



Parabolic Lipschitz truncation and caloric approximation

L. Diening¹ · S. Schwarzacher² · B. Stroffolini³ ·
A. Verde³

Received: 7 February 2017 / Accepted: 15 June 2017 / Published online: 17 July 2017
© Springer-Verlag GmbH Germany 2017

Abstract We develop an improved version of the parabolic Lipschitz truncation, which allows qualitative control of the distributional time derivative and the preservation of zero boundary values. As a consequence, we establish a new caloric approximation lemma. We show that almost p -caloric functions are close to p -caloric functions. The distance is measured in terms of spatial gradients as well as almost uniformly in time. Both results are extended to the setting of Orlicz growth.

Mathematics Subject Classification 35A35 · 35K55

1 Introduction

The purpose of the Lipschitz truncation is to regularize a given function by a Lipschitz continuous one by changing it only on a small *bad* set. It is crucial for the applications that the function is not changed globally, which rules out the possibility of convolutions. The

Communicated by J. Ball.

✉ B. Stroffolini
bstroffo@unina.it

L. Diening
lars.diening@uni-bielefeld.de

S. Schwarzacher
schwarz@karlin.mff.cuni.cz

A. Verde
anverde@unina.it

¹ Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany

² Mathematical Institute, Charles University in Prague, Sokolovska 83, 186 75 Praha 8-Karlin, Czechia

³ Dipartimento di Matematica, Università di Napoli, Federico II, Via Cintia, 80126 Naples, Italy

Lipschitz truncation technique was introduced by Acerbi–Fusco [2] to show lower semi-continuity of certain variational integrals.

Since then this technique has been successfully applied in many different areas. Let us provide a few examples. The Lipschitz truncation was used in the context of biting lemmas, existence theory and regularity results of non-linear elliptic PDE for example in [1, 11, 19, 20, 27] and [15].

It was also successfully applied in the framework of non-Newtonian fluids of power law type [4, 7, 16, 24] and even in the context of numerical analysis [14]. In [8–10] the Lipschitz truncation was used to develop an existence theory of vector valued very weak solutions of elliptic PDEs.

All of these application have in common that the desired test functions are a priori not admissible, but have to be approximated by Lipschitz functions. In order to preserve things like pointwise monotonicity of the system, it is important that the truncation takes place only on the small bad set. The bad set is usually defined in terms of the level sets of the maximal operator of the gradients.

During these years the Lipschitz truncation technique has been refined with respect to several aspects. In the stationary situation the picture is almost complete. It is now possible to preserve zero boundary value, obtain stability in all L^p -spaces and to apply the technique to sequences of functions. Moreover, the Lipschitz truncation can be interpreted as a Calderón–Zygmund decomposition in the Sobolev spaces of first order, see [3].

In the parabolic context the theory is much less developed. The parabolic Lipschitz truncation was introduced by Kinnunen–Lewis [25]. They used it to prove higher integrability for very weak solutions of the evolutive p -Laplacian systems. On the other hand, Diening–Ruzicka–Wolf [17] developed a parabolic Lipschitz truncation to show existence of fluids of power law type; i.e. the evolutive analogue to [24]. In [5, 7] a parabolic Lipschitz truncation was developed, which preserves the solenoidal structure of the given function and makes the truncation more suitable for problems from fluids dynamics.

The difficulty of the parabolic Lipschitz truncation in contrast to the stationary case is due to the fact, that the time-derivative of the solution is only defined in terms of negative Sobolev spaces or in the distributional sense. Therefore, the parabolic Lipschitz truncations mentioned above lacked the possibility to preserve zero boundary values and to obtain control on the time derivative of the truncation. In this paper we will overcome both of these problems.

In what follows we will introduce our parabolic Lipschitz truncation in the setting of p -growth assumptions. The full statement that holds for general Orlicz growth assumptions can be found in Theorem 2.3 in the next section.

Our standing assumption for the Lipschitz truncation, is that the given function w has a time derivative in the following sense:

$$\partial_t w = \operatorname{div} G \quad \text{in } \mathcal{D}'(J \times \Omega) \tag{1.1}$$

where J is a time interval and Ω is a bounded domain in \mathbb{R}^m , $m \geq 2$. We take as “bad set” a superlevel set of the maximal function of the spatial gradient and of the time derivative in the following way. Let

$$\mathcal{O}_\lambda^\alpha := \{\mathcal{M}^\alpha(\chi_{J \times \Omega} \nabla w) > \lambda\} \cup \{\alpha \mathcal{M}^\alpha(\chi_{J \times \Omega} G) > \lambda\},$$

where $\lambda > 0$ and the α -parabolic maximal function \mathcal{M}^α is defined using the (backwards in time) parabolic cylinders $Q_r^\alpha := (-\alpha r^2, 0) \times B_r$ in the following way:

$$(\mathcal{M}^\alpha g)(x) := \sup_{Q \in \mathcal{Q}^\alpha : x \in Q} \int_Q |g|. \tag{1.2}$$

where \mathcal{Q}^α is the family of cylinders $\mathcal{Q}_r^\alpha, r > 0$.

Here α is a scaling quantity, to allow different integrability assumptions on ∇w and G . Having collected the necessary notation we may state the theorem.

Theorem 1.1 *Let $G \in L^{p'}(J \times \Omega)$ and $w \in L^p(J, W_0^{1,p}(\Omega))$ satisfy (1.1). Then there exists an approximation $w_\lambda^\alpha \in L^p(J, W_0^{1,p}(\Omega))$ with the following properties:*

- (a) $w_\lambda^\alpha = w$ on $(\mathcal{O}_\lambda^\alpha)^c$.
- (b) $\mathcal{M}^\alpha(\nabla w_\lambda^\alpha) \leq c\lambda$, i.e. w_λ^α is Lipschitz continuous with respect to space.
- (c)

$$\int_{J \times \Omega} |\nabla(w_\lambda^\alpha - w)|^p dz \leq c \int_{\mathcal{O}_\lambda^\alpha} |\nabla w|^p + \lambda^p |\mathcal{O}_\lambda^\alpha|.$$

- (d) $\alpha \mathcal{N}^\alpha(\partial_t w_\lambda^\alpha) \leq c\lambda$ where \mathcal{N}^α is defined in (2.11).
- (e) w_λ^α is Lipschitz continuous with respect to the scaled, parabolic metric, i.e.

$$|w_\lambda^\alpha(t, x) - w_\lambda^\alpha(s, y)| \leq c\lambda \max \left\{ \frac{|t - s|^{\frac{1}{2}}}{\alpha^{\frac{1}{2}}}, |x - y| \right\}$$

for all $(t, x), (s, y) \in J \times \Omega$.

- (f) for $J = (t^-, t^+)$ and arbitrary $\eta \in W_0^{1,\infty}(-\infty, t^+)$ it holds::

$$\langle \partial_t w, w_\lambda^\alpha \eta \rangle = \frac{1}{2} \int_{\mathcal{Q}} (|w_\lambda^\alpha|^2 - 2w \cdot w_\lambda^\alpha) \partial_t \eta dz + \int_{\mathcal{O}_\lambda^\alpha} (\partial_t w_\lambda^\alpha)(w_\lambda^\alpha - w) \eta dz.$$

Observe, that (d) shows that our approximation does also approximate the distributional time-derivative. The maximal operator \mathcal{N}^α is defined in terms of the distributional time derivative. It seems to be a novel tool to quantify the distributional time derivative in such a way. In a way the boundedness of $\mathcal{N}^\alpha(\partial_t w_\lambda^\alpha)$ corresponds to $\partial_t w_\lambda^\alpha \in L^\infty(J, W^{-1,\infty}(\Omega))$.

As an application of our parabolic Lipschitz truncation, we present a new *caloric approximation lemma*. We show that every ‘‘almost p -caloric’’ function has a p -caloric approximation ‘‘close enough’’. The following theorem is the p -version of the more general result for Orlicz function, see Theorem 4.2.

Theorem 1.2 *Let $p \in (1, \infty)$ and \mathcal{Q} be a times-space cylinder, $\mathcal{Q} = I \times B = (t^-, t^+) \times B$. Let $\sigma \in (0, 1), q \in [1, \infty)$ and $\theta \in (0, 1)$. Moreover, let $\tilde{\mathcal{Q}}$ be such that $\mathcal{Q} \subset \tilde{\mathcal{Q}} \subset 2\mathcal{Q}$. Then, for all $\epsilon > 0$ there exists a $\delta > 0$ s.t. the following holds: if $u \in L^p(I, W_0^{1,p}(B)), u_t = \text{div } G, G \in L^{p'}(J \times \Omega)$, is almost p -caloric in the sense that for all $\xi \in C_0^\infty(\mathcal{Q})$,*

$$\left| \int_{\mathcal{Q}} u \partial_t \xi + |\nabla u|^{p-2} \nabla u \nabla \xi dz \right| \leq \delta \left(\int_{\tilde{\mathcal{Q}}} |\nabla u|^p + |G|^{p'} dz + \|\nabla \xi\|_\infty^p \right)$$

then there exists a p -caloric function h s.t. $h = u$ on $\partial_p \mathcal{Q}$ and

$$\begin{aligned} & \left(\int_I \left(\int_B \left(\frac{|u - h|^2}{|t^+ - t^-|} \right)^\sigma dx \right)^{\frac{q}{\sigma}} dt \right)^{\frac{1}{q}} + \left(\int_{\mathcal{Q}} |V(\nabla u) - V(\nabla h)|^{2\theta} dz \right)^{\frac{1}{\theta}} \\ & \leq \epsilon \int_{\tilde{\mathcal{Q}}} |\nabla u|^p + |G|^{p'} dz. \end{aligned}$$

where $V(z) = |z|^{\frac{p-2}{2}} z$.

If u would be p -caloric, then we could choose $\delta = 0$ in the assumption of Theorem 1.2 and $h = u$ as an approximation. The small parameter $\delta > 0$ indicates, that u behaves like a small perturbation of a p -caloric function. This smallness however is only needed in reaction to very regular test functions ξ . Nevertheless, Theorem 1.2 ensures that u is close to a p -caloric function h . The closeness is expressed up to a small loss in the exponent in the natural distance of the p -heat equation, which are $L^\infty(L^2)$ and $L^p(W^{1,p})$. In particular, we have control on the distance in the sense of space and time derivatives.

In the stationary case, the method is called *harmonic approximation lemma* and its idea goes back to De Giorgi. He used it in geometric measure theory to prove regularity of harmonic maps. See [22] for an overview on the harmonic approximation lemma. The closeness in the sense of gradients and the preservation of the boundary values was introduced in [19].

The p -caloric approximation method was developed by Bögelein, Duzaar and Mingione [6], (see also [21, 23]). We wish to quickly point the improvements of the approximation lemma here with respect to the one in [6]. First, our assumptions are weaker: we only assume (1.1) and we deduce the validity of a Poincaré inequality. Second, our proof is directly and completely avoids any argument by contradiction and the proof is much shorter. Second, this direct approach via the parabolic Lipschitz truncations gives us a much finer control on the quantities. Indeed, we can show closeness of the p -caloric approximation function both in $L^q(L^{2\sigma})$ and $L^{p\theta}(W^{1,p\theta})$ norms, (the last closeness is via the natural quantity $V(z) = |z|^{\frac{p-2}{2}}z$). This means that our estimates measure the closeness of the weak time derivatives and spatial gradients in a quantitative way. Third, we can preserve boundary values, which is very handy for applications. Forth, our technique is developed in the more general frame of Orlicz spaces.

Besides the potential use for applications that was described above, we will demonstrate the useability of our method in a forthcoming paper where we will consider parabolic systems with critical growth, [18].

2 Parabolic Lipschitz truncation

In this section we derive an improved version of the parabolic Lipschitz truncation. Earlier versions are due to [26] and [17].

We start by assuming that $w \in L^1(J, W_0^{1,1}(\Omega))$ is a distributional solution (possible vectorial) to

$$\begin{aligned} \partial_t w &= \operatorname{div} G && \text{in } \mathcal{D}'(J \times \Omega) \\ w &= 0 && \text{on } \partial_{\text{par}}(J \times \Omega). \end{aligned} \tag{2.1}$$

Here $J = (-t_0, 0)$ denotes the time interval. The space domain $\Omega \subset \mathbb{R}^m$ should have the fat complement property, see Remark 2.1. In particular, it suffices that Ω is a bounded open domain with Lipschitz boundary. In many applications it is enough to consider the case where Ω is a ball or a cube. By $\partial_{\text{par}}(J \times \Omega)$ we denote the parabolic boundary of $J \times \Omega = (\{-t_0\} \times \Omega) \cup (J \times \partial\Omega)$. The function G will at least be in $L^1(J \times \Omega)$. Note that the zero boundary values on the parabolic boundary are well defined due to $w \in L^1(J, W_0^{1,1}(\Omega))$ and $\partial_t w \in L^1(J, (W_0^{1,\infty}(\Omega))^*)$.

Remark 2.1 It is sufficient for us to consider domains Ω that have the *fat complement property*, i.e. there exists $A_1 \geq 1$ such that for all $x \in \Omega$

$$|B_{2 \operatorname{dist}(x, \Omega^c)}(x)| \leq A_1 |B_{2 \operatorname{dist}(x, \Omega^c)}(x) \cap \Omega^c|. \tag{2.2}$$

If $\Omega \subset \mathbb{R}^d$ is an open bounded set with Lipschitz boundary then Ω has the fat complement property.

Let us recall some definitions and results that are standard in the context of N-functions. A real function $\phi : \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}^{\geq 0}$ is said to be an N-function if it satisfies the following conditions: $\phi(0) = 0$ and there exists the derivative ϕ' of ϕ . This derivative is right continuous, non-decreasing and satisfies $\phi'(0) = 0, \phi'(t) > 0$ for $t > 0$, and $\lim_{t \rightarrow \infty} \phi'(t) = \infty$. Moreover, ϕ is convex.

We say that ϕ satisfies the Δ_2 -condition, if there exists $c > 0$ such that for all $t \geq 0$ holds $\phi(2t) \leq c \phi(t)$. We denote the smallest possible constant by $\Delta_2(\phi)$. Since $\phi(t) \leq \phi(2t)$ the Δ_2 condition is equivalent to $\phi(2t) \sim \phi(t)$.

By L^ϕ and $W^{1,\phi}$ we denote the classical Orlicz and Sobolev-Orlicz spaces, i.e. $f \in L^\phi$ iff $\int \phi(|f|) dx < \infty$ and $f \in W^{1,\phi}$ iff $f, \nabla f \in L^\phi$. By $W_0^{1,\phi}(\Omega)$ we denote the closure of $C_0^\infty(\Omega)$ in $W^{1,\phi}(\Omega)$.

By $(\phi')^{-1} : \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}^{\geq 0}$ we denote the function

$$(\phi')^{-1}(t) := \sup \{s \in \mathbb{R}^{\geq 0} : \phi'(s) \leq t\}.$$

If ϕ' is strictly increasing then $(\phi')^{-1}$ is the inverse function of ϕ' . Then $\phi^* : \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}^{\geq 0}$ with

$$\phi^*(t) := \int_0^t (\phi')^{-1}(s) ds$$

is again an N-function and $(\phi^*)'(t) = (\phi')^{-1}(t)$ for $t > 0$. It is the complementary function of ϕ . Note that $\phi^*(t) = \sup_{s \geq 0}(st - \phi(s))$ and $(\phi^*)^* = \phi$. For all $\delta > 0$ there exists c_δ (only depending on $\Delta_2(\phi, \phi^*)$ such that for all $t, s \geq 0$ holds

$$t s \leq \delta \phi(t) + c_\delta \phi^*(s), \tag{2.3}$$

This inequality is called *Young's inequality*. For all $t \geq 0$

$$\begin{aligned} \frac{t}{2} \phi' \left(\frac{t}{2} \right) &\leq \phi(t) \leq t \phi'(t), \\ \phi \left(\frac{\phi^*(t)}{t} \right) &\leq \phi^*(t) \leq \phi \left(\frac{2\phi^*(t)}{t} \right). \end{aligned} \tag{2.4}$$

Therefore, uniformly in $t \geq 0$

$$\phi(t) \sim \phi'(t) t, \quad \phi^*(\phi'(t)) \sim \phi(t), \tag{2.5}$$

where the constants only depend on $\Delta_2(\phi, \phi^*)$.

We will assume that ϕ satisfies the following assumption.

Assumption 2.2 Let ϕ be an N-function such that ϕ is C^1 on $[0, \infty)$ and C^2 on $(0, \infty)$. Further assume that

$$\phi'(t) \sim t \phi''(t) \tag{2.6}$$

uniformly in $t > 0$. The constants in (2.6) are called the *characteristics of ϕ* .

We remark that under these assumptions $\Delta_2(\phi, \phi^*) < \infty$ will be automatically satisfied, where $\Delta_2(\phi, \phi^*)$ depends only on the characteristics of ϕ .

For given ϕ we define the associated N-function ψ by

$$\psi'(t) := \sqrt{\phi'(t)t}. \tag{2.7}$$

We remark that if ϕ satisfies Assumption 2.2, then also ϕ^* , ψ , and ψ^* satisfy this assumption.

The idea of the parabolic Lipschitz truncation is to cut certain maximal functions of the gradient and the time derivative. Since the time derivative is only defined in the weak sense by $\partial_t w = \operatorname{div}G$, we will cut the maximal operator of G instead of $\partial_t w$.

The properties of the Lipschitz truncation are summarized in the following theorem.

Theorem 2.3 *Let $w \in L^1(J, W_0^{1,1}(\Omega))$ and $\nabla w \in L^\phi(J \times \Omega)$ satisfies (2.1). For $\lambda, \alpha > 0$ define the bad set $\mathcal{O}_\lambda^\alpha$ by*

$$\mathcal{O}_\lambda^\alpha := \{\mathcal{M}^\alpha(\chi_{J \times \Omega} \nabla w) > \lambda\} \cup \{\alpha \mathcal{M}^\alpha(\chi_{J \times \Omega} G) > \lambda\}, \tag{2.8}$$

Then there exists an approximation $w_\lambda^\alpha \in L^\phi(J, W_0^{1,\phi}(\Omega))$ with the following properties:

- (a) $w_\lambda^\alpha = w$ on $(\mathcal{O}_\lambda^\alpha)^c$.
- (b) $\mathcal{M}^\alpha(\nabla w_\lambda^\alpha) \leq c\lambda$, i.e. w_λ^α is Lipschitz with respect to space.
- (c)

$$\int_{J \times \Omega} \phi(|\nabla(w_\lambda^\alpha - w)|) \, dz \leq c \int_{\mathcal{O}_\lambda^\alpha} \phi(|\nabla w|) + \phi(\lambda)|\mathcal{O}_\lambda^\alpha|.$$

- (d) $\alpha \mathcal{N}^\alpha(\partial_t w_\lambda^\alpha) \leq c\lambda$ where \mathcal{N}^α is defined in (2.11).
- (e) w_λ^α is Lipschitz continuous with respect to the scaled, parabolic metric, i.e.

$$|w_\lambda^\alpha(t, x) - w_\lambda^\alpha(s, y)| \leq c\lambda \max \left\{ \frac{|t - s|^{\frac{1}{2}}}{\alpha^{\frac{1}{2}}}, |x - y| \right\}$$

for all $(t, x), (s, y) \in J \times \Omega$.

- (f) for $J = (t^-, t^+)$ and arbitrary $\eta \in W_0^{1,\infty}(-\infty, t^+)$ it holds:

$$\langle \partial_t w, w_\lambda^\alpha \eta \rangle = \frac{1}{2} \int_Q (|w_\lambda^\alpha|^2 - 2w \cdot w_\lambda^\alpha) \partial_t \eta \, dz + \int_{\mathcal{O}_\lambda^\alpha} (\partial_t w_\lambda^\alpha)(w_\lambda^\alpha - w) \eta \, dz$$

The proof will be achieved through several lemmas.

2.1 Parabolic Poincaré type inequality

The goal of this subsection is to derive a very weak form of the parabolic Poincaré inequality on parabolic cylinders, where the time derivative is just defined in a weak sense, see Theorem 2.8.

We start with some notations. By $B_r(x)$, resp. $I_r(t)$, we denote the standard euclidean ball with radius r and center $x \in \mathbb{R}^m$, resp. $t \in \mathbb{R}$. For $\alpha > 0$ define the α -parabolic metric $d_\alpha : \mathbb{R} \times \mathbb{R}^m \rightarrow [0, \infty)$ by

$$d_\alpha((t, x), (\tau, y)) := \max \{ \alpha^{-\frac{1}{2}} |t - \tau|^{\frac{1}{2}}, |x - y| \}.$$

The balls with radius r respect to d_α are called α -parabolic cylinders with radius r . Any α -parabolic cylinder Q can be represented in terms of euclidean balls, i.e.

$$Q = Q_r^\alpha(t, x) := I_{\alpha r^2}(t) \times B_r(x) = I \times B.$$

for some $(t, x) \in \mathbb{R}^{m+1}$, where r is the radius of Q .

By σQ (for $\sigma > 0$) we denote the parabolic scaled cylinder with the same center but σ -times the radius with respect to d_α . In particular, for $Q = I \times B$ we have $\sigma Q = (\sigma^2 I) \times (\sigma B)$. We denote by $|E|$ the Lebesgue measure of E for a measurable set E and by χ_E its characteristic function. We define

$$\int_E |f| dx =: \frac{1}{|E|} \int_E |f| dx.$$

For a non-negative integrable function η we define

$$\langle f \rangle_\eta := \frac{1}{\|\eta\|_1} \int f \eta dx$$

and for a measurable set E we define $\langle f \rangle_E := \langle f \rangle_{\chi_E}$. The integration is taken over the natural domain of f , so if f is defined on Q , then the integral is over Q .

We need the following version of the norm conjugate formula for $L^1_0(I)$.

Lemma 2.4 *Let $f \in L^1(I)$, then*

$$\int_I |f - \langle f \rangle_I| dt \leq 2 \sup_{\beta \in C^\infty_{0,0}(I), \|\beta\|_\infty \leq 1} \int_I f \beta dt \leq 2 \int_I |f - \langle f \rangle_I| dt.$$

Proof The second estimate is obvious, so we just need to prove the first one. It suffices to prove the case $I = (0, 1)$. Fix $\delta > 0$. Then due to the isometry $(L^1(I))^* = L^\infty(I)$, we can find $g \in L^\infty(I)$ with $\|g\|_\infty \leq 1$, such that

$$\int_I |f - \langle f \rangle_I| dt \leq \delta + \int_I (f - \langle f \rangle_I) g dt = \delta + \int_I f(g - \langle g \rangle_I) dt. \tag{2.9}$$

For $\epsilon \in (0, \frac{1}{4})$ define $I_\epsilon = (\epsilon, 1 - \epsilon)$. Let ψ_ϵ denote a standard mollifier with $\text{supp } \psi_\epsilon \subset B_\epsilon(0)$. Define

$$h_\epsilon := (\chi_{I_\epsilon}(g - \langle g \rangle_{I_\epsilon})) * \psi_{\epsilon/2}.$$

It is easy to see that $h_\epsilon \in C^\infty_{0,0}(I)$, (subspace of C^∞_0 whose elements have mean value zero), $h_\epsilon \rightarrow g - \langle g \rangle_I$ almost everywhere for $\epsilon \rightarrow 0$, $\|h_\epsilon\|_{L^\infty(I)} \leq 2 \|g\|_\infty$. In particular, it follows by the dominated convergence theorem that

$$\int_I f(g - \langle g \rangle_I) dt = \lim_{\epsilon \rightarrow 0} \int_I f h_\epsilon dt.$$

This and (2.9) imply

$$\int_I |f - \langle f \rangle_I| dt \leq \delta + \sup_{h_\epsilon \in C^\infty_{0,0}(I), \|h_\epsilon\|_\infty \leq 2} \int_I f h_\epsilon dt.$$

The claim follows, since $\delta > 0$ was arbitrary. □

Lemma 2.5 *Let $f \in L^1(I)$, then*

$$\int_I |f - \langle f \rangle_I| dt \leq 2 \sup_{\gamma \in C^\infty_0(I), \|\gamma'\|_\infty \leq 1} \left| \int_I f \gamma' dt \right| \leq 2 \int_I |f - \langle f \rangle_I| dt.$$

Proof This follows immediately from Lemma 2.4. Indeed, if $\beta \in C_{0,0}^\infty(I)$, then its primitive $\gamma(t) := \int_{-\infty}^t \beta(s) ds$ satisfies $\gamma \in C_{0,0}^\infty(I)$. On the other hand for every $\gamma \in C_{0,0}^\infty(I)$, we have $\gamma' \in C_{0,0}^\infty(I)$. □

For an α -parabolic cylinder $Q = Q_r = I_{\alpha r^2} \times B_r$ we define

$$\mathcal{F}_Q := \{ \xi \in C_0^\infty(Q) : \| \xi \|_{\mathcal{F}_Q} := \| \xi \|_\infty + r \| \nabla \xi \|_\infty + \alpha r^2 \| \partial_t \xi \|_\infty \leq 1 \},$$

Define

$$\begin{aligned} \mathcal{M}_Q(a) &:= \int_Q |a| dz, \\ \mathcal{M}_Q^{\sharp,1}(a) &:= \int_Q \frac{|a - \langle a \rangle_Q|}{r_Q} dz \end{aligned}$$

For a distribution $a \in \mathcal{D}'(Q)$ we define

$$\mathcal{N}_Q(a) := \sup_{\xi \in \mathcal{F}_Q} (r |Q|^{-1} | \langle a, \xi \rangle |).$$

We use the letter \mathcal{N} for “negative”, since we measure somehow the local information on $\partial_t a$ in a negative space. We can observe that

$$(\mathcal{M}^\alpha a)(x) = \sup_{Q \in \mathcal{Q}^\alpha : x \in Q} \mathcal{M}_Q(a), \tag{2.10}$$

We also define the maximal operator

$$(\mathcal{N}^\alpha a)(x) := \sup_{Q \in \mathcal{Q}^\alpha : x \in Q} \mathcal{N}_Q(a). \tag{2.11}$$

Remark 2.6 If $\partial_t a = \operatorname{div} G$ on Q , then

$$\begin{aligned} \mathcal{N}_Q(\partial_t a) &= \sup_{\xi \in \mathcal{F}_Q} (r |Q|^{-1} | \langle \partial_t a, \xi \rangle |) \\ &= \sup_{\xi \in \mathcal{F}_Q} (r |Q|^{-1} | \langle G, \nabla \xi \rangle |) \\ &\leq \int_Q |G| dz. \end{aligned}$$

We need the following version of parabolic Poincaré’s inequality with respect to time.

Lemma 2.7 *Let $\eta \in C_0^\infty(B)$ with $\eta \geq 0$, $\int_B \eta(x) dx > 0$ and $\| \eta \|_\infty + r \| \nabla \eta \|_\infty \leq c_0 |B|^{-1} \| \eta \|_1$. Then for every α -parabolic cube $Q = I \times B$ we have*

$$\int_I | \langle a(t) \rangle_\eta - \langle a \rangle_{\eta \times I} | dt \leq c r \alpha \mathcal{N}_Q(\partial_t a),$$

where c depends on η only through c_0 . Here we use the notation $\langle a \rangle_{\eta \times I} = \frac{1}{|I|} \int_I \langle a(t) \rangle_\eta dt$.

Proof We can assume without loss of generality that $\int_B \eta(x) dx = 1$. From Lemma 2.5 it follows that

$$\begin{aligned} \int_I |\langle a(t) \rangle_\eta - \langle a \rangle_{\eta \times I}| dt &\leq 2 \sup_{\gamma \in C_0^\infty(I), \|\gamma'\|_\infty \leq 1} \left| \int_I \langle a(t) \rangle_{\eta \gamma'}(t) dt \right| \\ &= 2 |B| \sup_{\gamma \in C_0^\infty(I), \|\gamma'\|_\infty \leq 1} \left| \int_Q a \partial_t(\eta \gamma) dz \right|. \end{aligned}$$

We want to estimate the integral in the last expression by means of $\mathcal{N}_Q(\partial_t a)$. Let $\gamma \in C_0^\infty(I)$ with $\|\gamma'\|_\infty \leq 1$. Then $\|\gamma\|_\infty \leq c|I|$. We estimate

$$\begin{aligned} \|\eta \gamma\|_\infty &\leq \|\eta\|_\infty \|\gamma\|_\infty &&\leq c_0 |B|^{-1} |I|, \\ r \|\nabla(\eta \gamma)\|_\infty &\leq r \|\nabla \eta\|_\infty \|\gamma\|_\infty &&\leq c_0 |B|^{-1} |I|, \\ \alpha r^2 \|\partial_t(\eta \gamma)\|_\infty &\leq \|\eta\|_\infty \alpha r^2 \|\partial_t \gamma\|_\infty &&\leq c_0 |B|^{-1} |I|. \end{aligned}$$

In particular, $\|\eta \gamma\|_{\mathcal{F}_Q} \leq c_0 |B|^{-1} |I| = c_0 |B|^{-1} \alpha r^2$. Therefore, using the definition of $\mathcal{N}_Q(\partial_t a)$ we have

$$|B| \left| \int_Q a \partial_t(\eta \gamma) dz \right| \leq c |B| r^{-1} \mathcal{N}_Q(\partial_t a) \|\eta \gamma\|_{\mathcal{F}_Q} \leq c \alpha r \mathcal{N}_Q(\partial_t a).$$

and the claim follows. □

We are now in a position to state the following Poincaré inequality :

Theorem 2.8 *Let $Q = I \times B$ be α -parabolic cube and let $\rho \in L^1(Q)$ be such that $\rho \geq 0$ and $\|\rho\|_\infty \leq c_0 |Q|^{-1} \|\rho\|_1$. Then*

$$\int_Q \left| \frac{a - \langle a \rangle_\rho}{r} \right| dz \leq c \int_Q |\nabla a| dz + c \alpha \mathcal{N}_Q(\partial_t a).$$

Recall that $\mathcal{N}_Q(\partial_t a) \leq \int_Q |G| dz$ if $\partial_t u = \operatorname{div} G$ with $G \in L^1(Q)$, due to Remark 2.6.

Proof We begin with the special case $\rho = \chi_I \eta$ with η as in Lemma 2.7.

$$\begin{aligned} \int_Q \left| \frac{a - \langle a \rangle_{\eta \times I}}{r} \right| dz &\leq \int_I \int_B \left| \frac{a - \langle a(t) \rangle_\eta}{r} \right| dx dt + \int_I \left| \frac{\langle a(t) \rangle_\eta - \langle a \rangle_{\eta \times I}}{r} \right| dt \\ &=: I + II. \end{aligned}$$

Now the claim follows by using Poincaré in space for the first term and Lemma 2.7 for the second term.

Now consider the case of arbitrary ρ as in the assumptions. Then

$$\int_Q |a - \langle a \rangle_\rho| dz \leq \int_Q |a - \langle a \rangle_{\eta \times I}| dz + |\langle a \rangle_\rho - \langle a \rangle_{\eta \times I}|.$$

Now Jensen’s inequality with respect to the integration of $\langle a \rangle_\rho$ together with the assumptions on ρ imply

$$|\langle a \rangle_\rho - \langle a \rangle_{\eta \times I}| \leq \frac{\|\rho\|_{L^\infty(Q)}}{\|\rho\|_{L^1(Q)}} \int_Q |a - \langle a \rangle_{\eta \times I}| dz \leq c_0 \int_Q |a - \langle a \rangle_{\eta \times I}| dz.$$

In particular, we have

$$\int_Q |a - \langle a \rangle_\rho| dz \leq (1 + c_0) \int_Q |a - \langle a \rangle_{\eta \times I}| dz,$$

so the general case follows from the special one. □

Since the above (weak) setting can not be applied to the Orlicz setting in modular form, we include the following classical space-time Poincaré in modular Orlicz form.

Lemma 2.9 *Let $Q = I \times B$ be α -parabolic cube and let $\rho \in L^1(Q)$ be such that $\rho \geq 0$ and $\|\rho\|_\infty \leq c_0 |Q|^{-1} \|\rho\|_1$.*

Moreover, let $\partial_t a = \operatorname{div} G$ with $G \in L^1(Q)$ in the sense of distributions. Let ϕ be an Orlicz function satisfying the Δ_2 -condition. Then for every α -parabolic cube $Q = I \times B$ we have

$$\int_Q \phi\left(\left|\frac{a - \langle a \rangle_\rho}{r}\right|\right) dz \leq c \int_Q \phi(|\nabla a|) dz + c \phi\left(\alpha \int_Q |G| dz\right).$$

Proof As in Theorem 2.8 we begin with $\rho = \chi_I \eta$ with η as in Lemma 2.7. Analogously to the proof of Theorem 2.8 we estimate

$$\begin{aligned} & \int_Q \phi\left(\left|\frac{a - \langle a \rangle_{\eta \times I}}{r}\right|\right) dz \\ & \leq \int_I \int_B \phi\left(\left|\frac{a - \langle a(t) \rangle_\eta}{r}\right|\right) dx dt + \int_I \phi\left(\left|\frac{\langle a(t) \rangle_\eta - \langle a \rangle_{\eta \times I}}{r}\right|\right) dt \\ & =: I + II. \end{aligned}$$

Now I can be estimated by $\int_Q \phi(|\nabla a|) dz$ by using Poincaré in space for Orlicz functions, see e.g. [13, Theorem 7]. For the second we estimate

$$\begin{aligned} |\langle a(t) \rangle_\eta - \langle a \rangle_{\eta \times I}| &= \left| \int_I \langle a(t) \rangle_\eta - \langle a(s) \rangle_\eta ds \right| = \left| \int_I \frac{1}{\|\eta\|_{L^1(B)}} \int_s^t \langle \partial_t a(\tau), \eta \rangle d\tau ds \right| \\ &= \left| \int_I \frac{1}{\|\eta\|_{L^1(B)}} \int_s^t \langle G, \nabla \eta \rangle ds \right| \\ &\leq c\alpha r \int_Q |G| dz. \end{aligned} \tag{2.12}$$

This can be used to estimate (II) and the claim follows for $\rho = \chi_I \eta$.

Now as in the proof of Theorem 2.8 we can change to general ρ by showing in the same manner

$$\int_Q \phi(|a - \langle a \rangle_\rho|) dz \leq (1 + c_0) \int_Q \phi(|a - \langle a \rangle_{\eta \times I}|) dz.$$

□

2.2 Extension

It is convenient for our purpose to use function which are defined on the whole space $\mathbb{R} \times \mathbb{R}^m$. Therefore, we will extend our function w from (2.1) to a function on $\mathbb{R} \times \mathbb{R}^m$ such that most of its properties are preserved.

We therefore extend G and w from $J \times \Omega$ to $(-\infty, 0] \times \mathbb{R}^m$ by zero. Since $w(-t_0) = 0$ in the sense of a $(W_0^{1,\infty}(\Omega))^*$ -trace, it is easy to see that $\partial_t w = \operatorname{div} G$ on $\mathcal{D}'((-\infty, 0), \mathbb{R}^m)$.

Next, we extend w to $\mathbb{R} \times \mathbb{R}^m$ by even reflection and G by odd reflection. Then it follows that

$$\begin{aligned} \partial_t w &= \operatorname{div} G && \text{in } \mathcal{D}'(\mathbb{R} \times \Omega) \\ w &= 0 && \text{outside of } (-t_0, t_0) \times \Omega, \\ G &= 0 && \text{outside of } (-t_0, t_0) \times \Omega. \end{aligned} \tag{2.13}$$

We will construct a Lipschitz truncation w_λ^α of w on $\mathbb{R} \times \mathbb{R}^m$, which is zero outside of $(-t_0, t_0) \times \Omega$. The restriction of w_λ^α back to $J \times \Omega$ will then provide the Lipschitz truncation for our Theorem 2.3.

2.3 Whitney covering

For $\alpha, \lambda > 0$ we define the bad set $\mathcal{O}_\lambda^\alpha$ as

$$\mathcal{O}_\lambda^\alpha := \{\mathcal{M}^\alpha(\nabla w) > \lambda\} \cup \{\alpha \mathcal{M}^\alpha(G) > \lambda\}. \tag{2.14}$$

Note that this differs slightly from the definition (2.8) in the Theorem 2.3, since we extend w and G partly by reflection. This increase the maximal function $\mathcal{M}^\alpha(\nabla w)$ and $\mathcal{M}^\alpha(G)$ but at most by a factor of two. Therefore, for the sake of readability we prefer to work with (2.14). The result certainly also holds for (2.8).

According to [17, Lemma 3.1] there exists an α -parabolic Whitney covering $\{Q_j^\alpha\} = \{I_j \times B_j\}$ of $\mathcal{O}_\lambda^\alpha$ in the following sense:

- (W1) $\bigcup_j \frac{1}{2} Q_j^\alpha = \mathcal{O}_\lambda^\alpha$,
- (W2) for all $j \in \mathbb{N}$ we have $8Q_j^\alpha \subset \mathcal{O}_\lambda^\alpha$ and $16Q_j^\alpha \cap (\mathbb{R}^{m+1} \setminus \mathcal{O}_\lambda^\alpha) \neq \emptyset$,
- (W3) if $Q_j^\alpha \cap Q_k^\alpha \neq \emptyset$ then $\frac{1}{2}r_k \leq r_j \leq 2r_k$,
- (W4) $\frac{1}{4} Q_j^\alpha \cap \frac{1}{4} Q_k^\alpha = \emptyset$ for all $j \neq k$,
- (W5) each $x \in \mathcal{O}_\lambda^\alpha$ belongs to at most 120^{m+2} of the sets $4Q_j^\alpha$,

where $r_j := r_{B_j}$, the radius of B_j and $Q_j^\alpha = I_j \times B_j$.

With respect to the covering $\{Q_j^\alpha\}$ there exists a partition of unity $\{\rho_j\} \subset C_0^\infty(\mathbb{R}^{m+1})$ such that

- (P1) $\chi_{\frac{1}{2} Q_j^\alpha} \leq \rho_j \leq \chi_{\frac{3}{4} Q_j^\alpha}$
- (P2) $\|\rho_j\|_\infty + r_j \|\nabla \rho_j\|_\infty + r_j^2 \|\nabla^2 \rho_j\|_\infty + \alpha r_j^2 \|\partial_t \rho_j\|_\infty \leq c$.

For each $k \in \mathbb{N}$ we define $A_k := \{j : \frac{3}{4} Q_k^\alpha \cap \frac{3}{4} Q_j^\alpha \neq \emptyset\}$. Then

- (P3) $\sum_{j \in A_k} \rho_j = 1$ on $\frac{3}{4} Q_k^\alpha$.

We get the following additional property

- (W6) If $j \in A_k$, then $|Q_j^\alpha \cap Q_k^\alpha| \geq 16^{-m-2} \max\{|Q_j^\alpha|, |Q_k^\alpha|\}$.
- (W7) If $j \in A_k$, then $|\frac{3}{4} Q_j^\alpha \cap \frac{3}{4} Q_k^\alpha| \geq \max\{|Q_j^\alpha|, |Q_k^\alpha|\}$.

(W8) If $j \in A_k$, then $\frac{1}{2}r_k \leq r_j < 2r_k$.

(W9) $\#A_k \leq 120^{m+2}$.

Now, we define w_j by

$$w_j^\alpha := \begin{cases} \langle w \rangle_{\rho_j} & \text{if } \frac{3}{4}Q_j^\alpha \subset J \times \Omega, \\ 0 & \text{else.} \end{cases}$$

We define our truncation w_λ^α via the formula

$$w_\lambda^\alpha := w - \sum_j \rho_j(w - w_j^\alpha). \tag{2.15}$$

Since the ρ_j are locally finite, the sum is pointwise well defined. We will see later that the sum converges also as a distribution and in a few function spaces.

Note that the sum $\sum_j \rho_j(w - w_j^\alpha)$ is zero outside of $(-t_0, t_0) \times \Omega$. So also w_λ^α is zero outside of $(-t_0, t_0) \times \Omega$. In fact, we have

$$\text{supp}(\rho_j(w - w_j^\alpha)) \subset \frac{3}{4}Q_j^\alpha \cap ((-t_0, t_0) \times \Omega). \tag{2.16}$$

Indeed, $\text{supp}\rho_j \subset \frac{3}{4}Q_j^\alpha$, so the case $\frac{3}{4}Q_j^\alpha \subset J \times \Omega$ is obvious. If $\frac{3}{4}Q_j^\alpha \not\subset J \times \Omega$, then $w_j^\alpha = 0$ and the claim follows by $\text{supp}\rho_j \subset \frac{3}{4}Q_j^\alpha$ and $\text{supp}w \subset J \times \Omega$.

2.4 Estimates on the Whitney cylinders

We need a few auxiliary results that allow to estimate $w - w_j^\alpha$ on our Whitney cylinders. The estimates are based on our parabolic Poincaré’s inequality of Sect. 2.1.

Since the equation $\partial_t w = \text{div}G$ only holds on $\mathbb{R} \times \Omega$, we need the following auxiliary result to deal with the case of cylinders that our also outside of this domain. We use the fact that w is zero outside of $\mathbb{R} \times \Omega$.

Lemma 2.10 *Let Q be an α -parabolic cylinder with radius r . If $\frac{4}{5}Q \not\subset \mathbb{R} \times \Omega$, then*

$$\alpha \mathcal{N}_Q(\partial_t w) \leq c \mathcal{M}_Q(\nabla w).$$

Proof We calculate

$$\alpha \mathcal{N}_Q(\partial_t w) = \alpha \sup_{\xi \in \mathcal{F}_Q} (r|Q|^{-1}|\langle w, \partial_t \xi \rangle|) \leq c \int_Q \frac{|w|}{r} dz.$$

Let $Q =: I \times B$. Since $\frac{4}{5}Q \not\subset \mathbb{R} \times \Omega$ and Ω has fat complement, we have $|B \setminus \Omega| \geq c|B|$. Thus, we can apply the space Poincaré (with $w = 0$ outside of $\mathbb{R} \times \Omega$) to get

$$\int_Q \frac{|w|}{r} dz \leq c \int_Q |\nabla w| dz.$$

This proves the claim. □

Lemma 2.11 *The following holds.*

(a) *If $\frac{3}{4}Q_j^\alpha \subset J \times \Omega$, then $w_j^\alpha = \langle w \rangle_{\eta_j \times I}$ and*

$$\int_{\frac{3}{4}Q_j^\alpha} \left| \frac{w - w_j^\alpha}{r_j} \right| dz \leq c \mathcal{M}_{\frac{3}{4}Q_j^\alpha}(\nabla w) + c \alpha \mathcal{N}_{\frac{3}{4}Q_j^\alpha}(\partial_t w).$$

(b) If $\frac{4}{5}Q_j^\alpha \subset \mathbb{R} \times \Omega$ and $\frac{3}{4}Q_j^\alpha \not\subset J \times \Omega$, then $w_j^\alpha = 0$ and

$$\int_{\frac{4}{5}Q_j^\alpha} \left| \frac{w - w_j^\alpha}{r_j} \right| dz \leq c \mathcal{M}_{\frac{4}{5}Q_j^\alpha}(\nabla w) + c \alpha \mathcal{N}_{\frac{4}{5}Q_j^\alpha}(\partial_t w).$$

(c) If $\frac{4}{5}Q_j^\alpha \not\subset \mathbb{R} \times \Omega$, then $w_j^\alpha = 0$ and

$$\int_{Q_j^\alpha} \left| \frac{w - w_j^\alpha}{r_j} \right| dz \leq c \mathcal{M}_{Q_j^\alpha}(\nabla w).$$

Proof Part (a) follows immediately from Theorem 2.8 with $\rho = \rho_j$.

Let us consider part (b). In this situation $\mathbb{R} \times \Omega \setminus (J \times \Omega)$ contains a large part of $\frac{4}{5}Q_j^\alpha$ so that we can find a function $\rho \in L^\infty$ with support in $\frac{4}{5}Q_j^\alpha \cap ((\mathbb{R} \times \Omega) \setminus (J \times \Omega))$ such that $\|\rho\|_\infty \leq c |Q_j^\alpha|^{-1} \|\rho\|_1$. Since $w = 0$ on $\text{supp}(\rho)$, we have $w_j^\alpha = 0 = \langle w \rangle_\rho$. Again the claim follows by Theorem 2.8.

Let us now prove (c). Since $\frac{4}{5}Q_j^\alpha \not\subset \mathbb{R} \times \Omega$, we can find a function ρ with support outside in $Q_j^\alpha \cap \mathbb{R} \times \Omega$ with $\|\rho\|_\infty \leq c |Q_j^\alpha|^{-1} \|\rho\|_1$. Since $w = 0$ on $\text{supp}(\rho)$, we have $w_j^\alpha = 0 = \langle w \rangle_\rho$. Now Theorem 2.8 proofs our claim with an additional $\mathcal{N}_{Q_j^\alpha}(\partial_t w)$ on term on the right hand side. Due to Lemma 2.10 this term can be controlled again by $\mathcal{M}_{Q_j^\alpha}(\nabla w)$, which proves our claim. \square

Lemma 2.12 *We have*

$$\int_{\frac{3}{4}Q_j^\alpha} \left| \frac{w - w_j^\alpha}{r_j} \right| dz \leq \int_{Q_j^\alpha} |\nabla w| dz + \alpha \mathcal{N}_{\frac{4}{5}Q_j^\alpha}(\partial_t w) + \alpha \int_{Q_j^\alpha} |G| dz \leq c \lambda.$$

Proof Since $16Q_j^\alpha \cap (\mathbb{R}^{m+1} \setminus O_\lambda^\alpha) \neq \emptyset$, it follows that $\mathcal{M}_{16Q_j^\alpha}(\nabla w) \leq \lambda$ and $\alpha \mathcal{M}_{16Q_j^\alpha}(G) \leq \lambda$. Thus also $\mathcal{M}_{Q_j^\alpha}(\nabla w) \leq c \lambda$ and $\alpha \mathcal{M}_{Q_j^\alpha}(G) \leq c \lambda$.

The estimate $\alpha \mathcal{N}_{Q_j^\alpha}(\partial_t w) \leq c \lambda$ follows from Remark 2.6 if $\frac{4}{5}Q_j^\alpha \subset \mathbb{R} \times \Omega$ and from Lemma 2.10 if $\frac{4}{5}Q_j^\alpha \not\subset \mathbb{R} \times \Omega$. \square

Lemma 2.13 *We have*

$$\int_{\frac{3}{4}Q_j^\alpha} \phi \left(\left| \frac{w - w_j^\alpha}{r_j} \right| \right) dz \leq \int_{Q_j^\alpha} \phi(|\nabla w|) dz + \phi \left(\alpha \int_{Q_j^\alpha} |G| dx \right).$$

Proof The proof is similar to the one of Lemma 2.11 and Lemma 2.12 by using Lemma 2.9 instead of Theorem 2.8. \square

2.5 Stability

In this subsection we will show the stability of the Lipschitz truncation with respect to some norms.

Lemma 2.14 *If $w \in L^1(J, W_0^{1,1}(\Omega))$ and $G \in L^1(J \times \Omega)$, then $w_\lambda^\alpha \in L^1(J, W_0^{1,1}(\Omega))$. Moreover,*

$$\int_{O_\lambda^\alpha} |\nabla(w - w_\lambda^\alpha)| dz \leq c \int_{O_\lambda^\alpha} |\nabla w| dz + \lambda |O_\lambda^\alpha|.$$

Proof It follows from the definition of w_λ^α that

$$w - w_\lambda^\alpha = \sum_j \rho_j(w - w_j^\alpha).$$

Due to (2.16) the sum is zero outside of $\mathcal{O}_\lambda^\alpha$. Using that $\sum_j \rho_j = 1$ on $\mathcal{O}_\lambda^\alpha$ we get

$$\nabla(w - w_\lambda^\alpha) = \nabla w + \sum_j \nabla \rho_j(w - w_j^\alpha). \tag{2.17}$$

Now it follows with the help of (P2), (W1), (W5) and (2.16) that

$$\int_{\mathcal{O}_\lambda^\alpha} |\nabla(w - w_\lambda^\alpha)| \, dz \leq \int_{\mathcal{O}_\lambda^\alpha} |\nabla w| \, dz + c \sum_k \int_{\frac{3}{4}\mathcal{Q}_j^\alpha} \left| \frac{w - w_j^\alpha}{r_j} \right| \, dz.$$

This and Lemma 2.12 implies

$$\int_{\mathcal{O}_\lambda^\alpha} |\nabla(w - w_\lambda^\alpha)| \, dz \leq \int_{\mathcal{O}_\lambda^\alpha} |\nabla w| \, dz + c \lambda |\mathcal{O}_\lambda^\alpha|,$$

which proves the lemma. □

Lemma 2.15 *We get*

$$\int_{\mathcal{O}_\lambda^\alpha} \phi(|\nabla(w - w_\lambda^\alpha)|) \, dz \leq c \int_{\mathcal{O}_\lambda^\alpha} \phi(|\nabla w|) + c |\mathcal{O}_\lambda^\alpha| \phi(\lambda).$$

Proof The proof is similar to Lemma 2.14. Starting with (2.17) and using (P2), (W1), (W5), (2.16), and the Δ_2 -condition we get

$$\int_{\mathcal{O}_\lambda^\alpha} \phi(|\nabla(w - w_\lambda^\alpha)|) \, dz \leq \int_{\mathcal{O}_\lambda^\alpha} \phi(|\nabla w|) \, dz + c \sum_k \int_{\frac{3}{4}\mathcal{Q}_j^\alpha} \phi\left(\left|\frac{w - w_j^\alpha}{r_j}\right|\right) \, dz.$$

By Lemma 2.13 we can estimate the summands of the second part by

$$\int_{\frac{3}{4}\mathcal{Q}_j^\alpha} \phi\left(\left|\frac{w - w_j^\alpha}{r_j}\right|\right) \, dz \leq c \int_{\frac{3}{4}\mathcal{Q}_j^\alpha} \phi(|\nabla w|) \, dz + c |\mathcal{Q}_j^\alpha| \phi\left(\alpha \int_{\frac{3}{4}\mathcal{Q}_j^\alpha} |G| \, dz\right).$$

Using Lemma 2.12, we see that the mean value integral in the last term is bounded by $c \lambda$, so overall we get

$$\int_{\mathcal{O}_\lambda^\alpha} \phi(|\nabla(w - w_\lambda^\alpha)|) \, dz \leq \int_{\mathcal{O}_\lambda^\alpha} \phi(|\nabla w|) \, dz + c |\mathcal{O}_\lambda^\alpha| \phi(\lambda),$$

which proves the lemma. □

3 Lipschitz property

In this section we show that the truncated function w_λ^α has some sort of Lipschitz properties. In particular, we used $\mathcal{M}^\alpha(\nabla w)$ and $\alpha \mathcal{N}^\alpha(\partial_t w)$ (more precisely its upper bound $\mathcal{M}^\alpha(G)$) to define the bad set, where we truncate the function. It turns out that $\mathcal{M}^\alpha(\nabla w_\lambda^\alpha) + \alpha \mathcal{N}^\alpha(\partial_t w_\lambda^\alpha) \leq c \lambda$.

Lemma 3.1

$$\sum_{j \in A_k} \frac{|w_j^\alpha - w_k^\alpha|}{r_j} \leq c \sum_{j \in A_k} \int_{\frac{3}{4}Q_j^\alpha} \frac{|w - w_j^\alpha|}{r_j} dz \leq c \lambda.$$

Proof Due to (W7) and (W8) for every $j \in A_k$ holds $|\frac{3}{4}Q_j^\alpha \cap \frac{3}{4}Q_k^\alpha| \geq \max\{|\mathcal{Q}_j^\alpha|, |\mathcal{Q}_k^\alpha|\}$ and $r_j \geq \frac{1}{2}r_k$. Thus we can estimate

$$\begin{aligned} \sum_{j \in A_k} \frac{|w_j^\alpha - w_k^\alpha|}{r_j} &\leq \sum_{j \in A_k} \int_{Q_j^\alpha \cap Q_k^\alpha} \frac{|w_j^\alpha - w_k^\alpha|}{r_j} dz \\ &\leq c \int_{Q_j^\alpha \cap Q_k^\alpha} \frac{|w - w_k^\alpha|}{r_k} dz + c \sum_{j \in A_k} \int_{Q_j^\alpha \cap Q_k^\alpha} \frac{|w - w_j^\alpha|}{r_j} dz \\ &\leq \sum_{j \in A_k} \int_{Q_j^\alpha} \frac{|w - w_j^\alpha|}{r_k} dz, \end{aligned}$$

where we also used $k \in A_k$. The rest follows by Lemma 2.12. □

We need the following geometric alternatives.

Lemma 3.2 *Let Q be an α -parabolic cylinder with radius r . Then at least one of the following alternatives holds.*

- (A1) *There exists $k \in \mathbb{N}$ such that $Q \cap \frac{1}{2}Q_k^\alpha \neq \emptyset$, $8r \leq r_k$ and $Q \subset \frac{3}{4}Q_k^\alpha$.*
- (A2) *For all $j \in \mathbb{N}$ with $Q \cap \frac{3}{4}Q_j^\alpha \neq \emptyset$, there holds $r_j \leq 16r$ and $|\mathcal{Q}_j^\alpha| \leq 8^{m+2}|\mathcal{Q}_j^\alpha \cap Q|$.*
Moreover, $137Q \cap (\mathbb{R}^{m+1} \setminus \mathcal{O}_\lambda^\alpha) \neq \emptyset$.

Proof If there exists $k \in \mathbb{N}$ such that $Q \cap \frac{1}{2}Q_k^\alpha \neq \emptyset$ and $8r \leq r_k$, then automatically $Q \subset Q_k^\alpha$. Assume now that such an k does not exist. Then for every $l \in \mathbb{N}$ with $Q \cap \frac{1}{2}Q_l \neq \emptyset$, there holds $r_l \leq 8r$. Suppose that $Q \cap \frac{3}{4}Q_j^\alpha \neq \emptyset$. Now let $x \in Q \cap \frac{3}{4}Q_j^\alpha$, then by (W1) there exists m such that $x \in \frac{1}{2}Q_m$. In particular, we have $Q \cap \frac{3}{4}Q_j^\alpha \neq \emptyset$ and $\frac{1}{2}Q_m \cap \frac{3}{4}Q_j^\alpha \neq \emptyset$, since both sets contain x . Now, our assumption and $Q \cap \frac{3}{4}Q_j^\alpha \neq \emptyset$ implies $r_m \leq 8r$. On the other hand $\frac{1}{2}Q_m \cap \frac{3}{4}Q_j^\alpha \neq \emptyset$ and (W3) imply $r_j \leq 2r_m$. Thus, $r_j \leq 16r$. Moreover, it follows from $8r \geq r_m$ that $137Q = (1 + 17 \cdot 8)Q \supset 16Q_m$. Since $16Q_m \cap (\mathbb{R}^{m+1} \setminus \mathcal{O}_\lambda^\alpha) \neq \emptyset$, we also get $137Q \cap (\mathbb{R}^{m+1} \setminus \mathcal{O}_\lambda^\alpha) \neq \emptyset$. Now, let $z_0 \in Q \cap \frac{3}{4}Q_j^\alpha$. It remains to prove $|\mathcal{Q}_j^\alpha| \leq 8^{m+2}|\mathcal{Q}_j^\alpha \cap Q|$. If $r \leq \frac{1}{8}r_j$, then $Q \subset Q_j^\alpha$ and the claim follows. If $r \geq \frac{1}{8}r_j$, then there exists an α -parabolic cylinder Q' with radius $\frac{1}{8}r_j$ such that $Q' \subset Q_j^\alpha \cap Q$. So in this case $|\mathcal{Q}_j^\alpha \cap Q| \geq |\mathcal{Q}'| \geq 8^{-m-2}|\mathcal{Q}_j^\alpha|$. □

Lemma 3.3 *There holds*

$$\mathcal{M}^\alpha(\nabla w_\lambda^\alpha) \leq c \lambda.$$

Proof Let Q be an α -parabolic cylinder with radius R . We use the alternatives of Lemma 3.2.

We begin with alternative (A1). In particular, there exists $k \in \mathbb{N}$ such that $Q \cap \frac{1}{2}Q_k^\alpha \neq \emptyset$, $8R \leq r_k$ and $Q \subset \frac{3}{4}Q_k^\alpha$.

Then $w = \sum_{j \in A_k} \rho_j w_j^\alpha$ on Q and therefore

$$\begin{aligned} \mathcal{M}_Q(\nabla w_\lambda^\alpha) &= \mathcal{M}_Q(\nabla(w_\lambda^\alpha - w_k^\alpha)) = \mathcal{M}_Q\left(\nabla\left(\sum_{j \in A_k} \rho_j(w_j^\alpha - w_k^\alpha)\right)\right) \\ &\leq \sum_{j \in A_k} \mathcal{M}_Q\left(\nabla(\rho_j(w_j^\alpha - w_k^\alpha))\right) \\ &\leq c \sum_{j \in A_k} \frac{|w_j^\alpha - w_k^\alpha|}{r_j}. \end{aligned}$$

Now, Lemma 3.1 implies $\mathcal{M}_Q(\nabla w_\lambda^\alpha) \leq c\lambda$.

We turn to alternative (A2). In particular, for all $j \in \mathbb{N}$ with $Q \cap \frac{3}{4}Q_j^\alpha \neq \emptyset$, there holds $r_j \leq 16r$ and $|Q_j^\alpha| \leq 8^{m+2}|Q_j^\alpha \cap Q|$. Moreover, $137Q \cap (\mathbb{R}^{m+1} \setminus \mathcal{O}_\lambda^\alpha) \neq \emptyset$. Using $w_\lambda^\alpha = w - \sum_j \rho_j(w - w_j^\alpha)$ we estimate

$$\begin{aligned} \mathcal{M}_Q(\nabla w_\lambda^\alpha) &\leq \mathcal{M}_Q(\nabla w) + \sum_{j: Q \cap \frac{3}{4}Q_j^\alpha \neq \emptyset} \mathcal{M}_Q(\nabla(\rho_j(w - w_j^\alpha))). \\ &\leq \mathcal{M}_Q(\nabla w) + c \sum_{j: Q \cap \frac{3}{4}Q_j^\alpha \neq \emptyset} \frac{|Q_j^\alpha|}{|Q|} \mathcal{M}_{\frac{3}{4}Q_j^\alpha}(\nabla(\rho_j(w - w_j^\alpha))). \\ &\leq \mathcal{M}_Q(\nabla w) + c \sum_{j: Q \cap \frac{3}{4}Q_j^\alpha \neq \emptyset} \frac{|Q_j^\alpha \cap Q|}{|Q|} \mathcal{M}_{\frac{3}{4}Q_j^\alpha}(\nabla(\rho_j(w - w_j^\alpha))). \end{aligned}$$

Due to Lemma 2.12 there holds

$$\mathcal{M}_{\frac{3}{4}Q_j^\alpha}(\nabla(\rho_j(w - w_j^\alpha))) \leq c \int_{\frac{3}{4}Q_j^\alpha} \frac{|w - w_j^\alpha|}{r_j} dz + c \int_{\frac{3}{4}Q_j^\alpha} |\nabla w| dz \leq c\lambda.$$

On the other hand

$$\mathcal{M}_Q(\nabla w) \leq \mathcal{M}_{137Q}(\nabla w) \leq c\lambda,$$

since $137Q \cap (\mathbb{R}^{m+1} \setminus \mathcal{O}_\lambda^\alpha) \neq \emptyset$. We summarize the above estimate to get

$$\mathcal{M}_Q(\nabla w_\lambda^\alpha) \leq c\lambda + c\lambda \sum_{j: Q \cap \frac{3}{4}Q_j^\alpha \neq \emptyset} \frac{|Q_j^\alpha \cap Q|}{|Q|} \leq c\lambda,$$

where we used that the Q_j^α are locally finite, see (W5). □

Lemma 3.4 *There holds*

$$\alpha \mathcal{N}^\alpha(\partial_t w_\lambda^\alpha) \leq c\lambda.$$

Proof Let Q be an α -parabolic cylinder with radius R and $Q \subset \mathbb{R} \times \Omega$. We have to show that $\alpha \mathcal{N}_Q(\partial_t w_\lambda^\alpha) \leq c\lambda$. If $137Q \not\subset \mathbb{R} \times \Omega$, then the claim follows from Lemma 2.10 and Lemma 3.3, so we can assume in the following $137Q \subset \mathbb{R} \times \Omega$. We use the alternatives of Lemma 3.2.

We begin with alternative (A1). In particular, there exists $k \in \mathbb{N}$ such that $Q \cap \frac{1}{2}Q_k \neq \emptyset$, $8R \leq r_k$ and $Q \subset \frac{3}{4}Q_k$.

Then $w_\lambda^\alpha = \sum_{j \in A_k} \rho_j w_j^\alpha$ on Q_k and therefore

$$\begin{aligned} \alpha \mathcal{N}_Q(\partial_t w_\lambda^\alpha) &= \alpha \mathcal{N}_Q(\partial_t (w_\lambda^\alpha - w_k^\alpha)) = \alpha \mathcal{N}_Q\left(\partial_t \left(\sum_{j \in A_k} \rho_j (w_j^\alpha - w_k^\alpha)\right)\right) \\ &= \sum_{j \in A_k} \alpha \mathcal{N}_Q(\partial_t (\rho_j (w_j^\alpha - w_k^\alpha))). \end{aligned}$$

We estimate

$$\begin{aligned} \alpha \mathcal{N}_Q(\partial_t (\rho_j (w_j^\alpha - w_k^\alpha))) &= \alpha \sup_{\xi \in \mathcal{F}_Q} \left(R |Q|^{-1} |\langle \partial_t (\rho_j (w_j^\alpha - w_k^\alpha)), \xi \rangle| \right) \\ &\leq \alpha \sup_{\xi \in \mathcal{F}_Q} \left(R \|\partial_t \rho_j\|_\infty |w_j^\alpha - w_k^\alpha| \|\xi\|_\infty \right) \\ &\leq c \alpha R |w_j^\alpha - w_k^\alpha| (\alpha r_j^2)^{-1} \\ &\leq c \frac{|w_j^\alpha - w_k^\alpha|}{r_j}, \end{aligned}$$

where we used $8R \leq r_k \leq 2r_j$ in the last step. This and Lemma 3.1 imply $\alpha \mathcal{N}_Q(\partial_t w_\lambda^\alpha) \leq c \lambda$.

We turn to alternative (A2). In particular, for all $j \in \mathbb{N}$ with $Q \cap \frac{3}{4}Q_j \neq \emptyset$, there holds $r_j \leq 16R$ and $|Q_j| \leq 8^{d+2}|Q_j \cap Q|$. Moreover, $137Q \cap (\mathbb{R}^{d+1} \setminus \mathcal{O}_\lambda^\alpha) \neq \emptyset$. Using $w_\lambda^\alpha = w - \sum_j \rho_j (w - w_j^\alpha)$ we estimate

$$\mathcal{N}_Q(\partial_t w_\lambda^\alpha) \leq \mathcal{N}_Q(\partial_t w) + \sum_{j: Q \cap \frac{3}{4}Q_j \neq \emptyset} \mathcal{N}_Q(\partial_t (\rho_j (w - w_j^\alpha))).$$

Recall that $137Q \subset \mathbb{R} \times \Omega$. So $137Q \cap (\mathbb{R}^{d+1} \setminus \mathcal{O}_\lambda^\alpha) \neq \emptyset$ implies $\alpha \mathcal{N}_Q(\partial_t w) \leq c \alpha \mathcal{M}_Q(G) \leq c \lambda$ using also Remark 2.6. On the other hand using $r_j \leq 16R$, Lemma 3.1 and $|Q_j| \leq 8^{d+2}|Q_j \cap Q|$ we estimate

$$\begin{aligned} \mathcal{N}_Q(\partial_t w_\lambda^\alpha) &\leq c \lambda + \sum_{j: Q \cap \frac{3}{4}Q_j \neq \emptyset} \alpha \mathcal{N}_Q(\partial_t (\rho_j (w - w_j^\alpha))) \\ &= c \lambda + \alpha \sum_{j: Q \cap \frac{3}{4}Q_j \neq \emptyset} \sup_{\xi \in \mathcal{F}_Q} \left(R |Q|^{-1} |\langle \rho_j (w - w_j^\alpha), \partial_t \xi \rangle| \right). \end{aligned}$$

Now for j with $Q \cap \frac{3}{4}Q_j \neq \emptyset$ and $\xi_j := \langle \xi \rangle_{Q_j}$ we have

$$\begin{aligned} \frac{\alpha R}{|Q|} |\langle \rho_j (w - w_j^\alpha), \partial_t \xi \rangle| &\leq \frac{\alpha R}{|Q|} |\langle w - w_j^\alpha, \partial_t (\rho_j (\xi - \xi_j)) \rangle| \\ &\quad + \frac{\alpha R}{|Q|} |\langle w - w_j^\alpha, (\partial_t \rho_j) (\xi - \xi_j) \rangle| \\ &=: I + II. \end{aligned}$$

We will now estimate $\|\rho_j(\xi - \xi_j)\|_{\mathcal{F}_{Q_j}}$. Using $\|\xi\|_\infty + R\|\nabla\xi\|_\infty + \alpha R^2\|\partial_t\xi\|_\infty \leq 1$, we get by parabolic Poincaré’s inequality

$$\begin{aligned} \|\xi - \xi_j\|_{L^\infty(Q_j)} &\leq c r_i \|\nabla\xi\|_{L^\infty(Q_j)} + c \alpha r_i^2 \|\partial_t\xi\|_{L^\infty(Q_j)} \leq c \frac{r_i}{R} + c \frac{r_i^2}{R^2} \leq c \frac{r_i}{R}, \\ \|\rho_j(\xi - \xi_j)\|_\infty &\leq \|\xi - \xi_j\|_{L^\infty(Q_j)} \leq c \frac{r_i}{R}, \\ \|\nabla(\rho_j(\xi - \xi_j))\|_\infty &\leq \|\xi - \xi_j\|_{L^\infty(Q_j)} + r_i \|\nabla\xi\|_{L^\infty(Q_j)} \leq c \frac{r_i}{R}, \\ \|\partial_t(\rho_j(\xi - \xi_j))\|_\infty &\leq \|\xi - \xi_j\|_{L^\infty(Q_j)} + \alpha r_i^2 \|\partial_t\xi\|_{L^\infty(Q_j)} \leq c \frac{r_i}{R} + c \frac{r_i^2}{R^2} \leq c \frac{r_i}{R}. \end{aligned}$$

In particular, $\|\rho_j(\xi - \xi_j)\|_{\mathcal{F}_{Q_j}} \leq c \frac{r_i}{R}$. This and Lemma 2.12 imply

$$\begin{aligned} I &= \frac{\alpha R}{|Q|} |\langle w - w_j^\alpha, \partial_t(\rho_j(\xi - \xi_j)) \rangle| \\ &= \frac{\alpha R}{|Q|} |\langle w, \partial_t(\rho_j(\xi - \xi_j)) \rangle| \\ &\leq c \frac{\alpha R}{|Q|} \frac{|Q_j|}{r_i} \mathcal{N}_{Q_j}(\partial_t w) \|\rho_j(\xi - \xi_j)\|_{\mathcal{F}_{Q_j}} \\ &\leq c \frac{\alpha R}{|Q|} \frac{|Q_j|}{r_i} \frac{\lambda r_i}{\alpha R} \\ &= c \lambda \frac{|Q_j|}{|Q|}. \end{aligned}$$

Moreover, also by Lemma 2.12

$$\begin{aligned} II &= \frac{\alpha R}{|Q|} |\langle w - w_j^\alpha, (\partial_t \rho_j)(\xi - \xi_j) \rangle| \\ &\leq \frac{\alpha R}{|Q|} |Q_j| \int_{Q_j} |w - w_j^\alpha| dz \frac{c}{\alpha r_j^2} \|\xi - \xi_j\|_{L^\infty(Q_j)} \\ &\leq \frac{\alpha R}{|Q|} |Q_j| r_j \lambda \frac{c}{\alpha r_j^2} \frac{r_j}{R} \\ &= c \frac{|Q_j|}{|Q|} \lambda. \end{aligned}$$

Summarized we have

$$\alpha \mathcal{N}_Q(\partial_t w_\lambda^\alpha) \leq c \lambda + \sum_{j: Q \cap \frac{3}{4} Q_j \neq \emptyset} c \frac{|Q_j|}{|Q|} \lambda \leq c \lambda + \sum_{j: Q \cap \frac{3}{4} Q_j \neq \emptyset} c \frac{|Q_j \cap Q|}{|Q|} \lambda \leq c \lambda,$$

This proves the claim. □

Lemma 3.5 *There holds*

$$\mathcal{M}^{\alpha, \sharp, 1}(w_\lambda^\alpha) \leq c \lambda.$$

Proof Due to Theorem 2.8, Lemma 3.3 and Lemma 3.4 we have

$$M_Q^{\sharp, 1}(w_\lambda^\alpha) \leq c M_Q(\nabla w_\lambda^\alpha) + c \alpha \mathcal{N}_Q(\partial_t w_\lambda^\alpha) \leq c \lambda.$$

for every α -parabolic cylinder Q . □

Corollary 3.6 w_λ^α is Lipschitz continuous with respect to d^α , i.e.

$$|w_\lambda^\alpha(t, x) - w_\lambda^\alpha(s, y)| \leq c \lambda \max \left\{ \frac{|t - s|^{\frac{1}{2}}}{\alpha^{\frac{1}{2}}}, |x - y| \right\}$$

Proof It follows from $\mathcal{M}^{\alpha, \sharp, 1}(w_\lambda^\alpha) \leq c \lambda$ and [12] that w_λ^α is Lipschitz continuous with respect to d^α . \square

Lemma 3.7 Let $J = (t^-, t^+)$. For all $\eta \in W_0^{1, \infty}(-\infty, t^+)$ the expression $\langle \partial_t w, w_\lambda^\alpha \eta \rangle$ is well defined and can be calculated as

$$\langle \partial_t w, w_\lambda^\alpha \eta \rangle = \frac{1}{2} \int_Q (|w_\lambda^\alpha|^2 - 2w \cdot w_\lambda^\alpha) \partial_t \eta \, dz + \int_{\mathcal{O}_\lambda^\alpha} (\partial_t w_\lambda^\alpha)(w_\lambda^\alpha - w) \eta \, dz. \tag{3.1}$$

Proof Let $0 < h < T$. For a function f defined in space and time denote the Steklov average of f by

$$f_h(x, t) := \frac{1}{h} \int_t^{t+h} f(x, s) \, ds.$$

Then we have $\partial_t f_h(x, t) = h^{-1}(f(x, t + h) - f(x, t))$. We calculate

$$\begin{aligned} (I)_h &:= \langle \partial_t w, ((w_\lambda^\alpha)_h \eta)_{-h} \rangle = - \int_Q w_h \cdot \partial_t ((w_\lambda^\alpha)_h \eta) \, dz \\ &= \int_Q (w_\lambda^\alpha - w)_h \cdot \partial_t ((w_\lambda^\alpha)_h \eta) \, dz - \int_Q (w_\lambda^\alpha)_h \cdot \partial_t ((w_\lambda^\alpha)_h \eta) \, dz \\ &= \int_Q (w_\lambda^\alpha - w)_h \cdot (\partial_t (w_\lambda^\alpha)_h) \eta \, dz + \int_Q (w_\lambda^\alpha - w)_h \cdot (w_\lambda^\alpha)_h \partial_t \eta \, dz \\ &\quad - \int_Q \frac{1}{2} |(w_\lambda^\alpha)_h|^2 \partial_t \eta \, dz = \int_Q (w_\lambda^\alpha - w)_h \cdot (\partial_t (w_\lambda^\alpha)_h) \eta \, dz + \frac{1}{2} \int_Q (|(w_\lambda^\alpha)_h|^2 \\ &\quad - 2w_h \cdot (w_\lambda^\alpha)_h) \partial_t \eta \, dz =: (II)_h + (III)_h. \end{aligned}$$

All of these expressions are well defined. It has been shown in [17] formula (3.33) that

$$\begin{aligned} (II)_h &\rightarrow \int_Q (w_\lambda^\alpha - w) (\partial_t w_\lambda^\alpha) \eta \, dz, \\ (III)_h &\rightarrow \frac{1}{2} \int_Q (|w_\lambda^\alpha|^2 - 2w \cdot w_\lambda^\alpha) \partial_t \eta \, dz. \end{aligned}$$

for $h \rightarrow 0$. Let us point out that $w_\lambda^\alpha - w$ is only non-zero on $\mathcal{O}_\lambda^\alpha$. On this set w_λ^α is locally C^∞ , so $\partial_t w_\lambda^\alpha$ is a classical time derivative on this set. This shows that the limit $(I)_h$ is also well defined and can be calculated by 3.1. \square

This was the last piece to get Theorem 2.3.

Proof of Theorem 2.3 The definition of Lipschitz truncation w_λ^α is given in (2.15) and property (a) follows by the definition. Property (b) is proven in Lemma 3.3, property (c) is proven in Lemma 2.15, property (d) is proven in Lemma 3.4, property (e) is proven in Corollary 3.6, property (f) follows by Lemma 3.7. \square

4 The ϕ -caloric approximation

In this Section we will concentrate to prove the ϕ -caloric approximation result i.e. Theorem 1.2 in the general case of φ -growth.

Let us start defining $\mathbf{A}, \mathbf{V} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ in the following way:

$$\mathbf{A}(\mathbf{Q}) = \phi'(|\mathbf{Q}|) \frac{\mathbf{Q}}{|\mathbf{Q}|}, \tag{4.1a}$$

$$\mathbf{V}(\mathbf{Q}) = \psi'(|\mathbf{Q}|) \frac{\mathbf{Q}}{|\mathbf{Q}|}. \tag{4.1b}$$

Another important set of tools are the shifted N-functions $\{\phi_a\}_{a \geq 0}$. We define for $t \geq 0$

$$\phi_a(t) := \int_0^t \phi'_a(s) ds \quad \text{with} \quad \phi'_a(t) := \phi'(a+t) \frac{t}{a+t}. \tag{4.2}$$

Note that $\phi_a(t) \sim \phi'_a(t)t$. The families $\{\phi_a\}_{a \geq 0}$ and $\{(\phi_a)^*\}_{a \geq 0}$ satisfy the Δ_2 -condition uniformly in $a \geq 0$. The connection between \mathbf{A}, \mathbf{V} (see [19]) is the following:

$$(\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})) \cdot (\mathbf{P} - \mathbf{Q}) \sim |\mathbf{V}(\mathbf{P}) - \mathbf{V}(\mathbf{Q})|^2 \sim \phi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|),$$

uniformly in $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{m \times n}$. Moreover,

$$\mathbf{A}(\mathbf{Q}) \cdot \mathbf{Q} \sim |\mathbf{V}(\mathbf{Q})|^2 \sim \phi(|\mathbf{Q}|),$$

uniformly in $\mathbf{Q} \in \mathbb{R}^{m \times n}$.

Now we begin to prove some Lemmas regarding the level sets of the maximal function. Let $w \in L^\phi(J, W_0^{1,\phi}(\Omega))$ and $G \in L^{\phi^*}(J \times \Omega)$ such that

$$\begin{cases} \partial_t w = \text{div} G, & \text{on } [-t_0, 0) \times \Omega \\ w(-t_0, \cdot) \equiv 0. \end{cases} \tag{4.3}$$

We define for $Q = [-t_0, 0) \times \Omega$

$$\phi(\gamma) := \int_Q \phi(|\nabla w|) dz + \int_Q \phi^*(|G|) dz. \tag{4.4}$$

We then have the following lemma.

Lemma 4.1 *For every $m_0 \in \mathbb{N}$ there exists a $\lambda \in [\gamma, 2^{m_0} \gamma]$, such that for $\alpha = \alpha(\lambda) := \frac{\lambda}{\phi'(\lambda)}$*

$$|\{\mathcal{M}^\alpha(\nabla w \chi_Q) > \lambda\}| + |\{\mathcal{M}^\alpha(G \chi_Q) > \phi'(\lambda)\}| \leq c \frac{\phi(\gamma)}{m_0 \phi(\lambda)} |Q|$$

with c independent of m_0 and γ .

Proof We will use the following maximal operator

$$\mathcal{M}^*(f)(z) := \sup_{\{I \times B \subset \mathbb{R}^{m+1} : z \in I \times B\}} \int_I \int_B f \, dx \, dt.$$

Certainly we have that $\mathcal{M}^\alpha(f)(x) \leq \mathcal{M}^*(f)(x)$, for almost all $x \in \mathbb{R}^{m+1}$. Therefore,

$$\mathcal{O}_\lambda^\alpha := \{\mathcal{M}^\alpha(\nabla w) > \lambda\} \cup \{\alpha \mathcal{M}^\alpha(G) > \lambda\} \subset \{\mathcal{M}^*(\nabla w) > \lambda\} \cup \{\mathcal{M}^*(G) > \frac{\lambda}{\alpha}\}.$$

Now we have by the continuity of \mathcal{M}^* and since $(\phi')^{-1} \sim (\phi^*)'$, that for $m_0 \in \mathbb{N}$ and $\alpha(t) := \frac{t}{\phi'(t)}$,

$$\begin{aligned} & m_0 \min_{m \in \{0, \dots, m_0\}} \phi(2^m \gamma) \left(|\{\mathcal{M}^\alpha(\nabla w \chi_Q) > 2^m \gamma\}| + |\{\mathcal{M}^\alpha(G \chi_Q) > \phi'(2^m \gamma)\}| \right) \\ & \leq \sum_{m=0}^{m_0} \left(\phi(2^m \gamma) (|\{\mathcal{M}^\alpha(\nabla w \chi_Q) > 2^m \gamma\}| + |\{\mathcal{M}^\alpha(G \chi_Q) > \phi'(2^m \gamma)\}|) \right) \\ & \leq \sum_{m=0}^{m_0} \left(\phi(2^m \gamma) (|\{\mathcal{M}^*(\nabla w \chi_Q) > 2^m \gamma\}| + |\{\mathcal{M}^*(G \chi_Q) > \phi'(2^m \gamma)\}|) \right) \\ & \leq \int \phi(\mathcal{M}^*(\nabla w \chi_Q)) + \phi((\phi')^{-1}(\mathcal{M}^*(G \chi_Q))) \, dz \\ & \leq c \int_Q \phi(|\nabla w|) + \phi^*(|G|) \leq c \phi(\gamma) |Q|. \end{aligned}$$

This concludes the proof. □

Let $u \in L^\phi(J, W_0^{1,\phi}(\Omega))$ be solution of

$$\partial_t u = \operatorname{div} H$$

on $Q = I \times B = (t^-, t^+) \times B$ with $H \in L^{\phi^*}(J \times \Omega)$ and h be the weak solution of

$$\partial_t h - \operatorname{div}(A(\nabla h)) = 0 \text{ in } Q$$

with $h = u$ on $\partial_p Q$. The function h is called the ϕ -caloric comparison function of u in Q . Define $w := u - h$. Then

$$\begin{aligned} \partial_t w - \operatorname{div}(A(\nabla u) - A(\nabla h)) &= \partial_t u - \operatorname{div}(A(\nabla u)) \\ &= \operatorname{div}(H - A(\nabla u)) = \operatorname{div}(G) \end{aligned}$$

and $w = 0$ on $\partial_p Q$, where $G = H - A(\nabla u)$.

Since w is a valid testfunction, we find by the standard methods, that

$$\sup_{t \in I} \int_B \frac{|w|^2}{t^+ - t^-} \, dx + \int_Q |V(\nabla u) - V(\nabla h)|^2 \, dz \leq c_0 \int_Q \phi(|\nabla u|) + \phi^*(|G|) \, dz, \quad (4.5)$$

where c_0 is a fixed constant only depending on the characteristics of ϕ .

Now we are in a position to prove the ϕ -caloric approximation Theorem.

Theorem 4.2 *Let $\sigma \in (0, 1)$, $q \in [1, \infty)$ and $\theta \in (0, 1)$ fixed. Moreover, let $\tilde{Q} = Q$ or to be more flexible let \tilde{Q} be such that $Q \subset \tilde{Q} \subset 2Q$. Then for $\epsilon > 0$ there exists $\delta > 0$ such that the following holds: if u is ‘‘almost ϕ -caloric’’ in the sense that for all $\xi \in C_0^\infty(Q)$,*

$$\left| \int_Q -u \partial_t \xi + A(\nabla u) \nabla \xi \, dz \right| \leq \delta \left(\int_Q \phi(|\nabla u|) \, dz + \int_Q \phi^*(|H|) \, dz + \phi(\|\nabla \xi\|_\infty) \right), \quad (4.6)$$

then

$$\left(\int_I \left(\int_B \left(\frac{|u-h|^2}{t^+ - t^-} \right)^\sigma dx \right)^{\frac{q}{\sigma}} dt \right)^{\frac{1}{q}} + \left| \int_Q |V(\nabla u) - V(\nabla h)|^{2\theta} dz \right|^{\frac{1}{\theta}} \leq \epsilon \left(\int_{\tilde{Q}} \phi(|\nabla u|) dz + \int_{\tilde{Q}} \phi^*(|H|) dz \right).$$

Proof Let $w := u - h$ and $G = H - A(\nabla u)$. Then

$$\partial_t w = \operatorname{div} G \quad \text{on } Q$$

and $w = 0$ on $\partial_p Q$. We define

$$\phi(\gamma) := \int_{\tilde{Q}} \phi(|\nabla u|) dz + \int_{\tilde{Q}} \phi^*(|H|) dz. \tag{4.7}$$

By Lemma 4.1 and (4.5) we find for every $m_0 \in \mathbb{N}$ a $\lambda \in [\gamma, 2^{m_0}\gamma]$, such that for $\alpha = \alpha(\lambda) := \frac{\lambda}{\phi'(\lambda)}$

$$|\{\mathcal{M}^\alpha(\nabla w \chi_Q) > \lambda\}| + |\{\alpha \mathcal{M}^\alpha(G \chi_Q) > \lambda\}| \leq \frac{c\phi(\gamma)}{\phi(\lambda)m_0} |Q|. \tag{4.8}$$

with c independent of m_0, γ and λ .

Now, let w_λ^α be the Lipschitz truncation of w as in Sect. 2, i.e.

$$\mathcal{O}_\lambda^\alpha := (\{\mathcal{M}^\alpha(\nabla w \chi_Q) > \lambda\} \cup \{\alpha \mathcal{M}^\alpha(G \chi_Q) > \lambda\}) \text{ and } \operatorname{supp}(w_\lambda^\alpha) \subset \overline{\mathcal{O}_\lambda^\alpha \cap Q}.$$

We use the test function $\xi = w_\lambda^\alpha \eta$, where $\eta = \max\{\frac{t^+ - t}{t^+ - t^-}, 0\} \in [0, 1]$ on $I = [t^-, t^+]$. Note that in general $\xi \notin C_0^\infty(Q)$. However, it follows by a simple convolution argument as in (4.1) of [19], that the validity of (4.6) for all $\xi \in C_0^\infty(Q)$ implies its validity under the assumption $\|\nabla \xi\|_\infty < \infty$. Thus, ξ is a valid test function.

Therefore, using the Theorem 2.3 (f) we find

$$\begin{aligned} (I_1) + (I) + (II) &:= \int_Q \frac{|w_\lambda^\alpha|^2}{2} (-\partial_t \eta) dz - \int_Q (w - w_\lambda^\alpha) \partial_t ((w_\lambda^\alpha) \eta) dz \\ &\quad + \int_Q \langle (A(\nabla u) - A(\nabla h)), (\nabla w_\lambda^\alpha) \eta \rangle dz \\ &\leq \delta \left(\int_{\tilde{Q}} \phi(|\nabla u|) dz ds + \int_{\tilde{Q}} \phi^*(|H|) dz + \phi(\|\eta \nabla w_\lambda^\alpha\|_\infty) \right) \\ &\leq \delta \left(\int_{\tilde{Q}} \phi(|\nabla u|) dz + \int_{\tilde{Q}} \phi^*(|H|) dz + c_{m_0} \phi(\gamma) \right) =: (III). \end{aligned}$$

using $\|\nabla w_\lambda^\alpha\|_\infty \leq c\lambda \leq c2^{m_0}\gamma$. As $-\partial_t \eta = \frac{1}{(t^+ - t^-)} \geq 0$, we have that $(I_1) > 0$. We estimate the other terms.

$$\begin{aligned} (I) &= -\int_Q (w - w_\lambda^\alpha) \partial_t w_\lambda^\alpha \eta \, dz - \int_Q (w - w_\lambda^\alpha) w_\lambda^\alpha \partial_t \eta \, dz \\ &= -\int_Q (w - w_\lambda^\alpha) \partial_t w_\lambda^\alpha \eta \, dz - \frac{1}{|Q|} \sum_i \int_{Q_i} (w - w_i^\alpha) \rho_i w_\lambda^\alpha \partial_t \eta \, dz \\ &= -\int_Q (w - w_\lambda^\alpha) \partial_t w_\lambda^\alpha \eta \, dz - \frac{1}{|Q|} \sum_i \int_{Q_i} (w - w_i^\alpha) \rho_i \sum_{j \in A_i} \rho_j w_j^\alpha \partial_t \eta \, dz \\ &= -\int_Q (w - w_\lambda^\alpha) \partial_t w_\lambda^\alpha \eta \, dz - \frac{1}{|Q|} \sum_{\{i: \exists j \in A_i: w_j^\alpha \neq 0\}} \int_{Q_i} (w - w_i^\alpha) \rho_i \sum_{j \in A_i} \rho_j w_j^\alpha \partial_t \eta \, dz \\ &= -(I_2) - (I_3). \end{aligned}$$

Using the fact, that $\text{supp}(\rho_j) \subset \frac{3}{4}Q_j^\alpha$, we estimate

$$\begin{aligned} (I_2) &\leq \int_Q \chi_{\mathcal{O}_\lambda^\alpha} |w - w_\lambda^\alpha| |\partial_t w_\lambda^\alpha \eta| \, dz \\ &= \frac{1}{|Q|} \int \chi_{Q \cap \mathcal{O}_\lambda^\alpha} \sum_i |\rho_i (w - w_i^\alpha)| \left| \sum_{j \in A_i} \partial_t \rho_i w_j^\alpha \right| \, dz \\ &= \frac{1}{|Q|} \sum_i \int_{\frac{3}{4}Q_i} \chi_{Q \cap \mathcal{O}_\lambda^\alpha} |\rho_i (w - w_i^\alpha)| \left| \sum_{j \in A_i} \partial_t \rho_i (w_i^\alpha - w_j^\alpha) \right| \, dz \end{aligned}$$

We estimate further using (P2), (P1), (W6), (W9), (W2) Lemma 3.1, Lemma 2.12 the fact that $\mathcal{O}_\lambda^\alpha$ is symmetric around t^+ and (4.8).

$$\begin{aligned} (I_2) &\leq \frac{c}{\alpha|Q|} \sum_i \sum_{j \in A_i} \int_{\frac{3}{4}Q_i} \frac{|w - w_i^\alpha|}{r_i} \frac{|w_j^\alpha - w_i^\alpha|}{r_j} \\ &\leq \frac{c}{\alpha|Q|} \sum_i \sum_{j \in A_i} |Q_i| \int_{\frac{3}{4}Q_i} \frac{|w - w_i^\alpha|}{r_i} \\ &\leq \frac{c\lambda^2}{\alpha} \frac{|\mathcal{O}_\lambda^\alpha|}{|Q|} \leq \frac{c\phi(\gamma)}{m_0}. \end{aligned}$$

To estimate (I_3) we making use of the fact that either $w_i^\alpha = 0$ or $\text{supp}(\rho_i) \subset \frac{3}{4}Q_i^\alpha \subset Q$ and then $\int_{Q_i} (w - w_i^\alpha) \rho_i \partial_t \eta \, dz = 0$. Since $\sum_i \rho_i = 1$ we find

$$(I_3) = \frac{1}{|Q|} \sum_{\{i: \exists j \in A_i: w_j^\alpha \neq 0\}} \int_{Q_i} (w - w_i^\alpha) \rho_i \sum_{j \in A_i} \rho_j (w_j^\alpha - w_i^\alpha) \partial_t \eta \, dz.$$

Next, observe that there exists $j \in A_i$, such that $w_j^\alpha \neq 0$ and hence $\frac{3}{4}Q_j \subset J \times B$. This however implies that $r_j^2 \alpha \leq 2(t^+ - t^-)$ and consequently by (W8) that $r_i^2 \alpha \leq c(t^+ - t^-)$.

Using this bound together with the argument that was used to estimate (I_2) implies

$$\begin{aligned} |(I_3)| &\leq \frac{c}{t^+ - t^-} \frac{1}{|Q|} \sum_{\{i: \exists j \in A_i: w_j^\alpha \neq 0\}} \int_{Q_i} |w - w_i^\alpha| \, dz \sum_{j \in A_i} |w_i^\alpha - w_j^\alpha| \\ &\leq \frac{c}{\alpha |Q|} \sum_i \int_{Q_i} \frac{|w - w_i^\alpha|}{r_i} \, dz \sum_{j \in A_i} \frac{|w_i^\alpha - w_j^\alpha|}{r_i} \\ &\leq \frac{c\lambda^2}{\alpha} \frac{|\mathcal{O}_\lambda^\alpha|}{|Q|} \leq \frac{c\phi(\gamma)}{m_0}. \end{aligned}$$

Now we continue by estimating (II) . Recall that $|\nabla w_\lambda^\alpha| \leq c\lambda$ and that $w_\lambda^\alpha = w = u - h$ on $Q \setminus \mathcal{O}_\lambda^\alpha$. This gives

$$\begin{aligned} (II) &= \int_Q \langle A(\nabla u) - A(\nabla h), \nabla w_\lambda^\alpha \eta \rangle \, dz \\ &\geq c \int_Q \chi_{Q \setminus \mathcal{O}_\lambda^\alpha} |V(\nabla u) - V(\nabla h)|^2 \eta \, dz - c \int_Q \chi_{\mathcal{O}_\lambda^\alpha} (|A(\nabla u)| + |A(\nabla h)|) \lambda \, dz \\ &=: (II_1) - (II_2). \end{aligned}$$

Using Young’s inequality with $\tilde{\delta}$, that can be chosen independent of m_0, γ, λ and (4.8), we find that

$$\begin{aligned} (II_2) &= \int_Q \chi_{\mathcal{O}_\lambda^\alpha} (|A(\nabla u)| + |A(\nabla h)|) \lambda \, dz \\ &\leq c_{\tilde{\delta}} \phi(\lambda) \frac{|\mathcal{Q} \cap \mathcal{O}_\lambda^\alpha|}{|Q|} + \tilde{\delta} \int_Q \chi_{\mathcal{O}_\lambda^\alpha} \phi(|\nabla u|) \, dz \leq c \left(\frac{c_{\tilde{\delta}}}{m} + \tilde{\delta} \right) \phi(\gamma). \end{aligned}$$

So far we have

$$(III) = (I) + (II) \geq (I_1) - (I_2) - (I_3) + (II_1) - (II_2)$$

which implies by that

$$\begin{aligned} (II_1) + (I_1) &\leq (II_2) + |(I_2) + (I_3)| + (III) \\ &\leq \left(c_{m_0} \delta + 3\tilde{\delta} + \frac{c_{\tilde{\delta}}}{m_0} \right) \phi(\gamma) \end{aligned} \tag{4.9}$$

Observe, that for $\beta \in (0, 1)$ we find

$$\begin{aligned} \left(\int_{t^-}^{t^+} \eta^{-\beta} \, dt \right)^{\frac{1}{\beta}} &= \left(- \int_{t^-}^0 \frac{(t^+ - t^-)^\beta}{s^\beta} \, ds \right)^{\frac{1}{\beta}} = (t^+ - t^-)^{\frac{\beta-1}{\beta}} \left(\int_0^{t^+ - t^-} s^{-\beta} \, ds \right)^{\frac{1}{\beta}} \\ &= (1 - \beta)^{\frac{-1}{\beta}}. \end{aligned}$$

Now we fix $\theta \in (0, \frac{1}{2})$, such that $\beta = \frac{1}{1-\theta} \in (0, 1)$. For θ closer to 1, we will later use an interpolation with (4.5). For this fixed $\theta \in (0, \frac{1}{2})$ we get by the above that

$$\begin{aligned}
 (IV) &:= \left(\int_Q |V(\nabla u) - V(\nabla h)|^{2\theta} dz \right)^{\frac{1}{\theta}} \\
 &= \left(\int_Q \chi_{\mathcal{O}_\lambda^\alpha} |V(\nabla u) - V(\nabla h)|^{2\theta} dz + \int_Q \chi_{(\mathcal{O}_\lambda^\alpha)^c} |V(\nabla u) - V(\nabla h)|^{2\theta} dz \right)^{\frac{1}{\theta}}. \\
 &\leq c \int_Q |V(\nabla u) - V(\nabla h)|^2 dz \left(\frac{|Q \cap \mathcal{O}_\lambda^\alpha|}{|Q|} \right)^{\frac{1-\theta}{\theta}} \\
 &\quad + c \int_Q \chi_{(\mathcal{O}_\lambda^\alpha)^c} |V(\nabla u) - V(\nabla h)|^2 \eta dz \left(\int_Q \chi_{(\mathcal{O}_\lambda^\alpha)^c} \eta^{-\frac{\theta}{1-\theta}} dz \right)^{\frac{1-\theta}{\theta}} \\
 &\leq c \int_Q |V(\nabla u) - V(\nabla h)|^2 dz \left(\frac{|Q \cap \mathcal{O}_\lambda^\alpha|}{|Q|} \right)^{\frac{1-\theta}{\theta}} + c \left(\frac{1-\theta}{1-2\theta} \right)^{\frac{1-\theta}{\theta}} (II_1) \\
 &=: c(V) + c(II_1).
 \end{aligned}$$

Now, by Lemma 4.1 we get

$$\begin{aligned}
 (V) &\leq c \int_Q \phi(|\nabla u|) + \phi(|\nabla h|) dz \left(\frac{|Q \cap \mathcal{O}_\lambda^\alpha|}{|Q|} \right)^{\frac{1-\theta}{\theta}}. \\
 &\leq \left(\frac{c\phi(\gamma)}{\phi(2^{m_0}\gamma)m_0} \right)^{\frac{1-\theta}{\theta}} \int_Q \phi(|\nabla u|) + \phi^*(|H|) dz \tag{4.10} \\
 &\leq \frac{c\phi(\gamma)}{2^{\frac{m_0(1-\theta)}{\theta}}}
 \end{aligned}$$

For the estimate from below for (I_1) we estimate similarly that

$$\begin{aligned}
 (VI) &:= \int_I \left(\int_B \frac{|w|}{\sqrt{t^+ - t^-}} dx \right)^2 dt \\
 &\leq \int_Q \chi_{(\mathcal{O}_\lambda^\alpha)^c} \frac{|w_\lambda^\alpha|^2}{t^+ - t^-} dz + \int_I \frac{|(\{t\} \times B) \cap \mathcal{O}_\lambda^\alpha|}{|B|} \int_{\{t\} \times B} \frac{|w|^2}{t^+ - t^-} dx dt \\
 &\leq c(I_1) + \frac{|Q \cap \mathcal{O}_\lambda^\alpha|}{|Q|} \sup_I \int_B \frac{|w|^2}{t^+ - t^-} dx.
 \end{aligned}$$

Now we use (4.5), (4.7) and Lemma 4.1 to find that

$$(VI) \leq c(I_1) + c \frac{\phi(\gamma)}{m_0 \phi(2^{m_0}\gamma)} \phi(\gamma) \leq \frac{\phi(\gamma)}{2^{m_0}}$$

This implies together with Lemma 4.1, (4.9) and (4.10)

$$(VI) + (IV) \leq \left(c_{m_0} \delta + 3\tilde{\delta} + \frac{c\tilde{\delta}}{m_0} + \frac{c}{2^{\frac{m_0(1-\theta)}{\theta}}} + \frac{c}{2^{m_0}} \right) \phi(\gamma).$$

Let us fix the auxiliary constant $\tilde{\epsilon} \in (0, 1)$. It shall be fixed at the very end of the proof. In the following order we choose $\tilde{\delta}$, m_0 and δ . We choose $\tilde{\delta} = \frac{\tilde{\epsilon}}{5}$. Then we choose m_0 large enough, such that $\frac{c\tilde{\delta}}{m_0} + \frac{c}{2^{\frac{m_0(1-\theta)}{\theta}}} + \frac{c}{2^{m_0}} \leq \frac{\tilde{\epsilon}}{5}$. Finally we fix δ small enough such that $c_{m_0} \delta \leq \frac{\tilde{\epsilon}}{5}$. These

choices imply the following estimate for a fixed $\theta \in (0, \frac{1}{2})$ (for example $\theta = \frac{1}{4}$)

$$\int_I \left(\int_B \frac{|w|}{\sqrt{t^+ - t^-}} dx \right)^2 dt + \left(\int_Q |V(\nabla u) - V(\nabla h)|^{2\theta} dz \right)^{\frac{1}{\theta}} \leq \epsilon \phi(\gamma).$$

Finally, by interpolation between the estimate above and estimate (4.5), we find the result. For the sake of completion we include the interpolation between $L^2(L^1)$ and the $L^\infty(L^2)$ estimate. The interpolation between L^2 and $L^{2\theta}$ for the gradient terms is similar but more straight forward, such that we omit the details. Let us fix $f = \frac{|w|}{\sqrt{t^+ - t^-}}$. Then we find for $b \in (2, \infty)$ and $a \in (1, 2)$ by Hölder, Jensen’s inequality, (4.5) and (4.7) that

$$\begin{aligned} & \left(\int_I \left(\int_B \left(\frac{|w|}{\sqrt{t^+ - t^-}} \right)^a dx \right)^{\frac{b}{a}} dt \right)^{\frac{2}{b}} = \left(\int_I \left(\int_B |f|^{2-a} |f|^{2(a-1)} dx \right)^{\frac{b}{a}} dt \right)^{\frac{2}{b}} \\ & \leq \left(\int_I \left(\int_B |f| dx \right)^{\frac{b(2-a)}{a}} \left(\int_B |f|^2 dx \right)^{\frac{b(a-1)}{a}} dt \right)^{\frac{2}{b}} \\ & \leq \left(\int_I \left(\int_B |f|^2 dx \right)^{\frac{b(a-1)}{a} + \frac{b(2-a)}{2a} - 1} \left(\int_B |f| dx \right)^2 dt \right)^{\frac{2}{b}} \\ & \leq \sup_I \left(\int_B |f|^2 dx \right)^{\frac{b-2}{2} \frac{2}{b}} \left(\int_I \left(\int_B |f| dx \right)^2 dt \right)^{\frac{2}{b}} \leq c_0^{\frac{b-2}{b}} \tilde{\epsilon}^{\frac{2}{b}} \phi(\gamma). \end{aligned}$$

Choosing $a = 2\sigma$ and $b = 2q$ the proof is completed by an appropriate choice of $\tilde{\epsilon}$. □

Acknowledgements These results were announced for the first time at the Mittag-Leffler Institute for the special program “Evolutionary problems” in 2013. We would like to thank the institute for the hospitality. S. Schwarzacher wishes to thank program PRVOUK P47, financed by Charles University in Prague. B. Stroffolini and A. Verde have been partially supported by the Italian M.I.U.R. Project “Calcolo delle Variazioni ” (2012).

References

1. Acerbi, E., Fusco, N.: Semicontinuity problems in the calculus of variations. Arch. Rational Mech. Anal. **86**(2), 125–145 (1984)
2. Acerbi, E., Fusco, N.: An Approximation Lemma for $W^{1,p}$ Functions, Material Instabilities in Continuum Mechanics (Edinburgh, 1985–1986). Oxford science publications, Oxford University Press, New York (1988)
3. Auscher, Pascal: On L^p estimates for square roots of second order elliptic operators on \mathbb{R}^n . Publ. Mat. **48**(1), 159–186 (2004)
4. Breit, D.: The \mathcal{A} -Stokes approximation for non-stationary problems. Q. J. Math. **67**, 201–231 (2016)
5. Breit, D., Diening, L., Fuchs, M.: Solenoidal Lipschitz truncation and applications in fluid mechanics. J. Differ. Equ. **253**(6), 1910–1942 (2012)
6. Bögelein, V., Duzaar, F., Mingione, G.: The regularity of general parabolic systems with degenerate diffusion. Mem. Am. Math. Soc. **221**(1041), vi+143 (2013)
7. Breit, D., Diening, L., Schwarzacher, S.: Solenoidal lipschitz truncation for parabolic pdes. Math. Models Methods Appl. Sci. **53**(14), 2671–2700 (2013)
8. Bulíček, M., Burczak, J., Schwarzacher, S.: A unified theory for some non Newtonian fluids under singular forcing. Siam. J. Math. Anal. **48**, 4241–4267 (2016)
9. Bulíček, M., Diening, L., Schwarzacher, S.: Existence, uniqueness and optimal regularity results for very weak solutions to nonlinear elliptic systems. Anal. PDE **9**, 1115–1151 (2016)
10. Bulíček, B., Schwarzacher, S.: Existence of Very Weak Solutions to Elliptic Systems of p-Laplacian Type. Calc. of Var & PDE. **55**, 52 (2016). doi:10.1007/100526-016-0986-7
11. Ball, J.M., Zhang, K.-W.: Lower semicontinuity of multiple integrals and the biting lemma. Proc. R. Soc. Edinburgh Sect. A **114**(3–4), 367–379 (1990)

12. Da Prato, G.: Spazi $\mathcal{L}^{(p,\theta)}(\Omega, \delta)$ e loro proprietà. *Ann. Math. Pura Appl.* (4) **69**, 383–392 (1965)
13. Diening, L., Ettwein, F.: Fractional estimates for non-differentiable elliptic systems with general growth. *Forum Mathematicum* **20**(3), 523–556 (2008)
14. Diening, L., Kreuzer, Ch., Süli, E.: Finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *SIAM J. Numer. Anal.* **51**(2), 984–1015 (2013)
15. Diening, L., Lengeler, D., Strohfolini, B., Verde, A.: Partial regularity for minimizers of quasi-convex functionals with general growth. *SIAM J. Math. Anal.* **44**(5), 3594–3616 (2012)
16. Diening, L., Málek, J., Steinhauer, M.: On Lipschitz truncations of Sobolev functions (with variable exponent) and their selected applications. *ESAIM Control Optim. Calc. Var.* **14**(2), 211–232 (2008)
17. Diening, L., Ružička, M., Wolf, J.: Existence of weak solutions for unsteady motions of generalized Newtonian fluids. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **9**(1), 1–46 (2010)
18. Diening, L., Schwarzacher, S., Strohfolini, B., Verde, A.: Partial regularity results for Parabolic Systems with Critical Growth (**in preparation**)
19. Diening, L., Strohfolini, B., Verde, A.: The ϕ -harmonic approximation and the regularity of ϕ -harmonic maps. *J. Differ. Equ.* **253**(7), 1943–1958 (2012)
20. Duzaar, F., Mingione, G.: The p -harmonic approximation and the regularity of p -harmonic maps. *Calc. Var. Partial Differ. Equ.* **20**(3), 235–256 (2004)
21. Duzaar, F., Mingione, G.: Second order parabolic systems, optimal regularity, and singular sets of solutions. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **22**(6), 705–751 (2005)
22. Duzaar, F., Mingione, G.: Harmonic type approximation lemmas. *J. Math. Anal. Appl.* **352**(1), 301–335 (2009)
23. Duzaar, A., Mingione, G., Steffen, K.: Parabolic systems with polynomial growth and regularity. *Mem. Am. Math. Soc.* **214**(1005), x+118 (2011)
24. Frehse, J., Málek, J., Steinhauer, M.: On analysis of steady flows of fluids with shear-dependent viscosity based on the Lipschitz truncation method. *SIAM J. Math. Anal.* **34**(5), 1064–1083 (2003). (electronic)
25. Kinnunen, J., Lewis, J.L.: Higher integrability for parabolic systems of p -Laplacian type. *Duke Math. J.* **102**(2), 253–271 (2000)
26. Kinnunen, J., Lewis, J.L.: Very weak solutions of parabolic systems of p -Laplacian type. *Ark. Mat.* **40**(1), 105–132 (2002)
27. Zhang, K.: Biting theorems for Jacobians and their applications. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **7**(4), 345–365 (1990)