Refined asymptotics for the Cauchy problem for the fast p-Laplace evolution equation

Matteo Bonforte[a], Iwona Chlebicka[b], Nikita Simonov[c]

[a] *Departamento de Matemáticas, Universidad Autónoma de Madrid, ICMAT – Instituto de Ciencias Matemáticas, CSIC-UAM-UC3M-UCM, Campus de Cantoblanco, 28049 Madrid, Spain*
[b] *Institute of Applied Mathematics and Mechanics, Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, ul. Banacha 2, 02-097 Warsaw, Poland*
[c] *Laboratoire Jacques Louis Lions (CNRS UMR n° 7598), Sorbonne Université, 4 place Jussieu, 75005 Paris, France*

Abstract

Our focus is on the fast diffusion equation $\partial_t u = \Delta_p u$ with $p < 2$ in the whole Euclidean space of dimension $N \geq 2$. The properties of the solutions to the p-Laplace Cauchy problem change in several special values of the parameter p.

In the range of p when mass is conserved, non-negative and integrable solutions behave like the Barenblatt (or fundamental) solutions for large times. By making use of the entropy method, we establish the polynomial rates of the convergence in the uniform relative error for a natural class of initial data. The convergence was established in the literature for p close to 2, but no rates were available. In particular, we allow for the values of p, for which the entropy is not displacement convex, as we do not apply the optimal transportation tools.

We approach the issue of long-term asymptotics of the gradients of solutions. In fact, in the case of the radial initial datum, we provide also polynomial rates of the uniform convergence in the relative error of radial derivatives of solutions for $\frac{2N}{N+2} < p < 2$.

Finally, providing an analysis of needed properties of solutions for the entropy method to work, we open the question on the full description of the basin of attraction of the Bareblatt solutions for p close to 1.

Keywords: Asymptotical behaviour of solutions, Cauchy problem, p-Laplacian
2020 MSC: 35B40 (35K15, 35J92)

Contents

Email addresses: matteo.bonforte@uam.es (Matteo Bonforte), i.chlebicka@mimuw.edu.pl (Iwona Chlebicka), nikita.simonov@sorbonne-universite.fr (Nikita Simonov)

1. Introduction

Nonlinear evolution equations involving the p-Laplace operator $\Delta_p u = \operatorname{div}\left(|\nabla u|^{p-2}\,\nabla u\right)$ attract attention of the mathematical community for over a half of century. The interest come from both – the preeminent role of the power-growth operators in the modelling of fluid dynamics [56] and intrinsic mathematical challenges, see .g. [36, 48, 23, 54] and references therein.

In this paper we are interested in the long-time behaviour of solutions to the following Cauchy problem

$$\begin{cases} \partial_t u = \operatorname{div}\left(|\nabla u|^{p-2}\,\nabla u\right) & x \in \mathbb{R}^N, \quad t > 0\,, \\ u(0,x) = u_0(x) \geq 0\,, \quad x \in \mathbb{R}^N\,, \end{cases} \tag{CPLE}$$

where the exponent $1 < p < 2$. The above equation is called the *p-Laplace evolution* equation, it is a nonlinear, gradient-driven diffusion equation which is singular in the considered regime. It is known that, if $u_0 \in \mathrm{L}^1(\mathbb{R}^N)$ then the solution exists, is unique and *conservation of mass* holds, i.e., for every $t > 0$ it holds

$$\int_{\mathbb{R}^N} u(t,x)\,\mathrm{d}x = \int_{\mathbb{R}^N} u_0(x)\,\mathrm{d}x \quad \forall t > 0\,, \quad \text{if} \quad p > p_c := \frac{2N}{N+1}\,. \tag{1}$$

Moreover, the solution is a continuous curve in $\mathrm{L}^1(\mathbb{R}^N)$, that is $u(t) \in C\left([0,\infty)\,, \mathrm{L}^1(\mathbb{R}^N)\right)$, it remains nonnegative once $u_0 \geq 0$, and $u(t,\cdot) \in C^{1,\alpha}(\mathbb{R}^N)$ for some $\alpha \in (0,1)$. When $p_c < p < 2$ the behaviour of solutions to (CPLE) for large times is described by the means of *fundamental solution* (also called the *Barenblatt* solution) defined as

$$\mathcal{B}_M(t,x) := t^{\frac{1}{2-p}} \left[b_1\, t^{\frac{\beta\, p}{p-1}}\, M^{\frac{\beta\, p\,(p-2)}{p-1}} + b_2\, |x|^{\frac{p}{p-1}}\right]^{\frac{1-p}{2-p}} \quad \text{where} \quad \beta = \frac{1}{p - N(2-p)} \tag{2}$$

and where M represents the (conserved) mass of $\mathcal{B}_M$, b_1 and b_2 are given numerical constants (see (40) and (39) for their definitions). The Barenblatt solution $\mathcal{B}_M$ has a Dirac delta $M\,\delta$ as initial datum. It will be useful for us to express $\mathcal{B}_M$ with the use of $V_D : \mathbb{R}^N \to [0,\infty)$ given by the following formula

$$V_D(y) := \left(D + \tfrac{2-p}{p}|y|^{\frac{p}{p-1}}\right)^{-\frac{p-1}{2-p}}\,. \tag{3}$$

Having $R_T : [0,\infty) \to [0,\infty)$ defined as

$$R_T(t) := \left(\tfrac{t+T}{\beta}\right)^{\beta}\,, \tag{4}$$

we have

$$\mathcal{B}_M(t+T,x) = R_T^{-N}(t) V_D\left(\frac{x}{R_T(t)}\right) \qquad \text{and} \qquad D := \beta^{N\beta\frac{2-p}{p-1}}\, \frac{b_1}{M^{\frac{\beta p(2-p)}{p-1}}}\,. \tag{5}$$

In the case of the heat equation ($p = 2$ in (CPLE)) the fundamental solution (also called the Gaussian profile in that case) decays exponentially in space. In contrast to this fact, in the present nonlinear case ($p < 2$) the Barenblatt solution develops a *fat tail*, i.e. it decays in space with a polynomial rate.

For nonnegative initial data $u_0 \in \mathrm{L}^1(\mathbb{R}^N)$, solutions to (CPLE) *relax to self-similarity* and the precise result can be stated as follows

$$\|u(t,\cdot) - \mathcal{B}_M(t,\cdot)\|_{\mathrm{L}^1(\mathbb{R}^N)} \to 0 \quad \text{and} \quad t^{N\beta}\|u(t,\cdot) - \mathcal{B}_M(t,\cdot)\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \to 0 \quad \text{as} \quad t \to \infty\,, \tag{6}$$

where $M = \int_{\mathbb{R}^N} u_0(x)\,\mathrm{d}x$, the factor $t^{N\beta}$ in front of the L^∞-norm is necessary to get a meaningful result since the L^∞ norm of solutions decays in time as $t^{-N\beta}$. We notice that, in order to speak of results as (6), we shall often use the term *convergence*, even if the Barenblatt function $\mathcal{B}_M$ is not a stationary profile. By interpolation, similar results can be obtained for L^q-norms, for $1 < q < \infty$. It is known that, without no further assumptions, such results are sharp in terms of strong norm of convergence and no rates are available. However, results as in (6) do not take into account neither the tail behaviour of the Barenblatt nor of the solution $u(t,x)$ itself, and one may ask whether we can obtain a finer description of the tail behaviour for solutions to (CPLE). This was done in [20] where solution with the same polynomial tail behaviour of the Barenblatt profile have been completely characterized. Indeed, in [20, Theorem 1.1], much stronger convergence result was proven with a fine description of the tail behaviour for solutions to (CPLE).

The main result of that paper is the proof of *uniform convergence in relative error* (UCRE), that is, for $N \geq 1$ and $p_c < p < 2$, we have

$$\left\| \frac{u(t,\cdot)}{\mathcal{B}_M(t,\cdot)} - 1 \right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \xrightarrow[t\to+\infty]{} 0 \quad \text{if and only if} \quad \|u_0\|_{\mathcal{X}_p} := \sup_{R>0} R^{\frac{p}{2-p}-N} \int_{|x|\geq R} u_0(x)\,\mathrm{d}x < \infty\,. \tag{7}$$

In view of the above result, one may ask whether or not one can obtain rates for the uniform relative error. This is the main question we want to address in this paper which we can restate as follows:

Is it possible to prescribe an explicit polynomial rate for the uniform converge in relative error of solutions? (Q1)

The fact that the convergence rate must be at most polynomial (at least in the class of initial data in $\mathcal{X}_p$) has been carefully shown in [49] for a different equation, but the same reasoning applies to (CPLE). It can be easily seen by analysing the case $u(t,x) = \mathcal{B}_M(t+T,x)$: we have that $|\mathcal{B}_M(t+T,x)/\mathcal{B}_M(t,x) - 1|$ is of order t^{-1}.

Before stating our main results, we want to stress that the behaviour of solutions to (CPLE) changes in the range $p \in (1,2)$. For instance, when $1 < p < p_c$, even if solutions exist they do not conserve mass any more. Another important parameter in our analysis is

$$p_M := \frac{3(N+1) + \sqrt{(N+1)^2+8}}{2(N+2)} \in (p_c, 2). \tag{8}$$

Indeed, for $p' = p/(p-1)$, solutions whose initial datum is in $\mathcal{X}_p$ have their weighted $|x|^{p'}$-moments finite along the flow (i.e., $\int_{\mathbb{R}^N} |x|^{p'} u_0(x)\,\mathrm{d}x < \infty$ implies that $\int_{\mathbb{R}^N} |x|^{p'} u(t,x)\,\mathrm{d}x < \infty$ for all $t > 0$), while this property is lost for $p \leq p_M$. This is an important distinction since our analysis is based on the *entropy method* for which moments of solutions play a crucial role. For an overview of the method we refer to [7] and references therein. In our main result, we are able to provide an explicit and uniform rate of convergence for the uniform relative error as long as $u_0 \in \mathcal{X}_p$ and $p_M < p < 2$, while we rely on a further assumption for lower values of p. Our main result in the range $p_c < p < 2$ reads as follows.

Theorem 1. *Let $N \geq 3$, $0 \leq u_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap \mathcal{X}_p$, and $M := \|u_0\|_{L^1(\mathbb{R}^N)} > 0$. Assume u is a weak solution to* (CPLE) *with initial datum u_0. Suppose one of the following holds:*

- *(i) $p_M < p < 2$;*
- *(ii) $p_c < p \leq p_M$ and there exists $M_2 > M_1 > 0$ and $T > 0$ such that*

$$\mathcal{B}_{M_1}(T,x) \leq u_0(x) \leq \mathcal{B}_{M_2}(T,x) \qquad \forall\, x \in \mathbb{R}^N\,. \tag{9}$$

Then there exists $T_\star = T_\star(p,N,M,\|u_0\|_{\mathcal{X}_p}) > 0$, $K_\star = K_\star(p,N,M,\|u_0\|_{\mathcal{X}_p}) > 0$ and $\sigma = \sigma(p,N) > 0$ such that

$$\left\| \frac{u(t,\cdot) - \mathcal{B}_M(t,\cdot)}{\mathcal{B}_M(t,\cdot)} \right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \leq K_\star\, t^{-\sigma} \qquad \forall\, t \geq T_\star\,. \tag{10}$$

Remark 1.1. We notice that in the case $N = 2$, Theorem 1 still holds as long as $p \neq \frac{3}{2}$.

We notice that in the case $p_M < p < 2$, we only need to assume that the initial datum u_0 belongs to $\mathcal{X}_p$. This is the minimal assumption for uniform convergence in relative error, as was shown in [20, Theorem 1.1]. Indeed, for initial data in $\mathrm{L}^1(\mathbb{R}^N) \setminus \mathcal{X}_p$, such property is simply false, see [20, Proposition 5.1]. At the same time, Theorem 1 gives (at least in the range $p_M < p < 2$) a uniform convergence rate for the whole class $\mathcal{X}_p$ which is independent of the initial datum. In the case $p_c < p \leq p_M$, we need to ask the stronger assumption (9). This is mainly asked to ensure that the *relative entropy* is also finite for smaller p. This condition may appear very strong; however, it is known that, under the assumption $u_0 \in \mathcal{X}_p$, a very similar condition holds (with different T_1 and T_2) along the flow: here, we are only asking for a more localized tail behaviour.

Let us comment on the related results from the literature. The existence and uniqueness of the Barenblatt solution defined in (2) has been shown in [48]. In the same article, the authors prove the convergence of non-negative and integrable solutions towards $\mathcal{B}_M$ in the L^1 topology, however, with no rates. Rates of convergence were obtained, in the case $p_c + \frac{1}{N+1} < p < 2$, in [26]. In the same range, in [1, 2], optimal rates of convergence were obtained (in the class of initial data with finite $|x|^{p'}$ moment) in the L^1 and Wasserstein topologies. Finally, in [3], by using the entropy method (which we will mimic here), the authors were able to obtain explicit convergence rates in the whole range $p_c < p < 2$, roughly under the same assumptions as those of Theorem 1. In fact, the first result that considers convergence in relative error is [3, Lemma 2.5]. They prove a convergence result uniformly on compact sets: we can notice that this also can be inferred from the L^∞ convergence of (6), see for instance [19, Theorem 2.4]. Apart from [20] no other article has addressed (Q1) neither convergence rates in the uniform relative error. This form of convergence is much stronger than previously known results from the literature (stated either in with the use of the L^q or Wasserstein convergence). We also recall that the basin of attraction towards the Barenblatt solution in uniform relative error has been characterized as the space $\mathcal{X}_p$ only recently in [20]. To sum up, the convergence (6) was known for p close to 2, but no rates of convergence were provided so far.

One of the main steps in the proof of Theorem 1, presented in Sections 5 and 6, is the computation of a convergence rate in the L^1-topology. We do not claim originality for this result since similar results have been obtained in [27, 2, 3]. However, we give, perhaps for the first time, a complete proof for such a result in the entire range $p_c < p < 2$ with a correction of a flaw contained in [3]. Our result reads as

Proposition 1.2. *Under the same assumptions of Theorem 1, there exists* $\widetilde{T} = \widetilde{T}(p, N, M, \|u_0\|_{\mathcal{X}_p}) > 0$, $\widetilde{K} = \widetilde{K}(p, N, M, \|u_0\|_{\mathcal{X}_p}) > 0$ *and* $\nu = \nu(p, N) > 0$ *such that*

$$\|u(t,\cdot) - \mathcal{B}_M(t,\cdot)\|_{\mathrm{L}^1(\mathbb{R}^N)} \le \widetilde{K}\, t^{-\nu} \quad \forall\, t \ge \widetilde{T}\,. \tag{11}$$

The above result is stated for the L^1 distance. However, from this result, one can easily infer similar results for the L^p-topology for $1 < p < \infty$, also considering the rescaling factor needed for these norms. On the other hand, once Theorem 1 is obtained, one can deduce a convergence rate in the L^∞-topology.

The (CPLE) is a gradient-driven diffusion, therefore, it is quite natural to investigate whether or not the convergence in relative error may hold for derivatives of solutions. This has been established in [20, Theorem 1.4] in the case of radial derivatives for a radially decreasing initial datum and under some assumptions on the space decay of the radial derivative. The result reads as

$$\left\|\frac{\partial_r u(t,\cdot)}{\partial_r \mathcal{B}_M(t,\cdot)} - 1\right\|_{L^\infty(\mathbb{R}^N)} \xrightarrow[t\to+\infty]{} 0\,. \tag{12}$$

Interestingly, at least in dimension $N = 1$, having a symmetric and decreasing initial datum is a necessary condition; counterexamples have been provided in [20, Remark 1.5]. Even if counterexamples were not constructed, we believe that for $N \ge 2$, having a radially symmetric and decreasing initial datum is also a necessary condition. We are interested in the following issue:

Is it possible to prescribe an explicit polynomial rate for the uniform converge in relative error of radial derivatives of radial solutions? (Q2)

While, a priori, there is no reason why the rate should be polynomial, this can be inferred from the same example which applies to question (Q1): just consider the case $u(t,x) = \mathcal{B}_M(t+T,x)$. We provide a first answer for $p > p_c$ as the following estimate.

Theorem 2 (Convergence in relative error for radial derivatives for $p > p_c$). *Let $N \ge 2$, $p_c < p < 2$, $0 \le u_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap C^2(\mathbb{R}^N)$ be radially symmetric and $M := \|u_0\|_{L^1(\mathbb{R}^N)} > 0$ and u is a solution to*(CPLE) *with datum u_0. Suppose one of the following holds:*

(i) $p_M < p < 2$ and there exist $A > 0$ and $R_0 > 0$ satisfying

$$\partial_r u_0(r) \le 0 \quad \textit{and} \quad |\partial_r u_0(r)| \le A\, r^{-\frac{2}{2-p}} \quad \forall r \ge R_0\,. \tag{13}$$

(ii) $p_c < p \le p_M$ and there exist $M_2 > M_1 > 0$ and $T > 0$ such that

$$\partial_r \mathcal{B}_{M_1}(T,r) \le \partial_r u_0(r) \le \partial_r \mathcal{B}_{M_1}(T,r) \quad \forall\, r \ge 0\,. \tag{14}$$

Then there exist $t_\star > 0$, $k_\star > 0$ and $\lambda = \lambda(p,N) > 0$ such that

$$\left\|\frac{\partial_r u(t,\cdot)}{\partial_r \mathcal{B}_M(t,\cdot)} - 1\right\|_{L^\infty(\mathbb{R}^N)} \le k_\star\, t^{-\lambda} \quad \forall\, t \ge t_\star\,, \tag{15}$$

where $\partial_r u$ (resp. $\partial_r \mathcal{B}_M$) is the radial derivative of u (resp. $\mathcal{B}_M$). When $p_M < p < 2$ then $t_\star = t_\star(u_0, A, R_0, M, p, N)$ and $k_\star = k_\star(u_0, A, R_0, M, p, N)$, while when $p_c < p \le p_M$ then $t_\star = t_\star(u_0, M_2, M_1, T, p, N)$ and $k_\star = k_\star(u_0, M_2, M_1, T, p, N)$.

Remark 1.3. We notice that the regularity assumption on the initial datum, i.e., $u_0 \in C^2(\mathbb{R}^N)$ is imposed for the sake of simplicity of exposition. More precisely, the regularity threshold for our proof to work is $C^{1,\alpha}(\mathbb{R}^N)$ for $\alpha > 0$ depending on p. The same remark applies also to Theorems 3 and 4.

This result is proven in Section 7. The above theorem's proof is based on analogies between equation (CPLE) and a weighted fast diffusion equation. It is well-known that (CPLE) shares several properties with the *Fast Diffusion Equation* (FDE for short), which reads as $\partial_t \Phi = \Delta \Phi^m$, see the monographs [62, 61]. Here, however, we exploit a stronger relation between radial derivatives of solutions to (CPLE) and radial solutions to a weighted version of the FDE, formally established in [45]. In fact, the radial formulation of (CPLE) can be rewritten for u being a function of $(r = |x|, t)$ as follows

$$\partial_t u = r^{1-N}\, \partial_r \left(r^{N-1}\, |\partial_r u|^{p-2}\, \partial_r u \right) . \tag{16}$$

Let us consider $\Phi : \mathbb{R}^N \times (0, \infty) \to \mathbb{R}$ being a *non-negative* function of $(\varrho = |x|, t)$, and a solution to

$$\partial_t \Phi = \varrho^{1-n}\, \partial_\varrho \left(\varrho^{n-1}\, \partial_\varrho \Phi^m \right) , \qquad m = p - 1 , \tag{17}$$

where $n = N - 2 - 2\frac{N}{p'}$ is a positive parameter. In Section 7 we point out in more details in what sense a derivative of a solution to (17) solves (16) and comment on related results.

We are now ready to address (Q2) in both the critical range $p = p_c$ and in a subrange of the subcritical range $1 < p \le p_c$. Let us mention two major properties of solutions to (CPLE) that change below the critical exponent p_c. The first one is that for $p < p_c$ the mass is not conserved anymore (i.e. (1) is not valid) and solutions extinguish in finite time. The second one is that for $p \le p_c$ the Barenblatt profile as defined (2) ceases to exist as a self-similar function that takes a Dirac δ_0 as its initial datum. At the same time, the long-time behaviour for solutions, in the range $1 < p \le p_c$ is far richer than in the supercritical fast diffusion range. However, a *pseudo*-Barenblatt solution can still be defined in this range as a self-similar profile by the formula

$$\mathcal{B}_{D,T}(t, x) = R_T^{-N}(t) V_D \left(\frac{x}{R_T(t)} \right) , \tag{18}$$

where V_D is as in (3) and R_T is given by

$$R_T(t) := \left(\frac{T - t}{|\beta|} \right)_+^{\beta} \quad \text{for} \;\; 1 < p < p_c \qquad \text{and} \qquad R_T(t) := \exp\{\ell(T + t)\} \quad \text{if} \;\; p = p_c , \tag{19}$$

where β is as in (2) (notice that $\beta < 0$ for $1 < p < p_c$), while $\ell > 0$ is a free parameter. We stress that we do not keep track of the dependence of R_T on the constant ℓ for $p = p_c$, but it is possible. Notice also that, since $\mathcal{B}_{D,T}(t, \cdot) \notin \mathrm{L}^1(\mathbb{R}^N)$ for any $0 < t < T$, the parameter D in V_D is a free parameter which does not represent the mass anymore. Furthermore, for $1 < p < p_c$ *pseudo*-Barenblatt solutions vanish in finite time T, but for $p = p_c$ they are defined for all times. It is an open (and hard) problem to fully describe the *basin of attraction* of the pseudo-Barenblatt solutions. Here we begin by giving some sufficient conditions for the initial datum of a solution whose long-time behaviour is represented by pseudo-Barenblatts. We present now our answer to (Q2) for the critical value of p, which is justified in Section 7.

Theorem 3 (Convergence in relative error for radial derivatives for $p = p_c$). *Let $N \ge 2$, $p = p_c$, and u be a solution to* (CPLE) *with an initial datum $0 \le u_0 \in C^2(\mathbb{R}^N)$, which is radial and decreasing. Suppose that there exist $D_1, D_2 > 0$ and $T > 0$, such that*

$$\partial_r \mathcal{B}_{D_1,T}(0, r) \le \partial_r u_0(r) \le \partial_r \mathcal{B}_{D_2,T}(0, r) \quad \forall r \ge 0 . \tag{20}$$

Then there exists $D = D(u_0) > 0$, $T_\diamond = T_\diamond(u_0) > 0$, $C_\diamond = C_\diamond(u_0, D_1, D_2, N, p) > 0$ and $\lambda = \lambda(N, p) > 0$ such that

$$\left\| \frac{\partial_r u(t, \cdot)}{\partial_r \mathcal{B}_{D,T}(t, \cdot)} - 1 \right\|_{L^\infty(\mathbb{R}^N)} \le C_\diamond\, t^{-\lambda} \quad \textit{and} \quad \left\| \frac{u(t, \cdot)}{\mathcal{B}_{D,T}(t, \cdot)} - 1 \right\|_{L^\infty(\mathbb{R}^N)} \le C_\diamond\, t^{-\lambda} \quad \forall\, t > T_\diamond . \tag{21}$$

Investigation on the strictly subcritical case $1 < p < p_c$, requires introducing more thresholds for p. The first one is

$$p_Y := \frac{2N}{N + 2} , \tag{22}$$

where the p_Y stays for Yamabe exponent. We decide to call this exponent after Yamabe since, in this case, the radial derivative of (CPLE) solves the equation $\partial_t \Phi = \Delta \Phi^{\frac{N-2}{N+2}}$ (see Section 7 for more information). This equation is related to the Yamabe flow for a conformally flat metric, see [28, 25, 69, 9]. The value p_Y is a sharp threshold for the gradient regularity of solutions when no extra assumptions are imposed, cf. Section 4.2 and [36, Section 21.3]. Another threshold appears when one considers integrability properties of the difference of two Barenblatt solutions $\mathcal{B}_{D_1,T} - \mathcal{B}_{D_2,T}$ (defined in (18)) for different D_1, D_2 and the same $T > 0$. As it is clear from Lemma 9.2, the difference is integrable when

$$1 < N < \frac{p}{(2-p)(p-1)}\,. \tag{23}$$

Note that the condition (23) is satisfied in low dimensions ($1 < N < 6$) for all $p \in (1,2)$, while in high dimensions ($N \geq 6$) for p close to 1 and close to 2, precisely for $p \in (1, p_1) \cup (p_2, 2)$ where

$$p_1 := \tfrac{3}{2} - \tfrac{1}{2N} - \tfrac{\sqrt{N^2-6N+1}}{2N} \qquad \text{and} \qquad p_2 := \tfrac{3}{2} - \tfrac{1}{2N} + \tfrac{\sqrt{N^2-6N+1}}{2N}\,. \tag{24}$$

We stress that $\max\{p_Y, p_2\} < p_c$, but relation between p_Y and p_2 depends on the dimension. In Section 3 we collect all special values of p and indicate their interplays.

We are in the position to state the following theorem describing a quite strong result for radially decreasing data and $p < p_c$ being our last answer to (Q2). See Section 7 for the proof.

Theorem 4 (**Convergence in relative error for radial derivatives for** $p < p_c$)**.** *Let $N \geq 2$, $1 < p < p_c$, and let u be a solution to* (CPLE) *with an initial datum $0 \leq u_0 \in C^2(\mathbb{R}^N)$ radial and decreasing. Suppose that there exist $D_1, D_2 > 0$ and $T > 0$ it holds* (20)*. Assume further that one of the following conditions (i)–(iv) is satisfied*

(i) $N = 2$ *and* $1 = p_Y < p < p_c$,

(ii) $2 < N < 6$ *and* $p_Y \leq p < p_c$,

(iii) $N \geq 6$ *and* $p_2 < p < p_c$,

(iv) $N \geq 6$, $p_Y \leq p \leq p_2$, *and there exist* $\widetilde{D} > 0$ *and* $f \in \mathrm{L}^1((0,\infty), r^{n-1}\,\mathrm{d}r)$ *with* $n = 2(1 + N/p')$, *such that*

$$\partial_r u_0(r) = \partial_r \mathcal{B}_{\widetilde{D},T}(0,r) + r^{\frac{1}{p-1}} f(r^{\frac{p}{2(p-1)}}) \quad \forall r \geq 0\,. \tag{25}$$

Then there exists $D = D(u_0) > 0$, $T_\diamond = T_\diamond(u_0) > 0$, $C_\diamond = C_\diamond(u_0, D_1, D_2, N, p) > 0$ and $\lambda = \lambda(N,p) > 0$ such that

$$\left\| \frac{\partial_r u(t,\cdot)}{\partial_r \mathcal{B}_{D,T}(t,\cdot)} - 1 \right\|_{L^\infty(\mathbb{R}^N)} \leq C_\diamond\,(T-t)^{-\lambda} \quad \textit{and} \quad \left\| \frac{u(t,\cdot)}{\mathcal{B}_{D,T}(t,\cdot)} - 1 \right\|_{L^\infty(\mathbb{R}^N)} \leq C_\diamond\,(T-t)^{-\lambda} \quad \forall\, t_\diamond < t < T\,. \tag{26}$$

Moreover, if $N \geq 6$ and $p_Y \leq p \leq p_2$, then $D = \widetilde{D}$.

At this point, it is pretty clear that the pseudo-Barenblatt solutions defined in (18) also describe the asymptotic behaviour of a class of solution to (CPLE) in the subcritical regime $1 < p \leq p_c$. It is, therefore, natural to address the following more general question.

What is the basin of attraction of the pseudo-Barenblatt solution in the sub-critical regime $1 < p \leq p_c$? (Q3)

In the same spirit of [9], we shall address this question using the entropy method and build a theory that gathers the main steps used to prove Theorem 1. Our main result (which actually will cover the range defined in (23)) will be given in Theorems 5 and 6. To the best of our knowledge, at least in the case of the (CPLE), this is the first paper that addresses (Q3). However, we prefer to postpone the discussion about (Q3) to Section 2 since we need to introduce several tools, which are also required for the proof of Theorem 1. In particular, since solutions may vanish in finite time, the convergence towards the pseudo-Barenblatt solutions is subtle. Therefore, it is convenient to introduce a natural rescaling which transforms this delicate problem into the study of the convergence to stationary solutions of a particular nonlinear Fokker–Planck equation.

Let us now focus on the organization of the present article.

Organization. In Section 2, we provide information on the rescaled problem, introduce the main ideas of the entropy method, and formulate a counterpart of Theorem 1 for small values of p in rescaled variables. Section 3 collects special values of p and exposes their role. We give basic information on the problems like (CPLE) in Section 4. Section 5 is devoted to proving convergence in relative error under the assumption that the solutions converge in L^1. The proof of our main accomplishment, i.e. Theorem 1, is included in Section 6. In Section 7, we pass to question (Q2) and concentrate on the results for radial solutions. For the convergence results for low values of p, see Section 8. In the end, we prove some lemmata and list the most important parameters in the appendix.

2. Nonlinear Fokker–Planck equation and results on the asymptotics for small values of p

The main goal of this section is twofold: first provide an introduction to the entropy method and introduce the needed tools, second, we give a partial answer to (Q3) posed in Introduction.

2.1. Nonlinear Fokker–Planck equation

The entropy method is much better understood when introducing new variables for which the Barenblatt and pseudo-Barenblatt functions are stationary profiles. We shall first consider the case $p_c < p < 2$. In this case, we recall that solutions "live" for all times and the same happens for the critical case $p = p_c$. For $1 < p < p_c$, solutions extinguish in finite time. Consider the following change of variables

$$v(\tau, y) := R_T^N(t) u(t, x) \qquad \text{and} \qquad v_0(y) := R_T^N(0) u_0(x) \tag{27}$$

with

$$\tau := \ln \frac{R_T(t)}{R_T(0)}\,, \qquad \text{and} \qquad y := \tfrac{x}{R_T(t)}\,. \tag{28}$$

where $R_T(t)$ is as in (4). Then, problem (CPLE) becomes a nonlinear equation of the Fokker–Planck type reading

$$\begin{cases} \partial_\tau v(\tau, y) &= \operatorname{div}_y\Big(|\nabla v(\tau, y)|^{p-2} \nabla v(\tau, y) + y\, v(\tau, y)\Big)\,, \\ v(x, 0) &= v_0(x)\,. \end{cases} \tag{29}$$

The main advantage is that problem (29) has a family of stationary solutions V_D given by (3). A simple computation shows that $|\nabla V_D(y)|^{p-2}\nabla V_D(y) = -y\, V_D(y)$, for any $y \in \mathbb{R}^N$. We notice that at least when $p_c < p < 2$, the parameter D is completely determined by the mass of the function V_D. In the regime $p_c < p < 2$, most of the time we shall choose $T = \beta$ in (4), which sets the initial datum $v_0 = u_0$ (since $R_\beta(0) = 1$) and makes most of the computations easier. Namely, we pick

$$R_\beta(t) = \left(1 + \tfrac{t}{\beta}\right)^\beta\,, \quad \text{such that} \quad V_D(y) = R_\beta^N(t)\, \mathcal{B}_M(t + \beta, x)\,. \tag{30}$$

We shall call the functions V_D *stationary Barenblatt profiles* or simply *Barenblatt profiles* when no confusion arises. We also notice that, among all the non-stationary Barenblatt solutions defined in (2), only $\mathcal{B}_M(t+\beta, x)$ is transformed to the stationary one through the change of variables defined above. All the other functions from the family do not become stationary after this change of variables.

In the critical case $p = p_c$, the situation does not change much. In (27) and (28) we shall use R_T defined in (19) instead of the one defined in (4). At this moment, we shall not specify the value of ℓ in (19) and will specify it when needed. Additionally let us notice that in this case it is the pseudo-Barenblatt profile defined in (18) is transformed into the stationary profile V_D.

The subcritical case $1 < p < p_c$ is essentially different than the abovementioned ones. In this case, R_T is defined as in the first formula of (19). The change of variables (27) makes sense only for solutions that vanish in time T. So, in order to make a proper use of this change of variables, one needs to know a priori when the solution under consideration vanishes. This is often assumed as a hypothesis, as in Theorem 4. We also notice that the vanishing time T has disappeared in the new equation (29) and solutions are defined for all $\tau \in (0, \infty)$.

Let us now introduce entropy functional and the relative Fisher information. In order to do so, we define the following parameter

$$\gamma := \frac{2p-3}{p-1}\,, \qquad \text{so} \qquad \gamma - 1 = -\frac{2-p}{p-1} < 0 \qquad \text{and} \qquad \gamma - 2 = \frac{p-2-p+1}{p-1} = -\frac{1}{p-1} < 0\,. \tag{31}$$

In what follows, we shall also need the following auxiliary expression

$$\mathbf{b}[\phi] := |\nabla\phi|^{p-2}\nabla\phi \qquad \text{and} \qquad \mathbf{a}[\phi] := \tfrac{1}{\phi}\mathbf{b}[\phi]\,. \tag{32}$$

Then, the *relative entropy* is given by the formula

$$\mathcal{E}[v(\tau,\cdot)|V_D] := \frac{1}{\gamma(\gamma-1)}\int_{\mathbb{R}^N}\Big\{v^\gamma(\tau,y) - V_D^\gamma(y) - \gamma V_D^{\gamma-1}(y)\big[v(\tau,y) - V_D(y)\big]\Big\}\,\mathrm{d}y\,. \tag{33}$$

In relation to it, recalling (32), we define *relative Fisher information* by

$$\mathcal{I}[v(\tau,\cdot)|V_D] := \frac{1}{|\gamma-1|^p}\int_{\mathbb{R}^N} v(\tau,y)\big(\nabla v^{\gamma-1}(\tau,y) - \nabla V_D^{\gamma-1}(y)\big)\cdot\Big(\mathbf{b}[v^{\gamma-1}(\tau,y)] - \mathbf{b}[V_D^{\gamma-1}(y)]\Big)\,\mathrm{d}y\,. \tag{34}$$

This quantity, in connection with Fokker–Planck-type equations, is commonly referred to as relative entropy production. We notice that, when no confusion arises, we shall write $\mathcal{E}[v(\tau)|V_D]$ ($\mathcal{I}[v(\tau)|V_D]$ resp.) instead of $\mathcal{E}[v(\tau,\cdot)|V_D]$ ($\mathcal{I}[v(\tau,\cdot)|V_D]$ resp.).

The *entropy method* consists in proving that the entropy functional converges to zero exponentially. Consequently, we will prove the convergence of $v(\tau)$ towards V_D in the L^1-norm. With this aim, we need to show that the Fisher information is the derivative of the entropy functional along the flow. Let us note that from the above definitions, it does not result that $\mathcal{E}$ or $\mathcal{I}$ are well-defined for every set of parameters under consideration. We shall prove that these quantities are finite under specified regimes. Further we prove that for a solution v to problem (29) it holds that

$$\tfrac{d}{d\tau}\mathcal{E}[v(\tau,\cdot)|V_D] = -\mathcal{I}[v(\tau,\cdot)|V_D]\,.$$

We shall clarify all these details in Subsection 6.1. Having this relation, it is to be proven that the Fisher information controls the entropy functional, at least along the flow. This is a key step in the *entropy method*. Indeed, after obtaining that, for a positive constant $c > 0$, we have

$$\mathcal{I}[v(\tau)|V_D] \geq c\,\mathcal{E}[v(\tau)|V_D]\,, \tag{35}$$

one can easily find that $\frac{d}{d\tau}\mathcal{E}[v(\tau)|V_D] \leq -c\,\mathcal{E}[v(\tau)|V_D]$. Then, by a version of the Gronwall's Lemma, we shall obtain the exponential convergence of the entropy towards zero, namely $\mathcal{E}[v(\tau)|V_D] \leq e^{-c\tau}\mathcal{E}[v_0|V_D]$. This is a major step for us, since by the Csiszár–Kullback inequality (Lemma 9.1), we can easily infer the L^1-convergence of $v(\tau)$ towards V_D. Of course, establishing inequality (35) (also sometimes called a *entropy–entropy production inequality*) is a major difficulty. This has already been done in the range $p_D \leq p < 2$, by techniques involving optimal transport, see [2] and Section 6.3. We stress that for the range $p_c \leq p < p_D$ such a clear result is missing since within this range $\mathcal{E}$ and $\mathcal{I}$ are well-defined only when the solution is close to the Barenblatt profile V_D. In [3], the authors opened the way for establishing the needed relations. In the range $p_c \leq p < 2$, we take into account recent results established in [20] in order to simplify some of the proofs. We take inspiration from their method.

In order to approach an answer to (Q3) we also analyse the range $1 < p \leq p_c$. We remark that in this case some key estimates for the entropy functional are lost. However, it is possible to carefully analyze the argument used in the previous range and understand the needed conditions for the entropy method to work. We resume these conditions in Theorems 5 and 6 presented in the following section. Let us stress that there are no convergence results known for $1 < p \leq p_c$ so far.

2.2. Asymptotics for small values of p and general class of initial data

Our further analysis is motivated by question (Q3) in the subcritical range. Of course, using the entropy method, the basin of attraction of a *pseudo*-Barenblatt solutions that can be found is narrowed to the class of solutions with finite relative entropy with respect to the Barenblatt solution. We have the following result aimed at relaxing the range on p from Theorem 1 under no assumption on the radiality of initial data. We formulate it for the rescaled problem (29). The proof can be found in Section 8. The notion of solution, its existence and uniqueness is as discussed in Section 4.2.

Theorem 5. *Suppose condition* (23) *holds and* $p \in (1,2)\setminus\{\frac{3}{2}\}$. *Let* v *be a non-negative weak solution to* (29) *with non-negative initial datum* $v_0 \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N)\cap\mathrm{L}^\infty(\mathbb{R}^N)$. *Moreover, we assume what follows.*

(i) Suppose that there exist $D_1, D_2 > 0$, *such that it holds*

$$V_{D_1}(y) \leq v_0(y) \leq V_{D_2}(y) \quad \forall y \in \mathbb{R}^N\,;$$

(ii) Let $D>0$ be such that for every $\tau>0$ it holds

$$\int_{\mathbb{R}^N} (v(\tau,y)-V_D(y))\,\mathrm{d}y = \int_{\mathbb{R}^N} (v_0(y)-V_D(y))\,\mathrm{d}y = 0\,;$$

(iii) Let $D>0$ be as in (ii). There is $\tau_0>0$, such that $\tau\mapsto \mathcal{I}[v(\tau\,,\cdot)|V_D]\in \mathrm{L}^1(\tau_0,\infty)$ and

$$\mathcal{E}[v(\tau_1)]-\mathcal{E}[v(\tau_2)] = \int_{\tau_1}^{\tau_2} \mathcal{I}[v(s)|V_D]\,\mathrm{d}s \quad \text{for all} \quad \tau_2\ge\tau_1\ge\tau_0>0\,;$$

(iv) Let τ_0 be as in (iii). There is $\alpha\in(0,1)$ such that for $\tau>\tau_0$ the solution v is $C^{1,\alpha}$-regular locally in space and time with an uniform constant, i.e., there exists $C=C(\alpha,\tau_0,v_0)>0$ such that $\|v\|_{C^{1,\alpha}([\tau,\tau+1]\times\mathbb{R}^N)}\le C$ for $\tau>\tau_0$.

Then, for D is as in (ii), and for any $q>N\frac{(p-1)}{p}$ the following limit holds

$$\left\|\frac{v(\tau)-V_D}{V_D}\right\|_{\mathrm{L}^q(\mathbb{R}^N)} \xrightarrow[\tau\to\infty]{} 0\,. \tag{36}$$

Remark 2.1. We notice that condition (23) is equivalent (for $N\ge 6$) to $p\in(1,p_1)\cup(p_2,2)$ where p_1, p_2 are as in (24). Therefore, Theorem 5 not only extends the range of admissible values of p close to 2 from $(p_c,2)$ to $(p_2,2)$, but also covers a very singular range close to 1. In the case $2\le N<6$, the entire range $p\in(1,2)\setminus\{\frac{3}{2}\}$ is covered.

We explain the role of the assumptions of Theorem 5.

Remark 2.2. Assumption *(i)* is needed to guarantee both the boundedness of the entropy functional and the integrability of the difference of the solution v and some pseudo-Barenblatt profile V_D.

Assumptions *(ii)* and *(iii)* are needed to select the limit profile. In the range $p\le p_c$, it is unclear to which profile a solution v converges, even if a large time limit exists. In the good range $p_c<p<2$, the selection principle is guaranteed by the conservation of mass. Assumption *(ii)* plays a surrogate of this property. While this property may appear strange at first sight, we consider it natural as in the case of the classical fast diffusion equation it holds, see [9, Theorem 1]. On the other hand, assumption *(iii)* is the basis of the entropy method, as explained in Section 2.1.

Assumption *(iv)* is used to guarantee the existence of a limit profile. The main tool here is the Ascoli–Arzelá Theorem. While this may seem a lot to ask, let us attempt to convince the reader that this is natural. The first argument is that *(iv)* holds true in the good range $p_c<p<2$. In this case, the proof relates to the regularity properties of solutions to (CPLE), which can be easily transferred to solutions to (29). Notice that, for $p_c<p<2$, the change of variable is not singular nor degenerate in time. Indeed, a complete proof can be found in [3, Lemmata 2.2 and 2.3]. In the range $1<p\le p_c$, solutions to (CPLE) also enjoy good regularity properties (at least upon the notion of solutions we study), but it is impossible to transfer those properties to solutions to (29) since the change of variables is singular in this case. However, we consider it a technical detail, not a strong impediment. Indeed, in order to obtain *(iv)*, one should provide regularity estimates for solutions to (CPLE) that depend on the extinction time. We believe this to be true but could not find a relevant reference in the literature.

We conclude the consideration on Theorem 5 by commenting on the norms used in the convergence result (36).

Remark 2.3. Let us point out that for $p_c<p<2$ Theorem 5 seems to be stronger than Theorem 1, as instead of the L^∞-norm of the relative error $\frac{v(\tau,\cdot)-V_D(\cdot)}{V_D(\cdot)}$ we consider some L^q-norms of this quantity. Nonetheless, this is not exactly the case. In fact, for $p_D<p<2$ assumption *(i)* of Theorem 5 is much stronger than *(ii)* of Theorem 1 due to Lemma 9.2. By the comparison principle, we can infer that $\frac{v(\tau)-V_D}{V_D}$ has an integrable tail in this range. This is impossible for $p_c<p<2$ since the relative error is generally not integrable. On the other hand, for $p_c<p\le p_D$ hypothesis *(i)* of Theorem 5 is exactly *(ii)* in Theorem 1 written after the change of variable (28). In this range, we do not need to impose *(ii)–(iv)* since they are automatically verified along the flow. In turn, in the range $p_c<p\le p_D$, the convergence (36) yields a slight improvement compared to Theorem 1. We find it interesting to state the result precisely in the current way to stress that *(i)–(iv)* are the only ingredients needed for the entropy method to work in the entire range (23) apart from $p=3/2$.

Once one obtains the convergence result of (36), the natural question is whether the convergence holds with a rate. We answer this question with the following result.

Theorem 6. *Under the same assumptions of Theorem 5, assume furthermore that, for D is as in (ii) of Theorem 5, there exist $\varepsilon_0, \tau_0, \kappa > 0$ such that*

$$|\nabla v^{\gamma-1}(\tau,y)| \le \kappa \left(\varepsilon_0 + |\nabla V_D^{\gamma-1}(y)|\right) \qquad \text{for all } y \in \mathbb{R}^N\ \tau \ge \tau_0\,. \tag{37}$$

Then for any $q \in \left(N\frac{(p-1)}{p}, \infty\right]$ exists $\tau_1 = \tau_1(v_0,q) > 0$ and $\mathfrak{K} = \mathfrak{K}(v_0,N,p,q) > 0$ such that

$$\left\| \frac{v(\tau,\cdot) - V_D(\cdot)}{V_D(\cdot)} \right\|_{\mathrm{L}^q(\mathbb{R}^N)} \le e^{-\mathfrak{K}\tau} \qquad \text{for all } \tau \ge \tau_1. \tag{38}$$

Remark 2.4. The reason we assume (37) is mostly technical. We actually believe that (37) holds true under the conditions of Theorem 5 as it does hold at the level of equation (CPLE), see inequality (50). However, in the range $p < p_c$, we were not able to obtain estimates like (37) from properties of solutions to (CPLE) due to the lack of the regularity estimates that depend on the extinction time.

3. The role of parameter p

It should be stressed that the properties of the solutions to the p-Laplace Cauchy problem change in several special values of the parameter $p > 1$ and the whole picture is quite complex. This is reflected in different properties of initial data needed for proceeding with the relative entropy method for subranges of $p \in (1,2)$. In this section we give a global picture and summarize important properties of solutions to (CPLE) according to different regimes. Relevant special values of $p \in (1,2)$ are presented on the axis on Figure 1 (for $N \ge 7$). As their order changes when N varies, we summed it up by the end of the section.

The first special value is p_c defined in (1), which accounts for the threshold of conservation of mass. Indeed, for $p \in (p_c, 2)$ and for initial data which is nonnegative and integrable, the solution is positive and bounded for $x \in \mathbb{R}^N$ and $t > 0$, and the mass of the solution is conserved (i.e., identity (1) holds, see [20, 48] for a proof). Maybe even more importantly, Barenblatt solutions $\mathcal{B}_M(t,\cdot) \in \mathrm{L}^1(\mathbb{R}^N)$. In Theorem 1 there appears another special value, p_M, given by (8). Recall that $1 < p_c < p_M < 2$. This value is related to the boundedness of the entropy functional introduced in (33) and, at the same time, finite $|x|^{p'}$-moments for solutions to (CPLE). In fact, for $p \in (p_M, 2)$ solutions whose initial datum is in $\mathcal{X}_p$ have finite $|x|^{p'}$-moments. Note that this equips us with enough information to ensure basic needed properties of main relative entropy tools under no extra restriction on the data (cf. Lemmata 6.2, 6.4, and 6.6). To obtain the rate of convergence for $p \in (p_c, p_M]$ we need to impose more restrictions on the initial data (see Theorem 1 *(ii)*). To conclude this paragraph, let us mention another special value related to the entropy function: p_D from (88) used in Proposition 6.11. For $p > p_D$ the entropy functional is displacement convex and hence the optimal transportation approach of [3] leads to a quick conclusion on the convergence rates. Note that $p_D \ge p_M$ and this value does not play any role neither in our main reasoning (concluded in Sections 6.4 and 6.5), nor for the radial results of Theorems 2, 3, and 4, nor in the general result for $p < p_c$ provided in Theorems 5 and 6.

As already mentioned, the value $p = p_c$ is critical for the conservation of mass for solutions that live in $\mathrm{L}^1(\mathbb{R}^N)$. It is known that in this case solutions whose initial datum is in $\mathrm{L}^1(\mathbb{R}^N)$ still conserve mass, see for instance the recent survey [63]. However, the pseudo-Barenblatt solutions introduced in (18) are not in $\mathrm{L}^1(\mathbb{R}^N)$ in this case and do not attract solutions that live in $\mathrm{L}^1(\mathbb{R}^N)$. So, at least in the present paper and in this range, the conservation of mass does not play any role.

Let us now focus on the range $1 < p < p_c$. Here, not only the conservation of mass does not hold, but also for large class of data (i.e., $u_0 \in L^r(\mathbb{R}^N)$ with $r = n(2-p)/p$ and/or $u_0(x) \le c\mathcal{B}_{D,T}(0,\cdot)$ for some $c, D, T \ge 0$) the solutions extinct in finite time T (i.e., $u(t) = 0$, for all $t > T$), cf. [62, Chapter 11]. In the radial case the special role is played also by the Yamabe exponent p_Y from (22). We recall that $p_Y < p_c$. Its meaning for the gradient regularity of solutions and connection with the Yamabe flow is exposed around its definition and in Section 4.2. Another critical value is $p = \frac{3}{2}$, where the entropy functional $\mathcal{E}$ given by (33) is not well-defined. In such situation some other form of entropy would need to be employed. Note that $p_c > \frac{3}{2}$. Lastly, let us recall p_1 and p_2 from (24). These values are related to the integrability of the pseudo-Barenblatt solutions. At the same time, the condition (23) makes valid one of the main ingredients of the proof of Theorems 1, 5, and 6. This tool is the Hardy–Poincaré inequality (Proposition 6.9 provided in [24, Example 3.1]), which is used to relate the entropy $\mathcal{E}$ with the entropy production $\mathcal{I}$. In the complementary case (namely when (23) does not hold), one is equipped with an inequality of Hardy type provided in [24, Corollary 1.1].

Summing up all the special values of the parameter p, we need to take into account that

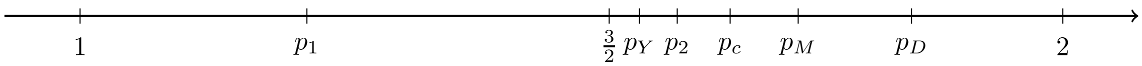

Figure 1: **Special values of parameter p when $N \geq 7$.** For $p > p_D$ the entropy is displacement convex, for $p > p_M$ Barenblatts have finite weighted $|x|^{p'}$-moments, for $p > p_c$ Barenblatt solutions are integrable and the mass is conserved, for $p \in (1, p_1) \cup (p_2, 2)$ difference of two Barenblatt solutions is integrable, for $p = \frac{3}{2}$ we have $\gamma = 0$ and the entropy functional is not defined.

1. if $N = 2$, then $1 = p_Y < p_c < \frac{3}{2} < p_M < p_D < 2$,
2. if $N = 3$, then $1 < p_Y < p_c = \frac{3}{2} < p_M < p_D < 2$,
3. if $N \in \{4, 5\}$, then $1 < p_Y < \frac{3}{2} < p_c < p_M < p_D < 2$,
4. if $N = 6$, then $1 < p_1 < \frac{3}{2} = p_Y = p_2 < p_c < p_M < p_D < 2$,
5. if $N \geq 7$, then $1 < p_1 < \frac{3}{2} < p_Y < p_2 < p_c < p_M < p_D < 2$.

4. Preliminary information

4.1. Notation

Following a usual custom, we denote by c a general positive constant. Different occurrences from line to line will be still denoted by c, while special occurrences will be denoted by $c_1, c_2, \tilde{c}$ or similar. Relevant dependencies on parameters will be emphasized using parentheses, i.e., $c = c(p, M)$ means that c depends on p and M. We define $(u)_+ := \max\{0, u\}$.

We also recall the definitions of the constants b_1 and b_2 which appear in the definition of the Barenblatt function (2), see also [20] for more information:

$$b_2 := \tfrac{2-p}{p} \left(p - N(2-p)\right)^{-\frac{1}{p-1}} , \tag{39}$$

while b_1 is such a positive constant that

$$\int_{\mathbb{R}^N} \left(b_1 + b_2 |x|^{p'}\right)^{-\frac{p-1}{2-p}} = 1 . \tag{40}$$

4.2. Existence and uniqueness

Let us first introduce the concept of nonnegative weak solutions that we shall use throughout the present work and comment on their well-posedness.

Definition 4.1. *We say that u is a non-negative weak solution to Problem* (CPLE) *on* $(0, \infty) \times \mathbb{R}^N$ *for* $1 < p < 2$ *with nonnegative initial data* $u_0 \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N)$ *if* $u \in \mathrm{L}^p((0, \infty); W^{1,p}_{\mathrm{loc}}(\mathbb{R}^N)) \cap C((0, \infty); \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N))$ *and*

$$\int_{\mathbb{R}^N} u(s, x)\phi(s, x)\,\mathrm{d}x = \int_{\mathbb{R}^N} u(t, x)\phi(t, x)\,\mathrm{d}x + \int_s^t \int_{\mathbb{R}^N} \big(- u(\tau, x)\partial_\tau \phi(\tau, x) + |\nabla u(\tau, x)|^{p-2} \nabla u(\tau, x) \cdot \nabla \phi(\tau, x)\big)\,\mathrm{d}x\,\mathrm{d}\tau\,, \tag{41}$$

for all $t > s > 0$ and for all functions $\phi \in C^\infty([0, +\infty) \times \mathbb{R}^N)$, such that the support of the maps $x \mapsto \phi(t, x)$ is compact for any $t \geq 0$. The initial data is attained in the following sense

$$\lim_{t \to 0} \int_{\mathbb{R}^N} u(t, x)\varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}^N} u_0(x)\varphi(x)\,\mathrm{d}x \quad \forall \varphi \in C^\infty_c(\mathbb{R}^N)\,.$$

When $u_t \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N)$ and u is bounded, then u is also a *distributional* solution, in the sense that equation (CPLE) holds in $\mathcal{D}'((0, \infty) \times \mathbb{R}^N)$, as proven in [37, Lemma I.1.2]. The space of test functions can be enlarged by less regular functions decaying fast enough for $|x| \to \infty$. The existence for (CPLE) has been settled in [37, Section III], in the whole range $1 < p < 2$ and for initial datum $u_0 \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N)$. Notice, however, that the concept of solution used in [37, I.1] is different from the one in Definition 4.1 (it is also weaker in some sense, our definition is similar to the one used in [37, III.4.3-III.4.5] or to one in [37, Theorem III.9.1]). The main reason is the singularity of (CPLE) when p is small. Indeed, if $|\nabla u|$ is not a function from $\mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^N)$, then the solution has to be interpreted in an appropriate way, see the discussion in [37, Introduction]. However, we remark that in the present work, we shall focus on solutions that are locally bounded and have $u_t \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N)$, which will be justified later. Therefore, the weaker notion used in [37] is equivalent to the notion used here. The uniqueness can be proven in the case when initial data are in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N)$ and solutions satisfy a particular inequality related to time derivatives, see [37, Section II]. These assumptions are satisfied for solutions to the Cauchy problem, cf. [37, Introduction]. Let us conclude this paragraph by saying a few words about the notion of solution used for problem (29). For $p_c \le p < 2$, the change of variables introduced in (27)-(28) is not singular. Therefore, to obtain a definition of weak solution for (29), it is enough to apply those variable changes to the solution concept used in Definition 4.1. The theory previously exposed guarantees the existence and uniqueness of weak solutions. In the case $1 < p < p_c$, the matter is more complicated, since the change of variables (27)-(28) is singular. One way to avoid the singularity and to ensure that the exposed theory also apply in this range, is to apply that change of variables to data that trapped between two pseudo-Barenblatt solutions with the same extinction time $T > 0$. Those are precisely the class of solution that we consider in Theorems 6 and 5 (notice that assumption *(i)* guarantees it). For the existence result for another notion of solutions, namely p-caloric ones, we refer to a recent paper [43].

4.3. Boundedness and regularity of solutions

Since the literature concerning the regularity properties of solutions to equations like (CPLE) is abundant, we do not aim to describe the state of art exhaustively. Instead, we shall restrict ourselves to presenting only the background needed for our study.

Here we shall mainly focus on the concept of local weak solution (not necessarily non-negative) which differ from our definition above by assumptions on the integrability properties of u, namely for a solution defined on $(0,T)\times\Omega$ one asks typically that $u \in \mathrm{L}^\infty_{\mathrm{loc}}((0,T);\mathrm{L}^2_{\mathrm{loc}}(\Omega))$ and $|\nabla u| \in \mathrm{L}^p_{\mathrm{loc}}((0,T);\mathrm{L}^p_{\mathrm{loc}}(\Omega))$. To the best of our knowledge, these kind of solutions have been studied first in [31] where continuity of ∇u has been proven, with an explicit modulus of continuity, for $p > \max\{1, p_Y\} = \max\{1, \frac{2N}{N+2}\}$. In the same range of p, the Hölder continuity of the gradient has been obtained in [33] (with some mistakes in the computations, as it was pointed out in [68], which have been fully solved in [32]), see also the review [65]. The threshold value $p_Y = \frac{2N}{N+1}$ is sharp for the gradient regularity under no extra assumptions on the solution. Indeed, when $p \le p_Y$ weak solutions are not bounded, see the discussion [36, Section 21.3]. However, in the whole range $1 < p < 2$ it is known, that *bounded* weak solutions are Hölder continuous, see [67, 66]. In our case, it is known that solutions to (CPLE) are bounded provided the initial datum $u_0 \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R}^N)$ for $q > N\frac{(2-p)}{p}$, see for instance [37, Section III] and [17, Theorem 2.1]. In conclusion, when $p \in (p_Y, 2)$ and the initial datum is integrable enough, weak solutions to (CPLE) are bounded and, therefore, the function $(t,x) \mapsto u(t,x)$ is $C^{1,\alpha}_{\mathrm{loc}}((0,\infty)\times\mathbb{R}^N)$. However, when $p < p_c$, the coefficient α may depend on the function itself other than on p and N.

In the present paper, we shall also consider the case $p \le p_Y$, so let us comment on how to obtain the $C^{1,\alpha}_{\mathrm{loc}}(\mathbb{R}^N)$ regularity for solutions in this case. Let us stress, however, that the following reasoning holds in the whole range $1 < p < 2$. The main idea is to use the concept of viscosity solutions (see [47] for a precise definition). In our case, when a weak solution is continuous, then it is also a viscosity solution; see, for instance , [47, 60, 40, 39]. Since we shall consider only bounded weak solutions (which are Hölder continuous), the above discussion proves that solutions to problem (CPLE) are, indeed, viscosity solutions. The main advantage of employing this notion of solutions is that, in the last decade, there has been a growing interest in obtaining regularity results for viscosity solutions to equations related to (CPLE), see for instance [40] for a detailed bibliography. For what concerns our investigation, the needed result is [46, Theorem 1.1], where the authors prove the function $(t,x) \mapsto u(t,x)$ is $C^{1,\alpha}_{\mathrm{loc}}((0,\infty)\times\mathbb{R}^N)$ for some $\alpha > 0$, which depends on p, N and the solution itself.

Several considerations on the integrability of time derivative u_t and of $\Delta_p u = \mathrm{div}\left(|\nabla u|^{p-2}\nabla u\right)$ are in order. It is known that for a continuous weak solution both u_t and $\mathrm{div}\left(|\nabla u|^{p-2}\nabla u\right)$ belong to $\mathrm{L}^2_{\mathrm{loc}}(Q)$, where Q is space-time cylinder. Furthermore, equation (CPLE) is satisfied almost everywhere in (t,x). We refer to [17, Corollary] and to [39]. We also remark that more is known about the integrability properties of derivatives of $(t,x) \mapsto |\nabla u|^{\frac{p-2+s}{2}}\nabla u$ (where s is chosen appropriately), for which we refer to [39, 40]. When it comes to explicit estimates of the continuity of the gradient of solutions, we refer to [50, 51, 52, 53]. In those results, the authors obtain the $C^{1,\alpha}$-regularity by exploiting very interesting connections with non-linear potential estimates. While they are valid mainly when $p > p_c$, those results also apply when the equation (CPLE) has a measure as right-hand-side. We refer to [54] for a general overview.

Since the uniform convergence in relative error is related to Harnack inequalities, let us conclude this subsection with some considerations on them. The problem of obtaining a precise form of those inequalities has been a longstanding quest. Indeed, in this nonlinear setting, the intrinsic cylinders depend on the solution itself, showing several differences between the case $p_c < p < 2$ and $1 < p \le p_c$. In the good range, we refer to the paper [34], while in the whole range $1 < p < 2$, it has been proven in [17]. The Harnack inequalities considered in those two papers are valid for local solutions, i.e., no assumption on boundary data is made. For boundary Harnack inequalities, we refer to [55]. Nowadays, several related results are available, see the monograph [35] and references therein.

4.4. Comparison principles

By the comparison principle, we mean that, in some sense, ordered data generate ordered solutions at all times. Such results for solutions to (CPLE) seem to be well known by experts in the field, cf. [41, Sections 3 and 4] or [10, Section 4.5]. Nonetheless, we could not find references with complete proofs in the case of the Cauchy problem within the whole range $1<p<2$. One of the main difficulties in the proofs of comparison principles is that, at least when $p<p_c$, in general solutions are not integrable. Thus, a priori the quantity $(u_1-u_2)_+$ (where u_1 and u_2 are two solutions to (CPLE)) cannot be used as a test function. However, in our case when solutions are regular and bounded, one is equipped with the following two comparison principles. We decided to include them with the sketches of the proofs for completeness.

The first comparison principle we present reads: for u_1 and u_2 being two solutions (that are regular enough, cf. Section 4.2) with initial data $u_{1,0}$ and $u_{2,0}$, respectively, we have

$$\text{if} \quad u_{1,0}\le u_{2,0}\,, \qquad \text{then} \quad u_1(t,x)\le u_2(t,x) \quad \forall\, t>0\ \forall\, x\in\mathbb{R}^N\,. \tag{42}$$

It can be proven via the construction inspired by [37, Chapter 3] where the authors constructed solutions to (CPLE) by approximation by solutions to the Dirichlet problem. The comparison principle for the Dirichlet problem goes back to [58]; for a more recent proof, we refer to [36, Chapter 7, Corollary 1.1] (cf. also references therein). We notice that for the Dirichlet problem, there is no restriction on $p>1$. We have u_1 and u_2 obtained by approximation with solutions to the Dirichlet problem with initial data $u_{1,0}\phi_R$ and $u_{2,0}\phi_R$ (where ϕ_R is a cut-off function supported in the ball of radius $R>0$). The approximate solutions are ordered on B_{2R}. This relation holds in the limit $R\to\infty$.

One is not deprived from comparison results in the class of less regular solutions. In fact the following $\mathrm{L}^1_{\rm loc}$-*comparison principle* holds: let u_1 and u_2 be two solutions (from the class the class $\Sigma^\star$, see [37, Chapter II]):

$$\text{if} \quad (u_1-u_2)_+\to 0 \text{ in } \mathrm{L}^1_{\rm loc}(\mathbb{R}^N) \ \text{ as } \ t\to 0\,, \qquad \text{then} \quad u_1(t,x)\le u_2(t,x) \quad \forall\, t>0\ \forall\, x\in\mathbb{R}^N\,. \tag{43}$$

The main advantage of this local comparison is that it avoids using global integrability, even if it assumes an order in $\mathrm{L}^1_{\rm loc}(\mathbb{R}^N)$. It might be proven by following the lines of [37, Proposition II.3.1and Theorem II.1.1]. Indeed, a careful inspection of the proof of [37, Proposition II.3.1] shows that under the assumption of (43) one gets that, for all $T>0$, $(u_1(t,x)-u_2(t,x))_+\in \mathrm{L}^\infty\big((0,T);\mathrm{L}^q_{\rm loc}(\mathbb{R}^N)\big)$ for all $q\in[1,\infty)$, and there exists $C=C(N,p,q)>0$ such that for all $R>0$, $t\in(0,T)$ and $\sigma\in(0,1)$ one has

$$\int_{B_R}\Big(u_1(t,x)-u_2(t,x)\Big)_+^q\,\mathrm{d}x \le \frac{C}{(\sigma\,R)^p}\int_0^t\int_{B_{(1+\sigma)R}}\Big(u_1(t,x)-u_2(t,x)\Big)_+^{q+p-2}\,\mathrm{d}x\,\mathrm{d}t\,. \tag{44}$$

Indeed, we also notice that, while originally inequality (44) is stated for the absolute value of the difference (i.e. $|u_1-u_2|$), its proof is done for the positive part of u_1-u_2. Once (44) is obtained, using the same argument as in the proof of [37, Theorem II.1.1, p. 257] one obtains that, for every $q>1$, all $t\in(0,T)$ and $C=C(N,p,q)$ independent of $R>0$:

$$\int_{B_R}\Big(u_1(t,x)-u_2(t,x)\Big)_+^q\,\mathrm{d}x \le C\,t^{\frac{q}{2-p}}\,R^{N-q\frac{p}{2-p}}\,.$$

By choosing q so large such that $qp/(2-p)>N$, one obtains in the limit $R\to 0$ that $\int_{\mathbb{R}^N}(u_1(t,x)-u_2(t,x))_+^q\,\mathrm{d}x=0$ for all $t>0$. Consequently, $u_1(t,x)\le u(t_2,x)$ for all $t>0$ and $x\in\mathbb{R}^N$. We acknowledge that, despite the comparison principle, it has not been stated in the form (43) in [37], it has been used in this form in [37, Proposition III.7.1]. Hence, we do not claim any originality for the above result.

5. Convergence in relative error under a priori convergence in Lebesgue space

The goal of this section is to obtain an explicit convergence rate towards the Barenblatt profile in uniform relative error, provided that we know a priori a convergence rate in a weaker norm. In what follows we shall use the L^1-norm. One can prove a counterpart of this result involving L^q-norm with $1\le q\le\infty$, by interpolation arguments.

Theorem 7. *Let $N \geq 1$, $p_c < p < 2$, $0 \leq u_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap \mathcal{X}_p$ and $M = \int_{\mathbb{R}^N} u_0(x)\,\mathrm{d}x > 0$. Assume u is a solution to* (CPLE) *with initial datum u_0. Suppose that for some $\widetilde{T} > 0$, $\widetilde{K} > 0$ and $N+1 \geq \nu > 0$, we have that*

$$\|u(t,\cdot) - \mathcal{B}_M(t,\cdot)\|_{\mathrm{L}^1(\mathbb{R}^N)} \leq \widetilde{K}\, t^{-\nu} \qquad \forall\, t \geq \widetilde{T}\,. \tag{45}$$

Then there exist $K_\star = K_\star(p, N, M, \|u_0\|_{\mathcal{X}_p}, \widetilde{K}) > 0$ and $T_\star = T_\star(p, N, M, \|u_0\|_{\mathcal{X}_p}, \widetilde{T}) > 0$ such that we have

$$\left\| \frac{u(t,\cdot) - \mathcal{B}_M(t,\cdot)}{\mathcal{B}_M(t,\cdot)} \right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \leq K_\star\, t^{-\frac{\nu(2-p)}{N+1}} \qquad \forall\, t \geq T_\star\,. \tag{46}$$

As we shall see, the strategy of the proof of the above theorem is to consider separately two different regions in the (t,x) plane: *inner cylinders*, i.e. $\{|x| \leq C\,t^\beta\}$ for a constant $C > 0$, and *outer cylinders*, i.e. $\{|x| \geq C\,t^\beta\}$. Assumption (45) plays a major rôle in the inner cylinders, while in the case of the outer cylinders it is the global Harnack principle (47) to imply the wanted result.

5.1. Properties of solutions to (CPLE) *for $p_c < p < 2$ and initial datum $u_0 \in \mathcal{X}_p$*

We stress that in a significant part of our paper (i.e., Sections 5, 6, and 7) we shall consider the exponent $p \in (p_c, 2)$. We notice that within this range of p the results of [20, Theorem 1.1 and 1.3] imply the following consequences having a fundamental meaning in our reasoning.

(i) If $0 \leq u_0 \in \mathcal{X}_p \setminus \{0\}$ (for $\mathcal{X}_p$ being defined in (7)), then for any $t_0 > 0$ there exist (explicit) constants τ_1, M_1, τ_2, M_2 such that for all $x \in \mathbb{R}^N$ and $t \geq t_0$ the following upper and lower bounds hold true

$$\mathcal{B}_{M_1}(t - \tau_1, x) \;\leq\; u(t,x) \;\leq\; \mathcal{B}_{M_2}(t + \tau_2, x)\,; \tag{47}$$

The above inequality is also known as the *global Harnack principle*.

(ii) Let $M = \|u_0\|_{\mathrm{L}^1(\mathbb{R}^N)} > 0$ and $u_0 \geq 0$. Then

$$\lim_{t\to\infty} \left\| \frac{u(t,\cdot)}{\mathcal{B}_M(t,\cdot)} - 1 \right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} = 0 \tag{48}$$

if and only if $u_0 \in \mathcal{X}_p \setminus \{0\}$.

(iii) If $0 \leq u_0 \in \mathrm{L}^1(\mathbb{R}^N)$, then L^∞-norm of the gradient decays in time. More precisely, there exists a constant $c_1 = c_1(p,N) > 0$ such that

$$\|\nabla u(t,\cdot)\|_{L^\infty(\mathbb{R}^N)} \leq c_1\, \frac{\|u_0\|^{2\beta}_{L^1(\mathbb{R}^N)}}{t^{(N+1)\beta}} \qquad \text{for any } t > 0\,. \tag{49}$$

(iv) If $0 \leq u_0 \in \mathcal{X}_p$, then we can say more about the spacial decay of the gradient. More precisely, there exists a constant $c_2 = c_2(N,p) > 0$ such that

$$|\nabla u(t,x)| \leq c_2\, \frac{\|u_0\|^{2\beta}_{\mathrm{L}^1(\mathbb{R}^N)} + \|u_0\|^{2\beta}_{\mathcal{X}_p} + t^{\frac{2\beta}{2-p}}}{(1+|x|)^{\frac{2}{2-p}}\, t^{(N+1)\beta}} \qquad \text{for any } x \in \mathbb{R}^N \text{ and } t > 0\,. \tag{50}$$

5.2. Convergence in outer cylinders

Proposition 5.1. *Let $N \geq 1$, $p_c < p < 2$, $0 \leq u_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap \mathcal{X}_p$ and $M = \int_{\mathbb{R}^N} u_0(x)\,\mathrm{d}x > 0$. Assume u is a weak solution to Problem* (CPLE) *with initial datum u_0. Then for any $\varepsilon \in (0,1)$ there exists $\underline{T}(\varepsilon) > 0$ and $\underline{\varrho}(\varepsilon) > 0$ such that then*

$$u(t,x) \geq (1-\varepsilon)\,\mathcal{B}_M(t,x) \quad \forall\, |x| \geq \underline{\varrho}(\varepsilon)\, t^\beta \quad \forall t > \underline{T}(\varepsilon)\,.$$

The strategy of the proof of the above proposition follows closely the proof of [15, Proposition 4.6].

Proof. By inequality (47), for $t \geq t_0 = 1$, we have $u(t,x) \geq \mathcal{B}_{M_1}(t-\tau_1, x)$. By integrating (47), we find as well that $M_1 \leq M$. If $M_1 = M$, then we conclude that $u = \mathcal{B}_M$ and the proposition is proven. Therefore restrict our attention to the case $M_1 < M$. Let us define $\underline{\varepsilon} \in (0,1)$ such that $(1-\underline{\varepsilon})M^\beta = M_1^\beta$ for β from (2) and let $0 < \varepsilon < \underline{\varepsilon}$. In order to prove the claim we need to prove that for $|x|$ and t large enough it holds

$$\frac{\mathcal{B}_{M_1}(t-\tau_1,x)}{\mathcal{B}_M(t,x)} \geq 1-\varepsilon\,. \tag{51}$$

Let us notice that for b_1, b_2 from (40) and (39), respectively, it holds that

$$\frac{\mathcal{B}_{M_1}(t-\tau_1,x)}{\mathcal{B}_M(t,x)} = \left(\frac{t-\tau_1}{t}\right)^{\frac{1}{2-p}} \left(\frac{b_1\left(\frac{t}{M^{2-p}}\right)^{\beta p'}}{b_1\left(\frac{t-\tau_1}{M_1{}^{2-p}}\right)^{\beta p'}}\right)^{\frac{p-1}{2-p}} \left(\frac{1+\frac{b_2}{b_1}M^{(2-p)\beta p'}|x|^{p'}t^{-\beta p'}}{1+\frac{b_2}{b_1}M_1{}^{(2-p)\beta p'}|x|^{p'}(t-\tau_1)^{-\beta p'}}\right)^{\frac{p-1}{2-p}},$$

where $p' = p/(p-1)$. Upon setting

$$\eta(t) := \left(\frac{t}{t-\tau_1}\right)^\beta, \qquad s(t,x) = |x|^{p'}t^{-\beta p'}, \qquad \text{and} \qquad c = \frac{b_2}{b_1}M^{(2-p)\beta p'} \tag{52}$$

and recalling that $\beta = (p - N(2-p))^{-1}$, the left-hand-side of (51) becomes

$$\frac{\mathcal{B}_{M_1}(t-\tau_1,x)}{\mathcal{B}_M(t,x)} = \left(\frac{t}{t-\tau_1}\right)^{N\beta}\left(\frac{1+c\,s(t,x)}{(1-\overline{\varepsilon})^{\frac{p-2}{p-1}}+c\,\eta^{p'}(t)\,s(t,x)}\right)^{\frac{p-1}{2-p}} = \eta^N(t)\left(\frac{1+cs(t,x)}{(1-\overline{\varepsilon})^{\frac{p-2}{p-1}}+c\eta^{p'}(t)s(t,x)}\right)^{\frac{p-1}{2-p}}.$$

Therefore, inequality (51) is equivalent to

$$s(t,x) \geq \frac{\left(\frac{1-\varepsilon}{1-\underline{\varepsilon}}\right)^{\frac{2-p}{p-1}}\eta^{-N\frac{2-p}{p-1}}(t)-1}{c\left(1-\eta^{\frac{1}{\beta(p-1)}}(t)(1-\varepsilon)^{\frac{2-p}{p-1}}\right)} =: \underline{s}(t,\varepsilon)\,, \tag{53}$$

provided that $1 > \eta^{\frac{1}{\beta(p-1)}}(t)(1-\varepsilon)^{\frac{2-p}{p-1}}$. We restrict our attention to $t > \underline{T}(\varepsilon)$, where

$$\underline{T}(\varepsilon) := \max\left\{\frac{\tau_1}{1-(1-\varepsilon)^{2-p}}\,, t_1\,, t_0\right\} \tag{54}$$

and $t_1 > 0$ is such that $\eta(t_1)^{\frac{1}{\beta(p-1)}} = 2(1+(1-\varepsilon)^{\frac{2-p}{p-1}})^{-1}$. Observe that for $t > \frac{\tau_1}{1-(1-\varepsilon)^{2-p}}$, we have $\eta^{\frac{1}{\beta(p-1)}}(t)(1-\varepsilon)^{\frac{2-p}{p-1}} < 1$. Then (51) will follow from (53) and $t > \underline{T}(\varepsilon)$. Inequality (53) holds true as long as $s(t,x) \geq \underline{\varrho}^{p'}(\varepsilon)$ for $\underline{\varrho}(\varepsilon)$ defined below. Indeed, since $\eta \geq 1$ is decreasing and, for $t \geq t_1$, it holds that $\eta(t) \leq \eta(t_1)$, we have

$$\underline{s}(t,\varepsilon) \leq \frac{(1-\underline{\varepsilon})^{\frac{p-2}{p-1}}(1-\varepsilon)^{\frac{2-p}{p-1}}}{c\left(1-(1-\varepsilon)^{\frac{2-p}{p-1}}\eta^{\frac{1}{\beta(p-1)}}(t_1)\right)} \leq \frac{(1-\underline{\varepsilon})^{\frac{p-2}{p-1}}(1-\varepsilon)^{\frac{2-p}{p-1}}}{c\left(1-\frac{2(1-\varepsilon)^{\frac{2-p}{p-1}}}{1+(1-\varepsilon)^{\frac{2-p}{p-1}}}\right)} = \frac{(1-\varepsilon)^{\frac{2-p}{p-1}}(1+(1-\varepsilon)^{\frac{2-p}{p-1}})}{c\,(1-\underline{\varepsilon})^{\frac{2-p}{p-1}}\left(1-(1-\varepsilon)^{\frac{2-p}{p-1}}\right)} =: \underline{\varrho}^{p'}(\varepsilon)\,. \tag{55}$$

□

Proposition 5.2. *Let $N \geq 1$, $p_c < p < 2$, $0 \leq u_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap \mathcal{X}_p$ and $M = \int_{\mathbb{R}^N} u_0(x)\,\mathrm{d}x > 0$. Assume u is a weak solution to Problem* (CPLE) *with initial datum u_0. Then for any $\varepsilon \in (0,1)$ there exists $\overline{T}(\varepsilon) > 0$ and $\overline{\varrho}(\varepsilon) > 0$ such that*

$$u(t,x) \leq (1+\varepsilon)\,\mathcal{B}_M(t,x) \quad \forall\,|x| \geq \overline{\varrho}(\varepsilon)\,t^\beta \quad \forall\,t \geq \overline{T}(\varepsilon)\,.$$

Proof. We shall proceed as in the proof of Proposition 5.1. By inequality (47), for $t \geq t_0 = 1$, we have $u(t,x) \leq \mathcal{B}_{M_2}(t+\tau_2,x)$. By integrating (47), we find as well that $M \leq M_2$. As previously, we can assume $M < M_2$. Let us define $\overline{\varepsilon} > 0$ such that $(1+\overline{\varepsilon})M^\beta = M_2^\beta$ and let $0 < \varepsilon < \min\{\overline{\varepsilon}, 1\}$. In order to prove the claim we need to prove that for $|x|$ and t large enough it holds

$$\frac{\mathcal{B}_{M_2}(t+\tau_2,x)}{\mathcal{B}_M(t,x)} \leq 1+\varepsilon\,. \tag{56}$$

Let us notice that for $\eta(t) := \left(\frac{t}{t+\tau_2}\right)^\beta$, s, and c as in (52), we have

$$\frac{\mathcal{B}_{M_2}(t+\tau_2,x)}{\mathcal{B}_M(t,x)} = \eta^N(t)\left(\frac{1+cs(t,x)}{(1+\overline{\varepsilon})^{\frac{p-2}{p-1}} + c\eta^{p'}(t)s(t,x)}\right)^{\frac{p-1}{2-p}},$$

and inequality (56) is equivalent to

$$s(t,x) \geq \frac{1-(1+\overline{\varepsilon})^{\frac{p-2}{p-1}}(1+\varepsilon)^{\frac{2-p}{p-1}}\eta^{-N\frac{2-p}{p-1}}(t)}{c\left(\eta^{\frac{1}{\beta(p-1)}}(t)(1+\varepsilon)^{\frac{2-p}{p-1}}-1\right)} =: \overline{s}(t,\varepsilon)\,, \tag{57}$$

provided $\eta^{\frac{1}{\beta(p-1)}}(t)(1+\varepsilon)^{\frac{2-p}{p-1}} > 1$. We restrict our attention to $t > \overline{T}(\varepsilon)$, where

$$\overline{T}(\varepsilon) := \max\left\{\frac{\tau_2}{(1+\varepsilon)^{2-p}-1},\, t_2,\, t_0\right\} \tag{58}$$

and $t_2 > 0$ is such that $\eta(t_2)^{\frac{1}{\beta(p-1)}} = 2(1+(1+\varepsilon)^{\frac{2-p}{p-1}})^{-1}$. We observe that for $t > \frac{\tau}{(1+\varepsilon)^{2-p}-1}$, we have $\eta^{\frac{1}{\beta(p-1)}}(t)(1+\varepsilon)^{\frac{2-p}{p-1}} > 1$. Then (56) will follow from (57) and $t > \overline{T}(\varepsilon)$. Inequality (57) holds true as long as $s(t,x) \geq \overline{\varrho}^{p'}(\varepsilon)$ (for $\overline{\varrho}$ defined below). Indeed, since $\eta \leq 1$ is increasing, and, for $t \geq t_2$, it holds that $\eta(t) \leq \eta(t_2)$, we find that

$$\overline{s}(t,\varepsilon) \leq \frac{1}{c\left(1-(1+\varepsilon)^{\frac{2-p}{p-1}}\eta^{\frac{1}{\beta(p-1)}}(t_2)\right)} = \frac{1}{c\left(1-\frac{2}{1+(1+\varepsilon)^{\frac{2-p}{p-1}}}\right)} = \frac{1+(1+\varepsilon)^{\frac{2-p}{p-1}}}{c\left((1+\varepsilon)^{\frac{2-p}{p-1}}-1\right)} =: \overline{\varrho}^{p'}(\varepsilon)\,. \tag{59}$$

□

5.3. Convergence in inner cylinders

Proposition 5.3. *Let $N \geq 1$, $p_c < p < 2$, $0 \leq u_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap \mathcal{X}_p$ and $M = \int_{\mathbb{R}^N} u_0(x)\,\mathrm{d}x > 0$. Assume u is a weak solution to Problem* (CPLE) *with initial datum u_0. Suppose that for some $\widetilde{T} > 0$ and $\widetilde{K} > 0$ inequality* (45) *holds. Then there exist $\varrho_0 = \varrho_0(p,N,M) > 0$, $\overline{K}_\star = \overline{K}_\star(p,N,M,\widetilde{K}) > 0$ such that for any $\varrho \geq \varrho_0$ we have*

$$\left|\frac{u(t,x)-\mathcal{B}_M(t,x)}{\mathcal{B}_M(t,x)}\right| \leq \overline{K}_\star\,\varrho^{\frac{p}{2-p}}\,t^{-\frac{\nu}{N+1}} \qquad \forall\,|x| \leq \varrho\,t^\beta \quad \forall\,t \geq \widetilde{T}\,, \tag{60}$$

where ν and $\widetilde{T}$ are as in (45).

Proof. For any $t > 0$ and $|x| \leq \varrho\,t^\beta$, the relative error satisfies the following inequality

$$\left|\frac{u(t,x)-\mathcal{B}_M(t,x)}{\mathcal{B}_M(t,x)}\right| \leq \|u(t,\cdot)-\mathcal{B}_M(t,\cdot)\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \sup_{\{|x|\leq\varrho\,t^\beta\}} \frac{1}{\mathcal{B}_M(t,x)}\,.$$

Since for any $t > 0$, the function $|x| \mapsto \mathcal{B}_M(t,|x|)$ is decreasing, we find that the supremum in the above inequality is attained at $|x| = \varrho\,t^\beta$. Using the expression of the Barenblatt profile (2) and b_1, b_2 from (40) and (39), respectively, a simple computation shows that, for any $\varrho \geq \varrho_0 := b_1^{\frac{p-1}{p}}/M^{\beta\,(2-p)}$, we have that

$$\sup_{\{|x|\leq\varrho\,t^\beta\}} \frac{1}{\mathcal{B}_M(t,x)} \leq \varrho^{\frac{p}{2-p}}\,t^{N\beta}\,(1+b_2)^{\frac{p-1}{2-p}} = C(p,N)\,\varrho^{\frac{p}{2-p}}\,t^{N\beta}\,.$$

Combining the two estimates we find

$$\left|\frac{u(t,x)-\mathcal{B}_M(t,x)}{\mathcal{B}_M(t,x)}\right| \leq C\,\varrho^{\frac{p}{2-p}}\,t^{N\beta}\,\|u(t,\cdot)-\mathcal{B}_M(t,\cdot)\|_{\mathrm{L}^\infty(\mathbb{R}^N)}\,.$$

Before continuing, let us recall the Gagliardo–Nirenberg inequality

$$\|f\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \le C_N \|\nabla f\|_{\mathrm{L}^\infty(\mathbb{R}^N)}^{\frac{N}{N+1}} \|f\|_{\mathrm{L}^1(\mathbb{R}^N)}^{\frac{1}{N+1}},$$

holding for any f regular enough for which all the involved quantities are finite, see [15, Lemma 3.5]. Combining the above inequality with (45), the time decay of gradient of solutions (49), and the triangle inequality, we get that for any $t \ge \widetilde{T}$:

$$\begin{aligned} t^{N\beta} \|u(t,\cdot) - \mathcal{B}_M(t,\cdot)\|_{\mathrm{L}^\infty(\mathbb{R}^N)} &\le C_N\, t^{N\beta} \|\nabla u(t,\cdot) - \nabla \mathcal{B}_M(t,\cdot)\|_{\mathrm{L}^\infty(\mathbb{R}^N)}^{\frac{N}{N+1}} \|u(t,\cdot) - \mathcal{B}_M(t,\cdot)\|_{\mathrm{L}^1(\mathbb{R}^N)}^{\frac{1}{N+1}} \\ &\le C_N\, t^{N\beta} c_1^{\frac{N}{N-1}} \left(2\, M t^{-(N+1)\beta}\right)^{\frac{N}{N+1}} \left(\widetilde{K}\, t^{-\nu}\right)^{\frac{1}{N+1}} \\ &\le C_\star\, t^{N\beta - \frac{N}{N+1}(N+1)\beta - \frac{\nu}{N+1}} \le C_\star\, t^{-\frac{\nu}{N+1}}, \end{aligned} \tag{61}$$

where $C_\star > 0$ is a constant depending on $N, p,$ M and $\widetilde{K}$. Combining all the above estimates, we can pick $\overline{K}_\star := C(p,N)\, C_\star$. □

5.4. Proof of convergence in relative error under a priori convergence in Lebesgue's space

After establishing the convergence inside and outside the cylinders in the last two sections, the only remained task is to link them.

Proof of Theorem 7. From Propositions 5.1, 5.2 and 5.3 we know that for fixed $\varepsilon \in (0, 1/2)$, there exist

$$T(\varepsilon) = \max\{\underline{T}(\varepsilon), \overline{T}(\varepsilon), \widetilde{T}\} > 0 \qquad \text{and} \qquad \varrho(\varepsilon) = \max\{\underline{\varrho}(\varepsilon), \overline{\varrho}(\varepsilon)\} > 0\,,$$

where $\underline{T}, \overline{T}$ are defined in (54) and (58), respectively, and $\underline{\varrho}, \overline{\varrho}$ are defined in (55) and (59), respectively, such that

$$\left|\frac{u(t,x) - \mathcal{B}_M(t,x)}{\mathcal{B}_M(t,x)}\right| \le \varepsilon \qquad \forall |x| \ge \varrho(\varepsilon)\, t^\beta \quad \forall t \ge T(\varepsilon)\,. \tag{62}$$

In the same way, using (60), we obtain that for $t \ge \widetilde{T}$ it holds

$$\left|\frac{u(t,x) - \mathcal{B}_M(t,x)}{\mathcal{B}_M(t,x)}\right| \le \overline{K}_\star\, \varrho(\varepsilon)^{\frac{p}{2-p}}\, t^{-\frac{\nu}{N+1}} \qquad \forall |x| \le 2\, \varrho(\varepsilon)\, t^\beta\,.$$

By a simple computation one finds that there exists a constant $\kappa_0(p) > 0$, such that $\varrho(\varepsilon) \le \kappa_0(p)\, \varepsilon^{\frac{1-p}{p}}$. Therefore, we have that

$$\overline{K}_\star\, \varrho(\varepsilon)^{\frac{p}{2-p}}\, t^{-\frac{\nu}{N+1}} \le \varepsilon \qquad \text{for} \quad t \ge C_p\, \varepsilon^{-\frac{1}{2-p}\frac{N+1}{\nu}}\,,$$

where the constant $C_p = C_p(N, p, \overline{K}_\star) > 0$ is independent of ε. Therefore

$$\left\|\frac{u(t,\cdot) - \mathcal{B}_M(t,\cdot)}{\mathcal{B}_M(t,\cdot)}\right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \le \varepsilon \qquad \text{for any} \qquad t \ge \max\left\{C_p\, \varepsilon^{-\frac{1}{2-p}\frac{N+1}{\nu}}, T(\varepsilon)\right\}. \tag{63}$$

From a careful analysis of the proofs of Propositions 5.1 and 5.2, we find that there exist κ_1, κ_2, independent of ε, such that $\kappa_1\, \varepsilon^{-1} \le T(\varepsilon) \le \kappa_2\, \varepsilon^{-1}$. Since $(N+1)(2-p)^{-1}\nu^{-1} > 1$, we find that the left inequality in (63) holds for any $t \ge C_p\, \varepsilon^{-\frac{N+1}{(2-p)\nu}}$, when ε is small enough. Let us take a positive integer m such that $\varepsilon \in \left[2^{-(m+1)}, 2^{-m}\right]$, then for any $s = \frac{t}{C_p} \in \left[2^{\frac{m(N+1)}{(2-p)\nu}}, 2^{\frac{(m+1)(N+1)}{(2-p)\nu}}\right]$ we have that

$$\left\|\frac{u(t,\cdot)}{\mathcal{B}_M(t,\cdot)} - 1\right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \le \varepsilon \le 2^{-m} \le 2\, 2^{-(m+1)} \le 2\, C_p^{\frac{(2-p)\nu}{N+1}}\, t^{-\frac{(2-p)\nu}{N+1}}\,.$$

The above computation holds for any $\varepsilon \in (0, 1/2)$, so that we conclude that inequality (10) holds for any $t \ge T_\star := 2^{\frac{N+1}{(2-p)\nu}} C_p$ and $K_\star := 2\, C_p^{\frac{(2-p)\nu}{N+1}}$. The proof is complete. □

6. Convergence in relative error with rates

The main goal of this section is to prove Theorem 1. Our strategy is to provide first a convergence rate in a weaker norm (i.e., a Lebesgue norm) and then apply Theorem 7. The *intermediate asymptotics*, i.e. the behaviour of solutions to (CPLE) for large t, is much better understood when the equation is written in a different set of variables. Hence, we shall study the behaviour of solutions to the rescaled problem (29). Our main tool will be the entropy functional introduced in Section 2.1.

6.1. Relation between the relative entropy and the Fisher information along the flow

In this section we shall pave the way for the use of the entropy method. The first step is to establish a relation between the entropy functional $\mathcal{E}$ defined in (33) and the relative Fisher information $\mathcal{I}$ given by (34). But prior to that we should prove that both quantities are well defined under our assumptions. In order to do so, we introduce a few conditions, on a given function v, which we shall prove to be sufficient to establish the finiteness of the entropy and the Fisher information.

(A0) Let $v : \mathbb{R}^N \to [0, \infty)$, $C_1, C_2 > 0$ and $D > 0$ be such that

$$C_1 \leq \frac{v(y)}{V_D(y)} \leq C_2 \quad \forall y \in \mathbb{R}^N ;$$

(A1) Let $v : \mathbb{R}^N \to [0, \infty)$, $\varepsilon \in (0, 1)$ and $D > 0$ be such that

$$1 - \varepsilon \leq \frac{v(y)}{V_D(y)} \leq 1 + \varepsilon \quad \forall y \in \mathbb{R}^N ;$$

(A2) Let $v : \mathbb{R}^N \to [0, \infty)$ and $D_1, D_2 > 0$ be such that

$$V_{D_1}(y) \leq v(y) \leq V_{D_2}(y) \quad \forall y \in \mathbb{R}^N .$$

If v is a solution to (29), then condition (A0) is the translation of the global Harnack principle (47) in the new variables (27). Condition (A1) is a way of quantifying the convergence in relative error (7). Lastly, condition (A2) is the translation of assumption (9) after the change of variables (27).

Here we collect some observation used in the sequel.

Remark 6.1. We suppose that $p \in (p_c, 2)$ and $0 \leq v_0 \in \mathrm{L}^1(\mathbb{R}^N)$ is an initial datum for v being a weak solution to (29).

(i) If $0 \leq v_0 \in \mathcal{X}_p \setminus \{0\}$, then for any $\tau_0 > 0$ there exists $C_1(\tau_0), C_2(\tau_0) > 0$ for which v satisfies (A0).

(ii) If $0 \leq v_0 \in \mathcal{X}_p \setminus \{0\}$, then for any $\varepsilon \in (0, 1)$ there exists $\tau_\varepsilon > 0$ such that the solution v satisfies (A1) (with the same ε and for some $D > 0$) for all $\tau > \tau_\varepsilon$.

(iii) If $0 \leq v_0 \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N)$ satisfies (A1) for some $\varepsilon \in (0, 1)$ and some $D > 0$, then there exists $\tau_\varepsilon > 0$ such that the solution v satisfies (A1) (with the same ε and D) for all $\tau > \tau_\varepsilon$.

(iv) If v_0 satisfies (A2) for some $D_1, D_2 > 0$, then by the comparison principle $v(\tau, \cdot)$ satisfies (A2) with D_1 and D_2 for all $\tau > 0$.

(v) Let $0 \leq v_0 \in \mathcal{X}_p \setminus \{0\}$ and $D > 0$ be such that $\int_{\mathbb{R}^N} v_0 \,\mathrm{d}y = \int_{\mathbb{R}^N} V_D \,\mathrm{d}y$. If $p_c < p \leq p_M$, we additionally suppose that v_0 satisfies (A2) for some $D_1, D_2 > 0$. Then, by applying the change of variable (27) to solutions to (CPLE) and [20, Theorem 1.1], we deduce that

$$\left\| \frac{v(\tau, \cdot) - V_D}{V_D} \right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \longrightarrow 0 \quad \text{as} \quad \tau \to \infty .$$

As a consequence we also have that $\mathcal{E}[v(\tau, \cdot)|V_D] \to 0$ as $\tau \to \infty$.

Let us now provide some sufficient conditions for the entropy functional to be finite. It will be clear then that such conditions are fulfilled by solutions to (29) and, consequently, the entropy functional is well-defined along the flow. Prior to that, let us recall the exponent $p_M = (3(N+1)+\sqrt{(N+1)^2+8})/(2(N+2))$ defined in (8). For $p_M < p < 2$, solutions to (29) (under the assumption $v_0 \in \mathcal{X}_p$) have finite weighted $|x|^{\frac{p}{p-1}}$-moments. This is a necessary condition for the entropy functional $\mathcal{E}$ to be well-defined. Contrary to this case, in the range $p_c < p \le p_M$, solutions to (29) do not have a finite weighted $|x|^{\frac{p}{p-1}}$-moment anymore and we need to invoke a stronger assumption to make the entropy functional finite along the flow.

Lemma 6.2. *Let $N \ge 1$, $p_c < p < 2$, $0 \le v \in \mathrm{L}^1(\mathbb{R}^N)$, and $D > 0$ be such that $\int_{\mathbb{R}^N} V_D(y)\,\mathrm{d}y = \int_{\mathbb{R}^N} v(y)\,\mathrm{d}y$. Suppose v satisfies (A0) for some $C_1, C_2 > 0$. In the case $p_c < p \le p_M$ we additionally assume that v satisfies (A2) for some $D_1, D_2 > 0$. Then*

$$\mathcal{E}[v|V_D] < \infty\,.$$

Before the proof, let us emphasize that, at least when p is close to 2, condition (A0) is not a necessary one to have a finite relative entropy. We refer more to [27, 2, 3] for further discussion. We have chosen to use condition (A0) since it is quite practical in our setting and solutions to (29) will fulfil it for any $\tau > 0$. Let us now give the proof.

Proof. Let us consider the case $p_M < p < 2$. We shall prove that $y \mapsto [v^\gamma(y) - V_D^\gamma(y)] - \gamma V_D^{\gamma-1}(y)[v(y) - V_D(y)] \in \mathrm{L}^1(\mathbb{R}^N)$. We start by proving that $V_D^{\gamma-1}[v - V_D] \in \mathrm{L}^1(\mathbb{R}^N)$. Notice that $V_D^{\gamma-1}(y) = D + \frac{p-2}{p}|y|^{\frac{p}{p-1}}$, so that it is sufficient to show that, under assumption (A0), both v and $|y|^{\frac{p}{p-1}}v$ are integrable. From (A0), we know that there exists a constant C_2 such that $v(y) \le C_2\, V_D(y)$ for all $y \in \mathbb{R}^N$. Since $p \in (p_M, 2)$, i.e. V_D has a finite weighted $|y|^{\frac{p}{p-1}}$-moment, the previous inequality shows that both $\int_{\mathbb{R}^N} v(y)\,\mathrm{d}y$ and $\int_{\mathbb{R}^N} |y|^{\frac{p}{p-1}} v(y)\,\mathrm{d}y$ are finite. We deduce than that $V_D^{\gamma-1}(y)[v(y) - V_D(y)] \in \mathrm{L}^1(\mathbb{R}^N)$. Consequently, to conclude we only need to prove that v^γ, V_D^γ are integrable. Since both v and V_D are integrable and have the weighted $|y|^{p'}$-moments finite, we are allowed to conclude by using the Carlson–Levin inequality (see [22, Lemma 5] and references therein):

$$\mathcal{C}_{s,q,N}\left(\int_{\mathbb{R}^N} |f(y)|^q\,\mathrm{d}y\right)^{\frac{1}{q}} \le \left(\int_{\mathbb{R}^N} |y|^s |f(y)|\,\mathrm{d}y\right)^{\frac{N(1-q)}{q\,s}} \left(\int_{\mathbb{R}^N} |f(y)|\,\mathrm{d}y\right)^{1-\frac{N(1-q)}{q\,s}}\,,$$

which holds for any $s > 0$ and q such that $N/(N+s) < q < 1$. We choose $s := p'$ and $q := \gamma$. Note that they satisfy the assumption of the above Carlson–Levin inequality, because $\gamma = \frac{2p-3}{p-1}$ and $p_M < p < 2$ (cf. (8)).

Let us consider now the case $p_c < p \le p_M$. Fix $y \in \mathbb{R}^N$ and consider identity (147), with $t = v(y)$, $s = V_D$ and $v \le \xi \le V_D$. We find from that inequality, by using assumption (A0), for all $y \in \mathbb{R}^N$,

$$\frac{V_D^{\gamma-2}(y)}{2C_2^{2-\gamma}}\,(v(y) - V_D(y))^2 \le \frac{v^\gamma(y) - V_D^\gamma(y) - V_D^{\gamma-1}(y)\,(v(y) - V_D(y))}{\gamma(\gamma-1)} \le \frac{V_D^{\gamma-2}(y)}{2C_1^{2-\gamma}}\,(v(y) - V_D(y))^2\,. \tag{64}$$

Under the stronger assumption (A2), by Lemma 9.2, we also have that $|v(y) - V_D(y)| \le C\,|y|^{-\frac{p}{(2-p)(p-1)}}$ for a constant $C > 0$ and $|y| \ge 1$. This decay is enough to prove that the first and last terms in inequality (64) are integrable and so $\mathcal{E}[v|V_D]$ is finite. □

In order to show that the Fisher information is the derivative in time of the entropy functional, we make use of the space decay rate of the gradient of solutions to rescaled problem (29). To get it we adapt inequality (50) for (CPLE).

Lemma 6.3. *Let $N \ge 1$, $p_c < p < 2$ and v be a weak solution to (29) with $0 \le v_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap \mathcal{X}_p(\mathbb{R}^N)$. Then there exists a constant $c_3 = c_3(\tau, p, N) > 0$ such that*

$$|\nabla_y v(\tau, y)| \le c_3 \max\left\{1, \|v_0\|^{2\beta}_{\mathrm{L}^1(\mathbb{R}^N)} + \|v_0\|^{2\beta}_{\mathcal{X}_p}\right\}|y|^{-\frac{2}{2-p}} \quad \textit{for any} \quad |y| \ge 1 \quad \textit{and} \quad \tau > 0\,. \tag{65}$$

Constant c_3 can be chosen in such a way that for all $\tau > 0$ large enough it holds $c_3(\tau, p, N) \le c(p, N)$ for some $c(p, N)$.

Proof. We recall that, the problems (CPLE) and (29) are related through the change of variables (27). Therefore, the estimate (50) became

$$|\nabla_y v(\tau, y)| = e^{(N+1)\tau}|\nabla_x u(t, ye^\tau)| \le c e^{(N+1)\tau}\frac{\left(\|u_0\|^{2\beta}_{\mathrm{L}^1(\mathbb{R}^N)} + \|u_0\|^{2\beta}_{\mathcal{X}_p} + e^{\tau\frac{2}{2-p}}\right)}{e^{\tau(N+1)}(1 + e^\tau|y|)^{\frac{2}{2-p}}} \le C\frac{\left(\|u_0\|^{2\beta}_{\mathrm{L}^1(\mathbb{R}^N)} + \|u_0\|^{2\beta}_{\mathcal{X}_p} + e^{\tau\frac{2}{2-p}}\right)}{1 + e^{\frac{2\tau}{2-p}}|y|^{\frac{2}{2-p}}}\,.$$

Notice that we change a little bit the estimate from (50) is slightly modified to be valid for $t > 1$, eventually with a different constant. The constants above are independent of the initial datum. Having the above estimates, (65) follows from a direct computation. $\square$

Let us now focus on proving rigorously that the Fisher information is the derivative in time of the entropy functional along the flow defined by (29). We concentrate now on the case $p_M < p < 2$.

Lemma 6.4. *Let $N \geq 1$, $p_M < p < 2$, $0 \leq v_0 \in \mathrm{L}^1(\mathbb{R}^N)$, and $D > 0$ be such that $\int_{\mathbb{R}^N} V_D(y)\,\mathrm{d}y = \int_{\mathbb{R}^N} v_0(y)\,\mathrm{d}y$. Let v be a solution to (29) with v_0 as its initial datum. If v_0 satisfy (A0) holds, then for any $\tau_0 > 0$ we have that $\tau \mapsto \mathcal{I}[v(\tau)|V_D] \in \mathrm{L}^\infty(\tau_0,\infty) \cap \mathrm{L}^1(\tau_0,\infty)$ and*

$$\tfrac{d}{d\tau}\mathcal{E}[v(\tau)|V_D] = -\mathcal{I}[v(\tau)|V_D] \quad \textit{for almost every } \tau > \tau_0\,. \tag{66}$$

We notice that it is not necessary that v_0 satisfies (A0), we could have just asked that $v(\tau)$ satisfies either (A0) or (A1) after some time $\tau_1 > 0$. Indeed, this condition will be satisfied by solutions to (29) once the initial datum is assumed to be in $\mathcal{X}_p$.

Proof. The formal proof goes through equation (29) and integration by parts, namely:

$$\begin{aligned}
-\frac{d}{d\tau}(\mathcal{E}[v(\tau)|V_D]) &= \frac{1}{\gamma-1}\int_{\mathbb{R}^N}\Big(v^{\gamma-1} - V_D^{\gamma-1}\Big)\,\mathrm{div}_y\Big(v(\tau)\cdot\mathbf{a}[v(\tau)] + v(\tau)\cdot y\Big)\,\mathrm{d}y \\
&= \frac{1}{|\gamma-1|^p}\int_{\mathbb{R}^N} v(\tau)\,\Big(\nabla v^{\gamma-1} - \nabla V_D^{\gamma-1}\Big)\cdot\Big(\mathbf{b}[v^{\gamma-1}(\tau)] - \mathbf{b}[V_D^{\gamma-1}]\Big)\,\mathrm{d}y\,,
\end{aligned}$$

where $\mathbf{a}$ and $\mathbf{b}$ are as in (32). Let us justify it rigorously. For a smooth cut-off function ϕ_R such that $\phi_R = 1$ in B_R and $\phi_R = 0$ outside B_{2R}, we define

$$\begin{aligned}
\mathcal{E}_{\phi_R}[v(\tau)|V_D] &:= \frac{1}{\gamma(\gamma-1)}\int_{\mathbb{R}^N}\phi_R(y)\Big\{[v^\gamma(\tau,y) - V_D^\gamma(y)] - \gamma V_D^{\gamma-1}(y)[v(\tau,y) - V_D(y)]\Big\}\,\mathrm{d}y\,,\\
\mathcal{I}_{\phi_R}[v(\tau)|V_D] &:= \frac{1}{|\gamma-1|^p}\int_{\mathbb{R}^N}\phi_R(y)\; v(\tau,y)\big(\nabla v^{\gamma-1}(\tau,y) - \nabla V_D^{\gamma-1}(y)\big)\cdot\Big(\mathbf{b}[v^{\gamma-1}(\tau,y)] - \mathbf{b}[V_D^{\gamma-1}(y)]\Big)\,\mathrm{d}y\,,\\
\mathcal{R}_{\phi_R}[v(\tau)|V_D] &:= -\frac{1}{\gamma-1}\int_{\mathbb{R}^N}\nabla\phi_R(y)\;(v^{\gamma-1}(y) - V_D^{\gamma-1}(y))(|\nabla v(y)|^{p-2}\nabla v(y) + yv(y))\,\mathrm{d}y.
\end{aligned}$$

It will be clear in a few lines that the above quantities are well defined. In what follows, we would like to test equation (29) against $\gamma(v^{\gamma-1} - V_D^{\gamma-1})\phi_R$. This is an admissible test function. Indeed, the assumption (A0) on v_0 implies (by the maximum principle) that $v(\tau)$ satisfies (A0) for all $\tau > 0$. Moreover, we know that for a fixed $R > 0$ function $y \mapsto v^{\gamma-1}\phi_R$ is C^1. Actually more is known: the function $y \mapsto v(\tau,y)$ is $C^{1,\alpha}$ locally in space for some $\alpha = \alpha(N,p)$, i.e., $\|\nabla v\|_{C^{0,\alpha}(B_R)}$ is finite for any $R > 0$, see [37, Theorem III.8.1]. Therefore, by a standard approximation procedure, we can test (29) against $\gamma(v^{\gamma-1} - V_D^{\gamma-1})\phi_R$ to get

$$L := \int_s^t\int_{\mathbb{R}^N}\gamma\phi_R\,\Big(v^{\gamma-1} - V_D^{\gamma-1}\Big)\,\partial_\tau v\,\mathrm{d}y\,\mathrm{d}\tau = -\int_s^t\int_{\mathbb{R}^N}\gamma\phi_R\,\Big(v^{\gamma-1} - V_D^{\gamma-1}\Big)\,\mathrm{div}_y\Big(v(\tau)\cdot\mathbf{a}[v(\tau)] + v(\tau)\,y\Big)\,\mathrm{d}y\,\mathrm{d}\tau =: K,$$

where

$$\begin{aligned}
L &= \int_{\mathbb{R}^N}\gamma v(t)(v^{\gamma-1}(t) - V_D^{\gamma-1})\phi_R\,\mathrm{d}y - \gamma\int_{\mathbb{R}^N} v(s)(v^{\gamma-1}(s) - V_D^{\gamma-1})\phi_R\,\mathrm{d}y - \int_s^t\gamma\int_{\mathbb{R}^N} v(\gamma-1)v^{\gamma-2}\partial_\tau v\,\phi_R\,\mathrm{d}y\,\mathrm{d}\tau\\
&=: L_1 + L_2 + L_3\,.
\end{aligned}$$

By Fubini's theorem we infer that

$$L_3 = (\gamma-1)\int_s^t\int_{\mathbb{R}^N}\partial_\tau(v^\gamma)\,\phi_R\,\mathrm{d}y\,\mathrm{d}\tau = (\gamma-1)\left[\int_{\mathbb{R}^N} v^\gamma(t)\phi_R\,\mathrm{d}y - \int_{\mathbb{R}^N} v^\gamma(s)\phi_R\,\mathrm{d}y\right]\,.$$

Let us notice that

$$
\begin{aligned}
L = &\int_{\mathbb{R}^N} v^\gamma(t)\phi_R \,\mathrm{d}y - \int_{\mathbb{R}^N} \gamma v(t) V_D^{\gamma-1}\phi_R \,\mathrm{d}y - \int_{\mathbb{R}^N} v^\gamma(s)\phi_R \,\mathrm{d}y + \int_{\mathbb{R}^N} \gamma v(s) V_D^{\gamma-1}\phi_R \,\mathrm{d}y \\
&- (\gamma-1)\left[\int_{\mathbb{R}^N} V_D^\gamma \phi_R \,\mathrm{d}y - \int_{\mathbb{R}^N} V_D^\gamma \phi_R \,\mathrm{d}y\right] \\
= &\int_{\mathbb{R}^N} \phi_R \;\left[v^\gamma(t,y) - V_D^\gamma(y)\right] - \gamma V_D^{\gamma-1}(y)[v(t,y) - V_D(y)] \,\mathrm{d}y \\
&- \int_{\mathbb{R}^N} \phi_R \;\left[v^\gamma(s,y) - V_D^\gamma(y)\right] - \gamma V_D^{\gamma-1}(y)[v(s,y) - V_D(y)] \,\mathrm{d}y \\
= &\,\gamma(\gamma-1)\left(\mathcal{E}_{\phi_R}[v(t)|V_D] - \mathcal{E}_{\phi_R}[v(s)|V_D]\right).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
K = &\,\gamma \int_s^t \int_{\mathbb{R}^N} \phi_R \; v(\tau)\left(\nabla v^{\gamma-1} - \nabla V_D^{\gamma-1}\right)\cdot\left(\mathbf{b}[v^{\gamma-1}(\tau)] - \mathbf{b}[V_D^{\gamma-1}]\right)\mathrm{d}y\,\mathrm{d}\tau \\
&+ \gamma \int_s^t \int_{\mathbb{R}^N} \nabla\phi_R \;(v^{\gamma-1} - V_D^{\gamma-1})(|\nabla v|^{p-2}\nabla v + yv)\,\mathrm{d}y\,\mathrm{d}\tau \\
= &-\gamma(\gamma-1)\int_s^t \mathcal{I}_{\phi_R}[v(\tau)|V_D]\,\mathrm{d}\tau + \gamma(\gamma-1)\int_s^t \mathcal{R}_{\phi_R}[v(\tau)|V_D]\,\mathrm{d}\tau\,.
\end{aligned}
$$

We find therefore, by posing $s=\tau_0$ and $t=\tau_0+h$, for $h>0$, that

$$
\mathcal{E}_{\phi_R}[v(\tau_0+h)|V_D] - \mathcal{E}_{\phi_R}[v(\tau_0)|V_D] = -\int_{\tau_0}^{\tau_0+h} \mathcal{I}_{\phi_R}[v(\tau)|V_D]\,\mathrm{d}\tau - \int_{\tau_0}^{\tau_0+h} \mathcal{R}_{\phi_R}[v(\tau)|V_D]\,\mathrm{d}\tau\,. \tag{67}
$$

We need to prove that, for almost every $\tau>0$, it holds

$$
\mathcal{E}_{\phi_R}[v(\tau)|V_D] \xrightarrow[R\to\infty]{} \mathcal{E}[v(\tau)|V_D]\,, \tag{68}
$$

$$
\int_{\tau_0}^{\tau_0+h} \mathcal{I}_{\phi_R}[v(\tau)|V_D]\,\mathrm{d}\tau \xrightarrow[R\to\infty]{} \int_{\tau_0}^{\tau_0+h} \mathcal{I}[v(\tau)|V_D]\,\mathrm{d}\tau\,, \tag{69}
$$

$$
\int_{\tau_0}^{\tau_0+h} \left|\mathcal{R}_{\phi_R}[v(\tau)|V_D]\right| \xrightarrow[R\to\infty]{} 0\,. \tag{70}
$$

This would allow to pass to the limit in (67), from which we will get that

$$
\mathcal{E}[v(\tau_0+h)|V_D] - \mathcal{E}[v(\tau_0)|V_D] = -\int_{\tau_0}^{\tau_0+h} \mathcal{I}[v(\tau)|V_D]\,\mathrm{d}\tau\,. \tag{71}
$$

Then, by using the Lebesgue Differentiation Theorem, we will conclude with (66) and the proof would be complete.

Let us first deal with the entropy $\mathcal{E}$. By identity (147) (with $t=v(\tau,y)$ and $s=V_D(y)$) we know that the integrand of $\mathcal{E}[v(\tau,V_D)]$ is positive, and, by the Monotone Convergence Theorem, we can say that (68) holds. Let us now focus on the relative Fisher information. In order to justify (69) and to prove that $\tau\mapsto\mathcal{I}[v(\tau)|V_D]\in \mathrm{L}^\infty(\tau_0,\infty)$ we will make use of the Dominated Convergence Theorem. Let us note that that the integrand of $\mathcal{I}_{\phi_R}[v(\tau)|V_D]$ can be estimated pointwise by using Young's inequality and assumption (A0) (on the solution $v(\tau)$) as follows

$$
\begin{aligned}
I_\tau(y) &:= v(\tau,y)\big(\nabla v^{\gamma-1}(\tau,y) - \nabla V_D^{\gamma-1}(y)\big)\cdot\left(\mathbf{b}[v^{\gamma-1}(\tau,y)] - \mathbf{b}[V_D^{\gamma-1}(y)]\right) \\
&\le c(p,\gamma)|v(\tau,y)|\left(|\nabla v^{\gamma-1}(\tau,y)| + |\nabla V_D^{\gamma-1}(y)|\right)^p \\
&\le c(p,\gamma)|v(\tau,y)|\left(|\nabla v^{\gamma-1}(\tau,y)|^p + |\nabla V_D^{\gamma-1}(y)|^p\right) \\
&\le c(p,\gamma)|v(\tau,y)|\left(v^{p(\gamma-2)}(\tau,y)|\nabla v(\tau,y)|^p + v_D^{p(\gamma-2)}(\tau,y)|\nabla V_D(y)|^p\right) \\
&\le c(p,\gamma,C_2)|V_D(y)|^{-\frac{1}{p-1}}\left(|\nabla v(\tau,y)|^p + |\nabla V_D(y)|^p\right),
\end{aligned}
$$

Notice that in the above computation the value of the constant may change from line to line and in the last step we have used assumption (A0) so that c depends on the value of the constant C_2 appearing in (A0). For $|y| \le 1$ we note that the right-hand side of the above inequality is bounded by a constant. Indeed, the only element that cannot be explicitly computed is $|\nabla v(\tau, y)|$. However notice that, by applying inequality (49) (after the change of variables (27)), we get that $|\nabla v(\tau,\cdot)| \in \mathrm{L}^\infty(\mathbb{R}^N)$ for any $\tau > 0$. On the other hand, for $|y| \ge 1$, we can estimate $|\nabla v(\tau, y)|$ by inequality (65), to obtain that:

$$|I_\tau(y)| \le c(u_0,p,\gamma)|V_D(\tau,y)||y|^{\frac{p^2}{(2-p)(p-1)}}\,|y|^{-\frac{2p}{2-p}} = c(u_0,p,\gamma)|V_D(\tau,y)||y|^{p'}\,,$$

where the right-hand side is integrable outside a ball as long as $p > p_M$. Consequently, we have proven that $\tau \mapsto \mathcal{I}[v(\tau)|V_D] \in \mathrm{L}^\infty(\tau_0,\infty)$. At the same time, it is clear from the above that (69) follows from the Dominated Convergence Theorem.

Lastly, let us now consider the error term $\mathcal{R}$. We notice that $\mathcal{R}_{\phi_R}[v(\tau)|V_D] = \int_{\mathbb{R}^N} F_R(y)\,\mathrm{d}y$ with $F_R(y) \to 0$ almost everywhere as $R \to \infty$, since the term $\nabla\phi_R(y)$ is supported only in $A_R := B_{2R} \setminus B_R$. So, to prove (70), we only need to find an integrable function G such that $|F_R(y)| \le G(y)$ (uniformly in R and y) and then to invoke the Dominated Convergence Theorem. We already know that $F_R = 0$ in A_R^c, so that we only need to estimate it in A_R. Since v_0 satisfies (A0), so it does $v(\tau,\cdot)$ for any $\tau > 0$. Therefore, by applying condition (A0) with inequality (65), we find that, for $|y|$ large enough

$$\begin{aligned}|F_R(y)| &\le \tfrac{C_1}{R}\,|V_D(y)|^{\gamma-1}\left(|\nabla v|^{p-1} + |y|\,|vV_D(y)|\right) \le \tfrac{C_2}{R}\,|vV_D(y)|^{\gamma-1}\,|y|\left(|y|^{-\frac{p}{2-p}} + |V_D(y)|\right)\\ &\le C_3\,|V_D(y)|^{\gamma-1}\left(|y|^{-\frac{p}{2-p}} + |V_D(y)|\right) \le C_4\,|V_D(y)|^{\gamma}\,,\end{aligned}$$

where in the third inequality we used the fact that $|y|/R \le 2$ on A_R, while in the fourth one simply applies the fact that $|y|^{-\frac{p}{2-p}} \le \kappa\, V_D(y)$, for a $\kappa > 0$ independent of y. We recall that in the proof of Lemma 6.2 we have already proven that the function $y \mapsto |v_D(y)|^\gamma$ is integrable whenever $p \in (p_M, 2)$. Therefore (70) holds by the Dominated Convergence Theorem. Consequently, identity (71) is proven.

It only remains to prove that $\tau \mapsto \mathcal{I}[v(\tau)|V_D]$ is in $\mathrm{L}^1(\tau_0,\infty)$ for any $\tau_0 > 0$. This is easily done by observing that, we have $\mathcal{E}[v(\tau+h)|V_D] \to 0$ as $h \to \infty$. Therefore, we find that $\int_{\tau_0}^\infty \mathcal{I}[v(\tau)|V_D]\,\mathrm{d}\tau = \mathcal{E}[v(\tau_0)|V_D]$. We can conclude by observing that $\mathcal{I}[v(\tau)|V_D] \ge 0$ and $\mathcal{E}[v(\tau_0)|V_D] < \infty$. □

We shall now look at the relation between the relative entropy and the Fisher information when $p_c < p \le p_M$. The main difficulty here is that the Fisher information might be an unbounded function of time. We can still establish that the Fisher information is the derivative of the entropy along the flow within this range, but in a weaker form and under stronger assumptions.

Lemma 6.5. *Let $N \ge 1$, $p_c < p \le p_M$, $0 \le v_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap \mathcal{X}_p$, and $D > 0$ be such that $\int_{\mathbb{R}^N} V_D(y)\,\mathrm{d}y = \int_{\mathbb{R}^N} v_0(y)\,\mathrm{d}y$. Let v be a solution to* (29) *with v_0 as its initial datum. If v_0 satisfies both (A1) and (A2), then $\tau \mapsto \mathcal{I}[v(\tau)|v_D] \in \mathrm{L}^1(\tau_0,\infty)$ for any $\tau_0 > 0$ and*

$$\mathcal{E}[v(\tau_0)|v_D] = \int_{\tau_0}^\infty \mathcal{I}[v(\tau)|v_D]\,\mathrm{d}\tau \quad \forall\,\tau_0 > 0\,. \tag{72}$$

Proof. Let us first notice that, under the current assumption we have that $\mathcal{E}[v(\tau)|V_D] < \infty$ for any $\tau > 0$. Indeed, if v_0 satisfies (A1) then $v_0 \in \mathcal{X}_p$ and, by Remark (6.1) (point i)) we know that, for any $\tau > \tau_0 > 0$, the solution v satisfies (A0) for some constants $C_0, C_1 > 0$ which depend on τ_0. Then, thanks to Lemma 6.2, we know that $\mathcal{E}[v(\tau)|V_D] < \infty$ for any $\tau > 0$.

We start with proceeding with along the lines of the proof of Lemma 6.4. What we need to motivate differently is that (71) still holds in the current regime. The Monotone Convergence Theorem ensures that (68) and (69) hold true. However, it does not directly imply that $\int_{\tau_0}^{\tau_0+h} \mathcal{I}[v(\tau)|V_D]\,\mathrm{d}\tau < \infty$ for any $h, \tau_0 > 0$, since (70) is not known yet. With the same notation as in the proof of Lemma 6.4, recall that $\mathcal{R}_{\phi_R}[v(\tau)|V_D] = \int_{\mathbb{R}^N} F_R(y)\,\mathrm{d}y$ with $F_R(y) = -\frac{1}{\gamma-1}\nabla\phi_R(y)\cdot(v^{\gamma-1} - V_D^{\gamma-1})(|\nabla v|^{p-2}\nabla v + y\,v)$ supported in $A_R = B_{2R} \setminus B_R$. We restrict attention to $R > 1$. By using assumption (A2) and the Mean Value Theorem we have that

$$|v^{\gamma-1}(y) - V_D^{\gamma-1}(y)| \le (\gamma-1)\xi^{\gamma-2}|v(y) - V_D(y)|\,,$$

where $\min\{V_{D_1}(y), V_D(y)\} \le \xi \le \max\{V_{D_2}(y), V_D(y)\}$. Note that $|\xi| \ge c_0\,|y|^{-\frac{p}{2-p}}$, for a constant $c_0 > 0$ independent of y, so $|\xi|^{\gamma-2} \le c_1\,|y|^{\frac{p}{(2-p)(p-1)}}$. By Lemma 9.2, for $|y| > 1$ we also have that $|v(y) - V_D(y)| \le c_2\,|y|^{-\frac{p}{(2-p)(p-1)}}$ for a constant $c_2 > 0$ independent of y. This is enough to conclude that for $|y| > 1$ it holds $|v^{\gamma-1}(y) - V_D^{\gamma-1}(y)| \le c_3$ for a constant $c_3 > 0$, independent of y. Since $\nabla\phi_R$ is supported in A_R and $|\nabla\phi_R \cdot y| \le c_4$ for a constant $c_4 > 0$ independent of y, we can write that for all $R > 1$ and a constant $C > 0$ it holds

$$\left|\mathcal{R}_{\phi_R}[v(\tau)|V_D]\right| \le C\left(\frac{1}{R}\int_{A_R} |\nabla v(\tau,y)|^{p-1}\,\mathrm{d}y + \int_{A_R} v(\tau,y)\,\mathrm{d}y\right).$$

On the right-hand side above, the second integral converges to 0, as $R \to \infty$, since $v(\tau,\cdot)$ is an integrable function and $|A_R| \subset \{|y| > R\}$. It only remains to justify the convergence of the first integral. By inequality (65), we have that for large enough $|y|$ and large enough $\tau > 0$ it holds $|\nabla v(\tau,y)|^{p-1} \le c_5\,|y|^{-\frac{2(p-1)}{2-p}}$ for a constant $c_5 > 0$ independent of y (a similar estimate holds also for small τ but with a constant c_5 depending on τ). Therefore, we are left with

$$\frac{1}{R}\int_{A_R} |\nabla v(\tau,y)|^{p-1}\,\mathrm{d}y \le \frac{c_6}{R}\int_R^\infty r^{N-1-\frac{2(p-1)}{2-p}}\,\mathrm{d}r \le c_7\,R^{N-\frac{p}{2-p}}\,,$$

for a constant $c_7 > 0$ independent of R. This justifies (70) and implies that $\int_{\tau_0}^{\tau_0+h} \mathcal{I}[v(\tau)|V_D]\,\mathrm{d}\tau < \infty$ for any $h, \tau_0 > 0$. Therefore, (71) holds true. The proof is complete. $\square$

6.2. Linearised entropy functional and Fisher information

Recall that the entropy functional $\mathcal{E}$ is given by (33). We define the *linearised relative entropy* by

$$\mathsf{E}[v] := \frac{1}{2}\int_{\mathbb{R}^N} |v - V_D|^2 V_D^{\gamma-2}\,\mathrm{d}y\,. \tag{73}$$

This functional will play an important role in the most challenging regime $p_c < p \le p_M$. Let us justify that this functional is well defined.

Lemma 6.6. *Let $N \ge 1$, $p_c < p < 2$, $0 \le v \in \mathrm{L}^1(\mathbb{R}^N)$, and $D > 0$ be such that $\int_{\mathbb{R}^N} V_D(y)\,\mathrm{d}y = \int_{\mathbb{R}^N} v(y)\,\mathrm{d}y$. Suppose v satisfies (A1) for some $\varepsilon \in (0,1)$. If $p_c < p \le p_M$, we additionally assume that v satisties (A2) for some $D_1, D_2 > 0$. Then $\mathsf{E}[v] < \infty$ and*

$$(1+\varepsilon)^{\gamma-2}\,\mathsf{E}[v] \le \mathcal{E}[v|V_D] \le (1-\varepsilon)^{\gamma-2}\,\mathsf{E}[v]\,.$$

Proof. We notice that under the current assumptions inequality (64) holds trues. By Lemma 9.2, we also have that $|v(y) - V_D(y)| \le C\,|y|^{-\frac{p}{(2-p)(p-1)}}$ for a constant $C > 0$ and $|y| \ge 1$. This decay is enough to prove that the first and last terms in inequality (64) are integrable. Then, by integrating inequality (64), we deduce the claim. $\square$

Let us now introduce two different quantities which, in the sequel, will play the rôle of a *linearised* version of the Fisher information $\mathcal{I}$ defined in (34). For any function $v : \mathbb{R}^N \to \mathbb{R}$, consider

$$\mathsf{I}_\gamma^{(\varepsilon)}[v] := \frac{1}{|\gamma-1|^p}\int_{\mathbb{R}^N} \left|\nabla v^{\gamma-1} - \nabla V_D^{\gamma-1}\right|^2 V_D\,\left(\varepsilon + |\nabla V_D^{\gamma-1}|\right)^{p-2}\,\mathrm{d}y \tag{74}$$

and

$$\mathsf{I}^{(\varepsilon)}[v] := \frac{1}{|\gamma-1|^p}\int_{\mathbb{R}^N} \left|\nabla\left(V_D^{\gamma-2}(v - V_D)\right)\right|^2 V_D\,\left(\varepsilon + |\nabla V_D^{\gamma-1}|\right)^{p-2}\,\mathrm{d}y\,. \tag{75}$$

Our first result is a control, along the flow, of the Fisher information $\mathcal{I}[v(\tau)|V_D]$ from below by the first quantity defined above, at least when the considered solution v is close to Barenblatt profile V_D in the sense of assumptions (A1) and (A2).

Lemma 6.7. *Let $N \ge 1$, $p_c < p < 2$, $0 \le v_0 \in \mathrm{L}^1(\mathbb{R}^N)$, and $D > 0$ be such that $\int_{\mathbb{R}^N} V_D(y)\,\mathrm{d}y = \int_{\mathbb{R}^N} v_0(y)\,\mathrm{d}y$. Let v be a solution to* (29) *with v_0 as its initial datum and assume that v_0 satisfies (A1) for some $\varepsilon \in (0,1)$. If $p_c < p < p_M$, we additionally assume that v_0 satisfies (A2) for some $D_1, D_2 > 0$. Then there exists $\tau_\varepsilon > 0$ and $C_\varepsilon = C_\varepsilon(p, N, v_0, \varepsilon, D) > 0$ such that for all $\tau > \tau_\varepsilon$ it holds*

$$\mathcal{I}[v(\tau)|V_D] \ge C_\varepsilon\,\mathsf{I}_\gamma^{(\varepsilon)}[v(\tau)]\,. \tag{76}$$

Proof. We recall that, by Remark 6.1 that, there exists $\tau_\varepsilon > 0$ such that $v(\tau,\cdot)$ satisfies (A1) for all $\tau > \tau_\varepsilon$. By applying Lemma 9.3 and assumption (A1) we find for all $\tau > \tau_\varepsilon$ that

$$\begin{aligned}\mathcal{I}[v(\tau)|V_D] &\geq \frac{\min\{1/2, p-1\}}{|\gamma-1|^p}\int_{\mathbb{R}^N} v(\tau,y)\big(|\nabla v^{\gamma-1}(\tau,y)| + |\nabla V_D^{\gamma-1}(y)|\big)^{p-2}\big|\nabla v^{\gamma-1}(\tau,y) - \nabla V_D^{\gamma-1}(y)\big|^2\,\mathrm{d}y\\ &\geq \frac{\min\{1/2, p-1\}}{|\gamma-1|^p}(1-\varepsilon)\int_{\mathbb{R}^N} V_D(y)\big(|\nabla v^{\gamma-1}(\tau,y)| + |\nabla V_D^{\gamma-1}(y)|\big)^{p-2}\big|\nabla v^{\gamma-1}(\tau,y) - \nabla V_D^{\gamma-1}(y)\big|^2\,\mathrm{d}y\,, \end{aligned} \tag{77}$$

where in the second inequality we use condition (A2) for all $\tau > \tau_\varepsilon$. We will to show that

$$|\nabla v^{\gamma-1}(y)| \leq C(v_0,p,N,\varepsilon,D)\left(\varepsilon + |\nabla V_D^{\gamma-1}(y)|\right)\,. \tag{78}$$

Let us notice that $|\nabla V_D^{\gamma-1}(y)| = (1-\gamma)|y|^{2-\gamma} \geq 0$ and $\nabla v^{\gamma-1} = (\gamma-1)\,v^{\gamma-2}\nabla v$. By the fact that $\|\nabla v\|_{\mathrm{L}^\infty(\mathbb{R}^N)} < \infty$ (resulting from (49)) and condition (A2) we have that, for $|y| \leq 1$ and $\tau > \tau_\varepsilon$ it holds

$$\begin{aligned}|\nabla v^{\gamma-1}(\tau,y)| &\leq (1-\gamma)(1-\varepsilon)^{\gamma-2}\,V_D^{\gamma-2}(y)\,\|\nabla v\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \leq \frac{(1-\gamma)}{(1-\varepsilon)^{2-\gamma}}\sup_{|y|\leq 1} V_D^{\gamma-2}\,\frac{\|\nabla v\|_{\mathrm{L}^\infty(\mathbb{R}^N)}}{\varepsilon}\,\left(\varepsilon + |\nabla V_D^{\gamma-1}(y)|\right)\\ &\leq C(v_0,p,N,\varepsilon,D)\left(\varepsilon + |\nabla V_D^{\gamma-1}(y)|\right)\,.\end{aligned}$$

On the other hand, for $|y| \geq 1$ and $\tau > \tau_\varepsilon$, by using (65) we infer that for a constant $C(v_0,p,N) > 0$ it holds

$$\begin{aligned}|\nabla v^{\gamma-1}(\tau,y)| &\leq (1-\gamma)\,(1-\varepsilon)^{\gamma-2}\,V_D^{\gamma-2}(y)\,C(v_0)|y|^{-\frac{2}{2-p}} = (1-\varepsilon)^{\gamma-2}\,C(v_0,p,N)\,|\nabla V_D^{\gamma-1}(y)|\,V_D^{\gamma-2}(y)\,|y|^{-\frac{p}{(2-p)(p-1)}}\\ &\leq C(v_0,p,N,\varepsilon,D)\,\left(\varepsilon + |\nabla V_D^{\gamma-1}(y)|\right)\,,\end{aligned}$$

where we used the fact that $V_D^{\gamma-2}(y)\,|y|^{-\frac{p}{(2-p)(p-1)}} \leq C(D,N,p)$ for $|y| \geq 1$ where $C(D,N,p) > 0$ depends only on D, N and p. Combining the two cases $|y| \leq 1$ and $|y| \geq 1$ together, we find that (78) holds for all $y \in \mathbb{R}^N$. Then (78) together with (77) implies (76) with the constant $C_\varepsilon = 2^{p-2}\,C(v_0,p,N,\varepsilon,D)^{p-2}\,c_2(1-\varepsilon)$. $\square$

The previous lemma show a relation between the Fisher information and one of its linearised versions. In what follows, we state an inequality that holds among all the linearised quantities introduced in this section. The following lemma is originally contained in [3, Claim 1 of Proposition 4.2]. In that paper, the authors also use some quantities very similar to ours $\mathsf{I}^{(\varepsilon)}$ and $\mathsf{I}_\gamma^{(\varepsilon)}$, however their definition is slightly different. Here is their result written in our notation. Notice slightly different constants and extended range of p in comparison with [3].

Lemma 6.8. *Let $N \geq 1$, $p \in (1,2)$, $N < \frac{p}{(2-p)(p-1)}$, $0 \leq v \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N)$ such that $\nabla v \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^N)$, and let $D > 0$ be such that $v - V_D \in \mathrm{L}^1(\mathbb{R}^N)$ and $\int_{\mathbb{R}^N}(V_D(y) - v(y))\,\mathrm{d}y = 0$. Suppose v satisfies (A1) for some $\varepsilon \in (0,1)$. If $1 < p \leq p_M$, we additionally assume that v satisties (A2) for some $D_1, D_2 > 0$. Then we have that*

$$\mathsf{I}^{(\varepsilon)}[v] \leq \kappa_1\,\mathsf{I}_\gamma^{(\varepsilon)}[v] + \kappa_2\,\mathsf{E}[v] \tag{79}$$

where

$$\kappa_1 = \frac{(1+\varepsilon)^{2(2-\gamma)}}{(1-\gamma)^2} \qquad \text{and} \qquad \kappa_2 = \frac{\mathcal{C}_{p,N}}{(1-\gamma)^{p-1}}\left(\frac{(1+\varepsilon)^{2(2-\gamma)}}{(1-\varepsilon)^{2(2-\gamma)}} - 1\right)\,,$$

where $\mathcal{C}_{p,N} > 0$ is a constant depending on N and p.

In what follows we only sketch the main steps of the reasoning, which is based on the proof of [2, Proposition 4.2].

Proof. We first explain the outline of the proof and then we justify the key technicality, which is an integration by parts. Let us introduce $h_k(s) := s^{k-1} - 1$, for $k \in \{2,\gamma\}$, and $\mu(y) := (\varepsilon + (1-\gamma)|y|^{2-\gamma})^{p-2}\,V_D(y)$ so that $\mathrm{d}\mu(y) := \mu(y)\,\mathrm{d}y$. In order to unify the notation, let us define

$$I_k^\varepsilon[w] := \frac{1}{|1-\gamma|^p}\int_{\mathbb{R}^N}\left|\nabla(V_D^{\gamma-1}h_k(w))\right|^2\,\mathrm{d}\mu(y)\,, \qquad \text{for } k \in \{2,\gamma\}\,.$$

Notice that with this definition and upon setting $w := v/V_D$, we have that $\mathsf{I}^{(\varepsilon)}[v] = I_2^\varepsilon[w]$ and $\mathsf{I}_\gamma^{(\varepsilon)}[v] = I_\gamma^\varepsilon[w]$, where $\mathsf{I}^{(\varepsilon)}[v]$ and $\mathsf{I}_\gamma^{(\varepsilon)}[v]$ are defined in (75) and in (74), respectively. Recall that $\nabla V_D^{\gamma-1} = (1-\gamma)|y|^{1-\gamma}y$. Then, by computing the gradient $\nabla(V_D^{\gamma-1}h_k(w))$, expanding a square, and recognizing the form of $\nabla h_k^2(w)$, for any k, we have

$$|1-\gamma|^p I_k^\varepsilon[w] = \int_{\mathbb{R}^N} |h_k'(w)|^2\,|\nabla w|^2\,V_D^{\gamma-1}\,\mathrm{d}\mu(y) + (1-\gamma)^2\int_{\mathbb{R}^N} |h_k(w)|^2\,|y|^{2(2-\gamma)}\,\mathrm{d}\mu(y) + \int_{\mathbb{R}^N} (\nabla h_k^2(w))\cdot(\nabla V_D^{\gamma-1})\,V_D^{\gamma-1}\,\mathrm{d}\mu(y)\,.$$

By integrating by parts the last term, that will be justified below, one gets that

$$\begin{aligned} |1-\gamma|^p I_k^\varepsilon[w] &= \int_{\mathbb{R}^N} |h_k'(w)|^2\,|\nabla w|^2\,V_D^{\gamma-1}\,\mathrm{d}\mu(y) + (1-\gamma)\int_{\mathbb{R}^N} |h_k(w)|^2\,|y|^{2(2-\gamma)}\,\mathrm{d}\mu(y) \\ &\quad - (1-\gamma)\int_{\mathbb{R}^N} h_k^2(w)\,V_D^\gamma\,\mathrm{div}\left(y\,|y|^{1-\gamma}\left(\varepsilon + (1-\gamma)|y|^{2-\gamma}\right)^{p-2}\right)\mathrm{d}y\,. \end{aligned} \tag{80}$$

Under the current assumptions and for any $s \in (1-\varepsilon, 1+\varepsilon)$, we can deduce the following relations between h_2 and h_γ

$$(1-\varepsilon)^{2(2-\gamma)}\,\frac{h_\gamma(s)^2}{(1-\gamma)^2} \le h_2(s)^2 \le (1+\varepsilon)^{2(2-\gamma)}\,\frac{h_\gamma(s)^2}{(1-\gamma)^2} \qquad \text{and} \qquad (h_2'(s))^2 \le \frac{(1+\varepsilon)^{2(2-\gamma)}}{(1-\gamma)^2}\,(h_\gamma'(s))^2\,. \tag{81}$$

Combining the above inequalities with identity (80) we obtain that

$$\begin{aligned} &|1-\gamma|^p I_2^\varepsilon[w] \\ \le&\frac{(1+\varepsilon)^{2(2-\gamma)}}{(1-\gamma)^2}\,|1-\gamma|^p I_\gamma^\varepsilon[w] + (1-\gamma)\int_{\mathbb{R}^N}\left(\frac{(1+\varepsilon)^{2(2-\gamma)}}{(1-\gamma)^2}h_\gamma^2(w) - h_2^2(w)\right)V_D^\gamma\,\mathrm{div}\left(y\,|y|^{1-\gamma}\left(\varepsilon+(1-\gamma)|y|^{2-\gamma}\right)^{p-2}\right)\mathrm{d}y \\ \le&\frac{(1+\varepsilon)^{2(2-\gamma)}}{(1-\gamma)^2}\,|1-\gamma|^p I_\gamma^\varepsilon[w] + (1-\gamma)\left(\frac{(1+\varepsilon)^{2(2-\gamma)}}{(1-\varepsilon)^{2(2-\gamma)}} - 1\right)\int_{\mathbb{R}^N} h_2^2(w)\,V_D^\gamma\,\left|\mathrm{div}\left(y\,|y|^{1-\gamma}\left(\varepsilon+(1-\gamma)|y|^{2-\gamma}\right)^{p-2}\right)\right|\,\mathrm{d}y\,, \end{aligned}$$

where in the last step we have used the first inequality in (81) in order to control h_γ^2 with h_2^2. It only remains to compute the divergence in the last term of the above inequality. By a simple, though long, computation, one finds that

$$\mathrm{div}\left(y\,|y|^{1-\gamma}\left(\varepsilon+(1-\gamma)|y|^{2-\gamma}\right)^{p-2}\right) = \frac{|y|^{1-\gamma}}{\left(\varepsilon+(1-\gamma)|y|^{2-\gamma}\right)^{3-p}}\,\left[\varepsilon(N+1-\gamma) + (1-\gamma)(N+p-\gamma-1)|y|^{2-\gamma}\right]\,.$$

Since both $(N+1-\gamma)$ and $(N+p-1-\gamma)$ are nonnegative, the divergence above has a sign. Moreover, we notice that there exists a constant $\mathcal{C}_{p,N} > 0$, depending on N and p but not on ε, such that $\left|\mathrm{div}\left(y\,|y|^{1-\gamma}\left(\varepsilon+(1-\gamma)|y|^{2-\gamma}\right)^{p-2}\right)\right| \le \mathcal{C}_{p,N}$ for all $y \in \mathbb{R}^N$. Combining all the above estimates and noting that $2\mathsf{E}[v] = \int_{\mathbb{R}^N} h_2^2(w)V_D^\gamma\,\mathrm{d}y$, one obtains (79).

Let us give some details for the justification of the above integration by parts applied to (80). At first we notice that if $\mathsf{I}_\gamma^{(\varepsilon)}[v] = \infty$, there is nothing to prove, so we may assume that $\mathsf{I}_\gamma^{(\varepsilon)}[v] < \infty$. To deal with the opposite case we notice that all the integrals in this proof are well defined while restricted to a ball B_R, for $R > 0$. What is more, all the computations are exactly the same up to an error term which comes from the integration by parts applied to (80). Therefore, what remains to show is that this error term indeed vanishes in the limit as $R \to \infty$. In order to do so, let us define $U(y) := y\,(1-\gamma)|y|^{1-\gamma}V_D^{\gamma-1}(y)\mu(y)$. Then, by Green's identity, we have that for any $R > 0$ it holds

$$\int_{B_R}(\nabla h_k^2(w))\cdot(\nabla V_D^{\gamma-1})\,V_D^{\gamma-1}\,\mathrm{d}\mu(y) = \int_{B_R}\nabla h_k^2(w)\cdot U\,\mathrm{d}y = -\int_{B_R} h_k^2(w)\,\mathrm{div}\,U\,\mathrm{d}y + \int_{\partial B_R} h_k^2(w)\,U\cdot\widehat{y}\,\mathrm{d}\sigma(y)\,, \tag{82}$$

where $\widehat{y} := y/|y|$ and $\mathrm{d}\sigma$ is the surface measure of ∂B_R. Notice that

$$\mathrm{div}\,(U) = V_D^\gamma\,\mathrm{div}\left(y\,|y|^{1-\gamma}\left(\varepsilon+(1-\gamma)|y|^{2-\gamma}\right)^{p-2}\right) - \gamma(1-\gamma)\,|y|^{2(2-\gamma)}\,\mu(y)\,.$$

Consequently, in the limit $R \to \infty$ one recovers the two last terms in (80). Hence, we only need to show that the remainder term $\int_{\partial B_R} h_k^2(w)\,U\cdot\widehat{y}\,\mathrm{d}\sigma(y)$ converges to zero as $R \to \infty$. When $p > p_M$, by assumption (A1), we have that $\left|h_k^2(w)\,U\cdot\widehat{y}\right| \le C|h_2(w)|^2\,|U| \le C|y|\,V_D^\gamma$. As shown in Lemma 6.2, V_D^γ is integrable as long as $p > p_M$. In turn, $|y|\,V_D^\gamma \le C|y|^{1-N-\delta}$ for some $\delta > 0$, which implies that the term $\int_{\partial B_R} h_k^2(w)\,U\cdot\widehat{y}\,\mathrm{d}\sigma(y)$ vanishes in the limit in the considered range since. Let us now concentrate on the case of $p \le p_M$ and assumption (A2). Thanks to Lemma 9.2 we know that $|v - V_D| \le C\,|y|^{-\frac{p}{(p-1)(2-p)}}$ for $|y| \ge 1$, so that $\left|h_k^2(w)\,U\cdot\widehat{y}\right| \le C\,|y|^{-\frac{p}{(p-1)(2-p)}+1}$. Since $p > N(p-1)(2-p)+\delta$ for some $\delta > 0$, we conclude that also in this case the remainder term vanishes in the limit $R \to \infty$. The proof is complete. $\square$

Before we establish the final entropy – entropy production inequality (35) along the flow, we shall prove its linearised version. With this aim we employ the Hardy–Poincaré inequality provided as [24, Example 3.1] for $q = 2$, $\gamma = 0$, $\beta = p/(p-1)$ and $\alpha = -1/(2-p)$.

Proposition 6.9 (Hardy–Poincaré inequality)**.** *Let $N \geq 2$, $p \in (1,2)$, and $N < \frac{p}{(2-p)(p-1)}$. Then there exists a finite constant $C_{HP} = C_{HP}(p,N) > 0$, such that for every compactly supported $\varphi \in W^{1,\infty}(\mathbb{R}^N)$ the following inequality holds true*

$$\int_{\mathbb{R}^N} |\varphi - \overline{\varphi}|^2 (1 + |y|^{\frac{p}{p-1}})^{-\frac{1}{2-p}} \, \mathrm{d}y \leq C_{HP} \int_{\mathbb{R}^N} |\nabla \varphi|^2 |y|^2 (1 + |y|^{\frac{p}{p-1}})^{-\frac{1}{2-p}} \, \mathrm{d}y, \tag{83}$$

where $\overline{\varphi}$ is the average of φ with respect to $(1 + |y|^{\frac{p}{p-1}})^{-\frac{1}{2-p}}$.

Before continuing, let us also define $M_\star := \beta^{\frac{N}{p}} b_1^{\frac{p-1}{\beta p(2-p)}}$, where b_1 is as in (40) and β as in (2). Then

$$M_\star^{\frac{\beta p(2-p)}{p-1}} = \beta^{N\beta\frac{2-p}{p-1}} \, b_1 \qquad \text{and} \qquad M_\star = \int_{\mathbb{R}^N} V_1(y) \, \mathrm{d}y \,, \tag{84}$$

with V_1 as in (3) and (5). In what follows, for the sake of simplicity of the exposition, we shall consider solutions of mass equal to $M_\star$.

Lemma 6.10. *Let $N \geq 2$, $p_c < p < p_2$ and $0 \leq v \in \mathrm{L}^1(\mathbb{R}^N)$, such that $M_\star = \int_{\mathbb{R}^N} v(y) \, \mathrm{d}y$. Assume v satisfies (A1) for some $\varepsilon \in (0,1)$. If $p_c < p < p_M$, we additionally assume that v satisfies (A2). Then there exists a constant $\mathcal{C} = \mathcal{C}(p,N) > 0$ such that*

$$\mathcal{C} \, \mathsf{E}[v] \leq \mathsf{I}^{(\varepsilon)}[v] \,. \tag{85}$$

Proof. Let φ be any function for which both sides of (83) are well defined. We notice that, even if (83) is stated for regular and compactly supported functions, the same inequality holds true for a larger class of functions through a standard approximation procedure. Let also $\overline{\varphi} = \int_{\mathbb{R}^N} \varphi(y) \, \mathrm{d}\mu_1(y) / \mu_1(\mathbb{R}^N)$ where $\mathrm{d}\mu_1(y) := (1 + |y|^{p'})^{-\frac{1}{2-p}} \, \mathrm{d}y$. Since $p \, V_1^{\gamma-1} \geq (2-p) \, (1 + |y|^{\frac{p}{p-1}})$, we obtain that

$$\left(\frac{2-p}{p} \right)^{\frac{1}{2-p}} \inf_{c \in \mathbb{R}} \int_{\mathbb{R}^N} |\varphi(y) - c|^2 \, V_1^{2-\gamma}(y) \, \mathrm{d}y \leq \int_{\mathbb{R}^N} |\varphi(y) - \overline{\varphi}|^2 \, \mathrm{d}\mu_1(y) \,. \tag{86}$$

It is known that the infimum on the left-hand side of the above inequality is achieved when $c = Z^{-1} \int_{\mathbb{R}^N} \varphi(y) \, V_1^{2-\gamma}(y) \, \mathrm{d}y$ where $Z = \int_{\mathbb{R}^N} V_1^{2-\gamma}(y) \, \mathrm{d}y$. We apply inequality (86) to the function $\varphi = (v - V_1) \, V_1^{\gamma-2}$ (which satisfies $0 = \int_{\mathbb{R}^N} \varphi(y) \, V_1^{2-\gamma}(y) \, \mathrm{d}y$) and we find

$$2 \left(\frac{2-p}{p} \right)^{\frac{1}{2-p}} \mathsf{E}[v] \leq \int_{\mathbb{R}^N} |\varphi(y) - \overline{\varphi}|^2 \, \mathrm{d}\mu_1(y) \,. \tag{87}$$

It only remain to estimate the right-hand side of (83) by the right-hand side of (85). In order to do this, we observe that for any $y \in \mathbb{R}^N$ we have

$$\frac{|y|^2}{\left(1 + |y|^{\frac{p}{p-1}}\right)^{\frac{1}{2-p}}} \frac{\left(\varepsilon + (1-\gamma)|y|^{2-\gamma}\right)^{2-p}}{V_1(y)} \leq \left(\frac{2}{p} \right)^{\frac{p-1}{2-p}} (2-\gamma)^{2-p} \,,$$

where we used the fact that $|\nabla V_1^{\gamma-1}(y)| = (1-\gamma)|y|^{2-\gamma}$. This is enough to prove that

$$\int_{\mathbb{R}^N} |\nabla \varphi|^2 |y|^2 \, \mathrm{d}\mu_1(y) \leq (1-\gamma)^p \left(\frac{2}{p} \right)^{\frac{p-1}{2-p}} (2-\gamma)^{2-p} \, \mathsf{I}^{(\varepsilon)}[v] \,.$$

We conclude, therefore, that inequality (85) holds with the constant

$$\mathcal{C} := \frac{2}{C_{HP} \, (2-\gamma)^{2-p}} \frac{p^{\frac{p-1}{2-p}}}{2^{\frac{p-1}{2-p}} \, (1-\gamma)^p} \left(\frac{2-p}{p} \right)^{\frac{1}{2-p}} \,.$$

□

6.3. Convergence in L^1*, case* $p_D \le p < 2$

There is a special value of parameter p, above which it is well-known that one is equipped with strong tools. This value is

$$p_D := \frac{2N+1}{N+1}\,. \tag{88}$$

For $p > p_D$ the relative entropy functional is displacement convex. This notion expresses the convexity of the entropy functional along geodesics in the space of probability densities equipped with the Wasserstein metric, cf. [59]. In such situation, using the relation between $\mathcal{E}$ and $\mathcal{I}$ established in [2], we can directly show an exponential rate of L^1-convergence of solutions to (29) with $p \in [p_D, 2)$ via a Gronwall's argument.

Before presenting this result let us recall that $p_c < p_D$ and let us refer to Section 3 for more comments on other special values of parameter p. The main accomplishment of our paper is establishing rates of convergence below p_D; additionally we give alternative proof above p_D. Since in the proof of Theorem 1 parameter p_D plays no role, we stress that it is a special value for the validity of the optimal transportation method, but it is not critical from the point of view of the of the asymptotics. We are in a position to present a very short reasoning for $p \in (p_D, 2)$ relying on the argument by [2] and not making use of linearisation of $\mathcal{E}$ or $\mathcal{I}$.

Proposition 6.11. *Let $N \ge 1$, $p_D \le p < 2$, $0 \le v_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap \mathcal{X}_p$, and $D > 0$ be such that $\int_{\mathbb{R}^N} V_D(y)\,\mathrm{d}y = \int_{\mathbb{R}^N} v_0(y)\,\mathrm{d}y$. Assume that v is a weak solution to* (29) *with v_0 as initial datum. Then there exists $c = c(v_0, p, N, D)$, such that for any $\tau > 0$*

$$\|v(\tau) - V_D\|_{\mathrm{L}^1(\mathbb{R}^N)} \le c\, e^{-\tau/2}\,. \tag{89}$$

Proof. Since $v_0 \in \mathcal{X}_p$, we have, thanks to Remark 6.1 *(i)*, that fixed any $\tau_0 > 0$ the solution $v(\tau,\cdot)$ satisfies (A0) for any $\tau > \tau_0$. Then, thanks to Lemma 6.2, we have that $\mathcal{E}[v(\tau)|V_D] < \infty$. Thus, by Lemma 9.1 we know that $\|v - V_D\|^2_{\mathrm{L}^1(\mathbb{R}^N)} \le c(p,N,D)\mathcal{E}[v(\tau)|V_D]$. Moreover, since $p_D > p_c$ Lemma 6.4 implies that $\frac{d}{d\tau}\mathcal{E}[v(\tau)|V_D] = -\mathcal{I}[v(\tau)|V_D]$ for almost every $\tau > 0$. By [2, Theorem 2.2], within the range $p_D \le p < 2$ and for $v_0 \in \mathrm{L}^1(\mathbb{R}^N)$, it holds

$$\mathcal{E}[v(\tau)|V_D] \le \mathcal{I}[v(\tau)|V_D] = -\tfrac{d}{d\tau}\mathcal{E}[v(\tau)|V_D]\,.$$

Then Gronwall's Lemma implies $\mathcal{E}[v(\tau)|V_D] \le ce^{-\tau}\mathcal{E}[v(\tau_0)|V_D]$ for all $\tau \ge \tau_0$. Collecting all information we get (89). $\square$

6.4. Convergence in L^1*, case* $p_M < p < 2$

In this section we shall provide a proof of the convergence in the L^1-norm in the range $p_M < p < 2$. Note that for $p_M < p_D$ this result is new, while for $p_D \le p < 2$ we give a proof different not involving the optimal transportation tools (applied in [3]). We stress again that the parameter p_D plays no role in this reasoning. For the simplicity of the exposition, we shall consider initial data with a fixed mass $\int_{\mathbb{R}^N} v_0(y)\,\mathrm{d}y = M_\star$ with $M_\star$ being defined in (84). We will recover full generality in the proof of Proposition 1.2.

Proposition 6.12. *Let $N \ge 2$, $p_M < p < p_2$, $0 \le v_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap \mathcal{X}_p$, be such that $\int_{\mathbb{R}^N} v_0(y)\,\mathrm{d}y = M_\star$. Assume that v is a weak solution to* (29) *with v_0 as initial datum. If v_0 satisfies (A1) for some $\varepsilon \in (0,1)$ and $D = 1$, then there exists $\tau_\varepsilon > 0$,$c = c(v_0, p, N, \varepsilon)$, and $\vartheta = \vartheta(p, N, \varepsilon) > 0$ such that for any $\tau > \tau_\varepsilon$ it holds*

$$\|v(\tau) - V_1\|_{\mathrm{L}^1(\mathbb{R}^N)} \le c\, e^{-\vartheta\,\tau/2}\,. \tag{90}$$

Proof. Thanks to Remark 6.1 *(iii)* and the fact that $M_\star = \int_{\mathbb{R}^N} V_1(y)\,\mathrm{d}y$, there exists $\tau_\varepsilon > 0$ such that the solution $v(\tau,\cdot)$ satisfies (A1) with $D = 1$ for any $\tau > \tau_\varepsilon$. Then, thanks to Lemmata 6.2 and 6.4, we have that $\mathcal{E}[v(\tau)|V_1], \mathcal{I}[v(\tau)|V_1] < \infty$ for all $\tau > \tau_\varepsilon$. In what follows, we assume that $\tau > \tau_\varepsilon$. By Lemma 6.8 for every ε there exist $\kappa_0, \kappa_2 > 0$ such that $\mathsf{I}^{(\varepsilon)}[v] \le \kappa_1\, \mathsf{I}^{(\varepsilon)}_\gamma[v] + \kappa_2\, \mathsf{E}[v]$ and $\kappa_2 \to 0$ as $\varepsilon \to 0$. Let $\mathcal{C}$ be the constant from Lemma 6.10 for which $\mathcal{C}\,\mathsf{E}[v] \le \mathsf{I}^{(\varepsilon)}[v]$. Moreover, we recall that by Lemma 6.7 we have $C_\varepsilon\, \mathsf{I}^{(\varepsilon)}_\gamma[v(\tau)] \le \mathcal{I}[v(\tau)|V_1]$. We restrict attention to $\varepsilon \in (0,1)$ small enough to ensure that $\kappa_2 \le \frac{\mathcal{C}}{2}$. Summing up, we get that

$$\tfrac{\mathcal{C}}{2}\mathsf{E}[v] \le \frac{\kappa_1}{C_\varepsilon}\mathcal{I}[v(\tau)|V_1]\,. \tag{91}$$

On the other hand, by using Lemmata 6.6 and 6.4 we obtain

$$\tfrac{d}{d\tau}\mathcal{E}[v(\tau)|V_1] \le -C_\varepsilon\, \tfrac{\mathcal{C}}{2\kappa_1}(1+\varepsilon)^{\gamma-2}\mathcal{E}[v(\tau)|V_1]\,.$$

By using Gronwall's Lemma we can deduce from the above inequality that $\mathcal{E}[v(\tau)|V_1] \le e^{-\vartheta\tau}\,\mathcal{E}[v(\tau_\varepsilon)|V_1]$ for some $\vartheta = \vartheta(p,N,\varepsilon) > 0$ and all $\tau \ge \tau_\varepsilon$. Then (90) follows from Lemma 9.1. $\square$

6.5. Convergence in relative error with rate, case $p_c < p \le p_M$

We finally address the case of $p_c < p \le p_M$. The main difference here is that the relative Fisher information is not bounded anymore, but it is merely an L^1-function of time. This technical difficulty will be overcome by the use of a different version of Gronwall's Lemma 9.4.

Proposition 6.13. *Let $N \ge 2$, $p_c < p \le p_M$, $0 \le v_0 \in \mathrm{L}^1(\mathbb{R}^N) \cap \mathcal{X}_p$, be such that $\int_{\mathbb{R}^N} v_0(y)\,\mathrm{d}y = M_\star$. Assume that v is a weak solution to* (29) *with v_0 as initial datum. If v_0 satisfies (A1), for some $\varepsilon \in (0,1)$ and $D = 1$, and v_0 satisfies (A2), for some $D_1, D_2 > 0$, then there exists $\tau_\varepsilon > 0$, $c = c(v_0, p, N, \varepsilon)$ and $\vartheta = \vartheta(p, N, \varepsilon) > 0$ such that for any $\tau > \tau_\varepsilon$ it holds*

$$\|v(\tau) - V_1\|_{\mathrm{L}^1(\mathbb{R}^N)} \le c\, e^{-\vartheta\,\tau/2} . \tag{92}$$

Proof. Thanks to Remark 6.1 (*iii*) and the fact that $M_\star = \int_{\mathbb{R}^N} V_1(y)\,\mathrm{d}y$, there exists $\tau_\varepsilon > 0$ such that the solution $v(\tau,\cdot)$ satisfies (A1) with $D = 1$ for any $\tau > \tau_\varepsilon$. Then, thanks to Lemma 6.2 we have that $\mathcal{E}[v(\tau)|V_1] < \infty$ for all $\tau > \tau_\varepsilon$. In this case, thanks to Lemma 6.5, we know that the function $\tau \mapsto \mathcal{I}[v(\tau)|V_1]$ is in $\mathrm{L}^1(\tau_0, \infty)$, for any $\tau_0 > 0$. This implies that $\mathcal{I}[v(\tau)|V_1] < \infty$ *only for almost every $\tau > \tau_\varepsilon$*. Our goal is to prove the following inequality for some $\vartheta > 0$:

$$2\,\vartheta \int_\tau^\infty \mathcal{E}[v(\tau)|V_1]\,\mathrm{d}\tau \le \mathcal{E}[v(\tau)|V_1] \qquad \forall\, \tau_0 > \tau_\varepsilon\,. \tag{93}$$

Indeed, by using (93) with Lemma 9.4, we can easily prove that $\mathcal{E}[v(\tau)|V_1] \le c e^{-2\,\vartheta\tau}$ for all $\tau \ge \tau_\varepsilon$. Then inequality (92) will follow from Lemma 9.1. In order to prove (93) let us start from identity (72), namely

$$\mathcal{E}[v(\tau)|V_1] = \int_\tau^\infty \mathcal{I}[v(\tau)|V_1]\,\mathrm{d}\tau \qquad \forall\, \tau > 0\,. \tag{94}$$

Let $A = \{\tau \in (\tau_\varepsilon, \infty) :\ \mathcal{I}[v(\tau)|V_1] = \infty\}$. Since the function $\tau \mapsto I[v(\tau)|V_1] \in \mathrm{L}^1(\tau_\varepsilon, \infty)$ the 1-dimensional measure of A is zero. We notice that, for every $\tau \in A$ and some $c > 0$ it holds $\mathcal{I}[v(\tau)|V_1] \ge c\,\mathcal{E}[v(\tau)|V_1]$, since $\mathcal{E}[v(\tau)|V_1]$ is finite for every $\tau > 0$. On the other hand, on the set A^c, by proceeding as in the case of Proposition 6.12, we can prove inequality (91) and, again with the use of Lemmata 6.6 and 6.4, we obtain

$$C_\varepsilon\, \tfrac{\mathcal{C}}{2\kappa_1}(1+\varepsilon)^{\gamma-2}\mathcal{E}[v(\tau)|V_1] \le \mathcal{I}[v(\tau)|V_1] \qquad \forall \tau \in A^c\,.$$

From the above inequality and, by considering what happens in A, one easily deduces (93) with $\vartheta := C_\varepsilon\, \frac{\mathcal{C}}{4\kappa_1}(1+\varepsilon)^{\gamma-2}$. □

6.6. Proofs of main results – Proposition 1.2 and Theorem 1

We focus now on the proof of polynomial rate of L^1-convergence of solutions to (CPLE) towards a Barenblatt profile.

Proof of Proposition 1.2. We observe that it is enough to consider solutions whose initial datum u_0 has initial mass equal to $M_\star$ for $M_\star$ as in (84), that is $\int_{\mathbb{R}^N} u_0(x)\,\mathrm{d}x = M_\star$. Indeed, let u be a weak solution to (CPLE) with u_0 as initial datum and let $M = \int_{\mathbb{R}^N} u_0(x)\,\mathrm{d}x$. Then, as previously observed in [20, Preliminaries], by defining

$$\widetilde{u}(t,x) := \frac{M_\star}{M}\, u\left(t\left(\frac{M_\star}{M}\right)^{2-p}, x\right) \tag{95}$$

one gets $\widetilde{u}$ being a solution to (CPLE) with initial datum $u_0\, M_\star/M$ and mass $M_\star$. Once (1.2) will be obtained for $\widetilde{u}$, we can rescale it back with the use of the identity (95) and obtain the same inequality for u. This follows since the same mass changing formula (95) applies also to the family of Barenblatt solutions. After this computation we get (1.2), where the constant $\widetilde{K}$ changes its value, but the rate ν remains the same.

In the same way, we observe that it is enough to consider inequality (11) for any among the Barenblatt solutions $\mathcal{B}_{M_\star}(t+T,\cdot)$ (for $T \ge 0$) and not necessarily $\mathcal{B}_{M_\star}(t,\cdot)$. Indeed, we may notice that there exists a constant $C = C(T, M_\star) > 0$ such that

$$\|\mathcal{B}_{M_\star}(t+T,\cdot) - \mathcal{B}_{M_\star}(t,\cdot)\|_{\mathrm{L}^1(\mathbb{R}^N)} \le C(T)t^{-1} \qquad \forall t > T\,. \tag{96}$$

Let us assume, for the moment, the above inequality. Then, we have that, from a convergence result with respect to the profile $\mathcal{B}_{M_\star}(t+T,\cdot)$, we retrieve

$$\begin{aligned} \|u(t,\cdot) - \mathcal{B}_{M_\star}(t,\cdot)\|_{\mathrm{L}^1(\mathbb{R}^N)} &\le \|u(t,x) - \mathcal{B}_{M_\star}(t+T,\cdot)\|_{\mathrm{L}^1(\mathbb{R}^N)} + \|\mathcal{B}_{M_\star}(t+T,\cdot) - \mathcal{B}_{M_\star}(t,\cdot)\|_{\mathrm{L}^1(\mathbb{R}^N)} \\ &\le \tfrac{\widetilde{K}}{t^\nu} + \tfrac{C(T)}{t} \le \tfrac{\widetilde{K}+C(T)}{t^\nu}\,, \end{aligned} \tag{97}$$

for any $t > \max\{1, T, \widetilde{T}\}$. We have used above that $\nu \le 1$. Let us justify it. The rate obtained in (96) is optimal, which can be proven by a direct computation. Since Proposition 1.2 also covers the case considered in (96), we conclude that $\nu \le 1$ by its optimality.

On the other hand, inequality (96) can be proven by considering the relative error between the two solutions $\mathcal{B}_{M_\star}(t+T,\cdot)$ and $\mathcal{B}_{M_\star}(t,\cdot)$. Indeed, for any fixed $t > 0$, the supremum in $|x|$ of the quotient $\mathcal{B}_{M_\star}(t+T,\cdot)/\mathcal{B}_{M_\star}(t,\cdot)$ is attained either at 0 or at ∞ (one can prove this through a simple, however lengthy, computation). From this observation one finds that

$$\left\| \frac{\mathcal{B}_{M_\star}(t+T,\cdot)}{\mathcal{B}_{M_\star}(t,\cdot)} - 1 \right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \le \frac{C(T, M_\star)}{t} \qquad \forall t > T\,. \tag{98}$$

It is also direct to see that the above inequality is optimal. Then, inequality (96) is obtained through the following computation

$$\int_{\mathbb{R}^N} |\mathcal{B}_{M_\star}(t+T,x) - \mathcal{B}_{M_\star}(t,x)|\,\mathrm{d}x \le \left\| \frac{\mathcal{B}_{M_\star}(t+T,\cdot)}{\mathcal{B}_{M_\star}(t,\cdot)} - 1 \right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \int_{\mathbb{R}^N} \mathcal{B}_{M_\star}(t,x)\,\mathrm{d}x \le M_\star\,\frac{C(T,M_\star)}{t}\,. \tag{99}$$

We also notice that the optimality in (96) can be deduced by a similar reasoning as above.

We will argue for ranges $p_M < p < 2$ and $p_c < p \le p_M$ separately. In both cases, as we have noticed before, it is sufficient to consider the mass of the initial datum being equal to $M_\star$.

Let us consider the case $p_M < p < 2$. As we explained, it is enough to consider to compute the rate of convergence towards the Barenblatt profile $\mathcal{B}_{M_\star}(t+\beta,\cdot)$. Under the current assumption, we can perform the change of variables (27) and consider a solution v to (29) with initial datum v_0 and mass $M_\star$. By Remark 6.1 (items *(ii)* and *(iii)*) we know that, for any $\varepsilon \in (0,1)$ there exists $\tau_\varepsilon > 0$ such that v satisfies (A1) (with ε and $D = 1$) for any $\tau > \tau_\varepsilon$. Therefore, all hypotheses of Proposition 6.12 are satisfied and inequality (90) holds true. By rescaling back inequality (90), one finds exactly (11). Hence, the claim is proven in this case.

Lastly, we are considering the case $p_c < p \le p_M$. Under the additional hypothesis, namely (9), we shall employ the profile $\mathcal{B}_{M_\star}(t+T,\cdot)$ for the convergence result. Therefore, in the change of variable (27), instead of using $R_\beta(t)$, we shall use $R_T(t)$ defined in (4) with T from (9). Let us define also

$$v(\tau, y) := R_T(t)^N\, u(t,x)\,, \tag{100}$$

where y and τ are as in (28). Then v is a solution to (29) with $v_0(y) = R_T(0)^N u_0(yR_T(0))$. We notice that $u_0 \in \mathcal{X}_p$, so by Remark 6.1 *(ii)* we know that, for any $\varepsilon \in (0,1)$, there exists $\tau_\varepsilon > 0$ such that $v(\tau,\cdot)$ satisfies (A1) for all $\tau > \tau_\varepsilon$. At the same time, thanks to assumption (9) and after the change of variables defined in, we get that the initial datum v_0 satisfies (A2). Therefore, $v(\tau,\cdot)$ satisfies both (A1) and (A2) for any $\tau > \tau_\varepsilon$. Eventually, by taking $v(\tau_\varepsilon,\cdot)$ as the initial datum, the assumptions of Proposition 6.13 are satisfied and, hence, (92) holds true. By re-scaling back, we find that for all t large enough it holds

$$\|u(t,x) - \mathcal{B}_{M_\star}(t+T,\cdot)\|_{\mathrm{L}^1(\mathbb{R}^N)} \le \widetilde{K} t^{-\nu}\,.$$

By inequality (99) we find the wanted result. □

We are in the position to justify our main accomplishment.

Proof of Theorem 1. Once Proposition 1.2 is proven, inequality (10) directly results from Theorem 7. □

7. Convergence of derivatives of radial solutions

This section is devoted to providing an exhaustive answer to (Q2) and prescribing an explicit polynomial rate for the uniform converge in relative error of radial derivatives of radial solutions in three cases separately. In particular we prove Theorem 2 for $p > p_c$, Theorem 3 for $p = p_c$, and Theorem 4 for $p < p_c$. We exploit here the relation between solutions to (16) and (17).

7.1. Information on related radial classical and weighted Fast Diffusion Equations

Below we present a radial equivalence due to [45, Theorem 1.2]. We point out, however, that we use a particular version of the result stated in [45]. The original result is indeed much stronger and valid also for sign-changing solutions. Notice also that we consider a slightly different equation, so the constant $\mathcal{D}$ defined below differs slightly comparing with the one of [45, Theorem 1.2]. In what follows we shall denote by $r = |x|$ in the (16) case and by $\varrho = |x|$ the coordinates for the (17) equation.

Theorem 8 ([45, Theorem 1.2])**.** *Suppose* $2 < n < \infty$*. If* u *is a radially symmetric and decreasing solution of equation* (16)*, then* Φ *being a non-negative solution of* (17) *is related to* u *through the following transformation:*

$$-\partial_r u(t,r) = \mathcal{D}\, \varrho^{\frac{2}{m+1}}\, \Phi(t,\varrho)\,, \quad \mathcal{D} = \left(\frac{2m}{m+1}\right)^{\frac{2}{m-1}}\,, \tag{101}$$

where $r = \varrho^{\frac{2m}{m+1}}$ *and the correspondence of the parameters is given by*

$$p = m+1, \quad N = (n-2)\frac{(m+1)}{2m}\,. \tag{102}$$

Let us note that in [45] the authors also analyze the case $0 < n < 2$, however, we shall not use those results. We also remark that, even if the solutions to (CPLE) are at least $C^{1,\alpha}$, a priori we do not know whether Φ is well-defined at the origin by transformation (101). We will address these issues below.

In fact, in Theorem 8 the following transformation is defined as

$$(p, N) \mapsto (m, n, \mathfrak{a}) := \left(p-1\,, 2+2\tfrac{N}{p'}\,, N-2-2\tfrac{N}{p'}\right)\,. \tag{103}$$

Note that $\mathfrak{a}+n = N$, so the above map is injective. Let us recall that $p_c = \frac{2N}{N+1}$ and $p_Y = \frac{2N}{N+2}$, and $p_Y < p_c$. We have the following ranges for different values of p:

(i) if $p_c < p < 2$ then
$$\tfrac{n-2}{n} < m < 1\,, \quad N+1 < n < N+2\,, \quad \text{and} \quad -2 < \mathfrak{a} < -1\,; \tag{104}$$

(ii) if $p = p_c$, then $m = \frac{N-1}{N+1} = \frac{n-2}{n}\,, \quad n = N+1\,, \quad$ and $\quad \mathfrak{a} = -1\,;$

(iii) if $p_Y < p < p_c$, then $\frac{n-2}{n+2} < m < \frac{n-2}{n}\,, \quad N < n < N+1\,, \quad$ and $\quad -1 < \mathfrak{a} < 0\,;$

(iv) if $p = p_Y$ (the Yamabe case), then $m = \frac{N-2}{N+2}\,, \quad n = N\,, \quad$ and $\quad \mathfrak{a} = 0\,.$

In general, the artificial dimension n is not an integer: this happens only in the limit case $p = p_c$ and $p = p_Y$. As already noticed in [45], the only case when $n = N$ (and also when the weight $\mathfrak{a} = 0$) is the case $p = p_Y$. We also remark that, when $N = 2$ the value $p_Y = 1$ and it is excluded from our analysis.

Note that the equation

$$\partial_t \Phi = |x|^{\mathfrak{a}}\, \mathrm{div}\left(|x|^{-\mathfrak{a}}\, \nabla \Phi^m\right)\,. \tag{WFDE}$$

written for radial solutions is exactly (17) for the choice of parameters from (103). This equation is sometimes referred to as the Weighted FDE with Caffarelli–Kohn–Nirenberg weights, see [13, 14, 19]. We also stress that radial initial data produce radial solutions for (WFDE). In order to infer the asymptotics of derivatives of radial solutions we exploit the relation between radial solution to (CPLE) and (WFDE) together with known properties of solutions to (WFDE). Therefore a solution to (17) is a radial solution to (WFDE), cf. [45] and Proposition 7.1. In the same spirit, we notice that Φ as a function of $(\varrho,\, t)$ is a radial solution to the original (unweighted) FDE when $n = N$. We stress that in (17), parameter n plays the role of an artificial dimension and is not an integer in general. It is unusual to consider equations in a continuous dimension, however, in the radial case, this allows us to unveil some unexpected features. Let us remark that equation (WFDE) shares many features with (CPLE), as it was already observed in [19]. For values of $m \in (0,1)$ there are three different ranges where solutions to (WFDE) behave differently. As for (CPLE), we call the interval $\frac{n-2}{n} < m < 1$ the *good range* corresponding to $p_c < p < 2$, as observed in (104). In this range solution to the Cauchy problem associated to (WFDE) conserve mass ($\int_{\mathbb{R}^N} \Phi(t,x)\, |x|^{-\mathfrak{a}}\, \mathrm{d}x = \int_{\mathbb{R}^N} \Phi(0,x)\, |x|^{-\mathfrak{a}}\, \mathrm{d}x$) once the initial datum $\Phi(0,x) \in \mathrm{L}^1(\mathbb{R}^N, |x|^{-\mathfrak{a}}\, \mathrm{d}x)$. Similarly, in this range (WFDE) also admits a family of self-similar solutions given by (105), that is commonly called Barenblatt solution, see [13, 14, 18, 19]. For parameters from (104) Barenblatt solutions are defined as follows

$$\mathfrak{B}_M(t,x) = t^{\frac{1}{1-m}} \left[\left(a_1\, t^{2\theta}\, M^{2\theta(m-1)} + a_2|x|^2\right)\right]^{\frac{1}{1-m}}, \quad \text{where} \quad \theta = \frac{1}{2-n(1-m)}\,. \tag{105}$$

The constants a_1 and a_2 given by

$$\int_{\mathbb{R}^N} (a_1 + a_2|x|^2)^{\frac{1}{1-m}}\, |x|^{-\mathfrak{a}}\, \mathrm{d}x = 1\,, \quad \text{while} \quad a_2 = \frac{1-m}{2m}\,\theta\,, \tag{106}$$

where θ is as in (105). We remark that the mass M of the profile $\mathfrak{B}_M$ is computed with respect to the measure $|x|^{-\mathfrak{a}}\,\mathrm{d}x$, that is

$$M = \int_{\mathbb{R}^N} \mathfrak{B}_M(t,x)\, |x|^{-\mathfrak{a}}\, \mathrm{d}x\,.$$

Lastly, as the reader may suspect, the Barenblatt profile $\mathcal{B}_M$ of equation (CPLE) and the one $\mathfrak{B}_{\overline{M}}$ of equation (WFDE) are related by formula (101), namely

$$-\partial_r \mathcal{B}_M(t,r) = \mathcal{D}\, \varrho^{\frac{2}{m+1}}\, \mathfrak{B}_{\mathfrak{C}M}(t,\varrho)\,, \tag{107}$$

where r, ϱ, and $\mathcal{D}$ are as in Theorem 8. We remark that the mass of $\mathfrak{B}$ is corrected by a multiplicative factor $\mathfrak{C} = \mathfrak{C}(N,p) > 0$ given by

$$\mathfrak{C} = \frac{pN}{2(p-1)\mathcal{D}}\,. \tag{108}$$

When $m \le \frac{n-2}{n}$, the Barenblatt solutions do not exist anymore as solutions generated by a δ_0 as initial datum. Nevertheless, a pseudo-Barenblatt profile is still available, for any $T > 0$ and $D > 0$ let us define

$$\mathfrak{B}_{D,T}(t,x) = \mathfrak{R}_T(t)^n\, \mathfrak{U}_D(x\,\mathfrak{R}_T(t)) \quad \text{where} \quad \mathfrak{U}_D(x) := \left(D + \frac{1-m}{2\,m}|x|^2\right)^{\frac{1}{1-m}}, \tag{109}$$

and where

$$\mathfrak{R}_T(t) := \left(\frac{T-t}{|\theta|}\right)_+^{\theta} \quad \text{if } 0 < m < \frac{n-2}{n} \quad \text{and} \quad \mathfrak{R}_T(t) := \exp\{\mathfrak{l}(t+T)\} \quad \text{if } m = \frac{n-2}{n}\,, \tag{110}$$

where θ is as in (105) (that is negative in the case $0 < m < \frac{n-2}{n}$) and $\mathfrak{l} > 0$ is a free parameter. We conclude this section by noticing that identity (107) also holds for the range $p_Y \le p \le p_c$, and we have that

$$-\partial_r \mathcal{B}_{D,T}(t,r) = \mathcal{D}\, \varrho^{\frac{2}{m+1}}\, \mathfrak{B}_{\overline{\mathfrak{C}}D,T}(t,\varrho)\,, \tag{111}$$

where $\overline{\mathfrak{C}} = \overline{\mathfrak{C}}(N,p) > 0$ is given by

$$\overline{\mathfrak{C}} = \left(\frac{1-m}{2m}\right)^{2\theta+1}. \tag{112}$$

Let us explain how solutions to (CPLE) are related to solutions to the Cauchy problem of (WFDE). While at the level of solutions this is given directly by the transformation (101), it is not clear what happens to the initial data and in what sense identity (101) should be understood. Before giving a complete answer, let us fix the notation which will be used in what follows. Let us denote by $u : (0,\infty) \times \mathbb{R}^N \to [0,\infty)$ the solution to (CPLE) with a radial initial datum $u_0(x)$. Since the solution $u(t,x)$ is radial, with an abuse of notation, we shall denote $u(t,x)$ by $u(t,r)$ and $u_0(x)$ by $u_0(r)$. The function $\Phi : (0,\infty) \times \mathbb{R}^N \to [0,\infty)$ will be a solution to (WFDE) with initial datum $\Phi_0(x)$. Again, in the case of radial initial datum $\Phi_0(x) = \Phi_0(r)$, we shall denote the solution $\Phi(t,x)$ by $\Phi(t,\varrho)$. The following proposition answers the main questions of this section.

Proposition 7.1. *Suppose that $N \ge 2$, $1 \le p < 2$, $\mathcal{D}$ is as in* (101)*, and n, m, $\mathfrak{a}$ are as in* (102)*. Let u be a radial solution to* (CPLE) *with a radial initial datum $u_0 \in C^2(\mathbb{R}^N)$ satisfying one of the following conditions:*

(i) if $p_c < p < 2$ we assume that (13) *holds for some $A > 0$ and $R_0 > 0$;*

(ii) if $N = 2$ and $1 = p_Y < p \le p_c$, or $3 \le N \le 6$ and $p_Y \le p \le p_c$, or $N > 6$ and $p_2 < p < p_c$, we assume that there exist D_1, $D_2 > 0$ and $T > 0$ such that (20) *holds;*

(iii) if $N > 6$ and $p_Y \le p \le p_2$, we assume that there exists D_1, $D_2 > 0$ and $T > 0$ such that (20) *holds and that there exist $\widetilde{D} > 0$ and $f \in \mathrm{L}^1((0,\infty), r^{n-1}\,\mathrm{d}r)$ with such that* (25) *holds.*

Then Φ_0 given by

$$\Phi_0(\varrho) := -\frac{1}{\mathcal{D}\varrho^{\frac{2}{1+m}}}\,(\partial_r u_0)\,(\varrho^{\frac{2m}{1+m}}) \qquad \forall \varrho > 0\,, \tag{113}$$

satisfies $0 \le \Phi_0 \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N, |x|^{-\mathfrak{a}}\,\mathrm{d}x)$ and the following Cauchy problem

$$\begin{cases} \partial_t \Phi = |x|^{\mathfrak{a}}\,\mathrm{div}\,(|x|^{-\mathfrak{a}}\,\nabla\Phi^m) & \textit{for} \;\; (t,x) \in (0,\infty)\times\mathbb{R}^N \\ \Phi(0,x) = \Phi_0(|x|) & \textit{for} \;\; x \in \mathbb{R}^N\,, \end{cases} \tag{114}$$

is solvable. Moreover, its solution $\Phi(t,\cdot)$ belongs to $\mathrm{L}^\infty_{\mathrm{loc}}(\mathbb{R}^N)$ for any $t > 0$ and it is related to u by transformation (101).

Proof. Let us start with justifying that for $p_Y \le p < 2$, the initial datum $0 \le \Phi_0 \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R}^N, |x|^{-\mathfrak{a}}\,\mathrm{d}x)$ for any $0 < q < \frac{n(1+m)}{2(1-m)}$. Simple computations show that $\frac{n(1+m)}{2(1-m)} > 1$ under the current assumptions. To motivate the abovementioned integrability of Φ_0, we recall that $u_0 \in C^2(\mathbb{R}^N)$, and u_0 is radial and decreasing, we have that $\partial_r u_r(0) = 0$ and $|\partial_r u_r(\varrho)| \le C\varrho$ close to the origin. Therefore, we find that

$$\varrho^{n-1}\,|\Phi_0(\varrho)|^q \le C\,\varrho^{\mathfrak{b}} \quad \text{where} \quad \mathfrak{b} = q\,\tfrac{2(p-2)}{p} + 2\,N\,\tfrac{(p-1)}{p} + 1 = n - 1 + q\tfrac{2(m-1)}{1+m}\,.$$

Note that $\mathfrak{b} > -1$ as long as $q < \frac{n(1+m)}{2(1-m)}$. We recall that $\int_{|x|\le 1} |\Phi_0(x)|^q |x|^{-\mathfrak{a}}\,\mathrm{d}x = \omega_N \int_0^1 |\Phi_0(\varrho)|^q\,\varrho^{N-1-\mathfrak{a}}\,\mathrm{d}\varrho = \omega_N \int_0^1 |\Phi_0(\varrho)|^q\,\varrho^{n-1}\,\mathrm{d}\varrho$, where ω_N is the area of the N-dimensional sphere.

Further we proceed case by case.

Case (i): $p_c < p < 2$. Let us consider the integrability of Φ_0. From the last inequality of (13) we deduce that for any $\varrho > R_0^{\frac{1+m}{2m}}$ we have

$$\Phi_0(\varrho) \le C\,\varrho^{-\frac{2}{1+m}-\frac{2}{1-m}\frac{2m}{1+m}} = C\,\varrho^{-\frac{2}{1-m}}\,. \tag{115}$$

Therefore, for any $\varrho > R_0^{\frac{1+m}{2m}}$ it holds that

$$\varrho^{n-1}\,\Phi_0(\varrho) \le C\,\varrho^{n-1-\frac{2}{1-m}}\,.$$

We notice that in this case ($p_c < p < 2$ equivalently to $\frac{n-2}{n} < m < 1$), we have that $\frac{2}{1-m} - n > 0$. Therefore, the quantity $|\varrho^{n-1}\,\Phi_0(\varrho)|$ is integrable with respect to the Lebesgue measure and the initial datum $\Phi_0 \in \mathrm{L}^1(\mathbb{R}^N, |x|^{-\mathfrak{a}}\,\mathrm{d}x)$.

We are in the position to pass to justification of solvability of (114). Let us notice that the equation in (114) is the same as in (17). As explained in (104), in the present range of parameters we always have $\mathfrak{a} < 0$. Solutions for problem (114) have been constructed in [14] in the same spirit as in [44, Theorem 2.1]. We stress that in [14, Proposition 7] the initial datum is assumed to be in $\mathrm{L}^\infty(\mathbb{R}^N)$. This assumption can be weakened to merely $\Phi_0 \in \mathrm{L}^1(\mathbb{R}^N, |x|^{-\mathfrak{a}}\,\mathrm{d}x)$ by a standard approximation procedure as it is done in the proof of [44, Theorem 2.1]. Since $\Phi_0 \ge 0$ and the comparison principle holds due to [14, Corollary 9], we know that the solution is non-negative. In this range of parameters solutions are bounded since $\Phi_0 \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^N, |x|^{-\mathfrak{a}}\,\mathrm{d}x)$, see [18, Theorem 1.2]. It is also known that solutions are at least C^α-regular close to the origin (see [18, Theorem 1.8]) and C^∞-smooth outside of the origin. This has been already remarked in [14, Lemma 11]. See also [36, Section 21.5.3] where the authors affirm that local analyticity in space and, at least, Lipschitz continuity in time holds for solutions to a general equation of the form (114). This considerations prove that the solution Φ to (114) exists and it has the wanted properties. It only remains to verify that a radial solution u to (CPLE) is related to the Cauchy problem (114) through the transformation (101). Despite this seems obvious, it is not since we need to find a relevant relation between initial data. Let us consider the auxiliary function

$$\overline{u}(t,r) := \mathcal{D}\int_r^\infty \Phi(t, s^{\frac{p}{2(p-1)}})\,s^{\frac{1}{p-1}}\;\mathrm{d}s\,, \tag{116}$$

where $\mathcal{D}$ is as in (101). By the global Harnack principle (see [19, Theorem 1.1] and cf. (47)) for solutions to the radial problem (114) and by (115) we infer that the following result holds: for any $\tau_0 > 0$ there exist $M_1(\tau_0), M_2(\tau_0) > 0$ and $\tau_1(\tau_0), \tau_2(\tau_0) > 0$, such that we have

$$\mathfrak{B}_{M_1}(t-\tau_1,\varrho) \le \Phi(t,\varrho) \le \mathfrak{B}_{M_2}(t+\tau_2,\varrho) \qquad \forall\varrho > 0 \quad \forall t \ge 2\,\tau_0\,. \tag{117}$$

From the above estimates, we deduce that there exists a constant $C = C(t, \tau_0) > 0$ such that for any $t \geq 2\tau_0$ we have

$$\Phi(t, s^{\frac{p}{2(p-1)}})\, s^{\frac{1}{p-1}} \leq C(t, \tau_0)\, \frac{s^{\frac{1}{p-1}}}{s^{\frac{2}{1-m}\frac{p}{2(p-1)}}} = C(t, \tau_0)\, s^{-\frac{2}{2-p}}\,, \quad \forall s > 0\,, \tag{118}$$

where $C(t, \tau) \lesssim t^{-n\,\theta}$ for any $t > 2\tau$ with θ as in (105). The exponent $2/(2-p) > 1$ since $2 > p > 1$, so we deduce from the above inequality that the function $\overline{u}$ is well defined. Furthermore, we have that $\overline{u}(t) \in \mathrm{L}^\infty(0, \infty)$ for any $t > 0$ (indeed, τ_0 is chosen arbitrarily). Let us now investigate the regularity of Φ. The validity of the inequality (117) allows us to use the regularity information resulting from the proof of [14, Lemma 11]. By those results we have that $\Phi \in C^\infty(0, \infty)^2$, and, for any $\tau > 0$, $\varepsilon > 0$ and $k > 0$ there exist $C_1 = C_1(t, \tau, \varepsilon) > 0$ and $C_2 = C_2(t, \tau, \varepsilon, k)$ such that

$$|\partial_t \Phi(t, \varrho)| \leq C_1(t) \varrho^{-\frac{2}{1-m}} \quad \text{and} \quad \left| \frac{\partial^k}{\partial \varrho^k} \big(\partial_t \Phi(t, \varrho)\big) \right| \leq C_2(t) \varrho^{-\frac{2}{1-m} - k} \quad \forall t \geq \tau \quad \forall \varrho \geq \varepsilon\,. \tag{119}$$

The above estimates allow us to differentiate in t and in ϱ under the sign of the integral in (116). Consequently, $\overline{u} \in C^2(0, \infty)^2$ and $\overline{u}$ solves equation (16) almost everywhere in $(t, r) \in (0, \infty)^2$. Furthermore, function $\overline{u}$ is a weak solution to the following Neumann problem

$$\begin{cases} \partial_t \overline{u} = r^{1-N}\, \partial_r \left(r^{N-1}\, |\partial_r \overline{u}|^{p-2}\, \partial_r \overline{u} \right) \quad \text{for} \ \ (t, r) \in (0, \infty)^2\ , \\ \partial_r \overline{u}(t, 0) = 0\,, \\ \overline{u}(0, r) = \overline{u}_0(r) \quad \text{for} \ \ r \in [0, \infty)\ . \end{cases} \tag{120}$$

We briefly comment on the literature for the above problem in Remark 7.2. Let us continue with the rest of the proof. By using the result of [37, Theorem III.8.1], we know that the function $(t, x) \mapsto \nabla u(t, x) \in C^\alpha_{\mathrm{loc}}((0, \infty) \times \mathbb{R}^N)$, which is enough to guarantee that $\partial_r u(t, 0) = 0$ for all $t > 0$. We conclude therefore that u also solves problem (120). Since $u_0 = \overline{u}_0$ we would like to conclude that $u = \overline{u}$ by using the uniqueness result for (120). This would be enough to conclude the proof, since by the construction we will have that Φ and u satisfy the relation (101). However, in order to apply the uniqueness result of [37, Theorem II.1] we need to ensure that

$$\partial_t \overline{u}(t, r) \leq \mathfrak{h}\, \overline{u}(t, r) \quad \text{a.e. } (t, r) \in (0, \infty)^2\,, \tag{121}$$

where $\mathfrak{h} = \mathfrak{h}(N, p, t)$ is independent of $\overline{u}$. We notice that, as observed in [37, p. 45], solutions to problem (CPLE) satisfy (121) by the construction. In order to prove (121) for solutions to (120) we shall use a modification of a trick due to Bénilan and Crandall [8] provided in [37, Lemma III.3.4]. By the comparison principle proven in [14], the uniqueness for (114) is guaranteed. Let us consider Ψ^λ being the unique solution to (114) with initial datum

$$\Psi_0^\lambda(\varrho) := \Psi^\lambda(0, \varrho) = \lambda^{\frac{1}{m-1}}\, \Phi_0(\varrho) \quad \text{for} \ \ \lambda > 0\,.$$

Notice that if $\lambda \geq 1$, then $\Psi_0^\lambda(\varrho) \leq \Phi_0(\varrho)$ for all $\varrho \geq 0$. The homogeneity of (114) implies that Ψ can be written as

$$\Psi^\lambda(t, \varrho) = \lambda^{\frac{1}{m-1}}\, \Phi(\lambda\, t, \varrho)\,.$$

Therefore, again by the comparison principle, we have that $\Psi^\lambda(t, \varrho) \leq \Phi(t, \varrho)$ for all $(t, \varrho) \in (0, \infty)^2$. By setting $\lambda = 1 + h/t$, for a small $h > 0$, we obtain that, for any $(t, \varrho) \in (0, \infty)^2$ it holds

$$\begin{aligned} \Phi(t + h, \varrho) - \Phi(t, \varrho) = \Phi(\lambda\, t, \varrho) - \Phi(t, \varrho) &= \lambda^{\frac{1}{1-m}}\, \lambda^{\frac{1}{m-1}} \Phi(\lambda\, t, \varrho) - \Phi(t, \varrho) = \lambda^{\frac{1}{1-m}}\, \Psi^\lambda(t, \varrho) - \Phi(t, \varrho) \\ &\leq \left(\lambda^{\frac{1}{1-m}} - 1 \right)\, \Phi(t, \varrho)\,. \end{aligned}$$

By the Mean Value Theorem we infer that for some $\zeta \in (0, \frac{h}{t})$ it holds

$$\Phi(t + h, \varrho) - \Phi(t, \varrho) \leq \frac{h}{(1-m)\, t}\, (1 + \zeta)^{\frac{m}{1-m}}\, \Phi(t, \varrho)\,.$$

By using the above inequality in $(t, s^{\frac{p}{2(p-1)}})$, multiplying by $\mathcal{D}$ (for $\mathcal{D}$ being as in (101)), and integrating with respect to the measure $s^{\frac{1}{p-1}}\, \mathrm{d}s$, we find that

$$\overline{u}(t + h, r) - \overline{u}(t, r) \leq \frac{\mathcal{D}\, h}{(1-m)\, t}\, (1 + \zeta)^{\frac{m}{1-m}}\, \overline{u}(t, r)\,. \tag{122}$$

We notice that, by the same computations with $\lambda \le 1$, we shall establish inequality (122) with a reversed sign and $h < 0$. Then we divide by h both sides of (122) and take the limit for $h \to 0$. Let us point out that $\overline{u} \in C^1(0,\infty)^2$, so the left-hand side of (122) converges to $\partial_t \overline{u}(t,r)$. In turn, we find estimate (121) with $t\,(1-m)\,\mathfrak{h} = \mathcal{D}$. The proof in the case $p_c < p < 2$ is complete.

Case (ii): $N = 2$ and $1 = p_Y < p \le p_c$, or $3 \le N \le 6$ and $p_Y \le p \le p_c$, or $N > 6$ and $p_2 < p < p_c$. We shall mainly explain the main differences between this case and the above one. Let us consider first the case $p < p_c$, we shall explain the main difference for $p = p_c$ in the end of the proof. First of all, we notice that, by assumption (20) and relation (111), there exists $\overline{D}_1, \overline{D}_2 > 0$ such that

$$\mathfrak{B}_{\overline{D}_1,T}(0,\varrho) \le \Phi_0(\varrho) \le \mathfrak{B}_{\overline{D}_2,T}(0,\varrho) \quad \forall \varrho \ge 0\,, \tag{123}$$

the case $\varrho = 0$ being obtained as a limit case. Notice also that $\overline{D}_i = \overline{\mathfrak{C}}\, D_i$, where D_i is as in (20) and $\overline{\mathfrak{C}}$ as in (112). As in the previous case, the result [14, Propostion 7] is enough to establish the existence of a non-negative solution $\Phi \in \mathrm{L}^\infty_{\mathrm{loc}}(\mathbb{R}^N)$. The comparison principle has been established in [14, Corollary 9] for initial data which satisfies (123) and the following assumption: there exists $\overline{D} > 0$ such that

$$\Phi_0(\varrho) = \mathfrak{B}_{\overline{D},T}(0,\varrho) + f(\varrho) \quad \forall \varrho \ge 0\,, \tag{124}$$

for $f \in \mathrm{L}^1((0,\infty), r^{n-1}\,\mathrm{d}r)$ (notice that parameter n defined in Theorem 4 is the same as in (102)). In the present range the assumption (124) easily follows from (123), since the difference of two Barenblatts $\mathfrak{B}_{\overline{D}_2,T} - \mathfrak{B}_{\overline{D}_1,T}$ is always integrable if $m > m_\star$, where

$$m_\star = \frac{n-4}{n-2} = \frac{2N(p-1) - 2p}{2N(p-1)}\,. \tag{125}$$

Two remarks are in order. First, the fact that the difference of two Barenblatt is integrable can be proven by technniques similar to those in Lemma 9.2, we also refer to [14, Section 2.1] and [9, Introduction] for a general discussion. Second, it is easy to see that $m > m_\star$ if and only if

$$2N(p-1)^2 - 2N(p-1) + 2p \ge 0 \quad \text{and} \quad p > 1\,. \tag{126}$$

A simple computation shows that the above condition always holds for $N < 6$. For $N \ge 6$ it is satisfied for $p \in (1,p_1) \cup (p_2, 2)$, which explains the appearance of the exponent p_2.

We deduce that, by the comparison principle, inequality (123) continues to hold for $t > 0$. More precisely

$$\mathfrak{B}_{\overline{D}_1,T}(t,\varrho) \le \Phi(t,\varrho) \le \mathfrak{B}_{\overline{D}_2,T}(t,\varrho) \quad \forall \varrho \ge 0 \quad \text{and} \quad 0 < t \le T\,, \tag{127}$$

which proves that $\Phi(t,\varrho) = 0$ for all $t \ge T$ and $\varrho \ge 0$. Inequality (127) plays the role of inequality (117) in this range of parameters. Indeed, from (127) one can deduce (118) for any $0 < t < T$ which is enough to establish that $\overline{u}$ is well-defined also in the present case. At the same time, and using again [14, Lemma 11], we have that $\Phi \in C^\infty(0,\infty)^2$ and inequalities from (119) hold also in the present case. This is enough to show that $\overline{u}$ is a weak solution to (120). At the same time, using the same argument as above, one can easily prove inequality (121), which is the missing condition to verify in order to use the uniqueness result of [37, Theorem II.1]. We have explained in Section 4.2 that, in our setting, for a solution u to (CPLE), the function $(t,x) \mapsto \nabla u(t,x) \in C^\alpha((0,\infty)\times\mathbb{R}^N)$: so we are in the position to guarantee that $\partial_r u(t,0) = 0$ (which means that u is also a solution to (120)) and therefore, by uniqueness, we have that $u = \overline{u}$, which concludes the proof in the case $p < p_c$.

In the case $p = p_c$, the above proof is also valid. The only thing that changes is that (127) holds for any $t > 0$. In this case, the solution lives for all $t > 0$ as for $p_c < p < 2$ and there is no extinction in finite time.

Case (iii): $N \ge 6$ and $p_Y \le p \le p_2$. The present case is very similar to *Case (ii)*. Indeed, the main difference is that identity (124) is not a consequence of (123) but instead it needs to be assumed from the very beginning. Nevertheless, notice that (124) is exactly assumption (25) rewritten after the use of transformation (101). The rest of the proof follows exactly the lines of *Case (ii)*. Therefore the proof is complete.

□

Remark 7.2 (On the Neumann problem)**.** To the best of our knowledge, the problem (120) has not been investigated yet. It seems, as well, that the Neumann problem for p-Laplacian type of equation has been much less studied. For more information we refer to [6, 5] in the case of a Neumann problem in bounded domains, to [4] for the Neumann problem for the Porous Medium Equation ($\partial_t u = \Delta u^m$, $m > 1$), and to [61, Chapter 11] for exposition of the background in detail. In dimension $N = 1$, the techniques used in the seminal paper [38] can be adapted (at least in the good range $p_c < p < 2$) in order to prove the existence, uniqueness and comparison principle. We also stress that problem (120) is very similar to (CPLE) and the techniques of [37] can be adapted in the whole generality for the entire range $1 < p < 2$.

7.2. Proof of the convergence in the relative error of the radial derivatives

It is convenient to rescale (WFDE) in the way we are able to consider at the same time the supercritical, critical and subcritical range. The following change of variables is very much in the same spirit of (27). Consider Φ to be a solution to (WFDE) and let us define Ψ as

$$\Psi(\tau, y) := \mathfrak{R}_T(t)^n\, \Phi(t, x) \quad \text{where} \quad \tau = \log \frac{\mathfrak{R}_T(t)}{\mathfrak{R}_T(0)} \quad \text{and} \quad y := \frac{x}{\mathfrak{R}_T(t)}\,, \tag{128}$$

where $\mathfrak{R}_T$ is as in (110). We recall that the definition of $\mathfrak{R}_T$ differ when $p = p_c$ ($m = m_c$) and $p < p_c$ ($m < m_c$). However, in both cases, if Φ satisfies (WFDE) then the problem satisfied by Ψ is the following

$$\begin{cases} \partial_\tau \Psi = |y|^{\mathfrak{a}}\, \mathrm{div}\big[|y|^{-\mathfrak{a}}\, (\nabla \Psi^m - y\, \Psi)\big] & \text{for} \ \ (\tau, y) \in (0, \infty) \times \mathbb{R}^N\,, \\ \Psi(0, y) = \Psi_0(|y|) \quad \text{for} \ \ y \in \mathbb{R}^N\,, \end{cases} \tag{129}$$

where the initial datum $\Psi_0(y) = \mathfrak{R}_T(0)^n\, \Phi_0(x\, \mathfrak{R}_T(0)^n)$.

There are two main advantages which justify the introduction of the change of variables of (128). The first reason is that, in the case $p < p_c$ ($m < m_c$), on the contrary to the solution to (WFDE) which extinguishes in finite time T (as does the Barenblatt function), the rescaled solution Ψ lives for any $0 < \tau < \infty$. The second reason is that (129) admits the stationary solution

$$\mathfrak{U}_D(x) := \left(D + \frac{1-m}{2\,m}|x|^2\right)^{\frac{1}{1-m}}\,, \qquad D > 0\,,$$

introduced in (109). We notice that, when $m > \frac{n-2}{n}$, the parameter D is related to the mass of $\mathfrak{U}_D$, i.e. $\int_{\mathbb{R}^N} \mathfrak{U}_D |x|^{-\mathfrak{a}}\, \mathrm{d}x$.

We notice that conditions (14), (20) and (123) imply the existence of $\overline{D}_1$, $\overline{D}_2 > 0$ such that

$$\mathfrak{U}_{\overline{D}_1}(\varrho) \le \Psi_0(\varrho) \le \mathfrak{U}_{\overline{D}_2}(\varrho) \quad \forall \varrho \ge 0\,, \tag{130}$$

while, condition (25) (or, equivalently, (124)) translates to the existence of $\overline{D} > 0$ such that

$$\Psi_0(\varrho) = \mathfrak{U}_{\overline{D}}(\varrho) + f(\varrho) \quad \forall \varrho \ge 0\,, \tag{131}$$

where $f \in \mathrm{L}^1((0, \infty)\,, r^{n-1}\, \mathrm{d}r)$.

In what follows, we refer to the result [14, Theorem 5 for $\gamma \le 0$], which can be stated in our language as follows. Under assumptions (130) and (131) there exist $D > 0$, $\tau_\bullet > 0$, $C_\bullet > 0$, and $\Lambda > 0$ such that

$$\left\| \frac{\Psi(\tau)}{\mathfrak{U}_{\overline{D}}} - 1 \right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \le C_\bullet\, e^{-2\,\frac{(1-m)^2}{2-m}\Lambda(\tau - \tau_\bullet)} \quad \forall \tau \ge \tau_\bullet \tag{132}$$

The convergence rate $\Lambda = \Lambda(n, m)$ is the optimal constant in a relevant Hardy–Poincaré inequality related to (WFDE). We refer to [14, Proposition 3] for more information, see also [24]. For the sake of completeness, we state here the different values of Λ for various parameters. If $m \le \frac{n}{n+2}$ (notice that $m = \frac{n}{n+2}$ means $p = p_M$), then

$$\Lambda = \Lambda_{\mathrm{ess}} := \frac{\big((n-2)(1-m) - 2\big)^2}{4(1-m)^2}\,,$$

while, when $m > \frac{n}{n+2}$ ($p > p_M$), then

$$\Lambda = \min\left\{\Lambda_{\mathrm{ess}}\,, \frac{2\,\eta}{1-m}\,, \frac{2(2 - n(1-m))}{1-m}\right\} \quad \text{where} \quad \eta := \sqrt{N - 1 + \left(\frac{N-2-\mathfrak{a}}{2}\right)^2} - \frac{N-2-\mathfrak{a}}{2}\,.$$

A detailed inspection of the proof reveals that the time $\tau_\bullet$ cannot be quantified a priori and depends on the initial datum Φ_0, see in particular [14, Proposition 14 and Section 3.2]. However, the constant $C_\bullet$ can be explicitly quantified and it depends on the initial datum Ψ_0 (through its entropy), the parameters $\overline{D}_1$, $\overline{D}_2$ and, of course, n and $\mathfrak{a}$, see in particular [14, Proof of Theorem 4].

Proof of Theorem 2. Since $p > p_c$ and we assume (13), we can make use of Proposition 7.1 *(i)*. Namely, the solution u is related to the solution Φ of (114) by the transformation (101). We notice that assumption (14) is more restrictive than (13), so also in this case Proposition 7.1 *(i)* applies. Then, by [20, Theorem 1.4], we know that

$$\left\|\frac{\partial_r u(t,\cdot)}{\partial_r \mathcal{B}_M(t,\cdot)} - 1\right\|_{\mathrm{L}^\infty(0,\infty)} = \left\|\frac{\Phi(t,\cdot)}{\mathfrak{B}_{\mathfrak{C}M}(t,\cdot)} - 1\right\|_{\mathrm{L}^\infty(0,\infty)} \longrightarrow 0 \quad \text{as} \quad t \longrightarrow \infty. \tag{133}$$

Since the mass of Φ is conserved in time due to [14, Proposition 10], the above display shows that

$$\mathfrak{C}M = \int_{\mathbb{R}^N} \Phi(t,x)\,|x|^{-\mathfrak{a}}\,\mathrm{d}x = \omega_N \int_0^\infty \Phi(t,\varrho)\,\varrho^{n-1}\,\mathrm{d}\varrho\,.$$

To conclude our proof, we only need to obtain a convergence rate towards zero for the uniform relative error which appears in the middle of (133). In the case $\frac{n}{n+2} \le m < 1$ (recall that $m = \frac{n}{n+2}$ means $p = p_M$), this can be inferred from [16]. More precisely, under a condition that we shall discuss below, from [16, Theorem 7] it follows that there exist explicit constants $\varepsilon_\star = \varepsilon_\star(m, N, \mathfrak{a}) > 0$, $C_\star = C_\star(m, N, \mathfrak{a}, \Phi_0) > 0$, and $\lambda = \lambda(m, N, \mathfrak{a})$ such that for any $0 < \varepsilon < \varepsilon_\star$ we have

$$\left\|\frac{\Phi(t,\cdot)}{\mathfrak{B}_{\mathfrak{C}M}(t,\cdot)} - 1\right\|_{\mathrm{L}^\infty(0,\infty)} \le \varepsilon \qquad \forall t \ge C_\star \varepsilon^{\frac{1}{\lambda}}\,. \tag{134}$$

While the result in [16, Theorem 7] is stated only for $\frac{n-1}{n} \le m < 1$, the method can be easily extended up to $m = \frac{n}{n+2}$ since it is based on a weaker form of assumption (130). Indeed, the proofs in [16] are based on the Global Harnack principle for equation (WFDE), namely inequality (117), and the fact that the second moment with respect to the measure $|x|^{-\mathfrak{a}}$ is finite, which holds exactly for $\frac{n}{n+2} < m < 1$.

Once inequality (134) is obtained, to establish the convergence rate of (12) it is enough to inverse the relation between ε and t. This has been done in detail in [15, Corollary 4.14]. In the case $\frac{n-2}{n} < m < \frac{n}{n+2}$, to infer the rate of convergence in (133), we invoke [14, Theorem 5], indeed, that result guarantee an explicit convergence rate for (133) (see inequality (132)) when the initial datum Ψ_0 satisfies both (130) and (131). We remark that assumption (130) is nothing than (14). Lastly, that in the current regime $\frac{n-2}{n} < m < \frac{n}{n+2}$, assumption (131) can be easily obtained from (130), since in this regime the difference of two Bareblatts profile is always integrable.

To conclude the proof, let us briefly comment on the last restriction of [16, Theorem 7]: the initial datum should satisfy

$$\|\Phi_0\|_{\mathcal{Y}_m} := \sup_{R>0} R^{\frac{2}{1-m}-n} \int_{|x|>R} \Phi_0(x)\,|x|^{-\mathfrak{a}}\,\mathrm{d}x\, < \infty\,. \tag{135}$$

We stress that the condition $\|\Phi_0\|_{\mathcal{Y}_m} < \infty$ plays the same role for (114) as (7) for (CPLE). This has already been pointed out in [19]. In our setting Φ_0 verifies (135) due to Proposition 7.1. We also remark that such a condition is satisfied uniformly in $\frac{n-2}{n} < m < 1$. $\square$

Proof of Theorems 3 and 4. We consider first the case $p < p_c$. By reversing the change of variables (128) and using the convergence rate (132) one gets inequality

$$\left\|\frac{\Phi(t,\cdot)}{\mathfrak{B}_{\overline{D},T}(t,\cdot)} - 1\right\|_{L^\infty(\mathbb{R}^N)} \le C_\diamond\,(T-t)^{-\lambda} \quad \forall\; t_\diamond < t < T\,, \tag{136}$$

where $t_\diamond$ is such that

$$\tau_\bullet = \log\frac{\mathfrak{R}_T(t_\diamond)}{\mathfrak{R}_T(0)}\,, \tag{137}$$

$$C_\diamond = \frac{C_\bullet\, e^{2\,\frac{(1-m)^2}{2-m}\,\Lambda\,\tau_\bullet}}{|\theta|^{\frac{2\,\theta\,(1-m)^2\,\Lambda}{2-m}}}\,\mathfrak{R}_T(0)^{\frac{2\,\theta\,(1-m)^2\,\Lambda}{2-m}}\,, \qquad \text{and} \qquad \lambda = \frac{2\,\theta\,(1-m)^2}{2-m}\,\Lambda\,. \tag{138}$$

In the case $p = p_c$, the only change in inequality (136) is the fact that the right-hand side is of the form $t^{-\lambda}$ and the inequality holds for any $t \ge t_\diamond$. Once (136) is obtained one can easily obtain the first inequality of (26) (respectively, the first inequality of (21)) by using the relations (111) and (101).

It only remains to prove the second inequality of (26) (respectively, of (21)). We notice that, inequality (136) can be rewritten in the following form. For any $T > t > t_\diamond$ and $\varrho \geq 0$ it holds

$$-\varepsilon(t)\,\mathfrak{B}_{\overline{D},T}(t,\varrho) \leq \Phi(t,\varrho) - \mathfrak{B}_{\overline{D},T}(t,\varrho) \leq \varepsilon(t)\,\mathfrak{B}_{\overline{D},T}(t,\varrho) \quad \text{where} \quad \varepsilon(t) = C_\diamond\,(T-t)^{-\lambda}\,.$$

By integrating the above inequality as in (116), using the relation between u and Φ (explained in Proposition 7.1) and the relation among the different Barenblatt solutions (exposed in (109)), one finds the following link between u and $\mathcal{B}_{D,T}$:

$$-\varepsilon(t)\,\mathcal{B}_{D,T}(t,r) \leq u(t,r) - \mathcal{B}_{D,T}(t,r) \leq \varepsilon(t)\,\mathcal{B}_{D,T}(t,r) \quad \forall\, r \geq 0 \quad \forall\, t_\diamond < t < T\,,$$

which is equivalent to the second inequality of (26) (respectively, of (21) upon choosing $\varepsilon(t) = C_\diamond\, t^{-\lambda}$). The proof is then concluded. □

8. Proofs of Theorem 5 and 6 capturing in particular the subcritical case $p \leq p_c$

Proof of Theorem 5. Through the proof we shall assume condition (23) and shall not distinguish the cases $p \leq p_c$ or $p_c < p < 2$. The proof works in both cases. We shall follow several steps from [3], which relies on the ideas of [9].

Step 1): identification of the limit when $\tau \to \infty$. For this step, we follow mainly [2, Lemma 2.5]. We will first prove that $v(\tau,\cdot)$ converges to V_D pointwise and in L^p-norms. As in [2, Lemma 2.5], let us define $v^h(\tau,y) := v(\tau+h,y)$, for any given $h > 0$ and $\tau \in [0,1]$. By the comparison principle $\{v^h\}$ is uniformly bounded thanks to assumption *(i)*. Furthermore, it is uniformly continuous in $[0,1]\times B_R$ thanks to assumption *(iv)*. By the Ascoli–Arzelá Theorem, for any sequence $h_n \to \infty$ (as $n \to \infty$) the sequence of functions $\{v^{h_n}\}$ converges uniformly (up to a subsequence) to a function v^∞ on compact subsets of $[0,1]\times\mathbb{R}^N$. Moreover, we infer that for any $R > 0$ it holds that $\|v\|_{C^{1,\alpha}([0,1]\times B_R)} < \infty$ and for $\tau \in [0,1]$ function $v^\infty(\tau,\cdot)$ satisfies *(i)*. Since $N < \frac{p}{(2-p)(p-1)}$ we know that $V_D^{\gamma-2}(v - V_D)^2 \in \mathrm{L}^1(\mathbb{R}^N)$ (cf. (64) and the end of the proof of Lemma 6.7). Therefore, by using the arguments of Lemma 6.6, we get that $\mathsf{E}[v], \mathcal{E}[v|V_D] < \infty$. Thanks to assumption *(iii)* the entropy functional $\mathcal{E}[v|V_D]$ is non-negative and $\tau \mapsto \mathcal{E}[v(\tau)|V_D]$ is decreasing in time. By the time monotonicity, $\mathcal{E}[v^{h_n}(\tau)|V_D]$ and $\mathcal{E}[v^{h_n+1}(\tau)|V_D]$ have the same limit for $h_n \to \infty$. Therefore we infer that

$$\int_0^1 \mathcal{I}[v^{h_n}(\tau)|V_D]\,\mathrm{d}\tau = \int_{h_n}^{h_n+1} \mathcal{I}[v(\tau)|V_D]\,\mathrm{d}\tau = \mathcal{E}[v(h_n)|V_D] - \mathcal{E}[v(h_n+1)|V_D] \xrightarrow[n\to\infty]{} 0\,.$$

By the positivity of $\mathcal{I}[v^\infty(\tau)|V_D]$ and Fatou's Lemma, we infer therefore that $0 \geq \int_0^1 \mathcal{I}[v^\infty(\tau)|V_D]\,\mathrm{d}\tau = 0$. Consequently, $\nabla(v^\infty)^{\gamma-1} = \nabla V_{D^\star}^{\gamma-1}$ for some $D^\star > 0$ and so $v^\infty = V_{D^\star}$. Up to now, we have proven that v^{h_n} converges pointwise towards $V_{D^\star}$ as $n \to \infty$. We only need to ensure that $D^\star = D$. By Lemma 9.2, we are in the position of using the Dominated Convergence Theorem to infer that v^{h_n} converges in $\mathrm{L}^1(\mathbb{R}^N)$ which implies that

$$\lim_{n\to\infty} \int_{\mathbb{R}^N} \left(v^{h_n}(y) - V_{D^\star}(y)\right)\,\mathrm{d}y = 0\,.$$

At the same time, the above identity implies that necessarily $D^\star = D$. Indeed, otherwise one would find that $\int_{\mathbb{R}^N}(V_{D^\star}(y) - V_D(y))\,\mathrm{d}y = 0$, which leads to a contradiction in the case $D \neq D^\star$. Lastly, we observe that the limit does not depend on the sequence $\{v^{h_n}\}$ since the above reasoning is true for any possible convergent subsequence. Therefore, we conclude that v converges to V_D in the L^1-topology as $\tau \to \infty$.

Step 2): from convergence in $\mathrm{L}^1(\mathbb{R}^N)$ to convergence in L^∞. In this step we follow the ideas of [2, Lemma 2.6] with a few differences. We notice that, by assumption *(i)*, the L^∞-norm of the function $y \mapsto |v(\tau,y) - V_D(\tau,y)|$ is bounded uniformly in τ, see again Lemma 9.2. Therefore, by interpolation, one obtain that $v(\tau)$ converges to V_D in the $\mathrm{L}^q(\mathbb{R}^N)$-topology, for any $1 \leq q < \infty$. The convergence in $\mathrm{L}^\infty(\mathbb{R}^N)$ is more subtle. We shall first prove this convergence on balls. Let $R > 0$ and $\mathfrak{d} \in (0,1)$. For any function $f \in C^{\mathfrak{d}}(B_{2R}) \cap \mathrm{L}^1(B_{2R})$ we have the following interpolation inequality whose proof can be found in [15]:

$$\|f\|_{\mathrm{L}^\infty(B_R)} \leq C_{N,\mathfrak{d}}\left(\|f\|_{C^{\mathfrak{d}}(B_{2R})}^{\frac{N}{N+\mathfrak{d}}}\,\|f\|_{\mathrm{L}^1(B_{2R})}^{\frac{\mathfrak{d}}{N+\mathfrak{d}}} + R^{-N}\|f\|_{\mathrm{L}^1(B_{2R})}\right)\,. \tag{139}$$

Let us fix $\varepsilon > 0$. By using (139), assumption *(iv)*, and the already proven L^1-convergence, we infer that there exists $\tilde{\tau} = \tilde{\tau}(\varepsilon, v_0) > 0$ such that

$$\|v(\tau) - V_D\|_{\mathrm{L}^\infty(B_R)} < \varepsilon \quad \forall\, \tau > \tilde{\tau}\,. \tag{140}$$

Assume further that $R > C^{-1}\,\varepsilon^{\frac{-(p-1)(2-p)}{p}}$ where C is as in Lemma 9.2. Thanks to assumption *(i)* and by Lemma 9.2, we infer that

$$|v(\tau, y) - V_D(y)| \le \varepsilon \quad \forall\, |y| \ge R\,. \tag{141}$$

Inequalities (140) and (141) imply that v converges to V_D in $\mathrm{L}^\infty(\mathbb{R}^N)$ as $\tau \to \infty$.

Step 3): from convergence in $\mathrm{L}^\infty(\mathbb{R}^N)$ *to convergence in uniform relative error.* We will prove first the convergence of the relative error in the L^∞-norm and then obtain the general result by an interpolation argument. By using again Lemma 9.2, one can infer that the relative error decays as follows

$$\left|\frac{v(\tau, y) - V_D(y)}{V_D(y)}\right| \le \kappa_1 \frac{|x|^{\frac{p}{2-p}}}{|x|^{\frac{p}{(p-1)(2-p)}}} = \frac{\kappa_1}{|x|^{\frac{p}{p-1}}} \qquad \forall\, |y| \ge 1\,. \tag{142}$$

On a ball of radius $R > 0$, from the $\mathrm{L}^\infty(\mathbb{R}^N)$-convergence, one can infer the following:

$$\sup_{|y|\le R}\left|\frac{v(\tau, y) - V_D(y)}{V_D(y)}\right| \le \kappa_2\, \|v(\tau, y) - V_D(y)\|_{\mathrm{L}^\infty(\mathbb{R}^N)}\, R^{\frac{p}{2-p}}\,. \tag{143}$$

For fixed $\varepsilon > 0$ there exist $R_\varepsilon > 0$ and $\tau_\star = \tau_\star(\varepsilon, v_0) > 0$ such that

$$\kappa_1\, R_\varepsilon^{\frac{-p}{p-1}} < \varepsilon \quad \text{ and } \quad \kappa_2\, \|v(\tau, y) - V_D(y)\|_{\mathrm{L}^\infty(\mathbb{R}^N)}\, R_\varepsilon^{\frac{p}{2-p}} < \varepsilon\,.$$

Combining together inequalities (142) and (143) we get that for any $\varepsilon > 0$ it holds

$$\left\|\frac{v(\tau, y) - V_D(y)}{V_D(y)}\right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \le \varepsilon \quad \forall\, \tau \ge \tau_\star\,, \tag{144}$$

which justifies the uniform convergence in the relative error. It only remains to prove the convergence of the relative error in $\mathrm{L}^q(\mathbb{R}^N)$. Notice that the relative error $\frac{v(\tau,\cdot)-V_D(\cdot)}{V_D(\cdot)}$ is uniformly bounded in space and, thanks to inequality (142), it is integrable for any $q > N\,\frac{(p-1)}{p}$. Indeed, for $\delta > 0$ such that $2\,\delta < q - N\,\frac{(p-1)}{p}$ we have the following inequality

$$\left\|\frac{v(\tau, y) - V_D(y)}{V_D(y)}\right\|_{\mathrm{L}^q(\mathbb{R}^N)} \le \left\|\frac{v(\tau, y) - V_D(y)}{V_D(y)}\right\|_{\mathrm{L}^\infty(\mathbb{R}^N)}^{q - N\frac{p-1}{p} - \frac{\delta}{2}} \int_{\mathbb{R}^N} \left|\frac{v(\tau, y) - V_D(y)}{V_D(y)}\right|^{N\frac{(p-1)}{p} + \frac{\delta}{2}} \mathrm{d}y \xrightarrow[\tau\to\infty]{} 0\,, \tag{145}$$

where we used the fact that $\int_{\mathbb{R}^N}\left|\frac{v(\tau,y)-V_D(y)}{V_D(y)}\right|^{N\frac{(p-1)}{p}+\frac{\delta}{2}} \mathrm{d}y$ is uniformly bounded in time. The proof is complete. $\square$

Proof of Theorem 6. The strategy of the proof is to obtain first a convergence rate of the convergence in the L^1-topology and then improve it to the final result (38). We remind that we stay under condition (23) and that the assumptions of Theorem 5 are satisfied. From the convergence result (36), and assumption *(i)*, we deduce that there exists $\tau_\bullet = \tau_\bullet(v_0, D_1, D_2) > 0$ such that (A0), (A1) (for some $\varepsilon > 0$), and (A2) hold for $v(\tau, y)$ with every $\tau > \tau_\bullet$ and $y \in \mathbb{R}^N$. Since we assume the decay condition (37), we can make use of the lines of the proof of Lemma 6.2 to justify that $\mathcal{E}[v(\tau)|V_D] < \infty$. Since $N < \frac{p}{(2-p)(p-1)}$ we know that $V_D^{\gamma-2}(v - V_D)^2 \in \mathrm{L}^1(\mathbb{R}^N)$. Analogously, using the arguments of Lemma 6.6, we get that $\mathsf{E}[v], \mathcal{E}[v|V_D] < \infty$, and $(1+\varepsilon)^{\gamma-2}\,\mathsf{E}[v] \le \mathcal{E}[v|V_D] \le (1-\varepsilon)^{\gamma-2}\,\mathsf{E}[v]\,.$

Let us now clarify the relation between the entropy and the Fisher information, both non-linear and linearised versions. We notice that under assumptions (A1), (A2) and *(ii)* we are able to repeat the proof of Lemma 6.10 that implies that $\mathcal{C}(p, D, \varepsilon)\mathsf{E}[v] \le \mathsf{I}^{(\varepsilon)}[v]$. At the same time, we get that $\mathcal{I}[v(\tau)|V_D] \ge C_\varepsilon\, \mathsf{I}_\gamma^{(\varepsilon)}[v(\tau)]$ via Lemma 6.8 and the reasoning of Lemma 6.7, where we make use of (37) in the place of (50). Collecting the above we infer that $\frac{d}{d\tau}\mathcal{E}[v(\tau)|V_D] \le -c\mathcal{E}[v(\tau)|V_D]$, which via the Gronwall Lemma allow to state that for all $\tau > 0$ it holds $\mathcal{E}[v(\tau)|V_D] \le e^{-\vartheta\,\tau}\mathcal{E}[v_0|V_D]$ for some $\mathfrak{K} = \mathfrak{K}(p, N, \varepsilon)$. On the other hand, due to *(i)* and *(ii)* and Csiszár–Kullback inequality provided in Lemma 9.1, we know that $\|v - V_D\|^2_{\mathrm{L}^1(\mathbb{R}^N)} \le c(V_D, p)\mathcal{E}[v(\tau)|V_D]$, where the right-hand side is finite. Therefore, we get that there exists $\widetilde{T} > 0$ and $\widetilde{K} > 0$, such that we have that

$$\|v(\tau,\cdot) - V_D(\cdot)\|_{\mathrm{L}^1(\mathbb{R}^N)} \le \widetilde{K}\, e^{-\mathfrak{K}\tau/2} \qquad \forall\, \tau \ge \widetilde{T}\,. \tag{146}$$

We can now get a convergence rate in the uniform relative error. Let $\varepsilon > 0$ and $R = \left(\frac{\kappa_1}{\varepsilon}\right)^{\frac{(p-1)}{p}}$ where κ_1 is as in (142). Then we obtain from inequality (142) that

$$\left\| \frac{v(\tau,y) - V_D(y)}{V_D(y)} \right\|_{\mathrm{L}^q(\mathbb{R}^N)} \le \varepsilon \quad \forall\, |y| > R \;\forall\, \tau > 0\,.$$

At the same time, by using an interpolation inequality between the L^∞, C^1, and L^1 norms on $\mathbb{R}^N$ (directly resulting from (139) by taking the limit $R \to \infty$) one finds that for any $\tau \ge \tau_0$ (where τ_0 is as in *(iv)*) we have

$$\|v(\tau) - V_D\|_{\mathrm{L}^\infty(\mathbb{R}^N)} \le C_N \|v(\tau,\cdot) - V_D(\cdot)\|^{\frac{N}{N+1}}_{C^1(\mathbb{R}^N)} \|v(\tau,\cdot) - V_D(\cdot)\|^{\frac{1}{N+1}}_{\mathrm{L}^1} \le C_N\, C(\alpha, \tau_0, v_0)\, \widetilde{K}\, e^{-\mathfrak{K}\tau/2}\,,$$

where $C(\alpha, \tau_0, v_0)$ is as in *(iv)*. Combining the above estimate with (143) one get (144) with $\tau_\star = \max\{\tau_0, -\frac{2}{\mathfrak{K}} \log(H\, \varepsilon^{\frac{3-2p}{2-p}})\}$, for a constant $H = H(\alpha, \tau_0, v_0) > 0$. We notice that for ε small enough we have that $-\frac{2}{\mathfrak{K}} \log(H\, \varepsilon^{\frac{3-2p}{2-p}}) > \max\{0, \tau_0\}$. Once inequality (144) is obtained with an explicit functional relation between ε and $\tau_\star(\varepsilon)$, one can compute the rate of convergence by inverting this relation, as it was done, for instance, in [15, Corollary 4.14]. This is enough to obtain the convergence result (38) for the L^∞-norm. The result in the L^q-norm is obtained by interpolation as it is done in the proof of Theorem 5. The proof is complete. □

9. Comments, Extensions, and Open Problems

In this paper we presented several results on the long-time behaviour for solutions to (CPLE). Let us summarize our results and open problems in the view of Questions (Q1), (Q2), and (Q3) from Introduction. They open several directions in which we our study might be extended.

(i) Theorem 1 and Proposition 1.2 give convergence rate towards the Barenblatt solutions. It is clear, from the examples in Introduction, that our rates are not sharp. It is, therefore, an interesting open problem to obtain the optimal rates.

(ii) In the range $p_c < p < p_M$, the entropy method requires an additional assumption (i.e. *(ii)* of Theorem 1) in order to give convergence rates (both for the L^1-norm and the uniform relative error). It is known, however, that when $u_0 \in \mathcal{X}_p$, solutions still converge to the Barenblatt in the uniform relative error. We pose a question: is it possible to obtain the convergence rate without the additional assumption *(ii)* of Theorem 1? This does not seem an easy task. It was done in the case of the fast diffusion equation by exploiting the (very good) regularity properties of solutions in that case, see [29] and the shortest version [30]. We notice, however, that solutions to (CPLE) do not enjoy the same regularity properties.

(iii) Let us have a closer look on the convergence results of Theorem 2. It is unclear how to extend the convergence result for radial derivatives to the non radial case. We shall expect the following. For an intial datum $u_0 \in \mathcal{X}_p$, and for $|x|/t$ large enough, the gradient of a solution to (CPLE) behaves as $|\nabla u(t,x)| \sim t^{\frac{1}{2-p}} |x|^{-\frac{2}{2-p}}$. Therefore we propose the following question: prove or disprove that, when $p_c < p < 2$, for an initial datum $u_0 \in \mathcal{X}_p$ with mass $M = \int_{\mathbb{R}^N} u_0 \,\mathrm{d}x$, we have that

$$\lim_{t\to\infty} t^{-\frac{1}{2-p}} \left\| (\nabla u(t) - \nabla \mathcal{B}_M(t)) \left(1 + |x|^{\frac{2}{2-p}}\right) \right\|_{\mathrm{L}^\infty(\mathbb{R}^N)} = 0\,.$$

Of course, the same question should be asked in $1 < p < p_c$ for solutions expected to converge to the pseudo-Barenblatt profile.

(iv) As a partial answer to (Q3) Theorem 6 provides sufficient conditions that has to be satisfied by solutions along time so that the entropy methods work. What is the full description of the basin of attraction of the Barenblatt solutions for p satisfying $1 < N < \frac{p}{(2-p)(p-1)}$? The most interesting information would be giving explicit convergence rates in the relative error under conditions imposed on the initial data only.

(v) Despite p_D used to be treated as an important threshold in the analysis of p-Laplace Cauchy problem (see Section 6.3), we have shown that is only a technical one restricting the use of the optimal transportation approach not the result itself. Are the special values we apply: p_c, p_M, and p_Y essential or technical thresholds?

Lastly, let us comment on two very natural directions that may arise after the present work: the doubly nonlinear equation and anisotropic p-Laplace evolution equation. By the doubly nonlinear diffusion equation we mean $\partial_t u = \Delta_p(u^m)$. The fast diffusion regime is when $p(m-1) < 1$. It is known, at least in the corresponding *good diffusion range*, that (non-negative and integrable) solutions to the Cauchy problem behave for large times as the corresponding Barenblatt profiles, see for instance [1, 2, 3]. Of course, the very natural question is how much of what has been proven in this work also applies to doubly non-linear case. We believe that the available regularity theory, see for instance [10, 11, 12, 57], allows to try to address questions (Q1), (Q2), and (Q3).

The second direction that we believe it is natural to explore are equations of the form $u_t = \sum_{i=1}^N \partial_i \left(|\partial_i u|^{p_i-2}\nabla u\right)$ for possibly different values of $p_i \in (0,1)$, $i = 1,\dots,N$. In these models, there are several difficulties, starting from the regularity theory to the existence and uniqueness of a fundamental solution. Nevertheless, these models seem to have attracted more and more attention, cf. [41, 42, 64]. In our analysis, the main difficulty would be understanding the right behaviour for large enough $|x|$. It is unclear whether a class as $\mathcal{X}_p$ can be found. At the same time, an interesting challenge is adapting the entropy method to those models as the fundamental solution, when it exists, is not explicit.

Acknowledgements

This work has been (partially) supported by the Project Conviviality ANR-23-CE40-0003 of the French National Research Agency. The authors are glad to acknowledge funding from Excellence Initiative – Research University (IDUB) at the University of Warsaw.

Appendix

We present here some general facts that do not rely strongly on our setting.

Let us present a Csiszár–Kullback-type inequality yielding that the relative entropy $\mathcal{E}[v|V_D]$ with respect to the Barenblatt profile of the same mass as v controls the L^1-distance to the Barenblatt profile. The proof we shall give is inspired from [15, Lemma 2.12], see also [23] for a previous contribution and more information about this inequality.

Lemma 9.1 (Csiszár–Kullback inequality). *Let $1 < p < 2$ and $v : \mathbb{R}^N \to [0,\infty)$ be a measurable function. Suppose that there exists $D > 0$ such that*

$$v - V_D \in \mathrm{L}^1(\mathbb{R}^N), \quad \int_{\mathbb{R}^N} (v - V_D)\,\mathrm{d}y = 0, \quad \textit{and} \quad \mathcal{E}[v|V_D] < \infty.$$

Then the following inequality holds true

$$\|v - V_D\|^2_{\mathrm{L}^1(\mathbb{R}^N)} \le 8\|V_D^{2-\gamma}\|_{\mathrm{L}^1(\mathbb{R}^N)}\,\mathcal{E}[v|V_D].$$

Proof. By the Mean Value Theorem, we know that for $0 \le t \le s$ it holds

$$t^\gamma - s^\gamma - \gamma\tfrac{1}{\gamma-1}s^{\gamma-1}(t-s) = \tfrac{\gamma(\gamma-1)}{2}\xi^{\gamma-2}(t-s)^2, \quad \text{with some } \ \xi \in [t,s]. \tag{147}$$

Since $\xi \le s$ and $\gamma - 2 \le 0$, we infer that

$$s - t \le \sqrt{2s^{2-\gamma}}\sqrt{\tfrac{1}{\gamma(\gamma-1)}t^\gamma - \tfrac{1}{\gamma(\gamma-1)}s^\gamma - \tfrac{1}{\gamma-1}s^{\gamma-1}(t-s)}. \tag{148}$$

From the assumption $\int_{\mathbb{R}^N}(v - V_D)\,\mathrm{d}y = 0$ we deduce that $\int_{\{v\le V_D\}}(V_D - v)\,\mathrm{d}y = \int_{\{V_D\le v\}}(v - V_D)\,\mathrm{d}y$, and hence

$$\frac{1}{2}\int_{\mathbb{R}^N}|v - V_D|\,\mathrm{d}y = \frac{1}{2}\left(\int_{\{v\le V_D\}}(V_D - v)\,\mathrm{d}y + \int_{\{V_D\le v\}}(v - V_D)\,\mathrm{d}y\right) = \int_{\{v\le V_D\}}(V_D - v)\,\mathrm{d}y.$$

Therefore, recalling the very definition of $\mathcal{E}$ and using inequality (148) with Cauchy-Schwarz's inequality we find

$$\frac{1}{4}\left(\int_{\mathbb{R}^N}|v(\tau) - V_D|\,\mathrm{d}y\right)^2 \le \|2V_D^{2-\gamma}\|_{\mathrm{L}^1(\mathbb{R}^N)}\,\mathcal{E}[v(\tau)|V_D].$$

□

Let us establish the decay of functions trapped between two Barenblatt profiles. Recall that V_D is defined in (3).

Lemma 9.2. *Let $p \in (1,2)$, $D_1 \geq D_2 > 0$ and let $v : \mathbb{R}^N \to [0,\infty)$ be a measurable function such that*

$$V_{D_1}(y) \leq v(y) \leq V_{D_2}(y) \qquad \text{for} \quad |y| \geq 1\,.$$

Then, for any $D \geq 0$, there exists a constant $C = C(D, D_1, D_2, p) > 0$ such that

$$|v(y) - V_D(y)| \leq C|y|^{-\frac{p}{(p-1)(2-p)}} \qquad \text{for} \quad |y| \geq 1\,.$$

Proof. Let us notice that

$$\frac{\partial}{\partial D} V_D(y) = \frac{\partial}{\partial D}\left[\left(D + \frac{2-p}{p}\,|y|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p-2}}\right] = \frac{-\frac{p-1}{2-p}\left(D + \frac{2-p}{p}\,|y|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p-2}-1}}{\left(D + \frac{2-p}{p}\,|y|^{\frac{p}{p-1}}\right)^{2\frac{p-1}{p-2}}} = -\frac{p-1}{2-p} V_D^{\frac{1}{p-1}}(y)\,.$$

Since $D_1 > D_2 > 0$ we can write, for any $y \in \mathbb{R}^N$

$$0 \leq V_{D_2}(y) - V_{D_1}(y) = \int_{D_1}^{D_2} \frac{\partial}{\partial D} V_D(y)\,\mathrm{d}D = -\frac{p-1}{2-p}\int_{D_1}^{D_2} V_D^{\frac{1}{p-1}}(y)\,\mathrm{d}D \leq \frac{p-1}{2-p}|D_2 - D_1|\,V_{D_1}^{\frac{1}{p-1}}(y)\,,$$

which is integrable for the prescribed range of p. The lower bound can be shown in the same way. Therefore

$$\tfrac{p-1}{2-p}|D_2 - D_1|\,V_{D_2}^{\frac{1}{p-1}}(y) \leq V_{D_2}(y) - V_{D_1}(y) \leq \tfrac{p-1}{2-p}|D_2 - D_1|\,V_{D_1}^{\frac{1}{p-1}}(y)\,.$$

In turn, for any D, we can estimate

$$\begin{aligned} |v(y) - V_D(y)| &\leq |V_{D_2}(y) - v(y)| + |V_D(y) - V_{D_2}(y)| \leq (V_{D_2}(y) - V_{D_1}(y)) + |V_D(y) - V_{D_2}(y)| \\ &\leq \tfrac{p-1}{2-p}(D_2 - D_1 + |D - D_2|) V_{D_0}^{\frac{1}{p-1}}(y)\,, \end{aligned} \tag{149}$$

where $D_0 = \min\{D, D_1, D_2\}$. By taking into account that for $|y| \geq 1$ it holds $V_{D_0}(y) \leq C|y|^{-\frac{p}{2-p}}$ we get the claim. □

Lemma 9.3 (Lemma 3.1, [21])**.** *Suppose $1 < p < 2$. Then there exist $c_1, c_2 > 0$ such that for all $\xi, \eta \in \mathbb{R}^N$ such that $\xi \neq 0$ we have*

$$\left\langle |\xi|^{p-2}\xi - |\eta|^{p-2}\eta, \xi - \eta \right\rangle \geq c_1 \frac{|\xi - \eta|^2}{|\xi|^{2-p} + |\eta|^{2-p}}\,, \tag{150}$$

where the optimal constant is achieved when $\langle \xi, \eta\rangle = |\xi|\,|\eta|$ and is given by $c_1 = \min\{1, 2(p-1)\}$, and

$$\left\langle |\xi|^{p-2}\xi - |\eta|^{p-2}\eta, \xi - \eta \right\rangle \geq c_2 \frac{|\xi - \eta|^2}{(|\xi| + |\eta|)^{2-p}}\,. \tag{151}$$

where $c_2 = c_1/2$.

We give here a modified version of the Gronwall-type lemma. We are sure it is known, but since we were not able to find a relevant reference, we present it with a proof.

Lemma 9.4. *Let $u : [0,\infty) \to [0,\infty)$ be bounded, decreasing and satisfying the inequality*

$$s\int_t^\infty u(\tau)\,\mathrm{d}\tau \leq u(t) \quad \forall\, t > 0\,,$$

where $s > 0$. Then, there exists $C = C(u_0, s) > 0$ such that $u(t) \leq C\,e^{-s\,t}$ for all $t > 0$. From the proof is clear that $C(u_0, s) = \frac{e^s}{s}\,u(0)$.

Proof. Let us define

$$v(t) := \int_t^\infty u(\tau)\,\mathrm{d}\tau\,.$$

By the properties of u, we infer that $v \in W^{1,\infty}(0,\infty)$. We may apply the classical version of the Gronwall lemma to v, since it satisfies the inequality

$$v'(t) = -u(t) \leq -s\int_t^\infty u(\tau)\,\mathrm{d}\tau = -s\,v(\tau)\,,$$

and thus $v(t) \leq v(0)e^{-s\,t}$. Since, by hypothesis, we have $s\,v(0) \leq u(0)$, we shall find $v(t) \leq v(0)e^{-s\,t} \leq \frac{u_0(0)}{s}\,e^{-s\,t}$. Notice that, by definition of $v(t)$, we have

$$v(t-1) - v(t) = \int_{t-1}^t u(\tau)\,\mathrm{d}\tau \geq 0\,.$$

At the same time, since u is nonincreasing, i.e., $u(t) \leq u(s)$ for any $s \in [t-1, t]$, we have that

$$u(t) = \int_{t-1}^t u(t)\,\mathrm{d}\tau \leq \int_{t-1}^t u(\tau)\,\mathrm{d}\tau = v(t-1) - v(t) \leq v(t-1)\,.$$

Combining all the above estimates we deduce that $u(t) \leq \frac{u(0)}{s}\,e^{-s(t-1)}$. □

9.1. Parameters

Thresholds for p. Their role is described in more details in Section 3.

symbol	introduced	info
$p_c = \frac{2N}{N+1} \in (\frac{3}{2}, 2)$	(1)	for $p > p_c$ solutions to (CPLE) conserve mass
$p_M \in (p_c, 2)$	(8)	for $p > p_M$ solutions to (CPLE) have finite weighted $\lvert x\rvert^{p'}$-moments
p_1, p_2	(24)	integrability threshold for $\lvert\mathcal{B}_{M_1} - \mathcal{B}_{M_2}\rvert$ defined if $N \geq 6$; $N < \frac{p}{(2-p)(p-1)} \iff p \in (1, p_1) \cup (p_2, 2)$
$p_Y = \frac{2N}{N+2}$	(22)	Yamabe exponent and gradient regularity threshold, cf. Section 4.2
$p_D = \frac{2N+1}{N+1} \in (p_M, 2)$	(88)	for $p > p_D$ the entropy functional is displacement convex, cf. Section 6.3

Main characters

symbol	introduced	info
u	(CPLE)	a solution to p-Laplace Cauchy problem with u_0 as initial datum; proven to converge to $\mathcal{B}_M$
v	(29)	a solution to Nonlinear Fokker–Planck problem with v_0 as initial datum
Φ	(17) or (114)	a solution to a radial FDE problem

Other symbols

symbol	introduced	info
V_D	(3)	stationary solution to the Fokker–Planck equation (29)
β	(2)	parameter for definition of Barenblatt profile; $\beta(p - p_c) \geq 0$
b_1, b_2	(40), (39)	parameter for definition of Barenblatt profile
ℓ	(19)	free parameter for definition of Barenblatt profile when $p = p_c$
$R_T(t)$	(4) or (19)	time rescaling for definition of Barenblatt profile $R_T(t) = \begin{cases} \{(T-t)_+/\lvert\beta\rvert\}^{\beta}, & 1 < p < p_c, \\ \exp\{\ell(T+t)\}, & p = p_c, \\ \{(T+t)/\lvert\beta\rvert\}^{\beta}, & p_c < p < 2. \end{cases}$
$\mathcal{B}_M(t+\beta, x) = R_\beta^{-N}(t) V_D(y)$	(5)	the Bareblatt solution to (CPLE) with D as in (5); for $p > p_c$
$\mathcal{B}_{D,T}(t+\beta, x) = R_T^{-N}(t) V_D(y)$	(5)	the Bareblatt solution to (CPLE); for $p \leq p_c$
$M_\star$	(84)	$M_\star = \lVert V_1 \rVert_{\mathrm{L}^1}$
$\gamma = \frac{2p-3}{p-1}$	(31)	parameter of the entropy
$\mathcal{E}$	(33)	relative entropy functional
$\mathcal{I}$	(34)	relative Fisher information
$m = p - 1 \in (0, 1)$	(103)	exponent for the radial FDE
$n > 0$	(103)	artificial dimension for the radial FDE
$\mathfrak{a} = N - 2 - 2\frac{N}{p'} \in (-2, 0]$	(103)	$\mathfrak{a} = N - n$
$a_1, a_2 > 0$	(106)	depend on $m, n, \mathfrak{a}$ or p, N
$\theta > 0$	(105)	FDE-Barenblatt parameter
$\mathfrak{B}_M$	(105)	the Bareblatt solution to the radial FDE (WFDE)

References

[1] M. Agueh. Asymptotic behavior for doubly degenerate parabolic equations. *C. R., Math., Acad. Sci. Paris*, 337(5):331–336, 2003.

[2] M. Agueh. Rates of decay to equilibria for p-Laplacian type equations. *Nonlinear Analysis: Theory, Methods and Applications*, 68(7):1909–1927, apr 2008.

[3] M. Agueh, A. Blanchet, and J. A. Carrillo. Large time asymptotics of the doubly nonlinear equation in the non-displacement convexity regime. *J. Evol. Equ.*, 10(1):59–84, 2010.

[4] N. Alikakos and R. Rostamian. Large time behavior of solutions of Neumann boundary value problem for the porous medium equation. *Indiana University Mathematics Journal*, 30(5):749, 1981.

[5] N. Alikakos and R. Rostamian. Gradient estimates for degenerate diffusion equations. *Mathematische Annalen*, 259(1):53–70, Mar. 1982.

[6] F. Andreu, J. M. Mazón, J. D. Rossi, and J. Toledo. Local and nonlocal weighted p-Laplacian evolution equations with Neumann boundary conditions. *Publicacions Matemàtiques*, 55:27–66, Jan. 2011.

[7] A. Arnold, J. A. Carrillo, L. Desvillettes, J. Dolbeault, A. Jüngel, C. Lederman, P. A. Markowich, G. Toscani, and C. Villani. Entropies and equilibria of many-particle systems: an essay on recent research. *Monatsh. Math.*, 142(1-2):35–43, 2004.

[8] P. Bénilan and M. G. Crandall. Regularizing effects of homogeneous evolution equations. Contributions to analysis and geometry, Suppl. Am. J. Math., 23-39, 1981.

[9] A. Blanchet, M. Bonforte, J. Dolbeault, G. Grillo, and J. L. Vázquez. Asymptotics of the fast diffusion equation via entropy estimates. *Arch. Ration. Mech. Anal.*, 191(2):347–385, 2009.

[10] V. Bögelein, F. Duzaar, U. Gianazza, N. Liao, and C. Scheven. Hölder continuity of the gradient of solutions to doubly non-linear parabolic equations, 05 2023.

[11] V. Bögelein, F. Duzaar, and N. Liao. On the Hölder regularity of signed solutions to a doubly nonlinear equation. *J. Funct. Anal.*, 281(9):Paper No. 109173, 58, 2021.

[12] V. Bögelein, F. Duzaar, and C. Scheven. Higher integrability for doubly nonlinear parabolic systems. *Partial Differ. Equ. Appl.*, 3(6):Paper No. 74, 41, 2022.

[13] M. Bonforte, J. Dolbeault, M. Muratori, and B. Nazaret. Weighted fast diffusion equations (Part I): Sharp asymptotic rates without symmetry and symmetry breaking in Caffarelli-Kohn-Nirenberg inequalities. *Kinetic and Related Models*, 10(1):33–59, 2017.

[14] M. Bonforte, J. Dolbeault, M. Muratori, and B. Nazaret. Weighted fast diffusion equations (Part II): Sharp asymptotic rates of convergence in relative error by entropy methods. *Kinetic and Related Models*, 10(1):61–91, 2017.

[15] M. Bonforte, J. Dolbeault, B. Nazaret, and N. Simonov. Stability in Gagliardo–Nirenberg–Sobolev inequalities, , 2020.

[16] M. Bonforte, J. Dolbeault, B. Nazaret, and N. Simonov. Constructive stability results in interpolation inequalities and explicit improvements of decay rates of fast diffusion equations. *Discrete and Continuous Dynamical Systems*, 43(3, 4):1070–1089, 2023.

[17] M. Bonforte, R. G. Iagar, and J. L. Vázquez. Local smoothing effects, positivity, and Harnack inequalities for the fast p-Laplacian equation. *Advances in Mathematics*, 224(5):2151–2215, 2010.

[18] M. Bonforte and N. Simonov. Quantitative a priori estimates for fast diffusion equations with Caffarelli–Kohn–Nirenberg weights. Harnack inequalities and Hölder continuity. *Advances in Mathematics*, 345:1075–1161, 2019.

[19] M. Bonforte and N. Simonov. Fine properties of solutions to the Cauchy problem for a fast diffusion equation with Caffarelli–Kohn–Nirenberg weights. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, 40(1):1–59, 2023.

[20] M. Bonforte, N. Simonov, and D. Stan. The Cauchy problem for the fast p-Laplacian evolution equation. Characterization of the global Harnack principle and fine asymptotic behaviour. *J. Math. Pures Appl. (9)*, 163:83–131, 2022.

[21] M. Borowski and I. Chlebicka. Controlling monotonicity of nonlinear operators. *Expo. Math.*, 40(4):1159–1180, 2022.

[22] J. A. Carrillo, M. G. Delgadino, J. Dolbeault, R. L. Frank, and F. Hoffmann. Reverse Hardy–Littlewood–Sobolev inequalities. *J. Math. Pures Appl. (9)*, 132:133–165, 2019.

[23] J. A. Carrillo, A. Jüngel, P. A. Markowich, G. Toscani, and A. Unterreiter. Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities. *Monatsh. Math.*, 133(1):1–82, 2001.

[24] I. Chlebicka and N. Simonov. Functional inequalities and applications to doubly nonlinear diffusion equations.

[25] P. Daskalopoulos and N. Sesum. On the extinction profile of solutions to fast diffusion. *J. Reine Angew. Math.*, 622:95–119, 2008.

[26] M. Del Pino and J. Dolbeault. Nonlinear diffusions and optimal constants in Sobolev type inequalities: asymptotic behaviour of equations involving the p-Laplacian. *C. R. Math. Acad. Sci. Paris*, 334(5):365–370, 2002.

[27] M. Del Pino and J. Dolbeault. Asymptotic behavior of nonlinear diffusions. *Math. Res. Lett.*, 10(4):551–557, 2003.

[28] M. del Pino and M. Sáez. On the extinction profile for solutions of $u_t = \Delta u^{(N-2)/(N+2)}$. *Indiana Univ. Math. J.*, 50(1):611–628, 2001.

[29] J. Denzler, H. Koch, and R. J. McCann. Higher-order time asymptotics of fast diffusion in Euclidean space: a dynamical systems approach. *Memoirs of the American Mathematical Society*, 234(1101):vi+81, 2015.

[30] J. Denzler, H. Koch, and R. J. McCann. Long-time asymptotic expansions for nonlinear diffusions in Euclidean space. In *Mathematical congress of the Americas. First mathematical congress of the Americas, Guanajuato, México, August 5–9, 2013*, pages 85–94. Providence, RI: American Mathematical Society (AMS), 2016.

[31] E. DiBenedetto and A. Friedman. Regularity of solutions of nonlinear degenerate parabolic systems. *J. Reine Angew. Math.*, 349:83–128, 1984.

[32] E. DiBenedetto and A. Friedman. Addendum to "Hölder estimates for nonlinear degenerate parabolic systems". *J. Reine Angew. Math.*, 363:217–220, 1985.

[33] E. DiBenedetto and A. Friedman. Hölder estimates for non-linear degenerate parabolic systems. *J. Reine Angew. Math.*, 357:1–22, 1985.

[34] E. DiBenedetto, U. Gianazza, and V. Vespri. Harnack estimates for quasi-linear degenerate parabolic differential equations. *Acta Math.*, 200(2):181–209, 2008.

[35] E. DiBenedetto, U. Gianazza, and V. Vespri. Harnack estimates for quasi-linear degenerate parabolic differential equations. *Acta Math.*, 200(2):181–209, 2008.

[36] E. DiBenedetto, U. Gianazza, and V. Vespri. *Harnack's inequality for degenerate and singular parabolic equations*. Springer Monographs in Mathematics. Springer, New York, 2012.

[37] E. DiBenedetto and M. A. Herrero. Nonnegative solutions of the evolution p-Laplacian equation. Initial traces and Cauchy problem when $1 < p < 2$. *Arch. Rational Mech. Anal.*, 111(3):225–290, 1990.

[38] J. R. Esteban and J. L. Vazquez. Homogeneous diffusion in $\mathbb{R}$ with power-like nonlinear diffusivity. *Archive for Rational Mechanics and Analysis*, 103(1):39–80, 1988.

[39] Y. Feng, M. Parviainen, and S. Sarsa. On the second-order regularity of solutions to the parabolic p-Laplace equation. *J. Evol. Equ.*, 22(1):Paper No. 6, 17, 2022.

[40] Y. Feng, M. Parviainen, and S. Sarsa. A systematic approach on the second order regularity of solutions to the general parabolic p-Laplace equation. *Calc. Var. Partial Differential Equations*, 62(7):Paper No. 204, 39, 2023.

[41] F. Feo, J. L. Vázquez, and B. Volzone. Anisotropic p-Laplacian evolution of fast diffusion type. *Advanced Nonlinear Studies*, 21(3):523–555, July 2021.

[42] F. Feo, J. L. Vázquez, and B. Volzone. Anisotropic fast diffusion equations. *Nonlinear Analysis*, 233:113298, Aug. 2023.

[43] R. K. Giri, J. Kinnunen, and K. Moring. Supercaloric functions for the parabolic p-Laplace equation in the fast diffusion case. *NoDEA Nonlinear Differential Equations Appl.*, 28(3):Paper No. 33, 21, 2021.

[44] M. A. Herrero and M. Pierre. The Cauchy Problem for $u_t = \Delta u^m$ when $0 < m < 1$. *Transactions of the American Mathematical Society*, 291(1):145, sep 1985.

[45] R. G. Iagar, A. Sánchez, and J. L. Vázquez. Radial equivalence for the two basic nonlinear degenerate diffusion equations. *Journal de Mathématiques Pures et Appliquées*, 89(1):1–24, jan 2008.

[46] C. Imbert, T. Jin, and L. Silvestre. Hölder gradient estimates for a class of singular or degenerate parabolic equations. *Advances in Nonlinear Analysis*, 8(1):845–867, Sept. 2017.

[47] P. Juutinen, P. Lindqvist, and J. J. Manfredi. On the equivalence of viscosity solutions and weak solutions for a quasi-linear equation. *SIAM Journal on Mathematical Analysis*, 33(3):699–717, Jan. 2001.

[48] S. Kamin and J. L. Vázquez. Fundamental solutions and asymptotic behaviour for the p-Laplacian equation. *Rev. Mat. Iberoamericana*, 4(2):339–354, 1988.

[49] Y. J. Kim and R. J. McCann. Potential theory and optimal convergence rates in fast nonlinear diffusion. *J. Math. Pures Appl. (9)*, 86(1):42–67, 2006.

[50] T. Kuusi and G. Mingione. Nonlinear potential estimates in parabolic problems. *Rendiconti Lincei - Matematica e Applicazioni*, pages 161–174, 2011.

[51] T. Kuusi and G. Mingione. New perturbation methods for nonlinear parabolic problems. *Journal de Mathématiques Pures et Appliquées*, 98(4):390–427, Oct. 2012.

[52] T. Kuusi and G. Mingione. Gradient regularity for nonlinear parabolic equations. *Annali Scuola Normale Superiore - Classe di Scienze*, pages 755–822, Dec. 2013.

[53] T. Kuusi and G. Mingione. Riesz potentials and nonlinear parabolic equations. *Archive for Rational Mechanics and Analysis*, 212(3):727–780, Dec. 2013.

[54] T. Kuusi and G. Mingione. Guide to nonlinear potential estimates. *Bull. Math. Sci.*, 4(1):1–82, 2014.

[55] T. Kuusi, G. Mingione, and K. Nyström. A boundary Harnack inequality for singular equations of p-parabolic type. *Proc. Amer. Math. Soc.*, 142(8):2705–2719, 2014.

[56] O. A. Ladyzhenskaya. *The mathematical theory of viscous incompressible flow*. Mathematics and its Applications, Vol. 2. Gordon and Breach Science Publishers, New York-London-Paris, 1969. Second English edition, revised and enlarged, Translated from the Russian by Richard A. Silverman and John Chu.

[57] N. Liao and L. Schätzler. On the Hölder regularity of signed solutions to a doubly nonlinear equation. Part III. *Int. Math. Res. Not. IMRN*, 3:2376–2400, 2022.

[58] J. L. Lions. Quelques méthodes de résolution des problèmes aux limites non linéaires. Études mathématiques. Paris: Dunod; Paris: Gauthier-Villars. xx, 554 p. (1969)., 1969.

[59] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

[60] J. Siltakoski. Equivalence of viscosity and weak solutions for a p-parabolic equation. *Journal of Evolution Equations*, 21(2):2047–2080, 2021.

[61] J. L. 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.

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

[63] J. L. Vázquez. A survey on mass conservation and related topics in nonlinear diffusion, 11 2023.

[64] J. L. Vázquez. The very singular solution for the anisotropic fast diffusion equation and its consequences, 2023.

[65] M. Wiegner. On C_α-regularity of the gradient of solutions of degenerate parabolic systems. *Ann. Mat. Pura Appl. (4)*, 145:385–405, 1986.

[66] C. Ya-Zhe and E. Di Benedetto. On the local behavior of solutions of singular parabolic equations. *Archive for Rational Mechanics and Analysis*, 103(4):319–345, Dec. 1988.

[67] C. Ya-Zhe and E. DiBenedetto. Hölder estimates of solutions of singular parabolic equations with measurable coefficients. *Archive for Rational Mechanics and Analysis*, 118(3):257–271, 1992.

[68] C. Yazhe. Hölder continuity of the gradient of solutions of nonlinear degenerate parabolic systems. *Acta Mathematica Sinica*, 2(4):309–331, Dec. 1986.

[69] R. Ye. Global existence and convergence of Yamabe flow. *J. Differ. Geom.*, 39(1):35–50, 1994.