Communications in Contemporary Mathematics

PARABOLIC SYSTEMS WITH MEASURABLE COEFFICIENTS IN WEIGHTED ORLICZ SPACES

SUN-SIG BYUN

Department of Mathematical Sciences and
Research Institute of Mathematics, Seoul National University,
Seoul 151-747, Korea
byun@snu.ac.kr

JIHOON OK

Department of Mathematical Sciences, Seoul National University,
Seoul 151-747, Korea
khjojh1@snu.ac.kr

DIAN K. PALAGACHEV*

Dipartimento di Meccanica, Matematica e Management, Politecnico di Bari,
Via E. Orabona 4, 70125 Bari, Italy
dian.palagachev@poliba.it

LUBOMIRA G. SOFTOVA

Department of Civil Engineering, Design, Construction Industry and
Environment, Second University of Naples,
Via Roma 29, 81031 Aversa, Italy
luba.softova@unina2.it

We consider a parabolic system in divergence form with measurable coefficients in a cylindrical space-time domain with nonsmooth base. The associated nonhomogeneous term is assumed to belong to a suitable weighted Orlicz space. Under possibly optimal assumptions on the coefficients and minimal geometric requirements on the boundary of the underlying domain, we generalize the Calderón–Zygmund theorem for such systems by essentially proving that the spatial gradient of the weak solution gains the same weighted Orlicz integrability as the nonhomogeneous term.

Keywords: Parabolic system; Measurable coefficients; Calderón-Zygmund theory; Muckenhoupt weight; Orlicz space; Reifenberg domain; BMO.

Mathematics Subject Classification 2010: 35K40, 35R05, 35B45, 35B65, 46E30

1. Introduction

In the present work we obtain a global gradient estimate for weak solutions of parabolic systems in divergence form with bounded measurable coefficients when

*Corresponding author

the associated nonhomogeneous term belongs to weighted Orlicz spaces. Let Ω be a bounded domain with a nonsmooth boundary $\partial\Omega$ and set Ω_T for the cylinder $\Omega\times(0,T)$ with a parabolic boundary $\partial_p\Omega_T=\partial\Omega\times[0,T)\cup\Omega\times\{0\}$.

Consider the following Cauchy–Dirichlet problem with zero boundary data

$$\begin{cases} u^i_t - D_\alpha\left(a^{\alpha\beta}_{ij}(x,t)D_\beta u^j\right) = D_\alpha f^\alpha_i(x,t) & \text{in } \Omega_T,\\ u^i=0 & \text{on } \partial_p\Omega_T \end{cases} \tag{1.1}$$

where $i=1,2,\dots,m$, and the standard summation convention on the repeated upper and lower indexes is adopted for $1\le\alpha,\beta\le n$ with $n\ge2$ and $1\le i,j\le m$ with $m\ge1$.

The tensor matrix of the coefficients

$$\mathbf{A}(x,t)=\left\{a^{\alpha\beta}_{ij}(x,t)\right\}:\ \Omega_T\to\mathbb{R}^{mn\times mn}$$

is assumed to be uniformly bounded and uniformly parabolic. Namely, we suppose that there exist positive constants L and ν such that

$$\|\mathbf{A}\|_{L^\infty(\Omega_T,\mathbb{R}^{mn\times mn})}\le L, \tag{1.2}$$

and

$$a^{\alpha\beta}_{ij}(x,t)\xi^i_\alpha\xi^j_\beta\ge\nu|\xi|^2 \tag{1.3}$$

for all matrices $\xi\in\mathbb{R}^{mn}$ and for almost every $(x,t)\in\Omega_T$. According to the standard existence and regularity theory for (1.1) (see [16, Chapter VII]) the Cauchy–Dirichlet problem (1.1) has a unique weak solution $\mathbf{u}=(u^1,\dots,u^m)$ when the nonhomogeneous term $\mathbf{F}=\{f^\alpha_i(x,t)\}$ belongs to $L^2(\Omega_T,\mathbb{R}^{mn})$. This means that

$$\mathbf{u}=(u^1,\dots,u^m)\in C^0([0,T];L^2(\Omega,\mathbb{R}^m))\cap L^2(0,T;H^1_0(\Omega,\mathbb{R}^m))$$

and it satisfies (1.1) in a weak sense

$$-\int_{\Omega_T}u^i\varphi^i_t\,dxdt+\int_{\Omega_T}a^{\alpha\beta}_{ij}D_\beta u^jD_\alpha\varphi^i\,dxdt=-\int_{\Omega_T}f^\alpha_iD_\alpha\varphi^i\,dxdt$$

for all $\varphi=(\varphi^1,\dots,\varphi^m)\in C^\infty_0(\Omega_T,\mathbb{R}^m)$ with $\varphi(\cdot,T)=0$. Moreover, the standard L^2-estimate

$$\max_{0\le t\le T}\|\mathbf{u}(\cdot,t)\|^2_{L^2(\Omega)}+\int_{\Omega_T}|D\mathbf{u}(x,t)|^2\,dxdt\le c\int_{\Omega_T}|\mathbf{F}(x,t)|^2\,dxdt \tag{1.4}$$

holds true with a constant $c=c(n,m,\nu,L,|\Omega_T|)>0$, where $D\mathbf{u}$ is the spatial gradient matrix of $\mathbf{u}$. In particular, the weak solution $\mathbf{u}$ of (1.1) belongs to

$$H^{\frac12}(0,T;L^2(\Omega,\mathbb{R}^m))\cap L^2(0,T;H^1_0(\Omega,\mathbb{R}^m)). \tag{1.5}$$

Our goal here is to obtain a gradient estimate for the weak solution in the framework of the weighted Orlicz spaces $L^\Phi_w(\Omega_T)$ where Φ is a Young function satisfying appropriate conditions and w is a weight of Muckenhoupt type.

The present work is a natural extension of the results obtained in the earlier papers [5,6,7,9] for the unweighted case. In particular, the article [5] deals with the problem (1.1) in classical (unweighted) Orlicz spaces. Later on, [4] studies (1.1) in the settings of the weighted Lebesgue spaces L^p_w and as a consequence Morrey regularity of the gradient follows. In contrast to [5,4] however, the main difficulties here come from the properties of the wider functional class containing the nonhomogeneous term $\mathbf{F}$. Although we use a similar analytic approach as in [4,5,6], refined analysis of the admissible Muckenhoupt classes of weights is required. For this goal, we employ the results of Fiorenza and Krbec [11] which gave a description of the index classes for the admissible Young functions Φ. Furthermore, in order to adopt the *maximal function approach* (see [5,7]) to our case, we need to estimate the level sets of the spatial gradient of the weak solution under the weight w and to get suitable power decay for their weighted measures. This is done by employing the L^Φ_w-maximal inequality proved by Kerman and Torchinsky in [13].

Let us point out that one can obtain the same results via the so-called *maximal function free technique.* This is a very influential approach introduced by Acerbi and Mingione in [1], and later developed for the regularity estimates for nonlinear parabolic problems when the invariance property under scaling and normalization is not available, see [3,8].

Regarding the coefficients $a^{\alpha\beta}_{ij}(x,t)$ of the parabolic operator considered, apart from (1.2) and (1.3), we assume these are *only measurable* with respect to one spatial variable and are averaged in the sense of *small bounded mean oscillation* (BMO) in the remaining space and time variables. This *partially BMO* assumption is fairly general and allows *arbitrary* discontinuity in one spatial direction which is often related to equilibrium problems for linear laminates or composite materials, while the behaviour with respect to the other directions and the time are controlled in terms of small-BMO, such as small factors of the Heaviside step function for example. Indeed, the situations when the coefficients $a^{\alpha\beta}_{ij}(x,t)$ are continuous, VMO or small-BMO with respect to *all variables* are particular cases of the problem here considered. For what concerns the underlying domain Ω, we assume that the *non-smooth boundary* $\partial\Omega$ is *sufficiently flat* in the sense of Reifenberg [19], that is, $\partial\Omega$ is well approximated by hyperplanes at each point and at each scale. This is a sort of minimal regularity of the boundary, guaranteeing validity in Ω of some natural properties of geometric analysis and partial differential equations such as Sobolev extension, nontangential accessibility property, measure density condition, the Poincaré inequality and so on. In particular, domains with C^1-smooth or Lipschitz continuous boundaries with small Lipschitz constant belong to that category, but the class of Reifenberg flat domains extends beyond these common examples and contains domains with rough fractal boundaries such as the Helge von Koch snowflake (cf. [22]).

The paper is organized as follows. Section 2 recalls the definitions and some properties of the Muckenhoupt weights and the Orlicz spaces. In Section 3 we introduce some notations, set down the basic assumptions and state the main result of the paper, while in Section 4 we prove the optimal gradient estimate for the problem

(1.1) in the Orlicz L^{Φ}_w-settings. Throughout, the letter c denotes a constant that can be explicitly computed in terms of known quantities such as L, ν, m, n, Φ, w and $|\Omega_T|$, as well as of δ and R which are related to the minimal regularity requirements imposed on the coefficients $a^{\alpha\beta}_{ij}$ and on the nonsmooth geometric structure of $\partial\Omega$ (cf. Definition 3.1). The exact value of c may vary from one occurrence to another.

2. Muckenhoupt Weights and Weighted Orlicz Spaces

To start with, let us recall the definitions and some basic properties of the Muckenhoupt weights and the Orlicz spaces, referring the readers to [2,12,14,15,18,20,21] for more details.

We endow $\mathbb{R}^{n+1}$ with the classical parabolic metric

$$\varrho(x,t) = \max\{|x|, \sqrt{|t|}\}$$

and in what follows we will use the following families of domains:

- The *parabolic cylinders* $\{\mathcal{I}_r\}_{r>0}$ in $\mathbb{R}^{n+1}$ centered at some point $(y,\tau) \in \mathbb{R}^n \times \mathbb{R}$ and of radius $r > 0$:

$$\mathcal{I} \equiv \mathcal{I}_r \equiv \mathcal{I}_r(y,\tau) = \left\{(x,t) \in \mathbb{R}^{n+1}\colon\ |x-y| < r, |t-\tau| < r^2\right\}$$

  with a Lebesgue measure $|\mathcal{I}_r| = c(n)r^{n+2}$.
- The *cylinders* $\{Q_r\}_{r>0}$ in $\mathbb{R}^{n+1}$ centered at $(y_1, y', \tau) \in \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R}$:

$$Q_r(y_1,y',\tau) = \left\{(x_1,x',t) \in \mathbb{R}^{n+1}\colon\ |x_1-y_1| < r, |x'-y'| < r, |t-\tau| < r^2\right\}$$

  with $|Q_r| = c(n)r^{n+2}$.
- The *cubes* $\{\mathcal{C}_r\}_{r>0}$ in $\mathbb{R}^n$ centered at $(y_1, y') \in \mathbb{R} \times \mathbb{R}^{n-1}$:

$$\mathcal{C}_r(y) = \{(x_1,x') \in \mathbb{R} \times \mathbb{R}^{n-1}\colon\ |x_1-y_1| < r, |x'-y'| < r\}$$

  with $|\mathcal{C}_r| = c(n)r^n$.
- The *n-cylinders* $\{Q'_r\}_{r>0}$ in $\mathbb{R}^{n-1} \times \mathbb{R}$ centered at $(y',\tau) \in \mathbb{R}^{n-1} \times \mathbb{R}$:

$$Q'_r(y',\tau) = \{(x',t) \in \mathbb{R}^{n-1} \times \mathbb{R}\colon\ |x'-y'| < r, |t-\tau| < r^2\}$$

  with $|Q'_r| = c(n)r^{n+1}$.

Consider a positive, locally integrable function $w(x,t)\colon \mathbb{R}^{n+1} \to \mathbb{R}_+$. We say that w is an A_q-weight of Muckenhoupt for some $1 < q < \infty$ if it satisfies the A_q-condition, i.e.,

$$[w]_q = \sup_{\mathcal{I}} \left(\frac{1}{|\mathcal{I}|}\int_{\mathcal{I}} w(x,t)\, dxdt\right)\left(\frac{1}{|\mathcal{I}|}\int_{\mathcal{I}} w(x,t)^{\frac{-1}{q-1}}\, dxdt\right)^{q-1} < \infty \tag{2.1}$$

where the supremum is taken over all parabolic cylinders $\mathcal{I} \subset \mathbb{R}^{n+1}$. There is an alternative way of defining the Muckenhoupt weights. For any nonnegative, locally integrable function f and any cylinder $\mathcal{I}$, the weight w belongs $A_q, 1 < q < \infty$ if and only if

$$\left(\frac{1}{|\mathcal{I}|}\int_{\mathcal{I}} f(x,t)\,dxdt\right)^q \leq \frac{A}{w(\mathcal{I})}\int_{\mathcal{I}} f^q(x,t)w(x,t)\,dxdt < \infty \tag{2.2}$$

for some constant $A = A(q,n) > 0$. Here $w(\mathcal{I})$ is the measure of $\mathcal{I}$ with respect to the weighted Lebesgue measure $w(x,t)\,dxdt$, that is,

$$w(\mathcal{I}) = \int_{\mathcal{I}} w(x,t)\,dxdt < \infty.$$

The smallest constant A for which (2.2) holds equals $[w]_q$. It is an immediate consequence of (2.2) that whenever $w \in A_q$, then it satisfies the doubling property, that is,

$$w(\mathcal{I}_{2r}(y,\tau)) \leq c(n,q)w(\mathcal{I}_r(y,\tau)). \tag{2.3}$$

In fact, applying (2.2) with $\mathcal{I} = \mathcal{I}_{2r}(y,\tau)$ and $f = \chi_{\mathcal{I}_r(y,\tau)}$ (the characteristic function of $\mathcal{I}_r(y,\tau)$) we get (2.3) with a constant $c = [w]_q 2^{q(n+2)}$. The doubling property of w shows that in the definition (2.1) we can replace the family of cylinders $\{\mathcal{I}_r\}_{r>0}$ by any equivalent set of domains and we shall do it when necessary without explicit reference.

A typical example of a Muckenhoupt weight in $\mathbb{R}^{n+1}$ is

$$w_\tau(x,t) = \rho(x,t)^\tau, \qquad \rho(x,t) = \sqrt{\frac{|x|^2 + \sqrt{|x|^4 + 4t^2}}{2}},$$

where $\rho(x,t)$ is a parabolic metric equivalent to ϱ (see [10,20]). The weight w_τ belongs to A_q if and only if $-(n+2) < \tau < (n+2)(q-1)$.

An exhaustive description of the classes A_p can be found in the classical monographs by Stein [20] or Torchinsky [21]. The following lemma collects some basic properties of these weights and the detailed proof can be found in [20, Section V.5.3] or [21, Section IX.4].

Lemma 2.1. *Assume $w \in A_q$ for some $q > 1$. Then*

(i) Increasing property: $w \in A_p$ *with* $p \geq q$ *and* $[w]_p \leq [w]_q$;

(ii) Self-improved property: $w \in A_{q-\varepsilon_0}$ *with small enough* $\varepsilon_0 > 0$ *depending on* $q, [w]_q$ *and* n;

(iii) Reverse doubling property: *There is* $\tau_1 \in (0,1)$ *such that for any open cylinder* $\mathcal{I}$ *and any measurable subset* $\mathcal{A} \subset \mathcal{I}$ *there is a positive constant* c_1 *depending only on* n *and* q *such that*

$$\frac{1}{[w]_q}\left(\frac{|\mathcal{A}|}{|\mathcal{I}|}\right)^q \leq \frac{w(\mathcal{A})}{w(\mathcal{I})} \leq c_1\left(\frac{|\mathcal{A}|}{|\mathcal{I}|}\right)^{\tau_1}. \tag{2.4}$$

For the purposes of this paper, a *Young function* will be any non-negative and strictly increasing convex function Φ defined on $[0,\infty)$ such that

$$\lim_{\rho\to 0^+}\frac{\Phi(\rho)}{\rho}=0,\qquad \lim_{\rho\to+\infty}\frac{\Phi(\rho)}{\rho}=+\infty. \tag{2.5}$$

Definition 2.1. [15,18] A Young function Φ is supposed to satisfy:

(i) Δ_2-*condition:* there exists a constant $\mu>1$ such that

$$\Phi(2\rho)\le\mu\Phi(\rho),\quad \text{for all } \rho\ge 0;$$

(ii) ∇_2-*condition:* there exists a constant $\rho_0>1$ such that

$$\Phi(\rho)\le\frac{1}{2\rho_0}\Phi(\rho_0\rho),\quad \text{for all } \rho>0.$$

Then we write $\Phi\in\Delta_2\cap\nabla_2$.

The limits (2.5), along with $\Delta_2\cap\nabla_2$, mean in particular that

$$0=\Phi(0)=\lim_{\rho\to 0^+}\Phi(\rho),\qquad \lim_{\rho\to+\infty}\Phi(\rho)=+\infty$$

and the limits are neither too slow nor too fast (cf. [15]).

A classical example of a continuous Young function is $\Phi(\rho)=\rho^q$, $1<q<\infty$ which satisfies Δ_2-condition with $\mu>2^q$ and ∇_2 for any $\rho_0\ge 2^{\frac{1}{q-1}}$. Let us note that the Δ_2-condition implies that there exists a constant $\mu(\lambda)>1$ such that

$$\Phi(\lambda\rho)\le\mu(\lambda)\Phi(\rho)\qquad \forall\ \rho>0,\ \lambda>1,$$

and it is also equivalent to the condition

$$\limsup_{\rho\to 0^+}\frac{\Phi(2\rho)}{\Phi(\rho)}<+\infty.$$

One more example of Young function satisfying the $\Delta_2\cap\nabla_2$-condition is $\Phi(\rho)=\rho^\alpha\log(1+\rho),\alpha>1$. As will be seen later, the condition $\Phi\in\Delta_2\cap\nabla_2$ is unavoidable for the type of regularity results we are going to derive here.

Consider now the function

$$h_\Phi(\lambda)=\sup_{\rho>0}\frac{\Phi(\lambda\rho)}{\Phi(\rho)},\quad \lambda>0$$

and define the *lower index* of Φ by

$$i(\Phi)=\lim_{\lambda\to 0^+}\frac{\log(h_\Phi(\lambda))}{\log\lambda}=\sup_{0<\lambda<1}\frac{\log(h_\Phi(\lambda))}{\log\lambda}.$$

We have $i(\Phi)>1$ as consequence of $\Phi\in\nabla_2$ (cf. [11]). On the other hand, $\Phi\in\Delta_2$ implies that there exist two exponents $q_1,q_2\in(1,\infty),q_1\le q_2$, such that

$$\frac{1}{c}\min\{\lambda^{q_1},\lambda^{q_2}\}\Phi(\rho)\le\Phi(\lambda\rho)\le c\max\{\lambda^{q_1},\lambda^{q_2}\}\Phi(\rho)\quad(\rho,\lambda>0) \tag{2.6}$$

with a constant c independent of ρ and λ. The supremum of those q_1 for which (2.6) holds true with $\lambda \geq 1$ being equal to $i(\Phi)$. If, for instance, $\Phi(\rho) = \rho^q$ with $q > 1$ then $i(\Phi) = q$.

Definition 2.2. Given a couple of functions $(w, \Phi) \in (A_q, \Delta_2 \cap \nabla_2)$, the *weighted Orlicz space* $L^\Phi_w(\Omega_T)$ consists of all Lebesgue measurable functions defined in Ω_T, for which there is a constant $\kappa > 0$ such that

$$\int_{\Omega_T} \Phi\left(\frac{|f(x,t)|}{\kappa}\right) w(x,t)\, dxdt \leq 1.$$

The norm of f in $L^\Phi_w(\Omega_T)$ is the infimum over all such κ and is called a *Luxemburg norm.* Precisely

$$\|f\|_{L^\Phi_w(\Omega_T)} = \inf\left\{\kappa > 0\colon \int_{\Omega_T} \Phi\left(\frac{|f(x,t)|}{\kappa}\right) w(x,t)\, dxdt \leq 1\right\} \tag{2.7}$$

and the equality sign occurs in (2.7) if $\Phi \in \Delta_2$ (cf. [15]).

To proceed further, we fix any $(x,t) \in \Omega_T$, and take $\lambda = \kappa$ and $\rho = \frac{|f(x,t)|}{\kappa}$ in (2.6). This gives the following two sides estimate for $\Phi(|f(x,t)|)$

$$\frac{1}{c}\min(\kappa^{q_1}, \kappa^{q_2})\Phi\left(\frac{|f(x,t)|}{\kappa}\right) \leq \Phi(|f(x,t)|) \leq c\max(\kappa^{q_1}, \kappa^{q_2})\Phi\left(\frac{|f(x,t)|}{\kappa}\right).$$

Integrating with respect to the weighted measure over Ω_T and taking the infimum over all $\kappa > 0$ for which $\int_{\Omega_T} \Phi\left(\frac{|f(x,t)|}{\kappa}\right) w(x,t) dxdt \leq 1$, we get

$$\frac{1}{c}\|f\|^\alpha_{L^\Phi_w(\Omega_T)} \leq \int_{\Omega_T} \Phi(|f(x,t)|) w(x,t)\, dxdt \leq c\|f\|^\beta_{L^\Phi_w(\Omega_T)} \tag{2.8}$$

where $\alpha = q_1, \beta = q_2$ if $\|f\|_{L^\Phi_w(\Omega_T)} \geq 1$ and $\alpha = q_2, \beta = q_1$ if $\|f\|_{L^\Phi_w(\Omega_T)} < 1$.

The well known result of Muckenhoupt [17] gives a characterization of the weight functions w for which the Hardy–Littlewood maximal operator is bounded from the weighted Lebesgue spaces L^q_w into itself. Recall that for a locally integrable function $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$, the Hardy–Littlewood maximal function $\mathcal{M}f$ is defined by

$$\mathcal{M}f(y,\tau) = \sup_{\mathcal{I}_r(y,\tau)} \fint_{\mathcal{I}_r(y,\tau)} |f(x,t)|\, dxdt,$$

where the supremum is taken over all parabolic cylinders $\mathcal{I}_r(y,\tau)$ in $\mathbb{R}^{n+1}$. If f is defined only in a bounded domain in $\mathbb{R}^{n+1}$, we set $\mathcal{M}f = \mathcal{M}\overline{f}$, where $\overline{f}$ is the zero extension of f in $\mathbb{R}^{n+1}$. It is well known (see [21]) that the A_q condition is necessary and sufficient for the maximal operator to map L^q_w into the *weak-*L^q_w space, $1 \leq q < \infty$, and the following weak-type estimate holds

$$w\big(\{(x,t) \in \mathbb{R}^{n+1}\colon\ \mathcal{M}f(x,t) > \lambda\}\big) \leq \frac{c}{\lambda^q}\int_{\mathbb{R}^{n+1}} |f(x,t)|^q w(x,t)\, dxdt$$

for any $\lambda > 0$, where c is a positive constant depending only on p and w. Moreover, $\mathcal{M}$ is a continuous operator from L^q_w into itself when $w \in A_q$ for some $q > 1$.

These results have been extended by Kerman and Torchinsky in [13] to the case of weighted reflexive Orlicz spaces, i.e., spaces for which $\Phi \in \Delta_2 \cap \nabla_2$. Precisely, conditions on w are obtained under which the Hardy–Littlewood maximal operator is bounded on L^Φ_w.

Lemma 2.2. *Suppose that $\Phi \in \Delta_2 \cap \nabla_2$ is a Young function and $w \in A_{i(\Phi)}$. Then there exists a constant $c = c(n, \Phi, w) > 1$ such that*

$$\int_{\mathbb{R}^{n+1}} \Phi(|f(x,t)|)w(x,t)\,dxdt \le \int_{\mathbb{R}^{n+1}} \Phi(\mathcal{M}f(x,t))w(x,t)\,dxdt \le c\int_{\mathbb{R}^{n+1}} \Phi(|f(x,t)|)w(x,t)\,dxdt$$

whenever $f \in L^\Phi_w$ with a compact support in $\mathbb{R}^{n+1}$.

As already mentioned, we are going to study the Calderón–Zygmund property for the parabolic operator (1.1) in the settings of the weighted Orlicz spaces. Precisely, we will show that under minimal regularity assumptions on $\mathbf{A}$ and a lower level of geometric requirements on $\partial\Omega$, the following implication holds true

$$|\mathbf{F}|^2 \in L^\Phi_w(\Omega_T) \implies |D\mathbf{u}|^2 \in L^\Phi_w(\Omega_T) \tag{2.9}$$

with $(w, \Phi) \in (A_{i(\Phi)}, \Delta_2 \cap \nabla_2)$. Let us point out that the regularity results derived in [4,5,9] are special cases of (2.9), when $\Phi(\rho) = \rho^{p/2}$ for some $2 < p < \infty$. On the other hand, [6] studies the problem in unweighted Orlicz spaces when the coefficients of the principal operator are only measurable in the time variable. In contrast to these works, we are dealing here with weighted Orlicz classes and we allow the coefficients to be measurable in one of the spatial variables.

3. Assumptions and Main Result

We will study the weak solution of the system (1.1), assuming the coefficients $a^{\alpha\beta}_{ij}(x,t)$ to be only *measurable* in one of the spatial variables, say x_1, and to have a *small mean oscillation* with respect to the remaining variables $(x', t) = (x_2, \dots, x_n, t)$ at each point and at each scale r.

For any fixed $x_1 \in \mathbb{R}$, the integral average of a function $a(x,t)$ with respect to (x', t)-variables on Q'_r will be denoted by

$$\overline{a}_{Q'_r}(x_1) = \fint_{Q'_r} a(x_1, x', t)\,dx'dt = \frac{1}{|Q'_r|}\int_{Q'_r} a(x_1, x', t)\,dx'dt.$$

In what follows we assume a smallness of the BMO norm of $\mathbf{A} = \left\{a^{\alpha\beta}_{ij}\right\}$ with respect to the (x', t)-variables supposing that $\mathbf{A}$ is *only* measurable in x_1. At the same time, the boundary $\partial\Omega$ will be assumed to be flat enough in the sense of Reifenberg. The precise meaning of these requirements is given in the following definition.

Definition 3.1. We say that the couple $(\mathbf{A}, \Omega)$ is (δ, R)-vanishing of codimension 1, if for every point $(y, \tau) \in \Omega \times \mathbb{R}_+$ and for every number $r \in \left(0, \frac{1}{3}R\right]$ with

$$\operatorname{dist}(y, \partial\Omega) > \sqrt{2}r,$$

there exists a coordinate system depending on (y, τ) and r, whose variables we still denote by (x, t), so that in this new coordinate system (y, τ) is the origin and

$$\fint_{Q_r} \left|\mathbf{A}(x,t) - \overline{\mathbf{A}}_{Q'_r}(x_1)\right|^2 \, dxdt \leq \delta^2.$$

Later on, for every point $(y, \tau) \in \Omega \times \mathbb{R}$ and for every number $r \in \left(0, \frac{1}{3}R\right]$ with

$$\operatorname{dist}(y, \partial\Omega) = \operatorname{dist}(y, x_0) \leq \sqrt{2}r$$

for some $x_0 \in \partial\Omega$, there exists a coordinate system depending on (y, τ) and r, whose variables we still denote by (x, t), such that in this new coordinate system (x_0, τ) is the origin,

$$\Omega \cap \{x \in \mathcal{C}_{3r} \colon x_1 > 3r\delta\} \subset \Omega \cap \mathcal{C}_{3r} \subset \Omega \cap \{x \in \mathcal{C}_{3r} \colon x_1 > -3r\delta\} \tag{3.1}$$

and

$$\fint_{Q_{3r}} \left|\mathbf{A}(x,t) - \overline{\mathbf{A}}_{Q'_{3r}}(x_1)\right|^2 \, dxdt \leq \delta^2. \tag{3.2}$$

Some comments regarding the Definition 3.1 are in order. Thanks to a scaling invariance property of the problem (1.1), one can take for simplicity $R = 1$ or any other constants greater than 1. On the other hand δ is a small positive constant, which is invariant under such a scaling argument.

If $(\mathbf{A}, \Omega)$ is (δ, R)-vanishing of codimension 1, then for each point and for each sufficiently small scale there is a coordinate system so that the coefficients have small oscillation in the (x', t)-variables while these might have arbitrary jumps along the x_1-variable. In addition, (3.1) means that the boundary of Ω is δ-Reifenberg flat and the coefficients $a_{ij}^{\alpha\beta}$ have small oscillations along the flat direction x' while these are only measurable along the "normal" direction x_1.

The number $\sqrt{2}r$ above is selected only for convenience. It comes from the reason that we need to take enough space of the cubes so that there is a room to have the rotation of $Q_r(y, \tau)$ in any spatial direction. By the same reason $3r$ appears in (3.1) and (3.2).

We are in a position to state the main result of the paper.

Theorem 3.1. *Given a Young function* $\Phi \in \Delta_2 \cap \nabla_2$ *and a weight* $w(x,t) \in A_{i(\Phi)}$, *suppose that*

$$\int_{\Omega_T} \Phi(|\mathbf{F}(x,t)|^2) w(x,t) \, dxdt < \infty.$$

Then there exists a small constant $\delta > 0$ depending on known quantities, such that if $(\mathbf{A}, \Omega)$ is (δ, R)-vanishing of codimension 1, then the unique weak solution $\mathbf{u}$ of (1.1) satisfies

$$\int_{\Omega_T} \Phi(|D\mathbf{u}(x,t)|^2) w(x,t)\; dxdt < \infty$$

and we have the estimate

$$\||D\mathbf{u}|^2\|_{L^\Phi_w(\Omega_T)} \le c \||\mathbf{F}|^2\|_{L^\Phi_w(\Omega_T)}$$

with a constant c independent of $\mathbf{u}$ and $\mathbf{F}$.

4. Gradient Estimate in Weighted Orlicz Spaces

We start with some auxiliary results.

Lemma 4.1. *Assume $|\mathbf{F}|^2 \in L^\Phi_w(\Omega_T)$ with $(w, \Phi) \in (A_{i(\Phi)}, \Delta_2 \cap \nabla_2)$. Then $|\mathbf{F}|^2 \in L^1(\Omega_T)$ and*

$$\int_{\Omega_T} |\mathbf{F}(x,t)|^2\; dxdt \le c \left(\||\mathbf{F}|^2\|^{\beta'}_{L^\Phi_w(\Omega_T)} + 1 \right) \tag{4.1}$$

for some positive constants c and β' depending only on n, Φ, w and Ω_T.

Proof. Because of the self-improving property of w, there exists a small $\varepsilon_0 > 0$ depending on n, w, and Φ such that $w \in A_{i(\Phi)-\varepsilon_0}$ and $i(\Phi) - \varepsilon_0 > 1$. Let us set $i_0 = i(\Phi) - \varepsilon_0$ for the sake of simplicity. By the Hölder inequality and the A_q-condition (2.1) we get

$$\begin{aligned}
\int_{\Omega_T} |\mathbf{F}(x,t)|^2\; dxdt &= \int_{\Omega_T} |\mathbf{F}(x,t)|^2 w(x,t)^{\frac{1}{i_0}} w(x,t)^{-\frac{1}{i_0}}\; dxdt \\
&\le \left(\int_{\Omega_T} |\mathbf{F}(x,t)|^{2i_0} w(x,t)\; dxdt \right)^{\frac{1}{i_0}} \left(\int_{\Omega_T} w(x,t)^{-\frac{1}{i_0-1}}\; dxdt \right)^{\frac{i_0-1}{i_0}} \\
&\le \frac{|\mathcal{I}_{\Omega_T}| [w]_{i_0}^{\frac{1}{i_0}}}{w(\mathcal{I}_{\Omega_T})^{\frac{1}{i_0}}} \left(\int_{\Omega_T} |\mathbf{F}(x,t)|^{2i_0} w(x,t)\; dxdt \right)^{\frac{1}{i_0}} \\
&\le c(n, w, \Omega_T) \Bigg(\underbrace{\int_{\Omega_T} |\mathbf{F}(x,t)|^{2i_0} w(x,t)\; dxdt}_{=:I} \Bigg)^{\frac{1}{i_0}},
\end{aligned}$$

where $\mathcal{I}_{\Omega_T}$ is a parabolic cylinder which contains Ω_T.

To estimate I, we first claim that

$$|\mathbf{F}(x,t)|^{2i_0} \le c(\Phi) \left(\Phi(|\mathbf{F}(x,t)|^2) + 1 \right) \quad (x,t) \in \Omega_T.$$

Indeed, if $|\mathbf{F}(x,t)|^2 \le 1$ for each $(x,t) \in \Omega_T$, then there is nothing to prove. On the other hand, if $|\mathbf{F}(x,t)|^2 \ge 1$ for some $(x,t) \in \Omega_T$, we apply (2.6) with $\lambda = |\mathbf{F}(x,t)|^2$,

$\rho = 1$ and $q_1 = i_0$ (recall that $i(\Phi) = \sup q_1$ for which (2.6) holds with $\lambda \geq 1$), to discover that

$$|\mathbf{F}(x,t)|^{2i_0} \leq |\mathbf{F}(x,t)|^{2q_1} \leq \frac{c}{\Phi(1)}\Phi(|\mathbf{F}(x,t)|^2) \leq c\ \Phi(|\mathbf{F}(x,t)|^2),$$

where the constant c is depending only on Φ. Therefore, we have

$$\begin{aligned} I &\leq c(\Phi)\int_{\Omega_T}\left(\Phi(|\mathbf{F}(x,t)|^2)+1\right)w(x,t)\ dxdt \\ &\leq c(\Phi,w,\Omega_T)\left(\int_{\Omega_T}\Phi(|\mathbf{F}(x,t)|^2)w(x,t)\ dxdt+1\right) \\ &\leq c(\Phi,w,\Omega_T)\left(\||\mathbf{F}|^2\|^{\beta}_{L^{\Phi}_w(\Omega_T)}+1\right), \end{aligned}$$

where we have used (2.8) for the last inequality. Hence we conclude

$$\begin{aligned} \int_{\Omega_T}|\mathbf{F}(x,t)|^2\ dxdt &\leq c(n,\Phi,w,\Omega_T)I^{\frac{1}{i_0}} \\ &\leq c(n,\Phi,w,\Omega_T)\left(\||\mathbf{F}|^2\|^{\frac{\beta}{i_0}}_{L^{\Phi}_w(\Omega_T)}+1\right) \end{aligned}$$

and this completes the proof. □

It should be noted that Lemma 4.1 ensures that for each $\mathbf{F}(x,t)$ with $|\mathbf{F}|^2 \in L^{\Phi}_w(\Omega_T)$ the problem (1.1) has a unique weak solution in (1.5) satisfying

$$\||D\mathbf{u}|^2\|_{L^1(\Omega_T)} \leq c\||\mathbf{F}|^2\|_{L^1(\Omega_T)} \leq c\left(\||\mathbf{F}|^2\|^{\beta'}_{L^{\Phi}_w(\Omega_T)}+1\right).$$

Our approach in the sequel is based on harmonic analysis tools such as the maximal function operator and a Vitali type covering lemma (cf. [20,21]). Precisely, we will use the following weighted version of the Vitali covering lemma.

Lemma 4.2. *Let Ω be a bounded, $(\delta,1)$-Reifenberg flat domain, $\mathfrak{C}$ and $\mathfrak{D}$ be measurable sets such that $\mathfrak{C} \subset \mathfrak{D} \subset \Omega_T$ and $w(x,t) \in A_q$ for some $q \in (1,\infty)$. Assume that there exists $\varepsilon \in (0,1)$ such that:*

(i) for any $(y,\tau) \in \Omega_T$

$$\frac{w(\mathfrak{C}\cap Q_1(y,\tau))}{w(Q_1(y,\tau))} < \varepsilon; \tag{4.2}$$

(ii) for each $(y,\tau) \in \Omega_T$ and some $r \in (0,1)$,

$$\frac{w(\mathfrak{C}\cap Q_r(y,\tau))}{w(Q_r(y,\tau))} \geq \varepsilon \quad \textit{implies} \quad \Omega_T \cap Q_r(y,\tau) \subset \mathfrak{D}. \tag{4.3}$$

Then

$$w(\mathfrak{C}) \leq \varepsilon[w]_q^2\left(\frac{10\sqrt{2}}{1-\delta}\right)^{q(n+2)} w(\mathfrak{D}). \tag{4.4}$$

Proof. Fix $(y,\tau)\in\mathfrak{C}$ and for $r>0$ define the function

$$\Theta(r)=\frac{\int_{\mathfrak{C}\cap Q_r(y,\tau)} w(x,t)\,dxdt}{\int_{Q_r(y,\tau)} w(x,t)\,dxdt}=\frac{w(\mathfrak{C}\cap Q_r(y,\tau))}{w(Q_r(y,\tau))}.$$

Then we note that $\Theta(r)\in C^0(0,\infty)$, $\Theta(1)<\varepsilon$ by (4.2) and $\Theta(0)=\lim_{r\to 0^+}\Theta(r)=1$ by the Lebesgue Differentiation Theorem. Hence, for each $(y,\tau)\in\mathfrak{C}$ there exists $r_{(y,\tau)}\in(0,1)$ such that

$$\Theta(r_{(y,\tau)})=\varepsilon,\qquad \Theta(r)<\varepsilon\quad \forall r>r_{(y,\tau)}. \tag{4.5}$$

Since the family of sets $\{Q_{r_{(y,\tau)}}(y,\tau)\}_{(y,\tau)\in\mathfrak{C}}$ is an open covering of $\mathfrak{C}$, by the Vitali lemma (cf. [20, Lemma I.3.1]) there exists a disjoint subcollection $\{Q_{r_i}(y_i,\tau_i)\}_{i\geq 1}$ with $r_i=r_{(y_i,\tau_i)}$, $(y_i,\tau_i)\in\mathfrak{C}$ such that $\Theta(r_i)=\varepsilon$ as in (4.5) and

$$\mathfrak{C}\subset\bigcup_{i\geq 1} Q_{5r_i}(y_i,\tau_i)$$

with a positive constant $c=c(n)$. Since $\Theta(5r_i)<\varepsilon$, by (2.4) we have

$$\begin{aligned} w\big(\mathfrak{C}\cap Q_{5r_i}(y_i,\tau_i)\big) &< \varepsilon w(Q_{5r_i}(y_i,\tau_i))\\ &\leq \varepsilon[w]_q\left(\frac{|Q_{5r_i}(y_i,\tau_i)|}{|Q_{r_i}(y_i,\tau_i)|}\right)^q w(Q_{r_i}(y_i,\tau_i))\\ &= \varepsilon[w]_q 5^{q(n+2)} w(Q_{r_i}(y_i,\tau_i)). \end{aligned}$$

We combine now the estimate

$$\sup_{0<r<1}\sup_{(y,\tau)\in\Omega_T}\frac{|Q_r(y,\tau)|}{|\Omega_T\cap Q_r(y,\tau)|}\leq\left(\frac{2\sqrt{2}}{1-\delta}\right)^{n+2}$$

obtained in [7] (see also [5]) with the doubling condition (2.4) in order to get

$$w(Q_{r_i}(y_i,\tau_i))\leq[w]_q\left(\frac{2\sqrt{2}}{1-\delta}\right)^{q(n+2)} w(\Omega_T\cap Q_{r_i}(y_i,\tau_i)).$$

This way, the weighted measure of $\mathfrak{C}$ is estimated as follows

$$\begin{aligned} w(\mathfrak{C}) &= w\left(\bigcup_{i\geq 1}\big(\mathfrak{C}\cap Q_{5r_i}(y_i,\tau_i)\big)\right)\\ &\leq \sum_{i\geq 1} w\big(\mathfrak{C}\cap Q_{5r_i}(y_i,\tau_i)\big)<\varepsilon\sum_{i\geq 1} w\big(Q_{5r_i}(y_i,\tau_i)\big)\\ &\leq \varepsilon[w]_q 5^{q(n+2)}\sum_{i\geq 1} w\big(Q_{r_i}(y_i,\tau_i)\big)\\ &\leq \varepsilon[w]_q^2\left(\frac{10\sqrt{2}}{1-\delta}\right)^{q(n+2)}\sum_{i\geq 1} w\big(\Omega_T\cap Q_{r_i}(y_i,\tau_i)\big). \end{aligned}$$

Since $\{Q_{r_i}(y_i,\tau_i)\}$ are mutually disjoint sets and $\Theta(r_i)=\varepsilon$, we get by (4.3) that $\bigcup_{i\geq 1}\Omega_T\cap Q_{r_i}(y_i,\tau_i)\subset\mathfrak{D}$ whence

$$\begin{aligned} w(\mathfrak{C}) &\leq \varepsilon[w]_q^2\left(\frac{10\sqrt{2}}{1-\delta}\right)^{q(n+2)} w\left(\bigcup_{i\geq 1}\Omega_T\cap Q_{r_i}(y_i,\tau_i)\right)\\ &\leq \varepsilon[w]_q^2\left(\frac{10\sqrt{2}}{1-\delta}\right)^{q(n+2)} w(\mathfrak{D}). \end{aligned}$$

□

Let us recall the next unweighted result which is proved in [5, Lemma 5.3] and [7, Lemma 5.5].

Lemma 4.3. *Let* $\mathbf{u}$ *be the weak solution of* (1.1) *and assume* (1.2) *and* (1.3). *Then there is a constant* $\lambda_1=\lambda_1(\nu,L,m,n)>1$ *such that for any fixed* $\varepsilon\in(0,1)$ *there exists* $\delta=\delta(\varepsilon,\nu,L,m,n)>0$ *such that if* $(\mathbf{A},\Omega)$ *is* $(\delta,1)$*-vanishing of codimension 1 and if* $Q_r(y,\tau)$ *satisfies*

$$\left|\left\{(x,t)\in\Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2)>\lambda_1^2\right\}\cap Q_r(y,\tau)\right|\geq\varepsilon|Q_r(y,\tau)|,$$

then we have

$$\Omega_T\cap Q_r(y,\tau)\subset\left\{(x,t)\in\Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2)>1\right\}\cup\left\{\mathcal{M}(|\mathbf{F}|^2)>\delta^2\right\}.$$

To go further, we need to establish a weighted version of that result. For, fix ε and take δ and λ_1 as given by Lemma 4.3. With the weak solution $\mathbf{u}$ of (1.1) at hand, we define the sets

$$\begin{aligned} \mathfrak{C} &= \{(x,t)\in\Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2)>\lambda_1^2\},\\ \mathfrak{D} &= \{(x,t)\in\Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2)>1\}\cup\{\mathcal{M}(|\mathbf{F}|^2)>\delta^2\}. \end{aligned}\tag{4.6}$$

Lemma 4.4. *Let* $w(x,t)\in A_q$ *for some* $q\in(1,\infty)$. *Assume that* $(\mathbf{A},\Omega)$ *is* $(\delta,1)$*-vanishing of codimension 1 and that for all* $(y,\tau)\in\Omega_T$ *and some* $r\in(0,1)$ *one has*

$$\Theta(r)=\frac{w(\mathfrak{C}\cap Q_r(y,\tau))}{w(Q_r(y,\tau))}\geq\varepsilon.$$

Then $\Omega_T\cap Q_r(y,\tau)\subset\mathfrak{D}$.

Proof. By the reverse doubling property (2.4) we have

$$\varepsilon\leq\frac{w(\mathfrak{C}\cap Q_r(y,\tau))}{w(Q_r(y,\tau))}\leq c_1\left(\frac{|\mathfrak{C}\cap Q_r(y,\tau)|}{|Q_r(y,\tau)|}\right)^{\tau_1}$$

for some $c_1>0$ and $\tau_1\in(0,1)$. Hence

$$|\mathfrak{C}\cap Q_r(y,\tau)|\geq\left(\frac{\varepsilon}{c_1}\right)^{\frac{1}{\tau_1}}|Q_r(y,\tau)|=\varepsilon_1|Q_r(y,\tau)|.$$

The assertion holds after applying Lemma 4.3. □

We are in a position now to derive the power decay estimate for the weighted measure of the level set $\mathfrak{C} = \{(x,t) \in \Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2) > \lambda_1^2\}$ with respect to the A_q-weight w.

Lemma 4.5. *In addition to the assumptions of Lemma 4.4, we suppose*

$$\Theta(1) = \frac{w(\mathfrak{C} \cap Q_1(y,\tau))}{w(Q_1(y,\tau))} < \varepsilon \qquad \forall\ (y,\tau) \in \Omega_T. \tag{4.7}$$

Then

$$\begin{aligned} w\big(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2) > \lambda_1^{2k}\}\big) & \\ \leq \varepsilon_1^k w\left(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2) > 1\}\right) & \\ + \sum_{i=1}^{k} \varepsilon_1^i w\left(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|\mathbf{F}|^2) > \delta^2 \lambda_1^{2(k-i)}\}\right) & \end{aligned} \tag{4.8}$$

for $k = 1, 2, \ldots,$ *where* $\varepsilon_1 = \varepsilon [w]_q^2 \left(\frac{10\sqrt{2}}{1-\delta}\right)^{q(n+2)}$.

Proof. Lemma 4.4 and (4.7) ensure the validity of the hypotheses of Lemma 4.2 for the sets $\mathfrak{C}$ and $\mathfrak{D}$ as defined by (4.6). Thus, (4.4) implies

$$\begin{aligned} w\big(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2) > \lambda_1^2\}\big) &\leq \varepsilon_1 w\left(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2) > 1\}\right) \\ &\quad + \varepsilon_1 w\left(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|\mathbf{F}|^2) > \delta^2\}\right), \end{aligned}$$

with $\varepsilon_1 = \varepsilon [w]_q^2 \left(\frac{10\sqrt{2}}{1-\delta}\right)^{q(n+2)}$. The last inequality is just (4.8) with $k = 1$. Further on, we proceed with the proof by induction. Suppose that (4.8) holds true for the weak solution of (1.1) and for some $k > 1$. Define the functions

$$\mathbf{u}_1 = \frac{\mathbf{u}}{\lambda_1} \quad \text{and} \quad \mathbf{F}_1 = \frac{\mathbf{F}}{\lambda_1}.$$

It is easy to see that $\mathbf{u}_1$ is a weak solution to the problem (1.1) with a right-hand side $\mathbf{F}_1$. Hence, (4.7) and Lemma 4.4 hold with the sets $\mathfrak{C}$ and $\mathfrak{D}$ corresponding to $\mathbf{u}_1$ as defined by (4.6). According to (4.8), the inductive assumption holds true for $\mathbf{u}_1$ with the same $k > 1$. The definition of $\mathbf{u}_1$ ensures the inductive passage from k

to $k+1$ for $\mathbf{u}$. Namely,

$$\begin{aligned} w\big(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2) > \lambda_1^{2(k+1)}\}\big) \\ &= w\left(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|D\mathbf{u}_1|^2) > \lambda_1^{2k}\}\right) \\ &\le \varepsilon_1^k w\left(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|D\mathbf{u}_1|^2) > 1\}\right) \\ &\quad + \sum_{i=1}^{k} \varepsilon_1^i w\left(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|\mathbf{F}_1|^2) > \delta^2 \lambda_1^{2(k-i)}\}\right) \\ &= \varepsilon_1^k w\left(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|D\mathbf{u}|^2) > \lambda_1^2\}\right) \\ &\quad + \sum_{i=1}^{k} \varepsilon_1^i w\left(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|\mathbf{F}|^2) > \delta^2 \lambda_1^{2(k-i)}\lambda_1^2\}\right) \\ &\le \varepsilon_1^{k+1} w\left(\{(x,t) \in \Omega_T\colon \mathcal{M}(|D\mathbf{u}|^2) > 1\}\right) \\ &\quad + \sum_{i=1}^{k+1} \varepsilon_1^i w\left(\{(x,t) \in \Omega_T\colon\ \mathcal{M}(|\mathbf{F}|^2) > \delta^2 \lambda_1^{2(k+1-i)}\}\right). \qquad \square \end{aligned}$$

We need also the following standard result from the classical measure theory regarding weighted Orlicz spaces, see [2,12,14].

Lemma 4.6. *Let $h = h(x,t)$ be a nonnegative and measurable function in Ω_T. Let $\theta > 0$ and $\lambda > 1$ be constants.*

Then for any $(w, \Phi) \in (A_q, \Delta_2 \cap \nabla_2)$, $q \in (1,\infty)$, one has

$$h \in L^\Phi_w(\Omega_T) \Longleftrightarrow \sum_{k\ge 1} \Phi(\lambda^k) w\left(\{(x,t) \in \Omega_T\colon\ |h(x,t)| > \theta\lambda^k\}\right) =: S < \infty$$

and

$$\frac{1}{c} S \le \int_{\Omega_T} \Phi(|h(x,t)|) w(x,t)\, dxdt \le c\left(w(\Omega_T) + S\right), \tag{4.9}$$

where $c > 0$ is a constant depending only on θ, λ and Φ.

Proof. Since Φ is strictly increasing we can write

$$\begin{aligned} \int_{\Omega_T} \Phi(|h(x,t)|) w(x,t)\, dxdt \\ &= \int_{\{(x,t)\in\Omega_T : |h(x,t)| \le \lambda\theta\}} \Phi(|h(x,t)|) w(x,t)\, dxdt \\ &\quad + \sum_{k\ge1} \int_{\{(x,t)\in\Omega_T : \theta\lambda^k < |h(x,t)| \le \theta\lambda^{k+1}\}} \Phi(|h(x,t)|) w(x,t)\, dxdt \\ &\le \Phi(\theta\lambda) w(\Omega_T) + \sum_{k\ge1} \Phi(\theta\lambda^{k+1}) w(\{(x,t) \in \Omega_T \colon |h(x,t)| > \theta\lambda^k\}). \end{aligned}$$

If $\theta\lambda < 1$ then $\theta\lambda^{k+1} \le \lambda^k$ and $\Phi(\theta\lambda^{k+1}) \le \Phi(\lambda^k)$. If $\theta\lambda \ge 1$, the Δ_2-condition ensures $\Phi(\theta\lambda^{k+1}) \le \mu(\theta\lambda)\Phi(\lambda^k)$. Hence the second inequality in (4.9) holds. On the other hand, by the Fubini theorem we have

$$\begin{aligned}
&\int_{\Omega_T} \Phi(|h(x,t)|)w(x,t)\,dxdt \\
&\qquad = \int_{\Omega_T} w(x,t)\left(\int_0^{|h(x,t)|} d\Phi(\tau)\right) dxdt \\
&\qquad = \int_0^{\infty}\left(\int_{\{(x,t)\in\Omega_T:|h(x,t)|>\tau\}} w(x,t)\,dxdt\right) d\Phi(\tau) \\
&\qquad \ge \sum_{k\ge 1}\int_{\theta\lambda^{k-1}}^{\theta\lambda^k}\left(\int_{\{(x,t)\in\Omega_T:|h(x,t)|>\tau\}} w(x,t)\,dxdt\right) d\Phi(\tau) \\
&\qquad \ge \sum_{k\ge 1}\left(\int_{\{(x,t)\in\Omega_T:|h(x,t)>\theta\lambda^k|\}} w(x,t)\,dxdt\right)\int_{\theta\lambda^{k-1}}^{\theta\lambda^k} d\Phi(\tau) \\
&\qquad = \big(\Phi(\theta\lambda^k)-\Phi(\theta\lambda^{k-1})\big)\, w\big(\{(x,t)\in\Omega_T:|h(x,t)|>\theta\lambda^k\}\big) \\
&\qquad \ge \sum_{k\ge 1}\left(1-\frac{1}{2\lambda}\right)\Phi(\theta\lambda^k)w\big(\{(x,t)\in\Omega_T:|h(x,t)|>\theta\lambda^k\}\big)
\end{aligned}$$

where we have used that $\Phi\in\nabla_2$. If $\theta<1$ then $\Phi(\theta\lambda^k)\ge\frac{1}{\mu(\frac{1}{\theta})}\Phi(\lambda^k)$, while for $\theta\ge 1$ there holds $\Phi(\theta\lambda^k)\ge\Phi(\lambda^k)$ by the properties of Φ. This implies the first inequality in (4.9). □

We are ready now to complete the proof of Theorem 3.1.

Assume, without loss of generality, that the L^Φ_w-norm of $|\mathbf{F}|^2$ is less than 1. In fact, by the scaling invariance property of (1.1) we can normalize the solution, making the norm of the right-hand side arbitrary small. Precisely, taking

$$\overline{\mathbf{u}} = \frac{\delta\mathbf{u}}{\sqrt{\||\mathbf{F}|^2\|_{L^\Phi_w(\Omega_T)}}} \quad \text{and} \quad \overline{\mathbf{F}} = \frac{\delta\mathbf{F}}{\sqrt{\||\mathbf{F}|^2\|_{L^\Phi_w(\Omega_T)}}}$$

instead of $\mathbf{u}$ and $\mathbf{F}$ in (1.1), we get by (4.1) that

$$\big\||\overline{\mathbf{F}}|^2\big\|_{L^\Phi_w(\Omega_T)} = \delta^2 \quad \text{and} \quad \int_{\Omega_T} |\overline{\mathbf{F}}(x,t)|^2\,dxdt \le c\delta^2. \tag{4.10}$$

In fact, the second inequality follows from

$$\begin{aligned}
\int_{\Omega_T} |\overline{\mathbf{F}}(x,t)|^2\,dxdt &\le \delta^2\int_{\Omega_T}\left|\frac{\mathbf{F}(x,t)}{\sqrt{\||\mathbf{F}|^2\|_{L^\Phi_w(\Omega_T)}}}\right|^2 dxdt \\
&\le c\delta^2\left(\left\|\frac{|\mathbf{F}|^2}{\||\mathbf{F}|^2\|_{L^\Phi_w(\Omega_T)}}\right\|^{\beta'}_{L^\Phi_w(\Omega_T)} + 1\right) \le c\delta^2.
\end{aligned}$$

From (2.8) and Lemma 2.2 we have

$$\begin{aligned}\||D\overline{\mathbf{u}}|^2\|^{\alpha}_{L^{\Phi}_w(\Omega_T)} &\le c\int_{\Omega_T}\Phi(|D\overline{\mathbf{u}}(x,t)|^2)w(x,t)\,dxdt\\ &\le c\int_{\Omega_T}\Phi(\mathcal{M}(|D\overline{\mathbf{u}}(x,t)|^2))w(x,t)\,dxdt.\end{aligned}$$

Hence, in order to obtain a gradient estimate in $L^{\Phi}_w(\Omega_T)$, it is enough to prove

$$\int_{\Omega_T}\Phi(\mathcal{M}(|D\overline{\mathbf{u}}(x,t)|^2))w(x,t)\,dxdt \le c. \tag{4.11}$$

To achieve this, we apply Lemma 4.6 with $h = \mathcal{M}(|D\overline{\mathbf{u}}|^2)$, $\lambda = \lambda_1^2$, $\theta = 1$ and Lemma 4.5. To check the condition (4.7), we make use of the doubling property (2.4) as well as (4.6), (1.4) and (4.10) in order to get

$$\begin{aligned}\Theta(1) &= \frac{w(\mathfrak{C}\cap Q_1(y,\tau))}{w(Q_1(y,\tau))} \le c\left(\frac{|\mathfrak{C}\cap Q_1(y,\tau)|}{|Q_1(y,\tau)|}\right)^{\tau_1} \le c|\mathfrak{C}|^{\tau_1}\\ &\le c\left(\int_{\Omega_T}\mathcal{M}(|D\overline{\mathbf{u}}(x,t)|^2)\,dxdt\right)^{\tau_1} \le c\left(\int_{\Omega_T}|D\overline{\mathbf{u}}(x,t)|^2\,dxdt\right)^{\tau_1}\\ &\le c\left(\int_{\Omega_T}|\overline{\mathbf{F}}(x,t)|^2\,dxdt\right)^{\tau_1} \le c\delta^{2\tau_1} < \varepsilon\end{aligned}$$

for small enough δ.

In the light of the power decay estimate (4.8) in Lemma 4.5, we have

$$\begin{aligned}S &:= \sum_{k\ge1}\Phi(\lambda_1^{2k})w\left(\{(x,t)\in\Omega_T\colon\ \mathcal{M}(|D\overline{\mathbf{u}}|^2) > \lambda_1^{2k}\}\right)\\ &\le \sum_{k\ge1}\Phi(\lambda_1^{2k})\varepsilon_1^k w\left(\{(x,t)\in\Omega_T\colon\ \mathcal{M}(|D\overline{\mathbf{u}}|^2) > 1\}\right)\\ &\quad + \sum_{k\ge1}\Phi(\lambda_1^{2k})\sum_{i=1}^{k}\varepsilon_1^i w\left(\{(x,t)\in\Omega_T\colon\ \mathcal{M}(|\overline{\mathbf{F}}|^2) > \delta^2\lambda_1^{2(k-i)}\}\right)\\ &=: S_1 + S_2.\end{aligned}$$

Employing once again the properties of Φ, we get that there exists a constant μ_1 depending on λ_1, and so ν, L, m and n, such that $\Phi(\lambda_1^2) \le \mu_1\Phi(1)$ and therefore $\Phi(\lambda_1^{2k}) \le \mu_1^k\Phi(1)$ by iteration. This way,

$$S_1 \le \Phi(1)w(\Omega_T)\sum_{k\ge1}(\mu_1\varepsilon_1)^k \le c\sum_{k\ge1}(\mu_1\varepsilon_1)^k$$

and (4.9) and (2.6) yield

$$\begin{aligned} S_2 &\le \sum_{i\ge 1}\varepsilon_1^i \sum_{k\ge i}\Phi(\lambda_1^{2(k-i)}\lambda_1^{2i})w\left(\{(x,t)\in\Omega_T:\ \mathcal{M}(|\overline{\mathbf{F}}|^2)>\delta^2\lambda_1^{2(k-i)}\}\right)\\ &\le \sum_{i\ge 1}(\varepsilon_1\mu_1)^i \sum_{k\ge i}\Phi(\lambda_1^{2(k-i)})w\left(\{(x,t)\in\Omega_T:\ \mathcal{M}\left(\frac{|\overline{\mathbf{F}}|^2}{\delta^2}\right)>\lambda_1^{2(k-i)}\}\right)\\ &\le c\sum_{i\ge 1}(\varepsilon_1\mu_1)^i\int_{\Omega_T}\Phi\left(\mathcal{M}\left(\frac{|\overline{\mathbf{F}}|^2}{\delta^2}\right)\right)w(x,t)\,dxdt\\ &\le c\sum_{i\ge 1}(\varepsilon_1\mu_1)^i\int_{\Omega_T}\Phi\left(\frac{|\overline{\mathbf{F}}|^2}{\delta^2}\right)w(x,t)\,dxdt\\ &\le c\sum_{i\ge 1}(\varepsilon_1\mu_1)^i\left\|\frac{|\overline{\mathbf{F}}|^2}{\delta^2}\right\|^{\beta}_{L^\Phi_w(\Omega_T)} \le c\sum_{i\ge 1}(\varepsilon_1\mu_1)^i \end{aligned}$$

since $\||\overline{\mathbf{F}}|^2\|_{L^\Phi_w(\Omega_T)}=\delta^2$. Unifying the both estimates we get

$$S\le c\sum_{k\ge 1}(\varepsilon_1\mu_1)^k$$

and the last series converges when ε_1 is small enough. Hence (4.11) holds true and

$$\||D\overline{\mathbf{u}}|^2\|_{L^\Phi_w(\Omega_T)}<c$$

with a constant depending on known quantities. Recalling the definition of $\overline{\mathbf{u}}$, we get the desired a priori estimate

$$\||D\mathbf{u}|^2\|_{L^\Phi_w(\Omega_T)}<c\||\mathbf{F}|^2\|_{L^\Phi_w(\Omega_T)}$$

and this completes the proof of Theorem 3.1.

Acknowledgments

S. Byun was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (MEST) (2012-047030). J. Ok was supported by TJ Park Science Fellowship of POSCO TJ Park Foundation. D.K. Palagachev and L.G. Softova are members of INdAM–GNAMPA.

References

[1] E. Acerbi and G. Mingione, Gradient estimates for a class of parabolic systems, *Duke Math. J.* **136** (2007), no. 2, 285–320.
[2] C. Bennett and R. Sharpley, *Interpolation of Operators*, Pure and Applied Mathematics, Vol. **129**, Academic Press, Inc., Boston, MA, 1988.
[3] V. Bogelein, F. Duzaar and G. Mingione, Degenerate problems with irregular obstacles, *J. Reine Angew. Math.* **650** (2011), 107–160.
[4] S. Byun, D.K. Palagachev and L.G. Softova, Global gradient estimates in weighted lebesgue spaces for parabolic operators, http://arxiv.org/abs/1309.6199v2.

[5] S. Byun, D.K. Palagachev and L. Wang, Parabolic systems with measurable coefficients in Reifenberg domains, *Int. Math. Res. Not. IMRN* **2013** (2013), no. 13, 3053–3086.

[6] S. Byun and S. Ryu, Global estimates in Orlicz spaces for the gradient of solutions to parabolic systems, *Proc. Amer. Math. Soc.* **138** (2010), no. 2, 641–653.

[7] S. Byun and L. Wang, L^p-regularity for fourth order parabolic systems with measurable coefficients, *Math. Z.* **272** (2012), no. 1-2, 515–530.

[8] F. Duzaar, G. Mingione and K. Steffen, Parabolic systems with polynomial growth and regularity, *Mem. Amer. Math. Soc.* **214** (2011), no. 1005.

[9] H. Dong and D. Kim, Parabolic and elliptic systems in divergence form with variably partially BMO coefficients, *SIAM J. Math. Anal.* **43** (2011), no. 3, 1075–1098.

[10] E.B. Fabes and N.M Rivière, Singular integrals with mixed homogeneity, *Studia Math.* **27** (1966), 19–38.

[11] A. Fiorenza and M. Krbec, Indices of Orlicz spaces and some applications, *Comment. Math. Univ. Carolin.* **38** (1997), no. 3, 433–451.

[12] J. García-Cuerva and J. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*, North-Holland Mathematics Studies, Vol. **116**, North-Holland Publishing Co., Amsterdam, 1985.

[13] R. Kerman and A. Torchinsky, Integral inequalities with weights for the Hardy maximal function, *Studia Math.* **71** (1981/82), no. 3, 277–284.

[14] V. Kokilashvili and M. Krbec, *Weighted Inequalities in Lorentz and Orlicz Spaces*, World Scientific Publishing Co., Inc., River Edge, NJ, 1991.

[15] M.A. Krasnoselskii and Ja.B. Rutickii, *Convex Functions and Orlicz Spaces*, P. Noordhoff Ltd., Groningen, 1961.

[16] O.A. Ladyzhenskaya, V.A. Solonnikov and N.N. Uraltseva, *Linear and Quasilinear Equations of Parabolic Type*, Translations of Mathematical Monographs, Vol. **23**, American Mathematical Society, Providence, RI, 1968.

[17] B. Muckenhoupt, Weighted norm inequalities for the Hardy maximal function, *Trans. Amer. Math. Soc.* **165** (1972), 207–226.

[18] M.M. Rao and Z.D. Ren, *Theory of Orlicz spaces*, Monographs and Textbooks in Pure and Applied Mathematics, Vol. **146**, Marcel Dekker, Inc., New York, 1991.

[19] E. Reifenberg, Solutions of the Plateau problem for m-dimensional surfaces of varying topological type, *Acta Math.* **104** (1960), 1–92.

[20] E. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality and Oscillatory Integrals,* Princeton Mathematical Series, Vol. **43**, Princeton Univ. Press, Princeton, NJ, 1993.

[21] A. Torchinsky, *Real-Variable Methods in Harmonic Analysis,* Pure and Applied Mathematics, Vol. **123**, Academic Press, Inc., Orlando, FL, 1986.

[22] T. Toro, Doubling and flatness: geometry of measures, *Notices Amer. Math. Soc.* **44** (1997), no. 9, 1087–1094.