Université de Nîmes

Equipe MIPA
Université de Nîmes, Site des Carmes
Place Gabriel Péri, 30021 Nîmes, France
http://mipa.unimes.fr

Global estimates
for nonlinear parabolic equations

by

Paolo Baroni, Agnese Di Castro
and Giampiero Palatucci

January 2013

Global estimates for nonlinear parabolic equations

**Paolo Baroni · Agnese Di Castro ·
Giampiero Palatucci**

Abstract We consider nonlinear parabolic equations of the type

$$u_t - \operatorname{div} a(x, t, Du) = f(x, t) \ \text{ on } \ \Omega_T = \Omega \times (-T, 0),$$

under standard growth conditions on a, with f only assumed to be integrable. We prove general decay estimates up to the boundary for level sets of the solutions u and the gradient Du which imply very general estimates in Lebesgue and Lorentz spaces. Assuming only that the involved domains satisfy a mild exterior capacity density condition, we provide global regularity results.

Keywords Nonlinear parabolic problems · Calderón-Zygmund theory · Lorentz regularity · Rearrangement invariant function spaces · Higher integrability

Mathematics Subject Classification (2000) 35K55 · 35B65 · 35K10 · 46E30

P. Baroni
Scuola Normale Superiore
Piazza dei Cavalieri, 7
56126 Pisa, Italy
E-mail: paolo.baroni@sns.it

A. Di Castro · G. Palatucci
Dipartimento di Matematica e Informatica
Università degli Studi di Parma
Parco Area delle Scienze, 53/A
43124 Parma, Italy
E-mail: agnese.dicastro@unipr.it · E-mail: giampiero.palatucci@unimes.fr

1 Introduction

The aim of the present paper is to establish *global* regularity results for the
solutions of the following class of Cauchy-Dirichlet problem

$$
\begin{cases}
u_t - \operatorname{div} a(x,t,Du) = f(x,t) & \text{in } \Omega_T = \Omega \times (-T,0), \\
u = 0 & \text{on } \partial_{\mathrm{par}} \Omega_T,
\end{cases}
\tag{1}
$$

where Ω is a bounded open set in $\mathbb{R}^n$, $n \geq 2$, $T > 0$, $\partial_{\mathrm{par}} \Omega_T$ is the usual
parabolic boundary of Ω_T. Here we assume that f, in the most general case, is
a summable function in Ω_T, and $a : \Omega_T \times \mathbb{R}^n \to \mathbb{R}^n$ is a Carathéodory regular
vector field satisfying standard monotonicity and growth conditions; i.e.,

$$
\begin{cases}
\nu |z_2 - z_1|^2 \leq \langle a(x,t,z_2) - a(x,t,z_1), z_2 - z_1 \rangle \\
|a(x,t,z)| \leq L|z|
\end{cases}
\tag{2}
$$

for every z_1, z_2, $z \in \mathbb{R}^n$ and $(x,t) \in \Omega_T$; the structure constants satisfy
$0 < \nu \leq 1 \leq L$.

   We will focus here on the case when f belongs to the Lebesgue space
$L^\gamma(\Omega_T)$ in a range of γ that does not necessarily permit to obtain the existence
of finite energy solutions $u \in L^2(-T,0; W_0^{1,2}(\Omega))$ to problem (1).

Remark 1 In this paper we are not going to consider any solution to the
problem (1), but rather we shall deal with the (very) weak solutions $u \in
L^1(-T,0; W_0^{1,1}(\Omega))$ obtained via the Boccardo-Gallouët by now standard ap-
proximation procedure ([9]; see forthcoming Section 2.1). These solutions are
called SOLA (Solutions Obtained by Limits of Approximations) and turn out
to be unique in their class in the case $f \in L^1(\Omega_T)$, see [12, Theorem 4.1]. For
such reasons, with some abuse of terminology, we shall call in the following
the unique SOLA to (1) "the solution" to problem (1).

   In the classical Lebesgue framework, Boccardo, Dall'Aglio, Gallouët and
Orsina (see [8, Theorem 1.9]) show the existence of a unique solution

$$
u \in L^q(-T,0; W_0^{1,q}(\Omega)) \text{ with } q = \frac{N\gamma}{N - \gamma}
\tag{3}
$$

to problem (1) under the assumption (2), provided that the datum f satisfies

$$
f \in L^\gamma(\Omega_T) \text{ for some } 1 < \gamma < \frac{2N}{N+2}.
$$

Here $N := n+2$ denotes the homogeneous parabolic dimension. Moreover, the
solution u belongs to $L^\sigma(\Omega_T)$ with $\sigma = N\gamma/(N-2\gamma)$. This result is optimal in
the scale of Lebesgue spaces; however it is natural to ask for a more accurate
scale to describe the regularity of Du in dependence on the datum f. In this
respect, recently, the results above were extended in all the most familiar
spaces of rearrangement (Lebesgue, Lorentz, Orlicz) and non-rearrangement
once, up to Lorentz-Morrey[1], by Habermann and the first author of this paper

---

[1] We refer to Section 2.2 for the definitions of the involved function spaces in the parabolic
setting.

in [5] (see, also, [25, 14, 26]). We would like to remark that a fine analysis in Lorentz spaces regularity is often decisive to establish higher regularity, as it has been shown, for instance, by recent developments in the theory of harmonic and biharmonic mappings (see, e.g., [41, 22, 35] and the references therein). Precisely, in [5, Theorem 6.6], starting from the datum f in the Lorentz space $L(\gamma, q)(\Omega_T)$, for $1 < \gamma \leq 2N/(N+2)$, the authors prove that for any $0 < q \leq \infty$

$$|Du| \in L\left(\frac{N\gamma}{N-\gamma}, q\right) \text{ locally in } \Omega_T,$$

extending to the parabolic setting a very general approach introduced by Mingione in [31, 32, 33] for the regularity of solutions to the elliptic counterpart of problem (1).

After [31], many important extensions and generalizations have been considered in the literature, in which a similar approach based on maximal operators techniques has been pursued to achieve regularity results in nonlinear degenerate problems: see, e. g., [33, 13, 15] for elliptic problems, [1, 5, 14, 17, 26, 27] for parabolic ones and the references therein. It is important to stress that the aforementioned papers, despite considering more and more general settings, provide *local* results not considering the regularity properties *up to the boundary* of the solutions.

Recently, some attention has been devoted to such global extension, as in the paper by Phuc [40], where Lorentz regularity in the whole domain Ω has been analyzed for the solutions to the elliptic counterpart of problem (1) for general structures of p-Laplacian type, again taking into account the original potential approach in [32][2].

In this paper, we will present a series of up-to-the-boundary regularity results aimed at giving, on one hand, an extension of the classical Lebesgue theory developed in [8] to a very general class of rearrangement invariant function spaces. For instance we will cover the case of Lorentz spaces, that do not seem to be treatable with the currently available methods. On the other hand, we will give a set of techniques allowing for treating domains with very poorly regular boundaries. In this respect we also extend to the parabolic case the recent results obtained in [40].

Indeed, since we are dealing with precise estimates up to the boundary, it is necessary to take some regularity condition for $\Omega_T = \Omega \times (-T, 0)$. For this, we assume that the domain Ω has only to satisfy a capacity density condition. Namely, in order to achieve the main results of the present paper, the complement of the set Ω is required to be a *thick domain* (see Section 2.4 below); and such assumption is both very natural (as, for instance, in the classic general Sobolev embedding in capacity density settings; and, also, in the elliptic framework studied in [40]), and very mild (it is fulfilled by any domain with Lipschitz boundary, or even of corkscrew type; also, in $\mathbb{R}^2$ by

---

[2] Although – as stated in [40], too – the possibility of extending such local results to global ones was already mentioned without proof in [32] and [34].

any proper simply connected domain). As a further matter, this condition is essentially sharp for higher integrability results, as pointed out by Kilpeläinen and Koskela in Remark 3.3 in the paper [23], where the authors developed a first analysis in global Lebesgue regularity of the gradient of solutions to degenerate elliptic equations in general thick domains.

All in all, we extend the potential theoretic strategy recently developed by Mingione in [32,33] to the parabolic framework up to the boundary, combining it with global higher integrability results (in Theorem 7 in Section 3), to obtain general gradient estimates on level sets (see Section 4), which in turn imply the following

Theorem 1 *Let Ω be such that $\mathbb{R}^n \setminus \Omega$ is a thick domain with constants c_0 and ρ_0 and assume (2). Let $f \in L(\gamma, q)(\Omega_T)$ with*

$$1 < \gamma \leq \frac{2N}{N+2} \qquad and \qquad 0 < q \leq \infty.$$

Then the solution $u \in L^1(-T, 0; W_0^{1,1}(\Omega))$ to (1) satisfies

$$|Du| \in L\left(\frac{N\gamma}{N-\gamma}, q\right) \ in \ \Omega_T.$$

Moreover, the following estimate holds

$$\|Du\|_{L\left(\frac{N\gamma}{N-\gamma}, q\right)(\Omega_T)} \leq c \, \|f\|_{L(\gamma, q)(\Omega_T)}, \tag{4}$$

where c is a constant depending only on $n, q, L, \nu, \gamma, c_0, \rho_0, \operatorname{diam}(\Omega)$.

Of course, by choosing $q = \infty$, we can also deduce regularity results on the Marcinkiewicz scale.

Remark 2 After stating Theorem 1, we want to make a more precise comparison between our results and the basic results in [8], where the authors treat the case of Lebesgue spaces; i. e., they take $q = \gamma$ in the statement above thereby obtaining the corresponding conclusion. The results in Theorem 1 is not a direct extension of the ones in [8], as in this paper no regularity assumptions are considered on Ω. On the other hand, the results in the present paper are global but they are obtained by covering arguments via estimates that are *local in nature* as it is easy to see that, using the methods developed here, one can obtain local estimates up to the boundary and not only in the interior. This corresponds to replace, in Theorem 1, Ω_T by cylinders of the type $(B(x_0, R) \cap \Omega) \times (-T, 0)$, where $x_0 \in \partial\Omega$. This is a valued consequence of the techniques approached here and the necessity of assuming the considered boundary capacity conditions stems exactly from this aspect of the involved techniques. Now, whereas it is likely that some Lorentz space estimates can be also obtained under milder assumption on Ω, it is worth mentioning that the methods here can be also used to get global estimates in Morrey and Lorentz-Morrey spaces (as those obtained in the interior case in [14,32]).

The proof of Theorem 1 relies on the fact that the integrability of the spatial gradient of the solutions to problem (1) is linked to a suitable choice of potential operators. In this sense, one of the key-points will consist into obtaining a decay estimate that involves the level sets of the parabolic Hardy-Littlewood maximal operator M_0^* of $|Du|$ in term of those of another maximal operator M_1^* of the datum f, up to a correction term which is negligible when considering the higher regularity. Indeed, we will prove an estimate of the following type

$$\left|\{M_0^*(|Du|) \geq S\lambda\}\right| \lesssim \frac{1}{S^{2\chi_1}}\left|\{M_0^*(|Du|) \geq \lambda\}\right| + \left|\{[M_1^*(|f|)] \geq \lambda\}\right|, \quad (5)$$

for every λ suitably large, and in which $S >> 1$ is a constant to be chosen, and the exponent $\chi_1 > 1$ is related to the higher integrability theory.

The estimate in (5) (whose precise version is given by forthcoming formula (50)) is fairly general and it will be relevant to deducing the Lorentz estimates stated in our theorems, also including the borderline case. In order to obtain such level sets estimate, we apply a parabolic covering lemma of Calderón-Zygmund type (see Lemma 6 below) together with the higher integrability theory which we will extend up to the boundary (see Theorem 7 below). Therefore, we will work on basic estimates of the solutions u to (1) in comparison to the solutions w to the corresponding homogeneous problem (see Section 3), by exploiting very recent contributions in the parabolic framework given in [27]. Clearly, the situation is complicated by the necessity to work in a global setting.

Analogously, we will show how to establish a modified version of the level sets estimate in (5) by replacing Du by u there. This will permit us to recover higher regularity estimates in the Lorentz scale for the solutions u themselves, as stated in the following

Theorem 2 *Let Ω be such that $\mathbb{R}^n \setminus \Omega$ is a thick domain with constants c_0 and ρ_0 and let $f \in L(\gamma, q)(\Omega_T)$ with*

$$1 < \gamma < \frac{N}{2} \qquad and \qquad 0 < q \leq \infty.$$

Then the solution $u \in L^1(-T, 0; W_0^{1,1}(\Omega))$ to (1) satisfies

$$u \in L\left(\frac{N\gamma}{N - 2\gamma}, q\right) \ in \ \Omega_T.$$

Moreover, the following estimate holds

$$\|u\|_{L\left(\frac{N\gamma}{N-2\gamma}, q\right)(\Omega_T)} \leq c\|f\|_{L(\gamma, q)(\Omega_T)},$$

where c is a constant depending only on $n, q, L, \nu, \gamma, c_0, \rho_0, \mathrm{diam}(\Omega)$.

It is worth pointing out that, although in the elliptic case the regularity of u can be recovered by plainly combining the regularity of the gradient Du with the classic Sobolev embeddings, here we need to work in a separate way, by means of sharp estimates also involving different maximal operators in (5), and higher regularity results up to the boundary for the solutions u.

Finally, we recall that all the results we obtained in the present paper hold for the very weak solutions given by the approximation method described in forthcoming Section 2.1. It would be interesting to understand whether these results can be extended to some other notion of solutions. In this respect, a positive answer can be given when dealing with notions of solutions to measure data problems holding uniqueness in case of integrable data, as in case of the renormalized solutions in [38] and [39]; see [7] for a first definition of renormalized solutions in this framework, and also [16].

Remark 3 In the present paper we are dealing with the case of operators with linear growth while in other papers, as for instance [32], level sets techniques are used to approach similar regularity results for p-Laplacian type operators, that is for equations of the type $-\operatorname{div}\left(|Du|^{p-2}Du\right) = f$. When considering the parabolic analog

$$u_t - \operatorname{div}\left(|Du|^{p-2}Du\right) = f,$$

the approach of the present paper is not sufficient if used alone, and has to be modified and extended using the concept of intrinsic geometry first pioneered by DiBenedetto to overcome the lack of homogeneity of the evolutionary p-Laplacian operator when $p \neq 2$ (see for instance [1] for a discussion of this concept in the relevant context). This is the object of forthcoming investigations ([6]).

The paper is organized as follows.

In Section 2, we fix notation; we give some details on the structure of the problem and we briefly recall the definitions and a few basic properties of the spaces and the operators we deal with, also discussing the capacity density condition and some properties of Sobolev functions in p-thick domains; we recall a parabolic covering lemma of Calderón-Zygmund type.

In Section 3, we state and prove global higher regularity results for the solutions to the homogeneous analog of problem (1), and other preliminary results.

In Section 4, we prove the results stated in Theorem 1, and the Lorentz space estimates for the solutions u.

2 Preliminaries

Throughout the paper we follow the usual convention of denoting by c a general positive constant *always greater or equal than one* that may vary from line to line. Relevant dependencies will be emphasized by using parentheses; special constants will be denoted by c_0, c_1, c_2,....

As usual, we denote by

$$B_R(x_0) = B(x_0; R) := \{x \in \mathbb{R}^n : |x - x_0| < R\}$$

the open ball centered in $x_0 \in \mathbb{R}^n$ with radius $R > 0$. When not important and clear from the context, we shall use the shorter notation $B_R = B_R(x_0)$. Let $(x_0, t_0) \in \mathbb{R}^{n+1}$, we denote by

$$\mathcal{Q}_R(x_0, t_0) = \mathcal{Q}(x_0, t_0; R) := B(x_0; R) \times (t_0 - R^2, t_0) \qquad (6)$$

the open parabolic cylinder with height R^2 and having a ball B_R as horizontal slice.

Finally, we mean $\partial_{\text{par}} \Omega_T = \partial_{\text{lat}} \Omega_T \cup \Omega_0$, where $\partial_{\text{lat}} \Omega_T := \partial\Omega \times (-T, 0)$ is the lateral boundary and $\Omega_0 = \Omega \times \{-T\}$ is the initial boundary.

2.1 Solvability of the problem

A measurable function u is a very weak solution to (1) if $u \in L^1(-T, 0; W_0^{1,1}(\Omega))$ and

$$-\int_{\Omega_T} u\phi_t \,\mathrm{d}x \,\mathrm{d}t + \int_{\Omega_T} \langle a(x, t, Du), D\phi \rangle \,\mathrm{d}x \,\mathrm{d}t = \int_{\Omega_T} f\phi \,\mathrm{d}x \,\mathrm{d}t, \qquad (7)$$

for any $\phi \in C_0^\infty(\Omega_T)$. Also, while the lateral boundary condition can be formulated by prescribing the belonging of u to $L^1(-T, 0; W_0^{1,1}(\Omega))$, the initial boundary condition is understood in the L^1-sense, that is

$$\frac{1}{h} \int_{-T}^{-T+h} \int_\Omega |u(x,t)| \,\mathrm{d}x \,\mathrm{d}t \to 0 \quad \text{as } h \to 0. \qquad (8)$$

As usual in the parabolic setting, one can provide a convenient "slice-wise" reformulation of equality (7) by mean of Steklov average. Indeed, for $h > 0$ and $t \in [-T, 0)$, we can define

$$u_h(x,t) := \begin{cases} \dfrac{1}{h} \displaystyle\int_t^{t+h} u(x, \tilde{t}) \,\mathrm{d}\tilde{t} & \text{if } t + h \le 0 \\ 0 & \text{if } t + h > 0 \end{cases}$$

and the following equality

$$\int_\Omega \left[(u_h)_t \phi + \langle [a(x, t, Du)]_h, D\phi \rangle \right] \,\mathrm{d}x = \int_\Omega f_h \phi \,\mathrm{d}x \qquad (9)$$

holds for any $\phi \in W_0^{1,2}(\Omega)$ and for a. e. $t \in (-T, 0)$. This formulation allows to use the solution u, which enjoys only very weak regularity properties in time, as a test function in (9), and therefore *throughout all the paper* we shall proceed formally in this sense, referring for instance to [4] for the details in similar situations.

The existence of such solution is obtained using an approximation method; this means that we shall consider a special class of solutions called SOLA (Solutions Obtained by Limits of Approximations). In particular, one considers a sequence of bounded function $\{f_k\} \subset L^\infty(\Omega_T)$ such that $f_k \to f$ in $L^1(\Omega_T)$ as $k \to \infty$. Then, by standard monotonicity arguments, for each k fixed, there exists a unique solution

$$u_k \in C^0([-T,0]; L^2(\Omega)) \cap L^2(-T,0; W_0^{1,2}(\Omega))$$

and

$$(u_k)_t \in C^0(-T,0; W^{-1,2}(\Omega))$$

to the Cauchy-Dirichlet problem

$$\begin{cases} (u_k)_t - \operatorname{div} a(x,t,Du_k) = f_k & \text{in } \Omega_T \\ u_k = 0 & \text{on } \partial_{\mathrm{par}}\Omega_T. \end{cases} \tag{10}$$

Finally, thanks to the arguments in [9,8], we can pass to the limit in the problem above and to prove the existence of a function $u \in L^1(-T,0; W_0^{1,1}(\Omega))$ which solves (1) in the distributional sense. For the rest of this paper we understand by $\{u_k\}$ the sequence obtained by solving (10) with

$$f_k(x,t) := \max\big\{-k,\ \min\{f(x,t),k\}\big\}, \quad k \in \mathbb{N}. \tag{11}$$

Notice that, as stated in Remark 1, this choice is not restrictive, since in this class of solutions uniqueness holds, in the sense that considering another approximating sequence $\{\bar{f}_k\}$ would lead to the same limit u.

2.2 Relevant parabolic function spaces

Let Ω_T be the space time cylinder $\Omega \times (-T,0)$, with Ω being a bounded open set in $\mathbb{R}^n$, $n \geq 2$ and $T > 0$.

Fix $q \in (0,\infty)$. A measurable map $g : \Omega_T \to \mathbb{R}^k$ belongs to the *Lorentz space* $L(\gamma,q)(\Omega_T)$ with $1 \leq \gamma < \infty$ if and only if

$$\|g\|_{L(\gamma,q)(\Omega_T)}^q := q \int_0^\infty \left(\lambda^\gamma \big|\{(x,t) \in \Omega_T : |g(x,t)| > \lambda\}\big|\right)^{\frac{q}{\gamma}} \frac{d\lambda}{\lambda} < +\infty. \tag{12}$$

In the case $q = \infty$, the Lorentz space $L(\gamma,\infty)$ with $\gamma \in [1,\infty)$ is the so-called *Marcinkiewicz space* and it is usually denoted by $\mathcal{M}^\gamma(\Omega_T)$. A measurable map $g : \Omega_T \to \mathbb{R}^k$ belongs to $\mathcal{M}^\gamma(\Omega_T)$ if and only if

$$\|g\|_{\mathcal{M}^\gamma(\Omega_T)}^\gamma \equiv \|g\|_{L(\gamma,\infty)(\Omega_T)}^\gamma := \sup_{\lambda>0} \lambda^\gamma \big|\{(x,t) \in \Omega_T : |g(x,t)| > \lambda\}\big| < +\infty. \tag{13}$$

Note that by Fubini's Theorem one can see that, for any $\gamma \in [1,\infty)$, $L(\gamma,\gamma)(\Omega_T)$ is nothing than the classic Lebesgue space $L^\gamma(\Omega_T)$. Also, by Fatou's Lemma, one can see that the functionals $\|\cdot\|_{L(\gamma,q)(\Omega_T)}$ defined in (12)-(13)

are lower semi-continuous with respect to the a.e. convergence; see [5, Section 3.2].

Moreover, the spaces defined above enjoy Hölder type inequalities. In particular, below we state a standard inequality for the Marcinkiewicz spaces $\mathcal{M}^\gamma(\Omega)$ in the form we will need it in the following of the paper.

Lemma 1 *Let $E \subseteq \mathbb{R}^{n+1}$ be a measurable set and let $g \in \mathcal{M}^\gamma(E)$ with $\gamma > 1$. Then, for any $q \in [1, \gamma)$, $g \in L^q(E)$ and*

$$\|g\|_{L^q(E)} \leq \left(\frac{\gamma}{\gamma - q} \right)^{\frac{1}{q}} |E|^{\frac{1}{q} - \frac{1}{\gamma}} \|g\|_{\mathcal{M}^\gamma(E)}.$$

For details and results about the theory of Lorentz spaces, we refer the interested reader to [2, 20].

2.3 Parabolic maximal operators

For any measurable function g, the restricted *fractional maximal operator* $M^*_{\alpha, \mathcal{Q}_0}$, with $\alpha \in [0, N]$, relative to a parabolic cylinder $\mathcal{Q}_0$, is defined by

$$M^*_{\alpha, \mathcal{Q}_0}(g)(x, t) := \sup_{\mathcal{Q} \subseteq \mathcal{Q}_0, \ (x,t) \in \mathcal{Q}} |\mathcal{Q}|^{\frac{\alpha}{N}} \fint_{\mathcal{Q}} |g(y, \tau)| \, \mathrm{d}y \, \mathrm{d}\tau,$$

where the cylinders $\mathcal{Q}$ have sides parallel to those of $\mathcal{Q}_0$.

The boundedness of maximal operators in Marcinkiewicz spaces is classical (see, for instance, [20]); i.e., for any $g \in L^\gamma(\mathcal{Q}_0)$,

$$\left| \left\{ (x, t) \in \mathcal{Q}_0 : M^*_{0, \mathcal{Q}_0}(g)(x, t) \geq \lambda \right\} \right| \leq \frac{\bar{c}}{\lambda^\gamma} \int_{\mathcal{Q}_0} |g|^\gamma \, \mathrm{d}x \, \mathrm{d}t \qquad (14)$$

holds for every $\lambda > 0$ and $\gamma \geq 1$; the constant c depending only on n and γ.

More in general, a standard embedding result for the maximal functions in Lorentz spaces does hold, as given by the following theorem, whose proof follows by the Marcinkiewicz theorem together with standard sublinear interpolation.

Theorem 3 ([32, Theorem 7]). *Let α in $[0, N]$, and $\gamma > 1$ be such that $\alpha\gamma < N$. Let $\mathcal{Q} \subset \mathbb{R}^{n+1}$ be a parabolic cylinder and $q \in (0, \infty]$. Then for every measurable function g in $L(\gamma, q)(\mathcal{Q})$ it holds*

$$\|M^*_{\alpha, \mathcal{Q}}(g)\|_{L\left(\frac{N\gamma}{N - \alpha\gamma}, q \right)(\mathcal{Q})} \leq c\|g\|_{L(\gamma, q)(\mathcal{Q})},$$

where c is a constant depending only on α, γ, n and q.

Also, in the borderline case $\gamma = 1$, we have the following

Theorem 4 ([5, Lemma 4.11 and Theorem 4.12]). *Let $\alpha \in [0, N]$, $\mathcal{Q} \subset \mathbb{R}^{n+1}$ be a parabolic cylinder and consider the concentric parabolic cylinder $\sigma\mathcal{Q}$ scaled by a factor $\sigma > 1$. Then there exists a constant $c = c(n, \alpha, \sigma)$ such that, for any measurable function g in $\sigma\mathcal{Q}$, it holds*

$$\|M^*_{\alpha, \mathcal{Q}}(g)\|_{L^{\frac{N}{N - \alpha}}(\mathcal{Q})} \leq c|\mathcal{Q}|^{1 - \frac{\alpha}{N}} \|g\|^{\frac{\alpha}{N}}_{L^1(\sigma\mathcal{Q})} \|g\|^{1 - \frac{\alpha}{N}}_{L \log L(\mathcal{Q})}.$$

2.4 Capacity density condition and thick domains

In the following we recall the definition and some properties of the p-thick domains. We need first to recall the definition of p-capacity of sets.

Definition 1 Let $K \subseteq \Omega$ be a compact set. For any $p \geq 1$ the *variational p-capacity of K on Ω* is defined by

$$\mathrm{cap}_p(K, \Omega) := \inf \left\{ \int_\Omega |D\varphi|^p \, \mathrm{d}x : \varphi \in C_0^\infty(\Omega), \varphi \geq \mathbb{1}_K \right\}.$$

As customary, the definition above can be extended to generic open sets $A \subseteq \Omega$ by taking the supremum over the capacities of the compact sets contained in A. In the case of balls, the definition simplifies considerably and we have the following explicit formula for B_ρ on $B_{2\rho}$

$$\mathrm{cap}_p(\overline{B_\rho}, B_{2\rho}) = c\,\rho^{n-p},$$

where c is a positive constant depending only on n and p. For further details about the variational p-capacity, we refer for instance to Chapter 2 in [21].

Definition 2 A set $E \subset \mathbb{R}^n$ is *uniformly p-thick* (equivalently, satisfies the *p-capacity uniform thickness* condition) if there exist constants c_0 and $\rho_0 > 0$ such that

$$\mathrm{cap}_p(E \cap \overline{B_\rho(x)}, B_{2\rho}(x)) \geq c_0 \, \mathrm{cap}_p(\overline{B_\rho(x)}, B_{2\rho}(x)) = c(n, p, c_0)\, \rho^{n-p}, \quad (15)$$

for any $x \in E$ and for any $0 < \rho < \rho_0$.

For the sake of simplicity, from now on we shall use the short notation *thick* to denote a 2-*thick* domain.

As stated in the introduction, one can prove that domains satisfying Definition 2 include those with Lipschitz boundaries or even those that satisfy a uniform corkscrew condition. Furthermore, the estimate in (15) remains valid for balls centered *outside* a uniformly p-thick domain (and near the boundary), as stated in the next lemma; see [37, Lemma 3.8]. This will be useful since we are going to deal with bounded sets *whose complementary* is thick.

Lemma 2 *Let Ω be a bounded domain such that $\mathbb{R}^n \setminus \Omega$ is uniformly p-thick with constants c_0 and ρ_0, and let $y \in \Omega$ be such that $B_{\rho/6}(y) \setminus \Omega \neq \emptyset$. Then there exists a constant $c_1 = c_1(n, p, \rho_0, c_0)$ such that*

$$\mathrm{cap}_p(\overline{B_{\rho/4}(y)} \setminus \Omega, B_{\rho/2}(y)) \geq c_1 \, \mathrm{cap}_p(\overline{B_{\rho/4}(y)}, B_{\rho/2}(y)) = c\,\rho^{n-p}.$$

Also, a uniform p-thick domain enjoys a self-improving property, as stated in the following theorem due to Lewis [28] (see also Section 8 of [30] for a complete survey on the boundary regularity).

Theorem 5 *Let Ω be a bounded domain such that $\mathbb{R}^n \setminus \Omega$ is uniformly p-thick with constants c_0 and ρ_0, and $1 < p \leq n$. Then there exists $q = q(n, p, c_0) \in (1, p)$ such that $\mathbb{R}^n \setminus \Omega$ is uniformly q-thick with constant $c_2 = c_2(n, p, c_0)$ and ρ_0. Moreover, q can be chosen near p so that $q \in (np/(n+p), p)$. Thus, we have*

$$\text{cap}_q(\overline{B_{\rho/4}(y)} \setminus \Omega, B_{\rho/2}(y)) \geq c_2 \, \text{cap}_q(\overline{B_{\rho/4}(y)}, B_{\rho/2}(y)) \geq c(n, p, c_0)\rho^{n-q}$$

for $y \in \Omega$ be such that $B_{\rho/6}(y) \setminus \Omega \neq \emptyset$.

Remark 4 On the other hand a p-thick domain is r-thick for every $r \geq p$. Indeed, for $x \in E$, $\rho < \rho_0$ and $\varphi \in C_0^\infty(B_{2\rho}(x))$, $\varphi \geq \mathbb{1}_{E \cap \overline{B_\rho(x)}}$, by the Hölder inequality, we have

$$\int_{B_{2\rho}} |D\varphi|^p \, \mathrm{d}x \leq c(n, p, r) \left(\int_{B_{2\rho}} |D\varphi|^r \, \mathrm{d}x \right)^{\frac{p}{r}} \rho^{n\left(1 - \frac{p}{r}\right)}.$$

Taking the infimum for such φ we get

$$c(n, p, c_0) \, \rho^{n-p} \leq \text{cap}_p(E \cap \overline{B_\rho}, B_{2\rho}) \leq c \, \rho^{n\left(1 - \frac{p}{r}\right)} \left[\text{cap}_r(E \cap \overline{B_\rho}, B_{2\rho}) \right]^{\frac{p}{r}}$$

and r-thickness follows by taking the r/p-power.

We conclude this section by recalling some capacity Sobolev type inequalities; see for example [23, Lemma 3.3] or [30, Lemma 8.11 and Remark 8.14].

Lemma 3 *Let $w \in W^{1,p}(B_{2\rho})$ and $K_\rho(w) = \{x \in \overline{B_\rho} : w(x) = 0\}$. Define $\kappa = n/(n-p)$ if $p < n$ and $\kappa = 2$ if $p \geq n$. Then there exists a positive constant $c = c(n, p)$ such that*

$$\left(\fint_{B_{2\rho}} |w|^{\kappa p} \, \mathrm{d}x \right)^{\frac{1}{\kappa p}} \leq \left(\frac{c}{\text{cap}_p(K_\rho(w), B_{2\rho})} \int_{B_{2\rho}} |Dw|^p \, \mathrm{d}x \right)^{\frac{1}{p}}.$$

The previous inequality also holds in the case stated below, with p-powers on both sides; that is, we have the following Poincaré type inequality

Lemma 4 *Let $w \in W^{1,p}(B_{2\rho})$ and $K_\rho(w)$ as before. Then there exists a positive constant $c \equiv c(n, p)$ such that*

$$\fint_{B_{2\rho}} |w|^p \, \mathrm{d}x \leq \frac{c}{\text{cap}_p(K_\rho(w), B_{2\rho})} \int_{B_{2\rho}} |Dw|^p \, \mathrm{d}x. \tag{16}$$

Now a parabolic Sobolev inequality; its proof requires only slight changes with respect to the proof of Lemma 3.17 in [37]. Essentially, it is enough to apply Sobolev's inequality Lemma 3 slicewise to the continuation to zero of w in $\mathcal{Q}_{\rho/2} \setminus \Omega_T$. Note moreover that the exponent q appearing therein is the same present in Theorem 5, for $p = 2$, and that here it is unavoidable such a self-improving property.

Lemma 5 *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain such that $\mathbb{R}^n \setminus \Omega$ is uniformly thick with constants c_0 and ρ_0, and let $\mathcal{Q}_{\rho/2} := B_{\rho/2}(x_1) \times (t_1 - (\rho/2)^2, t_1)$ be a parabolic cylinder such that $B_{\rho/6}(x_1) \setminus \Omega \neq \emptyset$. Let moreover w be a function such that*

$$w \in L^2(\mathcal{Q}_{\rho/2} \cap \Omega_T), \quad Dw \in L^2(\mathcal{Q}_{\rho/2} \cap \Omega_T),$$

$$w \in L^\infty(t_1 - (\rho/2)^2, t_1; L^2(B_{\rho/2}(x_1)))$$

and $w \equiv 0$ on $\mathcal{Q}_{\rho/2} \cap \partial_{\mathrm{lat}}\Omega_T$. Then there exist an exponent $q \equiv q(n, c_0) < 2$ and a constant $c = c(n, c_0, \rho_0)$ such that

$$\frac{1}{|\mathcal{Q}_{\rho/2}|} \int_{\mathcal{Q}_{\rho/2} \cap \Omega_T} |w|^2 \, \mathrm{d}x \, \mathrm{d}t \le c \left(\frac{1}{|\mathcal{Q}_{\rho/2}|} \int_{\mathcal{Q}_{\rho/2} \cap \Omega_T} |Dw|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{2n}{q(n+2)}}$$

$$\times \left(\sup_{\tau \in (t_1 - (\rho/2)^2, t_1) \cap (-T, 0)} \int_{B_{\rho/2} \cap \Omega} |w(\cdot, \tau)|^2 \, \mathrm{d}x \right)^{\frac{2}{n+2}} \quad (17)$$

and

$$\int_{\mathcal{Q}_{\rho/2} \cap \Omega_T} |w|^{2(1+\frac{2}{n})} \, \mathrm{d}x \, \mathrm{d}t \le c \int_{\mathcal{Q}_{\rho/2} \cap \Omega_T} |Dw|^2 \, \mathrm{d}x \, \mathrm{d}t$$

$$\times \left(\sup_{\tau \in (t_1 - (\rho/2)^2, t_1) \cap (-T, 0)} \int_{B_{\rho/2} \cap \Omega} |w(\cdot, \tau)|^2 \, \mathrm{d}x \right)^{\frac{2}{n}}. \quad (18)$$

2.5 A parabolic covering lemma

Below we present a technical lemma (see forthcoming Lemma 6) which is nothing that a special version of the classical Calderón-Zygmund-Krylov-Safonov decomposition. It will allow us to work with parabolic cylinders that have as horizontal slice a ball instead of a cube. We would recall that a parabolic version of the Calderón-Zygmund-Krylov-Safonov covering lemma was proved by one of the authors in [5], following the elliptic analog (which could be found, for instance, in [11, Lemma 1.2]). Precisely, in [5, Proposition 2.1] it is shown that a covering of a set by parabolic cylinders can be "arranged" in an optimal disjoint way so that in each of the relative horizontal slice cube the density is considerably small.

Since in the present paper we are interesting into analyzing some regularity properties of problem (1) up to the boundary, we need to work with parabolic cylinders having a ball as horizontal slice. For this, in order to prove Lemma 6 below, we will extend its elliptic counterpart given by Lemma 3.1 in [40] to a parabolic framework.

Lemma 6 *Let $\mathcal{Q}_{\bar{R}} \subset \mathbb{R}^{n+1}$. Assume that $\mathcal{X} \subset \mathcal{Y} \subset \mathcal{Q}_{\bar{R}}$ are measurable sets such that there exists $0 < \delta < 1$ for which the following properties hold:*

(i) $|\mathcal{X}| < \delta |\mathcal{Q}_{\bar{R}}|$;

(ii) for all $(x,t) \in \mathcal{Q}_{\bar{R}}$ and for all $\rho \in (0, \bar{R}]$ if $|\mathcal{X} \cap \mathcal{Q}_\rho(x,t)| \geq \delta|\mathcal{Q}_\rho(x,t)|$ then $\mathcal{Q}_\rho(x,t) \cap \mathcal{Q}_{\bar{R}} \subset \mathcal{Y}.$

Then we have

$$|\mathcal{X}| \leq c_3\,\delta|\mathcal{Y}|$$

where $c_3 = c_3(n)$.

Proof Since (i) holds, for almost every $(x,t) \in \mathcal{X}$ there exists a $\rho_{x,t} < \bar{R}$ so that $|\mathcal{X} \cap \mathcal{Q}_{\rho_{x,t}}(x,t)| = \delta|\mathcal{Q}_{\rho_{x,t}}|$ and $|\mathcal{X} \cap \mathcal{Q}_\rho(x,t)| < \delta|\mathcal{Q}_\rho|$ for all $\rho_{x,t} < \rho < \bar{R}$. Then, by the Vitali covering lemma we can extract a sequence of disjoint cylinders $\{Q_{\rho_{x_k,t_k}}(x_k,t_k)\}_k$ such that $\mathcal{X} \subset \left(\bigcup_k \mathcal{Q}_{5\rho_{x_k,t_k}}(x_k,t_k)\right) \cap \mathcal{Q}_{\bar{R}}$. In view of the choice of $\mathcal{Q}_{\rho_{x_k,t_k}}(x_k,t_k)$, we infer that

$$|\mathcal{X} \cap \mathcal{Q}_{5\rho_{x_k,t_k}}(x_k,t_k)| < \delta|\mathcal{Q}_{5\rho_{x_k,t_k}}| = 5^{n+2}\delta|\mathcal{Q}_{\rho_{x_k,t_k}}| = 5^{n+2}|\mathcal{X} \cap \mathcal{Q}_{\rho_{x_k,t_k}}(x_k,t_k)|.$$

Moreover, there exists a constant $c(n)$ such that

$$|\mathcal{Q}_{\rho_{x_k,t_k}}(x_k,t_k)| \leq c(n)|\mathcal{Q}_{\rho_{x_k,t_k}}(x_k,t_k) \cap \mathcal{Q}_{\bar{R}}|,$$

since $(x_k,t_k) \in \mathcal{Q}_{\bar{R}}$ and $\rho_{x_k,t_k} < \bar{R}$.

All in all, we have

$$
\begin{aligned}
|\mathcal{X}| = \left|\left(\cup_k \mathcal{Q}_{5\rho_{x_k,t_k}}(x_k,t_k)\right) \cap \mathcal{X}\right| &\leq \sum_k |\mathcal{Q}_{5\rho_{x_k,t_k}}(x_k,t_k) \cap \mathcal{X}| \\
&\leq \sum_k 5^{n+2}\delta|\mathcal{Q}_{\rho_{x_k,t_k}}(x_k,t_k)| \\
&\leq c(n)5^{n+2}\sum_k \delta|\mathcal{Q}_{\rho_{x_k,t_k}}(x_k,t_k) \cap \mathcal{Q}_{\bar{R}}| \\
&\leq c(n)5^{n+2}\delta|\mathcal{Y}|,
\end{aligned}
$$

where we used (ii) in the last inequality. This concludes the proof. $\qquad\square$

3 (Higher) regularity for the homogeneous problem and comparison results

One of the key-points in the proofs of the main results in the present paper relies on obtaining both local interior and boundary comparison estimates.

Let us first handle the interior ones by considering the following homogeneous Cauchy-Dirichlet problem.

$$
\begin{cases}
w_t - \operatorname{div} a(x,t,Dw) = 0 & \text{in } \mathcal{Q}_R \\
w = u & \text{on } \partial_{\mathrm{par}}\mathcal{Q}_R,
\end{cases}
\tag{19}
$$

with $\mathcal{Q}_R \equiv \mathcal{Q}_R(x_0,t_0) \subset \Omega_T$ and u being the unique solution to the regularized problem (10). We recall some basic results from the higher integrability theory of Gehring. For the proofs we refer to [29]; see also [5, Theorem 5.5].

Theorem 6 *Let $w \in C^0([t_0 - R^2, t_0]; L^2(B_R(x_0))) \cap L^2(t_0 - R^2, t_0; W^{1,2}(B_R(x_0)))$ be a weak solution to the parabolic equation $(19)_1$. Then there exists $\chi_1 = \chi_1(n, L, \nu) > 1$ such that $Dw \in L^{2\chi_1}_{\text{loc}}(Q_R)$ and for any $q \in (0, 2]$ it holds*

$$\left(\fint_{Q_{R/2}} |Dw|^{2\chi_1} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2\chi_1}} \leq c \left(\fint_{Q_R} |Dw|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{q}}. \tag{20}$$

Moreover, for any $q \in (0, 2]$ and for any $\chi > 1$ it holds

$$\left(\fint_{Q_{R/2}} |w|^{2\chi} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2\chi}} \leq c \left(\fint_{Q_R} |w|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{q}}. \tag{21}$$

The constants c in (20) and (21) depend only on n, L, ν and q.

Concerning local interior comparison estimates we can use those established in [5, Lemma 5.2] or [4, Lemma 6.4].

Lemma 7 *Let $u \in C^0([-T, 0]; L^2(\Omega)) \cap L^2(-T, 0; W_0^{1,2}(\Omega))$ be the unique solution to the regularized problem (10). Moreover, let w be the unique solution of the Cauchy-Dirichlet problem (19). Then there exists a constant $c = c(n, \nu)$ such that*

$$\int_{Q_R} \left(R^{-1} |u - w| + |Du - Dw| \right) \mathrm{d}x \, \mathrm{d}t \leq c R \int_{Q_R} |f| \, \mathrm{d}x \, \mathrm{d}t.$$

For the local boundary estimates we consider $Q_R = Q_R(x_0, t_0) \subset \mathbb{R}^{n+1}$ such that it intersects the lateral boundary of Ω_T, that is

$$Q_R \cap (\partial\Omega \times (-T, 0)) = Q_R \cap \partial_{\text{lat}}\Omega_T \neq \emptyset.$$

This will be enough for our purposes, see Remark 5 at the beginning of Section 4. Then, we consider the unique solution to the following homogeneous Cauchy-Dirichlet problem

$$\begin{cases} w_t - \operatorname{div} a(x, t, Dw) = 0 & \text{in } Q_R \cap \Omega_T \\ w = u & \text{on } \partial_{\text{par}}(Q_R \cap \Omega_T). \end{cases} \tag{22}$$

Here we need a higher integrability result for w, the counterpart of Theorem 6 up to the boundary, relying on some reverse Hölder inequalities. This kind of boundary higher integrability is already present in the literature (see for instance [37, Theorem 4.7], and also [3,10,36] for more regular boundaries), but we prefer to propose an independent statement (and subsequently the relative proof) fitting our context.

For this, we need a couple of technical lemmata, the first of which is an appropriate version of the Gehring lemma, see [18, Proposition 1.3]:

Lemma 8 *Let $A \subset \mathbb{R}^{n+1}$ an open set and let $g : A \to \mathbb{R}^k$ be an integrable map such that*

$$\fint_{\mathcal{Q}_{\rho/8}} |g|^2 \, \mathrm{d}x \, \mathrm{d}t \leq \hat{c} \left(\fint_{\mathcal{Q}_\rho} |g|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{2}{q}} + \vartheta \fint_{\mathcal{Q}_\rho} |g|^2 \, \mathrm{d}x \, \mathrm{d}t,$$

for some constant $q < 2$ and for all $\mathcal{Q}_\rho \subset A$. Then there exists a constant $\vartheta_0 = \vartheta_0(n, q)$ such that if $\vartheta < \vartheta_0$ then $g \in L^{2\chi_1}_{\mathrm{loc}}(A)$ with $\chi_1 > 1$ depending on $n, \hat{c}, q, \vartheta$; moreover

$$\left(\fint_{\mathcal{Q}_{\rho/2}} |g|^{2\chi_1} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{\chi_1}} \leq c \fint_{\mathcal{Q}_\rho} |g|^2 \, \mathrm{d}x \, \mathrm{d}t$$

for all $\mathcal{Q}_\rho \subset A$. The constant c also depends on $n, \hat{c}, q, \vartheta$.

The second one encodes the self-improving character of reverse Hölder inequalities. The proof follows [19, Remark 6.12] and uses a covering argument.

Lemma 9 *Let A, g as in* Lemma 8 *and let*

$$\left(\fint_{\mathcal{Q}_{\rho/2}} |g|^{2\chi_1} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{\chi_1}} \leq \hat{c} \fint_{\mathcal{Q}_\rho} |g|^2 \, \mathrm{d}x \, \mathrm{d}t$$

hold whenever $\mathcal{Q}_\rho \subset A$ is a cylinder, where $\chi_1 > 1$ and $\hat{c} \geq 1$. Then, for every $\sigma \in (0, 2]$, there exists a constant $c = c(n, \sigma, \hat{c})$ such that

$$\left(\fint_{\mathcal{Q}_{\rho/2}} |g|^{2\chi_1} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2\chi_1}} \leq c \left(\fint_{\mathcal{Q}_\rho} |g|^\sigma \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{\sigma}}$$

for every $\mathcal{Q}_\rho \subset A$.

Now, we are in a position to prove the following

Theorem 7 *Let Ω be a bounded domain such that $\mathbb{R}^n \setminus \Omega$ satisfies a uniform thick condition with constants c_0 and ρ_0, and let w be the unique solution to (22). Then there exists a constants $\chi_1 = \chi_1(n, \nu, L, \rho_0, c_0) > 1$ such that for every $q > 0$ there holds*

$$\left(\frac{1}{|\mathcal{Q}_{R/2}|} \int_{\mathcal{Q}_{R/2} \cap \Omega_T} |Dw|^{2\chi_1} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2\chi_1}} \leq c \left(\frac{1}{|\mathcal{Q}_R|} \int_{\mathcal{Q}_R \cap \Omega_T} |Dw|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{q}} \tag{23}$$

for a constant $c = c(n, \nu, L, \rho_0, c_0, q)$. Moreover we also have

$$\left(\frac{1}{|\mathcal{Q}_{R/2}|} \int_{\mathcal{Q}_{R/2} \cap \Omega_T} |w|^{2\chi} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2\chi}} \leq c \left(\frac{1}{|\mathcal{Q}_R|} \int_{\mathcal{Q}_R \cap \Omega_T} |w|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{q}} \tag{24}$$

for any $\chi = \chi(n, \nu, L, \rho_0, c_0) > 1$ and the same constant as above.

Proof The goal here is proving that if we consider the continuation of w to zero (denoted by $\tilde{w}$) in $\mathcal{Q}_R \setminus \Omega_T$, then for every parabolic cylinder $\mathcal{Q}_\rho \equiv \mathcal{Q}_\rho(x_1, t_1) \subset \mathcal{Q}_R$ we have a reverse Hölder's inequality

$$\left(\fint_{\mathcal{Q}_{\rho/8}} |D\tilde{w}|^2 \, dx \, dt \right)^{\frac{1}{2}} \le c \left(\fint_{\mathcal{Q}_\rho} |D\tilde{w}|^q \, dx \, dt \right)^{\frac{1}{q}} + \frac{\vartheta_0}{2} \left(\fint_{\mathcal{Q}_\rho} |D\tilde{w}|^2 \, dx \, dt \right)^{\frac{1}{2}},$$
(25)

for some $q < 2$ and for a constant depending on $n, \nu, L, \rho_0, c_0, \vartheta_0$, where ϑ_0 is the constant of Lemma 8 corresponding to q. Therefore we could apply the Gehring lemma in the form of Lemma 8, and then Lemma 9, to infer (23), after again restricting $\tilde{w}$ to $\mathcal{Q}_R \cap \Omega_T$. The point here is that we are not interested in the boundary higher integrability of w in the part of the same boundary inside Ω_T, but just in the higher integrability on cylinders centered in $\partial_{\text{lat}} \Omega_T$, where $w \equiv u \equiv 0$. In a very rough sense, this is similar to interior higher integrability. Hence now we fix a parabolic cylinder $\mathcal{Q}_\rho \subset \mathcal{Q}_R$; note that due to the assumption $\mathcal{Q}_R \cap \Omega_0 = \emptyset$, we also have $\mathcal{Q}_\rho \cap \Omega_0 = \emptyset$.

The scheme of the proof is the following: first we consider the case where $B_{\rho/6} \setminus \Omega \ne \emptyset$. Here we first use a Caccioppoli type inequality to, roughly speaking, control the gradient Dw over $B_{\rho/4} \cap \Omega_T$ with the integral of w over $B_{\rho/2} \cap \Omega_T$. By applying the Sobolev inequality over $B_{\rho/2} \cap \Omega_T$, we now need to estimate the supremum in (32) in terms of $|Dw|$ and this is again achieved with the help of a Caccioppoli type inequality and the Poincaré capacity inequality in Lemma 4. Finally, we get (25) by narrowing the set on the left-hand side. On the other hand, if $B_{\rho/6} \setminus \Omega = \emptyset$, estimate (25) between $B_{\rho/8}$ and $B_{\rho/6}$ is just an interior one, and then we can enlarge the domain on the right-hand side.

Suppose now hence $B_{\rho/6} \setminus \Omega \ne \emptyset$ (but clearly also $B_{\rho/6} \cap \Omega \ne \emptyset$, otherwise there would be nothing to prove) and consider a cut-off function $\varphi \in C_0^\infty(\mathcal{Q}_R)$, $0 \le \varphi \le 1$ such that $\varphi \equiv 1$ on $\mathcal{Q}_{\rho/4}$, $\text{supp } \varphi \subset \mathcal{Q}_{\rho/2}$ and $|D\varphi| + \rho|\varphi_t| \le c/\rho$. Note that from now on all the cylinders with radii multiple of ρ will have vertex (x_1, t_1). Up to a standard regularization of w in time by Steklov averages, we use as a test function in (22) the map $\phi := w\varphi^2 \zeta_\varepsilon \mathbb{1}_{\Omega_T}$, where, once fixed $\tau \in (t_1 - (\rho/2)^2, t_1)$, $\zeta_\varepsilon \equiv \zeta_{\varepsilon,\tau}(t) \in W^{1,\infty}(\mathbb{R})$ is a piecewise linear continuous map, for $\varepsilon \le (t_1 - \tau)/2$, such that $\zeta_\varepsilon \equiv 1$ in $(-\infty, \tau]$ and $\zeta_\varepsilon \equiv 0$ in $[\tau + \varepsilon, +\infty)$. Note that for a.e. $t \in (t_1 - (\rho/2)^2, t_1)$, we have $\phi(\cdot, t) \in W_0^{1,2}(\Omega)$ and thus can be used in (9). We therefore have, integrating over $(t_1 - (\rho/2)^2, t_1)$,

$$\int_{\mathcal{Q}_{\rho/2} \cap \Omega_T} w_t w \varphi^2 \zeta_\varepsilon \, dx \, dt + \int_{\mathcal{Q}_{\rho/2} \cap \Omega_T} \langle a(x, t, Dw), Dw \rangle \varphi^2 \zeta_\varepsilon \, dx \, dt$$
$$= -2 \int_{\mathcal{Q}_{\rho/2} \cap \Omega_T} \langle a(x, t, Dw), D\varphi \rangle w \varphi \zeta_\varepsilon \, dx \, dt. \quad (26)$$

Now, we treat the first term in the following way, using integration by parts and recalling that $\varphi \zeta_\varepsilon \equiv 0$ on $(B_{\rho/2} \cap \Omega) \times \{t_1 - (\rho/2)^2, t_1\}$,

$$\int_{\mathcal{Q}_{\rho/2}\cap\Omega_T} w_t w\varphi^2\zeta_\varepsilon \,\mathrm{d}x\,\mathrm{d}t = \frac{1}{2}\int_{\mathcal{Q}_{\rho/2}\cap\Omega_T} \left(|w|^2\right)_t \varphi^2\zeta_\varepsilon \,\mathrm{d}x\,\mathrm{d}t$$

$$= -\frac{1}{2}\int_{\mathcal{Q}_{\rho/2}\cap\Omega_T} |w|^2(\varphi^2\zeta_\varepsilon)_t \,\mathrm{d}x\,\mathrm{d}t$$

$$= -\int_{\mathcal{Q}_{\rho/2}\cap\Omega_T} |w|^2\varphi\varphi_t\zeta_\varepsilon \,\mathrm{d}x\,\mathrm{d}t + \frac{1}{2}\int_\tau^{\tau+\varepsilon}\int_{B_{\rho/2}\cap\Omega} |w|^2\varphi^2 \,\mathrm{d}x\,\mathrm{d}t. \tag{27}$$

Letting $\varepsilon \searrow 0$ and then taking the (essential) supremum with respect to $\tau \in (t_1 - (\rho/2)^2, t_1)$ we therefore get

$$\sup_{\tau\in(t_1-(\rho/2)^2,t_1)} \int_{\mathcal{Q}_{\rho/2,\tau}\cap\Omega_T} w_t w\varphi^2 \,\mathrm{d}x\,\mathrm{d}t \geq -c\int_{\mathcal{Q}_{\rho/2}\cap\Omega_T} \frac{|w|^2}{\rho^2} \,\mathrm{d}x\,\mathrm{d}t$$

$$+ \frac{1}{2}\sup_{\tau\in(t_1-(\rho/2)^2,t_1)} \int_{B_{\rho/2}\cap\Omega} \left[|w|^2\varphi^2\right](\cdot,\tau)\,\mathrm{d}x. \tag{28}$$

where we denoted for shortness with $\mathcal{Q}_{\rho/2,\tau}$ the set $\mathcal{Q}_{\rho/2}\cap\mathbb{1}_{\{t<\tau\}}$. The second term in the left-hand side of (26) is estimated from below

$$\int_{\mathcal{Q}_{\rho/2}\cap\Omega_T} \langle a(x,t,Dw),Dw\rangle\varphi^2\zeta_\varepsilon \,\mathrm{d}x\,\mathrm{d}t \geq \frac{\nu}{2}\int_{\mathcal{Q}_{\rho/2,\tau}\cap\Omega_T} |Dw|^2\varphi^2 \,\mathrm{d}x\,\mathrm{d}t \tag{29}$$

with $c = c(\nu, L)$, having at hand the following monotonicity condition, that can be deduced by the assumptions in (2):

$$\langle a(x,t,z),z\rangle \geq \frac{\nu}{2}|z|^2 \text{ for any } (x,t)\in\Omega_T \text{ and any } z\in\mathbb{R}^n.$$

For the latter term in the right-hand side of (26) we have, using again the growth condition $(2)_2$ together with the Young inequality,

$$2\int_{\mathcal{Q}_{\rho/2}\cap\Omega_T} \langle a(x,t,Dw),D\varphi\rangle w\varphi\zeta_\varepsilon \,\mathrm{d}x\,\mathrm{d}t$$

$$\leq \frac{\nu}{4}\int_{\Omega_T\cap\mathcal{Q}_{\rho/2,\tau+\varepsilon}} |Dw|^2\varphi^2 \,\mathrm{d}x\,\mathrm{d}t + \frac{c(\nu,L)}{\rho^2}\int_{\mathcal{Q}_{\rho/2}\cap\Omega_T} |w|^2 \,\mathrm{d}x\,\mathrm{d}t. \tag{30}$$

To conclude we put (27), (29) and (30) into (26), we let $\varepsilon \searrow 0$, we perform some algebraic manipulations and we take the supremum with respect to $\tau \in (t_1 - (\rho/2)^2, t_1)$. Using also (28) finally we get

$$\sup_{\tau\in(t_1-(\rho/4)^2,t_1)} \int_{B_{\rho/4}\cap\Omega} |w(\cdot,\tau)|^2 \,\mathrm{d}x + \int_{\mathcal{Q}_{\rho/4}\cap\Omega_T} |Dw|^2 \,\mathrm{d}x\,\mathrm{d}t$$

$$\leq \frac{c(\nu,L)}{\rho^2}\int_{\mathcal{Q}_{\rho/2}\cap\Omega_T} |w|^2 \,\mathrm{d}x\,\mathrm{d}t, \tag{31}$$

recalling that $\varphi \equiv 1$ on $\mathcal{Q}_{\rho/4}$. Note that also a completely similar estimate, where the cylinder $\mathcal{Q}_{\rho/2}$ appears on the left-hand side, and the cylinder $\mathcal{Q}_\rho$ on the right-hand side, follows by straightforward changes in the proof above.

We recall now we are supposing $B_{\rho/6} \setminus \Omega \neq \emptyset$; therefore we can use the Sobolev inequality in Lemma 5 in the right-hand side of (31) and, after dividing by $|\mathcal{Q}_{\rho/2}|$, we infer

$$
\frac{c}{|\mathcal{Q}_{\rho/4}|} \int_{\mathcal{Q}_{\rho/4} \cap \Omega_T} |Dw|^2 \, \mathrm{d}x \, \mathrm{d}t
$$

$$
\leq c \left(\frac{1}{|\mathcal{Q}_{\rho/2}|} \int_{\mathcal{Q}_{\rho/2} \cap \Omega_T} |Dw|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{2n}{q(n+2)}} \tag{32}
$$

$$
\times \left(\sup_{\tau \in (t_1 - (\rho/2)^2, t_1)} \frac{1}{|\mathcal{Q}_{\rho/2}|} \int_{B_{\rho/2} \cap \Omega} |w(\cdot, \tau)|^2 \, \mathrm{d}x \right)^{\frac{2}{n+2}}
$$

where c depends upon n, ν, L, ρ_0, c_0. Finally, we estimate the supremum on the right-hand side in the following way: for a. e. $\tau \in (t_1 - (\rho/2)^2, t_1)$ using (31) in its version over $\mathcal{Q}_{\rho/2}$ and $\mathcal{Q}_\rho$ and Poincaré's inequality (16) slice-wise, for $p = 2$, we have

$$
\int_{B_{\rho/2} \cap \Omega} |w(\cdot, \tau)|^2 \, \mathrm{d}x \leq \frac{c}{\rho^2} \int_{\mathcal{Q}_\rho \cap \Omega_T} |w|^2 \, \mathrm{d}x \, \mathrm{d}t
$$

$$
\leq \frac{c}{\rho^2} \int_{t_1 - \rho^2}^{t_1} \frac{\rho^n}{\mathrm{cap}_2(K_{\rho/2}(w), B_\rho)} \int_{B_\rho} |Dw|^2 \, \mathrm{d}x \, \mathrm{d}t
$$

where $K_{\rho/2}(w) = \{x \in \overline{B_{\rho/2}} : w(x) = 0\}$. We recall that we extended w to zero in $\mathcal{Q}_R \setminus \Omega_T$; note this gives $w \in W^{1,2}(B_R)$ for a. e. t since $w \equiv 0$ on $\partial_{\mathrm{lat}}\Omega_T$. Now, since $K_{\rho/2}(w) \supset \overline{B_{\rho/2}} \setminus \Omega$ and $B_{\rho/6} \setminus \Omega \neq \emptyset$ we can apply Lemma 2 to get

$$
\mathrm{cap}_2(K_{\rho/2}(w), B_\rho) \geq c_1 \, \mathrm{cap}_2(\overline{B_{\rho/2}}, B_\rho) = c \, \rho^{n-2},
$$

and therefore, again dividing by $|\mathcal{Q}_{\rho/2}|$,

$$
\sup_{\tau \in (t_1 - (\rho/2)^2, t_1)} \frac{1}{|\mathcal{Q}_{\rho/2}|} \int_{B_{\rho/2} \cap \Omega} |w(\cdot, \tau)|^2 \, \mathrm{d}x \leq \frac{c}{|\mathcal{Q}_{\rho/2}|} \int_{\mathcal{Q}_\rho \cap \Omega_T} |Dw|^2 \, \mathrm{d}x \, \mathrm{d}t.
$$

Putting the latter estimate into (32) and using the Young inequality with conjugate exponents $(n+2)/2$ and $(n+2)/n$, with ϑ_0 being the constant appearing

in Lemma 8 corresponding to the choice of q following from Lemma 5, we infer

$$\frac{c}{|\mathcal{Q}_{\rho/4}|} \int_{\mathcal{Q}_{\rho/4} \cap \Omega_T} |Dw|^2 \, \mathrm{d}x \, \mathrm{d}t$$

$$\leq c \left(\frac{1}{|\mathcal{Q}_{\rho/2}|} \int_{\mathcal{Q}_{\rho/2} \cap \Omega_T} |Dw|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{2n}{q(n+2)}} \left(\frac{1}{|\mathcal{Q}_{\rho/2}|} \int_{\mathcal{Q}_\rho \cap \Omega_T} |Dw|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{2}{n+2}}$$

$$\leq c(\vartheta_0) \left(\frac{1}{|\mathcal{Q}_{\rho/2}|} \int_{\mathcal{Q}_{\rho/2} \cap \Omega_T} |Dw|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{2}{q}}$$

$$+ \left(\frac{\vartheta_0}{100} \right)^2 \frac{1}{|\mathcal{Q}_\rho|} \int_{\mathcal{Q}_\rho \cap \Omega_T} |Dw|^2 \, \mathrm{d}x \, \mathrm{d}t.$$

The inequality above plainly yields (25) in the case $B_{\rho/6} \setminus \Omega \neq \emptyset$.

The case $B_{\rho/6} \setminus \Omega = \emptyset$ is easier, since now an estimate

$$\left(\fint_{\mathcal{Q}_{\rho/8}} |D\tilde{w}|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2}} \leq c \left(\fint_{\mathcal{Q}_{\rho/6}} |D\tilde{w}|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{q}},$$

q possibly different from the one appearing in (32), is an interior one and can be deduced following [18] or [24]. At this point we just enlarge the integral on the right-hand side to get (25). Note that actually higher integrability estimate (23) holds for every couple of cylinders $\mathcal{Q}_{\rho/4}$, $\mathcal{Q}_{\rho/2}$ such that $\mathcal{Q}_\rho \subset \mathcal{Q}_R$, and we shall use this fact in few lines.

Finally, the estimate (24) for w is a straightforward consequence of boundary sup-estimate for parabolic problems with null boundary datum. For this, we refer for instance to [29, Theorem 6.30] in a slightly different, linear setting. For the sake of the reader we also propose in the following lemma a proof adapted to our framework. $\qquad \square$

Lemma 10 *Let w be the solution of the Cauchy-Dirichlet problem (22). Then for every $q > 0$ it holds*

$$\sup_{\mathcal{Q}_{R/2} \cap \Omega_T} |w| \leq c \left(\frac{1}{|\mathcal{Q}_R|} \int_{\mathcal{Q}_R \cap \Omega_T} |w|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{q}}, \tag{33}$$

for a constant c depending $n, \nu, L, c_0, \rho_0, q$.

Proof Consider, for $R/2 \leq \rho \leq R$ and $h \in \mathbb{R}$, the quantities

$$U(h, \rho) := \int_{\mathcal{Q}_\rho \cap \Omega_T} \left| (w - h)_+ \right|^2 \, \mathrm{d}x \, \mathrm{d}t,$$

$$V(h, \rho) := |\mathcal{Q}_\rho \cap \Omega_T \cap \{w \geq h\}| =: |A(h, \rho)|; \tag{34}$$

here s_+ is the positive parts of s; here all the cylinders share the same center (x_0, t_0). For $\rho_1 < \rho_2$ both belonging to $[R/2, R]$ and with $h < k$, we have the relations

$$
V(k, \rho_1) \leq \frac{1}{(k-h)^2} U(h, \rho_2),
$$
$$
U(k, \rho_1) \leq c \frac{1}{(\rho_2 - \rho_1)^2} U(h, \rho_2) \big[V(h, \rho_2) \big]^{\frac{2}{n+2}}. \tag{35}
$$

The first inequality plainly follows by the involved definitions. Indeed,

$$
V(k, \rho_1)(k-h)^2 \leq \int_{A(k, \rho_1)} (w-h)_+^2 \, \mathrm{d}x \, \mathrm{d}t \leq \int_{A(h, \rho_1)} (w-h)_+^2 \, \mathrm{d}x \, \mathrm{d}t. \tag{36}
$$

For the second one, we need some work. First of all, minor modifications in the proof of Caccioppoli's inequality (31) give for all $i \in \mathbb{N}$

$$
\sup_{\tau \in (t_0 - (\rho_1 + \rho_2)^2/4, t_0)} \int_{B_{(\rho_1 + \rho_2)/2} \cap \Omega} \big| (w-k)_+ (\cdot, \tau) \big|^2 \, \mathrm{d}x
$$
$$
+ \int_{\mathcal{Q}_{(\rho_1 + \rho_2)/2} \cap \Omega_T} \big| D(w-k)_+ \big|^2 \, \mathrm{d}x \, \mathrm{d}t
$$
$$
\leq \frac{c(\nu, L)}{(\rho_2 - \rho_1)^2} \int_{\mathcal{Q}_{\rho_2} \cap \Omega_T} \big| (w-k)_+ \big|^2 \, \mathrm{d}x \, \mathrm{d}t \tag{37}
$$

for every $h \in \mathbb{R}$; in particular we need to choose here the test function

$$
\phi := (w-k)_+ \varphi^2 \zeta_\varepsilon \mathbb{1}_{\Omega_T},
$$

with $\varphi \in C_0^\infty(\mathcal{Q}_R)$, $0 \leq \varphi \leq 1$ such that $\varphi \equiv 1$ on $\mathcal{Q}_{(\rho_1 + \rho_2)/2}$, $\operatorname{supp} \varphi \subset \mathcal{Q}_{\rho_2}$ and $|D\varphi| + (\rho_2 - \rho_1)|\varphi_t| \leq c/(\rho_2 - \rho_1)$.

Now, consider a cut-off function $\eta \in C_0^\infty(\mathcal{Q}_{(\rho_1 + \rho_2)/2})$ such that $0 \leq \eta \leq 1$, $\eta \equiv 1$ on $\mathcal{Q}_{\rho_1}$ and $|D\eta| \leq c/(\rho_2 - \rho_1)$. By Hölder's inequality and the fact that $|\eta| \leq 1$

$$
\int_{\mathcal{Q}_{\rho_1} \cap \Omega_T} \big| (w-k)_+ \big|^2 \, \mathrm{d}x \, \mathrm{d}t \leq c(n) \int_{\mathcal{Q}_{(\rho_1 + \rho_2)/2} \cap \Omega_T} \big| (w-k)_+ \eta \big|^2 \, \mathrm{d}x \, \mathrm{d}t
$$
$$
\leq c(n) \left(\int_{\mathcal{Q}_{(\rho_1 + \rho_2)/2} \cap \Omega_T} \big| (w-k)_+ \eta \big|^{2(1+\frac{2}{n})} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{n}{n+2}} \big[V(\rho_2, k) \big]^{\frac{2}{n+2}}.
$$

To estimate the right-hand side we use Sobolev's embedding Lemma 5 and then twice Caccioppoli's inequality (37),

$$
\int_{\mathcal{Q}_{(\rho_1+\rho_2)/2}\cap\Omega_T} \left| (w-k)_+ \eta \right|^{2(1+\frac{2}{n})} \mathrm{d}x\,\mathrm{d}t
$$

$$
\leq c \left(\sup_{\tau\in(t_0-(\rho_1+\rho_2)^2/4,\,t_0)} \int_{B_{(\rho_1+\rho_2)/2}\cap\Omega} \left| (w-k)_+(\cdot,\tau) \right|^2 \mathrm{d}x \right)^{\frac{2}{n}}
$$

$$
\times \left(\int_{\mathcal{Q}_{\rho_2}\cap\Omega_T} \left| D(w-k)_+ \right|^2 \mathrm{d}x\,\mathrm{d}t + \frac{1}{(\rho_2-\rho_1)^2} \int_{\mathcal{Q}_{\rho_2}\cap\Omega_T} \left| (w-k)_+ \right|^2 \mathrm{d}x\,\mathrm{d}t \right)
$$

$$
\leq c \left(\frac{1}{(\rho_2-\rho_1)^2} \int_{\mathcal{Q}_{\rho_2}\cap\Omega_T} \left| (w-k)_+ \right|^2 \mathrm{d}x\,\mathrm{d}t \right)^{1+\frac{2}{n}};
$$

we also used here the estimates for $|\eta|$ and $|D\eta|$. We estimate the first part of the latter term from above, keeping in mind that $h < k$ and therefore $A(k,\rho_2) \subset A(h,\rho_2)$

$$
\int_{\mathcal{Q}_{\rho_2}\cap\Omega_T} \left| (w-k)_+ \right|^2 \mathrm{d}x\,\mathrm{d}t \leq \int_{A(k,\rho_2)} \left| (w-h)_+ \right|^2 \mathrm{d}x\,\mathrm{d}t
$$

$$
\leq \int_{\mathcal{Q}_{\rho_2}\cap\Omega_T} \left| (w-h)_+ \right|^2 \mathrm{d}x\,\mathrm{d}t.
$$

Finally, it suffices to merge all the estimates above together with (36) (by replacing ρ_1 with ρ_2 there) to obtain the inequality in (35).

At this point, we define the sequence of level sets and radii

$$
0 < k_i := 2\left(1-\frac{1}{2^i}\right)k_1, \qquad \rho_i := \frac{R}{2} + \frac{R}{2^i}
$$

where k_1 is a positive constant which will be defined in few lines. We moreover choose into (35), for $i \in \mathbb{N}$ fixed, $\rho_1 \equiv \rho_{i+1}$, $\rho_2 \equiv \rho_i$, $k_i \equiv h < k \equiv k_{i+1}$. Call $U_i := U(k_i,\rho_i)$ and $V_i := V(k_i,\rho_i)$. By taking the α-th power, $\alpha \geq 0$, of $(35)_2$ and the β-th power, $\beta \geq 0$, of $(35)_1$, and multiplying the resulting inequalities we finally have

$$
U_{i+1}^\alpha V_{i+1}^\beta \leq c(\alpha,\beta) \frac{2^{2(\alpha+\beta)i}}{k_1^{2\beta} R^{2\alpha}} U_i^{\alpha+\beta} V_i^{\frac{2\alpha}{n+2}},
$$

where $c(\alpha,\beta)$ depends also on n,ν,L,c_0,ρ_0. Looking for $\vartheta > 1$ such that $\alpha + \beta = \vartheta\alpha$, $2\alpha/(n+2) = \vartheta\beta$ gives the value

$$
\vartheta = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2}{n+2}};
$$

therefore, fixing $\alpha = (n+2)\vartheta/2$ (and therefore $\beta = 1$), we have

$$
U_{i+1}^\alpha V_{i+1} \leq \tilde{c} \frac{2^{2\vartheta\alpha i}}{|\mathcal{Q}_R|^\vartheta k_1^2} \left(U_i^\alpha V_i \right)^\vartheta.
$$

A well known iteration argument, see [19, Lemma 7.1], ensures that

$$\lim_{i \to \infty} U_i^\alpha V_i = 0 \qquad \text{if} \qquad \left[U_1^\alpha V_1\right]^{\vartheta-1} \le \frac{2^{-\frac{\vartheta \alpha}{\vartheta-1}}}{\tilde{c}} |\mathcal{Q}_R|^\vartheta k_1^2. \tag{38}$$

Since, recalling that $\rho_i \ge R/2$ and $k_i \le k_1$,

$$0 \le \left(\fint_{\mathcal{Q}_{R/2} \cap \Omega_T} \left|(w - k_1)_+\right|^2 \mathrm{d}x \, \mathrm{d}t\right)^\alpha |\mathcal{Q}_{R/2} \cap \Omega_T \cap \{w \ge k_1\}| \le U_i^\alpha V_i \longrightarrow 0;$$

this yields $w_+ \le k_1$ almost everywhere in $\mathcal{Q}_{R/2} \cap \Omega_T$. We now find an appropriate choice of k_1 ensuring that the condition in (38) is satisfied. Since

$$\vartheta - 1 = \sqrt{\frac{1}{4} + \frac{2}{n+2}} - \frac{1}{2} = \frac{2}{n+2}\frac{1}{\vartheta} = \frac{1}{\alpha},$$

recalling the definitions in (34) and the fact that $\rho_1 = R$, we have

$$\begin{aligned}
\left[U_1^\alpha V_1\right]^{\vartheta-1} &= \left(\fint_{\mathcal{Q}_R \cap \Omega_T} \left|(w - k_1)_+\right|^2 \mathrm{d}x \, \mathrm{d}t\right)^{\alpha(\vartheta-1)} |\mathcal{Q}_R \cap \Omega_T \cap \{w \ge k_1\}|^{\vartheta-1} \\
&\le |\mathcal{Q}_R \cap \Omega_T|^{\vartheta-1} \fint_{\mathcal{Q}_R \cap \Omega_T} (w_+)^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\le |\mathcal{Q}_R|^{\vartheta-1} \fint_{\mathcal{Q}_R \cap \Omega_T} (w_+)^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\le \frac{2^{-\frac{\vartheta \alpha}{\vartheta-1}}}{\tilde{c}} |\mathcal{Q}_R|^\vartheta k_1^2,
\end{aligned}$$

if we choose as starting level

$$k_1^2 := \frac{2^{\frac{\vartheta \alpha}{\vartheta-1}} \tilde{c}}{|\mathcal{Q}_R|} \int_{\mathcal{Q}_R \cap \Omega_T} |w|^2 \, \mathrm{d}x \, \mathrm{d}t = \frac{c}{|\mathcal{Q}_R|} \int_{\mathcal{Q}_R \cap \Omega_T} |w|^2 \, \mathrm{d}x \, \mathrm{d}t,$$

$c \equiv c(n, \nu, L, c_0, \rho_0)$. A similar argument, considering the negative part instead of the positive one, yields $w_- \le k_1$. Therefore we showed the validity of (33), but with the L^2 norm on the right-hand side. We show here how to lower this exponent, following [19, Chapter 7]. First notice that the estimate

$$\sup_{\mathcal{Q}_{r_1} \cap \Omega_T} |w| \le c \left(\frac{1}{(r_2 - r_1)^{n+2}} \int_{\mathcal{Q}_{r_2} \cap \Omega_T} |w|^2 \, \mathrm{d}x \, \mathrm{d}t\right)^{\frac{1}{2}},$$

for all $R \le r_1 < r_2 \le R$ holds as well; this can be seen with slight modifications of the previous proof, or adapting [19, Corollary 7.1] to our setting. Note that this last approach in our case would require a more involved argument: we should indeed distinguish the cases as in the first part of the proof of Theorem 7, and change a bit the radii. At this point, with this estimate at hand, we argue as follows: denote by $\mathcal{U}(r)$ the supremum of $|w|$ over $\mathcal{Q}_r \cap \Omega_T$,

for $R/2 \leq r \leq R$. Then for $R/2 \leq r_1 < r_2 \leq R$ and $q > 0$, using Young's inequality

$$\mathcal{U}(r_1) \leq c \left(\frac{1}{(r_2 - r_1)^{n+2}} \int_{\mathcal{Q}_{r_2} \cap \Omega_T} |w|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2}} \left[\mathcal{U}(r_2) \right]^{\frac{2-q}{2}}$$

$$\leq \frac{1}{2} \mathcal{U}(r_2) + c \left(\frac{1}{(r_2 - r_1)^{n+2}} \int_{\mathcal{Q}_R \cap \Omega_T} |w|^q \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{q}}.$$

A well known iteration lemma (see, e.g., [19, Lemma 6.1]) at this point finally gives (33). $\qquad\square$

We conclude this section by stating the counterpart of Lemma 7 up to the boundary. The proof plainly follows by that of Lemma 4.1 in [27] (see also [14, Lemma 3.3]), once taking into account the modifications needed in order to handle the fact that the cylinder intersects the boundary, see [40, Lemma 2.6] and the proof of previous Theorem 7.

Lemma 11 *Let $u \in C^0([-T,0]; L^2(\Omega)) \cap L^2(-T, 0; W_0^{1,2}(\Omega))$ be the unique solution to the regularized problem (10), and let w be the unique solution of the Cauchy-Dirichlet problem (22). Then there exists a positive constant depending only on n, ν such that*

$$\int_{\mathcal{Q}_R \cap \Omega_T} \left(R^{-1} |u - w| + |Du - Dw| \right) \, \mathrm{d}x \, \mathrm{d}t \ \leq \ c \, R \int_{\mathcal{Q}_R \cap \Omega_T} |f| \, \mathrm{d}x \, \mathrm{d}t.$$

4 Proof of the main results

This section is devoted to the proof of the main result, to the regularity analysis of the borderline case, and to precise Lorentz space estimates up to the boundary for the solution u. First, however, a couple of remarks allowing to simplify the proofs.

Remark 5 Note that we can extend a solution to problem (1) to $\Omega_{2T} := \Omega \times (-2T, 0)$, where we also extend the function f to zero in $\Omega \times (-2T, -T]$ and the vector field a in the following way: $\tilde{a}(x, t, z) := a(x, -2T - t, z)$ for $t \in (-2T, -T)$. Therefore, denoting respectively by $\tilde{u}, \tilde{f}$ the extension, we have that $\tilde{u}$ satisfies

$$\begin{cases} \tilde{u}_t - \operatorname{div} \tilde{a}(x, t, D\tilde{u}) = \tilde{f}(x, t) & \text{in } \Omega_{2T}, \\ \tilde{u} = 0 & \text{on } \partial_{\mathrm{par}} \Omega_{2T}. \end{cases}$$

Indeed, we briefly sketch this argument. Given a test function $\varphi \in C_0^\infty(\Omega_{2T})$, and supposing $\operatorname{supp} \varphi \cap \Omega \times \{-T\} \neq \emptyset$ in order to avoid trivialities, it is clearly enough to prove that the distributional formulation (7) is satisfied in $C_0^\infty(\Omega_T)$ by the restriction of φ in Ω_T, i.e.

$$- \int_{\Omega_T} u \varphi_t \, \mathrm{d}x \, \mathrm{d}t + \int_{\Omega_T} \langle a(x, t, Du), D\varphi \rangle \, \mathrm{d}x \, \mathrm{d}t = \int_{\Omega_T} f \varphi \, \mathrm{d}x \, \mathrm{d}t, \qquad (39)$$

with φ as above. To this aim, take a $C^\infty(\mathbb{R})$ function ζ_ε such that $\zeta_\varepsilon \equiv 1$ on $[-T + 2\varepsilon, +\infty)$, $\zeta_\varepsilon \equiv 0$ on $(-\infty, -T + \varepsilon]$ and $|\zeta_\varepsilon'| \leq c/\varepsilon$, and put as a test function into (7) the product $\phi := \varphi\zeta_\varepsilon \in C_0^\infty(\Omega_T)$. We get

$$-\int_{\Omega_T} u\varphi_t\zeta_\varepsilon \, \mathrm{d}x \, \mathrm{d}t - \int_{\Omega_T} u\varphi\zeta_\varepsilon' \, \mathrm{d}x \, \mathrm{d}t + \int_{\Omega_T} \langle a(x, t, Du), D\varphi\rangle\zeta_\varepsilon \, \mathrm{d}x \, \mathrm{d}t$$
$$= \int_{\Omega_T} f\varphi\zeta_\varepsilon \, \mathrm{d}x \, \mathrm{d}t.$$

Now, we let ε go to zero. Notice that both the first and the third integral on the left-hand side converge to the corresponding integrals in (39); the same happens for the integral on the right-hand side. It remains to consider the contribution involving the derivative of ζ_ε; for this, we have

$$\left|\int_{\Omega_T} u\varphi\zeta_\varepsilon' \, \mathrm{d}x \, \mathrm{d}t\right| \leq c \, \|\varphi\|_{L^\infty(\Omega_T)} \fint_{-T+\varepsilon}^{-T+2\varepsilon} \int_\Omega |u(x, t)| \, \mathrm{d}x \, \mathrm{d}t \xrightarrow[\varepsilon \to 0]{} 0$$

thanks to (8). This completes the claim.

This will allow to treat, by a simple argument, only cylinders not intersecting the initial boundary. Indeed for these, it would be enough to extend the solution as described above, and treat them as they were interior cylinders, or eventually cylinders just intersecting the lateral boundary.

Remark 6 We stress here that we shall also not take into account the terminal boundary. Indeed the arguments in [5] could be also performed using not symmetric, but backward parabolic cylinders as those in (6). This would apply in particular to parabolic cylinders of the form $\mathcal{Q}_\rho(x_0, 0)$, with $B_\rho(x_0) \subset \Omega$.

4.1 Integrability of Du

We firstly need a preliminary lemma, whose proof can be deduced by extending the original proof in the local case in [5, Lemma 6.2] (see, also, [40, Propositions 3.2 and 3.4]), which will permit us to apply the Calderón-Zygmund type Lemma 6 to the level sets of maximal operator functionals applied to the gradient of the solutions to our problem (1). Since the novelty here is given by the up to the boundary estimates, we prefer to state this result in the following form, that is in accordance with the analog in the elliptic case given by Proposition 3.4 in [40].

Lemma 12 *Let u be the solution to* (10). *Let $A > 1$ be an absolute constant depending only on n, L, ν, c_0, ρ_0, and let $\chi_1 > 1$ be the higher integrability exponent as in* Theorem 6 *and* 7[3] . *Then for every $S > 1$ and any $\lambda > 0$ we*

---

[3] Actually, we take the minimum between the higher integrability exponent χ_1 arising in both the aforementioned theorems. For the sake of simplicity, we keep the same symbol.

have the following: if for some cylinder $\mathcal{Q}_\rho$, with $\rho < \min\{\rho_0, R_0\}/12$, ρ_0 as in Definition 2 *and R_0 the radius of a fixed ball B_{R_0}, it holds*

$$|\{(x,t) \in \mathcal{Q}_\rho : M^*(\mathbb{1}_{\Omega_T}|Du|)(x,t) > AS\lambda\}| \geq S^{-2\chi_1}|\mathcal{Q}_\rho|,$$

then there exists $\varepsilon = \varepsilon(S, \chi_1)$ such that

$$\mathcal{Q}_\rho \subset \{M^*(\mathbb{1}_{\Omega_T}|Du|) > \lambda\} \cup \{M_1^*(\mathbb{1}_{\Omega_T}f) > \varepsilon\lambda\},$$

where we denoted by $M^* := M_{0,\mathcal{Q}_{R_0}}^*$ and $M_1^* := M_{1,\mathcal{Q}_{R_0}}^*$.

Proof First, we notice that the statement can be restated as follows:

Assume that for some cylinder $\mathcal{Q}_\rho$ with $\rho < \min\{\rho_0, R_0\}/12$ we have

$$\mathcal{Q}_\rho \cap \{M^*(\mathbb{1}_{\Omega_T}(|Du|) \leq \lambda\} \cap \{M_1^*(\mathbb{1}_{\Omega_T}f) \leq \varepsilon\lambda\} \neq \emptyset, \tag{40}$$

then

$$|\{(x,t) \in \mathcal{Q}_\rho : M^*(\mathbb{1}_{\Omega_T}|Du|)(x,t) > AS\lambda\}| < S^{-2\chi_1}|\mathcal{Q}_\rho|. \tag{41}$$

By the hypothesis (40) there exists $(y_0, s_0) \in \mathcal{Q}_\rho \equiv \mathcal{Q}_\rho(y_1, s_1)$ such that, for any $0 < r \leq R_0$, with $(y_0, s_0) \in \mathcal{Q}_r$,

$$\fint_{\mathcal{Q}_r} \mathbb{1}_{\Omega_T}|Du|\, dy\, ds \leq \lambda \tag{42}$$

and

$$r\fint_{\mathcal{Q}_r} \mathbb{1}_{\Omega_T}|f|\, dy\, ds \leq \varepsilon\lambda, \tag{43}$$

reminding the definition of maximal operator.

We want to prove that, for any $(x,t) \in \mathcal{Q}_\rho(y_1, s_1)$, the following inequality holds

$$M^*(\mathbb{1}_{\Omega_T}|Du|)(x,t) \leq \max\left\{M^*(\mathbb{1}_{\mathcal{Q}_{2\rho}(y_1,s_1)\cap\Omega_T}|Du|)(x,t), 3^N\lambda\right\}. \tag{44}$$

For this, we consider a generic cylinder $\mathcal{Q}_r$ such that $(x,t) \in \mathcal{Q}_r \subseteq \mathcal{Q}_{R_0}$; for any $r \leq \rho$, $\mathcal{Q}_r \subset \mathcal{Q}_{2\rho}(y_1, s_1)$, and so we have

$$\fint_{\mathcal{Q}_r} \mathbb{1}_{\Omega_T}|Du|\, dy\, ds \leq M^*(\mathbb{1}_{\mathcal{Q}_{2\rho}(y_1,s_1)\cap\Omega_T}|Du|)(x,t).$$

On the other hand, for any $\rho < r \leq R_0$ we plainly have, enlarging, if necessary, the cylinder $\mathcal{Q}_r$, and using (42)

$$\fint_{\mathcal{Q}_r} \mathbb{1}_{\Omega_T}|Du|\, dy\, ds \leq 3^N\lambda.$$

Thus, the inequality in (44) is proven.

Now, let $\mathcal{Q}_{4\rho} \equiv \mathcal{Q}_{4\rho}(y_1, s_1) \supset \mathcal{Q}_\rho$ and distinguish the following three cases: $\mathcal{Q}_{4\rho} \subset \mathbb{R}^{n+1} \setminus \Omega_T$, $\mathcal{Q}_{4\rho} \cap \partial_{\mathrm{par}}\Omega_T \neq \emptyset$ and $\mathcal{Q}_{4\rho} \subset \Omega_T$.

In the first case, the inequality in (41) does hold provided that $A \geq 3^N$, by taking into account (44).

In order to deal with the second case, that is where the cylinder $\mathcal{Q}_{4\rho}$ intersects $\partial_{\mathrm{par}}\Omega_T$, we extend to the parabolic setting the strategy in Proposition 3.2 in [40]. Precisely, take any $(y_2, s_2) \in \partial_{\mathrm{par}}\Omega_T$ such that $\mathrm{dist}_{\mathrm{par}}\big((y_1, s_1), \partial_{\mathrm{par}}\Omega_T\big) := |y_1 - y_2| + |s_1 - s_2|^{\frac{1}{2}}$ and consider the unique solution w to the following homogeneous Cauchy-Dirichlet problem

$$\begin{cases} w_t - \mathrm{div}\, a(x, t, Dw) = 0 & \text{in } \mathcal{Q}_{12\rho}(y_2, s_2) \cap \Omega_T \\ w = u & \text{on } \partial_{\mathrm{par}}(\mathcal{Q}_{12\rho}(y_2, s_2) \cap \Omega_T). \end{cases}$$

Also, we extend by zero the function u to $\mathbb{R}^{n+1} \setminus \Omega_T$ and by u the function w to $\mathbb{R}^{n+1} \setminus (\mathcal{Q}_{12\rho}(y_2, s_2) \cap \Omega_T)$.

Now, we are in a position to estimate the measure of the investigated level sets by using (14) twice with $\gamma = 2\chi_1$ and $\gamma = 1$. It follows

$$\big|\big\{(x, t) \in \mathcal{Q}_\rho : M^*(\mathbb{1}_{\mathcal{Q}_{2\rho}(y_1, s_1) \cap \Omega_T} |Du|)(x, t) > AS\lambda\big\}\big|$$

$$\leq \big|\big\{(x, t) \in \mathcal{Q}_\rho : M^*(\mathbb{1}_{\mathcal{Q}_{2\rho}(y_1, s_1) \cap \Omega_T} |Dw|)(x, t) > AS\lambda/2\big\}\big|$$

$$+ \big|\big\{(x, t) \in \mathcal{Q}_\rho : M^*(\mathbb{1}_{\mathcal{Q}_{2\rho}(y_1, s_1) \cap \Omega_T} |Du - Dw|)(x, t) > AS\lambda/2\big\}\big|$$

$$\leq c(AS\lambda)^{-2\chi_1} \int_{\mathcal{Q}_{2\rho}(y_1, s_1) \cap \Omega_T} |Dw|^{2\chi_1} \, \mathrm{d}y \, \mathrm{d}s \qquad (45)$$

$$+ c(AS\lambda)^{-1} \int_{\mathcal{Q}_{2\rho}(y_1, s_1) \cap \Omega_T} |Du - Dw| \, \mathrm{d}y \, \mathrm{d}s.$$

By Theorem 7, applied with $R = 12\rho < \rho_0$, the first term in the right-hand side of (45) is bounded from above by

$$c(AS\lambda)^{-2\chi_1} |\mathcal{Q}_{6\rho}| \left(\frac{1}{|\mathcal{Q}_{12\rho}|} \int_{\mathcal{Q}_{12\rho}(y_2, s_2) \cap \Omega_T} |Dw| \, \mathrm{d}y \, \mathrm{d}s \right)^{2\chi_1},$$

since $\mathcal{Q}_{2\rho}(y_1, s_1) \subset \mathcal{Q}_{6\rho}(y_2, s_2)$. Thus we get

$$\big|\big\{(x, t) \in \mathcal{Q}_\rho : M^*(\mathbb{1}_{\mathcal{Q}_{2\rho}(y_1, s_1) \cap \Omega_T} |Du|)(x, t) > AS\lambda\big\}\big|$$

$$\leq c(AS\lambda)^{-2\chi_1} |\mathcal{Q}_{6\rho}| \left(\frac{1}{|\mathcal{Q}_{12\rho}|} \int_{\mathcal{Q}_{12\rho}(y_2, s_2) \cap \Omega_T} |Du| \, \mathrm{d}y \, \mathrm{d}s \right)^{2\chi_1}$$

$$+ c(AS\lambda)^{-2\chi_1} |\mathcal{Q}_{6\rho}| \left(\frac{1}{|\mathcal{Q}_{12\rho}|} \int_{\mathcal{Q}_{12\rho}(y_2, s_2) \cap \Omega_T} |Du - Dw| \, \mathrm{d}y \, \mathrm{d}s \right)^{2\chi_1}$$

$$+ c(AS\lambda)^{-1} \frac{|\mathcal{Q}_{6\rho}|}{|\mathcal{Q}_{12\rho}|} \int_{\mathcal{Q}_{12\rho}(y_2, s_2) \cap \Omega_T} |Du - Dw| \, \mathrm{d}y \, \mathrm{d}s.$$

In view of the comparison estimates established in Lemma 11 (taking $R = 12\rho < \rho_0$ there), together with (42) and (43), since $(y_0, s_0) \in \mathcal{Q}_\rho(y_1, s_1) \subset \mathcal{Q}_{12\rho}(y_2, s_2) \subseteq \mathcal{Q}_{R_0}$, and choosing $\varepsilon = S^{-2\chi_1+1}$, we arrive at

$$\left|\left\{(x,t) \in \mathcal{Q}_\rho : M^*(\mathbb{1}_{\mathcal{Q}_{2\rho}(y_1,s_1) \cap \Omega_T}|Du|) > AS\lambda\right\}\right| \leq c(A^{-2\chi_1} + A^{-1})S^{-2\chi_1}|\mathcal{Q}_\rho|.$$

Now, it suffices to choose the constant A so that $A \geq 3^N$ and $2cA^{-1} \leq 1/2$; i.e., $A \geq \max\{3^N, 4c\}$. It follows

$$\left|\left\{(x,t) \in \mathcal{Q}_\rho : M^*(\mathbb{1}_{\mathcal{Q}_{2\rho}(y_1,s_1) \cap \Omega_T}|Du|)(x,t) > AS\lambda\right\}\right| \leq \frac{1}{2}S^{-2\chi_1}|\mathcal{Q}_\rho|,$$

which, again in view of (44), yields (41).

Finally, when $\mathcal{Q}_{4\rho} \subset \Omega_T$, the inequality in (41) plainly follows by Lemma 6.2 in [5], hence using the local interior results given by Theorem 6 and Lemma 7. $\qquad\square$

Before starting with the proof of Theorem 1, we want to emphasize that most of the differences between the problem we are dealing with and the analog in the elliptic case analyzed in [40] have arisen in the previous section. For this, the general strategy of the proof follows that of the proof of Theorem 6.1 in [5] (see, also, [14, Theorem 4.1]. We prefer to give some details for the reader's convenience, and we sketch the proof in three steps.

Proof of Theorem 1 We denote by $\mathcal{Q}_0 \equiv B_{R_0}(x_0) \times (t_0 - R_0^2, t_0)$ a fixed cylinder in $\mathbb{R}^{n+1}$ that contains Ω_T, with its basis $B_0 \equiv B_{R_0}$ being a ball of radius $R_0 \leq 2\mathrm{diam}(\Omega)$. Note that since $\mathcal{Q}_0 \supset \Omega_T$ we can assume $R_0 \geq \mathrm{diam}(\Omega)/2$. Now we consider the following maximal operators

$$M^*(\cdot) := M^*_{0,\mathcal{Q}_0}(\mathbb{1}_{\Omega_T} \cdot) \quad \text{and} \quad M^*_1(\cdot) := M^*_{1,\mathcal{Q}_0}(\mathbb{1}_{\Omega_T} \cdot).$$

Step 1 - Application of Calderón-Zygmund-Krylov-Safonov covering type lemma. We consider the following sets $\mathcal{X}$ and $\mathcal{Y}$ defined

$$\mathcal{X} := \left\{(x,t) \in \mathcal{Q}_0 : M^*(|Du|)(x,t) > (AS)^{k+1}\lambda_1\right\},$$

$$\mathcal{Y} := \Big\{(x,t) \in \mathcal{Q}_0 : M^*(|Du|)(x,t) > (AS)^k\lambda_1$$
$$\text{and } M^*_1(f)(x,t) > \varepsilon(AS)^k\lambda_1\Big\},$$

with

$$\lambda_1 := \bar{c}\, S^{2\chi_1} \frac{1}{|\mathcal{Q}_{\bar{R}}|} \int_{\mathcal{Q}_0 \cap \Omega_T} |Du| \, \mathrm{d}x\, \mathrm{d}t, \tag{46}$$

where $S > 1$, A, χ_1, ε are as in Lemma 12. We want to apply Lemma 6 with $\bar{R} = \min\{\rho_0, R_0\}/12$. Firstly, we notice that by Lemma 12 (choosing $\lambda = (AS)^k\lambda_1 > 0$ there) the assumption (ii) is satisfied with $\delta = S^{-2\chi_1}$. Thus, it remains to prove that $|\mathcal{X}| < S^{-2\chi_1}|\mathcal{Q}_{\bar{R}}|$. For this, it suffices to use the

boundedness of the maximal operators in Marcinkiewicz spaces together with the choice of λ_1 in (46). Indeed,

$$\left|\left\{(x,t) \in \mathcal{Q}_0 : M^*(|Du|)(x,t) > (AS)^{k+1}\lambda_1\right\}\right|$$

$$< \frac{\bar{c}}{(AS)^{k+1}\lambda_1} \int_{\mathcal{Q}_0 \cap \Omega_T} |Du| \, dx \, dt$$

$$< S^{-2\chi_1} |\mathcal{Q}_{\bar{R}}|,$$

where we also used the fact that $A, S > 1$. Hence, the application of Lemma 6 and the definitions of $\mathcal{X}$ and $\mathcal{Y}$ yield

$$\left|\left\{(x,t) \in \mathcal{Q}_0 : M^*(|Du|)(x,t) > (AS)^{k+1}\lambda_1\right\}\right|$$

$$\leq c_3 S^{-2\chi} \left|\left\{(x,t) \in \mathcal{Q}_0 : M^*(|Du|)(x,t) > (AS)^k\lambda_1\right\}\right|$$

$$+ c_3 \left|\left\{(x,t) \in \mathcal{Q}_0 : M_1^*(f)(x,t) > \varepsilon(AS)^k\lambda_1\right\}\right|,$$

for every $k \in \mathbb{N}$; with c_3 depending only on n, ρ_0, and $\text{diam}(\Omega)$.

Now, by taking into account a multiplication by a factor $(AS)^{\frac{(k+1)N\gamma}{N-\gamma}}$, from the previous estimate we deduce

$$(AS)^{\frac{(k+1)N\gamma}{N-\gamma}} \lambda_1^{\frac{N\gamma}{N-\gamma}} \mu_1((AS)^{k+1}\lambda_1)$$

$$\leq c_3(AS)^{\frac{N\gamma k}{N-\gamma}} A^{\frac{N\gamma}{N-\gamma}} S^{\frac{N\gamma}{N-\gamma}-2\chi} \lambda_1^{\frac{N\gamma}{N-\gamma}} \mu_1((AS)^k\lambda_1) \qquad (47)$$

$$+ c_3(AS)^{\frac{N\gamma k}{N-\gamma}} \left(\frac{AS}{\varepsilon}\right)^{\frac{N\gamma}{N-\gamma}} (\lambda_1\varepsilon)^{\frac{N\gamma}{N-\gamma}} \mu_2(\varepsilon(AS)^k\lambda_1),$$

where, for any $K \geq 0$, we denoted by

$$\mu_1(K) := \left|\left\{(x,t) \in \mathcal{Q}_0 : M^*(|Du|)(x,t) > K\right\}\right| \qquad (48)$$

and

$$\mu_2(K) := \left|\left\{(x,t) \in \mathcal{Q}_0 : M_1^*(f)(x,t) > K\right\}\right|. \qquad (49)$$

Now, we note that, since $\chi_1 > 1$ and $\gamma \leq 2N/(N+2)$, we have that the quantity

$$d := 2\chi_1 - \frac{N\gamma}{N-\gamma} \geq 2(\chi_1 - 1) > 0$$

is positive. Therefore, we can choose

$$S := \left(4A^{\frac{N\gamma}{N-\gamma}}\right)^{\frac{1}{d}}$$

and, by computations, inequality (47) provides the existence of a constant $c \equiv c(n, L, \nu, \text{diam}(\Omega), \rho_0)$ such that, for every $k \geq 0$,

$$(AS)^{\frac{N\gamma(k+1)}{N-\gamma}} \lambda_1^{\frac{N\gamma}{N-\gamma}} \mu_1((AS)^{k+1}\lambda_1) \qquad (50)$$

$$\leq \frac{1}{4}(AS)^{\frac{N\gamma k}{N-\gamma}} \lambda_1^{\frac{N\gamma}{N-\gamma}} \mu_1((AS)^k\lambda_1) + c(AS)^{\frac{N\gamma k}{N-\gamma}} (\varepsilon\lambda_1)^{\frac{N\gamma}{N-\gamma}} \mu_2((AS)^k\varepsilon\lambda_1).$$

Step 2 - Level sets estimates. In order to establish some Lorentz spaces estimates on level sets we proceed as follows. We take $0 < \beta < \infty$ and operate some computations to get

$$\int_0^\infty [\lambda^{\frac{N\gamma}{N-\gamma}} \mu_1(\lambda)]^{\frac{\beta(N-\gamma)}{N\gamma}} \frac{\mathrm{d}\lambda}{\lambda}$$
$$\leq \left(\frac{1}{\beta} + \tilde{c}^\beta (AS)^\beta \log(AS)\right) \lambda_1^\beta |\mathcal{Q}_0|^{\frac{\beta(N-\gamma)}{N\gamma}} + \tilde{c}^\beta (AS)^\beta \log(AS) J(\infty),$$

where

$$J(\infty) := \sum_{k=0}^\infty \left((AS)^{\frac{N\gamma k}{N-\gamma}} (\varepsilon \lambda_1)^{\frac{N\gamma}{N-\gamma}} \mu_2((AS)^k \varepsilon \lambda_1)\right)^{\frac{\beta(N-\gamma)}{N\gamma}}$$

and the constant $\tilde{c} > 1$ is increasing in the variables n, L, ν and decreasing in β, such that $\tilde{c} \to \infty$ as $\beta \to 0$, while it remains bounded when β is bounded away from zero. Thus, from this inequality, we can plainly deduce

$$\int_0^\infty [\lambda^{\frac{N\gamma}{N-\gamma}} \mu_1(\lambda)]^{\frac{\beta(N-\gamma)}{N\gamma}} \frac{\mathrm{d}\lambda}{\lambda} \leq \left(\frac{1}{\beta} + 2\tilde{c}^\beta (AS)^\beta \log(AS)\right) \lambda_1^\beta |\mathcal{Q}_0|^{\frac{\beta(N-\gamma)}{N\gamma}}$$
$$+ \tilde{c}^\beta (AS)^{2\beta} \int_0^\infty [\lambda^{\frac{N\gamma}{N-\gamma}} \mu_2(\lambda)]^{\frac{\beta(N-\gamma)}{N\gamma}} \frac{\mathrm{d}\lambda}{\lambda}. \quad (51)$$

We recall the definitions of μ_1 and μ_2 given in (48)-(49), and, now, we choose[4] $\beta = q \in (0, \infty)$, so that, by the definition of parabolic maximal operators, we get

$$\|M^*(|Du|)\|_{L\left(\frac{N\gamma}{N-\gamma}, q\right)(\mathcal{Q}_0)} \leq \tilde{c} \lambda_1 |\mathcal{Q}_0|^{\frac{N-\gamma}{N\gamma}} + \tilde{c} \|M_1^*(f)\|_{L\left(\frac{N\gamma}{N-\gamma}, q\right)(\mathcal{Q}_0)}, \quad (52)$$

up to relabeling the constant $\tilde{c}$, by keeping the same properties as before.

Now, thanks to a classical global estimate for the gradient of u established in [9], which we will use as in the form presented in [4, Lemma 5.1] where the exact dependence on the L^1-norm appears, we can estimate λ_1 from above, as follows

$$\lambda_1 \leq c(n, \rho_0, \mathrm{diam}(\Omega)) |\mathcal{Q}_0|^{\frac{1}{N}-1} \int_{\mathcal{Q}_0 \cap \Omega_T} |f| \, \mathrm{d}x \, \mathrm{d}t.$$

Therefore, for any $(x,t) \in \mathcal{Q}_0$ we have $M_1^*(f)(x,t) \geq \lambda_1 / c(n, \rho_0, \mathrm{diam}(\Omega))$, and thus

$$\lambda_1 |\mathcal{Q}_0|^{\frac{N-\gamma}{N\gamma}} \leq \|M_1^*(f)\|_{L\left(\frac{N\gamma}{N-\gamma}\right)(\mathcal{Q}_0)}.$$

Finally, in view of the standard inequalities $|Du(x,t)| \leq M^*(|Du|)(x,t)$, for a.e. $(x,t) \in \mathcal{Q}_0$, and $\|M_1^*(f)\|_{L\left(\frac{N\gamma}{N-\gamma}, q\right)(\mathcal{Q}_0)} \leq \|f\|_{L(\gamma, q)(\mathcal{Q}_0)}$, we arrive at

$$\|Du\|_{L\left(\frac{N\gamma}{N-\gamma}, q\right)(\Omega_T)} \leq c \|f\|_{L(\gamma, q)(\mathcal{Q}_0)}, \quad (53)$$

---

[4] We note that in the computations above we preferred to keep the generic exponent β, since it could be chosen arbitrarily in order to obtain estimates also in other rearrangement and non rearrangement spaces.

where the constant c depends only on $n, L, \nu, \gamma, q, \rho_0$ and $\mathrm{diam}(\Omega)$. Recalling that $\Omega_T \subset \mathcal{Q}_0$ and $f = 0$ in $\mathbb{R}^{n+1} \setminus \Omega_T$ we plainly deduce the desired estimate in (4).

Similarly, we can deal with the Marcinkiewicz case $q = \infty$ (see, e.g., Step 3 in the proof of [5, Theorem 3]) and we arrive at

$$\|Du\|_{\mathcal{M}^{\frac{N\gamma}{N-\gamma}}(\Omega_T)} \leq c\,\|f\|_{\mathcal{M}^\gamma(\Omega_T)}.$$

Step 3 - Conclusion of the proof. We recall that we proved the estimate in (53) for the approximating solutions $u \equiv u_k$ to problem (10) with $f \equiv f_k = T_k(f)$, where T_k is the truncation operator defined by (11). In order to conclude, it suffices to use the lower semicontinuity of the Lorentz norms together with the standard approximating arguments stated in Section 2.1. Hence, we have proven that

$$\|Du_k\|_{L\left(\frac{N\gamma}{N-\gamma}, q\right)(\Omega_T)} \leq c\,\|f_k\|_{L(\gamma, q)(\mathcal{Q}_0)}, \tag{54}$$

where the constant c does not depend on k. We also notice that in (54) we have used the fact that

$$\|f_k\|_{L(\gamma, q)(\mathcal{Q}_0)} \leq \|f\|_{L(\gamma, q)(\mathcal{Q}_0)},$$

since by the definition of f_k it holds that $|f_k| \leq |f|$. Thus, in order to pass to the limit on $k \to \infty$ in (54), it suffices to use the lower semicontinuity of the Lorentz norms (recall remark on page 9) together with the approximating arguments stated in Section 2.1 (recall, in particular, (11) there). $\qquad\square$

Remark 7 As in the classic case, Theorem 1 fails for the borderline choice $\gamma = 1$. Even in the elliptic case, one has to impose some further $L \log L$ integrability on the datum f in order to obtain the following implication

$$f \in L^1(-T, 0; L \log L(\Omega)) \implies |Du| \in L^{\frac{N}{N-1}}(\Omega_T); \tag{55}$$

see Theorem 1.8 in [8]. However, we could extend the potential approach in the proof of Theorem 1 to the analysis of the borderline case in the parabolic Orlicz space $L \log L(\Omega_T)$, to get, for any cylinder $\mathcal{Q}_\rho \subset \mathbb{R}^{n+1}$, the following estimate

$$\|Du\|_{L^{\frac{N}{N-1}}(\mathcal{Q}_{\rho/2} \cap \Omega_T)} \leq c\,\|f\|_{L^1(\mathcal{Q}_\rho \cap \Omega_T)}^{\frac{1}{N}}\|f\|_{L \log L(\mathcal{Q}_\rho \cap \Omega_T)}^{\frac{N-1}{N}}. \tag{56}$$

We stress that when considering the preceding estimate in the whole Ω_T, the by-now classic result (55) in [8] is stronger, since it does not require any regularity assumptions on the domain Ω and it does hold in the weaker assumption of f being in $L^1(-T, 0; L \log L(\Omega))$, despite (56) provides local estimates up to the boundary.

4.2 Integrability of u

This section is devoted to Lorentz space estimates for the solution u to (1). We will use the same techniques used to prove the spatial regularity of the gradient. Hence, we will obtain an estimate on the level sets of the maximal operator associated to u, in terms of the level sets of a maximal operator of the datum f, up to a correction term; that is, the equivalent to (52) for u. This will allow to deduce the proof of Theorem 2, stated in the introduction, which is close to the one of Theorem 1, but also a higher regularity result is needed, as stated in the second part of Theorem 7 proven at the end of Section 3.

Keeping in mind the notation used in Theorem 1, we have the analog of Lemma 12 for u. Again, we recall that in what is following we always deal with the approximating solutions $u \equiv u_k$ defined in Section 2.1, and with the datum $f \equiv f_k$.

Lemma 13 *Let $u \in C^0([-T, 0]; L^2(\Omega)) \cap L^2(-T, 0; W_0^{1,2}(\Omega))$ be the solution to (10). Let $A > 1$ be an absolute constant depending only on n, L, ν, and let $\chi > 1$ be the higher integrability exponent as in Theorems 6 and 7. Then for every $S > 1$ and any $\lambda > 0$ we have that exists $\varepsilon = \varepsilon(S, \chi)$ such that the following statement holds: If for some cylinder $\mathcal{Q}_\rho$ with $\rho < \min\{\rho_0, R_0\}/12$ it holds*

$$|\{(x,t) \in \mathcal{Q}_\rho : M^*(u) > AS\lambda\}| \geq S^{-2\chi}|\mathcal{Q}_\rho|$$

then

$$\mathcal{Q}_\rho \subset \{M^*(u) > \lambda\} \cup \{M_2^*(f) > \varepsilon\lambda\},$$

where we denoted by $M^*(\cdot) := M_{0,\mathcal{Q}_{R_0}}^*(\mathbb{1}_{\Omega_T}\cdot)$ and $M_2^*(\cdot) := M_{2,\mathcal{Q}_{R_0}}^*(\mathbb{1}_{\Omega_T}\cdot)$.

Proof The proof of this lemma is essentially the same as that of Lemma 12. It first suffices to substitute $M^*(|Du|)$ and $M_1^*(f)$ by $M^*(u)$ and $M_2^*(f)$, respectively. Then, we need to use the comparison estimates for u stated in Lemma 7 and Lemma 11, and the higher integrability, again for u, given in (21) and (24). This will permit to introduce the parameter $\chi > 1$ which can be chosen arbitrarily large; and this is another difference with respect to the estimates in Lemma 12, in which such a parameter (read $\chi_1 > 1$) was fixed. $\square$

Proof of Theorem 2. Since we closely follow the proof of Theorem 1, we prefer to keep the division of the proof in the same three steps.

Step 1 - Application of Calderón-Zygmund-Krylov-Safonov covering type lemma. We want to apply Lemma 6 to the sets

$$\mathcal{X} := \left\{(x,t) \in \mathcal{Q}_0 : M^*(u)(x,t) > (AS)^{k+1}\lambda_0\right\},$$

$$\mathcal{Y} := \left\{(x,t) \in \mathcal{Q}_0 : M^*(u)(x,t) > (AS)^k\lambda_0\right.$$

$$\left. \text{and } M_2^*(f)(x,t) > \varepsilon(AS)^k\lambda_0\right\},$$

$$\text{with } \lambda_0 := \bar{c}\, S^{2\chi} \frac{1}{|\mathcal{Q}_{\bar{R}}|} \int_{\mathcal{Q}_0 \cap \Omega_T} |u|\, \mathrm{d}x\, \mathrm{d}t \quad \text{and} \quad \bar{R} = \frac{\min\{\rho_0, R_0\}}{12}.$$

In view of Lemma 13, the assumption (ii) of Lemma 6 (choosing $\lambda = (AS)^k \lambda_0$ there) is satisfied with $\rho < \bar{R} = \min\{\rho_0, R_0\}/12$ and $\delta = S^{-2\chi}$. To prove (i), that is $|\mathcal{X}| < S^{-2\chi}|\mathcal{Q}_0|$, it suffices to use the boundedness of the maximal operators in Marcinkiewicz spaces together with the choice of λ_0 and the fact that $A, S > 1$. It follows

$$(AS)^{\frac{N\gamma(k+1)}{N-2\gamma}} \lambda_0^{\frac{N\gamma}{N-2\gamma}} \mu_1((AS)^{k+1}\lambda_0)$$

$$\leq (AS)^{\frac{N\gamma k}{N-2\gamma}} A^{\frac{N\gamma}{N-2\gamma}} S^{\frac{N\gamma}{N-2\gamma}-2\chi_0} \lambda_0^{\frac{N\gamma}{N-2\gamma}} \mu_1((AS)^k\lambda_0) \qquad (57)$$

$$+ (AS/\varepsilon)^{\frac{N\gamma}{N-2\gamma}} (AS)^{\frac{N\gamma k}{N-\gamma}} (\varepsilon\lambda_0)^{\frac{N\gamma}{N-\gamma}} \mu_2((AS)^k \varepsilon\lambda_0), \quad \text{for every } k \geq 0,$$

where, for any $K > 0$, we denoted by

$$\mu_1(K) := \left|\left\{(x,t) \in \mathcal{Q}_0 : M^*(u)(x,t) > K\right\}\right|$$

and

$$\mu_2(K) := \left|\left\{(x,t) \in \mathcal{Q}_0 : M_2^*(f)(x,t) > K\right\}\right|.$$

At this level, we can take advantage of the possibility to choose χ large enough to satisfy

$$d := 2\chi - \frac{N\gamma}{N-2\gamma} > 0,$$

and, using the $S := \left(4A^{\frac{N\gamma}{N-2\gamma}}\right)^{\frac{1}{d}}$, it follows $A^{\frac{N\gamma}{N-2\gamma}} S^{\frac{N\gamma}{N-2\gamma}-2\chi} \leq 1/4$, so that the estimate in (57) becomes

$$(AS)^{\frac{N\gamma(k+1)}{N-2\gamma}} \lambda_0^{\frac{N\gamma}{N-2\gamma}} \mu_1((AS)^{k+1}\lambda_0) \leq \frac{1}{4}(AS)^{\frac{N\gamma k}{N-2\gamma}} \lambda_0^{\frac{N\gamma}{N-2\gamma}} \mu_1((AS)^k\lambda_0)$$

$$+ c(AS)^{\frac{N\gamma k}{N-\gamma}} (\varepsilon\lambda_0)^{\frac{N\gamma}{N-\gamma}} \mu_2((AS)^k \varepsilon\lambda_0),$$

for every $k \geq 0$.

Step 2 - Level sets estimates. Arguing as in Step 2 in the proof of Theorem 1 we obtain the following Lorentz estimates for the solution u,

$$\|u\|_{L\left(\frac{N\gamma}{N-2\gamma}, q\right)(\Omega_T)} \leq c\, \|f\|_{L(\gamma, q)(\mathcal{Q}_0)},$$

for any $0 < q < \infty$. Similarly, in the case of Marcinkiewicz space, that is $q = \infty$, we arrive at

$$\|u\|_{\mathcal{M}^{\frac{N\gamma}{N-2\gamma}}(\Omega_T)} \leq c\, \|f\|_{\mathcal{M}^\gamma(\mathcal{Q}_0)},$$

Notice that at this stage we used the fact that M_2^* has a higher regularizing effect with respect to that of M_1^*. Moreover since we are assuming $f = 0$ in $\mathbb{R}^{n+1} \setminus \Omega_T$ and $\Omega_T \subset \mathcal{Q}_0$ we obtain exactly the desired estimate.

Step 3 - Conclusion of the proof. We can now conclude the proof using standard approximation methods, again by means of the lower semicontinuity of the Lorentz norms.   □

References

1. E. Acerbi, G. Mingione: Gradient estimates for a class of parabolic systems. *Duke Math. J.* **136** (2007), no. 2, 285–320.
2. D. R. Adams: A note on Riesz potentials. *Duke Math. J.* **42** (1975), 765–778.
3. A. A. Arkhipova: L_p-estimates of the gradients of solutions of initial/boundary-value problems for quasilinear parabolic systems. Differential and pseudodifferential operators. *J. Math. Sci.* **73 (6)** (1995), 609–617.
4. P. Baroni, J. Habermann: Calderón-Zygmund estimates for parabolic measure data equations. *J. Differential Equations* **252** (2012), 412–447.
5. P. Baroni, J. Habermann: New gradient estimates for parabolic equations. *Houston J. Math.* **38** (2012), 855–914.
6. P. Baroni: Marcinkiewicz estimates for degenerate parabolic equations with measure data. In preparation.
7. D. Blanchard, F. Murat: Renormalised solutions of nonlinear parabolic problems with L^1 data: existence and uniqueness. *Proc. Roy. Soc. Edinburgh Sect. A* **127** (1997), no. 6, 1137–1152.
8. L. Boccardo, A. Dall'Aglio, T. Gallouët, L. Orsina: Nonlinear parabolic equations with measure data. *J. Funct. Anal.* **147** (1997), 237–258.
9. L. Boccardo, T. Gallouët: Nonlinear elliptic and parabolic equations involving measure data. *J. Funct. Anal.* **87** (1989), 149–169.
10. V. Bögelein, F. Duzaar, G. Mingione: The boundary regularity of non-linear parabolic systems II. *Ann. Inst. Henri Poincaré Anal. Non Linéaire* **27** (2010), 145–200.
11. L. Caffarelli, I. Peral: On $W^{1,p}$ estimates for elliptic equations in divergence form. *Comm. Pure Appl. Math.* **51** (1989), 1–21.
12. A. Dall'Aglio: Approximated solutions of equations with L^1 data. Application to the H-convergence of quasi-linear parabolic equations. *Ann. Mat. Pura Appl. (IV)*, **170** (1996), 207–240.
13. A. Di Castro, G. Palatucci: Measure data problems, lower order terms and interpolation effects. *Ann. Mat. Pura Appl.*, DOI: 10.1007/s10231-012-0277-7
14. A. Di Castro, G. Palatucci: Nonlinear parabolic problems with lower order terms and related integral estimates. *Nonlinear Anal.*, **75** (2012), 4177–4197.
15. A. Di Castro, G. Palatucci: Fractional regularity for nonlinear elliptic problems with measure data. *J. Convex Anal.*, **20** (2013), no. 3.
16. J. Droniou, A. Porretta, A. Prignet: Parabolic capacity and soft measures for nonlinear equations. *Potential Anal.* **19** (2003), no. 2, 99–161.
17. F. Duzaar, G. Mingione: Gradient estimates via non-linear potentials. *Amer. J. Math.*, **133** (2011) 1093–1149.
18. M. Giaquinta, M. Struwe: On the partial regularity of weak solutions of nonlinear parabolic systems. *Math. Z.* **179** (1982), 437–451.
19. E. Giusti: *Direct Methods in the Calculus of Variations.* World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
20. L. Grafakos: *Classical and Modern Fourier Analysis.* Pearson Edu. Inc., Upper Saddle River, 2004.

21. J. Heinonen, T. Kilpeläinen, O. Martio: *Nonlinear potential theory of degenerate elliptic equations*. Oxford University Press, Oxford, 1993.
22. F. Hélein, J. Wood: Harmonic maps, in: *Handbook of Global Analysis*, vol. **1213**, Elsevier Sci. B.V., Amsterdam, 2008, 417–491.
23. T. Kilpeläinen, P. Koskela: Global integrability of the gradients of solutions to partial differential equations. *Nonlinear Anal.* **23** (1994), no. 7, 899–909.
24. J. Kinnunen, J.L. Lewis: Higher integrability for parabolic systems of p-laplacian type. *Duke Math. J.* **102** (2000), 253–271.
25. T. Kuusi, G. Mingione: Nonlinear potential estimates in parabolic problems. *Rend. Lincei, Mat. e Appl.* **22** (2011), 161–174.
26. T. Kuusi, G. Mingione: Gradient regularity for nonlinear parabolic equations. *Ann. Scu. Norm. Sup. Cl. Sci. (5)*, to appear.
27. T. Kuusi, G. Mingione: The Wolff gradient bound for degenerate parabolic equations. *J. Eur. Math. Soc. (JEMS)*, to appear.
28. J. L. Lewis: Uniformly fat sets. *Trans. Amer. Math. Soc.* **308** (1988), 177–196.
29. G. M. Lieberman: *Second order parabolic differential equations*. World Scientific Press, River Edge, 1996.
30. P. Mikkonen: On the Wolff potential and quasilinear elliptic equations involving measures. *Ann. Acad. Sci. Fenn., Ser AI, Math. Dissert.* **104** (1996), 1–71.
31. G. Mingione: The Calderón-Zygmund theory for elliptic problems with measure data. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **6** (2007), 195–261.
32. G. Mingione: Gradient estimates below the duality exponent. *Math. Ann.* **346** (2010), 571–627.
33. G. Mingione: Gradient potential estimates. *J. Eur. Math. Soc. (JEMS)* **13** (2011), 459–486.
34. G. Mingione: Nonlinear measure data problems. *Milan J. Math.* **79** (2011), 429–496.
35. G. Mingione (ed.): *Topics in modern regularity theory*. CRM Series **13**, Edizioni della Normale, Pisa, 2012.
36. R. Nikkta: Regularity of solutions of linear second order elliptic and parabolic boundary value problems on Lipschitz domains. *J. Differential Equations* **251** (2011), 860–880.
37. M. Parviainen: Global gradient estimates for degenerate parabolic equations in nonsmooth domains. *Ann. Mat. Pura Appl.* (4) **188** (2009), 333–358.
38. F. Petitta: Renormalized solutions of nonlinear parabolic equations with general measure data. *Ann. Mat. Pura Appl.* (4) **187** (2008), 563–604.
39. F. Petitta, A. Ponce, A. Porretta: Diffuse measures and nonlinear parabolic equations, *J. Evol. Equations* **11** (2011), no. 4, 861–905.
40. N. C. Phuc: Global integral gradient bounds for quasilinear equations below or near the natural exponent. *Ark. Mat.*, to appear.
41. T. Rivière: Conservation laws for conformally invariant variational problems. *Invent. Math.* **168** (2007), 1–22.