Local Gradient Estimate for Porous Medium and Fast Diffusion Equations by Martingale Method

Zichen Zhang
St Hugh's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Hilary 2014

To my parents

Acknowledgements

I would like to thank Dr Zhongmin Qian for his enormous encouragement and enlightenment to not only my research but also my life. It is my great pleasure to work with such a passionate scholar. With the same degree of passion, gaining insight about mathematics as deeply as he does will be my lifelong pursuit.

I wish to express my sincere gratitude to Professor Ying Hu. Every discussion with him was inspiring. It is a privilege to work with him.

I also wish to express my gratitude to Professor Ben Hambly, Dr Hanqing Jin and Dr Dmitry Belyzev for their kindness and help.

My thanks also go to Kun Ma, Liyun Situ, Gechun Liang, Yiqing Lin, Wei Pan, Yuan Xia, Danyu Yang, Hao Ni, Horatio Boedihardjo, Xi Geng and all my other friends at Oxford, without whom my life in the past three years would be finite dimensional.

Abstract

This thesis focuses on a certain type of nonlinear parabolic partial differential equations, i.e. PME and FDE, which can be written as $\partial_t u = \Delta u^m$ with $m > 0$. Our aim is to derive better gradient bound in the form of

$$|\nabla u|^2 (t, \cdot) \le C \frac{\|u(0, \cdot)\|_\infty^\kappa - u^\kappa (t, \cdot)}{t}, \text{ and } \frac{|\nabla u|^2}{u^{\kappa_1}} (t, \cdot) - \frac{\partial_t u}{u^{\kappa_2}} (t, \cdot) \le \frac{C_1}{t}.$$

Chapter 1 consists of a survey on results related to PME and FDE, and a short review on some works about deriving gradient estimates in probabilistic ways.

In Chapter 2 we estimate gradient on space variables of solutions to the heat equation on Euclidean space. The main idea is to construct two semimartingales by letting the solution and its gradient running backward on the path space of a diffusion process. Estimates derived from decompositions of those two semimartingales are then combined to give rise to an upper bound on gradient that only involves the maximum of the initial data and time variable. In particular, it is independent of the dimension.

In Chapter 3 we carry the idea in Chapter 2 onto the study of positive solutions to PME or FDE, and obtained a similar type of bound on $|\nabla u|$ for local solutions to PME or FDE on Euclidean space. In existing literature there have always been constraints on m. By considering a more general form of transformation on u and introducing a family of equivalent measures on path space, we add more flexibility to our method. Thus our result is valid for a larger range of m. For global solutions, when m violates our constraint, we need two-sided bound on u to control $|\nabla u|$.

In Chapter 4 we utilize maximum principle to derive Li-Yau type gradient estimate for PME on a compact Riemannian manifold with Ricci curvature bounded from below. Our result is able to yield a Harnack inequality possessing the right order in time variable when the lower bound of Ricci curvature is negative.

Contents

1 Introduction 1

1.1 Porous medium and fast diffusion equations . . . 1

1.2 Application of stochastic analysis to gradient estimate . . . 13

2 Gradient estimate for solutions to the heat equation 15

3 Gradient estimate for positive solutions to PME and FDE 22

3.1 Local gradient estimate . . . 22

3.1.1 From PDE to SDE . . . 25

3.1.2 Fundamental decompositions . . . 26

3.1.3 Gradient bound for solution u . . . 30

3.1.4 Gradient bound for pressure variable $v = \frac{u^{m-1}}{m-1}$ . . . 44

3.1.5 Gradient bound for u depending only on initial data . . . 48

3.2 The constraint on m . . . 57

3.2.1 Supermartingale property of $\frac{1}{H}$ . . . 57

3.2.2 More parameters in change of measures . . . 59

3.2.3 General transformation . . . 68

3.3 Global gradient estimate . . . 71

3.3.1 From local bound to global bound . . . 71

3.3.2 Negative finite variation part . . . 76

4 Gradient estimate of Li-Yau's type for positive solutions to PME 85

4.1 Differential inequalities and the curvature . . . 86

4.1.1 Basic differential inequalities . . . 87

4.1.2 Further differential inequality . . . 89

4.2 Conjecture about the gradient bound . . . 90

4.3 Applying maximum principle . . . 93

4.4 Harnack inequality . . . 96

A Preliminaries **102**
A.1 Basics of Martingale theory 102
A.2 Change of measures on Wiener space 105

Bibliography **109**

Chapter 1

Introduction

1.1 Porous medium and fast diffusion equations

This thesis is devoted to gradient estimates for partial differential equations (PDEs) of the form

$$\frac{\partial u}{\partial t} = \Delta u^m, \tag{1.1}$$

where $m \in (0, \infty)$ is a given number. The problem is posed on $(0, \infty) \times M$, for a Riemannian manifold M. Δ is the Laplace-Beltrami operator. When $m = 1$, it is the celebrated heat equation, which is linear and parabolic, hence enjoying many good properties. When $m \neq 1$, the story becomes quite different. Let us focus on non-negative solutions to the equation (1.1) on Euclidean space with initial condition

$$u\left(0, x\right) = u_0\left(x\right),$$

the so-called Cauchy problem related to (1.1), and collect some results and phenomena that do not appear in the case of the heat equation.

To begin with, let us point out that there are two critical points about the value of m, namely $m = 1$ and $m = \frac{n-2}{n}$. By rewriting (1.1) as

$$\frac{\partial u}{\partial t} = m u^{m-1} \Delta u + m \left(m - 1\right) u^{m-2} \left|\nabla u\right|^2 , \tag{1.2}$$

one can see that the coefficient matrix of the second order derivative is $m u^{m-1} I_{n \times n}$. If $m < 1$, this matrix takes the value of infinity when $u = 0$, which means the parabolicity is singular. While if $m > 1$, the matrix vanishes when $u = 0$, which means the parabolicity is degenerate. If we interpret (1.1) as a differential equation describing a diffusion, this means when $m < 1$, the diffusion is very fast at places where u is small. Therefore in this case we call (1.1) the fast diffusion equation (FDE). When $m > 1$, the diffusion slows down wherever u is small, and in this case we call it porous medium equation (PME). Notice that since we only consider non-negative solution u, from (1.2) it is clear that (1.1) is always formally parabolic.

Due to the different behaviors of PME and FDE at $u = 0$, theories about existence and uniqueness of Cauchy problems for these two types of equations have been studied separately. For PME, the commonly used framework is $L^1 \left(\mathbb{R}^n\right)$ space. By Theorem 9.12 and Proposition 9.13 in [38], if $u_0 \geq 0$ and $u_0 \in L^1 \left(\mathbb{R}^n\right)$, then there is a unique strong solution which is continuous on $(0, \infty) \times \mathbb{R}^n$. Moreover, if $u_0 \in L^1 \left(\mathbb{R}^n\right)$ and is strictly positive and continuous, then the solution will be smooth. If we move beyond the scope of $L^1(\mathbb{R}^n)$ setting and impose a weaker growth condition on u_0 which reads

$$\sup_{R\geq 1} R^{-\left(n+\frac{2}{m-1}\right)} \int_{|x|\leq R} |u_0(x)|\, dx < \infty, \tag{1.3}$$

then by [7] a unique solution in distribution sense exists on $(0, T(u_0)) \times \mathbb{R}^n$, where $T(u_0) \in (0, \infty]$ and depends on u_0. By Theorem 3.1 in [3], this growth condition is satisfied by any continuous non-negative solution. Therefore, condition (1.3) is optimal for the class of continuous non-negative solutions. When the initial data is allowed to be measure-valued, [15] gives a result which requires similar growth condition as (1.3). As for FDE, no requirement on the growth of initial data is needed. In fact, by Theorem 2.1 and 2.3 in [20], there exists a unique solution $u \in C([0,\infty); L^1_{loc}(\mathbb{R}^n))$ in distribution sense if $u_0 \in L^1_{loc}(\mathbb{R}^n)$. Alternatively, if we impose some growth and decay conditions on u_0, by Theorem 1 in [8] there will be a classical solution in $\left[0, \bar{T}\right] \times \mathbb{R}^n$, where $\bar{T}$ is finite.

The degeneracy of parabolicity of PME results in the finite propagation speed of its solution, which is one of the peculiar feature of PME. In particular, by Theorem 14.6 in [38], if $u(t_1, \cdot)$ is compactly supported in $\mathbb{R}^n$, so is $u(t_2, \cdot)$ for any $t_2 > t_1$. Consequently, for this kind of solutions, there is a set in $(0,\infty) \times \mathbb{R}^n$ that separates the region on which u is positive and the region where u is zero. According to Theorem 3.3 in [13], this set, or the so-called free boundary, is locally Hölder continuous on $(0,\infty) \times \mathbb{R}^n$. Moreover, as a family of boundaries in $\mathbb{R}^n$ indexed by $t \in (0,\infty)$, those boundaries expands to infinity as $t \to \infty$ [13]. If $t_\star$ is the time when the solution overflows the smallest ball where the support of the initial data is contained, then after time $t_\star$ there is an improvement in the regularity of the free boundary. More precisely, Theorem 3 in [12] asserts that the free boundary is locally Lipschitz continuous on $(t_\star, \infty) \times \mathbb{R}^n$. With extra conditions on the first and second order derivatives of the initial data, [16] proves that the free boundary is smooth.

Although $m = 1$ is a crucial point when talking about finite propagation and existence theories of equation (1.1), it is not a significant value in the study of extinction in finite time and smoothing effect, where the value $m = \frac{n-2}{n}$ becomes decisive. The extinction in finite time of a solution is a phenomenon that arises only when $m < \frac{n-2}{n}$. On page 174 in [6], it is proved that any solution with initial value $u_0 \in L^{p_\star}(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$, where $p_\star = \frac{n(1-m)}{2}$, becomes identically zero after a finite time. More generally, by Theorem 5.2 in [35], the same result holds for $u_0 \in \mathcal{M}^{p_\star}$, where

$$\mathcal{M}^{p_\star} = \left\{ f \in L^1_{loc}(\mathbb{R}^n) : \int_K |f(x)|\, dx \leq C\, |K|^{1-\frac{1}{p_\star}}, \forall K \text{ with } |K| < \infty \right\}.$$

By Lemma 5.6 in [35], this is already very close to the sufficient condition for a solution to extinct in finite time. One can see from these results that even a positive initial data may produce a solution that vanishes completely in finite time, which is quite striking. The reason behind, is the failure of conservation of mass when $m < \frac{n-2}{n}$, as explained in Section 5.5 in [35].

Next, let us discuss regularities of solutions in terms of boundedness, positivity and smoothness. In general, for any $m \in (0, \infty)$ and $p \in [1, \infty]$, the solution decreases in $L^p(\mathbb{R}^n)$ norm as it evolves in time, according to Theorem 7.2 in [36]. Moreover, if

$m > \frac{n-2}{n}$, then by Section 3.4 in [35], for any p and q such that $1 \leq p \leq q \leq \infty$,

$$\left\| u\left(t, \cdot\right) \right\|_q \leq c\left(m, n, p, q\right) \left\| u_0 \right\|_p^{\sigma} t^{-\alpha}, \tag{1.4}$$

with

$$\alpha = \frac{n\left(q-p\right)}{q\left(n\left(m-1\right)+2p\right)}, \ \ \sigma = \frac{p\left(n\left(m-1\right)+2q\right)}{q\left(n\left(m-1\right)+2p\right)}.$$

In particular, from this result we see that initial data in $L^1\left(\mathbb{R}^n\right)$ produce solutions $u\left(t, \cdot\right) \in L^{\infty}\left(\mathbb{R}^n\right)$ for any time $t > 0$, which is termed as smoothing effect. When $m = \frac{n-2}{n}$, this is no longer true. Appendix A.3 in [14] constructed a solution which is not bounded at any time while still having an initial value in $L^1\left(\mathbb{R}^n\right)$. Moreover, when $m < \frac{n-2}{n}$ and $n \geq 3$, Theorem 5.14 in [35] shows that (1.4) holds for $q = 1$ and $1 < p < p_{\star}$. This means that $L^p\left(\mathbb{R}^n\right)$ data yield solutions only in $L^1\left(\mathbb{R}^n\right)$, a somehow backward smoothing effect. Nevertheless, we are still able to get a bounded solution in the case of $m < \frac{n-2}{n}$ if the initial data belongs to a better space. By Theorem 6.7 in [35], if $m < \frac{n-2}{n}$, $p \geq p_{\star}$ and $u_0 \in L^p\left(\mathbb{R}^n\right) + L^{\infty}\left(\mathbb{R}^n\right)$, then $u\left(t, \cdot\right)$ is locally bounded and smooth for any $t > 0$. Quantitatively, Theorem 2.1 in [11] gives a local upper bound for $u\left(t, \cdot\right)$ in terms of the $L^p\left(\mathbb{R}^n\right)$ $\left(p > p_{\star}\right)$ norm of u_0 over a larger local region.

As for positivity and smoothness. Theorem 3.1 in [3], Theorem 1.1 in [10] and Theorem 1.1 in [11] give local lower bounds for solutions to (1.1) in terms of local $L^1\left(\mathbb{R}^n\right)$ norm of initial data in the cases of $m > 1$, $m \in \left(\frac{n-2}{n}, 1\right)$ and $m \in \left(0, \frac{n-2}{n}\right)$ respectively. These lead to results about the positivity of solutions. For example, Proposition 1.1 in [12] gave a necessary and sufficient condition for the positivity of $u(t,x)$ when $m>1$ which is read as

$$\sup_{R>0} R^{-\left(n+\frac{2}{m-1}\right)} \int_{|y-x|\le R} |u_0(y)|\, dy = \infty.$$

Besides positivity, if the solution is also locally bounded, then the standard quasilinear theory [25] implies the smoothness of the solution on that region. In particular, when $m \in \left(\frac{n-2}{n}, 1\right)$, non-negative locally integrable initial data always produce positive and smooth global solutions, according to the remark after Theorem 2.3 in [20]. This is not true when $m \in \left(0, \frac{n-2}{n}\right)$, as non-negative locally integrable initial data is not enough for locally boundedness of solutions, which can be seen from solution (0.2) in [11]. Moreover, when m is in this range, extinction in finite time may occur, which kills positivity of solutions. As for PME, in general, solutions are only locally Hölder continuous. Theorem 4.1 in [13] states that when u_0 is non-negative, bounded and belongs to $L^2(\mathbb{R}^n)$, u^m is uniformly Hölder continuous in every set $(\eta_0, \infty) \times \mathbb{R}^n$, $\eta_0 > 0$. Theorem 7.17 in [38] tells us that u is locally Hölder continuous on the region where u is bounded.

During the development of the above works, one of the main tools is the comparison principle for equation (1.1), which is established in Theorem 7.3 in [36]. Generally speaking, the comparison is in terms of mass concentration of radially symmetric functions. Once this comparison of mass concentration is obtained, comparison in terms of $L^p(\mathbb{R}^n)$ ($p \in [1,\infty]$) norm follows. There are several special explicit solutions to (1.1) that are often used in combination with the comparison principle. Let us only mention three of them here. The first one is the known as Zel'dovich-Kompanyeets-Barenblatt (ZKB) solution, which takes the form

$$\mathcal{U}(t,x) = t^{-\alpha}\left(c - k\left|x\right|^2 t^{-2\beta}\right)_+^{\frac{1}{m-1}},$$

with

$$\alpha = \frac{n}{n(m-1)+2},\ \beta = \frac{\alpha}{n},\ k = \frac{\alpha(m-1)}{2mn},$$

and a positive constant c. When $m > \frac{n-2}{n}$,

$$\lim_{t\to 0}\mathcal{U}(t,x) = C(c,n,m)\,\delta_0(x),$$

in the sense of measures. So it is the solution to (1.1) with Dirac delta as initial trace. One can see that it reproduces the heat kernel as $m \to 1$. Since the ZKB solution has compact support when $m > 1$, it plays an important role in the study of finite propagation speed. When $m \le \frac{n-2}{n}$, $\mathcal{U}(t,x)$ is no longer integrable in space variables. Therefore, although still solving (1.1), it is not a solution to any Cauchy problem related to (1.1). The second type of solution is the self similar solution, which has the form

$$\mathcal{U}(t,x) = t^{-\lambda_1} F\left(\left|x\right| t^{-\lambda_2}\right).$$

With a scaling argument, it is shown in section 3.2.1 of [37] that any solution to (1.1) on $(0,\infty)\times\mathbb{R}^n$ with initial data $\left|x\right|^{-\lambda_3}$ must have the above form. Notice that in fact the ZKB solution also belongs to this type. The third type of solution we would like to mention here is a variance of the self similar solution. It has the form

$$\mathcal{U}(t,x) = (T-t)^{-\lambda_1} \bar{F}\left(|x|(T-t)^{-\lambda_2}\right),$$

with $T > 0$. Usually it satisfies (1.1) only for $t < T$, after which it blows up or vanishes. For this type of solution, probably the most popular and explicit one is

$$U(x,t;T) = 2\left(n - \frac{2}{1-m}\right)\left(\frac{T-t}{|x|^2}\right)^{\frac{1}{1-m}},$$

with $m < \frac{n-2}{n}$, which is a good example of solutions that extinct in finite time.

As frequently seen in the field of PDE, a crucial step in the study of equation (1.1) is to derive various types of estimates for solutions. In [4], Aronson and Bénilan established the following gradient estimate for solutions to (1.1). If $m \in \left(1 - \frac{2}{n}, \infty\right)$, u is a positive smooth solution to (1.1) and $v = \frac{m}{m-1}u^{m-1}$, then

$$\Delta v \geq -\frac{\alpha}{(m-1)t},$$

where

$$\alpha = \frac{n(m-1)}{n(m-1)+2},$$

which is equivalent to

$$\frac{|\nabla v|^2}{(m-1)v} - \frac{v_t}{(m-1)v} \leq \frac{\alpha}{(m-1)t}. \tag{1.5}$$

This fundamental estimate is then employed in [7] for the study of existence theory, in [20] for $L^{\infty}_{loc}(\mathbb{R}^n)$ estimate for solutions, and in [12] for obtaining regularity results for the free boundary of solutions, to name but a few. Later in [27], a local version of the Aronson-Bénilan estimate was obtained by P. Lu, L. Ni, J. L. Vazquez and C. Villani. They studied the same problem posed on a local ball of a Riemannian manifold. Let $B(\mathcal{O},2R)$ denote a ball with centre $\mathcal{O}$ and radius $2R>0$. Assume that u is a positive solution to (1.1) on $B(\mathcal{O},2R)\times[0,T]$ and the Ricci curvature Ric$\geq -(n-1)K^2$ on $B(\mathcal{O},2R)$ for some $K\geq 0$. They showed that for any $m>1$ and $\beta>1$, it holds on $B(\mathcal{O},R)\times[0,T]$ that

$$\frac{|\nabla v|^2}{v}-\beta\frac{v_t}{v}\leq \alpha\beta^2\left(\frac{1}{t}+C_1K^2v_{\max}^{2R,T}\right)+\alpha\beta^2\frac{v_{\max}^{2R,T}}{R^2}C_2(\beta,KR,m,n), \tag{1.6}$$

where

$$v_{\max}^{2R,T}=\max_{B(\mathcal{O},2R)\times[0,T]}v,\ C_1=\frac{(m-1)(n-1)}{\beta-1}.$$

Based on this, the following Harnack inequality was obtained.

$$\frac{v(x_2,t_2)}{v(x_1,t_1)}\geq\left(\frac{t_1}{t_2}\right)^{\alpha\beta}\exp\left(-\frac{\beta d^2(x_1,x_2)}{4v_{\min}^{R/2,T}}-\alpha\beta(t_2-t_1)v_{\max}^{R,T}\left(C_1K^2+\frac{C_2}{R^2}\right)\right), \tag{1.7}$$

where

$$v_{\min}^{R/2,T}=\min_{B\left(\mathcal{O},\frac{R}{2}\right)\times[0,T]}v.$$

As for $m \in \left(1-\frac{2}{n}, 1\right)$, they proved that on $B(\mathcal{O}, R) \times [0,T]$, for any $\gamma \in (0,1)$,

$$\frac{|\nabla v|^2}{v} - \gamma \frac{v_t}{v} \geq \frac{\alpha\gamma^2}{C_3}\left(\frac{1}{t} + C_4\sqrt{C_3}K^2 \bar{v}_{\max}^{2R,T}\right) + \frac{\alpha\gamma^2}{C_3}\frac{\bar{v}_{\max}^{2R,T}}{R^2}C_5,$$

with

$$\bar{v}_{\max}^{2R,T} = \max_{B(\mathcal{O},2R)\times[0,T]} (-v).$$

Later in [21] several results of similar type were obtained by G. Huang, Z. Huang and H. Li. Note that these gradient bounds do not depend on the initial data. While in [40], S.T. Yau established a similar type of gradient bounds depending on derivatives of initial data for degenerate parabolic equations of the form

$$\frac{\partial u}{\partial t} = \Delta\left(F(u)\right),$$

with $F \in C^2(0,\infty)$ and $F' > 0$. In particular, as explained in [28], Yau's result implies that for any function $c(t) \in C^1(0,\infty)$ satisfying

$$\begin{cases} c(t) \leq 0 \\ c'(t) \geq 0 \\ |\nabla v|^2 - 2v_t + 2m\left(\frac{m-1}{m}v\right)^{\frac{m-2}{m-1}} \leq c(0) \quad \text{at } t=0, \end{cases}$$

it holds for all $t > 0$ that

$$|\nabla v|^2 - 2v_t + 2m\left(\frac{m-1}{m}v\right)^{\frac{m-2}{m-1}} \leq c(t).$$

Besides gradient estimates of the Aronson-Bénilan type, the Hamilton type estimate also plays an important role. It originates from Hamilton [19] where it was proved that a positive smooth solution u to the heat equation on a compact manifold without boundary and with Ric$\geq -k$, $k>0$ satisfies

$$\frac{|\nabla u|^2}{u^2} \leq \left(\frac{1}{t}+2k\right) \ln \frac{\|u\|_\infty}{u}. \tag{1.8}$$

This is an upper bound on the gradient of space variables only, hence leading to a different type of Harnack inequalities. As proved by Kotschwar [24], the same result holds for complete noncompact manifolds as well. A local version was obtained by Souplet and Zhang in [33]. As for PME and FDE, L. A. Caffarelli, J. L. Vázquez and N. I. Wolanski [12] discussed the case when the initial data is compactly supported. Namely, under the assumption that the initial data $u_0 \geq 0$ is integrable and compactly supported, they established that for $m>1$, there exists a time $T=T\left(u_0\right)>0$ and a constant $c=c\left(m,n\right)>0$ such that

$$|\nabla v\left(x,t\right)| \leq c\left(\left(\frac{v}{t}\right)^{\frac{1}{2}}+\frac{|x|}{t}\right),$$

for any $t>T$ and almost every $x\in\mathbb{R}^n$. Later X. Xu [39] derived a local result on a complete Riemannian manifold with Ric$\geq -k$ for some $k\geq 0$. For $m>1$, if there exists a constant $\delta \in \left(0, \frac{4}{n-1}\right]$ such that

$$1 \leq \frac{v_{\max}^{2R,T}}{v_{\min}^{2R,T}} < \frac{1}{1+\delta}\left(\frac{4m}{\left(n-1\right)\left(m-1\right)}+1\right),$$

then on $B\left(x_0,R\right)\times\left[t_0-\frac{T}{2},t_0\right]$

$$\frac{|\nabla v|}{v_{\max}^{2R,T}\left(1+\delta\right)-v}\leq C_6\left(m,n\right)\left(\frac{1+\delta}{2\rho\delta R}+\frac{1}{\sqrt{\frac{m-1}{m}v_{\max}^{2R,T}\delta\rho T}}+\sqrt{\frac{k}{\delta}}\right),$$

where

$$\rho=2m-\frac{\left(n-1\right)\left(m-1\right)}{2}\frac{v_{\max}^{2R,T}\left(1+\delta\right)-v_{\min}^{2R,T}}{v_{\min}^{2R,T}},$$

$$v_{\max}^{2R,T}=\sup_{B(x_0,2R)\times[t_0-T,t_0]}v,\ v_{\min}^{2R,T}=\inf_{B(x_0,2R)\times[t_0-T,t_0]}v.$$

When $m\in\left(1-\frac{4}{n+3},1\right)$, they obtained that on $B\left(x_0,R\right)\times\left[t_0-\frac{T}{2},t_0\right]$,

$$\frac{|\nabla v|}{-v}\leq C_7\left(m,n\right)\left(\frac{1}{2R}+\frac{1}{\sqrt{\frac{1-m}{m}\bar{v}_{\min}^{2R,T}T}}+\sqrt{k}\right),$$

where

$$\bar{v}_{\min}^{2R,T}=\inf_{B(x_0,2R)\times[t_0-T,t_0]}\left(-v\right).$$

This is a generalization of Li Ma et al. [28], where the same estimate was derived only for $n=2$ or 3 with $m\in\left(1-\frac{1}{\sqrt{n}},1\right)$. In X. Zhu [42], it was proved that for $m\in\left(1,1+\frac{1}{\sqrt{2n+1}}\right)$, on $B\left(x_0,R\right)\times\left[t_0-\frac{T}{2},t_0\right]$

$$v^{\frac{1}{4}\frac{2-m}{m-1}}\left|\nabla v\right|\leq C_8\left(v_{\max}^{2R,T}\right)^{1+\frac{1}{4}\frac{2-m}{m-1}}\left(\frac{1}{2R}+\frac{1}{\sqrt{T}}+\sqrt{k}\right),$$

with

$$v_{\max}^{2R,T}=\sup_{B(x_0,2R)\times[t_0-T,t_0]}v.$$

In X. Zhu [41], a gradient bound for $m \in \left(1-\frac{2}{n},1\right)$ was obtained. On $B\left(x_0,R\right)\times\left[t_0-\frac{T}{2},t_0\right]$

$$\frac{|\nabla v|}{\sqrt{-v}} \le C_9\sqrt{\bar{v}_{\max}^{2R,T}}\left(\frac{1}{2R}+\frac{1}{\sqrt{T}}+\sqrt{k}\right),$$

where

$$\bar{v}_{\max}^{2R,T} = \sup_{B(x_0,2R)\times[t_0-T,t_0]}\left(-v\right).$$

1.2 Application of stochastic analysis to gradient estimate

Employing various aspects of stochastic analysis to derive gradient estimate is not new, and there are a large number of papers devoted to the study of solutions to linear PDEs, for example [18], [1] and [31].

Denote by $P_t f\left(x\right)$ the positive solution to the heat equation with initial data f on a compact manifold with Ric$\ge -k$ for some $k\ge 0$, and by X a Brownian motion on that manifold. Among other interesting results, [1] observed the submartingale property of

$$\frac{T-t}{2\left(1+k\left(T-t\right)\right)}\frac{\left|\nabla P_{T-t}f\right|^2}{P_{T-t}f}\left(X_t\right)+\left(P_{T-t}f\log P_{T-t}f\right)\left(X_t\right),$$

and then proved

$$\left|\frac{\nabla P_T f}{P_T f}\right|^2\left(x\right)\le 2\left(\frac{1}{T}+k\right)P_T\left(\frac{f}{P_T f\left(x\right)}\log\frac{f}{P_T f\left(x\right)}\right)\left(x\right),$$

which is able to recover Hamilton's result (1.8).

In [31], the solution $P_t f(x)$ to a parabolic equation with an elliptic generator L was considered. They started from the fact that $\nabla(P_{t-s}f)(X_s)J_s$ is a martingale, where X is the diffusion with generator L, and J is the Jacobian of the stochastic flow of X. Then they applied the Burkholder-Davis-Gundy inequality by identifying $\int_0^t |\nabla(P_{t-s}f)(X_s)|^2 ds$ with the quadratic variation of $P_{t-s}(X_s)$. As a result, they deduced gradient bound in the form

$$|\nabla(P_t f)| \leq \frac{C}{\sqrt{t} \wedge 1} P_t (|f|^q)^{\frac{1}{q}},$$

with $q > 1$ and for any bounded function f.

In [18], Bismut's type formula was derived without Malliavin calculus, and it is then used to derive estimate in the form

$$|\nabla(P_t f)| \leq \frac{1}{\delta} \frac{1}{\alpha t} \sqrt{e^{\alpha t} - 1} \sup |f|.$$

On the other hand, to the best of our knowledge, there are few papers dealing with certain kind of nonlinear PDEs by using martingale methods.

Chapter 2

Gradient estimate for solutions to the heat equation

To explain our main idea, let us consider the heat equation in this chapter. Assume $u(t,x)$ solves

$$\begin{aligned} \frac{\partial u}{\partial t} &= \Delta u \text{ on } (0,\infty) \times \mathbb{R}^n \\ u(0,x) &= u_0(x) \text{ on } \mathbb{R}^n, \end{aligned} \tag{2.1}$$

with u_0 being bounded. Then the smoothing effect of the heat kernel ensures that u is smooth and that all the first and second order derivatives at each fixed time $t > 0$ are bounded. Our aim is to estimate the upper bound of ∇u in terms of the L^∞ norm of u_0.

It is known that there is a close link between a large class of parabolic PDEs and diffusion processes, in the sense that the differential operator for a PDE can be identified as a generator for a diffusion process. Once this one-to-one correspondence has been established, we are given a way to study a PDE through its diffusion process counterpart, or the other way around. For example, the solution to a PDE can be expressed in terms of the expectation of a diffusion process at a certain time. The transition probability density function of a diffusion process is the fundamental solution of a PDE. For a more comprehensive account of this area, we refer to the book [34] by Stroock and Varadhan. Here we also relates our PDE, (2.1), to a diffusion process in the way that we have just explained. For a given point $(T, x) \in (0, \infty) \times \mathbb{R}^n$, define an $n-$dimensional stochastic process X_t by solving the stochastic differential equation

$$\begin{aligned} dX_t &= \sqrt{2} dW_t \\ X_0 &= x, \end{aligned}$$

where $W = (W^1, \cdots, W^n)$ is a standard $n-$dimensional Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then we have a progressively measurable function X from $((0, \infty) \times \Omega, \mathcal{B}((0, \infty)) \otimes \mathcal{F})$ to $\mathbb{R}^n$. Bearing in mind that the aim is to get an upper bound for $|\nabla u|^2 (T, x)$, we consider the process $|\nabla u|^2 (T - t, X_t)$ with index $t \in [0, T)$, that is, the gradient running backward on diffusion process X. The way we compose $|\nabla u|^2$ with X is commonly seen. In terms of computation, this will lead to the disappearance of terms containing the time derivative when using Ito's formula. Intuitively, this is because on the one hand our diffusion process always starts at the deterministic point that we are interested in and then evolves in a certain random way as t increases, but on the other hand, due to the nature of parabolic PDEs, we need to use the information about the solution before T. So we need the time variable to decrease when t increases.

One can observe that $|\nabla u|^2 (T-t, X_t)$ is a semimartingale. Hence we are encouraged to turn to the theory of martingales, a concept introduced by Paul Lévy in 1930s and greatly developed by J.L Doob in his book [17]. First of all, let us decompose the semimartingale $\left(|\nabla u|^2 (T-t, X_t)\right)_{0\le t<T}$ into a sum of a local martingale and a process of finite variation. From now on, when there is no potential confusion, we omit the specification of variables in functions, as we always consider functions running backward on diffusion X. Taking derivatives with respect to x^α on both sides of (2.1) yields

$$\frac{\partial^2 u}{\partial t \partial x^\alpha} = \Delta \frac{\partial u}{\partial x^\alpha}. \tag{2.2}$$

Since u is smooth by our assumption, we can apply Ito's formula on $\frac{\partial u}{\partial x^\alpha}(T-t, X_t)$. From now on, we will adopt the Einstein summation convention. When an index occurs twice in one term, that term is implicitly summed up over it. From (2.2), it follows that

$$d\frac{\partial u}{\partial x^\alpha} = \sqrt{2}\frac{\partial^2 u}{\partial x^\alpha \partial x^\beta} dW_t^\beta.$$

Then, again by Ito's formula,

$$
\begin{aligned}
d\left|\nabla u\right|^2 &= 2\frac{\partial u}{\partial x^\alpha} d\frac{\partial u}{\partial x^\alpha} + \sum_\alpha d\left\langle \frac{\partial u}{\partial x^\alpha} \right\rangle \\
&= 2\sqrt{2}\frac{\partial u}{\partial x^\alpha}\frac{\partial^2 u}{\partial x^\alpha \partial x^\beta} dW_t^\beta \\
&\quad +2\sum_{\alpha,\beta=1}^{n}\left(\frac{\partial^2 u}{\partial x^\alpha \partial x^\beta}\right)^2 dt. \qquad (2.3)
\end{aligned}
$$

This is the decomposition we are looking for. It can be seen that the term of finite variation is non-negative, and the local martingale part is a true martingale under $\mathbb{P}$, a consequence of the boundedness of $\frac{\partial u}{\partial x^\alpha}$ and $\frac{\partial^2 u}{\partial x^\alpha \partial x^\beta}$ over $[0, T-\delta]$, for any $\delta > 0$ and $\alpha, \beta \in \{1, \cdots, n\}$. This means that $\left\{\left|\nabla u\right|^2 (T-t, X_t)\right\}_{0\leq t<T}$ is a submartingale[1], which implies that

$$
E\left[\left|\nabla u\right|^2 (T-t, X_t)\right] \geq E\left[\left|\nabla u\right|^2 (T, X_0)\right] = \left|\nabla u\right|^2 (T, x),
$$

for any $t \in [0, T)$, where the equality results from the fact that $X_0 = x$ $\mathbb{P}-$almost surely. By integrating both sides on t from 0 to T, we have

$$
\int_0^T E\left[\left|\nabla u\right|^2 (T-t, X_t)\right] dt \geq T\left|\nabla u\right|^2 (T, x). \qquad (2.4)
$$

On the other hand, through a scaling argument, one can see that a bound on $\left|\nabla u\right|^2 (T, x)$ involving T and $\|u\|_\infty$ should have the order $O\left(\frac{\|u\|_\infty^2}{T}\right)$. Therefore, let us consider the

---

[1] In fact, the submartingale property of $\left|\nabla u\right|^2$ follows directly from the martingale property of ∇u. What we wish to develop here is an argument that can be adopted for a more complicated case in the next chapter.

process $u^2(T-t, X_t)$. By Ito's formula,

$$u^2(0, X_T) - u^2(T, X_0) = 2\sqrt{2}\int_0^T u\frac{\partial u}{\partial x^\alpha}dW_t^\alpha + 2\int_0^T |\nabla u|^2 dt. \tag{2.5}$$

Since there exists a sequence of stopping times $\{T_n\}$ with $T_n \to \infty$ monotonically, such that $\int_0^{t\wedge T_n} u\frac{\partial u}{\partial x^\alpha}dW_t^\alpha$ is a true martingale under $\mathbb{P}$, we have

$$E\left[u^2(0, X_{T\wedge T_n})\right] - u^2(T, x) = 2E\left[\int_0^{T\wedge T_n} |\nabla u|^2 dt\right].$$

Then letting $n \to \infty$ and applying bounded convergence theorem and monotone convergence theorem on the left hand side and right hand side respectively yield

$$E\left[u^2(0, X_T)\right] - u^2(T, x) = 2E\left[\int_0^T |\nabla u|^2 dt\right]. \tag{2.6}$$

Remark 1. Since we did not assume any boundedness on derivatives of the initial data u_0, $|\nabla u|^2(t, x)$ may explode as $t \to 0$ with order $\frac{1}{t}$. So in general, $\int_0^T |\nabla u|^2(t, x)\,dt$ is not finite. However, (2.6) shows us that the averaged behavior of $\int_0^T |\nabla u|^2(t, x)\,dt$ on path space Ω is finite.

By Fubini's Theorem, $E\left[\int_0^T |\nabla u|^2 dt\right] = \int_0^T E\left[|\nabla u|^2\right] dt$. Therefore, (2.4) and (2.6) together yields

$$|\nabla u|^2(T, x) \quad \leq \quad \frac{E\left[u^2(0, X_T)\right] - u^2(T, x)}{2T}.$$

Since the law of X_T is absolutely continuous with respect to Lebesgue measure,

$u^2(0,X_T) \leq \|u_0\|_\infty^2$ $\mathbb{P}-$ almost surely, where $\|\cdot\|_\infty$ denotes the essential superemum under Lebesgue measure. Hence we have the following result.

Theorem 2. *If u solves the Cauchy problem* (2.1) *and u is smooth, bounded and has bounded derivatives with respect to space variables up to the second order, then for any $(T,x) \in (0,\infty) \times \mathbb{R}^n$, we have*

$$|\nabla u|^2 (T,x) \leq \frac{\|u_0\|_\infty^2 - u^2(T,x)}{2T} \leq \frac{\|u_0\|_\infty^2}{2T}. \tag{2.7}$$

One can see that this bound does not depend on dimension n. It also supports the fact that when function u attains its maximum, its gradient is zero. Also, we want to remark here that this method also applies to estimates of gradients of higher orders. To see this, first observe that by considering the function $\bar{u}(t,x) = u(t+\epsilon,x)$ where $\epsilon > 0$ in the above argument, we get

$$|\nabla u|^2 (T,x) \leq \frac{\|u(\epsilon,\cdot)\|_\infty^2}{2(T-\epsilon)}.$$

Since $\frac{\partial u}{\partial x^i}$ also satisfies the heat equation, it holds that

$$\left|\nabla \frac{\partial u}{\partial x^i}\right|^2 (T,x) \leq \frac{\left\|\frac{\partial u}{\partial x^i}(\epsilon,\cdot)\right\|_\infty^2}{2(T-\epsilon)}.$$

Summing over the index i yields

$$\sum_{i,j=1}^n \left|\frac{\partial^2 u}{\partial x^i \partial x^j}\right|^2 (T,x) \leq \frac{\left\||\nabla u|^2(\epsilon,\cdot)\right\|_\infty}{2(T-\epsilon)} \leq \frac{\|u_0\|_\infty^2}{4\epsilon(T-\epsilon)},$$

where the last inequality results from (2.7). Then we can minimize the right hand side by choosing $\epsilon = \frac{T}{2}$, hence obtaining

$$\sum_{i,j=1}^{n} \left| \frac{\partial^2 u}{\partial x^i \partial x^j} \right|^2 (T, x) \leq \frac{\|u_0\|_\infty^2}{T^2}.$$

Chapter 3

Gradient estimate for positive solutions to PME and FDE[1]

3.1 Local gradient estimate

Now we move our interest to equation (1.1), and use the same idea to produce a similar type of upper bounds on gradients of its solutions. We will only consider the case when the space $M = \mathbb{R}^n$. In this section we study the problem locally both in time and space. Denote by $B\left(x_0, R+\epsilon\right)$ a closed ball in $\mathbb{R}^n$ with centre x_0 and radius $R+\epsilon$, where R, $\epsilon > 0$. Let u be a positive and bounded solution to (1.1) on $[0, t_1] \times B\left(x_0, R+\epsilon\right)$. The positivity of u ensures that no degeneracy of parabolicity would happen in $[0, t_1] \times B\left(x_0, R+\epsilon\right)$. Hence we can use theory about non-degenerate quasilinear parabolic PDE to obtain that u is smooth in $[0, t_1] \times B\left(x_0, R+\epsilon\right)$. Note that at this moment u is defined only on $[0, t_1] \times B\left(x_0, R+\epsilon\right)$. This brings difficulty

---

[1] This chapter is based on a joint work with Ying Hu and Zhongmin Qian.

to our martingale method, as we will consider u running backward on a stochastic process, which takes values on the whole space $\mathbb{R}^n$ at any time. To get around this obstacle, let $\tilde{u}$ be a positive and smooth function with bounded derivatives of all orders defined on $[0, t_1] \times \mathbb{R}^n$, such that

$$u = \tilde{u} \text{ on } [0, t_1] \times B\left(x_0, R + \epsilon\right).$$

Note that as u is strictly positive and smooth on the compact set $[0, t_1] \times B\left(x_0, R + \epsilon\right)$, such $\tilde{u}$ exists. It is worthwhile to point out here that the behavior of the extended function $\tilde{u}$ outside $[0, t_1] \times B\left(x_0, R + \epsilon\right)$ will not enter into our computation in the sequel, as Li-Yau's localization technique will be adopted. Next, we take a transformation on $\tilde{u}$ by setting

$$f = \frac{\tilde{u}^h}{h}, \tag{3.1}$$

for some $h \in \mathbb{R} \setminus \{0\}$. This is a generalization of the transformation $v = \frac{u^{m-1}}{m-1}$, which repeatedly appears in literature concerning PME and FDE, such as [4], [27], [12], [39] and [41]. From (1.1) one can derive that the so-called pressure variable v satisfies

$$\frac{\partial v}{\partial t} = m\left(m - 1\right) v \Delta v + m \left|\nabla v\right|^2. \tag{3.2}$$

We can see that the exponent m in (1.1) comes down into coefficients in (3.2), and both terms on the right hand side of (3.2) are quadratic, which facilitates many computations. However, this feature is not crucial to our method. Hence we attempt to generalize this transformation with (3.1). It turns out that the flexibility in choosing

h results in an enlargement of the range of m that our gradient bound is valid for.

As we are looking for a gradient bound on a local scale, we are keen to only use the local information about f. For this purpose, we adopt the localization technique of Li and Yau [26] to introduce a cut-off function $\phi \in C^2(\mathbb{R}^n)$ satisfying

$$\phi(x) = \begin{cases} 1 & \text{on } B(x_0, R) \\ 0 & \text{on } B(x_0, R+\epsilon)^c, \end{cases}$$

$$|\Delta \phi| \leq L\phi^{\frac{1}{2}}, \tag{3.3}$$

and

$$|\nabla \phi|^2 \leq L\phi^{\frac{3}{2}}. \tag{3.4}$$

for some $L > 0$. Note that such a cut-off function exists. One choice is

$$\phi(x) = \begin{cases} 1 & \text{on } B(x_0, R) \\ 0 & \text{on } B(x_0, R+\epsilon)^c \\ \left(\left(\frac{R+\epsilon}{R}\right)^2 - 1\right)^{-4} \left(\left(\frac{R+\epsilon}{R}\right)^2 - \left|\frac{x-x_0}{R}\right|^2\right)^4 & \text{otherwise,} \end{cases} \tag{3.5}$$

and

$$L = \frac{8}{\epsilon^2}\left(\frac{n\epsilon}{2R+\epsilon} + \frac{8(R+\epsilon)^2}{(2R+\epsilon)^2}\right). \tag{3.6}$$

ϕf, the multiplication of functions ϕ and f, cuts all the information of f outside $[0, t_1] \times B(x_0, R+\epsilon)$, while faithfully preserving its behavior in $[0, t_1] \times B(x_0, R)$. It is this function that we are going to consider in the sequel.

3.1.1 From PDE to SDE

Just as in Chapter 2, let us begin by fixing a point $(T,x) \in [0,t_1] \times B(x_0,R)$. From (1.1) and the definition of f, we have on $[0,t_1] \times \mathbb{R}^n$,

$$\begin{aligned} \frac{\partial \phi f}{\partial t} &= m(hf)^{\frac{m-1}{h}} \phi \Delta f + m(m-h)(hf)^{\frac{m-h-1}{h}} \phi |\nabla f|^2 \\ &= m(hf)^{\frac{m-1}{h}} \Delta(\phi f) - m(hf)^{\frac{m-1}{h}} f \Delta \phi \\ &\quad -2m(hf)^{\frac{m-1}{h}} \nabla f \cdot \nabla \phi + m(m-h)(hf)^{\frac{m-h-1}{h}} \phi |\nabla f|^2 . \end{aligned} \tag{3.7}$$

Then let us link PDE (3.7) with the diffusion process $X = (X_t)_{0 \le t \le T}$, whose generator $\mathcal{L}$ is given by

$$\mathcal{L}_t w(y) = m(hf)^{\frac{m-1}{h}} (T-t,y) \Delta w(y), \ \forall w \in C_0^2(\mathbb{R}^n).$$

$\mathcal{L}$ corresponds to the principal part of the differential operator in (3.7). Note that by the definition of f, $(hf)^{\frac{m-1}{h}} > 0$ since $u > 0$. By [34] the way to obtain X is to solve the stochastic differential equation (SDE)

$$\begin{aligned} dX_t^\alpha &= \sqrt{2m}(hf)^{\frac{m-1}{2h}} (T-t,X_t) dW_t^\alpha && (3.8) \\ X_0 &= x, \end{aligned}$$

for $t \in [0,T]$, where $W = (W^1, \cdots, W^n)$ is a standard n-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and the stochastic integral is in Ito's sense. By our assumption on u, $(hf)^{\frac{m-1}{2h}}$ is bounded, smooth and has bounded derivatives. Therefore, by [22] this SDE has a unique strong solution on the time interval $[0,T]$.

3.1.2 Fundamental decompositions

The next step is to consider various functions running backward on process X. They are $f(T-t, X_t)$, $\phi(X_t)$, $|\nabla f|^2 (T-t, X_t)$ and so on. t will always take values in $[0,T]$, which is exactly the time interval where X lives. Since all the functions to appear in the remaining of this chapter are compositions on $(T-t, X_t)$, we will omit the specification of variables in the functions. One can see that these processes are all semimartingales, waiting for us to decompose and then releasing information. But unlike the heat equation, (3.7) is not linear. Therefore, we should be very careful in choosing semimartingales for decomposition. Let us describe our way of choosing semimartingales to be studied now. Recall that we always decompose two semimartingales, one involves the solution and the other one involves the gradient of the solution. To begin with, consider the localized solution

$$Y_t \triangleq (\phi f)(T-t, X_t).$$

It is readily seen that

$$Y_T = \phi f(0, X_T) = \phi \frac{u^h}{h}(0, X_T).$$

Since ϕf is smooth, Ito's formula can be applied to separate the local martingale part and finite variation part of Y. By (3.7) and (3.8), it holds that

$$\begin{aligned} dY_t &= \sqrt{2m}\,(hf)^{\frac{m-1}{2h}}\frac{\partial(f\phi)}{\partial x^\alpha}dW_t^\alpha - m(m-h)(hf)^{\frac{m-h-1}{h}}\phi\,|\nabla f|^2\,dt \\ &\quad+\left(\Delta\phi m(hf)^{\frac{m-1}{h}}f + 2m(hf)^{\frac{m-1}{h}}\frac{\partial f}{\partial x^\alpha}\frac{\partial \phi}{\partial x^\alpha}\right)dt. \end{aligned} \tag{3.9}$$

To find the second semimartingale for decomposition, observe that in the finite variation part of (3.9), the term with highest degree in $|\nabla f|$ is

$$-m(m-h)(hf)^{\frac{m-h-1}{h}}\phi\,|\nabla f|^2(T-t,X_t)\,dt.$$

As in our argument for the heat equation, with the help of (3.9) this term can be controlled by $\|u(0,\cdot)\|_\infty$. Therefore, it is this term that is worth investigation. Define

$$H_t = (hf)^{\frac{m-h-1}{h}}\phi\,|\nabla f|^2(T-t,X_t),\ t\in[0,T]\subseteq[0,t_1].$$

The function $(hf)^{\frac{m-h-1}{h}}\phi\,|\nabla f|^2$ is smooth on $[0,t_1]\times\mathbb{R}^n$. Its dynamics only depend on the behavior of u within $[0,t_1]\times B(x_0,R+\epsilon)$, which is governed by (1.1). Therefore, we can use Ito's formula to decompose H into local martingale part and finite variation part as follows.

$$
\begin{aligned}
dH_t &= d\left((hf)^{\frac{m-h-1}{h}} \phi \sum_{\alpha} \left| \frac{\partial f}{\partial x^{\alpha}} \right|^2 \right) \\
&= 2\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}} \phi \frac{\partial^2 f}{\partial x^{\alpha} \partial x^{\beta}} \frac{\partial f}{\partial x^{\alpha}} dW_t^{\beta} \\
&\quad + \sqrt{2m}\,(m-h-1)\,(hf)^{\frac{3m-4h-3}{2h}} \phi \left|\nabla f\right|^2 \frac{\partial f}{\partial x^{\beta}} dW_t^{\beta} \\
&\quad + \sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}} \frac{\partial \phi}{\partial x^{\beta}} \left|\nabla f\right|^2 dW_t^{\beta} \\
&\quad - m\,(2m-h+1)\,(m-h-1)\,(hf)^{\frac{2m-3h-2}{h}} \phi \left|\nabla f\right|^4 dt \\
&\quad + 2m\,(m-h-1)\,(hf)^{\frac{2m-2h-2}{h}} \frac{\partial \phi}{\partial x^{\alpha}} \frac{\partial f}{\partial x^{\alpha}} \left|\nabla f\right|^2 dt \\
&\quad + m\,(hf)^{\frac{2m-h-2}{h}} \Delta\phi \left|\nabla f\right|^2 dt \\
&\quad - 2m\,(m-1)\,(hf)^{\frac{2m-2-2h}{h}} \phi \Delta f \left|\nabla f\right|^2 dt \\
&\quad - 4m\,(hf)^{\frac{2m-2h-2}{h}} \phi \frac{\partial f}{\partial x^{\beta}} \frac{\partial^2 f}{\partial x^{\beta} \partial x^{\alpha}} \frac{\partial f}{\partial x^{\alpha}} dt \\
&\quad + 4m\,(hf)^{\frac{2m-2-h}{h}} \frac{\partial \phi}{\partial x^{\beta}} \frac{\partial^2 f}{\partial x^{\alpha} \partial x^{\beta}} \frac{\partial f}{\partial x^{\alpha}} dt \\
&\quad + 2m\,(hf)^{\frac{2m-h-2}{h}} \phi \sum_{\alpha,\beta} \left(\frac{\partial^2 f}{\partial x^{\alpha} \partial x^{\beta}} \right)^2 dt.
\end{aligned} \tag{3.10}
$$

This is the decomposition under measure $\mathbb{P}$ for semimartingale H. But in order to get more flexibility, we need to take advantage of an important tool in stochastic analysis- the change of measure. Let us introduce a family of probability measures on $(\Omega, \mathcal{F}_T)$ depending on a parameter $\lambda \in \mathbb{R}$. To remain the orders of all terms unchanged, we consider a measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}_T)$ in the form of

$$
\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = \exp\left(\int_0^T \sqrt{2m}\lambda\,(hf)^{\frac{m-2h-1}{2h}} \frac{\partial f}{\partial x^{\beta}} dW^{\beta} - m\lambda^2 \int_0^T (hf)^{\frac{m-2h-1}{h}} \left|\nabla f\right|^2 dt \right). \tag{3.11}
$$

Since $(hf)^{\frac{m-2h-1}{2h}}\frac{\partial f}{\partial x^\beta} = u^{\frac{m-3}{2}}\frac{\partial u}{\partial x^\beta}$, $\beta \in \{1, \cdots, n\}$ are bounded by our assumption, Novikov's condition [32] is met. So the right hand side of (3.11) is a true martingale under $\mathbb{P}$, which ensures that measure $\mathbb{Q}$ is a probability measure. According to the Girsanov's Theorem, under measure $\mathbb{Q}$, the process $\tilde{W} = \left(\tilde{W}^1, \cdots, \tilde{W}^n\right)$ given by

$$d\tilde{W}_t^\beta = dW_t^\beta - \sqrt{2m}\lambda\,(hf)^{\frac{m-2h-1}{2h}}\frac{\partial f}{\partial x^\beta}dt,$$

is an $n-$dimensional Brownian motion. So from (3.10), we easily get the decomposition for H under measure $\mathbb{Q}$, which is

$$\begin{aligned}
dH_t &= 2\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\,\phi\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}d\tilde{W}_t^\beta \\
&\quad +\sqrt{2m}\,(m-h-1)\,(hf)^{\frac{3m-4h-3}{2h}}\,\phi\,|\nabla f|^2\,\frac{\partial f}{\partial x^\beta}d\tilde{W}_t^\beta \\
&\quad +\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\,\frac{\partial \phi}{\partial x^\beta}\,|\nabla f|^2\,d\tilde{W}_t^\beta \\
&\quad -m\,(2m-h+1-2\lambda)\,(m-h-1)\,(hf)^{\frac{2m-3h-2}{h}}\,\phi\,|\nabla f|^4\,dt \\
&\quad +2m\,(m-h-1+\lambda)\,(hf)^{\frac{2m-2h-2}{h}}\,\frac{\partial \phi}{\partial x^\alpha}\frac{\partial f}{\partial x^\alpha}\,|\nabla f|^2\,dt \\
&\quad +m\,(hf)^{\frac{2m-h-2}{h}}\,\Delta\phi\,|\nabla f|^2\,dt \\
&\quad -2m\,(m-1)\,(hf)^{\frac{2m-2-2h}{h}}\,\phi\Delta f\,|\nabla f|^2\,dt \\
&\quad +4m\,(\lambda-1)\,(hf)^{\frac{2m-2h-2}{h}}\,\phi\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}\frac{\partial f}{\partial x^\beta}dt \\
&\quad +4m\,(hf)^{\frac{2m-2-h}{h}}\,\frac{\partial \phi}{\partial x^\beta}\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}dt \\
&\quad +2m\,(hf)^{\frac{2m-h-2}{h}}\,\phi\sum_{\alpha,\beta}\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\right)^2 dt. \qquad (3.12)
\end{aligned}$$

While from (3.9), it is easy to see that under measure $\mathbb{Q}$,

$$\begin{aligned} & dY_t \\ = & \sqrt{2m}\left(hf\right)^{\frac{m-1}{2h}}\frac{\partial\left(f\phi\right)}{\partial x^\alpha}d\tilde{W}_t^\alpha \\ & +m\left(2\lambda-m+h\right)\left(hf\right)^{\frac{m-h-1}{h}}\phi\left|\nabla f\right|^2dt \\ & +\Delta\phi mh^{-1}\left(hf\right)^{\frac{m-1+h}{h}}dt \\ & +2m\left(1+\lambda h^{-1}\right)\left(hf\right)^{\frac{m-1}{h}}\frac{\partial f}{\partial x^\alpha}\frac{\partial\phi}{\partial x^\alpha}dt. \end{aligned} \tag{3.13}$$

Now we have (3.13) and (3.12) at hand, which are the fundamental decompositions we have been looking for.

3.1.3 Gradient bound for solution u

In the case of the heat equation, we obtained two estimates from decompositions of the gradient and the solution. One is the submartingale property of $\left|\nabla u\right|^2\left(T-t,X_t\right)$. The other is a control on $E\left[\int_0^T\left|\nabla u\right|^2\left(T-t,X_t\right)dt\right]$. A combination of them yielded the result. Here, our goal is the same, i.e., the submartingale property of H and a control on $E\left[\int_0^T H_tdt\right]$. In contrast to (2.3), which appears in the case of global estimate for the heat equation, (3.12) is more complicated, thanks to the nonlinearity of PDE (1.1) and the introduction of the cut-off function ϕ. Since we aim to estimate the first order derivatives of f, it is reasonable to get rid off the second order derivatives of f appearing in the decomposition (3.12). This is done in the following lemma. For simplicity, denote by $\mathcal{A}_{x_0}^{R,\epsilon}$ the annulus $B\left(x_0,R+\epsilon\right)\setminus B\left(x_0,R\right)$.

Lemma 3. *H satisfies*

$$
\begin{aligned}
& dH_t \\
\geq\ & A_t d\tilde{W}_t \\
& -\frac{1}{2} mn\left(m-1\right)^2 \left(hf\right)^{\frac{2m-3h-2}{h}} \phi \left|\nabla f\right|^4 dt \\
& -m\left(\left(2m-h+1-2\lambda\right)\left(m-h-1\right)+2\left(\lambda-1\right)\left(\lambda-m\right)\right)\left(hf\right)^{\frac{2m-3h-2}{h}} \phi \left|\nabla f\right|^4 dt \\
& +2m\left(2m-h-\lambda\right)\left(hf\right)^{\frac{2m-2h-2}{h}} \frac{\partial \phi}{\partial x^\alpha}\frac{\partial f}{\partial x^\alpha}\left|\nabla f\right|^2 dt \\
& +m\left(hf\right)^{\frac{2m-h-2}{h}} \Delta\phi \left|\nabla f\right|^2 dt \\
& -2m\left(hf\right)^{\frac{2m-2-h}{h}} \phi^{-1} I_{\mathcal{A}_{x_0}^{R,\epsilon}} \left|\nabla \phi\right|^2 \left|\nabla f\right|^2 dt, \qquad (3.14)
\end{aligned}
$$

where

$$
\begin{aligned}
A^\beta =\ & 2\sqrt{2m}\left(hf\right)^{\frac{3m-2h-3}{2h}} \phi \frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha} \\
& +\sqrt{2m}\left(m-h-1\right)\left(hf\right)^{\frac{3m-4h-3}{2h}} \phi \left|\nabla f\right|^2 \frac{\partial f}{\partial x^\beta} \\
& +\sqrt{2m}\left(hf\right)^{\frac{3m-2h-3}{2h}} \frac{\partial \phi}{\partial x^\beta}\left|\nabla f\right|^2 . \qquad (3.15)
\end{aligned}
$$

Proof. From (3.12), let us write

$$dH_t \triangleq A_t d\tilde{W}_t + \left(B+C+D\right)dt,$$

where

$$\begin{aligned} A^{\beta} &= 2\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\,\phi\frac{\partial^2 f}{\partial x^{\alpha}\partial x^{\beta}}\frac{\partial f}{\partial x^{\alpha}} \\ &\quad +\sqrt{2m}\,(m-h-1)\,(hf)^{\frac{3m-4h-3}{2h}}\,\phi\,|\nabla f|^2\,\frac{\partial f}{\partial x^{\beta}} \\ &\quad +\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\,\frac{\partial \phi}{\partial x^{\beta}}\,|\nabla f|^2\,, \end{aligned}$$

and by the Cauchy-Schwartz inequality,

$$\begin{aligned} B &= \sum_{\alpha}\Bigg(2m\,(hf)^{\frac{2m-h-2}{h}}\,\phi\left(\frac{\partial^2 f}{\partial x^{\alpha}\partial x^{\alpha}}\right)^2 \\ &\quad -2m\,(m-1)\,(hf)^{\frac{2m-2-2h}{h}}\,\phi\frac{\partial^2 f}{\partial x^{\alpha}\partial x^{\alpha}}\,|\nabla f|^2 \\ &\quad +4m\,(\lambda-1)\,(hf)^{\frac{2m-2h-2}{h}}\,\phi\frac{\partial^2 f}{\partial x^{\alpha}\partial x^{\alpha}}\frac{\partial f}{\partial x^{\alpha}}\frac{\partial f}{\partial x^{\alpha}} \\ &\quad +4m\,(hf)^{\frac{2m-2-h}{h}}\,\frac{\partial \phi}{\partial x^{\alpha}}\frac{\partial^2 f}{\partial x^{\alpha}\partial x^{\alpha}}\frac{\partial f}{\partial x^{\alpha}}\Bigg) \\ &\geq -2m\,(hf)^{\frac{2m-2-h}{h}}\,\phi^{-1}I_{\mathcal{A}_{x_0}^{R,\epsilon}}\sum_{\alpha}\left(\frac{\partial \phi}{\partial x^{\alpha}}\right)^2\left(\frac{\partial f}{\partial x^{\alpha}}\right)^2 \\ &\quad -2m\,(\lambda-1)^2\,(hf)^{\frac{2m-3h-2}{h}}\,\phi\sum_{\alpha}\left(\frac{\partial f}{\partial x^{\alpha}}\right)^4 \\ &\quad -4m\,(\lambda-1)\,(hf)^{\frac{2m-2h-2}{h}}\sum_{\alpha}\frac{\partial \phi}{\partial x^{\alpha}}\left(\frac{\partial f}{\partial x^{\alpha}}\right)^3 \\ &\quad +2m\,(m-1)\,(hf)^{\frac{2m-2-2h}{h}}\,|\nabla f|^2\,\frac{\partial \phi}{\partial x^{\alpha}}\frac{\partial f}{\partial x^{\alpha}} \\ &\quad +\left(\lambda-1-\frac{1}{4}n\,(m-1)\right)2m\,(m-1)\,(hf)^{\frac{2m-2-3h}{h}}\,\phi\,|\nabla f|^4\,, \end{aligned}$$

$$
\begin{aligned}
C &= \sum_{\alpha\neq\beta}\Bigg(2m\,(hf)^{\frac{2m-h-2}{h}}\,\phi\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\right)^2 \\
&\quad +4m\,(\lambda-1)\,(hf)^{\frac{2m-2h-2}{h}}\,\phi\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}\frac{\partial f}{\partial x^\beta} \\
&\quad +4m\,(hf)^{\frac{2m-2-h}{h}}\,\frac{\partial \phi}{\partial x^\beta}\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}\Bigg) \\
&\geq \sum_{\alpha\neq\beta}\Bigg(-2m\,(\lambda-1)^2\,(hf)^{\frac{2m-3h-2}{h}}\,\phi\left(\frac{\partial f}{\partial x^\alpha}\frac{\partial f}{\partial x^\beta}\right)^2 \\
&\quad -2m\,(hf)^{\frac{2m-2-h}{h}}\,\phi^{-1} I_{\mathcal{A}^{R,\epsilon}_{x_0}}\left(\frac{\partial \phi}{\partial x^\beta}\frac{\partial f}{\partial x^\alpha}\right)^2 \\
&\quad -4m\,(\lambda-1)\,(hf)^{\frac{2m-2h-2}{h}}\,\frac{\partial \phi}{\partial x^\beta}\frac{\partial f}{\partial x^\alpha}\frac{\partial f}{\partial x^\alpha}\frac{\partial f}{\partial x^\beta}\Bigg),
\end{aligned}
$$

and

$$
\begin{aligned}
D &= -m\,(2m-h+1-2\lambda)\,(m-h-1)\,(hf)^{\frac{2m-3h-2}{h}}\,\phi\,|\nabla f|^4 \\
&\quad +m\,(hf)^{\frac{2m-h-2}{h}}\,\Delta\phi\,|\nabla f|^2 \\
&\quad +2m\,(m-h-1+\lambda)\,(hf)^{\frac{2m-2h-2}{h}}\,\frac{\partial \phi}{\partial x^\alpha}\frac{\partial f}{\partial x^\alpha}\,|\nabla f|^2 .
\end{aligned}
$$

Then adding these three inequalities gives us (3.14). □

To deal with the finite variation part in (3.14), let us assume that

$$
\frac{1}{2}n\,(m-1)^2 + (2m-h+1-2\lambda)\,(m-h-1) + 2\,(\lambda-1)\,(\lambda-m) < 0. \tag{3.16}
$$

This allows us to get the following estimate.

Lemma 4. *Under assumption* (3.16), *we have*

$$E^{\mathbb{Q}}\left[\int_0^T H_s ds\right] \geq H_0 T - \left(\frac{1}{2}\left|2m-h-\lambda\right| l^{-3} + \frac{3}{2} l^{-1}\right) m M_1 L^2 \frac{T^2}{2}, \tag{3.17}$$

where

$$l = \frac{\frac{1}{2} n \left(m-1\right)^2 + \left(2m-h+1-2\lambda\right)\left(m-h-1\right) + 2\left(\lambda-1\right)\left(\lambda-m\right)}{-\frac{3}{2}\left|2m-h-\lambda\right| - \frac{3}{2}}, \tag{3.18}$$

and

$$M_1 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} u^{2m+h-2}\left(t,y\right).$$

Proof. By the Cauchy-Schwartz inequality, for any positive l_1, l_2, l_3 and l_4,

$$\begin{aligned}
&\left|\left(hf\right)^{\frac{2m-2h-2}{h}} \frac{\partial \phi}{\partial x^\alpha}\frac{\partial f}{\partial x^\alpha}\left|\nabla f\right|^2\right| \\
\leq& \frac{1}{2} l_1 \left(hf\right)^{\frac{2m-3h-2}{h}} \phi \left|\nabla f\right|^4 + \frac{1}{2} l_1^{-1} \left(hf\right)^{\frac{2m-h-2}{h}} \phi^{-1} \left|\nabla \phi\right|^2 \left|\nabla f\right|^2 \\
\leq& \frac{1}{2} l_1 \left(hf\right)^{\frac{2m-3h-2}{h}} \phi \left|\nabla f\right|^4 + \frac{1}{4} l_1^{-1} l_2 \left(hf\right)^{\frac{2m-3h-2}{h}} \phi \left|\nabla f\right|^4 \\
&+ \frac{1}{4} l_1^{-1} l_2^{-1} \left(hf\right)^{\frac{2m+h-2}{h}} \phi^{-3} \left|\nabla \phi\right|^4,
\end{aligned}$$

$$\begin{aligned}
&\left|\left(hf\right)^{\frac{2m-h-2}{h}} \phi^{-1} \left|\nabla \phi\right|^2 \left|\nabla f\right|^2\right| \\
\leq& \frac{1}{2} l_3 \left(hf\right)^{\frac{2m-3h-2}{h}} \phi \left|\nabla f\right|^4 + \frac{1}{2} l_3^{-1} \left(hf\right)^{\frac{2m+h-2}{h}} \phi^{-3} \left|\nabla \phi\right|^4,
\end{aligned}$$

and

$$\begin{aligned}&\left|(hf)^{\frac{2m-h-2}{h}}\Delta\phi\left|\nabla f\right|^2\right|\\ \leq&\frac{1}{2}l_4\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4+\frac{1}{2}l_4^{-1}\left(hf\right)^{\frac{2m+h-2}{h}}\phi^{-1}\left|\Delta\phi\right|^2.\end{aligned}$$

Plugging them into (3.14) in the above lemma yields

$$\begin{aligned}&dH_t\\ \geq&A_t d\tilde{W}_t\\ &-\frac{1}{2}mn\left(m-1\right)^2\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4dt\\ &-m\left(\left(2m-h+1-2\lambda\right)\left(m-h-1\right)+2\left(\lambda-1\right)\left(\lambda-m\right)\right)\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4dt\\ &-2m\left|2m-h-\lambda\right|\left(\frac{1}{2}l_1+\frac{1}{4}l_1^{-1}l_2\right)\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4dt\\ &-ml_3\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4dt\\ &-\frac{1}{2}ml_4\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4dt\\ &-\frac{1}{2}ml_4^{-1}\left(hf\right)^{\frac{2m+h-2}{h}}\phi^{-1}\left|\Delta\phi\right|^2dt\\ &-\frac{1}{2}m\left|2m-h-\lambda\right|l_1^{-1}l_2^{-1}\left(hf\right)^{\frac{2m+h-2}{h}}\phi^{-3}\left|\nabla\phi\right|^4dt\\ &-ml_3^{-1}\left(hf\right)^{\frac{2m+h-2}{h}}\phi^{-3}\left|\nabla\phi\right|^4dt.\end{aligned}$$

For simplicity, set $l_1^2 = l_2 = l_3^2 = l_4^2 = l^2$. Then we have

$$
\begin{aligned}
& dH_t \\
\geq & A_t d\tilde{W}_t \\
& - \frac{1}{2} mn (m-1)^2 (hf)^{\frac{2m-3h-2}{h}} \phi |\nabla f|^4 dt \\
& - m \left((2m-h+1-2\lambda)(m-h-1) + 2(\lambda-1)(\lambda-m) \right) (hf)^{\frac{2m-3h-2}{h}} \phi |\nabla f|^4 dt \\
& - \frac{3}{2} m |2m-h-\lambda| \, l \, (hf)^{\frac{2m-3h-2}{h}} \phi |\nabla f|^4 dt \\
& - \frac{3}{2} ml (hf)^{\frac{2m-3h-2}{h}} \phi |\nabla f|^4 dt \\
& - \frac{1}{2} m \left(|2m-h-\lambda| \, l^{-3} + 2l^{-1} \right) (hf)^{\frac{2m+h-2}{h}} \phi^{-3} |\nabla \phi|^4 dt \\
& - \frac{1}{2} ml^{-1} (hf)^{\frac{2m+h-2}{h}} \phi^{-1} |\Delta \phi|^2 dt.
\end{aligned}
$$

By estimates for the cut-off function ϕ in (3.3) and (3.4),

$$
\begin{aligned}
& dH_t \\
\geq & A_t d\tilde{W}_t \\
& - \frac{1}{2} mn (m-1)^2 (hf)^{\frac{2m-3h-2}{h}} \phi |\nabla f|^4 dt \\
& - m \left((2m-h+1-2\lambda)(m-h-1) + 2(\lambda-1)(\lambda-m) \right) (hf)^{\frac{2m-3h-2}{h}} \phi |\nabla f|^4 dt \\
& - \frac{3}{2} m |2m-h-\lambda| \, l \, (hf)^{\frac{2m-3h-2}{h}} \phi |\nabla f|^4 dt \\
& - \frac{3}{2} ml (hf)^{\frac{2m-3h-2}{h}} \phi |\nabla f|^4 dt \\
& - \frac{1}{2} m \left(|2m-h-\lambda| \, l^{-3} + 3l^{-1} \right) (hf)^{\frac{2m+h-2}{h}} I_{\mathcal{A}_{x_0}^{R,\epsilon}} L^2 dt.
\end{aligned}
$$

Now, by assumption (3.16), we can choose a positive l small enough such that

$$\begin{aligned}&-\frac{1}{2}mn\left(m-1\right)^2-m\left(\left(2m-h+1-2\lambda\right)\left(m-h-1\right)+2\left(\lambda-1\right)\left(\lambda-m\right)\right)\\=&\left(\frac{3}{2}\left|2m-h-\lambda\right|+\frac{3}{2}\right)ml.\end{aligned}$$

Hence we have for any $t\in[0,T]$,

$$dH_t \;\geq\; A_t d\tilde{W}_t-\left(\frac{1}{2}\left|2m-h-\lambda\right|l^{-3}+\frac{3}{2}l^{-1}\right)mM_1L^2dt,$$

where

$$M_1=\sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}}\left|\left(hf\right)^{\frac{2m+h-2}{h}}\left(t,y\right)\right|=\sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}}u^{2m+h-2}\left(t,y\right).$$

This gives

$$H_s-H_0\geq\int_0^s A_t d\tilde{W}_t-\left(\frac{1}{2}\left|2m-h-\lambda\right|l^{-3}+\frac{3}{2}l^{-1}\right)mM_1L^2s, \tag{3.19}$$

for any $s\in[0,T]$. According to our assumption on u, A, defined by (3.15), is bounded. So $\int_0^\cdot A_s d\tilde{W}_s$ is a true martingale under $\mathbb{Q}$. Therefore we can take expectation on both sides of (3.19) and obtain

$$E^{\mathbb{Q}}\left[H_s\right]\geq H_0-\left(\frac{1}{2}\left|2m-h-\lambda\right|l^{-3}+\frac{3}{2}l^{-1}\right)mM_1L^2s. \tag{3.20}$$

Then, integrating both sides with respect to s on $[0,T]$ yields (3.17). $\square$

On the other hand, by using (3.9), we have an upper bound on the left hand side of (3.17).

Lemma 5. *Assume*

$$2\lambda - m + h > 0. \tag{3.21}$$

Then for any $\rho > 0$ *such that*

$$2\lambda - m + h - \rho\left|1 + \lambda h^{-1}\right| > 0, \tag{3.22}$$

we have

$$E^{\mathbb{Q}}\left[\int_0^T H_t dt\right] \leq \frac{E^{\mathbb{Q}}\left[Y_T\right] - Y_0 + \left(\left|h^{-1}\right| + \rho^{-1}\left|1 + \lambda h^{-1}\right|\right) mLM_2T}{m\left(2\lambda - m + h - \rho\left|1 + \lambda h^{-1}\right|\right)}, \tag{3.23}$$

with

$$M_2 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R;\epsilon}} u^{m+h-1}(t,y).$$

Proof. Recall (3.13), which is the fundamental decomposition of Y under measure $\mathbb{Q}$. Then by property (3.3) of the cut-off function ϕ and the Cauchy-Schwartz inequality, for any $\rho > 0$,

$$\begin{aligned}
& dY_t \\
\geq\ & \sqrt{2m}\,(hf)^{\frac{m-1}{2h}} \frac{\partial\,(f\phi)}{\partial x^\alpha} d\tilde{W}_t^\alpha \\
& +m\,(2\lambda - m + h)\,(hf)^{\frac{m-h-1}{h}}\,\phi\,|\nabla f|^2\,dt \\
& -L\phi^{\frac{1}{2}} I_{\mathcal{A}_{x_0}^{R,\epsilon}} m\,\left|h^{-1}\right|(hf)^{\frac{m-1+h}{h}}\,dt \\
& -\rho m\,\left|1+\lambda h^{-1}\right|(hf)^{\frac{m-1-h}{h}}\,\phi\,|\nabla f|^2\,dt \\
& -\rho^{-1} m\,\left|1+\lambda h^{-1}\right|(hf)^{\frac{m-1+h}{h}}\,I_{\mathcal{A}_{x_0}^{R,\epsilon}}\phi^{-1}\,|\nabla \phi|^2\,dt \\
\geq\ & \sqrt{2m}\,(hf)^{\frac{m-1}{2h}} \frac{\partial\,(f\phi)}{\partial x^\alpha} d\tilde{W}_t^\alpha \\
& +m\left(2\lambda - m + h - \rho\left|1+\lambda h^{-1}\right|\right) H_t dt \\
& -\left(\left|h^{-1}\right| + \rho^{-1}\left|1+\lambda h^{-1}\right|\right) mLM_2 dt, \qquad (3.24)
\end{aligned}$$

where

$$M_2 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} (hf)^{\frac{m+h-1}{h}}\,(t,y) = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} u^{m+h-1}\,(t,y)\,.$$

According to our assumption on u, $(hf)^{\frac{m-1}{2h}} \frac{\partial(f\phi)}{\partial x^\alpha}$, $\alpha \in \{1,\cdots,n\}$ are all bounded, which ensures that $\int_0^\cdot (hf)^{\frac{m-1}{2h}} \frac{\partial(f\phi)}{\partial x^\alpha} d\tilde{W}_t^\alpha$ is a martingale under $\mathbb{Q}$. Hence

$$\begin{aligned}
& E^{\mathbb{Q}}\left[Y_T\right] - Y_0 \\
\geq & m\left(2\lambda - m + h - \rho\left|1+\lambda h^{-1}\right|\right) E^{\mathbb{Q}}\left[\int_0^T H_t dt\right] \\
& -\left(\left|h^{-1}\right| + \rho^{-1}\left|1+\lambda h^{-1}\right|\right) mLM_2 T.
\end{aligned}$$

Dividing both sides by $m\,(2\lambda - m + h - \rho\,|1+\lambda h^{-1}|)$ yields (3.23). $\square$

Now, based on estimates (3.17) and (3.23), we are ready to obtain a gradient bound for u.

Proposition 6. *Let the assumptions in Lemma 4 and Lemma 5 be satisfied, then for any* $(T,x) \in (0,t_1) \times B(x_0,R)$,

$$
\begin{aligned}
& u^{m+h-3} |\nabla u|^2 (T,x) \\
\leq \ & \frac{E^{\mathbb{Q}}\left[\phi u_0^h (X_T)\right] - u^h(T,x)}{hm\left(2\lambda - m + h - \rho\left|1+\lambda h^{-1}\right|\right)T} + \frac{\left(\left|h^{-1}\right| + \rho^{-1}\left|1+\lambda h^{-1}\right|\right)LM_2}{2\lambda - m + h - \rho\left|1+\lambda h^{-1}\right|} \\
& + \left(\frac{1}{2}\left|2m-h-\lambda\right| l^{-3} + \frac{3}{2} l^{-1}\right) m M_1 L^2 \frac{T}{2}. \qquad (3.25)
\end{aligned}
$$

Proof. Combining (3.17) with (3.23), we have

$$
\begin{aligned}
& H_0 \\
\leq & \frac{E^{\mathbb{Q}}[Y_T] - Y_0 + \left(\left|h^{-1}\right| + \rho^{-1}\left|1+\lambda h^{-1}\right|\right) mLM_2T}{m\left(2\lambda - m + h - \rho\left|1+\lambda h^{-1}\right|\right)(T-t_1)} \\
& + \left(\frac{1}{2}\left|2m-h-\lambda\right| l^{-3} + \frac{3}{2} l^{-1}\right) m M_1 L^2 \frac{T}{2}.
\end{aligned}
$$

Then we arrive at (3.25) by noting that

$$
H_0 = (hf)^{\frac{m-h-1}{h}} \phi |\nabla f|^2 (T,x), \ Y_0 = \phi f(T,x) = \phi \frac{u^h}{h}(T,x),
$$

and

$$
E^{\mathbb{Q}}[Y_T] = E^{\mathbb{Q}}\left[\phi \frac{u^h}{h}(0,X_T)\right].
$$

□

Note that this result relies on assumptions (3.16) and (3.21). One can see that the consequence of these two assumptions is a restriction on the choice of m. To work out the restriction explicitly, let us look at (3.16) first. It is equivalent to

$$2\lambda^2 - 2(2m-h)\lambda + \frac{1}{2}n(m-1)^2 + (2m-h+1)(m-h-1) + 2m < 0. \tag{3.26}$$

To let the left hand side attain the minimum, we take $\lambda = \frac{2m-h}{2}$. Then (3.26) is reduced to

$$h^2 - 2mh + n(m-1)^2 + 2(m-1) < 0. \tag{3.27}$$

Notice that after setting $\lambda = \frac{2m-h}{2}$, condition (3.21) is automatically satisfied. Define

$$h_- = m - \sqrt{1-(n-1)(m-1)^2},\ h_+ = m + \sqrt{1-(n-1)(m-1)^2}.$$

We have our first main result in this section as follows.

Theorem 7. *Assume u is a positive and bounded solution to* (1.1) *on* $(0,t_1) \times B(x_0, R+\epsilon)$ *with* $m \in \left(1-\frac{1}{\sqrt{n-1}}, 1+\frac{1}{\sqrt{n-1}}\right)$. *Let the cut-off function ϕ and the constant L be defined according to* (3.5) *and* (3.55). *Furthermore, let the diffusion X and the measure $\mathbb{Q}$ be defined according to* (3.8) *and* (3.11) *with* $\lambda = \frac{2m-h}{2}$. *Then for* *any* $h \in (h_-, h_+)$*,* $\rho \in \left(0, m\left|\frac{2m+h}{2h}\right|^{-1}\right)$*, and* $(T,x) \in (0,t_1) \times B(x_0, R)$*, we have*

$$\begin{aligned}
& u^{m+h-3}|\nabla u|^2(T,x) \\
\leq\ & \frac{E^{\mathbb{Q}}\left[\phi u_0^h(X_T)\right] - u^h(T,x)}{hm\left(m-\rho\left|\frac{2m+h}{2h}\right|\right)T} + \frac{\left(\rho + h\left|\frac{2m+h}{2h}\right|\right)LM_2}{|h|\rho\left(m-\rho\left|\frac{2m+h}{2h}\right|\right)} \\
& + \left(|2m-h|\,l^{-3} + 6l^{-1}\right)mM_1L^2\frac{T}{8},
\end{aligned} \tag{3.28}$$

where

$$l = \frac{(h-h_-)(h_+-h)}{3\left|m-\frac{h}{2}\right|+3},$$

$$M_1 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} u^{2m+h-2},\ M_2 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} u^{m+h-1}.$$

Proof. It is easy to check that when $\lambda = \frac{2m-h}{2}$, $m \in \left(1-\frac{1}{\sqrt{n-1}}, 1+\frac{1}{\sqrt{n-1}}\right)$ and $h \in (h_-, h_+)$, conditions (3.16) and (3.21) are satisfied. By substituting $\lambda = \frac{2m-h}{2}$ into (3.22) and (3.18), we have

$$m - \rho\left|\frac{2m+h}{2h}\right| > 0 \text{ and } l = \frac{(h-h_-)(h_+-h)}{3\left|m-\frac{h}{2}\right|+3}.$$

Then the result follows from the proposition above. □

Let us spend some time to look at this result. First of all, except for the last term, our bound has the same order in u and t as results in [42], [41] and [39]. On the other hand, our result is essentially about estimating the gradient at time t_0 with the information of u during $[t_0-T, t_0]$, while the other results we listed in the preliminaries need the information of u during $[t_0-2T, t_0]$. This may be the reason why our bound has an extra term linear in T. Secondly, there is some flexibility in choosing h and λ, which are the parameters we introduced when taking transformation and changing the measure. This flexibility has contributed to the enlargement the range of m. One can see that when $n = 1$, m is allowed to take any value. While when $n > 1$, our constraint on m, $\left(1 - \frac{1}{\sqrt{n-1}}, 1 + \frac{1}{\sqrt{n-1}}\right)$, is looser for FDE comparing to all the existing results that we are aware of. More precisely, it requires in [28] that $m \in \left(1 - \frac{1}{\sqrt{n}}, 1\right)$, in [39] that $m \in \left(1 - \frac{4}{n+3}, 1\right)$ and in [4], [27] and [41] that $m \in \left(1 - \frac{2}{n}, 1\right]$. Moreover, one can obtain a gradient bound for solutions to the heat equation by choosing h properly (for example, $h = m$) such that (3.27) holds when $m = 1$. It is worthwhile to point out that in the local case we can not choose $h = 0$ when $m = 1$, because condition (3.16) does not hold, whatever λ is. Moreover when $h = m - 1$, condition (3.27) fails to hold for any m. So the widely used transformation $v = \frac{u^{m-1}}{m-1}$ in many literature is not suitable for our method.

Thirdly, although it is not allowed to take $h = 0$ when $m = 1$, we can let $h \to 0$ when m is away from 1, which corresponds to the transformation $f = \log u$. This will yields a gradient bound for the pressure variable $v = \frac{u^{m-1}}{m-1}$. We will see that it has the same order in v and imposes the same constraint on m as [4], [27] and [41] did for FDE. This range of m is also where the Aronson-Bénilan estimate and the ZKB solution are valid for in the case of FDE. This is done in the following subsection.

3.1.4 Gradient bound for pressure variable $v = \frac{u^{m-1}}{m-1}$

Since most of the existing results are in terms of $\frac{|\nabla v|^2}{v}$, $v = \frac{u^{m-1}}{m-1}$, we are tempted to work out a bound for it with our method. Notice that $\frac{|\nabla v|^2}{v} = (m-1)\,u^{m-3}\,|\nabla u|^2$, which corresponds to the left hand side of (3.28) in Theorem 7 when $h = 0$. However, the second term in the right hand side of (3.28) will explode when $h \to 0$. Therefore, we can not use the result in Theorem 7 directly.

Let us modify the estimate we obtained from the decomposition of Y, so that the bound does not explode when $h \to 0$. The idea is to decompose $Y_t - \frac{\phi}{h}(X_t)$ instead of Y_t. This is because

$$Y_t - \frac{\phi}{h} = \phi \frac{u^h - 1}{h} \to \phi \log u, \text{ as } h \to 0,$$

which is finite.

Lemma 8. *Assume*

$$2\lambda - m + h > 0. \tag{3.29}$$

Then for any $\rho > 0$ such that

$$2\lambda - m + h - \rho > 0, \tag{3.30}$$

we have

$$E^{\mathbb{Q}}\left[\int_0^T H_t dt\right] \leq \frac{E^{\mathbb{Q}}\left[Y_T - \frac{\phi}{h}(X_T)\right] - Y_0 + \frac{\phi}{h}(x) + Lm\rho^{-1}M_3T}{m\,(2\lambda - m + h - \rho)}, \tag{3.31}$$

where

$$M_3 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} \left(\rho\left|\frac{u^h-1}{h}\right|u^h + \left(\lambda\left(\frac{u^h-1}{h}\right)+u^h\right)^2\right)u^{m-h-1}(t,y).$$

Proof. From (3.13),

$$\begin{aligned}
& d\left(Y_t - \frac{\phi}{h}(X_t)\right) \\
= \; & \sqrt{2m}\,(hf)^{\frac{m-1}{2h}}\left(\phi\frac{\partial f}{\partial x^\alpha} + \left(f-\frac{1}{h}\right)\frac{\partial \phi}{\partial x^\alpha}\right)d\tilde{W}_t^\alpha \\
& + m(2\lambda - m + h)(hf)^{\frac{m-h-1}{h}}\phi\left|\nabla f\right|^2 dt \\
& + \Delta\phi m\left(f-\frac{1}{h}\right)(hf)^{\frac{m-1}{h}}dt \\
& + 2m\left(\lambda\left(f-\frac{1}{h}\right)+hf\right)(hf)^{\frac{m-h-1}{h}}\frac{\partial f}{\partial x^\alpha}\frac{\partial \phi}{\partial x^\alpha}dt. \qquad (3.32)
\end{aligned}$$

By the Cauchy-Schwartz inequality, for any $\rho > 0$,

$$\begin{aligned}
& 2m\left(\lambda\left(f-\frac{1}{h}\right)+hf\right)(hf)^{\frac{m-h-1}{h}}\frac{\partial f}{\partial x^\alpha}\frac{\partial \phi}{\partial x^\alpha} \\
\geq & -\rho m(hf)^{\frac{m-h-1}{h}}\phi\left|\nabla f\right|^2 dt \\
& -\rho^{-1}m\left(\lambda\left(f-\frac{1}{h}\right)+hf\right)^2(hf)^{\frac{m-h-1}{h}}I_{\mathcal{A}_{x_0}^{R,\epsilon}}\phi^{-1}\left|\nabla\phi\right|^2 dt.
\end{aligned}$$

Hence

$$
\begin{aligned}
& d\left(Y_t - \frac{\phi}{h}\left(X_t\right)\right) \\
\geq & \sqrt{2m}\left(hf\right)^{\frac{m-1}{2h}}\left(\phi\frac{\partial f}{\partial x^\alpha} + \left(f - \frac{1}{h}\right)\frac{\partial \phi}{\partial x^\alpha}\right) d\tilde{W}_t^\alpha \\
& + m\left(2\lambda - m + h - \rho\right) H_t dt \\
& - Lm\rho^{-1}\left(\rho\left|f - \frac{1}{h}\right| hf + \left(\lambda\left(f - \frac{1}{h}\right) + hf\right)^2\right)\left(hf\right)^{\frac{m-h-1}{h}} dt.
\end{aligned}
$$

Let

$$
M_3 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}}\left(\rho\left|\frac{u^h - 1}{h}\right| u^h + \left(\lambda\left(\frac{u^h - 1}{h}\right) + u^h\right)^2\right) u^{m-h-1}\left(t, y\right).
$$

According to our assumptions on u, $\left(hf\right)^{\frac{m-1}{2h}}\left(\phi\frac{\partial f}{\partial x^\alpha} + \left(f - \frac{1}{h}\right)\frac{\partial \phi}{\partial x^\alpha}\right)$, $\alpha \in \{1, \cdots, n\}$ are all bounded, which ensures that the local martingale part of $Y_t - \frac{\phi}{h}\left(X_t\right)$ is a true $\mathbb{Q}$ martingale. Hence

$$
\begin{aligned}
& E^{\mathbb{Q}}\left[Y_T - \frac{\phi}{h}\left(X_T\right)\right] - Y_0 + \frac{\phi}{h}\left(x\right) \\
\geq & m\left(2\lambda - m + h - \rho\right) E^{\mathbb{Q}}\left[\int_0^T H_t dt\right] - Lm\rho^{-1} M_3 T,
\end{aligned}
$$

which completes the proof. □

By combining (3.31) with (3.17) in the way as in the proof of Theorem 7 and also taking $\lambda = \frac{2m-h}{2}$, we are able to get the following gradient bound.

Proposition 9. *Assume u is a positive and bounded solution to* (1.1) *on* $(0, t_1) \times$

$B(x_0, R+\epsilon)$ *with* $m \in \left(1 - \frac{1}{\sqrt{n-1}}, 1 + \frac{1}{\sqrt{n-1}}\right)$. *Let the cut-off function* ϕ *and the constant* L *be defined according to* (3.5) *and* (3.55). *Furthermore, let the diffusion* X *and the measure* $\mathbb{Q}$ *be defined according to*(3.8) *and* (3.11) *with* $\lambda = \frac{2m-h}{2}$. *Then for any* $h \in (h_-, h_+)$, $\rho \in (0, m)$, *and* $(T, x) \in (0, t_1) \times B(x_0, R)$, *we have*

$$\begin{aligned} & u^{m+h-3} \left|\nabla u\right|^2 (T, x) \\ \leq \quad & \frac{E^{\mathbb{Q}}\left[\phi\left(u_0^h - 1\right)(X_T)\right] - \left(u^h(T, x) - 1\right)}{hm(m-\rho)T} \\ & + \frac{LM_3}{\rho(m-\rho)} + \left(|2m-h| l^{-3} + 6l^{-1}\right) m M_1 L^2 \frac{T^2}{8}, \end{aligned}$$

where

$$l = \frac{(h-h_-)(h_+-h)}{3\left|m - \frac{h}{2}\right| + 3}, \quad M_1 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} u^{2m+h-2},$$

$$M_3 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} \left(\rho\left|\frac{u^h-1}{h}\right| u^h + \left(\frac{2m-h}{2}\left(\frac{u^h-1}{h}\right) + u^h\right)^2\right) u^{m-h-1}(t, y).$$

By letting $h \to 0$, we immediately get a gradient bound for the pressure variable $v = \frac{u^{m-1}}{m-1}$ as follows.

Corollary 10. *Assume* u *is a positive and bounded solution to* (1.1) *on* $(0, t_1) \times B(x_0, R+\epsilon)$ *with* $m \in \left(1 - \frac{2}{n}, 1\right)$. *Then for any* $\rho \in (0, m)$, *and* $(T, x) \in (0, t_1) \times B(x_0, R)$, *we have*

$$\begin{aligned} \frac{\left|\nabla v\right|^2}{(m-1)v}(T, x) \quad \leq \quad & \frac{\left\|I_{B(x_0,R+\epsilon)} \log u_0\right\|_\infty - \log u(T, x)}{m(m-\rho)T} \\ & + \frac{LM_3}{\rho(m-\rho)} + \left(\frac{m}{2} l^{-3} + \frac{3}{2} l^{-1}\right) m M_1 L^2 \frac{T^2}{2}, \end{aligned}$$

where

$$l = \frac{-\frac{1}{2}n\left(m-1\right)^2-\left(m-1\right)}{\frac{3}{2}\left(m+1\right)},\ M_1 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} u^{2m-2},$$

$$M_3 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} \left(\rho\log u+\left(m\log u+1\right)^2\right)u^{m-1}\left(t,y\right).$$

3.1.5 Gradient bound for u depending only on initial data

In our main result, Theorem 7, one can see that the right hand side of (3.28) depends not only on the local information of the initial data, but also on the value of u on $[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}$. But according to the local smoothing effect of FDE in [35], the local upper bound for u over $[0,T]$ can be controlled by the local information of initial data. Therefore, we should expect a gradient bound independent of the behavior of u on $[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}$ when $m<1$. This becomes our aim in this subsection. The key ingredients are still estimates obtained from the decompositions of Y and H. For the decomposition of H, we will use the estimate of it that has been derived above, that is, (3.20). As for the decomposition of Y, we will deal with terms containing $\Delta\phi$ and $\nabla\phi$ in a different way such that when h takes the right value the degrees of u in extra terms are all zero. Then we can have a gradient bound only depending on the initial value of u, and a Harnack inequality of elliptic type follows.

Lemma 11. *Let $h=2-2m$. Assume*

$$2\lambda-3m+2>0.$$

Then for any $\rho_1 > 0$ *such that*

$$2\lambda - 3m + 2 - \left|1 + \lambda (2 - 2m)^{-1}\right| \rho_1 > 0,$$

we have

$$\begin{aligned} & E^{\mathbb{Q}}\left[\int_0^T e^{C_1 t} H_t dt\right] \\ \leq \; & \frac{E^{\mathbb{Q}}\left[e^{C_1 T} Y_T\right] - Y_0}{m\left(2\lambda - 3m + 2 - \left|1 + \lambda (2 - 2m)^{-1}\right| \rho_1\right)} \\ & + \frac{L}{2C_1} \frac{\rho_1^{-1} (2 - 2m)^{-2} + \rho_1^{-3} \left|1 + \lambda (2 - 2m)^{-1}\right|}{2\lambda - 3m + 2 - \left|1 + \lambda (2 - 2m)^{-1}\right| \rho_1} \left(e^{C_1 T} - 1\right), \end{aligned} \tag{3.33}$$

where

$$C_1 = \left(\rho_1 m + m \left|1 + \lambda (2 - 2m)^{-1}\right| \rho_1\right) L (1 - m). \tag{3.34}$$

Proof. By (3.3) in the definition of the cut-off function ϕ,

$$\Delta\phi \geq -L\phi^{\frac{1}{2}}.$$

Hence for any $\rho_1 > 0$,

$$\begin{aligned} & \Delta\phi m h^{-1} (hf)^{\frac{m-1+h}{h}} \\ \geq \; & -L\phi^{\frac{1}{2}} m h^{-1} (hf)^{\frac{m-1+h}{h}} I_{\mathcal{A}_{x_0}^{R,\epsilon}} \\ \geq \; & -\frac{1}{2} m \rho_1 L \phi h f - \frac{1}{2} m \rho_1^{-1} h^{-2} L (hf)^{\frac{2m-2+h}{h}} I_{\mathcal{A}_{x_0}^{R,\epsilon}}, \end{aligned}$$

and

$$
\begin{aligned}
& (hf)^{\frac{m-1}{h}} \frac{\partial f}{\partial x^\alpha} \frac{\partial \phi}{\partial x^\alpha} \\
\geq \; & -\frac{1}{2}\rho_1 (hf)^{\frac{m-1-h}{h}} \phi |\nabla f|^2 I_{\mathcal{A}_{x_0}^{R,\epsilon}} - \frac{1}{2}\rho_1^{-1} (hf)^{\frac{m-1+h}{h}} \phi^{-1} |\nabla \phi|^2 I_{\mathcal{A}_{x_0}^{R,\epsilon}} \\
\geq \; & -\frac{1}{2}\rho_1 (hf)^{\frac{m-1-h}{h}} \phi |\nabla f|^2 \\
& -\frac{1}{4}\rho_1 L\phi hf - \frac{1}{4}\rho_1^{-3} L^{-1} (hf)^{\frac{2m-2+h}{h}} \phi^{-3} |\nabla \phi|^4 I_{\mathcal{A}_{x_0}^{R,\epsilon}},
\end{aligned}
$$

where we have used the fact that $hf > 0$ always holds. Then, plugging these two inequalities into (3.13) yields

$$
\begin{aligned}
& dY_t \\
\geq & \sqrt{2m} (hf)^{\frac{m-1}{2h}} \frac{\partial (f\phi)}{\partial x^\alpha} d\tilde{W}_t^\alpha \\
& + m (2\lambda - m + h) H_t dt \\
& - \left(\frac{1}{2}\rho_1 m + \frac{1}{2} m \left|1 + \lambda h^{-1}\right| \rho_1\right) L h Y_t dt \\
& - \frac{1}{2} m \rho_1^{-1} h^{-2} L (hf)^{\frac{2m-2+h}{h}} I_{\mathcal{A}_{x_0}^{R,\epsilon}} dt \\
& - m \left|1 + \lambda h^{-1}\right| \rho_1 H_t dt \\
& - \frac{1}{2} m \left|1 + \lambda h^{-1}\right| \rho_1^{-3} L (hf)^{\frac{2m-2+h}{h}} I_{\mathcal{A}_{x_0}^{R,\epsilon}} dt.
\end{aligned}
$$

When

$$
h = 2 - 2m,
$$

it holds that

$$
\begin{aligned}
& dY_t \\
\geq & \sqrt{2m}\left(2\left(1-m\right)f\right)^{-\frac{1}{4}}\frac{\partial\left(f\phi\right)}{\partial x^{\alpha}}d\tilde{W}_t^{\alpha} \\
& + m\left(2\lambda-3m+2-\left|1+\lambda\left(2-2m\right)^{-1}\right|\rho_1\right)H_t dt \\
& - \frac{1}{2}m\left(\rho_1^{-1}\left(2-2m\right)^{-2}+\rho_1^{-3}\left|1+\lambda\left(2-2m\right)^{-1}\right|\right)LI_{\mathcal{A}_{x_0}^{R,\epsilon}}dt \\
& - \left(\rho_1 m+m\left|1+\lambda\left(2-2m\right)^{-1}\right|\rho_1\right)L\left(1-m\right)Y_t dt.
\end{aligned}
$$

Define

$$C_1=\left(\rho_1 m+m\left|1+\lambda\left(2-2m\right)^{-1}\right|\rho_1\right)L\left(1-m\right).$$

Then it follows that

$$
\begin{aligned}
& de^{C_1 t}Y_t \\
\geq & \sqrt{2m}\left(2\left(1-m\right)f\right)^{-\frac{1}{4}}\frac{\partial\left(f\phi\right)}{\partial x^{\alpha}}e^{C_1 t}d\tilde{W}_t^{\alpha} \\
& + me^{C_1 t}\left(2\lambda-3m+2-\left|1+\lambda\left(2-2m\right)^{-1}\right|\rho_1\right)H_t dt \\
& - \frac{1}{2}me^{C_1 t}\left(\rho_1^{-1}\left(2-2m\right)^{-2}+\rho_1^{-3}\left|1+\lambda\left(2-2m\right)^{-1}\right|\right)LI_{\mathcal{A}_{x_0}^{R,\epsilon}}dt.
\end{aligned}
$$

According to our assumption on u, $\left(2\left(1-m\right)f\right)^{-\frac{1}{4}}\frac{\partial(f\phi)}{\partial x^{\alpha}}$, $\alpha\in\{1,\cdots,n\}$ are all bounded. Hence $\int_0^{\cdot}(2(1-m)f)^{-\frac{1}{4}}\frac{\partial(f\phi)}{\partial x^{\alpha}}e^{C_1t}d\tilde{W}_t^{\alpha}$ is a $\mathbb{Q}$ martingale. Therefore,

$$\begin{aligned}
&E^{\mathbb{Q}}\left[e^{C_1T}Y_T\right]-Y_0\\
\geq& m\left(2\lambda-3m+2-\left|1+\lambda(2-2m)^{-1}\right|\rho_1\right)E^{\mathbb{Q}}\left[\int_0^T e^{C_1t}H_tdt\right]\\
&-\frac{1}{2}m\left(\rho_1^{-1}(2-2m)^{-2}+\rho_1^{-3}\left|1+\lambda(2-2m)^{-1}\right|\right)L\int_0^T e^{C_1t}dt.
\end{aligned}$$

The result is obtained by dividing both sides by $2\lambda-3m+2-\left|1+\lambda(2-2m)^{-1}\right|\rho_1$. $\square$

With this estimate at hand, together with (3.20) we are ready to derive a gradient bound which is independent of the value of u on $[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}$.

Theorem 12. *Assume u is a positive and bounded solution to (1.1) on $(0,t_1)\times B(x_0,R+\epsilon)$ with $m\in\left(1-\frac{6}{n+8},1\right)$. Let the cut-off function ϕ and the constant L be defined according to (3.5) and (3.55). Furthermore, let the diffusion X and the measure $\mathbb{Q}$ be defined according to (3.8) and (3.11) with $h=2-2m$ and $\lambda=2m-1$. Then for any $\rho_1\in(0,2m(1-m))$ and $(T,x)\in(0,t_1)\times B(x_0,R)$, it holds that*

$$\begin{aligned}
&\frac{|\nabla u|^2}{u^{m+1}}(T,x)\\
\leq&\frac{C_1e^{C_1T}E^{\mathbb{Q}}\left[\phi u_0^{2-2m}(X_T)\right]-C_1u^{2-2m}(T,x)}{m(2m(1-m)-\rho_1)(e^{C_1T}-1)}\\
&+\frac{L(\rho_1^2+2-2m)}{4\rho_1^3(1-m)(2m(1-m)-\rho_1)}\\
&+\left(|2m-1|l^{-3}+3l^{-1}\right)mL^2\frac{C_1T-1+e^{-C_1T}}{2C_1(1-e^{-C_1T})},
\end{aligned}\tag{3.35}$$

where

$$l = \frac{(8+n)(1-m)\left(m-\left(1-\frac{6}{n+8}\right)\right)}{3(|2m-1|+1)}, \ C_1 = \frac{m(1-m)(3-2m)}{2} L\rho_1.$$

Proof. Let

$$h = 2-2m, \ \lambda = 2m-1.$$

Then condition (3.16) is reduced to

$$(1-m)((8+n)(1-m)-6) < 0,$$

which is equivalent to

$$m \in \left(1-\frac{6}{n+8}, 1\right).$$

In this setting (3.20) becomes

$$E^{\mathbb{Q}}[H_s] \geq H_0 - \left(\frac{1}{2}|2m-1| l^{-3} + \frac{3}{2} l^{-1}\right) mL^2 s,$$

with

$$l = \frac{(1-m)((8+n)m-n-2)}{3(|2m-1|+1)}.$$

Multiplying both sides by $e^{C_1 s}$ and then integrating s from 0 to T yields

$$E^{\mathbb{Q}}\left[\int_0^T H_s e^{C_1 s} ds\right] \geq \frac{e^{C_1 T}-1}{C_1} H_0 - \left(\frac{1}{2}|2m-1| l^{-3} + \frac{3}{2} l^{-1}\right) mL^2 \frac{C_1 e^{C_1 T} T - e^{C_1 T} + 1}{C_1^2}.$$

Combining it with (3.33) leads to

$$\begin{aligned}
&H_0\\
\leq&\frac{C_1E^{\mathbb{Q}}\left[e^{C_1T}Y_T\right]-C_1Y_0}{m\left(m-\frac{\rho_1}{2(1-m)}\right)\left(e^{C_1T}-1\right)}\\
&+\frac{L\left(\rho_1^2+2-2m\right)}{4\rho_1^3\left(1-m\right)\left(2m\left(1-m\right)-\rho_1\right)}\\
&+\left(\left|2m-1\right|l^{-3}+3l^{-1}\right)mL^2\frac{C_1T-1+e^{-C_1T}}{2C_1\left(1-e^{-C_1T}\right)}.
\end{aligned}$$

where $\rho_1\in(0,2m\left(1-m\right))$. The proof is complete after writing H and Y in terms of u. $\square$

A direct consequence of this result is the following Harnack inequality.

Proposition 13. *Assume u is a positive and bounded solution to (1.1) on $(0,t_1)\times B\left(x_0,R+\epsilon\right)$ with $m\in\left(1-\frac{6}{n+8},1\right)$. Let the constant L be defined according to (3.55). Then for any $\rho\in(0,2m\left(1-m\right))$, x_1 and x_2 in $B\left(x_0,R\right)$, $T\in(0,t_1)$ and $\alpha>0$, we have*

$$\begin{aligned}
&u^{\frac{3}{2}(1-m)}\left(T,x_2\right)-u^{\frac{3}{2}(1-m)}\left(T,x_1\right)\\
\leq&\frac{\alpha}{2}d^2\left(x_1,x_2\right)+\frac{m^2\left(1-m\right)\left(e^{C_1T}-1\right)}{8\alpha C_1^2}\left(\frac{C_1e^{C_1T}\max_{B(x_0,R+\epsilon)}u_0^{2-2m}}{m^2\left(1-m\right)\left(e^{C_1T}-1\right)}\right.\\
&\left.+\frac{L\left(m^2\left(1-m\right)+2\right)}{4\left(1-m\right)^4m^4}+\left(\left|2m-1\right|l^{-3}+3l^{-1}\right)mL^2\frac{C_1T-1+e^{-C_1T}}{2C_1\left(1-e^{-C_1T}\right)}\right)^2.
\end{aligned}$$

where

$$l=\frac{\left(8+n\right)\left(1-m\right)\left(m-\left(1-\frac{6}{n+8}\right)\right)}{3\left(\left|2m-1\right|+1\right)},\ C_1=\frac{m^2\left(1-m\right)^2\left(3-2m\right)}{2}L.$$

Proof. Let $r(\theta)$, $\theta \in [0,1]$ be the shortest path on $\mathbb{R}^n$ joining x_1 and x_2. Then

$$\begin{aligned}
& u^{\frac{3}{2}(1-m)}(T,x_2) - u^{\frac{3}{2}(1-m)}(T,x_1) \\
=& \int_0^1 \frac{\nabla u}{u^{\frac{3m-1}{2}}}(T,r(\theta)) \cdot \frac{dr}{d\theta} d\theta \\
\leq& \int_0^1 \left(\frac{\alpha}{2} \left| \frac{dr}{d\theta} \right|^2 + \frac{|\nabla u|^2}{2\alpha u^{3m-1}}(T,r(\theta)) \right) d\theta,
\end{aligned} \tag{3.36}$$

for any $\alpha > 0$. Since the whole path of $r(\theta)$ lies in $B(x_0,R)$, we can use gradient estimate (3.35). For simplicity, let

$$\rho_1 = m(1-m).$$

Then (3.35) becomes

$$\begin{aligned}
& \frac{|\nabla u|^2}{u^{m+1}}(T,x) \\
\leq& \frac{C_1 e^{C_1 T} E^{\mathbb{Q}}\left[\phi u^{2-2m}(0,X_T)\right] - C_1 u^{2-2m}(T,x)}{m^2(1-m)\left(e^{C_1 T}-1\right)} \\
&+ \frac{L\left(m^2(1-m)+2\right)}{4(1-m)^4 m^4} \\
&+ \left(|2m-1| l^{-3} + 3l^{-1}\right) m L^2 \frac{C_1 T - 1 + e^{-C_1 T}}{2C_1\left(1-e^{-C_1 T}\right)},
\end{aligned}$$

with

$$l = \frac{(1-m)((8+n)m-n-2)}{3(|2m-1|+1)},\ C_1 = \frac{m^2(1-m)^2(3-2m)}{2}L.$$

This implies that

$$
\begin{aligned}
&\frac{2\sqrt{C_1}}{m\sqrt{(1-m)\left(e^{C_1T}-1\right)}}\frac{|\nabla u|}{u^{\frac{3m-1}{2}}}(T,x)\\
\leq&\frac{|\nabla u|^2}{u^{m+1}}(T,x)+\frac{C_1}{m^2(1-m)\left(e^{C_1T}-1\right)}u^{2-2m}(T,x)\\
\leq&\frac{C_1e^{C_1T}E^{\mathbb{Q}}\left[\phi u^{2-2m}(t_1,X_T)\right]}{m^2(1-m)\left(e^{C_1T}-1\right)}\\
&+\frac{L\left(m^2(1-m)+2\right)}{4(1-m)^4m^4}\\
&+\left(|2m-1|\,l^{-3}+3l^{-1}\right)mL^2\frac{C_1T-1+e^{-C_1T}}{2C_1\left(1-e^{-C_1T}\right)}\\
\leq&\frac{C_1e^{C_1T}\max_{B(x_0,R+\epsilon)}u_0^{2-2m}}{m^2(1-m)\left(e^{C_1T}-1\right)}\\
&+\frac{L\left(m^2(1-m)+2\right)}{4(1-m)^4m^4}\\
&+\left(|2m-1|\,l^{-3}+3l^{-1}\right)mL^2\frac{C_1T-1+e^{-C_1T}}{2C_1\left(1-e^{-C_1T}\right)}.
\end{aligned}
\tag{3.37}
$$

Therefore, (3.36) and (3.37) together gives

$$
\begin{aligned}
&u^{\frac{3}{2}(1-m)}(T,x_2)-u^{\frac{3}{2}(1-m)}(T,x_1)\\
\leq&\int_0^1\left(\frac{\alpha}{2}\left|\frac{dr}{d\theta}\right|^2+\frac{|\nabla u|^2}{2\alpha u^{3m-1}}(T,r(\theta))\right)d\theta\\
\leq&\frac{\alpha}{2}d^2(x_1,x_2)+\frac{m^2(1-m)\left(e^{C_1T}-1\right)}{8\alpha C_1^2}\left(\frac{C_1e^{C_1T}\max_{B(x_0,R+\epsilon)}u_0^{2-2m}}{m^2(1-m)\left(e^{C_1T}-1\right)}\right.\\
&\left.+\frac{L\left(m^2(1-m)+2\right)}{4(1-m)^4m^4}+\left(|2m-1|\,l^{-3}+3l^{-1}\right)mL^2\frac{C_1T-1+e^{-C_1T}}{2C_1\left(1-e^{-C_1T}\right)}\right)^2.
\end{aligned}
$$

□

3.2 The constraint on m

As we have seen, comparing to existing results, although our local gradient bound is valid for a larger range of m when $m < 1$, it has a more specific restriction when $m > 1$. Since a gradient estimate involving two-sided bound for solutions obtained in [39] is valid for all $m > 1$, we attempt in three ways in this section to generalize our method so that the constraint on m can be relaxed. One will see that none of these approaches helps, which may suggest that condition (3.27) is subtle for the existence of a gradient bound for $u^{m+h-3} \left|\nabla u\right|^2$ that only involves one-sided bound for u.

3.2.1 Supermartingale property of $\frac{1}{H}$

Recall that the constraint (3.27) came from the requirement of positivity of the term containing H^2 in the decomposition of H. Inspired by the basic formula $\frac{dt}{t^2} = d\left(\frac{1}{t}\right)$, let us study the decomposition of $\frac{1}{H}$ instead. Bear in mind that if M is a supermartingale, then by the Jensen's inequality $\frac{1}{M}$ is a submartingale. So if we can prove that $\frac{1}{H}$ is a supermartingale, then H is a submartingale. From (3.12)

$$\begin{aligned}
d\frac{1}{H_t} = & -\frac{1}{H_t^2}2\sqrt{2m}\left(hf\right)^{\frac{3m-2h-3}{2h}}\phi\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}d\tilde{W}_t^\beta \\
& -\frac{1}{H_t^2}\sqrt{2m}\left(m-h-1\right)\left(hf\right)^{\frac{3m-4h-3}{2h}}\phi\left|\nabla f\right|^2\frac{\partial f}{\partial x^\beta}d\tilde{W}_t^\beta \\
& -\frac{1}{H_t^2}\sqrt{2m}\left(hf\right)^{\frac{3m-2h-3}{2h}}\frac{\partial \phi}{\partial x^\beta}\left|\nabla f\right|^2 d\tilde{W}_t^\beta \\
& +\frac{1}{H_t^2}m\left(4m-3h-1-2\lambda\right)\left(m-h-1\right)\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4 dt \\
& +\frac{2m}{H_t^2}\left(m-h-1-\lambda\right)\left(hf\right)^{\frac{2m-2h-2}{h}}\frac{\partial \phi}{\partial x^\beta}\frac{\partial f}{\partial x^\beta}\left|\nabla f\right|^2 dt \\
& -\frac{m}{H_t^2}\left(hf\right)^{\frac{2m-h-2}{h}}\Delta\phi\left|\nabla f\right|^2 dt \\
& +\frac{2m}{H_t^2}\left(hf\right)^{\frac{2m-h-2}{h}}\phi^{-1}\left|\nabla \phi\right|^2\left|\nabla f\right|^2 dt \\
& +\frac{1}{H_t^2}2m\left(m-1\right)\left(hf\right)^{\frac{2m-2-2h}{h}}\phi\Delta f\left|\nabla f\right|^2 dt \\
& +\frac{4m}{H_t^2}\left(2m-2h-1-\lambda\right)\left(hf\right)^{\frac{2m-2h-2}{h}}\phi\frac{\partial f}{\partial x^\beta}\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}dt \\
& +\frac{1}{H_t^2}4m\left(hf\right)^{\frac{2m-2-h}{h}}\frac{\partial \phi}{\partial x^\beta}\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}dt \\
& -\frac{1}{H_t^2}2m\left(hf\right)^{\frac{2m-h-2}{h}}\phi\sum_{\alpha,\beta}\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\right)^2 dt \\
& +\frac{8m}{H_t^2}\left(hf\right)^{\frac{2m-h-2}{h}}\phi\sum_{\beta}\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}\right)^2\left|\nabla f\right|^{-2} dt.
\end{aligned}$$

If we use the Hölder inequality to obtain the estimate

$$\sum_{\beta}\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}\right)^2\left|\nabla f\right|^{-2} \leq \sum_{\beta}\sum_{\alpha}\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\right)^2,$$

then the term that contains $\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\right)^2$ has positive coefficient. So $d\frac{1}{H}$ does not have an upper bound, which forbids us to derive a supermartingale property of $\frac{1}{H}$ for any m.

3.2.2 More parameters in change of measures

Let us observe that when $n = 1$, the constraint vanishes. This means that our method works well for the one-dimensional case. In other words, some part of our argument may be essentially one-dimension oriented, which needs to be improved to adapt to high dimensions. Recall that when changing the measure, only a one-dimension parameter λ is introduced, which might have undermined some problems. In light of this observation, we now consider change of measures in a more general form. Define the parameter matrix $\Lambda = \left(\Lambda^i_j\right)_{i,j\in\{1,\cdots,n\}} \in \mathbb{R}^{n\times n}$. Then consider the matrix

$$P \triangleq \frac{m-h}{2} I_{n\times n} + \Lambda. \tag{3.38}$$

By the Girsanov's Theorem, we can construct a measure $\bar{\mathbb{Q}}$ on $(\Omega, \mathcal{F}_T)$ such that $\bar{W} = \left(\bar{W}^1, \cdots, \bar{W}^n\right)$ defined by

$$d\bar{W}^\beta_t = dW^\beta_t - \sqrt{2m}\,(hf)^{\frac{m-2h-1}{2h}}\, P^\beta_j \frac{\partial f}{\partial x^j} dt,$$

is a Brownian motion under $\bar{\mathbb{Q}}$. In view of the proof of Lemma 5, one can see that considering the parameter matrix in the form of (3.38) rather than Λ itself will facilitates computation when extracting information from the decomposition of Y. Moreover, from condition (3.21) in Lemma 5, we should require Λ to satisfy

$$\xi^T \Lambda \xi \geq c_1 \left|\xi\right|^2, \tag{3.39}$$

for some positive constant c_1.

From (3.10) we immediately get the decomposition of H under $\mathbb{Q}$ as follows.

$$
\begin{aligned}
dH_t \;=\;& 2\sqrt{2m}\left(hf\right)^{\frac{3m-2h-3}{2h}}\phi\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}d\bar{W}_t^\beta \\
&+\sqrt{2m}\left(m-h-1\right)\left(hf\right)^{\frac{3m-4h-3}{2h}}\phi\left|\nabla f\right|^2\frac{\partial f}{\partial x^\beta}d\bar{W}_t^\beta \\
&+\sqrt{2m}\left(hf\right)^{\frac{3m-2h-3}{2h}}\frac{\partial \phi}{\partial x^\beta}\left|\nabla f\right|^2 d\bar{W}_t^\beta \\
&+2m\left(m-h-1\right)\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\frac{\partial f}{\partial x^\beta}\Lambda_j^\beta\frac{\partial f}{\partial x^j}\left|\nabla f\right|^2 dt \\
&-m\left(m+1\right)\left(m-h-1\right)\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4 dt \\
&+2m\left(hf\right)^{\frac{2m-2h-2}{h}}\frac{\partial \phi}{\partial x^\beta}\Lambda_j^\beta\frac{\partial f}{\partial x^j}\left|\nabla f\right|^2 dt \\
&+m\left(3m-3h-2\right)\left(hf\right)^{\frac{2m-2h-2}{h}}\frac{\partial \phi}{\partial x^\alpha}\frac{\partial f}{\partial x^\alpha}\left|\nabla f\right|^2 dt \\
&+m\left(hf\right)^{\frac{2m-h-2}{h}}\Delta\phi\left|\nabla f\right|^2 dt \\
&-2m\left(m-1\right)\left(hf\right)^{\frac{2m-2-2h}{h}}\phi\Delta f\left|\nabla f\right|^2 dt \\
&+4m\left(hf\right)^{\frac{2m-2h-2}{h}}\phi\frac{\partial f}{\partial x^\alpha}\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\Lambda_j^\beta\frac{\partial f}{\partial x^j}dt \\
&-2m\left(h-m+2\right)\left(hf\right)^{\frac{2m-2h-2}{h}}\phi\frac{\partial f}{\partial x^\beta}\frac{\partial^2 f}{\partial x^\beta \partial x^\alpha}\frac{\partial f}{\partial x^\alpha}dt \\
&+4m\left(hf\right)^{\frac{2m-2-h}{h}}\frac{\partial \phi}{\partial x^\beta}\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}dt \\
&+2m\left(hf\right)^{\frac{2m-h-2}{h}}\phi\sum_{\alpha,\beta}\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\right)^2 dt. \qquad (3.40)
\end{aligned}
$$

Again, the next step is to deal with terms containing the second order derivatives.

Lemma 14. *H satisfies*

$$\begin{aligned}
& dH_t \\
\geq & A_t d\bar{W}_t \\
& + m\left(m-1\right)\left(m-h-2-\frac{mn}{2}+\frac{n}{2}\right)\left(hf\right)^{\frac{2m-2-3h}{h}}\phi\left|\nabla f\right|^4 dt \\
& - \frac{m}{2}\left(m-h-2\right)^2\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4 dt \\
& - m\left(m+1\right)\left(m-h-1\right)\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4 dt \\
& + 2m^2\left(hf\right)^{\frac{2m-2-3h}{h}}\phi\left|\nabla f\right|^2\frac{\partial f}{\partial x^\alpha}\Lambda_j^\alpha\frac{\partial f}{\partial x^j}dt \\
& - 2m\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^2\sum_\beta\left(\Lambda_j^\beta\frac{\partial f}{\partial x^j}\right)^2 dt \\
& + m\left(3m-h\right)\left(hf\right)^{\frac{2m-2h-2}{h}}\frac{\partial\phi}{\partial x^\alpha}\frac{\partial f}{\partial x^\alpha}\left|\nabla f\right|^2 dt \\
& - 2m\left(hf\right)^{\frac{2m-2h-2}{h}}\frac{\partial\phi}{\partial x^\beta}\Lambda_j^\beta\frac{\partial f}{\partial x^j}\left|\nabla f\right|^2 dt \\
& - 2m\left(hf\right)^{\frac{2m-2-h}{h}}\phi^{-1}I_{\mathcal{A}_{x_0}^{R,\epsilon}}\left|\nabla f\right|^2\left|\nabla\phi\right|^2 dt \\
& + m\left(hf\right)^{\frac{2m-h-2}{h}}\Delta\phi\left|\nabla f\right|^2 dt.
\end{aligned} \tag{3.41}$$

Proof. From (3.40), let us write

$$dH_t \triangleq A_t d\bar{W}_t + \left(\bar{B}+\bar{C}+\bar{D}\right)dt,$$

where

$$
\begin{aligned}
A^{\beta} &= 2\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\,\phi\frac{\partial^2 f}{\partial x^{\alpha}\partial x^{\beta}}\frac{\partial f}{\partial x^{\alpha}} \\
&\quad +\sqrt{2m}\,(m-h-1)\,(hf)^{\frac{3m-4h-3}{2h}}\,\phi\,|\nabla f|^2\,\frac{\partial f}{\partial x^{\beta}} \\
&\quad +\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\,\frac{\partial \phi}{\partial x^{\beta}}\,|\nabla f|^2,
\end{aligned}
$$

and by the Cauchy-Schwartz inequality,

$$
\begin{aligned}
\bar{B} \;=\;& \sum_{\alpha} \Bigg(2m\,(hf)^{\frac{2m-h-2}{h}}\,\phi \left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\alpha} \right)^2 \\
& -2m\,(m-1)\,(hf)^{\frac{2m-2-2h}{h}}\,\phi \frac{\partial^2 f}{\partial x^\alpha \partial x^\alpha}\,|\nabla f|^2 \\
& +4m\,(hf)^{\frac{2m-2h-2}{h}}\,\phi \frac{\partial f}{\partial x^\alpha} \frac{\partial^2 f}{\partial x^\alpha \partial x^\alpha} \Lambda^\alpha_j \frac{\partial f}{\partial x^j} \\
& -2m\,(h-m+2)\,(hf)^{\frac{2m-2h-2}{h}}\,\phi \frac{\partial f}{\partial x^\alpha} \frac{\partial^2 f}{\partial x^\alpha \partial x^\alpha} \frac{\partial f}{\partial x^\alpha} \\
& +4m\,(hf)^{\frac{2m-2-h}{h}}\,\frac{\partial \phi}{\partial x^\alpha} \frac{\partial^2 f}{\partial x^\alpha \partial x^\alpha} \frac{\partial f}{\partial x^\alpha} \Bigg) \\
\geq\;& m\,(m-1) \left(m-h-2-\frac{mn}{2}+\frac{n}{2} \right) (hf)^{\frac{2m-2-3h}{h}}\,\phi\,|\nabla f|^4 \\
& +2m\,(m-1)\,(hf)^{\frac{2m-2-3h}{h}}\,\phi\,|\nabla f|^2\,\frac{\partial f}{\partial x^\alpha} \Lambda^\alpha_j \frac{\partial f}{\partial x^j} \\
& -2m\,(m-h-2)\,(hf)^{\frac{2m-3h-2}{h}}\,\phi \left(\frac{\partial f}{\partial x^\alpha} \right)^3 \Lambda^\alpha_j \frac{\partial f}{\partial x^j} \\
& -2m\,(hf)^{\frac{2m-3h-2}{h}}\,\phi \left(\frac{\partial f}{\partial x^\alpha} \right)^2 \left(\Lambda^\alpha_j \frac{\partial f}{\partial x^j} \right)^2 \\
& -\frac{m}{2}\,(m-h-2)^2\,(hf)^{\frac{2m-3h-2}{h}}\,\phi \sum_{\alpha} \left(\frac{\partial f}{\partial x^\alpha} \right)^4 \\
& +2m\,(m-1)\,(hf)^{\frac{2m-2-2h}{h}}\,|\nabla f|^2\,\frac{\partial \phi}{\partial x^\alpha} \frac{\partial f}{\partial x^\alpha} \\
& -4m\,(hf)^{\frac{2m-2h-2}{h}}\,\frac{\partial \phi}{\partial x^\alpha} \left(\frac{\partial f}{\partial x^\alpha} \right)^2 \Lambda^\alpha_j \frac{\partial f}{\partial x^j} \\
& -2m\,(m-h-2)\,(hf)^{\frac{2m-2h-2}{h}} \left(\frac{\partial f}{\partial x^\alpha} \right)^3 \frac{\partial \phi}{\partial x^\alpha} \\
& -2m\,(hf)^{\frac{2m-2-h}{h}}\,\phi^{-1} I_{\mathcal{A}^{R,\epsilon}_{x_0}} \sum_{\alpha} \left(\frac{\partial \phi}{\partial x^\alpha} \right)^2 \left(\frac{\partial f}{\partial x^\alpha} \right)^2 ,
\end{aligned}
$$

$$\begin{aligned}
\bar{C} &= \sum_{\alpha\neq\beta}\Bigg(2m\,(hf)^{\frac{2m-h-2}{h}}\,\phi\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\right)^2 \\
&\quad +4m\,(hf)^{\frac{2m-2h-2}{h}}\,\phi\frac{\partial f}{\partial x^\alpha}\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\Lambda_j^\beta\frac{\partial f}{\partial x^j} \\
&\quad -2m\,(h-m+2)\,(hf)^{\frac{2m-2h-2}{h}}\,\phi\frac{\partial f}{\partial x^\beta}\frac{\partial^2 f}{\partial x^\beta \partial x^\alpha}\frac{\partial f}{\partial x^\alpha} \\
&\quad +4m\,(hf)^{\frac{2m-2-h}{h}}\,\frac{\partial \phi}{\partial x^\beta}\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}\Bigg) \\
&\geq \sum_{\alpha\neq\beta}\Bigg(-\frac{1}{2}m\,(m-h-2)^2\,(hf)^{\frac{2m-3h-2}{h}}\,\phi\left(\frac{\partial f}{\partial x^\beta}\right)^2\left(\frac{\partial f}{\partial x^\alpha}\right)^2 \\
&\quad -2m\,(hf)^{\frac{2m-3h-2}{h}}\,\phi\left(\frac{\partial f}{\partial x^\alpha}\right)^2\left(\Lambda_j^\beta\frac{\partial f}{\partial x^j}\right)^2 \\
&\quad -2m\,(m-h-2)\,(hf)^{\frac{2m-3h-2}{h}}\,\phi\left(\frac{\partial f}{\partial x^\alpha}\right)^2\frac{\partial f}{\partial x^\beta}\Lambda_j^\beta\frac{\partial f}{\partial x^j} \\
&\quad -4m\,(hf)^{\frac{2m-2h-2}{h}}\,\left(\frac{\partial f}{\partial x^\alpha}\right)^2\frac{\partial \phi}{\partial x^\beta}\Lambda_j^\beta\frac{\partial f}{\partial x^j} \\
&\quad -2m\,(m-h-2)\,(hf)^{\frac{2m-2h-2}{h}}\,\frac{\partial \phi}{\partial x^\beta}\frac{\partial f}{\partial x^\beta}\left(\frac{\partial f}{\partial x^\alpha}\right)^2 \\
&\quad -2m\,(hf)^{\frac{2m-2-h}{h}}\,\phi^{-1}I_{\mathcal{A}_{x_0}^{R,\epsilon}}\left(\frac{\partial \phi}{\partial x^\beta}\right)^2\left(\frac{\partial f}{\partial x^\alpha}\right)^2\Bigg),
\end{aligned}$$

and

$$\begin{aligned}
\bar{D} &= 2m\,(m-h-1)\,(hf)^{\frac{2m-3h-2}{h}}\,\phi\frac{\partial f}{\partial x^\beta}\Lambda_j^\beta\frac{\partial f}{\partial x^j}\,|\nabla f|^2 \\
&\quad -m\,(m+1)\,(m-h-1)\,(hf)^{\frac{2m-3h-2}{h}}\,\phi\,|\nabla f|^4 \\
&\quad +2m\,(hf)^{\frac{2m-2h-2}{h}}\,\frac{\partial \phi}{\partial x^\beta}\Lambda_j^\beta\frac{\partial f}{\partial x^j}\,|\nabla f|^2 \\
&\quad +m\,(3m-3h-2)\,(hf)^{\frac{2m-2h-2}{h}}\,\frac{\partial \phi}{\partial x^\alpha}\frac{\partial f}{\partial x^\alpha}\,|\nabla f|^2 \\
&\quad +m\,(hf)^{\frac{2m-h-2}{h}}\,\Delta\phi\,|\nabla f|^2.
\end{aligned}$$

Then adding these three inequalities gives us (3.41). $\square$

Now we need to choose the value of matrix Λ to allow for a computation similar to the one in Lemma 4, hence obtaining submartingale property of a process related to H. To do this, let us focus on terms in (3.41) with the highest degree in ∇f . That term is,

$$\begin{aligned}&\phi\left|\nabla f\right|^2(hf)^{\frac{2m-2-3h}{h}}\left(K\left|\nabla f\right|^2+2m^2\frac{\partial f}{\partial x^\alpha}\Lambda_j^\alpha\frac{\partial f}{\partial x^j}-2m\frac{\partial f}{\partial x^j}\Lambda_j^\beta\Lambda_\alpha^\beta\frac{\partial f}{\partial x^\alpha}\right)\\ =&\phi\left|\nabla f\right|^2(hf)^{\frac{2m-2-3h}{h}}\left(K\left|\nabla f\right|^2+2m^2\left(\nabla f\right)^T\Lambda\nabla f-2m\left(\nabla f\right)^T\Lambda^T\Lambda\nabla f\right), \quad (3.42)\end{aligned}$$

where

$$K=m\left(m-1\right)\left(m-h-\frac{mn}{2}+\frac{n}{2}\right)-\frac{m}{2}\left(m-h-2\right)^2-m\left(m+1\right)\left(m-h-1\right). \quad (3.43)$$

For simplicity, define the matrix

$$\Gamma=KI_{n\times n}+2m^2\Lambda-2m\Lambda^T\Lambda.$$

Hence

$$K\left|\nabla f\right|^2+2m^2\left(\nabla f\right)^T\Lambda\nabla f-2m\left(\nabla f\right)^T\Lambda^T\Lambda\nabla f=\left(\nabla f\right)^T\Gamma\nabla f.$$

Then it is clear that in order to ensure the positivity of (3.42), we should look for Λ such that

$$\left(\nabla f\right)^T\Gamma\nabla f\geq c_2\left|\nabla f\right|^2, \quad (3.44)$$

for some positive c_2. We will see that the existence of such Λ relies on the choice of m, it turns out that this constraint on m is exactly (3.27).

The argument contains two steps. First, let us observe that it is enough to only consider symmetric Λ. Assume Λ is not symmetric. Then we separate its symmetric part with antisymmetric part by setting

$$\Lambda_1 = \frac{1}{2}\left(\Lambda + \Lambda^T\right) \text{ and } \Lambda_2 = \frac{1}{2}\left(\Lambda - \Lambda^T\right).$$

It follows that

$$\begin{aligned}
\Gamma &= KI_{n\times n} + 2m^2\left(\Lambda_1 + \Lambda_2\right) - 2m\left(\Lambda_1 + \Lambda_2\right)^T\left(\Lambda_1 + \Lambda_2\right) \\
&= KI_{n\times n} + 2m^2\Lambda_1 + 2m^2\Lambda_2 - 2m\left(\Lambda_1 - \Lambda_2\right)\left(\Lambda_1 + \Lambda_2\right) \\
&= KI_{n\times n} + 2m^2\Lambda_1 + 2m^2\Lambda_2 - 2m\Lambda_1^2 + 2m\Lambda_2^2,
\end{aligned}$$

where we have used the fact that Λ_1 is symmetric and Λ_2 is antisymmetric. But

$$\begin{aligned}
\left(\nabla f\right)^T \Lambda_2\Lambda_2\nabla f &= -\left(\nabla f\right)^T \Lambda_2^T\Lambda_2\nabla f \\
&= -\left(\Lambda_2\nabla f\right)^T \Lambda_2\nabla f \\
&= -\left|\Lambda_2\nabla f\right|^2 \\
&\leq 0.
\end{aligned}$$

So it always holds that

$$(\nabla f)^T \Gamma \nabla f \le (\nabla f)^T \left(KI_{n\times n} + 2m^2\Lambda_1 - 2m\Lambda_1^2\right) \nabla f,$$

which means that the antisymmetric part of a matrix Λ will not contribute to increasing the left hand side of (3.44).

So in the second step, let us only consider a symmetric Λ that satisfies (3.44). We will show that the constraint on m is the same as (3.27). Since Λ is symmetric, we can rewrite Γ as a polynomial of Λ,

$$KI_{n\times n} + 2m^2\Lambda - 2m\Lambda\Lambda.$$

Then by the Spectral Mapping Theorem, its eigenvalues are all positive only when

$$K + 2m^2 z - 2mz^2 > 0, \tag{3.45}$$

for any eigenvalue of Λ, z. Since Λ is symmetric and satisfies (3.39), z should be a positive real number. If we treat the left hand side of (3.45) as a function in z, to ensure that there exists a positive value for z such that(3.45) holds, we need

$$4m^4 + 8mK > 0.$$

Substituting the definition of K (3.43) in it gives us

$$m^2+2\left(m-1\right)\left(m-h-\frac{mn}{2}+\frac{n}{2}\right)-\left(m-h-2\right)^2-2\left(m+1\right)\left(m-h-1\right)>0,$$

which is equivalent to (3.27).

3.2.3 General transformation

From the condition (3.27), one can see that the restriction on m depends on h, which is the parameter we introduced when transforming u to f. Therefore it is natural to ask what will happen when taking a transformation other than the form of $f=\frac{u^h}{h}$. We show in this subsection that a constraint still exists. Let F be an invertible and smooth function on $\mathbb{R}$ and G be the inverse of F. We take transformation $f=F\left(u\right)$. Set

$$Y_t=\left(\phi f\right)\left(T-t,X_t\right).$$

Then

$$\begin{aligned}dY_t &= \sqrt{2m}G\left(f\right)^{\frac{m-1}{2}}\frac{\partial\left(f\phi\right)}{\partial x^i}dW_t^i+J_1\left(\Delta\phi,\nabla\phi,f,\nabla f\right)dt\\ &\quad -m\left(\frac{\left(m-1\right)}{2}G\left(f\right)^{-1}G'\left(f\right)+\frac{G''\left(f\right)}{G'\left(f\right)}\right)G\left(f\right)^{m-1}\phi\left|\nabla f\right|^2dt.\end{aligned}$$

where J_1 consists of terms with zero and first degree in ∇f. Define a measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that under $\mathbb{Q}$ the process $\widetilde{W}$ defined by

$$d\widetilde{W}_t^{\alpha}=dW_t^{\alpha}+\sqrt{2m}\left(\lambda_1G\left(f\right)^{-1}G'\left(f\right)+\lambda_1\frac{G''\left(f\right)}{G'\left(f\right)}\right)G\left(f\right)^{\frac{m-1}{2}}\frac{\partial f}{\partial x^{\alpha}}dt,$$

is a Brownian motion. After a similar computation, we have

$$\begin{aligned}
&d\phi G\left(f\right)^{m-1}\left|\nabla f\right|^{2}\\
\geq&2\sqrt{2m}\phi G\left(f\right)^{m-1}\frac{\partial f}{\partial x^{j}}\left(\frac{m-1}{2}G\left(f\right)^{\frac{m-3}{2}}G'\left(f\right)\frac{\partial f}{\partial x^{j}}\frac{\partial f}{\partial x^{\alpha}}+G\left(f\right)^{\frac{m-1}{2}}\frac{\partial^{2}f}{\partial x^{\alpha}\partial x^{j}}\right)d\tilde{W}_{t}^{\alpha}\\
&-\frac{\bar{K}}{2}m\phi\left|\nabla f\right|^{4}dt+J_{2}\left(\Delta\phi,\nabla\phi,f,\nabla f\right)dt,
\end{aligned}$$

where

$$\begin{aligned}
\bar{K}=&\left[4\lambda_{1}^{2}+\left(6\lambda_{1}-2\right)\left(m-1\right)+\left(m-1\right)^{2}\left(n+\frac{9}{4}\right)\right]\left(G\left(f\right)\right)^{2m-4}\left(G'\left(f\right)\right)^{2}\\
&+4\left(G\left(f\right)\right)^{2m-2}\frac{G'''\left(f\right)}{G'\left(f\right)}+4\left(\lambda_{2}^{2}+2\lambda_{2}\right)\left(G\left(f\right)\right)^{2m-2}\left(\frac{G''\left(f\right)}{G'\left(f\right)}\right)^{2}\\
&+4\left[\left(\frac{3\left(m-1\right)}{2}+2\lambda_{1}\right)\left(1+\lambda_{2}\right)+\left(m-1\right)\right]\left(G\left(f\right)\right)^{2m-3}G''\left(f\right), \qquad (3.46)
\end{aligned}$$

and J_2 consists of terms with degree of less than 4 in ∇f. Similar to (3.16), we need $\bar{K} < 0$, $\mathbb{Q}-$almost surely. Will this still yield a constraint on m if we take a transformation that is not a polynomial?

Take $f = \exp(u)$ as an example. Then $G(f) = \log f$, and

$$\begin{aligned}
f^{2}\bar{K}=&\left[4\lambda_{1}^{2}+\left(6\lambda_{1}-2\right)\left(m-1\right)+\left(m-1\right)^{2}\left(n+\frac{9}{4}\right)\right]\left(\log f\right)^{2m-4}\\
&+4\left(\lambda_{2}^{2}+2\lambda_{2}+2\right)\left(\log f\right)^{2m-2}\\
&-4\left[\left(\frac{3\left(m-1\right)}{2}+2\lambda_{1}\right)\left(1+\lambda_{2}\right)+\left(m-1\right)\right]\left(\log f\right)^{2m-3}.
\end{aligned}$$

In this case, the constraint on m depends on the range of the solution. Let us treat $\log f$ in the above formula as a positive constant for the moment. Then to minimize

$\bar{K}$, we should take

$$\lambda_1 = -\frac{3(m-1)}{4} + (1+\lambda_2)\log f,\ \lambda_2 = -1 + \frac{3(m-1)+4\lambda_1}{4\log f}.$$

Then

$$f^2\bar{K} = \left[-6(m-1) + (m-1)^2 n + 4\right](\log f)^{2m-4},$$

$$f^2\bar{K} = \left(n(m-1)^2 - 2(m-1)\right)(\log f)^{2m-4} + 4(\log f)^{2m-2} - 4(m-1)(\log f)^{2m-3}.$$

Since the coefficient of $(m-1)^2$ is positive, $\bar{K}$ will be positive when $|m-1|$ is large enough, no matter what value $u \equiv \log f$ takes. Therefore, the negativity of $\bar{K}$ imposed a constraint on m.

Similarly, in general, for any transformation G, let us look at $\bar{K}$ at any fixed position (t, x), in order to minimize the right hand side of (3.46) we should take

$$\lambda_1 = -\frac{3}{4}(m-1) - \frac{(1+\lambda_2)G(f)G''(f)}{(G'(f))^2},\ \lambda_2 = -1 - \frac{(3(m-1)+4\lambda_1)(G'(f))^2}{4G(f)G''(f)}.$$

Then

$$\begin{aligned}\bar{K} = &\left(n(m-1)^2 - 2(m-1)\right)(G(f))^{2m-4}(G'(f))^2\\ &+ 4(G(f))^{2m-2}\frac{G'''(f)}{G'(f)} - 4(G(f))^{2m-2}\left(\frac{G''(f)}{G'(f)}\right)^2\\ &+ 4(m-1)(G(f))^{2m-3}G''(f).\end{aligned}$$

Again, the coefficient of $(m-1)^2$ is positive. Therefore, when viewed as a deterministic function on$(0,T)\times\mathbb{R}^n$, $\bar{K}$ will be positive at any fixed location (t,x) when $|m-1|$ is large enough. So there is always a constraint on m.

3.3 Global gradient estimate

3.3.1 From local bound to global bound

Now let us consider a positive and bounded solution u to (1.1) on $(0,\infty)\times\mathbb{R}^n$. First of all, as a direct consequence of the main result in the last section, we can get a global gradient bound from the local bound by letting the radius of the local ball tend to infinity.

Proposition 15. *If u is a positive and bounded solution to* (1.1) *on* $(0,\infty)\times\mathbb{R}^n$ *with* $m\in\left(1-\frac{1}{\sqrt{n-1}},1+\frac{1}{\sqrt{n-1}}\right)$, *then for any* $h\in[h_-,h_+]$ *and* $(T,x)\in(0,\infty)\times\mathbb{R}^n$, *it holds that*

$$u^{m+h-3}\left|\nabla u\right|^2(T,x)\leq\frac{\left\|u_0^h\right\|_\infty-u^h(T,x)}{Thm^2}. \tag{3.47}$$

Proof. The idea is to let the radius R in our local estimate tend to ∞. This is because u is positive and bounded on $(0,\infty)\times\mathbb{R}^n$, which ensures that the local result in Theorem 7 holds for any R and ϵ. Note that the term $E^{\mathbb{Q}}\left[\phi u_0^h(X_T)\right]$ in (3.28) behaves in a complicated way when R changes, as the diffusion X and consequently measure $\mathbb{Q}$ all depend on R. We deal with this by noticing $E^{\mathbb{Q}}\left[\phi u_0^h(X_T)\right]\leq\left\|u_0^h\right\|_\infty$.

From (3.28), it holds for any R and ϵ that,

$$\begin{aligned}
&u^{m+h-3}\left|\nabla u\right|^2(T,x)\\
\leq&\frac{\left\|u_0^h\right\|_\infty-u^h(T,x)}{hm\left(m-\rho\left|\frac{2m+h}{2h}\right|\right)T}+\frac{\left(\rho+h\left|\frac{2m+h}{2h}\right|\right)LM_2}{|h|\rho\left(m-\rho\left|\frac{2m+h}{2h}\right|\right)}\\
&+\left(|2m-h|\,l^{-3}+6l^{-1}\right)mM_1L^2\frac{T}{8}.
\end{aligned}$$

By substituting $\epsilon=R$ into it, and then taking $R\to\infty$ on both sides, we obtain

$$u^{m+h-3}\left|\nabla u\right|^2(T,x)\leq\frac{\left\|u_0^h\right\|_\infty-u^h(T,x)}{Thm\left(m-\rho\left|\frac{2m+h}{2h}\right|\right)},$$

where we have used the fact that when $\epsilon=R$

$$\lim_{R\to\infty}L=\lim_{R\to\infty}\frac{8}{R^2}\left(\frac{nR}{2R+R}+\frac{8\left(2R\right)^2}{\left(3R\right)^2}\right)=0\text{ and }\lim_{R\to\infty}\phi=1.$$

By letting $\rho=0$, we have

$$u^{m+h-3}\left|\nabla u\right|^2(T,x)\leq\frac{\left\|u_0^h\right\|_\infty-u^h(T,x)}{Thm^2}.\tag{3.48}$$

Note that in the local case, due to the existence of l^{-1} in (3.28), h is not allowed to touch the boundary of the open interval (h_-,h_+). But now we can take limit of h to h_- or h_+ on both sides of (3.48). Therefore, (3.48) holds also for $h=h_-$ and $h=h_+$. □

In terms of this result, first of all, by a scaling argument, this is a gradient bound with correct orders in u and t.

Proposition 16. *Assume that any solution u to* (1.1) *satisfies*

$$\left|\nabla u\right|^2(t,x) \le Ct^{\gamma_1}\left\|u\right\|_\infty^{\gamma_2}, \tag{3.49}$$

for any $(t,x) \in (0,\infty) \times \mathbb{R}^n$, then γ_1 and γ_2 have to satisfy

$$\gamma_1 = -1,\ \gamma_2 = 3-m.$$

Proof. Define $\bar{u}(t,x) = K_1 u(K_2 t, K_3 x)$ for positive real numbers K_1, K_2 and K_3. We want $\bar{u}$ to satisfy the PDE (1.1) as well. Since

$$\begin{aligned}
\Delta \bar{u}^m &= m\bar{u}^{m-1}\Delta\bar{u} + m(m-1)\bar{u}^{m-2}\left|\nabla\bar{u}\right|^2 \\
&= mK_1^{m-1}u^{m-1}K_1K_3^2\Delta u + m(m-1)K_1^{m-2}u^{m-2}K_1^2K_3^2\left|\nabla u\right|^2 \\
&= K_1^mK_3^2\left(mu^{m-1}\Delta u + m(m-1)u^{m-2}\left|\nabla u\right|^2\right) \\
&= K_1^mK_3^2\Delta u^m \\
&= K_1^mK_3^2\frac{\partial u}{\partial t},
\end{aligned}$$

and

$$\frac{\partial \bar{u}}{\partial t} = K_1K_2\frac{\partial u}{\partial t},$$

$\bar{u}$ will solve (1.1) if K_1, K_2 and K_3 satisfy the relation

$$K_1^{m-1}K_3^2 = K_2. \tag{3.50}$$

Since it is assumed that (3.49) holds for any solution to (1.1), we have, provided that (3.50) holds,

$$|\nabla \bar{u}|^2 (t,x) \le C t^{\gamma_1} \|\bar{u}\|_\infty^{\gamma_2},$$

which, by the definition of $\bar{u}$, is equivalent to

$$K_1^2 K_3^2 |\nabla u|^2 (K_2 t, K_3 x) \le C t^{\gamma_1} K_1^{\gamma_2} \|u\|_\infty^{\gamma_2}.$$

By (3.50) , let us substitute $K_1 = K_2^{\frac{1}{m-1}} K_3^{-\frac{2}{m-1}}$ into the above inequality. Then we get

$$|\nabla u|^2 (K_2 t, K_3 x) \le C t^{\gamma_1} K_2^{\frac{\gamma_2 - 2}{m-1}} K_3^{\frac{-2m-2\gamma_2+6}{m-1}} \|u\|_\infty^{\gamma_2}.$$

Define

$$\bar{t} = K_2 t,\ \ \bar{x} = K_3 x.$$

Then we have

$$|\nabla u|^2 (\bar{t},\bar{x}) \le C \bar{t}^{\gamma_1} K_2^{\frac{\gamma_2 - 2}{m-1} - \gamma_1} K_3^{\frac{-2m-2\gamma_2+6}{m-1}} \|u\|_\infty^{\gamma_2}. \tag{3.51}$$

Since (3.49) is assumed to hold for any $(t,x) \in (0,\infty) \times \mathbb{R}^n$, we can let K_2 and K_3 take any positive values while adjusting the values of t and x so that $(\bar{t},\bar{x})$ remains the same. Note that the left hand side is independent of K_2 and K_3. Hence it follows that

$$\gamma_1 = -1 \text{ and } \gamma_2 = 3 - m.$$

Otherwise there will be a way to let the right hand side of (3.51) tends to 0 by adjusting the values of K_2 and K_3 while remaining the left hand side to be a positive constant. $\square$

Secondly, this bound does not depend on the dimension n explicitly. This is because the information about dimension has been incorporated into the initial value and the constraint on h. In fact, [9] derived a gradient bound which is independent of the initial data, but dependent on dimension n. Thirdly, it naturally shows that when a function touches its maximum, its gradient vanishes. In this sense it is a tight gradient bound. Moreover, when $m \to 1$ and $h \to 0$, this result recovers Hamilton's gradient bound (1.8).

Corollary 17. *(Hamilton) If u is a solution to the heat equation on $(0,\infty) \times \mathbb{R}^n$ with positive and bounded initial data, then for any $(T,x) \in (0,\infty) \times \mathbb{R}^n$, it holds that*

$$\frac{|\nabla u|^2}{u^2}(T,x) \leq \frac{1}{T} \log \frac{\|u_0\|_\infty}{u(T,x)}. \tag{3.52}$$

Proof. When $m = 1$, result (3.47) reads as

$$u^{h-2} |\nabla u|^2 (T,x) \leq \frac{\left\|u_0^h\right\|_\infty - u^h(T,x)}{Th}, \ \forall h \in [0,2].$$

Taking positive $h \to 0$ on both sides yields

$$\begin{aligned} \frac{|\nabla u|^2}{u^2}(T,x) &\leq \lim_{h\to\infty} \frac{\left\|u_0^h\right\|_\infty - u^h(T,x)}{Th} \\ &= \lim_{h\to\infty} \frac{\|u_0\|_\infty^h - u^h(T,x)}{Th} \\ &\leq \frac{1}{T} \log \frac{\|u_0\|_\infty}{u(T,x)}. \end{aligned}$$

□

3.3.2 Negative finite variation part

In this subsection, we consider a solution u that is bounded not only from above, but also from below by a positive constant. Moreover, we assume $\nabla u\,(t,\cdot)$ is bounded for each $t > 0$. In this case, there is no degeneracy, and u is smooth. Hence we are able to do gradient estimate for a wider range of m. To begin with, let us redefine some notations that have been used in the local case. Let

$$f = \frac{u^h}{h}.$$

Define the diffusion X by solving

$$\begin{aligned} dX_t^\alpha &= \sqrt{2m}\,(hf)^{\frac{m-1}{2h}}\,(T-t,X_t)\,dW_t^\alpha \\ X_0 &= x, \end{aligned}$$

where W is a Brownian motion under $\mathbb{P}$. Note that although by our assumptions $(hf)^{\frac{m-1}{2h}}$ is bounded and $\nabla\,(hf)^{\frac{m-1}{2h}}\,(t,\cdot)$ is bounded for each t, $\nabla\,(hf)^{\frac{m-1}{2h}}\,(t,\cdot)$ may explode as $t \to 0$. Hence we can only solve this SDE on $[0, T-\delta]$, $\delta > 0$. Nevertheless, from the boundedness of $(hf)^{\frac{m-1}{2h}}$ we know $(X_t)_{0\le t<T}$ is a martingale. Moreover, since it has bounded quadratic variation, by the Burkholder–Davis–Gundy inequalities and the Martingale Convergence Theorem it converges almost surely as $t \to T$. Set $X_T = \lim_{t\to} X_t$ almost surely. Then diffusion X is well defined on $[0,T]$. Introduce a

measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}_T)$ by setting the Radon-Nikodym derivative to be

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\int_0^t \sqrt{2m}\lambda\,(hf)^{\frac{m-2h-1}{2h}}\,\frac{\partial f}{\partial x^\beta} dW^\beta - m\lambda^2 \int_0^t (hf)^{\frac{m-2h-1}{h}}\,|\nabla f|^2\,dt\right).$$

Let semimartingales Y and H be defined by

$$Y_t = f\,(T-t, X_t)\,,\ H_t = (hf)^{\frac{m-h-1}{h}}\,|\nabla f|^2\,(T-t, X_t)\,.$$

Then we have their fundamental decompositions as follows.

$$dY_t = \sqrt{2m}\,(hf)^{\frac{m-1}{2h}}\,\frac{\partial f}{\partial x^\alpha} d\tilde{W}_t^\alpha + m\,(2\lambda - m + h)\,(hf)^{\frac{m-h-1}{h}}\,|\nabla f|^2\,dt. \tag{3.53}$$

$$\begin{aligned}
&dH_t \\
&\geq \left(2\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\,\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha} + \sqrt{2m}\,(m-h-1)\,(hf)^{\frac{3m-4h-3}{2h}}\,|\nabla f|^2\,\frac{\partial f}{\partial x^\beta}\right) d\tilde{W}_t^\beta \\
&\quad - \frac{1}{2} mn\,(m-1)^2\,(hf)^{\frac{2m-3h-2}{h}}\,|\nabla f|^4\,dt \\
&\quad - m\,((2m-h+1-2\lambda)\,(m-h-1) + 2\,(\lambda-1)\,(\lambda-m))\,(hf)^{\frac{2m-3h-2}{h}}\,|\nabla f|^4\,dt.
\end{aligned} \tag{3.54}$$

For simplicity, set

$$A_t^\beta = 2\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\,\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha} + \sqrt{2m}\,(m-h-1)\,(hf)^{\frac{3m-4h-3}{2h}}\,|\nabla f|^2\,\frac{\partial f}{\partial x^\beta},$$

$$L_1 = m\left(\left(2m-h+1-2\lambda\right)\left(m-h-1\right)+2\left(\lambda-1\right)\left(\lambda-m\right)\right)+\frac{1}{2}mn\left(m-1\right)^2. \tag{3.55}$$

Then (3.54) becomes

$$dH_t \geq A_t^\beta d\tilde{W}_t^\beta - L_1\left(hf\right)^{\frac{2m-3h-2}{h}}\left|\nabla f\right|^4 dt = A_t^\beta d\tilde{W}_t^\beta - L_1 h^{-1}Y_t^{-1}H_t^2 dt. \tag{3.56}$$

Previously, in order to obtain the submartingale property of H, we always assumed $L_1 \leq 0$, which resulted in a constraint on m. In this section, we consider the situation when

$$L_1 > 0.$$

Instead of deriving a gradient bound directly, we seek for an integral inequality satisfied by the gradient bound as a function of t. The arguments are based on the following estimate.

Lemma 18. *Let u be a strictly positive and bounded solution to* (1.1) *on* $(0,\infty)\times\mathbb{R}^n$ *with* $\left|\nabla u\right|(t,\cdot)$ *bounded for each* $t>0$. *Let* L_1 *be defined by* (3.55). *If* $L_1>0$, *then for any* $0\leq s<t<T$

$$E^{\mathbb{Q}}\left[e^{\int_0^t L_1 h^{-1}Y_r^{-1}H_r dr}H_t\middle|\mathcal{F}_s\right] \geq e^{\int_0^s L_1 h^{-1}Y_r^{-1}H_r dr}H_s, \tag{3.57}$$

and

$$m\left(2\lambda-m+h\right)E^{\mathbb{Q}}\left[\int_s^T H_t dt\middle|\mathcal{F}_s\right] \leq E^{\mathbb{Q}}\left[Y_T\middle|\mathcal{F}_s\right]-Y_s. \tag{3.58}$$

Proof. From (3.56), we have that

$$dH_t e^{\int_0^t L_1 h^{-1} Y_r^{-1} H_r dr} \geq e^{\int_0^t L_1 h^{-1} Y_r^{-1} H_r dr} A_t^\beta d\tilde{W}_t^\beta.$$

Since $\int_0^t e^{\int_0^z L_1 h^{-1} Y_r^{-1} H_r dr} A_z^\beta d\tilde{W}_z^\beta$ is a local martingales, there exists a sequence of stopping times $\{T_n^1\}_{n\geq 1}$ with $T_n^1 \to \infty$ monotonically, such that for each n,

$$\int_0^{t\wedge T_n^1} e^{\int_0^z L_1 h^{-1} Y_r^{-1} H_r dr} A_z^\beta d\tilde{W}_z^\beta,$$

is a true martingale under $\mathbb{Q}$. Hence

$$E^{\mathbb{Q}}\left[e^{\int_0^{t\wedge T_n^1} L_1 h^{-1} Y_r^{-1} H_r dr} H_{t\wedge T_n^1} \middle| \mathcal{F}_s \right] \geq e^{\int_0^{s\wedge T_n^1} L_1 h^{-1} Y_r^{-1} H_r dr} H_{s\wedge T_n^1}.$$

According to our assumptions on u and ∇u, the process $e^{\int_0^t L_1 h^{-1} Y_r^{-1} H_r dr} H_t$ has a uniform bound on $[0, T-\delta] \times \Omega$, $\forall \delta > 0$. Therefore, by the Bounded Convergence Theorem we have (3.57). As for (3.58), there exists a sequence of stopping times $\{T_n^2\}_{n\geq 1}$ with $T_n^2 \to \infty$ monotonically, such that $\int_0^{t\wedge T_n^2} (hf)^{\frac{m-1}{2h}} \frac{\partial f}{\partial x^\alpha} d\tilde{W}_r^\alpha$ is a martingale under $\mathbb{Q}$. So from (3.53) we have

$$E^{\mathbb{Q}}\left[Y_{T\wedge T_n^2} \middle| \mathcal{F}_s \right] - Y_{s\wedge T_n^2} = m\left(2\lambda - m + h\right) E^{\mathbb{Q}}\left[\int_{s\wedge T_n^2}^{T\wedge T_n^2} H_t dt \middle| \mathcal{F}_s \right].$$

On one hand, since solution u is bounded, by the Bounded Convergence Theorem

$$\lim_{n\to\infty} E^{\mathbb{Q}}\left[Y_{T\wedge T_n^2} \middle| \mathcal{F}_s \right] = E^{\mathbb{Q}}\left[Y_T \middle| \mathcal{F}_s \right].$$

On the other hand,

$$\lim_{n\to\infty} E^{\mathbb{Q}}\left[\int_{s\wedge T_n^2}^{T\wedge T_n^2} H_t dt \middle| \mathcal{F}_s\right] \geq \lim_{n\to\infty} E^{\mathbb{Q}}\left[\int_{s}^{T\wedge T_n^2} H_t dt \middle| \mathcal{F}_s\right] = E^{\mathbb{Q}}\left[\int_{s}^{T} H_t dt \middle| \mathcal{F}_s\right],$$

where the second step is justified by the Monotone Convergence Theorem. Hence we have (3.58). □

Now let us define for each $t > 0$

$$g(t) = \left\|Y_t^{-1} H_t\right\|_\infty . \tag{3.59}$$

It is finite when $f^{-1}(hf)^{\frac{m-h-1}{h}} |\nabla f|^2 (t, \cdot)$ is bounded.

Lemma 19. *Let u be a strictly positive and bounded solution to (1.1) on $(0, \infty) \times \mathbb{R}^n$ with $|\nabla u|(t, \cdot)$ bounded for each $t > 0$. Let g and L_1 be defined according to (3.59) and (3.55). When $L_1 > 0$ and $2\lambda - m + h > 0$, it holds that*

$$g(s) \leq \frac{\left\|E^{\mathbb{Q}}\left[\frac{Y_T}{Y_s}\middle| \mathcal{F}_s\right] - 1\right\|_\infty}{m(2\lambda - m + h)\int_s^T e^{-\int_s^t L_1 h^{-1} g(u)du} dt}. \tag{3.60}$$

Proof. By (3.57) in the above lemma,

$$H_s \leq E^{\mathbb{Q}}\left[e^{\int_s^t L_1 h^{-1} Y_r^{-1} H_r dr} H_t \middle| \mathcal{F}_s\right] \leq E^{\mathbb{Q}}\left[e^{\int_s^t L_1 h^{-1} g(r) dr} H_t \middle| \mathcal{F}_s\right],$$

for any $0 \leq s < t < T$. Hence

$$E^{\mathbb{Q}}\left[H_t \middle| \mathcal{F}_s\right] \geq e^{-\int_s^t L_1 h^{-1} g(r) dr} H_s.$$

Then we can integrate both sides on t from s to T.

$$E^{\mathbb{Q}}\left[\left.\int_s^T H_t dt\right| \mathcal{F}_s\right] \geq H_s \int_s^T e^{-\int_s^t L_1 h^{-1} g(r) dr} dt.$$

Together with (3.58), it follows that

$$m\left(2\lambda - m + h\right) H_s \int_s^T e^{-\int_s^t L_1 h^{-1} g(r) dr} dt \leq E^{\mathbb{Q}}\left[\left. Y_T \right| \mathcal{F}_s\right] - Y_s,$$

which leads to

$$g\left(s\right) \leq \frac{\left\| E^{\mathbb{Q}}\left[\left.\frac{Y_T}{Y_s}\right| \mathcal{F}_s\right] - 1 \right\|_\infty}{m\left(2\lambda - m + h\right) \int_s^T e^{-\int_s^t L_1 h^{-1} g(u) du} dt}.$$

□

Let us solve this integral inequality. Set

$$G\left(s\right) = \int_s^T e^{-\int_s^t L_1 h^{-1} g(u) du} dt.$$

Then $G\left(T\right) = 0$, and

$$g\left(0\right) \leq \frac{E^{\mathbb{Q}}\left[\frac{Y_T}{Y_0}\right] - 1}{m\left(2\lambda - m + h\right) G\left(0\right)}. \tag{3.61}$$

Moreover, for any $s \in (0, T)$

$$\begin{aligned} G'\left(s\right) &= -1 + L_1 h^{-1} g\left(s\right) \int_s^T e^{-\int_s^t L_1 h^{-1} g(u) du} dt \\ &\leq -1 + L_1 h^{-1} \frac{\left\| E^{\mathbb{Q}}\left[\left.\frac{Y_T}{Y_s}\right| \mathcal{F}_s\right] - 1 \right\|_\infty}{m\left(2\lambda - m + h\right)}. \end{aligned}$$

As a consequence,

$$\begin{aligned}
G(0) &= G(T) - \int_0^T G(s)' \, ds \\
&\geq G(T) + \int_0^T \frac{m(2\lambda - m + h) - L_1 h^{-1} \left\| E^{\mathbb{Q}}\left[\left. \frac{Y_T}{Y_s} \right| \mathcal{F}_s \right] - 1 \right\|_\infty}{m(2\lambda - m + h)} ds \\
&= \frac{1}{m(2\lambda - m + h)} \int_0^T m(2\lambda - m + h) - L_1 h^{-1} \left\| E^{\mathbb{Q}}\left[\left. \frac{Y_T}{Y_s} \right| \mathcal{F}_s \right] - 1 \right\|_\infty ds.
\end{aligned}$$

This, together with (3.61) yields

$$g(0) \leq \frac{E^{\mathbb{Q}}\left[\frac{Y_T}{Y_0} \right] - 1}{\int_0^T m(2\lambda - m + h) - L_1 h^{-1} \left\| E^{\mathbb{Q}}\left[\left. \frac{Y_T}{Y_s} \right| \mathcal{F}_s \right] - 1 \right\|_\infty ds}, \tag{3.62}$$

as long as

$$\int_0^T m(2\lambda - m + h) - L_1 h^{-1} \left\| E^{\mathbb{Q}}\left[\left. \frac{Y_T}{Y_s} \right| \mathcal{F}_s \right] - 1 \right\|_\infty ds > 0. \tag{3.63}$$

To express this condition more explicitly, define

$$u_{\min}^{h,T} = \inf_{(t,x)\in[0,T]\times\mathbb{R}^n} u^h(t,x).$$

Then condition (3.63) is satisfied when

$$mh(2\lambda - m + h) - L_1 \left(\frac{\left\| u_0^h \right\|_\infty}{u_{\min}^{h,T}} - 1 \right) > 0. \tag{3.64}$$

To maximize $mh(2\lambda - m + h) - L_1 \left(\frac{\left\| u_0^h \right\|_\infty}{u_{\min}^{h,T}} - 1 \right)$, we take $\lambda = \frac{2m-h}{2} + \frac{h}{2} \left(\frac{\left\| u_0^h \right\|_\infty}{u_{\min}^{h,T}} - 1 \right)^{-1}$.

Then (3.64) becomes

$$mh-\frac{1}{2}\left(h-h_{-}\right)\left(h-h_{+}\right)\left(\frac{\left\|u_0^h\right\|_\infty}{u_{\min}^{h,T}}-1\right)+\frac{h^2}{2}\left(\frac{\left\|u_0^h\right\|_\infty}{u_{\min}^{h,T}}-1\right)^{-1}>0. \tag{3.65}$$

For simplicity, set

$$U_\pm=\frac{mh\pm|h|\sqrt{m^2+\left(h-h_{-}\right)\left(h-h_{+}\right)}}{\left(h-h_{-}\right)\left(h-h_{+}\right)}.$$

We have the following result.

Theorem 20. *Let u be a strictly positive and bounded solution to* (1.1) *on* $(0,\infty)\times\mathbb{R}^n$*. Suppose* $h\in\mathbb{R}$ *satisfies*

$$\left(h-h_{-}\right)\left(h-h_{+}\right)>0,$$

and

$$U\equiv\frac{\left\|u_0^h\right\|_\infty}{u_{\min}^{h,T}}-1\in\left(U_{-},U_{+}\right),$$

where

$$u_{\min}^{h,T}=\inf_{(t,x)\in[0,T]\times\mathbb{R}^n}u^h\left(t,x\right).$$

Then

$$u^{m+h-3}\left|\nabla u\right|^2(T,x)\leq 2\frac{\left\|u_0^h\right\|_\infty-u^h\left(T,x\right)}{mT\left(h-h_{-}\right)\left(h-h_{+}\right)U^{-1}\left(U-U_{-}\right)\left(U_{+}-U\right)}. \tag{3.66}$$

Proof. It only remains to check the conditions we have required on h, m and u before obtaining (3.66) from (3.62). They are (3.65) which appeared after solving the integral inequality, $L_1 > 0$ and $2\lambda - m + h > 0$ which appeared in the lemma above. Let $\lambda = \frac{2m-h}{2} + \frac{h}{2}\left(\frac{\left\|u_0^h\right\|_\infty}{u_{\min}^{h,T}} - 1\right)^{-1}$. Then $2\lambda - m + h > 0$ holds, and

$$L_1 = \frac{m}{2}\left(h - h_-\right)\left(h - h_+\right) + \frac{mh^2}{2}U^{-2},$$

which is positive since $\left(h - h_-\right)\left(h - h_+\right) > 0$. Moreover, U_- and U_+ are real valued since $\left(h - h_-\right)\left(h - h_+\right) > 0$. Finally (3.65) is satisfied since

$$\frac{\left\|u_0^h\right\|_\infty}{u_{\min}^{h,T}} - 1 \in \left(U_-, U_+\right).$$

$\square$

This result, together with our first result for the global case shows that when $m \in \left(1 - \frac{1}{\sqrt{n-1}}, 1 + \frac{1}{\sqrt{n-1}}\right)$, a gradient bound only depending on the maximum of initial data can be established. But when m is outside of this range, the gradient bound will depend on the minimum of u as well. A similar phenomena can be seen from the result in [39], but only for $m > 1$. Nevertheless, the meaning behind this phenomena is not clear to us yet, which is worthy of more exploring.

Chapter 4

Gradient estimate of Li-Yau's type for positive solutions to PME[1]

In this chapter, we work out another type of gradient estimate, which involves derivatives not only in space variables but also in time variable, i.e. of Li-Yau's type. The main idea is borrowed from [5], where the heat equation on a complete manifold was studied. In that paper, Bakry and Qian developed a way to conjecture the form of the best possible gradient bound under the guidance of a differential inequality for the gradient, and then linearize the bound in such a way that maximum principle can help to identify the region where that gradient bound is valid. It turns out that their method also works for the nonlinear equation PME, and our result improves the coefficient β of the time derivative in result (1.6) to be equal to 1, hence leading to a Harnack inequality with correct order in t.

---

[1] This chapter is based on a joint work with Zhongmin Qian.

4.1 Differential inequalities and the curvature

Let u be a positive solution to the PME

$$\partial_t u = \Delta u^m \text{ in } [0,\infty) \times M, \tag{4.1}$$

where M is a compact manifold, and $m > 1$. Then we adopt the same transformation as in [2] and [27], by setting

$$f = \frac{m(u^{m-1}-1)}{m-1}.$$

Then

$$\partial_t f = [(m-1)f + m] \triangle f + |\nabla f|^2. \tag{4.2}$$

Inspired by Aronson and Bénilan's result on $\mathbb{R}^n$ (1.5), define

$$X = \frac{|\nabla f|^2}{(m-1)f+m} \text{ and } Y = \frac{\partial_t f}{(m-1)f+m}.$$

Our aim is to derive an estimate in the form

$$X - Y \leq B(t, Y, f),$$

for some function B which can be either computed explicitly or determined by ordinary differential equations.

4.1.1 Basic differential inequalities

To begin with, let us derive the parabolic equation satisfied by X and Y, and then look for a suitable differential operator L to derive estimate on $(L - \partial_t) X$ and $(L - \partial_t) Y$. If no confusion is possible, we use f_t to denote $\partial_t f$ for simplicity. By chain rule,

$$\partial_t Y = \Delta f_t + 2 \langle \nabla f, \nabla Y \rangle + 2(m-1) YX + (m-1) Y \Delta f - (m-1) Y^2, \tag{4.3}$$

and

$$\begin{aligned} \Delta Y = &\frac{1}{(m-1) f + m} \Delta f_t - \frac{m-1}{(m-1) f + m} Y \Delta f \\ &- \frac{2(m-1)}{(m-1) f + m} \langle \nabla f, \nabla Y \rangle, \end{aligned} \tag{4.4}$$

combining equations (4.3) and (4.4), together with the relation that $\Delta f = Y - X$, we deduce that

$$(L - \partial_t) Y = -(m-1) Y^2. \tag{4.5}$$

where

$$L = [(m-1) f + m] \Delta + 2m \nabla f . \nabla, \tag{4.6}$$

which is the desired elliptic operator we are looking for.

We next compute $(L - \partial_t) X$. Define Γ_2 by

$$\Gamma_2 f = \frac{1}{2} \Delta |\nabla f|^2 - \langle \nabla f, \nabla \Delta f \rangle .$$

Then we have

$$2\Gamma_2 f = LX + (m-1)\,X\,(Y-X) - 2\,\langle \nabla f, \nabla Y\rangle\,.$$

That is

$$LX = 2\Gamma_2 f + 2\,\langle \nabla f, \nabla Y\rangle - (m-1)\,X\,(Y-X)\,. \tag{4.7}$$

On the other hand we have

$$\partial_t X = 2\,\langle \nabla f, \nabla Y\rangle + (m-1)\,YX.$$

Together with (4.7) we obtain

$$(L-\partial_t)\,X = 2\Gamma_2 f + (m-1)\,X^2 - 2\,(m-1)\,XY. \tag{4.8}$$

Next we assume that Ricci curvature is bounded from below by $-K$. Then by Bochner identity,

$$\Gamma_2 f = \|Hessf\|^2 + Ric\,(\nabla f, \nabla f) \geq \frac{1}{n}\,(\Delta f)^2 - K\,|\nabla f|^2\,.$$

Using this condition we can deduce that

$$(L-\partial_t)\,X \geq \frac{2}{n}\,(X-Y)^{\,2} - 2K\,((m-1)\,f + m)\,X + (m-1)\,X^2 - 2\,(m-1)\,XY. \tag{4.9}$$

Together with (4.5) we deduce that

$$(L-\partial_t)(X-Y) \geq \left(\frac{2}{n}+m-1\right)(X-Y)^2 - 2K((m-1)f+m)X, \tag{4.10}$$

which is the fundamental differential inequality we will use.

4.1.2 Further differential inequality

For any function $B(t,Y,f)$, consider $F = X - Y - B(t,Y,f)$ and $G = tF$. Then by using (4.10), (4.5) and (4.2) we have

$$\begin{aligned}
&(L-\partial_t)F \\
\geq&\left(\frac{2}{n}+m-1\right)(F+B)^2 + \partial_Y B(m-1)Y^2 \\
&-[(m-1)f+m]\partial_Y^2 B|\nabla Y|^2 \\
&-2[(m-1)f+m]\partial_{Yf}^2\langle\nabla Y,\nabla f\rangle \\
&-2K[(m-1)f+m](F+B+Y) \\
&-[(m-1)f+m]^2\partial_f^2 B(F+B+Y) \\
&-\partial_f B(2m-1)[(m-1)f+m](F+B+Y)+\partial_t B.
\end{aligned}$$

Hence

$$
\begin{aligned}
&\left(L-\partial_t\right)G\\
\geq & t\left(\frac{2}{n}+m-1\right)\left(\frac{G}{t}+B\right)^2+t\partial_Y B\left(m-1\right)Y^2\\
&-t\left[\left(m-1\right)f+m\right]\partial_Y^2 B\left|\nabla Y\right|^2\\
&-2t\left[\left(m-1\right)f+m\right]\partial_{Yf}^2\left\langle\nabla Y,\nabla f\right\rangle\\
&-2tK\left[\left(m-1\right)f+m\right]\left(\frac{G}{t}+B+Y\right)-\frac{G}{t}\\
&-t\left[\left(m-1\right)f+m\right]^2\partial_f^2 B\left(\frac{G}{t}+B+Y\right)\\
&-t\partial_f B\left(2m-1\right)\left[\left(m-1\right)f+m\right]\left(\frac{G}{t}+B+Y\right)+t\partial_t B. \qquad (4.11)
\end{aligned}
$$

Based on this, we can use maximum principle to test whether a function $B\left(t,Y,f\right)$ is a bound for $X-Y$.

4.2 Conjecture about the gradient bound

When the space M consists of only one point, differential inequality (4.10) for $X-Y$ becomes

$$
-\partial_t\left(X-Y\right)\geq\left(\frac{2}{n}+m-1\right)\left(X-Y\right)^2-2K\left(\left(m-1\right)f+m\right)X. \qquad (4.12)
$$

Therefore, we should expect the upper bound $c\left(t,Y,f\right)$ on $X-Y$ to satisfy

$$
-\partial_t c=\left(\frac{2}{n}+m-1\right)c^2-2K\left(\left(m-1\right)f+m\right)\left(c+Y\right).
$$

However, the existence of f in it makes the problem very complicated. Until now we are only able to deal with the simplified case when a bound for the solution is introduced. Define

$$R = \sup_{[0,\infty)\times M} \left| K\left[(m-1)f + m\right]\right|.$$

Also, for simplicity, we use

$$N = \frac{1}{\frac{1}{n} + \frac{1}{2}(m-1)}. \tag{4.13}$$

Then from (4.10) we have

$$(L - \partial_t)(X - Y) \geq \frac{2}{N}(X - Y)^2 - 2RX.$$

Let us assume

$$N > 0 \text{ and } Y_0 + \frac{NR}{4} > 0,$$

and consider the ordinary differential equation

$$\frac{dC}{dt} + \frac{2}{N}C^2 - 2R(C + Y_0) = 0, \tag{4.14}$$

with $C(0) = \infty$. Solving it yields

$$C(t, Y_0) = \frac{NR}{2} + \frac{N}{2t}\frac{b(t, Y_0)}{2}\coth\frac{b(t, Y_0)}{2}, \tag{4.15}$$

where

$$b(t, Y) = \frac{4t}{N}\sqrt{NR\left(Y_0 + \frac{NR}{4}\right)}.$$

In terms of the property of this solution, first of all, since

$$b \coth \frac{b}{2} \geq 2, \ \forall b > 0,$$

we have

$$\frac{4t}{N} C\left(t, Y_0\right) - 2tR > 2. \tag{4.16}$$

Secondly, $Y \to C\left(t, Y\right)$ is differentiable when $Y + \frac{NR}{4} > 0$. We can explicitly work out that

$$\partial_Y C\left(t, Y_0\right) = 2tRH, \tag{4.17}$$

where

$$H = H(t, Y) = \frac{e^{2b(t,Y)} - 1 - 2e^{b(t,Y)} b\left(t, Y\right)}{\left(e^{b(t,Y)} - 1\right)^2 b\left(t, Y\right)},$$

for simplicity. Note that by the elementary inequality

$$0 \leq \frac{e^{2b} - 1 - 2e^b b}{\left(e^b - 1\right)^2 b} \leq \frac{1}{3}, \ \forall b \geq 0, \tag{4.18}$$

we have an estimate on $\partial_Y C$, which reads

$$\partial_Y C\left(t, Y_0\right) \in \left[0, \frac{2}{3} tR\right]. \tag{4.19}$$

Define $B\left(t, Y\right)$ to be the linearization in the variable Y of C, that is

$$B\left(t, Y\right) = \partial_Y C\left(t, Y_0\right)\left(Y - Y_0\right) + C\left(t, Y_0\right), \tag{4.20}$$

where parameter Y_0 will be determined later. Then we conjecture that

$$X - Y \leq B\left(t, Y\right). \tag{4.21}$$

4.3 Applying maximum principle

It only remains to substitute expressions for B and $\partial_t B$ into (4.11) before using maximum principle to test whether (4.21) holds. By (4.14) and (4.20),

$$\begin{aligned}
\partial_t B\left(t, Y\right) &= \partial_t \partial_Y C\left(t, Y_0\right)\left(Y - Y_0\right) + \partial_t C\left(t, Y_0\right) \\
&= -\frac{4}{N} C\left(t, Y_0\right) \partial_Y C\left(t, Y_0\right)\left(Y - Y_0\right) \\
&\quad + 2R\left(\partial_Y C\left(t, Y_0\right) + 1\right)\left(Y - Y_0\right) \\
&\quad - \frac{2}{N} C\left(t, Y_0\right)^2 + 2R\left(C\left(t, Y_0\right) + Y_0\right).
\end{aligned}$$

This, together with (4.20) and (4.11) yields

$$\begin{aligned}
&(L-\partial_t)\,G\\
\geq & t\frac{2}{N}\left(\frac{G}{t}+B\right)^2+t\partial_Y B\,(m-1)\,Y^2\\
&-2tK\left[(m-1)\,f+m\right]\left(\frac{G}{t}+B+Y\right)-\frac{G}{t}+t\partial_t B\\
\geq & t\frac{2}{N}\left(\frac{G}{t}+B\right)^2+t\partial_Y B\,(m-1)\,Y^2\\
&-2tR\left(\frac{G}{t}+B+Y\right)-\frac{G}{t}+t\partial_t B\\
= & t\frac{2}{N}\left(\frac{G}{t}+\partial_Y C\,(t,Y_0)\,(Y-Y_0)\right)^2+t\partial_Y C\,(t,Y_0)\,(m-1)\,Y^2\\
&+\left(\frac{4}{N}C\,(t,Y_0)-2R-\frac{1}{t}\right)G.
\end{aligned}$$

By estimate (4.19) about $\partial_Y C$,

$$\partial_Y C\,(t,Y_0)\,(m-1)\,Y^2\geq 0,\ \text{ if } m\geq 1.$$

Therefor, let us assume $m\geq 1$. Then

$$(L-\partial_t)\,G\geq\left(\frac{4}{N}C\,(t,Y_0)-2R-\frac{1}{t}\right)G. \tag{4.22}$$

Since M is compact, smooth function G on $[0,T]\times M$ has a maximum point (t_0,x_0). Assume $t_0>0$ and $G\,(t_0,x_0)>0$. Then at (x_0,t_0), it holds that $\Delta G\leq 0$, $\nabla G=0$ and $\partial_t G\geq 0$. Hence the left hand side of (4.22) is nonpositive at (x_0,t_0). On the other hand, by estimate (4.16) about C,

$$\frac{4}{N}C\left(t,Y_0\right)-2R-\frac{1}{t}>\frac{1}{t},\ \forall Y_0>-\frac{NR}{4}.$$

So the right hand side of (4.22) is strictly positive at (x_0,t_0). We thus have a contradiction to (4.22). Hence we either have $G\le 0$ or $t_0=0$. But from the definition of G, $G\left(0,\cdot\right)=0$. Therefore, we always have $G\le 0$, which means

$$\frac{\left|\nabla f\right|^2}{\left(m-1\right)f+m}-\frac{f_t}{\left(m-1\right)f+m}\le\partial_Y C\left(t,Y_0\right)\left(\frac{f_t}{\left(m-1\right)f+m}-Y_0\right)+C\left(t,Y_0\right),\tag{4.23}$$

for any $Y_0>-\frac{NR}{4}$.

Theorem 21. *Let u be a positive solution to* (4.1) *on a compact manifold with Ric $\ge -K$, $f=\frac{m\left(u^{m-1}-1\right)}{m-1}$ and*

$$R=\sup_{[0,\infty)\times M}\left|K\left[\left(m-1\right)f+m\right]\right|.$$

Then

$$\frac{\left|\nabla f\right|^2}{\left(m-1\right)f+m}-\frac{f_t}{\left(m-1\right)f+m}\le C\left(t,\frac{f_t}{\left(m-1\right)f+m}\right),\ \ \textit{on}\ \frac{f_t}{\left(m-1\right)f+m}>-\frac{NR}{4},\tag{4.24}$$

$$\begin{aligned}&\frac{|\nabla f|^2}{(m-1)f+m}-\frac{f_t}{(m-1)f+m}\\ \le\ &\frac{2R}{3}t\left(\frac{f_t}{(m-1)f+m}+\frac{NR}{4}\right)+\frac{NR}{2}+\frac{N}{2t}\\ \le\ &\frac{N}{2t}+\frac{NR}{2},\ \textit{on}\ \frac{f_t}{(m-1)f+m}<-\frac{NR}{4},\end{aligned}\tag{4.25}$$

and

$$\frac{|\nabla f|^2}{(m-1)f+m}-\frac{f_t}{(m-1)f+m}\le\inf_{Y_0>-\frac{NR}{4}}\left(\partial_Y C(t,Y_0)\left(\frac{f_t}{(m-1)f+m}-Y_0\right)+C(t,Y_0)\right),\tag{4.26}$$

where N and C are defined according to (4.13) *and* (4.15).

Proof. We have proved (4.23), which directly implies (4.26). By taking $Y_0=\frac{f_t}{(m-1)f+m}$ in (4.23), we have (4.24). As for (4.25). We first compute from (4.15) and (4.17) that

$$\lim_{Y_0\to-\frac{NR}{4}}C(t,Y_0)=\frac{NR}{2}+\frac{N}{2t},\quad\lim_{Y_0\to-\frac{NR}{4}}\partial_Y C(t,Y_0)=\frac{2R}{3}t.$$

Then (4.25) follows from (4.23) by letting $Y_0\to-\frac{NR}{4}$. □

4.4 Harnack inequality

One can see that gradient bound (4.23) we obtained in the section above is in fact a bound on

$$\frac{|\nabla f|^2}{(m-1)f+m}-(1+\partial_Y C(t,Y_0))\frac{f_t}{(m-1)f+m},$$

and the principal term in t in our bound is $\frac{N}{2t}$. Recall from formula (4.17) for $\partial_Y C$ that

$$\lim_{t\to 0} \partial_Y C\left(t, Y_0\right) = 0,$$

which means that the coefficient of $\frac{f_t}{(m-1)f+m}$ in our bound tends to 1 as $t \to 0$. In contrast, in the result (1.6) by Villani et al, the bound explodes as that coefficient tends to 1. With this advantage, our bound implies a Harnack inequality that has correct order in the time variable.

For any x_1, x_2 in M and $t_2 > t_1 > 0$, let γ be a minimal geodesic joining x_1 and x_2 with constant speed, so that $\gamma\left(t_1\right) = x_1$, $\gamma\left(t_2\right) = x_2$ and $\left|\frac{d\gamma}{dt}\right| = \frac{d(x_1,x_2)}{t_2-t_1}$. Define

$$p\left(s\right) = \left(s, \gamma\left(s\right)\right).$$

Furthermore, define

$$u_{\min}^{m-1} = \inf u^{m-1}.$$

So

$$\inf\left(\left(m-1\right)f + m\right) = m u_{\min}^{m-1}.$$

Then

$$
\begin{aligned}
& \log\left(\frac{(m-1)\, f\,(t_2, x_2)+m}{(m-1)\, f\,(t_1, x_1)+m}\right) \\
= & \int_{t_1}^{t_2} \frac{d}{ds}\left(\log\left[(m-1)\, f\,(p\,(s))+m\right]\right) ds \\
= & \int_{t_1}^{t_2}\left((m-1)\frac{\nabla f}{(m-1)\, f+m}\cdot\frac{d\gamma}{ds}+(m-1)\frac{f_t}{(m-1)\, f+m}\right) ds \\
\geq & -(m-1)\int_{t_1}^{t_2}\left(\frac{(1+\partial_Y C\,(s, Y_0))}{4\left[(m-1)\, f+m\right]}\left|\frac{d\gamma}{ds}\right|^2\right) ds \\
& -(m-1)\int_{t_1}^{t_2}\left(\frac{1}{1+\partial_Y C\,(s, Y_0)}\frac{|\nabla f|^2}{(m-1)\, f+m}-\frac{f_t}{(m-1)\, f+m}\right) ds \\
\geq & -(m-1)\int_{t_1}^{t_2}\left(\frac{(1+\partial_Y C\,(s, Y_0))}{4\left[(m-1)\, f+m\right]}\left|\frac{d\gamma}{ds}\right|^2+\frac{C\,(s, Y_0)-\partial_Y C\,(s, Y_0)\, Y_0}{1+\partial_Y C\,(s, Y_0)}\right) ds \\
\geq & -\frac{m-1}{4mu_{\min}^{m-1}}\frac{d^2\,(x_1, x_2)}{t_2-t_1}-\frac{m-1}{4mu_{\min}^{m-1}}\frac{\int_{t_1}^{t_2}\partial_Y C\,(s, Y_0)\, ds}{(t_2-t_1)^2} d^2\,(x_1, x_2) \\
& -(m-1)\int_{t_1}^{t_2}\left(C\,(s, Y_0)-\frac{\partial_Y C\,(s, Y_0)}{1+\partial_Y C\,(s, Y_0)}\left(Y_0+C\,(s, Y_0)\right)\right) ds,
\end{aligned}
$$

where the second inequality results from our gradient bound (4.23). Let us work out explicitly the dominating term in time variable. Recall that

$$
\begin{aligned}
C\,(s, Y_0) &= \frac{NR}{2}+\frac{N}{2s}\frac{b\,(s, Y_0)}{2}\coth\frac{b\,(s, Y_0)}{2} \\
&= \frac{NR}{2}+\sqrt{NR}\sqrt{Y_0+\frac{NR}{4}}\coth\left(\frac{2s}{N}\sqrt{NR}\sqrt{Y_0+\frac{NR}{4}}\right).
\end{aligned}
$$

Hence

$$\begin{aligned}
\int_{t_1}^{t_2} C(s, Y_0)\, ds &= \frac{NR}{2}(t_2 - t_1) + \frac{N}{2}\int_{\frac{2t_1}{N}\sqrt{NR}\sqrt{Y_0+\frac{NR}{4}}}^{\frac{2t_2}{N}\sqrt{NR}\sqrt{Y_0+\frac{NR}{4}}} \coth(t)\, dt \\
&= \frac{NR}{2}(t_2 - t_1) + \frac{N}{2}\log\left(\frac{\sinh\left(\frac{2t_2}{N}\sqrt{NR}\sqrt{Y_0 + \frac{NR}{4}}\right)}{\sinh\left(\frac{2t_1}{N}\sqrt{NR}\sqrt{Y_0 + \frac{NR}{4}}\right)}\right) \\
&= \frac{NR}{2}(t_2 - t_1) + \frac{N}{2}\log\left(\frac{\sinh\left(\frac{2t_2}{N}\sqrt{NR}\sqrt{Y_0 + \frac{NR}{4}}\right)}{\sinh\left(\frac{2t_1}{N}\sqrt{NR}\sqrt{Y_0 + \frac{NR}{4}}\right)}\right).
\end{aligned}$$

For simplicity, define

$$\begin{aligned}
Q\left(u_{\min}^{m-1}, Y_0, t_1, t_2\right) &= \frac{1}{4mu_{\min}^{m-1}}\frac{\int_{t_1}^{t_2}\partial_Y C(s, Y_0)\, ds}{(t_2 - t_1)^2} d^2(x_1, x_2) \\
&\quad - \int_{t_1}^{t_2}\left(\frac{\partial_Y C(s, Y_0)}{1 + \partial_Y C(s, Y_0)}(Y_0 + C(s, Y_0))\right) ds \\
&\quad + \frac{NR}{2}(t_2 - t_1).
\end{aligned} \tag{4.27}$$

As a result we obtain the following Harnack inequality.

Theorem 22. *Let u be a positive solution to* (4.1) *on a compact manifold with Ric $\geq -K$, $f = \frac{m(u^{m-1}-1)}{m-1}$ and*

$$R = \sup_{[0,\infty)\times M} |K[(m-1)f + m]|, \; u_{\min}^{m-1} = \inf[(m-1)f + m].$$

Then for any $Y_0 > -\frac{NR}{4}$*, any* x_1*,* x_2 *in* M *and any* $t_2 > t_1 > 0$*,*

$$\begin{aligned}
&\frac{(m-1)\,f\left(t_2,x_2\right)+m}{(m-1)\,f\left(t_1,x_1\right)+m}\\
\geq&\left(\frac{\sinh\left(\frac{2t_1}{N}\sqrt{NR}\sqrt{Y_0+\frac{NR}{4}}\right)}{\sinh\left(\frac{2t_2}{N}\sqrt{NR}\sqrt{Y_0+\frac{NR}{4}}\right)}\right)^{(m-1)\frac{N}{2}}\\
&\cdot\exp\left((m-1)\left(\frac{-1}{4mu_{\min}^{m-1}}\frac{d^2\left(x_1,x_2\right)}{t_2-t_1}-Q\left(u_{\min}^{m-1},Y_0,t_1,t_2\right)\right)\right),
\end{aligned}$$

where N*,* C *and* Q *are defined according to* (4.13)*,* (4.15) *and* (4.27)*.*

Remark 23. When t_1 and t_2 are small, the order of magnitude in $\left(\frac{t_2}{t_1}\right)$ of the right hand side of the above inequality is $(m-1)\,\frac{N}{2}$, which improves (1.7) by Villani et al mentioned in our introduction.

With a similar argument, we can derive another Harnack inequality, which is able to recover the result for the heat equation by Bakry and Qian in [5]. Define

$$\begin{aligned}
&Q'\left(u_{\max}^{m-1},Y_0,t_1,t_2\right)\\
=&\frac{\int_{t_1}^{t_2}\partial_Y C\left(s,Y_0\right)ds}{\left(t_2-t_1\right)^2}\frac{d^2\left(x_1,x_2\right)}{4}\\
&-mu_{\max}^{m-1}\int_{t_1}^{t_2}\left(\frac{\partial_Y C\left(s,Y_0\right)}{1+\partial_Y C\left(s,Y_0\right)}\left(Y_0+C\left(s,Y_0\right)\right)\right)ds\\
&+\frac{NR}{2}mu_{\max}^{m-1}\left(t_2-t_1\right). \qquad (4.28)
\end{aligned}$$

Theorem 24. *Let* u *be a positive solution to* (4.1) *on a compact manifold with* $Ric \geq$

$-K$, $f = \frac{m(u^{m-1}-1)}{m-1}$ and

$$R = \sup_{[0,\infty)\times M} |K[(m-1)f+m]|, \; u_{\max}^{m-1} = \sup[(m-1)f+m].$$

Then for any $Y_0 > -\frac{NR}{4}$, any x_1, x_2 in M and any $t_2 > t_1 > 0$,

$$\begin{aligned}&f(t_2,x_2) - f(t_1,x_1)\\ \geq& m\frac{N}{2}u_{\max}^{m-1}\log\left(\frac{\sinh\left(\frac{2t_1}{N}\sqrt{NR}\sqrt{Y_0+\frac{NR}{4}}\right)}{\sinh\left(\frac{2t_2}{N}\sqrt{NR}\sqrt{Y_0+\frac{NR}{4}}\right)}\right) - \frac{d^2(x_1,x_2)}{4(t_2-t_1)} - Q'\left(u_{\max}^{m-1}, Y_0, t_1, t_2\right).\end{aligned}$$

where N, C and Q' are defined according to (4.13), (4.15) *and* (4.28).

Proof. The proof is completed by the following inequality.

$$\begin{aligned}&f(t_2,x_2) - f(t_1,x_1)\\ =& \int_{t_1}^{t_2}\left(\nabla f\cdot\frac{d\gamma}{ds} + f_t\right)ds\\ \geq& -\int_{t_1}^{t_2}\left(\frac{(1+\partial_Y C(s,Y_0))}{4}\left|\frac{d\gamma}{ds}\right|^2 + \frac{|\nabla f|^2}{1+\partial_Y C(s,Y_0)} - f_t\right)ds\\ \geq& -\int_{t_1}^{t_2}\left(\frac{(1+\partial_Y C(s,Y_0))}{4}\left|\frac{d\gamma}{ds}\right|^2 + \frac{C(s,Y_0) - \partial_Y C(s,Y_0)Y_0}{1+\partial_Y C(s,Y_0)}mu_{\max}^{m-1}\right)ds\\ =& -\frac{d^2(x_1,x_2)}{4(t_2-t_1)} - \frac{\int_{t_1}^{t_2}\partial_Y C(s,Y_0)\,ds}{(t_2-t_1)^2}\frac{d^2(x_1,x_2)}{4}\\ & -mu_{\max}^{m-1}\int_{t_1}^{t_2}\left(C(s,Y_0) - \frac{\partial_Y C(s,Y_0)}{1+\partial_Y C(s,Y_0)}(Y_0 + C(s,Y_0))\right)ds,\end{aligned}$$

where the second inequality results from multiplying both sides of (4.23) by $mu_{\max}^{m-1}$.

□

Appendix A

Preliminaries

A.1 Basics of Martingale theory

A stochastic process valued in $\mathbb{R}^n$ can be viewed as a sequence of random variables indexed by a time parameter $t \in [0, \infty)$. It is uniquely determined up to an equivalent class by all of its finite-dimensional distributions. By the Daniell-Kolmogorov existence theorem [22], as long as all the finite-dimensional distributions are consistent, there exists a measure on $\left((\mathbb{R}^n)^{[0,\infty)}, \mathcal{B}\left((\mathbb{R}^n)^{[0,\infty)} \right) \right)$ such that its canonical process have those prescribed finite-dimensional distributions. When a process X satisfies

$$E\left[|X_t - X_s|^r \right] \leq C \left| t - s \right|^{1+\alpha}, \ 0 \leq s \leq t \leq T,$$

for some $\alpha > 0$ and $r > 1 + \alpha$, by A. N. Kolmogorov [34] X has a version X' among its equivalent class such that $t \to X'_t(\omega)$: $[0, T] \to \mathbb{R}^n$ is continuous for $\mathbb{P}-$almost all $\omega \in \Omega$, which means it has continuous sample paths. For such processes, a good choice of $(\Omega, \mathcal{F}, \mathbb{P})$ is the so-called Wiener space. Here Ω is the linear space consisting of all continuous functions from $[0,\infty)$ to $\mathbb{R}^n$ starting from 0. It is equipped with the metric $d(\cdot,\cdot)$ given by

$$d(\omega,\omega') = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\sup_{0\leq t\leq n} |\omega(t) - \omega'(t)|}{1+\sup_{0\leq t\leq n} |\omega(t) - \omega'(t)|}, \ \forall\, \omega, \omega' \in \Omega,$$

which corresponds to the uniform convergence. $\mathcal{F}$ is the Borel $\sigma-$algebra associated with the topology generated by the metric $d(\cdot,\cdot)$. Then for each t, define a map W_t: $\Omega \to \mathbb{R}^n$ by

$$W_t(\omega) = \omega(t).$$

W_t is continuous from metric space (Ω, d) to $\mathbb{R}^n$, hence measurable to $\mathcal{F}$. $\mathbb{P}$ is the probability measure on $(\Omega, \mathcal{F})$ such that

1. For any $0 \leq t_0 < t_1 < \cdots < t_n$, random variables W_{t_0}, $W_{t_1} - W_{t_0}, \cdots, W_{t_n} - W_{t_{n-1}}$are independent.

2. For any $t > s \geq 0$, $W_t - W_s$ has a normal distribution $N(0, t-s)$.

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called Wiener space. $W = W_t(\omega)$ is called standard Brownian motion on $\mathbb{R}^n$. To characterize the information generated by W_t as t evolves, the filtration $\{\mathcal{F}_t^0\}_{t\geq 0}$ is introduced by defining $\mathcal{F}_t^0 = \sigma(W_s, s \leq t)$. It is readily seen that $\omega \to W_s(\omega)$ is $\mathcal{F}_t^0-$measurable for any $s \leq t$, which means W is adapted to $\{\mathcal{F}_t^0\}_{t\geq 0}$. Moreover, $t \to W_t(\omega)$ is continuous for each $\omega \in \Omega$. Therefore, W is progressively measurable to $\{\mathcal{F}_t^0\}_{t\geq 0}$ in the sense that the mapping $(s,\omega) \to W_s(\omega)$: $([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t^0) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ is measurable for each $t > 0$.

The progressive measurability of a process enables one to take integration of it on $([0,t]\times\Omega,\mathcal{B}([0,t])\otimes\mathcal{F}_t^0,\mu\times\mathbb{P})$. To get a right continuous filtration, we enlarge $\mathcal{F}_t^0$ by defining $\mathcal{F}_t=\sigma\left\{\mathcal{F}_t^0,\mathcal{N}\right\}$, where $\mathcal{N}$ is the collection of all subsets of measure zero sets in Ω. Then the strong Markov property of Brownian motion W ensures that $\mathcal{F}_t=\cap_{\epsilon>0}\mathcal{F}_{t+\epsilon}$. On the filtered probability space $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathbb{P}\right)$, a stopping time T is a random variable such that $\{T\leq t\}$ is $\mathcal{F}_t-$measurable for any $t>0$. The right continuity of the augmented filtration $\{\mathcal{F}_t\}_{t\geq 0}$ ensures that for any right continuous and adapted process X on $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathbb{P}\right)$ and open set O in $\mathbb{R}^n$, the first hitting time $\tau=\inf\{t:X_t\in O\}$ is a stopping time (problem 1.2.6 in [22]).

On $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathbb{P}\right)$, a submartingale (respectively, a supermartingale) M is a stochastic process valued in $\mathbb{R}$ satisfying $E\left[M_t|\,\mathcal{F}_s\right]\geq M_s$ (respectively, $E\left[M_t|\,\mathcal{F}_s\right]\leq M_s$) for any $0\leq s\leq t$. A martingale is both a submartingale and a supermartingale. The martingale property implies a close relation between the expected behavior of the process in the future and its value in the present. In fact this relation holds not only on deterministic times but also on stopping times.

Theorem 25. *(Doob's optional sampling theorem [32]) Let M be a submartingale and $S\leq T$ be two bounded stopping times of $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathbb{P}\right)$. Then*

$$E\left[M_T|\,\mathcal{F}_S\right]\geq M_S.$$

With this powerful tool, the pathwise convergence of martingales can be tested by the following criteria.

Theorem 26. *(Martingale convergence theorem [22]) Let M be a right continuous* *submartingale with* $\sup_{t\geq 0} E\left[M_t^+\right] < \infty$. *Then* $\lim_{t\to\infty} M_t(\omega)$ *exists for almost all* $\omega \in \Omega$, *and this limit is integrable.*

An important property of martingales is that their running maximums can be controlled by their final values in terms of L^p norm for any $p > 1$.

Theorem 27. *(Doob's* L^p-*inequality [32]) If* M *is a submartingale, then for* $p > 1$

$$E\left[\left(\sup_{t\leq T} M_t^+\right)^p\right] \leq \left(\frac{p}{p-1}\right)^p E\left[\left(M_T^+\right)^p\right].$$

Another celebrated inequality is due to Donald Burkholder, Burgess Davis and Richard F. Gundy, which shows that one can estimate the running maximum of a martingale by its quadratic variation, or the other way around, in terms of L^p for any $p > 0$.

Theorem 28. *(Burkholder–Davis–Gundy inequalities) Let* $T > 0$ *and* $(M_t)_{0\leq t\leq T}$ *be a continuous local martingale such that* $M_0 = 0$. *For every* $0 < p < \infty$, *there exist universal constants* c_p *and* C_p, *independent of* T *and* $(M_t)_{0\leq t\leq T}$ *such that*

$$c_p \mathbb{E}\left(\langle M\rangle_T^{\frac{p}{2}}\right) \leq \mathbb{E}\left(\left(\sup_{0\leq t\leq T} |M_t|\right)^p\right) \leq C_p \mathbb{E}\left(\langle M\rangle_T^{\frac{p}{2}}\right).$$

A.2 Change of measures on Wiener space

For the Wiener space $(\Omega, \mathcal{F}, \mathbb{P})$, if there is another measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$, such that it is equivalent to $\mathbb{P}$, then Girsanov's theorem tells us that any semimartingale under $\mathbb{P}$ is still a semimartingale under $\mathbb{Q}$. In particular, it specifies the relation between the law of the canonical process W under the new measure $\mathbb{Q}$ and the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$.

Theorem 29. *(Girsanov's theorem [22])*

Let M_t be a continuous local martingale on $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\right)$ with $M_0 = 0$ and its stochastic exponential

$$\exp\left(M_t - \frac{1}{2}\langle M\rangle_t\right)$$

being a martingale under $\mathbb{P}$. Define a measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ such that

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left(M_t - \frac{1}{2}\langle M\rangle_t\right). \tag{A.1}$$

Then

$$W_t - \langle M, W\rangle_t ,$$

is a Brownian motion under measure $\mathbb{Q}$.

For any measure $\mathbb{Q}$ that is equivalent to $\mathbb{P}$, their Radon-Nikodym derivative ξ must be strictly positive. So $\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t}$ can always be written in the form of (A.1). By using Ito's formula on $\log\left(\frac{1}{E^{\mathbb{P}}[\xi|\mathcal{F}_t]}\right)$, one can see that

$$M_t = \int_0^t \frac{1}{E^{\mathbb{P}}\left[\xi|\,\mathcal{F}_t\right]} d\left(E^{\mathbb{P}}\left[\xi|\,\mathcal{F}_t\right]\right).$$

Therefore, Girsanov's theorem is applicable to any equivalent change of measure.

There are a lot of literature devoted to the study about sufficient conditions on M under which $\exp\left(M_t - \frac{1}{2}\langle M\rangle_t\right)$ is a martingale. A well-known result is due to A.

A. Novikov [30].

Theorem 30. *(Novikov) If M is a continuous local martingale such that*

$$M_0 = 0, \ E^{\mathbb{P}}\left[\exp\left(\frac{1}{2}\left\langle M\right\rangle_T\right)\right] < \infty,$$

Then $\exp\left(M_t - \frac{1}{2}\left\langle M\right\rangle_t\right)$ *is a martingale up to time T under $\mathbb{P}$.*

If we restrict M to be Gaussian, i.e. $M_t = \int_0^t h(s) \cdot dW_s$ for some deterministic function h, then by Theorem 2.2 in Chapter 8 of [32], Novikov's condition is also a necessary condition. This means that if we define a new measure $\mathbb{P} \circ \tau_g^{-1}$ on $(\Omega, \mathcal{F})$, where $g \in \Omega$ and τ_g is the translation on Ω given by

$$\tau_g(\omega) = \omega + g,$$

then this measure is equivalent to $\mathbb{P}$ if and only if $g(t) = \int_0^t h(s)\, ds$, $h \in L^2$. The collection of such g is called Cameron-Martin space. Therefore, although a random variable is only almost surely defined, its partial derivative along the direction of an element in the Cameron-Martin space is still well defined. This observation is the starting point of one way to build up the theory of Malliavin calculus.

Another sufficient condition for the martingale property of (A.1) is

$$E^{\mathbb{P}}\left[\sup_t M_t^2\right] < \infty, \text{ and } E^{\mathbb{P}}\left[\left\langle M\right\rangle_\infty - \left\langle M\right\rangle_s \middle| \mathcal{F}_S\right] \leq c, \ \forall \text{ stopping time } S. \qquad \text{(A.2)}$$

With this property, a martingale is called a BMO martingale, which is a probabilistic version of a function with bounded mean oscillation.

Theorem 31. *([23]) If M is a martingale satisfying* (A.2)*, then its stochastic exponential*

$$\mathcal{E}\left(M\right)_t \triangleq \exp\left(M_t - \frac{1}{2}\left\langle M\right\rangle_t\right),$$

is a uniformly integrable martingale.

Proof. Since $E\left[\sup_t M_t^2\right] < \infty$, M_∞ exists and $E\left[M_\infty \middle| \mathcal{F}_S\right] = M_S$, $\forall$ stopping time S. By definition $\mathcal{E}\left(M\right)_t$ is a positive local martingale, hence a positive supermartingale. By martingale convergence theorem, $\mathcal{E}\left(M\right)_\infty \triangleq \lim_{t\to\infty} \mathcal{E}\left(M\right)_t$ exists. Also, Fatou's Lemma

$$E\left[\mathcal{E}\left(M\right)_\infty\right] = E\left[\liminf_{t\to\infty} \mathcal{E}\left(M\right)_t\right] \leq \liminf_{t\to\infty} E\left[\mathcal{E}\left(M\right)_t\right] \leq E\left[\mathcal{E}\left(M\right)_0\right] = 1.$$

By the Jensen's Inequality, for any stopping time S,

$$\begin{aligned} E\left[\frac{\mathcal{E}\left(M\right)_\infty}{\mathcal{E}\left(M\right)_S}\middle| \mathcal{F}_S\right] &= E\left[\exp\left(M_\infty - M_S - \frac{1}{2}\left(\left\langle M\right\rangle_\infty - \left\langle M\right\rangle_S\right)\right)\middle| \mathcal{F}_S\right] \\ &\geq \exp\left(E\left[M_\infty - M_S - \frac{1}{2}\left(\left\langle M\right\rangle_\infty - \left\langle M\right\rangle_S\right)\middle| \mathcal{F}_S\right]\right) \\ &= \exp\left(E\left[-\frac{1}{2}\left(\left\langle M\right\rangle_\infty - \left\langle M\right\rangle_S\right)\middle| \mathcal{F}_S\right]\right) \\ &\geq \exp\left(-\frac{1}{2}c\right), \end{aligned}$$

which implies that

$$\mathcal{E}\left(M\right)_S \leq \exp\left(\frac{1}{2}c\right) E\left[\mathcal{E}\left(M\right)_\infty \middle| \mathcal{F}_S\right].$$

Since $\mathcal{E}(M)_\infty$ is integrable,

$$\{\mathcal{E}(M)_S\,,\ S \text{ is a stopping time}\}\,,$$

is uniformly integrable. This improves the local martingale $\mathcal{E}(M)$ to be a uniformly integrable martingale. $\square$

BMO martingales are commonly seen in solutions to backward stochastic differential equations with bounded terminal values [29].

Bibliography

[1] Marc Arnaudon and Anton Thalmaier. Li-Yau type gradient estimates and Harnack inequalities by stochastic analysis. In *Probabilistic approach to geometry*, volume 57 of *Adv. Stud. Pure Math.*, pages 29–48. Math. Soc. Japan, Tokyo, 2010.

[2] D. G. Aronson. Regularity propeties of flows through porous media. *SIAM J. Appl. Math.*, 17:461–467, 1969.

[3] D. G. Aronson and L. A. Caffarelli. The initial trace of a solution of the porous medium equation. *Trans. Amer. Math. Soc.*, 280(1):351–366, 1983.

[4] Donald G. Aronson and Philippe Bénilan. Régularité des solutions de l'équation des milieux poreux dans $\mathbf{R}^N$. *C. R. Acad. Sci. Paris Sér. A-B*, 288(2):A103–A105, 1979.

[5] Dominique Bakry and Zhongmin M. Qian. Harnack inequalities on a manifold with positive or negative Ricci curvature. *Rev. Mat. Iberoamericana*, 15(1):143–179, 1999.

[6] Philippe Bénilan and Michael G. Crandall. The continuous dependence on φ of solutions of $u_t - \Delta\varphi(u) = 0$. *Indiana Univ. Math. J.*, 30(2):161–177, 1981.

[7] Philippe Bénilan, Michael G. Crandall, and Michel Pierre. Solutions of the porous medium equation in $\mathbf{R}^N$ under optimal conditions on initial values. *Indiana Univ. Math. J.*, 33(1):51–87, 1984.

[8] Guy Bernard. Existence theorems for fast diffusion equations. *Nonlinear Anal.*, 43(5, Ser. A: Theory Methods):575–590, 2001.

[9] Michiel Bertsch and Maura Ughi. Positivity properties of viscosity solutions of a degenerate parabolic equation. *Nonlinear Anal.*, 14(7):571–592, 1990.

[10] Matteo Bonforte and Juan Luis Vazquez. Global positivity estimates and Harnack inequalities for the fast diffusion equation. *J. Funct. Anal.*, 240(2):399–428, 2006.

[11] Matteo Bonforte and Juan Luis Vázquez. Positivity, local smoothing, and Harnack inequalities for very fast diffusion equations. *Adv. Math.*, 223(2):529–578, 2010.

[12] L. A. Caffarelli, J. L. Vázquez, and N. I. Wolanski. Lipschitz continuity of solutions and interfaces of the N-dimensional porous medium equation. *Indiana Univ. Math. J.*, 36(2):373–401, 1987.

[13] Luis A. Caffarelli and Avner Friedman. Regularity of the free boundary of a gas flow in an n-dimensional porous medium. *Indiana Univ. Math. J.*, 29(3):361–391, 1980.

[14] Emmanuel Chasseigne and Juan Luis Vazquez. Theory of extended solutions for

fast-diffusion equations in optimal classes of data. Radiation from singularities. *Arch. Ration. Mech. Anal.*, 164(2):133–187, 2002.

[15] Björn E. J. Dahlberg and Carlos E. Kenig. Nonnegative solutions of generalized porous medium equations. *Rev. Mat. Iberoamericana*, 2(3):267–305, 1986.

[16] P. Daskalopoulos and R. Hamilton. The free boundary for the n-dimensional porous medium equation. *Internat. Math. Res. Notices*, (17):817–831, 1997.

[17] J. L. Doob. *Stochastic processes.* John Wiley & Sons Inc., New York, 1953.

[18] K. D. Elworthy and X.-M. Li. Formulae for the derivatives of heat semigroups. *J. Funct. Anal.*, 125(1):252–286, 1994.

[19] Richard S. Hamilton. A matrix Harnack estimate for the heat equation. *Comm. Anal. Geom.*, 1(1):113–126, 1993.

[20] Miguel A. Herrero and Michel Pierre. The Cauchy problem for $u_t = \Delta u^m$ when $0 < m < 1$. *Trans. Amer. Math. Soc.*, 291(1):145–158, 1985.

[21] Guangyue Huang, Zhijie Huang, and Haizhong Li. Gradient estimates for the porous medium equations on riemannian manifolds. *Journal of Geometric Analysis*, 2012.

[22] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1991.

[23] Norihiko Kazamaki. *Continuous exponential martingales and BMO*, volume 1579 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1994.

[24] Brett L. Kotschwar. Hamilton's gradient estimate for the heat kernel on complete manifolds. *Proc. Amer. Math. Soc.*, 135(9):3013–3019 (electronic), 2007.

[25] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Uraltseva. *Linear and quasilinear equations of parabolic type*. Izdat. "Nauka", Moscow, 1967.

[26] Peter Li and Shing-Tung Yau. On the parabolic kernel of the Schrödinger operator. *Acta Math.*, 156(3-4):153–201, 1986.

[27] Peng Lu, Lei Ni, Juan-Luis Vázquez, and Cédric Villani. Local Aronson-Bénilan estimates and entropy formulae for porous medium and fast diffusion equations on manifolds. *J. Math. Pures Appl. (9)*, 91(1):1–19, 2009.

[28] Li Ma, Lin Zhao, and Xianfa Song. Gradient estimate for the degenerate parabolic equation $u_t = \Delta F(u) + H(u)$ on manifolds. *J. Differential Equations*, 244(5):1157–1177, 2008.

[29] Marie-Amélie Morlais. Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem. *Finance Stoch.*, 13(1):121–150, 2009.

[30] A. A. Novikov. On moment inequalities and identities for stochastic integrals. In *Proceedings of the Second Japan-USSR Symposium on Probability Theory (Kyoto, 1972)*, pages 333–339. Lecture Notes in Math., Vol. 330, Berlin, 1973. Springer.

[31] Jean Picard. Gradient estimates for some diffusion semigroups. *Probab. Theory Related Fields*, 122(4):593–612, 2002.

[32] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[33] Philippe Souplet and Qi S. Zhang. Sharp gradient estimate and Yau's Liouville theorem for the heat equation on noncompact manifolds. *Bull. London Math. Soc.*, 38(6):1045–1053, 2006.

[34] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979.

[35] Juan Luis Vázquez. *Smoothing and decay estimates for nonlinear parabolic equations of porous medium type.* 2005.

[36] Juan Luis Vázquez. Symmetrization and mass comparison for degenerate nonlinear parabolic and related elliptic equations. *Adv. Nonlinear Stud.*, 5(1):87–131, 2005.

[37] Juan Luis Vázquez. *Smoothing and decay estimates for nonlinear diffusion equations*, volume 33 of *Oxford Lecture Series in Mathematics and its Applications.* Oxford University Press, Oxford, 2006. Equations of porous medium type.

[38] Juan Luis Vázquez. *The porous medium equation*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, Oxford, 2007. Mathematical theory.

[39] Xiangjin Xu. Gradient estimates for $u_t = \Delta F(u)$ on manifolds and some Liouville-type theorems. *J. Differential Equations*, 252(2):1403–1420, 2012.

[40] Shing-Tung Yau. On the Harnack inequalities of partial differential equations. *Comm. Anal. Geom.*, 2(3):431–450, 1994.

[41] Xiaobao Zhu. Hamilton's gradient estimates and Liouville theorems for fast diffusion equations on noncompact Riemannian manifolds. *Proc. Amer. Math. Soc.*, 139(5):1637–1644, 2011.

[42] Xiaobao Zhu. Hamilton's gradient estimates and liouville theorems for porous medium equations on noncompact riemannian manifolds. *Journal of Mathematical Analysis and Applications*, 402(1):201 – 206, 2013.