

Gradient estimates for porous medium and fast diffusion equations by martingale method

Ying Hu, Zhongmin Qian, Zichen Zhang

▶ **To cite this version:**

Ying Hu, Zhongmin Qian, Zichen Zhang. Gradient estimates for porous medium and fast diffusion equations by martingale method. Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques, 2017, 53 (4), pp.1793-1820. 10.1214/16-AIHP771 . hal-00705097v2

HAL Id: hal-00705097
https://hal.science/hal-00705097v2

Submitted on 1 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Gradient estimates for porous medium and fast diffusion equations by martingale method

Ying Hu[*], Zhongmin Qian[†], Zichen Zhang[‡]

Contents

1 Introduction **1**

2 Gradient estimate for solutions to the heat equation **10**

3 Gradient estimate for positive solutions to PME and FDE **13**
3.1 Local gradient estimate 13
3.1.1 From PDE to SDE 15
3.1.2 Fundamental decompositions 15
3.1.3 Gradient bound for solution u 18
3.1.4 Gradient bound for pressure variable $v = \frac{u^{m-1}}{m-1}$ 26
3.2 Global gradient estimate 28
3.2.1 From local bound to global bound 28
3.2.2 Negative finite variation part 29

Abstract

In this paper, we establish several local and global gradient estimates for the positive solution of Porous Medium Equations (PMEs) and Fast Diffusion Equations (FDEs). Our proof is probabilistic and uses martingale techniques.

1 Introduction

The present paper is devoted to a study of gradient estimates for solutions of partial differential equations (PDEs) of the form

$$\frac{\partial u}{\partial t} = \Delta u^m \tag{1}$$

[*]IRMAR, Université Rennes 1, 35042 Rennes Cedex, France. Email: ying.hu@univ-rennes1.fr.

[†]Mathematical Institute, University of Oxford, Oxford OX2 6GG, England. Email: qianz@maths.ox.ac.uk.

[‡]Mathematical Institute, University of Oxford, Oxford OX2 6GG, England. Email: zhangz@maths.ox.ac.uk.

where $m \in (0,\infty)$ is a given exponent. The problem is posed on $(0,\infty)\times\mathbb{R}^n$, and Δ is the Laplace operator with respect to space variables. When $m=1$, it is the celebrated heat equation, which is linear and parabolic, hence enjoying many nice properties. When $m \neq 1$, the story becomes quite different. Let us focus on non-negative solutions to the equation (1) with initial condition

$$u(0,x) = u_0(x),$$

the so-called Cauchy problem related to (1), and collect some results and discuss some phenomena that do not appear in the case of the heat equation.

There are two critical values of the exponent m, namely $m=1$ and $m=\frac{n-2}{n}$. By rewriting (1) as

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

one can see that the coefficient matrix of the second order derivative is $mu^{m-1}I_{n\times n}$. If $m<1$, this matrix takes the value of infinity where $u=0$, 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) 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) the fast diffusion equation (FDE). If $m>1$, the diffusion slows down wherever u is small, and in this case we call it the porous medium equation (PME). Notice that since we only consider non-negative solution u, from (2) it is clear that (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. The commonly used framework for PME is the $L^1(\mathbb{R}^n)$ space. By Theorem 9.12 and Proposition 9.13 in [32], if $u_0 \in L^1(\mathbb{R}^n)$, then there is a unique strong solution which is continuous on $(0,\infty)\times\mathbb{R}^n$. Moreover, if $u_0 \in L^1(\mathbb{R}^n)$ which is strictly positive and continuous, then the solution must be smooth. If we move beyond the scope of $L^1(\mathbb{R}^n)$ setting and impose a weaker growth condition on u_0 such as

$$\sup_{R\geq 1} R^{-\left(n+\frac{2}{m-1}\right)} \int_{|x|\leq R} |u_0(x)|\,dx < \infty, \tag{3}$$

then by [5] a unique solution in distribution sense exists on $(0,T(u_0))\times\mathbb{R}^n$, where $T(u_0)\in(0,\infty]$ depending on u_0. By Theorem 3.1 in [2], the previous growth condition is satisfied by any continuous non-negative solution. Therefore, condition (3) is optimal for the class of continuous non-negative solutions. If the initial data is allowed to be measure-valued, [13] gives a result which requires similar growth condition as (3). As for FDE, no requirement on the growth of initial data is needed. In fact, by Theorem 2.1 and 2.3 in [16], there exists a unique solution $u \in C\left([0,\infty); L^1_{loc}(\mathbb{R}^n)\right)$ 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 [6] there will be a classical solution in $\left[0,\bar{T}\right]\times\mathbb{R}^n$, and $\bar{T}$ is finite.

The degeneracy of parabolicity of PME leads to finite propagation of its solution, which is one of the special feature of PME. In particular, by Theorem 14.6 in [32], 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 [11], this set, or the so-called free boundary, is locally Hölder continuous on $(0,\infty)\times\mathbb{R}^n$. Moreover, in viewing it as a family of boundaries in $\mathbb{R}^n$ indexed by $t\in(0,\infty)$, those boundaries expands to infinity as $t\to\infty$ [11]. When the solution overflows the support of the initial data at some finite time $t_\star$, we see an improvement in the regularity of the free boundary. More precisely, Theorem 3 in [10] asserts that the free boundary is locally Lipschitz continuous on $(t_\star,\infty)\times\mathbb{R}^n$.

Although $m=1$ is a crucial value when talking about finite propagation and existence theories of equation (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 the phenomenon that arises only when $m<\frac{n-2}{n}$. On page 174 in [4], 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 [29], the same result holds for $u_0\in\mathcal{M}^{p_\star}$, where

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

According to Lemma 5.6 in [29], 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 [29].

Next, let us describe 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 [30]. Moreover, if $m>\frac{n-2}{n}$, then by Section 3.4 in [29], for any p and q such that $1\le p\le q\le\infty$,

$$\|u(t,\cdot)\|_q\le c(m,n,p,q)\,\|u_0\|_p^{\sigma}\,t^{-\alpha} \tag{4}$$

where

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

In particular, from this result we see that initial data in $L^1(\mathbb{R}^n)$ produce solutions $u(t,\cdot)\in L^\infty(\mathbb{R}^n)$ 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 [12] constructed a solution which is not bounded at any time while still having an initial value in $L^1(\mathbb{R}^n)$. Moreover, when $m<\frac{n-2}{n}$ and $n\ge3$, Theorem 5.14 in [29] shows that (4) holds for $q=1$ and $1<p<p_\star$. This means that $L^p(\mathbb{R}^n)$ data yield

solutions only in $L^1(\mathbb{R}^n)$, 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 [29], if $m < \frac{n-2}{n}$, $p \geq p_\star$ and $u_0 \in L^p(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$, then $u(t,\cdot)$ is locally bounded and smooth for any $t > 0$. Quantitatively, Theorem 2.1 in [9] gives a local upper bound for $u(t,\cdot)$ in terms of the $L^p(\mathbb{R}^n)$ $(p > p_\star)$ norm of u_0 over a larger local region.

As for positivity and smoothness, Theorem 3.1 in [2], Theorem 1.1 in [8] and Theorem 1.1 in [9] give local lower bounds for solutions to (1) in terms of local $L^1(\mathbb{R}^n)$ 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 [10] gave a necessary and sufficient condition for the positivity of $u(t,x)$ when $m > 1$ which read as

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

Besides positivity, if the solution is also locally bounded, then standard quasilinear theory [21] 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 [16]. 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 local boundedness of solutions, which can be seen from solution (0.2) in [9]. Moreover, when m is in this range, extinction in finite time may occur, which kills positivity of solutions in a global scale. As for PME, in general, solutions are only locally Hölder continuous. Theorem 4.1 in [11] 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 [32] 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), which is established in Theorem 7.3 in [30]. 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) that are often used in combination with the comparison principle. Let us only mention three of them here. The first one is source-type solution, also known as Zel'dovich-Kompanyeets-Barenblatt (ZKB) solution, which takes the form

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

with

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

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

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

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

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

By a scaling argument, it is shown in section 3.2.1 of [31] that any solution to (1) on $(0,\infty) \times \mathbb{R}^n$ with initial data $|x|^{-\lambda_3}$ must have the above form. Notice that in fact the ZKB solution also belongs to this type. The third type of solutions we would like to mention is a variance of the self similar solutions. 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) only for $t < T$, after which it blows up or vanishes. For this type of solutions, 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) is to derive various types of estimates for solutions. In [3], Aronson and Bénilan established the following gradient estimate for solutions to (1). If $m \in \left(\frac{n-2}{n}, \infty\right)$, u is a positive smooth solution to (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

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

This fundamental estimate is then employed in [5] for the study of existence theory, in [16] for $L^{\infty}_{loc}(\mathbb{R}^n)$ estimate for solutions, and in [10] for obtaining regularity results for the free boundary of solutions, to name but a few. Later in [23], a local version of 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 center $\mathcal{O}$ and radius $2R > 0$. Assume that u is a positive solution to (1) on $B(\mathcal{O}, 2R) \times [0,T]$ and the Ricci curvature $\mathrm{Ric} \geq -(n-1) K^2$ 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

$$\frac{|\nabla v|^2}{v} - \beta \frac{v_t}{v} \leq \alpha\beta^2 \left(\frac{1}{t} + C_2 K^2 v_{\max}^{2R,T}\right) + \alpha\beta^2 \frac{v_{\max}^{2R,T}}{R^2} C_1$$

where

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

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

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

where

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

Later in [18] 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 [34], S.T. Yau established a similar type of gradient bounds depending on derivatives of initial data for degenerate parabolic equations of the form

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

with $F \in C^2(0,\infty)$ and $F' > 0$. In particular, as explained in [24], Yau's result implies that for any function $c(t) \in C^1(0,\infty)$ satisfying

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

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 Aronson-Bénilan type, the Hamilton type estimate also plays an important role. It originates from Hamilton [15] where it was proved that a positive smooth solution u to the heat equation on a compact manifold without boundary and with Ric$\geq -k$, $k > 0$, we have

$$\frac{|\nabla u|^2}{u^2} \leq \left(\frac{1}{t} + 2k\right)\ln\frac{\|u\|_\infty}{u}. \tag{5}$$

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 [20], the same result holds for complete noncompact manifolds as well. Later a local version was obtained by Souplet and Zhang in [27]. As for PME and FDE, L. A. Caffarelli, J. L. Vázquez and N. I. Wolanski [10] discussed the case where 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 $T=T(u_0)>0$ and a constant $c=c(m,n)>0$ such that

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

for any $t>T$ and almost every $x\in\mathbb{R}^n$. Later X. Xu [33] derived a local result on a complete Riemannian manifold with Ric$\ge -k$ for some $k\ge 0$. For $m>1$, if there exists a constant $\delta\in\left(0,\frac{4}{n-1}\right]$ such that

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

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

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

where

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

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

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

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

where

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

This is a generalization of Li Ma et al. [24], 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 [36], it was proved that for $m\in\left(1,1+\frac{1}{\sqrt{2n}+1}\right)$, on $B(x_0,R)\times\left[t_0-\frac{T}{2},t_0\right]$

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

with

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

In X. Zhu [35], a gradient bound for $m\in\left(1-\frac{2}{n},1\right)$ was obtained. On $B(x_0,R)\times\left[t_0-\frac{T}{2},t_0\right]$

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

where

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

In this paper, let us denote by $\mathcal{A}_{x_0}^{R,\epsilon}$ the annulus $B(x_0, R+\epsilon) \setminus B(x_0, R)$. Set

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

$$h_- = m - \sqrt{1-(n-1)(m-1)^2},\ h_+ = m + \sqrt{1-(n-1)(m-1)^2}. \quad (6)$$

We derived several local and global gradient estimates as follows.

Theorem 1. *If u is a positive and bounded solution to (1) on $(0,t_1)\times B(x_0,R+\epsilon)$ with $m \in \left(1-\frac{1}{\sqrt{n-1}}, 1+\frac{1}{\sqrt{n-1}}\right)$, 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*

$$\begin{aligned} & u^{m+h-3}|\nabla u|^2(T,x) \\ \leq\ & \frac{\left\|I_{B(x_0,R+\epsilon)}u_0^h\right\|_\infty - u^h(T,x)}{hm\left(m-\rho\left|\frac{2m+h}{2h}\right|\right)T} + \frac{\left(\rho+\left|\frac{2m+h}{2}\right|\right)LM_2}{|h|\,\rho\left(m-\rho\left|\frac{2m+h}{2h}\right|\right)} \\ & + \left(|2m-h|\,l^{-3}+6l^{-1}\right)mM_1L^2\frac{T}{8} \end{aligned} \quad (7)$$

where

$$l = \frac{(h-h_-)(h_+-h)}{3\left|m-\frac{h}{2}\right|+3},\ M_1 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} u^{2m+h-2},\ M_2 = \sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}} u^{m+h-1}.$$

Comparing to the existing work, 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 above are using information of u during $[t_0-2T,t_0]$. Therefore, we have an extra term linear in T. In terms of the constraint on m, one can see that when the dimension $n=1$, there is no constraint. When $n>1$, comparing to all the existing results we are aware of, our bound is valid for a wider range of m when $m<1$. As for the case when $m>1$, our range of m is larger than that in [36]. If we look for a bound on the pressure variable, by taking $h=0$, the restriction on m becomes identical to the one in [35].

Corollary 2. *If u is a positive and bounded solution to (1) on $(0,t_1)\times B(x_0,R+\epsilon)$ with $m \in \left(1-\frac{2}{n},1\right)$, and the pressure variable $v = \frac{u^{m-1}}{m-1}$, then for any $\rho \in (0,m)$, and $(T,x) \in (0,t_1)\times B(x_0,R)$, we have*

$$\begin{aligned} \frac{|\nabla v|^2}{(m-1)v}(T,x) \leq\ & \frac{\left\|I_{B(x_0,R+\epsilon)}\log u_0\right\|_\infty - \log u(T,x)}{m(m-\rho)T} \\ & + \frac{LM_3}{\rho(m-\rho)} + \left(\frac{m}{2}l^{-3}+\frac{3}{2}l^{-1}\right)mM_1L^2\frac{T}{2} \end{aligned}$$

where

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

$$M_{3}=\sup_{[0,T]\times\mathcal{A}_{x_{0}}^{R,\epsilon}}\left(\rho\log u+\left(m\log u+1\right)^{2}\right)u^{m-1}$$

As for the global case, by letting $\epsilon=R$, then $R\to\infty$ and $\rho\to0$ in (7), we obtain a very neat gradient bound.

Corollary 3. *If u is a positive and bounded solution to (1) on $(0,\infty)\times\mathbb{R}^{n}$ and $m\in\left(1-\frac{1}{\sqrt{n-1}},1+\frac{1}{\sqrt{n-1}}\right)$, then for any $h\in[h_{-},h_{+}]$ and $(T,x)\in(0,\infty)\times\mathbb{R}^{n}$, it holds that*

$$u^{m+h-3}\left|\nabla u\right|^{2}(T,x)\leq\frac{\left\|u_{0}^{h}\right\|_{\infty}-u^{h}(T,x)}{Thm^{2}}.\tag{8}$$

First of all, 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, [7] derived a gradient bound which is independent of the initial data, but dependent on dimension n. Secondly, it naturally shows that when a function touches its maximum, its gradient vanishes. In this sense it is a tight gradient bound. Moreover, when $m=1$ and $h\to0^{+}$, our results recovers Hamilton's gradient bound for heat equation.

Corollary 4. *(Hamilton's gradient estimate) 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*

$$\frac{\left|\nabla u\right|^{2}}{u^{2}}(T,x)\leq\frac{1}{T}\log\frac{\left\|u_{0}\right\|_{\infty}}{u(T,x)}.\tag{9}$$

Our last result is an extension of Corollary 3, where the condition that

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

is essential.

We next consider the case where m and h satisfy the condition that

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

The previous inequality is equivalent to that $m\in\left(0,1-\frac{1}{\sqrt{n-1}}\right)\cup\left(1+\frac{1}{\sqrt{n-1}},\infty\right)$ and $h\in\mathbb{R}$, or $m\in\left[1-\frac{1}{\sqrt{n-1}},1+\frac{1}{\sqrt{n-1}}\right]$ but $h\in(-\infty,h_{-})\cup(h_{+},\infty)$. For such m and h, let us define

$$U_{\pm}=\frac{mh\pm\left|h\right|\sqrt{m^{2}+\left(h-h_{-}\right)\left(h-h_{+}\right)}}{\left(h-h_{-}\right)\left(h-h_{+}\right)}.\tag{10}$$

Theorem 5. *Let u be a positive and bounded solution to (1) on $(0,\infty)\times\mathbb{R}^n$. Suppose $h\in\mathbb{R}$ satisfying the condition that*

$$(h-h_-)(h-h_+)>0,$$

and

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

where

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

Then

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

This result, together with Corollary 3 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. While when m is outside of this range, the gradient bound will depend on the minimum of u as well. A similar phenomena has been shown in [33] , but only for $m>1$. Nevertheless, the meaning behind this phenomena is not clear to us yet, which is worthy of exploring.

We would like to mention that employing martingale theory to derive gradient estimate is not new, and there is a large number of papers devoted to the study of solutions of PDEs by using stochastic differential equations, for example [25], [1] and literature therein for a small sample. On the other hand, to the best of authors' knowledge, there are few papers dealing with the kind of nonlinear PDEs by using martingale methods.

The rest of the paper is organized as follows. In Section 2, we illustrate our main idea by establishing the gradient estimate for solution to the heat equation. In Section 3, we prove our local and global estimates.

2 Gradient estimate for solutions to the heat equation

To illustrate the main idea, let us consider the heat equation in this chapter. Assume u solves

$$\begin{aligned}\frac{\partial u}{\partial t}&=\Delta u \text{ on } (0,\infty)\times\mathbb{R}^n\\ u(x,0)&=u_0(x) \text{ on } \mathbb{R}^n.\end{aligned}\tag{11}$$

To avoid technical difficulties, we also assume u is smooth, bounded and has bounded derivatives with respect to space variables up to the second order. 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 so-called path integration, and the transition probability density function of a diffusion process is the kernel of a PDE. For a more comprehensive account of this area, we refer to the book [28] by Stroock and Varadhan. Here we also relates our PDE, (11), 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

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

where $\alpha \in \{1, \cdots, n\}$ and $W = \left(W^1, \cdots, W^n\right)$ 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 $|\nabla u|^2 (T, x)$, let us 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 composite $|\nabla u|^2$ with X is commonly seen. In terms of computation, this will lead to the disappearance of terms containing the time derivative when using Ito 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 [14]. First of all, let us decompose the semimartingale $\left(|\nabla u|^2 (T - t, X_t)\right)_{0 \le t \le T}$ into a sum of a local martingale and a process with 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 (11) yields

$$\frac{\partial^2 u}{\partial t \partial x^\alpha} = \Delta \frac{\partial u}{\partial x^\alpha}. \tag{12}$$

Since u is smooth by our assumption, we can apply Ito formula on $\frac{\partial u}{\partial x^\alpha} (T - t, X_t)$

and obtain

$$\begin{aligned} d\frac{\partial u}{\partial x^\alpha} &= -\frac{\partial^2 u}{\partial t \partial x^\alpha}dt + \Delta\frac{\partial u}{\partial x^\alpha}dt + \sqrt{2}\frac{\partial^2 u}{\partial x^\alpha \partial x^\beta}dW_t^\beta \\ &= \sqrt{2}\frac{\partial^2 u}{\partial x^\alpha \partial x^\beta}dW_t^\beta \end{aligned}$$

where the last equality results from (12). Then, again by Ito formula,

$$\begin{aligned} d\left|\nabla u\right|^2 &= 2\frac{\partial u}{\partial x^\alpha}d\frac{\partial u}{\partial x^\alpha} + \sum_\alpha d\left\langle \frac{\partial u}{\partial x^\alpha}\right\rangle \\ &= 2\sqrt{2}\frac{\partial u}{\partial x^\alpha}\frac{\partial^2 u}{\partial x^\alpha \partial x^\beta}dW_t^\beta + 2\sum_{\alpha,\beta=1}^n \left(\frac{\partial^2 u}{\partial x^\alpha \partial x^\beta}\right)^2 dt. \end{aligned} \tag{13}$$

This is the decomposition we are looking for. It can be seen that the finite variation term 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 $\frac{\partial^2 u}{\partial x^\alpha \partial x^\beta}$, $\forall \alpha, \beta \in \{1, \cdots, n\}$. This means that $\left|\nabla u\right|^2(T-t, X_t)$ is a submartingale, 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

$$\int_0^T E\left[\left|\nabla u\right|^2(T-t, X_t)\right] dt \geq T\left|\nabla u\right|^2(T, x). \tag{14}$$

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 $\left\|u\right\|_\infty$ should have the order $O\left(\frac{\left\|u\right\|_\infty^2}{T}\right)$. Therefore, let us consider the process $u^2(T-t, X_t)$. By Ito formula,

$$u^2(0, X_T) - u^2(T, X_0) = 2\sqrt{2}\int_0^T u\frac{\partial u}{\partial x^\alpha}dW_t^\alpha + 2\int_0^T \left|\nabla u\right|^2 dt. \tag{15}$$

Since u and $\frac{\partial u}{\partial x^\alpha}$, $\alpha \in \{1, \cdots, n\}$, are bounded by our assumption, $\int_0^\cdot u\frac{\partial u}{\partial x^\alpha}dW_t^\alpha$ is a true martingale. Hence $E\left[\int_0^T u\frac{\partial u}{\partial x^\alpha}dW_t^\alpha\right] = 0$. It then follows from (15) that

$$E\left[u^2(0, X_T)\right] - u^2(T, x) = 2E\left[\int_0^T \left|\nabla u\right|^2 dt\right]. \tag{16}$$

By Fubini's theorem, $E\left[\int_0^T \left|\nabla u\right|^2 dt\right] = \int_0^T E\left[\left|\nabla u\right|^2\right] dt$. Therefore, (14) and (16) together yields

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

Since the law of X_T is absolutely continuous with respect to Lebesgue measure, $u^2(0, X_T) \leq \left\|u_0\right\|_\infty^2$ $\mathbb{P}-$ almost surely, where $\left\|\cdot\right\|_\infty$ denotes the essential supremum. Hence we have the following result.

Theorem 6. *If u solves the Cauchy problem (11) 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*

$$|\nabla u|^2 (T,x) \leq \frac{\|u_0\|_\infty^2 - u^2 (T,x)}{2T} \leq \frac{\|u_0\|_\infty^2}{T}. \tag{17}$$

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

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

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

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

Summing over the index i yields

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

where the last inequality results from (17). Then we can minimize the right hand side by choosing $\epsilon = \frac{T}{2}$, hence obtaining

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

Remark. An approach based on the link between Backward Stochastic Differential Equation (BSDE) and PDE to establish gradient estimates for positive solutions to the heat equation of elliptic or subelliptic operators on Euclidean spaces or on Riemannian manifolds is developed in [17].

3 Gradient estimate for positive solutions to PME and FDE

3.1 Local gradient estimate

Now we move to the equation (1). In this section we study the problem on a local scale. Denote by $B(x_0, R+\epsilon)$ a closed ball in $\mathbb{R}^n$ with center x_0 and radius $R+\epsilon$, where R, $\epsilon > 0$. Let u be a positive and bounded solution to (1) on $[0,t_1] \times B(x_0, R+\epsilon)$. The positivity of u ensures that no degeneracy of parabolicity would happen in $[0,t_1] \times B(x_0, R+\epsilon)$. Hence we can

use theory about non-degenerate quasilinear parabolic PDE to obtain that u is smooth in $[0,t_1]\times B(x_0,R+\epsilon)$. Note that at this moment u is defined only on $[0,t_1]\times B(x_0,R+\epsilon)$. This brings difficulty 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

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

Note that as u is strictly positive and smooth on the compact set $[0,t_1]\times B(x_0,R+\epsilon)$, such $\tilde{u}$ exists. It is worthwhile to point out here that the behavior of the extended function $\tilde{u}$ outside $[0,t_1]\times B(x_0,R+\epsilon)$ 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

$$f=\frac{\tilde{u}^h}{h} \tag{18}$$

for some $h\in\mathbb{R}\setminus\{0\}$. This is a generalization of the transform $v=\frac{u^{m-1}}{m-1}$, which repeatedly appears in literature concerning PME and FDE, such as [3], [23], [10], [33] and [35]. From (1) one can derive that the so-called pressure variable v satisfies

$$\frac{\partial v}{\partial t}=m(m-1)v\Delta v+m|\nabla v|^2. \tag{19}$$

We can see that the exponent m in (1) comes down into coefficients in (19), and both terms on the right hand side of (19) are quadratic, which facilitates many computations. However, this feature is not crucial to our method. Hence we attempt to generalize this transform with (18). 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 [22] to introduce a cut-off function $\phi\in C^2(\mathbb{R}^n)$ satisfying

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

$$|\Delta\phi|\le L\phi^{\frac{1}{2}} \tag{20}$$

and

$$|\nabla\phi|^2\le L\phi^{\frac{3}{2}}, \tag{21}$$

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

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

and

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

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

3.1.1 From PDE to SDE

Just as the case of heat equation, let us begin by fixing a point $(T,x)\in[0,t_1]\times B\left(x_0,R\right)$. From (1) and the definition of f, we have on $[0,t_1]\times\mathbb{R}^n$,

$$\begin{aligned}\frac{\partial\phi f}{\partial t} &= m\left(hf\right)^{\frac{m-1}{h}}\phi\Delta f + m\left(m-h\right)\left(hf\right)^{\frac{m-h-1}{h}}\phi\left|\nabla f\right|^2\\ &= m\left(hf\right)^{\frac{m-1}{h}}\Delta\left(\phi f\right) - m\left(hf\right)^{\frac{m-1}{h}}f\Delta\phi\\ &\quad -2m\left(hf\right)^{\frac{m-1}{h}}\nabla f\cdot\nabla\phi + m\left(m-h\right)\left(hf\right)^{\frac{m-h-1}{h}}\phi\left|\nabla f\right|^2. \quad (22)\end{aligned}$$

Then let us link PDE (22) with the diffusion process $X=\left(X_t\right)_{0\le t\le T}$, whose generator $\mathcal{L}$ is given by

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

$\mathcal{L}$ corresponds to the principle part of the differential operator in (22). Note that by the definition of f, $\left(hf\right)^{\frac{m-1}{h}}\ge 0$ since $u\ge 0$. By [28] the way to obtain X is to solve the stochastic differential equation (SDE)

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

for $t\in[0,T]$, where $W=\left(W^1,\cdots,W^n\right)$ 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, $\left(hf\right)^{\frac{m-1}{2h}}$ is bounded, smooth and has bounded derivatives. Therefore, by [19] 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 $f\left(T-t,X_t\right)$, $\phi\left(X_t\right)$, $\left|\nabla f\right|^2\left(T-t,X_t\right)$ and so on. In fact, we will omit the specification of variables in functions, as all the functions below are compositions on $\left(T-t,X_t\right)$. t will always take values in $[0,T]$, which is exactly the time interval where X lives. One can see that these processes are all semimartingales, waiting for us to decompose and then releasing information.

But unlike the heat equation, (22) is not linear, a consequence of the nonlinearity of (1). Therefore, we should be very careful in choosing semimartingales for decomposition. To begin with, let us consider

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

By Ito formula,

$$\begin{aligned} dY_t &= \sqrt{2m}\,(hf)^{\frac{m-1}{2h}} \frac{\partial(f\phi)}{\partial x^\alpha} dW_t^\alpha - m(m-h)(hf)^{\frac{m-h-1}{h}} \phi |\nabla f|^2 dt \\ &\quad + \left(\Delta\phi m (hf)^{\frac{m-1}{h}} f + 2m (hf)^{\frac{m-1}{h}} \frac{\partial f}{\partial x^\alpha} \frac{\partial \phi}{\partial x^\alpha} \right) dt. \end{aligned} \tag{23}$$

It is readily seen that

$$Y_T = \phi f(0, X_T) = \phi \frac{u_0^h}{h}(X_T). \tag{24}$$

Observe that in the finite variation part of (23), the term with highest degree in ∇f is

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

This means that it can be controlled in terms of $\|u_0\|_\infty$ with the help of (23). Therefore, it is this term that worth investigation. Define

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

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

$$
\begin{aligned}
dH_t &= d\left((hf)^{\frac{m-h-1}{h}}\phi\sum_{\alpha}\left|\frac{\partial f}{\partial x^{\alpha}}\right|^2\right)\\
&= 2\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\phi\frac{\partial^2 f}{\partial x^{\alpha}\partial x^{\beta}}\frac{\partial f}{\partial x^{\alpha}}dW_t^{\beta}\\
&\quad+\sqrt{2m}\,(m-h-1)\,(hf)^{\frac{3m-4h-3}{2h}}\phi\left|\nabla f\right|^2\frac{\partial f}{\partial x^{\beta}}dW_t^{\beta}\\
&\quad+\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\frac{\partial\phi}{\partial x^{\beta}}\left|\nabla f\right|^2dW_t^{\beta}\\
&\quad-m\,(2m-h+1)\,(m-h-1)\,(hf)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4dt\\
&\quad+2m\,(m-h-1)\,(hf)^{\frac{2m-2h-2}{h}}\frac{\partial\phi}{\partial x^{\alpha}}\frac{\partial f}{\partial x^{\alpha}}\left|\nabla f\right|^2dt\\
&\quad+m\,(hf)^{\frac{2m-h-2}{h}}\Delta\phi\left|\nabla f\right|^2dt\\
&\quad-2m\,(m-1)\,(hf)^{\frac{2m-2-2h}{h}}\phi\Delta f\left|\nabla f\right|^2dt\\
&\quad-4m\,(hf)^{\frac{2m-2h-2}{h}}\phi\frac{\partial f}{\partial x^{\beta}}\frac{\partial^2 f}{\partial x^{\beta}\partial x^{\alpha}}\frac{\partial f}{\partial x^{\alpha}}dt\\
&\quad+4m\,(hf)^{\frac{2m-2-h}{h}}\frac{\partial\phi}{\partial x^{\beta}}\frac{\partial^2 f}{\partial x^{\alpha}\partial x^{\beta}}\frac{\partial f}{\partial x^{\alpha}}dt\\
&\quad+2m\,(hf)^{\frac{2m-h-2}{h}}\phi\sum_{\alpha,\beta}\left(\frac{\partial^2 f}{\partial x^{\alpha}\partial x^{\beta}}\right)^2dt. \qquad (25)
\end{aligned}
$$

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 $(\Omega,\mathcal{F})$ depending on a parameter $\lambda\in\mathbb{R}$. Define a measure $\mathbb{Q}$ on $(\Omega,\mathcal{F})$ by

$$
\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t}=\exp\left(\int_0^t\sqrt{2m}\lambda\,(hf)^{\frac{m-2h-1}{2h}}\frac{\partial f}{\partial x^{\beta}}dW^{\beta}-m\lambda^2\int_0^t(hf)^{\frac{m-2h-1}{h}}\left|\nabla f\right|^2dt\right). \qquad (26)
$$

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 [26] is met. So the right hand side of (26) is a true martingale under $\mathbb{P}$, which ensures that measure $\mathbb{Q}$ is well defined. According to Girsanov's theorem [19], 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 (25), we easily get the decomposition for H under measure $\mathbb{Q}$, which is

$$\begin{aligned}
dH_t &= 2\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\,\phi\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}d\tilde{W}_t^\beta \\
&\quad +\sqrt{2m}\,(m-h-1)\,(hf)^{\frac{3m-4h-3}{2h}}\,\phi\,|\nabla f|^2\,\frac{\partial f}{\partial x^\beta}d\tilde{W}_t^\beta \\
&\quad +\sqrt{2m}\,(hf)^{\frac{3m-2h-3}{2h}}\,\frac{\partial \phi}{\partial x^\beta}\,|\nabla f|^2\,d\tilde{W}_t^\beta \\
&\quad -m\,(2m-h+1-2\lambda)\,(m-h-1)\,(hf)^{\frac{2m-3h-2}{h}}\,\phi\,|\nabla f|^4\,dt \\
&\quad +2m\,(m-h-1+\lambda)\,(hf)^{\frac{2m-2h-2}{h}}\,\frac{\partial \phi}{\partial x^\alpha}\frac{\partial f}{\partial x^\alpha}\,|\nabla f|^2\,dt \\
&\quad +m\,(hf)^{\frac{2m-h-2}{h}}\,\Delta\phi\,|\nabla f|^2\,dt \\
&\quad -2m\,(m-1)\,(hf)^{\frac{2m-2-2h}{h}}\,\phi\Delta f\,|\nabla f|^2\,dt \\
&\quad +4m\,(\lambda-1)\,(hf)^{\frac{2m-2h-2}{h}}\,\phi\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}\frac{\partial f}{\partial x^\beta}dt \\
&\quad +4m\,(hf)^{\frac{2m-2-h}{h}}\,\frac{\partial \phi}{\partial x^\beta}\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\frac{\partial f}{\partial x^\alpha}dt \\
&\quad +2m\,(hf)^{\frac{2m-h-2}{h}}\,\phi\sum_{\alpha,\beta}\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta}\right)^2 dt. \qquad (27)
\end{aligned}$$

While from (23), it is easy to see that under measure $\mathbb{Q}$,

$$\begin{aligned}
&dY_t \\
=\ &\sqrt{2m}\,(hf)^{\frac{m-1}{2h}}\,\frac{\partial\,(f\phi)}{\partial x^\alpha}d\tilde{W}_t^\alpha \\
&+m\,(2\lambda-m+h)\,(hf)^{\frac{m-h-1}{h}}\,\phi\,|\nabla f|^2\,dt \\
&+\Delta\phi m h^{-1}\,(hf)^{\frac{m-1+h}{h}}\,dt \\
&+2m\left(1+\lambda h^{-1}\right)(hf)^{\frac{m-1}{h}}\,\frac{\partial f}{\partial x^\alpha}\frac{\partial \phi}{\partial x^\alpha}dt. \qquad (28)
\end{aligned}$$

Now we have (28) and (27) at hand, which are the fundamental decompositions we are looking for.

3.1.3 Gradient bound for solution u

In contrast to (13) which appears in the case of global estimate for the heat equation, (27) is more complicated, thanks to the nonlinearity of PDE (1) and the introduction of cut-off function ϕ. To make use of it, recall that we aim to estimate the first order derivatives of f. Therefore, it is reasonable to get rid off the second order derivatives of f appearing in the decomposition (27). This is done in the following lemma. For simplicity, denote by $\mathcal{A}_{x_0}^{R,\epsilon}$ the annulus $B\,(x_0, R+\epsilon)\setminus B\,(x_0, R)$.

Lemma 7. *H satisfies*

$$\begin{aligned}
& dH_t \\
\geq\ & A_t d\tilde{W}_t \\
& -\frac{1}{2}mn\left(m-1\right)^2\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4 dt \\
& -m\left(\left(2m-h+1-2\lambda\right)\left(m-h-1\right)+2\left(\lambda-1\right)\left(\lambda-m\right)\right)\left(hf\right)^{\frac{2m-3h-2}{h}}\phi\left|\nabla f\right|^4 dt \\
& +2m\left(2m-h-\lambda\right)\left(hf\right)^{\frac{2m-2h-2}{h}}\frac{\partial\phi}{\partial x^\alpha}\frac{\partial f}{\partial x^\alpha}\left|\nabla f\right|^2 dt \\
& +m\left(hf\right)^{\frac{2m-h-2}{h}}\Delta\phi\left|\nabla f\right|^2 dt \\
& -2m\left(hf\right)^{\frac{2m-2-h}{h}}\phi^{-1}I_{\mathcal{A}_{x_0}^{R,\epsilon}}\left|\nabla\phi\right|^2\left|\nabla f\right|^2 dt
\end{aligned} \tag{29}$$

where

$$\begin{aligned}
A &= 2\sqrt{2m}\left(hf\right)^{\frac{3m-2h-3}{2h}}\phi\frac{\partial^2 f}{\partial x^\alpha\partial x^\beta}\frac{\partial f}{\partial x^\alpha} \\
&\quad +\sqrt{2m}\left(m-h-1\right)\left(hf\right)^{\frac{3m-4h-3}{2h}}\phi\left|\nabla f\right|^2\frac{\partial f}{\partial x^\beta} \\
&\quad +\sqrt{2m}\left(hf\right)^{\frac{3m-2h-3}{2h}}\frac{\partial\phi}{\partial x^\beta}\left|\nabla f\right|^2.
\end{aligned} \tag{30}$$

Proof. From (27), let us write

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

where

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

and by Cauchy-Schwartz inequality,

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

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

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 (29). □

To deal with the finite variation part in (29), let us assume that

$$\frac{1}{2}n\left(m-1\right)^{2}+\left(2m-h+1-2\lambda\right)\left(m-h-1\right)+2\left(\lambda-1\right)\left(\lambda-m\right)<0. \quad (31)$$

This allows us to get the following estimate.

Lemma 8. *Under assumption* (31)*, we have*

$$E^{\mathbb{Q}}\left[\int_{0}^{T}H_{s}ds\right]\geq H_{0}T-\left(\frac{1}{2}\left|2m-h-\lambda\right|l^{-3}+\frac{3}{2}l^{-1}\right)mM_{1}L^{2}\frac{T^{2}}{2} \quad (32)$$

where

$$l=\frac{\frac{1}{2}n\left(m-1\right)^{2}+\left(2m-h+1-2\lambda\right)\left(m-h-1\right)+2\left(\lambda-1\right)\left(\lambda-m\right)}{-\frac{3}{2}\left|2m-h-\lambda\right|-\frac{3}{2}} \quad (33)$$

and

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

Proof. By 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

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

Plugging them into (29) in the above lemma yields

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

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

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

By estimates for the cut-off function ϕ in (20) and (21),

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

Now, by assumption (31), 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 $t\in[0,T]$,

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

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

$$H_s-H_0\geq\int_0^s A_t d\tilde{W}_t-\left(\frac{1}{2}\left|2m-h-\lambda\right|l^{-3}+\frac{3}{2}l^{-1}\right)mM_1L^2s \tag{34}$$

for any $s\in[0,T]$. According to our assumption on u, A, defined by (30), 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 (34) and obtain

$$E^{\mathbb{Q}}\left[H_s\right]\geq H_0-\left(\frac{1}{2}\left|2m-h-\lambda\right|l^{-3}+\frac{3}{2}l^{-1}\right)mM_1L^2s. \tag{35}$$

Integrating both sides with respect to s on $[0,T]$ yields (32). □

On the other hand, by using (23), we have an upper bound on the left hand side of (32).

Lemma 9. *Assume*

$$2\lambda - m + h > 0. \tag{36}$$

Then for any $\rho > 0$ such that

$$2\lambda - m + h - \rho \left|1 + \lambda h^{-1}\right| > 0, \tag{37}$$

we have

$$E^{\mathbb{Q}}\left[\int_0^T H_t dt\right] \leq \frac{E^{\mathbb{Q}}\left[Y_T\right] - Y_0 + \left(\left|h^{-1}\right| + \rho^{-1}\left|1 + \lambda h^{-1}\right|\right) mLM_2 T}{m\left(2\lambda - m + h - \rho\left|1 + \lambda h^{-1}\right|\right)} \tag{38}$$

with

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

Proof. Recall (28), which is the fundamental decomposition of Y under measure $\mathbb{Q}$. Then by property (20) of cut-off function ϕ and Cauchy-Schwartz inequality, for any $\rho > 0$,

$$\begin{aligned}
& dY_t \\
\geq\ & \sqrt{2m}\,(hf)^{\frac{m-1}{2h}} \frac{\partial(f\phi)}{\partial x^\alpha} d\tilde{W}_t^\alpha \\
& + m\left(2\lambda - m + h\right)(hf)^{\frac{m-h-1}{h}} \phi\left|\nabla f\right|^2 dt \\
& - L\phi^{\frac{1}{2}} I_{\mathcal{A}_{x_0}^{R,\epsilon}} m\left|h^{-1}\right|(hf)^{\frac{m-1+h}{h}} dt \\
& - \rho m\left|1 + \lambda h^{-1}\right|(hf)^{\frac{m-1-h}{h}} \phi\left|\nabla f\right|^2 dt \\
& - \rho^{-1} m\left|1 + \lambda h^{-1}\right|(hf)^{\frac{m-1+h}{h}} I_{\mathcal{A}_{x_0}^{R,\epsilon}} \phi^{-1}\left|\nabla\phi\right|^2 dt \\
\geq\ & \sqrt{2m}\,(hf)^{\frac{m-1}{2h}} \frac{\partial(f\phi)}{\partial x^\alpha} d\tilde{W}_t^\alpha \\
& + m\left(2\lambda - m + h - \rho\left|1 + \lambda h^{-1}\right|\right) H_t dt \\
& - \left(\left|h^{-1}\right| + \rho^{-1}\left|1 + \lambda h^{-1}\right|\right) mLM_2 dt
\end{aligned} \tag{39}$$

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 $\mathbb{Q}$ martingale. Hence

$$\begin{aligned}
& E^{\mathbb{Q}}\left[Y_{T-t_1}\right] - Y_0 \\
\geq\ & m\left(2\lambda - m + h - \rho\left|1 + \lambda h^{-1}\right|\right) E^{\mathbb{Q}}\left[\int_0^{T-t_1} H_t dt\right] \\
& - \left(\left|h^{-1}\right| + \rho^{-1}\left|1 + \lambda h^{-1}\right|\right) mLM_2\left(T - t_1\right).
\end{aligned}$$

Dividing both sides by $m\left(2\lambda - m + h - \rho\left|1 + \lambda h^{-1}\right|\right)$ yields (38). ∎

Now, based on (32) and (38), we are ready to obtain a gradient bound for u.

Proposition 10. *Let the assumptions in Lemma 8 and Lemma 9 be satisfied, then for any* $(T,x)\in(0,t_1)\times B(x_0,R)$,

$$\begin{aligned} & u^{m+h-3}\left|\nabla u\right|^2(T,x) \\ \leq\ & \frac{E^{\mathbb{Q}}\left[\phi u_0^h(X_T)\right]-u^h(T,x)}{hm\left(2\lambda-m+h-\rho\left|1+\lambda h^{-1}\right|\right)T}+\frac{\left(\left|h^{-1}\right|+\rho^{-1}\left|1+\lambda h^{-1}\right|\right)LM_2}{2\lambda-m+h-\rho\left|1+\lambda h^{-1}\right|} \\ & +\left(\frac{1}{2}\left|2m-h-\lambda\right|l^{-3}+\frac{3}{2}l^{-1}\right)mM_1L^2\frac{T}{2}. \end{aligned} \tag{40}$$

Proof. Combining (32) with (38), we have

$$\begin{aligned} & H_0 \\ \leq\ & \frac{E^{\mathbb{Q}}\left[Y_T\right]-Y_0+\left(\left|h^{-1}\right|+\rho^{-1}\left|1+\lambda h^{-1}\right|\right)mLM_2T}{m\left(2\lambda-m+h-\rho\left|1+\lambda h^{-1}\right|\right)T} \\ & +\left(\frac{1}{2}\left|2m-h-\lambda\right|l^{-3}+\frac{3}{2}l^{-1}\right)mM_1L^2\frac{T}{2}. \end{aligned}$$

Then we arrive at (40) by noting that

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

$$Y_0=\phi f(T,x)=\phi\frac{u^h}{h}(T,x)$$

and

$$E^{\mathbb{Q}}\left[Y_T\right]=E^{\mathbb{Q}}\left[\phi\frac{u_0^h}{h}(X_T)\right].$$

□

Let us revise the assumptions we have made, that is, (31) and (36). They impose a restriction on the choice of m. Let us look at (31) first. It is equivalent to

$$2\lambda^2-2(2m-h)\lambda+\frac{1}{2}n(m-1)^2+(2m-h+1)(m-h-1)+2m<0 \tag{41}$$

To let the left hand side attain the minimum, we take $\lambda=\frac{2m-h}{2}$. Then (41) is reduced to

$$h^2-2mh+n(m-1)^2+2(m-1)<0. \tag{42}$$

Notice that after setting $\lambda=\frac{2m-h}{2}$, condition (36) is automatically satisfied. Moreover, (42) holds if and only if $m\in\left(1-\frac{1}{\sqrt{n-1}},1+\frac{1}{\sqrt{n-1}}\right)$, and

$$h\in\left(m-\sqrt{1-(n-1)(m-1)^2},m+\sqrt{1-(n-1)(m-1)^2}\right).$$

Then we arrive at Theorem 1 by substituting $\lambda=\frac{2m-h}{2}$ into (37) and (33), and noting that $E^{\mathbb{Q}}\left[\phi u_0^h(X_T)\right]\leq\left\|I_{B(x_0,R+\epsilon)}u_0^h\right\|_\infty$.

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 $\frac{|\nabla v|^2}{v}$, we are tempted to work out a bound on it with our method. Notice that $\frac{|\nabla v|^2}{v} = (m-1)\, u^{m-3} |\nabla u|^2$, which corresponds to the LHS of (7) in Theorem 1 when $h = 0$. However, the RHS of (7) will explode when $h \to 0$. Therefore, we can not use the result in Theorem 1 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

$$Y_t - \frac{\phi}{h}(X_t) = \phi \frac{u^h - 1}{h} \to \phi \log u,$$

which is finite.

Lemma 11. *Assume*

$$2\lambda - m + h > 0. \tag{43}$$

Then for any $\rho > 0$ such that

$$2\lambda - m + h - \rho > 0, \tag{44}$$

we have

$$E^{\mathbb{Q}}\left[\int_0^T H_t dt\right] \leq \frac{E^{\mathbb{Q}}\left[Y_T - \frac{\phi}{h}(X_T)\right] - Y_0 + \frac{\phi}{h}(x) + Lm\rho^{-1} M_3 T}{m(2\lambda - m + h - \rho)} \tag{45}$$

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 (28),

$$\begin{aligned}
& d\left(Y_t - \frac{\phi}{h}(X_t)\right) \\
= \; & \sqrt{2m}\,(hf)^{\frac{m-1}{2h}} \left(\phi \frac{\partial f}{\partial x^\alpha} + \left(f - \frac{1}{h}\right)\frac{\partial \phi}{\partial x^\alpha}\right) d\tilde{W}_t^\alpha \\
& + m(2\lambda - m + h)(hf)^{\frac{m-h-1}{h}} \phi |\nabla f|^2 dt \\
& + \Delta\phi m \left(f - \frac{1}{h}\right)(hf)^{\frac{m-1}{h}} dt \\
& + 2m\left(\lambda\left(f - \frac{1}{h}\right) + hf\right)(hf)^{\frac{m-h-1}{h}} \frac{\partial f}{\partial x^\alpha}\frac{\partial \phi}{\partial x^\alpha} dt.
\end{aligned} \tag{46}$$

Then by Cauchy-Schwartz inequality, for any $\rho > 0$,

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

Let

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

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

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

which completes the proof. □

By combining (45) with (32) in the way as in the proof of Theorem 1 and also taking $\lambda = \frac{2m-h}{2}$, we are able to get the following gradient bound.

Theorem 12. *If u is a positive and bounded solution to (1) on $(0,t_1)\times B(x_0, R+\epsilon)$ with $m \in \left(1 - \frac{1}{\sqrt{n-1}}, 1 + \frac{1}{\sqrt{n-1}}\right)$, then for any*

$$h \in \left(m - \sqrt{1 - (n-1)(m-1)^2}, m + \sqrt{1 - (n-1)(m-1)^2}\right),$$

$\rho \in (0,m)$, and $(T,x) \in (0,t_1) \times B(x_0,R)$, we have

$$\begin{aligned} & u^{m+h-3}\left|\nabla u\right|^2(T,x) \\ \leq\ & \frac{E^{\mathbb{Q}}\left[\phi\left(u_0^h-1\right)(X_T)\right]-\left(u^h(T,x)-1\right)}{hm(m-\rho)T} \\ & +\frac{LM_3}{\rho(m-\rho)}+\left(\frac{1}{4}\left|2m-h\right|l^{-3}+\frac{3}{2}l^{-1}\right)mM_1L^2\frac{T}{2} \end{aligned}$$

where

$$l=\frac{-\frac{1}{2}n(m-1)^2+(h-1)(m-1)-\frac{1}{2}h^2+h}{\frac{3}{4}\left|2m-h\right|+\frac{3}{2}},\ M_1=\sup_{[0,T]\times\mathcal{A}_{x_0}^{R,\epsilon}}u^{2m+h-2},$$

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

and L is a constant depending on the cut-off function ϕ by (20) and (21).

By letting $h \to 0$, we immediately get Corollary 2.

3.2 Global gradient estimate

In this section, we consider a positive and bounded solution u to (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.

3.2.1 From local bound to global bound

Since u is positive and bounded on $(0,\infty)\times\mathbb{R}^n$, the local result in Theorem 1 holds for any t_1, R and ϵ. By substituting $\epsilon = R$ into (7), and then taking $R\to\infty$ on both sides, we obtain

$$u^{m+h-3}\left|\nabla u\right|^2(T,x) \leq \frac{E^{\mathbb{Q}}\left[u_0^h(X_T)\right]-u^h(T,x)}{Thm\left(m-\rho\left|\frac{2m+h}{2h}\right|\right)}$$

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

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

By letting $\rho = 0$, we have

$$u^{m+h-3}\left|\nabla u\right|^2(T,x) \leq \frac{E^{\mathbb{Q}}\left[u_0^h(X_T)\right]-u^h(T,x)}{Thm^2}. \tag{47}$$

Note that in the local case, due to the existence of l^{-1} in (7), h is not allowed to touch the boundary of the open interval

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

But now we can take limit of h to

$$m-\sqrt{1-(n-1)(m-1)^2} \text{ or } m+\sqrt{1-(n-1)(m-1)^2}$$

on both sides of (47). Since

$$E^{\mathbb{Q}}\left[u_0^h(X_T)\right] = E^{\mathbb{Q}}\left[u_0^h(X_T) I_{\{u_0(X_T)\geq 1\}}\right] + E^{\mathbb{Q}}\left[u_0^h(X_T) I_{\{u_0(X_T)<1\}}\right],$$

we can apply monotone convergence theorem and bounded convergence theorem on these two terms respectively. Therefore, (47) holds also for

$$h = m-\sqrt{1-(n-1)(m-1)^2} \text{ and } h = m+\sqrt{1-(n-1)(m-1)^2}.$$

Hence we obtained Corollary 3.

Note that when $m=1$ in Corollary 3, h is allowed to approach to 0^+. Since $u_0>0$, by Jensen's inequality, it holds that

$$\begin{aligned}
\lim_{h\to 0^+}\frac{E^{\mathbb{Q}}\left[u_0^h(X_T)\right]-u^h(T,x)}{Th} &= \lim_{h\to 0^+}\frac{E^{\mathbb{Q}}\left[u_0^h(X_T)-1\right]}{Th} - \lim_{h\to 0}\frac{u^h(T,x)-1}{Th} \\
&= \frac{E^{\mathbb{Q}}\left[\log u_0(X_T)\right]-\log u(T,x)}{T} \\
&= \frac{E^{\mathbb{Q}}\left[\log \frac{u_0(X_T)}{u(T,x)}\right]}{T} \\
&\leq \frac{1}{T}\log\frac{E^{\mathbb{Q}}\left[u_0(X_T)\right]}{u(T,x)} \\
&\leq \frac{1}{T}\log\frac{\|u_0\|_\infty}{u(T,x)}
\end{aligned}$$

Therefore, by letting $h\to 0$ in (8), we have Corollary 4.

3.2.2 Negative finite variation part

Now we proceed from fundamental decompositions in a different way in order to get a global gradient bound valid for a wider class of m. From (29) and (28), we have the global version of two fundamental decompositions.

$$\begin{aligned}
& dH_t \\
\geq\ & A_t d\tilde{W}_t \\
& -\frac{1}{2}mn(m-1)^2(hf)^{\frac{2m-3h-2}{h}}|\nabla f|^4\,dt \qquad (48) \\
& -m\left((2m-h+1-2\lambda)(m-h-1)+2(\lambda-1)(\lambda-m)\right)(hf)^{\frac{2m-3h-2}{h}}|\nabla f|^4\,dt.
\end{aligned}$$

$$dY_t = \sqrt{2m}\,(hf)^{\frac{m-1}{2h}}\,\frac{\partial f}{\partial x^\alpha}d\tilde{W}_t^\alpha + m\,(2\lambda - m + h)\,(hf)^{\frac{m-h-1}{h}}\,|\nabla f|^2\,dt. \quad (49)$$

For simplicity, set

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

Then (49) becomes

$$\begin{aligned} dH_t &\geq A_t d\tilde{W}_t - L_1\,(hf)^{\frac{2m-3h-2}{h}}\,|\nabla f|^4\,dt \\ &= A_t d\tilde{W}_t - L_1 h^{-1} Y_t^{-1} H_t^2 dt. \end{aligned} \quad (50)$$

Previously, in order to obtain the submartingale property of H, we always assume $L_1 \leq 0$, resulting 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 13. *$H_\cdot e^{\int_0^\cdot L_1 h^{-1} Y_r^{-1} H_r dr}$ is a submartingale. Moreover,*

$$m\,(2\lambda - m + h)\,E^{\mathbb{Q}}\left[\int_s^T H_t \middle| \mathcal{F}_s\right] = E^{\mathbb{Q}}\left[Y_T \middle| \mathcal{F}_s\right] - Y_s. \quad (51)$$

Now let us define

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

Lemma 14. *Let u be a positive and bounded solution to (1) on $(0,\infty)\times\mathbb{R}^n$. Assume $L_1 \geq 0$ and $2\lambda - m + h > 0$.* Then

$$g\,(s) \leq \frac{\left\|E^{\mathbb{Q}}\left[\frac{Y_T}{Y_s}\middle|\mathcal{F}_s\right] - 1\right\|_\infty}{m\,(2\lambda - m + h)\int_s^T e^{-\int_s^t L_1 h^{-1} g(u)du} dt}. \quad (52)$$

Proof. Since $\left(e^{\int_0^t L_1 h^{-1} Y_r^{-1} H_r dr} H_t\right)_{t\in[0,T]}$ is a submartingale by Lemma 13, for any $0 \leq s < t \leq T$

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

which implies

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

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 have

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

Together with (51), it follows that

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

which leads to

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

□

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

$$g\left(0\right) \leq \frac{E^{\mathbb{Q}}\left[\frac{Y_T}{Y_0}\right] - 1}{m\left(2\lambda - m + h\right) G\left(0\right)}, \tag{53}$$

$G\left(T\right) = 0$. Moreover, for any $s \in (0, T)$

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

As a consequence,

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

This, together with (53) yields

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

as long as

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

This means

$$u^{m+h-3}\left|\nabla u\right|^2(T,x) \leq \frac{E^{\mathbb{Q}}\left[u_0^h\left(X_T\right)\right] - u^h(T,x)}{\int_0^T mh\left(2\lambda - m + h\right) - L_1 \left\| E^{\mathbb{Q}}\left[\frac{u_0^h(X_T)}{u^h(T-s,X_s)}\middle| \mathcal{F}_s\right] - 1\right\|_\infty ds}.$$

Define

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

It then follows that when

$$mh\left(2\lambda - m + h\right) - L_1\left(\frac{\left\|u_0^h\right\|_\infty}{u_{\min}^{h,T}} - 1\right) > 0, \tag{54}$$

we have

$$u^{m+h-3}\left|\nabla u\right|^2(T,x) \leq \frac{\left\|u_0^h\right\|_\infty - u^h(T,x)}{T\left(mh\left(2\lambda - m + h\right) - L_1\left(\frac{\left\|u_0^h\right\|_\infty}{u_{\min}^{h,T}} - 1\right)\right)}. \tag{55}$$

To maximize $mh\left(2\lambda - m + h\right) - L_1\left(\frac{\left\|u_0^h\right\|_\infty}{u_{\min}^{h,T}} - 1\right)$, we should 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 (55) becomes

$$\begin{aligned} & u^{m+h-3}\left|\nabla u\right|^2(T,x) \\ \leq\ & \frac{\left\|u_0^h\right\|_\infty - u^h(T,x)}{mT\left(mh + \left(mh - m + 1 - \frac{h^2}{2} - \frac{1}{2}n\left(m-1\right)^2\right)\left(\frac{\left\|u_0^h\right\|_\infty}{u_{\min}^{h,T}} - 1\right) + \frac{h^2}{2}\left(\frac{\left\|u_0^h\right\|_\infty}{u_{\min}^{h,T}} - 1\right)^{-1}\right)}, \end{aligned}$$

and the constraint (54) becomes

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

By rewriting them in terms of U, $U_\pm$ and $h_\pm$, we have Theorem 5.

References

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

[2] D. G. Aronson and L. A. Caffarelli. The initial trace of a solution of the porous medium equation. *Trans. Amer. Math. Soc.* **280** (1983), no. 1, 351–366.

[3] D. G. Aronson and P. Bénilan. Régularité des solutions de l'équation des milieux poreux dans $\mathbb{R}^N$. *C. R. Acad. Sci. Paris Sér. A-B* **288** (1979), no. 2, A103–A105.

[4] P. Bénilan and M. G. Crandall. The continuous dependence on φ of solutions of $u_t - \Delta\varphi(u) = 0$. *Indiana Univ. Math. J.* **30** (1981), no. 2, 161–177.

[5] P. Bénilan, M. G. Crandall and M. Pierre. Solutions of the porous medium equation in $\mathbb{R}^N$ under optimal conditions on initial values. *Indiana Univ. Math. J.* **33** (1984), no. 1, 51–87.

[6] G. Bernard. Existence theorems for fast diffusion equations. *Nonlinear Anal.* **43** (2001), no. 5, Ser. A: Theory Methods, 575–590.

[7] M. Bertsch and M. Ughi. Positivity properties of viscosity solutions of a degenerate parabolic equation. *Nonlinear Anal.* **14** (1990), no. 7, 571–592.

[8] M. Bonforte and J. L. Vazquez. Global positivity estimates and Harnack inequalities for the fast diffusion equation. *J. Funct. Anal.* **240** (2006), no. 2, 399–428.

[9] M. Bonforte and J. L. Vázquez. Positivity, local smoothing, and Harnack inequalities for very fast diffusion equations. *Adv. Math.* **223** (2010), no. 2, 529–578.

[10] 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** (1987), no. 2, 373–401.

[11] L. A. Caffarelli and A. Friedman. Regularity of the free boundary of a gas flow in an n-dimensional porous medium. *Indiana Univ. Math. J.* **29** (1980), no. 3, 361–391.

[12] E. Chasseigne and J. L. Vazquez. Theory of extended solutions for fast-diffusion equations in optimal classes of data. Radiation from singularities. *Arch. Ration. Mech. Anal.* **164** (2002), no. 2, 133–187.

[13] B. E. J. Dahlberg and C. E. Kenig. Nonnegative solutions of generalized porous medium equations. *Rev. Mat. Iberoamericana* **2** (1986), no. 3, 267–305.

[14] J. L. Doob. Stochastic processes. *John Wiley & Sons, Inc., New York; Chapman & Hall, Limited, London*, 1953.

[15] R. S. Hamilton. A matrix Harnack estimate for the heat equation. *Comm. Anal. Geom.* **1** (1993), no. 1, 113–126.

[16] M. A. Herrero and M. Pierre. The Cauchy problem for $u_t = \Delta u^m$ when $0 < m < 1$. *Trans. Amer. Math. Soc.* **291** (1985), no. 1, 145–158.

[17] Y. Hu and Z. Qian. BMO Martingales and Positive Solutions of Heat Equations. *Math. Control Relat. Fields*. To appear.

[18] G. Huang, Z. Huang and H. Li. Gradient Estimates for the Porous Medium Equations on Riemannian Manifolds. *J. Geom. Anal.* **23** (2013), no. 4, 1851–1875.

[19] I. Karatzas and S. E. Shreve. Brownian motion and stochastic calculus. Second edition. Graduate Texts in Mathematics, **113**. *Springer-Verlag, New York*, 1991.

[20] B. L. Kotschwar. Hamilton's gradient estimate for the heat kernel on complete manifolds. *Proc. Amer. Math. Soc.* **135** (2007), no. 9, 3013–3019.

[21] O. A. Ladyzenskaja, V. A. Solonnikov and N. N. Ural'ceva. Linear and quasi-linear equations of parabolic type. *Izdat. "Nauka", Moscow* 1967

[22] P. Li and S. T. Yau. On the parabolic kernel of the Schrödinger operator. *Acta Math.* **156** (1986), no. 3–4, 153–201.

[23] P. Lu, L. Ni, J. L. Vázquez and C. Villani. Local Aronson-Bénilan estimates and entropy formulae for porous medium and fast diffusion equations on manifolds. *J. Math. Pures Appl.* (9) **91** (2009), no. 1, 1–19.

[24] L. Ma, L. Zhao and X. Song. Gradient estimate for the degenerate parabolic equation $u_t = \Delta F(u) + H(u)$ on manifolds. *J. Differential Equations* **244** (2008), no. 5, 1157–1177.

[25] J. Picard. Gradient estimates for some diffusion semigroups. *Probab. Theory Related Fields* **122** (2002), no. 4, 593–612.

[26] D. Revuz and M. Yor. Continuous martingales and Brownian motion. Third edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **293**. *Springer-Verlag, Berlin*, 1999.

[27] P. Souplet and Q. S. Zhang. Sharp gradient estimate and Yau's Liouville theorem for the heat equation on noncompact manifolds. *Bull. London Math. Soc.* **38** (2006), no. 6, 1045–1053.

[28] D. W. Stroock and S. R. S. Varadhan. Multidimensional diffusion processes. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **233**. *Springer-Verlag, Berlin-New York*, 1979.

[29] J. L. Vázquez. Smoothing and decay estimates for nonlinear parabolic equations of porous medium type. 2005.

[30] J. L. Vázquez. Symmetrization and mass comparison for degenerate nonlinear parabolic and related elliptic equations. *Adv. Nonlinear Stud.* **5** (2005), no. 1, 87–131.

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

[32] J. L. Vázquez. The porous medium equation. Mathematical theory. Oxford Mathematical Monographs. *The Clarendon Press, Oxford University Press, Oxford*, 2007.

[33] X. Xu. Gradient estimates for $u_t = \Delta F(u)$ on manifolds and some Liouville-type theorems. *J. Differential Equations* **252** (2012), no. 2, 1403–1420.

[34] S. T. Yau. On the Harnack inequalities of partial differential equations. *Comm. Anal. Geom.* **2** (1994), no. 3, 431–450.

[35] X. Zhu. Hamilton's gradient estimates and Liouville theorems for fast diffusion equations on noncompact Riemannian manifolds. *Proc. Amer. Math. Soc.* **139** (2011), no. 5, 1637–1644.

[36] X. Zhu. Hamilton's gradient estimates and Liouville theorems for porous medium equations on noncompact Riemannian manifolds. *J. Math. Anal. Appl.* **402** (2013), no. 1, 201–206.