PUBLISHED IN HOUSTON J. MATH. 38 (3): 855-914, 2012

NEW GRADIENT ESTIMATES FOR PARABOLIC EQUATIONS

PAOLO BARONI AND JENS HABERMANN

ABSTRACT. We prove sharp Lorentz- and Morrey-space estimates for the gradient of solutions u to nonlinear parabolic equations of the type

$$u_t - \operatorname{div} a(z, Du) = g, \qquad \text{on } \Omega_T = \Omega \times (-T, 0),$$

where the vector field a is assumed to satisfy classical growth and ellipticity conditions and where the inhomogeneity g is only assumed to be integrable to some power $\gamma > 1$. In particular we investigate the case where γ stays below the exponent allowing for weak solutions $u \in L^2(-T, 0; W^{1,2}(\Omega))$.

1. INTRODUCTION AND RESULTS

In this paper we investigate regularity properties of solutions to Cauchy-Dirichlet problems

$$\begin{cases} u_t - \operatorname{div} a(z, Du) = g & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial_{\text{par}}\Omega_T. \end{cases} \tag{1.1}$$

We assume that $\Omega \subset \mathbb{R}^n$ is a bounded open subset, $n \geq 2$ and $T > 0$. Here, $\Omega_T := \Omega \times (-T, 0)$ denotes the usual parabolic cylinder. The vector-field $a\colon \Omega_T \times \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be a Carathéodory map which satisfies the following classical growth and monotonicity conditions:

$$\begin{cases} \nu|w_1 - w_2|^2 \leq \langle a(z, w_1) - a(z, w_2), w_1 - w_2\rangle \\ |a(z, w)| \leq L(1 + |w|) \end{cases} \tag{1.2}$$

for every choice of $z \in \Omega_T$, $w_1, w_2 \in \mathbb{R}^n$. In particular $z \mapsto a(z, w)$ is a measurable map for every $w \in \mathbb{R}^n$ and $w \mapsto a(z, w)$ is continuous for a.e. $z \in \Omega_T$. Unless otherwise stated the structure constants ν and L are assumed to fulfill

$$0 < \nu \leq 1 \leq L. \tag{1.3}$$

We will focus mainly on the situation $g \in L^\gamma(\Omega_T), \gamma > 1$ in which the right hand side does not necessarily belong to the Lebesgue space which allows to obtain existence of energy solutions $u \in L^2(-T, 0; W_0^{1,2}(\Omega))$ to problem (1.1). However, a by now classical approach towards "subdual" problems as mentioned above is to set up a suitable approximation scheme to obtain a unique so-called SOLA (Solution Obtained by Limits of Approximations, see Section 2.1) for which holds $u \in L^1(-T, 0; W_0^{1,1}(\Omega))$. Starting from such a unique solution, we are interested in finding optimal integrability estimates for solutions to equations of the type (1.1) depending on the regularity of the inhomogeneity. A by now classical result of Boccardo, Dall'Aglio, Gallouët and Orsina [9, Theorem 1.9] asserts the existence of a (unique) solution

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

to the Cauchy-Dirichlet problem (1.1) under the assumption (1.2), provided the datum g satisfies

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

Here $N := n+2$ denotes the homogeneous dimension. Moreover, the solution u belongs to $L^\sigma(\Omega_T)$ with σ given by $\sigma = \frac{N\gamma}{N-2\gamma}$. This result is optimal in the scale of Lebesgue spaces. However, one may ask for a more accurate scale to describe regularity of Du in dependence on the inhomogeneity g. Let us focus for a moment on **elliptic equations**. For those ones, Mingione presented in his paper [48] a non-linear potential theory version of the fundamental papers of Adams [4] and Adams & Lewis [5], giving optimal regularity results on the Morrey and also Lorentz-Morrey scale. Since we are dealing in this paper with equations with linear growth, we recall the results of [48] not in their whole generality but only for the special case $p = 2$. Mingione proved for solutions $u \in W_0^{1,1}(\Omega)$ of elliptic equations that

$$g \in L^{\gamma,\theta}(\Omega) \Longrightarrow Du \in L_{\mathrm{loc}}^{\frac{\theta\gamma}{\theta-\gamma},\theta}(\Omega), \quad 1 < \gamma \leq \frac{2\theta}{\theta+2}, \;\; 2 < \theta \leq n,$$

and where the definition of the (elliptic) Morrey spaces $L^{\gamma,\theta}(\Omega)$ can be adapted from (1.6), by replacing parabolic cylinders $C_R \subset \Omega_T$ of radius $R > 0$ by balls $B_R \subset \Omega$. Note at this stage that $L^{\gamma,n} \equiv L^\gamma$ and therefore the mentioned result covers the classical implication

$$g \in L^\gamma(\Omega) \Longrightarrow Du \in L_{\mathrm{loc}}^{\frac{n\gamma}{n-\gamma}}(\Omega), \qquad 1 < \gamma \leq \frac{2n}{n+2}, \tag{1.4}$$

going back to Talenti [56] and in the non-linear situation $p \neq 2$ to Boccardo & Gallouët [10]. In fact, the above mentioned implication is a special case of the more general result in [48] which provides estimates on the scale of Lorentz-Morrey spaces of the type

$$g \in L^\theta(\gamma, q) \Longrightarrow Du \in L^\theta\left(\tfrac{\theta\gamma}{\theta-\gamma}, \tfrac{\theta q}{\theta-\gamma}\right) \qquad \text{locally in } \Omega, \tag{1.5}$$

for exponents $1 < \gamma \leq \frac{2\theta}{\theta+2}$, $2 < \theta \leq n$ and $0 < q \leq \infty$. For the definition of the Lorentz-Morrey spaces we refer the reader to Section 3 and we mention again that the elliptic version of the spaces can be obtained by replacing in the parabolic definition the cylinders $C_R \subset \Omega_T$ by balls $B_R \subset \Omega$.

On the other hand, classical counter examples show that even in the linear case (1.4) fails for the borderline choice $\gamma = 1$. Indeed imposing some further $L\log L$-integrability on the inhomogeneity, the sharp implication

$$g \in L^{1,\theta}(\Omega) \cap L\log L(\Omega) \Longrightarrow Du \in L_{\mathrm{loc}}^{\frac{\theta}{\theta-1}}(\Omega),$$

holds true (see [10, 48]), where the space $L\log L(\Omega)$ is defined analogously to the parabolic one in (3.5).

The aim of this paper is to extend the above mentioned results to the situation of non-linear **parabolic equations** of the type (1.1) fulfilling structure conditions of the form (1.2) and therefore in particular to provide a non-linear parabolic analogue of the classical Theorems of Adams [4] and Adams & Lewis [5]. Our results in detail will be presented in the following section. As we already mentioned, we finally point out that the results presented in this paper hold for SOLAs, and therefore by writing of *weak solution to* (1.1) we will always mean the *solution obtained via the approximation procedure described in Section 2*. It is therefore natural to wonder whether these results can be extended to some other notion of solutions to (1.1). This becomes trivially true for the notions of solutions to measure data problems holding uniqueness in the case of L^1 data. In particular, in [51] a suitable definition of *renormalized solution* is given for nonlinear parabolic problems with measure data, and this definition provides uniqueness in the case of data in L^1. All our regularity results could therefore also be stated in terms of renormalized solution.

1.1. **Morrey space estimates.** We start by considering a Morrey-type condition of the form

$$R^{\theta-N}\int_{\mathcal{C}_R} |g|^\gamma \, dz \leq M^\gamma \quad \text{and} \quad \theta \in [2, N], \tag{1.6}$$

whenever $\mathcal{C}_R \subset \Omega_T$ is a parabolic cylinder with radius R. For the definition of parabolic cylinders see Section 2 below. Functions g satisfying (1.6) are said to belong to the Morrey-space $L^{\gamma,\theta}(\Omega_T)$ and one sets

$$\|g\|^{\gamma}_{L^{\gamma,\theta}(\Omega_T)} := \sup_{\mathcal{C}_R\subset\Omega_T} R^{\theta-N} \int_{\mathcal{C}_R} |g|^{\gamma}\, dz.$$

The range of exponents on which we will put our main focus in the sequel is

$$1 < \gamma \le \frac{2\theta}{\theta+2}\,, \qquad \text{and} \qquad 2 < \theta \le N\,. \tag{1.7}$$

Our first result is concerned with Morrey-space regularity for Du and it represents the parabolic counter part (for $p = 2$) of [48, Theorem 1]. The theorem is a special case of Theorem 6.1 which is the main theorem of our paper and is concerned with the more general Lorentz-Morrey space regularity, the parabolic extension of (1.5). Indeed, Theorem 1.1 follows from Theorem 6.1 by the special choice $q = \gamma$.

Theorem 1.1 (Non-linear parabolic Adams theorem)**.** *Under the assumptions (1.2) and $g \in L^{\gamma,\theta}(\Omega_T)$ with γ, θ as in (1.7) the solution $u \in L^1(-T, 0; W^{1,1}_0(\Omega))$ to the Cauchy-Dirichlet problem (1.1) is such that*

$$Du \in L^{\frac{\theta\gamma}{\theta-\gamma},\theta}_{\mathrm{loc}}(\Omega_T, \mathbb{R}^n).$$

Moreover, the quantitative local estimate

$$\|Du\|_{L^{\frac{\theta\gamma}{\theta-\gamma},\theta}(\mathcal{C}_{R/2})} \le c\, R^{\frac{\theta-\gamma}{\gamma}-N} \|1 + |Du|\|_{L^1(\mathcal{C}_R)} + c\, \|g\|_{L^{\gamma,\theta}(\mathcal{C}_R)}$$

holds for any parabolic cylinder $\mathcal{C}_R \subset \Omega_T$ with a constant $c = c(n, L, \nu, \gamma)$.

Here we note that the special choice $\theta = N$ in the above theorem – by the identity $L^{\frac{N\gamma}{N-\gamma},N} \equiv L^{\frac{N\gamma}{N-\gamma}}$ – gives back the classical result of Boccardo, Dall'Aglio, Gallouët and Orsina [9] on the Lebesgue scale. On the other hand, Theorem 1.1 fails in the borderline case $\gamma = 1$. Here, analogously to the elliptic case we have to impose some further $L\log L$ integrability on the inhomogeneity g and we obtain the following

Theorem 1.2 (Borderline parabolic Adams theorem)**.** *Under the assumptions (1.2) and $g \in L^{1,\theta}(\Omega_T) \cap L\log L(\Omega_T)$ with $2 \le \theta \le N$, the solution $u \in L^1(-T, 0;$ $W^{1,1}_0(\Omega))$ to the Cauchy-Dirichlet problem (1.1) is such that*

$$Du \in L^{\frac{\theta}{\theta-1}}_{\mathrm{loc}}(\Omega_T, \mathbb{R}^n).$$

Moreover, the quantitative local estimate

$$\begin{aligned}\left[\fint_{\mathcal{C}_{R/2}} |Du|^{\frac{\theta}{\theta-1}}\, dz\right]^{\frac{\theta-1}{\theta}} &\le c \fint_{\mathcal{C}_R} (1 + |Du|)\, dz \\ &\quad + c\, \|g\|^{\frac{1}{\theta}}_{L^{1,\theta}(\mathcal{C}_R)} \left[\fint_{\mathcal{C}_R} |g| \log\left(e + \frac{g}{\fint_{\mathcal{C}_R} |g(\tilde z)|\, d\tilde z}\right) dz\right]^{\frac{\theta-1}{\theta}},\end{aligned} \tag{1.8}$$

holds for any parabolic cylinder $\mathcal{C}_R \subset \Omega_T$ with a constant $c = c(n, L, \nu)$.

Indeed, also Theorem 1.2 can be seen as the particular case $\theta = N$ of the more general Theorem 6.5 which provides Morrey-space regularity of the following type

$$g \in L\log L^{\theta}(\Omega_T), \quad 2 \le \theta \le N \Longrightarrow Du \in L^{\frac{\theta}{\theta-1},\theta}_{\mathrm{loc}}(\Omega_T, \mathbb{R}^n).$$

Moreover we mention that the particular choice $\theta = 2$ is allowed in the above theorem, since we are in the case $\gamma = 1$. With this particular choice we reach the maximal regularity, that is $g \in L^{1,2}(\Omega_T) \cap L\log L(\Omega_T) \Longrightarrow Du \in L^2_{\mathrm{loc}}(\Omega_T, \mathbb{R}^n)$.

Remark 1.3. In the case we don't impose a $L\log L$ condition on g, still an estimate in Marcinkiewicz spaces holds true:

$$g \in L^{1,\theta}(\Omega_T), \quad 2 \le \theta \le N \Longrightarrow Du \in \mathcal{M}^{\frac{\theta}{\theta-1}}_{\mathrm{loc}}(\Omega_T, \mathbb{R}^n).$$

This is the parabolic analogue of [46, Theorem 1.8].

On the other hand, in the case of $\gamma > \frac{2\theta}{\theta+2}$ the techniques applied for Theorems 1.1 and 1.2 also provide Morrey-Gehring regularity in the following sense:

Theorem 1.4 (Morrey-Gehring regularity)**.** *Under the assumptions (1.2) and with $g \in L^{\gamma,\theta}(\Omega_T)$ for γ, θ satisfying $\frac{2\theta}{\theta+2} < \gamma$ and $2 \le \theta \le N$ the solution $u \in L^1(-T, 0; W^{1,1}_0(\Omega))$ to the Cauchy-Dirichlet problem (1.1) is such that*

$$Du \in L^{h,\theta}_{\mathrm{loc}}(\Omega_T, \mathbb{R}^n) \quad \textit{for some } h = h(n, L, \nu, \gamma, \theta) > 2. \tag{1.9}$$

Moreover, for a constant $c = c(n, L, \nu, \gamma)$ the quantitative local estimate

$$\|Du\|_{L^{h,\theta}(\mathcal{C}_{R/2})} \le c\, R^{\frac{\theta}{h}-N} \|1 + |Du|\|_{L^1(\mathcal{C}_R)} + c\, \|g\|_{L^{\gamma,\theta}(\mathcal{C}_R)} \tag{1.10}$$

holds for any parabolic cylinder $\mathcal{C}_R \subset \Omega_T$ with radius $R \le 1$.

1.2. **Lorentz-Morrey space regularity.** As we have already mentioned above, the results of Theorems 1.1 to 1.4 are particular cases of more general results in Lorentz-Morrey spaces. For the further discussion in this section let us briefly give the definitions of the involved spaces. A more detailed discussion on these spaces involving also basic properties and embeddings we refer the reader to Section 3. Letting $\Omega_T := \Omega \times (-T, 0)$ be the space time cylinder, where $\Omega \subset \mathbb{R}^n$ ($n \ge 2$) denotes a bounded open subset, a measurable map $g : \Omega_T \to \mathbb{R}^k$ is said to belong to the **Lorentz space** $L(p, q)(\Omega_T, \mathbb{R}^k)$, with $1 \le p < \infty$ and $0 < q < \infty$, iff

$$\|g\|^q_{L(p,q)(\Omega_T,\mathbb{R}^k)} := p \int_0^\infty \Big(\lambda^p \big|\{z \in \Omega_T : |g(z)| > \lambda\}\big|\Big)^{\frac{q}{p}} \frac{d\lambda}{\lambda} < \infty.$$

Here we refer also to Section 3 for the definition in the case $q = \infty$. The **parabolic Lorentz-Morrey spaces** are then defined in the following way: We say that a measurable function $g : \Omega_T \to \mathbb{R}^k$ belongs to $L^\theta(p, q)(\Omega_T, \mathbb{R}^k)$ for $1 \le p < \infty$, $0 < q < \infty$ and $\theta \in [0, N]$, iff

$$\|g\|_{L^\theta(p,q)(\Omega_T,\mathbb{R}^k)} := \sup_{C_\varrho \subset \Omega_T} \varrho^{\frac{\theta-N}{p}} \|g\|_{L(p,q)(C_\varrho,\mathbb{R}^k)} < \infty.$$

The main theorem of our paper – which contains also Theorem 1.1 for a particular choice of the parameters – is Theorem 6.1 and asserts the following implication for solutions $u \in L^1(-T, 0; W^{1,1}(\Omega))$ of the Cauchy-Dirichlet problem (1.1) under the structural assumptions (1.2) and for exponents γ, θ as in (1.7):

$$g \in L^\theta(\gamma, q)(\Omega_T), \quad 0 < q \le \infty \Longrightarrow |Du| \in L^\theta\Big(\tfrac{\theta\gamma}{\theta-\gamma}, \tfrac{\theta q}{\theta-\gamma}\Big) \quad \text{locally in } \Omega_T.$$

For the special choice $q = \gamma$, having in mind that $L^\theta(p, p)(\Omega_T) \equiv L^{p,\theta}(\Omega_T)$, we obtain the statement of Theorem 1.1. Moreover, in the borderline case $\theta = N$, we arrive at the following sharp estimate in Lorentz spaces, which is the content of Theorem 6.6: For exponents $1 < \gamma \le \frac{2N}{N+2}$ and $0 < q \le \infty$ the following implication holds for the solution $u \in L^1(-T, 0; W^{1,1}_0(\Omega))$ of the problem (1.1) under the condition (1.2):

$$g \in L(\gamma, q)(\Omega_T) \Longrightarrow |Du| \in L\Big(\tfrac{N\gamma}{N-\gamma}, q\Big) \quad \text{locally in } \Omega_T.$$

1.3. **Equations with more regular coefficients.** We would now like to focus on the situation where the vector-field $a(z,w)$ in (1.1) satisfies stronger assumptions, especially more regularity with respect to the variable z. We therefore consider weak solutions $u \in L^1(-T,0;W^{1,1}(\Omega))$ to the equation (1.1) under either one of the following two settings:

- The vector-field $a(z,w) = a(x,t,w)$ satisfies the structure assumptions
$$\begin{cases} |a(x,t,w)| + (1+|w|)|\partial_w a(x,t,w)| \le L(1+|w|) \\ \langle \partial_w a(x,t,w)\tilde{w}, \tilde{w}\rangle \ge \nu |\tilde{w}|^2 \\ |a(x,t,w) - a(x_0,t,w)| \le L\omega\left(|x-x_0|\right)\left(1+|w|\right), \end{cases} \tag{1.11}$$
for any choice of $x, x_0 \in \Omega, t \in (-T,0)$ and $w, \tilde{w} \in \mathbb{R}^n$. Moreover we assume that the structure constants ν, L satisfy (1.3). Finally, we assume that $\omega : [0,\infty) \to [0,\infty)$ is a bounded, concave modulus of continuity satisfying $\lim_{\varrho\downarrow 0}\omega(\varrho) = 0 = \omega(0)$ and $\omega \le 2$ on $[0,\infty)$, and that $(x,t,w) \mapsto a(x,t,w)$ and $(x,t,w) \mapsto \partial_w a(x,t,w)$ are Carathéodory maps.
- The vector-field has the structure $a(z,w) = a(x,t,w) := c(x)\bar{a}(t,w)$, where $\bar{a}\colon (-T,0)\times\mathbb{R}^n \to \mathbb{R}^n$ satisfies the growth and ellipticity conditions:
$$\begin{cases} |\bar{a}(t,w)| + (1+|w|)|\partial_w \bar{a}(t,w)| \le L(1+|w|) \\ \langle \partial_w \bar{a}(t,w)\tilde{w}, \tilde{w}\rangle \ge \nu |\tilde{w}|^2\,, \end{cases} \tag{1.12}$$
for any choice of $t \in (-T,0)$ and $w, \tilde{w} \in \mathbb{R}^n$. For the structure constants ν, L we assume (1.3) and we also assume that $(t,w) \mapsto \bar{a}(t,w)$ and $(t,w) \mapsto \partial_w \bar{a}(t,w)$ are Carathéodory maps. For the function $c : \Omega \to \mathbb{R}$ we shall assume that
$$0 < \nu \le c(x) \le L < \infty\,, \qquad \forall\, x \in \Omega\,, \tag{1.13}$$
and VMO-regularity, which means that the function c satisfies
$$\lim_{R\downarrow 0}\omega(R) = 0\,, \qquad \text{where} \qquad \omega(R) := \sup_{\substack{B_\varrho \Subset \Omega \\ 0<\varrho\le R}} \fint_{B_\varrho} |c(x) - (c)_{B_\varrho}|\,dx\,. \tag{1.14}$$

For solutions $u \in L^1(-T,0;W_0^{1,1}(\Omega))$ to the Cauchy-Dirichlet problem (1.1) under one of the above mentioned settings, i.e. either the structure assumptions (1.11) or the conditions (1.12) to (1.14) the implication
$$g \in L^\theta(\gamma,q)(\Omega_T) \Longrightarrow |Du| \in L^\theta\left(\tfrac{\theta\gamma}{\theta-\gamma}, \tfrac{\theta q}{\theta-\gamma}\right) \quad \text{locally in } \ \Omega_T,$$
holds true for all $0 < q \le \infty$ and all pairs of exponents (θ,γ) satisfying the condition
$$1 < \gamma < \theta \le N. \tag{1.15}$$
The concrete statement is the content of Theorem 6.7 and the proof can be performed using the same technique as in the case of Theorems 1.1, 6.1, respectively but exploiting the stronger estimates, proved by Duzaar, Mingione & Steffen [28] for solutions to homogeneous equations, fulfilling the structure conditions (1.11) or (1.12) to (1.14).

1.4. **Integrability of u.** The technique of establishing Calderón-Zygmund type estimates for the maximal function for the spatial gradient Du of the solution leading to the statements of Theorems 1.1 and 6.1 can also be applied on the level of the solution u itself and provides – under certain modifications – also Lorentz-Morrey space estimates for the solution u. More precisely, in Theorem 7.1 we prove for solutions $u \in L^1(-T,0;W_0^{1,1}(\Omega))$ to the problem (1.1) under the conditions (1.2) the implication
$$g \in L^\theta(\gamma,q)(\Omega_T) \Longrightarrow u \in L^\theta\left(\tfrac{\theta\gamma}{\theta-2\gamma}, \tfrac{\theta q}{\theta-2\gamma}\right) \quad \text{locally in } \Omega_T,$$
for all $0 < q \le \infty$ and all pairs of (θ,γ) satisfying the conditions
$$1 < \gamma < \tfrac{\theta}{2}, \qquad 2 < \theta \le N. \tag{1.16}$$

Also here, we may establish a "borderline" estimate in Lorentz spaces, coming up in the special case $\theta = N$, in the sense that

$$g \in L(\gamma, q)(\Omega_T) \Longrightarrow u \in L\Big(\tfrac{N\gamma}{N-2\gamma}, q\Big) \quad \text{locally in } \Omega_T,$$

for any $0 < q \leq \infty$ and $1 < \gamma < \frac{N}{2}$. On the other hand, the borderline case $\gamma = \theta/2, q = \infty$ provides the following BMO-estimate for the solution $u \in L^1(-T, 0; W_0^{1,1}(\Omega))$ of (1.1) under the structure assumptions (1.2):

$$g \in \mathcal{M}^{\theta/2,\theta}(\Omega_T), \quad 2 < \theta \leq N \Longrightarrow u \in \mathrm{BMO}_{\mathrm{loc}}(\Omega_T).$$

Here $\mathcal{M}^{\theta/2,\theta}(\Omega_T) \equiv L^\theta(\theta/2, \infty)(\Omega_T)$ denotes the Marcinkiewicz-Morrey space and the above statement is the content of Theorem 7.4.

1.5. **Some notes about the techniques of the proofs.** The proof of our theorems is based on the method developed in [48] for elliptic equations, which we carry over to the framework of parabolic equations. The key point to the proof of Theorem 6.1 is an estimate which allows to control the level set of the Hardy-Littlewood maximal function of the spatial gradient $|Du|$ locally by the level sets of a suitable parabolic Riesz potential operator. More precisely (see (6.20) together with Lemma 4.3 for the exact estimate) we establish for some exponent $\chi > 1$ an estimate of the type

$$\big|\{M(|Du|) \geq T\lambda\}\big| \lesssim T^{-2\chi}\big|\{M(|Du|) \geq \lambda\}\big| + c(T)\big|\{I_1(|g|) \geq \lambda\}\big|, \tag{1.17}$$

where $M(|Du|)$ denotes the maximal operator of $|Du|$, λ is a number large enough and $T \gg 1$ is a constant. Here, the parabolic Riesz potential operator for $\beta \in (0, N)$ is defined as

$$I_\beta(|g|)(z) := \int_{\mathbb{R}^{n+1}} \frac{|g(\tilde{z})|}{d_{\mathrm{par}}(z, \tilde{z})^{N-\beta}}\, d\tilde{z}, \quad z \in \mathbb{R}^{n+1}.$$

and d_{par} denotes the parabolic metric (see (4.3)). The precise definitions of the other involved quantities can be found in Section 4.1. All the Lorentz- and Lorentz-Morrey estimates and also the borderline cases can then be derived with the help of (1.17). In order to prove the decay estimate (1.17), we apply the classical Calderón-Zygmund covering Lemma in the parabolic setting to suitable level sets of the maximal function. Indeed, to verify the conditions necessary to apply the Calderón-Zygmund Lemma in our setting, we take use of a comparison strategy to the solution v to an associated homogeneous problem. For the solution to homogeneous problems, well known Hölder continuity and higher integrability results coming up from the De Giorgi-Nash-Moser theory provide suitable reference estimates. Since u is merely found to be of the class $L^1(-T, 0; W_0^{1,1}(\Omega))$, comparison estimates have to be established on the level of L^1-norms, involving not more than the L^1-norm of the inhomogeneity. The proof of such a comparison estimate is established via certain truncation techniques, which already go back to the contributions of Boccardo & Gallouët [10] and are used also in the more recent papers [26, 7].

In the case of general equations, as (1.1), fulfilling the structure conditions (1.2), the decay estimate of the type (1.17) can be established for some fixed exponent $\chi > 1$, depending on the data of the equation. This, in turn leads to restrictions on the range of exponents which are allowed in the Lorentz and Lorentz-Morrey estimates, and we come up with the desired estimates for exponents fulfilling (1.7). On the other hand, assuming more regularity on the data, especially having certain continuity conditions on the vector-field $a(x, t, w)$ with respect to the space variable x, integrability and Hölder continuity estimates for solutions to homogeneous equations have been established by Duzaar, Mingione & Steffen in [28] in a stronger form (in particular there holds Hölder continuity to any exponent $\alpha \in (0, 1)$) and therefore the decay estimate (1.17) can be found to hold true for any exponent $\chi > 1$. As a consequence, we derive Lorentz and Lorentz-Morrey estimates in this case for the full range of exponents, as in (1.15).

Concerning the level of the solution u itself, the decay estimate (1.17) can be substituted by an estimate of the form

$$\big|\{M(u) \geq T\lambda\}\big| \lesssim T^{-2\chi}\big|\{M(u) \geq \lambda\}\big| + c(T)\big|\{I_2(|g|) \geq \lambda\}\big|,$$

which holds for any λ large enough and $T \gg 1$, and for any exponent $\chi > 1$. Here, instead of the Riesz potential $I_1(|g|)$, we have the potential $I_2(|g|)$ involved on the right hand side. This finally allows to establish the desired Lorentz-Morrey estimate for u for the range of exponents declared in (1.16).

1.6. **Other recent developments on measure data problems.** Let us finally focus on further developments which have recently been made concerning regularity for equations with right hand sides below the duality exponent. In very recent contributions, Duzaar & Mingione [26] established for elliptic and also parabolic equations with measure data right hand side pointwise estimates for the spatial gradient Du of solutions. The results extend the well known potential estimates by Kilpeläinen & Malý [40] for solutions to the gradient of solutions. The authors in [26] prove their results under slightly stronger assumptions on the vector field $a(t, x, w)$ and they involve a local version of the Riesz potential of the measure μ on the right hand side. In this way, for equations satisfying these slightly stronger structural assumptions, our results can be recovered naturally from their pointwise results. See also [43] for an overview on recent potential results for parabolic equations.

2. Preliminaries, Notation

Throughout the paper we denote by c a general constant that may vary from line to line. In general we shall have $c \geq 1$. Peculiar dependencies on parameters will be emphasized in parentheses when needed. Special constants will be denoted by $\tilde{c}$, c_0, $c_1 \dots$. Points in Euclidean n-space $\mathbb{R}^n$ are denoted by $x = (x_1, \dots, x_n)$, while points in $\mathbb{R}^{n+1}$ are denoted by $z = (x, t) \in \mathbb{R}^n \times \mathbb{R}$. With $x_0 \in \mathbb{R}^n$ we denote by $B_R(x_0) = B(x_0, R) := \{x \in \mathbb{R}^n : |x - x_0| < R\}$ respectively $Q_R(x_0) = Q(x_0, R) := \{x \in \mathbb{R}^n : \max_i |x_i - x_{0,i}| < R\}$ the open ball and cube, respectively, with center x_0 and radius R, respectively sidelength $2R$ in the case of the cube. We shall often use the short hand notation $B_R = B(x_0, R)$ and $Q_R = Q(x_0, R)$, when no ambiguity will arise and all the balls/cubes considered have the same center. With $z_0 = (x_0, t_0) \in \mathbb{R}^{n+1}$, we denote by

$$\mathcal{C}_R(z_0) = \mathcal{C}(z_0, R) := B(x_0, R) \times (t_0 - R^2, t_0 + R^2)$$

the open (symmetric) parabolic cylinder with center z_0 having a ball with center x_0 of radius R as horizontal slice and height $2R^2$, while

$$\mathcal{Q}_R(z_0) = \mathcal{Q}(z_0, R) := Q(x_0, R) \times (t_0 - R^2, t_0 + R^2)$$

denotes the open parabolic cylinder with center z_0 having a cube $Q(x_0, R)$ with center x_0 of sidelength $2R$ as horizontal slice and height $2R^2$. Moreover, with B and Q being balls and cubes respectively, by γB, γQ we shall denote the concentric balls and cubes with radius/sidelength scaled by a non-negative factor $\gamma > 0$. Finally, with $\mathcal{C}$ and $\mathcal{Q}$ being parabolic cylinders with horizontal slice being a ball or cube respectively we shall denote by $\gamma\mathcal{C}$ and $\gamma\mathcal{Q}$ the concentric parabolic cylinders scaled by the factor $\gamma > 0$; i.e. $\gamma\mathcal{C}(z_0, R) = B(x_0, \gamma R) \times (t_0 - (\gamma R)^2, t_0 + (\gamma R)^2)$ and $\gamma\mathcal{Q}(z_0, R) = Q(x_0, \gamma R) \times (t_0 - (\gamma R)^2, t_0 + (\gamma R)^2)$. Throughout the paper all the cubes considered will have sides parallel to the coordinate axes in $\mathbb{R}^n$ and will have positive sidelength.

For a measurable set $A \subset \mathbb{R}^k$ with finite positive measure and an integrable function $g\colon A \to \mathbb{R}^\ell$ the average of g over A is

$$(g)_A = \fint_A g(x)\,dx := \frac{1}{|A|}\int_A g(x)\,dx\,.$$

2.1. **Regularized problems, and solvability of (1.1).** By a solution u to (1.1) we understand a function $u \in L^1(-T,0;W_0^{1,1}(\Omega))$ solving $(1.1)_1$ in the distributional sense

$$\int_{\Omega_T} \big(u\varphi_t - a(z,Du)D\varphi\big)\,dz = -\int_{\Omega_T} g\varphi\,dz\,, \qquad \forall\,\varphi \in C_0^\infty(\Omega_T)\,. \tag{2.1}$$

The existence of such a solution is obtained in [9, 10] by an approximation argument, which is by now standard in the theory of measure data problems. For convenience of the reader we briefly sketch the strategy: One considers a sequence $g_k \in L^\infty(\Omega_T)$, $k \in \mathbb{N}$, such that $g_k \to g$ in $L^1(\Omega_T)$ when $k \to \infty$. Then by standard monotonicity arguments one finds, for each fixed k, a unique solution $u_k \in C^0([-T,0];L^2(\Omega)) \cap L^2(-T,0;W_0^{1,2}(\Omega))$ to the Cauchy-Dirichlet problem

$$\begin{cases} (u_k)_t - \operatorname{div} a(z,Du_k) = g_k & \text{in } \Omega_T,\\ u_k = 0 & \text{on } \partial_{\mathrm{par}}\Omega_T. \end{cases} \tag{2.2}$$

The arguments from [9, 10] yield the existence of a solution $u \in L^1(-T,0;W_0^{1,1}(\Omega))$ such that for a not relabeled subsequence

$$u_k \to u \quad \text{strongly in } L^1(-T,0;W_0^{1,1}(\Omega)) \text{ and a.e.}$$

and (1.1) is solved in the distributional sense (2.1). For the rest of the paper we understand by $\{u_k\} \subset C^0([-T,0];L^2(\Omega)) \cap L^2(-T,0;W_0^{1,2}(\Omega))$ the sequence obtained by solving (2.2) for the specific choice

$$g_k(z) := \max\{-k, \min\{g(z),k\}\}, \qquad k \in \mathbb{N}. \tag{2.3}$$

2.2. **Calderón-Zygmund coverings, inner and outer parabolic cylinders.** Let $\mathcal{Q}_0 = \mathcal{Q}(z_0,R) = Q(x_0,R)\times(t_0-R^2,t_0+R^2)$ be a parabolic cylinder in $\mathbb{R}^{n+1}$ with horizontal cross section being a cube. By $\mathcal{D}(\mathcal{Q}_0)$ we shall denote the class of all dyadic parabolic cylinders obtained from $\mathcal{Q}_0$ by a finite number of dyadic subdivisions. The construction of a dyadic subdivision is as follows: If $\mathcal{Q}_0$ is as above then we subdivide Q_0 into 2^n congruent sub-cubes Q' having sides parallel to Q_0 and $(t_0 - R^2, t_0 + R^2)$ into four disjoint intervals I' of equal length $R^2/2$. Then, the set of all parabolic sub-cylinders obtained by this dyadic subdivision consist of all cylinders of the form $Q' \times I'$. The total number of sub-cylinders obtained from a parabolic cylinder $\mathcal{Q}$ by one dyadic subdivision is 2^N, $N = n+2$. We note that $\mathcal{Q}_0 \notin \mathcal{D}(\mathcal{Q}_0)$. For later use we mention a few simple facts of the class $\mathcal{D}(\mathcal{Q}_0)$: First, if $\mathcal{Q}_1, \mathcal{Q}_2 \in \mathcal{D}(\mathcal{Q}_0)$ then either $\mathcal{Q}_1 \cap \mathcal{Q}_2 = \emptyset$, or one of the parabolic cylinders contains the other one, i.e. $\mathcal{Q}_1 \subset \mathcal{Q}_2$ or $\mathcal{Q}_2 \subset \mathcal{Q}_1$. We shall denote $\widetilde{\mathcal{Q}} \in \mathcal{D}(\mathcal{Q}_0)$ **the predecessor** of $\mathcal{Q}$ if $\mathcal{Q}$ has been obtained by exactly one dyadic subdivision from the parabolic cylinder $\tilde{\mathcal{Q}}$. The following is a Calderón-Zygmund-Krylov-Safanov type covering lemma in the parabolic setting; for the elliptic (classical) version we refer to [14].

Proposition 2.1. *Let $\mathcal{Q}_0 \subset \mathbb{R}^{n+1}$ be a parabolic cylinder. Assume that $X \subset Y \subset \mathcal{Q}_0$ are measurable sets such that the following properties (i) and (ii) hold:*

(i) there exist $\delta > 0$ such that $|X| < \delta|\mathcal{Q}_0|$;

(ii) if $\mathcal{Q} \in \mathcal{D}(\mathcal{Q}_0)$, then $|X \cap \mathcal{Q}| > \delta|\mathcal{Q}|$ implies $\widetilde{\mathcal{Q}} \subset Y$, where $\widetilde{\mathcal{Q}}$ denotes the predecessor of $\mathcal{Q}$.

Then there holds $|X| < \delta|Y|$.

The proof of the preceding proposition can be inferred using arguments from [14]. For convenience of the reader we give the simple adaptation to our parabolic set up. The starting point is the following version of the classical Calderón-Zygmund type covering lemma.

Lemma 2.2. *Let $\mathcal{Q}_0 \subset \mathbb{R}^{n+1}$ be a parabolic cylinder and X a measurable subset of $\mathcal{Q}_0$ satisfying*

$$0 < |X| < \delta|\mathcal{Q}_0|$$

for some $0 < \delta < 1$. Then there exists a sequence $(\mathcal{Q}_i)_{i\in\mathbb{N}}$ of disjoint dyadic sub-cylinders of $\mathcal{Q}_0$ such that there holds:

(i) $\big|X \setminus \bigcup_{i=1}^{\infty} \mathcal{Q}_i\big| = 0$,
(ii) $|X \cap \mathcal{Q}_i| \geq \delta|\mathcal{Q}_i|$ *and*
(iii) $|X \cap \widetilde{\mathcal{Q}}| < \delta|\widetilde{\mathcal{Q}}|$ *if* $\widetilde{\mathcal{Q}} \in \mathcal{D}(\mathcal{Q}_0)$ *and* $\mathcal{Q}_i \subsetneqq \widetilde{\mathcal{Q}}$.

Proof. We divide $\mathcal{Q}_0$ into 2^N dyadic sub-cylinders $\mathcal{Q}_1^{(j)}$ and select those satisfying

$$\big|X \cap \mathcal{Q}_1^{(j)}\big| \geq \delta\big|\mathcal{Q}_1^{(j)}\big|.$$

Now, we take those cylinders that were not chosen, divide each of them again into 2^N dyadic sub-cylinders and repeat the selection argument from above. Proceeding iteratively in this way we obtain a sequence of disjoint dyadic cylinders $\mathcal{Q}_i \in \mathcal{D}(\mathcal{Q}_0)$, $i \in \mathbb{N}$. By construction each of these cylinders satisfies (ii) and (iii). For $z \in \mathcal{Q}_0 \setminus \bigcup_{i=1}^{\infty} \mathcal{Q}_i$ we have a sequence of dyadic cylinders $\mathcal{P}_k$ with $|\mathcal{P}_k| \to 0$ as $k \to \infty$, each of them containing z, such that

$$|\mathcal{P}_k \cap X| < \delta|\mathcal{P}_k|,$$

or equivalently

$$⨍_{\mathcal{P}_k} \chi_X(\tilde{z})\, d\tilde{z} = \frac{|\mathcal{P}_k \cap X|}{|\mathcal{P}_k|} < \delta < 1.$$

By Lebesgue's differentiation theorem the left-hand side of the preceding inequality converges to $\chi_X(z)$ for a.e. z as $k \to \infty$, and therefore we have $z \in \mathcal{Q}_0 \setminus X$ for a.e. $z \in \mathcal{Q}_0 \setminus \bigcup_{i=1}^{\infty} \mathcal{Q}_i$. Hence $\big|X \setminus \bigcup_{i=1}^{\infty} \mathcal{Q}_i\big| = 0$, proving finally (i). □

Proof of Proposition 2.1. We apply Lemma 2.2 to have a sequence of disjoint dyadic cylinders $(\mathcal{Q}_i)_{i\in\mathbb{N}}$ covering almost all of X. By (ii) of Lemma 2.2 we have $|X \cap \mathcal{Q}_i| > \delta|\mathcal{Q}_i|$; therefore by assumption (ii) the predecessor $\widetilde{\mathcal{Q}}_i$ of $\mathcal{Q}_i$ is contained in Y. Now, from the sequence of predecessors $(\widetilde{\mathcal{Q}}_i)_{i\in\mathbb{N}}$ we can extract a sub-covering $(\widetilde{\mathcal{Q}}_i)_{i\in\mathfrak{K}}$ of X, where the $\widetilde{\mathcal{Q}}_i$ are pairwise disjoint and $\mathfrak{K} \subset \mathbb{N}$. Then, using Lemma 2.2, (iii), the fact that $\widetilde{\mathcal{Q}}_i \subset Y$, as well as the disjointness of the $\widetilde{\mathcal{Q}}_i$ for $i \in \mathfrak{K}$ we obtain

$$|X| \leq \sum_{i\in\mathfrak{K}} |X \cap \widetilde{\mathcal{Q}}_i| < \delta \sum_{i\in\mathfrak{K}} |\widetilde{\mathcal{Q}}_i| \leq \delta|Y|,$$

proving the claim of Proposition 2.1. □

For a given ball $B \subset \mathbb{R}^n$ we denote by $Q_{\mathrm{inn}}(B)$ and $Q_{\mathrm{out}}(B)$ the largest and the smallest cubes with sides parallel to the coordinate axes concentric to B contained in B or containing B, respectively; i.e. if $B = B(x_0, R)$ we have $Q_{\mathrm{inn}}(B) = Q(x_0, R/\sqrt{n})$ and $Q_{\mathrm{out}}(B) = Q(x_0, R)$. These cubes we shall call inner and outer cubes. Moreover, for a given parabolic cylinder $\mathcal{C} = B(x_0, R) \times (t_0 - R^2, t_0 + R^2) \subset \mathbb{R}^{n+1}$ we will denote by $\mathcal{Q}_{\mathrm{out}}(\mathcal{C})$ the smallest parabolic cylinder with horizontal cross section a cube with sides parallel to the coordinate axes containing $\mathcal{C}$, i.e. $\mathcal{Q}(x_0, R) \times (t_0 - R^2, t_0 + R^2)$. Similarly, the largest parabolic cylinder with cross section a cube contained in $\mathcal{C}$ is denoted by $\mathcal{Q}_{\mathrm{inn}}$ and given by $Q(x_0, R/\sqrt{n}) \times (t_0 - (R/\sqrt{n})^2, t_0 + (R/\sqrt{n})^2)$. Note that due to the parabolic scaling we have to decrease the time interval in the case of $\mathcal{Q}_{\mathrm{inn}}(\mathcal{C})$. Without abuse of confusion we will call $\mathcal{Q}_{\mathrm{inn}}(\mathcal{C})$ and $\mathcal{Q}_{\mathrm{out}}(\mathcal{C})$ **inner** and **outer parabolic cylinder** (associated to $\mathcal{C}$). Finally, we mention the following classical iteration lemma:

Lemma 2.3. *Let $\varphi\colon [0, R_0] \to [0, \infty)$ be a non-decreasing function such that*

$$\varphi(\varrho) \leq A\left[\left(\frac{\varrho}{R}\right)^{\delta_0} + \varepsilon\right]\varphi(R) + \mathcal{B}R^{\delta_1} \qquad \textit{for every } 0 < \varrho \leq R \leq R_0, \tag{2.4}$$

where $A, \mathcal{B} \geq 0$ *and* $0 < \delta_1 < \delta_0$. *Then there exist* $\varepsilon_0 = \varepsilon_0(\delta_0, \delta_1, A) > 0$ *and* $c_1 = c_1(\delta_0, \delta_1, A)$ *such that whenever (2.4) holds for some* $0 < \varepsilon \leq \varepsilon_0$ *then*

$$\varphi(\varrho) \leq c_1 \left[\left(\frac{\varrho}{R} \right)^{\delta_1} \varphi(R) + \mathcal{B}\varrho^{\delta_1} \right] \qquad \text{for every } 0 < \varrho \leq R \leq R_0.$$

3. Function Spaces

3.1. **Parabolic spaces measuring size.** Throughout this section Ω denotes a bounded open subset in $\mathbb{R}^n$ and $T > 0$. By Ω_T we denote the space time cylinder $\Omega \times (-T, 0)$. A measurable map $g\colon \Omega_T \to \mathbb{R}^k$ is said to belong to the **Lorentz-space** $L(p,q)(\Omega_T, \mathbb{R}^k)$ with $1 \leq p < \infty$ and $0 < q \leq \infty$ iff

$$\|g\|^q_{L(p,q)(\Omega_T,\mathbb{R}^k)} := p \int_0^\infty \Big(\lambda^p |\{z \in \Omega_T : |g(z)| > \lambda\}| \Big)^{\frac{q}{p}} \frac{d\lambda}{\lambda} < \infty, \tag{3.1}$$

when $q < \infty$, while for $q = \infty$ it is imposed that

$$\sup_{\lambda > 0} \lambda^p |\{z \in \Omega_T : |g(z)| > \lambda\}| =: \|g\|^p_{\mathcal{M}^p(\Omega_T,\mathbb{R}^k)} < \infty \,. \tag{3.2}$$

The latter is the so called **Marcinkiewicz-**, or **weak-L^p-space**. Since we always assume Ω to have finite measure the spaces $L(p,q)(\Omega_T, \mathbb{R}^k)$ decrease in the first parameter p, which means that for $1 \leq \tilde{p} < p \leq \infty$ and $0 < q \leq \infty$ we have a continuous embedding $L(p,q)(\Omega_T) \hookrightarrow L(\tilde{p},q)(\Omega_T)$ with the estimate $\|g\|_{L(\tilde{p},q)(\Omega_T)} \leq |\Omega_T|^{\frac{1}{\tilde{p}} - \frac{1}{p}} \|g\|_{L(p,q)(\Omega_T)}$. On the other hand the Lorentz-spaces increase in the second parameter q, i.e. we have for $0 < q < \tilde{q} \leq \infty$ the continuous embedding $L(p,q)(\Omega_T) \hookrightarrow L(p,\tilde{q})(\Omega_T)$ with the estimate $\|g\|_{L(p,\tilde{q})(\Omega_T)} \leq c(p,q,\tilde{q}) \|g\|_{L(p,q)(\Omega_T)}$.

The so-called **parabolic Lorentz-Morrey-spaces** are obtained by coupling definition (3.1) with a density condition in the following sense: A measurable map $g\colon \Omega_T \to \mathbb{R}^k$ belongs to $L^\theta(p,q)(\Omega_T, \mathbb{R}^k)$ for $1 \leq p < \infty$, $0 < q < \infty$ and $\theta \in [0, N]$ iff

$$\begin{aligned}\|g\|_{L^\theta(p,q)(\Omega_T,\mathbb{R}^k)} &:= \sup_{\mathcal{C}_\varrho \subset \Omega_T} \varrho^{\frac{\theta - N}{p}} \|g\|_{L(p,q)(\mathcal{C}_\varrho,\mathbb{R}^k)} \\ &= \sup_{\mathcal{C}_\varrho \subset \Omega_T} \left[p \int_0^\infty \Big(\lambda^p \varrho^{\theta - N} |\{z \in \mathcal{C}_\varrho : |g(z)| > \lambda\}| \Big)^{\frac{q}{p}} \frac{d\lambda}{\lambda} \right]^{\frac{1}{q}} < \infty \,,\end{aligned} \tag{3.3}$$

while $g \in L^\theta(p,\infty)(\Omega_T, \mathbb{R}^k) = \mathcal{M}^{p,\theta}(\Omega_T, \mathbb{R}^k)$ iff

$$\begin{aligned}\|g\|_{\mathcal{M}^{p,\theta}(\Omega_T,\mathbb{R}^k)} &:= \sup_{\mathcal{C}_\varrho \subset \Omega_T} \varrho^{\frac{\theta - N}{p}} \|g\|_{\mathcal{M}^p(\mathcal{C}_\varrho,\mathbb{R}^k)} \\ &= \sup_{\mathcal{C}_\varrho \subset \Omega_T} \varrho^{\frac{\theta - N}{p}} \sup_{\lambda > 0} \Big(\lambda^p |\{z \in \mathcal{C}_\varrho : |g(z)| > \lambda\}| \Big)^{\frac{1}{p}} < \infty \,.\end{aligned} \tag{3.4}$$

Note that the supremum is taken over all parabolic cylinders $\mathcal{C}_\varrho = \mathcal{C}(z_0, \varrho)$ contained in Ω_T.

Remark 3.1. By Fubini's Theorem we have

$$\|g\|^p_{L^p(\Omega_T)} = p \int_0^\infty \lambda^p |\{z \in \Omega_T : |g(z)| > \lambda\}| \frac{d\lambda}{\lambda} = \|g\|^p_{L(p,p)(\Omega_T)} \,,$$

so that $L^p(\Omega_T) = L(p,p)(\Omega_T)$. As an immediate consequence we also have $L^{p,\theta}(\Omega_T) = L^\theta(p,p)(\Omega_T)$ with $\|g\|_{L^{p,\theta}(\Omega_T)} = \|g\|_{L^\theta(p,p)(\Omega_T)}$.

A measurable map g defined on Ω_T belongs to the space $L \log L(\Omega_T)$ iff

$$\|g\|_{L \log L(\Omega_T)} := \inf \left\{ \lambda > 0 : \fint_{\Omega_T} \left| \frac{g}{\lambda} \right| \log \left(e + \left| \frac{g}{\lambda} \right| \right) dz \leq 1 \right\} < \infty \,. \tag{3.5}$$

Note that we have incorporated in the preceding definition a dependence on the measure $|\Omega_T|$, by considering an averaged integral in (3.5). The reason for this will become clear in few lines, when we introduce a Morrey-type variant of the $L\log L$-spaces. Due to a remarkable result by T. Iwaniec we have

$$\|g\|_{L\log L(\Omega_T)} \approx |g|_{L\log L(\Omega_T)} := \fint_{\Omega_T} |g| \log\Big(e + \frac{g}{\fint_{\Omega_T}|g(\tilde z)|\,d\tilde z}\Big)\,dz\,. \tag{3.6}$$

The constant connecting the Luxemburg-norm $\|\cdot\|_{L\log L}$ with $|\cdot|_{L\log L}$ is independent of Ω_T and g. Moreover, and this is the striking fact of Iwaniec's result, $|\cdot|_{L\log L}$ defines a true norm on $L\log L(\Omega_T)$.

In the light of Definition (3.3) for $\theta\in[0,N]$ (3.3) the **parabolic Morrey-Orlicz-space** $L\log L^\theta(\Omega_T)$ is defined as the space of measurable function g defined on Ω_T satisfying

$$\begin{aligned}\|g\|_{L\log L^\theta(\Omega_T)} &:= \sup_{\mathcal{C}_\varrho\subset\Omega_T} \varrho^\theta \|g\|_{L\log L(\mathcal{C}_\varrho)}\\ &\approx \sup_{\mathcal{C}_\varrho\subset\Omega_T} \varrho^{\theta-N}\int_{\mathcal{C}_\varrho}|g|\log\Big(e+\frac{g}{\fint_{\mathcal{C}_\varrho}|g(\tilde z)|\,d\tilde z}\Big)\,dz<\infty\,.\end{aligned} \tag{3.7}$$

The following Lemma concerning the scaling properties of $\|\cdot\|_{L^\theta(p,q)}$ respectively $\|\cdot\|_{L\log L^\theta}$ is an immediate consequence of the definitions (3.3) resp. (3.7).

Lemma 3.2. *Let $g\in L^\theta(p,q)(\mathcal{C}(z_0,\varrho))$ with $1\le p<\infty$ and $0<q\le\infty$. Then, the map $\tilde g(y,\tau):=g(x_0+\varrho y,t_0+\varrho^2\tau)$, $(y,\tau)\in\mathcal{C}_1=\mathcal{C}(0,1)$, belongs to $L^\theta(p,q)(\mathcal{C}_1)$ and*

$$\|\tilde g\|_{L^\theta(p,q)(\mathcal{C}_1)} = \varrho^{-\frac{\theta}{p}}\|g\|_{L^\theta(p,q)(\mathcal{C}(z_0,\varrho))}.$$

Similarly, if $g\in L\log L^\theta(\mathcal{C}(z_0,\varrho))$ then $\tilde g\in L\log L^\theta(\mathcal{C}_1)$ and

$$\|\tilde g\|_{L\log L^\theta(\mathcal{C}_1)} = \varrho^{-\theta}\|g\|_{L\log L^\theta(\mathcal{C}(z_0,\varrho))}\,.$$

3.2. **Lower semi-continuity of quasi-norms.** As we have pointed out before the quantity $\|\cdot\|_{L^\theta(p,q)(\Omega_T)}$ is only a quasi-norm. Nevertheless, the mapping $g\mapsto\|g\|_{L^\theta(p,q)(\Omega_T)}$ is lower semi-continuous with respect to a.e. convergence. This can be seen as follows: Take $g_k\in L^\theta(p,q)(\Omega_T)$ with $g_k(z)\to g(z)$ a.e. on Ω_T as $k\to\infty$. Then by Fatou's Lemma we have

$$|\{z\in\Omega_T:|g(z)|>\lambda\}|\le\liminf_{k\to\infty}|\{z\in\Omega_T:|g_k(z)|>\lambda\}|, \tag{3.8}$$

whenever $\lambda\ge0$. For $q<\infty$ we use (3.8) and again Fatou's Lemma in (3.1) to have that

$$\begin{aligned}\|g\|_{L(p,q)(\Omega_T)} &= \Big[p\int_0^\infty\Big(\lambda^p|\{z\in\Omega_T:|g(z)|>\lambda\}|\Big)^{\frac{q}{p}}\frac{d\lambda}{\lambda}\Big]^{\frac1q}\\ &\le\Big[p\int_0^\infty\Big(\lambda^p\liminf_{k\to\infty}|\{z\in\Omega_T:|g_k(z)|>\lambda\}|\Big)^{\frac{q}{p}}\frac{d\lambda}{\lambda}\Big]^{\frac1q}\\ &=\liminf_{k\to\infty}\Big[p\int_0^\infty\Big(\lambda^p|\{z\in\Omega_T:|g_k(z)|>\lambda\}|\Big)^{\frac{q}{p}}\frac{d\lambda}{\lambda}\Big]^{\frac1q}\\ &=\liminf_{k\to\infty}\|g_k\|_{L(p,q)(\Omega_T)}.\end{aligned}$$

When $q=\infty$ we recall the definition of the Marcinkiewicz norm, i.e $\|g\|_{\mathcal{M}^p(\Omega_T)}=\sup_{\lambda>0}\big(\lambda|\{z\in\Omega_T:|g(z)|>\lambda\}|\big)^{\frac1p}$. Hence, by (3.8) for fixed $\lambda>0$ each of the functionals $g\mapsto\big(\lambda|\{z\in\Omega_T:|g(z)|>\lambda\}|\big)^{\frac1p}$ is lower semi-continuous with respect to a.e. convergence. The lower semi-continuity of the $\mathcal{M}^p$-norm now follows from the general fact that the supremum of an arbitrary family of lower semi-continuous functionals is still lower semi-continuous. The same argument also implies the lower semi-continuity of the quantities $\|\cdot\|_{L^\theta(p,q)(\Omega_T)}$ and $\|\cdot\|_{L\log L^\theta(\Omega_T)}$ since they are defined as the supremum over a family of balls of lower semi-continuous functionals.

3.3. **Additivity of quasi-norms.** The following elementary inequality holds

$$\Big(\sum_{i=1}^m a_i\Big)^\beta \le \max\{1, m^{\beta-1}\}\sum_{i=1}^m a_i^\beta \tag{3.9}$$

whenever $\beta > 0$ and a_i, $i = 1, \dots, m$ are non-negative numbers. We assume now that $\Omega_T \subset \bigcup_{i=1}^m \omega_i$, where $\omega_i = \Omega_i \times (\tau_i, t_i)$. Then, from (3.1) and (3.9) we infer that

$$\|g\|_{L(p,q)(\Omega_T,\mathbb{R}^k)} \le G(m,p,q)\sum_{i=1}^m \|g\|_{L(p,q)(\omega_i,\mathbb{R}^k)}, \tag{3.10}$$

holds for every $0 < q \le \infty$, where $G(m,p,q) = 1$ if $1 \le q \le p$ or $q = \infty$, while $G(m,p,q) = m^{1/p-1/q}$ if $q > p$ and $G(m,p,q) = m^{1/p-1}$ if $0 < q < 1$.

4. Parabolic Maximal Operators and Riesz Potentials

4.1. **Maximal Operators.** For fixed $\beta \in [0, N]$ we consider the (restricted) fractional maximal function operator relative to a symmetric parabolic cylinder $\mathcal{Q}_0 = \mathcal{Q}(z_0, R) \subset \mathbb{R}^{n+1}$ which is defined by

$$M^*_{\beta,\mathcal{Q}_0}(f)(z) := \sup_{\mathcal{Q}\subset\mathcal{Q}_0,\, z\in\mathcal{Q}} |\mathcal{Q}|^{\frac{\beta}{N}} \fint_{\mathcal{Q}} |f(\tilde z)|\, d\tilde z\,, \tag{4.1}$$

where the sup is taken with respect to all parabolic cylinders $\mathcal{Q}$ contained in $\mathcal{Q}_0$ having sides parallel to those of $\mathcal{Q}_0$ and containing the point z. When $\beta = 0$ we write $M^*_{\mathcal{Q}_0}$ instead of $M^*_{\beta,\mathcal{Q}_0}$. Moreover, in the case $\mathcal{Q}_0 = \mathbb{R}^{n+1}$ we abbreviate $M_\beta \equiv M^*_{\beta,\mathbb{R}^{n+1}}$ respectively $M \equiv M^*_{\mathbb{R}^{n+1}}$. Completely similar definitions and notations are given when cylinders with a cube as horizontal slice are replaced by those ones with a ball as horizontal slices:

$$M^*_{\beta,\mathcal{C}_0}(f)(z) := \sup_{\mathcal{C}\subset\mathcal{C}_0,\, z\in\mathcal{C}} |\mathcal{C}|^{\frac{\beta}{N}} \fint_{\mathcal{C}} |f(\tilde z)|\, d\tilde z\,,$$

where $\mathcal{C}_0 = \mathcal{C}(z_0, R)$ is a fixed parabolic cylinder and $\mathcal{C}$ is any other parabolic cylinder contained in $\mathcal{C}_0$ containing the point z. From [13, 32] we recall the boundedness of the maximal operators in Marcinkiewicz spaces, i.e. if $g \in L^q(\mathcal{Q}_0)$ then

$$|\{z \in \mathcal{Q}_0 : M^*_{\mathcal{Q}_0}(g)(z) \ge \lambda\}| \le \frac{c_0(n,q)}{\lambda^q}\int_{\mathcal{Q}_0} |g|^q\, dz \tag{4.2}$$

holds for every $\lambda > 0$ and $q \ge 1$. Moreover, we premise the following standard Hölder type inequality for Marcinkiewicz spaces.

Lemma 4.1. *Let $g \in \mathcal{M}^p(A)$ with $p > 1$ and $A \subset \mathbb{R}^{n+1}$ a measurable subset with finite measure $|A| < \infty$. Then $g \in L^q(A)$ for any $1 \le q < p$. Moreover, we have the estimate*

$$\|g\|_{L^q(A)} \le \Big(\frac{p}{p-q}\Big)^{\frac{1}{q}} |A|^{\frac{1}{q}-\frac{1}{p}} \|g\|_{\mathcal{M}^p(A)}\,.$$

The next theorem is a standard embedding theorem for the maximal function in Lorentz spaces. It can be easily inferred from [48, Theorem 7].

Theorem 4.2. *Let $\beta \in [0, N)$ and $p > 1$ such that $\beta p < N$; moreover let $q \in (0, \infty]$ and $\mathcal{C}$ a parabolic cylinder in $\mathbb{R}^{n+1}$. Then there exists a constant $c = c(n, p, \beta, q)$ such that for every map $g \in L(p, q)(\mathcal{C})$ there holds*

$$\|M^*_{\beta,\mathcal{C}}(g)\|_{L\left(\frac{Np}{N-\beta p},q\right)(\mathcal{C})} \le c\, \|g\|_{L(p,q)(\mathcal{C})}\,.$$

4.2. **Parabolic Riesz Potentials.** For $\beta \in (0, N)$ the fractional integral operator $I_\beta(\cdot)$, also called parabolic Riesz potential, is the linear operator defined by

$$I_\beta(g)(z) := \int_{\mathbb{R}^{n+1}} \frac{g(\tilde{z})}{d_{\text{par}}(z,\tilde{z})^{N-\beta}}\, d\tilde{z}\,, \qquad z \in \mathbb{R}^{n+1}, \tag{4.3}$$

for all measurable functions $g\colon \mathbb{R}^{n+1} \to \mathbb{R}$. This specializes the definition given in [39, p.24, (31)] for a doubling metric space (X, d, μ), i.e.

$$I_\beta(g)(z) := \int_X \frac{g(\tilde{z})d^\beta(z,\tilde{z})}{\mu(B(z, d(z,\tilde{z})))}\, d\tilde{z}\,, \qquad z \in X,$$

to the case $(\mathbb{R}^{n+1}, d_{\text{par}}, \mathcal{P}^N)$. We recall that for $z = (x,t)$, $\tilde{z} = (y,s) \in \mathbb{R}^{n+1}$ we have set

$$d_{\text{par}}(z,\tilde{z}) := \max\left\{|x-y|, \sqrt{|t-s|}\right\}.$$

Moreover, the parabolic Hausdorff-measure $\mathcal{P}^N$ is equivalent to the Lebesgue measure in $\mathbb{R}^{n+1}$. The following Lemma is an immediate consequence of the definitions of the fractional Riesz potential and the fractional maximal operators.

Lemma 4.3. *For every non-negative measurable function g defined on $\mathbb{R}^{n+1}$ there holds*

$$I_\beta(g)(z) \ge 2^{\beta-N}|\mathcal{Q}(0,1)|^{1-\frac{\beta}{N}} M_\beta(g)(z) \qquad \textit{for every } z \in \mathbb{R}^{n+1}.$$

Proof. Let $\mathcal{Q}(z_0,\varrho) \subset \mathbb{R}^{n+1}$ be an arbitrary but fixed symmetric parabolic cylinder containing the point z. Then $\mathcal{Q}(z_0,\varrho) \subset \mathcal{Q}(z, 2\varrho)$ and therefore

$$\begin{aligned}
I_\beta(g)(z) &\ge \int_{\mathcal{Q}(z,2\varrho)} \frac{g(\tilde{z})}{d_{\text{par}}(z,\tilde{z})^{N-\beta}}\, d\tilde{z} \\
&\ge (2\varrho)^{\beta-N} \int_{\mathcal{Q}(z_0,\varrho)} g(\tilde{z})\, d\tilde{z} \\
&= 2^{\beta-N}|\mathcal{Q}(0,1)|^{1-\frac{\beta}{N}} |\mathcal{Q}(z_0,\varrho)|^{\frac{\beta}{N}} \fint_{\mathcal{Q}(z_0,\varrho)} g(\tilde{z})\, d\tilde{z}\,.
\end{aligned}$$

Taking the sup with respect to all parabolic cylinders $\mathcal{Q}(z_0,\varrho)$ containing z then yields the result. □

Lemma 4.4. *Let $0 < \beta < N$, $p \ge 1$, $\theta > 0$ be such that $\beta < \theta/p \le N$, and let g be a non-negative measurable function on $\mathbb{R}^{n+1}$. Then the pointwise estimate*

$$I_\beta(g)(z) \le c[M_{\theta/p}(g)(z)]^{\frac{\beta p}{\theta}} [M(g)(z)]^{1-\frac{\beta p}{\theta}}$$

holds for every $z \in \mathbb{R}^{n+1}$ with a constant $c = c(n,p,\theta,\beta)$.

Proof. Without loss of generality we may assume that $g \not\equiv 0$. Let $z \in \mathbb{R}^{n+1}$. For $\delta > 0$ to be chosen later we decompose $\mathbb{R}^{n+1}$ into $\mathcal{Q}(z,\delta)$ and $\mathbb{R}^{n+1} \setminus Q(z,\delta)$ and write

$$I_\beta(g)(z) = \int_{\mathcal{Q}(z,\delta)} \dots d\tilde{z} + \int_{\mathbb{R}^{n+1}\setminus\mathcal{Q}(z,\delta)} \dots d\tilde{z} =: I_1 + I_2,$$

with the obvious meaning of I_1 and I_2. Moreover, for $k \in \mathbb{Z}$ we let

$$A_k(z) := \{\tilde{z} \in \mathbb{R}^{n+1} : 2^k\delta \le d_{\text{par}}(\tilde{z}, z) < 2^{k+1}\delta\}\,.$$

We first treat the integral I_1:

$$\begin{aligned}
I_1 &\le \sum_{k=1}^{\infty} \int_{A_{-k}(z)} \frac{g(\tilde z)}{d_{\mathrm{par}}(\tilde z, z)^{N-\beta}} d\tilde z \\
&\le \sum_{k=1}^{\infty} (2^{-k}\delta)^{\beta-N} \int_{\mathcal{Q}(z,2^{-k+1}\delta)} g(\tilde z) d\tilde z \\
&\le \sum_{k=1}^{\infty} (2^{-k}\delta)^{\beta-N} |\mathcal{Q}(z,2^{-k+1}\delta)| \fint_{\mathcal{Q}(z,2^{-k+1}\delta)} g(\tilde z) d\tilde z \\
&\le \alpha(n)\delta^{\beta} \sum_{k=1}^{\infty} 2^{-\beta k} M(g)(z) \\
&= \frac{\alpha(n)}{2^{\beta}-1}\, \delta^{\beta} M(g)(z)\,.
\end{aligned}$$

On the other hand we have for the integral I_2:

$$\begin{aligned}
I_2 &\le \sum_{k=0}^{\infty} \int_{A_{k}(z)} \frac{g(\tilde z)}{d_{\mathrm{par}}(\tilde z, z)^{N-\beta}} d\tilde z \\
&\le \sum_{k=0}^{\infty} (2^{k}\delta)^{\beta-N} |\mathcal{Q}(z,2^{k+1}\delta)| \fint_{\mathcal{Q}(z,2^{k+1}\delta)} g(\tilde z) d\tilde z \\
&\le \sum_{k=0}^{\infty} (2^{k}\delta)^{\beta-N} |\mathcal{Q}(z,2^{k+1}\delta)|^{1-\frac{\theta}{Np}} M_{\theta/p}(g)(z) \\
&= [2^{N+1}\alpha(n)]^{1-\frac{\theta}{Np}} \delta^{\beta-\frac{\theta}{p}} \sum_{k=0}^{\infty} 2^{k(\beta-\frac{\theta}{p})} M_{\theta/p}(g)(z) \\
&= \frac{[2^{N+1}\alpha(n)]^{1-\frac{\theta}{Np}}}{1-2^{\beta-\frac{\theta}{p}}} \delta^{\beta-\frac{\theta}{p}} M_{\theta/p}(g)(z)\,.
\end{aligned}$$

Having arrived at this stage we choose

$$\delta = \delta(z) := \left[\frac{M_{\theta/p}(g)(z)}{M(g)(z)}\right]^{\frac{p}{\theta}},$$

and finally obtain

$$\begin{aligned}
I_\beta(g)(z) &\le c\Bigg[\left(\frac{M_{\theta/p}(g)(z)}{M(g)(z)}\right)^{\frac{\beta p}{\theta}} M(g)(z) + \left(\frac{M_{\theta/p}(g)(z)}{M(g)(z)}\right)^{\frac{p}{\theta}(\beta-\frac{\theta}{p})} M_{\theta/p}(g)(z)\Bigg] \\
&\le 2c\,[M_{\theta/p}(g)(z)]^{\frac{\beta p}{\theta}} [M(g)(z)]^{1-\frac{\beta p}{\theta}},
\end{aligned}$$

where $c = c(n,p,\theta,\beta)$. □

Remark 4.5. We note that the constant in Lemma 4.4 blows up either when $\beta \downarrow 0$ - the case of singular integrals - or when $\beta p \uparrow \theta$ - the limiting case in the Sobolev-embedding. This settles in a certain sense the dependencies of the constants in all following results. Moreover, as an immediate consequence of Lemma 4.4 we obtain that for any measurable function g defined on $\mathbb{R}^{n+1}$ the pointwise Hedberg-type-estimate

$$I_\beta(|g|)(z) \le c\,[M_{\theta/p}(g)(z)]^{\frac{\beta p}{\theta}} [M(g)(z)]^{1-\frac{\beta p}{\theta}}$$

holds true for every $z \in \mathbb{R}^{n+1}$ with a constant $c = c(n,p,\theta,\beta)$.

Corollary 4.6. *Let* $0 < \beta < N$, $0 < \theta \le N$, $1 < p < \frac{\theta}{\beta}$, $1 \le q \le \infty$, $g \in L^p(\mathbb{R}^{n+1})$, $E \subset \mathbb{R}^{n+1}$ *and* $M_{\theta/p}(g) \in L^q(E)$. *Then*

$$\|I_\beta(g)\|_{L^r(E)} \le c\,\|M_{\theta/p}(g)\|_{L^q(E)}^{\frac{\beta p}{\theta}} \|g\|_{L^p(\mathbb{R}^{n+1})}^{1-\frac{\beta p}{\theta}},$$

where

$$\frac{1}{r} = \frac{1}{p} - \frac{\beta}{\theta} + \frac{\beta p}{\theta q}. \tag{4.4}$$

Proof. The case $q < \infty$: Integrating the Hedberg-Type inequality from Lemma 4.4 over E, using Hölder's inequality and (4.4) we infer

$$\begin{aligned}\int_E [I_\beta(|g|)]^r\,dz &\le c \int_E [M_{\theta/p}(g)]^{r\frac{\beta p}{\theta}} [M(g)]^{r(1-\frac{\beta p}{\theta})}\,dz \\ &\le c\bigg(\int_E [M_{\theta/p}(g)]^q\,dz\bigg)^{\frac{r\beta p}{\theta q}} \bigg(\int_E [M(g)]^{r\frac{1-\beta p/\theta}{1-r\beta p/\theta q}}\,dz\bigg)^{1-\frac{r\beta p}{\theta q}} \\ &\le c\,\|M_{\theta/p}(g)\|_{L^q(E)}^{\frac{r\beta p}{\theta}} \bigg(\int_{\mathbb{R}^{n+1}} [M(g)]^p\,dz\bigg)^{1-\frac{r\beta p}{\theta q}}.\end{aligned}$$

This leads to the estimate

$$\begin{aligned}\bigg(\int_E [I_\beta(|g|)]^r\,dz\bigg)^{\frac{1}{r}} &\le c\,\|M_{\theta/p}(g)\|_{L^q(E)}^{\frac{\beta p}{\theta}} \bigg(\int_{\mathbb{R}^{n+1}} [M(g)]^p\,dz\bigg)^{\frac{1}{r}-\frac{\beta p}{\theta q}} \\ &\le c\,\|M_{\theta/p}(g)\|_{L^q(E)}^{\frac{\beta p}{\theta}} \|g\|_{L^p(\mathbb{R}^{n+1})}^{1-\frac{\beta p}{\theta}},\end{aligned}$$

where we have used the boundedness of the maximal operator between L^p-spaces and the identity (4.4). This proves $I_\beta(|g|) \in L^r(E)$ and the desired estimate follows easily.

In the case $q = \infty$, instead of using Hölder's inequality in the first step, we have the trivial estimate

$$\int_E [I_\beta(|g|)]^r\,dz \le c\,\|M_{\theta/p}(g)\|_{L^\infty(E)}^{\frac{r\beta p}{\theta}} \int_{\mathbb{R}^{n+1}} [M(g)]^p\,dz$$

and we immediately obtain the desired estimate, taking into account (4.4) and $\frac{\beta p}{\theta q} = 0$. □

Lemma 4.7. *Let* $0 < \beta < N$, $p \ge 1$ *such that* $\beta p < N$, *and assume* $g \in L^p(\mathbb{R}^{n+1})$. *Then*

$$\bigg(\int_{\mathbb{R}^{n+1}} |I_\beta(g)|^{\frac{Np}{N-\beta p}}\,dz\bigg)^{\frac{N-\beta p}{Np}} \le c(n,p,\beta)\bigg(\int_{\mathbb{R}^{n+1}} |g|^p\,dz\bigg)^{\frac{1}{p}},$$

i.e. $I_\beta \colon L^p(\mathbb{R}^{n+1}) \hookrightarrow L^{\frac{Np}{N-\beta p}}(\mathbb{R}^{n+1})$ *is a continuous embedding.*

Proof. We apply Corollary 4.6 with $E \equiv \mathbb{R}^{n+1}$, $\theta = N$ and $q = \infty$. Note that

$$\begin{aligned}M_{\theta/p}(g)(z) = M_{N/p}(g)(z) &= \sup_{\mathcal{Q}\subset\mathbb{R}^{n+1},\,z\in\mathcal{Q}} |\mathcal{Q}|^{\frac{1}{p}} ⨍_{\mathcal{Q}} |g(\tilde z)|\,d\tilde z \\ &\le \sup_{\mathcal{Q}\subset\mathbb{R}^{n+1},\,z\in\mathcal{Q}} \bigg(\int_{\mathcal{Q}} |g(\tilde z)|^p\,d\tilde z\bigg)^{\frac{1}{p}} \le \|g\|_{L^p(\mathbb{R}^{n+1})},\end{aligned}$$

so that $\|M_{N/p}(g)\|_{L^\infty(\mathbb{R}^{n+1})} \le \|g\|_{L^p(\mathbb{R}^{n+1})}$. Moreover, $1/r = 1/p - \beta/N = \frac{N-\beta p}{Np}$. Inserting this in the estimate of Corollary (4.6) then yields the result. □

Remark 4.8. The preceding lemma yields in combination with the pointwise estimate from Lemma 4.3 that also the fractional maximal operator M_β is a continuous embedding from $L^p(\mathbb{R}^{n+1})$ into $L^{\frac{Np}{N-\beta p}}(\mathbb{R}^{n+1})$. Moreover, we have the estimate

$$\left(\int_{\mathbb{R}^{n+1}} |M_\beta(g)|^{\frac{Np}{N-\beta p}}\,dz\right)^{\frac{N-\beta p}{Np}} \le c(n,p,\beta)\left(\int_{\mathbb{R}^{n+1}} |g|^p\,dz\right)^{\frac{1}{p}},$$

whenever $0<\beta<N$ and $p\ge 1$ such that $\beta p<N$.

Lemma 4.9. *Let $g\in L^\theta(p,q)(\mathbb{R}^{n+1})$ and $\mathcal{C}_R\subset\mathbb{R}^{n+1}$ be a parabolic cylinder with radius $R>0$. Then, for any $s>1$ there holds*

$$\|g\chi_{\mathcal{C}_R}\|_{L^\theta(p,q)(\mathbb{R}^{n+1})} \le \max\left\{1,\left[(s-1)/2\right]^{\frac{\theta-N}{p}}\right\}\|g\|_{L^\theta(p,q)(\mathcal{C}_{sR})}\,.$$

Proof. We consider $\mathcal{C}_\varrho$ such that $\mathcal{C}_\varrho\cap\mathcal{C}_R\neq\emptyset$ and remark that

$$|\{z\in\mathcal{C}_\varrho : |(g\chi_{\mathcal{C}_R})(z)|>\lambda\}| \le |\{z\in\mathcal{C}_\varrho : |g(z)|>\lambda\}|\,.$$

In the case $\mathcal{C}_\varrho\subset\mathcal{C}_{sR}$ we have

$$\begin{aligned}
\varrho^{\theta-N}\|g\chi_{\mathcal{C}_R}\|^p_{L(p,q)(\mathcal{C}_\varrho)} &= \left[p\int_0^\infty \left(\lambda^p\varrho^{\theta-N}|\{z\in\mathcal{C}_\varrho : |(g\chi_{\mathcal{C}_R}(z)|>\lambda\}|\right)^{\frac{q}{p}}\frac{d\lambda}{\lambda}\right]^{\frac{p}{q}}\\
&\le \left[p\int_0^\infty \left(\lambda^p\varrho^{\theta-N}|\{z\in\mathcal{C}_\varrho : |g(z)|>\lambda\}|\right)^{\frac{q}{p}}\frac{d\lambda}{\lambda}\right]^{\frac{p}{q}}\\
&\le \left[p\int_0^\infty \left(\lambda^p\varrho^{\theta-N}|\{z\in\mathcal{C}_{sR} : |g(z)|>\lambda\}|\right)^{\frac{q}{p}}\frac{d\lambda}{\lambda}\right]^{\frac{p}{q}}\\
&\le \|g\|^p_{L^\theta(p,q)(\mathcal{C}_{sR})}\,.
\end{aligned}$$

In the remaining case $\mathcal{C}_\varrho\not\subset\mathcal{C}_{sR}$ (taking also into account $\mathcal{C}_\varrho\cap\mathcal{C}_{sR}\neq\emptyset$) we have $2\varrho>(s-1)R$. This implies

$$\varrho^{\theta-N}|\{z\in\mathcal{C}_\varrho : |(g\chi_{\mathcal{C}_R})(z)|>\lambda\}| \le [(s-1)/2]^{\theta-N}R^{\theta-N}|\{z\in\mathcal{C}_R : |g(z)|>\lambda\}|\,,$$

and similarly to the first case this leads us now to the estimate

$$\varrho^{\theta-N}\|g\chi_{\mathcal{C}_R}\|^p_{L(p,q)(\mathcal{C}_\varrho)} \le [(s-1)/2]^{\theta-N}\|g\|^p_{L^\theta(p,q)(\mathcal{C}_R)}\,.$$

Combining the two cases yields the desired estimate. □

Remark 4.10. Let $\mathcal{C}_\varrho$ be a parabolic cylinder with radius $\varrho>0$. Then, from the definition of the Lorentz-Morrey-norm (see (3.3)) we infer the bound

$$\begin{aligned}
\|g\|_{L(p,q)(\mathcal{C}_\varrho)} &= \varrho^{\frac{N-\theta}{p}}\left[p\int_0^\infty \left(\lambda^p\varrho^{\theta-N}|\{z\in\mathcal{C}_\varrho : |g(z)|>\lambda\}|\right)^{\frac{q}{p}}\frac{d\lambda}{\lambda}\right]^{\frac{1}{q}}\\
&\le \max\left\{1,\varrho^{\frac{N-\theta}{p}}\right\}\|g\|_{L^\theta(p,q)(\mathcal{C}_\varrho)}\,.
\end{aligned}$$

Lemma 4.11. *Let $\mathcal{C}$ be a parabolic cylinder in $\mathbb{R}^{n+1}$. Then for every $g\in L\log L(\mathcal{C})$ with support in $\mathcal{C}$ we have*

$$⨍_{\mathcal{C}} M(g)\,dz \le c(n)\|g\|_{L\log L(\mathcal{C})} \approx c(n)|g|_{L\log L(\mathcal{C})}\,.$$

Proof. We define

$$h := \frac{g}{\|g\|_{L\log L(\mathcal{C})}}\,.$$

Then, we have $\|h\|_{L\log L(\mathcal{C})}=1$ and therefore

$$⨍_{\mathcal{C}} |h|\log\left(e+|h|\right)dz \le 1\,.$$

Applying (the parabolic analogue of) [21, Theorem 2.15] we conclude

$$\int_{\mathcal{C}} M(h)\,dz \le c(n)|\mathcal{C}| + c(n)\int_{\mathcal{C}} |h|\log\big(e+|h|\big)\,dz \le c(n)|\mathcal{C}|\,,$$

so that $\fint_{\mathcal{C}} M(h)\,dz \le c(n)$. Re-scaling back from h to g then yields the desired estimate. □

Theorem 4.12. *Let $0<\beta<\theta\le N$, $\mathcal{C}$ a parabolic cylinder in $\mathbb{R}^{n+1}$ and $s>1$. Then there exists a constant $c=c(n,\beta,\theta,s)$ such that the estimate*

$$\|M^*_{\beta,\mathcal{C}}(g)\|_{L^{\frac{\theta}{\theta-\beta}}(\mathcal{C})} \le c\,|\mathcal{C}|^{1-\frac{\beta}{\theta}}\,\|g\|^{\frac{\beta}{\theta}}_{L^{1,\theta}(s\mathcal{C})}\|g\|^{1-\frac{\beta}{\theta}}_{L\log L(\mathcal{C})}$$

holds, whenever $g\colon s\mathcal{C}\to\mathbb{R}$ is measurable.

Proof. Without loss of generality we may assume that $g\ge 0$. Then, $I_\beta(g)\ge 0$. Let $\tilde g = g\chi_{\mathcal{C}}$. Applying Lemma 4.4 with $p=1$ we obtain that

$$[I_\beta(\tilde g)(z)]^{\frac{\theta}{\theta-\beta}} \le c\,[M_\theta(\tilde g)(z)]^{\frac{\beta}{\theta-\beta}}M(\tilde g)(z) \le c\,\|\tilde g\|^{\frac{\beta}{\theta-\beta}}_{L^{1,\theta}(\mathbb{R}^{n+1})}M(\tilde g)(z)$$

holds for every $z\in\mathbb{R}^{n+1}$. Here we have also used the obvious estimate $M_\theta(\tilde g)(z)\le c\,\|\tilde g\|_{L^{1,\theta}(\mathbb{R}^{n+1})}$ which follows directly from the definitions of the fractional maximal operators given in (4.1) and the usual one of a Morrey-space. Using the pointwise bound from below for $I_\beta(\tilde g)$ from Lemma 4.3 we infer that

$$[M_\beta(\tilde g)(z)]^{\frac{\theta}{\theta-\beta}} \le c\,\|\tilde g\|^{\frac{\beta}{\theta-\beta}}_{L^{1,\theta}(\mathbb{R}^{n+1})}M(\tilde g)(z)$$

holds for every $z\in\mathbb{R}^{n+1}$. Integrating the preceding inequality on $\mathcal{C}$ and using Lemma 4.11 yields

$$\|M_\beta(\tilde g)\|^{\frac{\theta}{\theta-\beta}}_{L^{\frac{\theta}{\theta-\beta}}(\mathcal{C})} \le c\,|\mathcal{C}|\|\tilde g\|^{\frac{\beta}{\theta-\beta}}_{L^{1,\theta}(\mathbb{R}^{n+1})}\fint_{\mathcal{C}} M(\tilde g)\,dz \le c\,|\mathcal{C}|\|\tilde g\|^{\frac{\beta}{\theta-\beta}}_{L^{1,\theta}(\mathbb{R}^{n+1})}\|\tilde g\|_{L\log L(\mathcal{C})}\,.$$

Recalling the obvious inequality $M^*_{\beta,\mathcal{C}}(g)\le M_\beta(\tilde g)$ in order to estimate the left-hand side from below and Lemma 4.9 to estimate the right-hand side from above, i.e. the fact that $\|\tilde g\|_{L^{1,\theta}(\mathbb{R}^{n+1})}\le c\|g\|_{L^{1,\theta}(s\mathcal{C})}$, we conclude the assertion of the lemma. □

Theorem 4.13. *Let $\beta,\theta\in(0,N]$, $p>1$, such that $\beta p<\theta$, and let $q\in(0,\infty]$. Furthermore, let $\mathcal{C}$ be a parabolic cylinder in $\mathbb{R}^{n+1}$ and $s>1$. Then there exists a constant $c=c(n,p,q,\beta,\theta,s)$ such that*

$$\|M^*_{\beta,\mathcal{C}}(g)\|_{L(\frac{\theta p}{\theta-\beta p},\frac{\theta q}{\theta-\beta p})(\mathcal{C})} \le c\,\|g\|^{\frac{\beta p}{\theta}}_{L^\theta(p,q)(s\mathcal{C})}\|g\|^{1-\frac{\beta p}{\theta}}_{L(p,q)(\mathcal{C})}$$

holds whenever g is a measurable map defined on $s\mathcal{C}$. Moreover, if $|s\mathcal{C}|\le 100^N$ we have

$$\|M^*_{\beta,\mathcal{C}}(g)\|_{L(\frac{\theta p}{\theta-\beta p},\frac{\theta q}{\theta-\beta p})(\mathcal{C})} \le c\,\|g\|_{L^\theta(p,q)(s\mathcal{C})}\,.$$

The constant c blows up, i.e. $c\to\infty$, when $q\downarrow 0$ or $p\downarrow 1$.

Proof. In the case $q=\infty$ we let $\frac{\theta q}{\theta-\beta p}:=\infty$. Once again we may assume without loss of generality that $g\ge 0$. We define $\tilde g := g\chi_{\mathcal{C}}$. Then for $\mathcal{C}_R\subset\mathbb{R}^{n+1}$ we have

$$\begin{aligned}\int_{\mathcal{C}_R}|\tilde g|\,dz &\le \frac{p}{p-1}|\mathcal{C}_R|^{1-\frac{1}{p}}\|\tilde g\|_{\mathcal{M}^p(\mathcal{C}_R)} = \frac{cp}{p-1}\,R^{N(1-\frac{1}{p})}\|\tilde g\|_{\mathcal{M}^p(\mathcal{C}_R)}\\ &= \frac{cp}{p-1}\,R^{N-\frac{\theta}{p}}R^{\frac{\theta-N}{p}}\|\tilde g\|_{\mathcal{M}^p(\mathcal{C}_R)} \le \frac{cp}{p-1}\,R^{N-\frac{\theta}{p}}\|\tilde g\|_{\mathcal{M}^{p,\theta}(\mathbb{R}^{n+1})},\end{aligned}$$

where $c=c(n,p)$. This implies in particular that $M_{\theta/p}(\tilde g)(z)\le\frac{cp}{p-1}\|\tilde g\|_{\mathcal{M}^{p,\theta}(\mathbb{R}^{n+1})}$ holds for every $z\in\mathbb{R}^{n+1}$. Using this in the Hedberg-type inequality from Lemma 4.4 yields

$$[I_\beta(\tilde g)(z)]^{\frac{\theta}{\theta-\beta p}} \le c\,[M_{\theta/p}(\tilde g)(z)]^{\frac{\beta p}{\theta-\beta p}}M(\tilde g)(z) \le c\,\|\tilde g\|^{\frac{\beta p}{\theta-\beta p}}_{\mathcal{M}^{p,\theta}(\mathbb{R}^{n+1})}M(\tilde g)(z)\,,$$

for every $z \in \mathbb{R}^{n+1}$. In the preceding inequality we want to replace the Marcinkiewicz-norm of $\tilde{g}$ by an appropriate Lorentz-Morrey norm. For this we recall that for $q > 0$ we have $\|\tilde{g}\|_{\mathcal{M}^p(\mathcal{C}_R)} \le (q/p)^{\frac{1}{q}} \|\tilde{g}\|_{L(p,q)(\mathcal{C}_R)}$, so that $\|\tilde{g}\|_{\mathcal{M}^{p,\theta}(\mathbb{R}^{n+1})} \le c\, \|\tilde{g}\|_{L^\theta(p,q)(\mathbb{R}^{n+1})}$. Inserting this above we immediately find

$$[I_\beta(\tilde{g})(z)]^{\frac{\theta}{\theta-\beta p}} \le c\, \|\tilde{g}\|^{\frac{\beta p}{\theta-\beta p}}_{L^\theta(p,q)(\mathbb{R}^{n+1})} M(\tilde{g})(z)\,,$$

which leads after integrating in an appropriate way over $\mathbb{R}^{n+1}$ to

$$\|[I_\beta(\tilde{g})]^{\frac{\theta}{\theta-\beta p}}\|^{\frac{\theta-\beta p}{\theta}}_{L(p,q)(\mathbb{R}^{n+1})} \le c\, \|\tilde{g}\|^{\frac{\beta p}{\theta}}_{L^\theta(p,q)(\mathbb{R}^{n+1})} \|M(\tilde{g})\|^{1-\frac{\beta p}{\theta}}_{L(p,q)(\mathbb{R}^{n+1})}\,. \tag{4.5}$$

Using definition (3.1) and a simple change-of-variable argument we find for the left-hand side of (4.5) the identity

$$\begin{aligned}
&\|[I_\beta(\tilde{g})]^{\frac{\theta}{\theta-\beta p}}\|_{L(p,q)(\mathbb{R}^{n+1})} \\
&\quad = \left[p \int_0^\infty \left(\lambda^p |\{z \in \mathbb{R}^{n+1} : [I_\beta(\tilde{g})(z)]^{\frac{\theta}{\theta-\beta p}} > \lambda\}| \right)^{\frac{q}{p}} \frac{d\lambda}{\lambda} \right]^{\frac{1}{q}} \\
&\quad = \left[p \int_0^\infty \left(\lambda^p |\{z \in \mathbb{R}^{n+1} : I_\beta(\tilde{g})(z) > \lambda^{\frac{\theta-\beta p}{\theta}}\}| \right)^{\frac{q}{p}} \frac{d\lambda}{\lambda} \right]^{\frac{1}{q}} \\
&\quad = \left[\frac{\theta p}{\theta-\beta p} \int_0^\infty \left(\mu^{\frac{\theta p}{\theta-\beta p}} |\{z \in \mathbb{R}^{n+1} : I_\beta(\tilde{g})(z) > \mu\}| \right)^{\frac{\theta q/(\theta-\beta p)}{\theta p/(\theta-\beta p)}} \frac{d\mu}{\mu} \right]^{\frac{1}{q}} \\
&\quad = \|I_\beta(\tilde{g})\|^{\frac{\theta}{\theta-\beta p}}_{L\left(\frac{\theta p}{\theta-\beta p}, \frac{\theta q}{\theta-\beta p}\right)(\mathbb{R}^{n+1})}\,.
\end{aligned} \tag{4.6}$$

On the other hand the boundedness of the maximal operator in Lorentz-spaces allows us to estimate the second term on the right-hand side of (4.5) from above; to be precise we have

$$\|M(\tilde{g})\|_{L(p,q)(\mathbb{R}^{n+1})} \le c(n,p,q) \|\tilde{g}\|_{L(p,q)(\mathbb{R}^{n+1})}\,.$$

Using this and (4.6) in (4.5) we arrive at

$$\|I_\beta(\tilde{g})\|_{L\left(\frac{\theta p}{\theta-\beta p}, \frac{\theta q}{\theta-\beta p}\right)(\mathbb{R}^{n+1})} \le c\, \|\tilde{g}\|^{\frac{\beta p}{\theta}}_{L^\theta(p,q)(\mathbb{R}^{n+1})} \|\tilde{g}\|^{1-\frac{\beta p}{\theta}}_{L(p,q)(\mathbb{R}^{n+1})}\,.$$

The first term in the right-hand side of the preceding inequality is estimated by Lemma 4.9, i.e $\|\tilde{g}\|_{L^\theta(p,q)(\mathbb{R}^{n+1})} \le c(s)\, \|g\|_{L^\theta(p,q)(s\mathcal{C})}$, while the second term is equal to $\|g\|_{L(p,q)(\mathcal{C})}$. On the other hand from Lemma 4.3 and the definition of the restricted maximal operator in (4.1) we infer the pointwise estimate $c^{-1} I_\beta(\tilde{g})(z) \ge M_\beta(\tilde{g})(z)$. Inserting this above yields

$$\|M_\beta(\tilde{g})\|_{L\left(\frac{\theta p}{\theta-\beta p}, \frac{\theta q}{\theta-\beta p}\right)(\mathcal{C})} \le c\, \|g\|^{\frac{\beta p}{\theta}}_{L^\theta(p,q)(s\mathcal{C})} \|g\|^{1-\frac{\beta p}{\theta}}_{L(p,q)(\mathcal{C})}\,.$$

Combining this with $M_\beta(\tilde{g})(z) \ge M_{\beta,\mathcal{C}}(g)(z)$, $z \in \mathcal{C}$, leads to the first asserted estimate of the theorem. In order to obtain the second assertion we use Remark 4.10 and the assumption $|s\mathcal{C}| \le 100^N$ to estimate the right-hand side from above. □

5. Basic Regularity

5.1. **L^1-regularity for regularized problems.** In this section we will consider

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

defined as the unique solution to the regularized Cauchy-Dirichlet problems

$$\begin{cases} u_t - \operatorname{div} a(z, Du) = g \in L^\infty(\Omega_T) & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial_{\text{par}}\Omega_T, \end{cases} \tag{5.1}$$

for some fixed g. Associated to a fixed symmetric parabolic cylinder $\mathcal{C}_R = \mathcal{C}(z_0, R) \subset \Omega_T$ we consider the unique solution

$$v \in C^0([t_0 - R^2, t_0 + R^2]; L^2(B_R(x_0))) \cap L^2(t_0 - R^2, t_0 + R^2; W^{1,2}(B_R(x_0)))$$

to the following homogeneous Cauchy-Dirichlet problem:

$$\begin{cases} v_t - \operatorname{div} a(z, Du) = 0 & \text{in } \mathcal{C}_R, \\ v = u & \text{on } \partial_{\text{par}}\mathcal{C}_R. \end{cases} \tag{5.2}$$

Remark 5.1. At certain points in the proofs of our results it is useful to scale from an arbitrary parabolic cylinder $C(z_0, R)$ to $\mathcal{C} = \mathcal{C}(0, 1)$ via the following scaling procedure: For $\tilde{z} = (y, s) \in \mathcal{C}$ we define

$$\begin{cases} \tilde{u}(\tilde{z}) := R^{-1}u(x_0 + Ry, t_0 + R^2 s), & \tilde{v}(\tilde{z}) := R^{-1}v(x_0 + Ry, t_0 + R^2 s), \\ \tilde{a}(\tilde{z}, w) := a(x_0 + Ry, t_0 + R^2 s, w), & \tilde{g}(\tilde{z}) := Rg(x_0 + Ry, t_0 + R^2 s). \end{cases}$$

Then it is easy to verify that $\tilde{u}_s - \operatorname{div} \tilde{a}(\tilde{z}, D\tilde{u}) = \tilde{g}$ and $\tilde{v}_s - \operatorname{div} \tilde{a}(\tilde{z}, D\tilde{v}) = 0$ in $\mathcal{C}$, and $\tilde{u} = \tilde{v}$ on $\partial_{\text{par}}\mathcal{C}$. Furthermore it is easy to check that the new vector field $\tilde{a}$ satisfies the growth and and monotonicity properties described in (1.2).

The following comparison lemma can be inferred from [26, Lemma 4.1], see also [7, Lemma 6.4].

Lemma 5.2. *Let $u \in C^0([-T, 0]; L^2(\Omega)) \cap L^2(-T, 0; W^{1,2}(\Omega))$ be a solution to (5.1) under the assumption (1.2) and $\mathcal{C}(z_0, R)$ a parabolic cylinder in Ω_T. Moreover, let $v \in C^0([t_0 - R^2, t_0 + R^2]; L^2(B_R(x_0))) \cap L^2(t_0 - R^2, t_0 + R^2; W^{1,2}(B_R(x_0)))$ be a solution to the Cauchy-Dirichlet problem (5.2). Then there exists a constant $c = c(n)$ such that*

$$\int_{\mathcal{C}(z_0,R)} R^{-1}|u - v| + |Du - Dv| \, dz \le c \, \nu^{-1} \, R \int_{\mathcal{C}(z_0,R)} |g| \, dz \,. \tag{5.3}$$

Lemma 5.3. *Let $u \in C^0([-T, 0]; L^2(\Omega)) \cap L^2(-T, 0; W^{1,2}(\Omega))$ be a solution to (5.1) where the structure conditions (1.2) are in force and $\mathcal{C}_R = \mathcal{C}(z_0, R)$ a parabolic cylinder in Ω_T. Moreover, let $v \in C^0([t_0 - R^2, t_0 + R^2]; L^2(B_R(x_0))) \cap L^2(t_0 - R^2, t_0 + R^2; W^{1,2}(B_R(x_0)))$ be the unique solution to the Cauchy-Dirichlet problem (5.2) in $\mathcal{C}(z_0, R)$ and $g \in L^\theta(\gamma, q)(\mathcal{C}_R)$ for some $\gamma > 1$. Then there exists a constant $c = c(n, \nu, \gamma)$ such that*

$$\int_{\mathcal{C}(z_0,R)} R^{-1}|u - v| + |Du - Dv| \, dz \le c \, R^{N - \frac{\theta - \gamma}{\gamma}} \|g\|_{L^\theta(\gamma,q)(\mathcal{C}(z_0,R))} \,. \tag{5.4}$$

Proof. Using Lemma 4.1 and the embedding $\|g\|_{\mathcal{M}^\gamma(\mathcal{C}_R)} \le (q/\gamma)^{\frac{1}{q}} \|g\|_{L(\gamma,q)(\mathcal{C}_R)}$ we can conclude for the right-hand side in (5.3) that

$$\begin{aligned} R \int_{\mathcal{C}_R} |g| \, dz &\le c \left(\frac{\gamma}{\gamma - 1} \right) R^{1 + N(1 - \frac{1}{\gamma})} \|g\|_{\mathcal{M}^\gamma(\mathcal{C}_R)} \\ &\le c \left(\frac{\gamma}{\gamma - 1} \right) \left(\frac{q}{\gamma} \right)^{\frac{1}{q}} R^{1 + N(1 - \frac{1}{\gamma})} \|g\|_{L(\gamma,q)(\mathcal{C}_R)} \\ &\le c \, R^{N - \frac{\theta - \gamma}{\gamma}} R^{\frac{\theta - N}{\gamma}} \|g\|_{L(\gamma,q)(\mathcal{C}_R)} \\ &\le c \, R^{N - \frac{\theta - \gamma}{\gamma}} \|g\|_{L^\theta(\gamma,q)(\mathcal{C}_R)}, \end{aligned}$$

where $c = c(n, \nu, \gamma)$. Using the preceding inequality in (5.3) yields the result. □

Lemma 5.4. *Let $u \in C^0([-T, 0]; L^2(\Omega)) \cap L^2(-T, 0; W^{1,2}(\Omega))$ be a solution to (5.1) where the structure conditions (1.2) are in force and $\mathcal{C}_R = \mathcal{C}(z_0, R)$ a parabolic cylinder in Ω_T. Moreover, let $v \in C^0([t_0 - R^2, t_0 + R^2]; L^2(B_R(x_0))) \cap L^2(t_0 - R^2, t_0 + R^2; W^{1,2}(B_R(x_0)))$ be the unique solution to the Cauchy-Dirichlet problem (5.2) in $\mathcal{C}(z_0, R)$ and $g \in L^{1,\theta}(\mathcal{C}_R)$. Then there exists a constant $c = c(n, \nu, \gamma)$ such that*

$$\int_{\mathcal{C}(z_0,R)} R^{-1}|u - v| + |Du - Dv| \, dz \le c \, R^{N - (\theta - 1)} \|g\|_{L^{1,\theta}(\mathcal{C}(z_0,R))} \,. \tag{5.5}$$

Proof. For the proof it is sufficient to note that $\|g\|_{L^1(\mathcal{C}_R)} \le R^{N - \theta} \|g\|_{L^{1,\theta}(\mathcal{C}_R)}$. □

5.2. **Homogeneous problems.** The results of this chapter summarize the basic Hölder regularity results from the De Giorgi-Nash-Moser theory of solutions to non-linear, homogeneous parabolic equations as well as the higher integrability theory.

Theorem 5.5. *Let $v \in C^0([-T,0];L^2(\Omega)) \cap L^2(-T,0;W^{1,2}(\Omega))$ be a weak solution to the parabolic equation*

$$v_t - \operatorname{div} a(z,Dv) = 0 \qquad \text{in } \Omega_T, \tag{5.6}$$

under the assumptions

$$|a(z,w)| \le L(1+|w|)\,, \qquad \nu|w|^2 - L^2/\nu \le \langle a(z,w),w\rangle, \tag{5.7}$$

for every choice of $z \in \Omega_T$ and $w \in \mathbb{R}^n$ where $0<\nu\le 1\le L<\infty$ and $a\colon \Omega_T \times \mathbb{R}^n \to \mathbb{R}^n$ is a Carathéodory vector field. Then, there exists $\alpha \in (0,\frac12]$ depending only on n and L/ν, such that for every $q \in (0,2]$ there exists a constant $c = c(n,L,\nu,q)$ such that the following holds: Whenever $\mathcal{C}_R \subset \Omega_T$ and $0<\varrho\le R$ there holds

$$\int_{\mathcal{C}_\varrho} (|Dv|^q+1)\,dz \le c\left(\frac{\varrho}{R}\right)^{N-q+\alpha q}\int_{\mathcal{C}_R}(|Dv|^q+1)\,dz \tag{5.8}$$

and

$$\int_{\mathcal{C}_\varrho} (|v|^q+\varrho^q)\,dz \le c\left(\frac{\varrho}{R}\right)^{N}\int_{\mathcal{C}_R}(|v|^q+R^q)dz\,. \tag{5.9}$$

Furthermore, there exists $\chi = \chi(n,L,\nu) > 1$ such that $Dv \in L^{2\chi}_{\mathrm{loc}}(\Omega_T,\mathbb{R}^n)$ and

$$\left(\fint_{\mathcal{C}_{R/2}} (|Dv|+1)^{2\chi}\,dz\right)^{\frac{1}{2\chi}} \le c\left(\fint_{\mathcal{C}_R}(|Dv|+1)^q\,dz\right)^{\frac1q} \tag{5.10}$$

holds for every $q\in(0,2]$, while for every $\chi_0>1$ there holds

$$\left(\fint_{\mathcal{C}_{R/2}} (|v|+R)^{2\chi_0}\,dz\right)^{\frac{1}{2\chi_0}} \le c\left(\fint_{\mathcal{C}_R}(|v|+R)^q\,dz\right)^{\frac1q}. \tag{5.11}$$

In both (5.10) and (5.11) we have the following dependence of the constant c from the structural constants: $c = c(n,L,\nu,q)$

Proof. The statement is a direct consequence of De Giorgi-Nash-Moser's theory. We give a very brief hint how to retrieve the estimates (5.8) to (5.11). (5.10) for $q=2$ can be inferred for instance from [50], and we refer the reader also to [12, Lemma 3.1], where the statement is directly proved at the boundary. From this, the reduction of the exponent 2 on the right hand side to any exponent $q\in(0,2]$ follows by a standard result on reverse Hölder inequalities. For details, we refer the reader for example to [26, Chapter 4] and Lemma 3.1. The estimates (5.8), (5.9) and also (5.11) with $q=2$ follow for instance from [45], Chapter 6, where De Giorgi's proof is performed in the parabolic setting. Then, again the arguments of [26] allow to reduce the exponent 2 on the right hand side to any exponent $q\in(0,2]$. Note here, that the arguments in [45] are worked out for linear parabolic equations, but as mentioned in the notes in Chapter 6, the linearity of the equation is actually irrelevant for the estimates and they hold also for quasi-linear equations fulfilling the structure conditions (5.7). □

Remark 5.6. *Theorem 5.5 actually holds for a much larger class of equations, involving also p-growth conditions with exponent $p\neq 2$, and possibly being degenerate. Indeed the estimates can be retrieved also from* [19, Chapter III,V].

The next theorem is the homogeneous case of a much more general result concerning Calderón-Zygmund estimates for weak solutions to non-linear parabolic equations (systems); see [28, Theorem 1.8, 1.9] for the specific form of the statement. We consider weak solutions $v \in C^0([-T,0];L^2(\Omega)) \cap L^2(-T,0;W^{1,2}(\Omega))$ to the following homogeneous non-linear parabolic equation

$$v_t - \operatorname{div} a(z, Dv) = 0 \qquad \text{in } \Omega_T, \tag{5.12}$$

where the vector-field $a(z,w)$ satisfies either the structure conditions (1.11) or the vector field has the special form $a(z,w) = a(x,t,w) = c(x)\bar{a}(t,w)$, where $c(x)$ and $\bar{a}(t,w)$ satisfy the conditions (1.12), (1.13) and (1.14). Then the following theorem holds:

Theorem 5.7. *Let $v \in C^0([-T,0];L^2(\Omega)) \cap L^2(-T,0;W^{1,2}(\Omega))$ be a weak solution to the homogeneous non-linear parabolic equation (5.12) where either the structure assumptions (1.11) or the conditions (1.12) to (1.14) are in force. Then for any $\alpha \in (0,1)$ and $q \in (0,2]$ there exists a constant $c = c(n,L,\nu,\alpha,q)$ such that the following holds: Whenever $\mathcal{C}_R \subset \Omega_T$ and $0 < \varrho \leq R$ there holds*

$$\int_{\mathcal{C}_\varrho} (|Dv|^q + 1)\, dz \leq c \left(\frac{\varrho}{R}\right)^{N-q+\alpha q} \int_{\mathcal{C}_R} (|Dv|^q + 1)\, dz\,. \tag{5.13}$$

Furthermore, $Dv \in L^{2\chi_0}_{\mathrm{loc}}(\Omega_T,\mathbb{R}^n)$ for any $\chi_0 > 1$. Moreover, for any given $\chi_0 > 1$ and $q \in (0,2]$ there exits a constant $c = c(n,\nu,L,\chi_0,\omega(\cdot),q)$ such that for any $\mathcal{C}_R \Subset \Omega_T$ there holds

$$\left(\fint_{\mathcal{C}_{R/2}} (|Dv|+1)^{2\chi_0}\, dz\right)^{\frac{1}{2\chi_0}} \leq c\left(\fint_{\mathcal{C}_R} (|Dv|+1)^q\, dz\right)^{\frac{1}{q}}. \tag{5.14}$$

Proof. Estimate (5.14) is the statement of [28, Theorems 1.8, 1.9]. Then, once having (5.13) for the case $q = 2$, the general estimate for $q \in (0,2]$ can be retrieved by a simple application of Hölder's inequality to pass from exponent $q < 2$ to exponent 2, then exploiting (5.13) for $q = 2$, and subsequently using (5.14) to reduce the exponent 2 again to exponent $q < 2$. However, (5.13) for the special case $q = 2$ is a consequence of the Hölder continuity to any Hölder exponent $\alpha \in (0,1)$ for solutions to parabolic equations with linear growth. On the other hand, Hölder continuity to every exponent $\alpha \in (0,1)$ is a standard consequence of the fact that the vector field a is sufficiently regular with respect to x. In this case, Hölder continuity can be shown via suitable comparison procedures to differentiable or constant coefficient equations (see [28, Chapter 8] for comparison estimates in the case of VMO-regular vector fields as well as continuous ones). We note here, that actually the estimates in [28] are shown for much more general possibly degenerate p growth equations and systems. Standard references for Hölder regularity in the constant coefficient case are for example [44, 45]. □

Remark 5.8. In the estimates (5.10) and (5.14) we can replace $R/2, R$ by $\sigma R, R$ for any $\sigma \in [\frac{1}{2}, 1)$ as long as we enlarge the constant by factor $\approx (1-\sigma)^{N(\frac{1}{2\chi}-\frac{1}{q})}$. This can be inferred along the arguments form [36, Remark 6.12]. On the other hand inequalities (5.8)–(5.14) continue to hold when replacing the parabolic cylinders $\mathcal{C}$ with a ball as horizontal slice by the cylinders $\mathcal{Q}$ having a cube as horizontal cross section.

6. Integrability of Du

6.1. **Parabolic Lorentz space estimates.** Theorem 6.1 below can be considered as as the non-linear parabolic analogue of a classical result of Adams & Lewis [5]. The corresponding non-linear elliptic version has been obtained in [48]. Moreover the **Proof of Theorem 1.1** follows directly from the more general Theorem 6.1 by the choice $q = \gamma$.

Theorem 6.1. *Let $u \in L^1(-T,0;W^{1,1}_0(\Omega))$ be the solution to (1.1) where the structure conditions (1.2) are in force. Moreover, assume $g \in L^\theta(\gamma,q)(\Omega_T)$ with γ,θ as in (1.7) and $0 < q \leq \infty$. Then*

$$|Du| \in L^\theta\left(\tfrac{\theta\gamma}{\theta-\gamma}, \tfrac{\theta q}{\theta-\gamma}\right) \quad \textit{locally in } \Omega_T. \tag{6.1}$$

Furthermore, we have the local estimate

$$\|Du\|_{L^{\theta}(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_{R/2})} \le c\, R^{\frac{\theta-\gamma}{\gamma}-N}\|1+|Du|\|_{L^1(\mathcal{C}_R)} + c\|g\|_{L^{\theta}(\gamma,q)(\mathcal{C}_R)}, \quad (6.2)$$

for any parabolic cylinder $\mathcal{C}_R \subset \Omega_T$, where the constant c depends only on n, L, ν, γ, q.

Proof. The proof is divided into several steps.

Step 1: Level sets decay. On a fixed parabolic cylinder $\mathcal{Q}_0$ satisfying $n^2\mathcal{Q}_0 \Subset \Omega_T$ and $|\mathcal{Q}_0| \le 1$ we consider the following maximal operators

$$M^* := M^*_{0,n^2\mathcal{Q}_0} = M^*_{n^2\mathcal{Q}_0} \qquad \text{and} \qquad M^*_1 := M^*_{1,n^2\mathcal{Q}_0}\,.$$

For the definitions of these restricted maximal function operators we refer the reader to Section 4.1, especially to (4.1).

Lemma 6.2. *Let $u(=u_k) \in C^0([-T,0];L^2(\Omega)) \cap L^2(-T,0;W^{1,2}_0(\Omega))$ be a weak solution to (5.1) where the assumptions (1.2) are in force and $g \in L^\infty(\Omega_T)$. Then, for every $S > 1$ there exists a constant $\varepsilon = \varepsilon(n,L,\nu,S) \in (0,1)$ such that if $\lambda > 1$ and $\mathcal{Q}$ is a dyadic sub-cylinder of $\mathcal{Q}_0$ such that*

$$\big|\mathcal{Q} \cap \big\{z \in \mathcal{Q}_0 : M^*(1+|Du|)(z) > AS\lambda \ \textit{and}\ \ M^*_1(g) \le \varepsilon\lambda\big\}\big| > \frac{|\mathcal{Q}|}{S^{2\chi}}, \quad (6.3)$$

then the predecessor $\tilde{\mathcal{Q}}$ of $\mathcal{Q}$ satisfies

$$\tilde{\mathcal{Q}} \subset \big\{z \in \mathcal{Q}_0 : M^*(1+|Du|)(z) > \lambda\big\}\,. \quad (6.4)$$

Here $\chi = \chi(n, L/\nu) > 1$ is the higher integrability exponent introduced in Theorem 5.5, while $A = A(n, L/\nu) > 1$ is an absolute constant.

Proof of Lemma 6.2. We shall prove the assertion of the lemma by a contradiction argument. Therefore we assume that (6.3) holds but (6.4) fails. Hence we can find $\tilde{z}$ such that there holds

$$M^*(1+|Du|)(\tilde{z}) \le \lambda \qquad \text{and} \qquad \tilde{z} \in \tilde{\mathcal{Q}}\,. \quad (6.5)$$

Since $\tilde{\mathcal{Q}} \subset 3\mathcal{Q} \subset n^2\mathcal{Q}_0$, and trivially $\tilde{z} \in 3\mathcal{Q}$ we have

$$\fint_{3\mathcal{Q}} (1+|Du|)\,dz \le M^*(1+|Du|)(\tilde{z}) \le \lambda\,. \quad (6.6)$$

Moreover, from (6.3) we infer the existence of $\bar{z}$ satisfying

$$M^*_1(g)(\bar{z}) \le \varepsilon\lambda \qquad \text{and} \qquad \bar{z} \in \mathcal{Q}\,. \quad (6.7)$$

Now, let $\mathcal{C}$ denote – in the sense of Section 2.2 – the unique parabolic cylinder having $3\mathcal{Q}$ as inner cylinder, i.e. $\mathcal{Q}_{\text{inn}}(\mathcal{C}) = 3\mathcal{Q}$. If $\mathcal{Q} = Q(x_1,\varrho) \times (t_1-\varrho^2, t_1+\varrho^2)$ then $\mathcal{C}$ is given by $B(x_1, 3\sqrt{n}\varrho) \times (t_1 - (3\sqrt{n}\varrho)^2, t_1 + (3\sqrt{n}\varrho)^2)$. It is easy to check that $\mathcal{C} \subset n^2\mathcal{Q}_0$. Next we denote by

$$\begin{aligned} v \in\, & C^0\big([t_1-(3\sqrt{n}\varrho)^2, t_1+(3\sqrt{n}\varrho)^2]; L^2(B(x_1,3\sqrt{n}\varrho)\big) \\ & \cap L^2\big(t_1-(3\sqrt{n}\varrho)^2, t_1+(3\sqrt{n}\varrho)^2; W^{1,2}(B(x_1,3\sqrt{n}\varrho)\big) \end{aligned}$$

the unique solution to the homogeneous Cauchy-Dirichlet problem (5.2)

$$\begin{cases} v_t - \operatorname{div} a(z,Dv) = 0 & \text{in } \mathcal{C}, \\ v = u & \text{on } \partial_{\text{par}}\mathcal{C}. \end{cases} \quad (6.8)$$

We consider the outer parabolic cylinder to $\mathcal{C}$, i.e. $\mathcal{Q}_{\text{out}}(\mathcal{C}) = Q(x_1, 3\sqrt{n}\varrho) \times (t_1 - (3\sqrt{n}\varrho)^2, t_1 + (3\sqrt{n}\varrho)^2)$, which satisfies also $\mathcal{Q}_{\text{out}}(\mathcal{C}) \subset n^2\mathcal{Q}_0$. Then the definition of the fractional maximal operator M^*_1 and (6.7), i.e. $\bar{z} \in \mathcal{Q} \subset \mathcal{Q}_{\text{out}}(\mathcal{C})$ and $M^*_1(g)(\bar{z}) \le \varepsilon\lambda$, yield that

$$|\mathcal{C}|^{\frac{1}{N}} \fint_{\mathcal{C}} |g|\,dz \le \left(\frac{|\mathcal{Q}_{\text{out}}(\mathcal{C})|}{|\mathcal{C}|}\right)^{1-\frac{1}{N}} |\mathcal{Q}_{\text{out}}(\mathcal{C})|^{\frac{1}{N}} \fint_{\mathcal{Q}_{\text{out}}(\mathcal{C})} |g|\,dz \le c(n)\varepsilon\lambda. \quad (6.9)$$

Combining (6.9) with the universal comparison estimate from (5.3) we obtain

$$\int_{3\mathcal{Q}} |Du - Dv| \, dz \le \int_{\mathcal{C}} |Du - Dv| \, dz \le c\, \nu^{-1} |\mathcal{C}|^{\frac{1}{N}} \int_{\mathcal{C}} |g| \, dz \le c\, \varepsilon \lambda |\mathcal{C}|,$$

for a constant $c = c(n)/\nu$. Using $|\mathcal{C}| = c(n)|3\mathcal{Q}|$ in the preceding inequality we arrive at

$$\fint_{3\mathcal{Q}} |Du - Dv| \, dz \le c(n,\nu)\varepsilon\lambda. \tag{6.10}$$

Next, we observe that the hypothesis of Theorem 5.5 are fulfilled for the solution v to the homogeneous Cauchy-Dirichlet problem (5.2) on $\mathcal{C}$. Therefore, we have the local higher integrability of Dv on $2\mathcal{Q} \subset 3\mathcal{Q} \subset \mathcal{C}$ with the estimate

$$\left(\fint_{2\mathcal{Q}} (1 + |Dv|)^{2\chi} \, dz \right)^{\frac{1}{2\chi}} \le c(n,\nu,L) \fint_{3\mathcal{Q}} (1 + |Dv|) \, dz, \tag{6.11}$$

where $\chi = \chi(n,\nu,L)$ is the higher integrability exponent introduced in Theorem 5.5. Using the comparison estimate (6.10), (6.6) and $0 < \varepsilon \le 1$ the right-hand side of the preceding inequality is estimated as follows:

$$\fint_{3\mathcal{Q}} (1 + |Dv|) \, dz \le \fint_{3\mathcal{Q}} (1 + |Du|) \, dz + \fint_{3\mathcal{Q}} |Du - Dv| \, dz \le \lambda + c\,\varepsilon\lambda \le c\,\lambda\,,$$

with a constant $c = c(n,\nu)$. Combining the preceding estimate with (6.11) yields

$$\fint_{2\mathcal{Q}} (1 + |Dv|)^{2\chi} \, dz \le c(n,\nu,L)\lambda^{2\chi}\,. \tag{6.12}$$

In order to proceed further we use the restricted maximal operator on $2\mathcal{Q}$ and here we abbreviate $M^{**} := M^*_{0,2\mathcal{Q}}$. Using (4.2) twice, (6.12) and (6.10) we obtain

$$\begin{aligned}
&\big|\{z \in \mathcal{Q} : M^{**}(1 + |Du|)(z) > AS\lambda\}\big| \\
&\quad \le \big|\{z \in \mathcal{Q} : M^{**}(1 + |Dv|)(z) > \tfrac{1}{2}AS\lambda\}\big| \\
&\qquad\qquad + \big|\{z \in \mathcal{Q} : M^{**}(|Du - Dv|)(z) > \tfrac{1}{2}AS\lambda\}\big| \\
&\quad \le \frac{c(n,\chi)}{(AS\lambda)^{2\chi}} \int_{2\mathcal{Q}} (1 + |Dv|)^{2\chi} \, dz + \frac{c(n)}{AS\lambda} \int_{2\mathcal{Q}} |Du - Dv| \, dz \\
&\quad \le \frac{c(n,\nu,L)}{(AS)^{2\chi}} |2\mathcal{Q}| + \frac{c(n,\nu)\varepsilon}{AS} |3\mathcal{Q}| \\
&\quad = \Big[\frac{c_1(n,\nu,L)}{(AS)^{2\chi}} + \frac{c_2(n,\nu)\varepsilon}{AS} \Big] |\mathcal{Q}|.
\end{aligned} \tag{6.13}$$

Having arrived at this stage we perform the following choices of A and ε: We first choose $A = A(n,\nu,L) > 1$ such that

$$A = 4 \cdot 10^N \big[1 + c_1(n,\nu,L)\big] \quad \Longrightarrow \quad \frac{c_1}{(AS)^{2\chi}} \le \frac{1}{4S^{2\chi}}. \tag{6.14}$$

Then we the choose $\varepsilon = \varepsilon(n,L,\nu,S) \in (0,1)$ such that

$$\varepsilon = \frac{1}{4S^{2\chi-1}[1 + c_2]} \quad \Longrightarrow \quad \frac{c_2\varepsilon}{AS} \le \frac{1}{4S^{2\chi}}. \tag{6.15}$$

Using these choices in (6.13) we find that

$$\big|\{z \in \mathcal{Q} : M^{**}(1 + |Du|)(z) > AS\lambda\}\big| < S^{-2\chi} |\mathcal{Q}|\,. \tag{6.16}$$

At this stage it remains to replace in (6.16) the restricted maximal operator $M^{**} = M^*_{0,2\mathcal{Q}}$ by the restricted maximal operator $M^* = M^*_{0,n^2\mathcal{Q}_0}$. Let $\ell = 2\varrho$ be the side-length of $\mathcal{Q}$

and $z \in \mathcal{Q}$ arbitrary. Moreover, let $\widehat{\mathcal{Q}}$ denote an arbitrary parabolic cylinder with side-length $\hat{\ell} = 2\hat{\varrho}$ contained in $n^2\mathcal{Q}_0$ and containing the point z. We distinguish two cases: **In the case $\hat{\ell} \leq \frac{1}{2}\ell$** we have $\widehat{\mathcal{Q}} \subset 2\mathcal{Q} \subset n^2\mathcal{Q}_0$ and therefore

$$\fint_{\widehat{\mathcal{Q}}} (1 + |Du|)\, dz \leq M^{**}(1 + |Du|)(z)\,. \tag{6.17}$$

In **the other case $2\hat{\ell} > \ell$** or equivalently $2\hat{\varrho} > \varrho$, it is possible to enlarge the cylinder $\widehat{\mathcal{Q}}$ to another cylinder $\mathcal{Q}'$ in such that $\widehat{\mathcal{Q}} \subset \mathcal{Q}' \subset n^2\mathcal{Q}_0$, $|\mathcal{Q}'| \leq 5^N|\widehat{\mathcal{Q}}|$ and finally $\widetilde{\mathcal{Q}} \subset \mathcal{Q}'$, where $\widetilde{\mathcal{Q}}$ is the predecessor of $\mathcal{Q}$. In particular we have $\tilde{z} \in \mathcal{Q}'$. Therefore, we find

$$\fint_{\widehat{\mathcal{Q}}} (1 + |Du|)\, dz \leq 5^N \fint_{\mathcal{Q}'} (1 + |Du|)\, dz \leq 5^N\lambda,$$

where we have also used (6.6). Since $\widetilde{\mathcal{Q}}$ is an arbitrary cylinder in $n^2\mathcal{Q}_0$ we have shown

$$M^*(1 + |Du|)(z) \leq \max\left\{M^{**}(1 + |Du|)(z),\ 5^N\lambda\right\} \qquad \forall\, z \in \mathcal{Q}.$$

Combining the preceding inequality with (6.16) and the particular choice of A in (6.14) leads us to the estimate

$$\left|\{z \in \mathcal{Q} : M^*(1 + |Du|)(z) > AS\lambda\}\right| < S^{-2\chi}|\mathcal{Q}|, \tag{6.18}$$

which contradicts (6.3) and therefore proves the assertion of the Lemma. □

Step 2: Application of Proposition 2.1. Let $\mathcal{Q}_0$ as in Step 1. Then, we define

$$\lambda_0 := 2c_0(n)n^{2N}S^{2\chi} \fint_{n^2\mathcal{Q}_0} (1 + |Du|)\, dz\,, \tag{6.19}$$

where c_0 is taken from (4.2). Obviously we have that $\lambda_0 > 0$. The strategy of proof is now to apply Lemma 6.2 for the choice $\lambda := (AS)^k\lambda_0$ for $k \in \mathbb{N}_0$. We first show that the hypotheses of Lemma 6.2 are fulfilled for every $k \in \mathbb{N}_0$. Using (4.2) and (6.19) we infer that

$$\begin{aligned}
&\left|\{z \in \mathcal{Q}_0 : M^*(1 + |Du|)(z) > (AS)^k\lambda_0\}\right| \\
&\qquad\qquad \leq \left|\{z \in \mathcal{Q}_0 : M^*(1 + |Du|)(z) > \lambda_0\}\right| \\
&\qquad\qquad \leq \frac{c_0(n)}{\lambda_0}\int_{\mathcal{Q}_0} (1 + |Du|)\, dz < S^{-2\chi}|\mathcal{Q}_0|\,.
\end{aligned}$$

In the light of Lemma 6.2 we can therefore apply Proposition 2.1 with $\delta := S^{-2\chi}$,

$$X := \left\{z \in \mathcal{Q}_0 : M^*(1 + |Du|)(z) > (AS)^{k+1}\lambda_0 \ \text{ and } \ M_1^*(g)(z) \leq (AS)^k\varepsilon\lambda_0\right\}$$

and

$$Y := \left\{z \in \mathcal{Q}_0 : M^*(1 + |Du|)(z) > (AS)^k\lambda_0\right\}.$$

The application of Proposition 2.1 and the definition of X and Y yield that

$$\begin{aligned}
&\left|\left\{z \in \mathcal{Q}_0 : M^*(1 + |Du|)(z) > (AS)^{k+1}\lambda_0\right\}\right| \\
&\qquad \leq S^{-2\chi}\left|\left\{z \in \mathcal{Q}_0 : M^*(1 + |Du|)(z) > (AS)^k\lambda_0\right\}\right| \\
&\qquad\qquad + \left|\left\{z \in \mathcal{Q}_0 : M_1^*(g)(z) > (AS)^k\varepsilon\lambda_0\right\}\right|
\end{aligned} \tag{6.20}$$

holds for every $k \in \mathbb{N}_0$. With

$$\begin{aligned}
\mu_1(H) &:= \left|\left\{z \in \mathcal{Q}_0 : M^*(1 + |Du|)(z) > H\right\}\right|, \\
\mu_2(H) &:= \left|\left\{z \in \mathcal{Q}_0 : M_1^*(g)(z) > H\right\}\right|,
\end{aligned}$$

the preceding inequality turns into

$$\mu_1\left((AS)^{k+1}\lambda_0\right) \leq S^{-2\chi}\mu_1\left((AS)^k\lambda_0\right) + \mu_2\left((AS)^k\varepsilon\lambda_0\right). \tag{6.21}$$

Multiplying (6.21) by $(AS)^{\frac{(k+1)\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}$ we find

$$\begin{aligned}(AS)^{\frac{(k+1)\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1\left((AS)^{k+1}\lambda_0\right)&\\ \leq A^{\frac{\theta\gamma}{\theta-\gamma}}S^{\frac{\theta\gamma}{\theta-\gamma}-2\chi}(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1\left((AS)^k\lambda_0\right)&\\ +\left(AS/\varepsilon\right)^{\frac{\theta\gamma}{\theta-\gamma}}(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}\left(\varepsilon\lambda_0\right)^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2\left((AS)^k\varepsilon\lambda_0\right).&\end{aligned}\tag{6.22}$$

Note that $\gamma<\theta$ by (1.7) and therefore $\frac{\theta\gamma}{\theta-\gamma}>0$. On the other hand, condition $(1.7)_1$, i.e. $\gamma\leq\frac{2\theta}{\theta+2}$, is equivalent to require $\frac{\theta\gamma}{\theta-\gamma}\leq 2$, and therefore, since $\chi>1$, we have

$$d:=2\chi-\frac{\theta\gamma}{\theta-\gamma}\geq 2(\chi-1)>0.\tag{6.23}$$

We now choose

$$S:=\left[4A^{\frac{\theta\gamma}{\theta-\gamma}}\right]^{\frac{1}{d}},\tag{6.24}$$

where A has been determined in (6.14). Note that $y^{\frac{\theta\gamma}{\theta-\gamma}}\leq y^2$, whenever $y\geq 1$, and therefore $S\leq\left[4A^2\right]^{\frac{1}{d}}\leq\left[4A^2\right]^{\frac{1}{2(\chi-1)}}=[2A]^{\frac{1}{\chi-1}}$. Recalling the dependencies of A and χ, i.e. $A=A(n,L,\nu)$ and $\chi=\chi(n,L,\nu)$ we easily infer that S from (6.24) is bounded by a universal constant depending on n,L,ν. On the other hand, we have the estimate

$$AS/\varepsilon=4\,[1+c_2]AS^{2\chi}\leq 2\,[1+c_2](2A)^{1+\frac{2\chi}{\chi-1}}=\tfrac{1}{2}c_*(n,L,\nu)\,,$$

so that $\left(AS/\varepsilon\right)^{\frac{\theta\gamma}{\theta-\gamma}}\leq(c_*/2)^{\frac{\theta\gamma}{\theta-\gamma}}$. Using this and (6.24) in (6.22) we conclude that

$$\begin{aligned}(AS)^{\frac{(k+1)\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1\left((AS)^{k+1}\lambda_0\right)&\\ \leq\frac{1}{4}(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1\left((AS)^k\lambda_0\right)&\\ +\left(\tfrac{c_*}{2}\right)^{\frac{\theta\gamma}{\theta-\gamma}}(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}\left(\varepsilon\lambda_0\right)^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2\left((AS)^k\varepsilon\lambda_0\right)&\end{aligned}\tag{6.25}$$

holds for every $k\in\mathbb{N}_0$.

Step 3: Parabolic Lorentz spaces estimates on level sets. We take $\tau\in(0,\infty)$ and raise the terms appearing in (6.25) to the power $\frac{\tau(\theta-\gamma)}{\theta\gamma}$. This leads us to

$$\begin{aligned}\left[(AS)^{\frac{(k+1)\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1\left((AS)^{k+1}\lambda_0\right)\right]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}&\\ \leq\max\left\{(1/4)^{\frac{\tau(\theta-\gamma)}{\theta\gamma}},(1/2)\right\}\left[(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1\left((AS)^k\lambda_0\right)\right]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}&\\ +c_*^\tau\left[(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}\left(\varepsilon\lambda_0\right)^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2\left((AS)^k\varepsilon\lambda_0\right)\right]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}.&\end{aligned}$$

We note that here we have also used that $\frac{1}{2}\leq\frac{\theta-\gamma}{\theta\gamma}<1$ and therefore $2^{\frac{\tau(\theta-\gamma)}{\theta\gamma}-1}\leq 2^\tau$ in order to obtain the constant c_*^τ in the second term. Now, we sum up the preceding inequality for $k=0,\dots,H$ and finally add the quantity $\lambda_0^\tau|\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}$ to both sides. In this way we obtain:

$$I_1(H)\leq\lambda_0^\tau|\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}+\max\left\{(1/4)^{\frac{\tau(\theta-\gamma)}{\theta\gamma}},(1/2)\right\}I_1(H)+c_*^\tau I_2(\infty)\,,$$

where we have defined

$$I_1(H) := \sum_{k=0}^{H+1} \left[(AS)^{\frac{k\theta\gamma}{\theta-\gamma}} \lambda_0^{\frac{\theta\gamma}{\theta-\gamma}} \mu_1\left((AS)^k\lambda_0\right)\right]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} \quad \text{and}$$

$$I_2(\infty) := \sum_{k=0}^{\infty} \left[(AS)^{\frac{k\theta\gamma}{\theta-\gamma}} (\varepsilon\lambda_0)^{\frac{\theta\gamma}{\theta-\gamma}} \mu_2\left((AS)^k\varepsilon\lambda_0\right)\right]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} .$$

We note that $I_2 = I_2(\infty)$ is finite since $g(= g_k) \in L^\infty(\Omega_T)$, but this fact is not needed here. Re-absorbing the second therm appearing in the right-hand side of the preceding inequality on the left and then letting $H \to \infty$ yields

$$I_1(\infty) \le c_1 c_*^\tau \left[\lambda_0^\tau |\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + I_2(\infty)\right], \tag{6.26}$$

where we have abbreviated

$$c_1 := \left[1 - \max\left\{(1/4)^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}, (1/2)\right\}\right]^{-1}.$$

Using the fact $\gamma \le \frac{2\theta}{\theta+2}$ we see that $(1/4)^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} \le (1/2)^\tau$ so that the constant c_1 in (6.26) can be replaced by the larger quantity $[1 - \max\{(1/2)^\tau, (1/2)\}]^{-1}$. This implies in particular that $c_1 c_*^\tau$ can be replaced by a constant of the form c^τ where $c(n, L, \nu, \tau) := c_*[1-\max\{(1/2)^\tau, (1/2)\}]^{-1/\tau}$. We note that c is a decreasing function of τ with $c \to \infty$ when $\tau \downarrow 0$ and $c \to c_*$ when $\tau \to \infty$. Therefore, $c = c(\tau)$ stays bounded on any interval $[\tau_0, \infty)$ with $\tau_0 > 0$. We will keep this kind of dependence for the rest of the proof. Now, if $k \ge 0$ and $\lambda \in [(AS)^k\lambda_0, (AS)^{k+1}\lambda_0)$ then

$$\lambda^{\frac{\theta\gamma}{\theta-\gamma}} \mu_1(\lambda) \le (AS)^{\frac{(k+1)\theta\gamma}{\theta-\gamma}} \lambda_0^{\frac{\theta\gamma}{\theta-\gamma}} \mu_1\big((AS)^k\lambda_0\big), \tag{6.27}$$

and similarly when $k \ge 1$ and $\lambda \in [(AS)^{k-1}\varepsilon\lambda_0, (AS)^k\varepsilon\lambda_0)$ we have

$$(AS)^{\frac{(k-1)\theta\gamma}{\theta-\gamma}} (\varepsilon\lambda_0)^{\frac{\theta\gamma}{\theta-\gamma}} \mu_2\big((AS)^k\varepsilon\lambda_0\big) \le \lambda^{\frac{\theta\gamma}{\theta-\gamma}} \mu_2(\lambda). \tag{6.28}$$

Using (6.27) we find

$$\begin{aligned}
&\int_0^\infty \left[\lambda^{\frac{\theta\gamma}{\theta-\gamma}} \mu_1(\lambda)\right]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} \frac{d\lambda}{\lambda} \\
&\quad = \int_0^{\lambda_0} \dots \frac{d\lambda}{\lambda} + \sum_{k=0}^{\infty} \int_{(AS)^k\lambda_0}^{(AS)^{k+1}\lambda_0} \dots \frac{d\lambda}{\lambda} \\
&\quad \le \frac{\lambda_0^\tau}{\tau} |\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + (AS)^\tau \log(AS) \sum_{k=0}^{\infty} \left[(AS)^{\frac{k\theta\gamma}{\theta-\gamma}} \lambda_0^{\frac{\theta\gamma}{\theta-\gamma}} \mu_1\left((AS)^k\lambda_0\right)\right]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} \\
&\quad = \frac{\lambda_0^\tau}{\tau} |\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + (AS)^\tau \log(AS) I_1(\infty),
\end{aligned}$$

and similarly using (6.28) and the fact that we have chosen $\varepsilon \le 1$ we infer that

$$\begin{aligned}
I_2(\infty) &= (\lambda_0\varepsilon)^\tau |\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + \sum_{k=1}^{\infty} \left[(AS)^{\frac{k\theta\gamma}{\theta-\gamma}} (\varepsilon\lambda_0)^{\frac{\theta\gamma}{\theta-\gamma}} \mu_2\left((AS)^k\varepsilon\lambda_0\right)\right]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} \\
&\le \lambda_0^\tau |\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + \frac{(AS)^\tau}{\log(AS)} \sum_{k=1}^{\infty} \int_{(AS)^{k-1}\varepsilon\lambda_0}^{(AS)^k\varepsilon\lambda_0} \left[\lambda^{\frac{\theta\gamma}{\theta-\gamma}} \mu_2(\lambda)\right]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} \frac{d\lambda}{\lambda} \\
&= \lambda_0^\tau |\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + \frac{(AS)^\tau}{\log(AS)} \int_0^\infty \left[\lambda^{\frac{\theta\gamma}{\theta-\gamma}} \mu_2(\lambda)\right]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} \frac{d\lambda}{\lambda}.
\end{aligned}$$

Combining the preceding estimates with (6.26) we conclude

$$\begin{aligned}
\int_0^\infty & \Big[\lambda^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1(\lambda)\Big]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}\frac{d\lambda}{\lambda} \\
&\le \frac{\lambda_0^\tau}{\tau}|\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + c^\tau (AS)^\tau \log(AS)\Big[\lambda_0^\tau|\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + I_2(\infty)\Big] \\
&= \Big[\frac{1}{\tau} + c^\tau (AS)^\tau \log(AS)\Big]\lambda_0^\tau|\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + c^\tau (AS)^\tau \log(AS) I_2(\infty) \\
&\le c_2^\tau\lambda_0^\tau|\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + c_2^\tau\int_0^\infty \Big[\lambda^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2(\lambda)\Big]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}\frac{d\lambda}{\lambda}\,, \qquad (6.29)
\end{aligned}$$

where we have abbreviated $c_2 = \max\big\{\big[\frac{1}{\tau} + 2c^\tau (AS)^\tau \log(AS)\big]^{\frac{1}{\tau}}, c\,(AS)^2\big\}$. As for the constant c the constant $c_2 = c_2(n,L,\nu,\tau)$ blows up when $\tau\downarrow 0$, while c_2 remains bounded when τ is bounded away from zero. Taking into account the definition (3.1) and (3.9) the preceding inequality turns into

$$\begin{aligned}
\|M^*&(1+|Du|)\|_{L(\frac{\theta\gamma}{\theta-\gamma},\tau)(\mathcal{Q}_0)} \\
&= \bigg(\frac{\theta\gamma}{\theta-\gamma}\int_0^\infty \Big[\lambda^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1(\lambda)\Big]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}\frac{d\lambda}{\lambda}\bigg)^{\frac{1}{\tau}} \\
&\le \bigg[\frac{\theta\gamma}{\theta-\gamma}c_2^\tau\lambda_0^\tau|\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + \frac{\theta\gamma}{\theta-\gamma}c_2^\tau\int_0^\infty \Big[\lambda^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2(\lambda)\Big]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}\frac{d\lambda}{\lambda}\bigg]^{\frac{1}{\tau}} \\
&\le c(\tau)c_2\bigg[\Big(\frac{\theta\gamma}{\theta-\gamma}\Big)^{\frac{1}{\tau}}\lambda_0|\mathcal{Q}_0|^{\frac{\theta-\gamma}{\theta\gamma}} + \bigg(\frac{\theta\gamma}{\theta-\gamma}\int_0^\infty \Big[\lambda^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2(\lambda)\Big]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}}\frac{d\lambda}{\lambda}\bigg)^{\frac{1}{\tau}}\bigg] \\
&\le c(\tau)c_2\Big[\lambda_0|\mathcal{Q}_0|^{\frac{\theta-\gamma}{\theta\gamma}} + \|M_1^*(g)\|_{L(\frac{\theta\gamma}{\theta-\gamma},\tau)(\mathcal{Q}_0)}\Big]\,,
\end{aligned}$$

where $c(\tau) = 4^{1/\tau}$. Here we have used in the last line that $\big(\frac{\theta\gamma}{\theta-\gamma}\big)^{\frac{1}{\tau}} \le 2^{1/\tau}$. With the obvious inequality $|Du(z)| \le M^*(1+|Du|)(z)$ for almost every $z\in\mathcal{Q}_0$ we conclude from the preceding inequality that

$$\|Du\|_{L(\frac{\theta\gamma}{\theta-\gamma},\tau)(\mathcal{Q}_0)} \le c\lambda_0|\mathcal{Q}_0|^{\frac{\theta-\gamma}{\theta\gamma}} + c\|M_1^*(g)\|_{L(\frac{\theta\gamma}{\theta-\gamma},\tau)(\mathcal{Q}_0)}, \qquad (6.30)$$

where $c = c(n,L,\nu,\tau)$ stays bounded as long τ is bounded away from zero, and $c\to\infty$ when $\tau\downarrow 0$. For $0<q<\infty$ the choice $\tau = \frac{\theta q}{\theta-\gamma}$ in (6.30) yields

$$\|Du\|_{L(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{Q}_0)} \le c\lambda_0|\mathcal{Q}_0|^{\frac{\theta-\gamma}{\theta\gamma}} + c\|M_1^*(g)\|_{L(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{Q}_0)}\,. \qquad (6.31)$$

Having arrived at this stage we can apply Theorem 4.13 with $\beta = 1$ and $p=\gamma$ (note that $\beta p = \gamma < \theta$) passing to the outer parabolic cylinder $\mathcal{C}$, i.e. $\mathcal{Q}_{\mathrm{inn}}(\mathcal{C}) = \mathcal{Q}_0$ and choosing $s=2$. Note that, if $\mathcal{Q}_0 = Q_R\times(-R^2,R^2)$ then $\mathcal{C} = B_{\sqrt{n}R}\times(-nR^2,nR^2)$ and $2\mathcal{C} = B_{2\sqrt{n}R}\times(-4nR^2,4nR^2)\subset Q_{n^2R}\times(-n^4R^2,n^4R^2) = n^2\mathcal{Q}_0$. Furthermore, $|\mathcal{Q}_0|\le 1$ implies $R<1$, so that $\sqrt{n}R<\sqrt{n}$. Hence, the application of Remark 4.10 at the end of the proof of Theorem 4.13 yields a constant $\max\{1,\sqrt{n}\}^{(N-\theta)\frac{\theta-\gamma}{\theta\gamma}} \le \sqrt{n}^{\frac{n(\theta-\gamma)}{\theta\gamma}} \le c(n,\gamma)$. Applying Theorem 4.13 with the choice of the parameters described before leads to

$$\|M_1^*(g)\|_{L(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{Q}_0)} \le c(n,\gamma,q)\|g\|_{L^\theta(\gamma,q)(n^2\mathcal{Q}_0)}. \qquad (6.32)$$

Combining (6.32) and (6.19), i.e. the choice of λ_0, with (6.31) and noting that $S^{2\chi}\le c(n,L,\nu)$ by the choice in (6.24), we finally arrive at

$$\|Du\|_{L(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{Q}_0)} \le c\bigg(\fint_{n^2\mathcal{Q}_0}(1+|Du|)\,dz\bigg)|\mathcal{Q}_0|^{\frac{\theta-\gamma}{\theta\gamma}} + c\|g\|_{L^\theta(\gamma,q)(n^2\mathcal{Q}_0)}, \qquad (6.33)$$

where $c = c(n, L, \nu, \gamma, q)$.

We now show how the previous inequality, i.e. (6.33), can be extended to the case $q = \infty$. We proceed as follows: We first go back to (6.25) and obtain for $H \in \mathbb{N}$ that

$$\begin{aligned} I_3(H) &:= \sup_{0\le k\le H+1} (AS)^{\frac{k\theta\gamma}{\theta-\gamma}} \lambda_0^{\frac{\theta\gamma}{\theta-\gamma}} \mu_1\left((AS)^k\lambda_0\right) \\ &\le \lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}|\mathcal{Q}_0| + \frac{1}{4}I_3(H) + c_*^2 \sup_{k\ge 0}(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}(\varepsilon\lambda_0)^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2\left((AS)^k\varepsilon\lambda_0\right). \end{aligned}$$

Here we have used that $c_*^{\frac{\theta\gamma}{\theta-\gamma}} \le c_*^2$, since $\frac{\theta\gamma}{\theta-\gamma} \le 2$ and $c_* \ge 1$. Re-absorbing as usual $\frac{1}{4}I_3(H)$ in the left-hand side, and then letting $H \to \infty$ we deduce

$$\begin{aligned} I_3(\infty) &\le (4/3)\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}|\mathcal{Q}_0| + c\sup_{k\ge 0}(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}(\varepsilon\lambda_0)^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2\left((AS)^k\varepsilon\lambda_0\right) \\ &\le (4/3)\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}|\mathcal{Q}_0| + c\sup_{\lambda>0}\lambda^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2(\lambda). \end{aligned}$$

Here we have used (6.28) in the last line. On the other hand using (6.27) we can bound the left-hand side of the preceding inequality from below and obtain

$$\sup_{\lambda>0}\lambda^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1(\lambda) \le \lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}|\mathcal{Q}_0| + (AS)^{\frac{\theta\gamma}{\theta-\gamma}}I_3(\infty).$$

We note that $(AS)^{\frac{\theta\gamma}{\theta-\gamma}}$ can be bounded by a constant $c = c(n, L, \nu)$. Combining the last two inequalities we have

$$\sup_{\lambda>0}\lambda^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1(\lambda) \le c\,\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}|\mathcal{Q}_0| + c\sup_{\lambda>0}\lambda^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2(\lambda),$$

where $c = c(n, L, \nu)$. Taking into account the definition of the Marcinkiewicz space from (3.2) and again the obvious a.e. estimate $|Du(z)| \le M^*(1+|Du|)(z)$ we conclude that

$$\|Du\|_{\mathcal{M}^{\frac{\theta\gamma}{\theta-\gamma}}(\mathcal{Q}_0)} \le c\,\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}|\mathcal{Q}_0| + c\,\|M_1^*(g)\|_{\mathcal{M}^{\frac{\theta\gamma}{\theta-\gamma}}(\mathcal{Q}_0)}. \tag{6.34}$$

Similarly to (6.32) we use Theorem 4.13, now with the choice $q = \infty$, in order to get

$$\|M_1^*(g)\|_{\mathcal{M}^{\frac{\theta\gamma}{\theta-\gamma}}(\mathcal{Q}_0)} \le c(n,\gamma)\,\|g\|_{\mathcal{M}^{\gamma,\theta}(n^2\mathcal{Q}_0)}.$$

Connecting the last two inequalities and recalling the choice of λ_0 from (6.19) we infer that (6.33) extends to the case $q = \infty$.

Proof of Remark 1.3. Note that (6.34) holds also for $\gamma = 1$, when $\theta \ge 2$. In this case it is enough to estimate the Marcinkiewicz norm on the right-hand side with the same norm of $I_1(g)$ and then recall the classical result of Adams [3]: $I_1 : L^{1,\theta} \to \mathcal{M}^{\frac{\theta}{\theta-1}}$. See also Step 6. □

Step 4: Intermediate parabolic Morrey-space regularity of Du.

Proposition 6.3. *Let $u (= u_k) \in C^0([-T, 0]; L^2(\Omega)) \cap L^2(-T, 0; W_0^{1,2}(\Omega))$ be a weak solution to (5.1) where the assumptions (1.2) are in force and let $g \in L^\theta(\gamma, q)(\Omega_T)$ with $2 < 2\gamma \le \theta \le N$. Then, for every pair of concentric parabolic cylinders $\mathcal{C}_\sigma \Subset \mathcal{C}_\varrho \subset \Omega_T$ there holds*

$$\|1+|Du|\|_{L^{1,\frac{\theta-\gamma}{\gamma}}(\mathcal{C}_\sigma)} \le c\,(\varrho-\sigma)^{\frac{\theta-\gamma}{\gamma}-N}\|1+|Du|\|_{L^1(\mathcal{C}_\varrho)} + c\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_\varrho)}, \tag{6.35}$$

where $c = c(n, L, \nu, \gamma, q)$.

Proof. Let $\mathcal{C}_\sigma \Subset \mathcal{C}_\varrho \subset \Omega_T$ be two fixed concentric cylinders and let z_0 be a point in $\mathcal{C}_\sigma$. Moreover let $\mathcal{C}_R(z_0)$ be a parabolic cylinder with $0 < R \leq d_{\rm par}(z_0, \partial\mathcal{C}_\varrho)$, i.e. $\mathcal{C}(z_0,R) \subset \mathcal{C}_\varrho$. Moreover, let $v \in C^0([t_0-R^2,t_0+R^2];L^2(B_R(x_0)) \cap L^2(t_0-R^2,t_0+R^2;W^{1,2}(B_R(x_0))$ be the unique solution to the Cauchy-Dirichlet problem (5.2) in $C_R(z_0)$. Then, using (5.8) for the choice $q=1$ we infer that for any $0<r\leq R$ we have

$$
\begin{aligned}
\int_{\mathcal{C}(z_0,r)} (1+|Du|)\,dz &\leq \int_{\mathcal{C}(z_0,r)} (1+|Dv|)\,dz + \int_{\mathcal{C}(z_0,R)} |Du-Dv|\,dz \\
&\leq c\left(\frac{r}{R}\right)^{N-1+\alpha} \int_{\mathcal{C}(z_0,R)} (1+|Dv|)\,dz + \int_{\mathcal{C}(z_0,R)} |Du-Dv|\,dz\,,
\end{aligned}
$$

where $c=c(n,L,\nu)$ and $\alpha=\alpha(n,L,\nu)\in(0,1/2]$. Now, the second integral appearing on the right-hand side of the preceding inequality is estimated by (5.4) from Lemma 5.3, i.e. we have

$$
\int_{\mathcal{C}(z_0,R)} |Du-Dv|\,dz \leq c\,R^{N-\frac{\theta-\gamma}{\gamma}}\|g\|_{L^\theta(\gamma,q)(\mathcal{C}(z_0,R))},
$$

where $c=c(n,\nu,\gamma)$. Inserting this above yields

$$
\begin{aligned}
&\int_{\mathcal{C}(z_0,r)} (1+|Du|)\,dz \\
&\qquad \leq c\left(\frac{r}{R}\right)^{N-1+\alpha} \int_{\mathcal{C}(z_0,R)} (1+|Du|)\,dz + c\,\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_\varrho)} R^{N-\frac{\theta-\gamma}{\gamma}}\,,
\end{aligned}
$$

for any choice of $0<r\leq R$. We remark that $2\gamma\leq\theta$ implies $\frac{\theta-\gamma}{\gamma}\geq 1$ and therefore $N-\frac{\theta-\gamma}{\gamma}\leq N-1<N-1+\alpha$. This allows us to apply the iteration Lemma 2.3 with

$$
\varphi(r) := \int_{\mathcal{C}(z_0,r)} (1+|Du|)\,dz, \qquad A := c, \qquad \mathcal{B} := c\,\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_\varrho)},
$$

$R_0 := d_{\rm par}(z_0,\partial\mathcal{C}_\varrho)$ and

$$
\delta_0 := N-1+\alpha > N - \frac{\theta-\gamma}{\gamma} =: \delta_1 .
$$

Applying (2.4) and taking the choice $R=R_0>\varrho-\sigma$ we obtain in particular

$$
\begin{aligned}
&\int_{\mathcal{C}(z_0,r)} (1+|Du|)\,dz \\
&\qquad \leq c\Bigg[(\varrho-\sigma)^{\frac{\theta-\gamma}{\gamma}-N} \int_{\mathcal{C}_\varrho} (1+|Du|)\,dz + \|g\|_{L^\theta(\gamma,q)(\mathcal{C}_\varrho)}\Bigg] r^{N-\frac{\theta-\gamma}{\gamma}},
\end{aligned}
\tag{6.36}
$$

whenever $\mathcal{C}(z_0,r)\subset\mathcal{C}_\sigma$, and with a constant $c=c(n,L,\nu,\gamma)$. Here we have used the dependence $\alpha=\alpha(n,L,\nu)$. Inequality (6.36) immediately implies (6.35), and this completes the proof of the Lemma. □

Remark 6.4 (Extensions)**.** Proposition 6.3 holds under the assumption $2\gamma\leq\theta$ and this is implied by the assumptions of Theorem 6.1. In fact, $\gamma\leq\frac{2\theta}{\theta+2}$ and $2<\theta\leq N$ imply that $2\gamma<\theta$. Therefore, the assumption $2\gamma\leq\theta$ can be replaced by the weaker assumption $\gamma(2-\frac{\alpha}{2})<\theta$, where $\alpha>0$ is the exponent from (5.8). This condition serves for $\frac{\theta-\gamma}{\gamma}>1-\frac{\alpha}{2}$ and therefore also for $\delta_1=N-\frac{\theta-\gamma}{\gamma}<N-1+\frac{\alpha}{2}<N-1+\alpha=\delta_0$, which was needed in the proof of Proposition 6.3. When $\gamma=1$ the proof of Proposition 6.3 still works, provided $2\leq\theta\leq N$ and that we use the comparison estimate (5.5) instead of (5.4).

Note that in this case $N-1+\alpha > N-(\theta-1)$. The final outcome is instead of (6.36) the following estimate:

$$\int_{\mathcal{C}_r} (1+|Du|)\,dz \\ \leq c\bigg[(\varrho-\sigma)^{\theta-1-N}\int_{\mathcal{C}_\varrho}(1+|Du|)\,dz + \|g\|_{L^{1,\theta}(\mathcal{C}_\varrho)}\bigg] r^{N-(\theta-1)}\,, \tag{6.37}$$

where now $c=c(n,L,\nu)$.

With respect to u (instead of Du) we have a statement similar to (6.35), assuming

$$1<\gamma<\frac{\theta}{2} \qquad \text{and} \qquad 2<\theta\leq N$$

instead of (1.7). This can be seen as follows: Keeping in mind the notation introduced at the beginning of the proof of Proposition 6.3 we obtain using (5.9) instead of (5.8) and again (5.4) the following decay estimate:

$$\int_{\mathcal{C}(z_0,r)}(r+|u|)\,dz \\ \leq c\Big(\frac{r}{R}\Big)^N\int_{\mathcal{C}(z_0,R)}(R+|u|)\,dz + c\,\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_\varrho)}R^{N-\frac{\theta-2\gamma}{\gamma}}\,,$$

for any $0<r\leq R$. We note that $2\gamma<\theta$ implies that $\frac{\theta-2\gamma}{\gamma}>0$. Therefore we can apply Lemma 2.3 to the quantity $\varphi(r):=\int_{\mathcal{C}(z_0,r)}(r+|u|)\,dz$. The final outcome, that follows along the lines of the proof of (6.36), is

$$\|u\|_{L^{1,\frac{\theta-2\gamma}{\gamma}}(\mathcal{C}_\sigma)} \leq c\,(\varrho-\sigma)^{\frac{\theta-2\gamma}{\gamma}-N}\|\varrho+|u|\|_{L^1(\mathcal{C}_\varrho)} + c\,\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_\varrho)}. \tag{6.38}$$

We note that $c\to\infty$ when $\gamma\uparrow\theta/2$.

Step 5: Full Morrey space regularity of Du. In this section we prove (6.2) for the approximating solutions $u=u_k$. We consider a parabolic cylinder $\mathcal{C}_\varrho\subset\Omega_T$ and scale the problem as in Remark 5.1 to $\mathcal{C}_1=\mathcal{C}(0,1)$, switching from u,g,a to $\tilde u,\tilde g,\tilde a$. Applying (6.33) with $\tilde u,\tilde g$ with $\mathcal{Q}_0:=\mathcal{Q}_{1/n^4}$ (note, with this choice of $\mathcal{Q}_0$ we have $n^2\mathcal{Q}_0=\mathcal{Q}_{1/n^2}\subset\mathcal{C}_{1/n}$) we conclude that

$$\begin{aligned}\|D\tilde u\|_{L(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_{1/n^4})} &\leq c\,\|1+|D\tilde u|\|_{L^1(\mathcal{Q}_{1/n^2})} + c\,\|\tilde g\|_{L^\theta(\gamma,q)(\mathcal{Q}_{1/n^2})}\\ &\leq c\,\|1+|D\tilde u|\|_{L^1(\mathcal{C}_{1/n})} + c\,\|\tilde g\|_{L^\theta(\gamma,q)(\mathcal{C}_{1/n})}\\ &\leq c\,\|1+|D\tilde u|\|_{L^{1,\frac{\theta-\gamma}{\gamma}}(\mathcal{C}_{1/n})} + c\,\|\tilde g\|_{L^\theta(\gamma,q)(\mathcal{C}_{1/n})}\\ &\leq c\,\|1+|D\tilde u|\|_{L^{1,\frac{\theta-\gamma}{\gamma}}(\mathcal{C}_{9/10})} + c\,\|\tilde g\|_{L^\theta(\gamma,q)(\mathcal{C}_1)}.\end{aligned}$$

At this stage we scale back to $\mathcal{C}_\varrho$ and find

$$\|Du\|_{L(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_{\varrho/n^4})} \leq c(n,L,\nu,\gamma,q)\;\Psi(\mathcal{C}_\varrho)\varrho^{(N-\theta)\frac{\theta-\gamma}{\theta\gamma}}\,, \tag{6.39}$$

where we have defined

$$\Psi(\mathcal{C}_\varrho) := \|1+|Du|\|_{L^{1,\frac{\theta-\gamma}{\gamma}}(\mathcal{C}_{9\varrho/10})} + \|g\|_{L^\theta(\gamma,q)(\mathcal{C}_\varrho)}$$

for every choice of $\mathcal{C}_\varrho\subset\Omega_T$. For a general parabolic cylinder $\mathcal{C}_R\subset\Omega_T$ we conclude the proof by means of a covering argument. Let $\mathcal{C}_\varrho\subset\mathcal{C}_{R/2}$ be a parabolic cylinder not necessary concentric to $\mathcal{C}_R$. If $\mathcal{C}_{n^4\varrho}\subset\mathcal{C}_{R/2}$ then applying (6.39) we have

$$\varrho^{(\theta-N)\frac{\theta-\gamma}{\theta\gamma}}\|Du\|_{L(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_\varrho)} \leq c\,\Psi(\mathcal{C}_{n^4\varrho}) \leq c\,\Psi(\mathcal{C}_{R/2}).$$

On the other hand, if $\mathcal{C}_{n^4\varrho}\not\subset\mathcal{C}_{R/2}$ we cover $\mathcal{C}_\varrho$ with a finite number of parabolic cylinders $\mathcal{C}_i$ of radius $\varrho/(8n^4)$ and center in $\mathcal{C}_\varrho$. Note, that the total number of these cylinders is bounded by a constant $m(n)$ independently on the radius ϱ. Moreover, for each i we have $n^4\mathcal{C}_i \subset \mathcal{C}_{3R/4}$. Therefore, (3.10) and (6.39) imply

$$\varrho^{(\theta-N)\frac{\theta-\gamma}{\theta\gamma}}\|Du\|_{L(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_\varrho)} \le c(n,q)\sum_i \varrho^{(\theta-N)\frac{\theta-\gamma}{\theta\gamma}}\|Du\|_{L(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_i)}$$
$$\le c(n,L,\nu,\gamma,q)\sum_i \Psi(n^4\mathcal{C}_i) \le c\, m\Psi(\mathcal{C}_{3R/4}).$$

Together the last two inequalities (recall definition (3.3)) imply that

$$\|Du\|_{L^\theta(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_{R/2})} \le c(n,L,\nu,\gamma,q)\ \Psi(\mathcal{C}_{3R/4})$$
$$= c\left[\|1+|Du|\|_{L^{1,\frac{\theta-\gamma}{\gamma}}(\mathcal{C}_{27R/40})} + \|g\|_{L^\theta(\gamma,q)(\mathcal{C}_{3R/4})}\right]. \tag{6.40}$$

It remains to estimate the first term on the right-hand side of (6.40). For this we proceed by applying (6.35), i.e. Proposition 6.3 (note that the conditions imposed in (1.7), i.e. $1<\gamma\le\frac{2\theta}{\theta+2}$ and $2<\theta\le N$, yield $1<\gamma<\frac{\theta}{2}$ so that the hypothesis $2\gamma\le\theta$ from Proposition 6.3 is fulfilled) with $\sigma := 27R/40$ and $\varrho := R$ and conclude that

$$\|1+|Du|\|_{L^{1,\frac{\theta-\gamma}{\gamma}}(\mathcal{C}_{27R/40})} \le c\,R^{\frac{\theta-\gamma}{\gamma}-N}\|1+|Du|\|_{L^1(\mathcal{C}_R)} + c\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_R)},$$

and inserting this into (6.40) yields

$$\|Du\|_{L^\theta(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_{R/2})} \le c\,R^{\frac{\theta-\gamma}{\gamma}-N}\|1+|Du|\|_{L^1(\mathcal{C}_R)} + c\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_R)},$$

where $c = c(n,L,\nu,\gamma,q)$. Recalling that in the preceding estimate we have $u\equiv u_k$, $g\equiv g_k$, where u_k is the approximating solution from (2.2) with right-hand side $g_k\in L^\infty(\Omega_T)$ we see that we have established

$$\|Du_k\|_{L^\theta(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_{R/2})} \le c\,R^{\frac{\theta-\gamma}{\gamma}-N}\|1+|Du_k|\|_{L^1(\mathcal{C}_R)} + c\|g_k\|_{L^\theta(\gamma,q)(\mathcal{C}_R)}\,,$$

with a constant $c = c(n,L,\nu,\gamma,q)$ independent of k. Since the right-hand side $g_k\in L^\infty(\Omega_T)$ of the approximating problems is constructed in such a way that it satisfies $|g_k|\le|g|$ we have $\|g_k\|_{L^\theta(\gamma,q)(\mathcal{C}_R)} \le \|g\|_{L^\theta(\gamma,q)(\mathcal{C}_R)}$ (see (2.3)). Therefore the preceding estimate can be replaced by

$$\|Du_k\|_{L^\theta(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_{R/2})} \le c\,R^{\frac{\theta-\gamma}{\gamma}-N}\|1+|Du_k|\|_{L^1(\mathcal{C}_R)} + c\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_R)}, \tag{6.41}$$

and this is exactly the estimate for the approximating solutions we were looking for.

Step 6: Approximation and conclusion. The proof of (6.2) and therefore the one of Theorem 6.1 follows by the use of the lower semi-continuity of the Lorentz-Morrey-norm with respect to a.e. convergence. In fact, the approximating solutions u_k converge as $k\to\infty$ to the solution u in $L^1(-T,0;W_0^{1,1}(\Omega))$ and a.e. on Ω_T (see Section 2.1). Therefore, we can pass to the limit $k\to\infty$ in (6.41) using the lower semi-continuity of the Lorentz-Morrey-norms from Section 3.2. This finishes the proof of Theorem 6.1 and therefore also of Theorem 1.1. □

Proof of Theorem 1.4. We first recall that $\gamma = \frac{2\theta}{\theta+2}$ is equivalent to $\frac{\theta\gamma}{\theta-\gamma} = 2$. Since $\theta\ge 2$ we have $\frac{2\theta}{\theta+2}\le\frac{\theta}{2}<\frac{\theta}{2-\alpha/2}$. Therefore the assumption $\frac{2\theta}{\theta+2}<\gamma$ yields the existence of γ_0 such that

$$\frac{2\theta}{\theta+2}<\gamma_0\le\min\left\{\gamma,\frac{\theta}{2-\alpha/2}\right\} \quad\text{and}\quad d := 2\chi-\frac{\theta\gamma_0}{\theta-\gamma_0}\ge\chi-1, \tag{6.42}$$

where $\chi = \chi(n, L, \nu) > 1$ is the higher integrability exponent from Theorem 5.5 and $\alpha = \alpha(n, L, \nu) \in (0, \frac{1}{2}]$ is the Hölder exponent from the same Theorem. Using Hölder's inequality we easily obtain the following embedding for parabolic Morrey spaces:

$$\|g\|_{L^{\gamma_0,\theta}(\mathcal{C}_R)} \le c(n) R^{\frac{\theta(\gamma-\gamma_0)}{\gamma\gamma_0}} \|g\|_{L^{\gamma,\theta}(\mathcal{C}_R)} \tag{6.43}$$

for any parabolic cylinder $\mathcal{C}_R \subset \Omega_T$. With these preliminaries we proceed along the lines of the proof of Theorem 6.1 taking into account the following changes: We replace γ by γ_0 and choose $q = \gamma_0$. Then everything can be carried out, since (6.23) holds with d defined in (6.42). Moreover, from the definition of γ_0 we see that $\gamma_0(2 - \frac{\alpha}{2}) \le \theta$, so that (6.37) is applicable with γ_0 instead of γ (see Remark 6.4). On the other hand we can also apply Theorem 4.13 in this setting with $p = \gamma_0$ and $\beta = 1$ as in (6.32), since $\beta p = \gamma_0 \le \frac{\theta}{2-\alpha/2} < \theta$ by (6.42). Having arrived at this stage we let

$$h := \frac{\theta\gamma_0}{\theta - \gamma_0},$$

and note that by (6.42) we have $h > 2$. Then (1.9) follows from (6.1) specialized to $\gamma = q = \gamma_0$, and the quantitative estimate (1.10) follows from (6.2) and (6.43) for radii $R \le 1$ as follows:

$$\begin{aligned}
\|Du\|_{L^{h,\theta}(\mathcal{C}_{R/2})} &\le c\, R^{\frac{\theta-\gamma_0}{\gamma_0}-N} \|1 + |Du|\|_{L^1(\mathcal{C}_R)} + c\|g\|_{L^{\gamma_0,\theta}(\mathcal{C}_R)} \\
&\le c\, R^{\frac{\theta-\gamma_0}{\gamma_0}-N} \|1 + |Du|\|_{L^1(\mathcal{C}_R)} + c\, R^{\frac{\theta(\gamma-\gamma_0)}{\gamma\gamma_0}} \|g\|_{L^{\gamma,\theta}(\mathcal{C}_R)} \\
&\le c\, R^{\frac{\theta}{h}-N} \|1 + |Du|\|_{L^1(\mathcal{C}_R)} + c\, \|g\|_{L^{\gamma,\theta}(\mathcal{C}_R)}.
\end{aligned}$$

This completes the proof of Theorem 1.4. □

6.2. **Borderline estimates.** Here we consider the cases that $g \in L\log L(\Omega_T)$, resp. $g \in L\log L^\theta(\Omega_T)$. We start with the case $g \in L^{1,\theta}(\Omega_T) \cap L\log L(\Omega_T)$ and the

Proof of Theorem 1.2. We proceed as in the proof of Theorem 6.1 taking $\gamma = 1$; we note that the proof works with this choice up to (6.29), i.e.

$$\int_0^\infty \Big[\lambda^{\frac{\theta\gamma}{\theta-\gamma}} \mu_1(\lambda)\Big]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} \frac{d\lambda}{\lambda} \quad \le c_2^\tau \lambda_0^\tau |\mathcal{Q}_0|^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} + c_2^\tau \int_0^\infty \Big[\lambda^{\frac{\theta\gamma}{\theta-\gamma}} \mu_2(\lambda)\Big]^{\frac{\tau(\theta-\gamma)}{\theta\gamma}} \frac{d\lambda}{\lambda}.$$

Taking $\tau = \frac{\theta}{\theta-1}$ and recalling that $\gamma = 1$ the preceding inequality turns into the following analogue of (6.30):

$$\int_{\mathcal{Q}_0} |Du|^{\frac{\theta}{\theta-1}}\, dz \le \big(c_2\lambda_0\big)^{\frac{\theta}{\theta-1}} |\mathcal{Q}_0| + c_2^{\frac{\theta}{\theta-1}} \int_{\mathcal{Q}_0} |M_1^*(g)|^{\frac{\theta}{\theta-1}}\, dz.$$

In order to bound the integral appearing on the right-hand side of the preceding inequality we argue as in the proof of Theorem 6.1 (the paragraph before (6.32)) passing to the outer parabolic cylinder $\mathcal{C}$ and then applying Theorem 4.12 for the choice $\beta = 1$ and $s = 2$ instead of Theorem 4.13. Proceeding in this way we have

$$\begin{aligned}
\|M_1^*(g)\|^{\frac{\theta}{\theta-1}}_{L^{\frac{\theta}{\theta-1}}(\mathcal{Q}_0)} &\le c^{\frac{\theta}{\theta-1}}\, |\mathcal{C}|\, \|g\|^{\frac{1}{\theta-1}}_{L^{1,\theta}(2\mathcal{C})} \|g\|_{L\log L(\mathcal{C})} \\
&\le c^{\frac{\theta}{\theta-1}}\, |n^2\mathcal{Q}_0|\, \|g\|^{\frac{1}{\theta-1}}_{L^{1,\theta}(n^2\mathcal{Q}_0)} \|g\|_{L\log L(n^2\mathcal{Q}_0)} \\
&\le c^{\frac{\theta}{\theta-1}}\, |\mathcal{Q}_0|\, \|g\|^{\frac{1}{\theta-1}}_{L^{1,\theta}(n^2\mathcal{Q}_0)} \|g\|_{L\log L(n^2\mathcal{Q}_0)},
\end{aligned}$$

where $c = c(n, L, \nu) \geq 1$. Inserting this in the second last inequality and recalling the definition of λ_0 from (6.19) we find

$$\begin{aligned}
&\bigg(\fint_{\mathcal{Q}_0} |Du|^{\frac{\theta}{\theta-1}} \, dz \bigg)^{\frac{\theta-1}{\theta}} \\
&\quad \leq \Big[(c_2\lambda_0)^{\frac{\theta}{\theta-1}} + (c_2 c)^{\frac{\theta}{\theta-1}} \, \|g\|^{\frac{1}{\theta-1}}_{L^{1,\theta}(n^2\mathcal{Q}_0)} \|g\|_{L\log L(n^2\mathcal{Q}_0)} \Big]^{\frac{\theta-1}{\theta}} \\
&\quad \leq c \Big[\lambda_0 + \|g\|^{\frac{1}{\theta}}_{L^{1,\theta}(n^2\mathcal{Q}_0)} \|g\|^{\frac{\theta-1}{\theta}}_{L\log L(n^2\mathcal{Q}_0)} \Big] \\
&\quad \leq c \fint_{n^2\mathcal{Q}_0} (1 + |Du|) \, dz + c \, \|g\|^{\frac{1}{\theta}}_{L^{1,\theta}(n^2\mathcal{Q}_0)} \|g\|^{\frac{\theta-1}{\theta}}_{L\log L(n^2\mathcal{Q}_0)}, \qquad (6.44)
\end{aligned}$$

with a constant $c = c(n, L, \nu)$. Apart from the fact that the preceding estimate (6.44) holds for the approximating solutions $u = u_k$ on the concentric parabolic cylinders $\mathcal{Q}_0$, $n^2\mathcal{Q}_0$, having in mind (3.6), it has exactly the structure of (1.8) from Theorem 1.2. Therefore, all assertions of Theorem 1.2 follow by a standard covering argument combined with the usual approximation argument. □

Theorem 6.5. *Assume that (1.2) and $g \in L\log L^\theta(\Omega_T)$ with $2 \leq \theta \leq N$ hold. Then the solution $u \in L^1(-T, 0; W_0^{1,1}(\Omega_T))$ to (1.1) satisfies*

$$Du \in L^{\frac{\theta}{\theta-1},\theta}_{\mathrm{loc}}(\Omega_T, \mathbb{R}^n).$$

Moreover, the local quantitative estimate

$$\|Du\|_{L^{\frac{\theta}{\theta-1},\theta}(\mathcal{C}_{R/2})} \leq c \, R^{\theta-1-N} \|1 + |Du|\|_{L^1(\mathcal{C}_R)} + \|g\|_{L\log L^\theta(\mathcal{C}_R)}$$

holds for every parabolic cylinder $\mathcal{C}_R \subset \Omega_T$, with a constant $c = c(n, L, \nu)$.

Proof. Since $g \in L\log L^\theta(\Omega_T)$ we have $g \in L^{1,\theta}(\Omega_T) \cap L\log L(\Omega_T)$. Therefore, the arguments from the proof of Theorem 1.2 apply and we initially end up with (6.44) from above. At this stage we go on using the strategy from the proof of Theorem 6.1, Step 5, and scale everything back to $\mathcal{C}_1$. Using the thereby introduced notation, in particular passing to inner and outer parabolic cylinders, we obtain for the re-scaled function $\tilde{u}$ the following estimate:

$$\begin{aligned}
\|D\tilde{u}\|_{L^{\frac{\theta}{\theta-1}}(\mathcal{C}_{1/n^4})} &\leq c \, \|1 + |D\tilde{u}|\|_{L^1(\mathcal{Q}_{1/n^2})} + c \, \|\tilde{g}\|^{\frac{1}{\theta}}_{L^{1,\theta}(\mathcal{Q}_{1/n^2})} \|\tilde{g}\|^{\frac{\theta-1}{\theta}}_{L\log L(\mathcal{Q}_{1/n^2})} \\
&\leq c \, \|1 + |D\tilde{u}|\|_{L^1(\mathcal{C}_{9/10})} + c \, \|\tilde{g}\|^{\frac{1}{\theta}}_{L^{1,\theta}(\mathcal{C}_1)} \|\tilde{g}\|^{\frac{\theta-1}{\theta}}_{L\log L(\mathcal{C}_1)} \\
&\leq c \, \|1 + |D\tilde{u}|\|_{L^{1,\theta-1}(\mathcal{C}_{9/10})} + c \, \|\tilde{g}\|^{\frac{1}{\theta}}_{L^{1,\theta}(\mathcal{C}_1)} \|\tilde{g}\|^{\frac{\theta-1}{\theta}}_{L\log L(\mathcal{C}_1)} \\
&\leq c \, \|1 + |D\tilde{u}|\|_{L^{1,\theta-1}(\mathcal{C}_{9/10})} + c \, \|\tilde{g}\|_{L\log L^\theta(\mathcal{C}_1)} \, ,
\end{aligned}$$

where $c = c(n, L, \nu)$. Here we have used the simple fact that $\|\tilde{g}\|_{L^{1,\theta}(\mathcal{C}_1)} \lesssim \|\tilde{g}\|_{L\log L^\theta(\mathcal{C}_1)}$ in the last line. Scaling back to $\mathcal{C}_\varrho$ we have

$$\|Du\|_{L^{\frac{\theta}{\theta-1}}(\mathcal{C}_{\varrho/n^4})} \leq c(n, L, \nu) \; \Psi(\mathcal{C}_\varrho) \varrho^{(N-\theta)\frac{\theta-1}{\theta}},$$

where this time we have defined

$$\Psi(\mathcal{C}_\varrho) := \|1 + |Du|\|_{L^{1,\theta-1}(\mathcal{C}_{9\varrho/10})} + \|g\|_{L\log L^\theta(\mathcal{C}_\varrho)}.$$

Having arrived at this stage we can use the covering argument from the proof of Theorem 6.1, Step 5; more precisely, the argument leading us to (6.40) now yields

$$\begin{aligned}
\|Du\|_{L^{\frac{\theta}{\theta-1},\theta}(\mathcal{C}_{R/2})} &\leq c(n, L, \nu) \; \Psi(\mathcal{C}_{3R/4}) \\
&= c \, \Big[\|1 + |Du|\|_{L^{1,\theta-1}(\mathcal{C}_{27R/40})} + \|g\|_{L\log L^\theta(\mathcal{C}_R)} \Big], \qquad (6.45)
\end{aligned}$$

whenever $\mathcal{C}_R \subset \Omega_T$ is a parabolic cylinder. Now, as observed in Remark 6.4 Proposition 6.3 works also when $\gamma = 1$ and $g \in L^{1,\theta}(\Omega_T)$. Therefore, we apply (6.37) with $\sigma := 27R/40$ and $\varrho := R$ in order to bound $\|1 + |Du|\|_{L^{1,\theta-1}(\mathcal{C}_{27R/40})}$. This leads us to the estimate

$$\begin{aligned}\|1 + |Du|\|_{L^{1,\theta-1}(\mathcal{C}_{27R/40})} &\le c\left[R^{\theta-1-N}\|1+|Du|\|_{L^1(\mathcal{C}_R)} + \|g\|_{L^{1,\theta}(\mathcal{C}_R)}\right]\\ &\le c\left[R^{\theta-1-N}\|1+|Du|\|_{L^1(\mathcal{C}_R)} + \|g\|_{L\log L^{\theta}(\mathcal{C}_R)}\right],\end{aligned}$$

where we have used once again the trivial bound $\|\tilde g\|_{L^{1,\theta}(\mathcal{C}_1)} \lesssim \|\tilde g\|_{L\log L^{\theta}(\mathcal{C}_1)}$ in the second line. Using the preceding inequality in (6.45) we finally arrive at

$$\|Du\|_{L^{\frac{\theta}{\theta-1},\theta}(\mathcal{C}_{R/2})} \le c\left[R^{\theta-1-N}\|1+|Du|\|_{L^1(\mathcal{C}_R)} + \|g\|_{L\log L^{\theta}(\mathcal{C}_R)}\right], \tag{6.46}$$

with a constant $c = c(n, L, \nu)$. Note that in the preceding inequality we have $u \equiv u_k$, where u_k are the approximating solutions from (2.2) with right-hand side $g_k \in L^\infty(\Omega_T)$ satisfying $|g_k| \le |g|$. From the definition of the $L\log L^\theta$-norm we easily have that $\|g_k\|_{L\log L^\theta(\Omega_T)} \le \|g\|_{L\log L^\theta(\Omega_T)}$, so that (6.46) turns into

$$\|Du_k\|_{L^{\frac{\theta}{\theta-1},\theta}(\mathcal{C}_{R/2})} \le c\left[R^{\theta-1-N}\|1+|Du_k|\|_{L^1(\mathcal{C}_R)} + \|g\|_{L\log L^{\theta}(\mathcal{C}_R)}\right],$$

where again $c = c(n, L, \nu)$. This is the desired estimate for the approximating solutions we were looking for and the final result follows again by passing to the limit $k \to \infty$ in the right-hand side and the lower-semicontinuity on the left-hand side. □

6.3. **Further estimates in parabolic Lorentz spaces and a borderline case.**

Theorem 6.6. *Assume that (1.2) and $g \in L(\gamma, q)(\Omega_T)$ with $1 < \gamma \le \frac{2N}{N+2}$ and $0 < q \le \infty$ hold. Then the solution $u \in L^1(-T, 0; W_0^{1,1}(\Omega_T))$ to (1.1) satisfies*

$$|Du| \in L\Big(\frac{N\gamma}{N-\gamma}, q\Big) \quad \textit{locally in } \Omega_T.$$

Moreover, the local quantitative estimate

$$\|Du\|_{L(\frac{N\gamma}{N-\gamma},q)(\mathcal{C}_{R/2})} \le c\, R^{\frac{N-\gamma}{\gamma}-N}\|1+|Du|\|_{L^1(\mathcal{C}_R)} + \|g\|_{L(\gamma,q)(\mathcal{C}_R)} \tag{6.47}$$

holds for every parabolic cylinder $\mathcal{C}_R \subset \Omega_T$, with a constant $c = c(n, L, \nu, \gamma, q)$.

Proof. Once again we refer to the proof of Theorem 6.1. We start with (6.30) with the choices $\tau = q$ and $\theta = N$ and obtain

$$\|Du\|_{L(\frac{N\gamma}{N-\gamma},q)(\mathcal{Q}_0)} \le c\lambda_0|\mathcal{Q}_0|^{\frac{N-\gamma}{N\gamma}} + c\|M_1^*(g)\|_{L(\frac{N\gamma}{N-\gamma},q)(\mathcal{Q}_0)}, \tag{6.48}$$

with λ_0 from (6.19). The second term appearing on the right-hand side of (6.48) is treated via Theorem 4.2 (again switching to outer and inner cylinders)

$$\|M_1^*(g)\|_{L\left(\frac{N\gamma}{N-\beta\gamma},q\right)(\mathcal{Q}_0)} \le c\,\|g\|_{L(\gamma,q)(n^2\mathcal{Q}_0)}.$$

Combining this with (6.48) and recalling the definition (6.19) of λ_0 yields the following analogue of (6.33):

$$\|Du\|_{L(\frac{N\gamma}{N-\gamma},q)(\mathcal{Q}_0)} \le c\bigg(\fint_{n^2\mathcal{Q}_0}(1+|Du|)\,dz\bigg)|\mathcal{Q}_0|^{\frac{N-\gamma}{N\gamma}} + c\|g\|_{L(\gamma,q)(n^2\mathcal{Q}_0)}.$$

Modulo a standard covering argument and the additivity of quasi-norms from Remark 3.3 the preceding inequality is essentially equivalent to (6.47). The conclusion of the Theorem then follows by approximation. □

6.4. **Parabolic equations with more regular coefficients.** In this section we consider parabolic equations where the vector field a satisfies either the structure assumptions (1.12) to (1.14) – the VMO-case – or (1.11) – the case of a continuous vector-field. In these cases we can weaken the assumption (1.7). As we will see below, we can assume that

$$1 < \gamma < \theta \leq N \tag{6.49}$$

holds. The reason for this comes from the fact that the corresponding solutions to homogeneous Cauchy-Dirichlet problems satisfy reverse Hölder-type inequalities for arbitrarily large integrability exponents; see Theorem 5.7.

Theorem 6.7. *Let $u \in L^1(-T,0;L^1(\Omega))$ be the solution to (1.1) where either the structure conditions (1.12) to (1.14) or (1.11) are in force. Moreover, assume $g \in L^\theta(\gamma,q)(\Omega_T)$ with γ,θ as in (6.49) and $0 < q \leq \infty$. Then*

$$|Du| \in L^\theta\Big(\tfrac{\theta\gamma}{\theta-\gamma},\tfrac{\theta q}{\theta-\gamma}\Big) \quad \textit{locally in } \Omega_T.$$

Furthermore, we have the local estimate

$$\|Du\|_{L^\theta(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{C}_{R/2})} c\, R^{\frac{\theta-\gamma}{\gamma}-N}\|1+|Du|\|_{L^1(\mathcal{C}_R)} + c\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_R)},$$

for any parabolic cylinder $\mathcal{C}_R \subset \Omega_T$, where the constant c depends only on n, L, ν, θ, γ, q.

Since the ideas in the proof of Theorem 6.7 are very close to the ones of Theorem 6.1 we confine ourselves to sketch the necessary modifications. We deal with the approximating solutions $u \equiv u_k$ introduced in Section 2.1, where $g \equiv g_k \in L^\infty(\Omega_T)$. Keeping in mind the notation introduced in the proof of Theorem 6.1 we must replace Lemma 6.2 by

Lemma 6.8. *Let $u(=u_k) \in C^0([-T,0];L^2(\Omega)) \cap L^2(-T,0;W_0^{1,2}(\Omega))$ be a weak solution to (5.1) where the assumptions (1.2) are in force and $g \in L^\infty(\Omega_T)$. Then, for every choice of $\chi_0, S > 1$ there exist constants $\varepsilon \in (0,1)$ and $A > 1$ depending on $n, L, \nu, \omega(\cdot), S$ and χ_0 such that if $\lambda > 1$ and $\mathcal{Q}$ a dyadic sub-cylinder of $\mathcal{Q}_0$ such that if*

$$\big|\mathcal{Q} \cap \big\{z \in \mathcal{Q}_0 : M^*(1+|Du|)(z) > AS\lambda \ \textit{and}\ M_1^*(g) \leq \varepsilon\lambda\big\}\big| > \frac{|\mathcal{Q}|}{S^{2\chi_0}}, \tag{6.50}$$

then the predecessor $\tilde{\mathcal{Q}}$ of $\mathcal{Q}$ satisfies

$$\tilde{\mathcal{Q}} \subset \big\{z \in \mathcal{Q}_0 : M^*(1+|Du|)(z) > \lambda\big\}.$$

Proof. Due to the fact that we are dealing with parabolic equations with more regular vector-fields we can use the better higher integrability result from Theorem 5.7 instead of the ones from Theorem 5.5. Again we shall prove the assertion by a contradiction argument. We proceed as in the proof of Lemma 6.2 until (6.10). Having arrived at this stage we observe that the hypothesis of Theorem 5.7 are fulfilled for the solution v of the homogeneous Cauchy-Dirichlet problem (6.8) (see (5.12)) on $\mathcal{C}$. Therefore, we have the local higher integrability (5.14) of Dv on $2\mathcal{Q} \subset 3\mathcal{Q} \subset \mathcal{C}$. This means that for any given $\chi_0 > 1$ there exist a constant $c = c(n,L,\nu,\omega(\cdot),\chi_0)$ such that the estimate

$$\bigg(\fint_{2\mathcal{Q}} (1+|Dv|)^{2\chi_0}\,dz\bigg)^{\frac{1}{2\chi_0}} \leq c\fint_{3\mathcal{Q}} (1+|Dv|)\,dz,$$

holds. Exactly as in the proof of Lemma 6.2 this leads us to

$$\fint_{2\mathcal{Q}} (1+|Dv|)^{2\chi_0}\,dz \leq c(n,L,\nu,\omega(\cdot),\chi_0)\,\lambda^{2\chi_0}\,.$$

We proceed further using again the restricted maximal operator $M^{**} := M^*_{0,2\mathcal{Q}}$ on $2\mathcal{Q}$ and obtain the following analogue of (6.13):

$$\big|\{z \in \mathcal{Q} : M^{**}(1+|Du|)(z) > AS\lambda\}\big| \leq \Big[\frac{c_1(n,L,\nu,\omega(\cdot),\chi_0)}{(AS)^{2\chi_0}} + \frac{c_2(n,\nu)\varepsilon}{AS}\Big]|\mathcal{Q}|. \quad (6.51)$$

Now we perform the following choices of A and ε: first we choose $A = A(n,L,\nu,\omega(\cdot),\chi_0) > 1$ such that

$$A = 4\cdot 10^N\big[1+c_1\big] \quad \Longrightarrow \quad \frac{c_1}{(AS)^{2\chi_0}} \leq \frac{1}{4S^{2\chi_0}}\,, \quad (6.52)$$

and then choose $\varepsilon = \varepsilon(n,\nu,S,\chi_0) \in (0,1)$ accordingly to

$$\varepsilon = \frac{1}{4S^{2\chi_0-1}[1+c_2]} \quad \Longrightarrow \quad \frac{c_2\varepsilon}{AS} \leq \frac{1}{4S^{2\chi_0}}\,.$$

These choices in (6.51) yield the following analogue of (6.16):

$$\big|\{z \in \mathcal{Q} : M^{**}(1+|Du|)(z) > AS\lambda\}\big| < S^{-2\chi_0}|\mathcal{Q}|.$$

Having arrived at this stage we can argue exactly as in the proof of Proposition 6.2 after (6.16) to derive the analogue of (6.18), i.e.

$$\big|\{z \in \mathcal{Q} : M^{*}(1+|Du|)(z) > AS\lambda\}\big| < S^{-2\chi_0}|\mathcal{Q}|,$$

which contradicts (6.50). This proves the assertion of the proposition. □

To proceed with the proof of Theorem 6.7 we choose λ_0 accordingly to

$$\lambda_0 := 2c_0(n)n^{2N}S^{2\chi_0}\fint_{n^2\mathcal{Q}_0}(1+|Du|)\,dz.$$

With the arguments form the proof of Theorem 6.1, Step 3, replacing χ by χ_0 everywhere we arrive at the following proper version of (6.22) (the only change here is the replacement of χ by χ_0):

$$\begin{aligned}(AS)^{\frac{(k+1)\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1\big((AS)^{k+1}\lambda_0\big) &\leq A^{\frac{\theta\gamma}{\theta-\gamma}}S^{\frac{\theta\gamma}{\theta-\gamma}-2\chi_0}(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1\big((AS)^k\lambda_0\big)\\ &\quad + \big(AS/\varepsilon\big)^{\frac{\theta\gamma}{\theta-\gamma}}(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}\big(\varepsilon\lambda_0\big)^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2\big((AS)^k\varepsilon\lambda_0\big)\,.\end{aligned}$$

Since $\gamma < \theta$ by assumption, the quantity $\frac{\theta\gamma}{\theta-\gamma}$ can be arbitrarily large. Nevertheless, since $\chi_0 > 0$ is at our disposal, we can choose χ_0 large enough to have

$$d := 2\chi_0 - \frac{\theta\gamma}{\theta-\gamma} > 0, \quad (6.53)$$

a relation playing the same role as (6.23) before. Note, that here we really need the possibility of taking χ_0 large. This fixes $\chi_0 = \chi_0(\theta,\gamma)$ (for example we could choose $\chi_0 = \frac{\theta\gamma}{\theta-\gamma}$), $d = d(\theta,\gamma)$ and also $A = A(n,L,\nu,\omega(\cdot),\theta,\gamma)$ by (6.52). Having fixed χ_0 we choose

$$S := \Big[4A^{\frac{\theta\gamma}{\theta-\gamma}}\Big]^{\frac{1}{d}}\,, \quad (6.54)$$

where A has been determined in (6.52). Then S admits the same dependencies as A, i.e. $S = S(n,L,\nu,\omega(\cdot),\theta,\gamma)$, and therefore we can write $AS/\varepsilon =: c_*(n,L,\nu,\omega(\cdot),\theta,\gamma)$. In view of (6.53) and (6.54) we find that the analogue of (6.25), i.e.

$$\begin{aligned}(AS)^{\frac{(k+1)\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1\left((AS)^{k+1}\lambda_0\right) &\le \frac{1}{4}(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-\gamma}}\mu_1\left((AS)^k\lambda_0\right)\\ &\quad+\left(\tfrac{c_*}{2}\right)^{\frac{\theta\gamma}{\theta-\gamma}}(AS)^{\frac{k\theta\gamma}{\theta-\gamma}}\left(\varepsilon\lambda_0\right)^{\frac{\theta\gamma}{\theta-\gamma}}\mu_2\left((AS)^k\varepsilon\lambda_0\right), \qquad (6.55)\end{aligned}$$

holds for every $k\in\mathbb{N}_0$ with a constant c_*. The preceding estimate for the level sets allows us to proceed as in the proof of Theorem 6.1 after (6.25); i.e. we first sum up (6.55) upon $k\in\mathbb{N}$ and then re-absorb the intermediate sum in the left-hand side. Arguing exactly as in (6.25)–(6.33) we arrive at the following analogue of (6.33):

$$\|Du\|_{L(\frac{\theta\gamma}{\theta-\gamma},\frac{\theta q}{\theta-\gamma})(\mathcal{Q}_0)} \le c\bigg(\fint_{n^2\mathcal{Q}_0}(1+|Du|)\,dz\bigg)|\mathcal{Q}_0|^{\frac{\theta-\gamma}{\theta\gamma}} + c\|g\|_{L^\theta(\gamma,q)(n^2\mathcal{Q}_0)}, \qquad (6.56)$$

where now $c=c(n,L,\nu,\theta,\gamma,q)$. In the next step we have to replace Proposition 6.3 by an appropriate version valid under the weaker assumption $1<\gamma<\theta\le N$. This is achieved in the following

Proposition 6.9. *Let $u(=u_k)\in C^0([-T,0];L^2(\Omega))\cap L^2(-T,0;W_0^{1,2}(\Omega))$ be a weak solution to (5.1) where either the structure conditions (1.12) to (1.14) or the condition (1.11) are in force and $g\in L^\theta(\gamma,q)(\Omega_T)$ with $1<\gamma\le\theta\le N$. Then, for every pair of concentric parabolic cylinders $\mathcal{C}_\sigma\subset\mathcal{C}_\varrho\subset\Omega_T$ there holds*

$$\|1+|Du|\|_{L^{1,\frac{\theta-\gamma}{\gamma}}(\mathcal{C}_\sigma)} \le c\,(\varrho-\sigma)^{\frac{\theta-\gamma}{\gamma}-N}\|1+|Du|\|_{L^1(\mathcal{C}_\varrho)} + c\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_\varrho)},$$

where $c=c(n,L,\nu,\gamma,\theta,q)$.

Proof. We first note that by Proposition 6.3 we only have to treat the case $\gamma\in(\theta/2,\theta)$. Let z_0 be a point in $\mathcal{C}_\sigma$ and $\mathcal{C}(z_0,R)$ a parabolic cylinder with $0<R\le d_{\mathrm{par}}(z_0,\partial\mathcal{C}_\varrho)$, i.e. $\mathcal{C}(z_0,R)\subset\mathcal{C}_\varrho$. Moreover, let $v\in C^0([t_0-R^2,t_0+R^2];L^2(B_R(x_0))\cap L^2(t_0-R^2,t_0+R^2;W^{1,2}(B_R(x_0))$ be the unique solution to the Cauchy-Dirichlet problem (6.8) in $C(z_0,R)$. Then, using (5.13) for the choice $q=1$ and with $\alpha\in(0,1)$ to be fixed later we infer by the argument from the beginning of the proof of Proposition 6.3 that

$$\begin{aligned}\int_{\mathcal{C}(z_0,r)}(1+|Du|)\,dz \le c\Big(\frac{r}{R}\Big)^{N-1+\alpha}\int_{\mathcal{C}(z_0,R)}(1+|Du|)\,dz + c\int_{\mathcal{C}(z_0,R)}|Du-Dv|\,dz,\end{aligned}$$

holds for any $0<r\le R$ where $c=c(n,L,\nu,\alpha)$. Using (5.4) from Lemma 5.3 the previous inequality leads us to

$$\begin{aligned}\int_{\mathcal{C}(z_0,r)}(1+|Du|)\,dz \le c\Big(\frac{r}{R}\Big)^{N-1+\alpha}\int_{\mathcal{C}(z_0,R)}(1+|Du|)\,dz + c\,\|g\|_{L^\theta(\gamma,q)(\mathcal{C}_\varrho)}R^{N-1+(1-\frac{\theta-\gamma}{\gamma})},\end{aligned}$$

where $c=c(n,L,\nu,q,\alpha)$. At this stage we remark that $\gamma\in(\theta/2,\theta)$ yields $1-\frac{\theta-\gamma}{\gamma}\in(0,1)$. Therefore we choose $\alpha\in(0,1)$ such that $1>\alpha>1-\frac{\theta-\gamma}{\gamma}>0$. Note, that here we really need the possibility of taking α close to 1 at our disposal. For example we could choose $\alpha:=1-\frac{\theta-\gamma}{2\gamma}$, fixing $\alpha=\alpha(\theta,\gamma)$. Now, we can finish the proof exactly as in the proof of Proposition 6.3 by the application of Lemma 2.4. □

Having arrived at this stage the local Lorentz integrability of Du from (6.56) can be turned into the desired Lorentz-Morrey space estimate via the scaling argument along the lines of the proof of Theorem 6.1, Step 5, combined with the intermediate Morrey space information for Du from Proposition 6.9.

7. INTEGRABILITY OF u

Theorem 7.1. *Under the assumptions (1.2), $g \in L^\theta(\gamma, q)(\Omega_T)$ with $0 < q \le \infty$ and*

$$1 < \gamma < \frac{\theta}{2} \qquad \text{and} \qquad 2 < \theta \le N, \tag{7.1}$$

the solution $u \in L^1(-T, 0, W_0^{1,1}(\Omega))$ to (1.1) is such that

$$u \in L^\theta\Big(\frac{\theta\gamma}{\theta - 2\gamma}, \frac{\theta q}{\theta - 2\gamma}\Big) \quad \text{locally in } \Omega_T.$$

Moreover, the following quantitative estimate

$$\|u\|_{L^\theta\left(\frac{\theta\gamma}{\theta-2\gamma}, \frac{\theta q}{\theta-2\gamma}\right)(\mathcal{C}_{R/2})} \le c\, R^{\frac{\theta-2\gamma}{\gamma} - N} \||u| + R\|_{L^1(\mathcal{C}_R)} + c\, \|g\|_{L^\theta(\gamma,q)(\mathcal{C}_R)} \tag{7.2}$$

holds for any $\mathcal{C}_R \subset \Omega_T$, where $c = c(n, L, \nu, \gamma, \theta, q)$.

Theorem 7.2. *Under the assumptions (1.2) and being $g \in L(\gamma, q)(\Omega_T)$ with $1 < \gamma < \frac{N}{2}$, $0 < q \le \infty$, then the solution $u \in L^1(-T, 0, W_0^{1,1}(\Omega))$ to (1.1) is such that*

$$u \in L\Big(\frac{N\gamma}{N - 2\gamma}, q\Big) \quad \text{locally in } \Omega_T.$$

Moreover, the following quantitative estimate

$$\|u\|_{L\left(\frac{N\gamma}{N-2\gamma}, q\right)(\mathcal{C}_{R/2})} \le c\, R^{\frac{N-2\gamma}{\gamma} - N} \||u| + R\|_{L^1(\mathcal{C}_R)} + c\, \|g\|_{L(\gamma,q)(\mathcal{C}_R)}$$

holds for any $\mathcal{C}_R \subset \Omega_T$, where $c = c(n, L, \nu, \gamma, q)$.

Since the proofs of Theorems 7.1, 7.2 are very close to the one of Theorem 6.1 we confine ourselves to outline the necessary modifications only. Again, we deal with the approximating solutions $u \equiv u_k$ defined in Section 2.1, abbreviating again $g \equiv g_k$. Now, we go back to the proof of Theorem 6.1 and keep in mind the notation introduced thereby. Then, Lemma 6.2 must be replaced by

Lemma 7.3. *Let $u (= u_k) \in C^0([-T, 0]; L^2(\Omega)) \cap L^2(-T, 0; W_0^{1,2}(\Omega))$ be a weak solution to (5.1) where the assumptions (1.2) are in force and $g \in L^\infty(\Omega_T)$. Then, there exists an absolute constant $A = A(n, L, \nu) > 1$ such that: For every $S > 1$ and $\chi_0 > 1$ there exists a constant $\varepsilon = \varepsilon(n, L, \nu, S, \chi_0) \in (0, 1)$ such that if $\lambda > 1$ and $\mathcal{Q}$ a dyadic sub-cylinder of $\mathcal{Q}_0$ such that*

$$\big|\mathcal{Q} \cap \big\{z \in \mathcal{Q}_0 : M^*(1 + |u|)(z) > AS\lambda \;\; \text{and} \;\; M_2^*(g) \le \varepsilon\lambda\big\}\big| > \frac{|\mathcal{Q}|}{S^{2\chi_0}}, \tag{7.3}$$

then the predecessor $\tilde{\mathcal{Q}}$ of $\mathcal{Q}$ satisfies

$$\tilde{\mathcal{Q}} \subset \big\{z \in \mathcal{Q}_0 : M^*(1 + |u|)(z) > \lambda\big\}.$$

The main changes in the statement of Lemma 7.3 are essentially the replacement of $M_1^*(g) = M_{1,n^2\mathcal{Q}_0}^*(g)$ by $M_2^*(g) = M_{2,n^2\mathcal{Q}_0}^*(g)$ and the introduction of the parameter $\chi_0 > 1$ which is at our disposal, i.e. χ_0 can be picked large at will, while the quantity χ in Lemma 6.2 was fixed. In principle, the proof of Lemma 7.3 follows the one of Lemma 6.2 replacing $M^*(1 + |Du|)$, $M_1^*(g)$ by $M^*(1 + |u|)$, $M_2^*(g)$. But for convenience of the reader we describe the main differences. Since we are dealing with $M_2^*(g)$, (6.7) has to be replaced by $M_2^*(g)(\bar z) \le \varepsilon\lambda$ for some $\bar z \in \mathcal{Q}$, yielding in turn the following analogue of (6.9):

$$|\mathcal{C}|^{\frac{2}{N}} \fint_{\mathcal{C}} |g|\, dz \le c(n)\varepsilon\lambda.$$

Therefore, by (5.3) we have

$$\int_{3\mathcal{Q}} |u-v|\, dz \le c(n,\nu)|\mathcal{C}|^{\frac{2}{N}} \int_{\mathcal{C}} |g|\, dz \le c(n,\nu)\varepsilon\lambda.$$

At this stage the proof proceeds exactly along the lines of Lemma 6.2, starting with the preceding inequality instead of (6.9). This leads us to the following analogue of (6.12):

$$\int_{2\mathcal{Q}} (1+|v|)^{2\chi_0}\, dz \le c(n,L,\nu)\, \lambda^{2\chi_0}.$$

Here we have taken into account that the side-length of the cylinders is bounded by 1. Now, as in the proof of Lemma 6.2 we compare the level set of $M^{**}(1+|u|)$ in the cylinder $\mathcal{Q}$ with the one of $M^{**}(1+|v|)$, where $M^{**} := M^*_{0,2\mathcal{Q}}$. This procedure implies that

$$\big|\{z \in \mathcal{Q} : M^{**}(1+|u|)(z) > AS\lambda\}\big| \le \Big[\frac{c_1(n,L,\nu)}{(AS)^{2\chi_0}} + \frac{c_2(n,\nu)\varepsilon}{AS}\Big]|\mathcal{Q}|.$$

The choices of A and ε are performed exactly as in (6.14), (6.15), but everywhere replacing χ by χ_0. This fixes $A = A(n,L,\nu) > 1$ and $\varepsilon = \varepsilon(n,L,\nu,S,\chi_0) \in (0,1)$ and leads first to the analogue of (6.16), and secondly with the arguments from the proof of Lemma 6.2 to the following analogue of (6.18):

$$\big|\{z \in \mathcal{Q} : M^*(1+|u|)(z) > AS\lambda\}\big| < S^{-2\chi_0}|\mathcal{Q}|\,,$$

contradicting (7.3). This completes the proof of Lemma 7.3.

We now proceed with the proof of Theorem 7.1 along the lines of Theorem 6.1, starting at Step 2. We initially choose

$$\lambda_0 := 2c_0(n)n^{2N}S^{2\chi_0} \fint_{n^2\mathcal{Q}_0} (1+|u|)\, dz\,, \tag{7.4}$$

and define $\mu_1(\cdot)$, $\mu_2(\cdot)$ by

$$\mu_1(H) := \big|\big\{z \in \mathcal{Q}_0 : M^*(1+|u|)(z) > H\big\}\big|,$$

and

$$\mu_2(H) := \big|\big\{z \in \mathcal{Q}_0 : M_2^*(g)(z) > H\big\}\big|,$$

respectively, for $H \ge 0$. At this stage we start replacing $\frac{\theta}{\theta-\gamma}$ by $\frac{\theta}{\theta-2\gamma}$ everywhere. Applying Lemma 7.3 and Proposition 2.1 at levels $H = (AS)^{k+1}\lambda_0$, $(AS)^k\lambda_0$ for $k = 0,1,2,\dots$ we arrive at the following analogue of (6.22):

$$\begin{aligned}(AS)^{\frac{(k+1)\theta\gamma}{\theta-2\gamma}} &\lambda_0^{\frac{\theta\gamma}{\theta-2\gamma}} \mu_1\left((AS)^{k+1}\lambda_0\right)\\ &\le A^{\frac{\theta\gamma}{\theta-2\gamma}} S^{\frac{\theta\gamma}{\theta-2\gamma}-2\chi_0}(AS)^{\frac{k\theta\gamma}{\theta-2\gamma}}\lambda_0^{\frac{\theta\gamma}{\theta-2\gamma}}\mu_1\left((AS)^k\lambda_0\right)\\ &\quad + \left(AS/\varepsilon\right)^{\frac{\theta\gamma}{\theta-2\gamma}}(AS)^{\frac{k\theta\gamma}{\theta-2\gamma}}\left(\varepsilon\lambda_0\right)^{\frac{\theta\gamma}{\theta-2\gamma}}\mu_2\left((AS)^k\varepsilon\lambda_0\right).\end{aligned} \tag{7.5}$$

We observe that this time we can choose χ_0 large enough to have $2\chi_0 - \frac{\theta\gamma}{\theta-2\gamma} > 0$; see (6.23) for the corresponding relation in the proof of Theorem 6.1. Note also that at this stage we need the possibility for choosing χ_0 large at will, since 2γ can be arbitrarily close θ making $\frac{\theta\gamma}{\theta-2\gamma}$ large. This motivates the following definitions

$$d := 2\chi_0 - \frac{\theta\gamma}{\theta-2\gamma} > 0, \qquad S := \left[4A^{\frac{\theta\gamma}{\theta-2\gamma}}\right]^{\frac{1}{d}}\,,$$

so that $A^{\frac{\theta\gamma}{\theta-2\gamma}}S^{\frac{\theta\gamma}{\theta-2\gamma}-2\chi_0} \le \frac{1}{4}$. Using this in (7.5) allows us to proceed as in the proof of Theorem 6.1 after (6.25); i.e. we first sum up (7.5) upon $k \in \mathbb{N}$ and then re-absorb the intermediate sum in the left-hand side. Arguing exactly as in (6.25)–(6.30) we arrive at the following analogue of (6.30):

$$\|u\|_{L(\frac{\theta\gamma}{\theta-2\gamma},\tau)(\mathcal{Q}_0)} \le c\lambda_0|\mathcal{Q}_0|^{\frac{\theta-2\gamma}{\theta\gamma}} + c\|M_2^*(g)\|_{L(\frac{\theta\gamma}{\theta-2\gamma},\tau)(\mathcal{Q}_0)}, \tag{7.6}$$

which holds for every $\tau > 0$, and where $c = c(n, L, \nu, \gamma, \theta, \tau)$. At this stage we take $\tau = \frac{\theta q}{\theta-2\gamma}$ in (7.6) and apply Theorem 4.13 with $\beta = 2$ and $p = \gamma$ (note that $\beta p = 2\gamma < \theta$ by assumption (7.1)); this leads us to

$$\|M_2^*(g)\|_{L(\frac{\theta\gamma}{\theta-2\gamma},\frac{\theta q}{\theta-2\gamma})(\mathcal{Q}_0)} \le c\,\|g\|_{L^\theta(\gamma,q)(n^2\mathcal{Q}_0)}.$$

It is worth to remark that this is exactly the point where we use the fact that M_2^* admits a higher regularizing effect then M_1^*. Combining the preceding inequality with (7.6) and recalling the definition of λ_0 from (7.4) we obtain

$$\|u\|_{L(\frac{\theta\gamma}{\theta-2\gamma},\frac{\theta q}{\theta-2\gamma})(\mathcal{Q}_0)} \le c\bigg(\fint_{n^2\mathcal{Q}_0}(1+|u|)\,dz\bigg)|\mathcal{Q}_0|^{\frac{\theta-2\gamma}{\theta\gamma}} + c\,\|g\|_{L^\theta(\gamma,q)(n^2\mathcal{Q}_0)}.$$

Having arrived at this stage the local Lorentz integrability of u can be turned into the desired Lorentz-Morrey space estimate via a scaling argument along the lines of the proof of Theorem 6.1, Step 5, combined with the intermediate Morrey space information for u from Remark 6.4, (6.38). Consider $\mathcal{C}_\varrho \subset \Omega_T$. Scaling back to $\mathcal{C}_1$ as in Remark 5.1 and arguing along the lines of Step 5 we find

$$\|\tilde u\|_{L(\frac{\theta\gamma}{\theta-2\gamma},\frac{\theta q}{\theta-2\gamma})(\mathcal{C}_{1/n^4})} \le c\,\|1+|\tilde u|\|_{L^{1,\frac{\theta-2\gamma}{\gamma}}(\mathcal{C}_{9/10})} + c\,\|\tilde g\|_{L^\theta(\gamma,q)(\mathcal{C}_1)}\,.$$

Scaling back to $\mathcal{C}_\varrho$ via Lemma 3.2, we find for every parabolic cylinder $\mathcal{C}_\varrho \subset \Omega_T$ that

$$\|u\|_{L(\frac{\theta\gamma}{\theta-2\gamma},\frac{\theta q}{\theta-2\gamma})(\mathcal{C}_{\varrho/n^4})} \le c\,\Psi(\mathcal{C}_\varrho)\,\varrho^{(N-\theta)\frac{\theta-2\gamma}{\theta\gamma}},$$

where this time we have set

$$\Psi(\mathcal{C}_\varrho) := \|\varrho+|u|\|_{L^{1,\frac{\theta-2\gamma}{\gamma}}(\mathcal{C}_{9\varrho/10})} + \|g\|_{L^\theta(\gamma,q)(\mathcal{C}_\varrho)}.$$

Having arrived at this stage we follow exactly the proof of Theorem 6.1, Step 5, after (6.39). The only difference occurs when using the intermediate Morrey-space estimate (6.38) instead of (6.35). The desired estimate (7.2) then follows by the approximation argument from the proof of Theorem 6.1, Step 6.

The proof of Theorem 7.2 follows similarly to the one of Theorem 6.6, taking into account Theorem 4.2 for the choice $\beta = 2$.

Theorem 7.4. *Under the assumption (1.2) and $g \in \mathcal{M}^{\theta/2,\theta}(\Omega_T)$ with $2 < \theta \le N$ the solution $u \in L^1(-T,0;W_0^{1,1}(\Omega))$ to (1.1) belongs to $\mathrm{BMO}_{\mathrm{loc}}(\Omega_T)$. Moreover, there exists a constant $c = c(n, L, \nu, \theta)$ such that for any parabolic cylinder $\mathcal{C}_R \subset \Omega_T$ holds*

$$[u]_{\mathrm{BMO}(\mathcal{C}_{R/4})} \le c\,R^{1-N}\|1+|Du|\|_{L^1(\mathcal{C}_R)} + c\,\|g\|_{\mathcal{M}^{\theta/2,\theta}(\Omega_T)(\mathcal{C}_R)}$$

Proof. Once again we consider the approximating solutions $u \equiv u_k \in C^0([-T,0];L^2(\Omega)) \cap L^2(-T,0;W_0^{1,2}(\Omega))$ to (2.2). From [11, Lemma 4.3] we recall that the following Poincaré-type inequality holds

$$\int_{\mathcal{C}_{\varrho/2}} |u-(u)_{\mathcal{C}_{\varrho/2}}|\,dz \le c\,\varrho\int_{\mathcal{C}_\varrho}|Du|\,dz + c\,\varrho^2\int_{\mathcal{C}_\varrho}|g|\,dz,$$

for any parabolic cylinder $\mathcal{C}_\varrho \subset \Omega_T$, with a constant $c = c(n, L, \nu)$. Therefore, we have

$$\begin{aligned}[u]_{\mathrm{BMO}(\mathcal{C}_{R/4})} &= \sup_{\mathcal{C}_\varrho\subset\mathcal{C}_{R/4}}\fint_{\mathcal{C}_\varrho}|u-(u)_{\mathcal{C}_\varrho}|\,dz\\ &\le c\left[\|Du\|_{L^{1,1}(\mathcal{C}_{R/2})} + \sup_{\mathcal{C}_\varrho\subset\mathcal{C}_{R/2}}\varrho^2\fint_{\mathcal{C}_\varrho}|g|\,dz\right].\end{aligned}$$

The first term appearing on the right-hand side of the preceding inequality can be estimated with Proposition 6.3 for the choice $\gamma = \theta/2$ and $q = \infty$ (note that $\frac{\theta-\theta/2}{\theta/2} = 1$); we infer that

$$\|Du\|_{L^{1,1}(\mathcal{C}_{R/2})} \leq c\, R^{1-N} \|1 + |Du|\|_{L^1(\mathcal{C}_R)} + c\, \|g\|_{\mathcal{M}^{\theta/2,\theta}(\mathcal{C}_R)},$$

where $c = c(n, L, \nu, \theta)$. On the other hand, the second term can be treated by use of Lemma 4.1 as follows:

$$\begin{aligned} \varrho^2 \fint_{\mathcal{C}_\varrho} |g|\, dz &\leq \frac{\theta/2}{\theta/2 - 1} |\mathcal{C}_\varrho|^{-2/\theta} \varrho^2 \|g\|_{\mathcal{M}^{\theta/2}(\mathcal{C}_\varrho)} \\ &\leq \frac{\theta/2}{\theta/2 - 1} [2\alpha(n)]^{-2/\theta} \varrho^{\frac{\theta-N}{\theta/2}} \|g\|_{\mathcal{M}^{\theta/2}(\mathcal{C}_\varrho)} \\ &= c(n,\theta) \varrho^{\frac{\theta-N}{\theta/2}} \|g\|_{\mathcal{M}^{\theta/2}(\mathcal{C}_\varrho)}. \end{aligned}$$

Therefore, we have

$$\sup_{\mathcal{C}_\varrho \subset \mathcal{C}_{R/2}} \varrho^2 \fint_{\mathcal{C}_\varrho} |g|\, dz \leq c(n,\theta) \|g\|_{\mathcal{M}^{\theta/2,\theta}(\mathcal{C}_{R/2})}.$$

Combining the preceding estimates leads us to

$$[u]_{\mathrm{BMO}(\mathcal{C}_{R/4})} \leq c\, R^{1-N} \|1 + |Du|\|_{L^1(\mathcal{C}_R)} + c\, \|g\|_{\mathcal{M}^{\theta/2,\theta}(\mathcal{C}_R)},$$

where $c = c(n, L, \nu, \theta)$ we note that the constant c blows up, i.e. $c \to \infty$, when $\theta \downarrow 2$. Again the desired result follows by approximation. □

Acknowledgements. This research is supported by the ERC grant 207573 "Vectorial Problems". The authors thank A. Di Castro and G. Palatucci for remarks on a preliminary version of the paper.

REFERENCES

[1] E. Acerbi, G. Mingione, G.A. Seregin, Regularity results for parabolic systems related to a class of non-newtonian fluids. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 21, No.1:25–60, 2004.

[2] E. Acerbi, G. Mingione, Gradient estimates for a class of parabolic systems. *Duke Math. J.*, 136:285–320, 2007.

[3] R.A. Adams, *Sobolev Spaces*. Academic Press, New York, 1978.

[4] D.R. Adams, A note on Riesz potentials. *Duke Math. J.*, 42:765-778, 1975.

[5] D.R. Adams, J. L. Lewis, On Morrey-Besov inequalities. *Studia Math.*, 74:169–182, 1982.

[6] A.A. Arkhipova, On a partial regularity up to the boundary of weak solutions to quasilinear parabolic systems with quadratic growth. *Zap. Nauchn. Semin. POMI*, 249(5):20–39, 1997.

[7] P. Baroni, J. Habermann, Calderón-Zygmund estimates for parabolic measure data equations. *J. Differential Equations*, 252 (1):412–447, 2012.

[8] L. Beck, Partial regularity for weak solutions of nonlinear elliptic systems: the subquadratic case. *Manuscr. Math.*, 123(4):453–491, 2007.

[9] L. Boccardo, A. Dall'Aglio, T. Gallouët, L. Orsina, Nonlinear parabolic equations with measure data. *J. Funct. Anal.*, 147:237–258, 1997.

[10] L. Boccardo, T. Gallouët, Non-linear elliptic and parabolic equations involving measure data. *J. Funct. Anal.*, 87:149–169, 1989.

[11] V. Bögelein, V, Partial regularity and singular sets of solutions of higher order parabolic systems. *Ann. Mat. Pura Appl.*, 188(1):61–122, 2009.

[12] V. Bögelein, F. Duzaar, G. Mingione, The boundary regularity of non-linear parabolic systems II *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 27:145–200, 2010.

[13] B. Bojarski, T. Iwaniec, Analytical foundations of the theory of quasiconformal mappings in $\mathbb{R}^n$. *Ann. Acad. Sci. Fenn.*, Ser. A I 8:257–324, 1983.

[14] L. Caffarelli, I. Peral, On $W^{1,p}$ estimates for elliptic equations in divergence form. *Comm. Pure Appl. Math.*, 51:1–21, 1989.

[15] S. Campanato, On the nonlinear parabolic systems in divergence form. Hölder continuity and partial Hölder continuity of the solutions. *Ann. Mat. Pura Appl.*, 137 (4):83–122, 1984.

[16] Y.Z. Chen, E. DiBenedetto, Boundary estimates for solutions of nonlinear degenerate parabolic systems, *J. reine ang. Math. (Crelles J.)*, 395:102–131, 1989.

[17] E. De Giorgi, Frontiere orientate di misura minima. *Sem. Scuola Normale Superiore Pisa*, 1960–1961.

[18] E. De Giorgi, Un esempio di estremali discontinue per un problema variazionale di tipo ellittico. *Boll. UMI*, 4:135–137, 1968.

[19] E. DiBenedetto, *Degenerate parabolic equations*. Universitext. Springer-Verlag, XV 387, New York, NY, 1993.

[20] A. Domokos, Differentiability of solutions for the non-degenerate p-laplacian in the Heisenberg group. *J. Differential Equations*, 204:439–470, 2004.

[21] J. Duan, *Fourier analysis*. Graduate Studies in Mathematics, 29. AMS, Providence, 2001.

[22] F. Duzaar, J.F. Grotowski, Optimal interior partial regularity for nonlinear elliptic systems: The method of a-harmonic approximation. *Manuscr. Math.*, 103:267–298, 2000.

[23] F. Duzaar, J.F. Grotowski, M. Kronz, Partial and full boundary regularity for minimizers of functionals with nonquadratic growth. *J. Convex Anal.*, 11:437–476, 2004.

[24] F. Duzaar, J. Kristensen, G. Mingione, The existence of regular boundary points for non-linear elliptic systems. *J. reine ang. Math. (Crelles J.)*, 602:17–58, 2007.

[25] F. Duzaar, G. Mingione, Second order parabolic systems, optimal regularity, and singular sets of solutions. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 22:705–751, 2005.

[26] F. Duzaar, G. Mingione, Gradient estimates via non-linear potentials. *Amer. J. Math.*, 133(4):1093–1149, 2011.

[27] F. Duzaar., G. Mingione, Gradient continuity estimates. *Calc. Var. Partial Differential Equations*, 39:379–418, 2010.

[28] F. Duzaar, G. Mingione, K. Steffen, Parabolic systems with polynomial growth and regularity. *Mem. Amer. Math. Soc.*, 214, n.1005, 2011.

[29] F. Duzaar, G. Mingione, Gradient estimates via linear and nonlinear potentials. *J. Funct. Anal.*, 259:2961–2998, 2010.

[30] F. Duzaar, K. Steffen, Optimal interior and boundary regularity for almost minimizers to elliptic variational integrals. *J. reine ang. Math. (Crelles J.)*, 546:73–138, 2002.

[31] C. Fefferman, E.M. Stein, H^p spaces of several variables. *Acta Math.*, 129:137–193, 1972.

[32] L. Grafakos, *Classical and Modern Fourier Analysis*. Pearson Edu. Inc., Upper Saddle River, 2004.

[33] M. Giaquinta, A counter-example to the boundary regularity of solutions to quasilinear systems. *Manuscr. Math.*, 24:217–220, 1978.

[34] M. Giaquinta, *Multiple Integrals in the Calculus of Variations and Nonlinear Elliptic Systems*. Princeton University Press, Princeton, New Jersey, 1983.

[35] M. Giaquinta, *Introduction to regularity theory for nonlinear elliptic systems*. Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1993.

[36] E. Giusti, *Direct Methods in the Calculus of Variations*. World Scientific Publishing Company, Tuck Link, Singapore, 2003.

[37] J. Grotowski, *Boundary regularity results for non-linear elliptic systems in divergence form*. Habilitationsschrift, 2000.

[38] J. Grotowski, Boundary regularity for nonlinear elliptic systems. *Calc. Var. Part. Diff. Equ.*, 15:353–388, 2002.

[39] P. Hajłasz, P. Koskela, Sobolev Met Poincaré. *Mem. Amer. Math. Soc.*, 145, n.688, 2000.

[40] T. Kilpeläinen, J. Malý, The Wiener test and potential estimates for quasilinear elliptic equations. *Acta Math.*, 172:137–161, 1994.

[41] J. Kristensen, G. Mingione, The singular set of ω-minima. *Arch. Ration. Mech. Anal.*, 177(1):93–114, 2005.

[42] M. Kronz, Quasimonotone systems of higher order. *Boll. UMI*, 6-B (8):459–480, 2003.

[43] T. Kuusi, G. Mingione, Nonlinear potential estimates in parabolic problems. *Rendiconti Lincei, Matematica e Applicazioni*, 22(2):161–174.

[44] O. Ladyzhenskaya, V. Solonnikov, N. Uraltseva, *Linear and quasi-linear equations of parabolic type*. Translations of Mathematical Monographs, 23, AMS, 1968.

[45] G.M. Lieberman, *Second order parabolic differential equations*. World Scientific press, River Edge, 1996.

[46] G. Mingione, The Calderón-Zygmund theory for elliptic problems with measure data. *Ann Scu. Norm. Sup. Pisa Cl. Sci.*, 5(6):195–261, 2007.

[47] G. Mingione, Gradient potential estimates. *J. Eur. Math. Soc.*, 13:459–486, 2011.

[48] G. Mingione, Gradient estimates below the duality exponent, *Math. Ann.*, 346:571–627, 2010.

[49] G. Mingione, Nonlinear aspects of Calderón-Zygmund theory. *Jahresbericht der DMV*, 112:159–191, 2010.

[50] J. Naumann, J. Wolf, Interior integral estimates on weak solutions of nonlinear parabolic systems. *Preprint 94-11 Inst. für Math.*, Humboldt Universität, Bonn, 1994.

[51] F. Petitta, Renormalized solutions of nonlinear parabolic equations with general measure data. *Ann. Mat. Pura Appl., IV.*, 187(4):563–604, 2008.

[52] J. Simon, Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl., IV.*, 146:65–96, 1987.

[53] L. Simon, *Theorems on Regularity and Singularity of Energy Minimizing Maps*. Lectures in Math., ETH Zürich, Birkhäuser, Basel, 1996.

[54] J. Stará, O. John, and J. Malý, Counterexamples to the regularity of weak solutions of the quasilinear parabolic system. *Commentat. Math. Univ. Carol.*, 27:123–136, 1986.

[55] E.W. Stredulinsky, Higher integrability from reverse Hölder inequalities. *Indiana Univ. Math. J.*, 29:407–413, 1980.

[56] G. Talenti, Elliptic equations and rearrangements. *Ann. Scu. Norm. Sup. Pisa Cl. Sci.*, 3(4):697–717, 1976.

PAOLO BARONI, SCUOLA NORMALE SUPERIORE, PIAZZA DEI CAVALIERI 7, I-56126 PISA, ITALY
E-mail address: paolo.baroni@sns.it

JENS HABERMANN, DEPARTMENT MATHEMATIK, UNIVERSITÄT ERLANGEN, CAUERSTR. 11, 90158 ERLANGEN, GERMANY
E-mail address: habermann@math.fau.de