

Parabolic Equations and Diffusion Processes with Divergence-free Vector Fields



Guangyu Xi
Queen's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Hilary term 2018

Abstract

The study of this thesis is motivated by the stochastic Lagrangian representations of solutions to the Navier-Stokes equations. The stochastic Lagrangian formulation to the Navier-Stokes equations is described by stochastic differential equations, which essentially represent the diffusions under divergence-free velocity fields. The associated stochastic differential equations are closely related to a class of parabolic equations and these two types of equations are the central objects of this thesis. The difficulty of the problem mainly comes from the low regularity of the velocity field. The key point is that we use the divergence-free condition to relax the regularity assumptions.

The thesis is divided into two parts. The first part is the Aronson-type estimate which is an *a priori* estimate on the fundamental solutions (transition probability) independent of the smoothness of the coefficients. In the critical case, we obtain the Aronson estimate in its classical form, while in supercritical cases we obtain a weaker Aronson-type estimate. In the second part, we use approximation arguments to apply the Aronson estimate to the construction of solutions to the parabolic equations and the stochastic differential equations, and further regularity theory of the solutions is obtained for the critical case. Under the supercritical conditions, we will focus on the uniqueness of solutions to the parabolic equations and their relation to the construction of the diffusion processes.

Acknowledgements

I would like to thank my supervisor Professor Zhongmin Qian for his guidance. It has been a great time meeting him every week to discuss. He has always guided me patiently using his acute mind for the past four years. I would also like to thank my secondary supervisor Professor Gui-Qiang G. Chen for meeting me frequently and giving me invaluable advice on both academy and career. Their passion for mathematics greatly motivated me, and their vision and experience keep inspiring me.

In the past four years I also benefited a lot from many mathematicians here in Oxford and those who were visiting Oxford. A gratitude is to our cohort mentor Professor Yves Capdeboscq for helping me in many aspects. I would like to thank Professor Ben Hambly, Professor Jan Kristensen and Professor Terry Lyons for assessing my progress for transfer and confirmation and giving me many suggestions. I also want to thank Professor Gregory Seregin and Professor Elton P. Hsu for agreeing to be my examiners.

I gratefully acknowledge the support from the EPSRC CDT-PDE program, the Mathematical Institute and Queen's College. Also I want to thank all cohort members and all my friends for having lots of happy time together and helping me during the most struggling first year in Oxford. In particular, I want to thank Aleksander Klimek, Guy Flint and Ilya Chevyrev for sharing the office with me and discussing with me. I would also like to acknowledge my gratitude to Siran Li, who has been giving me great advice. Their encouragement and credible ideas have been great contributions in the completion of this thesis.

Finally I want to thank my parents for supporting me from all aspects and encouraging me constantly. Their attitude towards life affected me profoundly which makes me patient, strong and optimistic. I also want to thank my girlfriend Shuman for always staying by my side whenever I need and supporting me wholeheartedly.

Contents

1	Introduction	1
1.1	Stochastic Lagrangian representation	3
1.2	Linearized equations and diffusion processes	4
1.3	Divergence-free condition	7
1.4	Main results	10
2	The Aronson estimate	21
2.1	Technical facts	21
2.1.1	BMO space and compensated compactness	21
2.1.2	Poincaré-Wirtinger inequality	27
2.1.3	A Riccati differential inequality	29
2.2	A critical condition on the drift	31
2.2.1	The upper bound	32
2.2.2	The lower bound	39
2.3	Supercritical conditions on the drift	46
2.3.1	The upper bound	46
2.3.2	The lower bound	61
3	Weak solutions and diffusion processes: critical cases	66
3.1	Hölder regularity of the solutions	67
3.2	Uniqueness of weak solutions	74
3.3	Diffusion processes	82

4	Weak solutions and uniqueness: supercritical cases	93
4.1	Tightness of the fundamental solutions	93
4.2	Uniqueness with time-homogeneous coefficient	96
4.3	Renormalized solutions	104
5	Conclusions	109
5.1	Parabolic equations	109
5.2	Diffusion processes	110
	Bibliography	111

Chapter 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. [44, 48, 60, 78] and etc.) due to the fundamental work by Leray [49]. The Navier-Stokes equations are partial differential equations of second-order

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

$$\nabla \cdot u = 0, \quad (1.2)$$

subject to the no-slip boundary condition if the domain of fluid is finite, where $u(t, x)$ is the velocity vector field of the fluid flow and $p(t, x)$ is the pressure at time t and location x . Leray [49] demonstrated the existence of a weak solution u which belongs to the space $L^\infty(0, \infty; L^2(\mathbb{R}^n))$ and also to the space $L^2(0, \infty; H^1(\mathbb{R}^n))$. The vorticity ω exists in $L^2_{t,x}$ space and formally, by differentiating the Navier-Stokes equations, solves the vorticity equation

$$\frac{\partial}{\partial t}\omega + u \cdot \nabla \omega = \nu \Delta \omega + \omega \cdot \nabla u. \quad (1.3)$$

Here u is the velocity and ω is the vorticity, which is also a time dependent vector field $\omega(t, x)$, and they are related by the definition $\omega = \nabla \times u$. The resolution of the three di-

mensional Navier-Stokes equations remains to be an open mathematical problem (see e.g. [48, 85, 73, 84]). 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 written in the following form

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

and

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

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

$$\partial_t - L = \frac{\partial}{\partial t} - \nu \Delta + u \cdot \nabla. \quad (1.6)$$

The elliptic operator $\nu \Delta - u \cdot \nabla$ is the generator of the so-called Taylor diffusion (see Taylor [82, 83]) 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 fluid flow (see e.g. [58, 59]). 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 [85]) 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 $L = \nu \Delta - u \cdot \nabla$, where u is a Leray-Hopf weak solution of the Navier-Stokes equations.

The following sections are devoted to explaining these ideas in detail. In section 1.1, we present stochastic representations to the Navier-Stokes equations and the vorticity equation, and discuss the motivation from the probability aspect. Section 1.2 is a brief review of the classical results on the parabolic equations and related diffusion processes. In section 1.3 we discuss the divergence-free condition and its advantages. Finally in section 1.4, we state our main results of this thesis and review the related literature with more details.

1.1 Stochastic Lagrangian representation

There are two mathematical descriptions of fluid flow. The first one is the Eulerian coordinates which are fixed coordinates in the ambient space. For example, equations (1.1) to (1.3) are written under the Eulerian coordinates. The second description is the Lagrangian coordinates which follow an individual fluid parcel as it moves through space and time. In inviscid flow, the coordinate X_t is governed by the ODE

$$dX_t = u(X_t, t)dt, \quad X_0 = x_0,$$

where u is the velocity field. This coordinate has been used to study the Euler equations extensively (see e.g. [10]). In terms of viscous flow with the viscosity modeled by $(\nu\Delta)$, we have a stochastic Lagrangian coordinate described by the SDE

$$dX_t = \sqrt{2\nu}dB_t + u(t, X_t)dt, \quad X_0 = x, \quad (1.7)$$

where B_t is the Brownian motion and ν is the coefficient of viscosity. This X_t is the diffusion process corresponding to the operator L in (1.6). This also shows the importance of studying these diffusion processes and related parabolic equations.

Under the stochastic Lagrangian coordinates, it is natural to obtain the following stochastic Lagrangian representation of the vorticity

$$\omega(t, x) = \mathbb{E}^x \left[((\nabla X_t) \omega_0) \circ X_t^{-1} \right], \quad (1.8)$$

where X_t is the same as in (1.7) determined by the velocity u . Then the velocity can be recovered from the vorticity using the Biot-Savart law. Actually, this representation is essentially the Feynman-Kac representation. This stochastic representation of the vorticity has been used in [7]. For the Navier-Stokes equations, the following stochastic

Lagrangian representation of the velocity

$$u(t, x) = \mathbb{E}\mathbf{P}[(\nabla^T X_t^{-1})(u_0 \circ X_t^{-1})]$$

is obtained in Constantin and Iyer [11, 12] and Zhang [91]. Here $\mathbf{P}$ is the Leray-Hodge projection onto divergence free vector fields and ∇^T is the transpose of ∇ .

In all these papers [7, 11, 12, 91], they use these stochastic representations to prove the uniqueness of the short time strong solutions, which is to choose an appropriate function space so that these representations form contraction mappings for small time t . Different from their work, our motivation here is to consider the regularity results of the weak solutions through these representations, which imposes the need of constructing such coordinates.

1.2 Linearized equations and diffusion processes

As mentioned in the previous section, we would like to solve SDEs of the form (1.7), which is closely related to a class of parabolic equations. Motivated by this, we consider, in a more general form, parabolic equations of second order with singular divergence-free drift

$$Lu = \partial_t u(t, x) - \sum_{i,j=1}^n \partial_{x_i}(a_{ij}(t, x) \partial_{x_j} u(t, x)) + \sum_{i=1}^n b_i(t, x) \partial_{x_i} u(t, x) = 0, \quad (1.9)$$

where (a_{ij}) is a symmetric matrix-valued and Borel measurable function on $\mathbb{R}^n$. Throughout this thesis, we always assume that there exists a number $\lambda > 0$ such that

$$\lambda |\xi|^2 \leq \sum_{i,j=1}^n a_{ij} \xi_i \xi_j \leq \frac{1}{\lambda} |\xi|^2 \quad \mathbf{E}$$

for all $\xi \in \mathbb{R}^n$, and that $b = (b_i)$ is a divergence-free vector field, i.e.

$$\sum_{i=1}^n \partial_{x_i} b_i(t, x) = 0 \quad \mathbf{S}$$

in the sense of distributions for all t .

Equation (1.9) has been well-studied without the divergence-free condition (S). Let us first consider the regular case where (a, b) are smooth, bounded and possess bounded derivatives of all orders on $[0, \infty) \times \mathbb{R}^n$. It is known that (see Friedman [28], Theorem 11 and 12, Chapter 1), under condition (E) and smoothness assumptions on (a, b) , there is a unique positive *fundamental solution* $\Gamma(t, x; \tau, \xi)$ to (1.9), and it is smooth in $(t, x; \tau, \xi)$ on $0 \leq \tau < t < \infty$ and $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$. Recall that the following properties are satisfied.

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

$$\partial_t \Gamma(t, x; \tau, \xi) - \sum_{i,j=1}^n \partial_{x_i} (a_{ij}(t, x) \partial_{x_j} \Gamma(t, x; \tau, \xi)) + \sum_{i=1}^n b_i(t, x) \partial_{x_i} \Gamma(t, x; \tau, \xi) = 0. \quad (1.10)$$

- 3) Chapman-Kolmogorov's equation holds

$$\Gamma(t, x; \tau, \xi) = \int_{\mathbb{R}^n} \Gamma(t, x; s, z) \Gamma(s, z; \tau, \xi) dz. \quad (1.11)$$

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

$$\lim_{t \downarrow \tau} \int_{\mathbb{R}^n} f(\xi) \Gamma(t, x; \tau, \xi) d\xi = f(x) \quad (1.12)$$

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

These results are crucial to our *a priori* estimates that we are going to state later since they allow us to apply many calculations without concerning integrability and differentiability. For the corresponding diffusion processes, it is well known that Lipschitz coefficients with linear growth give a unique strong solution to the SDE.

For measurable coefficients, classical solutions to the PDE no longer exist and the concept of weak solution is introduced. A classical monograph on weak solutions is [44] by Ladyzhenskaya, Solonnikov, and Ural'ceva, in which if b is assumed to be in

$L^l(0, T; L^q(\mathbb{R}^n))$ with

$$\gamma = \frac{2}{l} + \frac{n}{q} \leq 1, \quad l \in [2, \infty) \text{ and } q \in (n, \infty],$$

then for any initial data $u_0 \in L^2$, there is a unique weak solution with Hölder regularity satisfying that $\partial_t u \in L^2(0, T; H^{-1}(\mathbb{R}^n))$, $u \in C([0, T], L^2(\mathbb{R}^n))$ and the energy identity

$$\frac{1}{2} \|u(T)\|_{L^2}^2 + \int_0^T \int_{\mathbb{R}^n} \langle a(x) \cdot \nabla u(t, x), \nabla u(t, x) \rangle dx dt = \frac{1}{2} \|u_0\|_{L^2}^2. \quad (1.13)$$

If $\gamma < 1$, in [3], Aronson proved that there exist Gaussian upper and lower bounds for the fundamental solution, from which the Hölder continuity of weak solutions can be deduced. We call such estimate on fundamental solutions the Aronson estimate. Similar conditions are also obtained in Krylov and Röckner [42], which proved that the SDE

$$dX_t = dB_t + b(t, X_t)dt$$

has a unique strong solution if $\gamma < 1$.

The reason why we require such conditions on γ can be easily seen from the natural scaling property of parabolic equations. Under the following scaling transformation

$$u^{(\rho)}(t, x) = u(\rho^2 t, \rho x), \quad a^{(\rho)}(t, x) = a(\rho^2 t, \rho x), \quad b^{(\rho)}(t, x) = \rho b(\rho^2 t, \rho x)$$

for $\rho > 0$, if u is a solution to (1.9) with ellipticity constant λ , then $u^{(\rho)}$ is still a solution to the parabolic equation with $(a^{(\rho)}, b^{(\rho)})$, and condition (E) is still satisfied with the same λ .

If $b \in B$, where B is a Banach space, then based on $\|\rho b(\rho^2 t, \rho x)\|_B$ converging to 0, $+\infty$ or being bounded as $\rho \rightarrow 0$, we call them subcritical, supercritical and critical respectively.

We can deduce that estimates depending on $\|b\|_B$ vary accordingly on finer scales. The critical and subcritical conditions on b imply that estimates are uniform on all finer scales and hence solutions have better regularity, while the supercritical case does not. This

classification allows us to say that $b \in L_t^l L_x^q$ is critical (or subcritical and supercritical) if $\frac{2}{l} + \frac{n}{q} = 1$ (or < 1 and > 1 respectively). In the supercritical case, we are unable to obtain the Harnack inequality uniformly for small scales. But in the critical and subcritical cases, we still can control solutions for small scales to obtain the Harnack inequality, and hence obtain Hölder regularity. However, an exceptional case is $L^\infty(0, T; L^n(\mathbb{R}^n))$, which is critical, but the Harnack inequality fails.

Moreover, many regularity results regarding the Navier-Stokes equations also reflects this observation from scaling. In Prodi [67], Serrin [75] and Ladyzhenskaya [43], it was proved that if Leray-Hopf weak solution u to the 3D Navier-Stokes equations satisfies the Ladyzhenskaya-Prodi-Serrin condition, i.e. $u \in L^l(0, T; L^q(\mathbb{R}^3))$ with $\frac{2}{l} + \frac{3}{q} = 1$ and $q \in (3, \infty]$, then u is the unique solution and it is smooth. Later the results are extended in Escauriaza, Seregin and Šverák [19] to the marginal case $u \in L^3(0, T; L^\infty(\mathbb{R}^3))$. This result is closely related to the divergence-free condition, which we will discuss in the next section.

1.3 Divergence-free condition

In this section, we would like to discuss why the divergence-free condition on b is so important to our problem. As shown in the last section, without the divergence-free condition, all results are confined to subcritical or critical conditions on b . Divergence-free is the key condition that allows us to extend the classical results mentioned in the last section. Recall that a vector field $b = (b_i)$ being divergence-free, i.e. $\nabla \cdot b = 0$, implies that there exists a skew-symmetric tensor (d_{ij}) such that $b = \operatorname{div} d$ and

$$b \cdot \nabla = \sum_{i,j=1}^3 \frac{\partial d_{ij}}{\partial x_i} \frac{\partial}{\partial x_j}.$$

If we denote $A = a - d$, equation (1.9) can be put into a divergence form

$$\partial_t u(t, x) - \sum_{i,j=1}^n \partial_{x_i} (A_{ij}(t, x) \partial_{x_j} u(t, x)) = 0, \quad (1.14)$$

where $A = (A_{ij})$ is not necessarily symmetric. The symmetric part (a_{ij}) is uniformly elliptic, and the skew-symmetric part (d_{ij}) determines the divergence-free drift vector field b . (1.14) is also an important equation in our work.

Another important equation is the adjoint equation of (1.9). Because of the divergence-free condition, the adjoint equation essentially has the same form up to a sign. Hence their fundamental solutions share essentially the same properties. Consider equation (1.9) on $[0, T] \times \mathbb{R}^n$ and denote its fundamental solution as $\Gamma(t, x; \tau, \xi)$ with $0 \leq \tau < t \leq T$, $x, \xi \in \mathbb{R}^n$. For the adjoint equation

$$\partial_t u(t, x) - \sum_{i,j=1}^n \partial_{x_i} (a_{ij}(T-t, x) \partial_{x_j} u(t, x)) - \sum_{i=1}^n b_i(T-t, x) \partial_{x_i} u(t, x) = 0 \quad (1.15)$$

and its fundamental solution $\Gamma_T^*(t, x; \tau, \xi)$, we have $\Gamma(t, x; \tau, \xi) = \Gamma_T^*(T - \tau, \xi; T - t, x)$.

Writing (1.15) into the same form as (1.14), we obtain

$$\partial_t u(t, x) - \sum_{i,j=1}^n \partial_{x_i} (A_{ji}(T-t, x) \partial_{x_j} u(t, x)) = 0, \quad (1.16)$$

where (A_{ji}) is the transpose of (A_{ij}) . Moreover, if we define operator

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

and define its adjoint operator $\Gamma_{r,t}^\perp$ by $\langle \Gamma_{\tau,t} f, g \rangle = \langle f, \Gamma_{\tau,t}^\perp g \rangle$, then we have

$$\Gamma_{r,t}^\perp g(x) = \int_{\mathbb{R}^n} g(\xi) \Gamma_T^*(T - \tau, x; T - t, \xi) d\xi. \quad (1.17)$$

This means that the adjoint operator $\Gamma_{r,t}^\perp$ solves equation (1.15) and hence any estimate on $\Gamma_{r,t}$ also applies to $\Gamma_{r,t}^\perp$.

Recall that we are interested in those b which are the Leray-Hopf weak solutions of the Navier-Stokes equations. For the 3D Navier-Stokes equations, the weak solutions are in space $L^\infty(0, T; L^2(\mathbb{R}^3)) \cap L^2(0, T; L^6(\mathbb{R}^3))$ by Sobolev embedding, and they both give $\frac{2}{l} + \frac{3}{q} = \frac{3}{2}$, which is supercritical. All previous results with critical conditions on b do not hold here. With the additional features brought by the divergence-free condition on b , we expect to relax the assumption on b . We note that the divergence-free condition implies the conservation of volume and energy. Such uniform estimate guarantees the existence of weak solutions for any $b \in L^2(0, T; L^2_{loc}(\mathbb{R}^n))$ and forms the ground on which we can discuss uniqueness and regularity of the weak solutions. We can also consider the meaning of the corresponding diffusion processes.

Here we give a simple proof of the existence of weak solutions to (1.9) when (a, b) satisfies conditions (E), (S) and $b \in L^2(0, T; L^2_{loc}(\mathbb{R}^n))$.

Definition 1.1. A function $u \in L^\infty(0, T; L^2(\mathbb{R}^n)) \cap L^2(0, T; H^1(\mathbb{R}^n))$ is a weak solution to (1.9) corresponding to (a, b) and initial data u_0 if

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^n} u(t, x) \partial_t \varphi(t, x) \, dx dt - \int_0^T \int_{\mathbb{R}^n} \langle a(t, x) \cdot \nabla u(t, x), \nabla \varphi(t, x) \rangle \, dx dt \\ - \int_0^T \int_{\mathbb{R}^n} \langle b(t, x), \nabla u(t, x) \rangle \varphi(t, x) \, dx dt = - \int_{\mathbb{R}^n} u_0(x) \varphi(0, x) \, dx \end{aligned}$$

for any $\varphi \in C_0^\infty([0, T) \times \mathbb{R}^n)$.

Theorem 1.2. Suppose conditions (E) and (S) are satisfied and $b \in L^2(0, T; L^2_{loc}(\mathbb{R}^n))$, there exists a weak solution to (1.9) with initial data $u_0 \in L^2(\mathbb{R}^n)$.

Proof. Denote by u_k the weak solution corresponding to (a, b_k) , where $b_k \in C_0^\infty((0, T) \times \mathbb{R}^n)$ are divergence-free and $b_k \rightarrow b$ in $L^2(0, T; L^2_{loc}(\mathbb{R}^n))$. Then $\{u_k\}$ is uniformly bounded in $L^\infty(0, T; L^2(\mathbb{R}^n)) \cap L^2(0, T; H^1(\mathbb{R}^n))$ and hence has a sub-sequence which converges weakly to some u . This weak convergence allows us to take limit as $k \rightarrow \infty$ in the equation:

$$\int_0^T \int_{\mathbb{R}^n} u_k(t, x) \partial_t \varphi(t, x) \, dx dt - \int_0^T \int_{\mathbb{R}^n} \langle a(t, x) \cdot \nabla u_k(t, x), \nabla \varphi(t, x) \rangle \, dx dt$$

$$-\int_0^T \int_{\mathbb{R}^n} \langle b_k(t, x), \nabla u_k(t, x) \rangle \varphi(t, x) \, dx dt = -\int_{\mathbb{R}^n} u_0(x) \varphi(0, x) \, dx$$

to obtain that

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^n} u(t, x) \partial_t \varphi(t, x) \, dx dt - \int_0^T \int_{\mathbb{R}^n} \langle a(t, x) \cdot \nabla u(t, x), \nabla \varphi(t, x) \rangle \, dx dt \\ & - \int_0^T \int_{\mathbb{R}^n} \langle b(t, x), \nabla u(t, x) \rangle \varphi(t, x) \, dx dt = -\int_{\mathbb{R}^n} u_0(x) \varphi(0, x) \, dx. \end{aligned}$$

□

1.4 Main results

In this section, we review related literature and summarize our main results. With the assumption that b is divergence-free, we approach equation (1.9) by estimating its fundamental solution, which is called the Aronson estimate. We will discuss critical conditions and supercritical conditions on b separately. The first result we mention below is an Aronson estimate of its classical form under a critical condition on b , or equivalently a critical condition on the anti-symmetric part $d = A - A^T$ in equation (1.14).

Theorem 1.3. *Suppose Γ is the fundamental solution to (1.14) satisfying condition (E). There is a constant $C > 0$ depending only on the dimension n , the elliptic constant $\lambda > 0$, and the $L^\infty(0, \infty; BMO)$ norm of the skew-symmetric part $d_{ij} = \frac{1}{2} (A_{ij} - A_{ji})$ such that*

$$\frac{1}{C(t - \tau)^{n/2}} \exp \left[-\frac{C|x - \xi|^2}{t - \tau} \right] \leq \Gamma(t, x; \tau, \xi) \leq \frac{C}{(t - \tau)^{n/2}} \exp \left[-\frac{|x - \xi|^2}{C(t - \tau)} \right] \quad (1.18)$$

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

$$\|d\|_{L^\infty(BMO)} = \sup_{t \geq 0} \sqrt{\sum_{i < j} \|d_{ij}(t, \cdot)\|_{BMO}^2}.$$

The $L^\infty(0, \infty; BMO)$ ($L_t^\infty BMO_x$ for short) condition on d here is equivalent to the $L^\infty(0, \infty; BMO^{-1})$ ($L_t^\infty BMO_x^{-1}$ for short) condition on b because of the equivalence of

equation (1.9) and equation (1.14). As mentioned above, a marginal case $L^\infty(0, T; L^n(\mathbb{R}^n))$ is excluded in [44], although it is critical. However, with the divergence-free condition and $L^n(\mathbb{R}^n) \subset \text{BMO}^{-1}(\mathbb{R}^n)$, our results actually cover the marginal case $L^\infty(0, T; L^n(\mathbb{R}^n))$. The BMO^{-1} condition on the velocity has been well-studied in many Navier-Stokes equations literatures. We refer to [48] for more details about its importance in the study of the Navier-Stokes equations.

The heat kernel estimate (1.18) for parabolic equations has a long history. Two sided estimate (1.18) was first established in Aronson [1, 2] for uniformly elliptic operators in divergence form where (A_{ij}) is symmetric (so that $(d_{ij}) \equiv 0$), in which his constant C depends only on the elliptic constant λ and the dimension n . The estimate (1.18) is therefore referred to as the Aronson estimate. A weaker but global estimate similar to (1.18) under the same assumption as in Aronson [2] already appeared in the Appendix of Nash [64]. Aronson [2, 3] indicated that his estimate can be established for a general elliptic operator, and a written proof is available in Fabes and Stroock [21], Stroock [79], Norris and Stroock [65] too. In these papers, the Aronson estimate (1.18) was established for the following type of uniformly elliptic operator

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

where (a_{ij}) is symmetric and uniformly elliptic. In this case, the constant C depends on the dimension, the elliptic constant λ and the $L_{t,x}^\infty$ -norms of $b, \hat{b}$ and c .

The Aronson estimate is related to the regularity of solutions to the parabolic equation $(\partial_t - L)u = 0$ (see [44] for a complete survey of classical results). If the elliptic operator is symmetric and is in divergence form, it was Nash [64] 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 [64], in 1964, Moser [62] established the Harnack inequality for positive solutions of the parabolic equation $(\partial_t - L)u = 0$, based on which Aronson was able to derive his estimate (1.18).

Fabes and Stroock [21] showed that Moser's Harnack inequality could be derived from the Aronson estimate together with Nash's idea, and Stroock [79] further demonstrated that both the Hölder continuity of classical solutions and Moser's Harnack inequality for positive solutions could be established by utilizing the two sided Aronson estimate (1.18). Nash's idea in [64] and the techniques in Moser [61, 62, 63] 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 [29]. See for example Grigor'yan [32], Davies [15] and Stroock [80] for a small sample of references, and also the literature therein.

The case that (A_{ij}) 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 [66], the Aronson estimate (1.18) was obtained for an elliptic operator in divergence form, where (A_{ij}) may not be symmetric. His constant C in (1.18) however depends on the dimension n , the elliptic constant λ and the $L_{t,x}^\infty$ -norm of the skew-symmetric part (d_{ij}) . In recent works by Seregin, Silvestre, Šverák, and Zlatoš [74] and Friedlander, Vicol [27], they both relaxed the condition on (d_{ij}) from $L_{t,x}^\infty$ to $L_t^\infty(\text{BMO}_x)$ using De Giorgi's technique. In [27], the Harnack inequality is proved under assumptions that (d_{ij}) is in $L_t^\infty(\text{BMO}_x)$ and satisfies $\text{div } d \in L_t^2 L_x^2$. In [74], the main result is the Harnack inequality for weak solutions satisfying an additional local energy inequality. The relaxation from $L_{t,x}^\infty$ to $L_t^\infty(\text{BMO}_x)$ here is non-trivial, which is also exposed in [27] and [74]. It is worth mentioning that the Harnack inequalities proved in [27] and [74] are both local, i.e. they are true for any local weak solution. We take an approach different from the techniques used in [27] and [74], which is to estimate the fundamental solution and relies on several versions of compensated compactness results. In the proof of the upper bound of the Aronson estimate, we use essentially Proposition 2.2, which is in the same spirit as the classical compensated compactness results and is new to the knowledge of the author.

Moreover, it is mentioned in [74] 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 is motivated by the observation made by Seregin et al. [74] to obtain a better estimate of Γ , and the approach put forward by Davies [14], Fabes and Stroock [21], and Stroock [79]. We follow the approach in Davies and Stroock to work on the non-symmetric case, and adapt their arguments to our case by overcoming the difficulties arising from the singularities of the skew-symmetric part (d_{ij}) .

As applications of the Aronson estimate, we have the following continuity theorem and the Harnack inequality as in Stroock [79].

Theorem 1.4. *There exist $C > 0$ and $\alpha \in (0, 1)$ depending only on the dimension n , the elliptic constant λ and the $L_t^\infty BMO_x$ -norm of the skew-symmetric part (d_{ij}) , such that for every $\delta > 0$*

$$|\Gamma(t, x; \tau, \xi) - \Gamma(t', x'; \tau, \xi')| \leq \frac{C}{\delta^n} (|t' - t| \vee |x' - x| \vee |\xi' - \xi|)^\alpha \quad (1.19)$$

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

Theorem 1.5. *[Harnack Inequality] There exists a constant $C > 0$ depending only on n, λ and $\|d\|_{L^\infty(BMO)}$ such that given any non-negative $v \in L^2(\mathbb{R}^n)$ with $v \geq 0$ and set $u(t, x) = \Gamma_{\tau, t} v(x)$, we have*

$$\sup_{[s, s+R^2] \times B(x_0, R)} u(t, x) \leq C \inf_{[s+3R^2, s+4R^2] \times B(x_0, R)} u(t, y) \quad (1.20)$$

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

For supercritical conditions, we focus on the cases that b belongs to Lebesgue spaces $L^l(0, T; L^q(\mathbb{R}^n))$ ($L_t^l L_x^q$ for short) for $l, q \in [1, \infty]$, and we denote

$$\Lambda := \|b\|_{L_t^l L_x^q} = \left(\int_0^T \left(\int_{\mathbb{R}^n} |b(t, x)|^q dx \right)^{\frac{l}{q}} dt \right)^{\frac{1}{l}}.$$

Since the divergence-free condition on drift b prevents the formation of local blow up, we have the following upper bound which is of exponential decay.

Theorem 1.6. *Suppose conditions **(E)**, **(S)** hold and $b \in L^l(0, T; L^q(\mathbb{R}^n))$ for some $n \geq 3$, $l > 1$, $q > \frac{n}{2}$ such that $1 \leq \gamma < 2$. In addition, we assume that a and b are smooth with bounded derivatives of all orders. Let $\Gamma(t, x; \tau, \xi)$ be the fundamental solution of (1.9). Then*

$$\Gamma(t, x; \tau, \xi) \leq \frac{C}{(t - \tau)^{n/2}} \exp(m(t - \tau, x - \xi)), \quad (1.21)$$

where

$$m(t, x) = \min_{\alpha \in \mathbb{R}^n} (C(|\alpha|^2 t + |\alpha|^\mu \Lambda^\mu t^\nu) + \alpha \cdot x)$$

with $\Lambda = \|b\|_{L^l(0, T; L^q(\mathbb{R}^n))}$, $\mu = \frac{2}{2 - \gamma + \frac{2}{l}}$, $\nu = \frac{2 - \gamma}{2 - \gamma + \frac{2}{l}}$ and $C = C(l, q, n, \lambda)$.

The point here is that the upper bound above only depends on n , λ and Λ , but not on the estimates of the derivatives of a or b . One restriction of Nash's scheme is that its iteration procedure requires the bound on b to be uniform in time, i.e. $l = \infty$. Therefore in general, we use Moser's iteration scheme instead and use cut-off functions. A more explicit form of the upper bound can be obtained as a corollary of Theorem 1.6.

Corollary 1.7. *Under the same assumptions and notations as in Theorem 1.6, if $\mu \equiv \frac{2}{2 - \gamma + \frac{2}{l}} > 1$, the fundamental solution has an upper bound*

$$\Gamma(t, x; \tau, \xi) \leq \begin{cases} \frac{C_1}{(t - \tau)^{n/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|x - \xi|^2}{t - \tau}\right)\right) & \frac{|x|^{\mu-2}}{t^{\mu-\nu-1}} < 1 \\ \frac{C_1}{(t - \tau)^{n/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|x - \xi|^\mu}{(t - \tau)^\nu}\right)^{\frac{1}{\mu-1}}\right) & \frac{|x|^{\mu-2}}{t^{\mu-\nu-1}} \geq 1, \end{cases} \quad (1.22)$$

where $\Lambda = \|b\|_{L^l(0, T; L^q(\mathbb{R}^n))}$, $C_1 = C_1(l, q, n, \lambda)$, $C_2 = C_2(l, q, n, \lambda, \Lambda)$. If $\mu = 1$, which

implies $q = \infty$, we can solve for $m(t, x)$ explicitly and obtain

$$\Gamma(t, x; \tau, \xi) \leq \frac{C_1}{(t - \tau)^{n/2}} \exp \left(- \frac{(C_1 \Lambda(t - \tau)^\nu - |x - \xi|)^2}{4C_1(t - \tau)} \right). \quad (1.23)$$

Similar estimates on the fundamental solution are obtained for several entropy conditions on b , which are supercritical under scalling. In [88], Zhang obtained the exponential decay upper bound for the fundamental solution when $m \in (1, 2]$ and b satisfies the following entropy condition

$$\int_0^T \int_{\mathbb{R}^n} |b|^m \varphi^2 dx dt \leq C \int_0^T \int_{\mathbb{R}^n} |\nabla \varphi|^2 dx dt \quad (1.24)$$

for every smooth function φ on $[0, T] \times \mathbb{R}^n$ with compact support in space. Such condition can be traced back to [41], in which the previous entropy condition was first introduced for the time independent case in order to construct a semigroup theory. The Sobolev embedding allows us to deduce from the entropy condition (1.24) that $b \in L^\infty(0, T; L^{\frac{mn}{2}}(\mathbb{R}^n))$. Therefore, the entropy condition is effectively scaling invariant when $m = 2$, i.e. it is a critical case, while it is supercritical if $m \in (1, 2)$. For the supercritical case, Zhang [89] further considered the following entropy condition:

$$\int_0^T \int_{\mathbb{R}^n} |b|(\ln(1 + |b|))^2 \varphi^2 dx dt \leq C \int_0^T \int_{\mathbb{R}^n} |\nabla \varphi|^2 dx dt, \quad (1.25)$$

and proved the existence of a bounded weak solution in this case. Under supercritical conditions, in [37], assuming that $b \in L_{t,x}^q \cap L_t^\infty L_x^2$ with $q \in (\frac{n}{2} + 1, n + 2]$, Ignatova, Kukavica and Ryzhik proved a weak Harnack inequality. The constant in the weak Harnack inequality explodes as the radius of the parabolic ball goes to zero, which is consistent with our observation through scaling. Hence it fails to yield the Hölder continuity of weak solutions.

In the supercritical case $\gamma \in (1, 2)$, using the upper bound (1.21), we still can derive a lower bound. Actually, we know that the fundamental solution is conservative, and in

fact, for fixed $t > \tau$, $\Gamma(t, x; \tau, \xi)$ is a probability density in x (and in ξ as well due to the divergence-free condition). Because the upper bound decays exponentially, we can find a radius $\tilde{R}(t)$ such that Γ has a lower bound inside the ball of radius $\tilde{R}(t)$. However, we can not hope too much for the lower bound for supercritical case. Using current techniques, we establish the following theorem.

Theorem 1.8. *Assume that a and b are smooth with bounded derivatives. Suppose conditions **(E)**, **(S)** hold and $b \in L^l(0, T; L^q(\mathbb{R}^n))$ for some $n \geq 3$, $l \geq 2$, $q \geq 2$ such that $1 < \gamma < 2$. For any $\kappa > 0$, there is a constant $C > 0$ depending only on κ, l, q, n, λ and $\Lambda = \|b\|_{L^l(0, T; L^q(\mathbb{R}^n))}$ such that*

$$\Gamma(t, x; 0, \xi) \geq \exp \left[-C t^{(\frac{n}{2}+1)(1-\gamma)} \left(\ln \frac{1}{t} \right)^{n+2} \right] \quad (1.26)$$

for $x, \xi \in B(0, \kappa \tilde{R}(t))$ and small enough t , where $\tilde{R}(s) = C s^{(2-\gamma)/2} \ln \frac{1}{s}$ for $s > 0$.

Such form of lower bounds also appear in [4], but in a rather different setting of Dirichlet forms. Although we only deal with the lower bound in the cone $B(0, \tilde{R}(t))$ for small $t > 0$, by the Chapman-Kolmogorov equation, we can extend this lower bound to the whole space. Therefore this form of lower bound is actually the essence of lower bound in the heat kernel estimate, and it determines the local behavior of solutions to the parabolic equation. In all cases the upper bound looks stronger than the lower bound, and the lower bound in the supercritical case fails to yield Hölder continuity of weak solutions. Therefore, in the supercritical case, the regularity theory for this kind of linear parabolic equations remains an open problem.

Not just the regularity, uniqueness is also an important result desired. Recall that the divergence-free condition on b formally gives us the uniform energy identity

$$\frac{1}{2} \|u(T)\|_{L^2}^2 + \int_0^T \int_{\mathbb{R}^n} \langle \nabla u(t, x), a(x) \cdot \nabla u(t, x) \rangle dx dt = \frac{1}{2} \|u_0\|_{L^2}^2,$$

which is independent of the b . The uniqueness of the weak solutions can be instantly obtained if they satisfy this energy identity. With this observation, we first attempted the

case when the coefficients are independent of time.

We consider the time-homogeneous parabolic equation

$$\partial_t u(t, x) - \sum_{i,j=1}^n \partial_{x_i} (a_{ij}(x) \partial_{x_j} u(t, x)) + \sum_{i=1}^n b_i(x) \partial_{x_i} u(t, x) = 0, \quad (1.27)$$

where (a, b) satisfies conditions **(E)** and **(S)**. We use $\Gamma^{(a,b)}(t, x, \xi)$ ($t > 0$) to denote the fundamental solution (recall that a, b are independent of t) which is defined by $\Gamma^{(a,b)}(t - \tau, x, \xi) = \Gamma^{(a,b)}(t, x; \tau, \xi)$. We study the Markov semi-group associated with $\Gamma^{(a,b)}$ for $b \in L^2(\mathbb{R}^n) \cap L^q(\mathbb{R}^n)$, $q > \frac{n}{2}$. The corresponding bi-linear form

$$\mathcal{E}(u, v) = \int_{\mathbb{R}^n} [\langle \nabla u, a \cdot \nabla v \rangle + (b \cdot \nabla u) v] \, dx \quad (1.28)$$

is not sectorial in general in the sense defined in [56] and the theory of non-symmetric Dirichlet forms does not apply in this case. On the other hand, due to the divergence-free condition **(S)**, the symmetric part of the bi-linear form is given by

$$\mathcal{E}_s(u, v) = \int_{\mathbb{R}^n} \langle \nabla u, a \cdot \nabla v \rangle \, dx,$$

which is however sectorial, and $(\mathcal{E}_s, D(\mathcal{E}_s))$ is a Dirichlet form. See for example [30, 56].

We are now in a position to state the result of uniqueness.

Theorem 1.9. *Suppose conditions **(E)**, **(S)** hold and $b \in L^2(\mathbb{R}^n) \cap L^q(\mathbb{R}^n)$ for $q > \frac{n}{2}$. There is a unique Markov semi-group $(P_t)_{t \geq 0}$ on $L^2(\mathbb{R}^n)$ associated with the bi-linear form (1.28) which has transition probability kernel $\Gamma(t, x, y)$ for $t > 0$, $x, y \in \mathbb{R}^n$. Moreover, the uniqueness of weak solutions holds for the Cauchy initial problem to (1.27) and is given by the representation*

$$u(t, x) = \int_{\mathbb{R}^n} \Gamma(t, x, y) u_0(y) \, dy \quad (1.29)$$

for any initial data $u_0 \in L^2(\mathbb{R}^n)$.

When the dimension $n = 3$, the condition of the theorem above is satisfied if $b \in L^2(\mathbb{R}^3)$. As a consequence of Theorem 1.9, we have the following result which is interesting by its own.

Corollary 1.10. *Let b be a C^1 -vector field in $\mathbb{R}^n$ with $n \geq 3$ such that $\nabla \cdot b = 0$. If $b \in L^q(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ where $n \geq q > \frac{n}{2}$, then the diffusion process defined by solving*

$$dX_t = dB_t + b(X_t)dt, \quad X_0 = x$$

is conservative, i.e. its transition density function $\Gamma(t, x, y)$ (with respect to the Lebesgue measure) satisfies that

$$\int_{\mathbb{R}^n} \Gamma(t, x, y) dy = 1, \quad (1.30)$$

where $\Gamma(t, x, y) = \mathbb{P}[X_t = y | X_0 = x]$ formally.

The closest conditions in literature to ensure the stochastic completeness (1.30) for unbounded b are those on the symmetric tensor $\nabla^s b$ (Ricci curvature or Bakry-Émery condition) and $\nabla \cdot b$ is the trace of $\nabla^s b$. This was first proved by Yau [86] that the Brownian motion on a complete Riemannian manifold with Ricci curvature bounded from below is stochastic complete. Later various conditions for the stochastic completeness of a complete Riemannian manifold are studied. For example, one natural idea is to control the volume of the geodesic balls from below to control the speed that the Brownian motion goes to infinity. We refer to Hsu [35, 36] and Grigor'yan [33] for more details. Hence the divergence-free condition imposes a constrain on the “scalar curvature” of the operator $L = \Delta - b \cdot \nabla$. Together with an integrability condition, they implies the stochastic completeness of the process.

There have been many works on the construction of Markov semi-groups from non-sectorial bi-linear forms, which is an important topic in stochastic analysis. In [41], Kovalenko and Semenov proved the existence of a semi-group on L^p for p larger than a certain number under an entropy condition on b . Their entropy condition is still a critical condition on b . Using ideas from Dirichlet form, it is proved in [53] that there exists a C_0 -

semigroup if the drift b is form bounded. Later, Sobol and Vogt [76] proved the existence of a strong continuous semi-group on L^p for any $p \in [1, \infty)$ if the space $Q(b^2) \cap D(\mathcal{E}_s)$ is core for $\mathcal{E}_s$, where $Q(b^2) = \{u \in L^2 : u^2 b^2 \in L^1\}$, and $\mathcal{E}$ is accretive. Their idea is to use the continuity argument. They first add a potential V to the bi-linear form in order to remove the singularity appearing from the drift, and then send the potential to zero. Therefore the association of the semi-group e^{tL} and bi-linear form $\mathcal{E}$ is established through the correspondence of $e^{t(L-V)}$ and $\mathcal{E} + V$. Our approach is to directly approximate b by smooth b_k , which gives the existence and conservative of the kernel. The proof is inspired by Zhikov [98] in which he considered the time-homogeneous parabolic equations

$$\partial_t u - \operatorname{div}((a + d) \cdot \nabla u) = 0, \quad (1.31)$$

and constructed the unique approximation semi-group for periodic $d \in L^\infty(\mathbb{R}^n)$, $\operatorname{div} d \in L^2_{loc}(\mathbb{R}^n)$ and $\sup_{r \geq 1} \frac{1}{r^n} \|d\|_{L^n(B(0,r))}^n < \infty$. Later in [54], Liskevich and Sobol further proved the heat kernel estimate of these semi-groups under additional functional conditions on the bi-linear form, by using the idea developed in [14], which is similar to proving upper bound in time-inhomogeneous cases in [68, 70]. Recently, there are works by Flandoli et al. [26], Zhang and Zhao [94] in which distributional drift b in Sobolev space $L^\infty(0, \infty; W^{-\alpha, p})$ with $\alpha \in (0, \frac{1}{2})$ and $p \in (\frac{n}{1-\alpha}, \frac{n}{\alpha})$ has been studied. They proved uniqueness and a heat kernel estimate. By dimensional analysis, their condition is still subcritical and the scaling property of the Sobolev space implies the uniform control of solutions on finer scales, while in our case where b is supercritical and these uniform control are no longer expected to be true.

To obtain the uniqueness of solutions, another approach arises from the idea of renormalized solutions. Following the work by Le Bris and Lions [46, 47], which first proved the uniqueness of renormalized solutions to (1.9) under singular conditions on b , there are many papers dealing with the related parabolic equations and SDEs (see e.g. [8, 25, 50, 91, 92]). The idea of renormalized solution relaxed the condition on b greatly, which

can include our condition that $b \in L^2(\mathbb{R}^n) \cap L^q(\mathbb{R}^n)$, while it needs more regularity on the diffusion term a in addition to uniform elliptic. It is worth mentioning that the uniqueness of the SDEs derived from the uniqueness of renormalized solutions is only true for almost every initial data $x \in \mathbb{R}^n$ under the Lebesgue measure (see e.g. Figalli [25] and Zhang [91, 92]).

The rest of this thesis is organized as follows. In Chapter 2, we prove several Aronson-type estimates for different critical and supercritical conditions including Theorem 1.3, 1.6 and 1.8. In Chapter 3, we apply the Aronson estimate in the critical case to study the parabolic equations and diffusion processes with singular drift using approximation argument. We show that in the critical case, the Aronson estimate implies Hölder continuity of the weak solutions, and also the uniqueness of the weak solutions. In Chapter 4, we discuss the supercritical case, for which we prove Theorem 1.9 through analyzing the convergence of resolvents. We also present the idea of renormalized solutions, which is very useful in obtaining the uniqueness of solutions. The supercritical case in general is not as good as critical case and we will also discuss the difficulties present. Finally, in chapter 5, we summarize previous chapters and discuss interesting problems which are left open.

Chapter 2

The Aronson estimate

This chapter is devoted to the proof of several Aronson-type estimates. It is worth mentioning that the Aronson estimate is an *a priori* estimate, and hence we will always assume that our coefficients are regular enough through this chapter. The important point is that all constants in these estimates do not depend on the smoothness of the coefficients, which allow us to use approximation argument later to work on singular coefficients.

2.1 Technical facts

In this section, we prove all the technical facts which are used for the proof of the Aronson estimate.

2.1.1 BMO space and compensated compactness

The first result we need is a variation of Coifman-Meyer's compensated compactness Theorem (see e.g. [9, 48]) which highlights the importance of the Hardy spaces in the study of partial differential equations.

We first recall some facts on BMO functions [38, 78]. A function f is in $\text{BMO}(\mathbb{R}^n)$ if

$$\|f\|_{\text{BMO}} = \sup_{B \subset \mathbb{R}^n} \frac{1}{|B|} \int_B |f(x) - f_B| dx < \infty, \quad (2.1)$$

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 we define another norm

$$\|f\|_{\text{BMO}_p}^p = \sup_{B \subset \mathbb{R}^n} \frac{1}{|B|} \int_B |f(x) - f_B|^p dx < \infty \quad (2.2)$$

for any $1 \leq p < \infty$, the John-Nirenberg inequality [38] (see also for example, Appendix in Stroock and Varadhan [81]) implies that $\|\cdot\|_{\text{BMO}_p}$ are equivalent for different p .

The original version of the compensated compactness theorem in [9], which we use in our proof of the lower bound of Aronson estimate, can be stated as follows.

Proposition 2.1. *Let vector fields E, B satisfy that $E \in L^p(\mathbb{R})$, $B \in L^q(\mathbb{R})$ with $\frac{1}{p} + \frac{1}{q} = 1$ ($p \geq 1, q \geq 1$) and $\nabla \cdot E = 0, \nabla \times B = 0$. Then $E \cdot B \in \mathcal{H}^1$ where $\mathcal{H}^1$ is the Hardy space, and*

$$\|E \cdot B\|_{\mathcal{H}^1} \leq C \|E\|_{L^p} \|B\|_{L^q}. \quad (2.3)$$

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

$$\|\nabla f \times \nabla g\|_{\mathcal{H}^1} \leq C \|\nabla f\|_{L^p} \|\nabla g\|_{L^q} \quad (2.4)$$

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

$$\left| \int_{\mathbb{R}^n} \langle \nabla f(x), d(x) \cdot \nabla g(x) \rangle dx \right| \leq C \|d\|_{\text{BMO}} \|\nabla f\|_{L^2} \|\nabla g\|_{L^2} \quad (2.5)$$

for any $f, g \in H^1(\mathbb{R}^n)$ and for any $d = (d_{ij}) \in \text{BMO}$ which is skew-symmetric ($d_{ij} = -d_{ji}$).

In the energy estimate of equation (1.14), the anti-symmetric part d gives a term similar to $\int_{\mathbb{R}^n} \langle \nabla f(x), d(x) \cdot \nabla g(x) \rangle dx$, which appears in the bilinear form. The compensated compactness result above will be the key to obtain energy estimate with $d \in L^\infty(0, \infty; \text{BMO})$. 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. Actually a term like $\int_{\mathbb{R}^n} \langle \nabla f(x), d(x) \cdot \xi \rangle f(x) dx$ will also appear in our bilinear

form. Here ∇f is curl-free, while ξf is not divergence-free and we can not apply the compensated compactness result above. So this leads us to prove the following result.

Proposition 2.2. *There is a universal constant C depending only on the dimension n , such that*

$$\|\xi f \cdot \nabla f\|_{\mathcal{H}^1} \leq C \|\xi\| \|\nabla f\|_{L^2} \|f\|_{L^2} \quad (2.6)$$

for any $f \in H^1(\mathbb{R}^n) = W^{1,2}(\mathbb{R}^n)$ and $\xi \in \mathbb{R}^n$, where $\|\cdot\|_{\mathcal{H}^1}$ 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 $\xi f \cdot \nabla f = \frac{1}{2} \nabla \cdot (f^2 \xi)$ in $L^1(\mathbb{R}^n)$, so

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

where $\int_{B_t(x)}$ 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

$$|I_1| \leq C \left[\int_{B_t(x)} |\xi f|^\alpha \right]^{\frac{1}{\alpha}} \left[\int_{B_t(x)} \left| \left(f(y) - \int_{B_t(x)} f \right) t^{-1} \right|^{\alpha'} \right]^{\frac{1}{\alpha'}}, \quad (2.7)$$

where $\alpha \in [1, 2)$, $\frac{1}{\alpha} + \frac{1}{\alpha'} = 1$. Choose α, β such that $1 \leq \alpha, \beta < 2$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1 + \frac{1}{n}$.

Then by the Sobolev-Poincaré inequality, we have

$$\left[\int_{B_t(x)} \left| \left(f - \int_{B_t(x)} f \right) t^{-1} \right|^{\alpha'} \right]^{\frac{1}{\alpha'}} \leq C \left(\int_{B_t(x)} |\nabla f|^\beta \right)^{\frac{1}{\beta}}. \quad (2.8)$$

For the second term on the right-hand side, we integrate by part again to obtain

$$\begin{aligned}
|I_2| &= \left| \int_{B_t(x)} h\left(\frac{x-y}{t}\right) \frac{1}{t^n} \cdot \operatorname{div}(\xi f(y)) \left(\int_{B_t(x)} f \right) dy \right| \\
&\leq C|\xi| \int_{B_t(x)} |\nabla f(y)| dy \left(\int_{B_t(x)} f \right) dy.
\end{aligned} \tag{2.9}$$

By using these estimates we thus conclude that

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

where $M(f)$ is the maximal function. Since $1 \leq \alpha < 2$, $1 < \beta < 2$, we have

$$\|M(|f|^\alpha)^{\frac{1}{\alpha}}\|_{L^2} \leq C\|f\|_{L^2}, \quad \|M(|\nabla f|^\beta)^{\frac{1}{\beta}}\|_{L^2} \leq C\|\nabla f\|_{L^2},$$

and similarly

$$\|M(|f|)\|_{L^2} \leq C\|f\|_{L^2}, \quad \|M(|\nabla f|)\|_{L^2} \leq C\|\nabla f\|_{L^2}.$$

So $\sup_{t>0} |h_t * (\xi f \cdot \nabla f)| \in L^1$ and

$$\|\xi f \cdot \nabla f\|_{\mathcal{H}^1} \leq C|\xi| \|\nabla f\|_{L^2} \|f\|_{L^2}. \tag{2.10}$$

□

Given a function $d \in L^\infty(0, \infty; \operatorname{BMO}(\mathbb{R}^n))$, we want to approximate it by a mollified sequence, which is not trivial as it looks. A simple example is a vector field $d(t)$ which depends only on t , not on the space variables. Then it may not be in L^1_{loc} and there is no approximation by mollifying sequences. However, the problem considered here allows us

to add a constant to it, i.e. consider $d(t, x) + f(t)$, where $f(t)$ is skew-symmetric so that it will not alter the weak solution formulation of the corresponding parabolic equations. By subtracting the mean value of d on a unit ball, we may assume that

$$d_{B_1(0)}(t) = \oint_{B_1(0)} d(t, x) dx = 0 \quad (2.11)$$

for all $t \in [0, \infty)$. Then for any $r > 0$

$$|d_{B_r(0)}(t)| = |d_{B_r(0)}(t) - d_{B_1(0)}(t)| = \left| \oint_{B_1(0)} d_{B_r(0)}(t) - d(t, x) dx \right| \quad (2.12)$$

$$\leq r^n \oint_{B_r(0)} |d_{B_r(0)}(t) - d(t, x)| dx \leq r^n \|d\|_{L^\infty(\text{BMO}(\mathbb{R}^n))}. \quad (2.13)$$

By the definition of BMO functions, we have

$$\oint_{B_r(0)} |d(t, x) - d_{B_r(0)}(t)|^p dx \leq C \|d\|_{L^\infty(\text{BMO}(\mathbb{R}^n))}^p, \quad (2.14)$$

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

Proposition 2.3. Take $\Phi \in C_0^\infty(B_1(0))$, $\eta \in C_0^\infty((-1, 1))$ 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 $d \in L^\infty(\text{BMO}(\mathbb{R}^n))$ and satisfies (2.11).

Define

$$d_\varepsilon(t, x) = \int_{-\varepsilon}^{+\varepsilon} \int_{B_\varepsilon(0)} \Phi_\varepsilon(y) \eta_\varepsilon(s) d(t-s, x-y) dy ds. \quad (2.15)$$

Then $d_\varepsilon \rightarrow d$ locally in L^p for any $1 \leq p < \infty$, and

$$\|d_\varepsilon\|_{L^\infty(\text{BMO}(\mathbb{R}^n))} \leq \|d\|_{L^\infty(\text{BMO}(\mathbb{R}^n))}. \quad (2.16)$$

Proof. Let x_0 and $r > 0$ be fixed. Let $\oint$ denotes the average integral over $B(x_0, r)$, that is,

$$\oint \phi(y) dy = \oint_{B(x_0, r)} \phi(y) dy.$$

For $x \in B(x_0, r)$ we have

$$\begin{aligned} & \left| d_\varepsilon(t, x) - \oint d_\varepsilon(t, y) dy \right| \\ &= \left| \int_{-\varepsilon}^{+\varepsilon} \int_{B(0, \varepsilon)} \Phi_\varepsilon(y) \eta_\varepsilon(s) d(t-s, x-y) dy ds \right. \\ & \quad \left. - \oint \int_{-\varepsilon}^{+\varepsilon} \int_{B(0, \varepsilon)} \Phi_\varepsilon(z) \eta_\varepsilon(s) d(t-s, y-z) dz ds dy \right| \\ &= \left| \int_{-\varepsilon}^{+\varepsilon} \int_{B(0, \varepsilon)} \Phi_\varepsilon(y) \eta_\varepsilon(s) \left[d(t-s, x-y) - \oint d(t-s, z-y) dz \right] dy ds \right| \\ &\leq \int_{-\varepsilon}^{+\varepsilon} \int_{B(0, \varepsilon)} \Phi_\varepsilon(y) \eta_\varepsilon(s) \left| d(t-s, x-y) - \oint d(t-s, z-y) dz \right| dy ds, \end{aligned}$$

so that

$$\begin{aligned} & \oint \left| d_\varepsilon(t, x) - \oint d_\varepsilon(t, y) dy \right| dx \\ &\leq \oint \int_{-\varepsilon}^{+\varepsilon} \int_{B(0, \varepsilon)} \Phi_\varepsilon(y) \eta_\varepsilon(s) \left| d(t-s, x-y) - \oint d(t-s, z-y) dz \right| dy ds dx \\ &= \int_{-\varepsilon}^{+\varepsilon} \int_{B(0, \varepsilon)} \Phi_\varepsilon(y) \eta_\varepsilon(s) \oint \left| d(t-s, x-y) - \oint d(t-s, z-y) dz \right| dx dy ds \\ &\leq \int_{-\varepsilon}^{+\varepsilon} \int_{B(0, \varepsilon)} \Phi_\varepsilon(y) \eta_\varepsilon(s) \|d\|_{L^\infty(\text{BMO}(\mathbb{R}^n))} dy ds \\ &= \|d\|_{L^\infty(\text{BMO}(\mathbb{R}^n))}. \end{aligned}$$

Now we proved that $\|d_\varepsilon\|_{L^\infty(\text{BMO}(\mathbb{R}^n))} \leq \|d\|_{L^\infty(\text{BMO}(\mathbb{R}^n))}$. □

The lattice property of the BMO space in the proposition below should be well known.

We include a proof here for completeness.

Proposition 2.4. *Suppose $f, g \in \text{BMO}(\mathbb{R}^n)$, then $f \wedge g$ and $f \vee g \in \text{BMO}(\mathbb{R}^n)$. Moreover,*

we have

$$\|f \wedge g\|_{BMO} \leq C(\|f\|_{BMO} + \|g\|_{BMO}), \quad (2.17)$$

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

$$|a \wedge b - c \wedge d| \leq |a - c| + |b - d|. \quad (2.18)$$

Hence for any ball B ,

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

and the proof is complete. □

2.1.2 Poincaré-Wirtinger inequality

We will need the following Poincaré-Wirtinger inequality for the Gaussian measures [6, Corollary 1.7.3] in the proof of the Aronson-type estimate. For the completeness, we will give a short proof here. In the sequel, we shall write $C_b^1(\mathbb{R}^n)$ to be the space of functions with bounded continuous first order derivatives.

Lemma 2.5. *Let μ be the standard Gaussian measure on $\mathbb{R}^n$, i.e. $\mu(dx) = \mu(x)dx$ with $\mu(x) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{|x|^2}{2}\right)$. Then for every $p \geq 1$*

$$\int_{\mathbb{R}^n} |f(x) - \bar{f}|^p \mu(dx) \leq M(p) \left(\frac{\pi}{2}\right)^p \int_{\mathbb{R}^n} |\nabla f(x)|^p \mu(dx), \quad (2.19)$$

for any $f \in C_b^1(\mathbb{R}^n)$, where $\bar{f} = \int_{\mathbb{R}^n} f(x) \mu(dx)$ and

$$M(p) = \int_{-\infty}^{\infty} |\xi|^p \frac{1}{(2\pi)^{1/2}} \exp\left(-\frac{1}{2}|\xi|^2\right) d\xi.$$

Further, setting $\mu_r(x) = \frac{1}{r^{n/2}} \exp\left(-\frac{\pi|x|^2}{r}\right)$ and $\bar{f}_r = \int_{\mathbb{R}^n} f(x) \mu_r(dx)$, one has

$$\int_{\mathbb{R}^n} |f(x) - \bar{f}_r|^p \mu_r(dx) \leq M(p) \left(\frac{\pi}{2}\right)^p \left(\frac{r}{2\pi}\right)^{p/2} \int_{\mathbb{R}^n} |\nabla f(x)|^p \mu_r(dx). \quad (2.20)$$

Proof. Suppose X_0 and X_1 are independent random vectors of n -dimensional normal distribution $\mathcal{N}_n(0, I_n)$, i.e. they have mean zero and covariance matrix I_n . Then their density functions are both $\mu(x)$. Consider

$$Y(\theta) = X_0 \sin \theta + X_1 \cos \theta$$

with $\theta \in [0, \frac{\pi}{2}]$, which is also of distribution $\mathcal{N}_n(0, I_n)$. Given $f \in C_b^1(\mathbb{R}^n)$, we will have

$$f(Y(\frac{\pi}{2})) - f(Y(0)) = \int_0^{\frac{\pi}{2}} \nabla f(Y(\theta)) \cdot \frac{d}{d\theta} Y(\theta) d\theta,$$

where $\frac{d}{d\theta} Y(\theta) = X_0 \cos \theta - X_1 \sin \theta$ is of distribution $\mathcal{N}_n(0, I_n)$ and independent of $Y(\theta)$.

Now we take the L^p -norm under the expectation on both sides to obtain that

$$\begin{aligned} \mathbb{E} \left[\left| f(Y(\frac{\pi}{2})) - f(Y(0)) \right|^p \right] &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x) - f(y)|^p \mu(dy) \mu(dx) \\ &\geq \int_{\mathbb{R}^n} |f(x) - \bar{f}|^p \mu(dx) \end{aligned}$$

by Jensen's inequality, and

$$\begin{aligned} &\mathbb{E} \left[\left| \int_0^{\frac{\pi}{2}} \nabla f(Y(\theta)) \cdot \frac{d}{d\theta} Y(\theta) d\theta \right|^p \right] \\ &\leq \mathbb{E} \left[\left(\frac{\pi}{2} \right)^{\frac{p}{q}} \int_0^{\frac{\pi}{2}} \left| \nabla f(Y(\theta)) \cdot \frac{d}{d\theta} Y(\theta) \right|^p d\theta \right] \end{aligned}$$

$$\begin{aligned}
&= \left(\frac{\pi}{2}\right)^{\frac{p}{q}} \int_0^{\frac{\pi}{2}} \mathbb{E} \left[\left| \nabla f(Y(\theta)) \cdot \frac{d}{d\theta} Y(\theta) \right|^p \right] d\theta \\
&= \left(\frac{\pi}{2}\right)^p \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |\nabla f(x) \cdot y|^p \mu(dy) \mu(dx). \\
&= \left(\frac{\pi}{2}\right)^p \int_{\mathbb{R}^n} |\nabla f(x)|^p \int_{\mathbb{R}^n} |y|^p \cos \langle y, \nabla f(x) \rangle \mu(dy) \mu(dx),
\end{aligned}$$

where $\frac{1}{q} + \frac{1}{p} = 1$ and $\langle y, \nabla f(x) \rangle$ is the angle between y and $\nabla f(x)$. Since the Gaussian measure μ is symmetric, we actually have that

$$\int_{\mathbb{R}^n} |y|^p \cos \langle y, x \rangle \mu(dy) = \int_{\mathbb{R}^n} |y_1|^p \mu(dy) = \int_{-\infty}^{\infty} |y_1|^p \frac{1}{(2\pi)^{1/2}} \exp\left(-\frac{1}{2}|y_1|^2\right) dy_1$$

for any $x \in \mathbb{R}^n$. Plug this in and the proof of (2.19) is complete. Finally, to prove inequality (2.20), we just need to scale the coordinate by $y = r^{-\frac{1}{2}}x$ and it follows from inequality (2.19), so we will omit the detail here. $\square$

Remark 2.6. Lemma 2.5 can be extended to any function with weak derivative such that both sides of the Poincaré-Wirtinger inequality are well-defined using a truncation and approximation argument. We refer to [6] for more details.

2.1.3 A Riccati differential inequality

We will need the following lemma on a Riccati differential inequality for proving the lower bound of the Aronson estimate. The first lemma below can be found in [79] and we include the proof here for the sake of completeness. The second lemma is new to the knowledge of the author.

Lemma 2.7. *Suppose a non-positive valued function u is continuous and differentiable on $[\frac{T}{2}, T]$, where $T > 0$ is a constant. If u satisfies the Riccati differential inequality*

$$u'(t) \geq -\alpha + \beta u(t)^2 \tag{2.21}$$

for $t \in [\frac{T}{2}, T]$, where $\alpha, \beta > 0$ are two constants, then

$$u(T) \geq \min \left\{ -\alpha T - 2\sqrt{\frac{\alpha}{\beta}}, -\frac{8}{3\beta T} \right\}.$$

Proof. If $u(T) \geq -\alpha T - 2\sqrt{\frac{\alpha}{\beta}}$, then the proof is done. Otherwise, integrating the differential inequality (2.21) from $T/2$ to T , we have $u(T) - u(t) \geq -\frac{\alpha T}{2}$ for any $t \in [\frac{T}{2}, T]$. In other words, we have

$$u(t) \leq u(T) + \frac{\alpha T}{2} \leq -\alpha T - 2\sqrt{\frac{\alpha}{\beta}} + \frac{\alpha T}{2},$$

which in turn yields that $u(t) \leq -2\sqrt{\frac{\alpha}{\beta}}$. Notice that $u(t)$ is negative on $[\frac{T}{2}, T]$ and therefore $u(t)^2 \geq 4\frac{\alpha}{\beta}$. Hence differential inequality (2.21) implies that

$$u'(t) \geq \beta \left(-\frac{\alpha}{\beta} + u(t)^2 \right) \geq \beta \left(-\frac{1}{4}u(t)^2 + u(t)^2 \right) = \frac{3\beta}{4}u(t)^2$$

for every $t \in [\frac{T}{2}, T]$. Dividing both sides by $u(t)^2$ and integrating from $t \in [\frac{T}{2}, T]$ to T , we obtain that

$$\frac{1}{u(T)} \leq -\frac{3\beta T}{8} + \frac{1}{u(t)} \leq -\frac{3\beta T}{8}.$$

In particular, $u(T) \geq -\frac{8}{3\beta T}$ and the proof is complete. $\square$

If α is a function depending on t , which is integrable and non-negative, then we still can derive a lower bound on u .

Lemma 2.8. *Let $T > 0$. Suppose that non-positive function u is continuous on $[\frac{T}{2}, T]$ and satisfies the following integral inequality*

$$u(t_2) - u(t_1) \geq \int_{t_1}^{t_2} (-\alpha(t) + \beta u(t)^2) dt \quad \text{for all } \frac{T}{2} \leq t_1 < t_2 \leq T, \quad (2.22)$$

where α is non-negative and integrable on $[\frac{T}{2}, T]$, and $\beta > 0$ is a constant, then

$$u(T) \geq - \int_{\frac{T}{2}}^T \alpha(t) dt - C\beta^{-1}T^{-1}$$

for some $C > 0$.

Proof. Let $C_2 = \int_{\frac{T}{2}}^T \alpha(t) dt$ and $C_1 > C_2$ be a constant to be determined later. Suppose $u(T) < -C_1$. Then for any $t \in [\frac{T}{2}, T]$ it holds that

$$u(t) \leq u(T) + \int_t^T \alpha(s) ds < -C_1 + C_2 =: -C_3,$$

where $C_3 > 0$ since $C_1 > C_2$. Since $u(t)$ is negative, this implies that $u(t)^2 \geq C_3^2$. Now (2.22) gives us

$$u(t) \leq u(T) + \int_t^T \alpha(s) ds - \int_t^T \beta u(s)^2 ds < - \int_t^T \beta u(s)^2 ds, \quad (2.23)$$

which implies that $u(t) \leq -\beta C_3^2(T-t)$ for all $t \in [\frac{T}{2}, T]$. Repeating the procedure of using the old bound on $u(t)$ and (2.23) to obtain a new bound, we deduce that

$$u(t) \leq -\beta^{2^m-1} C_3^{2^m} (T-t)^{2^m-1} \prod_{k=1}^m \left(\frac{1}{2^k-1} \right)^{2^{m-k}} \leq -C \left[\beta C_3 (T-t) \prod_{k=1}^m (2^{-\frac{k}{2^k}}) \right]^{2^m}$$

after m times. Since $\inf_m \prod_{k=1}^m (2^{-\frac{k}{2^k}}) = C_4 > 0$, the right-hand side can be arbitrarily small at time $\frac{T}{2}$ if $\beta C_3 (T - \frac{T}{2}) C_4 > 1$, which contradicts to the fact that u is finite. Now if we take $C_3 > \frac{2}{\beta T C_4}$, i.e. $C_1 = C_2 + C\beta^{-1}T^{-1}$ for some constant C , then $u(T) \geq -C_1$. $\square$

2.2 A critical condition on the drift

In this section, we study the divergence form equation (1.14) and prove Theorem 1.3. The proof follows the main lines as in Stroock [79] and Davies [14] from which a clever use of the h -transforms from harmonic analysis is borrowed, while we need to overcome several

difficulties since A is non-symmetric and the skew-symmetry part d is singular. The ideas used are mainly due to Nash [64] and Moser [61, 62, 63].

2.2.1 The upper bound

In this section we show the upper bound in Theorem 1.3:

$$\Gamma(t, x; \tau, \xi) \leq \frac{C}{(t - \tau)^{\frac{n}{2}}} \exp \left[-\frac{|x - \xi|^2}{C(t - \tau)} \right] \quad (2.24)$$

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

The main idea of Davies [14] is to consider the h -transform of the fundamental solution $\Gamma(t, x; \tau, \xi)$ and apply Nash and Moser's iteration to the h -transform 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

$$\|u\|_{L^2}^{2+\frac{4}{n}} \leq C_n \|\nabla u\|_{L^2}^2 \|u\|_{L^1}^{\frac{4}{n}}, \quad \forall u \in L^1(\mathbb{R}^n) \cap H^1(\mathbb{R}^n), \quad (2.25)$$

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

The Nash iteration is neatly described as follows (Stroock [79], Lemma I.1.14).

Lemma 2.9. *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:*

$$u'(t) \leq -\frac{c_1}{p} \frac{t^{(p-2)} u(t)^{1+\beta p}}{w(t)^{\beta p}} + c_2 p u(t), \quad t \geq 0. \quad (2.26)$$

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

$$t^{(p-1)/\beta p} u(t) \leq \left(\frac{K p^2}{\delta} \right)^{\frac{1}{\beta p}} e^{\frac{c_2 \delta t}{p}} w(t), \quad t \geq 0 \quad (2.27)$$

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

The iteration procedure above works in a very general setting, which has been explored since the publication of Nash' paper [64], and it is still the major ingredient in our proof. It is surprising that it works well even in our setting where the diffusion is very singular.

Fortunately as well, Davies' idea [14, 15] also works well for our parabolic equations. Following Davies [14] and Stroock [79], given a smooth function ψ on $\mathbb{R}^n$, we consider

$$\Gamma^\psi(t, x; \tau, \xi) = e^{-\psi(x)} \Gamma(t, x; \tau, \xi) e^{\psi(\xi)}, \quad (2.28)$$

and the linear operator

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

which is defined for non-negative Borel measurable function f , and for f which is smooth with a compact support. Similar to representation (1.17), it is easy to see that the adjoint operator of $\Gamma_{\tau, t}^\psi$ can be identified as the following integral operator

$$\begin{aligned} \Gamma_{\tau, t}^{\psi \perp} f(x) &= \int_{\mathbb{R}^n} f(\xi) \exp(-\psi(\xi)) \Gamma(t, \xi; \tau, x) \exp(\psi(x)) d\xi \\ &= \int_{\mathbb{R}^n} f(\xi) \exp(-\psi(\xi)) \Gamma_T^*(T - \tau, x; T - t, \xi) \exp(\psi(x)) d\xi, \end{aligned}$$

that is

$$\langle \Gamma_{\tau, t}^\psi f, g \rangle = \langle f, \Gamma_{\tau, t}^{\psi \perp} g \rangle$$

for any smooth functions f and g with compact supports.

Lemma 2.10. *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

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

for $t > \tau$, and

$$\begin{aligned} f_t^\perp(x) &= \Gamma_{\tau,t}^{\Psi^\perp} f(x) = \int_{\mathbb{R}^n} f(\xi) e^{\Psi(x) - \Psi(\xi)} \Gamma(t, \xi; \tau, x) d\xi \\ &= \int_{\mathbb{R}^n} f(\xi) e^{\Psi(x) - \Psi(\xi)} \Gamma_T^*(T - \tau, x; T - t, \xi) d\xi \end{aligned}$$

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

There is a constant $C > 0$ depending only on n , λ and $\|d\|_{L_t^\infty BMO_x}$ such that for any $p \geq 1$,

$$\frac{d}{dt} \|f_t\|_{L^{2p}}^{2p} \leq -\lambda \|\nabla f_t^p\|_{L^2}^2 + C_B \frac{p^2 |\alpha|^2}{\lambda} \|f_t\|_{L^{2p}}^{2p} \quad (2.29)$$

for $t > \tau$, and

$$\frac{d}{dt} \|f_t^\perp\|_{L^{2p}}^{2p} \leq -\lambda \|\nabla f_t^{\perp p}\|_{L^2}^2 + C_B \frac{p^2 |\alpha|^2}{\lambda} \|f_t^\perp\|_{L^{2p}}^{2p} \quad (2.30)$$

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

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

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

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

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

and by equation (1.10) we have

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

Similarly, the adjoint equation (1.16) implies that

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

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

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

and similarly we have

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

Since $\frac{d}{dt} \|f_t\|_{L^{2p}}^{2p}$ and $\frac{d}{dt} \|f_t^\perp\|_{L^{2p}}^{2p}$ are similar, we only need to prove (2.29).

Each term I_j on the right-hand side of (2.29) can be dominated as the following. The first three terms I_1 , I_2 and I_3 can be handled exactly as in Davies [14] and Stroock [79]. Recall that $\nabla\psi = \alpha$ is a constant vector. Hence we have

$$I_1 \leq \frac{|\alpha|^2}{\lambda} \|f_t\|_{L^{2p}}^{2p}. \quad (2.31)$$

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

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

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

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

and

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

we deduce that

$$-I_2 - I_3 \leq -\frac{\lambda}{p} \|\nabla f_t^p\|_{L^2}^2 + \frac{p-1}{\lambda} |\alpha|^2 \|f_t\|_{L^{2p}}^{2p}. \quad (2.32)$$

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 estimate (2.6) in Proposition 2.2 to obtain

$$|I_4| = \left| 2 \int_{\mathbb{R}^n} f_t(x)^p \langle \nabla f_t(x)^p, d(t, x) \cdot \alpha \rangle dx \right|$$

$$\leq C \|d\|_{L^\infty(\text{BMO})} |\alpha| \|f_t^p\|_{L^2} \|\nabla f_t^p\|_{L^2}, \quad (2.33)$$

where C is a constant depending only on n . Therefore

$$|I_4| \leq \frac{\lambda}{2p} \|\nabla f_t^p\|_{L^2}^2 + C \|d\|_{L^\infty(\text{BMO})}^2 \frac{p}{\lambda} |\alpha|^2 \|f_t\|_{L^{2p}}^{2p}. \quad (2.34)$$

Putting these estimates on I_1 to I_4 together, we thus obtain (2.29). $\square$

Now we can follow arguments in Stroock [79] 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 in Stroock [79].

Lemma 2.11. *There is a constant $C > 0$ depending only on n and the $L_t^\infty(\text{BMO}_x)$ -norm of the skew-symmetric part of $(A_{ij}(t, x))$ such that*

$$\|\Gamma_{\tau, t}^\psi f\|_{L^\infty} \leq \frac{C}{(t - \tau)^{n/4}} e^{\frac{C|\alpha|^2(t-\tau)}{\lambda}} \|f\|_{L^2}, \quad (2.35)$$

and

$$\|\Gamma_{\tau, t}^{\psi^\perp} f\|_{L^\infty} \leq \frac{C}{(t - \tau)^{n/4}} e^{\frac{C|\alpha|^2(t-\tau)}{\lambda}} \|f\|_{L^2} \quad (2.36)$$

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 (2.35) for the case that $0 = \tau < t$. The proof of (2.36) is similar, which uses the inequality (2.30) instead and the fact that the constant appears in that inequality is independent of $T > 0$.

To show (2.35), we apply Nash's inequality (2.25) to the first term on the right-hand side of (2.29) to deduce that

$$\frac{d}{dt} \|f_t\|_{L^{2p}} \leq -\frac{\lambda}{2pC_n} \frac{\|f_t\|_{L^{2p}}^{1+4p/n}}{\|f_t\|_{L^p}^{4p/n}} + C_B \frac{|\alpha|^2}{\lambda} p \|f_t\|_{L^{2p}} \quad (2.37)$$

for every $p > 1$. Let $u_p(t) = \|f_t\|_{L^{2p}}$ and $w_p(t) = \sup_{0 \leq s \leq t} s^{n(p-2)/4p} u_{p/2}(s)$. Then (2.37)

can be written as

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

so that, according to Lemma 2.9, we have

$$\begin{aligned} w_{2p}(t) &= \sup_{0 \leq s \leq t} s^{n(p-1)/4p} u_p(s) \\ &\leq \sup_{0 \leq s \leq t} \left(\frac{Kp^2}{\delta} \right)^{\frac{n}{4p}} \exp \left(\frac{C|\alpha|^2 \delta s}{p\lambda} \right) w_p(s) \\ &= \left(\frac{Kp^2}{\delta} \right)^{\frac{n}{4p}} \exp \left(\frac{C|\alpha|^2 \delta t}{p\lambda} \right) w_p(t). \end{aligned}$$

According to (2.37), if take $p = 1$, we have

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

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

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

Taking $m \rightarrow \infty$, we therefore obtain that

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

which completes the proof. □

Proof of the upper bound (2.24). Let us use the same notations as in the proof of Lemma (2.11). By (2.36) and the fact that $\Gamma_{\tau,t}^{\Psi\perp}$ is the adjoint operator of $\Gamma_{\tau,t}^\Psi$, we have

$$\|f_t\|_{L^2} \leq \frac{C}{t^{n/4}} \exp \left(\frac{C|\alpha|^2 t}{\lambda} \right) \|f\|_{L^1}.$$

Since $\Gamma_{0,2t}^\Psi = \Gamma_{t,2t}^\Psi \circ \Gamma_{0,t}^\Psi$, we thus deduce that

$$\|f_{2t}\|_{L^\infty} \leq \frac{C}{t^{n/2}} \exp\left(\frac{C|\alpha|^2 t}{\lambda}\right) \|f\|_{L^1},$$

which is equivalent to

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

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

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

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

2.2.2 The lower bound

In this section, we prove the lower bound in Theorem 1.3:

$$\Gamma(t, x; \tau, \xi) \geq \frac{1}{C(t - \tau)^{\frac{n}{2}}} \exp\left[-\frac{C|x - \xi|^2}{t - \tau}\right] \quad (2.38)$$

following the idea due to Nash [64], where C depends only on n , λ and $\|d\|_{L^\infty(\text{BMO})}$.

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

Lemma 2.12. *There is a constant $C_0 > 0$ depending only on the dimension n , $\lambda > 0$ and the $L^\infty(\text{BMO})$ norm of (d_{ij}) , such that*

$$\int_{\mathbb{R}^n} \ln(\Gamma(1, x; 0, \xi)) \mu(d\xi) \geq -C_0 \quad \forall x \in B(0, 2), \quad (2.39)$$

and

$$\Gamma(2, x; 0, \xi) \geq e^{-2C_0} \quad x, \xi \in B(0, 2), \quad (2.40)$$

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

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

Proof. The proof follows the same ideas as in Nash [64], as explained in Stroock [64]. We have to overcome difficulties from the additional non-symmetric part $d(t, x) = (d_{ij}(t, x))$. The idea is to consider for any $x \in B(0, 2)$ the following function

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

where $t \in (0, 1]$ and $x \in B(0, 2)$. Since (A_{ij}) is uniformly elliptic with bounded derivatives, $\Gamma(t, x; \tau, \xi)$ is a probability density in x (when other variables are fixed) and also in ξ (as other variables are fixed). Hence $\int_{\mathbb{R}^n} \Gamma(t, x; 0, \xi) d\xi = 1$ for every $t \in (0, 1]$. We have $G(t) \leq 0$ according to Jensen's inequality. 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

$$\begin{aligned} G'(t) &= \int_{\mathbb{R}^n} \langle \xi, a(1-t, \xi) \cdot \nabla_{\xi} \ln \Gamma(1, x; 1-t, \xi) \rangle \mu(d\xi) \\ &\quad + \int_{\mathbb{R}^n} \langle \nabla_{\xi} \ln \Gamma(1, x; 1-t, \xi), a(1-t, \xi) \cdot \nabla_{\xi} \ln \Gamma(1, x; 1-t, \xi) \rangle \mu(d\xi) \quad (2.41) \\ &\quad + \frac{1}{\delta} \int_{\mathbb{R}^n} \left\langle \nabla_{\xi} \mu(\xi)^{\delta}, d(1-t, \xi) \cdot \nabla_{\xi} \left(\mu(\xi)^{1-\delta} \ln \Gamma(1, x; 1-t, \xi) \right) \right\rangle d\xi \end{aligned}$$

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

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

as b is skew-symmetric. Using the Cauchy-Schwartz inequality and the compensated

compactness inequality (2.5) we deduce that

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

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

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

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

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

and

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

which imply that

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

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

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

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

$$G'(t) \geq -C + \frac{\lambda}{3} \|\nabla \ln \Gamma(1, x; 1-t, \cdot)\|_{L^2(\mu)}^2 - C \|\ln \Gamma(1, x; 1-t, \cdot)\|_{L^2(\mu)} \quad (2.42)$$

for all $t \in (0, 1)$, for some constant $C > 0$ depending only on n and the L^∞ (BMO) norm of the skew-symmetric part $d(t, x)$.

The remaining arguments of the proof are more or less the same as in Stroock [79]. Firstly, by the Poincaré-Wirtinger inequality for the Gaussian measure, we obtain

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

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

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

$$= \frac{1}{2} G(t)^2 \mu \{ \xi \in \mathbb{R}^n : \ln \Gamma(1, x; 1-t, \xi) \geq -K \} - K^2.$$

According to the upper bound

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

we have

$$\ln \Gamma(1, x; 1-t, \xi) \leq \ln C - \frac{n}{2} \ln t - \frac{|x-\xi|^2}{Ct}$$

for $x \in B(0, 2)$ and $t \in (\frac{1}{2}, 1)$. Hence

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

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

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

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

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

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

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

Using this estimate, we deduce that

$$\begin{aligned}
G'(t) &\geq -C + \frac{\lambda}{6}J(t, x) - C \|\ln \Gamma(1, x; 1-t, \cdot)\|_{L^2(\mu)} \\
&\geq -C + \frac{\lambda}{6}J(t, x) - C \left[\sqrt{J(t, x)} + |G(t)| \right] \\
&\geq -C(\varepsilon, \lambda) + \frac{\lambda}{12}J(t, x) - \varepsilon G(t)^2 \\
&\geq -C(\varepsilon, \lambda) + \frac{\lambda}{24}G(t)^2 \mu \{ \xi \in \mathbb{R}^n : \ln \Gamma(1, x; 1-t, \xi) \geq -K \} - \frac{\lambda}{12}K^2 - \varepsilon G(t)^2 \\
&\geq -C(\varepsilon, K, \lambda) + \left(\frac{\lambda}{12}\kappa(r) - \varepsilon \right) G(t)^2
\end{aligned}$$

for $\varepsilon > 0$ such that $\frac{\lambda}{12}\kappa(r) - \varepsilon > 0$. Now we obtain

$$G'(t) \geq -C_1 + C_2 G(t)^2 \quad (2.43)$$

for any $t \in [\frac{1}{2}, 1]$, where $C_1 > 0$, $C_2 \in (0, 1]$. The previous inequality (2.43) 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

$$G(1) \geq \min \left\{ -C_1 - 2\sqrt{\frac{C_1}{C_2}}, -\frac{8}{3C_2} \right\} = -C_0. \quad (2.44)$$

The lower bound in (2.39) follows from the Chapman-Kolmogorov equation and Jensen's inequality. In fact

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

which yields (2.40). □

Proof of the lower bound (2.38). By the scaling invariant properties, i.e. for any $r > 0$ and $z \in \mathbb{R}^n$,

$$\Gamma(r^2t, rx + z; 0, r\xi + z) = r^{-n} \Gamma^{A_{r,z}}(t, x; 0, \xi) \quad (2.45)$$

where $A_{r,z}(t, x) = A(r^2t, rx + z)$ and Γ^A is the fundamental solution associated with A in equation (1.14). The transformation $A \rightarrow A^{r,z}$ preserves the elliptic constant λ and more importantly the $L^\infty(\text{BMO})$ norms. So we may apply (2.40) to $\Gamma^{A_{r,z}}$ to deduce that

$$\Gamma(2t, x; 0, \xi) \geq \frac{e^{-2A}}{t^{n/2}}, \quad \text{for } |\xi - x| < 4t^{\frac{1}{2}}. \quad (2.46)$$

Finally, to extend it to all $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$, we use the Chapman-Kolmogorov equation again. Suppose that $k \leq |x - \xi|^2 < k + 1$, set $x_m = \xi + \frac{m}{k+1}(x - \xi)$ and $B_m = B(x_m, \frac{1}{k^{1/2}})$ for $0 \leq m \leq k + 1$. Notice that if $y_m \in B_m$, then $|y_m - y_{m+1}| < \frac{3}{k^{1/2}}$ and therefore

$$\Gamma\left(\frac{2m}{k+1}, y_m; \frac{2(m-1)}{k+1}, y_{m-1}\right) = \Gamma^{A_m}\left(\frac{2}{k+1}, y_m; 0, y_{m-1}\right) \geq k^{n/2} e^{-2A},$$

where $A_m(t, x) = A\left(t + \frac{2(m-1)}{k+1}, x\right)$. Hence by the Chapman-Kolmogorov equation

$$\begin{aligned} & \Gamma(2, x; 0, \xi) \\ & \geq \int_{B_1} \cdots \int_{B_k} \Gamma\left(\frac{2}{k+1}, y_1; 0, \xi\right) \Gamma\left(\frac{4}{k+1}, y_2; \frac{2}{k+1}, y_1\right) \cdots \Gamma\left(2, x; \frac{2k}{k+1}, y_k\right) dy_k \cdots dy_1 \\ & \geq (k^{n/2} e^{-2A})^{k+1} (|B(0, 1)| k^{-\frac{n}{2}})^k. \end{aligned}$$

Choose β such that $e^{-\beta} \leq |B(0, 1)| e^{-2A}$, then we have $\Gamma(2, x; 0, \xi) \geq e^{-2A} e^{-\beta|x-\xi|^2}$.

Again apply the scaling argument to Γ , we obtain the lower bound:

$$\frac{1}{Ct^{\frac{n}{2}}} \exp\left(-\frac{C|x-\xi|^2}{t}\right) \leq \Gamma(t, x; 0, \xi).$$

□

2.3 Supercritical conditions on the drift

In this section, we will prove Theorem 1.6 and 1.8, which are *a priori* estimates of the fundamental solution to equation (1.9). We will assume that $a \in C^\infty([0, T] \times \mathbb{R}^n)$ and $b \in C^\infty([0, T], C_0^\infty(\mathbb{R}^n))$ so that there exists a unique regular fundamental solution.

2.3.1 The upper bound

In this section, we prove the upper bound, which essentially use the h -transform of the fundamental solution. But here we will use Moser's approach instead of Nash's to prove the upper bound because it has the potential of applicability to more general cases where

$$b \in L^l(0, T; L^q(\mathbb{R}^n)), \quad 1 \leq \frac{2}{l} + \frac{n}{q} < 2 \quad \mathbf{A}$$

for $n \geq 3$, $l > 1$ and $q > \frac{n}{2}$. Recall that we denote

$$\Lambda = \|b\|_{L_t^l L_x^q} = \left(\int_0^T \left(\int_{\mathbb{R}^n} |b(t, x)|^q dx \right)^{\frac{l}{q}} dt \right)^{\frac{1}{l}}.$$

Same as in the divergence form (1.14), we apply the h -transform to equation (1.9). Given a function ψ on $\mathbb{R}^n$ which is smooth and has bounded derivatives, we define the operator

$$\begin{aligned} A_t^\psi u(x) &= \exp(-\psi(x)) \sum_{i,j=1}^n \partial_{x_i}(a_{ij}(t, x) \partial_{x_j} [\exp(\psi(x)) u(x)]) \\ &\quad - \exp(-\psi(x)) \sum_{i=1}^n b_i(t, x) \partial_{x_i} [\exp(\psi(x)) u(x)]. \end{aligned}$$

Then its corresponding fundamental solution is

$$\Gamma^\Psi(t, x; \tau, \xi) = \exp(-\psi(x)) \Gamma(t, x; \tau, \xi) \exp(\psi(\xi)),$$

and the adjoint operator is

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

as defined in the critical case. Clearly they satisfy

$$\langle \Gamma_{\tau, t}^{\Psi} f, g \rangle_{L^2(\mathbb{R}^n)} = \langle f, \Gamma_{\tau, t}^{\Psi \perp} g \rangle_{L^2(\mathbb{R}^n)}. \quad (2.47)$$

Lemma 2.13. *Suppose (a, b) satisfies conditions **(E)**, **(S)** and **(A)**. Given $\alpha \in \mathbb{R}^n$, and $\psi(x) = \alpha \cdot x$, set*

$$f_t(x) = \Gamma_{0, t}^{\Psi} f(x) = \int_{\mathbb{R}^n} f(\xi) \Gamma^\Psi(t, x; 0, \xi) d\xi$$

for $f \in C_0^\infty(\mathbb{R}^n)$. Then there exists a constant C depending on (n, l, q) such that

$$\|f_t\|_{L_x^2}^2 \leq \exp\left(\frac{2|\alpha|^2}{\lambda} t + 2C\lambda^{-\frac{1+\theta}{1-\theta}} |\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu t^\nu\right) \cdot \|f\|_{L_x^2}^2, \quad (2.48)$$

where $\theta = \frac{n}{q} - 1$, $\mu = \frac{2}{2-\gamma+\frac{2}{l}}$, $\nu = \frac{2-\gamma}{2-\gamma+\frac{2}{l}}$, $\gamma = \frac{2}{l} + \frac{n}{q}$ and $\Lambda = \|b\|_{L^l(0, T; L^q(\mathbb{R}^n))}$.

Proof. We begin with the fact that f_t satisfies

$$\frac{d}{dt} \|f_t\|_{L_x^2}^2 = 2 \langle A_t^\Psi f_t, f_t \rangle_{L^2(\mathbb{R}^n)}.$$

It follows that

$$\begin{aligned} \frac{1}{2} \left(\|f_t\|_{L_x^2}^2 - \|f\|_{L_x^2}^2 \right) &= \frac{1}{2} \int_0^t \frac{d}{ds} \|f_s\|_{L_x^2}^2 ds = \int_0^t \int_{\mathbb{R}^n} A_s^\Psi f_s(x) \cdot f_s(x) dx ds \\ &= - \int_0^t \int_{\mathbb{R}^n} \sum_{i, j=1}^n a_{ij}(s, x) \partial_{x_j} [\exp(\psi(x)) f_s(x)] \partial_{x_i} [\exp(-\psi(x)) f_s(x)] dx ds \end{aligned}$$

$$\begin{aligned}
& - \int_0^t \int_{\mathbb{R}^n} \sum_{i=1}^n b_i(s, x) \partial_{x_i} [\exp(\psi(x)) f_s(x)] [\exp(-\psi(x)) f_s(x)] dx ds \\
& = \int_0^t \int_{\mathbb{R}^n} \langle \alpha \cdot a(s, x), \alpha \rangle f_s^2(x) dx ds - \int_0^t \int_{\mathbb{R}^n} \langle \nabla f_s(x) \cdot a(s, x), \nabla f_s(x) \rangle dx ds \\
& \quad - \int_0^t \int_{\mathbb{R}^n} \langle \alpha \cdot a(s, x), \nabla f_s(x) \rangle f_s(x) dx ds + \int_0^t \int_{\mathbb{R}^n} \langle \nabla f_s(x) \cdot a(s, x), \alpha \rangle f_s(x) dx ds \\
& \quad - \int_0^t \int_{\mathbb{R}^n} \langle b(s, x), \alpha \rangle f_s^2(x) dx ds - \int_0^t \int_{\mathbb{R}^n} \langle b(s, x), \nabla f_s(x) \rangle f_s(x) dx ds.
\end{aligned}$$

Since b is divergence-free, we have for any s that

$$\int_{\mathbb{R}^n} \langle b(s, x), \nabla f_s(x) \rangle f_s(x) dx = 0.$$

The third and fourth terms cancel each other and condition **(E)** gives

$$\begin{aligned}
& \frac{1}{2} \left(\|f_t\|_{L_x^2}^2 - \|f\|_{L_x^2}^2 \right) \\
& = \int_0^t \int_{\mathbb{R}^n} \langle \alpha \cdot a(s, x), \alpha \rangle f_s^2(x) dx ds - \int_0^t \int_{\mathbb{R}^n} \langle \nabla f_s(x) \cdot a(s, x), \nabla f_s(x) \rangle dx ds \\
& \quad - \int_0^t \int_{\mathbb{R}^n} \langle b(s, x), \alpha \rangle f_s^2(x) dx ds \\
& \leq \int_0^t \frac{|\alpha|^2}{\lambda} \|f_s\|_{L_x^2}^2 ds - \int_0^t \lambda \|\nabla f_s\|_{L_x^2}^2 ds - \int_0^t \int_{\mathbb{R}^n} \langle b(s, x), \alpha \rangle f_s^2(x) dx ds.
\end{aligned}$$

For the last term, one obtains the following estimate

$$\begin{aligned}
\left| \int_0^t \int_{\mathbb{R}^n} \langle b(s, x), \alpha \rangle f_s^2(x) dx ds \right| & \leq \int_0^t |\alpha| \|b(s, \cdot)\|_{L_x^q} \|f_s^{1+\theta}\|_{L_x^{r_1}} \|f_s^{1-\theta}\|_{L_x^{r_2}} ds \\
& = \int_0^t |\alpha| \|b(s, \cdot)\|_{L_x^q} \|f_s\|_{L_x^{(1+\theta)r_1}}^{1+\theta} \|f_s\|_{L_x^{(1-\theta)r_2}}^{1-\theta} ds,
\end{aligned}$$

where

$$\theta = \frac{n}{q} - 1, \quad (1 + \theta)r_1 = \frac{2n}{n-2}, \quad (1 - \theta)r_2 = 2.$$

By Sobolev's embedding and Young's inequality, we can further control it as follows

$$\begin{aligned}
& \left| \int_0^t \int_{\mathbb{R}^n} \langle b(s, x), \alpha \rangle f_s^2(x) dx ds \right| \\
& \leq \int_0^t C |\alpha| \|b(s, \cdot)\|_{L_x^q} \|f_s\|_{L_x^2}^{1-\theta} \|\nabla f_s\|_{L_x^2}^{1+\theta} ds \\
& = \int_0^t \left(\left(\frac{2}{\lambda} \right)^{\frac{1+\theta}{2}} C |\alpha| \|b(s, \cdot)\|_{L_x^q} \|f_s\|_{L_x^2}^{1-\theta} \right) \left(\left(\frac{\lambda}{2} \right)^{\frac{1+\theta}{2}} \|\nabla f_s\|_{L_x^2}^{1+\theta} \right) ds \\
& \leq \int_0^t \frac{1-\theta}{2} \left(\frac{2}{\lambda} \right)^{\frac{1+\theta}{1-\theta}} (C |\alpha| \|b(s, \cdot)\|_{L_x^q})^{\frac{2}{1-\theta}} \|f_s\|_{L_x^2}^2 + \frac{1+\theta}{2} \frac{\lambda}{2} \|\nabla f_s\|_{L_x^2}^2 ds \\
& \leq \int_0^t C \left(\frac{1}{\lambda} \right)^{\frac{1+\theta}{1-\theta}} (|\alpha| \|b(s, \cdot)\|_{L_x^q})^{\frac{2}{1-\theta}} \|f_s\|_{L_x^2}^2 + \frac{\lambda}{2} \|\nabla f_s\|_{L_x^2}^2 ds.
\end{aligned}$$

Combining all the estimates above, one has

$$\begin{aligned}
\|f_t\|_{L_x^2}^2 & \leq \|f\|_{L_x^2}^2 + 2 \int_0^t \frac{|\alpha|^2}{\lambda} \|f_s\|_{L_x^2}^2 - \lambda \|\nabla f_s\|_{L_x^2}^2 ds \\
& \quad + \int_0^t C \left(\frac{1}{\lambda} \right)^{\frac{1+\theta}{1-\theta}} (|\alpha| \|b(s, \cdot)\|_{L_x^q})^{\frac{2}{1-\theta}} \|f_s\|_{L_x^2}^2 + \frac{\lambda}{2} \|\nabla f_s\|_{L_x^2}^2 ds \\
& \leq \|f\|_{L_x^2}^2 + 2 \int_0^t \left(\frac{|\alpha|^2}{\lambda} + C \left(\frac{1}{\lambda} \right)^{\frac{1+\theta}{1-\theta}} (|\alpha| \|b(s, \cdot)\|_{L_x^q})^{\frac{2}{1-\theta}} \right) \|f_s\|_{L_x^2}^2 ds.
\end{aligned}$$

Recall $\frac{2}{l} + \frac{n}{q} = \gamma$ with $1 \leq \gamma < 2$ and $\Lambda = \|b\|_{L^l([0, T], L^q(\mathbb{R}^n))}$. Hölder's inequality implies that

$$\begin{aligned}
\int_0^t \|b(s, \cdot)\|_{L_x^q}^{\frac{2}{1-\theta}} ds & = \int_0^t \|b(s, \cdot)\|_{L_x^q}^{\frac{2}{(2-\gamma+\frac{2}{l})}} ds \leq \left(\int_0^t \|b(s, \cdot)\|_{L_x^q}^l ds \right)^{\frac{\frac{2}{1-\theta}}{2-\gamma+\frac{2}{l}}} t^{\frac{2-\gamma}{2-\gamma+\frac{2}{l}}} \\
& = \Lambda^{\frac{2}{2-\gamma+\frac{2}{l}}} t^{\frac{2-\gamma}{2-\gamma+\frac{2}{l}}},
\end{aligned}$$

where we set $\frac{c}{l} = 0$ if $l = \infty$. For simplicity, we denote $\mu = \frac{2}{2-\gamma+\frac{2}{l}}$ and $\nu = \frac{2-\gamma}{2-\gamma+\frac{2}{l}}$. Hence, by Grönwall's inequality and

$$\|f_t\|_{L_x^2}^2 \leq \|f\|_{L_x^2}^2 + 2 \int_0^t \left(\frac{|\alpha|^2}{\lambda} + C \left(\frac{1}{\lambda} \right)^{\frac{1+\theta}{1-\theta}} (|\alpha| \|b(s, \cdot)\|_{L_x^q})^{\frac{2}{1-\theta}} \right) \|f_s\|_{L_x^2}^2 ds,$$

we deduce that

$$\begin{aligned}\|f_t\|_{L_x^2}^2 &\leq \exp\left(2\int_0^t\left(\frac{|\alpha|^2}{\lambda}+C\left(\frac{1}{\lambda}\right)^{\frac{1+\theta}{1-\theta}}(|\alpha|\|b(s,\cdot)\|_{L_x^q})^{\frac{2}{1-\theta}}\right)ds\right)\|f\|_{L_x^2}^2 \\ &\leq \exp\left(\frac{2|\alpha|^2}{\lambda}t+2C\left(\frac{1}{\lambda}\right)^{\frac{1+\theta}{1-\theta}}|\alpha|^{\frac{2}{1-\theta}}\Lambda^\mu t^\nu\right)\|f\|_{L_x^2}^2.\end{aligned}$$

Now the proof is complete. $\square$

Lemma 2.14. *Suppose that (a, b) , ψ and f_t are defined as in Lemma 2.13. For any $p \geq 1$ and any smooth non-negative function η on $[0, T]$ satisfying $\eta(0) = 0$, we have*

$$\begin{aligned}\|f_t^p \eta^\sigma\|_{L_t^{2\chi} L_x^{2\chi}}^2 &\leq C|\alpha|^2 p^2 \|f_t^p \eta^\sigma\|_{L_t^2 L_x^2}^2 + C(|\alpha|p)^{\frac{2}{2-\gamma}} \|b\|_{L_t^l L_x^q}^{\frac{2}{2-\gamma}} \|f_t^p \eta^{\frac{1}{2-\gamma}}\|_{L_t^2 L_x^2}^2 \\ &\quad + C \int_0^T \int_{\mathbb{R}^n} \sigma f_t^{2p}(x) |\partial_t \eta(t)| \eta^{2\sigma-1}(t) dx dt.\end{aligned}$$

where $\chi = \frac{n+2}{n}$, $\sigma = \frac{1}{2-\gamma}$ and $C > 0$ is a constant depending only on l, q, n, λ .

Proof. For any $p \geq 1$, we have

$$\frac{d}{dt} \|f_t\|_{L_x^{2p}}^{2p} = 2p \langle A_t^\psi f_t, f_t^{2p-1} \rangle_{L^2(\mathbb{R}^n)}.$$

Next we multiply both sides by $\eta^{2\sigma}$ and integrate on $[0, T]$ to obtain

$$\begin{aligned}&\int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \partial_t f_t(x) f_t(x)^{2p-1} dx dt \\ &= - \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle \nabla(\exp(\psi(x)) f_t(x)) \cdot a(t, x), \nabla(\exp(-\psi(x)) f_t^{2p-1}(x)) \rangle dx dt \\ &\quad - \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle b(t, x), \nabla(\exp(\psi(x)) f_t(x)) \rangle \exp(-\psi(x)) f_t^{2p-1}(x) dx dt \\ &= \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle \alpha \cdot a(t, x), \alpha \rangle f_t^{2p}(x) dx dt \\ &\quad - (2p-1) \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle \nabla f_t(x) \cdot a(t, x), \nabla f_t(x) \rangle f_t^{2p-2}(x) dx dt \\ &\quad - (2p-1) \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle \alpha \cdot a(t, x), \nabla f_t(x) \rangle f_t^{2p-1}(x) dx dt \\ &\quad + \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle \nabla f_t(x) \cdot a(t, x), \alpha \rangle f_t^{2p-1}(x) dx dt\end{aligned}$$

$$\begin{aligned}
& - \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle b(t, x), \alpha \rangle f_t^{2p}(x) dx dt \\
& - \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle b(t, x), \nabla f_t(x) \rangle f_t^{2p-1}(x) dx dt.
\end{aligned} \tag{2.49}$$

Condition **(S)** implies that

$$\int_{\mathbb{R}^n} \langle b(t, x), \nabla f_t(x) \rangle f_t^{2p-1}(x) dx = 0$$

for any t , and hence the last term vanishes. Set $g_t = f_t^p$ for simplicity, then the left-hand side becomes

$$\begin{aligned}
& \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \partial_t f_t(x) f_t(x)^{2p-1} dx dt = \int_0^T \int_{\mathbb{R}^n} \eta^{2\sigma}(t) \frac{1}{2p} \partial_t (g_t^2(x)) dx dt \\
& = \int_{\mathbb{R}^n} \frac{1}{2p} \eta^{2\sigma}(t) g_t^2(x) dx \Big|_0^T - \int_0^T \int_{\mathbb{R}^n} \frac{\sigma}{p} g_t^2(x) (\partial_t \eta(t)) \eta^{2\sigma-1}(t) dx dt.
\end{aligned}$$

Multiplying by p on both sides of equation (2.49), we obtain

$$\begin{aligned}
& \int_{\mathbb{R}^n} \frac{1}{2} \eta^{2\sigma}(t) g_t^2(x) dx \Big|_0^T - \int_0^T \int_{\mathbb{R}^n} \sigma g_t^2(x) (\partial_t \eta(t)) \eta^{2\sigma-1}(t) dx dt \\
& = p \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle \alpha \cdot a(t, x), \alpha \rangle g_t^2(x) dx dt \\
& \quad - \frac{(2p-1)}{p} \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle \nabla g_t(x) \cdot a(t, x), \nabla g_t(x) \rangle dx dt \\
& \quad - (2p-2) \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle \alpha \cdot a(t, x), \nabla g_t(x) \rangle g_t(x) dx dt \\
& \quad - p \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle b(t, x), \alpha \rangle g_t^2(x) dx dt \\
& = I_1 - I_2 - I_3 - I_4.
\end{aligned}$$

Now we estimate each term individually as follows

$$I_1 \leq \int_0^T \eta^{2\sigma}(t) \frac{|\alpha|^2}{\lambda} p \|g_t\|_{L_x^2}^2 dt,$$

$$\begin{aligned}
-I_2 - I_3 &\leq \int_0^T \eta^{2\sigma}(t) \frac{|\alpha|^2}{\lambda} (p-1)p \|g_t\|_{L_x^2}^2 dt - \int_0^T \eta^{2\sigma}(t) \lambda \|\nabla g_t\|_{L_x^2}^2 dt, \\
|I_4| &= \left| p \int_0^T \eta^{2\sigma}(t) \int_{\mathbb{R}^n} \langle b(t, x), \alpha \rangle g_t^2(x) dx dt \right| \\
&\leq \int_0^T \int_{\mathbb{R}^n} p |b(t, x)| |g_t \eta^\sigma|^\gamma |g_t|^{2-\gamma} (|\alpha| \eta) dx dt \\
&\leq |\alpha| p \|b\|_{L_t^l L_x^q} \|g_t \eta^\sigma\|_{L_t^s L_x^r}^\gamma \|g_t \eta^{\frac{1}{2-\gamma}}\|_{L_t^2 L_x^2}^{2-\gamma}
\end{aligned}$$

since $\sigma\gamma = 2\sigma - 1$ and

$$\frac{1}{l} + \frac{\gamma}{s} + \frac{2-\gamma}{2} = 1, \quad \frac{1}{q} + \frac{\gamma}{r} + \frac{2-\gamma}{2} = 1.$$

From this relation, it is easy to see

$$\frac{2}{s} + \frac{n}{r} = \frac{n}{2},$$

which yields the interpolation inequality

$$\|f\|_{L_t^s L_x^r} \leq C \|f\|_{L_t^\infty L_x^2}^{1-\beta} \|\nabla f\|_{L_t^2 L_x^2}^\beta, \quad \beta = \frac{n}{2} - \frac{n}{r}.$$

Together with Young's inequality, we deduce the following estimate

$$\|f\|_{L_t^s L_x^r} \leq C_1 \|f\|_{L_t^\infty L_x^2} + C_2 \|\nabla f\|_{L_t^2 L_x^2}. \quad (2.50)$$

Now we choose $\varepsilon > 0$ small enough such that, by Young's inequality, we have

$$\begin{aligned}
|I_4| &\leq \varepsilon \|g_t \eta^\sigma\|_{L_t^s L_x^r}^2 + C(\varepsilon) (|\alpha| p)^{\frac{2}{2-\gamma}} \|b\|_{L_t^l L_x^q}^{\frac{2}{2-\gamma}} \|g_t \eta^{\frac{1}{2-\gamma}}\|_{L_t^2 L_x^2}^2 \\
&\leq \frac{\lambda \wedge 1}{4} (\|g_t \eta^\sigma\|_{L_t^\infty L_x^2}^2 + \|\nabla g_t \eta^\sigma\|_{L_t^2 L_x^2}^2) + C (|\alpha| p)^{\frac{2}{2-\gamma}} \|b\|_{L_t^l L_x^q}^{\frac{2}{2-\gamma}} \|g_t \eta^{\frac{1}{2-\gamma}}\|_{L_t^2 L_x^2}^2.
\end{aligned}$$

Combining these together, we conclude that

$$\int_{\mathbb{R}^n} \frac{1}{2} \eta^{2\sigma}(t) g_t^2(x) dx \Big|_0^T - \int_0^T \int_{\mathbb{R}^n} \sigma g_t^2(x) (\partial_t \eta(t)) \eta^{2\sigma-1}(t) dx dt$$

$$\begin{aligned}
&\leq \int_0^T \eta^{2\sigma}(t) \frac{|\alpha|^2}{\lambda} p^2 \|g_t\|_{L_x^2}^2 dt - \int_0^T \eta^{2\sigma}(t) \lambda \|\nabla g_t\|_{L_x^2}^2 dt \\
&\quad + \frac{\lambda \wedge 1}{4} (\|g_t \eta^\sigma\|_{L_t^\infty L_x^2}^2 + \|\nabla g_t \eta^\sigma\|_{L_t^2 L_x^2}^2) + C(|\alpha|p)^{\frac{2}{2-\gamma}} \|b\|_{L_t^q L_x^q}^{\frac{2}{2-\gamma}} \|g_t \eta^{\frac{1}{2-\gamma}}\|_{L_t^2 L_x^2}^2.
\end{aligned}$$

If we set $\eta(0) = 0$, then the inequality above implies that

$$\begin{aligned}
&\frac{1}{2} \|g_T \eta^\sigma(T)\|_{L_x^2}^2 + \frac{\lambda}{2} \|\nabla g_T \eta^\sigma\|_{L_t^2 L_x^2}^2 \\
&\leq \frac{|\alpha|^2 p^2}{\lambda} \|g_t \eta^\sigma\|_{L_t^2 L_x^2}^2 + \frac{1}{4} \|g_t \eta^\sigma\|_{L_t^\infty L_x^2}^2 + C(|\alpha|p)^{\frac{2}{2-\gamma}} \|b\|_{L_t^q L_x^q}^{\frac{2}{2-\gamma}} \|g_t \eta^{\frac{1}{2-\gamma}}\|_{L_t^2 L_x^2}^2 \\
&\quad + \int_0^T \int_{\mathbb{R}^n} \sigma g_t^2(x) |\partial_t \eta(t)| \eta^{2\sigma-1}(t) dx dt,
\end{aligned}$$

and the same is true if we replace T by any $t \in [0, T]$. Hence

$$\begin{aligned}
&\frac{1}{4} \|g_t \eta^\sigma\|_{L_t^\infty L_x^2}^2 + \frac{\lambda}{2} \|\nabla g_t \eta^\sigma\|_{L_t^2 L_x^2}^2 \\
&\leq \frac{|\alpha|^2 p^2}{\lambda} \|g_t \eta^\sigma\|_{L_t^2 L_x^2}^2 + C(|\alpha|p)^{\frac{2}{2-\gamma}} \|b\|_{L_t^q L_x^q}^{\frac{2}{2-\gamma}} \|g_t \eta^{\frac{1}{2-\gamma}}\|_{L_t^2 L_x^2}^2 \\
&\quad + \int_0^T \int_{\mathbb{R}^n} \sigma g_t^2(x) |\partial_t \eta(t)| \eta^{2\sigma-1}(t) dx dt.
\end{aligned}$$

Applying the interpolation inequality (2.50) with $s = r = \chi = \frac{n+2}{n}$, we deduce that

$$\begin{aligned}
\|g_t \eta^\sigma\|_{L_t^{2\chi} L_x^{2\chi}}^2 &\leq C|\alpha|^2 p^2 \|g_t \eta^\sigma\|_{L_t^2 L_x^2}^2 + C(|\alpha|p)^{\frac{2}{2-\gamma}} \|b\|_{L_t^q L_x^q}^{\frac{2}{2-\gamma}} \|g_t \eta^{\frac{1}{2-\gamma}}\|_{L_t^2 L_x^2}^2 \\
&\quad + C \int_0^T \int_{\mathbb{R}^n} \sigma g_t^2(x) |\partial_t \eta(t)| \eta^{2\sigma-1}(t) dx dt,
\end{aligned}$$

and the proof is complete. $\square$

Now we can use Moser's iteration to prove Theorem 1.6.

Proof of Theorem 1.6. Define open intervals $I_k = ((\frac{1}{2} - \frac{1}{2^{k+1}})T, T)$ and choose η_k as cut-off functions such that $\eta_k = 1$ on I_k , $\eta_k = 0$ on $I_0 \setminus I_{k-1}$ and $|\partial_t \eta_k| \leq 4^k T^{-1}$. Denote $L_{I_k \times \mathbb{R}^n}^p$ the L^p space on the space-time domain $I_k \times \mathbb{R}^n$. Then

$$\|g_t\|_{L_{I_k \times \mathbb{R}^n}^{2\chi}}^2 \leq \|g_t \eta_k^\sigma\|_{L_t^{2\chi} L_x^{2\chi}}^2$$

$$\begin{aligned}
&\leq C|\alpha|^2 p^2 \|g_t \eta_k^\sigma\|_{L_t^2 L_x^2}^2 + C(|\alpha|p)^{\frac{2}{2-\gamma}} \|b\|_{L_t^1 L_x^q}^{\frac{2}{2-\gamma}} \|g_t \eta_k^{\frac{1}{2-\gamma}}\|_{L_t^2 L_x^2}^2 \\
&\quad + C \int_0^T \int_{\mathbb{R}^n} \sigma g_t^2(x) |\partial_t \eta_k(t)| \eta_k^{2\sigma-1}(t) dx dt \\
&\leq C|\alpha|^2 p^2 \|g_t\|_{L_{I_{k-1} \times \mathbb{R}^n}^2}^2 + C(|\alpha|p)^{\frac{2}{2-\gamma}} \|b\|_{L_t^1 L_x^q}^{\frac{2}{2-\gamma}} \|g_t\|_{L_{I_{k-1} \times \mathbb{R}^n}^2}^2 + C\sigma \frac{4^k}{T} \|g_t\|_{L_{I_{k-1} \times \mathbb{R}^n}^2}^2 \\
&\leq C \left(|\alpha|^2 p^2 + p^{\frac{2}{2-\gamma}} \|b\|_{L_t^1 L_x^q}^{\frac{2}{2-\gamma}} |\alpha|^{\frac{2}{2-\gamma}} + \sigma \frac{4^k}{T} \right) \|g_t\|_{L_{I_{k-1} \times \mathbb{R}^n}^2}^2.
\end{aligned}$$

Recall that $g_t = f_t^p$. Let $p_0 = 1$ and $p_k = \chi^k = (\frac{n+2}{n})^k$ for $k = 1, 2, \dots$. Then

$$\|f_t^{p_{k-1}}\|_{L_{I_k \times \mathbb{R}^n}^{2\chi}}^2 \leq C \left(|\alpha|^2 p_{k-1}^2 + p_{k-1}^{\frac{2}{2-\gamma}} \|b\|_{L_t^1 L_x^q}^{\frac{2}{2-\gamma}} |\alpha|^{\frac{2}{2-\gamma}} + \sigma \frac{4^k}{T} \right) \|f_t^{p_{k-1}}\|_{L_{I_{k-1} \times \mathbb{R}^n}^2}^2,$$

or equivalently,

$$\|f_t\|_{L_{I_k \times \mathbb{R}^n}^{2p_k}} \leq C^{\frac{1}{2p_{k-1}}} \left(|\alpha|^2 p_{k-1}^2 + p_{k-1}^{\frac{2}{2-\gamma}} \|b\|_{L_t^1 L_x^q}^{\frac{2}{2-\gamma}} |\alpha|^{\frac{2}{2-\gamma}} + \sigma \frac{4^k}{T} \right)^{\frac{1}{2p_{k-1}}} \|f_t\|_{L_{I_{k-1} \times \mathbb{R}^n}^{2p_{k-1}}}.$$

Iterate the procedure above to get that

$$\|f_t\|_{L_{(\frac{T}{2}, T) \times \mathbb{R}^n}^\infty} \leq \left(\prod_{k=1}^\infty C^{\frac{1}{2p_{k-1}}} (|\alpha|^2 p_{k-1}^2 + p_{k-1}^{\frac{2}{2-\gamma}} \|b\|_{L_t^1 L_x^q}^{\frac{2}{2-\gamma}} |\alpha|^{\frac{2}{2-\gamma}} + \sigma \frac{4^k}{T})^{\frac{1}{2p_{k-1}}} \right) \|f_t\|_{L_{I_0 \times \mathbb{R}^n}^2}.$$

Since $p_k = (\frac{n+2}{n})^k \leq 2^k$ and $2 - \gamma \leq 1$, we have

$$\begin{aligned}
\|f_t\|_{L_{(\frac{T}{2}, T) \times \mathbb{R}^n}^\infty} &\leq \left(\prod_{k=1}^\infty C^{\frac{1}{2p_{k-1}}} (|\alpha|^2 p_{k-1}^2 + p_{k-1}^{\frac{2}{2-\gamma}} \Lambda^{\frac{2}{2-\gamma}} |\alpha|^{\frac{2}{2-\gamma}} + \sigma \frac{4^k}{T})^{\frac{1}{2p_{k-1}}} \right) \|f_t\|_{L_{I_0 \times \mathbb{R}^n}^2} \\
&\leq \left(\prod_{k=1}^\infty C^{\frac{1}{2p_{k-1}}} (|\alpha|^2 + \Lambda^{\frac{2}{2-\gamma}} |\alpha|^{\frac{2}{2-\gamma}} + \frac{\sigma}{T})^{\frac{1}{2p_{k-1}}} (4^{\frac{k}{2-\gamma}})^{\frac{1}{2p_{k-1}}} \right) \|f_t\|_{L_{I_0 \times \mathbb{R}^n}^2} \\
&\leq C(|\alpha|^2 + \Lambda^{\frac{2}{2-\gamma}} |\alpha|^{\frac{2}{2-\gamma}} + \frac{\sigma}{T})^{\frac{n+2}{4}} \|f_t\|_{L_{I_0 \times \mathbb{R}^n}^2} \\
&= C(|\alpha|^2 T + \Lambda^{\frac{2}{2-\gamma}} |\alpha|^{\frac{2}{2-\gamma}} T + \sigma)^{\frac{n+2}{4}} T^{-\frac{n+2}{4}} \|f_t\|_{L_{I_0 \times \mathbb{R}^n}^2}. \tag{2.51}
\end{aligned}$$

We already proved inequality (2.48), which implies

$$\|f_t\|_{L_{I_0 \times \mathbb{R}^n}^2} \leq T^{\frac{1}{2}} \exp \left(C(|\alpha|^2 T + |\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu T^\nu) \right) \|f\|_{L_x^2}.$$

Inserting this into (2.51), we derive that

$$\begin{aligned} \|f_t\|_{L^\infty_{(\frac{T}{2}, T) \times \mathbb{R}^n}} &\leq C(|\alpha|^2 T + \Lambda^{\frac{2}{2-\gamma}} |\alpha|^{\frac{2}{2-\gamma}} T + \sigma)^{\frac{n+2}{4}} T^{-\frac{n}{4}} \\ &\quad \times \exp\left(C(|\alpha|^2 T + |\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu T^\nu)\right) \|f\|_{L^2_x}. \end{aligned}$$

Notice that $1 - \theta = 2 - \frac{n}{q} = 2 - \gamma + \frac{2}{l}$ and recall that $\mu = \frac{2}{2-\gamma+\frac{2}{l}}$, $\nu = \frac{2-\gamma}{2-\gamma+\frac{2}{l}}$. Hence

$$|\alpha|^{\frac{2}{2-\gamma}} \Lambda^{\frac{2}{2-\gamma}} T = (|\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu T^\nu)^{\frac{2-\gamma+\frac{2}{l}}{2-\gamma}},$$

and $(|\alpha|^2 T + \Lambda^{\frac{2}{2-\gamma}} |\alpha|^{\frac{2}{2-\gamma}} T + \sigma)^{\frac{n+2}{4}}$ can be viewed as a polynomial of $(|\alpha|^2 T, |\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu T^\nu)$, which can be dominated by

$$C \exp\left(C(|\alpha|^2 T + |\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu T^\nu)\right).$$

Therefore we have

$$\|f_T\|_{L^\infty_x} \leq C T^{-\frac{n}{4}} \exp\left(C(|\alpha|^2 T + |\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu T^\nu)\right) \|f\|_{L^2_x}.$$

By duality, i.e. equation (2.47)

$$\|f_T\|_{L^2_x} \leq C T^{-\frac{n}{4}} \exp\left(C(|\alpha|^2 T + |\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu T^\nu)\right) \|f\|_{L^1_x}.$$

Using the Chapman-Kolmogorov equation, one has

$$\|f_{2T}\|_{L^\infty_x} \leq C T^{-\frac{n}{2}} \exp\left(C(|\alpha|^2 T + |\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu T^\nu)\right) \|f\|_{L^1_x}.$$

Recall that

$$f_{2T}(x) = \int_{\mathbb{R}^n} f(\xi) \exp(-\psi(x)) \Gamma(2T, x; 0, \xi) \exp(\psi(\xi)) d\xi$$

for any $f \in C_0^\infty(\mathbb{R}^n)$ and that $\psi(x) = \alpha \cdot x$. Replacing $2T$ by t and dividing both sides by $\exp(-\psi(x)) \exp(\psi(\xi))$, then we have a point-wise upper bound on Γ as follows

$$\Gamma(t, x; 0, \xi) \leq \frac{C}{t^{n/2}} \exp \left(C(|\alpha|^2 t + |\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu t^\nu) + \alpha \cdot (x - \xi) \right)$$

for any $\alpha \in \mathbb{R}^n$, where C depends only on (l, q, n, λ) . Set $m(t, x) = \min_{\alpha \in \mathbb{R}^n} (C(|\alpha|^2 t + |\alpha|^{\frac{2}{1-\theta}} \Lambda^\mu t^\nu) + \alpha \cdot x)$. Taking the minimum of the right-hand side over all $\alpha \in \mathbb{R}^n$, we can conclude that

$$\Gamma(t, x; 0, \xi) \leq \frac{C}{t^{n/2}} \exp(m(t, x - \xi)).$$

Finally, we shift $\Gamma(t - \tau, x, 0, \xi)$ by τ to obtain estimate (1.21). Now the proof is complete. $\square$

We may give an elementary and explicit estimate for the function m in the theorem we just proved, which also gives a more explicit form of this upper bound.

Corollary 2.15. *Under the same assumptions and notations as in Theorem 1.6, if $\mu \equiv \frac{2}{2-\gamma+\frac{2}{l}} > 1$, the fundamental solution has upper bound*

$$\Gamma(t, x; \tau, \xi) \leq \begin{cases} \frac{C_1}{(t-\tau)^{n/2}} \exp \left(-\frac{1}{C_2} \left(\frac{|x-\xi|^2}{t-\tau} \right) \right) & \frac{|x|^\mu}{t^{\mu-\nu-1}} < 1 \\ \frac{C_1}{(t-\tau)^{n/2}} \exp \left(-\frac{1}{C_2} \left(\frac{|x-\xi|^\mu}{(t-\tau)^\nu} \right)^{\frac{1}{\mu-1}} \right) & \frac{|x|^\mu}{t^{\mu-\nu-1}} \geq 1 \end{cases} \quad (2.52)$$

where $\Lambda = \|b\|_{L^l(0,T;L^q(\mathbb{R}^n))}$, $C_1 = C_1(l, q, n, \lambda)$, $C_2 = C_2(l, q, n, \lambda, \Lambda)$. If $\mu = 1$, which implies $q = \infty$, we can solve for $m(t, x)$ explicitly and obtain

$$\Gamma(t, x; \tau, \xi) \leq \frac{C_1}{(t-\tau)^{n/2}} \exp \left(-\frac{(C_1 \Lambda (t-\tau)^\nu - |x-\xi|)^2}{4C_1(t-\tau)} \right). \quad (2.53)$$

Proof. Clearly, it is enough to estimate function $m(t, x)$. In this proof, we denote C_1 as a constant depending only on (l, q, n, λ) and C_2 a constant depending on $(l, q, n, \lambda, \Lambda)$. Their values may be different throughout the proof. Notice that $\mu \geq 1$. When $\mu > 1$, by

taking $\alpha = -\frac{x}{4C_2t}$, we have

$$m(t, x) \leq \frac{C_1|x|^2}{16C_2^2t} + \frac{C_1\Lambda^\mu|x|^\mu}{4^\mu C_2^\mu t^{\mu-\nu}} - \frac{|x|^2}{4C_2t} = \frac{C_1|x|^2}{16C_2^2t} + \frac{C_1\Lambda^\mu|x|^2}{4^\mu C_2^\mu t^1} \cdot \frac{|x|^{\mu-2}}{t^{\mu-\nu-1}} - \frac{|x|^2}{4C_2t} \leq -\frac{|x|^2}{8C_2t}$$

if $\frac{|x|^{\mu-2}}{t^{\mu-\nu-1}} < 1$. When $\frac{|x|^{\mu-2}}{t^{\mu-\nu-1}} \geq 1$, we take $\alpha = -\frac{1}{4C_2} \left(\frac{|x|}{t^\nu}\right)^{\frac{1}{\mu-1}} \frac{x}{|x|}$. Then one has

$$\begin{aligned} m(t, x) &\leq \frac{C_1|x|^{\frac{2}{\mu-1}}}{16C_2^2t^{\frac{2\nu}{\mu-1}-1}} + \frac{C_1\Lambda^\mu|x|^{\frac{\mu}{\mu-1}}}{4^\mu C_2^\mu t^{\frac{\nu}{\mu-1}}} - \frac{|x|^{\frac{\mu}{\mu-1}}}{C_2t^{\frac{\nu}{\mu-1}}} \\ &= \frac{C_1|x|^{\frac{\mu}{\mu-1}}}{16C_2^2t^{\frac{\nu}{\mu-1}}} \cdot \frac{|x|^{\frac{2-\mu}{\mu-1}}}{t^{\frac{\nu-\mu+1}{\mu-1}}} + \frac{C_1\Lambda^\mu|x|^{\frac{\mu}{\mu-1}}}{4^\mu C_2^\mu t^{\frac{\nu}{\mu-1}}} - \frac{|x|^{\frac{\mu}{\mu-1}}}{C_2t^{\frac{\nu}{\mu-1}}} \leq -\frac{|x|^{\frac{\mu}{\mu-1}}}{2C_2t^{\frac{\nu}{\mu-1}}}. \end{aligned}$$

Now consider the case that $\mu = 1$. To obtain $m(t, x) = \min_{\alpha \in \mathbb{R}^n} (C_1(|\alpha|^2t + |\alpha|\Lambda t^\nu) + \alpha \cdot x)$, it is easy to see that α must be in opposite direction of x , i.e. $\frac{\alpha}{|\alpha|} = -\frac{x}{|x|}$. So we only need to find the minimum of the polynomial $C_1t|\alpha|^2 + (C_1\Lambda t^\nu - |x|)|\alpha|$, which is obtained at $|\alpha| = -\frac{C_1\Lambda t^\nu - |x|}{2C_1t}$ and the value is

$$m(t, x) = -\frac{(C_1\Lambda t^\nu - |x|)^2}{4C_1t}.$$

Now the proof is complete. □

Recall that in dimension three, any Leray-Hopf weak solution to Navier-Stokes equations satisfies

$$u \in L^\infty(0, T; L^2(\mathbb{R}^3)) \cap L^2(0, T; H^1(\mathbb{R}^3)).$$

Clearly $L^2(0, T; H^1(\mathbb{R}^3)) \subset L^2(0, T; L^6(\mathbb{R}^3))$, thus $\gamma = \frac{3}{2}$ for both function spaces. Notice that by interpolation, $u \in L^l(0, T; L^q(\mathbb{R}^3))$ for any $l \in [2, \infty]$ and $q \in [2, 6]$ satisfying $\frac{2}{l} + \frac{3}{q} = \frac{3}{2}$. This is an interesting case for which we have the following theorem.

Theorem 2.16. *Suppose $n = 3$, and conditions **(E)** and **(S)** hold for $b \in L^l(0, T; L^q(\mathbb{R}^3))$ satisfying $\frac{2}{l} + \frac{3}{q} = \frac{3}{2}$. Then the fundamental solution Γ to (1.9) has the upper bound*

$$\Gamma(t, x; \tau, \xi) \leq \begin{cases} \frac{C_1}{(t-\tau)^{3/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|x-\xi|^2}{t-\tau}\right)\right) & \frac{|x|^{l-4}}{t^{l-2}} < 1 \\ \frac{C_1}{(t-\tau)^{3/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|x-\xi|^4}{t-\tau}\right)^{\frac{1}{3-\frac{4}{l}}}\right) & \frac{|x|^{l-4}}{t^{l-2}} \geq 1, \end{cases} \quad (2.54)$$

where $\Lambda = \|b\|_{L^l(0,T;L^q(\mathbb{R}^3))}$, $C_1 = C_1(l, q, n, \lambda)$, $C_2 = C_2(l, q, \lambda, \Lambda)$. Here we set $\frac{|x|^{l-4}}{t^{l-2}} = \frac{|x|}{t}$ when $l = \infty$.

Since we know that $\int_{\mathbb{R}^n} \Gamma(t, x; \tau, \xi) d\xi = 1$ and we have proved the upper bound in Corollary 2.15, which is of exponential decay in space, we can derive a lower bound for Γ in the following form. This proposition will be used in the proof of a pointwise lower bound later.

Proposition 2.17. *Take the fundamental solution of (1.9) satisfying conditions (E) and (S), for any $\delta \in (0, 1)$ and $t - \tau$ small enough, we have*

$$\int_{B(x, \tilde{R}(t-\tau))} \Gamma(t, x; \tau, \xi) d\xi \geq \delta,$$

where $\tilde{R}(\cdot)$ is a function defined as follows

$$\tilde{R}(t) = \begin{cases} Ct^{1/2} & \text{if } \gamma = 1 \\ Ct^{(2-\gamma)/2} \ln \frac{1}{t} & \text{if } \gamma > 1, \end{cases}$$

$B(x, r)$ is the ball of radius r and center x , and C depends only on $(\delta, l, q, n, \lambda, \Lambda)$.

Proof. Firstly, when $\mu > 1$, we have

$$\Gamma(t, x; \tau, \xi) \leq h_1(t - \tau, x - \xi) + h_2(t - \tau, x - \xi), \quad (2.55)$$

where

$$h_1(t, x) = \frac{C_1}{t^{n/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|x|^2}{t}\right)\right), \quad h_2(t, x) = \frac{C_1}{t^{n/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|x|^\mu}{t^\nu}\right)^{\frac{1}{\mu-1}}\right).$$

Thus it is enough to prove that

$$\int_{B(x, \tilde{R}(t-\tau))^c} h_1(t-\tau, x-\xi) + h_2(t-\tau, x-\xi) d\xi \leq 1-\delta.$$

Without loss of generality, we can assume $\tau = 0$ and $x = 0$. By the following change of variable

$$\int_{B(0, RC_2^{1/2}t^{1/2})^c} \frac{C_1}{t^{n/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|\xi|^2}{t}\right)\right) d\xi = C \int_{B(0, R)^c} \exp(-|\xi|^2) d\xi,$$

we have

$$\int_{B(x, R_1(t))^c} \frac{C_1}{t^{n/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|\xi|^2}{t}\right)\right) d\xi \leq \frac{1-\delta}{2}$$

with $R_1(t) = Ct^{1/2}$ for some sufficiently large constant $C > 0$. For the second term, since $\frac{\nu}{\mu} = \frac{2-\gamma}{2} \leq \frac{1}{2}$, it follows that

$$\begin{aligned} \int_{B(0, RC_2^{(\mu-1)\nu/\mu}t^{\nu/\mu})^c} \frac{C_1}{t^{n/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|\xi|^\mu}{t^\nu}\right)^{\frac{1}{\mu-1}}\right) d\xi = \\ \frac{C}{t^{n(\frac{1}{2}-\frac{\nu}{\mu})}} \int_{B(0, R)^c} \exp\left(-|\xi|^{\frac{\mu}{\mu-1}}\right) d\xi. \end{aligned}$$

Setting $\Phi(R) = \int_{B(0, R)^c} \exp\left(-|\xi|^{\frac{\mu}{\mu-1}}\right) d\xi$, then one has $\Phi(R) \leq Ce^{-R}$ by $\frac{\mu}{\mu-1} > 1$. So if set $R = C(1 - (\frac{1}{2} - \frac{\nu}{\mu}) \ln t)$ for some C , we obtain that

$$\int_{B(0, R)^c} \exp\left(-|\xi|^{\frac{\mu}{\mu-1}}\right) d\xi \leq Ce^{-R} \leq \frac{t^{n(\frac{1}{2}-\frac{\nu}{\mu})}}{C_1} \cdot \frac{1-\delta}{2},$$

and therefore

$$\int_{B(0, R_2(t))^c} \frac{C_1}{t^{n/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|\xi|^\mu}{t^\nu}\right)^{\frac{1}{\mu-1}}\right) d\xi \leq \frac{1-\delta}{2}$$

with

$$R_2(t) = Ct^{\nu/\mu} \left(1 + \left(\frac{1}{2} - \frac{\nu}{\mu}\right) \ln \frac{1}{t}\right) = Ct^{(2-\gamma)/2} \left(1 + \left(\frac{1}{2} - \frac{2-\gamma}{2}\right) \ln \frac{1}{t}\right)$$

for some constant C . When t is small enough, $R_1(t) \leq R_2(t)$ and we obtain the radius

$$\tilde{R}(t) = R_2(t).$$

When $\mu = 1$, we have $\nu = \frac{2-\gamma}{2} \leq \frac{1}{2}$ and, by using the elementary inequality $(a-b)^2 + b^2 \geq \frac{a^2}{2}$,

$$\begin{aligned} & \int_{B(0, R_2 t^\nu)^c} \frac{C_1}{t^{n/2}} \exp\left(-\frac{(C_1 \Lambda t^\nu - |\xi|)^2}{4C_1 t}\right) d\xi \\ &= \frac{C}{t^{n(\frac{1}{2}-\nu)}} \int_{B(0, R)^c} \exp\left(-\frac{(C_1 \Lambda - |\xi|)^2}{C_1 t^{1-2\nu}}\right) d\xi \\ &\leq \frac{C}{t^{n(\frac{1}{2}-\nu)}} \int_{B(0, R)^c} \exp\left(-\frac{(C_1 \Lambda - |\xi|)^2}{C_1}\right) d\xi \\ &\leq \frac{C}{t^{n(\frac{1}{2}-\nu)}} \int_{B(0, R)^c} \exp\left(-\frac{|\xi|^2}{2C_1} + C_1^2 \Lambda^2\right) d\xi. \end{aligned}$$

Let $\Phi(R) = \int_{B(0, R)^c} \exp(-|\xi|^2) d\xi$. Then $\Phi(R) \leq C e^{-R}$ for some universal constant.

Thus we still take

$$\tilde{R}(t) = C t^{(2-\gamma)/2} \left(1 + \left(\frac{1}{2} - \frac{2-\gamma}{2}\right) \ln \frac{1}{t}\right)$$

to obtain

$$\int_{B(0, \tilde{R}(t))^c} \frac{C_1}{t^{n/2}} \exp\left(-\frac{(C_1 \Lambda t^\nu - |\xi|)^2}{4C_1 t}\right) d\xi \leq 1 - \delta.$$

Clearly, when $\gamma = 1$, $\tilde{R}(t)$ is just $C t^{1/2}$. We only need $\tilde{R}(t)$ for small t , and under this condition $\ln \frac{1}{t} \gg 1$. Thus we can $\tilde{R}(t) = C t^{(2-\gamma)/2} \ln \frac{1}{t}$ when $\gamma > 1$ and the proof is now complete. $\square$

Remark. Although estimate (2.55) seems to be better than (2.52), actually it can be shown that this observation will not affect the result. Therefore, based on Corollary 2.15, this $\tilde{R}(t)$ is the smallest cone radius such that we can derive a lower bound of this form inside the cone.

2.3.2 The lower bound

Let $T > 0$ and $x \in \mathbb{R}^n$. To prove the lower bound, the idea is to consider the quantity

$$G_r(t, x) = \int_{\mathbb{R}^n} \ln(\Gamma(T, x; T - t, \xi)) \mu_r(d\xi) = \int_{\mathbb{R}^n} \ln(\Gamma_T^*(t, \xi; 0, x)) \mu_r(d\xi)$$

for $t \in [0, T]$, where $\mu_r(x) = \frac{1}{r^{n/2}} \exp\left(-\frac{\pi|x|^2}{r}\right)$ as defined in Lemma 2.5. Then Jensen's inequality implies that $G_r(t, x) \leq 0$. We will write it as $G(t, x)$ if $r = 1$. If we have $G_r(T, x) > -C$ for some positive constant C , then we can derive a lower bound for $\Gamma(T, x; 0, \xi)$. Consider the time derivative of $G_r(t, x)$

$$\begin{aligned} G'_r(t, x) &= \int_{\mathbb{R}^n} \partial_t \ln(\Gamma(T, x; T - t, \xi)) \mu_r(d\xi) \\ &= \int_{\mathbb{R}^n} \left\langle \frac{2\pi\xi}{r}, a(T - t, \xi) \cdot \nabla_\xi \ln \Gamma(T, x; T - t, \xi) \right\rangle \mu_r(d\xi) \\ &\quad + \int_{\mathbb{R}^n} \langle \nabla_\xi \ln \Gamma(T, x; T - t, \xi), a(T - t, \xi) \cdot \nabla_\xi \ln \Gamma(T, x; T - t, \xi) \rangle \mu_r(d\xi) \\ &\quad + \int_{\mathbb{R}^n} \langle b(T - t, \xi), \nabla_\xi \ln \Gamma(T, x; T - t, \xi) \rangle \mu_r(d\xi). \end{aligned} \quad (2.56)$$

In the rest of this subsection, we will estimate $G_r(t, x)$ under various conditions on b and hence obtain a lower bound of Γ .

For $\gamma = 1$, or in particular $l = \infty, q = n$, which is the only case that regularity theory is missing, the argument is the same to the BMO case. Since in the critical case, $\|b\|_{L_t^\infty L_x^n}$ is invariant under scaling, we do not need to worry about explicitly how the constant depends on Λ . Hence we will only need to estimate $G(1, x)$ and obtain the estimate of $G(t, x)$ for all t by scaling. In supercritical case $1 < \gamma < 2$, in order to use scaling, we need to find out how the constants appearing in lower bounds depend on Λ , and therefore it is not a good idea to use the scaling argument. So we will alter the strategy to estimate $G_r(t, x)$ for all t directly.

Lemma 2.18. *Suppose $q \geq 2$, $l \geq 2$, $1 < \gamma < 2$ and $\tilde{R}(t)$ is defined as in Proposition 2.17. For any $\kappa > 0$, $x \in B(0, \kappa \tilde{R}(t))$ and $t > 0$ small enough, there is a constant $C > 0$*

depending only on $\kappa, l, q, n, \lambda, \Lambda = \|b\|_{L^l(0, T; L^q(\mathbb{R}^n))}$, such that

$$G_r(t, x) \geq -C(\lambda) \left(\frac{t}{r} + r^{-n/q} t^{\frac{l-2}{l}} \Lambda^2 \right) - C \left(\frac{r}{t} \right)^{n/2+1} \exp \left(\frac{\pi |\tilde{R}(t)|^2}{Cr} \right). \quad (2.57)$$

Proof. We fix a $T > 0$. By the definition of $G_r(t, x)$, for $0 < t_1 < t_2 \leq T$, we can deduce that

$$\begin{aligned} & G_r(t_2, x) - G_r(t_1, x) \\ &= \int_{t_1}^{t_2} \int_{\mathbb{R}^n} \frac{\partial_s \Gamma(T, x; T-s, \xi)}{\Gamma(T, x; T-s, \xi)} \mu_r(d\xi) ds \\ &= \int_{t_1}^{t_2} \int_{\mathbb{R}^n} \left\langle \frac{2\pi\xi}{r}, a(T-s, \xi) \cdot \nabla_\xi \ln \Gamma(T, x; T-s, \xi) \right\rangle \mu_r(d\xi) ds \\ &\quad + \int_{t_1}^{t_2} \int_{\mathbb{R}^n} \langle \nabla_\xi \ln \Gamma(T, x; T-s, \xi), a(T-s, \xi) \cdot \nabla_\xi \ln \Gamma(T, x; T-s, \xi) \rangle \mu_r(d\xi) ds \\ &\quad + \int_{t_1}^{t_2} \int_{\mathbb{R}^n} \langle b(T-s, \xi), \nabla_\xi \ln \Gamma(T, x; T-s, \xi) \rangle \mu_r(d\xi) ds \\ &\geq - \int_{t_1}^{t_2} \frac{2\pi}{\lambda r} \|\xi\|_{L^2(\mu_r)} \|\nabla_\xi \ln \Gamma(T, x; T-s, \cdot)\|_{L^2(\mu_r)} ds \\ &\quad + \int_{t_1}^{t_2} \lambda \|\nabla_\xi \ln \Gamma(T, x; T-s, \cdot)\|_{L^2(\mu_r)}^2 ds \\ &\quad - \int_{t_1}^{t_2} \|b(T-s, \cdot)\|_{L^q(\mathbb{R}^n)} \|\mu_r^{\frac{1}{2}}\|_{L^{\frac{2q}{q-2}}(\mathbb{R}^n)} \|\nabla_\xi \ln \Gamma(T, x; T-s, \cdot)\|_{L^2(\mu_r)} ds \\ &\geq - \int_{t_1}^{t_2} \frac{C(\lambda)}{r^2} \|\xi\|_{L^2(\mu_r)}^2 + C(\lambda) \|b(T-s, \cdot)\|_{L^q(\mathbb{R}^n)}^2 \|\mu_r^{\frac{1}{2}}\|_{L^{\frac{2q}{q-2}}(\mathbb{R}^n)}^2 ds \\ &\quad + \frac{\lambda}{2} \int_{t_1}^{t_2} \|\nabla_\xi \ln \Gamma(T, x; T-s, \cdot)\|_{L^2(\mu_r)}^2 ds. \end{aligned}$$

Here we set $\frac{2q}{q-2} = \infty$ when $q = 2$. Since $l \geq 2$, we have

$$\int_0^T \|b(T-s, \cdot)\|_{L^q(\mathbb{R}^n)}^2 ds < \infty.$$

For the last term in the equation above, we use the Poincaré-Wirtinger inequality to obtain

that

$$\begin{aligned} \int_{t_1}^{t_2} \|\nabla_{\xi} \ln \Gamma(T, x; T-s, \cdot)\|_{L^2(\mu_r)}^2 ds &\geq \\ &Cr^{-1} \int_{t_1}^{t_2} \int_{\mathbb{R}^n} |\ln \Gamma(T, x; T-s, \xi) - G_r(s, x)|^2 \mu_r(d\xi) ds. \end{aligned}$$

Since $G_r(t, x) \leq 0$, using $(a-b)^2 \geq \frac{a^2}{2} - b^2$, the right-hand side can be estimated as

$$\begin{aligned} &\int_{t_1}^{t_2} \int_{\mathbb{R}^n} |\ln \Gamma(T, x; T-s, \xi) - G_r(s, x)|^2 \mu_r(d\xi) ds \\ &\geq \int_{t_1}^{t_2} \int_{\{\ln \Gamma(T, x; T-s, \xi) \geq -1\}} (\ln \Gamma(T, x; T-s, \xi) - G_r(s, x) - 1 + 1)^2 \mu_r(d\xi) ds \\ &\geq \frac{1}{2} \int_{t_1}^{t_2} \int_{\{\ln \Gamma(T, x; T-s, \xi) \geq -1\}} (\ln \Gamma(T, x; T-s, \xi) + 1 - G_r(s, x))^2 \mu_r(d\xi) - 1 ds \\ &\geq \frac{1}{2} \int_{t_1}^{t_2} \int_{\{\ln \Gamma(T, x; T-s, \xi) \geq -1\}} G_r(s, x)^2 \mu_r(d\xi) - 1 ds \\ &= \frac{1}{2} \int_{t_1}^{t_2} G_r(s, x)^2 \mu_r\{\ln \Gamma(T, x; T-s, \xi) \geq -1\} - 1 ds. \end{aligned}$$

By Proposition 2.17, for any $x \in B(0, \kappa \tilde{R}(t))$, we have

$$\int_{B(0, (C+\kappa)\tilde{R}(t))} \Gamma(T, x; T-t, \xi) d\xi \geq \frac{1}{2},$$

which implies that

$$\mu_r\{\ln \Gamma(T, x; T-t, \xi) \geq -1\} \frac{Cr^{n/2}}{T^{n/2}} \exp\left(\frac{\pi|\tilde{R}(T)|^2}{Cr}\right) + |B(0, (C+\kappa)\tilde{R}(t))|e^{-1} \geq \frac{1}{2}$$

for $t \in [\frac{T}{2}, T]$. So if we take $T > 0$ small enough such that $|B(0, (C+\kappa)\tilde{R}(t))|e^{-1} \leq \frac{1}{4}$ for all $t \in [0, T]$, then

$$\mu_r\{\ln \Gamma(T, x; T-t, \xi) \geq -1\} \geq C\left(\frac{T}{r}\right)^{n/2} \exp\left(-\frac{\pi|\tilde{R}(T)|^2}{Cr}\right).$$

Also it is easy to calculate that $\|\xi\|_{L^2(\mu_r)}^2 = r$ and

$$\|\mu_r^{\frac{1}{2}}\|_{L^{\frac{2q}{q-2}}(\mathbb{R}^n)}^2 = \|\mu_r\|_{L^{\frac{q}{q-2}}(\mathbb{R}^n)} = C(q)r^{-n/q}.$$

We now can conclude that

$$\begin{aligned} G_r(t_2, x) - G(t_1, x) &\geq - \int_{t_1}^{t_2} \frac{C(\lambda)}{r} + C(\lambda)r^{-n/q} \|b(T-s, \cdot)\|_{L^q(\mathbb{R}^n)}^2 + Cr^{-1} ds \\ &\quad + Cr^{-1} \left(\frac{T}{r}\right)^{n/2} \exp\left(-\frac{\pi|\tilde{R}(T)|^2}{Cr}\right) \int_{t_1}^{t_2} G_r(s, x)^2 ds, \end{aligned}$$

for $\frac{T}{2} \leq t_1 < t_2 \leq T$. By Lemma 2.8 and $l \geq 2$, we have

$$G_r(T, x) \geq -C(\lambda) \left(\frac{T}{r} + r^{-n/q} T^{\frac{l-2}{l}} \Lambda^2\right) - C\left(\frac{r}{T}\right)^{n/2+1} \exp\left(\frac{\pi|\tilde{R}(T)|^2}{Cr}\right).$$

□

Proof of Theorem 1.8. For $x, \xi \in B(0, \kappa\tilde{R}(t))$, by using the Chapman-Kolmogorov equation we obtain that

$$\begin{aligned} \ln \Gamma(2T, x; 0, \xi) &= \ln \int_{\mathbb{R}^n} \Gamma(2T, x; T, z) \Gamma(T, z; 0, \xi) dz \\ &\geq \ln \int_{\mathbb{R}^n} r^{n/2} \Gamma(2T, x; T, z) \Gamma(T, z; 0, \xi) \mu_r(dz) \\ &\geq \frac{n}{2} \ln r + \int_{\mathbb{R}^n} \ln \Gamma(2T, x; T, z) \Gamma(T, z; 0, \xi) \mu_r(dz) \\ &= \frac{n}{2} \ln r + \int_{\mathbb{R}^n} \ln \Gamma(2T, x; T, z) \mu_r(dz) + \int_{\mathbb{R}^n} \ln \Gamma(T, z; 0, \xi) \mu_r(dz) \\ &\geq \frac{n}{2} \ln r - C(\lambda) \left(\frac{T}{r} + r^{-n/q} T^{\frac{l-2}{l}} \Lambda^2\right) - C\left(\frac{r}{T}\right)^{n/2+1} \exp\left(\frac{\pi|\tilde{R}(T)|^2}{Cr}\right), \end{aligned}$$

i.e.

$$\Gamma(2T, x; 0, \xi) \geq r^{\frac{n}{2}} \exp \left[-C \left(\frac{T}{r} + r^{-n/q} T^{\frac{l-2}{l}} \Lambda^2 \right) - C \left(\frac{r}{T} \right)^{n/2+1} \exp \left(\frac{\pi |\tilde{R}(T)|^2}{Cr} \right) \right].$$

Now we can take maximum of the right-hand side over all positive r . Recall $\tilde{R}(t) = Ct^{(2-\gamma)/2} \ln \frac{1}{t}$, if we take $r = \tilde{R}(T)^2$, then the right-hand side becomes

$$T^{\frac{n}{2}(2-\gamma)} \left(\ln \frac{1}{T} \right)^{\frac{n}{2}} \exp \left[-CT^{\theta_1} \left(\ln \frac{1}{T} \right)^{-2} - CT^{\theta_2} \left(\ln \frac{1}{T} \right)^{-2n/q} - CT^{\theta_3} \left(\ln \frac{1}{T} \right)^{(n+2)} \right],$$

where $\theta_1 = \gamma - 1$, $\theta_2 = 1 - \frac{2}{l} - \frac{n}{q}(2 - \gamma)$, and $\theta_3 = (\frac{n}{2} + 1)(1 - \gamma) < 0$. Clearly $\theta_3 = \min\{\theta_1, \theta_2, \theta_3\} < 0$. Because we consider only for small t , the dominant term will be

$$\Gamma(2T, x; 0, \xi) \geq \exp \left[-CT^{\theta_3} \left(\ln \frac{1}{T} \right)^{(n+2)} \right],$$

and the proof is complete. □

We can use the Chapman-Kolmogorov equation to obtain a positive lower bound on the whole space, but we prefer to omit the details of computations.

Remark 2.19. Using full power of the Poincaré-Wirtinger inequality and following similar arguments as above, we can actually drop the assumptions that $q \geq 2$ and $l \geq 2$. We only need to assume that $1 \leq \gamma < 2$ to obtain a lower bound.

Chapter 3

Weak solutions and diffusion processes: critical cases

Using the *a priori* Aronson-type estimates proved in Chapter 2, in this chapter, we will study the weak solutions and the related diffusion processes in the critical case. In Section 3.1, we will show that under the critical condition $d \in L^\infty(0, T; \text{BMO}(\mathbb{R}^n))$ (or equivalently $b \in L^\infty(0, \infty; \text{BMO}^{-1}(\mathbb{R}^n))$), the Aronson estimate implies Hölder continuity of weak solutions following the ideas from Stroocks [79]. We also obtain the uniqueness of weak solutions in Section 3.2 through approximation. Moreover, we can construct a unique diffusion process from the results in first two sections. In Section 3.3, we further construct a strong solution to the SDE when the diffusion part is Brownian motion. We show that there is a unique almost everywhere defined strong solution if we in addition assume that $b \in L^2(0, T; H^1)$. The $L^2(0, T; H^1)$ condition can be generalized with current method. But we stick to this special condition here because it is satisfied by the Leray-Hopf weak solution to the Navier-Stokes equations.

3.1 Hölder regularity of the solutions

In the previous chapter, if $d \in L^\infty(0, T; \text{BMO}(\mathbb{R}^n))$, we have proved the Aronson estimate

$$\frac{1}{Ct^{n/2}} \exp(-C \frac{|x - \xi|^2}{t}) \leq \Gamma(t, x; 0, \xi) \leq \frac{C}{t^{n/2}} \exp(-\frac{|x - \xi|^2}{Ct}),$$

where C only depends on n , λ and $\|d\|_{L^\infty(\text{BMO})}$. In this section, we will prove the regularity results to the parabolic equations using the Aronson estimate. We will still assume that the coefficients a and d are smooth and the key point is that the constants in the continuity theorem do not depend on the smoothness. Then the same estimate will still be true when we approximate singular coefficients, because of the point-wise convergence of weak solutions and fundamental solutions obtained in the next section.

Recall the parabolic equation

$$\sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left[A_{ij}(t, x) \frac{\partial}{\partial x_j} u(t, x) \right] - \frac{\partial}{\partial t} u(t, x) = 0, \quad (3.1)$$

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

Let us consider Cauchy's initial and Dirichlet boundary problem associated with (3.1). Let $D \subset \mathbb{R}^n$ be an open subset with a smooth boundary. Given $\tau > 0$, $u(t, x)$, which is a locally integrable and Borel measurable function in $(t, x) \in [\tau, T] \times D$, is a weak solution to the Dirichlet boundary problem of (3.1) with initial data $u(\tau, \cdot) = f \in L^2(D)$, if $u \in L^2(\tau, T; H^1(D)) \cap L^\infty(\tau, T; L^2(D))$ and

$$\begin{aligned} - \int_{\tau}^T \int_D \langle \nabla \varphi(t, x), A(t, x) \cdot \nabla u(t, x) \rangle dx dt + \int_{\tau}^T \int_D u(s, x) \frac{\partial}{\partial s} \varphi(s, x) ds dx \\ + \int_D f(x) \varphi(\tau, x) dx = 0 \end{aligned} \quad (3.2)$$

for any smooth function $\varphi(s, x)$ which has compact support in $(\tau, T) \times D$. Let $\Gamma^D(t, x; \tau, \xi)$ denote the corresponding fundamental solution. Then we recall the following result in [44, Chapter IV, Section 15] when the coefficients are smooth.

Lemma 3.1. *Suppose that A_{ij} are smooth, so that there exists a smooth fundamental solution $\Gamma(t, x; \tau, \xi)$ satisfying the Aronson estimate, and therefore*

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

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

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

Moreover, we have the energy inequality

$$\|u(t, \cdot)\|_{L^2}^2 + 2\lambda \int_{\tau}^t \|\nabla u(s, \cdot)\|_{L^2}^2 \leq \|f\|_{L^2}^2 \quad (3.3)$$

for all $t \geq \tau$, and $u(t, x)$ is also a weak solution to (3.2).

Proof. This is a well known result in the theory of parabolic equations. Suppose f is smooth with compact support in D , then $u(t, x) = \Gamma_{\tau, t}^D f(x)$ is a classical solution to the parabolic equation (3.1), so that

$$\sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(A_{ij}(t, x) \frac{\partial}{\partial x_j} u(t, x) \right) - \frac{\partial}{\partial t} u(t, x) = 0$$

for all $x \in D$ and $t \geq \tau$. It follows that

$$-\int_D \sum_{i,j=1}^n A_{ij}(t, x) \frac{\partial}{\partial x_i} u(t, x) \frac{\partial}{\partial x_j} u(t, x) dx - \frac{1}{2} \frac{\partial}{\partial t} \int_D u(t, x)^2 dx = 0$$

for all $t > \tau$, and therefore we have the energy inequality

$$\|u(t, \cdot)\|_{L^2}^2 + 2\lambda \int_{\tau}^t \|\nabla u(s, \cdot)\|_{L^2}^2 ds \leq \|f\|_{L^2}^2 \quad (3.4)$$

for all $t > \tau$. From the energy inequality above, we deduce that for every $f \in L^2(D)$, $u(t, x) = \Gamma_{\tau, t}^D f(x)$ belongs to $L^\infty(\tau, T; L^2(D))$ and also to $L^2(\tau, T; H^1(D))$, and the energy inequality remains true. Therefore for any $\varphi(t, x)$ which is smooth with compact support in $[\tau, t) \times D$, we have

$$\int_{\tau}^t \int_D u(s, x) \frac{\partial}{\partial s} \varphi(s, x) dx ds - \int_{\tau}^t \int_D \langle \nabla \varphi(s, x), A(s, x) \cdot \nabla u(s, x) \rangle dx ds + \int_D f(x) \varphi(\tau, x) dx = 0, \quad (3.5)$$

which is true for smooth f with compact support, so that it remains true for $f \in L^2(D)$ by the energy inequality above. Thus $u(t, x) = \Gamma_{\tau, t}^D f(x)$ is a weak solution with initial data $f \in L^2(D)$. $\square$

This result allows us to approximate A by $\{A_m\}$ and there exists a unique fundamental solution Γ^m and strong solution u^m corresponding to A_m , which satisfy the claims in Lemma 3.1.

We denote the parabolic ball as $Q((t_0, x_0), R) = (t_0 - R^2, t_0) \times B(x_0, R)$ and

$$Osc_{Q((t_0, x_0), R)} u = \max_{Q((t_0, x_0), R)} u - \min_{Q((t_0, x_0), R)} u.$$

Firstly, we will prove Nash's continuity theorem as follows.

Theorem 3.2. *Suppose $u \in C^{1,2}(Q((t_0, x_0), R))$ is a solution to equation (3.1), then for any $\delta \in (0, 1)$, there are $\alpha \in (0, 1]$ and $C > 0$ depending only on $(\delta, n, \lambda, \|d\|_{L^\infty(BMO)})$ such that*

$$|u(t_1, x_1) - u(t_2, x_2)| \leq C \left(\frac{|t_1 - t_2|^{1/2} \vee |x_1 - x_2|}{R} \right)^\alpha Osc_{Q((t_0, x_0), R)} u$$

for any $(t_1, x_1), (t_2, x_2) \in Q((t_0, x_0), \delta R)$.

Applying this theorem to the fundamental solution, we have the following corollary.

Corollary 3.3. *There exist $\alpha \in (0, 1]$ and $C > 0$ depending only on $(n, \lambda, \|d\|_{L^\infty(BMO)})$ such that for any $\delta > 0$, we have*

$$|\Gamma(t_1, x_1; 0, \xi_1) - \Gamma(t_2, x_2; 0, \xi_2)| \leq \frac{C}{\delta^n} \left(\frac{|t_1 - t_2|^{1/2} \vee |x_1 - x_2| \vee |\xi_1 - \xi_2|}{\delta} \right)^\alpha$$

for all $(t_1, x_1, \xi_1), (t_2, x_2, \xi_2) \in [\delta^2, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$ with $|x_1 - x_2| \vee |\xi_1 - \xi_2| \leq \delta$.

Here we prove Nash's continuity theorem. The proof is inspired by [79], which was originally written in probability language and relies heavily on the strong Markov property of the diffusion process. Here we rewrite it using a PDE approach instead.

Since we still assume (a, d) to be smooth, it implies that equation (3.1) is equivalent to

$$\partial_t u - \operatorname{div}(a \cdot \nabla u) + b \cdot \nabla u = 0 \quad (3.6)$$

with $b = \operatorname{div} d$. Clearly this equation satisfies the maximum principle. We consider the Dirichlet problem on $[0, T] \times B(x_0, R)$ for any fixed x_0 and $R > 0$ with $u(0, x) = f(x)$ and $u(t, x) = 0$ for $x \in \partial B(x_0, R)$. Then there is a unique regular fundamental solution $\Gamma^{x_0, R}(t, x; \tau, \xi)$ with $x, \xi \in B(x_0, R)$. So for any $f \in C_0^\infty(B(x_0, R))$ satisfying $f \geq 0$,

$$\Gamma_t^{x_0, R} f(x) = \int_{B(x_0, R)} \Gamma^{x_0, R}(t, x; 0, \xi) f(\xi) d\xi$$

is the unique strong solution to Dirichlet problem. We will prove the following lower bound for $\Gamma^{x_0, R}(t, x; \tau, \xi)$, which is also interesting by its own.

Theorem 3.4. *For any $\delta \in (0, 1)$, there exists a constant $C = C(\delta, n, \lambda, \|d\|_{L^\infty(BMO)})$ such that*

$$\Gamma^{x_0, R}(t, x; \tau, \xi) \geq \frac{1}{C(t - \tau)^{n/2}} \exp\left(-C \frac{|x - \xi|^2}{t - \tau}\right)$$

for any $t - \tau \in (0, R^2]$ and $x, \xi \in B(x_0, \delta R)$.

Proof. Without loss of generality, we take $\tau = 0$. For any $t > 0$, given $f \in C_0^\infty(B(x_0, R))$ satisfying $f \geq 0$, consider $w(s, x) = \Gamma_s f(x) - \Gamma_s^{x_0, R} f(x) - M$ for $s \in [0, t]$ where

$$M = \sup_{s \in [0, t], z \in B(x_0, R)^c} \Gamma_s f(z).$$

Then we notice that w solves (3.6) in $(0, t] \times B(x_0, R)$ with the initial-boundary condition that $w(0, x) \leq 0$ for $x \in B(x_0, R)$ and $w(s, x) \leq 0$ for $s \in (0, t]$, $x \in \partial B(x_0, R)$. So the maximum principle implies that $w \leq 0$ in $(0, t] \times B(x_0, R)$, which means $\Gamma_t^{x_0, R} f(x) \geq \Gamma_t f(x) - M$. Since this is true for any $f \in C_0^\infty(B(x_0, \delta R))^+$ with $\delta \in (0, 1)$, we have

$$\begin{aligned} \Gamma^{x_0, R}(t, x; 0, \xi) &\geq \Gamma(t, x; 0, \xi) - \sup_{s \in [0, t], z \in B(x_0, R)^c, y \in B(x_0, \delta R)} \Gamma(s, z; 0, y) \\ &\geq \frac{1}{Ct^{n/2}} \exp\left(-C \frac{|x - \xi|^2}{t}\right) - \sup_{s \in [0, t]} \frac{C}{s^{n/2}} \exp\left(-\frac{(1 - \delta)^2 R^2}{Cs}\right) \end{aligned}$$

for any $x, \xi \in B(x_0, \delta R)$. Consider the second term, and set $\tilde{t} = t/R^2$ and $\tilde{s} = s/R^2$

$$\begin{aligned} &\sup_{s \in [0, t]} \frac{C}{s^{n/2}} \exp\left(-\frac{(1 - \delta)^2 R^2}{Cs}\right) \\ &= \frac{1}{2Ct^{n/2}} \exp\left(-C \frac{|x - \xi|^2}{t}\right) \sup_{s \in [0, t]} \frac{2C^2 t^{n/2}}{s^{n/2}} \exp\left(-\frac{(1 - \delta)^2 R^2}{Cs} + C \frac{|x - \xi|^2}{t}\right) \\ &= \frac{1}{2Ct^{n/2}} \exp\left(-C \frac{|x - \xi|^2}{t}\right) \sup_{\tilde{s} \in [0, \tilde{t}]} \frac{2C^2 \tilde{t}^{n/2}}{\tilde{s}^{n/2}} \exp\left(-\frac{(1 - \delta)^2}{C\tilde{s}} + C \frac{|x - \xi|^2}{\tilde{t}R^2}\right). \end{aligned}$$

If $|x - \xi|^2 \leq \frac{(1 - \delta)^2 R^2}{2C^2}$ and $t \leq R^2$, it implies

$$\sup_{\tilde{s} \in [0, \tilde{t}]} \frac{2C^2 \tilde{t}^{n/2}}{\tilde{s}^{n/2}} \exp\left(-\frac{(1 - \delta)^2}{C\tilde{s}} + C \frac{|x - \xi|^2}{\tilde{t}R^2}\right) \leq \sup_{\tilde{s} \in [0, \tilde{t}]} \frac{2C^2}{\tilde{s}^{n/2}} \exp\left(-\frac{(1 - \delta)^2}{2C\tilde{s}}\right),$$

where $\frac{2C^2}{\tilde{s}^{n/2}} \exp\left(-\frac{(1 - \delta)^2}{2C\tilde{s}}\right) \rightarrow 0$ as $\tilde{s} \rightarrow 0$. So we can take $\tilde{t}$ small enough so that $RHS \leq 1$

and hence we have

$$\Gamma^{x_0, R}(t, x; 0, \xi) \geq \frac{1}{2Ct^{n/2}} \exp\left(-C \frac{|x - \xi|^2}{t}\right)$$

where $\max\{t, |x - \xi|^2\} \leq \varepsilon^2 R^2$ for some small ε depending on $(\delta, n, \lambda, \Lambda)$.

Now we use the Chapman-Kolmogorov equation to extend this to any $x, \xi \in B(x_0, \delta R)$ and $t \in (0, R^2]$. First consider $|x - \xi| \geq \varepsilon R$ and any t , we set $\xi_m = \xi + \frac{m}{k+1}(x - \xi)$, $B_m = B(x_0, \delta R) \cap B(\xi_m, \frac{|x - \xi|}{k+1})$, $t_m = \frac{mt}{k+1}$. Then for any $z_m \in B_m$, we have $|z_m - z_{m-1}| \leq \frac{3|x - \xi|}{k+1}$. So, to obtain $|z_m - z_{m-1}| \leq \varepsilon R$ and $|t_m - t_{m-1}| \leq \varepsilon^2 R^2$, we just need to choose $k \geq \frac{3}{\varepsilon^2}$. Now one has

$$\begin{aligned} \Gamma^{x_0, R}(t, x; 0, \xi) &\geq \int_{B_1} \cdots \int_{B_k} \prod_{m=0}^k \Gamma^{x_0, R}(t_{m+1}, z_{m+1}; t_m, z_m) dz_k \cdots dz_1 \\ &\geq C \left(\frac{|x - \xi|}{k+1} \right)^{nk} \left(\frac{(k+1)^{n/2}}{2Ct^{n/2}} \exp\left(-C \frac{|x - \xi|^2}{(k+1)t}\right) \right)^{k+1} \\ &\geq C \frac{|x - \xi|^{nk}}{t^{nk/2}} \frac{1}{t^{n/2}} \exp\left(-C \frac{|x - \xi|^2}{t}\right) \\ &\geq \frac{1}{Ct^{n/2}} \exp\left(-C \frac{|x - \xi|^2}{t}\right). \end{aligned}$$

The only case left now is the case where $|x - \xi| \leq \varepsilon R$ and $t \geq \varepsilon^2 R^2$. Set t_m as before, then

$$\begin{aligned} \Gamma^{x_0, R}(t, x; 0, \xi) &\geq C \left(\frac{\varepsilon R}{k+1} \right)^{nk} \left(\frac{(k+1)^{n/2}}{2Ct^{n/2}} \exp\left(-C \frac{(k+1)|x - \xi|^2}{t}\right) \right)^{k+1} \\ &\geq \frac{1}{Ct^{n/2}} \exp\left(-C \frac{|x - \xi|^2}{t}\right), \end{aligned}$$

and the proof is complete. □

Now we give the proof of Nash's continuity theorem. First consider a non-negative

solution on a parabolic ball $u \in C^{1,2}([t_0 - R^2, t_0] \times \overline{B(x_0, R)})$, clearly we have

$$u(t, x) \geq \int_{B(x_0, R)} \Gamma^{x_0, R}(t, x; t_0 - R^2, \xi) u(t_0 - R^2, \xi) d\xi$$

by the maximum principle. Then by Theorem 3.4

$$u(t, x) \geq \frac{1}{C|B(x_0, \delta_2 R)|} \int_{B(x_0, \delta_2 R)} u(t_0 - R^2, \xi) d\xi \quad (3.7)$$

for any $(t, x) \in [t_0 - \delta_1^2 R^2, t_0] \times \overline{B(x_0, \delta_2 R)}$, $\delta_1, \delta_2 \in (0, 1)$, and C depending only on $\delta_1, \delta_2, n, \lambda$ and $\|d\|_{L^\infty(BMO)}$. This estimate is called the super-mean value property.

Lemma 3.5. *Suppose $u \in C^{1,2}(Q((t_0, x_0), R))$ is a solution to equation (3.1), then for any $\delta \in (0, 1)$, there is a $\theta = \theta(\delta, n, \lambda, \|d\|_{L^\infty(BMO)}) \in (0, 1)$ such that*

$$\text{Osc}_{Q((t_0, x_0), \delta R)} u \leq \theta \text{Osc}_{Q((t_0, x_0), R)} u.$$

Proof. Let

$$M(r) = \max_{Q((t_0, x_0), r)} u, \quad m(r) = \min_{Q((t_0, x_0), r)} u,$$

and consider $M(R) - u$ and $u - m(R)$, which are non-negative solutions. Inequality (3.7) implies that

$$M(R) - M(\delta R) \geq \frac{1}{C|B(x_0, \delta R)|} \int_{B(x_0, \delta R)} M(R) - u(t_0 - R^2, \xi) d\xi,$$

and

$$m(\delta R) - m(R) \geq \frac{1}{C|B(x_0, \delta R)|} \int_{B(x_0, \delta R)} u(t_0 - R^2, \xi) - m(R) d\xi.$$

The sum of these two inequalities gives us

$$[M(R) - m(R)] - [M(\delta R) - m(\delta R)] \geq \frac{1}{C} [M(R) - m(R)],$$

which completes the proof. □

Proof of Theorem 3.2. Denote $l = |t_1 - t_2|^{1/2} \vee |x_1 - x_2|$. If $\frac{l}{R} \geq 1 - \delta$, then it is easy to find C and the proof is done. If $\frac{l}{R} < 1 - \delta$, we choose integer K such that $(1 - \delta)^{K+1} \leq \frac{l}{R} < (1 - \delta)^K$. Assume $t_1 \leq t_2$. Then

$$\begin{aligned} |u(t_1, x_1) - u(t_2, x_2)| &\leq \underset{Q((t_2, x_2), (1-\delta)^K R)}{Osc} u \leq \theta^{K-1} \underset{Q((t_2, x_2), (1-\delta)R)}{Osc} u \\ &\leq \theta^{K-1} \underset{Q((t_0, x_0), R)}{Osc} u = \theta^{-2} (\theta^{K+1}) \underset{Q((t_0, x_0), R)}{Osc} u. \end{aligned}$$

Now we can find α such that $\theta = ((1 - \delta) \wedge \theta)^\alpha$, which implies $\theta^{K+1} \leq (1 - \delta)^{(K+1)\alpha} \leq (\frac{l}{R})^\alpha$ and the proof is complete. $\square$

Remark 3.6. Another important consequence of the Aronson estimate is the Harnack inequality as in Theorem 1.5. It is a simple consequence of the super-mean value property and Lemma 3.5. We will omit the proof here and a complete proof can be found in [79].

3.2 Uniqueness of weak solutions

In this section, we prove the uniqueness of weak solutions. To this end we need the following fact from Evans [20, Theorem 5.9.3].

Lemma 3.7. *Let $T > \tau \geq 0$. Let $u \in L^2(\tau, T; H^1(\mathbb{R}^n))$ and $\frac{\partial}{\partial t} u \in L^2(\tau, T; H^{-1}(\mathbb{R}^n))$, then we have $u \in C([\tau, T], L^2(\mathbb{R}^n))$ and*

$$\|u(T, \cdot)\|_{L^2}^2 - \|u(0, \cdot)\|_{L^2}^2 = 2 \left\langle \frac{\partial}{\partial t} u, u \right\rangle_{L^2(\tau, T; H^{-1}(\mathbb{R}^n)), L^2(\tau, T; H^1(\mathbb{R}^n))}, \quad (3.8)$$

where $\langle \cdot, \cdot \rangle_{W^*, W}$ denotes the pairing between a Banach space W and its dual Banach space W^* .

Proof. If $u(t, x) = \sum \varphi_i(x) \eta_i(t)$ where $\varphi_i \in H^1(\mathbb{R}^n)$ and η_i are smooth with compact support in (τ, T) , then

$$\int_{\mathbb{R}^n} u(t, x)^2 dx = \sum_{i,j} \eta_i(t) \eta_j(t) \int_{\mathbb{R}^n} \varphi_i(x) \varphi_j(x) dx$$

so that

$$\frac{d}{dt} \int_{\mathbb{R}^n} u(t, x)^2 dx = 2 \sum_{i,j} \eta'_i(t) \eta_j(t) \int_{\mathbb{R}^n} \varphi_i(x) \varphi_j(x) dx,$$

and therefore

$$\begin{aligned} \int_{\tau}^T \frac{d}{dt} \int_{\mathbb{R}^n} u(t, x)^2 dx dt &= 2 \int_{\tau}^T \sum_{i,j} \eta'_i(t) \eta_j(t) \int_{\mathbb{R}^n} \varphi_i(x) \varphi_j(x) dx dt \\ &= 2 \left\langle \frac{\partial}{\partial t} u, u \right\rangle_{L^2(\tau, T; H^{-1}(\mathbb{R}^n)), L^2(\tau, T; H^1(\mathbb{R}^n))}. \end{aligned}$$

Hence

$$\|u(T, \cdot)\|_{L^2}^2 - \|u(\tau, \cdot)\|_{L^2}^2 = 2 \left\langle \frac{\partial}{\partial t} u, u \right\rangle_{L^2(\tau, T; H^{-1}(\mathbb{R}^n)), L^2(\tau, T; H^1(\mathbb{R}^n))}.$$

By the density property, this equation remains true for any $u \in L^2(\tau, T; H^1(\mathbb{R}^n))$ such that $\frac{\partial}{\partial t} u \in L^2(\tau, T; H^{-1}(\mathbb{R}^n))$, and we can deduce that $u \in C([\tau, T], L^2(\mathbb{R}^n))$. $\square$

Now we are in the position to state the following uniqueness theorem.

Theorem 3.8. *Suppose $A = a + d$ satisfies conditions stated at the beginning of the section, i.e. $\lambda \leq (a_{ij}(t, x)) \leq \lambda^{-1}$ and $\|d\|_{L^\infty(BMO)} < \infty$. Let $\tau \geq 0$. Suppose $u(t, x) \in L^\infty(\tau, T; L^2(\mathbb{R}^n)) \cap L^2(\tau, T; H^1(\mathbb{R}^n))$, and satisfies*

$$\int_{\tau}^T \int_{\mathbb{R}^n} u(t, x) \frac{\partial}{\partial t} \varphi(t, x) dx dt = \int_{\tau}^T \int_{\mathbb{R}^n} \langle \nabla \varphi(t, x), A(t, x) \cdot \nabla u(t, x) \rangle dx dt \quad (3.9)$$

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

$$\frac{\partial u}{\partial t} \in L^2(\tau, T; H^{-1}(\mathbb{R}^n)). \quad (3.10)$$

Hence the following energy inequality holds:

$$\|u(T, \cdot)\|_{L^2}^2 + 2\lambda \int_{\tau}^T \int_{\mathbb{R}^n} |\nabla u(t, x)|^2 dx dt \leq \|u(\tau, \cdot)\|_{L^2}^2, \quad (3.11)$$

and the uniqueness of weak solutions holds for the initial problem of (3.1) in space $L^\infty(\tau, T; L^2(\mathbb{R}^n)) \cap L^2(\tau, T; H^1(\mathbb{R}^n))$.

Proof. Consider the linear functional

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

for $\psi \in H^1(\mathbb{R}^n)$. By the compensated compactness inequality (2.5) we have

$$F_t(\psi) \leq \left(\|a\|_{L^\infty([0, T] \times \mathbb{R}^n)} + C \|d\|_{L^\infty(\text{BMO})} \right) \|\nabla u(t, \cdot)\|_{L^2(\mathbb{R}^n)} \|\nabla \psi\|_{L^2(\mathbb{R}^n)} \quad (3.12)$$

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

$$F_t(\psi) = \int_{\mathbb{R}^n} (\nabla w(t, x) \cdot \nabla \psi(x) + w(t, x) \psi(x)) dx, \quad (3.13)$$

where

$$\|w(t, \cdot)\|_{H^1(\mathbb{R}^n)} \leq \left(\|a\|_{L^\infty([0, T] \times \mathbb{R}^n)} + C \|d\|_{L^\infty(\text{BMO})} \right) \|\nabla u(t, \cdot)\|_{L^2(\mathbb{R}^n)},$$

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

In terms of $w(t, x)$, (3.9) becomes

$$\int_{\tau}^T \int_{\mathbb{R}^n} u(t, x) \varphi(x) \eta'(t) dx dt = \int_{\tau}^T \int_{\mathbb{R}^n} (\nabla w(t, x) \cdot \nabla \varphi(x) + w(t, x) \varphi(x)) \eta(t) dx dt \quad (3.14)$$

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

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

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

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

we obtain

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

which implies that

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

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

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

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

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

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

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

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 $L^2(\tau, T; H^{-1}(\mathbb{R}^n))$ and its dual space $L^2(\tau, T; H^1(\mathbb{R}^n))$.

Since

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

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

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

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

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

Now, by combining with Lemma 3.7, we deduce that

$$\|u(T, \cdot)\|_{L^2}^2 - \|u(\tau, \cdot)\|_{L^2}^2 = 2 \left\langle \frac{\partial}{\partial t} u, u \right\rangle_{L^2(H^{-1}), L^2(H^1)} \leq -2\lambda \int_{\tau}^T \int_{\mathbb{R}^n} |\nabla u(t, x)|^2 dx dt,$$

which in turn yields the energy inequality (3.11). Other conclusions of the theorem follow easily. $\square$

Finally we prove the existence and uniqueness of the fundamental solution.

Theorem 3.9. *Suppose $(A_{ij}) = (a_{ij}) + (d_{ij})$, where a and d are symmetric and skew-symmetric parts of A respectively, is uniformly elliptic: $\lambda \leq a(t, x) \leq \lambda^{-1}$ in matrix sense for some constant $\lambda > 0$, and $\|d\|_{L^\infty(BMO)} < \infty$. Then there is a unique positive function $\Gamma(t, x; \tau, \xi)$ defined for $t > \tau \geq 0$ and $x, \xi \in \mathbb{R}^n$, which possesses the following properties.*

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

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

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

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

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

2) There is a constant $M > 0$ depending only on n , λ and $\|d\|_{L^\infty(\text{BMO})}$ such that

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

for all $t > \tau$.

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

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

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

$$\|\varepsilon \log(|x|)\|_{\text{BMO}} \leq \|d\|_{L^\infty(\text{BMO})}.$$

Define

$$U^{(m)}(x) = (-\varepsilon \log(|x|) + m) \wedge m \vee 0, \quad L^{(m)}(x) = (\varepsilon \log(|x|) - m) \wedge 0 \vee (-m), \quad (3.15)$$

which are compactly supported BMO functions with

$$\|U^{(m)}\|_{\text{BMO}} = \|L^{(m)}\|_{\text{BMO}} \leq C \|d\|_{L^\infty(\text{BMO})},$$

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

$$d^{(m)}(t, x) = d(t, x) \wedge U^{(m)}(x) \vee L^{(m)}(x). \quad (3.16)$$

By Proposition 2.3, we can mollify it to define $d_{\frac{1}{m}}^{(m)}$. Then, there is a C independent of d and m , such that $\|d_{\frac{1}{m}}^{(m)}\|_{L^\infty(\text{BMO})} \leq C\|d\|_{L^\infty(\text{BMO})}$. Each $d_{\frac{1}{m}}^{(m)}$ is smooth with compact support and $d_{\frac{1}{m}}^{(m)} \rightarrow d$ in $L_{loc}^p([0, T] \times \mathbb{R}^n)$ for any $1 \leq p < \infty$. For simplicity denote $d_{\frac{1}{m}}^{(m)}$ by d_m . Similarly a_m denotes the mollified approximation of a for $m = 1, 2, \dots$. $a_m(t, x)$ and $d_m(t, x)$ are smooth, bounded and have bounded derivatives of all orders, and $a_m \rightarrow a$ and $d_m \rightarrow d$ in $L_{loc}^p([0, T] \times \mathbb{R}^n)$ for every $p \in [1, \infty)$.

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

$$\|d_m\|_{L^\infty(\text{BMO})} \leq C\|d\|_{L^\infty(\text{BMO})}$$

for some constant depending only on the dimension n , thus there is a unique fundamental solution $\Gamma^m(t, x; \tau, \xi)$ which satisfies the Aronson estimate with the same constant. According to Theorem 1.4, $\Gamma^m(t, x; \tau, \xi)$ 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 $m = 1, 2, \dots$. 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 $\Gamma(t, x; \tau, \xi)$ for $t > \tau \geq 0$ and $x, \xi \in \mathbb{R}^n$. Clearly $\Gamma(t, x; \tau, \xi)$ still satisfies 1) and 2).

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

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

point-wisely. According to Lemma 3.1, u^m (actually $u^m(t, x)$ is Hölder continuous too in $t > \tau$ 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:

$$\|u^m(t, \cdot)\|_{L^2}^2 + \lambda \int_{\tau}^t \|\nabla u^m(s, \cdot)\|_{L^2}^2 \leq \|f\|_{L^2}^2, \quad (3.17)$$

which implies that $\{u^m\}$ is uniformly bounded in $L^2(\tau, T; H^1(\mathbb{R}^n))$. Hence there is a

sub-sequence which converges weakly, whose limit must be u and

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

Next we prove that u also satisfies the energy inequality (3.17) as u^m . For each m , we have

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

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

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

for any $\varphi \in C_0^\infty((\tau, T) \times \mathbb{R}^n)$. That is, u is a weak solution to the Cauchy problem of the parabolic equation (3.1) with initial data f . Now according to Lemma 3.8, $\frac{\partial u}{\partial t} \in L^2(\tau, T; H^{-1}(\mathbb{R}^n))$ and therefore

$$\|u(t, \cdot)\|_{L^2}^2 + \lambda \int_\tau^t \|\nabla u(s, \cdot)\|_{L^2}^2 ds \leq \|f\|_{L^2}^2. \quad (3.18)$$

The uniqueness of the fundamental solution Γ follows from the energy inequality easily. In fact, suppose there is another sub-sequence of Γ^m converges to $\tilde{\Gamma}$. Then $\tilde{u}(t, x) = \tilde{\Gamma}_{\tau, t} f(x)$ satisfies all the results above. Especially they both satisfy the energy inequality. Therefore $w = u - \tilde{u}$ is also a weak solution. By Theorem 3.8, we deduce that

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

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

3.3 Diffusion processes

Clearly the fundamental solution obtained above in Theorem 3.9 can be regarded as the transition probability of a diffusion process and hence we can construct a diffusion process which is unique in law. In this section, we will study in detail the diffusion processes corresponding to equation (1.9). In particular, we are interested in solutions to the SDE

$$dX_t = b(t, X_t)dt + dB_t \quad (3.19)$$

for its application to the Navier-Stokes equations. It is well known that for b which is Lipschitz continuous and has linear growth, SDE (3.19) has a unique strong solution for any initial data $x_0 \in \mathbb{R}^n$. For measurable drift b , this SDE has weak solution which is unique in law when $b \in L^l(0, T; L^q(\mathbb{R}^n))$ with $\gamma = \frac{2}{l} + \frac{n}{q} \leq 1$, $l \in [2, \infty)$ and $q \in (n, \infty]$. In Krylov and Röckner [42], it was shown that there exists a unique strong solution to (3.19) when $b \in L^l(0, T; L^q(\mathbb{R}^n))$ with $\frac{2}{l} + \frac{n}{q} < 1$. Later it was further proved in Fedrizzi and Flandoli [23] that the strong solution to SDE (3.19) is Hölder continuous and differentiable in space variable with $\nabla_x X \in L^2(\Omega \times [0, T])$. However, these results are still not known when b is critical, i.e. $\frac{2}{l} + \frac{n}{q} = 1$. The main difficulty can be observed from the Girsanov transform theorem, which requires that the exponential martingale

$$\exp \left\{ \int_0^T b(t, B_t) dB_t - \frac{1}{2} \int_0^T |b(t, B_t)|^2 dt \right\}$$

is a true martingale. By the Novikov condition, we can see a sufficient condition is that

$$\sup_{x \in \mathbb{R}^n} \mathbb{E} \left[\int_0^T |b(t, B_t)|^2 dt \right] \leq \infty,$$

where the supremum is taken over all initial data x of the Brownian motion B_t . By the transition probability of the Brownian motion, this condition is satisfied for $b \in L^l(0, T; L^q(\mathbb{R}^n))$ with $\frac{2}{l} + \frac{n}{q} < 1$.

With the divergence-free condition on b , we use another approach to work on the

existence of strong solutions and its differentiability. The approach is first used by Crippa and De Lellis [13] on ODEs, which is to estimate the difference between two processes with different initial data or different drift term b . The key tool in the estimate is inequality (3.22) below. To use this inequality, these two processes have to be controlled to stay in a finite ball, which is easier to obtain for ODEs. For SDEs, because of the diffusion term, the control of the process within a ball is much harder and most works impose boundedness or linear g-growth condition on b (see e.g. [25, 90]). In our work, we will use the Aronson estimate to control the processes to stay within a ball.

Proposition 3.10. *Suppose a diffusion process X_t has transition probability Γ which satisfies the Aronson estimate*

$$\frac{1}{C_0(t-\tau)^{n/2}} \exp \left[-\frac{C_0|x-\xi|^2}{t-\tau} \right] \leq \Gamma(t, x; \tau, \xi) \leq \frac{C_0}{(t-\tau)^{n/2}} \exp \left[-\frac{|x-\xi|^2}{C_0(t-\tau)} \right], \quad (3.20)$$

then

$$P \left(\sup_{0 \leq s \leq t} |X_s - x| \geq R \right) \leq 2C_0^{2+n} P(Z \geq \frac{R}{t}),$$

where x is the initial data of X_t and Z is of normal distribution $\mathcal{N}(0, I_n)$.

Proof. Without loss of generality, we assume $x = 0$ and set

$$M_t = \sup_{0 \leq s \leq t} |X_s|, \quad T_R = \min \{t : |X_t| \geq R\}.$$

Then we have

$$\begin{aligned} P(|X_t| \geq R) &= P(|X_t| \geq R, M_t \geq R) \\ &= P(|X_t| \geq R | M_t \geq R) P(M_t \geq R) \\ &= P(|X_t| \geq R | T_R \leq t) P(M_t \geq R), \end{aligned}$$

in which

$$\begin{aligned} P(|X_t| \geq R | T_R \leq t) &\geq \frac{1}{2} \int_{\mathbb{R}^n} \frac{1}{C_0(t - T_R)^{n/2}} \exp \left[-\frac{C_0|x|^2}{t - T_R} \right] dx \\ &= \frac{1}{2} C_0^{-1-\frac{n}{2}} \end{aligned}$$

by the lower bound. We plug this in to obtain $P(M_t \geq R) \leq 2C_0^{1+\frac{n}{2}} P(|X_t| \geq R)$. Now using the upper bound and we proved that $P(M_t \geq R) \leq 2C_0^{2+n} P(Z \geq \frac{R}{t})$. $\square$

Remark 3.11. This estimate is not optimal, but it is enough for the application below to obtain the integrability of the supremum process and also to control the path. The estimate (3.20) here is expected to be relaxed to supercritical condition on b .

Corollary 3.12. *Under the same assumptions as in Proposition 3.10, we have that*

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s - x|^p \right] < \infty$$

for any $1 \leq p < \infty$

Proof. The proof is a straightforward calculation using Proposition 3.10

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s - x|^p \right] &= \int_0^\infty P \left(\sup_{0 \leq s \leq t} |X_s - x|^p > y \right) dy \\ &\leq \int_0^\infty 2C_0^{2+n} P(Z \geq \frac{y^{\frac{1}{p}}}{t}) < \infty, \end{aligned}$$

which is essentially due to the exponential decay of the density function of the normal distribution. $\square$

In addition, we use the following feature that when b is smooth and divergence-free, the strong solution X_t preserves the Lebesgue measure in the sense that

$$P[\omega \in \Omega : |X_t(A, \omega)| = |A|] = 1 \quad (3.21)$$

where $X_t(A, \omega)$ is the image of any Borel set $A \in \mathbb{R}^n$ under the mapping $x \mapsto X_t(x, \omega)$.

Now we start to present the construction of the strong solutions based on the argument in [90]. First, we recall the following lemma in [13, Appendix, Lemma A.3].

Lemma 3.13. *Let $M_R f$ be the local maximal function of locally integrable function f defined as*

$$M_R f(x) = \sup_{0 < r < R} \frac{1}{|B_r|} \int_{B_r(x)} f(y) dy.$$

Suppose $f \in BV_{loc}(\mathbb{R}^n)$, then

$$|f(x) - f(y)| \leq C|x - y|(M_r|\nabla f|(x) + M_r|\nabla f|(y)) \quad (3.22)$$

for $x, y \in \mathbb{R}^n \setminus N$, where N is a negligible set in $\mathbb{R}^n$, $R = |x - y|$ and constant C depends only on the dimension n .

We denote Mf as the maximal function

$$Mf(x) = \sup_{0 < r < \infty} \frac{1}{|B_r|} \int_{B_r(x)} f(y) dy$$

and clearly inequality (3.22) is also satisfied with the maximal function on the right-hand side.

Lemma 3.14. *Suppose $X_t(x)$ and $\tilde{X}_t(x)$ are strong solutions to SDE (3.19) driven by the same Brownian motion, with initial data x and smooth drift b and $\tilde{b}$ respectively. Then*

$$\mathbb{E} \left[\int_{B_r} \log \left(\frac{\sup_{0 \leq s \leq t} |X_s(x) - \tilde{X}_s(x)|^2}{\theta^2} + 1 \right) dx \right] \leq C(\|\nabla b\|_{L_t^1 L_x^2} + \frac{1}{\theta} \|b - \tilde{b}\|_{L_t^1 L_x^2}),$$

where the constant C depends on r and n .

Proof. Consider

$$\frac{d}{dt} \log \left(\frac{|X_t(x) - \tilde{X}_t(x)|^2}{\theta^2} + 1 \right) \leq \frac{|X_t(x) - \tilde{X}_t(x)| |b(t, X_t(x)) - \tilde{b}(t, \tilde{X}_t(x))|}{|X_t(x) - \tilde{X}_t(x)|^2 + \theta^2}$$

$$\begin{aligned}
&\leq \frac{|b(t, X_t(x)) - b(t, \tilde{X}_t(x))|}{\sqrt{|X_t(x) - \tilde{X}_t(x)|^2 + \theta^2}} + \frac{|b(t, \tilde{X}_t(x)) - \tilde{b}(t, \tilde{X}_t(x))|}{\sqrt{|X_t(x) - \tilde{X}_t(x)|^2 + \theta^2}} \\
&= g_1(x) + g_2(x).
\end{aligned}$$

Integrate both sides on $B_r(0)$ and take expectation, then by (3.21) and Lemma 3.13 we have that

$$\begin{aligned}
\mathbb{E} \left[\int_{B_r} g_1(x) dx \right] &\leq \mathbb{E} \left[\int_{B_r} \frac{C|X_t(x) - \tilde{X}_t(x)| (M|\nabla b|(t, X_t(x)) + M|\nabla b|(t, \tilde{X}_t(x)))}{\sqrt{|X_t(x) - \tilde{X}_t(x)|^2 + \theta^2}} dx \right] \\
&\leq C \int_{\Omega} \int_{X_t(B_r, \omega)} M|\nabla b|(t, x) dx + \int_{\tilde{X}_t(B_r, \omega)} M|\nabla b|(t, x) dx dP(\omega) \\
&\leq C \int_{\Omega} 2|B_r|^{\frac{1}{2}} \|\nabla b\|_{L^2(\mathbb{R}^n)} dP(\omega) \\
&= 2C|B_r|^{\frac{1}{2}} \|\nabla b\|_{L^2(\mathbb{R}^n)}
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E} \left[\int_{B_r} g_2(x) dx \right] &\leq \frac{1}{\theta} \mathbb{E} \left[\int_{B_r} |b(t, \tilde{X}_t(x)) - \tilde{b}(t, \tilde{X}_t(x))| dx \right] \\
&\leq \frac{1}{\theta} \int_{\Omega} \int_{\tilde{X}_t(B_r, \omega)} |b - \tilde{b}|(t, x) dx dP(\omega) \\
&\leq \frac{1}{\theta} |B_r|^{\frac{1}{2}} \|b - \tilde{b}\|_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

Finally we take supremum in time t on $\log \left(\frac{|X_t(x) - \tilde{X}_t(x)|^2}{\theta^2} + 1 \right)$ and the proof is complete. $\square$

Now we are ready to prove our main result in this section.

Theorem 3.15. *Given a divergence-free vector field $b \in L^2(0, T; H^1(\mathbb{R}^n))$ such that there exists a sequence of divergence-free vector fields $b^{(n)} \in C([0, T], C_0^\infty(\mathbb{R}^n))$ converging to b in $L^2(0, T; H^1(\mathbb{R}^n))$, and the corresponding diffusion processes $(X_t^{(n)})$ to SDE (3.19) satisfies uniform Aronson estimate (3.20). Then we have that $(X_t^{(n)})$ is a Cauchy sequence in the function space $L^p(\Omega \times B_r; C([0, t]))$ for any $p \in [1, 2)$ and $r > 0$. Moreover, the limit*

X_t is the unique strong solution to

$$dX_t = b(t, X_t)dt + dB_t$$

in space $L^P(\Omega \times B_r; C([0, t]))$.

Proof. Step 1: We first prove that $(X_t^{(n)})$ is a Cauchy sequence. The estimate of

$$\mathbb{E} \left[\int_{B_r} \log \left(\frac{\sup_{0 \leq s \leq t} |X_s^{(n)}(x) - X_s^{(m)}(x)|^2}{\theta^2} + 1 \right) dx \right]$$

in Lemma 3.14 does not imply the estimate of

$$\mathbb{E} \left[\int_{B_r} \sup_{0 \leq s \leq t} |X_s^{(n)}(x) - X_s^{(m)}(x)|^2 dx \right],$$

and that is why we need Lemma 3.10 to estimate it when $\sup_{0 \leq s \leq t} |X_s^{(n)}(x) - X_s^{(m)}(x)|^2$ is large. Denote

$$O_{n,m}^R(\omega) = \left\{ x \in \mathbb{R}^n : \sup_{0 \leq s \leq t} |X_s^{(n)}(x, \omega)| < R, \sup_{0 \leq s \leq t} |X_s^{(m)}(x, \omega)| < R \right\}$$

and by Lemma 3.10 we have that for any fixed $r > 0$,

$$\sup_{x \in B_r} P(\omega : x \notin O_{n,m}^R(\omega)) \rightarrow 0, \quad \text{as } R \rightarrow \infty. \quad (3.23)$$

Set $S_t^{(n,m)}(x) = \sup_{0 \leq s \leq t} |X_s^{(n)}(x) - X_s^{(m)}(x)|^2$, then for any fixed $\delta > 0$ we have

$$\begin{aligned} P \left(\omega : \int_{B_r} S_t^{(n,m)}(x) dx \geq 2\delta \right) &\leq P \left(\omega : \int_{B_r \cap O_{n,m}^R(\omega)} S_t^{(n,m)}(x) dx \geq \delta \right) \\ &\quad + P \left(\omega : \int_{B_r \setminus O_{n,m}^R(\omega)} S_t^{(n,m)}(x) dx \geq \delta \right) \\ &= I_1^{(n,m)} + I_2^{(n,m)}. \end{aligned}$$

We first estimate the second term $I_2^{(n,m)}$

$$\begin{aligned}
I_2^{(n,m)} &\leq \frac{1}{\delta} \mathbb{E} \left[\int_{B_r \setminus O_{n,m}^R(\omega)} \sup_{0 \leq s \leq t} |X_s^{(n)}(x) - X_s^{(m)}(x)|^2 dx \right] \\
&\leq \frac{1}{\delta} \int_{B_r} \int_{\Omega} \sup_{0 \leq s \leq t} |X_s^{(n)}(x) - X_s^{(m)}(x)|^2 1_{(\omega: x \notin O_{n,m}^R(\omega))} dP(\omega) dx \\
&\leq \frac{1}{\delta} \int_{B_r} 2 \mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^{(n)}(x) - x|^4 + \sup_{0 \leq s \leq t} |X_s^{(m)}(x) - x|^4 \right]^{\frac{1}{2}} P(\omega: x \notin O_{n,m}^R(\omega))^{\frac{1}{2}} dx \\
&\leq \varepsilon
\end{aligned}$$

for large enough R by (3.23) and Corollary 3.12, and this estimate is independent of (n, m) .

This estimate also helps us to fix an R . To estimate $I_1^{(n,m)}$, we separate $B_r \cap O_{n,m}^R(\omega)$ into two parts

$$\{S_t^{(n,m)}(x) \geq \theta^2(e^{L^2} - 1)\}, \quad \{S_t^{(n,m)}(x) < \theta^2(e^{L^2} - 1)\}$$

such that if $\int_{B_r} \log \left(\frac{S_t^{(n,m)}(x)}{\theta^2} + 1 \right) dx \leq L$, we have that $|\{S_t^{(n,m)}(x) \geq \theta^2(e^{L^2} - 1)\}| \leq \frac{1}{L}$, which implies

$$\begin{aligned}
\int_{B_r \cap O_{n,m}^R(\omega)} S_t^{(n,m)}(x) dx &= \int_{B_r \cap O_{n,m}^R(\omega)} S_t^{(n,m)}(x) 1_{\{S_t^{(n,m)}(x) \geq \theta^2(e^{L^2} - 1)\}} dx \\
&\quad + \int_{B_r \cap O_{n,m}^R(\omega)} S_t^{(n,m)}(x) 1_{\{S_t^{(n,m)}(x) < \theta^2(e^{L^2} - 1)\}} dx \\
&\leq \theta^2(e^{L^2} - 1) |B_r| + 4R^2 \frac{1}{L}.
\end{aligned}$$

Now we set $\theta^{(n,m)} = \|b^{(n)} - b^{(m)}\|_{L_t^1 L_x^2}$ to obtain

$$\sup_{n,m} \mathbb{E} \left[\int_{B_r} \log \left(\frac{S_t^{(n,m)}}{(\theta^{(n,m)})^2} + 1 \right) dx \right] \leq C$$

by Lemma 3.14, which gives us that

$$P \left(\int_{B_r} \log \left(\frac{S_t^{(n,m)}}{(\theta^{(n,m)})^2} + 1 \right) dx \geq L \right) \leq \frac{C}{L}.$$

Thus for fixed $\delta > 0$, we can choose L large enough and (n, m) large enough ($\theta^{(n, m)}$ is small enough) such that $(\theta^{(n, m)})^2(e^{L^2} - 1)|B_r| + 4R^2 \frac{1}{L} < \delta$ and $\frac{C}{L} \leq \varepsilon$. Hence

$$P\left(\omega : \int_{B_r \cap O_{n, m}^R(\omega)} S_t^{(n, m)}(x) dx \geq \delta, \int_{B_r} \log\left(\frac{S_t^{(n, m)}(x)}{\theta^2} + 1\right) dx \leq L\right) = 0,$$

and

$$\begin{aligned} I_1^{(n, m)} &= P\left(\omega : \int_{B_r \cap O_{n, m}^R(\omega)} S_t^{(n, m)}(x) dx \geq \delta, \int_{B_r} \log\left(\frac{S_t^{(n, m)}(x)}{\theta^2} + 1\right) dx > L\right) \\ &\leq P\left(\omega : \int_{B_r} \log\left(\frac{S_t^{(n, m)}(x)}{\theta^2} + 1\right) dx > L\right) \leq \varepsilon. \end{aligned}$$

Now we proved that

$$P\left(\omega : \int_{B_r} S_t^{(n, m)}(x) dx \geq 2\delta\right) \rightarrow 0$$

as $n, m \rightarrow \infty$ for any fixed δ , i.e. convergence in probability. Recall that

$$\sup_{n, x \in B_r} \mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^{(n)}(x)|^2 \right] < \infty$$

by Corollary 3.12, which means that $\sup_{0 \leq s \leq t} |X_s^{(n)}(x)|^p$ is uniformly integrable for any $p \in [1, 2)$. Finally we can deduce that for any fixed $r > 0$, $(X_t^{(n)})$ is a Cauchy sequence in $L^p(\Omega \times B_r; C([0, t]))$ for any $p \in [1, 2)$ and we denote the limit as X_t .

Step 2: Now we verify that the limit X_t is a solution corresponding to b , i.e. if we define

$$Y_t(x) = x + \int_0^t b(s, X_s(x)) ds + \int_0^t dB_s,$$

then $X_t = Y_t$ in space $L^1(\Omega \times B_r; C([0, t]))$. Consider

$$\sup_{0 \leq s \leq t} |X_s^{(n)}(x) - Y_s(x)| \leq \int_0^t |b^{(n)}(s, X_s^{(n)}(x)) - b(s, X_s(x))| ds.$$

Then integrate both sides on B_r and take expectation, we have that

$$\begin{aligned}
& \mathbb{E} \left[\int_{B_r} \sup_{0 \leq s \leq t} |X_s^{(n)}(x) - Y_s(x)| dx \right] \\
& \leq \mathbb{E} \left[\int_{B_r} \int_0^t |b^{(n)}(s, X_s^{(n)}(x)) - b(s, X_s^{(n)}(x))| ds dx \right] \\
& \quad + \mathbb{E} \left[\int_{B_r} \int_0^t |b(s, X_s^{(n)}(x)) - b(s, X_s(x))| ds dx \right] \\
& = I_1 + I_2.
\end{aligned}$$

Again using (3.21), we have that $I_1 \leq |B_r|^{\frac{1}{2}} \|b - \tilde{b}\|_{L_t^1 L_x^2}$. For the second term, we will find another smooth b_ε such that $\|b_\varepsilon - b\|_{L_{t,x}^2} \leq \varepsilon$ and then separate I_2 into three parts

$$\begin{aligned}
I_2 & \leq \mathbb{E} \left[\int_{B_r} \int_0^t |b_\varepsilon(s, X_s^{(n)}(x)) - b_\varepsilon(s, X_s(x))| ds dx \right] \\
& \quad + \mathbb{E} \left[\int_{B_r} \int_0^t |b_\varepsilon(s, X_s(x)) - b(s, X_s(x))| ds dx \right] \\
& \quad + \mathbb{E} \left[\int_{B_r} \int_0^t |b_\varepsilon(s, X_s^{(n)}(x)) - b(s, X_s^{(n)}(x))| ds dx \right].
\end{aligned}$$

For the second and the third part, we will control them just as I_1 and the first term converges to 0 as $n \rightarrow \infty$ since $X_t^{(n)} \rightarrow X_t$ in $L^p(\Omega \times B_r; C([0, t]))$ for any $p \in [1, 2)$. Now we proved that $X_t^{(n)} \rightarrow Y_t$ in $L^1(\Omega \times B_r; C([0, t]))$ and concluded that $X_t = Y_t$.

Step 3: Finally we prove that the limit is unique. Suppose we have two solutions X_t and $\tilde{X}_t$ and we apply Lemma 3.14 to deduce that

$$\mathbb{E} \left[\int_{B_r} \log \left(\frac{\sup_{0 \leq s \leq t} |X_s(x) - \tilde{X}_s(x)|^2}{\theta^2} + 1 \right) dx \right] \leq C \|\nabla b\|_{L_t^1 L_x^2},$$

which is uniform for all $\theta > 0$. Hence we can take $\theta \rightarrow 0$ and now the proof is complete. $\square$

Remark 3.16. The solution X_t in the theorem above is in the space $L^p(\Omega \times B_r; C([0, t]))$, which means that we have a unique strong solution to the SDE (3.19) for almost every

initial data $x \in \mathbb{R}^n$ under the Lebesgue measure. It is worth mentioning that for ODEs, a unique solution is also obtained for almost every initial data $x \in \mathbb{R}^n$ in DiPerna and Lions [17] using the idea of renormalized solutions.

Recall the Aronson estimate obtained in Theorem 1.3, we have the following corollary.

Corollary 3.17. *For a divergence-free vector field $b \in L^2(0, T; H^1(\mathbb{R}^n)) \cap L^\infty(0, T; BMO^{-1})$, the SDE (3.19) has a unique strong solution for almost every initial data $x \in \mathbb{R}^n$ under the Lebesgue measure.*

Following similar argument in Lemma 3.14, we can deduce the approximately differentiability of the solution X_t obtained in Theorem 3.15.

Proposition 3.18. *Suppose X_t is the solution obtained in Theorem 3.15, then*

$$\mathbb{E} \left[\int_{B_r} \int_{B_r} \log \left(\frac{\sup_{0 \leq s \leq t} |X_s(x) - X_s(y)|^2}{\theta^2} + 1 \right) dx dy \right] \leq C |B_r|^{\frac{3}{2}} \|\nabla b\|_{L_t^1 L_x^2},$$

where constant C depends on r and n . Moreover, the solution X_t is P -a.s. approximately differentiable in the space variable x .

Proof. Under similar argument as in Lemma 3.14, we have

$$\frac{d}{dt} \log \left(\frac{|X_t(x) - X_t(y)|^2}{\theta^2} + 1 \right) \leq \frac{|X_t(x) - X_t(y)| |b(t, X_t(x)) - b(t, X_t(y))|}{|X_t(x) - X_t(y)|^2 + \theta^2}.$$

Integrate both sides on $B_r(0) \times B_r(0)$ and take expectation, then by (3.21) and Lemma 3.13 we have that the right-hand side is dominated by

$$\begin{aligned} & \mathbb{E} \left[\int_{B_r} \int_{B_r} \frac{C |X_t(x) - X_t(y)|^2 (M |\nabla b|(t, X_t(x)) + M |\nabla b|(t, X_t(y)))}{|X_t(x) - X_t(y)|^2 + \theta^2} dx dy \right] \\ & \leq C \int_{\Omega} \int_{B_r} \int_{X_t(B_r, \omega)} M |\nabla b|(t, x) dx dy + \int_{B_r} \int_{X_t(B_r, \omega)} M |\nabla b|(t, y) dy dx dP(\omega) \\ & \leq C \int_{\Omega} 2 |B_r|^{\frac{3}{2}} \|\nabla b\|_{L^2(\mathbb{R}^n)} dP(\omega) \\ & = 2C |B_r|^{\frac{3}{2}} \|\nabla b\|_{L^2(\mathbb{R}^n)}. \end{aligned}$$

Fix r , for any $\varepsilon > 0$, there exists $\Omega_\varepsilon \subset \Omega$ with $P(\Omega \setminus \Omega_\varepsilon) < \varepsilon$ such that for any $\omega \in \Omega_\varepsilon$

$$\int_{B_r} \int_{B_r} \log \left(\frac{|X_s(x) - X_s(y)|^2}{\theta^2} + 1 \right) dx dy \leq \frac{C|B_r|^{\frac{3}{2}} \|\nabla b\|_{L_t^1 L_x^2}}{1 - \varepsilon},$$

which again implies that there exists $K_\varepsilon \subset B_r$ with $|B_r \setminus K_\varepsilon| < \varepsilon$ such that for any $x \in K_\varepsilon$

$$\int_{B_r} \log \left(\frac{|X_s(x) - X_s(y)|^2}{\theta^2} + 1 \right) dy \leq \frac{C|B_r|^{\frac{3}{2}} \|\nabla b\|_{L_t^1 L_x^2}}{(1 - \varepsilon)^2}.$$

Now for any $x, y \in K_\varepsilon$, we set $\theta = |x - y|$ and consider

$$\begin{aligned} \log \left(\frac{|X_s(x) - X_s(y)|^2}{\theta^2} + 1 \right) &= \int_{B_\theta(x) \cap B_\theta(y)} \log \left(\frac{|X_s(x) - X_s(y)|^2}{\theta^2} + 1 \right) dz \\ &\leq C \int_{B_\theta(x)} \log \left(\frac{|X_s(x) - X_s(z)|^2}{\theta^2} + 1 \right) dz \\ &\quad + C \int_{B_\theta(y)} \log \left(\frac{|X_s(y) - X_s(z)|^2}{\theta^2} + 1 \right) dz \\ &\leq \frac{C|B_r|^{\frac{1}{2}} \|\nabla b\|_{L_t^1 L_x^2}}{(1 - \varepsilon)^2}. \end{aligned}$$

This implies that $X_t(x)$ is Lipschitz on K_ε and hence is approximately differentiable by Theorem 3.19 below. $\square$

Finally, to complete the proof above, we state the result below whose proof can be found in [22, Theorem 3.1.16].

Theorem 3.19. *For functions $f : D \rightarrow \mathbb{R}^n$, if there exists a sequence of subsets $D_n \subset D$ for $n \in \mathbb{Z}$ such that $D_{n-1} \subset D_n$ and $|D \setminus D_n| \rightarrow 0$, then f is approximately differentiable on Ω .*

Remark 3.20. We would like to emphasize here again that the key point in this section is the usage of the Aronson estimate to estimate the supremum process instead of the growth conditions previous used by other works. We only deal with a very special case here for divergence-free vector fields $b \in L^2(0, T; H^1(\mathbb{R}^n))$, which is the regularity of the weak solutions to the Navier-Stokes equations. Similar result can be obtained for $b \in L^1(0, T; W_{loc}^{1,p})$ with $p > 1$ and with bounded divergence as in [13, 90].

Chapter 4

Weak solutions and uniqueness: supercritical cases

In this chapter, we will discuss the cases when b is supercritical. The Aronson-type estimate is weaker under supercritical conditions, and all the regularity results including the Harnack inequality and Hölder continuity are unknown. Hence we will need to impose more conditions to achieve uniqueness and other properties of the weak solutions and related processes.

4.1 Tightness of the fundamental solutions

Given $b \in L^l(0, T; L^q(\mathbb{R}^n))$ with $\frac{2}{l} + \frac{n}{q} \in [1, 2)$ satisfying condition **(S)**, we can always find a sequence $b_m \rightarrow b$ in $L^l(0, T; L^q(\mathbb{R}^n))$ (when $l, q \neq \infty$) such that b_m are smooth, bounded with bounded derivatives of all orders and still satisfy condition **(S)**. Moreover, we can have $\|b_m\|_{L^l(0, T; L^q(\mathbb{R}^n))} \leq 2\|b\|_{L^l(0, T; L^q(\mathbb{R}^n))}$. For these b_m , the fundamental solution to **(1.9)**, denoted by Γ^m , are unique and have uniform upper bound as proved in Chapter 2. In this section, we show several tightness results of $\{\Gamma^m\}$ relying on the upper bound we

proved, i.e

$$\Gamma(t, x; \tau, \xi) \leq \begin{cases} \frac{C_1}{(t-\tau)^{n/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|x-\xi|^2}{t-\tau}\right)\right) & \frac{|x|^{\mu-2}}{t^{\mu-v-1}} < 1 \\ \frac{C_1}{(t-\tau)^{n/2}} \exp\left(-\frac{1}{C_2} \left(\frac{|x-\xi|^\mu}{(t-\tau)^v}\right)^{\frac{1}{\mu-1}}\right) & \frac{|x|^{\mu-2}}{t^{\mu-v-1}} \geq 1 \end{cases}$$

and the divergence-free condition on b . We first recall the definition of tightness.

Definition 4.1. (Tightness) Given a family of probability measures $\{P_i\}_{i \in I}$ on a metric space, if for every $\varepsilon > 0$, there is a compact set K such that $\sup_{i \in I} P_i(K) > 1 - \varepsilon$, then we call this family of measures tight.

Clearly for fixed (t, τ, ξ) , the family of fundamental solutions $\{\Gamma^m(t, x; \tau, \xi)\}$ decays exponentially in space variables x , which implies the tightness of $\{\Gamma^m(t, x; \tau, \xi)\}$. Actually, we have tightness for the finite dimensional distributions

$$\prod_{i=1}^n \Gamma^m(t_i, x_i; t_{i-1}, x_{i-1}) dx_0 \cdots dx_n$$

for fixed $s \leq t_0 < t_1 < \cdots < t_n$. This allows us to take $m \rightarrow \infty$ to obtain a measure Γ as limit and this measure actually has a density as shown in the following proposition.

Proposition 4.2. *Given a sequence of probability measures $\{P_n\}$ on $\mathbb{R}^n$ which have densities $\{f_n\}$ uniformly bounded from above by a continuous function h . Suppose h satisfies*

$$\lim_{R \rightarrow \infty} \int_{B(0, R)^c} h(x) dx = 0,$$

where $B(0, R)$ is the open ball in $\mathbb{R}^n$ centered at 0 with radius R . Then $\{P_n\}$ is weakly compact in the space of probability measures. Suppose we take a convergent sub-sequence, then its limit P has density f which is also bounded from above by h .

Proof. It is easy to see that $\{P_n\}$ is tight, which implies that it is weakly compact by Prohorov's theorem. So we just need to show that P has density f which is bounded by h . Firstly, we show that P is absolutely continuous with respect to the Lebesgue measure m .

Suppose $A \subset \mathbb{R}^n$ such that $m(A) = 0$, then there is a decreasing sequence of open sets $\{O_i\}$ containing A such that $\lim_{i \rightarrow \infty} m(O_i) = 0$. Therefore $\lim_{i \rightarrow \infty} P_n(O_i) \rightarrow 0$ uniformly for all P_n . By Portmanteau theorem [81, Theorem 1.1.1], we have $P(O_i) \leq \limsup_{n \rightarrow \infty} P_n(O_i)$, which implies that $\lim_{i \rightarrow \infty} P(O_i) = 0$ and hence $P(A) = 0$. So P has a density f by Radon–Nikodym’s theorem.

Next we show that this f is bounded by h . If not, we can find a bounded set A such that $m(A) > 0$ and $f > h$ a.e. on A . Since h is continuous, we can find an open set O small enough such that it contains A and $P(O) > \int_O h \geq P_n(O)$ for all n . Clearly this contradicts to that $P_n \rightarrow P$ weakly in measure. $\square$

The tightness allows us to construct $\Gamma(t, x; \tau, \xi)$ for Borel measurable a and b which satisfy (E), (S) and $b \in L^l(0, T; L^q(\mathbb{R}^n))$ with $\frac{2}{l} + \frac{n}{q} \in [1, 2)$. However, the weak convergence for measures is too weak to ensure the Chapman–Kolmogorov equation:

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

for the limit, which means that we do not have the convergence of measures on the path space. To study the convergence of measures on the path space, we turn to the tightness criteria by Meyer and Zheng [95, 57]. We need to emphasize that here a is the identity matrix.

Theorem 4.3. *Let X_t^m be a diffusion governed by operators $\frac{1}{2}\Delta + b_m(t, x) \cdot \nabla$, such that for some $p > 1$*

$$\sup_m \mathbb{E}^m \left[\int_0^T |b_m(X_s^m, s)|^p ds \right] \leq \infty. \quad (4.1)$$

Assuming that the sequence (X_0^m) is tight in $\mathbb{R}^n$, then the sequence (X_t^m) is tight. Moreover, X is a semimartingale with canonical decomposition

$$X_t = X_0 + \int_0^t H_s ds + B_t, \quad E \left[\int_0^T |H_s|^p ds \right] < \infty,$$

where B is the Brownian motion.

Recall that when b_m is smooth and divergence-free, the strong solution X_t^m preserves the Lebesgue measure in the sense that

$$P[\omega \in \Omega : |X_t^m(A, \omega)| = |A|] = 1$$

where $X_t^m(A, \omega)$ is the image of any Borel set $A \in \mathbb{R}^n$ under the mapping $x \mapsto X_t^m(x, \omega)$.

If the initial data X_0 has density $\mu_0 \in L^\infty(\mathbb{R}^n)$, then we have

$$\begin{aligned} \mathbb{E}^m \left[\int_0^T |b_m(X_s^m, s)|^p ds \right] &= \int_0^T \int_{\mathbb{R}^n} \mu_0(x) \mathbb{E}^m [|b_m(X_s^m(x), s)|^p] dx ds \\ &= \int_0^T \mathbb{E}^m \left[\int_{\mathbb{R}^n} \mu_0(x) |b_m(X_s^m(x), s)|^p dx \right] ds \\ &\leq \|b_m\|_{L_{t,x}^p}^p \|\mu_0\|_{L^\infty} \end{aligned}$$

for any $p > 1$. By Meyer-Zheng's criteria, we have the tightness of the family of approximation processes (X_t^m) with bounded initial distribution. However, the uniqueness of the limit X_t and its relation with b are not clear here. Motivated by this, we study the uniqueness and the Chapman–Kolmogorov equation in the following sections.

4.2 Uniqueness with time-homogeneous coefficient

In the supercritical case, we do not have the Chapman–Kolmogorov equation through the tightness argument above, so in this section we look at a special case when the coefficients are time-homogeneous.

In order to establish the existence and uniqueness of a Markov semi-group associated with parabolic equation (1.27), which also defines the unique weak solution, we use an idea from [98]. For $b \in L^2(\mathbb{R}^n) \cap L^q(\mathbb{R}^n)$ with $q > \frac{n}{2}$, there are divergence-free vector fields $b_k \in C_0^\infty(\mathbb{R}^n)$ for $k = 1, 2, \dots$ such that $b_k \rightarrow b$ in $L^2(\mathbb{R}^n) \cap L^q(\mathbb{R}^n)$. For the existence of such an approximation sequence to divergence-free vector fields, see Section 1.5 in [73]. Throughout this section, L denotes the elliptic operator $\operatorname{div}(a \cdot \nabla) - b \cdot \nabla$, and its adjoint

operator is

$$L^* = \operatorname{div}(a \cdot \nabla) + b \cdot \nabla$$

as b is divergence-free. The fact that the dual operator has the same form will be of great importance to our arguments in what follows.

Recall that we proved the existence of weak solutions in Theorem 1.2 using an approximation argument. We call such solutions the approximation solutions. Next, we show that every weak solution is an approximation solution in a weaker sense. This result follows from a similar argument in [88].

Proposition 4.4. *Suppose $b \in L^2(\mathbb{R}^n)$ and $b_k \in C_0^\infty(\mathbb{R}^n)$ are divergence-free such that $b_k \rightarrow b$ in $L^2(\mathbb{R}^n)$. Let u and $\{u_k\}$ be the weak solutions to (1.27) on $[0, T] \times \mathbb{R}^n$ with initial data u_0 , and drifts b and $\{b_k\}$ respectively. Then u is the $L^\infty(0, T; L^1(\mathbb{R}^n))$ limit of functions $\{u_k\}$.*

Proof. Choose a sequence $b_k \rightarrow b$ in $L^2(\mathbb{R}^n)$. Consider the Cauchy problem

$$\partial_t u_k - \operatorname{div}(a \cdot \nabla u_k) + b_k \cdot \nabla u_k = 0$$

with initial data $u_k(x, 0) = u(x, 0) = u_0(x)$. Clearly $u_k - u$ is a weak solution to

$$\partial_t(u_k - u) - \operatorname{div}(a \cdot \nabla(u_k - u)) + b_k \cdot \nabla(u_k - u) = (b - b_k) \cdot \nabla u$$

with 0 as the initial value. By assumption, $\|(b - b_k) \cdot \nabla u\|_{L^2(0, T; L^1(\mathbb{R}^n))} \rightarrow 0$ as $k \rightarrow \infty$.

Since $b_k \in C_0^\infty(\mathbb{R}^n)$, we have a representation given by

$$(u_k - u)(t, x) = \int_0^t \int_{\mathbb{R}^n} \Gamma_k(t - \tau, x, \xi) (b - b_k) \cdot \nabla u(\xi, \tau) d\xi d\tau,$$

where Γ_k is the fundamental solution corresponding to b_k on $\mathbb{R}^n$. Then $\Gamma_k^*(t, \xi, x) := \Gamma_k(t, x, \xi)$ is the fundamental solution to $(\partial_t - L_k^*)u = 0$, which is of the same form as the

original equation (1.9) up to a sign on the drift. Hence

$$\int_{\mathbb{R}^n} \Gamma_k(t - \tau, x, \xi) dx = 1 \quad (4.2)$$

for any fixed (t, τ, ξ) . This implies that

$$\int_{\mathbb{R}^n} |u_k - u|(t, x) dx \leq \int_0^t \int_{\mathbb{R}^n} |b - b_k| |\nabla u| d\xi d\tau \rightarrow 0$$

and the proof is done. $\square$

The proposition above implies that any weak solution is an approximation solution. Here the divergence-free condition is the key to having the dual operator being conservative to obtain (4.2).

Now we start to prove our main result Theorem 1.9. The idea is to construct a unique approximation Markov semi-group corresponding to generator $L = \operatorname{div}(a \cdot \nabla) - b \cdot \nabla$. Since a is only Borel measurable, the generator L is not well-defined as a differential operator. Hence we will construct L in the following, while we still use formal expression $L = \operatorname{div}(a \cdot \nabla) - b \cdot \nabla$, if no confusion may arise, for simplicity of notations. We start with the bi-linear form

$$\mathcal{E}(u, v) = \int_{\mathbb{R}^n} \langle \nabla u, a \cdot \nabla v \rangle + (b \cdot \nabla u) v dx.$$

Naturally we consider the elliptic problem and its weak solutions. The approach is standard in literature.

Definition 4.5. Let (a, b) satisfies (E), (S) and $b \in L^2(\mathbb{R}^n)$. For $f \in L^2(\mathbb{R}^n)$, if there exists a $u \in H^1(\mathbb{R}^n)$ such that

$$\int_{\mathbb{R}^n} \langle \nabla u, a \cdot \nabla \varphi \rangle + (b \cdot \nabla u) \varphi + \alpha u \varphi dx = \int_{\mathbb{R}^n} f \varphi dx$$

for all $\varphi \in C_0^\infty(\mathbb{R}^n)$, we call u a weak solution to the elliptic problem $(\alpha - L, f)$, where $\alpha \geq 0$.

For $b \in C_0^\infty(\mathbb{R}^n)$, the bi-linear form is actually a Dirichlet form. We recall the following result on Dirichlet forms in [56, Chapter 1].

Theorem 4.6. *Let (a, b) satisfies **(E)**, **(S)** and $b \in C_0^\infty(\mathbb{R}^n)$. Then $(\mathcal{E}, H^1(\mathbb{R}^n))$, where*

$$\mathcal{E}(u, v) = \int_{\mathbb{R}^n} \langle \nabla u, a \cdot \nabla v \rangle + (b \cdot \nabla u) v \, dx$$

for $u, v \in H^1(\mathbb{R}^n)$, is a (non-symmetric) Dirichlet form. We still use L together with its domain $D(L)$ to denote the generator associated with the Dirichlet form $(\mathcal{E}, H^1(\mathbb{R}^n))$. The resolvent $R_\alpha = (\alpha - L)^{-1}$ for $\alpha > 0$ is a bounded linear operator from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$ with $\|(\alpha - L)^{-1}\|_{L^2 \rightarrow L^2} \leq \alpha^{-1}$, and it satisfies

$$\mathcal{E}(R_\alpha f, v) + \alpha(R_\alpha f, v) = (f, v). \quad (4.3)$$

Thus for $b \in C_0^\infty(\mathbb{R}^n)$, $\operatorname{div}(a \cdot \nabla) - b \cdot \nabla$ is understood as the generator L defined as in Theorem 4.6 above. Clearly, for any $f \in L^2(\mathbb{R}^n)$, $(\alpha - L)^{-1} f$ is the unique weak solution to $(\alpha - L, f)$. We can take $v = (\alpha - L)^{-1} f$ and derive that

$$\|(\alpha - L)^{-1} f\|_{H^1} \leq \frac{1}{\min\{\lambda, \alpha\}} \|f\|_{L^2} \quad \text{and} \quad \|(\alpha - L)^{-1} f\|_{L^2} \leq \frac{1}{\alpha} \|f\|_{L^2} \quad (4.4)$$

for all $\alpha > 0$ and $f \in L^2(\mathbb{R}^n)$. The following estimate on R_α , which follows from [98], plays an important role in proving our main result.

Lemma 4.7. *Suppose $b \in C_0^\infty(\mathbb{R}^n)$ and L as in Theorem 4.6, set $u = (1 - L)^{-1} f$ for $f \in C_0^\infty(\mathbb{R}^n)$. Then for $n \geq 3$, we have*

$$\int_{\mathbb{R}^n} [\ln(|x|^2 + e)]^{2\gamma} u^2(x) \, dx \leq C_0 \int_{\mathbb{R}^n} [\ln(|x|^2 + e)]^{2\gamma} f^2(x) \, dx$$

with sufficiently small positive γ and constant C_0 depending only on n, λ, γ and $\|b\|_{L^q(\mathbb{R}^n)}$ with $q > \frac{n}{2}$.

Proof. Let $\psi = \gamma \psi_0$, $\psi_0 = \ln \ln(|x|^2 + e)$, for $\gamma > 0$, and consider the operator $L_\psi =$

$e^\psi L e^{-\psi}$. For $v = e^\psi u$, we have $L_\psi v - v = g = e^\psi f$ and

$$\int_{\mathbb{R}^n} -\langle \nabla(e^\psi v), a \cdot \nabla(e^{-\psi} v) \rangle - b \cdot \nabla(e^{-\psi} v) e^\psi v - v^2 \, dx = \int_{\mathbb{R}^n} g v \, dx.$$

It follows, together with (E) and (S), that

$$\int_{\mathbb{R}^n} \lambda |\nabla v|^2 - \frac{1}{\lambda} \gamma^2 |\nabla \psi_0|^2 v^2 - \gamma (b \cdot \nabla \psi_0) v^2 + v^2 \, dx \leq - \int_{\mathbb{R}^n} g v \, dx.$$

Notice that

$$|\nabla \psi_0| \leq \frac{2|x|}{(|x|^2 + 1) \ln(|x|^2 + e)},$$

which is bounded. Hence we have

$$\begin{aligned} \int_{\mathbb{R}^n} (b \cdot \nabla \psi_0) v^2 \, dx &\leq C \|b\|_{L^q} \|\nabla \psi_0\|_{L^\infty} \|v\|_{L^2}^{1-\theta} \|\nabla v\|_{L^2}^{1+\theta} \\ &\leq C \|b\|_{L^q} \|\nabla \psi_0\|_{L^\infty} C(\theta) (\|v\|_{L^2}^2 + \|\nabla v\|_{L^2}^2) \end{aligned}$$

where $\theta = \frac{n}{q} - 1$ and C depends on n, q . Now we can take γ small enough such that $\|v\|_{L^2} \leq C_0 \|g\|_{L^2}$ and the proof is complete. $\square$

Given a divergence-free $b \in L^q(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and a sequence of smooth functions $b_k \rightarrow b$ in $L^q(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, this lemma implies that for each fixed $f \in C_0^\infty(\mathbb{R}^n)$,

$$\lim_{r \rightarrow \infty} \int_{|x| > r} |(1 - L_k)^{-1} f|^2 = 0$$

uniformly in k . Using them, we prove the compactness of resolvent operators $\{(1 - L_k)^{-1}\}$ as follows.

Lemma 4.8. *Given divergence-free $b \in L^q(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, smooth approximations $b_k \rightarrow b$ in $L^q(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, and $f \in L^2(\mathbb{R}^n)$, the sequence $\{(1 - L_k)^{-1} f\}$ is strongly compact in $L^2(\mathbb{R}^n)$ and weakly compact in $H^1(\mathbb{R}^n)$.*

Proof. Since

$$\|(1 - L_k)^{-1}f\|_{H^1} \leq \frac{1}{\min\{\lambda, 1\}} \|f\|_{L^2} \quad (4.5)$$

the sequence $\{(1 - L_k)^{-1}f\}$ is weakly compact in $H^1(\mathbb{R}^n)$. To prove the strong compactness in $L^2(\mathbb{R}^n)$, recall that we have proved $\|(1 - L_k)^{-1}\|_{L^2 \rightarrow L^2} \leq 1$ for all k . Since the convergence of bounded linear operators is determined by its convergence on a dense subset (see Theorem 6 in [45, Ch15]), it is sufficient to establish the compactness of $\{(1 - L_k)^{-1}f\}$ for f in a dense subset of $L^2(\mathbb{R}^n)$. For $f \in C_0^\infty(\mathbb{R}^n)$, by Lemma 4.7 and the inequality (4.5), the compactness of $\{(1 - L_k)^{-1}f\}$ in $L^2(\mathbb{R}^n)$ follows from the Fréchet–Kolmogorov theorem [87, Chapter X, Section 1]. $\square$

The previous lemma allows us to take limit as $k \rightarrow \infty$ and to define the generator L for singular b .

Lemma 4.9. *Given L_k defined as in Theorem 4.6 corresponding to b_k which converges to b in $L^q(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, after a possible selection of a sub-sequence (denoted as L_k again), there exists a closed operator L defined on a dense subset of $L^2(\mathbb{R}^n)$ such that $\|(\alpha - L)^{-1}\|_{L^2 \rightarrow L^2} \leq \alpha^{-1}$ for all $\alpha > 0$ and*

$$(\alpha - L_k)^{-1}f \rightarrow (\alpha - L)^{-1}f \quad \text{in } L^2(\mathbb{R}^n)$$

as $k \rightarrow \infty$, for all $\alpha > 0$ and $f \in L^2(\mathbb{R}^n)$.

Proof. We first consider the case when $\alpha = 1$. We apply Lemma 4.8 to f in a countable dense subset of $L^2(\mathbb{R}^n)$, by Theorem 6 in [45, Ch15] and Cantor’s diagonal argument, we can find a sub-sequence of $(1 - L_k)^{-1}$ that converges strongly. We still denote the sub-sequence as $(1 - L_k)^{-1}$ and denote its limit as S , i.e.

$$(1 - L_k)^{-1}f \rightarrow Sf$$

strongly in $L^2(\mathbb{R}^n)$ for $f \in L^2(\mathbb{R}^n)$. Since $(1 - L_k)^{-1}f$ is weakly compact in H^1 , it also converges to Sf weakly in H^1 . It is easy to see that the limit Sf is a weak solution to

$(1 - L, f)$. Since S is a bounded linear operator from $L^2(\mathbb{R}^n)$ to itself, we can define its adjoint operator S^* by $\langle Sf, g \rangle = \langle f, S^*g \rangle$ for all $f, g \in L^2(\mathbb{R}^n)$. We already know that

$$\lim_{k \rightarrow \infty} \langle (1 - L_k)^{-1} f, g \rangle = \langle Sf, g \rangle$$

for all $f, g \in L^2(\mathbb{R}^n)$ and

$$\langle (1 - L_k)^{-1} f, g \rangle = \langle f, (1 - L_k^*)^{-1} g \rangle.$$

Hence we can see that S^*g is a weak solution to $(1 - L^*, g)$. Proposition 4.10 implies that both S and S^* have kernels $K(S) = K(S^*) = 0$ and hence they have dense range in $L^2(\mathbb{R}^n)$ due to the equality that $K(S^*) = R(S)^\perp$. Now we can define $L = 1 - S^{-1}$, which has dense domain $D(L)$ and $D(L) \subset H^1$. Since $S = R_1 = (1 - L)^{-1}$ is the resolvent, we also have that L is a closed operator. Clearly, for each $u \in D(L)$, it is the weak solution to $(-L, -Lu)$. Hence $(\alpha - L)^{-1}f$ is a weak solution to $(\alpha - L, f)$ for f in the range of $(\alpha - L)$, i.e. $f \in R(\alpha - L)$. From last theorem, we already know that $R_\alpha = (\alpha - L)^{-1}$ is bounded linear operator. We therefore need to show that $R(\alpha - L) = L^2(\mathbb{R}^n)$. We can show that for each $f \in L^2(\mathbb{R}^n)$, there is a unique weak solution $u \in D(L)$ to $(\alpha - L, f)$. This is because for each $u \in D(L)$, $f = (1 - L)u + (\alpha - 1)u \in L^2$ and u is the weak solution to $(\alpha - L, f)$. Finally we can apply Theorem 1.3 in [39, Ch.8] to the approximation sequence L_k to obtain that

$$(\alpha - L_k)^{-1}f \rightarrow (\alpha - L)^{-1}f \quad \text{in } L^2(\mathbb{R}^n)$$

as $k \rightarrow \infty$, for all $\alpha > 0$ and $f \in L^2(\mathbb{R}^n)$. □

Let u be the limit of $\{(1 - L_k)^{-1}f\}$ weakly in $H^1(\mathbb{R}^n)$. Then it is easy to check that u is a weak solution to $(1 - L, f)$. Next we show that for $b \in L^q(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, there is a unique S defined as in Lemma 4.9. The uniqueness of S implies that the definition of L is independent of the choice of the convergent sub-sequence.

Proposition 4.10. *Suppose (a, b) satisfies conditions (E) and (S). For any $f \in L^2$, there*

exists a unique weak solution $u \in H^1$ to the elliptic problem $(\alpha - L, f)$ for $n \geq 3$, $b \in L^q(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and $\alpha > 0$.

Proof. We already showed the existence of weak solution by an approximation approach. Given a weak solution u where $f = 0$, actually we can take a test function as $h = \bar{u}\varphi$ with $\bar{u} = u \wedge N \vee (-N)$ and $\varphi \in C_0^\infty$, because $b \cdot \bar{u} \in L^2$. We let

$$\varphi_r = \begin{cases} 1 & |x| \leq \frac{r}{2} \\ 0 & |x| \geq r \end{cases}, \quad |\nabla \varphi| \leq \frac{4}{r}$$

for any $r > 0$ and $0 \leq \varphi_r \leq 1$. Then we have

$$\int_{\mathbb{R}^n} \langle \nabla u, a \cdot \nabla(\bar{u}\varphi_r) \rangle + b \cdot \nabla u(\bar{u}\varphi_r) + \alpha u(\bar{u}\varphi_r) dx = 0.$$

Because $\bar{u}\varphi_r \rightarrow \bar{u}$ in $H^1(\mathbb{R}^n)$ and almost everywhere, by taking $r \rightarrow \infty$, we obtain that

$$\int_{\mathbb{R}^n} \langle \nabla u, a \cdot \nabla(\bar{u}) \rangle + b \cdot \nabla u(\bar{u}) + \alpha u(\bar{u}) dx = 0.$$

Next we consider the second term in the equation above. Since $\int_{\mathbb{R}^n} b \cdot \nabla \bar{u} \bar{u} dx = 0$, we have

$$\begin{aligned} \int_{\mathbb{R}^n} b \cdot \nabla u \bar{u} dx &= \int_{\mathbb{R}^n} b \cdot (\nabla u - \nabla \bar{u}) \bar{u} dx \\ &= N \int_{\{u > N\}} b \cdot (\nabla u - \nabla \bar{u}) dx - N \int_{\{u < -N\}} b \cdot (\nabla u - \nabla \bar{u}) dx = 0, \end{aligned}$$

and therefore

$$\int_{\mathbb{R}^n} \langle \nabla u, a \cdot \nabla u \rangle + \alpha u^2 dx = 0$$

by taking $N \rightarrow \infty$. Now we obtained that $u = 0$ and the proof is complete. $\square$

Finally, to prove the representation (1.29), we also need the convergence of the fundamental solution.

Now we are in a position to complete the proof of Theorem 1.9.

Proof. By the fundamental approximation theorem of semi-groups in [39, Cp 9, Theorem 2.16], the convergence of resolvents in Theorem 4.9 implies that $e^{tL_k} \rightarrow e^{tL}$ as bounded linear operators from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$ and are uniform for t in any finite interval $[0, T]$. Further, Proposition 4.4 yields that e^{tL} is the unique semi-group which generates the unique weak solution. Let $\Gamma_k(t, x, y)$ be the fundamental solution to $(\partial_t - L_k)u = 0$. Then

$$u_k(t, x) = \int_{\mathbb{R}^n} \Gamma_k(t, x, y) u_0(y) dy = e^{tL_k} u_0$$

for any $u_0 \in L^2(\mathbb{R}^n)$ and $k = 1, 2, \dots$. By Corollary 2.15 and Proposition 4.2, we have that for each fixed (t, x) (and (t, y)), the family of transition probabilities $\{\Gamma_k(t, x, y) dy\}$ (and also the family $\{\Gamma_k(t, x, y) dx\}$) is tight and hence converges weakly in measure to some $\Gamma(t, x, y) dy$ which has the same upper bound as that of $\Gamma_k(t, x, y)$. Define

$$u(t, x) = \int_{\mathbb{R}^n} \Gamma(t, x, y) u_0(y) dy$$

for $u_0 \in C_0^\infty(\mathbb{R}^n)$, then $u_k(t, x) \rightarrow u(t, x)$ by the weak convergence of measure. As we have proved above that $u_k \rightarrow e^{tL} u_0$ in $L^2(\mathbb{R}^n)$, so $u = e^{tL} u_0$ in $L^2(\mathbb{R}^n)$. Since $C_0^\infty(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, we can extend it to conclude that operator e^{tL} has a kernel $\Gamma(t, x, y)$. $\square$

4.3 Renormalized solutions

In this section, we will introduce the concept of renormalized solutions [17]. This section is a brief review of the uniqueness theorems of the renormalized solutions as in Le Bris and Lions [46, 47], because this theory allows us to define a unique solution under supercritical conditions.

For simplicity, we assume that the diffusion coefficient a is constant here, which can be generalized (see e.g. [46, 47]). Also the divergence-free condition on the drift b can be relaxed to bounded convergence condition. Motivated for applications to the incompress-

ible fluid, we will stick to the special case here. Suppose u is a weak solution, then we have

$$\partial_t u - \operatorname{div}(a \cdot \nabla u) + b \cdot \nabla u = 0 \quad (4.6)$$

in the distribution sense. Let ρ be a non-negative smooth function supported on the unit ball $B(0, 1)$ and $\rho_\varepsilon(x) = \frac{1}{\varepsilon^n} \rho(\frac{x}{\varepsilon})$. If we mollify the equation in space by applying ρ_ε to each term and denote $u_\varepsilon = \rho_\varepsilon * u$, then we have that

$$\partial_t u_\varepsilon - \operatorname{div}(a \cdot \nabla u_\varepsilon) + b \cdot \nabla u_\varepsilon = b \cdot \nabla u_\varepsilon - \rho_\varepsilon * (b \cdot \nabla u).$$

We notice that the right-hand side is a commutator

$$[\rho_\varepsilon, b \cdot \nabla](u) = b \cdot \nabla(\rho_\varepsilon * u) - \rho_\varepsilon * (b \cdot \nabla u).$$

It is easy to check that when b and u are in $C_0^\infty((0, T) \times \mathbb{R}^n)$, commutator $[\rho_\varepsilon, b \cdot \nabla](u) \rightarrow 0$ in $L^\infty((0, T) \times \mathbb{R}^n)$ as $\varepsilon \rightarrow 0$. So we want to have the convergence with more singular b and u . This is the key observation made first in DiPerna and Lions [17] for defining renormalized solutions to transport equations and then adapted to parabolic equations in Le Bris and Lions [46, 47].

Lemma 4.11. *Suppose $b \in L^2(0, T; L^2(\mathbb{R}^n))$, $u \in L^2(0, T; H^1(\mathbb{R}^n))$ and ρ_ε are mollifiers on space variables, then the commutator $[\rho_\varepsilon, b \cdot \nabla](u) \rightarrow 0$ in $L^1(0, T; L^1(\mathbb{R}^n))$.*

Proof. The idea of the proof here is similar to proving $\|\rho_\varepsilon * f - f\|_{L^p} \rightarrow 0$ for $f \in L^p$, which is to consider approximations $f_m \rightarrow f$ and compare $\rho_\varepsilon * f_m - \rho_\varepsilon * f$. As mentioned above, this lemma is true for smooth and compactly supported b and u . Let $b_m, u_m \in C_0^\infty((0, T) \times \mathbb{R}^n)$ with $b_m \rightarrow b$ in $L^2(0, T; L^2(\mathbb{R}^n))$ and $u_m \rightarrow u$ in $L^2(0, T; H^1(\mathbb{R}^n))$. Then we have

$$\begin{aligned} [\rho_\varepsilon, b_m \cdot \nabla](u_m)(x) - [\rho_\varepsilon, b \cdot \nabla](u)(x) &= \int_{\mathbb{R}^n} (b_m(y) - b_m(x)) \cdot \nabla u_m(y) \rho_\varepsilon(x - y) dy \\ &\quad - \int_{\mathbb{R}^n} (b(y) - b(x)) \cdot \nabla u(y) \rho_\varepsilon(x - y) dy \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}^n} b_m(y) \cdot \nabla(u_m(y) - u(y)) \rho_\varepsilon(x - y) dy \\
&\quad + \int_{\mathbb{R}^n} (b_m(y) - b(y)) \cdot \nabla u(y) \rho_\varepsilon(x - y) dy \\
&\quad - b_m(x) \cdot \int_{\mathbb{R}^n} \nabla(u_m(y) - u(y)) \rho_\varepsilon(x - y) dy \\
&\quad - (b_m(x) - b(x)) \cdot \int_{\mathbb{R}^n} \nabla u(y) \rho_\varepsilon(x - y) dy \\
&= I_1 + I_2 - I_3 - I_4.
\end{aligned}$$

Now we estimate each term. Firstly, by Fubini's theorem

$$\begin{aligned}
\int_0^T \int_{\mathbb{R}^n} |I_1(x)| dx dt &= \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |b_m(y) \cdot \nabla(u_m(y) - u(y))| \rho_\varepsilon(x - y) dy dx dt \\
&= \int_0^T \int_{\mathbb{R}^n} |b_m(y) \cdot \nabla(u_m(y) - u(y))| \int_{\mathbb{R}^n} \rho_\varepsilon(x - y) dx dy dt \\
&= \int_0^T \int_{\mathbb{R}^n} |b_m(y) \cdot \nabla(u_m(y) - u(y))| dy dt \\
&\leq \|b_m\|_{L^2_{t,x}} \|\nabla(u_m(y) - u(y))\|_{L^2_{t,x}} \rightarrow 0,
\end{aligned}$$

and similar estimate applies to I_2 . Now for I_3 , we have

$$\begin{aligned}
\int_0^T \int_{\mathbb{R}^n} |I_3(x)| dx dt &= \int_0^T \int_{\mathbb{R}^n} |\nabla(u_m(y) - u(y))| \cdot \int_{\mathbb{R}^n} |b_m(x)| \rho_\varepsilon(x - y) dx dy dt \\
&\leq \|b_m\|_{L^2_{t,x}} \|\nabla(u_m(y) - u(y))\|_{L^2_{t,x}} \rightarrow 0,
\end{aligned}$$

and again similar estimate applies to I_4 . Since $[\rho_\varepsilon, b_m \cdot \nabla](u_m) \rightarrow 0$ in $L^1(0, T; L^1(\mathbb{R}^n))$, we now deduce that $[\rho_\varepsilon, b \cdot \nabla](u) \rightarrow 0$ in $L^1(0, T; L^1(\mathbb{R}^n))$. $\square$

Now we are ready to prove the uniqueness of weak solution to bounded initial data.

Theorem 4.12. Suppose a, b satisfy conditions (E), (S) and a is constant in equation (1.9).

If we assume that $b \in L^2(0, T; L^2(\mathbb{R}^n))$, then for any initial data $u_0 \in L^2 \cap L^\infty$, there is a unique weak solution $u \in L^2(0, T; H^1) \cap L^\infty(0, T; L^\infty)$ corresponding to u_0 . In addition,

the weak solution satisfies the energy inequality

$$\|u(T)\|_{L^2} + \lambda \|\nabla u\|_{L^2} \leq \|u(0)\|_{L^2}.$$

Proof. The idea is to prove the energy inequality for all weak solutions. Suppose that u is a weak solution and $u_\varepsilon = \rho_\varepsilon * u$ satisfies

$$\partial_t u_\varepsilon - \operatorname{div}(a \cdot \nabla u_\varepsilon) + b \cdot \nabla u_\varepsilon = b \cdot \nabla u_\varepsilon - \rho_\varepsilon * (b \cdot \nabla u).$$

Then we multiply by u_ε and integrate it on $[0, T] \times \mathbb{R}^n$ to obtain that

$$\begin{aligned} \|u_\varepsilon(T)\|_{L^2} - \|u_\varepsilon(0)\|_{L^2} + \int_0^T \int_{\mathbb{R}^n} \langle \nabla u_\varepsilon, a \cdot \nabla u_\varepsilon \rangle + b \cdot \nabla u_\varepsilon u_\varepsilon dx dt \\ = \int_0^T \int_{\mathbb{R}^n} [\rho_\varepsilon, b \cdot \nabla](u) u_\varepsilon dx dt, \end{aligned}$$

in which $\int_0^T \int_{\mathbb{R}^n} b \cdot \nabla u_\varepsilon u_\varepsilon dx dt = 0$ because of condition **(S)** and that u_ε is smooth and bounded. We take $\varepsilon \rightarrow 0$ and the right-hand side converges to 0 due to Lemma 4.11, Now we obtain

$$\|u(T)\|_{L^2} - \|u(0)\|_{L^2} + \int_0^T \int_{\mathbb{R}^n} \langle \nabla u, a \cdot \nabla u \rangle dx dt = 0,$$

which gives us

$$\|u(T)\|_{L^2} + \lambda \|\nabla u\|_{L^2} \leq \|u(0)\|_{L^2}$$

using condition **(E)**. This energy inequality now implies the uniqueness of the weak solution. $\square$

The condition that u_0 is bounded implies that the weak solution u is also bounded, which is crucial to the convergence of right-hand side when taking $\varepsilon \rightarrow 0$. When the initial data is not bounded, we can not expect the solution to be bounded and hence impose difficulty to us. To deal with this, the concept of renormalized solution is introduced by considering $h(u)$ with bounded smooth h instead of u . The idea is first introduced in [17] and then generalized to parabolic equations. A first expose of the definition of

renormalized solution to parabolic equation is in Lions [52].

Definition 4.13. We say that u is a renormalized solution to

$$\partial_t u - \operatorname{div}(a \cdot \nabla u) + b \cdot \nabla u = 0$$

if for any bounded smooth function h with compactly supported h' , $h(u)$ satisfies

$$\partial_t h(u) - \operatorname{div}(a \cdot \nabla h(u)) + b \cdot \nabla h(u) + h''(u) \langle \nabla u, a \cdot \nabla u \rangle = 0$$

in the distribution sense.

Theorem 4.14. Suppose a, b satisfy conditions (E), (S) and a is constant in equation (1.9).

If we assume that $b \in L^2(0, T; L^2(\mathbb{R}^n))$, then for any initial data $u_0 \in L^2$, there is a unique renormalized solution $u \in L^2(0, T; H^1) \cap C([0, T]; L^2)$ corresponding to u_0 .

Since the unboundedness of the solution is the main difficulty, the definition of the renormalized solution actually truncates it by imposing boundedness condition on h . The proof of this theorem in essence follows the idea used for initial data $u_0 \in L^\infty$. A detailed proof of this theorem can be found in [52, Appendix E] and we will omit it here.

Chapter 5

Conclusions

In this chapter we summarize what have been obtained using current approaches and some thoughts on further problems. We split the discussion into two parts regarding the parabolic equations and the diffusion processes respectively.

5.1 Parabolic equations

With the divergence-free condition, the existence of weak solutions to the parabolic equations can be obtained for $b \in L^2(0, T; L^2_{loc})$. Regarding the uniqueness of the solution, the concept of renormalized solution is introduced and the solution is unique for $b \in L^2(0, T; L^2(\mathbb{R}^n))$ (this condition here can be generalized, see [46, 47]). Compared with existence and uniqueness, regularity of the solution remains an open problem. The Hölder continuity of the weak solution is obtained for the critical case that $b \in L^\infty(0, T; \text{BMO}^{-1})$ which is a marginal condition. The regularity theory can be derived from the Aronson estimate obtained in Chapter 2. The Aronson estimate shows that the diffusion term in the equation is dominant over the drift term and hence the fundamental solution is comparable with Gaussian functions. For the supercritical case, it appears that in short time the drift term is dominant over the diffusion term and singularity may appear.

Regarding the linear parabolic equations under supercritical conditions, examples of singular solutions would be of great interest. Also partial regularity result of the weak

solution can be considered as well for the purpose of understanding the solutions. For the general theory of parabolic equations, the bounded divergence condition can be considered instead of the divergence-free condition to obtain the Aronson-type estimate, because bounded divergence is also an important assumption in the theorem of renormalized solutions. Regarding fluid dynamics, the idea of stochastic Lagrangian representation can be used to study various of equations describing viscous incompressible fluid, including axisymmetric Navier-Stokes equations and Prandtl's boundary layer problem.

5.2 Diffusion processes

Applying the Aronson estimate in the critical case, a unique diffusion process (a weak solution to the SDE satisfying the strong Markov property) can be immediately obtained. Further, since the Aronson estimate controls the speed of the process X_t going to infinity, we can prove that if in addition $b \in L^2(0, T; H^1)$ (this condition can be generalized) there actually exists a unique strong solution to the SDE for almost every initial data $x_0 \in \mathbb{R}^n$. Moreover, the solution is approximately differentiable in space variable x . Recall that the stochastic representation of velocity and vorticity fields both involve ∇X_t , which represents the deformation of the stochastic flow along the stochastic Lagrangian coordinates. Therefore, it is also interesting to estimate the gradient. It is worth mentioning that under the divergence free condition, we always have $\det \nabla X_t = 1$. In Krylov and Röckner [42], there is a unique strong solution X_t if $b \in L^l(0, T; L^q(\mathbb{R}^n))$ with $\frac{2}{l} + \frac{n}{q} < 1$. Moreover, X_t is Hölder continuous and differentiable. Similar result when $\frac{2}{l} + \frac{n}{q} \geq 1$ is still largely open and we can only obtain approximately differentiability in this thesis. We know that the differentiability of the flow is closely related to the regularity of the flow, while more understanding is required to reveal its relation with the regularity theory of the parabolic equations.

In the supercritical case, failing to obtain the Chapman–Kolmogorov equation for the transition probability becomes a major difficulty in constructing a diffusion process

satisfying the strong Markov property. Although the Aronson estimate is not Gaussian, its power on controlling the speed of the process X_t going to infinity is expected, which may potentially extend the previous construction of strong solutions from critical cases to supercritical cases. From the point of view of renormalized solutions, intuitively it should determine a weak solution to the SDE uniquely in law for almost every initial data $x_0 \in \mathbb{R}^n$ (see [46, 47]), while technically it has not been achieved yet due to the appearance of the diffusion part. In general, more understanding is needed for the renormalized solution and what properties of the diffusion processes it implies.

Bibliography

- [1] D. G. Aronson. Uniqueness of positive weak solutions of second order parabolic equations. In *Annales Polonici Mathematici*, volume 3, pages 285–303, 1965.
- [2] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bulletin of the American Mathematical society*, 73(6):890–896, 1967.
- [3] D. G. Aronson. Non-negative solutions of linear parabolic equations. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 22(4):607–694, 1968.
- [4] M. T. Barlow, A. Grigor’yan, and T. Kumagai. On the equivalence of parabolic Harnack inequalities and heat kernel estimates. *J. Math. Soc. Japan*, 64(4):1091–1146, 2012.
- [5] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.
- [6] V. I. Bogachev. *Gaussian measures*, volume 62. American Mathematical Society Providence, 1998.
- [7] B. Busnello, F. Flandoli, and M. Romito. A probabilistic representation for the vorticity of a three-dimensional viscous fluid and for general systems of parabolic equations. *Proceedings of the Edinburgh Mathematical Society*, 48(2):295–336, 2005.
- [8] N. Champagnat and P. Jabin. Strong solutions to stochastic differential equations with rough coefficients. *arXiv preprint arXiv:1303.2611*, 2013.

- [9] R. R. Coifman, S. Semmes, P.-L. Lions, and Y. Meyer. Compacité par compensation et espaces de Hardy. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, 309(18):945–949, 1989.
- [10] P. Constantin. An Eulerian-Lagrangian approach for incompressible fluids: local theory. *J. Amer. Math. Soc.*, 14(2):263–278, 2001.
- [11] P. Constantin and G. Iyer. A stochastic Lagrangian representation of the three-dimensional incompressible Navier-Stokes equations. *Communications on Pure and Applied Mathematics*, 61(3):330–345, 2008.
- [12] P. Constantin and G. Iyer. A stochastic-Lagrangian approach to the Navier-Stokes equations in domains with boundary. *The Annals of Applied Probability*, 21(4):1466–1492, 2011.
- [13] G. Crippa and C. De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2008(616):15–46, 2008.
- [14] E. B. Davies. Explicit constants for Gaussian upper bounds on heat kernels. *American Journal of Mathematics*, 109(2):319–333, 1987.
- [15] E. B. Davies. *Heat kernels and spectral theory*, volume 92. Cambridge university press, 1990.
- [16] E. De Giorgi. Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)*, 3:25–43, 1957.
- [17] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Inventiones mathematicae*, 98(3):511–547, 1989.
- [18] H. Dong and S. Kim. Fundamental solutions for second order parabolic systems with drift terms. *arXiv preprint arXiv:1707.09162*, 2017.

- [19] L. Escauriaza, G. Seregin, and V. Šverák. L_3^∞ -solutions of the Navier-Stokes equations and backward uniqueness. *Russian Mathematical Surveys*, 58(2):211, 2003.
- [20] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.
- [21] E. B. Fabes and D. W. Stroock. A new proof of Moser’s parabolic Harnack inequality using the old ideas of Nash. In *Analysis and Continuum Mechanics*, pages 459–470. Springer, 1989.
- [22] H. Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.
- [23] E. Fedrizzi and F. Flandoli. Hölder flow and differentiability for SDEs with nonregular drift. *Stoch. Anal. Appl.*, 31(4):708–736, 2013.
- [24] H. Feng and V. Šverák. On the Cauchy problem for axi-symmetric vortex rings. *Archive for Rational Mechanics and Analysis*, 215(1):89–123, 2015.
- [25] A. Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, 254(1):109–153, 2008.
- [26] F. Flandoli, E. Issoglio, and F. Russo. Multidimensional stochastic differential equations with distributional drift. *Transactions of the American Mathematical Society*, 369(3):1665–1688, 2017.
- [27] S. Friedlander and V. Vicol. Global well-posedness for an advection–diffusion equation arising in magneto-geostrophic dynamics. In *Annales de l’Institut Henri Poincaré (C) Non Linear Analysis*, volume 28, pages 283–301. Elsevier, 2011.
- [28] A. Friedman. *Partial differential equations of parabolic type*. Courier Corporation, 2013.
- [29] M. Fukushima. *Dirichlet forms and Markov processes*. Elsevier Science & Technology, 1980.

- [30] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19. Walter de Gruyter, 2010.
- [31] T. Gallay and V. Šverák. Uniqueness of axisymmetric viscous flows originating from circular vortex filaments. *arXiv preprint arXiv:1609.02030*, 2016.
- [32] A. Grigor'yan. The heat equation on noncompact Riemannian manifolds. *Matematicheskii Sbornik*, 182(1):55–87, 1991.
- [33] A. Grigor'yan. *Heat kernel and analysis on manifolds*, volume 47. American Mathematical Soc., 2009.
- [34] S. Hofmann and S. Mayboroda. Hardy and BMO spaces associated to divergence form elliptic operators. *Mathematische Annalen*, 344(1):37–116, 2009.
- [35] E. P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.
- [36] P. Hsu. Heat semigroup on a complete Riemannian manifold. *Ann. Probab.*, 17(3):1248–1254, 1989.
- [37] M. Ignatova, I. Kukavica, and L. Ryzhik. The Harnack inequality for second-order parabolic equations with divergence-free drifts of low regularity. *Communications in Partial Differential Equations*, 41(2):208–226, 2016.
- [38] F. John and L. Nirenberg. On functions of bounded mean oscillation. *Communications on pure and applied Mathematics*, 14(3):415–426, 1961.
- [39] T. Kato. *Perturbation theory for linear operators*, volume 132. Springer Science and Business Media, 2013.
- [40] H. Koch and D. Tataru. Well-posedness for the Navier–Stokes equations. *Advances in Mathematics*, 157(1):22–35, 2001.

- [41] V. F. Kovalenko and Y. A. Semenov. C_0 -Semigroups in $L^p(\mathbb{R}^d)$ and $C(\mathbb{R}^d)$ spaces generated by the differential expression $\Delta + b \cdot \nabla$. *Theory of Probability & Its Applications*, 35(3):443–453, 1991.
- [42] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probability theory and related fields*, 131(2):154–196, 2005.
- [43] O. A. Ladyzhenskaya. Uniqueness and smoothness of generalized solutions of Navier-Stokes equations. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 5:169–185, 1967.
- [44] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Ural’ceva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988.
- [45] P. D. Lax. *Functional analysis*. New York, Wiley-Interscience, 2002.
- [46] C. Le Bris and P.-L. Lions. Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications. *Annali di Matematica pura ed applicata*, 183(1):97–130, 2004.
- [47] C. Le Bris and P.-L. Lions. Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients. *Communications in Partial Differential Equations*, 33(7):1272–1317, 2008.
- [48] P. G. Lemarie-Rieusset. *The Navier-Stokes problem in the 21st century*. CRC Press, 2016.
- [49] J. Leray. Sur le mouvement d’un liquide visqueux emplissant l’espace. *Acta mathematica*, 63(1):193–248, 1934.
- [50] H. Li and D. Luo. Quantitative stability estimates for solutions of Fokker-Planck equations. *arXiv preprint arXiv:1701.00566*, 2017.
- [51] G. M. Lieberman. *Second order parabolic differential equations*. World scientific, 1996.

- [52] P.-L. Lions. *Mathematical topics in fluid mechanics. Vol. 1*, volume 3 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1996.
- [53] V. Liskevich. On C_0 -semigroups generated by elliptic second order differential expressions on L^p -spaces. *Differential and Integral Equations*, 9(4):811–826, 1996.
- [54] V. Liskevich and Z. Sobol. Estimates of integral kernels for semigroups associated with second-order elliptic operators with singular coefficients. *Potential Analysis*, 18(4):359–390, 2003.
- [55] V. Liskevich and Q. Zhang. Extra regularity for parabolic equations with drift terms. *Manuscripta Mathematica*, 113(2):191–209, 2004.
- [56] Z. Ma and M. Röckner. *Introduction to the theory of (non-symmetric) Dirichlet forms*. Springer Science & Business Media, 2012.
- [57] P. A. Meyer and W. A. Zheng. Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincaré Probab. Statist.*, 20(4):353–372, 1984.
- [58] A. S. Monin and A. M. Yaglom. *Statistical Fluid Mechanics, Volume I: Mechanics of Turbulence*. Courier Corporation, 2007.
- [59] A. S. Monin and A. M. Yaglom. *Statistical Fluid Mechanics, Volume II: Mechanics of Turbulence*. Courier Corporation, 2013.
- [60] C. B. Morrey Jr. *Multiple integrals in the calculus of variations*. Springer Science & Business Media, 2009.
- [61] J. Moser. A new proof of De Giorgi’s theorem concerning the regularity problem for elliptic differential equations. *Communications on Pure and Applied Mathematics*, 13(3):457–468, 1960.
- [62] J. Moser. A Harnack inequality for parabolic differential equations. *Communications on pure and applied mathematics*, 17(1):101–134, 1964.

- [63] J. Moser. On a pointwise estimate for parabolic differential equations. *Communications on Pure and Applied Mathematics*, 24(5):727–740, 1971.
- [64] J. Nash. Continuity of solutions of parabolic and elliptic equations. *American Journal of Mathematics*, 80(4):931–954, 1958.
- [65] J. R. Norris and D. W. Stroock. Estimates on the fundamental solution to heat flows with uniformly elliptic coefficients. *Proceedings of the London Mathematical Society*, 3(2):373–402, 1991.
- [66] H. Osada. Diffusion processes with generators of generalized divergence form. *J. Math. Kyoto Univ*, 27(4):597–619, 1987.
- [67] G. Prodi. Un teorema di unicità per le equazioni di Navier-Stokes. *Annali di Matematica pura ed applicata*, 48(1):173–182, 1959.
- [68] Z. Qian and G. Xi. Parabolic equations with singular divergence-free drift vector fields. *arXiv preprint arXiv:1612.07727*, 2016.
- [69] Z. Qian and G. Xi. Markov semi-groups generated by elliptic operators with divergence-free drift. *arXiv preprint arXiv:1706.06317*, 2017.
- [70] Z. Qian and G. Xi. Parabolic equations with divergence-free drift in space $L_t^l L_x^q$. *To appear in Indiana University Mathematics Journals*, 2018.
- [71] F. Rezakhanlou. Regular flows for diffusions with rough drifts. *arXiv preprint arXiv:1405.5856*, 2014.
- [72] Y. A. Semenov. Regularity theorems for parabolic equations. *Journal of Functional Analysis*, 231(2):375–417, 2006.
- [73] G. Seregin. *Lecture notes on regularity theory for the Navier-Stokes equations*. World Scientific, 2015.

- [74] G. Seregin, L. Silvestre, V. Šverák, and A. Zlatoš. On divergence-free drifts. *Journal of Differential Equations*, 252(1):505–540, 2012.
- [75] J. Serrin. The initial value problem for the Navier-Stokes equations. In *Nonlinear Problems (Proc. Sympos., Madison, Wis., 1962)*, pages 69–98. Univ. of Wisconsin Press, Madison, Wis., 1963.
- [76] Z. Sobol and H. Vogt. On the L_p -theory of C_0 -semigroups associated with second-order elliptic operators. *Journal of Functional Analysis*, 193(1):24–54, 2002.
- [77] W. Stannat. Dirichlet forms and markov processes: a generalized framework including both elliptic and parabolic cases. *Potential Analysis*, 8(1):21–60, 1998.
- [78] E. M. Stein. *Singular integrals and differentiability properties of functions (PMS-30)*, volume 30. Princeton university press, 2016.
- [79] D. W. Stroock. Diffusion semigroups corresponding to uniformly elliptic divergence form operators. In *Séminaire de Probabilités XXII*, pages 316–347. Springer, 1988.
- [80] D. W. Stroock. *Partial differential equations for probabilists*. Cambridge Univ. Press, 2008.
- [81] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*. Springer, 2007.
- [82] G. I. Taylor. Diffusion by continuous movements. *Proceedings of the london mathematical society*, 2(1):196–212, 1922.
- [83] G. I. Taylor. Statistical theory of turbulence. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 151, pages 421–444. The Royal Society, 1935.
- [84] R. Temam. *Navier-Stokes equations*, volume 2. North-Holland Amsterdam, 1984.

- [85] W. Von Wahl. *The equations of Navier-Stokes and abstract parabolic equations*. Springer, 1985.
- [86] S. T. Yau. On the heat kernel of a complete Riemannian manifold. *J. Math. Pures Appl.* (9), 57(2):191–201, 1978.
- [87] K. Yosida. *Functional analysis*. Springer-verlag, 1980.
- [88] Q. Zhang. A strong regularity result for parabolic equations. *Communications in mathematical physics*, 244(2):245–260, 2004.
- [89] Q. Zhang. Local estimates on two linear parabolic equations with singular coefficients. *Pacific journal of mathematics*, 223(2):367–396, 2006.
- [90] X. Zhang. Stochastic flows of SDEs with irregular coefficients and stochastic transport equations. *Bull. Sci. Math.*, 134(4):340–378, 2010.
- [91] X. Zhang. A stochastic representation for backward incompressible Navier-Stokes equations. *Probability theory and related fields*, 148(1-2):305–332, 2010.
- [92] X. Zhang. Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients. *Rev. Mat. Iberoam.*, 29(1):25–52, 2013.
- [93] X. Zhang. Stochastic differential equations with Sobolev diffusion and singular drift and applications. *The Annals of Applied Probability*, 26(5):2697–2732, 2016.
- [94] X. Zhang and G. Zhao. Heat kernel and ergodicity of SDEs with distributional drifts. *arXiv preprint arXiv:1710.10537*, 2017.
- [95] W. A. Zheng. Tightness results for laws of diffusion processes application to stochastic mechanics. *Ann. Inst. H. Poincaré Probab. Statist.*, 21(2):103–124, 1985.
- [96] V. Zhikov. Estimates of the Nash-Aronson type for degenerating parabolic equations. *Journal of Mathematical Sciences*, 190(1), 2013.

- [97] V. V. Zhikov. Remarks on the uniqueness of a solution of the Dirichlet problem for second-order elliptic equations with lower-order terms. *Functional Analysis and Its Applications*, 38(3):173–183, 2004.
- [98] V. V. Zhikov. Estimates of Nash-Aronson type for a diffusion equation with asymmetric matrix and their applications to homogenization. *Russian Academy of Sciences Sbornik Mathematics*, 197, 12 2006.