|_{BMO} + \|f\|_{\widetilde{L^2(I;BMO)}}\Big), \tag{1.14}$$

where $C_M > 0$ is independent of $\nu > 0$ and $T > 0$.

Corollary 1.3 (Global maximal regularity in BMO). *All the statements in Theorem 1.2 hold for the case where $\nabla u_0 \in VMO(\mathbb{R}^n)$ and $f \in \widetilde{L^2(\mathbb{R}_+; VMO(\mathbb{R}^n))}$. Namely, it holds that*

$$\|\partial_t u\|_{\widetilde{L^2(\mathbb{R}_+};VMO)} + \nu\|\Delta u\|_{\widetilde{L^2(\mathbb{R}_+};VMO)} \le C_M\Big(\|\nabla u_0\|_{VMO} + \|f\|_{\widetilde{L^2(\mathbb{R}_+};VMO)}\Big)$$

and for $\nabla u_0 \in BMO(\mathbb{R}^n)$ and $f \in \widetilde{L^2(\mathbb{R}_+; BMO(\mathbb{R}^n))}$,

$$\|\partial_t u\|_{\widetilde{L^2(\mathbb{R}_+};BMO)} + \nu\|\Delta u\|_{\widetilde{L^2(\mathbb{R}_+};BMO)} \le C_M\Big(\|\nabla u_0\|_{BMO} + \|f\|_{\widetilde{L^2(\mathbb{R}_+};BMO)}\Big).$$

The uniqueness of the solution presented in the above statements is up to constant, since we admit the spaces BMO and VMO are understood to be Banach spaces, where any difference of constants is identified.

We should like to note that an analogous maximal regularity estimate for the heat equation is partially shown by Iwabuchi–Nakamura [27].

It is interesting to consider whether maximal regularity for the heat equation holds in the Bochner class $f \in L^2(I; BMO)$ or not. Indeed, the homogeneous estimate (for the initial data) does not hold in BMO since one can derive that

$$\|\partial_t u\|_{L^2(I;BMO)} + \nu\|\Delta u\|_{L^2(I;BMO)} \le C_M\|\nabla u_0\|_{\dot{B}^0_{\infty,2}} \tag{1.15}$$

holds for the solution of the heat equation (1.1) with $f \equiv 0$ (cf. [36], [37]), where $\dot{B}^s_{p,\sigma}(\mathbb{R}^n)$ denotes the homogeneous Besov space (see [5]). It is known that the homogeneous estimate (1.15) above is sharp in the sense that the corresponding trace estimate holds, i.e.,

$$\|\nabla u(\cdot)\|_{BUC\left(I;\dot{B}^0_{\infty,2}\right)} \le C\big(\|\partial_t u\|_{L^2(I;BMO)} + \|\Delta u\|_{L^2(I;BMO)}\big). \tag{1.16}$$

We give a sketch of proof for (1.16) in the Appendix below. By the strict inclusion

$$\dot{B}^0_{\infty,2}(\mathbb{R}^n) \subsetneq BMO(\mathbb{R}^n),$$

the homogeneous estimate by $\|\nabla u_0\|_{BMO}$ instead of $\|\nabla u_0\|_{\dot{B}^0_{\infty,2}}$ in (1.15) fails in general. [3] This is also provided by the trace estimate below. The analogous estimate for the Stokes equation is also possible (cf. [39]).

One can generalize the maximal regularity estimate into a non-maximal but useful estimate for the semi-linear problem:

Theorem 1.4 (Generalized maximal regularity in BMO). *There exists a constant $C_M > 0$ independent of T such that for any $f \in \widetilde{L^\theta\Big(I;\dot{BMO}^{-1+2/\theta}(\mathbb{R}^n)\Big)}$ and $\nabla u_0 \in BMO(\mathbb{R}^n)$, then the initial value problem (1.1) admits a unique solution $u \in \widetilde{\dot{W}^{1,2}}(I;BMO(\mathbb{R}^n)) \cap \widetilde{L^2\Big(\mathbb{R}_+;\dot{BMO}^2(\mathbb{R}^n)\Big)}$ which satisfies the following estimate:*

$$\|\partial_t u\|_{\widetilde{L^2(I;BMO)}} + \nu\|\Delta u\|_{\widetilde{L^2(I;BMO)}} \le C_M\left(\|\nabla u_0\|_{BMO} + \|f\|_{\widetilde{L^\theta\left(\mathbb{R}_+;\dot{BMO}^{-1+2/\theta}\right)}}\right).$$

The homogeneous estimate for maximal regularity gives the optimality conditions for the initial data taken in BMO. This is important to the claim that the Koch–Tataru space is crucial to treating the semi-linear problem in BMO-based spaces.

Theorem 1.5 (Time trace estimates). *For $0 < T \le \infty$, let $I = (0,T)$. Then the following estimate holds:*

1. *For any $f \in \widetilde{\dot{W}^{1,2}}(I;VMO) \cap \widetilde{L^2\Big(I;\dot{VMO}^2\Big)}$ with $f(0) \equiv 0$, there exists a constant $C > 0$ independent of f such that*

$$\|\nabla f\|_{BUC(I;VMO)} \le C\left(\|\partial_t f\|_{\widetilde{L^2(I;VMO)}} + \|\Delta f\|_{\widetilde{L^2(I;VMO)}}\right) \tag{1.17}$$

   holds. In particular, this estimate is sharp.

2. *For any $f \in \widetilde{\dot{W}^{1,2}}(I;BMO) \cap \widetilde{L^2\Big(I;\dot{BMO}^2\Big)}$ which $f(0) \equiv 0$, there exists a constant $C > 0$ independent of f such that*

$$\|\nabla f\|_{L^\infty(I;BMO)} \le C\left(\|\partial_t f\|_{\widetilde{L^2(I;BMO)}} + \|\Delta f\|_{\widetilde{L^2(I;BMO)}}\right). \tag{1.18}$$

   In particular, from (1.12),

$$\|\nabla f\|_{L^\infty(I;BMO)} \le C\left(\|\partial_t f\|_{L^2(I;BMO)} + \|\Delta f\|_{L^2(I;BMO)}\right). \tag{1.19}$$

Remark 1.6. The condition on $f(0) = 0$ is assumed to avoid the case when f has an uncertainty of adding a linear function such as $f + ax + b$ for $a, b \in \mathbb{R}/\{0\}$. We also note that the above trace estimate holds for a function in

$$\dot{W}^{1,p}(I;BMO) \cap L^p\left(I;\dot{BMO}^2\right),$$

but the regularity should be naturally arranged such as $|\nabla|^{1-1/p}$.

Remark 1.7. After completing this work, we noticed that a related estimate for maximal regularity (1.13) was reported by Auscher–Frey [3]. They introduced the tent space function $T^{\infty,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ and established a maximal regularity estimate for the heat equation of the Navier–Stokes type. Their elegant proof is based on an operator theoretical method using a T-T^* argument. Indeed, the Koch–Tataru estimate is dependent on the initial data estimate, which is understood as the homogeneous space-time estimate. If one can apply a suitable duality argument, then the T-T^* argument yields a maximal regularity estimate, such as (1.13) and (1.14). The basis of our proof is different from theirs, however it is implicitly related to the T-T^* argument via integration by parts.

We should also like to mention that the properties on the heat and Stokes semi-group on a smooth domain are considered by Bolkart–Giga–Suzuki–Tsutsui [6], where they consider the equivalence of various norms of bounded mean oscillation in a smooth domain and establish the boundedness and analyticity of the semi-groups in BMO.

In what follows, we use the following notation. $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$ denotes the rapidly decreasing functions. $\{\phi_j\}_{j\in\mathbb{Z}}$ denotes the Littlewood–Paley dyadic decomposition of unity. $A \simeq B$ stands for the equivalent quantities that are bounded by a constant each other. Various constants are simply denoted by C unless otherwise noted.

2 | GENERALIZATION OF THE HOMOGENEOUS ESTIMATE

We now turn our focus onto the case $T = \infty$. The general case $T < \infty$ can be treated in a similar way. We also assume that $\nu = 1$ in (1.1) for simplicity. The general case easily follows by scaling.

2.1 | Homogeneous estimate in *BMO*

The following estimate is obtained by Koch–Tataru [32] using the Stein estimate [50].

Proposition 2.1 [32], [50]. *Let $e^{t\Delta}$ be the heat kernel and let $u_0 \in BMO(\mathbb{R}^n)$. Then*

$$\|\nabla e^{t\Delta}u_0\|_{\widetilde{L^2(I;BMO)}} \le C_0\|u_0\|_{BMO}, \tag{2.1}$$

where C_0 is independent of $T > 0$.

It is also shown by Stein [50] that the left-hand side of the estimate (2.1) is equivalent to $\|u_0\|_{BMO}$.

Corollary 2.2. *Let $I = (0, T)$, let $e^{t\Delta}$ be the heat kernel and let $u_0 \in BMO(\mathbb{R}^n)$. Then*

$$\sup_{x_0, R>0} \int_{I\cap(0,R^2)} \left(\frac{1}{|B_R|} \int_{B_R(x_0)} \left| \nabla e^{t\Delta}u_0 - \left(\overline{\nabla e^{t\Delta}u_0}\right)_{B_R} \right| dx \right)^2 dt \le C_0\|u_0\|^2_{BMO}, \tag{2.2}$$

where C_0 is independent of $T > 0$.

Corollary 2.2 follows from Proposition 2.1 directly. We give the proof of Proposition 2.1 without using the Stein argument in the Appendix (Subsection A.4).

We then extend the estimate in Proposition 2.1 into the inhomogeneous terms of the heat equation.

Proposition 2.3. *Let $e^{t\Delta}$ be the heat kernel and let $u_0 \in BMO(\mathbb{R}^n)$. Then*

$$\begin{aligned}\sup_{x_0, R>0} &\left(\int_0^{R^2} \frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left(\int_s^{R^2} \left| \nabla_x e^{(t-s)\Delta_x} f(s,x) - \nabla_y e^{(t-s)\Delta_y} f(s,y) \right|^2 dt \right) dx\,dy\,ds \right)^{1/2} \\ &\le C_0\|f\|_{\widetilde{L^2(I;BMO)}},\end{aligned} \tag{2.3}$$

where C_0 is independent of $T > 0$.

The Stein estimate (see Theorem A.11 below) can be modified into an estimate in the Chemin–Lerner space $\widetilde{L^2(I;BMO)}$. For that purpose, we prepare the following lemma:

Lemma 2.4. *Let $1 \le \theta \le 2$. For any $x_0 \in \mathbb{R}^n$ and $R > 0$, there exists a constant $C > 0$ independent of them such that for any $f \in \widetilde{L^\theta(I;BMO)}$,*

$$\begin{aligned}\int_0^{R^2} &\left(\frac{1}{|B_R|^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} |f(t,x) - f(t,y)|^2 \frac{dx}{\left(1 + R^{-1}(x - x_0)\right)^{n+1}} \frac{dy}{\left(1 + R^{-1}(y - x_0)\right)^{n+1}} \right)^{\theta/2} dt \\ &\le C\|f\|^\theta_{\widetilde{L^\theta(I;BMO)}}.\end{aligned} \tag{2.4}$$

Especially,

$$\int_0^{R^2}\left(\int_{\mathbb{R}^n}\left|f(t,x)-\overline{f_{B_R}}(t)\right|\frac{R}{\left(R+|x-x_0|\right)^{n+1}}\,dx\right)^{\theta}dt\le C\|f\|^{\theta}_{\widetilde{L^\theta(I;BMO)}},\tag{2.5}$$

where $\overline{f_{B_R}}$ *denotes an average of* f *over a ball centered* x_0 *with radius* $R>0$.

The estimate (2.4) is shown in the following way. By setting $B_k = B_{2^kR}(x_0)$, we notice

$$\begin{aligned}\int_0^{2^{2(k+1)}R^2}\left|\overline{f_{B_{k+1}}}-\overline{f_{B_k}}\right|^{\theta}dt &= \int_0^{2^{2(k+1)}R^2}\left|\frac{1}{|B_k|}\int_{B_k}\left(f(t,x)-\overline{f_{B_{k+1}}}\right)dx\right|^{\theta}dt\\
&=\int_0^{2^{2(k+1)}R^2}\left|\frac{1}{|B_k|}\int_{B_k}\frac{1}{|B_{k+1}|}\int_{B_{k+1}}(f(t,x)-f(t,y))\,dy\,dx\right|^{\theta}dt\\
&\le\int_0^{2^{2(k+1)}R^2}\left(\frac{1}{2^n|B_k|^2}\iint_{B_k\times B_{k+1}}|f(t,x)-f(t,y)|^2\,dx\,dy\right)^{\theta/2}dt\\
&\le\int_0^{2^{2(k+1)}R^2}\left(\frac{2^n}{|B_{k+1}|^2}\iint_{B_{k+1}\times B_{k+1}}|f(t,x)-f(t,y)|^2\,dx\,dy\right)^{\theta/2}dt\\
&\le C\|f\|^{\theta}_{\widetilde{L^\theta(I;BMO)}}.\end{aligned}\tag{2.6}$$

Then for any $k\in\mathbb{N}$, apply (2.6) $(k-2)$-times to have

$$\begin{aligned}&\left(\int_0^{R^2}\left|\frac{1}{|B_k|}\int_{B_k}\left|f(t,x)-\overline{f_{B_1}}\right|^2dx\right|^{\theta/2}dt\right)^{1/\theta}\\
&\le\left(\int_0^{R^2}\left(\frac{1}{|B_k|}\int_{B_k}\left|f(t,x)-\overline{f_{B_{k-1}}}\right|^2dx\right)^{\theta/2}dt\right)^{1/\theta}\\
&\quad+\underbrace{\left(\int_0^{R^2}\left|\overline{f_{B_{k-1}}}-\overline{f_{B_{k-2}}}\right|^{\theta}dt\right)^{1/\theta}+\cdots+\left(\int_0^{R^2}\left|\overline{f_{B_2}}-\overline{f_{B_1}}\right|^{\theta}dt\right)^{1/\theta}}_{k-2\text{ terms}}\\
&\le\left(\int_0^{2^{2k}R^2}\left(\frac{1}{|B_k|}\int_{B_k}\left|f(t,x)-\overline{f_{B_{k-1}}}\right|^2dx\right)^{\theta/2}dt\right)^{1/\theta}\\
&\quad+\underbrace{\left(\int_0^{2^{2(k-1)}R^2}\left|\overline{f_{B_{k-1}}}-\overline{f_{B_{k-2}}}\right|^{\theta}dt\right)^{1/\theta}+\cdots+\left(\int_0^{2^{2\cdot2}R^2}\left|\overline{f_{B_2}}-\overline{f_{B_1}}\right|^{\theta}dt\right)^{1/\theta}}_{k-2\text{ terms}}\\
&\le Ck\|f\|_{\widetilde{L^\theta(I;BMO)}}.\end{aligned}\tag{2.7}$$

Therefore, applying the estimate (2.7) and noticing that

$$\int_{\mathbb{R}^n} \frac{R\,dx}{\left(R+|x-x_0|\right)^{n+1}} = \frac{1}{R^n}\int_{\mathbb{R}^n} \frac{dx}{\left(1+R^{-1}|x-x_0|\right)^{n+1}} \le C_0,$$

$$\begin{aligned}
&\int_0^{R^2}\left(\frac{1}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}|f(t,x)-f(t,y)|^2\frac{dx}{\left(1+R^{-1}|x-x_0|\right)^{n+1}}\frac{dy}{\left(1+R^{-1}|y-x_0|\right)^{n+1}}\right)^{\theta/2}dt\\
&\le 2C_\theta\int_0^{R^2}\left(\frac{1}{|B_R|}\sum_{k\ge 0}\int_{B_{k+1}\setminus B_k}\left|f(t,x)-\overline{f_{B_R}}(t)\right|^2\frac{dx}{\left(1+R^{-1}|x-x_0|\right)^{n+1}}\right)^{\theta/2}dt\\
&\quad+2C_\theta\int_0^{R^2}\left(\frac{1}{|B_R|}\int_{B_R(x_0)}\left|f(t,x)-\overline{f_{B_R}}(t)\right|^2dx\right)^{\theta/2}dt\\
&\le 4C\int_0^{R^2}\left(\frac{1}{|B_R|}\sum_{k\ge 0}2^{-k(n+1)}\int_{B_{k+1}}\left|f(t,x)-\overline{f_{B_R}}(t)\right|^2dx\right)^{\theta/2}dt\\
&\quad+4C\int_0^{R^2}\left(\frac{1}{|B_R|}\int_{B_R(x_0)}\left|f(t,x)-\overline{f_{B_R}}(t)\right|^2dx\right)^{\theta/2}dt\\
&\le C\int_0^{R^2}\sum_{k\ge 0}2^{-\theta k/2}\left(\frac{1}{|B_{2^kR}|}\int_{B_{2^kR}(x_0)}\left|f(t,x)-\overline{f_{B_R}}(t)\right|^2dx\right)^{\theta/2}dt\\
&\quad+C\int_0^{R^2}\left(\frac{1}{|B_R|}\int_{B_R(x_0)}\left|f(t,x)-\overline{f_{B_R}}(t)\right|^2dx\right)^{\theta/2}dt\\
&\le C\sum_{k\ge 0}2^{-\theta k/2}k^\theta\|f\|^\theta_{\widetilde{L^\theta(I;BMO)}}+C\|f\|^\theta_{\widetilde{L^\theta(I;BMO)}}\\
&\le C\|f\|^\theta_{\widetilde{L^\theta(I;BMO)}}.
\end{aligned}\tag{2.8}$$

The proof of Proposition 2.3 essentially relies on the Stein argument [50, p. 158], which is based on the expression of *BMO* by the Carleson measure.

Theorem 2.5 (Stein). *Let $\Phi(x)$ be a smooth potential and satisfy for some constant $C>0$*

$$|\Phi(x)|\le\frac{C}{(1+|x|)^{n+1}}$$

with $\|\Phi\|_1=1$ and set $\Phi_\lambda(x)\equiv\lambda^{-n}\Phi(\cdot/\lambda)$ for any $\lambda>0$. Then there exists a constant $C>0$ such that for any $f\in\widetilde{L^2(I;BMO)}$, it holds that

$$\begin{aligned}
&\sup_{x_0,R>0}\int_0^{R^2}\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}\left(\int_0^R|\Phi_\lambda(x)*\Phi_\lambda(y)*(f(t,x)-f(t,y))|^2\frac{d\lambda}{\lambda}\right)dx\,dy\,dt\\
&\le C\|f\|^2_{\widetilde{L^2(I;BMO)}},
\end{aligned}\tag{2.9}$$

where the double convolution stands for convolutions by each of the variables.

Theorem 2.5 is an extension of the Stein estimate (see Theorem A.11 in the Appendix). The potential is relaxed without posing the zero-average condition because we formulate the estimate with the difference of function f. For the sake of completeness, we show the outlined proof here.

Proof of Theorem 2.5. Let x_0 and R be chosen arbitrarily and fix them. Then noting $\|\Phi_\lambda\|_1 = 1$ we have

$$
\begin{aligned}
&\left|\Phi_\lambda(x) * \Phi_\lambda(y) * \big(f(t,x) - f(t,y)\big)\right| \\
&\le \left|\Phi_\lambda(x) * \big(f(t,x) - \overline{f_{B_R}}(t)\big)\right| + \left|\Phi_\lambda(y) * \big(\overline{f_{B_R}}(t) - f(t,y)\big)\right| \\
&\le \left|\int_{B_{2R}(x_0)} \lambda^{-n}\Phi\left(\frac{x-z}{\lambda}\right)\big(f(t,z) - \overline{f_{B_R}}(t)\big)\,dz\right| + \left|\int_{B_{2R}(x_0)^c} \lambda^{-n}\Phi\left(\frac{x-z}{\lambda}\right)\big(f(t,z) - \overline{f_{B_R}}(t)\big)\,dz\right| \\
&\quad + \left|\int_{B_{2R}(x_0)} \lambda^{-n}\Phi\left(\frac{y-w}{\lambda}\right)\big(f(t,w) - \overline{f_{B_R}}(t)\big)\,dw\right| + \left|\int_{B_{2R}(x_0)^c} \lambda^{-n}\Phi\left(\frac{y-w}{\lambda}\right)\big(f(t,w) - \overline{f_{B_R}}(t)\big)\,dw\right|. \qquad (2.10)
\end{aligned}
$$

The first term can be treated by applying the L^2-boundedness for the square function. Namely,

$$
\left\|\left(\int_0^\infty |\Phi_\lambda * F|^2 \frac{d\lambda}{\lambda}\right)^{1/2}\right\|_2 \le C\|F\|_2. \qquad (2.11)
$$

Indeed, such an estimate can be derived by setting $A_j = [2^j, 2^{j+1})$ as a dyadic interval,

$$
\begin{aligned}
\left\|\left(\int_0^\infty |\Phi_\lambda * F|^2 \frac{d\lambda}{\lambda}\right)^{1/2}\right\|_2^2 &\simeq \int_{\mathbb{R}^n}\left(\sum_{j\in\mathbb{Z}}\int_{A_j} |\phi_j * F|^2 \frac{d\lambda}{\lambda}\right) dx \\
&\simeq 2\log 2 \int_{\mathbb{R}^n}\left(\sum_{j\in\mathbb{Z}} |\phi_j * F|^2\right)^{1/2} dx \\
&\simeq C\|F\|_2^2,
\end{aligned}
$$

where $\{\phi_j\}_{j\in\mathbb{Z}}$ denotes the Littlewood–Paley dyadic decomposition of unity and $\simeq$ stands for the iquivarent quantities that are bounded both from above and below by constants. The first and the third terms of (2.10) are both estimated from (2.11) as follows: Let $\chi_{B_R}(x)$ be a characteristic function of the ball $B_R(x_0)$,

$$
\begin{aligned}
&\int_0^{R^2} \frac{1}{|B_R|^2}\iint_{B_R\times B_R}\left(\int_0^R \left(\int_{B_{2R}(x_0)} \Phi_\lambda(x-z)\left|f(t,z) - \overline{f_{B_R}}(t)\right| dz\right)^2 \frac{d\lambda}{\lambda}\right) dx\,dy\,dt \\
&\quad + \int_0^{R^2} \frac{1}{|B_R|^2}\iint_{B_R\times B_R}\left(\int_0^R \left(\int_{B_{2R}(x_0)} \Phi_\lambda(y-w)\left|f(t,y) - \overline{f_{B_R}}(t)\right| dw\right)^2 \frac{d\lambda}{\lambda}\right) dx\,dy\,dt \\
&\le 2\int_0^{R^2} \frac{1}{|B_R|}\int_{\mathbb{R}^n}\left(\int_0^\infty \left|\Phi_\lambda(x) * \left(\chi_{B_{2R}}\left|f(t,x) - \overline{f_{B_R}}(t)\right|\right)\right|^2 \frac{d\lambda}{\lambda}\right) dx\,dt \qquad (2.12) \\
&\le C\int_0^{R^2} \frac{1}{|B_{2R}|}\int_{\mathbb{R}^n} \chi_{B_{2R}}(x)\left|f(t,x) - \overline{f_{B_R}}(t)\right|^2 dx\,dt \\
&\le C\|f\|^2_{\widetilde{L^2(I;BMO)}}.
\end{aligned}
$$

In contrast, the second and the forth terms in (2.10) are symmetric to each other, so we only treat the second term. Since $x \in B_R(x_0)$ and $z \in B^c_{2R}(x_0)$,

$$\lambda + |x - z| \geq |x_0 - z| - |x - x_0| \geq \frac{1}{3}\big(R + |x_0 - z|\big) \tag{2.13}$$

and for $0 < \lambda < R$,

$$\begin{aligned}
\left| \int_{B^c_{2R}(x_0)} \lambda^{-n} \Phi\Big(\frac{x - z}{\lambda}\Big) \big(f(t,z) - \overline{f_{B_R}}(t)\big)\, dz \right|^2 &\leq C \left(\int_{B^c_{2R}(x_0)} \big|f(t,z) - \overline{f_{B_R}}(t)\big| \frac{dz}{\lambda^n \Big(1 + \frac{|x-z|}{\lambda}\Big)^{n+1}} \right)^2 \\
&\leq C \frac{\lambda^2}{R^2} \left(\int_{B^c_{2R}(x_0)} \big|f(t,z) - \overline{f_{B_R}}(t)\big| \frac{R}{\big(R + |x_0 - z|\big)^{n+1}}\, dz \right)^2 .
\end{aligned} \tag{2.14}$$

On the other hand, by Lemma 2.4,

$$\begin{aligned}
&C \int_0^{R^2} \left(\int_{B^c_{2R}(x_0)} \big|f(t,z) - \overline{f_{B_R}}(t)\big| \frac{R}{\big(R + |x_0 - z|\big)^{n+1}}\, dz \right)^2 dt \\
&\leq C \int_0^{R^2} \left(\int_{\mathbb{R}^n} \big|f(t,z) - \overline{f_{B_R}}(t)\big| \frac{R}{\big(R + |x_0 - z|\big)^{n+1}}\, dz \right)^2 dt \\
&\leq C \|f\|^2_{\widetilde{L^2(I;BMO)}} .
\end{aligned} \tag{2.15}$$

Hence, applying (2.13), (2.14) and (2.15), we see by cancelling y-integration for the second term from (2.10) that

$$\begin{aligned}
&\int_0^{R^2} \frac{1}{|B_R|^2} \iint_{B_R(x_0) \times B_R(x_0)} \left(\int_0^R \left(\int_{B^c_{2R}(x_0)} \Phi_\lambda(x - z) \big|f(t,z) - \overline{f_{B_R}}(t)\big|\, dz \right)^2 \frac{d\lambda}{\lambda} \right) dx\, dy\, dt \\
&\leq C \int_0^{R^2} \frac{1}{|B_R|} \int_{B_R(x_0)} \int_0^R \left(\int_{B^c_{2R}(x_0)} \big|f(t,z) - \overline{f_{B_R}}(t)\big| \frac{\lambda}{\big(\lambda + |x - z|\big)^{n+1}}\, dz \right)^2 \frac{d\lambda}{\lambda}\, dx\, dt \\
&\leq C \frac{1}{|B_R|} \int_{B_R(x_0)} \int_0^R \int_0^{R^2} \left(\int_{B^c_{2R}(x_0)} \big|f(t,z) - \overline{f_{B_R}}(t)\big| \frac{R}{\big(R + |x_0 - z|\big)^{n+1}}\, dz \right)^2 dt\, \frac{\lambda}{R^2}\, d\lambda\, dx \\
&\leq C \int_0^R \left(\frac{1}{|B_R|} \int_{B_R(x_0)} \|f\|^2_{\widetilde{L^2(I;BMO)}}\, dx \right) \frac{\lambda}{R^2}\, d\lambda \\
&\leq C \|f\|^2_{\widetilde{L^2(I;BMO)}} \frac{1}{R^2} \int_0^R \lambda\, d\lambda \\
&\leq C \|f\|^2_{\widetilde{L^2(I;BMO)}} .
\end{aligned} \tag{2.16}$$

The estimate (2.9) is shown after taking the supremum on $R > 0$ and $x_0 \in \mathbb{R}^n$ in both the left-hand sides of (2.12) and (2.16). □

Proof of Proposition 2.3. In Theorem 2.5, let $\Phi_\lambda(x)$ as $\sqrt{t-s}\,\nabla_x e^{(t-s)\Delta_x}$, then it is an integration due to the Haar measure $\frac{d\sqrt{t-s}}{\sqrt{t-s}}$, $\sqrt{t-s}$ over (s^2, R^2), and noting

$$\nabla_x e^{(t-s)\Delta_x}\nabla_y e^{(t-s)\Delta_y}\big(f(s,x)-f(s,y)\big) = \nabla_x e^{(t-s)\Delta_x} f(s,x) - \nabla_y e^{(t-s)\Delta_y} f(s,y),$$

it is a direct consequence of (2.9) in Theorem 2.5. □

Remark 2.6. One can obtain a similar estimate for maximal regularity directly from Theorem 2.5. Setting $\Phi_\lambda(x) \equiv t\Delta e^{-t\Delta}$ with $\lambda = \sqrt{t}$, we directly obtain

$$\begin{aligned}
&\left\| \int_0^t \sqrt{t-s}\Delta_x e^{(t-s)\Delta_x} f(s,x)\,ds \right\|^2_{\widetilde{L^2}(I:BMO)} \\
&= \sup_{x_0, R>0} \int_0^{R^2} \frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \bigg(\int_0^t \Big(\sqrt{t-s}\Delta_x e^{(t-s)\Delta_x} f(s,x) \\
&\qquad\qquad - \sqrt{t-s}\Delta_y e^{(t-s)\Delta_y} f(s,y) \Big)\, ds \bigg)^2 dx\,dy\,dt \\
&\le C\|f\|^2_{\widetilde{L^2(I;BMO)}}.
\end{aligned} \tag{2.17}$$

This estimate (2.17) does not seem to imply the maximal regularity estimate directly.

3 | PROOF OF MAXIMAL REGULARITY IN *BMO*

Proof of Theorem 1.2. The homogeneous estimate for the initial value problem with $f \equiv 0$ have seen in Proposition 2.1 and it suffices to show that the inhomogeneous estimate with $u_0 \equiv 0$. Recall that we have assumed that $\nu = 1$. Let $0 < T \le \infty$, let $I = (0,T)$ and let $f, g \in L^2(I; \mathcal{S})$, where $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$ denotes the rapidly decreasing functions. In order to apply Lemma A.2, we introduce a cut-off function. For $R > 0$, let

$$\eta_R(x) = \begin{cases} 1, & |x-x_0| \le \frac{1}{2}R, \\ \text{smooth, radially decreasing,} & \frac{1}{2}R < |x-x_0| \le \frac{4}{5}R, \\ 0, & \frac{4}{5}R < |x-x_0|, \end{cases} \tag{3.1}$$

be a smooth cut-off function around $x_0 \in \mathbb{R}^n$ and set

$$v(t,x) = \int_0^t e^{(t-s)\Delta} f(s,x)\,ds. \tag{3.2}$$

Then

$$\begin{aligned}
&\frac{d}{dt}\frac{1}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} |v(t,x)-v(t,y)|^2 \eta_R(x)\eta_R(y)\,dx\,dy \\
&= \frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} (v(t,x)-v(t,y))\cdot \partial_t(v(t,x)-v(t,y))\eta_R(x)\eta_R(y)\,dx\,dy
\end{aligned}$$

$$
\begin{aligned}
&= \frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} (v(t,x)-v(t,y))\cdot\big(\Delta_x v(t,x)-\Delta_y v(t,y)\big)\eta_R(x)\eta_R(y)\,dx\,dy \\
&\quad + \frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} (v(t,x)-v(t,y))\cdot(f(t,x)-f(t,y))\eta_R(x)\eta_R(y)\,dx\,dy \\
&\equiv I + II.
\end{aligned} \tag{3.3}
$$

The first term of the right-hand side of (3.3) can be treated by noticing

$$
\nabla_x v(t,y) = \nabla_y v(t,x) = 0
$$

and integrate by parts of each derivative for x and y variables, respectively to see

$$
\begin{aligned}
I =& \frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \big(v(t,x)-v(t,y)\big)\cdot \mathrm{div}_x\big(\nabla_x v(t,x)-\nabla_y v(t,y)\big)\,\eta_R(x)\eta_R(y)\,dx\,dy \\
&+ \frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \big(v(t,x)-v(t,y)\big)\cdot \mathrm{div}_y\big(\nabla_x v(t,x)-\nabla_y v(t,y)\big)\,\eta_R(x)\eta_R(y)\,dx\,dy \\
=& -\frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \nabla_x v(t,x)\big(\nabla_x v(t,x)-\nabla_y v(t,y)\big)\,\eta_R(x)\eta_R(y)\,dx\,dy \\
&+ \frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \nabla_y v(t,y)\big(\nabla_x v(t,x)-\nabla_y v(t,y)\big)\,\eta_R(x)\eta_R(y)\,dx\,dy \\
&- \frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \big(v(t,x)-v(t,y)\big)\big(\nabla_x v(t,x)-\nabla_y v(t,y)\big)\cdot\nabla_x\eta_R(x)\eta_R(y)\,dx\,dy \\
&- \frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \big(v(t,x)-v(t,y)\big)\big(\nabla_x v(t,x)-\nabla_y v(t,y)\big)\cdot\eta_R(x)\nabla_y\eta_R(y)\,dx\,dy \\
=& -\frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \big|\nabla_x v(t,x)-\nabla_y v(t,y)\big|^2\,\eta_R(x)\eta_R(y)\,dx\,dy \\
&- \frac{4}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \big(v(t,x)-v(t,y)\big)\big(\nabla_x v(t,x)-\nabla_y v(t,y)\big)\cdot\nabla_x\eta_R(x)\eta_R(y)\,dx\,dy,
\end{aligned} \tag{3.4}
$$

where we note that the last term in the right-hand side of (3.4) is obtained from the one line above by symmetry of x-y exchanging. Let $0 \le \widetilde{\eta_R(x)} \le 1$ denote a smooth cutoff supported in $B_R(x)$ and that is identically 1 over the support of $\eta_R(x)$. Combining (3.3) and (3.4) and by using $\widetilde{\eta_R(x)}\nabla\eta_R(x) = \nabla\eta_R(x)$ we integrate it over $t \in \big(0, R^2\big)$ to see

$$
\begin{aligned}
&2\int_0^{R^2}\left(\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n} \big|\nabla_x v(t,x)-\nabla_y v(t,y)\big|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)dt \\
&\le -\left[\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n} \big|v(t,x)-v(t,y)\big|^2\eta_R(x)\eta_R(y)\,dx\,dy\right]_{t=0}^{R^2}
\end{aligned}
$$

(this term is negative and we drop it)

$$
\begin{aligned}
&+\varepsilon\int_0^{R^2}\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\nabla_x v(t,x)-\nabla_y v(t,y)\right|^2\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\,dt\\
&+C(\varepsilon)\int_0^{R^2}\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|v(t,x)-v(t,y)\right|^2|\nabla_x\eta_R(x)\eta_R(y)|^2\,dx\,dy\,dt+\left|\int_0^{R^2}II(t)\,dt\right|\\
\le{}&\varepsilon\int_0^{R^2}\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\nabla_x v(t,x)-\nabla_y v(t,y)\right|^2\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\,dt\\
&+C(\varepsilon)\int_0^{R^2}\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|v(t,x)-v(t,y)\right|^2|\nabla_x\eta_R(x)\eta_R(y)|^2\,dx\,dy\,dt+\left|\int_0^{R^2}II(t)\,dt\right|\\
\equiv{}&J_1+J_2+J_3.
\end{aligned}\tag{3.5}
$$

J_1 can be cancelled with the left-hand side by choosing $\varepsilon>0$ small.

Since

$$
\eta_R(x)=\widetilde{\eta_R}(x)\eta_R(x),\quad \nabla_x\eta_R(x)=\widetilde{\eta_R}(x)\nabla_x\eta_R(x),\quad \frac{1}{2^n}|B_R|\le\|\eta_R\|_1\le\|\widetilde{\eta_R}\|_1\le|B_R|,\tag{3.6}
$$

and noticing that $|\nabla\eta_R|\le\frac{C}{R}$ and (3.2), J_2 is estimated as

$$
\begin{aligned}
J_2\le{}&C\int_0^{R^2}\left(\frac{1}{R^2\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\int_0^t\int_0^t\left(e^{(t-s)\Delta}f(s,x)-e^{(t-s)\Delta}f(s,y)\right)\right.\right.\\
&\left.\left.\times\left(e^{(t-r)\Delta}f(r,x)-e^{(t-r)\Delta}f(r,y)\right)dr\,ds\right)\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\right)dt\\
={}&C\frac{1}{R^2\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\int_0^{R^2}\int_0^{R^2}\left(\int_{\max(s,r)}^{R^2}\left(e^{(t-s)\Delta}f(s,x)-e^{(t-s)\Delta}f(s,y)\right)\right.\right.\\
&\left.\left.\times\left(e^{(t-r)\Delta}f(r,x)-e^{(t-r)\Delta}f(r,y)\right)dt\right)dr\,ds\right)\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\\
\le{}&C\int_0^{R^2}\int_0^{R^2}\frac{1}{R^2\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\int_s^{R^2}\left|e^{(t-s)\Delta}f(s,x)-e^{(t-s)\Delta}f(s,y)\right|^2dt\right.\\
&\left.\times\int_r^{R^2}\left|e^{(t-r)\Delta}f(r,x)-e^{(t-r)\Delta}f(r,y)\right|^2dt\right)^{1/2}\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\,dr\,ds.
\end{aligned}\tag{3.7}
$$

Here the integral can be split into r integration and s integration, the right-hand side is expressed by a multiple integration and using the Cauchy–Schwartz inequality and $\widetilde{\eta_R}(x)\le\chi_{B_R(x_0)}(x)$, $2^{-n}|B_R|\le\|\eta_R\|_1$ to have the following: Let the Riesz operator be defined by

$$
R_jf=\mathcal{F}^{-1}\left[\frac{\xi_j}{i|\xi|}\hat{f}\right].\tag{3.8}
$$

Then it follows

$$
\begin{aligned}
J_2 &\le C\left[\frac{1}{R}\int_0^{R^2}\left(\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\int_s^{R^2}\left|e^{(t-s)\Delta}f(s,x)-e^{(t-s)\Delta}f(s,y)\right|^2dt\right)\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\right)^{1/2}ds\right]^2\\
&\le C\int_0^{R^2}\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\int_s^{R^2}\left|e^{(t-s)\Delta}f(s,x)-e^{(t-s)\Delta}f(s,y)\right|^2dt\right)\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\,ds\\
&\le C\sup_{x_0,R>0}\int_0^{R^2}\frac{1}{|B_R|^2}\\
&\quad\times\iint_{B_R(x_0)\times B_R(x_0)}\left(\int_s^{R^2}|\nabla_x\cdot e^{(t-s)\Delta}R_x|\nabla|^{-1}f(s,x)-\nabla_y\cdot e^{(t-s)\Delta}R_y|\nabla|^{-1}f(s,y)|^2\,dt\right)dx\,dy\,ds\\
&\le C\||\nabla|^{-1}f\|^2_{\widetilde{L^2(I;BMO)}},
\end{aligned}\tag{3.9}
$$

where we used the fact that the Riesz operator is bounded in the Chemin–Lerner space (cf. Proposition A.3 in the Appendix). Here we have also used the estimate in Proposition 2.3.

We then consider the term J_3. Regarding

$$
f(t,x)=\nabla_x\cdot R_x|\nabla_x|^{-1}f(t,x)
$$

again it follows that

$$
\begin{aligned}
II &= \frac{2}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}(v(t,x)-v(t,y))\nabla_x\cdot\left(R_x|\nabla_x|^{-1}f(t,x)-R_y|\nabla_y|^{-1}f(t,y)\right)\eta_R(x)\eta_R(y)\,dx\,dy\\
&\quad-\frac{2}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}(v(t,x)-v(t,y))\nabla_y\cdot\left(R_y|\nabla_y|^{-1}f(t,y)-R_x|\nabla_x|^{-1}f(t,x)\right)\eta_R(x)\eta_R(y)\,dx\,dy\\
&=-\frac{2}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}(\nabla_xv(t,x)-\nabla_yv(t,y))\cdot\left(R_x|\nabla_x|^{-1}f(t,x)-R_y|\nabla_y|^{-1}f(t,y)\right)\eta_R(x)\eta_R(y)\,dx\,dy\\
&\quad-\frac{4}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}(v(t,x)-v(t,y))\left(R_x|\nabla_x|^{-1}f(t,x)-R_y|\nabla_y|^{-1}f(t,y)\right)\cdot\nabla_x\eta_R(x)\eta_R(y)\,dx\,dy\\
&\equiv II_1+II_2.
\end{aligned}\tag{3.10}
$$

Integrating the both sides of (3.10) over $t\in(0,R^2)$,

$$
\begin{aligned}
\left|\int_0^{R^2}II(t)\,dt\right| &\le \int_0^{R^2}(|II_1|+|II_2|)\,dt\\
&\le \varepsilon\int_0^{R^2}\left(\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}|\nabla_xv(t,x)-\nabla_yv(t,y)|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)dt\\
&\quad+4\varepsilon^{-1}\int_0^{R^2}\left(\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|R_x|\nabla_x|^{-1}f(t,x)-R_y|\nabla_y|^{-1}f(t,y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)dt
\end{aligned}
$$

$$+\int_0^{R^2} \frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} |v(t,x)-v(t,y)|^2 \big|\nabla_x \eta_R(x)\big|^2 \eta_R(y)\,dx\,dy\,dt$$

$$+\int_0^{R^2} \frac{2}{\|\eta_R\|_1^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \Big|R_x\big|\nabla_x\big|^{-1} f(t,x) - R_y\big|\nabla_y\big|^{-1} f(t,y)\Big|^2 \widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\,dt$$

$$\equiv \varepsilon L_1 + 4\varepsilon^{-1}L_2 + L_3 + L_4. \tag{3.11}$$

The first term L_1 can be cancelled by choosing $\varepsilon > 0$ small enough.

$$4\varepsilon^{-1}L_2 + L_4 \le C \sup_{x_0,R>0} \int_0^{R^2} \left(\frac{1}{\|\eta_R\|_1^2} \int_{B_{2R}(x_0)} \int_{B_{2R}(x_0)} \Big|R_x\big|\nabla_x\big|^{-1} f(t,x) - R_y\big|\nabla_y\big|^{-1} f(t,y)\Big|^2 dx\,dy \right) dt$$

$$\le C_\varepsilon \big\||\nabla_x|^{-1} f\big\|^2_{\widetilde{L^2(I;BMO)}}. \tag{3.12}$$

Using $|\nabla \eta_R| \le \frac{C}{R}$, we estimate the third term L_3 by changing the order of the time integration to see

$$L_3 \le C \int_0^{R^2} \left(\frac{1}{R^2|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left(\int_0^t \big(e^{(t-s)\Delta} f(s,x) - e^{(t-s)\Delta} f(s,y)\big)\,ds \right) \right.$$
$$\left. \times \left(\int_0^t \big(e^{(t-r)\Delta} f(r,x) - e^{(t-r)\Delta} f(r,y)\big)\,dr \right) dx\,dy \right) dt$$

$$= C \int_0^{R^2} \left(\frac{1}{R^2|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left(\int_0^t \int_0^t \big(e^{(t-s)\Delta} f(s,x) - e^{(t-s)\Delta} f(s,y)\big) \right.\right.$$
$$\left.\left. \times \big(e^{(t-r)\Delta} f(r,x) - e^{(t-r)\Delta} f(r,y)\big)\,dr\,ds \right) dx\,dy \right) dt$$

$$\le \frac{1}{R^2} \int_0^{R^2}\int_0^{R^2} \frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left(\int_s^{R^2} \big|e^{(t-s)\Delta} f(s,x) - e^{(t-s)\Delta} f(s,y)\big|^2 dt \right.$$
$$\left. \times \int_r^{R^2} \big|e^{(t-r)\Delta} f(r,x) - e^{(t-r)\Delta} f(r,y)\big|^2 dt \right)^{1/2} dx\,dy\,dr\,ds$$

$$\le C \left[\frac{1}{R} \int_0^{R^2} \left(\frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left(\int_s^{R^2} \big|e^{(t-s)\Delta} f(s,x) - e^{(t-s)\Delta} f(s,y)\big|^2 dt \right) dx\,dy \right)^{1/2} ds \right]^2$$

$$\le C \int_0^{R^2} \frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left(\int_s^{R^2} \big|e^{(t-s)\Delta} f(s,x) - e^{(t-s)\Delta} f(s,y)\big|^2 dt \right) dx\,dy\,ds$$

$$\le C \sup_{x_0,R>0} \int_0^{R^2} \frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left(\int_s^{R^2} \Big|\nabla_x e^{(t-s)\Delta} R_x |\nabla|^{-1} f(s,x) - \nabla_y e^{(t-s)\Delta} R_y |\nabla|^{-1} f(s,y)\Big|^2 dt \right) dx\,dy\,ds$$

$$\le C \big\||\nabla|^{-1} f\big\|^2_{\widetilde{L^2(I;BMO)}}. \tag{3.13}$$

Here we have again used Proposition 2.3. Gathering the estimates (3.5), (3.9), (3.11)–(3.13), we obtain

$$\sup_{x_0,R>0}\int_0^{R^2}\left(\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\nabla_x v(t,x)-\nabla_y v(t,y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)dt \le C\||\nabla|^{-1}f\|^2_{\widetilde{L^2(I;BMO)}}.$$

In particular, by exchanging f into ∇f and noting (A.4) in Lemma A.2, we have

$$\left\|\nabla_x^2\int_0^t e^{(t-s)\Delta}f(s)\,ds\right\|_{\widetilde{L^2(I;BMO)}} \le C\|f\|_{\widetilde{L^2(I;BMO)}}.$$

Combining Proposition 2.1, we conclude the desired estimate. □

We then show the useful estimate of generalized maximal regularity in order to solve the initial value problem of semi-linear type such as the Navier–Stokes equations.

Proposition 3.1. *Let $e^{t\Delta}$ be the heat kernel and let $f\in\widetilde{L^1(I;BMO(\mathbb{R}^n))}$. Then*

$$\left\|\nabla\int_0^t e^{(t-s)\Delta}f(s)\,ds\right\|_{\widetilde{L^2(I;BMO)}} \le C_0\|f\|_{\widetilde{L^1(I;BMO)}}, \tag{3.14}$$

where C_0 is independent of $T>0$.

Corollary 3.2. *Let $e^{t\Delta}$ be the heat kernel, let $1\le\theta\le 2$ and let $f\in\widetilde{L^\theta\Big(I;\dot{BMO}^{-1+2/\theta}(\mathbb{R}^n)\Big)}$. Then*

$$\left\|\nabla\int_0^t e^{(t-s)\Delta}f(s)\,ds\right\|_{\widetilde{L^2(I;BMO)}} \le C_0\|f\|_{\widetilde{L^\theta\left(I;\dot{BMO}^{-1+2/\theta}\right)}}, \tag{3.15}$$

where C_0 is independent of $T>0$.

Corollary 3.2 follows by interpolating between the proofs from Theorem 1.2 and Proposition 3.1.

Proof of Proposition 3.1. For $R>0$ let $B_{2R}=B_{2R}(x_0)$ denote the ball centered at $x_0\in\mathbb{R}^n$ and radius $R>0$. In the similar way to the proof of Theorem 1.2, it follows from (3.3) and (3.4), and regarding $\widetilde{\eta_R(x)}\nabla\eta_R(x)=\nabla\eta_R(x)$ that

$$\begin{aligned}
&2\int_0^{R^2}\left(\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\nabla_x v(t,x)-\nabla_y v(t,y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)dt\\
&\le -\left[\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|v(t,x)-v(t,y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy\right]_{t=0}^{R^2}\\
&\quad+\varepsilon\int_0^{R^2}\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\nabla_x v(t,x)-\nabla_y v(t,y)\right|^2\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\,dt\\
&\quad+C(\varepsilon)\int_0^{R^2}\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|v(t,x)-v(t,y)\right|^2\left|\nabla_x\eta_R(x)\eta_R(y)\right|^2dx\,dy\,dt
\end{aligned}$$

$$
\begin{aligned}
&+\int_0^{R^2}\frac{2}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}(v(t,x)-v(t,y))\cdot(f(t,x)-f(t,y))\eta_R(x)\eta_R(y)\,dx\,dy\,dt\\
&\le \varepsilon\int_0^{R^2}\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\nabla_x v(t,x)-\nabla_y v(t,y)\right|^2\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\,dt\\
&+C(\varepsilon)\int_0^{R^2}\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|v(t,x)-v(t,y)\right|^2\left|\nabla_x\eta_R(x)\eta_R(y)\right|^2\,dx\,dy\,dt\\
&+C\sup_{0<t<R^2}\left(\frac{1}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|v(t,x)-v(t,y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)^{1/2}\\
&\times\int_0^{R^2}\left(\frac{1}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|f(t,x)-f(t,y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)^{1/2}dt\\
&\equiv K_1+K_2+K_3.
\end{aligned}
\tag{3.16}
$$

The term K_1 cancels with the right-hand side. Let $\widetilde{\eta_R}(x)$ be a characteristic function satisfying (3.6). Then the same estimate from (3.7), it follows from $|\nabla\eta_R|\le\frac{C}{R}$ that

$$
\begin{aligned}
K_2&\le C\int_0^{R^2}\left(\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|v(t,x)-v(t,y)\right|^2\left|\nabla\eta_R(x)\eta_R(y)\right|^2\,dx\,dy\right)dt\\
&\le C\left[\int_0^{R^2}\left(\frac{1}{R^2\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\int_s^{R^2}\left|e^{(t-s)\Delta_x}e^{(t-s)\Delta_y}\big(f(s,x)-f(s,y)\big)\right|^2dt\right)\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\right)^{1/2}ds\right]^2.
\end{aligned}
\tag{3.17}
$$

Here the integrant in the right-hand side of (3.17) is estimated by

$$
\begin{aligned}
&\left(\frac{1}{R^2\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\int_s^{R^2}\left|e^{(t-s)\Delta_x}e^{(t-s)\Delta_y}\big(f(s,x)-f(s,y)\big)\right|^2dt\right)\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\right)^{1/2}\\
&\le C\left(\frac{1}{R^2|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}\left(\int_s^{R^2}\left|e^{(t-s)\Delta_x}e^{(t-s)\Delta_y}\chi_{B_{2R}}(x)\big(f(s,x)-\overline{f_{B_R}}(s)\big)\right|^2dt\right)dx\,dy\right)^{1/2}\\
&+C\left(\frac{1}{R^2|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}\left(\int_s^{R^2}\left|e^{(t-s)\Delta_x}e^{(t-s)\Delta_y}\chi_{B_{2R}^c}(x)\big(f(s,x)-\overline{f_{B_R}}(s)\big)\right|^2dt\right)dx\,dy\right)^{1/2}\\
&\le C\left(\frac{1}{R^2|B_R|}\int_{B_R(x_0)}\left(\int_s^{R^2}\left|e^{(t-s)\Delta_x}\chi_{B_{2R}}(x)\big(f(s,x)-\overline{f_{B_R}}(s)\big)\right|^2dt\right)dx\right)^{1/2}\\
&+C\left(\frac{1}{R^2|B_R|}\int_{B_R(x_0)}\left(\int_s^{R^2}\left|e^{(t-s)\Delta_x}\chi_{B_{2R}^c}(x)\big(f(s,x)-\overline{f_{B_R}}(s)\big)\right|^2dt\right)dx\right)^{1/2}\\
&\equiv K_{2,1}+K_{2,2}.
\end{aligned}
\tag{3.18}
$$

The first term of the right-hand side of (3.18) can be estimated by the L^2-bound of the square function (cf. (2.11)). Regarding the heat kernel G_t as Φ_t in (2.11), it follows by $0 < s = t - r \le R^2$ that

$$K_{2,1} \le C\left(\frac{1}{R^2|B_R|}\int_{B_R(x_0)}\int_s^{R^2}\left|e^{(t-s)\Delta_x}\chi_{B_{2R}}(x)\big(f(s,x)-\overline{f_{B_R}}(s)\big)\right|^2 dt\,dx\right)^{1/2}$$
$$\le C\left(\frac{1}{|B_R|}\int_{B_R(x_0)}\int_s^{R^2}\frac{1}{R^2}\left|\int_{\mathbb{R}^n} G_{t-s}(y)\chi_{B_{2R}}(x)\big(f(s,x-y)-\overline{f_{B_R}}(s)\big)\,dy\right|^2 dt\,dx\right)^{1/2}$$
$$\le C\left(\frac{1}{|B_R|}\int_{B_R(x_0)}\int_0^{R^2-s}\frac{r}{R^2}\left|\int_{\mathbb{R}^n} G_r(y)\chi_{B_{2R}}(x)\big(f(s,x-y)-\overline{f_{B_R}}(s)\big)\,dy\right|^2 \frac{dr}{r}\,dx\right)^{1/2}$$
$$\le C\left(\frac{1}{|B_R|}\int_{B_{2R}(x_0)}\left|f(s,x)-\overline{f_{B_R}}(s)\right|^2 dx\right)^{1/2}, \tag{3.19}$$

while for $K_{2,2}$, by changing the variable $r = t - s$,

$$K_{2,2} = C\left(\frac{1}{R^2|B_R|}\int_{B_R(x_0)}\left(\int_s^{R^2}\left|e^{(t-s)\Delta_x}\chi_{B_{2R}^c}(x)\big(f(s,x)-\overline{f_{B_R}}(s)\big)\right|^2 dt\right)dx\right)^{1/2}$$
$$\le C\left(\frac{1}{R^2|B_R|}\int_{B_R(x_0)}\left(\int_s^{R^2}\left(\int_{\mathbb{R}^n} G_{t-s}(x-y)\chi_{B_{2R}^c}(y)\left|f(s,y)-\overline{f_{B_R}}(s)\right|dy\right)^2 dt\right)dx\right)^{1/2}$$
$$\le C\left(\frac{1}{R^2|B_R|}\int_{B_R(x_0)}\left(\int_0^{R^2}\left(\int_{\mathbb{R}^n}\frac{r^{-n/2}}{(1+r^{-1/2}(x-y))^{n+1}}\chi_{B_{2R}^c}(y)\left|f(s,y)-\overline{f_{B_R}}(s)\right|dy\right)^2 dr\right)dx\right)^{1/2}. \tag{3.20}$$

Applying (2.5) in Lemma 2.4 with $\theta = 1$ for $K_{2,2}$, we conclude from (3.19) and (3.20) that

$$K_2 \le C\left[\int_0^{R^2}\left(\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}|f(s,x)-f(s,y)|^2\widetilde{\eta_R}(x)\widetilde{\eta_R}(y)\,dx\,dy\right)^{1/2} ds\right]^2$$
$$\le C\|f\|^2_{\widetilde{L^1(I;BMO)}}. \tag{3.21}$$

Lastly for the estimate of K_3, it suffices to show that

$$\sup_{t\in[0,R^2]}\left(\frac{1}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}|v(t,x)-v(t,y)|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)^{1/2}$$
$$\le C\sup_{x_0,R>0}\int_0^{R^2}\left(\frac{1}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}|f(t,x)-f(t,y)|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)^{1/2} dt. \tag{3.22}$$

Indeed by the Minkowski inequality,

$$
\begin{aligned}
&\left(\frac{1}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}|v(t,x)-v(t,y)|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)^{1/2}\\
&=\left(\frac{1}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\int_0^t\left(e^{(t-s)\Delta_x}f(s,x)-e^{(t-s)\Delta_y}f(s,y)\right)ds\right|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)^{1/2}\\
&\le 2\int_0^t\left(\frac{1}{|B_R|}\int_{\mathbb{R}^n}\left|e^{(t-s)\Delta_x}\left(f(s,x)-\overline{f_{B_R}}(s)\right)\right|^2\eta_R(x)\,dx\right)^{1/2}ds\\
&=2\int_0^t\left(\frac{1}{|B_R|}\int_{\mathbb{R}^n}\left|e^{(t-s)\Delta_x}\left(\chi_{B_{2R}}(x_0)\left(f(s,x)-\overline{f_{B_R}}(s)\right)\right)\right|^2\eta_R(x)\,dx\right)^{1/2}ds\\
&\quad+2\int_0^t\left(\frac{1}{|B_R|}\int_{\mathbb{R}^n}\left|e^{(t-s)\Delta_x}\left(\chi_{B^c_{2R}}(x_0)\left(f(s,x)-\overline{f_{B_R}}(s)\right)\right)\right|^2\eta_R(x)\,dx\right)^{1/2}ds\\
&=K_{3,1}+K_{3,2}.
\end{aligned}
\tag{3.23}
$$

Then L^2 boundedness for the heat semi-group implies that

$$
\begin{aligned}
K_{3,1}&\le 2\int_0^t\left(\frac{1}{|B_R|}\int_{\mathbb{R}^n}\left|e^{(t-s)\Delta_x}\chi_{B_{2R}}(x_0)\left(f(s,x)-\overline{f_{B_R}}(s)\right)\right|^2\eta_R(x)\,dx\right)^{1/2}ds\\
&\le 2\int_0^t\left(\frac{1}{|B_R|}\int_{\mathbb{R}^n}\left|e^{(t-s)\Delta_x}\chi_{B_{2R}}(x_0)\left(f(s,x)-\overline{f_{B_R}}(s)\right)\right|^2dx\right)^{1/2}ds\\
&\le 2\int_0^t\left(\frac{1}{|B_R|}\int_{\mathbb{R}^n}\left|\chi_{B_{2R}}(x_0)\left(f(t,x)-\overline{f_{B_R}}(t)\right)\right|^2dx\right)^{1/2}ds\\
&\le C\sup_{x_0,R>0}\int_0^{R^2}\left(\frac{1}{|B_R|^2}\iint_{B_{2R}(x_0)\times B_R(x_0)}|f(t,x)-f(t,y)|^2\,dx\,dy\right)^{1/2}ds\\
&\le C\|f\|_{\widetilde{L^1(I;BMO)}}.
\end{aligned}
\tag{3.24}
$$

On the other hand, $K_{3,2}$ can be estimated by changing the variables $t-s=r$, and noticing $0<r<R^2$, it follows on the annulus $B_{2^{k+1}R}(x_0)\setminus B_{2^kR}(x_0)$ that for any $x\in B_R(x_0)$,

$$
\frac{r^{1/2}}{\left(r^{1/2}+|x-y|\right)^{n+1}}\le\frac{r^{1/2}}{((2^k-1)R)^{n+1}}\le\frac{C_n r^{1/2}}{(2^kR)^{n+1}}.
$$

Therefore

$$
\begin{aligned}
&\left(\frac{1}{|B_R|}\int_{\mathbb{R}^n}\left|e^{r\Delta_x}\left(\chi_{B^c_R}(x_0)\left(f(t-r,x)-\overline{f_{B_R}}(t-r)\right)\right)\right|^2\eta_R(x)\,dx\right)^{1/2}\\
&=\left(\frac{1}{|B_R|}\int_{B_R(x_0)}\left|\int_{\mathbb{R}^n}\frac{1}{(4\pi r)^{n/2}}\exp\left(-\frac{|x-y|^2}{4r}\right)\left(\chi_{B^c_R}(x_0)\left(f(t-r,y)-\overline{f_{B_R}}(t-r)\right)dy\right|^2dx\right)^{1/2}\\
&\le C\left(\frac{1}{|B_R|}\int_{B_R(x_0)}\left(\int_{\mathbb{R}^n}\frac{r^{-n/2}}{\left(1+r^{-1/2}|x-y|\right)^{n+1}}\left|\chi_{B^c_R}(x_0)\left(f(t-r,y)-\overline{f_{B_R}}(t-r)\right|dy\right)^2dx\right)^{1/2}
\end{aligned}
$$

$$= C\left(\frac{1}{|B_R|}\int_{B_R(x_0)}\left(\sum_{k\ge 0}\int_{B_{2^{k+1}R}(x_0)\setminus B_{2^kR}(x_0)}\frac{r^{1/2}}{(r^{1/2}+|x-y|)^{n+1}}\left|f(t-r,y)-\overline{f_{B_R}}(t-r)\right|dy\right)^2 dx\right)^{1/2}$$

$$\le C\left(\frac{1}{|B_R|}\int_{B_R(x_0)}\left(\sum_{k\ge 0}\frac{2^{-k}r^{1/2}}{R|B_{2^kR}|}\int_{B_{2^{k+1}R}(x_0)\setminus B_{2^kR}(x_0)}\left|f(t-r,y)-\overline{f_{B_R}}(t-r)\right|dy\right)^2 dx\right)^{1/2}$$

$$\le C\sum_{k\ge 1}2^{-k}\frac{r^{1/2}}{R}\left(\frac{1}{|B_{2^kR}|}\int_{B_{2^kR}(x_0)}\left|f(t-r,y)-\overline{f_{B_R}}(t-r)\right|^2 dy\right)^{1/2}. \tag{3.25}$$

Integrating with respect to time of both sides of (3.25), we see by $0\le r\le t\le R^2$ and

$$\left|f(t,x-y)-\overline{f_{B_R}}(t)\right| \le \left|f(t,x-y)-\overline{f_{B_{2^kR}}}(t)\right| + \left|\overline{f_{B_{2^kR}}}(t)-\overline{f_{B_{2^{k-1}R}}}(t)\right| + \cdots + \left|\overline{f_{B_{2R}}}(t)-\overline{f_{B_R}}(t)\right|$$

that

$$K_{3,2} = 2\int_0^t\left(\frac{1}{|B_R|}\int_{\mathbb{R}^n}\left|e^{(t-s)\Delta_x}\left(\chi_{B_{2R}^c}(x_0)\left(f(s,x)-\overline{f_{B_R}}(s)\right)\right)\right|^2\eta_R(x)\,dx\right)^{1/2} ds$$

$$\le C\int_0^t\sum_{k\ge 1}k2^{-k}\frac{r^{1/2}}{R}\left(\frac{1}{|B_{2^kR}|}\int_{B_{2^kR}(x_0)}\left|f(t-r,y)-\overline{f_{B_{2^kR}}}(t-r)\right|^2 dy\right)^{1/2} dr$$

$$\le C\sum_{k\ge 1}k2^{-k}\int_0^t\left(\frac{1}{|B_{2^kR}|}\int_{B_{2^kR}(x-x_0)}\left|f(s,y)-\overline{f_{B_{2^kR}}}(s)\right|^2 dy\right)^{1/2} ds$$

$$\le C\sum_{k\ge 1}k2^{-k}\sup_{x-x_0,R>0}\int_0^{(2^kR)^2}\left(\frac{1}{|B_{2^kR}|}\int_{B_{2^kR}(x-x_0)}\left|f(t,y)-\overline{f_{B_{2^kR}}}(t)\right|^2 dy\right)^{1/2} dt$$

$$\le C\|f\|_{\widetilde{L^1(I;BMO)}}. \tag{3.26}$$

From (3.16), (3.21) and (3.23), (3.24), (3.26), it follows that

$$\sup_{x_0,R>0}\int_0^{R^2}\left(\frac{1}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\nabla_x v(t,x)-\nabla_y v(t,y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy\right)dt \le C\|f\|^2_{\widetilde{L^1(I;BMO)}}.$$

Namely

$$\left\|\nabla_x\int_0^t e^{(t-s)\Delta}f(s)\,ds\right\|_{\widetilde{L^2(I;BMO)}} \le C\|f\|_{\widetilde{L^1(I;BMO)}}. \qquad \square$$

4 | PROOF OF THE SHARP TRACE ESTIMATE

Proof of Theorem 1.5. For the simplicity we show the proof for the case $T=\infty$. Let $f\in L^\infty\cap C^1(\mathbb{R}_+;\mathcal{S}(\mathbb{R}^n))$ with $f(0)\equiv 0$. Let $\eta_R(x)$ be a smooth cut-off function defined in (3.1). Then

$$\frac{d}{dt}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\nabla_x f(t,x)-\nabla_y f(t,y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy$$

$$= 2\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\nabla_x f(t,x)-\nabla_y f(t,y)\right)\cdot\nabla_x\left(\partial_t f(t,x)-\partial_t f(t,y)\right)\eta_R(x)\eta_R(y)\,dx\,dy$$

$$
\begin{aligned}
&-2\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\nabla_x f(t,x)-\nabla_y f(t,y)\right)\cdot\nabla_y\left(\partial_t f(t,y)-\partial_t f(t,x)\right)\eta_R(x)\eta_R(y)\,dx\,dy\\
&=-2\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\Delta_x f(t,x)-\Delta_y f(t,y)\right)\left(\partial_t f(t,x)-\partial_t f(t,y)\right)\eta_R(x)\eta_R(y)\,dx\,dy\\
&\quad-4\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(\nabla_x f(t,x)-\nabla_y f(t,y)\right)\left(\partial_t f(t,x)-\partial_t f(t,y)\right)\nabla_x\eta_R(x)\eta_R(y)\,dx\,dy\\
&=I+II.
\end{aligned}\tag{4.1}
$$

Dividing by $|B_{R/2}|^2$, restricting the integral region on $B_{R/2}$, integrating it over time and then taking a supremum on $R>0$ and x_0, we obtain that the left-hand side is the BMO-norm. Noting $f(0,x)=0$, we have for any $t_0\le R^2$ that

$$
\begin{aligned}
&\frac{1}{|B_{R/2}|^2}\int_{B_{R/2}(x_0)}\int_{B_{R/2}(x_0)}\left|\nabla_x f(t_0,x)-\nabla_y f(t_0,y)\right|^2dx\,dy\\
&\le\int_0^{t_0}\frac{1}{|B_{R/2}|^2}\frac{d}{dt}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\nabla_x f(t,x)-\nabla_y f(t,y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy\,dt\\
&\le\left(\int_0^{t_0}\frac{1}{|B_{R/2}|^2}I\,dt+\int_0^{t_0}\frac{1}{|B_{R/2}|^2}II\,dt\right).
\end{aligned}\tag{4.2}
$$

The first term of the right-hand side is

$$
\begin{aligned}
\int_0^{t_0}\frac{1}{|B_{R/2}|^2}I\,dt\le{}&2^{2n}\int_0^{R^2}\frac{1}{|B_R|^2}\int_{B_R(x_0)}\int_{B_R(x_0)}\left|\Delta_x f(t,x)-\Delta_y f(t,y)\right|^2dx\,dy\,dt\\
&+2^{2n}\left(\int_0^{R^2}\frac{1}{|B_R|^2}\int_{B_R(x_0)}\int_{B_R(x_0)}\left|\partial_t f(t,x)-\partial_t f(t,y)\right|^2dx\,dy\,dt\right).
\end{aligned}\tag{4.3}
$$

Furthermore, the second term of the right-hand side is

$$
\begin{aligned}
\int_0^{t_0}\frac{1}{|B_{R/2}|^2}II\,dt&=4\cdot2^{2n}\int_0^{R^2}\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}\left|\left(\nabla_x f(t,x)-\nabla_y f(t,y)\right)\left(\partial_t f(t,x)-\partial_t f(t,y)\right)\nabla_x\eta_R(x)\eta_R(y)\right|dx\,dy\,dt\\
&\le C\int_0^{R^2}\left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}\left|\nabla_x f(t,x)-\nabla_y f(t,y)\right|^2dx\,dy\right.\\
&\qquad\left.\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}\left|\partial_t f(t,x)-\partial_t f(t,y)\right|^2\left|\nabla_x\eta_R(x)\eta_R(y)\right|^2dx\,dy\right)^{1/2}dt.
\end{aligned}
$$

Applying the Cauchy–Schwartz inequality in t variable and using $|\nabla_x\eta_R(x)|\le\frac{C}{R}$, it follows

$$
\begin{aligned}
\int_0^{t_0}\frac{1}{|B_{R/2}|^2}II\,dt\le{}&4\cdot2^{2n}\left(R^2\int_0^{R^2}\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}\left|\partial_t f(t,x)-\partial_t f(t,y)\right|^2\left|\nabla_x\eta_R(x)\eta_R(y)\right|^2dx\,dy\,dt\right)^{1/2}\\
&\times\sup_{t\in(0,R^2)}\left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)}\left|\nabla_x f(t,x)-\nabla_y f(t,y)\right|^2dx\,dy\right)^{1/2}
\end{aligned}
$$

$$\leq 2C_\varepsilon \int_0^{R^2} \frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left|\partial_t f(t,x) - \partial_t f(t,y)\right|^2 dx\,dy\,dt$$

$$+2\varepsilon \sup_{t\in(0,R^2)} \left(\frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left|\nabla_x f(t,x) - \nabla_y f(t,y)\right|^2 dx\,dy \right). \tag{4.4}$$

In particular, from (4.2)–(4.4)

$$\begin{aligned}\|\nabla f(t_0)\|_{VMO}^2 \leq C_n \sup_{x_0,R>0} \int_0^{R^2} \frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left|\Delta_x f(t,x) - \Delta_y f(t,y)\right|^2 dx\,dy\,dt \\ + C \sup_{x_0,R>0} \int_0^{R^2} \frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left|\partial_t f(t,x) - \partial_t f(t,y)\right|^2 dx\,dy\,dt \\ + 2\varepsilon \sup_{x_0,R>0} \sup_{t\in(0,\infty)} \left(\frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left|\nabla_x f(t,x) - \nabla_y f(t,y)\right|^2 dx\,dy \right).\end{aligned} \tag{4.5}$$

After taking the supremum in $t_0 > 0$ in the left-hand side of (4.5) and then choosing $\varepsilon < 1/4$ and passing the right end term into the left-hand side, we obtain

$$\|\nabla f\|_{BUC(I;VMO)}^2 \leq C\left(\|\Delta f\|_{\widetilde{L^2(I;VMO)}}^2 + \|\partial_t f\|_{\widetilde{L^2(I;VMO)}}^2 \right).$$

Note that the above estimate remains valid for the finite interval case $I = (0,T)\cap(0,R^2)$ with $T < \infty$, since in the inequality (4.4)

$$\min(T^2,R^2)\left|\nabla\eta_R(x)\eta_R(y)\right|^2 \leq \frac{C\min(T^2,R^2)}{R^2} \leq C$$

independent of $R > 0$. Then since $t_0 \in I = (0,T)$ can be taken arbitrary in (4.5), we conclude the estimate (1.17) holds.

To obtain (1.18), we employ the Hahn–Banach extension theorem. The right-hand side of (1.18) can be changed into $L^2(I;BMO)$ as is seen in (1.12) but it is a weaker estimate (which is (1.19)). From (1.18), it holds that

$$\|\nabla f\|_{L^\infty(I;BMO)} \leq C\left(\|\partial_t f\|_{L^2(I;BMO)} + \|\Delta f\|_{L^2(I;BMO)} \right). \qquad \square$$

ACKNOWLEDGMENTS

The authors would like to thank the referees for their careful reading the original manuscript and helpful suggestion on some details. The work of T. Ogawa is partially supported by JSPS Grant in aid for Scientific Research (S) JP19H05597 and Challenging Research (Pioneering) JP20K20284. The work of S. Shimizu is partially supported by JSPS Grant in aid for Scientific Research (B) JP16H03945, (B) JP21H00992 and Fostering Joint International Research JP18KK0072.

ORCID

Senjo Shimizu https://orcid.org/0000-0003-1220-0627

ENDNOTES

[1] Koch–Tataru denotes $\widetilde{L^2}(I : BMO(\mathbb{R}^n))$ by V.
[2] Hence, the measurability remains consistent.
[3] This observation is an indirect proof for the strict inclusion shown in (1.11).

REFERENCES

[1] H. Amann, *Linear and quasilinear parabolic problems. Vol I: Abstract linear theory*, Monogr. Math., vol. 89, Birkhäuser Verlag, Basel–Boston–Berlin, 1995.

[2] H. Amann, *Linear and quasilinear parabolic problems. Vol II: Function spaces*, Monogr. Math., vol. 106, Birkhäuser Verlag, Basel–Boston–Berlin, 2019.
[3] P. Auscher and D. Frey, *On well-posedness of parabolic equations Navier–Stokes type with BMO^{-1} data*, J. Inst. Math. Jussieu **16** (2017), 947–985.
[4] A. Benedek, A. P. Calderón, and R. Panzone, *Convolution operators on Banach space valued functions*, Proc. Natl. Acad. Sci. USA **48** (1962), 356–365.
[5] J. Bergh and J. Löfström, *Interpolation spaces: an introduction*, Springer-Verlag, Berlin, 1976.
[6] M. Bolkart, Y. Giga, T. Suzuki, and Y. Tsutsui, *Equivalence of BMO-type norms with applications to the heat and Stokes semigroup*, Potential Anal. **49** (2018), 105–130.
[7] J. Bourgain and N. Pavlović, *Ill-posedness of the Navier–Stokes equations in a critical space in 3D*, J. Funct. Anal. **255** (2008), 2233–2247.
[8] H. Brezis and L. Nirenberg, *Degree theory and BMO; part I: Compact manifolds without boundaries*, Selecta Math. (N.S.) **1** (1995), no. 2, 197–263.
[9] M. Cannone, *Ondelettes, paraproduits et Navier–Stokes*, Diderot, Arts et Sciences, Paris–New York–Amsterdam, 1995. (French).
[10] M. Cannone and F. Planchon, *Self-similar solutions for Navier–Stokes equations in $\mathbb{R}^3$*, Comm. Partial Differential Equations **21** (1996), 179–193.
[11] J.-Y. Chemin and N. Lerner, *Flot de champ de vecteurs non lipschitziens et équations de Naveir–Stokes*, J. Differential Equations **121** (1995), 314–328. (French).
[12] Ph. Clément and J. Prüss, *Global existence for a semilinear parabolic Volterra equation*, Math. Z. **209** (1992), 17–26.
[13] T. Coulhon and D. Lamberton, *Régularité L^p pour les équations dévolution*, Séminaire d'analyse fonctionelle 1984–85, Publications Mathématiques de l'Université Paris VII, tome 26, 1987, pp. 141–153. (French).
[14] G. Da Prato and P. Grisvard, *Sommes d'opérateurs linéaires et équations différentielles opérationelles*, J. Math. Pures Appl. (9) **54** (1975), 305–387. (French).
[15] R. Denk, M. Hieber, and J. Prüss, *$\mathcal{R}$-boundedness, Fourier multipliers and problems of elliptic and parabolic type*, Mem. Amer. Math. Soc., vol. 788, 2003.
[16] R. Denk, M. Hieber, and J. Prüss, *Optimal L_p-L_q-regularity for parabolic problems with inhomogeneous boundary data*, Math. Z. **257** (2007), 193–224.
[17] L. De Simon, *Un'applicazione della teoria degli integrali allo studio delle equazioni differenziali astratta del primo ordine*, Rend. Semin. Mat. Univ. Padova **34** (1964), 157–162. (Italian).
[18] G. Dore, *L^p regularity for abstract differential equations*, Functional Analysis and Related Topics (H. Komatsu, ed.), Lecture Notes in Math., vol. 1540, Springer, 1993, pp. 25–38.
[19] G. Dore and A. Venni, *On the closedness of the sum of two closed operators*, Math. Z. **196** (1987), 189–201.
[20] X. T. Duong, *H_∞ functional calculus of second order elliptic partial differential operators on L^p spaces*, Miniconference on Operators in Analysis, 1989, Proc. Centre Math. Appl. Austral. Nat. Univ., Canberra, vol. 24, 1990, pp. 91–102.
[21] C. Fefferman and E. M. Stein, *H^p spaces of several variables*, Acta Math. **129** (1972), 137–193.
[22] M. Frazier and B. Jawerth, *A discreate transform and decomposition of distribution spaces*, J. Funct. Anal. **93** (1990), 34–170.
[23] H. Fujita and T. Kato, *On Navier–Stokes initial value problem. I*, Arch. Rational Mech. Anal. **46** (1964), 269–315.
[24] Y. Giga, *Solutions for semilinear parabolic equations in L^p and regularity of weak solutions of the Navier–Stokes system*, J. Differential Equations **61** (1986), 186–212.
[25] Y. Giga and H. Sohr, *Abstract L^p estimates for the Cauchy problem with applications to the Navier–Stokes equations in exterior domains*, J. Funct. Anal. **102** (1991), 72–94.
[26] M. Hieber and J. Prüss, *Heat kernels and maximal L^p-L^q estimates for parabolic evolution equations*, Comm. Partial Differential Equations **22** (1997), 1674–1669.
[27] T. Iwabuchi and M. Nakamura, *Small solutions for nonlinear heat equations, the Navier–Stokes equations, and Keller–Segel system in Besov and Triebel–Lizorkin spaces*, Adv. Differential Equations **7–8** (2013), 687–736.
[28] F. John and L. Nirenberg, *On functions of bounded mean oscillation*, Comm. Pure Appl. Math. **14** (1961), 415–426.
[29] N. J. Kalton and G. Lancien, *A solution to the problem of the L^p-maximal regularity*, Math. Z. **235** (2000), 559–568.
[30] N. J. Kalton and L. Weis, *The H^∞-calculus and sums of closed operators*, Math. Ann. **321** (2001), 319–345.
[31] T. Kato, *Strong L^p-solution of the Navier–Stokes equation in $\mathbb{R}^m$ with applications to weak solutions*, Math. Z. **187** (1984), 471–480.
[32] H. Koch and D. Tataru, *Well-posedness foir the Navier–Stokes equations*, Adv. Math. **154** (2001), 22–35.
[33] P. C. Kunstmann and L. Weis, *Maximal L_p regularity for parabolic equations Fourier multiplier theorems and H^∞-functional calculus*, Lecture Note in Math., vol. 1855, Springer, 2004, pp. 65–311.
[34] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Ural'tseva, *Linear and quasilinear equations of parabolic type*, Transl. Math. Monogr. Amer. Math. Soc., Providence, RI, 1968.
[35] H. Miura and O. Sawada, *On the regularizing rate estimates of Koch–Tataru's solution to the Navier–Stokes equations*, Asymptot. Anal. **49** (2006), 1–15.
[36] T. Ogawa and S. Shimizu, *The drift-diffusion system in the two dimensional critical Hardy space*, J. Funct. Anal. **255** (2008), 1107–1138.
[37] T. Ogawa and S. Shimizu, *End-point maximal regularity and wellposedness of the two dimensional Keller–Segel system in a critical Besov space*, Math. Z. **264** (2010), 601–628.

[38] T. Ogawa and S. Shimizu, *End-point maximal L^1-regularity for the Cauchy problem to a parabolic equation with variable coefficients*, Math. Ann. **365** (2016), 661–705.
[39] T. Ogawa and S. Shimizu, *Maximal regularity for the Stokes equation in BMO and well-posedness of incompressible magnetrohydrodynamics equations with the Hall effect* (in preparation).
[40] T. Ohyama, *Interior regularity of weak solutions of the time-dependent Navier–Stokes equation*, Proc. Japan Acad. **36** (1960), no. 5, 273–277.
[41] F. Planchon, *Global strong solutions in Sobolev or Lebesgue spaces to the incompressible Navier–Stokes equations in* $\mathbf{R}^3$, Ann. Inst. H. Poincaré Anal. Non Linéaire **13** (1996), 319–336.
[42] J. Peetre, *On convolution operators leaving $L^{p,\lambda}$ invariant*, Ann. Mat. Pura Appl. (4) **72** (1966), 295–304.
[43] J. Peetre, *On spaces of Triebel–Lizorkin type*, Ark. Mat. **11** (1975), no. 1–2, 123–130.
[44] G. Prodi, *Un teorema di unicità per le equazioni di Navier–Stokes*, Ann. Mat. Pura. Appl. (4) **48** (1959), 173–182. (Italian).
[45] J. Prüss and G. Simonett, *Moving interfaces and quasi-linear parabolic differential equations*, Monogr. Math., vol. 105, Birkhäuser, Basel, 2016.
[46] J. L. Rubio de Francia, *Martingale and integral transforms of Banach space valued functions*, Probability and Banach Spaces (J. Bastero and M. San Miguel, eds.), Lecture Notes in Math., vol. 1221, Springer Verlag, Berlin, 1986, pp. 195–222.
[47] J. Serrin, *On the interior regularity of weak solutions of the Navier–Stokes equations*, Arch. Rational Mech. Anal. **9** (1962), 187–195.
[48] P. E. Sobolevskii, *Fractional powers of coercively positive sums of operators*, Sibirsk. Mat. Zh., **18** (1977), no. 3, 637–657; English transl. in: Siberian Mat. J., **18** (1977), no. 3, 1323–1358.
[49] V. A. Solonnikov, *On boundary value problems for linear parabolic systems of differential equations of general form*, Trudy Mat. Inst. Steklov, **83** (1965), 3–163. (Russian); Engl. transl. in: Proc. Steklov Inst. Math. **83** (1965), 1–184.
[50] E. Stein, *Harmonic analysis*, Princeton Univ. Press, Princeton, New Jersey 1990.
[51] H. Triebel, *Theory of function spaces*, Birkhäuser, Basel, 1983.
[52] B. Wang, *Ill-posedness for the Navier–Stokes equations in critical Besov spaces* $\dot{B}^{-1}_{\infty,q}$, Adv. Math. **268** (2015), 350–372.
[53] L. Weis, *Operator-valued Fourier multiplier theorems and maximal L_p-regularity*, Math. Ann. **319** (2001), 735–758.
[54] T. Yoneda, *Ill-posedness of the 3D Navier–Stokes equations in a generalized Besov space near* BMO^{-1}, J. Funct. Anal. **258** (2010), 3376–3387.

How to cite this article: T. Ogawa and S. Shimizu, *Maximal regularity for the Cauchy problem of the heat equation in BMO*, Mathematische Nachrichten **295** (2022), 1406–1442. https://doi.org/10.1002/mana.201900506

APPENDIX A

We summarize basic properties of the function class BMO and its Chemin–Lerner class $\widetilde{L^2(I; BMO)}$.

A.1 | A class of the bounded mean oscillation

We mention that the equivalent expression of norms on bounded mean oscillations. First we notice the elementary relation:

Lemma A.1. *Let $1 \le p < \infty$ and let $f \in L^p_{loc}(\mathbb{R}^n; \mathbb{R}^n)$. Then it holds that*

$$\begin{aligned}\frac{1}{|B_R|}\int_{B_R(x_0)} \left|f(x) - \overline{f_{B_R(x_0)}}\right|^p dx \le& \frac{1}{|B_R|^2}\int_{B_R(x_0)}\int_{B_R(x_0)} |f(x) - f(y)|^p \, dx\, dy \\ \le& \frac{4}{|B_R|}\int_{B_R(x_0)} \left|f(x) - \overline{f_{B_R(x_0)}}\right|^p dx,\end{aligned}$$

where

$$\overline{f_{B_R(x_0)}} = \frac{1}{|B_R|}\int_{B_R(x_0)} f(y)\, dy$$

denotes the average of f over $B_R(x_0)$.

Proof of Lemma A .1. Applying the Minkowski and Hölder inequalities, for $f = (f_1(x), f_2(x), \dots, f_n(x))$, it follows

$$\begin{aligned}\frac{1}{|B_R|}\int_{B_R(x_0)} \left|f(x) - \overline{f_{B_R(x_0)}}\right|^p dx &= \frac{1}{|B_R|}\int_{B_R(x_0)} \left(\sum_{k=1}^n \left(f_k(x) - \overline{f_{B_R(x_0)}}\right)^2\right)^{p/2} dx \\ &= \frac{1}{|B_R|}\int_{B_R(x_0)} \left|\sum_{k=1}^n \left(f_k(x) - \frac{1}{|B_R|}\int_{B_R(x_0)} f_k(y)\,dy\right)^2\right|^{p/2} dx \\ &= \frac{1}{|B_R|}\int_{B_R(x_0)} \left|\sum_{k=1}^n \left(\frac{1}{|B_R|}\int_{B_R(x_0)} (f_k(x) - f_k(y))dy\right)^2\right|^{p/2} dx \\ &\le \frac{1}{|B_R|}\int_{B_R(x_0)} \left|\frac{1}{|B_R|}\int_{B_R(x_0)} \left(\sum_{k=1}^n (f_k(x) - f_k(y))^2\right)^{1/2} dy\right|^p dx \\ &= \frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)} |f(x) - f(y)|^p\,dx\,dy.\end{aligned}$$

The other inequality follows by the triangle inequality,

$$\begin{aligned}&\left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)} |f(x) - f(y)|^p\,dx\,dy\right)^{1/p} \\ &\le \left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)} \left|f(x) - \overline{f_{B_R(x_0)}}\right|^p dx\,dy\right)^{1/p} + \left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)} \left|\overline{f_{B_R(x_0)}} - f(y)\right|^p dx\,dy\right)^{1/p}.\end{aligned}$$

□

Let $\eta_R(x)$ be a smooth cut-off function satisfying

$$\chi_{B_{R/2}}(x - x_0) \le \eta_R(x) \le \chi_{B_R}(x - x_0). \tag{A.1}$$

Here $\chi_A(x)$ denotes the characteristic function whose support is a set A. It holds

$$|B_{R/2}| \le \|\eta_R\|_2^2 \le \|\eta_R\|_1 \le |B_R|. \tag{A.2}$$

Then from (A.2), it follows

$$\begin{aligned}\frac{1}{2^n|B_{R/2}|}\int_{B_{R/2}(x_0)} \left|f(x) - \overline{f_{B_R}}\right|^2 dx &= \frac{1}{|B_R|}\int_{B_{R/2}(x_0)} \left|f(x) - \bar{f}_{B_R}\right|^2 dx \\ &\le \frac{1}{\|\eta_R\|_1}\int_{\mathbb{R}^n} \left|f(x) - \overline{f_{B_R}}\right|^2 \eta_R(x)\,dx \\ &\le \frac{1}{|B_{R/2}|}\int_{B_R(x_0)} \left|f(x) - \overline{f_{B_R}}\right|^2 dx \\ &\le 2^n \frac{1}{|B_R|}\int_{B_R(x_0)} \left|f(x) - \overline{f_{B_R}}\right|^2 dx.\end{aligned} \tag{A.3}$$

Thus under the setting (A.1)–(A.2), we have the following:

Lemma A.2. *It holds that*

$$\left(2^{n/2}\right)^{-1}\|f\|_{BMO} \le \sup_{x_0,R>0}\left(\frac{1}{\|\eta_R\|_1}\int_{\mathbb{R}^n}\left|f(x)-\overline{f_{B_R}}\right|^2\eta_R(x)\,dx\right)^{1/2} \le 2^{n/2}\|f\|_{BMO}.$$

Similarly,

$$\begin{aligned}\|f\|_{\widetilde{L^2(I;BMO)}} &\simeq \sup_{x_0,R>0}\left(\int_0^{R^2}\frac{1}{\|\eta_R\|_1}\int_{\mathbb{R}^n}\left|f(x)-\overline{f_{B_R}}\right|^2\eta_R(x)\,dx\,dt\right)^{1/2}\\ &\simeq \sup_{x_0,R>0}\left(\int_0^{R^2}\frac{1}{\|\eta_R\|_1^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|f(t,x)-f(t,y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy\,dt\right)^{1/2}. \end{aligned} \tag{A.4}$$

The most of the properties on BMO are adopted into the class $\widetilde{L^2(I:BMO)}$. For instance the singular integral operator is bounded on this space. In particular, we have the following:

Proposition A.3. *Let R_j be the Riesz operator defined in (3.8). Then there exists a constant $C>0$ such that for any $f\in\widetilde{L^2(I;BMO)}$,*

$$\|R_jf\|_{\widetilde{L^2(I;BMO)}} \le C\|f\|_{\widetilde{L^2(I;BMO)}}.$$

The proof of Proposition A.3 can be shown very much similar way as in the case of BMO. See for the details Peetre [42].

A.2 | Extended John–Nirenberg estimate

We recall the well-known John–Nirenberg estimate.

Definition A.4. A measurable function f on $\mathbb{R}^n$ is *bounded mean oscillation* $BMO(\mathbb{R}^n)$ if

$$\|f\|_{BMO}\equiv \sup_{x\in\mathbb{R}^n,R>0}\frac{1}{|B_R|}\int_{B_R(x)}\left|f(y)-\overline{f_{B_R(x)}}\right|dy<\infty,$$

where $\overline{f_{B_R(x)}}$ is the average of f over the ball B_R;

$$\overline{f_{B_R(x)}}=\frac{1}{|B_R|}\int_{B_R(x)}f(y)\,dy.$$

A typical example of a function belonging to $BMO(\mathbb{R}^n)$ is $\log|x|$. This example and an easy observation show that $L^\infty(\mathbb{R}^n)\subsetneq BMO(\mathbb{R}^n)$.

The following fact is well-known due to John–Nirenberg [28]. For any $1<p<\infty$, let $f\in BMO_p(\mathbb{R}^n)$ stand for

$$\|f\|_{BMO_p}=\sup_{x\in\mathbb{R}^n,R>0}\left(\frac{1}{|B_R|}\int_{B_R(x)}\left|f(y)-\overline{f_{B_R(x)}}\right|^p dy\right)^{1/p}<\infty.$$

Then there exists a constant $C_p>0$ such that

$$\|f\|_{BMO}\le\|f\|_{BMO_p}\le C_p\|f\|_{BMO}.$$

We extend this fact to the Chemin–Lerner space.

Definition A.5. For any $1 \le p < \infty$, $f \in \widetilde{L^2(I;BMO_p(\mathbb{R}^n))}$ if

$$\|f\|_{\widetilde{L^2(I;BMO_p)}} \equiv \sup_{x_0\in\mathbb{R}^n, R>0} \left(\int_0^{R^2} \left(\frac{1}{|B_R|} \int_{B_R(x_0)} \left| f(t,x) - \overline{f_{B_R(x_0)}} \right|^p dx \right)^{2/p} dt \right)^{1/2} < \infty.$$

As we see in Lemma A.1 above, the norms $\|f\|_{\widetilde{L^2(I;BMO_p)}}$ are equivalent to

$$\|f\|_{\widetilde{L^2(I;BMO_{(p)})}} \equiv \sup_{x_0\in\mathbb{R}^n, R>0} \left(\int_0^{R^2} \left(\frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} |f(t,x) - f(t,y)|^p \, dx\, dy \right)^{2/p} dt \right)^{1/2}.$$

We recall the Bochner space $BMO_p(\mathbb{R}^n; L^2(I))$ given by

$$\|f\|_{BMO_p(\mathbb{R}^n;L^2)} \equiv \sup_{x_0\in\mathbb{R}^n, R>0} \left(\frac{1}{|B_R|} \int_{B_R(x_0)} \left(\int_0^{R^2} \left| f(t,x) - \overline{f_{B_R(x_0)}} \right|^2 dt \right)^{p/2} dx \right)^{1/p} < \infty,$$

where $I = (0,T)$ for any $T < \infty$.

Proposition A.6 (John–Nirenberg). *Let* $1 \le p < \infty$.

1. *For* $f \in BMO_p(\mathbb{R}^n; L^2)$*, there exists a constant* $C_p > 0$ *such that*

$$C_p^{-1} \|f\|_{BMO_p(\mathbb{R}^n;L^2)} \le \|f\|_{BMO_1(\mathbb{R}^n;L^2)} \le \|f\|_{BMO_p(\mathbb{R}^n;L^2)}, \tag{A.5}$$

where $C_p \simeq O(p)$ *as* $p \to \infty$.

2. *For* $f \in \widetilde{L^2(I;BMO_p)}$ *with* $1 \le p \le 2$*, there exists a constant* $C_p > 0$ *such that*

$$C_p^{-1} \|f\|_{\widetilde{L^2(I;BMO_p)}} \le \|f\|_{\widetilde{L^2(I;BMO_1)}} \le \|f\|_{\widetilde{L^2(I;BMO_p)}}. \tag{A.6}$$

The right-hand side inequality in (A.6) also holds for all $1 \le p < \infty$.

Remark A.7. If $2 \le p$,

$$\|f\|_{BMO_p(\mathbb{R}^n;L^2)} \le \|f\|_{\widetilde{L^2(I;BMO_p)}}$$

and if $1 \le p \le 2$,

$$\|f\|_{\widetilde{L^2(I;BMO_p)}} \le \|f\|_{BMO_p(\mathbb{R}^n;L^2)},$$

and particularly

$$\|f\|_{\widetilde{L^2(I;BMO_2)}} \simeq \|f\|_{BMO_2(\mathbb{R}^n;L^2)}.$$

Lemma A.8 (Vector valued John–Nirenberg's inequality). *Let* $f \in \widetilde{L^2(I;BMO_1)}$ *and assume that* $\|f\|_{\widetilde{L^2(I;BMO_1)}} \le 1$. *For any cube* $Q^0 \subset \mathbb{R}^n$ *and* $\lambda > 0$*, there exist constants* $\alpha_n > 0$ *and* $C > 0$ *independent of* f *such that*

$$\mu\left(\left\{ x \in Q^0 ; \left\| f(\cdot,x) - \overline{f_{Q^0}} \right\|_{L^2_I} > \lambda \right\} \right) \le C e^{-\alpha_n \lambda}, \tag{A.7}$$

where

$$\|f(\cdot,x)-\bar f_{Q^0}\|_{L^2_I} = \left(\int_0^{R^2} \left|f(t,x)-\overline{f_{Q^0}(t)}\right|^2 dt\right)^{1/2}.$$

In particular, for any cube $Q^0 \subset \mathbb{R}^n$ and $0<\alpha<\alpha_n$,

$$\int_{Q^0}\left(e^{\alpha\left\|f(x)-\overline{f_{Q^0}}\right\|_{L^2_I}}-1\right)dx \le C|Q^0|. \tag{A.8}$$

Remark A.9. The assumption on the function $f\in \widetilde{L^2(I;BMO_1)}$ can be exchanged into f belongs to $BMO_1(\mathbb{R}^n;L^2_t(I))$ and the following proof remains valid with a minor modification.

Proof of Lemma A .8. We show the proof by following the original proof of John–Nirenberg [28]. For $f\in\widetilde{L^2(I;BMO_1)}$, we assume without losing generality that $\|f\|_{\widetilde{L^2(I;BMO_1)}}\le 1$. Let a family of hyper-cubes $\{Q^0_k\}_k$ satisfies

$$\left(\int_0^{R^2}\left(\frac{1}{|Q^0|}\int_{Q^0_{k_0}}\left|f(t,x)-\overline{f_{Q^0}}\right|dx\right)^2 dt\right)^{1/2}\le 1.$$

We then apply the Caldéron–Zygmund decomposition and selection to each of hyper-cube. We consider some Q^0 and find a sequence of hyper-cubes $\left\{Q^m_{k_m}\right\}_{k_m}$ (where m denotes the volume of cubes and k_m denotes the number of the cubes of the same volume). Then let $G_{m-1}=\bigcup_{k_m}Q^m_{k_m}$ and

- For a good region G_m, every cube $Q^m_{k'_m}$ satisfies (where $\left|Q^m_{k'_m}\right|=|Q^m|$)

$$\int_0^{R^2}\left(\frac{1}{|Q^m|}\int_{Q^m_{k'_m}}\left|f(t,x)-\overline{f_{Q^0}}\right|dx\right)^2 dt\le 2^2,\qquad Q^m_{k'_m}\subset G_m, \tag{A.9}$$

  decompose $Q^m_{k'_m}$ further and reduce it into $G=\bigcap_m G_m$.
- For a bad region B_m, every hyper-cube $Q^m_{k_m}$ satisfies

$$\int_0^{R^2}\left(\frac{1}{|Q^m|}\int_{Q^m_{k_m}}\left|f(t,x)-\overline{f_{Q^0}}\right|dx\right)^2 dt > 2^2,\qquad Q^m_{k_m}\subset B_m. \tag{A.10}$$

Then we see that

1. By passing a limit $m\to\infty$ in (A.9), it follows by the Fatou lemma and the Lebesgue differential theorem that for almost all $x\in G=\bigcap_m G_m$,

$$\left(\int_0^{R^2}\left|f(t,x)-\overline{f_{Q^0}}\right|^2 dt\right)^{1/2}\le 2. \tag{A.11}$$

2. For any $x\in B=G^c$, there exists a double sides hyper-cube which is in a good region, namely, for any $Q^m_{k_m}\subset B$, there exists some $Q^{m-1}_{k'_{m-1}}\subset G_{m-1}$ such that $Q^m_{k_m}\subset Q^{m-1}_{k'_{m-1}}$ and from (A.9), (A.10),

$$
\begin{aligned}
2|Q^m| &\le \left(\int_0^{R^2}\left(\int_{Q^m_{k_m}}\left|f(t,x)-\overline{f_{Q^0}}\right|dx\right)^2 dt\right)^{1/2}\\
&\le \left(\int_0^{R^2}\left(\int_{Q^{m-1}_{k'_{m-1}}}\left|f(t,x)-\overline{f_{Q^0}}\right|dx\right)^2 dt\right)^{1/2}\\
&\le 2\left|Q^{m-1}\right| = 2\cdot 2^n|Q^m|.
\end{aligned}
\tag{A.12}
$$

3. For any hyper-cube in the bad region $Q^m_{k_m}$, it follows from (A.10) and $\sum_k a_k^2 \le \left(\sum_k a_k\right)^2$ that

$$
\begin{aligned}
\left(\sum_{m,k_m}\left|Q^m_{k_m}\right|^2\right)^{1/2} &\le \frac{1}{2}\left(\sum_{m,k_m}\int_0^{R^2}\left(\int_{Q^m_{k_m}}\left|f(t,x)-\overline{f_{Q^0}}\right|dx\right)^2 dt\right)^{1/2}\\
&\le \frac{1}{2}\left(\int_0^{R^2}\left(\int_{\bigcup_{m,k_m}Q^m_{k_m}}\left|f(t,x)-\overline{f_{Q^0}}\right|dx\right)^2 dt\right)^{1/2}\\
&\le \frac{1}{2}\left(\int_0^{R^2}\left(\int_{Q^0}\left|f(t,x)-\overline{f_{Q^0}}\right|dx\right)^2 dt\right)^{1/2}.
\end{aligned}
\tag{A.13}
$$

Then we claim that for sufficiently large $\lambda > 0$,

$$
\left\{x\in Q^0;\left\|f(\cdot,x)-\overline{f_{Q^0}}\right\|_{L^2_I} > \lambda\right\} \subset \bigcup_{m,k_m}\left\{x\in Q^m_{k_m};\left\|f(x)-\overline{f_{Q^m_{k_m}}}\right\|_{L^2_I} > \lambda - 2\cdot 2^n\right\}.
\tag{A.14}
$$

Indeed, from (A.10) it suffices to show the case $x\in\bigcup_{k_m,m}Q^m_{k_m}$ in the bad region. Then from (A.11), there exists a cube $Q^m_{k_m}\ni x$ such that from (A.12),

$$
\left\|\overline{f_{Q^m_{k_m}}}-\overline{f_{Q^0}}\right\|_{L^2(I)} \equiv \left(\int_0^{R^2}\left(\frac{1}{\left|Q^m_{k_m}\right|}\int_{Q^m_{k_m}}\left|f(t,x)-\overline{f_{Q^0}}\right|dx\right)^2 dt\right)^{1/2} \le 2\cdot 2^n.
\tag{A.15}
$$

Thus for any $x\in\left\{x\in Q^0;\left\|f(\cdot,x)-\overline{f_{Q^0}}\right\|_{L^2_I} > \lambda\right\}\cap Q^m_{k_m}$,

$$
\begin{aligned}
&\left\|f(x)-\overline{f_{Q^m_{k_m}}}\right\|_{L^2_I} + \left\|\overline{f_{Q^m_{k_m}}}-\overline{f_{Q^0}}\right\|_{L^2_I} \ge \left\|f(x)-\overline{f_{Q^0}}\right\|_{L^2_I} > \lambda,\\
&\left\|f(x)-\overline{f_{Q^m_{k_m}}}\right\|_{L^2_I} > \lambda - 2\cdot 2^n.
\end{aligned}
$$

Hence (A.14) follows. Following the argument due to John–Nirenberg [28], let $F(\lambda)$ defined by

$$
F(\lambda) \equiv \sup_Q \frac{\mu\left(\left\{x\in Q;\left\|f(x)-\overline{f_Q}\right\|_{L^2_I} > \lambda\right\}\right)}{\int_0^{R^2}\left(\int_Q\left|f(t,x)-\overline{f_Q(t)}\right|dx\right)^2 dt}.
$$

Then by (A.14),

$$
\begin{aligned}
&\mu\left(\left\{x \in Q^0; \left\|f(\cdot,x)-\overline{f_{Q^0}}\right\|_{L_I^2} > \lambda\right\}\right)\\
&\le \sum_{m,k_m} \mu\left(\left\{x \in Q^m_{k_m}; \left\|f(x)-\overline{f_{Q^m_{k_m}}}\right\|_{L_I^2} > \lambda - 2\cdot 2^n\right\}\right)\\
&\le \sum_{m,k_m} F\left(\lambda - 2\cdot 2^n\right) \int_0^{R^2} \left(\int_{Q^m_{k_m}} \left|f(t,x)-\overline{f_{Q^m_{k_m}}}\right| dx\right)^2 dt\\
&\le \sum_{m,k_m} F\left(\lambda - 2\cdot 2^n\right) \left|Q^m_{k_m}\right|^2 \int_0^{R^2} \left(\frac{1}{\left|Q^m_{k_m}\right|}\int_{Q^m_{k_m}} \left|f(t,x)-\overline{f_{Q^m_{k_m}}}\right| dx\right)^2 dt\\
&\left(\text{since we assumed that } \|\mathrm{f}\|_{\widetilde{L^2(I;BMO_1)}} \le 1\right)\\
&\le F\left(\lambda - 2\cdot 2^n\right) \sum_{m,k_m} \left|Q^m_{k_m}\right|^2\\
&= F\left(\lambda - 2\cdot 2^n\right) \frac{1}{4} \int_0^{R^2} \left(\int_{Q^0} \left|f(t,x)-\overline{f_{Q^0}}\right| dx\right)^2 dt, \qquad \text{(A.16)}
\end{aligned}
$$

where we use (A.13). The inequality (A.16) holds for all hyper-cube Q^0 which implies

$$
F(\lambda) \le \frac{1}{4} F\left(\lambda - 2\cdot 2^n\right).
$$

For $\alpha_n = 2^{-n}$ and $\lambda > 0$, let $A > 0$ satisfy

$$
F(\lambda) \le A 2^{-\alpha_n \lambda}.
$$

Then we see that

$$
F(\lambda + 2\cdot 2^n) \le \frac{1}{4} F(\lambda) \le A 2^{-\alpha_n(\lambda + 2\cdot 2^n)}, \quad 2^{n+1} \le \lambda < 2\cdot 2^{n+1}.
$$

Then for any integer $m >> 1$, $m \in \mathbb{N}$, for any parmeter $\lambda \in \left[2^{n+1}m, 2^{n+1}(m+1)\right)$,

$$
F(\lambda) \le A 2^{-\alpha_n \lambda}, \quad 2^{n+1} m \le \lambda < 2^{n+1}(m+1).
$$

Varying m, it follows for any large $\lambda >> 1$ that

$$
\begin{aligned}
\mu\left(\left\{x \in Q^0; \left\|f(\cdot,x)-\overline{f_{Q^0}}\right\|_{L_I^2} > \lambda\right\}\right) &\le F(\lambda) \int_0^{R^2} \left(\int_{Q^0} \left|f(t,x)-\overline{f_{Q^0}}\right|^2 dx\right)^2 dt\\
&\le A e^{-\alpha_n' \lambda} \int_0^{R^2} \left(\int_{Q^0} \left|f(t,x)-\overline{f_{Q^0}}\right|^2 dx\right)^2 dt,
\end{aligned}
$$

where $\alpha_n' = 2^{-n} \log 2$. This shows (A.7). The estimate (A.8) follows from (A.7) directly. In fact, for $\alpha < \alpha_n'$.

Hence for any $1 < p \le 2$, we have from (A.8) that

$$
\begin{aligned}
\frac{1}{|Q^0|}\int_{Q^0}\left(e^{\alpha\|f(\cdot,x)-\bar f_{Q^0}\|_{L^2_I}}-1\right)dx &= \frac{1}{|Q^0|}\int_{Q^0}\left(\int_0^{\|f(\cdot,x)-\bar f_{Q^0}\|_{L^2_I}}\frac{d}{d\lambda}e^{\alpha\lambda}\,d\lambda\right)dx\\
&= \frac{1}{|Q^0|}\int_{Q^0}\int_0^\infty \chi_{\left\{x\in Q^0;\|f(\cdot,x)-\bar f_{Q^0}\|_{L^2_I}>\lambda\right\}}(t,x)\alpha e^{\alpha\lambda}\,d\lambda\,dx\\
&= \frac{1}{|Q^0|}\int_0^\infty \mu\left(\left\{x\in Q^0;\left\|f(\cdot,x)-\bar f_{Q^0}\right\|_{L^2_I}>\lambda\right\}\right)\alpha e^{\alpha\lambda}\,d\lambda\\
&\le \frac{1}{|Q^0|}\int_0^\infty \alpha A e^{-(\alpha_n'-\alpha)\lambda}\,d\lambda\\
&\le \frac{\alpha A}{\alpha_n'-\alpha}.
\end{aligned}
\tag{A.17}
$$

□

Proof of Proposition A.6. We only show the left inequality (A.6). The inequality (A.5) follows almost in the same way including $1 \le p < \infty$ after we modify Lemma A.8. Let $1 < p \le 2$. For any $f \in \widetilde{L^2(I;BMO_1)}$, letting

$$
\tilde f = \|f\|^{-1}_{\widetilde{L^2(I;BMO_1)}} f
\tag{A.18}
$$

and without losing generality, we may assume that $\|f\|_{\widetilde{L^2(I;BMO_1)}} = 1$. Using $\lambda^p \le \frac{[p]!}{\alpha^p}(e^{\alpha\lambda}-1) \simeq C_p^p(e^{\alpha\lambda}-1)$,

$$
\left\|f(\cdot,x)-\overline{f_{Q^0}}\right\|^m_{L^2(I)} \le \alpha^{-m} m!\exp\left(\alpha\left\|f(\cdot,x)-\overline{f_{Q^0}}\right\|_{L^2(I)}\right)
$$

with $m = 2\ldots$. From (A.8) in Lemma A.8, the Minkowski inequality yields that

$$
\begin{aligned}
\left(\int_0^{R^2}\left(\frac{1}{|Q^0|}\int_{Q^0}\left|f(t,x)-\overline{f_{Q^0}}\right|^p dx\right)^{2/p}dt\right)^{1/2} &\le \left(\frac{1}{|Q^0|}\int_{Q^0}\left\|f(\cdot,x)-\overline{f_{Q^0}}\right\|^p_{L^2(I)}dx\right)^{1/p}\\
&\le 2\alpha^{-m/p}m^{m/p}\left(\frac{1}{|Q^0|}\int_{Q^0}\exp\left(\alpha\left\|f(\cdot,x)-\overline{f_{Q^0}}\right\|_{L^2(I)}\right)dx\right)^{1/p}\\
&\le 2\alpha^{-m/p}m^{m/p}K^{1/p}.
\end{aligned}
\tag{A.19}
$$

In general, if $\|f\|_{\widetilde{L^2(I;BMO_1)}} \ne 1$ then by (A.18) and (A.19),

$$
\begin{aligned}
\left(\int_0^{R^2}\left(\frac{1}{|Q^0|}\int_{Q^0}\left|f(\cdot,x)-\overline{f_{Q^0}}\right|^p dx\right)^{2/p}\right)^{1/2} &\le C\nu^{-1}K^{1/p}p\sup_{R>0,x_0}\left(\int_0^{R^2}\left(\frac{1}{|Q^0|}\int_{Q^0}\left|f(\cdot,x)-\overline{f_{Q^0}}\right|dx\right)^2\right)^{1/2}\\
&= C_p\|f\|_{\widetilde{L^2(I:BMO_1)}},
\end{aligned}
\tag{A.20}
$$

which implies for all $1 \le p \le 2$,

$$
\|f\|_{\widetilde{L^2(I;BMO_p)}} \le C_p\|f\|_{\widetilde{L^2(I;BMO_1)}}.
\tag{A.21}
$$

This proves (A.6). □

Remark A.10. If we modify the proof of Proposition A.6 into the case $f \in BMO_1(\mathbb{R}^n; L^2_t(I))$, the estimate (A.21) remains valid for all $1 \le p < \infty$ and the constant C_p in (A.21) behaves $C_p \simeq O(p)$ as $p \to \infty$.

A.3 Carleson measure and *BMO*

It is well-understood that a characterization of a function in BMO by the Carleson measure (cf. Stein [50, p. 159, Theorem 3]). We first recall the following estimate:

Let $\Phi(x) \in \mathcal{S}$ be

$$\int_{\mathbb{R}^n} \Phi(x)\,dx = 0$$

and set $\Phi_\lambda(x) \equiv \lambda^{-n}\Phi(\cdot/\lambda)$ and consider

$$\Phi_\lambda * f(x) = \int_{\mathbb{R}^n} \Phi_\lambda(x-y) f(y)\,dy.$$

Theorem A.11 (Stein, Frazier–Jawerth). *For any* $\Phi(x) \in \mathcal{S}$ *with*

$$\int_{\mathbb{R}^n} \Phi(x)\,dx = 0,$$

and $\lambda > 0$, *set* $\Phi_\lambda(x) \equiv \lambda^{-n}\Phi(\cdot/\lambda)$. *Then there exists a constant* $C > 0$ *such that for any* $f \in BMO$

$$\sup_{x_0, R>0} \int_0^R \frac{1}{|B_R|} \int_{B_R} |\Phi_\lambda * f(x)|^2\,dx\,\frac{d\lambda}{\lambda} \le C\|f\|_{BMO}^2. \tag{A.22}$$

This theorem is a characterization of a function of BMO by a Carleson measure. In particular, choosing Φ_λ as $\lambda \simeq 2^{-j}$, Φ_λ can be chosen as the Littlewood–Paley dyadic decomposition of unity $\Phi_\lambda = \phi_j$, then the estimate (A.22) implies the relation between the Lizorkin–Triebel space and BMO. For any $\lambda \in I_j = [2^{-j}, 2^{-j+1})$,

$$\begin{aligned}
\|f\|_{\dot F^0_{\infty,2}} &= \left\| \sup_{k\in\mathbb{Z}} \frac{1}{|B_{2^k}|} \int_{B_{2^k}} \left(\sum_{j\ge -k} |\phi_j * f(x)|^2 \right)^{(1/2)\cdot 2} dx \right\|_\infty \\
&= \sup_{x_0, R>0} \sum_{j\ge -k} \int_{\lambda\in I_j} \frac{1}{|B_{2^k}|} \int_{B_{2^k}} |\phi_j * f(x)|^2\,dx\,\frac{d\lambda}{\lambda} \\
&= \sup_{x_0, R>0} \int_0^R \frac{1}{|B_R|} \int_{B_R} |\Phi_\lambda * f(x)|^2\,dx\,\frac{d\lambda}{\lambda} \\
&\le C\|f\|_{BMO}^2.
\end{aligned} \tag{A.23}$$

For the proof of Theorem A.11, see Stein [50, p. 159, Theorem 3].

As a corollary of the above theorem, one can consider the Koch–Tataru [32] class:

Corollary A.12 [32]. *Let* $e^{t\Delta}$ *be the heat semi-group. Then there exists a constant* $C > 0$ *such that for any* $u_0 \in BMO$,

$$\sup_{x_0, R>0} \int_0^\infty \frac{1}{|B_R|} \int_{B_R} |\nabla e^{t\Delta} u_0|^2\,dx\,dt \le C\|u_0\|_{BMO}^2. \tag{A.24}$$

Proof of Corollary A .12. Let the potential in Theorem A.11 be

$$\Phi_\lambda(x) = \sqrt{t}\nabla G_t(x) \equiv \left(\frac{1}{4\pi t}\right)^{n/2} \left(\frac{x}{2\sqrt{t}}\right) \exp\left(-\frac{|x|^2}{4t}\right).$$

Regarding $\lambda = \sqrt{t}$, $G_t(x)$ is rapidly decaying smooth function and

$$\int_{\mathbb{R}^n} \sqrt{t}\nabla G_t(y)\,dy = \int_{\mathbb{R}^n} \nabla G_1(y)\,dy = 0.$$

Then

$$\frac{1}{2}\sup_{x_0,R>0}\int_0^{R^2}\frac{1}{|B_R|}\int_{B_R}\left|\nabla G_t * u_0(x)\right|^2 dx\,dt = \sup_{x_0,R>0}\int_0^{R^2}\frac{1}{|B_R|}\int_{B_R}\left|\sqrt{t}\nabla G_t * u_0(x)\right|^2 dx\,\frac{dt}{2\sqrt{t}^2}$$
$$= \sup_{x_0,R>0}\int_0^{R}\frac{1}{|B_R|}\int_{B_R}\left|\Phi_\lambda * u_0(x)\right|^2 dx\,\frac{d\lambda}{\lambda} \tag{A.25}$$
$$\leq C\|u_0\|_{BMO}^2. \qquad \square$$

In Theorem A.11, the condition of vanishing average ∇G is essential. If one try to extend such an estimate into a fractional order Laplacian, for example, $s > 0$,

$$\int_{\mathbb{R}^n}|\nabla|^s G_t(x)\,dx = c_n^{-1}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}|\nabla|^s G_t(x)\,dx\bigg|_{\xi=0} = c_n^{-1}|\xi|^s\widehat{G_t}(\xi)\Big|_{\xi=0} = 0.$$

Hence it can be transformed into the fractional power of derivative. We expect that

$$\sup_{x_0,R>0}\frac{1}{|B_R|}\int_{B_R(x_0)}\left(\int_0^\infty \left||\nabla|^{\frac{2}{\sigma}}e^{t\Delta}u_0\right|^\sigma dt\right)^{2/\sigma} dx \leq C\|u_0\|_{BMO}^2$$

and hence

$$\|\Delta e^{t\Delta}u_0\|_{\widetilde{L^\sigma(I;BMO)}} \leq C\||\nabla|^{2/\sigma'}u_0\|_{BMO}$$

are valid.

A.4 | Proof of Proposition 2.1

We give an alternative proof of Koch–Tataru estimate

Proof of Proposition 2.1. Let $\eta_R(x) = 1$ for $|x - x_0| < R$ and $\eta_R(x) = 0$ for $|x - x_0| > 2R$ be a smooth cut-off function. Applying Lemma A.1, we have

$$\frac{1}{|B_R|^2}\frac{d}{dt}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy$$
$$= \frac{2}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right)\left(\partial_t e^{\nu t\Delta}u_0(x) - \partial_t e^{\nu t\Delta}u_0(y)\right)\eta_R(x)\eta_R(y)\,dx\,dy$$
$$= \frac{2\nu}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right)\nabla_x\cdot\left(\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right)\eta_R(x)\eta_R(y)\,dx\,dy$$
$$- \frac{2\nu}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right)\nabla_y\cdot\left(\nabla_y e^{\nu t\Delta}u_0(y) - \nabla_x e^{\nu t\Delta}u_0(x)\right)\eta_R(x)\eta_R(y)\,dx\,dy$$
$$= -\frac{2\nu}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left|\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right|^2\eta_R(x)\eta_R(y)\,dx\,dy$$
$$- \frac{4\nu}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n}\left(e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right)\left(\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right)\cdot\nabla_x\eta_R(x)\eta_R(y)\,dx\,dy$$
$$\equiv -J_1 + J_2. \tag{A.26}$$

For $x \in \operatorname{supp} \nabla\eta_R(x)$, $y \in \operatorname{supp} \eta_R(x)$, $\frac{|x-y|}{R} \leq 2$ and $|\nabla\eta_R(x)| \leq C/R$, integrating the both sides of (A.26) over $t \in (0, R^2)$, we obtain

$$\begin{aligned}
\left|\int_0^t J_2(t)\,dt\right| &= \left|\int_0^t \frac{4\nu}{|B_R|^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \left(e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right)\right. \\
&\qquad \left.\times \left(\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right)\cdot \nabla_x\eta_R(x)\eta_R(y)\,dx\,dy\,dt\right| \\
&\leq 4\nu \int_0^t \left(\frac{1}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n} \left|e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right|^2 \left|\nabla_x\eta_R(x)\eta_R(y)\right|^2 dx\,dy\right)^{1/2} \\
&\qquad \times \left(\frac{1}{|B_R|^2}\iint_{\mathbb{R}^n\times\mathbb{R}^n} \left|\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right|^2 \widetilde{\eta_R}(x)\eta_R(y)\,dx\,dy\right)^{1/2} dt \\
&\leq 4\nu \sup_{t\in(0,R^2)} \left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)} \left|e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right|^2 dx\,dy\right)^{1/2} \\
&\qquad \times \int_0^{R^2} \left(\frac{1}{R^2|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)} \left|\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right|^2 dx\,dy\right)^{1/2} dt \\
&\leq C_\varepsilon \nu \sup_{t\in(0,R^2)} \left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)} \left|e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right|^2 dx\,dy\right) \\
&\qquad + \varepsilon\nu \left(\int_0^{R^2} \frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)} \left|\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right|^2 dx\,dy\,dt\right) \\
&\equiv J_2^1 + \varepsilon \int_0^{R^2} J_1\,dt.
\end{aligned} \tag{A.27}$$

Then

$$\begin{aligned}
J_2^1 &= C_\varepsilon \nu \sup_{t\in(0,R^2)} \left(\frac{1}{|B_R|^2}\iint_{B_R(x_0)\times B_R(x_0)} \left|e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right|^2 dx\,dy\right) \\
&= C_\varepsilon \nu \sup_{t\in(0,R^2)} \left\|e^{\nu t\Delta}u_0\right\|_{BMO}^2 \\
&\leq C_\varepsilon \nu \|u_0\|_{BMO}^2.
\end{aligned} \tag{A.28}$$

The last estimate follows from the Fefferman–Stein $\mathcal{H}^1$-BMO-duality [21] and heat kernel can be regarded as the $\mathcal{H}^1$ test as is shown in

$$\left|\left(e^{\nu t\Delta}u_0, \phi\right)\right| = \left(u_0, e^{\nu t\Delta}\phi\right) \leq C\|u_0\|_{BMO}\left\|e^{\nu t\Delta}\phi\right\|_{\mathcal{H}^1} \leq C\|u_0\|_{BMO}\|\phi\|_{\mathcal{H}^1},$$

$$\sup_{\phi\in\mathcal{H}^1} \frac{\left|\left(e^{\nu t\Delta}u_0, \phi\right)\right|}{\|\phi\|_{\mathcal{H}^1}} \leq C\|u_0\|_{BMO}.$$

Hence

$$\sup_{x_0}\sup_{R>0} J_2^1 = C\nu \left\|e^{\nu t\Delta}u_0\right\|_{BMO}^2 \leq C\nu\|u_0\|_{BMO}^2 \tag{A.29}$$

holds. From (A.26), (A.27) and (A.28), we obtain

$$
\begin{aligned}
&2\nu \int_0^{R^2} \left(\frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left|\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right|^2 dx\, dy \right) dt \\
&= -\left[\frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left|e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right|^2 \eta_R(x)\eta_R(y)\, dx\, dy \right]_{t=0}^{R^2} \\
&\quad - \frac{4\nu}{|B_R|^2} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \left(e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right) \\
&\qquad \times \left(\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right)\cdot \nabla_x \eta_R(x)\eta_R(y)\, dx\, dy \\
&\le \sup_{x_0,R>0} \frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} |u_0(x) - u_0(y)|^2\, dx\, dy \\
&\quad + C_\varepsilon \nu \sup_{t\in(0,R^2)} \left(\frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left|e^{\nu t\Delta}u_0(x) - e^{\nu t\Delta}u_0(y)\right|^2 dx\, dy \right) \\
&\quad + \varepsilon\nu \left(\int_0^{R^2} \frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left|\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right|^2 dx\, dy\, dt \right).
\end{aligned}
\tag{A.30}
$$

Thus the BMO-bound (A.29) for the heat kernel it follows that

$$
\sup_{x_0,R>0} \nu \int_0^{R^2} \left(\frac{1}{|B_R|^2} \iint_{B_R(x_0)\times B_R(x_0)} \left|\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right|^2 dx\, dy \right) dt \le C\|u_0\|_{BMO}^2. \tag{A.31}
$$

□

The left-hand side of the above estimate is a form of the Chemin–Lerner type and one can regard that

$$
\left\| \|\nabla e^{\nu t\Delta}u_0\|_{L^2(\mathbb{R}_+)} \right\|_{BMO}^2 = \sup_{x_0,R>0} \int_0^\infty \frac{1}{|B_R|^2} \int_{B_R(x_0)} \int_{B_R(x_0)} \left|\nabla_x e^{\nu t\Delta}u_0(x) - \nabla_y e^{\nu t\Delta}u_0(y)\right|^2 dx\, dy\, dt.
$$

A.5 | Sharp trace estimate in the Bochner space

We show the outline of the proof for (1.16). For $t \in I = (0, \infty)$, let $f \in W^{1,2}(I; BMO) \cap L^2\left(I; BMO^2\right)$. Then there exists a sequence of smooth functions $\{f_n\} \subset C^1(I; BMO) \cap C^1\left(I; BMO^2\right)$ such that for any $\varepsilon > 0$ and $t \in I$ there exists $N \in \mathbb{N}$ such that for all $n \ge N$,

$$
\|\partial_t f_n - \partial_t f\|_{L^2(I;BMO)} + \|\Delta f_n - \Delta f\|_{L^2(I;BMO)} < \varepsilon, \quad n \ge N_0, \tag{A.32}
$$

and $\lim_{t\to\infty} f_n(t, x) = 0$. Then it suffices to show that

$$
\|\nabla f_n(t)\|_{\dot{B}^0_{\infty,2}} \le C\left(\|\partial_t f_n\|_{L^2(I;BMO)} + \|\Delta f_n\|_{L^2(I;BMO)} \right). \tag{A.33}
$$

Since

$$
\begin{aligned}
\nabla f_n(r, x) &= -\int_r^\infty \partial_s \nabla e^{2(s-r)\Delta} f_n(s, x)\, ds \\
&= -\int_r^\infty \nabla e^{2(s-r)\Delta} \partial_s f_n(s, x)\, ds - \int_r^\infty \nabla e^{2(s-r)\Delta} \Delta f_n(s, x)\, ds
\end{aligned}
$$

and

$$\left(\int_0^\infty \|e^{r\Delta}u\|_1^2 \frac{dr}{r}\right)^{1/\sigma} \le C\|u\|_{\dot{B}^0_{1,2}}$$

(see for instance [38], Proposition 3.3) we choose for any $g \in C_0^\infty(\mathbb{R}^n)$, it holds that

$$\begin{aligned}
|\langle \nabla f_n(r), g\rangle| &\le \left|\left\langle \int_r^\infty \nabla e^{2(s-r)\Delta}\partial_s f_n(s,x)\,ds, g\right\rangle\right| + \left|\left\langle \int_r^\infty \nabla e^{2(s-r)\Delta}\Delta f_n(s,x)\,ds, g\right\rangle\right| \\
&\le \left|\int_r^\infty \left\langle \nabla e^{(s-r)\Delta}\partial_s f_n(s,x), e^{(s-r)\Delta}g\right\rangle ds\right| + \left|\int_r^\infty \left\langle \nabla e^{(s-r)\Delta}\Delta f_n(s,x), e^{(s-r)\Delta}g\right\rangle ds\right| \\
&\le \left|\int_0^\infty \left\langle \nabla e^{s\Delta}\partial_s f_n(s+r,x), e^{s\Delta}g\right\rangle ds\right| + \left|\int_0^\infty \left\langle \nabla e^{s\Delta}\Delta f_n(s+r,x), e^{s\Delta}g\right\rangle ds\right| \\
&\le \left(\int_0^\infty \left\|\sqrt{s}\nabla e^{s\Delta}\partial_s f_n(s,x)\right\|_\infty^2 ds\right)\left(\int_r^\infty s^{-1}\|e^{s\Delta}g\|_1^2\,ds\right) \\
&\quad + \left(\int_0^\infty \left\|\sqrt{s}\nabla e^{s\Delta}\Delta f_n(s,x)\right\|_\infty^2 ds\right)\left(\int_0^\infty s^{-1}\|e^{s\Delta}g\|_1^2\,ds\right) \\
&\le \left(\int_0^\infty \left(\left\|\sqrt{s}\nabla G_s\right\|_{\mathcal{H}^1}\|\partial_s f_n(s)\|_{BMO}\right)^2 ds\right)\left(\int_0^\infty \left(s^{-1/2}\|e^{s\Delta}g\|_1\right)^2 ds\right) \\
&\quad + \left(\int_0^\infty \left(\left\|\sqrt{s}\nabla G_s\right\|_{\mathcal{H}^1}\|\Delta f_n(s)\|_{BMO}\right)^2 ds\right)\left(\int_0^\infty \left(s^{-1/2}\|e^{s\Delta}g\|_1\right)^2 ds\right) \\
&\le C\left(\|\partial_s f_n\|_{L^2(I;BMO)} + \|\Delta f_n\|_{L^2(I;BMO)}\right)\|g\|_{\dot{B}^0_{1,2}}.
\end{aligned}$$

Dividing the both side of above by $\|g\|_{\dot{B}^0_{1,2}}$,

$$\sup_{g\in C_0^\infty\setminus\{0\}} \frac{|\langle \nabla f_n(r), g\rangle|}{\|g\|_{\dot{B}^0_{1,2}}} \le C\left(\|\partial_s f_n\|_{L^2(I;BMO)} + \|\Delta f_n\|_{L^2(I;BMO)}\right)$$

and we conclude that (A.33) holds.