Parabolic equations with singular divergence-free drift vector fields

Zhongmin Qian* and Guangyu Xi†

Abstract

In this paper, we study an elliptic operator in divergence-form but not necessarily symmetric. In particular, our results can be applied to elliptic operator $L = \nu\Delta + u(x,t)\cdot\nabla$, where $u(\cdot,t)$ is a time-dependent vector field in $\mathbb{R}^n$, which is divergence-free in the distributional sense, i.e. $\nabla\cdot u = 0$. Suppose $u \in L^\infty(0,\infty;\mathrm{BMO}^{-1}(\mathbb{R}^n))$. We show the existence of the fundamental solution $\Gamma(x,t;\xi,\tau)$ of the parabolic operator $L-\partial_t$, and show that Γ satisfies the Aronson estimate with a constant depending only on the dimension n, the elliptic constant λ and the norm $\|u\|_{L^\infty_t(\mathrm{BMO}^{-1}_x)}$. Therefore the existence and uniqueness of the parabolic equation $(L-\partial_t)v = 0$ are established for initial data in L^2-space, and their regularity is obtained too. In fact, we establish these results for a general non-symmetric elliptic operator in divergence form.

Key words: Aronson estimate, divergence-free vector field, Harnack inequality, parabolic equation, weak solution

AMS classification: 35K08

1 Introduction

The analysis of the Navier-Stokes equations, which are non-linear partial differential equations describing the motion of incompressible fluids confined in certain spaces, has inspired a large portion of the mathematical analysis of non-linear partial differential equations (see e.g. [19, 20, 24, 33] and etc.) since the fundamental work by Leray [21]. The Navier-Stokes equations are partial differential equations of second-order

$$\frac{\partial}{\partial t}u + u\cdot\nabla u = \nu\Delta u - \nabla p, \tag{1.1}$$

$$\nabla\cdot u = 0, \tag{1.2}$$

subject to the no-slip boundary condition if the domain of fluid is finite, where $u = (u^1,u^2,u^3)$ is the velocity vector field of the fluid flow, $p(x,t)$ is the pressure at the instance t and location x. Leray [21] constructed a weak solution u which belongs to the space $L^\infty\left(0,\infty;L^2(\mathbb{R}^n)\right)$ and also to the space $L^2\left(0,\infty;H^1(\mathbb{R}^n)\right)$. The vorticity ω exists in $L^2_{t,x}$ space and formally, by differentiating the Navier-Stokes equations, solves the vorticity equation

$$\frac{\partial}{\partial t}\omega + u\cdot\nabla\omega = \nu\Delta\omega + \omega\cdot\nabla u, \tag{1.3}$$

*Research supported partly by an ERC grant Esig ID 291244. Mathematical Institute, University of Oxford, OX2 6GG, England. Email: qianz@maths.ox.ac.uk

†This work was supported by the Engineering and Physical Sciences Research Council [EP/L015811/1]. Mathematical Institute, University of Oxford, OX2 6GG, England. Email: guangyu.xi@maths.ox.ac.uk

where the velocity u and the vorticity ω, which is also a time-dependent vector field $\omega = (\omega^1, \omega^2, \omega^3)$, are related by the definition that $\omega = \nabla \times u$. The resolution of the three dimensional Navier-Stokes equations remains to be an open mathematical problem (see e.g. [20, 39]). Most literature in this research area concentrates on the understanding of related partial differential equations and numerical solutions.

The Navier-Stokes equations and the vorticity equation may be put into the following form

$$\left(\frac{\partial}{\partial t} - \nu\Delta + u \cdot \nabla\right) u = -\nabla p, \tag{1.4}$$

and

$$\left(\frac{\partial}{\partial t} - \nu\Delta + u \cdot \nabla\right) \omega = \omega \cdot \nabla u \tag{1.5}$$

respectively, where the diffusion part is the same and involves the following parabolic operator

$$L = \frac{\partial}{\partial t} - \nu\Delta + u \cdot \nabla. \tag{1.6}$$

The elliptic operator $\nu\Delta - u \cdot \nabla$ is the generator of the so-called Taylor diffusion (see Taylor [37, 38]) of the flow of fluids. There are two non-linear terms appearing in the Navier-Stokes equations and the vorticity equation, which determine the turbulent nature of the flow of fluids (see e.g. [22, 23]). The parabolic operator L has the capability of covering the so-called non-linear convection mechanism – the rate-of-strain (for the Navier-Stokes equations [39]) and the vorticity (in the case of the vorticity equation) can be amplified even more rapidly by an increase of the velocity. It is therefore important to study the parabolic equations associated with the elliptic operator $A = \nu\Delta - u \cdot \nabla$, where u is a weak solution of the Navier-Stokes equations. The main feature here is that $u(x,t)$ is a time-dependent vector field with little regularity, which however is solenoidal, that is, for every t, $\nabla \cdot u(\cdot, t) = 0$ in the distributional sense. Hence the formal adjoint $A^* = \nu\Delta + u \cdot \nabla$ is also a diffusion generator. These special features have significance, and have been explored in several recent articles (see e.g. [17, 31, 32]). In this paper we give a thorough study for a class of such parabolic equations.

A vector field $u = (u^i)$ is divergence-free, i.e. $\nabla \cdot u = 0$, implies that its corresponding $(n-1)$-form (with respect to the Hodge star operation) $\star u$ is closed, that is, $d \star u = 0$. In fact the divergence operator $\nabla \cdot$ coincides with the Hodge dual $\star d \star$ up to a sign, where d is the exterior differentiation. Therefore, according to the Poincaré lemma, $\star u$ is exact, that is, there is a $(n-2)$-form b so that $\star u = db$. Hence u coincides with $\star db$ up to a sign. $\star b$ is a form with components b^{ij} where (b^{ij}) is skew-symmetric, and

$$u \cdot \nabla = \sum_{i,j=1} \frac{\partial b^{ij}}{\partial x^i} \frac{\partial}{\partial x^j}.$$

The elliptic operator $\nu\Delta - u \cdot \nabla$ can thus be put into a divergence form

$$\sum_{i,j} \frac{\partial}{\partial x^i} \left(\nu\delta^{ij} - b^{ij}\right) \frac{\partial}{\partial x^j} =: \sum_{i,j} \frac{\partial}{\partial x^i} \left(A^{ij} \frac{\partial}{\partial x^j}\right),$$

where $A = (A^{ij})$ is not necessarily symmetric. The symmetric part $(\nu\delta^{ij})$ is uniformly elliptic, and the skew-symmetric part (b^{ij}) determines the divergence-free drift vector field u.

In the present paper we develop a theory for the linear parabolic equation

$$\sum_{i,j} \frac{\partial}{\partial x^i} \left(A^{ij} \frac{\partial u}{\partial x^j}\right) - \frac{\partial}{\partial t} u = 0$$

under very weak assumptions that $\left(A^{ij}\right)$ is uniformly elliptic and its anti-symmetric part only belongs to the $L^\infty(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$ space (or $L^\infty_t(\mathrm{BMO}_x)$ for short).

The paper is organized as follows. In Section 2, we describe the main result, that is the Aronson estimate which depends only on the elliptic constant and the $L^\infty(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$ norm of the anti-symmetric part of A. The Aronson estimate is the key tool to the study of weak solutions. In Section 3, we provide several results which will be used to prove the Aronson estimate in our setting. These results are interesting by their own, including several versions of the compensated compactness theorem, and a density result of the BMO space. In Section 4, we give the details of the proof of the Aronson estimate, and finally in Section 5, we study the weak solutions to the linear parabolic equations in divergence form (but not necessarily symmetric) under weak assumptions. In particular, we prove the existence and uniqueness of weak solutions to the parabolic equation associated with a non-symmetric diffusion matrix $A=\left(A^{ij}\right)$.

2 Aronson's estimate for non-symmetric parabolic equations

Let us begin with the description of our framework. We consider the following type of linear parabolic equations of second order

$$\frac{\partial}{\partial t}u(x,t)-[A_t u](x,t)=0 \quad \text{on } \mathbb{R}^n\times(0,\infty), \tag{2.1}$$

where

$$A_t=\sum_{i,j=1}^n\frac{\partial}{\partial x^i}\left(A^{ij}(\cdot,t)\frac{\partial}{\partial x^j}\right) \tag{2.2}$$

is in divergence form but *not necessarily symmetric*, and their associated diffusion processes in terms of fundamental solutions defined by (2.1). There is a unique decomposition $A(x,t)=a(x,t)+b(x,t)$ such that $a(x,t)=\left(a^{ij}(x,t)\right)$ is symmetric, while $b(x,t)=\left(b^{ij}(x,t)\right)$ is skew-symmetric. We assume that A is uniformly elliptic in the following sense: there exists a constant $\lambda>0$ such that

$$\lambda|\xi|^2\le\sum_{i,j=1}^n a^{ij}(x,t)\xi_i\xi_j=\sum_{i,j=1}^n A^{ij}(x,t)\xi_i\xi_j\le\frac{1}{\lambda}|\xi|^2 \tag{2.3}$$

for any $\xi=(\xi_i)\in\mathbb{R}^n$, $x\in\mathbb{R}^n$ and $t\ge0$.

Let us first consider the regular case where A^{ij} are smooth, bounded and possess bounded derivatives of all orders on $\mathbb{R}^n\times[0,\infty)$.

Let $L=A_t-\frac{\partial}{\partial t}$ be the parabolic linear operator associated with $\left(A^{ij}(x,t)\right)$. The formal adjoint of L is again a parabolic operator (with vanished zero order term) given by

$$L^\star=\sum_{i,j=1}^n\frac{\partial}{\partial x^i}\left(A^{ji}(\cdot,t)\frac{\partial}{\partial x^j}\right)+\frac{\partial}{\partial t}, \tag{2.4}$$

where $A^{ji}=a^{ij}-b^{ij}$ is the transpose of $\left(A^{ij}\right)$. It is known that (see Friedman [14], Theorem 11 and 12, Chapter 1), under the elliptic condition and smoothness assumptions on $A^{ij}(x,t)$, there is a unique positive *fundamental solution* $\Gamma(x,t;\xi,\tau)$ of the parabolic operator L, and it is smooth in (x,t,ξ,τ) on $0\le\tau<t<\infty$ and $(x,\xi)\in\mathbb{R}^n\times\mathbb{R}^n$. Recall that the following properties are satisfied.

1) $\Gamma(x,t;\xi,\tau)>0$ for any $0\le\tau<t$ and $x,\xi\in\mathbb{R}^n$.

2) For every $\xi \in \mathbb{R}^n$ and $\tau \in [0,\infty)$, as a function of $(x,t) \in \mathbb{R}^n \times (\tau,\infty)$, $u(x,t) \equiv \Gamma(x,t;\xi,\tau)$ solves the parabolic equation $Lu = 0$ on $\mathbb{R}^n \times (\tau,\infty)$:

$$\sum_{i,j=1}^{n} \frac{\partial}{\partial x^i}\left(A^{ij}(x,t)\frac{\partial}{\partial x^j}\Gamma(x,t;\xi,\tau)\right) - \frac{\partial}{\partial t}\Gamma(x,t;\xi,\tau) = 0 \qquad \text{on } \mathbb{R}^n \times (\tau,\infty). \tag{2.5}$$

3) Chapman-Kolmogorov's equation holds

$$\Gamma(x,t;\xi,\tau) = \int_{\mathbb{R}^n} \Gamma(x,t;z,s)\Gamma(z,s;\xi,\tau)dz. \tag{2.6}$$

4) For any bounded continuous function f and $\tau \in [0,\infty)$, it holds that

$$\lim_{t\downarrow\tau} \int_{\mathbb{R}^n} f(\xi)\Gamma(x,t;\xi,\tau)d\xi = f(x) \tag{2.7}$$

for every $x \in \mathbb{R}^n$.

For $0 \le \tau < t$, let $\Gamma_{\tau,t}$ denote the corresponding linear operator defined by

$$\Gamma_{\tau,t}f(x) = \int_{\mathbb{R}^n} f(\xi)\Gamma(x,t;\xi,\tau)d\xi, \tag{2.8}$$

where f is Borel measurable, either non-negative, or/and bounded. By (2.6)

$$\Gamma_{s,t} \circ \Gamma_{\tau,s} = \Gamma_{\tau,t} \tag{2.9}$$

for any $0 \le \tau < s < t$.

Define $\Gamma^\star(x,s;y,t) = \Gamma(y,t;x,s)$ for $t > s \ge 0$. Then $\Gamma^\star$ is the fundamental solution to $L^\star v = 0$ in the sense that for every fixed (y,t), as a function of (x,s), $\Gamma^\star$ solves the *backward* parabolic equation

$$\sum_{i,j=1}^{n} \frac{\partial}{\partial x^i}\left(A^{ji}(x,s)\frac{\partial}{\partial x^j}\Gamma^\star(x,s;y,t)\right) + \frac{\partial}{\partial s}\Gamma^\star(x,s;y,t) = 0 \tag{2.10}$$

on $(x,s) \in \mathbb{R}^n \times [0,t)$. It follows that the fundamental solution Γ also solves the backward parabolic equation:

$$\sum_{i,j=1}^{n} \frac{\partial}{\partial \xi^i}\left(A^{ji}(\xi,\tau)\frac{\partial}{\partial \xi^j}\Gamma(x,t;\xi,\tau)\right) + \frac{\partial}{\partial \tau}\Gamma(x,t;\xi,\tau) = 0, \tag{2.11}$$

which holds for any $\xi, x \in \mathbb{R}^n$ and $0 < \tau < t$.

We are now in a position to state the main result of the present paper.

Theorem 2.1 *There is a constant $M > 0$ depending only on the dimension n, the elliptic constant $\lambda > 0$, and the $L^\infty(0,\infty;BMO(\mathbb{R}^n))$ norm of the skew-symmetric part $b^{ij} = \frac{1}{2}\left(A^{ij} - A^{ji}\right)$ such that*

$$\frac{1}{M(t-\tau)^{n/2}} \exp\left[-\frac{M|x-\xi|^2}{t-\tau}\right] \le \Gamma(x,t;\xi,\tau) \le \frac{M}{(t-\tau)^{n/2}} \exp\left[-\frac{|x-\xi|^2}{M(t-\tau)}\right] \tag{2.12}$$

for any $0 \le \tau < t < \infty$ and $(x,\xi) \in \mathbb{R}^n \times \mathbb{R}^n$, where the $L^\infty(0,\infty;BMO(\mathbb{R}^n))$ norm of b is defined by

$$\|b\|_{L_t^\infty(BMO_x)} = \sup_{t\ge 0} \sqrt{\sum_{i<j} \|b^{ij}(\cdot,t)\|_{BMO}^2}.$$

The heat kernel estimate (2.12) for parabolic equations has a long history. Two sided estimate (2.12) was first established in Aronson [1, 2] for uniformly elliptic operators in divergence form where A^{ij} is symmetric (so that $b^{ij} \equiv 0$) and his constant M depends only on the elliptic constant λ and the dimension n. Estimate (2.12) is therefore referred to as the Aronson estimate. A weaker but global estimate similar to (2.12) under the same assumption as in Aronson [2] already appeared in the appendix of Nash [29]. Aronson [2, 3] indicated that his estimate can be established for a general elliptic operator, and a proof is available in Fabes and Stroock [12], Stroock [34] and Norris and Stroock [30] too. In these papers, the Aronson estimate (2.12) was established for the following type of uniformly elliptic operator

$$\sum_{i,j=1}^{n} \frac{\partial}{\partial x^i} a^{ij}(x,t) \frac{\partial}{\partial x^j} + \sum_{i,j=1}^{n} a^{ij}(x,t) b_j(x,t) \frac{\partial}{\partial x^i} - \frac{\partial}{\partial x^i} \left(a^{ij}(x,t) \hat{b}_j(x,t) \right) + c(x,t),$$

where $\left(a^{ij}\right)$ is symmetric and uniformly elliptic. In this case, the constant M depends on the dimension, the elliptic constant λ and the $L^\infty_{t,x}$-norms of b, $\hat{b}$ and c.

A related topic to the Aronson estimate is the regularity of solutions to the parabolic equation $Lu = 0$ (see [19] for a complete survey of classical results). If the elliptic operator is symmetric and is in divergence form, it was Nash [29] who proved the Hölder continuity of bounded solutions and also proved that the Hölder exponent depends only on the dimension and the elliptic constant λ. Under the same setting as that of Nash [29], in 1964, Moser [26] established the Harnack inequality for positive solutions of the parabolic equation $Lu = 0$, based on which Aronson was able to derive his estimate (2.12). Fabes and Stroock [12] showed that Moser's Harnack inequality can be derived from Aronson estimate together with Nash's idea, and Stroock [34] further demonstrated that both the Hölder continuity of classical solutions and Moser's Harnack inequality for positive solutions can be established by utilizing the two sided Aronson estimate (2.12). Nash's idea in [29] and the techniques in Moser [26, 27, 28] have been investigated intensively during the past decades. Many excellent results have been obtained in more general settings, but mainly under the symmetric setting of Dirichlet forms [15]. See for example Grigor'yan [16], Davies [8] and Stroock [35] for a small sample of references, and see also the literature therein.

The case that $\left(A^{ij}\right)$ is non-symmetric has received intensive study only recently, due to the connection with the Navier-Stokes equations and the blow-up behavior of their solutions. In Osada [31], the Aronson estimate (2.12) was obtained for an elliptic operator in divergence form, where $\left(A^{ij}\right)$ may not be symmetric, his constant M in (2.12) however depends on the dimension n, the elliptic constant λ and the $L^\infty_{t,x}$-norm of the skew-symmetric part (b^{ij}). In recent works by Seregin, Silvestre, Šverák, and Zlatoš [32] and Friedlander, Vicol [13], they both relaxed the condition on (b^{ij}) from $L^\infty_{t,x}$ to $L^\infty(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$ using De Giorgi's technique. In [13], the Harnack inequality is proved under assumptions that (b^{ij}) is in $L^\infty(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$ and satisfies $\nabla \cdot b \in L^2_t L^2_x$. In [32], the main result is the Harnack inequality for weak solutions satisfying an additional local energy inequality. Also motivated by incompressible fluid, another closely related problem is regarding the family of operators with fractional Laplacian $(-\Delta)^\alpha$ (instead of Laplacian) and divergence free drift u, i.e.

$$L = \frac{\partial}{\partial t} + (-\Delta)^\alpha + u \cdot \nabla.$$

Caffarelli and Vasseur [4] proved that when $\alpha = \frac{1}{2}$, the corresponding parabolic equations with $u \in L^\infty(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$ has Hölder continuous solution for any $L^2(\mathbb{R}^n)$ initial data. Later in Delgadino and Smith [10], similar results are proved for $\alpha \in (\frac{1}{2}, 1)$ with $u \in L^q(0,\infty;\mathrm{BMO}^{-\gamma}(\mathbb{R}^n)$,

$\frac{2\alpha}{q}+\gamma=2\alpha-1$, $\gamma\in(0,2\alpha-1]$. Notice that when $\alpha=1$, the diffusion part is Laplacian and $u\in L^\infty(0,\infty;\mathrm{BMO}^{-1}(\mathbb{R}^n))$ can be regarded as a limit case. By the time of writing up this paper, this is the first work that proved the Aronson estimate together with the Harnack inequality and the existence of a unique weak solution only assuming that $b\in L^\infty(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$. The relaxation from $L^\infty_{t,x}$ to $L^\infty(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$ here is non-trivial, which is also exposed in [13] and [32]. We take a different approach from the techniques used in [13] and [32], which is to estimate the fundamental solution and relies on several versions of compensated compactness results. In proving the upper bound of the Aronson estimate, we used essentially Proposition 3.2, which is in the same spirit as the classical compensated compactness results and is new to the knowledge of the authors.

Moreover, it is mentioned in [32] that the fundamental solution Γ of operator

$$\Delta+u(x,t)\cdot\nabla-\frac{\partial}{\partial t}$$

satisfies the diagonal decay estimate

$$\Gamma(x,t;x,\tau)\leq\frac{C}{(t-\tau)^{\frac{n}{2}}}$$

for all $t>\tau\geq 0$. Our work was motivated by the observation made by Seregin and etc. [32] to obtain a better estimate of Γ, and the approach put forward by Davies [7], Fabes and Stroock [12], and Stroock [34]. We follow the approach in Davies and Stroock to work on the non-symmetric case, and adopt their arguments to our case by overcoming the difficulties arising from the singularities of the skew-symmetric part (b^{ij}).

As consequences of the Aronson estimate, we have the following continuity theorem and the Harnack inequality as in Stroock [34].

Theorem 2.2 *There exist $C>0$ and $\alpha\in(0,1)$ depending only on the dimension n, the elliptic constant λ and the $L^\infty(0,\infty;BMO(\mathbb{R}^n))$-norm of the skew-symmetric part (b^{ij}), such that for every $\delta>0$*

$$|\Gamma(x,t;\xi,\tau)-\Gamma(x',t';\xi',\tau)|\leq\frac{C}{\delta^n}(|t'-t|\vee|x'-x|\vee|\xi'-\xi|)^\alpha \tag{2.13}$$

for all $\tau\geq 0$, $(t',x',\xi'),(t,x,\xi)\in[\tau+\delta^2,\infty)\times\mathbb{R}^n\times\mathbb{R}^n$ with $|t'-t|\vee|x'-x|\vee|y'-y|\leq\delta$.

Theorem 2.3 *[The Harnack Inequality] There exists a constant $C>0$ depending only on n,λ and $\|b\|_{L^\infty_t(BMO_x)}$ such that given any non-negative $v\in L^2(\mathbb{R}^n)$ and set $u(t,x)=\Gamma_{\tau,t}v(x)$, we have*

$$\sup_{[s,s+R^2]\times B(x_0,R)}u(t,x)\leq C\inf_{[s+3R^2,s+4R^2]\times B(x_0,R)}u(t,y) \tag{2.14}$$

for any $R>0$, $(x,s)\in\mathbb{R}^n\times[\tau,\infty)$.

3 Several technical facts

In this and next several sections, we are going to prove the main result Theorem 2.1. In this section, we prove several technical facts which will be needed in the proof of the main result.

The first result we need is a variation of Coifman-Meyer's compensated compactness theorem [5, 6, 20] which highlights the importance of the Hardy spaces in the study of partial differential equations.

We first recall some facts on BMO functions [18, 33]. A function f is in $\mathrm{BMO}(\mathbb{R}^n)$ if

$$\|f\|_{\mathrm{BMO}} = \sup_{B\subset\mathbb{R}^n} \frac{1}{|B|}\int_B |f(x)-f_B|\,dx < \infty, \tag{3.1}$$

where $f_B = \frac{1}{|B|}\int_B f(x)\,dx$ and the supremum is taken over all open balls $B\in\mathbb{R}^n$ (in what follows, $B_r(x)$ or $B(x,r)$ denotes the ball centered at x with radius r). If define another norm

$$\|f\|^p_{\mathrm{BMO}_p} = \sup_{B\subset\mathbb{R}^n} \frac{1}{|B|}\int_B |f(x)-f_B|^p\,dx < \infty \tag{3.2}$$

for any $1\le p<\infty$, the John-Nirenberg inequality [18] (see e.g. Appendix in Stroock and Varadhan [36]) implies that $\|\cdot\|_{\mathrm{BMO}_p}$ are equivalent for different p.

The original version of the compensated compactness theorem, which will be used in our proof of the lower bound in the Aronson estimate, can be stated as follows.

Proposition 3.1 *Let vector fields E,B satisfy that $E\in L^p(\mathbb{R}^n)^n$, $B\in L^q(\mathbb{R}^n)^n$ with $\frac{1}{p}+\frac{1}{q}=1$ ($p\ge 1$, $q\ge 1$) and $\nabla\cdot E=0$, $\nabla\times B=0$. Then $E\cdot B\in\mathscr{H}^1$ where $\mathscr{H}^1$ is the Hardy space, and*

$$\|E\cdot B\|_{\mathscr{H}^1}\le C\|E\|_p\|B\|_q. \tag{3.3}$$

In particular, there is a constant C depending on the dimension n and $p>1$ such that

$$\|\nabla f\times\nabla g\|_{\mathscr{H}^1}\le C\,\|\nabla f\|_p\,\|\nabla g\|_q \tag{3.4}$$

for any $f,g\in C_0^\infty(\mathbb{R}^n)$, where $\frac{1}{p}+\frac{1}{q}=1$. Hence

$$\left|\int_{\mathbb{R}^n}\langle\nabla f(x), b(x)\cdot\nabla g(x)\rangle\,dx\right|\le C\,\|b\|_{BMO}\,\|\nabla f\|_2\,\|\nabla g\|_2 \tag{3.5}$$

for any $f,g\in H^1(\mathbb{R}^n)$ and for any $b=\left(b^{ij}\right)\in BMO$ which is skew-symmetric, i.e. $b^{ij}=-b^{ji}$.

To prove the upper bound of the Aronson estimate, we need the following estimate, which is in the same spirit as the compensated compactness theorem above.

Proposition 3.2 *There is a universal constant C depending only on the dimension n, such that*

$$\|f\nabla_\xi f\|_{\mathscr{H}^1}\le C|\xi|\|\nabla f\|_2\|f\|_2. \tag{3.6}$$

for any $f\in H^1(\mathbb{R}^n)=W^{1,2}(\mathbb{R}^n)$ and $\xi\in\mathbb{R}^n$, where $\|\cdot\|_{\mathscr{H}^1}$ denotes the Hardy norm.

Proof. Let h be any smooth non-negative function on $\mathbb{R}^n$, with its support in the unit ball $B_1(0)$ such that $\int_{\mathbb{R}^n}h(x)dx=1$, and $h_t(x)=t^{-n}h(x/t)$ for $t>0$. Notice that $f\nabla_\xi f=\frac{1}{2}\nabla\cdot(f^2\xi)$ in $L^1(\mathbb{R}^n)$, so

$$\begin{aligned}
h_t*(f\nabla_\xi f)(x) &= \frac{1}{2}\int_{B_t(x)}\nabla h_t(x-y)\cdot\xi f^2(y)dy\\
&= \frac{1}{2}\int_{B_t(x)}\frac{1}{t^{n+1}}\nabla h\left(\frac{x-y}{t}\right)\cdot\xi f^2(y)dy\\
&= \frac{1}{2}\int_{B_t(x)}\frac{1}{t^{n+1}}\nabla h\left(\frac{x-y}{t}\right)\cdot\xi f(y)\left[f(y)-\fint_{B_t(x)}f\right]dy
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}\int_{B_t(x)}\frac{1}{t^{n+1}}\nabla h\left(\frac{x-y}{t}\right)\cdot\xi f(y)\left(\fint_{B_t(x)} f\right)dy\\
=\;& I_1+I_2,
\end{aligned}
$$

where $\fint_{B_t(x)}$ denotes the average integral over the ball $B_t(x)$, that is, $|B_t(x)|^{-1}\int_{B_t(x)}$. For the first term on the right-hand side, we have

$$
|I_1|\le C\left[\fint_{B_t(x)}|\xi f|^{\alpha}\right]^{\frac{1}{\alpha}}\left[\fint_{B_t(x)}\left|\left(f-\fint_{B_t(x)}f\right)t^{-1}\right|^{\alpha'}\right]^{\frac{1}{\alpha'}}, \tag{3.7}
$$

where $\alpha\in[1,2)$, $\frac{1}{\alpha}+\frac{1}{\alpha'}=1$. Choose α,β such that $1\le\alpha,\beta<2$ and $\frac{1}{\alpha}+\frac{1}{\beta}=1+\frac{1}{n}$. Then by the Sobolev-Poincaré inequality, we have

$$
\left[\fint_{B_t(x)}\left|\left(f-\fint_{B_t(x)}f\right)t^{-1}\right|^{\alpha'}\right]^{\frac{1}{\alpha'}}\le C\left(\fint_{B_t(x)}|\nabla f|^{\beta}\right)^{\frac{1}{\beta}}. \tag{3.8}
$$

For the second term on the right-hand side, we integrate by part and recall that $\frac{1}{t^{n+1}}\nabla h(\frac{x-y}{t})=\nabla h_t(x-y)$ to obtain

$$
|I_2|=\left|\frac{1}{2}\int_{B_t(x)}\frac{1}{t^n}h\left(\frac{x-y}{t}\right)\cdot\mathrm{div}(\xi f(y))\left(\fint_{B_t(x)}f\right)dy\right|\le C|\xi|\fint_{B_t(x)}|\nabla f(y)|dy\left(\fint_{B_t(x)}|f|\right). \tag{3.9}
$$

By using these estimates, we thus conclude that

$$
\begin{aligned}
\sup_{t>0}|\{h_t*(\xi f\cdot\nabla f)\}(x)| \;&\le\; C|\xi|\sup_{t>0}\left(\fint_{B_t(x)}|f|^{\alpha}\right)^{\frac{1}{\alpha}}\sup_{t>0}\left(\fint_{B_t(x)}|\nabla f|^{\beta}\right)^{\frac{1}{\beta}}\\
&\quad+C|\xi|\sup_{t>0}\left(\fint_{B_t(x)}|f|\right)\sup_{t>0}\left(\fint_{B_t(x)}|\nabla f|\right)\\
&=\; C|\xi|[M(|f|^{\alpha})^{\frac{1}{\alpha}}M(|\nabla f|^{\beta})^{\frac{1}{\beta}}+M(|f|)M(|\nabla f|)],
\end{aligned}
$$

where $M(f)$ is the maximal function. Since $1\le\alpha<2$, $1<\beta<2$, we have $\|M(|f|^{\alpha})^{\frac{1}{\alpha}}\|_2\le C\|f\|_2$, $\|M(|\nabla f|^{\beta})^{\frac{1}{\beta}}\|_2\le C\|\nabla f\|_2$ and similarly $\|M(|f|)\|_2\le C\|f\|_2$, $\|M(|\nabla f|)\|_2\le C\|\nabla f\|_2$. So $\sup_{t>0}|h_t*(\xi f\cdot\nabla f)|\in L^1$ and

$$
\|f\cdot\nabla_{\xi}f\|_{\mathscr{H}^1}\le C|\xi|\|\nabla f\|_2\|f\|_2. \tag{3.10}
$$

■

Given a function $b\in L^{\infty}(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$, we want to approximate it by a mollified sequence, which is not trivial as it looks. A simple example is a vector field $b(t)$ which depends only on t, not on the space variables. Then it may not be in L^1_{loc} and there is no approximation of b by mollifying sequences. However, the problem considered here allows us to add a constant to it, i.e. consider $b(t,x)+f(t)$, where $f(t)$ is skew-symmetric so that it will not alter the weak solution formulation of the corresponding parabolic equations. So by subtracting the mean value of b on the unit ball $B_1(0)$, we may assume that the vector field $b \in L^\infty(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$ in addition satisfies that

$$b_{B_1(0)}(t) = \fint_{B_1(0)} b(t,x)\,dx = 0 \qquad \text{for all } t \in [0,\infty). \tag{3.11}$$

Then for any $r \geq 1$

$$\begin{aligned} |b_{B_r(0)}(t)| &= |b_{B_r(0)}(t) - b_{B_1(0)}(t)| = |\fint_{B_1(0)} b_{B_r(0)}(t) - b(t,x)\,dx| && (3.12)\\ &\leq r^n \fint_{B_r(0)} |b_{B_r(0)}(t) - b(t,x)|\,dx \leq r^n \|b\|_{L_t^\infty(\mathrm{BMO}_x)}. && (3.13) \end{aligned}$$

By the definition of BMO functions, we have

$$\fint_{B_r(0)} |b(t,x) - b_{B_r(0)}(t)|^p\,dx \leq C\|b\|^p_{L_t^\infty(\mathrm{BMO}_x)}, \tag{3.14}$$

which implies that $b \in L^p_{loc}([0,\infty)\times\mathbb{R}^n)$ for any $1 \leq p < \infty$.

The integrability of b allows us to approximate it by mollified functions. Take $\Phi \in C_0^\infty(B_1(0))$ and $\eta \in C_0^\infty((-1,1))$ with $\Phi,\eta \geq 0$ and

$$\int_{B_1(0)} \Phi(x)\,dx = \int_{-1}^{1} \eta(t)\,dt = 1.$$

Let $\Phi_\varepsilon(x) = \frac{1}{\varepsilon^n}\Phi(\frac{x}{\varepsilon})$ and $\eta_\varepsilon(t) = \frac{1}{\varepsilon}\eta(\frac{t}{\varepsilon})$. Suppose $b \in L^\infty(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$ and satisfies (3.11). Define

$$b_\varepsilon(t,x) = \int_{-\varepsilon}^{+\varepsilon}\int_{B_\varepsilon(0)} \Phi_\varepsilon(y)\eta_\varepsilon(s)b(t-s,x-y)\,dyds. \tag{3.15}$$

Then $b_\varepsilon \to b$ locally in L^p for any $1 \leq p < \infty$. Morover, using Minkowski's integral inequality, it is easy to obtain

$$\|b_\varepsilon\|_{L_t^\infty(\mathrm{BMO}_x)} \leq \|b\|_{L_t^\infty(\mathrm{BMO}_x)}. \tag{3.16}$$

The lattice property of the BMO space in the proposition below should be well known. We include a proof here for completeness.

Proposition 3.3 *Suppose $f,g \in BMO(\mathbb{R}^n)$, then $f\wedge g$ and $f\vee g \in BMO(\mathbb{R}^n)$. Moreover, we have*

$$\|f\wedge g\|_{BMO} \leq C\left(\|f\|_{BMO} + \|g\|_{BMO}\right), \tag{3.17}$$

where C only depends on n, and the same is true for $f\vee g$.

Proof. Here we only prove it for $f\wedge g$ and $f\vee g$ follows similar proof. Observe that for any $a,b,c,d \in \mathbb{R}$, we have

$$|a\wedge b - c\wedge d| \leq |a-c| + |b-d|. \tag{3.18}$$

Hence for any ball B,

$$\begin{aligned} \frac{1}{|B|}\int_B |f\wedge g(x) - (f\wedge g)_B|^2\,dx &\leq \frac{1}{|B|}\int_B |f\wedge g(x) - f_B\wedge g_B|^2\,dx\\ &\leq \frac{2}{|B|}\int_B |f(x)-f_B|^2\,dx + \frac{2}{|B|}\int_B |g(x)-g_B|^2\,dx\\ &\leq C(\|f\|^2_{\mathrm{BMO}} + \|g\|^2_{\mathrm{BMO}}), \end{aligned}$$

and the proof is now complete. ■

4 Proof of Aronson's estimate

The proof follows the main lines as in Stroock [34] and in particular Davies [7] from which a clever use of the h-transform from harmonic analysis is borrowed, while we need to overcome several difficulties since A is non-symmetric and the skew-symmetric part b is singular. These ideas are mainly due to Nash [29] and Moser [25, 26, 27, 28].

Let us begin with the proof of the upper bound.

4.1 Proof of the upper bound

In this part we show the upper bound:

$$\Gamma(x,t;\xi,\tau) \leq \frac{C}{(t-\tau)^{\frac{n}{2}}} \exp\left[-\frac{|x-\xi|^2}{C(t-\tau)}\right] \tag{4.1}$$

for any $t > \tau$ and $x, \xi \in \mathbb{R}^n$, where C depends only on n, λ and $\|b\|_{L_t^\infty(\mathrm{BMO}_x)}$.

The main idea of Davies [7] is to consider the h-transform of the fundamental solution $\Gamma(x,t;\xi,\tau)$ and apply Nash and Moser's iteration to the h-transforms of the fundamental solution Γ. Nash's idea is to iterate the L^p-norms of solutions to parabolic equations, and to control the growth of the L^p-norms. The main ingredient in Nash's argument is the clever use of the Nash inequality

$$\|u\|_2^{2+\frac{4}{n}} \leq C_n \|\nabla u\|_2^2 \|u\|_1^{\frac{4}{n}} \qquad \forall u \in L^1(\mathbb{R}^n) \cap H^1(\mathbb{R}^n), \tag{4.2}$$

where $C_n > 0$ is a constant depending only on the dimension n.

The Nash iteration is neatly described as the following (Stroock [34], Lemma I.1.14).

Lemma 4.1 *Given positive numbers c_1, c_2, β and $p \geq 2$. Let w be positive, non-decreasing and continuous on $[0,\infty)$, and u be positive with continuous derivatives on $[0,\infty)$. Suppose the following differential inequality holds:*

$$u'(t) \leq -\frac{c_1}{p} \frac{t^{(p-2)} u(t)^{1+\beta p}}{w(t)^{\beta p}} + c_2 p u(t), \quad t \geq 0. \tag{4.3}$$

Then there exists a $K(c_1,\beta) > 0$ such that

$$t^{(p-1)/\beta p} u(t) \leq \left(\frac{Kp^2}{\delta}\right)^{\frac{1}{\beta p}} e^{\frac{c_2 \delta t}{p}} w(t), \quad t \geq 0 \tag{4.4}$$

for every $\delta \in (0,1]$.

The iteration procedure above works in a very general setting. It has been explored since the publication of Nash' paper [29] and is still the major ingredient in our proof. It is surprising that they work well even in our setting where the diffusion is very singular.

Fortunately as well, Davies' idea [7, 8] also works well for our parabolic equations. Following Davies [7] and Stroock [34], given a smooth function ψ on $\mathbb{R}^n$, we consider

$$\Gamma^\psi(x,t;\xi,\tau) = e^{-\psi(x)} \Gamma(x,t;\xi,\tau) e^{\psi(\xi)}, \tag{4.5}$$

and the linear operator

$$\Gamma_{\tau,t}^\psi f(x) = \int_{\mathbb{R}^n} f(\xi) e^{-\psi(x)} \Gamma(x,t;\xi,\tau) e^{\psi(\xi)} d\xi,$$

which is defined for non-negative Borel measurable f, and for f which is smooth with a compact support. It is easy to see that the adjoint operator of $\Gamma^{\psi}_{\tau,t}$ can be identified as the following integral operator

$$\Gamma^{\psi\dagger}_{\tau,t} f(x) = \int_{\mathbb{R}^n} f(\xi) e^{\psi(x)} \Gamma(\xi, t; x, \tau) e^{-\psi(\xi)} d\xi.$$

That is

$$\left\langle \Gamma^{\psi}_{\tau,t} f, g \right\rangle = \left\langle f, \Gamma^{\psi\dagger}_{\tau,t} g \right\rangle$$

for any smooth functions f and g with compact supports.

Lemma 4.2 *Let $T > 0, \tau \geq 0$. Let $f \in C_0^\infty(\mathbb{R}^n)$ be non-negative, and $\psi(x) = \alpha \cdot x$ for some $\alpha \in \mathbb{R}^n$. Define*

$$f_t(x) = \int_{\mathbb{R}^n} f(\xi) e^{\psi(\xi) - \psi(x)} \Gamma(x, t; \xi, \tau) d\xi = \Gamma^{\psi}_{\tau,t} f(x)$$

for $t > \tau$, and

$$f_t^\dagger(x) = \int_{\mathbb{R}^n} f(\xi) e^{\psi(x) - \psi(\xi)} \Gamma(\xi, T; x, T - t) d\xi = \Gamma^{\psi\dagger}_{T-t,T} f(x)$$

for $t \in (0, T]$.

There is a constant $C > 0$ depending only on n, such that for any $p \geq 1$,

$$\frac{d}{dt} \|f_t\|_{2p}^{2p} \leq -\lambda \left\| \nabla f_t^p \right\|_2^2 + C_B \frac{p^2 |\alpha|^2}{\lambda} \|f_t\|_{2p}^{2p} \tag{4.6}$$

$t > \tau$, and

$$\frac{d}{dt} \left\| f_t^\dagger \right\|_{2p}^{2p} \leq -\lambda \left\| \nabla f_t^{\dagger p} \right\|_2^2 + C_B \frac{p^2 |\alpha|^2}{\lambda} \left\| f_t^\dagger \right\|_{2p}^{2p} \tag{4.7}$$

for all $t \in (0, T]$, where $C_B = C \|b\|^2_{L_t^\infty(BMO_x)} + 2$.

Proof. We may assume that $\tau = 0$ without loss of generality, so that

$$\begin{aligned} \|f_t\|_{2p}^{2p} &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(\xi) \Gamma^{\psi}(x, t; \xi, 0) d\xi \right)^{2p} dx \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(\xi) e^{-\psi(x) + \psi(\xi)} \Gamma(x, t; \xi, 0) d\xi \right)^{2p} dx. \end{aligned}$$

Differentiating $\|f_t\|_{2p}^{2p}$ to obtain

$$\frac{d}{dt} \|f_t\|_{2p}^{2p} = 2p \int_{\mathbb{R}^n} f_t(x)^{2p-1} \left(\int_{\mathbb{R}^n} f(\xi) e^{-\psi(x) + \psi(\xi)} \frac{\partial}{\partial t} \Gamma(x, t; \xi, 0) d\xi \right) dx,$$

and by equation (2.5) we have

$$\frac{d}{dt} \|f_t\|_{2p}^{2p} = 2p \int_{\mathbb{R}^n} \left[f_t(x)^{2p-1} \int_{\mathbb{R}^n} f(\xi) e^{-\psi(x) + \psi(\xi)} \nabla_x \cdot (A(x,t) \nabla_x \Gamma(x, t; \xi, 0)) d\xi \right] dx.$$

Similarly, the backward equation (2.11) implies that

$$\frac{d}{dt} \left\| f_t^\dagger \right\|_{2p}^{2p} = 2p \int_{\mathbb{R}^n} \left[f_t^\dagger(x)^{2p-1} \int_{\mathbb{R}^n} f(\xi) e^{\psi(x) - \psi(\xi)} \nabla_x \cdot \left(A^T(x, T-t) \nabla_x \Gamma(\xi, T; x, T-t) \right) d\xi \right] dx,$$

where A^T is the transpose of A. By the Fubini theorem, then performing integration by parts we therefore have

$$\begin{aligned}
\frac{1}{2p}\frac{d}{dt}\|f_t\|_{2p}^{2p} &= \int_{\mathbb{R}^n}\left(e^{\psi(\xi)}f(\xi)\int_{\mathbb{R}^n}e^{-\psi(x)}f_t(x)^{2p-1}\nabla_x\cdot(A(x,t)\nabla_x\Gamma(x,t;\xi,0))\,dx\right)d\xi \\
&= \int_{\mathbb{R}^n}f_t(x)^{2p}\langle\nabla\psi,a(x,t)\cdot\nabla\psi\rangle\,dx \\
&\quad -\frac{2p-1}{p^2}\int_{\mathbb{R}^n}\langle\nabla f_t(x)^p,a(x,t)\cdot\nabla f_t(x)^p\rangle\,dx \\
&\quad -\frac{2(p-1)}{p}\int_{\mathbb{R}^n}f_t(x)^p\langle\nabla f_t(x)^p,a(x,t)\cdot\nabla\psi\rangle\,dx \\
&\quad -2\int_{\mathbb{R}^n}f_t(x)^p\langle\nabla f_t(x)^p,b(x,t)\cdot\nabla\psi\rangle\,dx \\
&= I_1-I_2-I_3-I_4, \qquad (4.8)
\end{aligned}$$

and similarly we have

$$\begin{aligned}
\frac{1}{2p}\frac{d}{dt}\left\|f_t^\dagger\right\|_{2p}^{2p} &= \int_{\mathbb{R}^n}f_t^\dagger(x)^{2p}\langle\nabla\psi,a(x,T-t)\cdot\nabla\psi\rangle\,dx \\
&\quad -\frac{2p-1}{p^2}\int_{\mathbb{R}^n}\left\langle\nabla f_t^\dagger(x)^p,a(x,T-t)\cdot\nabla f_t^\dagger(x)^p\right\rangle dx \\
&\quad -\frac{2(p-1)}{p}\int_{\mathbb{R}^n}f_t^\dagger(x)^p\left\langle\nabla\psi,a(x,T-t)\cdot\nabla f_t^\dagger(x)^p\right\rangle dx \\
&\quad -2\int_{\mathbb{R}^n}f_t^\dagger(x)^p\left\langle\nabla\psi,b(x,T-t)\cdot\nabla f_t^\dagger(x)^p\right\rangle dx.
\end{aligned}$$

Since $\frac{d}{dt}\|f_t\|_{2p}^{2p}$ and $\frac{d}{dt}\left\|f_t^\dagger\right\|_{2p}^{2p}$ are similar, we only need to prove (4.6).

Each term I_j on the right-hand side of (4.8) can be dominated as the following. The first three terms I_1, I_2 and I_3 can be handled exactly as in Davies [7] and Stroock [34]. Recall that $\nabla\psi=\alpha$ is a constant vector. Hence we have

$$I_1\le\frac{|\alpha|^2}{\lambda}\|f_t\|_{2p}^{2p}. \qquad (4.9)$$

While for I_2 and I_3, by completing squares we first rewrite the terms of I_2+I_3 as follows

$$\begin{aligned}
-I_2-I_3 &= -\frac{2p-1}{p^2}\int_{\mathbb{R}^n}\langle\nabla f_t(x)^p,a(x,t)\cdot\nabla f_t(x)^p\rangle\,dx \\
&\quad -2\frac{p-1}{p}\int_{\mathbb{R}^n}f_t(x)^p\langle\nabla f_t(x)^p,a(x,t)\cdot\alpha\rangle\,dx \\
&= -\frac{1}{p}\int_{\mathbb{R}^n}\langle\nabla f_t(x)^p,a(x,t)\cdot\nabla f_t(x)^p\rangle\,dx+(p-1)\int_{\mathbb{R}^n}f_t(x)^{2p}\langle\alpha,a(x,t)\cdot\alpha\rangle\,dx \\
&\quad -\frac{p-1}{p^2}\int_{\mathbb{R}^n}\langle(\nabla f_t(x)^p-pf_t(x)^p\alpha),a(x,t)\cdot(\nabla f_t(x)^p-pf_t(x)^p\alpha)\rangle\,dx.
\end{aligned}$$

The last term on the right-hand side is non-positive as $a(x,t)$ is positive definite, so by using inequalities

$$\langle\nabla f_t(x)^p,a(x,t)\cdot\nabla f_t(x)^p\rangle\ge\lambda\,|\nabla f_t(x)^p|^2$$

and

$$\langle \alpha, a(x,t)\cdot\alpha\rangle \leq \frac{1}{\lambda}|\alpha|^2,$$

we deduce that

$$-I_2 - I_3 \leq -\frac{\lambda}{p}\left\|\nabla f_t^p\right\|_2^2 + \frac{p-1}{\lambda}|\alpha|^2\|f_t\|_{2p}^{2p}. \tag{4.10}$$

The main innovation in our proof is the handling of the skew-symmetric part I_4 which does not appear in the symmetric case. The idea is to apply the estimate (3.6) in Proposition 3.2 to obtain

$$\begin{aligned} |I_4| &= \left|2\int_{\mathbb{R}^n} f_t(x)^p \langle \nabla f_t(x)^p, b(x,t)\cdot\alpha\rangle\, dx\right| \\ &\leq C\|b\|_{L_t^\infty(\mathrm{BMO}_x)}|\alpha|\left\|f_t^p\right\|_2\left\|\nabla f_t^p\right\|_2, \end{aligned} \tag{4.11}$$

where C is a constant depending only on n. Therefore

$$|I_4| \leq \frac{\lambda}{2p}\left\|\nabla f_t^p\right\|_2^2 + C\|b\|^2_{L_t^\infty(\mathrm{BMO}_x)}\frac{p}{\lambda}|\alpha|^2\|f_t\|_{2p}^{2p}. \tag{4.12}$$

Putting these estimates together we thus obtain (4.6). ■

Now we can follow arguments in Stroock [34] to obtain the upper bound, yet again by using the special feature of our elliptic operator. We include the major steps only for completeness.

First we can prove the following by exactly the same argument as in Stroock [34].

Lemma 4.3 *There is a constant $C>0$ depending only on n and the $L^\infty(0,\infty;BMO(\mathbb{R}^n))$-norm of the skew-symmetric part of $\left(A^{ij}(x,t)\right)$ such that*

$$\left\|\Gamma^{\psi}_{\tau,t}f\right\|_\infty \leq \frac{C}{(t-\tau)^{n/4}}e^{\frac{C|\alpha|^2(t-\tau)}{\lambda}}\|f\|_2 \tag{4.13}$$

and

$$\left\|\Gamma^{\psi\dagger}_{\tau,t}f\right\|_\infty \leq \frac{C}{(t-\tau)^{n/4}}e^{\frac{C|\alpha|^2(t-\tau)}{\lambda}}\|f\|_2 \tag{4.14}$$

for every $f\in L^2(\mathbb{R}^n)$, $0\leq\tau<t$ and $\alpha\in\mathbb{R}^n$, where $\psi(x)=\alpha\cdot x$.

Proof. We only need to prove (4.13) for the case that $0=\tau<t$. The proof of (4.14) is similar, and uses the inequality (4.7) instead and uses the fact that the constant appears in that inequality is independent of $T>0$.

To show (4.13), we apply Nash's inequality (4.2) to the first term on the right-hand side of (4.6) to deduce that

$$\frac{d}{dt}\|f_t\|_{2p} \leq -\frac{\lambda}{2pC_n}\frac{\|f_t\|_{2p}^{1+4p/n}}{\|f_t\|_p^{4p/n}} + C_B\frac{|\alpha|^2}{\lambda}p\|f_t\|_{2p} \tag{4.15}$$

for every $p>1$. Let $u_p(t)=\|f_t\|_{2p}$ and $w_p(t)=\sup_{0\leq s\leq t}s^{n(p-2)/4p}u_{p/2}(s)$. Then (4.15) may be written as

$$u_p'(t) \leq -\frac{\lambda}{2pC_n}\frac{t^{p-2}u_p(t)^{1+4p/n}}{(w_p(t))^{4p/n}} + C\frac{|\alpha|^2}{\lambda}pu_p(t)$$

so that, according to Lemma 4.1, we have

$$w_{2p}(t) = \sup_{0\leq s\leq t} s^{n(p-1)/4p}u_p(s)$$

$$\begin{aligned} &\le \sup_{0\le s\le t}(\frac{Kp^2}{\delta})^{\frac{n}{4p}}\exp(\frac{C|\alpha|^2\delta s}{p\lambda})w_p(s) \\ &= (\frac{Kp^2}{\delta})^{\frac{n}{4p}}\exp(\frac{C|\alpha|^2\delta t}{p\lambda})w_p(t). \end{aligned}$$

According to (4.15), if take $p=1$, we have

$$w_2(t) = \sup_{0\le s\le t}\|f_s\|_2 \le e^{C|\alpha|^2 t/\lambda}\|f\|_2.$$

Now we set $\delta = 1$ and iterate it to get

$$w_{2^m}(t) \le C\exp\left(\frac{C|\alpha|^2 t}{\lambda}\right)w_2(t) \le C\exp\left(\frac{C|\alpha|^2 t}{\lambda}\right)\|f\|_2.$$

Letting $m\to\infty$, we therefore obtain that

$$\|f_t\|_\infty \le \frac{C}{t^{n/4}}\exp\left(\frac{C|\alpha|^2 t}{\lambda}\right)\|f\|_2,$$

which completes the proof. ■

Proof of the upper bound (4.1). Let us use the same notations as in the proof of the above lemma. By (4.14) and the fact that $\Gamma^{\psi\dagger}_{\tau,t}$ is the adjoint operator of $\Gamma^{\psi}_{\tau,t}$, we have

$$\|f_t\|_2 \le \frac{C}{t^{n/4}}\exp\left(\frac{C|\alpha|^2 t}{\lambda}\right)\|f\|_1.$$

Since $\Gamma^{\psi}_{0,2t} = \Gamma^{\psi}_{t,2t}\circ\Gamma^{\psi}_{0,t}$, we thus deduce that

$$\|f_{2t}\|_\infty \le \frac{C}{t^{n/2}}\exp\left(\frac{C|\alpha|^2 t}{\lambda}\right)\|f\|_1,$$

which is equivalent to

$$\Gamma(x,2t;\xi,0) \le \frac{C}{t^{n/2}}\exp\left[\frac{C|\alpha|^2 t}{\lambda}+\alpha\cdot(\xi-x)\right].$$

Let $\alpha = \frac{\lambda}{2Ct}(x-\xi)$ and adjust $2t$ to t and 0 to τ, we therefore derive the upper bound

$$\Gamma(x,t;\xi,\tau) \le \frac{C}{(t-\tau)^{\frac{n}{2}}}\exp\left(-\frac{|x-\xi|^2}{C(t-\tau)}\right)$$

for any $t>\tau\ge 0$. This completes the proof of the upper bound. ■

4.2 Proof of the lower bound

In this part, we prove the lower bound

$$\Gamma(x,t;\xi,\tau) \ge \frac{1}{C(t-\tau)^{\frac{n}{2}}}\exp\left[-\frac{C|x-\xi|^2}{t-\tau}\right] \tag{4.16}$$

following the idea due to Nash [29], where C depends only on n, λ and $\|b\|_{L^\infty_t(\mathrm{BMO}_x)}$.

According to Nash's arguments, the lower bound is local in nature, and follows easily from the following.

Lemma 4.4 *There is a constant $C_0 > 0$ depending only on the dimension n, $\lambda > 0$ and the $L^\infty(0,\infty;BMO(\mathbb{R}^n))$ norm of $\left(b^{ij}\right)$, such that*

$$\int_{\mathbb{R}^n} \ln(\Gamma(x,1;\xi,0))\,\mu(d\xi) \geq -C_0 \qquad \forall x \in B(0,2), \tag{4.17}$$

and

$$\Gamma(x,2;\xi,0) \geq e^{-2C_0} \qquad \forall x,\xi \in B(0,2), \tag{4.18}$$

where μ denotes the standard Gaussian measure in $\mathbb{R}^n$, i.e.

$$\mu(d\xi) = \mu(\xi)\,d\xi, \qquad \textit{where } \mu(\xi) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{|\xi|^2}{2}}.$$

Proof. The proof follows the same ideas as in Nash, as explained in Stroock [29]. We have to overcome difficulties arising from the additional non-symmetric part $b(x,t) = \left(b^{ij}(x,t)\right)$. The idea is to consider for any $x \in B(0,2)$ the following function

$$G(t) = \int_{\mathbb{R}^n} \ln(\Gamma(x,1;\xi,1-t))\,\mu(d\xi),$$

where $t \in (0,1]$ and $x \in B(0,2)$. Since $\left(A^{ij}\right)$ is uniformly elliptic with bounded derivatives, $\Gamma(x,t;\xi,\tau)$ is a probability density in x (when others variables are fixed) and also in ξ (as other variables are fixed). Hence $\int_{\mathbb{R}^n} \Gamma(x,t;\xi,0)\,d\xi = 1$ for every $t \in (0,1]$. According to Jensen's inequality $G(t) \leq 0$ and what we want to show is that $G(1)$ is bounded from below uniformly for $x \in B(0,2)$. To this end we consider the derivative of G. By a simple calculation with integration by parts we obtain

$$\begin{aligned} G'(t) = & \int_{\mathbb{R}^n} \left\langle \xi, a(\xi,1-t)\cdot\nabla_\xi \ln\Gamma(x,1;\xi,1-t)\right\rangle \mu(d\xi) \\ & + \int_{\mathbb{R}^n} \left\langle \nabla_\xi \ln\Gamma(x,1;\xi,1-t), a(\xi,1-t)\cdot\nabla_\xi \ln\Gamma(x,1;\xi,1-t)\right\rangle \mu(d\xi) \\ & + \frac{1}{\delta}\int_{\mathbb{R}^n} \left\langle \nabla\mu(\xi)^\delta, b(\xi,1-t)\cdot\nabla_\xi\left(\mu(\xi)^{1-\delta}\ln\Gamma(x,1;\xi,1-t)\right)\right\rangle d\xi \end{aligned} \tag{4.19}$$

for any $\delta \in (0,1)$. We have used the facts that $\nabla \ln\mu(\xi) = -2\xi$, the backward equation for the fundamental solution $\Gamma(x,t;\xi,\tau)$ and the following fact that

$$\langle \nabla\mu, b\cdot\nabla\ln\Gamma\rangle = \frac{1}{\delta}\left\langle \nabla\mu^\delta, b\cdot\nabla\left(\mu^{1-\delta}\ln\Gamma\right)\right\rangle$$

as b is skew-symmetric. Using the Cauchy-Schwartz inequality and the compensated compactness theorem 3.1, we deduce that

$$\begin{aligned} G'(t) \geq & -\frac{C}{\varepsilon} + (1-\varepsilon)\int_{\mathbb{R}^n} \left\langle \nabla_\xi \ln\Gamma(x,1;\xi,1-t), a(\xi,1-t)\cdot\nabla_\xi \ln\Gamma(x,1;\xi,1-t)\right\rangle \mu(d\xi) \\ & -\frac{C}{\delta}\|b\|_{\mathrm{BMO}}\left\|\nabla\mu^\delta\right\|_2\left\|\nabla_\xi\left(\mu^{1-\delta}\ln\Gamma(x,1;\cdot,1-t)\right)\right\|_2 \\ \geq & -\frac{C}{\varepsilon} + (1-\varepsilon)\lambda\|\nabla\ln\Gamma(x,1;\cdot,1-t)\|_{L^2(\mu)}^2 \\ & -\frac{C}{\delta}\|b\|_{\mathrm{BMO}}\left\|\nabla\mu^\delta\right\|_2\left\|\nabla\left(\mu^{1-\delta}\ln\Gamma(x,1;\cdot,1-t)\right)\right\|_2 \end{aligned}$$

for any $\varepsilon, \delta \in (0,1)$. Choose $\delta \in (0, \frac{1}{2})$, then we have

$$\left\|\nabla \mu^{\delta}\right\|_{2} = \delta \left\|\mu^{\delta} \nabla \ln \mu\right\|_{2} < \infty.$$

Moreover, for $\delta \in (0, \frac{1}{2})$,

$$\sup_{\xi} \left|\mu(\xi)^{\frac{1}{2}-\delta} \nabla \ln \mu(\xi)\right| < \infty$$

and

$$\sup_{\xi} \left|\mu(\xi)^{\frac{1}{2}-\delta}\right| < \infty,$$

so that

$$\begin{aligned}
\left\|\nabla \left(\mu^{1-\delta} \ln \Gamma(x,1;\cdot,1-t)\right)\right\|_{2} &\leq \left\|(\nabla \mu^{1-\delta}) \ln \Gamma(x,1;\cdot,1-t)\right\|_{2} \\
&\quad + \left\|\mu^{1-\delta} \nabla \ln \Gamma(x,1;\cdot,1-t)\right\|_{2} \\
&= (1-\delta) \left\|(\mu^{\frac{1}{2}-\delta} \nabla \ln \mu) \ln \Gamma(x,1;\cdot,1-t)\right\|_{L^2(\mu)} \\
&\quad + \left\|\mu^{\frac{1}{2}-\delta} \nabla \ln \Gamma(x,1;\cdot,1-t)\right\|_{L^2(\mu)} \\
&\leq C \left(\|\ln \Gamma(x,1;\cdot,1-t)\|_{L^2(\mu)} + \|\nabla \ln \Gamma(x,1;\cdot,1-t)\|_{L^2(\mu)}\right)
\end{aligned}$$

for some constant C depending only on n and $\delta \in (0, \frac{1}{2})$. By substituting this estimate into the inequality for G' we obtain

$$\begin{aligned}
G'(t) &\geq -\frac{C}{\varepsilon} + (1-\varepsilon)\lambda \|\nabla \ln \Gamma(x,1;\cdot,1-t)\|^2_{L^2(\mu)} \\
&\quad - C\|b\|_{\mathrm{BMO}} \left(\|\ln \Gamma(x,1;\cdot,1-t)\|_{L^2(\mu)} + \|\nabla \ln \Gamma(x,1;\cdot,1-t)\|_{L^2(\mu)}\right) \\
&\geq -\frac{C}{\varepsilon} - \frac{1}{4\lambda^2\varepsilon^2} C^2 \|b\|^2_{\mathrm{BMO}} + (1-2\varepsilon)\lambda \|\nabla \ln \Gamma(x,1;\cdot,1-t)\|^2_{L^2(\mu)} \\
&\quad - C\|b\|_{\mathrm{BMO}} \|\ln \Gamma(x,1;\cdot,1-t)\|_{L^2(\mu)}
\end{aligned}$$

for any $\varepsilon \in (0, \frac{1}{2})$. By choosing $\varepsilon = 1/3$, we thus have the following differential inequality:

$$G'(t) \geq -C + \tfrac{\lambda}{3} \|\nabla \ln \Gamma(x,1;\cdot,1-t)\|^2_{L^2(\mu)} - C\|\ln \Gamma(x,1;\cdot,1-t)\|_{L^2(\mu)} \tag{4.20}$$

for all $t \in (0,1)$, for some constant $C > 0$ depending only on n and the $L^{\infty}(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$ norm of the skew-symmetric part $b(x,t)$.

The remaining arguments of the proof are more or less the same as in Stroock [34]. Firstly, by the Poincaré-Wirtinger inequality for the Gaussian measure, we obtain

$$J(x,t) \equiv \|\ln \Gamma(x,1;\cdot,1-t) - G(t)\|^2_{L^2(\mu)} \leq 2\|\nabla \ln \Gamma(x,1;\cdot,1-t)\|^2_{L^2(\mu)}.$$

On the other hand, since $G(t) < 0$, we have

$$\begin{aligned}
J(x,t) &= \|\ln \Gamma(x,1;\cdot,1-t) - G(t)\|^2_{L^2(\mu)} \\
&= \int_{\mathbb{R}^n} (\ln \Gamma(x,1;\xi,1-t) - G(t))^2 \mu(d\xi)
\end{aligned}$$

$$
\begin{aligned}
&\geq \int_{\{\ln\Gamma(x,1;\xi,1-t)\geq -K\}} (\ln\Gamma(x,1;\xi,1-t)-G(t))^2\,\mu(d\xi) \\
&= \int_{\{\ln\Gamma(x,1;\xi,1-t)\geq -K\}} (\ln\Gamma(x,1;\xi,1-t)+K-G(t)-K)^2\,\mu(d\xi) \\
&\geq \frac{1}{2}\int_{\{\ln\Gamma(x,1;\xi,1-t)\geq -K\}} (\ln\Gamma(x,1;\xi,1-t)+K-G(t))^2\,\mu(d\xi) - K^2 \\
&\geq \frac{1}{2}\int_{\{\ln\Gamma(x,1;\xi,1-t)\geq -K\}} G(t)^2\mu(d\xi) - K^2 \\
&= \frac{1}{2}G(t)^2\mu\left\{\xi\in\mathbb{R}^n : \ln\Gamma(x,1;\xi,1-t)\geq -K\right\} - K^2.
\end{aligned}
$$

Recall the upper bound

$$
\Gamma(x,1;\xi,1-t) \leq \frac{C}{t^{n/2}}\exp\left[-\frac{|x-\xi|^2}{Ct}\right]
$$

proved before. Hence

$$
\begin{aligned}
\int_{\{|\xi|>r\}} \Gamma(x,1;\xi,1-t)d\xi &\leq \int_{\{|\xi|>r\}} \frac{C}{t^{n/2}}\exp\left[-\frac{|x-\xi|^2}{Ct}\right]d\xi \\
&= C_1\mu\left[\left|\sqrt{\frac{C}{2}t}\xi + x\right| > r\right] \\
&\leq C_1\mu\left[|\xi| > \frac{r-2}{\sqrt{\frac{C}{2}t}}\right] \leq C_1\mu\left[|\xi| > \frac{r-2}{\sqrt{\frac{C}{2}}}\right],
\end{aligned}
$$

and therefore, there is a positive number R depending on C such that for any $r > R$

$$
\int_{\{|\xi|>r\}} \Gamma(x,1;\xi,1-t)\,d\xi < \frac{1}{4} \qquad \text{for all } t\in(0,1], x\in B(0,2).
$$

Thus for any $t\in[\frac{1}{2},1]$, there is some $M>0$ such that $\Gamma(x,1;\xi,1-t)\leq M$, so that

$$
\begin{aligned}
\frac{3}{4} &\leq \int_{B(0,r)} \Gamma(x,1;\xi,1-t)\,d\xi \\
&\leq |B(0,r)|e^{-K} + (2\pi)^{n/2}Me^{r^2/2}\mu\left\{\xi\in\mathbb{R}^n : \ln\Gamma(x,1;\xi,1-t)\geq -K\right\}.
\end{aligned}
$$

Choose $K>0$ such that $|B(0,r)|e^{-K} = \frac{1}{4}$, we obtain

$$
\mu\left\{\xi\in\mathbb{R}^n : \ln\Gamma(x,1;\xi,1-t)\geq -K\right\} \geq \frac{1}{2(2\pi)^{n/2}Me^{r^2/2}} =: \kappa(r) > 0.
$$

By using this estimate we deduce that

$$
\begin{aligned}
G'(t) &\geq -C + \frac{\lambda}{6}J(x,t) - C\left\|\ln\Gamma(x,t;\cdot,1-t)\right\|_{L^2(\mu)} \\
&\geq -C + \frac{\lambda}{6}J(x,t) - C\left[\sqrt{J(x,t)} + |G(t)|\right]
\end{aligned}
$$

$$
\begin{aligned}
&\geq -C(\varepsilon,\lambda)+\frac{\lambda}{12}J(x,t)-\varepsilon G(t)^2 \\
&\geq -C(\varepsilon,\lambda)+\frac{\lambda}{24}G(t)^2\mu\left\{\xi\in\mathbb{R}^n : \ln\Gamma(x,1;\xi,1-t)\geq -K\right\}-\frac{\lambda}{12}K^2-\varepsilon G(t)^2 \\
&\geq -C(\varepsilon,K,\lambda)+\left(\frac{\lambda}{24}\kappa(r)-\varepsilon\right)G(t)^2
\end{aligned}
$$

for $\varepsilon>0$ such that $\frac{\lambda}{24}\kappa(r)-\varepsilon>0$. Now we obtain

$$
G'(t)\geq -C_1+C_2G(t)^2 \tag{4.21}
$$

for any $t\in[\frac{1}{2},1]$, where $C_1>0$, $C_2\in(0,1]$. The previous inequality (4.21) may be written as

$$
G'(t)\geq C_2\left(G-\sqrt{\frac{C_1}{C_2}}\right)\left(G+\sqrt{\frac{C_1}{C_2}}\right)
$$

together with the fact that $G<0$. It follows that

$$
G(1)\geq\min\left\{-C_1-2\sqrt{\frac{C_1}{C_2}},-\frac{8}{3C_2}\right\}=-C_0. \tag{4.22}
$$

The lower bound in (4.17) follows from the Chapman-Kolmogrov equation and Jensen's inequality. In fact

$$
\begin{aligned}
\ln\Gamma(x,2;\xi,0) &= \ln\left(\int_{\mathbb{R}^n}\Gamma(x,2;z,1)\Gamma(z,1;\xi,0)dz\right) \\
&= \ln\left(\int_{\mathbb{R}^n}(2\pi)^{n/2}e^{|z|^2/2}\Gamma(x,2;z,1)\Gamma(z,1;\xi,0)\mu(dz)\right) \\
&\geq \ln\left(\int_{\mathbb{R}^n}\Gamma(x,2;z,1)\Gamma(z,1;\xi,0)\mu(dz)\right) \\
&\geq \int_{\mathbb{R}^n}\ln\Gamma(x,2;z,1)\mu(dz)+\int_{\mathbb{R}^n}\ln\Gamma(z,1;\xi,0)\mu(dz) \\
&\geq -2C_0,
\end{aligned}
$$

which yields (4.18). ■

Proof of the lower bound (4.16). By using the scaling invariant property, i.e. for any $r>0$ and $z\in\mathbb{R}^n$,

$$
\Gamma(rx+z,r^2t;r\xi+z,0)=r^{-n}\Gamma^{A_{r,z}}(x,t;\xi,0), \tag{4.23}
$$

where $A_{r,z}(x,t)=A(rx+z,r^2t)$ and Γ^A is the fundamental solution associated with A. The transformation $A\to A^{r,z}$ preserves the elliptic constant λ and more importantly the $L^\infty(0,\infty;\mathrm{BMO}(\mathbb{R}^n))$ norms. So we may apply (4.18) to $\Gamma^{A_{r,z}}$ to deduce that

$$
\Gamma(x,2t;\xi,0)\geq\frac{e^{-2A}}{t^{n/2}},\quad |\xi-x|<4t^{\frac{1}{2}}. \tag{4.24}
$$

Together with the same chain argument as in Stroock [34] by using the Chapman-Kolmogrov equation, we obtain the lower bound accordingly. ■

5 Weak solutions for non-symmetric parabolic equations

Let us consider the parabolic equation

$$\sum_{i,j=1}^{n} \frac{\partial}{\partial x^i}\left[A^{ij}(x,t)\frac{\partial}{\partial x^j}u(x,t)\right] - \frac{\partial}{\partial t}u(x,t) = 0, \tag{5.1}$$

where $A^{ij} = a^{ij} + b^{ij}$, $\left(a^{ij}\right)$ is symmetric satisfying the uniform elliptic condition that $\lambda \leq (a^{ij}) \leq \lambda^{-1}$ in the matrix sense, and $\left(b^{ij}\right)$ is skew-symmetric. We only assume that A^{ij} are Borel measurable in (x,t), and $b^{ij}(x,t)$ belong to the BMO space for every $t \geq 0$, such that the BMO norms $t \to \|b(\cdot,t)\|_{\text{BMO}}$ is bounded, whose supremum norm is denoted by $\|b\|_{L_t^\infty(\text{BMO}_x)}$ as before.

Let us consider Cauchy's initial and Dirichlet boundary problem associated with (5.1). Let $D \subset \mathbb{R}^n$ be an open subset with smooth boundary. Given $\tau > 0$, $u(x,t)$, which is a locally integrable and Borel measurable function in $(x,t) \in D \times [\tau,\infty)$, is a weak solution to the Dirichlet boundary problem of (5.1) with initial data $u(\cdot,\tau) = f \in L^2(D)$, if $u \in L^2\left(\tau,T;H^1(D)\right) \cap L^\infty\left(\tau,T;L^2(D)\right)$ and

$$-\int_\tau^\infty \int_D \langle \nabla\varphi(x,t), A(x,t)\cdot\nabla u(x,t)\rangle\, dxdt + \int_\tau^\infty \int_D u(x,t)\frac{\partial}{\partial t}\varphi(x,t)dtdx + \int_D f(x)\varphi(x,\tau)dx = 0 \tag{5.2}$$

for any smooth function $\varphi(x,t)$ which has a compact support in $D \times [\tau,\infty)$. Let $\Gamma^D(x,t;\xi,\tau)$ denote the corresponding fundamental solution. Then we recall the following result in [19, Chapter IV, Section 15] when the coefficients are smooth.

Lemma 5.1 *Suppose in addition that A^{ij} are smooth, so that the fundamental solution $\Gamma(x,t;\xi,\tau)$ exists, is smooth, and satisfies the Aronson estimate, and therefore*

$$0 < \Gamma^D(x,t;\xi,\tau) \leq \Gamma(x,t;\xi,\tau) \leq \frac{C}{(t-\tau)^{n/2}}\exp\left(-\frac{|x-\xi|^2}{C(t-\tau)}\right)$$

for all $t > \tau$ and $x,\xi \in D$. If $f \in L^2(D)$, then $u(x,t) = \Gamma_{\tau,t}^D f(x)$ belongs to

$$C([\tau,\infty),L^2(D)) \cap L^\infty(\tau,\infty;L^2(D)) \cap L^2(\tau,\infty;H^1(D)).$$

Moreover, we have the energy inequality

$$||u(\cdot,t)||_2^2 + 2\lambda\int_\tau^t \|\nabla u(\cdot,s)\|_2^2 \leq ||f||_2^2 \tag{5.3}$$

for all $t \geq \tau$, and $u(x,t)$ is also a weak solution to (5.2).

This allows us to approximate A^{ij} by $\{A_m^{ij}\}$ and there exists unique fundamental solution Γ^m and strong solution u^m corresponding to A_m^{ij}, which satisfy the claims in Lemma 5.1.

We are now in a position to show the following uniqueness theorem for weak solutions.

Theorem 5.2 *Suppose $A = a + b$ satisfies conditions stated at the beginning of the section, i.e. $\lambda \leq \left(a^{ij}(x,t)\right) \leq \lambda^{-1}$ and $\|b\|_{L_t^\infty(BMO_x)} < \infty$. Let $\tau \geq 0$. Suppose $u(x,t) \in L^\infty(\tau,\infty;L^2(\mathbb{R}^n)) \cap L^2(\tau,\infty;H^1(\mathbb{R}^n))$, and satisfies*

$$\int_\tau^\infty \int_{\mathbb{R}^n} u(x,t)\frac{\partial}{\partial t}\varphi(x,t)\, dxdt = \int_\tau^\infty \int_{\mathbb{R}^n} \langle \nabla\varphi(x,t), A(x,t)\cdot\nabla u(x,t)\rangle\, dxdt \tag{5.4}$$

for any $\varphi(x,t)$ that is smooth with compact support in $\mathbb{R}^n \times (\tau,\infty)$, then

$$\frac{\partial u}{\partial t} \in L^2\left(\tau,\infty;H^{-1}(\mathbb{R}^n)\right). \tag{5.5}$$

Hence the following energy inequality holds:

$$\|u(\cdot,T)\|_2^2 + 2\lambda \int_\tau^T \int_{\mathbb{R}^n} |\nabla u(x,t)|^2 \, dxdt \le \|u(\cdot,\tau)\|_2^2 \tag{5.6}$$

for every $T > \tau$, and the uniqueness of weak solutions holds for the initial problem of (5.1) in space $L^\infty(\tau,\infty;L^2(\mathbb{R}^n)) \cap L^2(\tau,\infty;H^1(\mathbb{R}^n))$.

Proof. Consider the linear functional

$$F_t(\psi) = \int_{\mathbb{R}^n} \langle \nabla\psi(x,t), A(x,t)\cdot\nabla u(x,t)\rangle \, dx$$

for $\psi \in H^1(\mathbb{R}^n)$. By using the compensated compactness inequality (3.5) we have

$$F_t(\psi) \le \left(\|a\|_{L^\infty(\mathbb{R}^n\times[0,\infty))} + C\|b\|_{L_t^\infty(\mathrm{BMO}_x)}\right) \|\nabla u(\cdot,t)\|_2 \|\nabla\psi\|_2 \tag{5.7}$$

for any $\psi \in H^1(\mathbb{R}^n)$. Hence by the Riesz representation theorem, there exists a unique $w(\cdot,t) \in H^1(\mathbb{R}^n)$ for every t such that

$$F_t(\psi) = \int_{\mathbb{R}^n} \left(\nabla w(x,t)\cdot\nabla\psi(x) + w(x,t)\psi(x)\right) dx, \tag{5.8}$$

where

$$\|w(\cdot,t)\|_{H^1(\mathbb{R}^n)} \le \left(\|a\|_{L^\infty(\mathbb{R}^n\times[0,\infty))} + C\|b\|_{L_t^\infty(\mathrm{BMO}_x)}\right) \|\nabla u(\cdot,t)\|_{L^2(\mathbb{R}^n)},$$

which implies that $w \in L^2(\tau,\infty;H^1(\mathbb{R}^n))$.

In terms of $w(x,t)$, (5.4) becomes

$$\int_\tau^T \int_{\mathbb{R}^n} u(x,t)\varphi(x)\eta'(t)\, dxdt = \int_\tau^T \int_{\mathbb{R}^n} \left(\nabla w(x,t)\cdot\nabla\varphi(x) + w(x,t)\varphi(x)\right)\eta(t)\, dxdt \tag{5.9}$$

for any $\eta \in C_0^\infty((\tau,T))$ and $\varphi \in C_0^\infty(\mathbb{R}^n)$, which can be written as

$$\int_\tau^T \langle u(\cdot,t),\varphi\rangle_{L^2} \eta'(t)dt = \int_\tau^T \langle w(\cdot,t),\varphi\rangle_{H^1} \eta(t)dt,$$

and can be extended to any $\varphi \in H^1(\mathbb{R}^n)$. Since

$$\begin{aligned}\left|\int_\tau^T \langle w(\cdot,t),\varphi\rangle_{H^1}\eta(t)dt\right| &\le \int_\tau^T \|w(\cdot,t)\|_{H^1}\|\varphi\|_{H^1}\eta(t)dt \\ &= \|\varphi\|_{H^1}\int_\tau^T \|w(\cdot,t)\|_{H^1}\eta(t)dt \\ &\le \|\varphi\|_{H^1}\sqrt{\int_\tau^T \|w(\cdot,t)\|_{H^1}^2\, dt}\, \|\eta\|_{L^2([\tau,T])},\end{aligned}$$

we obtain

$$\left|\int_{\tau}^{T} \langle u(\cdot,t),\varphi\rangle_{L^2}\,\eta'(t)dt\right| \le \|\varphi\|_{H^1}\sqrt{\int_{\tau}^{T}\|w(\cdot,t)\|_{H^1}^2\,dt}\,\|\eta\|_{L^2([\tau,T])},$$

which implies that

$$\frac{d}{dt}\langle u(\cdot,t),\varphi\rangle_{L^2}\in L^2([\tau,T])$$

for every $\varphi\in H^1(\mathbb{R}^n)$. Moreover, according to the Riesz representation theorem

$$\left\|\frac{d}{dt}\langle u(\cdot,t),\varphi\rangle_{L^2}\right\|_{L^2[\tau,T]} \le \|\varphi\|_{H^1}\sqrt{\int_{\tau}^{T}\|w(\cdot,t)\|_{H^1}^2\,dt}$$

for any $\varphi\in H^1(\mathbb{R}^n)$. Therefore, there is $\frac{\partial}{\partial t}u\in L^2\left(\tau,T;H^{-1}(\mathbb{R}^n)\right)$ such that

$$\int_{\tau}^{T}\left\langle \frac{\partial}{\partial t}u(\cdot,t),\varphi\right\rangle_{H^{-1},H^1}\eta(t)dt = -\int_{\tau}^{T}\langle u(\cdot,t),\varphi\rangle_{L^2}\,\eta'(t)dt$$

for every $\varphi\in H^1(\mathbb{R}^n)$ and $\eta\in C_0^\infty(\tau,T)$. The equation above can be written as

$$\begin{aligned}\left\langle \frac{\partial}{\partial t}u,\varphi\otimes\eta\right\rangle_{L^2(H^{-1}),L^2(H^1)} &= -\int_{\tau}^{T}\int_{\mathbb{R}^n}u(x,t)\frac{\partial}{\partial t}(\varphi(x)\eta(t))\,dxdt\\ &= -\int_{\tau}^{T}\int_{\mathbb{R}^n}\langle\nabla(\varphi(x)\eta(t)),A(x,t)\cdot\nabla u(x,t)\rangle\;dxdt,\end{aligned}$$

where and in the remaining part of the proof, for simplicity, we use $\langle\cdot,\cdot\rangle_{L^2(H^{-1}),L^2(H^1)}$ to denote the pairing between $L^2\left(\tau,T;H^1(\mathbb{R}^n)\right)$ and its dual space $L^2\left(\tau,T;H^{-1}(\mathbb{R}^n)\right)$. Since

$$\operatorname{span}\left\{\varphi\otimes\eta:\varphi\in H^1(\mathbb{R}^n)\text{ and }\eta\in C_0^\infty(\tau,T)\right\}$$

is dense in $L^2\left(\tau,T;H^1(\mathbb{R}^n)\right)$, we have

$$\left\langle \frac{\partial}{\partial t}u,\psi\right\rangle_{L^2(H^{-1}),L^2(H^1)} = -\int_{\tau}^{T}\int_{\mathbb{R}^n}\langle\nabla\psi(x,t),A(x,t)\cdot\nabla u(x,t)\rangle\;dxdt$$

for any $\psi\in L^2\left(\tau,T;H^1(\mathbb{R}^n)\right)$. In particular,

$$\begin{aligned}\left\langle \frac{\partial}{\partial t}u,u\right\rangle_{L^2(H^{-1}),L^2(H^1)} &= -\int_{\tau}^{T}\int_{\mathbb{R}^n}\langle\nabla u(x,t),A(x,t)\cdot\nabla u(x,t)\rangle\;dxdt\\ &\le -\lambda\int_{\tau}^{T}\int_{\mathbb{R}^n}|\nabla u(x,t)|^2dxdt.\end{aligned}$$

Now, by Theorem 5.9.3 in Evans [11], we deduce that $u\in C\left([\tau,T],L^2(\mathbb{R}^n)\right)$ and

$$\|u(\cdot,T)\|_2^2-\|u(\cdot,\tau)\|_2^2 = 2\left\langle\frac{\partial}{\partial t}u,u\right\rangle \le -2\lambda\int_{\tau}^{T}\int_{\mathbb{R}^n}|\nabla u(x,t)|^2dxdt,$$

which in turn yields the energy inequality (5.6). The other conclusions of the theorem follow easily. ■

We are now in a position to state and to prove the following theorem.

Theorem 5.3 *Suppose $\left(A^{ij}\right) = \left(a^{ij}\right) + \left(b^{ij}\right)$, where a and b are symmetric and skew-symmetric parts of A respectively, is uniformly elliptic: $\lambda \le a(x,t) \le \lambda^{-1}$ in matrix sense for some constant $\lambda > 0$, and $\|b\|_{L_t^\infty(BMO_x)} < \infty$. Then there is a unique positive function $\Gamma(x,t;\xi,\tau)$ defined for $t > \tau \ge 0$ and $x, \xi \in \mathbb{R}^n$, which possesses the following properties.*

1) Γ is a Markov transition density: $\Gamma(x,t;\xi,\tau) > 0$,

$$\int_{\mathbb{R}^n} \Gamma(x,t;\xi,\tau)d\xi = 1 \quad \text{and} \quad \int_{\mathbb{R}^n} \Gamma(x,t;\xi,\tau)dx = 1$$

for any $t > \tau \ge 0$, and

$$\Gamma(x,t;\xi,\tau) = \int_{\mathbb{R}^n} \Gamma(x,t;z,s)\Gamma(z,s;\xi,\tau)dz$$

for any $t > s > \tau \ge 0$.

2) There is a constant $M > 0$ depending only on n, λ and $\|b\|_{L_t^\infty(BMO_x)}$ such that

$$\frac{1}{M(t-\tau)^{n/2}} \exp\left(-\frac{M|x-\xi|^2}{t-\tau}\right) \le \Gamma(x,t;\xi,\tau) \le \frac{M}{(t-\tau)^{n/2}} \exp\left(-\frac{|x-\xi|^2}{M(t-\tau)}\right)$$

for all $t > \tau$.

3) For every $f \in L^2(\mathbb{R}^n)$, $u(x,t) = \int_{\mathbb{R}^n} f(\xi)\Gamma(x,t;\xi,\tau)d\xi$ (for any $t \ge \tau$) is the unique weak solution with initial data f, which belongs to

$$C([\tau,\infty), L^2(\mathbb{R}^n)) \cap L^\infty(\tau,\infty; L^2(\mathbb{R}^n)) \cap L^2(\tau,\infty; H^1(\mathbb{R}^n)).$$

Proof. Since $b \in L^\infty(0,\infty; \mathrm{BMO}(\mathbb{R}^n))$, we can choose a $\varepsilon > 0$ such that

$$\|\varepsilon \log(|x|)\|_{\mathrm{BMO}} \le \|b\|_{L_t^\infty(\mathrm{BMO}_x)}.$$

Define

$$U^{(m)}(x) = (-\varepsilon \log(|x|) + m) \wedge m \vee 0, \qquad L^{(m)}(x) = (\varepsilon \log(|x|) - m) \wedge 0 \vee (-m), \tag{5.10}$$

which are compactly supported BMO functions with

$$\|U^{(m)}\|_{\mathrm{BMO}} = \|L^{(m)}\|_{\mathrm{BMO}} \le C\|b\|_{L_t^\infty(\mathrm{BMO}_x)},$$

where constant $C > 0$ depends only on the dimension n. Let

$$b^{(m)}(x,t) = b(x,t) \wedge U^{(m)}(x) \vee L^{(m)}(x). \tag{5.11}$$

By inequality (3.16), we can mollify it to define $b^{(m)}_{\frac{1}{m}}$ and there is a C independent of b and m, such that $\|b^{(m)}_{\frac{1}{m}}\|_{L_t^\infty(\mathrm{BMO}_x)} \le C\|b\|_{L_t^\infty(\mathrm{BMO}_x)}$. Each $b^{(m)}_{\frac{1}{m}}$ is smooth with compact support and $b^{(m)}_{\frac{1}{m}} \to b$ in $L^p_{loc}([0,\infty) \times \mathbb{R}^n)$ for any $1 \le p < \infty$. For simplicity denote $b^{(m)}_{\frac{1}{m}}$ by b_m. Similarly a_m denotes the mollified approximation of a for $m = 1, 2, \cdots$. $a_m(x,t)$ and $b_m(x,t)$ are smooth, bounded and have bounded derivatives of all orders, and $a_m \to a$ and $b_m \to b$ in $L^p_{loc}([0,\infty) \times \mathbb{R}^n)$ for every $p \in [1,\infty)$.

Now for each $A_m(x,t) = a_m(x,t) + b_m(x,t)$, a_m is uniformly elliptic with elliptic constant 2λ and

$$\|b_m\|_{L_t^\infty(\mathrm{BMO}_x)} \le C\|b\|_{L_t^\infty(\mathrm{BMO}_x)}$$

for some constant depending only on the dimension n, thus there is a unique fundamental solution $\Gamma^m(x,t;\xi,\tau)$ which satisfies the Aronson estimate with the same constant M. According to Theorem 2.2, $\Gamma^m(x,t;\xi,\tau)$ are Hölder continuous in any compact sub-set of $t>\tau\geq 0$ and $x,\xi\in\mathbb{R}^n$ with the same Hölder exponent and the same Hölder constant for all $m=1,2,\cdots$. Therefore by the Arzela-Ascoli Theorem, there is a sub-sequence of Γ^m, for simplicity the sub-sequence is still denoted by Γ^m, which converges locally uniformly to some $\Gamma(x,t;\xi,\tau)$ for $t>\tau\geq 0$ and $x,\xi\in\mathbb{R}^n$. Clearly $\Gamma(x,t;\xi,\tau)$ still satisfies 1) and 2).

We now prove 3). By our construction, if $\tau>0$ and $f\in L^2(\mathbb{R}^n)$,

$$u^m(x,t)=\Gamma^m_{\tau,t}f(x)\to u(x,t)=\Gamma_{\tau,t}f(x)$$

pointwise. According to Lemma 5.1, u^m (actually $u^m(x,t)$ is Hölder continuous too for $t>\tau$ and x) is a strong solution to the Cauchy problem of the parabolic equation associated with the diffusion matrix A_m, so that the energy inequality holds:

$$||u^m(\cdot,t)||_2^2+\lambda\int_\tau^t\|\nabla u^m(\cdot,s)\|_2^2\leq||f||_2^2, \tag{5.12}$$

which implies that $\{u^m\}$ is uniformly bounded in $L^\infty(\tau,\infty;L^2(\mathbb{R}^n))\cap L^2(\tau,\infty;H^1(\mathbb{R}^n))$. Hence there is a sub-sequence which converges weakly, whose limit must be u and

$$u\in L^\infty(\tau,\infty;L^2(\mathbb{R}^n))\cap L^2(\tau,\infty;H^1(\mathbb{R}^n)).$$

Next we prove that u also satisfies the energy inequality (5.12) as u^m. For each m, we have

$$\int_\tau^\infty\int_{\mathbb{R}^n}u^m(x,t)\frac{\partial}{\partial t}\varphi(x,t)\,dxdt-\int_\tau^\infty\int_{\mathbb{R}^n}\langle\nabla\varphi(x,t),A_m(x,t)\cdot\nabla u^m(x,t)\rangle\,dxdt=0$$

for any $\varphi\in C_0^\infty((\tau,\infty)\times\mathbb{R}^n)$. Since $A_m\to A$ in $L^p_{loc}([\tau,\infty)\times\mathbb{R}^n)$ for any $1\leq p<\infty$ and $u_m\to u$ weakly in $L^2(\tau,\infty;H^1(\mathbb{R}^n))$. By taking $m\to\infty$ in the equation above, we obtain that

$$\int_\tau^\infty\int_{\mathbb{R}^n}u(x,t)\frac{\partial}{\partial t}\varphi(x,t)\,dxdt-\int_\tau^\infty\int_{\mathbb{R}^n}\langle\nabla\varphi(x,t),A(x,t)\cdot\nabla u(x,t)\rangle\,dxdt=0$$

for any $\varphi\in C_0^\infty((\tau,\infty)\times\mathbb{R}^n)$. That is, u is a weak solution to the Cauchy problem of the parabolic equation (5.1) with initial data f. Now according to Theorem 5.2, we have that $\frac{\partial u}{\partial t}\in L^2(\tau,\infty;H^{-1}(\mathbb{R}^n))$, $u\in C([\tau,\infty),L^2(\mathbb{R}^n))$ and therefore

$$||u(\cdot,t)||_2^2+\lambda\int_\tau^t\|\nabla u(\cdot,s)\|_2^2\leq||f||_2^2. \tag{5.13}$$

The uniqueness of the fundamental solution Γ follows from the energy inequality easily. In fact, suppose that there is another sub-sequence of Γ^m converges to $\tilde\Gamma$. Then $\tilde u(x,t)=\tilde\Gamma_{\tau,t}f(x)$ satisfies all the results above. Especially they both satisfy the energy inequality. Therefore $w=u-\tilde u$ is also a weak solution. By Theorem 5.2, we deduce that

$$\int_{\mathbb{R}^n}w(x,t)^2\,dx+\lambda\int_\tau^t\int_{\mathbb{R}^n}|\nabla w(x,t)|^2\,dxdt\leq 0$$

for any $t>\tau$ and we have $w=0$. This implies $u=\tilde u$ and hence $\Gamma=\tilde\Gamma$. The proof is complete. ■

References

[1] D. G. Aronson: Uniqueness of positive solutions of second order parabolic equations, *Ann. Polon. Math.*, 16 (1965).

[2] D. G. Aronson: Bounds for the fundamental solution of a parabolic equation, *Bull. Am. Math. Soc*. 73, 890-896 (1967).

[3] D. G. Aronson: Non-negative solutions of linear parabolic equations, *Ann. Sci. Norm. sup. Pisa* (3) 22, 607-94 (1968).

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

[5] R. Coifman, P.-L. Lions, Y. Meyer and S. Semmes: Compensated compactness and Hardy spaces, *J. Math. Pures et Appl.,* 72, 247-286 (1992).

[6] R. Coifman, G. Weiss: Analyse harmonique non-commutative sur certains sepaces homogènes, *Lecture Notes in Math*., Springer-Verlag (1971).

[7] E. B. Davies: Explicit Constants for Gaussian Upper Bounds on Heat Kernels, *American J. of Mathematics*, Vol. 109, No. 2, pp. 319-333 (Apr., 1987).

[8] E. B. Davies: *Heat Kernels and Spectral Theory*, Cambridge Tracts in Math. 92, Cambridge University Press (1989).

[9] E. De Giorgi: Sulla differenziabilita e l'analicita delle estremali degli integrali multipli regolari, *Mem. Accad. Sci. Toino. Cl. Sci. Mat. Nat.,* Ser. 3, 3 (1957).

[10] M. G. Delgadino and S. Smith: Hölder estimates for fractional parabolic equations with critical divergence free drifts, *Annales de l'Institut Henri Poincare (C) Non Linear Analysis.* Vol. 35. No. 3. Elsevier Masson, (2018).

[11] L. C. Evans: *Partial Differential Equations*, Graduate studies in mathematics. 19. American Mathematical Society (1998).

[12] E. B. Fabes and D. W. Stroock: A New Proof of Moser's Parabolic Harnack Inequality Using the old Ideas of Nash, *Arch. for Ratl. Mech. and Anal.* 96, no. 4, 327-338 (1986).

[13] S. Friedlander, V. Vicol: Global well-posedness for an advection–diffusion equation arising in magneto-geostrophic dynamics, *Annales de l'Institut Henri Poincare (C) Non Linear Analysis.* Elsevier Masson, 28(2): 283-301 (2011).

[14] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, INC. (1964).

[15] M. Fukushima: *Dirichlet Forms and Markov Processes*, North-Holland Pub. Amsterdam, Oxford, New York (1980).

[16] A. A. Grigor'yan: The heat equation on noncompact Riemannian manifolds, *Math. USSR Sbornik* Vol. 72, No.1, 47-77 (1992).

[17] S. Hofmann and S. Mayboroda: Hardy and BMO spaces associated to divergence form elliptic operators, arXiv:math/0611804v2 (2007).

[18] F. John, L. Nirenberg: On functions of bounded mean oscillation, *Comm. Pure Appl. Math.* 14, 785-799 (1961).

[19] O. A. Ladyženskaja, V. A. Solonnikov, N. N. Uralćeva: *Linear and Quasilinear Equations of Parabolic Type*, Translations of Math. Monographs 23, AMS Providence RI (1968).

[20] P. G. Lemarié-Rieusset: *The Navier-Stokes Problem in the 21st Century*, CRC Press (2016).

[21] J. Leray: Sur le mouvement d'un liquide visqueux emplissant l'espace, *Acta Math*. 63, 193 - 248 (1934).

[22] A. S. Monin, A. M. Yaglom: *Statistical Fluid Mechanics: Mechanics of Turbulence* Volume 1, The MIT Press (1971).

[23] A. S. Monin, A. M. Yaglom: *Statistical Fluid Mechanics: Mechanics of Turbulence* Volume 2, The MIT Press (1975).

[24] C. B. Morrey, Jr.: *Multiple Integrals in the Calculus of Variations,* Springer-Verlag New York Inc. (1966).

[25] J. Moser: A new proof of de Giorgi's theorem cocerning the regularity problem for elliptic differential equations, *Comm. Pure Appl. Math.*, Vol XIII, 457-469 (1960).

[26] J. Moser: A Harnack inequality for parabolic differential equations, *Comm. Pure Appl. Math.*, Vol XVII, 101-134 (1964).

[27] J. Moser: A Harnack inequality for parabolic differential equations, *Comm. Pure Appl. Math.*, Vol. 17, 1964, pp. 101-134, and correction in Comm. Pure Appl. Math., Vol. 20, 1967, pp. 231-236.

[28] J. Moser: On a Pointwise Estimate for Parabolic Differential Equations, *Comm. Pure Appl. Math.*, VOL. XXIV, 727-740 (1971).

[29] J. Nash: Continuity of Solutions of Parabolic and Elliptic Equations, *American J. of Mathematics*, Vol. 80, No. 4. (Oct., 1958), pp. 931-954.

[30] J. R. Norris and D. W. Stroock: Estimates on the fundamental solution to heat flows with uniformly elliptic coefficients, *Proc. London. Math. Soc*. (3) 62 (1991), 373-402.

[31] H. Osada: Diffusion processes with generators of generalized divergence form, *J. Math. Kyoto Univ.* (JMKYAZ), 27-4, 597-619 (1987).

[32] G. Seregin, L. Silvestre, V. Šverák, A. Zlatoš: On divergence-free drifts, *J. Differential Equations* 252 (2012) 505–540.

[33] E. M. Stein: *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, New Jersey (1970).

[34] D. W. Stroock: Diffusion semigroups corresponding to uniformly elliptic divergence form operators, *Seminaire de Probabilites*, tome 22 (1988), p. 316 - 347.

[35] D. W. Stroock: *Partial Differential Equations for Probabilists*, Cambridge University Press (2008).

[36] D. W. Stroock and S. R. S. Varadhan: *Multidimensional Diffusion Processes*, Grundlehren Series 233, Springer-Verlag, New York and Heidelberg (1979).

[37] G. I. Taylor: Diffusion by continuous movements, *Proc. London Math. Soc.* 20, 196 - 211 (1921).

[38] G. I. Taylor: Statistical theory of turbulence, *Proc. London Math. Soc.* A 151, 421 - 478 (1935).

[39] W. von Wahl: *The Equations of Navier-Stokes and Abstract Parabolic Equations*, Aspects of Mathematics, Friedr. Vieweg & Sohn (1985).