Ark. Mat., **54** (2016), 85–123
DOI: 10.1007/s11512-015-0221-3

Higher integrability for vector-valued parabolic quasi-minimizers on metric measure spaces

Jens Habermann

Abstract. We establish local higher integrability estimates for upper gradients of vector-valued parabolic quasi-minimizers in metric measure spaces, satisfying a doubling property and supporting a weak Poincaré inequality.

1. Introduction

We are concerned with regularity issues for parabolic quasi-minimizers on metric measure spaces. More precisely, we consider quasi-minimizers of integral functionals which are related in the Euclidean setting to vector-valued functions $u\colon (0,T)\times\Omega\to\mathbb{R}^N$, $N>1$, of the variational inequality

$$-p\iint \langle u,\partial_t\Phi\rangle\,dx\,dt+\iint |Du|^p\,dx\,dt\le \mathcal{Q}\iint |Du+D\Phi|^p\,dx\,dt, \tag{1.1}$$

for all testing-functions, or in the case of real minimizers ($\mathcal{Q}=1$), to solutions of the parabolic p-Laplace system

$$u_t-\operatorname{div}\big(|Du|^{p-2}Du\big)=0,\qquad \frac{2n}{n+2}<p<\infty. \tag{1.2}$$

Here $\Omega\subset\mathbb{R}^n$, $n\geq 2$ is a bounded domain.

Replacing the Euclidean space $\mathbb{R}^n$ by a metric measure space $(\mathcal{X},\mathrm{d},\mu)$ with metric d and measure μ, we can no longer speak of a gradient of a function u, but we have to introduce the notion of upper gradients, which we denote by g_u. The manuscript at hand deals with parabolic quasi-minimizers on parabolic cylinders $\Omega_T:=(0,T)\times\Omega$, with $T>0$, and $\Omega\subset\mathcal{X}$ open, bounded, and where $\mathcal{X}$ denotes a metric measure space with a doubling property of the measure and a suitable Poincaré inequality. We refer to Chapter 2 for the exact setting and definitions of the relevant spaces. Our aim is to establish integrability properties for upper gradients g_u of quasi-minimizers, and therefore to generalize results, which have recently by proven in the Euclidean setting in [16], to this framework.

In the past fifteen years, doubling metric measure spaces have been investigated quite intensively, see for example [9], [13], [17], [18], [20], [22], [32], [42] and [43] and in particular [2] for an overview and further references. First regularity results for elliptic minimizers and quasi-minimizers in the metric space setting were obtained in 2000, by Kinnunen and Shanmugalingam [29]. They studied scalar problems and adapted DeGiorgi's method to prove maximum principles for p-harmonic functions and quasi-minimizers and also local Hölder continuity. Later, Björn and Marola [3] showed that also Moser's iteration method is available in the metric setting. Further contributions on regularity for elliptic problems were done in [34] and [43].

The investigation of parabolic problems on metric measure spaces started very recently with a contribution of Kinnunen, Marola, Miranda and Paronetto [30]. The authors study scalar parabolic quasi-minimizers in the sense of Definition 2.3 in the case $p{=}2$ which corresponds to linear parabolic PDEs. Following the approaches of DiBenedetto, Gianazza, Vespri [10], [11] and [12] and Wieser [44] they introduce parabolic DeGiorgi classes of order 2 and prove a Harnack inequality for scalar quasi-minimizers. Their contribution generalizes a previous result by Grigor'yan [15] and Saloff-Coste [40] on Harnack inequalities for solutions to the heat equation on Riemannian manifolds. The advantage of the approach in [30] is that it is a purely variational technique which does not use any differential structure, such as Dirichlet spaces or the Cheeger differential. In [37], Masson and Siljander generalized these results to the super-quadratic case $p{\ge}2$ and a very recent contribution of Marola and Masson [35] is dedicated to Harnack estimates for parabolic minimizers.

Higher integrability properties for solutions of parabolic problems in the non-linear case were first proven by Kinnunen and Lewis [27], who studied the model case of the parabolic p-Laplace equation for $p{\neq}2$. Later, Bögelein [6] and Bögelein and Parviainen [8] generalized these results to higher-order parabolic systems, also up to the parabolic boundary. In the metric measure space setting, there have also been made some efforts very recently: Masson, Miranda, Paronetto and Parviainen [38] and Masson and Parviainen [36] investigated parabolic quasi-minimizers with quadratic growth $p{=}2$, and showed local and global higher integrability properties for upper gradients.

Our aim is to prove local $L^{p+\varepsilon}$-regularity for the spatial minimal weak upper gradient g_u for parabolic $\mathcal{Q}$-minimizers with general polynomial p-growth, $p{\neq}2$. Roughly speaking, we show that for a local parabolic quasi-minimizer of a functional

$$\mathcal{F}[w,\Omega_T] := \frac{1}{p}\iint_{\Omega_T} f(g_w)\, d\mu\, dt,$$

which satisfies a polynomial growth condition of the type

$$\nu\zeta^p - L_1 \le f(\zeta) \le L\zeta^p + L_1,$$

for all $\zeta\in[0,\infty)$, with constants $0<\nu\le L<\infty$ and $L_1\ge 0$, and with growth exponent $p>\frac{2\mathrm{n}}{\mathrm{n}+2}$, where n denotes the 'dimension' related to the doubling constant of the measure space, see (2.2) and (3.1), there holds the implication

$$u\in L^p_{\mathrm{loc}}\big(0,T;\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega)\big) \Longrightarrow u\in L^{p+\varepsilon}_{\mathrm{loc}}\big(0,T;\mathcal{N}^{1,p+\varepsilon}_{\mathrm{loc}}(\Omega)\big),$$

for some $\varepsilon>0$, depending only on the structural data of the problem. In this context, g_w denotes the minimal p-weak upper gradient of w and $\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega)$ denotes the local Newtonian space. We refer the reader to Chapter 2 for the exact definitions of the appearing objects, and in particular find the exact statement in Theorem 2.4. Higher integrability for upper gradients is an important first step towards regularity and stability discussions in the metric measure space setting.

For general p-growth functionals, higher integrability statements are much more difficult to obtain than for functionals with quadratic growth. This is essentially due to the non-linear structure of the problem. To come up with homogeneous estimates, which are essential for higher integrability issues, we need to deal with the intrinsic geometry, introduced by DiBenedetto and Friedman in the setting of the degenerate p-Laplace equation. It turns out that the techniques, applied in [7] and [16] to prove higher integrability in the framework of non-linear parabolic equations or parabolic quasi-minimizers on Euclidean spaces, respectively, are flexible enough to be applied also in the setting of metric measure spaces.

On the other hand, additional difficulties are caused by the non-linear behaviour of upper gradients. In contrast to a 'real' gradient, upper gradients are merely sub-linear, which leads to a series of difficulties in the estimates. In particular, smoothing procedures become much more involved and cannot be treated by standard arguments. Instead, a careful analysis of the approximation arguments and limit procedures has to be made.

2. Setting and statements of results

Let $(\mathcal{X},\mathrm{d},\mu)$ be a separable, connected metric measure space, which means that $(\mathcal{X},\mathrm{d})$ is a complete, separable and connected metric space and μ denotes a Borel measure on $\mathcal{X}$. The measure μ is assumed to fulfill a *doubling property*: There exists a constant $c\ge 1$ such that

$$0<\mu\big(B_{2r}(x)\big)\le c\cdot\mu\big(B_r(x)\big)<+\infty, \tag{2.1}$$

for all radii $r>0$ and all $x\in\mathcal{X}$. Here $B_r(x):=\{y\in\mathcal{X}\colon \mathrm{d}(y,x)<r\}$ denotes the open ball of radius r and center x with respect to the metric d. We define the doubling constant

$$c_d := \inf\big\{c\in(1,\infty)\colon \text{(2.1) holds}\big\}. \tag{2.2}$$

We follow the concept of Cheeger [9], Heinonen and Koskela [22], calling a Borel-function $g\colon \mathcal{X}\to[0,\infty]$ an 'upper gradient' for an extended real-valued function $u\colon \mathcal{X}\to[-\infty,+\infty]$, if for all rectifiable curves $\gamma\colon [0,\ell_\gamma]\to\mathcal{X}$ there holds

$$\big|u(\gamma(0)-u\big(\gamma(\ell_\gamma)\big)\big| \le \int_\gamma g\,ds. \tag{2.3}$$

Moreover, if g is nonnegative and measurable on $\mathcal{X}$ and (2.3) holds for p-almost every path, which means that it fails only for a path family of zero p-modulus, then g is a p-weak upper gradient of u.

We define for $1\le p<\infty$ and for a fixed open subset $\Omega\subset\mathcal{X}$ the vector space

$$\widetilde{\mathcal{N}}^{1,p}(\Omega) := \big\{u\in L^p(\Omega)\colon \exists\ p\text{-integrable } p\text{-weak upper gradient of } u\big\}.$$

This space can be endowed with a norm

$$\|u\|_{\widetilde{\mathcal{N}}^{1,p}(\Omega)} := \|u\|_{L^p(\Omega)} + \inf_g \|g\|_{L^p(\Omega)} = \|u\|_{L^p(\Omega)} + \|g_u\|_{L^p(\Omega)},$$

where the infimum is taken over all p-integrable p-weak upper gradients of u, and g_u denotes the minimal p-weak upper gradient. Introducing the equivalence relation

$$u\sim v \quad :\Longleftrightarrow \quad \|u-v\|_{\widetilde{\mathcal{N}}^{1,p}(\Omega)}=0, \tag{2.4}$$

we define the *Newtonian space* $\mathcal{N}^{1,p}(\Omega)$ as the quotient space

$$\mathcal{N}^{1,p}(\Omega) := \widetilde{\mathcal{N}}^{1,p}(\Omega)/\sim. \tag{2.5}$$

Obviously this definition depends on the metric d and the measure μ so it would be more clear to write $\mathcal{N}^{1,p}(\Omega,\mathrm{d},\mu)$ instead of $\mathcal{N}^{1,p}(\Omega)$. However, since the measure μ and the metric d are fixed, we omit these two parameters in the notation. We refer the reader to [2] for more details on Newtonian spaces.

According to [2], the Sobolev *p-capacity* of a set $E\subset\mathcal{X}$ with respect to the space $\mathcal{N}^{1,p}(\mathcal{X})$ is defined by

$$C_p(E) := \inf_u \|u\|^p_{\mathcal{N}^{1,p}},$$

where the infimum is taken over all functions $u\in\mathcal{N}^{1,p}(\mathcal{X})$ such that $u|_E\ge 1$. We say that a property holds *p-almost everywhere on* $\mathcal{X}$ or *quasi everywhere on* $\mathcal{X}$, if the subset $E\subset\mathcal{X}$ on which the property fails to hold is of p-capacity zero.

We follow the method of [26] (see also [29]) and define for an arbitrary subset $E\subset\mathcal{X}$ the space $\widetilde{\mathcal{N}}^{1,p}_o(E)$ as the set of all functions $u\colon E\to[-\infty,+\infty]$ for which there exists a function $\tilde{u}\in\widetilde{\mathcal{N}}^{1,p}(\mathcal{X})$ such that $\tilde{u}=u$ μ-almost everywhere in E and $\tilde{u}=0$ p-almost everywhere on $\mathcal{X}\setminus E$. Then, we define an equivalence relation on $\widetilde{\mathcal{N}}^{1,p}_o(E)$ by saying that $u\sim v$ if $u=v$ μ-almost everywhere on E and we set

$$\mathcal{N}^{1,p}_o(E)=\widetilde{\mathcal{N}}^{1,p}_o(E)/\sim,$$

equipped with the norm

$$\|u\|_{\mathcal{N}^{1,p}_o(E)}=\|\tilde{u}\|_{\widetilde{\mathcal{N}}^{1,p}(\mathcal{X})}.$$

Parabolic Newtonian spaces

Since we are dealing with time-dependent problems, we have to introduce the parabolic Newtonian space $L^p(0,T;\mathcal{N}^{1,p}(\Omega))$ or its local version, respectively. The parabolic space

$$L^p\big(0,T;\mathcal{N}^{1,p}(\Omega)\big)$$

consists in all functions $u\colon\Omega\times(0,T)\to\mathbb{R}$ for which $u(\cdot,\cdot)\colon(0,T)\to\mathcal{N}^{1,p}(\Omega)$ is strongly measurable and the function $(0,T)\ni t\mapsto\|u(\cdot,t)\|_{\mathcal{N}^{1,p}(\Omega)}$ is contained in $L^p((0,T))$. Here, strongly measurable means that there exists a sequence of simple functions $u_k\colon(0,T)\to\mathcal{N}^{1,p}(\Omega)$ such that $\|u(\cdot,t)-u_k(\cdot,t)\|_{\mathcal{N}^{1,p}(\Omega)}\to0$ as $k\to\infty$ for a.e. $t\in(0,T)$.

To define the local version of these parabolic Newtonian spaces, we first introduce, according to [2, Chapter 2.6], the space $L^p_{\mathrm{loc}}(\mathcal{X})$ as the set of all functions $u\colon\mathcal{X}\to\mathbb{R}$, such that for every $x\in\mathcal{X}$ there exists $r_x>0$ such that $u\in L^p(B_{r_x}(x))$. Similarly we can define $\mathcal{N}^{1,p}_{\mathrm{loc}}(\mathcal{X})$. If $u\in\mathcal{N}^{1,p}_{\mathrm{loc}}(\mathcal{X})$, then for every subset $\Omega\Subset\mathcal{X}$ there holds $u\in\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega)$ and moreover, since $\mathcal{X}$ is doubling, the equivalence

$$u\in\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega)\quad\Longleftrightarrow\quad u\in\mathcal{N}^{1,p}\big(\Omega'\big)\quad\text{for all }\Omega'\Subset\Omega,$$

holds true, for every open subset $\Omega\subset\mathcal{X}$. Once having these local spaces at hand, we can understand the local parabolic space $L^p_{\mathrm{loc}}(0,T;\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega))$ as the space of all functions u, such that $u(\cdot,t)\in\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega)$ for almost every $t\in(0,T)$ and

$$\int_{t_1}^{t_2}\big\|u(\cdot,t)\big\|^p_{\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega)}\,dt<\infty,$$

for all $0<t_1<t_2<T$.

Vector-valued functions

We are studying vector-valued problems with $u\equiv(u^{(1)},...,u^{(N)})\colon\Omega\to\mathbb{R}^N$, $N\geq 1$ having components $u^{(i)}$, $i=1,...,N$. To define upper gradients for such functions u, we follow the approach in [24], saying that $g\colon\mathcal{X}\to[0,\infty]$ is a p-weak upper gradient for $u\colon\mathcal{X}\to\mathbb{R}^N$, if for p-almost all rectifiable paths $\gamma\colon[0,\ell_\gamma]\to\mathcal{X}$ there holds

$$\big\|u\big(\gamma(0)\big)-u\big(\gamma(\ell_\gamma)\big)\big\|\leq\int_\gamma g\,ds. \tag{2.6}$$

We may define the Newtonian space $\mathcal{N}^{1,p}(\Omega;\mathbb{R}^N)$ as the space of all equivalence classes of L^p-functions $u\colon\Omega\to\mathbb{R}^N$ such that there exists a p-weak upper gradient in the sense of (2.6) which is in $L^p(\Omega)$. Moreover, it is not difficult to see that this is equivalent to the fact that every component function $u_i\colon\Omega\to\mathbb{R}$ is in $\mathcal{N}^{1,p}(\Omega)$. In this way, we can also extend our definition of spaces with zero-boundary values and the local Newtonian spaces to the vector-valued case, obtaining $\mathcal{N}^{1,p}_o(\Omega,\mathbb{R}^N)$ and $\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega,\mathbb{R}^N)$.

Parabolic quasi-minimizers

Starting from a metric measure space $(\mathcal{X},\mathrm{d},\mu)$, we consider the product space $\mathcal{X}\times(0,T)$, $T>0$, which we endow with the product measure $\nu:=\mu\otimes\mathcal{L}^1$, where $\mathcal{L}^1$ denotes the 1-dimensional Lebesgue measure. Following the approach in [30], let us denote $\mathcal{K}(\Omega\times(0,T)):=\{K\subset\Omega\times(0,T)\colon K \text{ compact}\}$ and consider the functional

$$\begin{aligned}&\mathcal{F}\colon L^p_{\mathrm{loc}}(0,T;\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R}^N))\times\mathcal{K}(\Omega\times(0,T))\to\mathbb{R},\\ &\mathcal{F}[w,K]:=\frac{1}{p}\iint_K f(g_w)\,d\mu\,dt,\end{aligned} \tag{2.7}$$

where g_w denotes the minimal p-weak upper gradient of w and the integrand $f\colon[0,\infty]\to\mathbb{R}$ satisfies the growth condition

$$\nu\zeta^p-L_1\leq f(\zeta)\leq L\zeta^p+L_1, \tag{2.8}$$

for all $\zeta\in[0,\infty)$, with constants $0<\nu\leq L<\infty$ and $L_1\geq 0$ and with growth exponent $p>\frac{2\mathrm{n}}{\mathrm{n}+2}$, where $\mathrm{n}:=\log_2 c_d$ and c_d denotes the doubling constant (2.2).

Remark 2.1. Since we are dealing with parabolic problems, the solution we are handling with is a function $u\equiv u(x,t)$ depending on the variable x in the metric space $\mathcal{X}$ and on the time variable t in $\mathbb{R}$. Therefore we have to understand the concept of weak upper gradients in the sense that for a function $u\in L^p_{\mathrm{loc}}(0,T;\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R}^N))$ the parabolic minimal p-weak upper gradient of u is defined as

$$g_u(x,t):=g_{u(\cdot,t)}(x), \tag{2.9}$$

for ν-almost every $(x,t)\in\Omega\times(0,T)$.

Remark 2.2. (On the notation of upper gradients) In the course of this manuscript, the symbol g_u will always denote a minimal p-weak upper gradient of u. In case we deal with p-weak upper gradients in general (not necessarily minimal ones), we use the notation $\tilde g_u$.

Let us now give the notion of a parabolic quasi-minimizer on a metric space:

Definition 2.3. For an open set $\Omega\subset\mathcal{X}$ the function $u\in L^p_{\mathrm{loc}}(0,T;\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R}^N))$, $N\geq 1$, is said to be a *parabolic* $\mathcal{Q}$*-minimizer*, $\mathcal{Q}\geq 1$, if

$$-\iint_{\operatorname{spt}\Phi}\langle u,\partial_t\Phi\rangle\,d\mu\,dt+\mathcal{F}\big[u,\operatorname{spt}(\Phi)\big]\leq\mathcal{Q}\cdot\mathcal{F}\big[u-\Phi,\operatorname{spt}(\Phi)\big], \tag{2.10}$$

for all Lipschitz functions Φ with compact support in $\Omega\times(0,T)$, $\Phi\in\mathrm{Lip}_c(\Omega\times(0,t);\mathbb{R}^N)$. Here $\langle\cdot,\cdot\rangle$ denotes the standard scalar product on $\mathbb{R}^N$, and $\partial_t\Phi\equiv\frac{\partial\Phi}{\partial t}$.

Poincaré inequality

We demand that the metric measure space $(\mathcal{X},\mathrm{d},\mu)$ supports a weak $(1,p)$-Poincaré inequality in the sense that there exist constants $c_P>0$ and $\Gamma>1$ such that for all open balls $B_\varrho(x_o)\subset B_{\Gamma\varrho}(x_o)\subset\mathcal{X}$, for all p-integrable functions u on X and all upper gradients $\tilde g_u$ of u there holds

$$\fint_{B_\varrho(x_o)}|u-u_{\varrho,x_o}|\,d\mu\leq c_P\,\varrho\left[\fint_{B_{\Gamma\varrho}(x_o)}\tilde g_u^p\,d\mu\right]^{\frac{1}{p}}, \tag{2.11}$$

where the symbol

$$u_{\varrho,x_o}:=\fint_{B_\varrho(x_o)}u\,d\mu:=\frac{1}{\mu(B_\varrho(x_o))}\int_{B_\varrho(x_o)}u\,d\mu$$

denotes the mean value integral of the function u on the ball $B_\varrho(x_o)$ with respect to the measure μ. In the sequel we will refer to this inequality simply as 'Poincaré inequality', omitting the word 'weak'. Poincaré inequalities on metric measure spaces have been studied quite extensively in the literature, see for example [1], [4], [23], [25], [28], [31], [33], [40] and [41].

The main result

Our aim is to show a local higher integrability result for a parabolic quasi-minimizer. Our main result is therefore the following:

Theorem 2.4. *Let* $(\mathcal{X}, \mathrm{d}, \mu)$ *be a complete metric space with a doubling measure* μ *and doubling constant* $c_d{=:}2^{\mathrm{n}}$, *which supports a weak* $(1,p)$*-Poincaré inequality with parameter* $\Gamma \geq 1$, *and let, for an open set* $\Omega \subset \mathcal{X}$ *and* $p > \frac{2\mathrm{n}}{\mathrm{n}+2}$, $u \in L^p_{\mathrm{loc}}(0,T;\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R}^N))$, $N \geq 1$, *be a vector-valued local parabolic* $\mathcal{Q}$*-minimizer in the sense of* (2.10), *where the functional* $\mathcal{F}$ *fulfills the growth condition* (2.8).

Then there exists $\varepsilon_o \equiv \varepsilon_o(\mathrm{n}, N, L, \nu, L_1, p, \mathcal{Q}, \Gamma) > 0$ *such that*

$$u \in L^{p+\varepsilon_o}_{\mathrm{loc}}\big(0,T;\mathcal{N}^{1,p+\varepsilon_o}_{\mathrm{loc}}\big(\Omega;\mathbb{R}^N\big)\big).$$

Moreover, there exists a constant $c \equiv c(\mathrm{n}, N, L, \nu, L_1, p, \mathcal{Q}, \Gamma)$ *such that for every* $\varepsilon \in (0, \varepsilon_o]$ *and for every parabolic cylinder* Q_o *with* $2Q_o \subset \Omega_T$, *the minimal upper gradient* g_u *of* u *satisfies the estimate*

$$\iint_{Q_o} g_u^{p+\varepsilon}\, d\mu\, dt \leq c\left[\iint_{2Q_o}(1+g_u)^p\, d\mu\, dt\right]^{\varepsilon\frac{d}{p}} \iint_{2Q_o} g_u^p\, d\mu\, dt,$$

where d *denotes the parabolic deficit*

$$d \equiv \begin{cases} \frac{p}{2}, & \text{if } p \geq 2,\\ \frac{2p}{p(\mathrm{n}+2)-2\mathrm{n}}, & \text{if } p < 2. \end{cases} \tag{2.12}$$

Remark 2.5. The result of Theorem 2.4 requires merely the p-growth condition (2.8) on the integrand and can therefore be easily generalized to functionals of the type

$$\mathcal{F}[v,K] \equiv \int_K f(z, v, g_v)\, d\mu,$$

with a Carathéodory integrand $f\colon \Omega_T \times \mathbb{R}^N \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ and a growth condition

$$\nu\zeta^p - L_1 \leq f(z,\xi,\zeta) \leq L\zeta^p + L_1,$$

for all $z \in \Omega_T$, $\xi \in \mathbb{R}^N$ and $\zeta \in \mathbb{R}_{\geq 0}$, and with constants $0 < \nu \leq L < \infty$ and $L_1 \geq 0$ and with growth exponent $p > \frac{2\mathrm{n}}{\mathrm{n}+2}$.

3. Preliminaries

In this chapter we collect some basic properties of metric measure spaces supporting a Poincaré inequality, and we provide the tools which we need to prove our main theorem.

3.1. Properties of metric measure spaces

As a direct consequence of the doubling property (2.1) we have for every ball $B_R(x)\subset\mathcal{X}$, $y\in B_R(x)$ and $0<r\le R<\infty$ that

$$\frac{\mu(B_R(x))}{\mu(B_r(y))}\le C\left(\frac{R}{r}\right)^{\mathrm{n}}, \tag{3.1}$$

where $\mathrm{n}=\log_2 c_d$ and $C=c_d^2$. Moreover, since $\mathcal{X}$ is connected there exists a constant $\tilde{c}$ and an exponent $\sigma>0$, such that

$$\frac{\mu(B_r(x))}{\mu(B_R(x))}\le\tilde{c}\left(\frac{r}{R}\right)^{\sigma}.$$

According to (3.1), the number n plays the role of a 'dimension from below', related to the measure μ; however we point out that n in general is not an integer.

Since $\mathcal{X}$ is a doubling space, it supports the Vitali covering theorem and therefore also the differentiation theorem of Lebesgue, i.e. for every nonnegative locally μ-integrable function on $\mathcal{X}$ we have that

$$\lim_{r\to 0}⨍_{B_r(x)} f\,d\mu=f(x),$$

for μ-almost all $x\in\mathcal{X}$.

An important fact, which we will use later to construct suitable test functions, is the following

Remark 3.1. ([2, Proposition 1.14]) If $f\colon\mathcal{X}\to\mathbb{R}$ is a locally Lipschitz continuous function, then the pointwise dilation

$$\operatorname{lip} f(x)=\liminf_{r\to 0}\sup_{y\in B_r(x)}\frac{|f(y)-f(x)|}{r}$$

is an upper gradient of f.

Now we collect some calculus rules for upper gradients in a metric measure space. The proofs of the following facts can be found in [2], where they are formulated for real-valued function on metric measure spaces. However, they also hold true if we replace the target space by an arbitrary Banach space F, in particular $F=\mathbb{R}^N$ with $N\ge 1$.

Lemma 3.2. (p-weak upper gradients) *Let $u \in \mathcal{N}^{1,p}(\mathcal{X};\mathbb{R}^N)$ and let $g \in L^p(\mathcal{X})$ be a p-weak upper gradient of u. Then for p-almost every curve $\gamma\colon [0,\ell_\gamma] \to \mathcal{X}$ there holds*

$$\big\|(u\circ\gamma)'(\xi)\big\| \le g\big(\gamma(\xi)\big), \quad \text{for almost every } \xi \in [0,\ell_\gamma]. \tag{3.2}$$

Conversely, if $g \ge 0$ is measurable, $u \in \mathcal{N}^{1,p}(\mathcal{X};\mathbb{R}^N)$ and (3.2) holds for p-almost every curve $\gamma\colon [0,\ell_\gamma] \to \mathcal{X}$, then g is a p-weak upper gradient of u.

Lemma 3.3. (Leibniz rule [2, Theorem 2.15]) *Let $u,v \in \mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R}^N)$ and let $g_u, g_v \in L^p_{\mathrm{loc}}(\Omega)$ be p-weak upper gradients of u and v, respectively. Then the functions $g_u + g_v$ and $|u|g_v + |v|g_u$ are p-weak upper gradients for $u+v$ and uv, respectively.*

Remark 3.4. We note here that in [2, Theorem 2.15] the Leibniz-rule is formulated for minimal p-weak upper gradients of u and v. However, a look at the proof of the statement shows, that at no point the minimality of g_u or g_v is needed. So the result applies also when we replace the minimal upper gradients g_u and g_v by arbitrary p-weak upper gradients $\tilde{g}_u$ and $\tilde{g}_v$. On the other hand, even in the case where one takes the minimal weak upper gradients g_u and g_v in the above statement, the function $|u|g_v + |v|g_u$ is not necessarily minimal. Counter examples can be found in the book [2].

Lemma 3.5. (Chain rule [2, Theorem 2.16]) *Let $u \in \mathcal{N}^{1,p}(\Omega;\mathbb{R}^N)$ and $\varphi\colon \mathbb{R} \to \mathbb{R}$ be locally Lipschitz. Then the function $|\varphi' \circ u| g_u$ is a minimal p-weak upper gradient of $\varphi \circ u$.*

A direct consequence of Lemmas 3.2 and 3.5 is the following rule which we will use frequently when calculating upper gradients of test functions:

Lemma 3.6. ([2, Theorem 2.18]) *Let $u,v \in \mathcal{N}^{1,p}(\Omega;\mathbb{R}^N)$ and η be Lipschitz on Ω, such that $0 \le \eta \le 1$. Set $w := u + \eta(v-u) = (1-\eta)u + \eta v$. Then*

$$g := (1-\eta)g_u + \eta g_v + |v-u| g_\eta$$

is a p-weak upper gradient of w.

Remark 3.7. Note that all these rules can be understood also for time-dependent functions $u \in L^p(0,T;\mathcal{N}^{1,p}(\Omega;\mathbb{R}^N))$, if we define the (minimal) p-weak upper gradients of a time-dependent function $u(x,t)$ as in (2.9).

Lemma 3.8. (L^p-convergence [2, Proposition 2.3]) *Let $f_j\in\mathcal{N}^{1,p}(\mathcal{X})$ $(j=1, 2, ...)$ be a sequence of functions with p-weak upper gradients $g_j\in L^p(\mathcal{X})$, such that $f_j\to f$ and $g_j\to g$ in $L^p(\mathcal{X})$. Moreover let $f\in\mathcal{N}^{1,p}(\mathcal{X})$. Then g is a p-weak upper gradient of f.*

One very important property of minimal p-weak upper gradients is that they are local in the sense that if two functions coincide on a set, then also their minimal upper gradients coincide on this set. We will need this property at many stages of our regularity proof, in particular when constructing suitable test functions. The result is taken from [2, Chapter 2.4].

Lemma 3.9. *Let u, v be functions on $\mathcal{X}$ with minimal upper gradients g_u and g_v. Then is holds that*

$$g_u = g_v \quad \textit{almost everywhere on the set} \quad \big\{x\in\mathcal{X}:\, u(x)=v(x)\big\}.$$

Moreover, if $c\in\mathbb{R}$ is a constant, then $g_u=0$ almost everywhere on the set $\{x\in \mathcal{X}:\, u(x)=c\}$.

3.2. Poincaré inequalities and Sobolev embedding

By Hölder's inequality it directly follows that if a metric space supports a $(1,p)$-Poincaré inequality, then it supports a $(1,q)$-Poincaré inequality for all $q\ge p$. On the other hand it was shown in [25] that if a complete metric space is endowed with a doubling measure and supports a $(1,p)$-Poincaré inequality, then it supports also a $(1,p-\varepsilon)$-Poincaré inequality for some $\varepsilon\equiv\varepsilon(c_P,\Lambda,c_d,p)>0$, and therefore also a $(1,q)$-Poincaré inequality for all $q\in[p-\varepsilon,p]$. Moreover, from [20] we know that if we assume a weak $(1,p)$-Poincaré inequality, then the Sobolev embedding theorem holds and hence a weak (q,p)-Poincaré inequality holds for all $q\le p^*$, with

$$p^* := \begin{cases} \frac{p\mathrm{n}}{\mathrm{n}-p}, & p<\mathrm{n}, \\ +\infty, & p\ge \mathrm{n}. \end{cases} \tag{3.3}$$

On the other hand it was shown in [29], see also [19], [20] and [22], that in this case for every $u\in\mathcal{N}^{1,p}(B_{2\Gamma\varrho}(x_o))$ with $B_{2\Gamma\varrho}(x_o)\subset\mathcal{X}$ the following Sobolev-type inequality holds:

$$\left[\fint_{B_\varrho(x_o)} |u-u_{\varrho,x_o}|^q\,d\mu\right]^{\frac{1}{q}} \le c_*\varrho\left[\fint_{B_{2\Gamma\varrho}(x_o)} g_u^p\,d\mu\right]^{\frac{1}{p}}, \quad \text{for all } 1\le q\le p^*. \tag{3.4}$$

The constant c_* in the above inequality depends only on c_d and on the constant c_P in the weak $(1,p)$-Poincaré inequality.

Poincaré and Sobolev inequalities hold also on more general domains. More precisely, the Poincaré inequality holds on bounded measurable subsets E of the metric space $\mathcal{X}$ such that the p-capacity of the complement $\mathcal{X}\setminus E$ does not vanish. In detail we have

$$\int_E |u|^p \, d\mu \le C_E \int_E g_u^p \, d\mu,$$

for every function $u\in\mathcal{N}_o^{1,p}(E)$ and for every bounded measurable set $E\subset\mathcal{X}$ with $\mathrm{Cap}_p(\mathcal{X}\setminus E)>0$. The constant C_E depends on c_P, c_d, p and E.

A consequence of the Poincaré and Sobolev inequality, respectively, is a Gagliardo–Nirenberg type interpolation theorem on metric measure spaces. It is not difficult to prove such a theorem, since it follows directly from the Sobolev inequality by a standard L^p-interpolation argument.

Lemma 3.10. *Let $\mathcal{X}$ be a doubling metric measure space with doubling constant $c_d=2^{\mathrm{n}}$, supporting a weak $(1,p)$-Poincaré inequality, with constant c_P and dilatation constant $\Gamma\ge 1$. Then for all exponents $1\le r\le\sigma\le\tau^*$ with $1\le\tau<\mathrm{n}$ and $\tau^*:=\frac{\mathrm{n}\tau}{\mathrm{n}-\tau}$ denoting the Sobolev conjugate of τ according to (3.3), and for $\theta\in[0,1]$ such that*

$$-\frac{\mathrm{n}}{\sigma}\le\theta\left(1-\frac{\mathrm{n}}{\tau}\right)-(1-\theta)\frac{\mathrm{n}}{r}, \tag{3.5}$$

we have the interpolation estimate

$$\fint_{B_\varrho(x_o)}\left|\frac{u-u_{\varrho,x_o}}{\varrho}\right|^\sigma d\mu\le c\left[\fint_{B_{2\Gamma\varrho}(x_o)} g_u^\tau \, d\mu\right]^{\frac{\theta\sigma}{\tau}}\left[\fint_{B_\varrho(x_o)}\left|\frac{u-u_{\varrho,x_o}}{\varrho}\right|^r d\mu\right]^{\frac{(1-\theta)\sigma}{r}}, \tag{3.6}$$

which holds for every $u\in\mathcal{N}^{1,\tau}(B_{2\Gamma\varrho}(x_o))$. The constant c depends on n, τ, ϑ, σ and c_P.

Having in mind the discussion on upper gradients and arguing on component functions, the Poincaré inequality can be extended to vector-valued functions in a straight forward way and we obtain a Poincaré inequality for functions $u\colon\mathcal{X}\to\mathbb{R}^N$:

$$\fint_{B_\varrho(x)}\|u-u_\varrho\| \, d\mu\le N\cdot c_P\,\varrho\left[\fint_{B_{\Gamma\varrho}(x)} g^p \, d\mu\right]^{\frac{1}{p}},$$

for every p-weak upper gradient g of u. From this point it is obvious that also the Sobolev inequality (3.4) can be extended to vector-valued functions, if we replace the Sobolev constant c_* by a constant $\tilde{c}_*\equiv\tilde{c}_*(c_*,N,p)$.

3.3. Smoothing procedures

Having a look at the previous definition of a parabolic $\mathcal{Q}$-minimizer, we see that we are faced with the same problem as in the Euclidean case: It is not possible to test the inequality (2.10) by $\Phi \approx u$, since u is a priori not regular enough with respect to time. However, since we deal with functions on the product space $\mathcal{X}\times(0,T)$, where $(0,T)$ is endowed with the usual $\mathcal{L}^1$-measure, we can follow the standard strategy by smoothing the test functions with respect to the time variable. In particular we define for $\Phi\colon \Omega\times(0,T)\to\mathbb{R}$ the smoothed function

$$\Phi_\varepsilon(x,t) := \int_{\mathbb{R}} \Phi(x,t-s)\varphi_\varepsilon(s)\,ds, \tag{3.7}$$

with a standard smoothing kernel φ_ε with $\operatorname{spt}\varphi_\varepsilon \subset (-\varepsilon,\varepsilon)$. Using such a smoothed function in the inequality (2.10), we obtain by a change of variable and Fubini's theorem the inequality

$$-\iint_{\operatorname{spt}\Phi_\varepsilon} \langle u_\varepsilon, \partial_t\Phi_\varepsilon\rangle\,d\mu\,dt + \mathcal{F}[u,\operatorname{spt}\Phi_\varepsilon] \le \mathcal{Q}\cdot\mathcal{F}[u-\Phi_\varepsilon,\operatorname{spt}\Phi_\varepsilon], \tag{3.8}$$

where u_ε denotes the regularization of u with respect to time, according to (3.7). The advantage of this concept is, that we can now proceed as in the Euclidean case, testing (2.10) by $\Phi\approx u_\varepsilon$, then perform an integration by parts with respect to the time variable in order to move the time derivative from Φ onto the function u_ε. By the growth conditions on $\mathcal{F}$ and the properties of the smoothing kernel, one may conclude strong convergence of the appearing integrals as $\varepsilon\to 0$. We will carry out this argument later in the proofs.

One basic observation on smoothed functions with respect to time, which we will need later at various points, is the following

Lemma 3.11. *Let $u\in L^p_{\mathrm{loc}}(0,T;\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R}^N))$ and $\tilde g_u$ be a p-weak upper gradient of u. Moreover let u_ε be the mollification of $u(x,t)$ with respect to time in the sense of (3.7) and let g_{u_ε} be a minimal p-weak upper gradient of u_ε. Then we have*

$$g_{u_\varepsilon}(x,t) \le (\tilde g_u)_\varepsilon(x,t) = \int_{\mathbb{R}} \tilde g_u(x,t-\tau)\varphi_\varepsilon(\tau)\,d\tau.$$

Proof. We fix a set $K\times(t_1,t_2)$ such that $K\Subset\Omega$ and $0<t_1<t_2<T$. First we note that we can argue on upper gradients instead of p-weak upper gradients. This is for the following reason. We fix $\delta>0$. Then by Lemma 1.46 in [2] we find $\tilde g_u^\delta$ such that $\tilde g_u^\delta(\cdot,t)$ is an upper gradient for $u(\cdot,t)$ for almost all $t\in(0,T)$ and $\|\tilde g_u - \tilde g_u^\delta\|_{L^p(K\times(t_1,t_2))}<\delta$. Hence, we can replace $\tilde g_u$ by $\tilde g_u^\delta$ in our argument. (2.3) with the upper gradient $\tilde g_u^\delta$ holds true for all curves $\gamma\colon [0,\ell_\gamma]\to\mathcal{X}$. By (3.7) and Fubini's theorem we then get

$$\begin{aligned}\big\|u_\varepsilon\big(\gamma(0),t\big)-u_\varepsilon\big(\gamma(\ell_\gamma),t\big)\big\| &\le \int_{\mathbb{R}}\varphi_\varepsilon(\tau)\big\|u\big(\gamma(0),t-\tau\big)-u\big(\gamma(\ell_\gamma),t-\tau\big)\big\|\,d\tau\\ &\le \int_{\mathbb{R}}\varphi_\varepsilon(\tau)\int_0^{\ell_\gamma}\tilde g_u^\delta\big(\gamma(s),t-\tau\big)\,ds\,d\tau\\ &=\int_0^{\ell_\gamma}\!\!\int_{\mathbb{R}}\varphi_\varepsilon(\tau)\tilde g_u^\delta\big(\gamma(s),t-\tau\big)\,d\tau\,ds=\int_\gamma\big(\tilde g_u^\delta\big)_\varepsilon(\cdot,t)\,ds.\end{aligned}$$

So $(\tilde g_u^\delta)_\varepsilon$ is an upper gradient of u_ε. The statement follows now with the L^p-convergence of $(\tilde g_u^\delta)_\varepsilon\to(\tilde g_u)_\varepsilon$, since g_{u_ε} is a minimal p-weak upper gradient. □

The following lemma is a key ingredient for the smoothing procedure with respect to time. It ensures that the upper gradient g_{u-u_ε} tends to zero, when $\varepsilon\downarrow 0$, if u_ε denotes the time-regularized function with parameter $\varepsilon>0$. The proof is performed by Masson and Siljander in [37] and uses the Cheeger differential calculus to deduce the strong L^p-convergence of g_{u-u_ε}.

Lemma 3.12. *Let $u\in L^p_{\mathrm{loc}}(0,T;\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R}^N))$ and let g_{u-u_ε} be a minimal p-weak upper gradient of $u-u_\varepsilon$, where u_ε denotes the mollification of u with respect to the time variable. Then, as $\varepsilon\to 0$ we have that $g_{u-u_\varepsilon}\to 0$ in $L^p_{\mathrm{loc}}(\Omega_T)$ and also pointwise $\mu\otimes\mathcal{L}^1$-almost everywhere on Ω_T.*

3.4. The parabolic geometry, cut-off functions and weighted means

We consider the product space $\mathcal{X}\times\mathbb{R}$ endowed with the product measure $\mu\otimes\mathcal{L}^1$ and denote the *parabolic cylinder* $\Omega_T:=\Omega\times(0,T)$. For a point $z_o=(x_o,t_o)\in\Omega_T$, a radius $\varrho>0$ and a parameter $\lambda>0$ we write

$$Q_\varrho^{(\lambda)}(z_o):=B_\varrho(x_o)\times\Lambda_\varrho^{(\lambda)}(t_o),\quad \text{with } \Lambda_\varrho^{(\lambda)}(t_o):=\big(t_o-\lambda^{2-p}\varrho^2,t_o+\lambda^{2-p}\varrho^2\big).$$

Here, $B_\varrho(x_o)$ is the open ball of radius ϱ and center x_o with respect to the metric d. We define the *parabolic distance* between two points $z=(x,t)$ and $\tilde z=(\tilde x,\tilde t)$ in the cylinder Ω_T as

$$\mathrm{d}_{\mathrm{par}}(z,\tilde z):=\max\big\{\mathrm{d}(x,\tilde x),\sqrt{|t-\tilde t|}\big\}.$$

Parabolic cylinders $Q_\varrho(z_o)\subset\Omega_T$ are the 'balls' with respect to the parabolic metric, i.e.

$$Q_\varrho(z_o)=\big\{z\in\Omega_T\colon \mathrm{d}_{\mathrm{par}}(z,z_o)<\varrho\big\}.$$

Moreover, we use the notation

$$\big(\mu\otimes\mathcal{L}^1\big)\big(Q_\varrho^{(\lambda)}(z_o)\big)\equiv\mu\big(B_\varrho(x_o)\big)\cdot\mathcal{L}^1\big(\Lambda_\varrho^{(\lambda)}(t_o)\big).$$

For a function $\xi\in\mathrm{Lip}(\mathcal{X})$, a ball $B_\varrho(x_o)\subset\mathcal{X}$ with $\|\xi\|_{L^1(B_\varrho(x_o))}\neq 0$ we define the *weighted mean* of u with respect to ξ by

$$\begin{aligned} u^{\xi}_{\varrho,x_o}(t) &:= \frac{1}{\|\xi\|_{L^1(B_\varrho(x_o))}}\int_{B_\varrho(x_o)}\xi(x)u(x,t)\,d\mu(x) \\ &= \bigg[\int_{B_\varrho(x_o)}\xi(x)\,d\mu(x)\bigg]^{-1}\int_{B_\varrho(x_o)}\xi(x)u(x,t)\,d\mu(x). \end{aligned}$$

Moreover, we consider cut-off functions η of the type

$$\eta(x,t):=\xi(x)\zeta(t), \tag{3.9}$$

where $\xi\colon\mathcal{X}\to\mathbb{R}$ and $\zeta\colon\mathbb{R}\to\mathbb{R}$ have the following properties: For a pair of parabolic cylinders $(Q_\sigma^{(\lambda)}(z_o),Q_\tau^{(\lambda)}(z_o))$ with radii $\frac{\varrho}{2}\leq\sigma<\tau\leq\varrho$, center $z_o=(x_o,t_o)$ and $Q_\varrho^{(\lambda)}(z_o)\subset\Omega_T$ we demand that

- $\xi(x)$ is a cut-off function in space of the following form:

$$\xi(x):=\min\bigg\{\frac{\tau-\mathrm{d}(x_o,x)}{\tau-\sigma},1\bigg\}_+\equiv\begin{cases}1 & \text{on } B_\sigma(x_o),\\ \frac{\tau-\mathrm{d}(x_o,x)}{\tau-\sigma} & \text{on } B_\tau(x_o)\setminus B_\sigma(x_o)\\ 0 & \text{on } \mathcal{X}\setminus B_\tau(x_o).\end{cases} \tag{3.10}$$

This function is clearly an element of $\mathcal{N}_o^{1,p}(B_\tau(x_o))$ and it is Lipschitz. Moreover, by Remark 3.1, and the local nature of minimal upper gradients, Lemma 3.9, we deduce that on $B_\sigma(x_o)$ and on $\mathcal{X}\setminus B_\tau(x_o)$ the function $\tilde g_\xi\equiv 0$ is an upper gradient for ξ, and on $B_\tau(x_o)\setminus B_\sigma(x_o)$, an upper gradient is given by the Lipschitz constant $1/(\tau-\sigma)$, so we have for the minimal p-weak upper gradient:

$$g_\xi\leq\frac{1}{\tau-\sigma}\chi_{B_\tau\setminus B_\sigma}\equiv\begin{cases}\frac{1}{\tau-\sigma} & \text{on } B_\tau(x_o)\setminus B_\sigma(x_o)\\ 0 & \text{on } B_\sigma(x_o) \text{ and } \mathcal{X}\setminus B_\tau(x_o).\end{cases} \tag{3.11}$$

On the other hand we easily calculate that

$$\begin{aligned} \fint_{B_\tau(x_o)}\xi(x)\,d\mu(x) &= \frac{1}{\mu(B_\tau(x_o))}\bigg[\mu\big(B_\sigma(x_o)\big)+\int_{B_\tau(x_o)\setminus B_\sigma(x_o)}\frac{\tau-\mathrm{d}(x_o,x)}{\tau-\sigma}\,d\mu(x)\bigg] \\ &\geq\frac{\mu(B_\sigma(x_o))}{\mu(B_\tau(x_o))}\geq c_d^{-2}\Big(\frac{\sigma}{\tau}\Big)^{\mathrm{n}}\geq c_d^{-2}2^{-\mathrm{n}}=c_d^{-3}. \end{aligned}$$

Here we first used that $\mathrm{d}(x_o,x)\le\tau$ for $x\in B_\tau\setminus B_\sigma$, thereafter the doubling property (3.1) together with $\frac{\varrho}{2}\le\sigma<\tau\le\varrho$ and finally $\mathrm{n}=\log_2 c_d$. Since on the other hand we obviously have $\sup_{B_\tau(x_o)}\xi=1$, we deduce

$$\sup_{B_\tau(x_o)}\xi\le c^*\fint_{B_\tau(x_o)}\xi(x)\,d\mu(x), \tag{3.12}$$

with $c^*:=c_d^3$.

- $\zeta(t)$ is a smooth cut-off function in time with

$$\operatorname{spt}\zeta\subset\Lambda_\tau^{(\lambda)}(t_o),\quad 0\le\zeta\le 1,\ \zeta\equiv 1 \text{ on } \Lambda_\sigma^{(\lambda)}(t_o). \tag{3.13}$$

We recall that $\Lambda_\tau^{(\lambda)}(t_o)=(t_o-\lambda^{2-p}\tau^2,t_o+\lambda^{2-p}\tau^2)\subset(0,T)$. Moreover we demand

$$\lambda^{2-p}|\zeta'|\le\frac{2}{(\tau-\sigma)^2}. \tag{3.14}$$

Defining

$$u_{\varrho,x_o}(t):=\fint_{B_\varrho(x_o)}u(x,t)\,d\mu(x),$$

we have for weight functions ξ, which satisfy (3.12), the following

Lemma 3.13. *For $B_\varrho(x_o)\subset\Omega$ and $u(\cdot,t)\in L^p_{\mathrm{loc}}(\Omega)$, $p\ge1$, radii $0<\tau\le\varrho$ and a weight function ξ of the form (3.10), there exists a constant $c\equiv c(p,c^*)\ge1$ such that*

$$\begin{aligned}c^{-1}\fint_{B_\tau(x_o)}\left|u(\cdot,t)-u^\xi_{\tau,x_o}(t)\right|^p\,d\mu&\le\fint_{B_\tau(x_o)}\left|u(\cdot,t)-u_{\tau,x_o}(t)\right|^p\,d\mu\\&\le c\fint_{B_\tau(x_o)}\left|u(\cdot,t)-u^\xi_{\tau,x_o}(t)\right|^p\,d\mu,\end{aligned}$$

for almost all $t\in(0,T)$.

Proof. The proof follows line by line the argument in the Euclidean setting, see [39]. □

Moreover, we note that we have a quasi-minimizing property of the mean value in the following sense:

Lemma 3.14. ([2, Lemma 4.17]) *For $1\le p<\infty$, $A\subset\mathcal{X}$ μ-measurable and $u\in L^p(\mathcal{X})$ there holds*

$$\left[\fint_A|u-u_A|^p\,d\mu\right]^{1/p}\le2\left[\fint_A|u-a|^p\,d\mu\right]^{1/p},$$

for every $a\in\mathbb{R}$.

3.5. Technical lemmas

We will use the following two standard lemmas:

Lemma 3.15. *For $R_1 < R_2$ let $\Phi : [R_1, R_2] \to [0, \infty)$ be a bounded function. We assume that for all radii $R_1 \le \sigma < \varrho \le R_2$ there holds*

$$\Phi(\sigma) \le \vartheta \Phi(\varrho) + \frac{A}{(\varrho - \sigma)^\alpha} + \frac{B}{(\varrho - \sigma)^\beta} + C,$$

where $\vartheta \in (0,1)$ and $A, B, C \ge 0$ are fixed parameters, and moreover $\alpha, \beta \ge 0$. Then there exists a constant $c \equiv c(\vartheta, \alpha, \beta)$ such that

$$\Phi(r) \le c \left[\frac{A}{(R-r)^\alpha} + \frac{B}{(R-r)^\beta} + C \right],$$

for all $R_1 \le r < R \le R_2$.

Lemma 3.16. (see [14, Lemma 8.3]) *Let $\xi, \eta \in \mathbb{R}^n$. Then for any $s > -1$ and $r > 0$ there exist constants $c_1 \equiv c_1, c_2 \equiv c_1, c_2(s,r)$ such that*

$$c_1 \big(1 + |\xi|^2 + |\eta|^2\big)^{s/2} \le \int_0^1 (1-t)^r \big(1 + \big|(1-t)\xi + t\eta\big|^2\big)^{s/2} \, dt \le c_2 \big(1 + |\xi|^2 + |\eta|^2\big)^{s/2}.$$

4. Proof of the higher integrability property

4.1. Caccioppoli inequality

In the first step, we prove a (Pre-) Caccioppoli type inequality for quasi-minimizers.

Proposition 4.1. (Preliminary Caccioppoli type inequality) *On the metric measure space $(\mathcal{X}, \mathrm{d}, \mu)$ from Theorem 2.4, let $u \in L^p_{\mathrm{loc}}(0, T; \mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega; \mathbb{R}^N))$ be a local parabolic $\mathcal{Q}$-minimizer with $\mathcal{Q} \ge 1$ in the sense of (2.10), under the growth assumption (2.8). For fixed $\lambda > 0$ we consider the parabolic cylinder $\Omega_R^{(\lambda)} \subset \Omega$. Denoting with g_u the minimal upper gradient of u, there exists a constant $c \equiv c(\mathrm{n}, p) > 0$ such that for every pair of radii (σ, τ) with $\varrho \le \sigma < \tau \le R$ and every cut-off function η on $(Q_\sigma^{(\lambda)}(z_o), Q_\tau^{(\lambda)}(z_o))$ of the form (3.9), $\eta(x,t) = \xi(x)\zeta(t)$, with the properties (3.10)* *to* (3.14) *the following estimate holds true*:

$$\sup_{t\in\Lambda_\sigma^{(\lambda)}(t_o)}\int_{B_\sigma(x_o)}\big|u(\cdot,t)-u^\xi_{\tau,x_o}(t)\big|^2\,d\mu+\nu\iint_{Q_\sigma^{(\lambda)}(z_o)}g_u^p\,d\mu\,dt$$
$$\le c\mathcal{Q}L\iint_{Q_\tau^{(\lambda)}(z_o)\setminus Q_\sigma^{(\lambda)}(z_o)}g_u^p\,d\mu\,dt+cL_1(\mathcal{Q}+1)\big|Q_\tau^{(\lambda)}(z_o)\big|$$
$$c\mathcal{Q}L\iint_{Q_\tau^{(\lambda)}(z_o)}\bigg|\frac{u-u_{\tau,x_o}(t)}{\tau-\sigma}\bigg|^p+\lambda^{p-2}\bigg|\frac{u-u_{\tau,x_o}(t)}{\tau-\sigma}\bigg|^2\,d\mu\,dt. \tag{4.1}$$

Proof. We follow basically the proof in [16] in the Euclidean setting. However, several difficulties have to be overcome due to the non-linear behaviour of upper gradients. For a fixed time $t_1\in\Lambda_\sigma^{(\lambda)}(t_o)$ let χ^h_{0,t_1} be the piecewise affine approximation of the characteristic function $\chi_{[0,t_1]}$ with the following properties.

$$\chi^h_{0,t_1}(t):=1 \text{ for } h\le t\le t_1-h,\quad \chi^h_{0,t_1}(t)=0 \text{ for } t\le\frac{h}{10},\ \text{resp. } t\ge t_1-\frac{h}{10}.$$

and moreover

$$\big|\partial_t\chi^h_{0,t_1}(t)\big|=\begin{cases}\frac{10}{9h}, & \text{for } t\in(\frac{h}{10},h) \text{ and } t\in(t_1-h,t_1-\frac{h}{10}),\\ 0, & \text{else.}\end{cases}$$

We use as test function in the formulation (3.8) of the parabolic $\mathcal{Q}$-minimality the function

$$\Phi^h_\varepsilon(x,t)\equiv\Phi^h_\varepsilon(x,t;x_o):=\eta(x,t)\big(u_\varepsilon(x,t)-\big(u^\xi_{\tau,x_o}\big)_\varepsilon(t)\big)\chi^{h,\varepsilon}_{0,t_1}(t),$$

where u_ε, $(u^\xi_{\tau,x_o})_\varepsilon$ and $\chi^{h,\varepsilon}_{0,t_1}$ denote the mollifications of the functions u, u^ξ_{τ,x_o} and χ^h_{0,t_1} with respect to time, according to (3.7) with a parameter $\varepsilon<\frac{h}{50}$. In particular we have

$$\big(u^\xi_{\tau,x_o}\big)_\varepsilon(t)=\int_{\mathbb{R}}\frac{1}{\|\xi\|_{B_\tau(x_o)}}\bigg[\fint_{B_\tau(x_o)}u(x,t-s)\xi(x)\,d\mu(x)\bigg]\varphi_\varepsilon(s)\,ds$$
$$=\frac{1}{\|\xi\|_{B_\tau(x_o)}}\fint_{B_\tau(x_o)}\bigg[\int_{\mathbb{R}}u(x,t-s)\varphi_\varepsilon(s)\,ds\bigg]\xi(x)\,d\mu(x)=(u_\varepsilon)^\xi_{\tau,x_o}(t).$$

Furthermore, $\eta(x,t)=\xi(x)\zeta(t)$ denotes the cut-off function defined in (3.9) on the pair of cylinders $(Q_\sigma^{(\lambda)}(z_o),Q_\tau^{(\lambda)}(z_o))$. We get on the left hand side of (3.8) the expression

$$\mathcal{L}^h_\varepsilon:=-\iint_{\operatorname{spt}\Phi^h_\varepsilon}\big\langle u_\varepsilon,\partial_t\Phi^h_\varepsilon\big\rangle\,d\mu\,dt.$$

In a first step we write

$$\mathcal{L}^h_\varepsilon = -\iint_{\mathrm{spt}\,\Phi^h_\varepsilon} \big\langle u_\varepsilon - \big(u^\xi_{\tau,x_o}\big)_\varepsilon, \partial_t \Phi^h_\varepsilon \big\rangle \,d\mu\,dt - \iint_{\mathrm{spt}\,\Phi^h_\varepsilon} \big\langle \big(u^\xi_{\tau,x_o}\big)_\varepsilon, \partial_t \Phi^h_\varepsilon \big\rangle \,d\mu\,dt.$$

We consider the second term on the right-hand side of the preceding identity. Since

$$\int_{B_\tau(x_o)} \eta\big(u_\varepsilon(x,t) - \big(u^\xi_{\tau,x_o}\big)_\varepsilon(t)\big)\,d\mu = 0,$$

for almost all $t\in(0,t_1)$, an integration by parts gives

$$\begin{aligned} &\iint_{\mathrm{spt}\,\Phi^h_\varepsilon} \big\langle \big(u^\xi_{\tau,x_o}\big)_\varepsilon, \partial_t \Phi^h_\varepsilon \big\rangle \,d\mu\,dt \\ &\qquad = -\int_0^{t_1} \Big\langle \chi^{h,\varepsilon}_{0,t_1} \partial_t \big(u^\xi_{\tau,x_o}\big)_\varepsilon, \int_{B_\tau(x_o)} \eta\big(u_\varepsilon - \big(u^\xi_{\tau,x_o}\big)_\varepsilon\big)\,d\mu \Big\rangle\,dt = 0, \end{aligned}$$

and hence the second term on the right-hand side of the above identity vanishes. It remains to consider the first term on the right-hand side. By an integration by parts we deduce

$$-\iint_{\mathrm{spt}\,\Phi^h_\varepsilon} \big\langle u_\varepsilon - \big(u^\xi_{\tau,x_o}\big)_\varepsilon, \partial_t \Phi^h_\varepsilon \big\rangle \,d\mu\,dt = -\frac{1}{2} \iint_{\mathrm{spt}\,\Phi^h_\varepsilon} \big|u_\varepsilon - \big(u^\xi_{\tau,x_o}\big)_\varepsilon\big|^2 \partial_t\big(\eta\chi^{h,\varepsilon}_{0,t_1}\big)\,d\mu\,dt.$$

Now we use that $\chi^{h,\varepsilon}_{0,t_1} \to \chi^h_{0,t_1}$ uniformly as $\varepsilon\downarrow 0$ and moreover $(u-u^\xi_{\tau,x_o})_\varepsilon \to u-u^\xi_{\tau,x_o}$ strongly in $L^p(0,t_1;\mathcal{N}^{1,p}(B_\tau(x_o);\mathbb{R}^N))$ and by the Sobolev-embedding also strongly in $L^2(B_\tau(x_o)\times(0,t_1))$ (note also that in order to see that $u-u^\xi_{\tau,x_o}\in L^2$ one first replaces u^ξ_{τ,x_o} by u_{τ,x_o} with the help of Lemma 3.13 and may then use the Sobolev embedding since $u-u_{\tau,x_o}\in\mathcal{N}^{1,p}_o(B_\tau(x_o))$ for almost every $t\in(0,t_1)$). We conclude that

$$\mathcal{L}^h_\varepsilon \xrightarrow{\varepsilon\downarrow 0} -\frac{1}{2} \iint_{B_\tau(x_o)\times(0,t_1)} \big|u - u^\xi_{\tau,x_o}(t)\big|^2 \partial_t\big(\eta\chi^h_{0,t_1}\big)\,d\mu\,dt =: \mathcal{L}^h. \tag{4.2}$$

In a next step, we consider the right hand side of (2.10), in particular the expression

$$\mathcal{R}^h_\varepsilon := \mathcal{Q}\mathcal{F}\big[u - \big(\Phi^h_\varepsilon\big)_\varepsilon, \mathrm{spt}\big(\Phi^h_\varepsilon\big)_\varepsilon\big] - \mathcal{F}\big[u, \mathrm{spt}\big(\Phi^h_\varepsilon\big)_\varepsilon\big].$$

Our aim is to perform the limit procedures $\varepsilon\downarrow 0$ and $h\downarrow 0$ and afterwards establish an appropriate estimate for this expression. We first concentrate on the limit $\varepsilon\downarrow 0$ and therefore proceed as follows: In a first step we use the growth condition (2.8) to obtain

$$\begin{aligned}\mathcal{R}^h_\varepsilon \leq -\nu \iint_{\operatorname{spt}(\Phi^h_\varepsilon)_\varepsilon} g^p_u \, d\mu \, dt \\ + \mathcal{Q} L \iint_{\operatorname{spt}(\Phi^h_\varepsilon)_\varepsilon} g^p_{u-(\Phi^h_\varepsilon)_\varepsilon} \, d\mu \, dt + L_1(\mathcal{Q}+1) \big|\operatorname{spt}\big(\Phi^h_\varepsilon\big)_\varepsilon\big|,\end{aligned} \tag{4.3}$$

where $g_{u-(\Phi^h_\varepsilon)_\varepsilon}$ denotes a minimal p-weak upper gradient to $u-(\Phi^h_\varepsilon)_\varepsilon$. Now we focus on the second integral on the right-hand-side, which we abbreviate

$$I := \iint_{\operatorname{spt}(\Phi^h_\varepsilon)_\varepsilon} g^p_{u-(\Phi^h_\varepsilon)_\varepsilon} \, d\mu \, dt.$$

Here we are faced with the problem, that upper gradients of differences of functions cannot be estimated by the difference of the upper gradients. Therefore, we will perform the limit procedure in a series of steps. First we split the domain of integration as follows:

$$\operatorname{spt}\big(\Phi^h_\varepsilon\big)_\varepsilon = S^{h,2\varepsilon}_{t_1} \cup \mathcal{C} S^{h,2\varepsilon}_{t_1},$$

with

$$S^{h,2\varepsilon}_{t_1} := B_\tau \times (h+2\varepsilon, t_1-h-2\varepsilon) \quad \text{and} \quad \mathcal{C} S^{h,2\varepsilon}_{t_1} = \operatorname{spt}\big(\Phi^h_\varepsilon\big)_\varepsilon \setminus \overline{S_{h,2\varepsilon}}.$$

and hence write

$$I \leq \iint_{S^{h,2\varepsilon}_{t_1}} g^p_{u-(\Phi^h_\varepsilon)_\varepsilon} \, d\mu \, dt + \iint_{\mathcal{C} S^{h,2\varepsilon}_{t_1}} g^p_{u-(\Phi^h_\varepsilon)_\varepsilon} \, d\mu \, dt =: I_1 + I_2,$$

with the obvious labeling of the expressions I_1 and I_2. With the help of the subadditivity of the upper gradient we then deduce for I_1:

$$I_1 \leq c(p) \iint_{S^{h,2\varepsilon}_{t_1}} g^p_{u-(u_\varepsilon)_\varepsilon} \, d\mu \, dt + c(p) \iint_{S^{h,2\varepsilon}_{t_1}} g^p_{(u_\varepsilon)_\varepsilon-(\Phi^h_\varepsilon)_\varepsilon} \, d\mu \, dt.$$

Now we observe the following: On the time interval $[h+\varepsilon, t_1-h-\varepsilon]$ we have $\chi^{h,\varepsilon}_{0,t_1} \equiv 1$ and therefore there holds $u_\varepsilon - \Phi^h_\varepsilon \equiv (1-\eta) u_\varepsilon + \eta (u^\xi_{\tau,x_o})_\varepsilon$. On the other hand, we conclude that

$$\tilde{g}_{u_\varepsilon - \Phi^h_\varepsilon} \equiv (1-\eta) g_{u_\varepsilon} + g_\eta \big| u_\varepsilon - \big(u^\xi_{\tau,x_o}\big)_\varepsilon \big| \tag{4.4}$$

is a p-weak upper gradient of $u_\varepsilon - \Phi^h_\varepsilon$, not necessarily the minimal one. The reason for this is the following: The function $\zeta(t)$ is Lipschitz and therefore, since

$\eta(x,t)=\xi(x)\zeta(t)$, by Remark 3.1, g_η is a p-weak upper gradient to η. Moreover, by Lemma 3.9 we have $g_{(u^\eta_{\tau,x_o})_\varepsilon}\equiv 0$ since $(u^\eta_{\tau,x_o})_\varepsilon$ is a function only in the time variable. Consequently, the assertion (4.4) follows by Lemma 3.6. Applying Lemma 3.11, we therefore conclude that

$$\begin{aligned}\iint_{S^{h,2\varepsilon}_{t_1}} g^p_{(u_\varepsilon)_\varepsilon-(\Phi^h_\varepsilon)_\varepsilon}\,d\mu\,dt &\le \iint_{S^{h,2\varepsilon}_{t_1}} (g_{u_\varepsilon-\Phi^h_\varepsilon})^p_\varepsilon\,d\mu\,dt\\ &\le \iint_{S^{h,\varepsilon}_{t_1}} g^p_{u_\varepsilon-\Phi^h_\varepsilon}\,d\mu\,dt\\ &\le c\iint_{S^{h,\varepsilon}_{t_1}} (1-\eta)^p g^p_{u_\varepsilon}+g^p_\eta\big|u_\varepsilon-\big(u^\xi_{\tau,x_o}\big)_\varepsilon\big|^p\,d\mu\,dt.\end{aligned}$$

To estimate I_2, we first remark that

$$\mathcal{C}S^{h,2\varepsilon}_{t_1}\subset B_\tau\times\left(\left[\frac{h}{10}-2\varepsilon,h+2\varepsilon\right]\cup\left[t_1-h-2\varepsilon,t_1-\frac{h}{10}+2\varepsilon\right]\right)$$

and use the fact that by Lemma 3.3 holds

$$g_{u-(\Phi^h_\varepsilon)_\varepsilon}\le g_u+g_{(\Phi^h_\varepsilon)_\varepsilon}.$$

In order to estimate the integral containing the minimal upper gradient $g_{(\Phi^h_\varepsilon)_\varepsilon}$, we want to show that

$$\tilde g_{\Phi^h_\varepsilon}:=g_\eta\big|u_\varepsilon-\big(u^\eta_{\tau,x_o}\big)_\varepsilon\big|\chi^{h,\varepsilon}_{0,t_1}+\eta g_{u_\varepsilon}\chi^{h,\varepsilon}_{0,t_1}\tag{4.5}$$

is a p-weak upper gradient of Φ^h_ε. For this aim, we use Lemma 3.2. Let Ω' be an open set such that $\overline{\Omega}'\subset\Omega$. First, we note that $\tilde g_{\Phi^h_\varepsilon}\ge 0$ and moreover $\tilde g_{\Phi^h_\varepsilon}(\cdot,t)\in L^p(\Omega')$ for almost every $t\in(0,T)$, since $u_\varepsilon(\cdot,t)\in L^p_{\mathrm{loc}}(\Omega)$, $g_{u_\varepsilon}\in L^p_{\mathrm{loc}}(\Omega)$, $\chi^{h,\varepsilon}_{0,t_1}\le 1$ and $g_\eta\le 2/(\tau-\sigma)$. In order to conclude by Lemma 3.2 that $\tilde g_{\Phi^h_\varepsilon}$ is a p-weak upper gradient of Φ^h_ε, we have to show that for almost every t and for p-almost every curve $\gamma\colon[0,\ell_\gamma]\to\Omega'$, there holds

$$\big|\big(\Phi^h_\varepsilon\circ\gamma\big)'(s,t)\big|\le\tilde g_{\Phi^h_\varepsilon}\big(\gamma(s),t\big),\quad\text{for a.e. } s\in[0,\ell_\gamma].\tag{4.6}$$

We denote by $g_\eta(x,t)$ and $g_{u_\varepsilon}(x,t)$ the minimal p-weak upper gradients of η and u_ε, respectively. Then by Lemma 3.2 we know that for almost every $t\in(0,T)$ and almost every curve $\gamma\colon[0,\ell_\gamma]\to\Omega'$ there holds

$$\big|(\eta\circ\gamma)'(s,t)\big|\le g_\eta\big(\gamma(s),t\big)\quad\text{and}\quad\big|(u_\varepsilon\circ\gamma)'(s,t)\big|\le g_{u_\varepsilon}\big(\gamma(s),t\big),\ \text{for a.e. } s\in[0,\ell_\gamma],$$

and, since both $u_\varepsilon, \eta \in \mathcal{N}^{1,p}(\Omega')$, the mappings $u_\varepsilon \circ \gamma$ and $\eta \circ \gamma$ are absolutely continuous. Then, we calculate

$$\begin{aligned}(\Phi_\varepsilon^h \circ \gamma)'(s,t) &= \frac{d}{ds}\big[\eta\big(\gamma(s),t\big)\big[u_\varepsilon\big(\gamma(s),t\big) - \big(u^\xi_{\tau,x_o}\big)_\varepsilon(t)\big]\chi^{h,\varepsilon}_{0,t_1}\big] \\ &= \chi^{h,\varepsilon}_{0,t_1}(t)\big[(\eta\circ\gamma)'(s,t)\big[u_\varepsilon\big(\gamma(s),t\big) - \big(u^\xi_{\tau,x_o}\big)_\varepsilon(t)\big] \\ &\quad + \eta\big(\gamma(s),t\big)(u_\varepsilon\circ\gamma)'(s,t)\big].\end{aligned}$$

As a consequence we get

$$\begin{aligned}\big|\big(\Phi_\varepsilon^h\circ\gamma\big)'(s,t)\big| &\le \chi^{h,\varepsilon}_{0,t_1}(t)\big[\big|(\eta\circ\gamma)'(s,t)\big|\big|(u_\varepsilon\circ\gamma)(s,t) - \big(u^\xi_{\tau,x_o}\big)_\varepsilon(t)\big| \\ &\quad + \big|(\eta\circ\gamma)(s,t)\big|\big|(u_\varepsilon\circ\gamma)'(s,t)\big|\big] \\ &\le \chi^{h,\varepsilon}_{0,t_1}(t)\big(g_\eta\big|u_\varepsilon - \big(u^\xi_{\tau,x_o}\big)_\varepsilon\big| + \eta g_{u_\varepsilon}\big)\circ\gamma(s,t).\end{aligned}$$

Hence, we conclude (4.6) and therefore (4.5) again by Lemma 3.2. Having this at hand, we get by the properties of the mollification and with Lemma 3.11:

$$\begin{aligned}&\iint_{\mathcal{C}S^{h,2\varepsilon}_{t_1}} g^p_{u-(\Phi^h_\varepsilon)_\varepsilon}\,d\mu\,dt \\ &\quad \le c\iint_{\mathcal{C}S^{h,2\varepsilon}_{t_1}}\big(g^p_u + g^p_{(\Phi^h_\varepsilon)_\varepsilon}\big)\,d\mu\,dt \\ &\quad \le c\iint_{\mathcal{C}S^{h,2\varepsilon}_{t_1}} g^p_u\,d\mu\,dt + c\iint_{B_\tau\times[\frac{h}{10}-4\varepsilon,h+4\varepsilon]} g^p_{\Phi^h_\varepsilon}\,d\mu\,dt \\ &\qquad + c\iint_{B_\tau\times[t_1-h-4\varepsilon,t_1-\frac{h}{10}+4\varepsilon]} g^p_{\Phi^h_\varepsilon}\,d\mu\,dt \\ &\quad \le c\iint_{\mathcal{C}S^{h,2\varepsilon}_{t_1}} g^p_u\,d\mu\,dt + c\iint_{B_\tau\times([\frac{h}{10}-4\varepsilon,h+4\varepsilon]\cup[t_1-h-4\varepsilon,t_1-\frac{h}{10}+4\varepsilon])} g^p_{\Phi^h_\varepsilon}\,d\mu\,dt.\end{aligned}$$

Now we combine all the estimates from before and additionally note that

$$\operatorname{spt}\big(\Phi^h_\varepsilon\big)_\varepsilon \subset Q^{(\lambda)}_\tau(z_o)\cap\{h-2\varepsilon\le t\le t_1-h+2\varepsilon\}\subset Q^{(\lambda)}_\tau(z_o)\cap\{t\le t_1\},$$

to conclude that

$$\begin{aligned}&-\mathcal{L}^h_\varepsilon + \nu\iint_{Q^{(\lambda)}_\sigma(z_o)\cap\{h-2\varepsilon\le t\le t_1-h+2\varepsilon\}} g^p_u\,d\mu\,dt \\ &\qquad \le L_1(\mathcal{Q}+1)\big|Q^{(\lambda)}_\tau(z_o)\cap\{t\le t_1\}\big| + c(p)L\mathcal{Q}\iint_{S^{h,2\varepsilon}_{t_1}} g^p_{u-(u_\varepsilon)_\varepsilon}\,d\mu\,dt\end{aligned}$$

$$
\begin{aligned}
&+c(p)L\mathcal{Q}\iint_{S_{t_1}^{h,\varepsilon}}(1-\eta)g_{u_\varepsilon}^p+g_\eta^p\big|u_\varepsilon-\big(u_{\tau,x_o}^{\xi}\big)_\varepsilon\big|^p\,d\mu\,dt\\
&+c(p)L\mathcal{Q}\iint_{\mathcal{C}S_{t_1}^{h,2\varepsilon}}g_u^p\,d\mu\,dt\\
&+c(p)L\mathcal{Q}\iint_{B_\tau\times([\frac{h}{10}-4\varepsilon,h+4\varepsilon]\cup[t_1-h-4\varepsilon,t_1-\frac{h}{10}+4\varepsilon])}\tilde g_{\Phi_\varepsilon^h}^p\,d\mu\,dt. \qquad (4.7)
\end{aligned}
$$

The integral

$$
\iint_{S_{t_1}^{h,2\varepsilon}}g_{u-(u_\varepsilon)_\varepsilon}^p\,d\mu\,dt
$$

tends to zero in the limit $\varepsilon\downarrow 0$ since

$$
0\le g_{u-(u_\varepsilon)_\varepsilon}\le g_{u-u_\varepsilon}+g_{u_\varepsilon-(u_\varepsilon)_\varepsilon}\le g_{u-u_\varepsilon}+(g_{u-u_\varepsilon})_\varepsilon\to 0,
$$

in $L^p(B_\tau\times(0,t_1))$, and here we have used the subadditivity of the upper gradient, then Lemma 3.11 and finally Lemma 3.12. Next, we consider the second integral on the right-hand-side of (4.7). Since $\eta\equiv 0$ on $Q_\tau^{(\lambda)}(z_o)\setminus Q_\sigma^{(\lambda)}(z_o)$, we obtain by the properties of the mollification end subsequently enlarging the domain of integration:

$$
\begin{aligned}
\int_{S_{t_1}^{h,\varepsilon}}(1-\eta)g_{u_\varepsilon}^p\,d\mu\,dt&\le\iint_{S_{t_1}^{h,\varepsilon}\cap(Q_\tau^{(\lambda)}(z_o)\setminus Q_\sigma^{(\lambda)}(z_o))}g_{u_\varepsilon}^p\,d\mu\,dt\\
&\le\iint_{S_{t_1}^{h,\varepsilon}\cap(Q_\tau^{(\lambda)}(z_o)\setminus Q_\sigma^{(\lambda)}(z_o))}(g_u)_\varepsilon^p\,d\mu\,dt\\
&\le\iint_{S_{t_1}^{h}\cap(Q_\tau^{(\lambda)}(z_o)\setminus Q_\sigma^{(\lambda)}(z_o))}g_u^p\,d\mu\,dt,
\end{aligned}
$$

where we introduced the short-hand notation $S_{t_1}^h\equiv S_{t_1}^{h,0}=B_\tau(x_o)\times[h,t_1-h]$. Moreover, by the strong convergence $u_\varepsilon\to u$, $u_{\tau,\varepsilon}^\eta\to u_\tau^\eta$ in $L^p(S_h)$, we deduce

$$
\iint_{S_{t_1}^{h,\varepsilon}}g_\eta^p\big|u_\varepsilon-\big(u_{\tau,x_o}^{\xi}\big)_\varepsilon\big|^p\,d\mu\,dt\longrightarrow\int_{S_{t_1}^h}g_\eta^p\big|u-u_{\tau,x_o}^{\xi}\big|^p\,d\mu\,dt,
$$

as $\varepsilon\downarrow 0$ and hence, combining the last two observations, we get

$$
\begin{aligned}
&\limsup_{\varepsilon\downarrow 0}\iint_{S_{t_1}^{h,\varepsilon}}(1-\eta)g_{u_\varepsilon}^p+g_\eta^p\big|u_\varepsilon-\big(u_{\tau,x_o}^{\xi}\big)_\varepsilon\big|^p\,d\mu\,dt\\
&\qquad\le\int_{S_{t_1}^{h}\cap(Q_\tau^{(\lambda)}(z_o)\setminus Q_\sigma^{(\lambda)}(z_o))}g_u^p\,d\mu\,dt+\iint_{S_{t_1}^h}g_\eta^p\big|u-u_{\tau,x_o}^{\xi}\big|^p\,d\mu\,dt.
\end{aligned}
$$

Finally we investigate the last term on the right-hand-side of (4.7). Therefore we recall that

$$\tilde{g}_{\Phi^h_\varepsilon} = g_\eta \big| u_\varepsilon - \big(u^\xi_{\tau,x_o}\big)_\varepsilon \big| \chi^{h,\varepsilon}_{0,t_1} + \eta g_{u_\varepsilon} \chi^{h,\varepsilon}_{0,t_1}.$$

The first term converges strongly in $L^p(B_\tau(x_o)\times(0,t_1))$ to $g_\eta|u-u^\xi_{\tau,x_o}|\chi^h_{0,t_1}$, since $u_\varepsilon-(u^\xi_{\tau,x_o})_\varepsilon \to u-u^\xi_{\tau,x_o}$ strongly in L^p, and $\chi^{h,\varepsilon}_{0,t_1}\to\chi^h_{0,t_1}$ uniformly. Moreover, since $\chi^{h,\varepsilon}_{0,t_1}\le 1$ and $\eta\le 1$, we get for the second term, abbreviating for a short moment $\mathcal{B}:=B_\tau(x_o)\times([h/10-4\varepsilon,h+4\varepsilon]\cup[t_1-h-4\varepsilon,t_1-h/10+4\varepsilon])$:

$$\begin{aligned}\iint_{\mathcal{B}} \eta^p g^p_{u_\varepsilon} \big(\chi^{h,\varepsilon}_{0,t_1}\big)^p \,d\mu\,dt &\le \iint_{\mathcal{B}} g^p_{u_\varepsilon}\,d\mu\,dt \\ &\le \iint_{B_\tau(x_o)\times([\frac{h}{10}-5\varepsilon,h+5\varepsilon]\cup[t_1-h-5\varepsilon,t_1-\frac{h}{10}+5\varepsilon])} g^p_u\,d\mu\,dt,\end{aligned}$$

so that

$$\begin{aligned}\limsup_{\varepsilon\downarrow 0} &\iint_{B_\tau(x_o)\times([\frac{h}{10}-4\varepsilon,h+4\varepsilon]\cup[t_1-h-4\varepsilon,t_1-\frac{h}{10}+4\varepsilon])} \tilde{g}^p_{\Phi^h_\varepsilon}\,d\mu\,dt \\ &\le \iint_{B_\tau(x_o)\times([\frac{h}{10},h]\cup[t_1-h,t_1-\frac{h}{10}])} \big(g^p_\eta \big|u-u^\xi_{\tau,x_o}\big|^p + g^p_u\big)\,d\mu\,dt.\end{aligned}$$

Combining all these estimates and letting $\varepsilon\downarrow 0$, we therefore conclude the following estimate:

$$\begin{aligned}\mathcal{L}^h + \nu &\iint_{Q^{(\lambda)}_\sigma(z_o)\cap\{h\le t\le t_1-h\}} g^p_u\,d\mu\,dt \\ &\le cL\mathcal{Q}\iint_{S^h_{t_1}\cap(Q^{(\lambda)}_\tau(z_o)\setminus Q^{(\lambda)}_\sigma(z_o))} g^p_u\,d\mu\,dt \\ &\quad + cL\mathcal{Q}\iint_{S^h_{t_1}} g^p_\eta\big|u-u^\xi_\tau\big|^p\,d\mu\,dt \\ &\quad + cL\mathcal{Q}\iint_{B_\tau(x_o)\times([\frac{h}{10},h]\cup[t_1-h,t_1-\frac{h}{10}])} \big(g^p_u + g^p_\eta\big|u-u^\xi_\tau\big|^p\big)\,d\mu\,dt \\ &\quad + L_1(\mathcal{Q}+1)\big|Q^{(\lambda)}_\tau(z_o)\cap\{t\le t_1\}\big|, \end{aligned} \tag{4.8}$$

with $c\equiv c(p)$. Finally we perform the limit $h\downarrow 0$. We first consider the term

$$\begin{aligned}\mathcal{L}^h = &-\frac{1}{2}\iint_{B_\tau(x_o)\times(0,t_1)} \big|u-u^\xi_{\tau,x_o}\big|^2 \partial_t\eta\chi^h_{0,t_1}\,d\mu\,dt \\ &-\frac{1}{2}\iint_{B_\tau(x_o)\times(0,t_1)} \big|u-u^\xi_{\tau,x_o}\big|^2 \eta\partial_t\chi^h_{0,t_1}\,d\mu\,dt.\end{aligned}$$

For the first expression on the right-hand-side we get

$$-\frac{1}{2}\iint_{B_\tau(x_o)\times(0,t_1)}\big|u-u^{\xi}_{\tau,x_o}\big|^2\partial_t\eta\chi^h_{0,t_1}\,d\mu\,dt \geq -\frac{1}{2}\iint_{B_\tau(x_o)\times(0,t_1)}\big|u-u^{\xi}_{\tau,x_o}\big|^2|\partial_t\eta|\,d\mu\,dt.$$

For the second expression we have

$$-\frac{1}{2}\iint_{B_\tau(x_o)\times(0,t_1)}\big|u-u^{\xi}_{\tau,x_o}\big|^2\eta\partial_t\chi^h_{0,t_1}\,d\mu\,dt = -\frac{5}{9h}\int_{\frac{h}{10}}^{h}\int_{B_\tau(x_o)}\big|u-u^{\xi}_{\tau,x_o}\big|^2\eta\,d\mu\,dt+\frac{5}{9h}\int_{t_1-h}^{t_1-\frac{h}{10}}\int_{B_\tau(x_o)}\big|u-u^{\xi}_{\tau,x_o}\big|^2\eta\,d\mu\,dt.$$

Since $\operatorname{spt}\eta\subset Q^{(\lambda)}_\tau(z_o)$, the first term is identically zero, if h is small enough. The second term converges to

$$\frac{1}{2}\int_{B_\tau(x_o)}\big|u(\cdot,t_1)-u^{\xi}_{\tau,x_o}(t_1)\big|^2\eta(\cdot,t_1)\,d\mu,$$

for almost all $t_1\in\Lambda^{(\lambda)}_\sigma(t_o)$. Alltogether we have seen:

$$\liminf_{h\downarrow 0}\mathcal{L}^h \geq -\frac{1}{2}\iint_{B_\tau(x_o)\times(0,t_1)}\big|u-u^{\xi}_{\tau,x_o}\big|^2|\partial_t\eta|\,d\mu\,dt + \frac{1}{2}\int_{B_\tau(x_o)}\big|u(\cdot,t_1)-u^{\xi}_{\tau,x_o}(t_1)\big|^2\eta(\cdot,t_1)\,d\mu.$$

It remains to consider the right-hand-side of (4.8). Here we argue by the continuous dependence on the domain of integration, since the integrands do not depend on h and therefore we finally obtain

$$\begin{aligned}\int_{B_\tau(x_o)}\big|u(\cdot,t_1)-u^{\xi}_{\tau,x_o}(t_1)\big|^2\eta(\cdot,t_1)\,d\mu+\nu\iint_{Q^{(\lambda)}_\sigma(z_o)\cap\{t\leq t_1\}} g_u^p\,d\mu\,dt \\ \leq c(p)L\mathcal{Q}\iint_{(Q^{(\lambda)}_\tau(z_o)\setminus Q^{(\lambda)}_\sigma(z_o))\cap\{t\leq t_1\}} g_u^p\,d\mu\,dt \\ +c(p)L\mathcal{Q}\iint_{Q^{(\lambda)}_\tau(z_o)} g_\eta^p\big|u-u^{\xi}_{\tau,x_o}\big|^p\,d\mu\,dt \\ +c(p)\iint_{Q^{(\lambda)}_\tau(z_o)}|\partial_t\eta|^p\big|u-u^{\xi}_{\tau,x_o}\big|^2\,d\mu\,dt \\ +L_1(\mathcal{Q}+1)\big|Q^{(\lambda)}_\tau(z_o)\cap\{t\leq t_1\}\big|,\end{aligned}$$

for almost all $t_1 \in \Lambda_\sigma^{(\lambda)}(t_o)$. Now we use the properties of the cut-off functions (3.11) and (3.14) together with $\xi, \zeta \le 1$ to conclude the estimate

$$\begin{aligned}
\int_{B_\tau(x_o)} \big|u(\cdot,t_1)-u^\xi_{\tau,x_o}(t_1)\big|^2 \eta(\cdot,t_1)\,d\mu + \nu \iint_{Q_\sigma^{(\lambda)}(z_o)\cap\{t\le t_1\}} g_u^p\,d\mu\,dt \\
\le c(p) L\mathcal{Q} \iint_{(Q_\tau^{(\lambda)}(z_o)\setminus Q_\sigma^{(\lambda)}(z_o))} g_u^p\,d\mu\,dt + L_1(\mathcal{Q}+1)\big|Q_\tau^{(\lambda)}(z_o)\big| \\
+ L\mathcal{Q}\iint_{Q_\tau^{(\lambda)}(z_o)} \left|\frac{u-u^\xi_{\tau,x_o}}{\tau-\sigma}\right|^p d\mu\,dt + c\lambda^{p-2}\iint_{Q_\tau^{(\lambda)}(z_o)} \left|\frac{u-u^\xi_{\tau,x_o}}{\tau-\sigma}\right|^2 d\mu\,dt,
\end{aligned}$$

which holds for almost every $t_1 \in \Lambda_\sigma^{(\lambda)}(t_o)$, with a constant $c \equiv c(p)$. Since the right hand side of the estimate does not depend on t_1, we may pass on the left hand side to the supremum over $t_1 \in \Lambda_\sigma^{(\lambda)}$. On the other hand we may replace the weighted mean values $u_\tau^\eta(t)$ by the standard mean values $u_\tau(t)$ with respect to the space variable by Lemma 3.13, and this is the point where the dependence of the constant on c_d and therefore on n comes into play. Hence, the proof is complete. □

The Pre-Caccioppoli inequality implies a Caccioppoli inequality for weak upper gradients. The proof of this consequence works exactly as in the Euclidean case and we therefore refer the reader to [16].

Lemma 4.2. (Caccioppoli inequality) *Under the assumptions of Proposition 4.1 there exist two constants $c_1 \equiv c_1(\mathrm{n}, L/\nu, \mathcal{Q})$ and $c_2 \equiv c_2(\mathrm{n}, L/\nu, L_1/\nu, \mathcal{Q})$ such that for every concentric pair of scaled parabolic cylinders $Q_\varrho^{(\lambda)}(z_o) \subset Q_R^{(\lambda)}(z_o) \subset \Omega_T$ with scaling parameter $\lambda > 0$ such that the following estimate for every p-weak minimal upper gradient g_u of u holds*:

$$\iint_{Q_\varrho^{(\lambda)}(z_o)} g_u^p\,d\mu\,dt \le c_1 \iint_{Q_R^{(\lambda)}(z_o)} \left|\frac{u-u_R(t)}{R-\varrho}\right|^p + \lambda^{p-2}\left|\frac{u-u_R(t)}{R-\varrho}\right|^2 d\mu\,dt + c_2\big|Q_R^{(\lambda)}(z_o)\big|. \tag{4.9}$$

Secondly, Proposition 4.1 implies a Poincaré type inequality on intrinsic parabolic cylinders. However for this estimate to hold, it is necessary that the metric measure space itself supports a $(1,p)$-Poincaré inequality.

Lemma 4.3. (Poincaré inequality on intrinsic cylinders) *Under the hypothesis of Proposition 4.1 on the metric measure space which supports a $(1,p)$-Poincaré inequality, for every fixed $\varkappa \ge 1$ there exists a constant $c \equiv c(\mathrm{n}, N, \nu, L, L_1, p, \varkappa, \mathcal{Q}, \Lambda)$*

such that for every parabolic cylinder $Q^{(\lambda)}_{2\Gamma\varrho}\subset\Omega_T$ *with scaling parameter* $\lambda>0$*, on which*

$$\tag{4.10} \fint\!\!\!\fint_{Q^{(\lambda)}_{2\Gamma\varrho}(z_o)} (1+g_u)^p \, d\mu\, dt \le \varkappa\lambda^p,$$

is satisfied, the estimate

$$\tag{4.11} \sup_{t\in\Lambda^{(\lambda)}_\varrho(t_o)} \fint_{B_\varrho(x_o)} \left|\frac{u(\cdot,t)-u_\varrho(t)}{\varrho}\right|^2 d\mu \le c\lambda^2$$

holds true.

Proof. We basically follow the strategy in [16]. In the case $p\ge 2$, we apply the estimate of Proposition 4.1 with the choices $\sigma:=\varrho$ and $\tau:=2\varrho$ to obtain

$$\begin{aligned}\sup_{t\in\Lambda^{(\lambda)}_\varrho(t_o)} \fint_{B_\varrho} \left|u(\cdot,t)-u^{\xi}_{2\varrho,x_o}(t)\right|^2 d\mu \\ &\le c \fint\!\!\!\fint_{Q^{(\lambda)}_{2\varrho}(z_o)} (1+g_u)^p \, d\mu\, dt \\ &\quad + c \fint\!\!\!\fint_{Q^{(\lambda)}_{2\varrho}(z_o)} \left|\frac{u-u_{2\varrho,x_o}(t)}{\varrho}\right|^p + \lambda^{p-2}\left|\frac{u-u_{2\varrho,x_o}(t)}{\varrho}\right|^2 d\mu\, dt. \end{aligned}\tag{4.12}$$

Now we recall that the metric space supports a weak $(1,p)$-Poincaré inequality, and by the Sobolev embedding theorem also a weak $(1,q)$-Poincaré inequality for all exponents $1\le q\le p^*=\frac{\mathrm{n}p}{\mathrm{n}-p}$. In particular, we have

$$\fint_{B_{2\varrho}(x_o)} |u-u_{2\varrho,x_o}|^p \, d\mu \le c_P\varrho^p \fint_{B_{2\Gamma\varrho}(x_o)} g_u^p \, d\mu,$$

on every time slice t. Hence, using also (4.10), we can estimate the integral on the right hand side above in the following way:

$$\begin{aligned}\iint_{Q^{(\lambda)}_{2\varrho}(z_o)} \left|\frac{u-u_{2\varrho,x_o}(t)}{\varrho}\right|^p d\mu\, dt &= \int_{\Lambda^{(\lambda)}_{2\varrho}(t_o)} \mu\big(B_{2\varrho}(x_o)\big) \fint_{B_{2\varrho}(x_o)} \left|\frac{u-u_{2\varrho,x_o}}{\varrho}\right|^p d\mu\, dt \\ &\le c_P\mu\big(B_{2\varrho}(x_o)\big) \int_{\Lambda^{(\lambda)}_{2\varrho}(t_o)} \fint_{B_{2\Gamma\varrho}(x_o)} g_u^p \, d\mu\, dt \\ &\le c_P\Gamma^2\varkappa\big(\mu\otimes\mathcal{L}^1\big)\big(Q^{(\lambda)}_{2\varrho}(z_o)\big)\lambda^p.\end{aligned}$$

On the other hand we have, using Hölder's inequality:

$$\iint_{Q_{2\varrho}^{(\lambda)}(z_o)} \left|\frac{u-u_{2\varrho,x_o}(t)}{\varrho}\right|^2 d\mu\, dt$$
$$\le \left(\mu\otimes\mathcal{L}^1\right)\left(Q_{2\varrho}^{(\lambda)}(z_o)\right)\left[\fint_{Q_{2\varrho}^{(\lambda)}(z_o)} \left|\frac{u-u_{2\varrho,x_o}(t)}{\varrho}\right|^p d\mu\, dt\right]^{2/p}$$
$$\le c\left(\mu\otimes\mathcal{L}^1\right)\left(Q_{2\varrho}^{(\lambda)}(z_o)\right)\Gamma^{4/p}\varkappa\lambda^2,$$

with $c\equiv c(N,p,c_P)$. Combining this with (4.12) we arrive at

$$\sup_{t\in\Lambda_\varrho^{(\lambda)}(t_o)} \int_{B_\varrho} \left|u(\cdot,t)-u_{2\varrho,x_o}^{\xi}(t)\right|^2 d\mu \le c\varkappa\left(\mu\otimes\mathcal{L}^1\right)\left(Q_{2\varrho}^{(\lambda)}(z_o)\right)\lambda^p$$
$$= c\varkappa\mu\left(B_{2\varrho}(x_o)\right)\varrho^2\lambda^2,$$

with a constant c which depends on N, c_P, p, $\mathcal{Q}$, L, L_1 and Γ. Now, we use the doubling property of the measure to get $\mu(B_{2\varrho}(x_o))\le 2^{\mathrm{n}}\mu(B_\varrho(x_o))$. Moreover we use the quasi-minimizing property of the mean value $u_{\varrho,x_o}(t)$ in terms of Lemma 3.14 to replace $u_{2\varrho,x_o}^{\xi}(t)$ by $u_{\varrho,x_o}(t)$ on the left hand side of the preceding inequality, then decide the resulting expression by $\mu(B_\varrho(x_o))\varrho^2$ to conclude the desired estimate in the case $p\ge 2$.

In the case $p<2$ the main additional ingredient for the proof is the Gagliardo–Nirenberg inequality (3.10), which allows us to 'reduce' the L^2-norm appearing on the right hand side to the L^p-norm. We proceed as in the Euclidean case and therefore sketch only the basic steps. In a first step, the application of Proposition 4.1 for radii $0<\varrho\le\sigma<\tau\le 2\varrho$ and the quasi-minimizing property of the mean value $u_\sigma(t)$ on every time slice t gives

$$\sup_{t\in\Lambda_\sigma^{(\lambda)}(t_o)} \int_{B_\sigma(x_o)} \left|u(\cdot,t)-u_{\sigma,x_o}(t)\right| d\mu \le c\iint_{Q_{2\Gamma\varrho}^{(\lambda)}(z_o)} (1+g_u)^p\, d\mu\, dt + cI_p + cI_2,$$

with

$$I_q := \lambda^{p-q}\iint_{Q_\tau^{(\lambda)}(z_o)} \left|\frac{u-u_{\tau,x_o}(t)}{\tau-\sigma}\right|^q d\mu\, dt,$$

for $q=p$ and $q=2$. For $q=p$ the expression I_p is estimated with the help of the Poincaré inequality and (4.10) as follows:

$$I_p = \frac{\tau^p}{(\tau-\sigma)^p}\iint_{Q_\tau^{(\lambda)}(z_o)} \left|\frac{u-u_{\tau,x_o}(t)}{\tau}\right|^q d\mu\, dt$$
$$\le \frac{2^{-p}\tau^p}{(\tau-\sigma)^p}\int_{\Lambda_\tau^{(\lambda)}(t_o)}\fint_{B_\tau(x_o)} \left|\frac{u-u_{\tau,x_o}(t)}{\tau}\right|^q d\mu\, dt$$

$$
\begin{aligned}
&\le c_P \frac{2^{-p}\tau^p}{(\tau-\sigma)^p}\frac{\mu(B_\tau(x_o))}{\mu(B_{\Gamma\tau}(x_o))}\big(\mu\otimes\mathcal{L}^1\big)\big(Q^{(\lambda)}_{2\Gamma\tau}(z_o)\big)\fint_{Q^{(\lambda)}_{2\Gamma\tau}(z_o)} g_u^p\,d\mu\,dt\\
&\le c_P \frac{2^{-p}\tau^p}{(\tau-\sigma)^p}2^{\mathrm{n}}c_d 2\lambda^{2-p}(2\Gamma\tau)^2\varkappa\lambda^p\\
&\le c_P \frac{2^{-p+3+\mathrm{n}}\tau^{p+2}}{(\tau-\sigma)^p}\mu\big(B_\tau(x_o)\big)\varkappa\Gamma^2\lambda^2 \le c\frac{\varrho^{p+2}}{(\tau-\sigma)^p}\mu\big(B_\varrho(x_o)\big)\varkappa\Gamma^2\lambda^2,
\end{aligned}
$$

with $c\equiv c(\mathrm{n},p,c_P,N)$. Here we have also used twice the doubling property of the measure.

For $q=2$, we apply the Gagliardo–Nirenberg estimate in the version of Lemma 3.10 with the exponents $(\sigma,\tau,r,\theta)=(2,p,2,p/2)$ and arrive at

$$
\begin{aligned}
I_2 &= \frac{\tau^2\lambda^{p-2}(\mu\otimes\mathcal{L}^1)(Q^{(\lambda)}_\tau(z_o))}{(\tau-\sigma)^2}\fint_{\Lambda^{(\lambda)}_\tau(t_o)}\fint_{B_\tau(x_o)}\left|\frac{u-u_{\tau,x_o}(t)}{\tau}\right|^2 d\mu\,dt\\
&\le c\frac{\tau^2\lambda^{p-2}(\mu\otimes\mathcal{L}^1)(Q^{(\lambda)}_\tau(z_o))}{(\tau-\sigma)^2}\\
&\quad\times\fint_{\Lambda^{(\lambda)}_\tau(t_o)}\left[\fint_{B_{2\Gamma\tau}(x_o)} g_u^p\,d\mu\right]\left[\fint_{B_\tau(x_o)}\left|\frac{u-u_{\tau,x_o}(t)}{\tau}\right|^2 d\mu\right]^{1-\frac{p}{2}}dt\\
&\le c\frac{\tau^2\lambda^{p-2}(\mu\otimes\mathcal{L}^1)(Q^{(\lambda)}_\tau(z_o))}{(\tau-\sigma)^2}\frac{|\Lambda^{(\lambda)}_{2\Gamma\tau}|}{|\Lambda^{(\lambda)}_\tau|}\left[\sup_{t\in\Lambda^{(\lambda)}_\tau(t_o)}\fint_{B_\tau(x_o)}\left|\frac{u-u_{\tau,x_o}(t)}{\tau}\right|^2 d\mu\right]^{1-\frac{p}{2}}\varkappa\lambda^p\\
&\le c\frac{\varrho^{p+2}\mu(B_\tau(x_o))^{p/2}}{(\tau-\sigma)^2}\Gamma^2\left[\sup_{t\in\Lambda^{(\lambda)}_\tau(t_o)}\big|u-u_{\tau,x_o}(t)\big|^2 d\mu\right]^{1-\frac{p}{2}}\varkappa\lambda^p\\
&\le \frac{1}{2}\sup_{t\in\Lambda^{(\lambda)}_\tau(t_o)}\int_{B_\tau(x_o)}\big|u-u_{\tau,x_o}(t)\big|^2 d\mu + c\frac{\varrho^{2+4/p}}{(\tau-\sigma)^{4/p}}\Gamma^{4/p}\mu(B_\tau)\varkappa^{2/p}\lambda^2,
\end{aligned}
$$

with a constant $c\equiv c(\mathrm{n},N,p,c_P)$. Here we have used several times the doubling property of the measure and $\varrho\le\tau\le 2\varrho$. Combining now the estimates for I_p and I_2 and defining

$$
\Phi(s):=\sup_{t\in\Lambda^{(\lambda)}_s(t_o)}\int_{B_s(x_o)}\big|u(\cdot,t)-u_{s,x_o}(t)\big|^2 d\mu,
$$

we get

$$
\Phi(\sigma)\le\frac{1}{2}\Phi(\tau)+c\frac{\mu(B_\varrho(x_o))\varrho^{2+p}}{(\tau-\sigma)^p}\lambda^2+c\frac{\mu(B_\varrho(x_o))\varrho^{2+4/p}}{(\tau-\sigma)^{4/p}}\lambda^2+c\mu\big(B_\varrho(x_o)\big)\varrho^2\lambda^2.
$$

The Iteration Lemma 3.15 and a subsequent division of the resulting inequality by $\varrho^2\mu(B_\varrho)$ then provides the desired Poincaré inequality also in the case $p<2$. $\square$

4.2. Reverse Hölder inequality

We consider for parameters $\lambda\geq 1$ and $\varkappa\geq 1$ concentric parabolic cylinders $Q_\varrho^{(\lambda)}(z_o)\subset Q_{4\Gamma\varrho}^{(\lambda)}(z_o)\subset\Omega_T=\Omega\times(0,T)$, such that

$$\varkappa^{-1}\lambda^p\leq \fint\!\!\!\fint_{Q_\varrho^{(\lambda)}(z_o)}(1+g_u)^p\,d\mu\,dt, \tag{4.13}$$

and

$$\fint\!\!\!\fint_{Q_{4\Gamma\varrho}^{(\lambda)}(z_o)}(1+g_u)^p\,d\mu\,dt\leq\varkappa\lambda^p. \tag{4.14}$$

On cylinders on which this kind of intrinsic coupling holds true, we have the following reverse Hölder inequality:

Lemma 4.4. *On the metric measure space $(\mathcal{X},\mathrm{d},\mu)$ from Theorem 2.4, let $p>\frac{2\mathrm{n}}{\mathrm{n}+2}$ and $u\in L^p_{\mathrm{loc}}(0,T;\mathcal{N}^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R}^N))$ be a parabolic $\mathcal{Q}$-minimizer in the sense of (2.10) under the growth condition (2.8). Let $Q_\varrho^{(\lambda)}(z_o)\subset Q_{4\Gamma\varrho}^{(\lambda)}(z_o)\subset\Omega_T$ be concentric parabolic cylinders for which the intrinsic coupling in terms of (4.13) and (4.14) holds. Then there exists a constant c depending on n, N, L, ν, L_1, p, $\mathcal{Q}$, $\varkappa$ and Γ, such that for all $0<\vartheta\leq\vartheta_1:=\min\{2,p\}\frac{\mathrm{n}+2}{2\mathrm{n}}$ the reverse Hölder estimate*

$$\fint\!\!\!\fint_{Q_\varrho^{(\lambda)}(z_o)}g_u^p\,d\mu\,dt\leq c\left[\fint\!\!\!\fint_{Q_{4\Gamma\varrho}^{(\lambda)}(z_o)}g_u^{p/\vartheta}\,d\mu\,dt+1\right]^\vartheta$$

holds true.

Proof. We start with Lemma 4.2 applied with radii $(\varrho,R)\equiv(\varrho,2\varrho)$, which provides after dividing the inequality by $\mu(B_\varrho(x_o))$ and using $\mu(B_{2\varrho}(x_o))\leq 2^{\mathrm{n}}\mu(B_\varrho(x_o))$, the following estimate:

$$\begin{aligned}\fint\!\!\!\fint_{Q_\varrho^{(\lambda)}(z_o)}g_u^p\,d\mu\,dt &\leq c\left[\fint\!\!\!\fint_{Q_{2\varrho}^{(\lambda)}(z_o)}\left|\frac{u-u_{2\varrho,x_o}(t)}{\varrho}\right|^p+\lambda^{p-2}\left|\frac{u-u_{2\varrho,x_o}(t)}{\varrho}\right|^2d\mu\,dt+1\right]\\ &=: c(I_p+I_2+1).\end{aligned} \tag{4.15}$$

Here we define

$$I_q:=\lambda^{p-q}\fint\!\!\!\fint_{Q_{2\varrho}^{(\lambda)}(z_o)}\left|\frac{u-u_{2\varrho}(t)}{\varrho}\right|^p d\mu\,dt.$$

Now, we apply the Gagliardo–Nirenberg estimate with the choices $(\sigma,\tau,r,\theta)\equiv(q,p/\vartheta,2,p/(q\vartheta))$ (which is possible by the restriction on ϑ_1) and subsequently the Poincaré inequality in terms of Lemma 4.3 on the pair of cylinders $(Q^{(\lambda)}_{2\varrho}(z_o), Q^{(\lambda)}_{4\Gamma\varrho}(z_o))$ to obtain

$$\begin{aligned} I_q &\le c\lambda^{p-q} \fint_{\Lambda^{(\lambda)}_{2\varrho}(t_o)} \bigg[\fint_{B_{4\Gamma\varrho}(x_o)} g_u^{p/\vartheta}\, d\mu\bigg]\bigg[\fint_{B_{2\varrho}(x_o)}\bigg|\frac{u-u_{2\varrho,x_o}(t)}{\varrho}\bigg|^2 d\mu\bigg]^{\frac{q}{2}(1-\frac{p}{q\vartheta})} dt \\ &\le c\lambda^{p-q}\lambda^{q(1-\frac{p}{q\vartheta})}\frac{|\Lambda^{(\lambda)}_{4\Gamma\varrho}(t_o)|}{|\Lambda^{(\lambda)}_{2\varrho}(t_o)|} \fint\!\!\fint_{Q^{(\lambda)}_{4\Gamma\varrho}(z_o)} g_u^{p/\vartheta}\, d\mu\, dt \\ &\le c\lambda^{p(1-\frac{1}{\vartheta})} \fint\!\!\fint_{Q^{(\lambda)}_{4\Gamma\varrho}(z_o)} g_u^{p/\vartheta}\, d\mu\, dt, \end{aligned}$$

for a constant $c\equiv c(\mathrm{n}, N, L, L_1, p, \mathcal{Q}, \varkappa, \Gamma)$. Inserting the estimate for I_p and I_2 into (4.15) and then applying Young's inequality we arrive at

$$\fint\!\!\fint_{Q^{(\lambda)}_{\varrho}(z_o)} (1+g_u)^p\, d\mu\, dt \le \frac{1}{2\varkappa}\lambda^p + c\bigg[\fint\!\!\fint_{Q^{(\lambda)}_{4\Gamma\varrho}(z_o)} g_u^{p/\vartheta}\, d\mu\, dt + 1\bigg]^{\vartheta}.$$

Using once again (4.13) and absorbing the term $\frac{1}{2}\lambda^p$ into the left hand side, the proof of Lemma 4.4 is complete. □

4.3. Proof of the gradient estimate

Let us fix a cylinder $Q_o\equiv Q_{2r}(\bar z)\subset\Omega_T$ with radius $r>0$ and center $\bar z$ and consider concentric parabolic cylinders of the type

$$\frac{1}{2}Q_o(\bar z)\equiv Q_r(\bar z)\subseteq Q_{r_1}(\bar z)\subset Q_{r_2}(\bar z)\subseteq Q_{2r}(\bar z)\equiv Q_o(\bar z)$$

with radii $r\le r_1<r_2\le 2r$. Furthermore, we define the quantity λ_o by

$$\lambda_o^{\frac{p}{d}} := \fint\!\!\fint_{Q_{2r}(\bar z)} (1+g_u)^p\, d\mu\, dt, \tag{4.16}$$

and note that $\lambda_o\ge 1$. Here we recall the definition of d from (2.12). In the sequel we will often omit the center of the cylinders at points where this is clear from the context. Let $z_o\in Q_{r_1}(\bar z)$ be an arbitrary point and we consider now for a scaling parameter $\lambda\ge 1$ parabolic cylinders $Q^{(\lambda)}_s(z_o)=B_s(x_o)\times\Lambda^{(\lambda)}_s(t_o)$ with radii

$$s\le \min\big\{1, \lambda^{\frac{p-2}{2}}\big\}(r_2-r_1). \tag{4.17}$$

With this restriction, we have $Q^{(\lambda)}_s(z_o)\subseteq Q_{r_2}(\bar z)$, which we can check easily as follows: For any point $z\in Q^{(\lambda)}_s(z_o)$ we have with the notation $z\equiv(x,t)$ that

$$\mathrm{d}(x,x_o)<\min\big\{1,\lambda^{\frac{p-2}{2}}\big\}(r_2-r_1) \quad\text{and}\quad |t-t_o|^{\frac12}<\lambda^{\frac{2-p}{2}}\min\big\{1,\lambda^{\frac{p-2}{2}}\big\}(r_2-r_1).$$

In the case $p \geq 2$ we have that $\min\{1, \lambda^{\frac{p-2}{2}}\} = 1$ and $\lambda^{\frac{2-p}{2}} \leq 1$, because $\lambda \geq 1$, whereas in the case $p<2$ it holds $\min\{1, \lambda^{\frac{p-2}{2}}\} = \lambda^{\frac{p-2}{2}}$ and $\lambda^{\frac{p-2}{2}} \leq 1$. In turn we get in any case $p>1$ that

$$\mathrm{d}_{\mathrm{par}}(z, z_o) \leq \max\big\{\mathrm{d}(x, x_o), \sqrt{|t-t_o|}\big\} < r_2 - r_1,$$

and thus

$$\mathrm{d}_{\mathrm{par}}(z, \bar{z}) \leq \mathrm{d}_{\mathrm{par}}(z, z_o) + \mathrm{d}_{\mathrm{par}}(z_o, \bar{z}) < r_2 - r_1 + r_1 = r_2.$$

So, $z \in Q_{r_2}(\bar{z})$.

We next have to find parabolic cylinders of the above type on which the intrinsic coupling (4.13), (4.14) is satisfied. Since the measure μ is doubling, the Lebesgue differentiation theorem holds on $\mathcal{X} \times \mathbb{R}$ and as $g_u \in L^p_{\mathrm{loc}}(\Omega_T)$ we have for μ-almost every point $z_o \in Q_{r_1}(\bar{z})$ with $g_u(z_o) > \lambda$ that

$$\lim_{s\downarrow 0} \fint\!\!\!\fint_{Q_s^{(\lambda)}(z_o)} (1+g_u)^p \, d\mu \, dt \geq \lim_{s\downarrow 0} \fint\!\!\!\fint_{Q_s^{(\lambda)}(z_o)} g_u^p \, d\mu \, dt = g_u(z_o)^p > \lambda^p. \tag{4.18}$$

To see the equality in (4.18), we have to use once again the doubling property of the measure as follows: For $\tau := \max\{\lambda^{\frac{2-p}{2}}, 1\}$ we have that $Q_s^{(\lambda)}(z_o) \subset Q_{\tau s}(z_o)$. Since $Q_{\tau s}(z_o)$ is a usual parabolic cylinder, the Lebesgue theorem holds on it and gives

$$\lim_{s\downarrow 0} \fint\!\!\!\fint_{Q_{\tau s}(z_o)} g_u^p \, d\mu \, dt = g_u(z_o)^p.$$

Moreover we have

$$\begin{aligned}
&\lim_{s\downarrow 0} \fint\!\!\!\fint_{Q_s^{(\lambda)}(z_o)} g_u^p \, d\mu \, dt \\
&\qquad \leq g_u(z_o)^p + \lim_{s\downarrow 0} \fint\!\!\!\fint_{Q_s^{(\lambda)}(z_o)} \big|g_u^p - g_u(z_o)^p\big| \, d\mu \, dt \\
&\qquad \leq g_u(z_o)^p + \lim_{s\downarrow 0} \frac{(\mu\otimes\mathcal{L}^1)(Q_{\tau s}(z_o))}{(\mu\otimes\mathcal{L}^1)(Q_s^{(\lambda)}(z_o))} \fint\!\!\!\fint_{Q_{\tau s}(z_o)} \big|g_u^p - g_u(z_o)^p\big| \, d\mu \, dt \\
&\qquad \leq g_u(z_o)^p + \lim_{s\downarrow 0} \frac{(\tau s)^2}{s^2 \lambda^{2-p}} \cdot c_d^2 \left(\frac{\tau s}{s}\right)^{\mathrm{n}} \fint\!\!\!\fint_{Q_{\tau s}(z_o)} \big|g_u^p - g_u(z_o)^p\big| \, d\mu \, dt \\
&\qquad = g_u(z_o)^p.
\end{aligned}$$

Now we define

$$\mathcal{B}^{p/d} := 2^{2\mathrm{n}} \left(\frac{80\Gamma r}{r_2 - r_1}\right)^{\mathrm{n}+2} > 1. \tag{4.19}$$

Then for pairs (λ, s) with $\lambda{>}\mathcal{B}\lambda_o$ and

$$\frac{1}{40\Gamma}\min\big\{\lambda^{\frac{p-2}{2}},1\big\}(r_2-r_1)\le s\le \frac{1}{2}\min\big\{\lambda^{\frac{p-2}{2}},1\big\}(r_2-r_1) \tag{4.20}$$

we estimate, using also the doubling property of the measure:

$$\begin{aligned}\iint_{Q_s^{(\lambda)}(z_o)}(1+g_u)^p\,d\mu\,dt &\le \frac{\mu(B_{2r}(x_o))|\Lambda_{2r}(t_o)|}{\mu(B_s(x_o))|\Lambda_s^{(\lambda)}(t_o)|}\iint_{Q_{2r}^{(\lambda)}(z_o)}(1+g_u)^p\,d\mu\,dt\\ &\le 2^{2\mathrm{n}}\left(\frac{2r}{s}\right)^{\mathrm{n}+2}\lambda^{p-2}\lambda_o^{\frac{p}{d}}<\lambda^p.\end{aligned}$$

Here we have used in the very last step the definition of λ_o as follows: In the case $p{\ge}2$ we have that $d{=}p/2$, therefore $p/d{=}2$ and $\min\{\lambda^{p-2},1\}{=}1$. Thus, we get by the definition of $\mathcal{B}$ and with the conditions on s and λ:

$$2^{2\mathrm{n}}\left(\frac{2r}{s}\right)^{\mathrm{n}+2}\lambda^{p-2}\lambda_o^{p/d}\le 2^{2\mathrm{n}}(2r)^{\mathrm{n}+2}\lambda^{p-2}\lambda_o^2\left(\frac{40\Gamma}{r_2-r_1}\right)^{\mathrm{n}+2}\le \mathcal{B}^2\lambda^{p-2}\lambda_o^2<\lambda^p.$$

In the case $p{<}2$, there holds $p/d{=}\frac{p(n+2)}{2}-n$ and $\min\{\lambda^{\frac{p-2}{2}},1\}{=}\lambda^{\frac{p-2}{2}}$ and we therefore get

$$2^{2\mathrm{n}}\left(\frac{2r}{s}\right)^{\mathrm{n}+2}\lambda^{p-2}\lambda_o^{p/d}\le \mathcal{B}^{p/d}\lambda^{\frac{2-p}{2}(n+2)}\lambda^{p-2}\lambda_o^{\frac{p}{2}(n+2)-n}<\lambda^{n\frac{2-p}{2}+2}\lambda_o^{n\frac{p-2}{2}+p-2}<\lambda^p.$$

Combining this estimate with the estimate (4.18) above and having in mind that the integral depends continuously on the radius of the cylinder, there must be at least one radius $0{<}\varrho_{z_o}{<}\frac{1}{40\Gamma}\min\{\lambda^{\frac{p-2}{2}},1\}(r_2-r_1)$ such that

$$\iint_{Q_{\varrho_{z_o}}^{(\lambda)}(z_o)}(1+g_u)^p\,d\mu\,dt=\lambda^p \tag{4.21}$$

holds. In case that there are several such radii we let ϱ_{z_o} be the maximal one, which means that for all radii $\varrho{>}\varrho_{z_o}$ there holds '$<$' instead of '$=$' in (4.21). Hence, the parabolic cylinder $Q_{\varrho_{z_o}}^{(\lambda)}(z_o)$ with $\lambda{>}\mathcal{B}\lambda_o$ is one on which the intrinsic coupling (4.13), (4.14) holds with $\varkappa{=}1$ and consequently the reverse Hölder inequality Lemma 4.4 holds on the pair $(Q_{\varrho_{z_o}}^{(\lambda)}(z_o), Q_{4\Gamma\varrho_{z_o}}^{(\lambda)}(z_o))$ of parabolic cylinders:

$$\iint_{Q_{\varrho_{z_o}}^{(\lambda)}(z_o)}g_u^p\,d\mu\,dt\le c\left[\iint_{Q_{4\Gamma\varrho_{z_o}}^{(\lambda)}(z_o)}(1+g_u)^{\frac{p}{\vartheta}}\,d\mu\,dt\right]^{\vartheta}, \tag{4.22}$$

for all $\vartheta{\le}\vartheta_1$, with $\vartheta_1{=}\min\{2,p\}\frac{\mathrm{n}+2}{2\mathrm{n}}$ and for a constant $c{\equiv}c(\mathrm{n},N,L,\nu,L_1,p,\mathcal{Q},\Gamma)$.

Introducing for $s{>}0$ and $\lambda{\ge}0$ the levelset

$$\mathfrak{S}^s(\lambda):=\big\{z\in Q_s(\bar z)\colon g_u(z)>\lambda\big\},$$

we have up to now shown that in the case $\lambda>\mathcal{B}\lambda_o$ for almost all $z_o\in\mathfrak{S}^{r_1}(\lambda)$ there exists a parabolic cylinder $Q^{(\lambda)}_{\varrho_{z_o}}(z_o)$ on which the intrinsic couplings (4.13) and (4.14) as well as the estimate (4.22) hold true. Moreover, by (4.20) and (4.17) we have that $Q^{(\lambda)}_{20\Lambda\varrho_{z_o}}(z_o)\subset Q_{r_2}(\bar{z})$.

In a next step we want to estimate the L^p-norm of g_u on the cylinder $Q^{(\lambda)}_{20\Gamma\varrho_{z_o}}(z_o)$. For this aim, let $\eta\in(0,1)$ be a parameter which we fix later. We abbreviate for a moment $Q\equiv Q^{(\lambda)}_{\varrho_{z_o}}(z_o)$ and $\alpha Q\equiv Q^{(\lambda)}_{\alpha\varrho_{z_o}}(z_o)$. Then, with (4.22) and assuming that $\eta\lambda\geq 1$, we obtain

$$\begin{aligned}\fint\!\!\fint_{Q}(1+g_u)^p\,d\mu\,dt &\leq c\left[\fint\!\!\fint_{4\Gamma Q}(1+g_u)^{p/\vartheta_1}\,d\mu\,dt\right]^{\vartheta_1}\\ &\leq c\left[\left(\frac{1}{(\mu\otimes\mathcal{L}^1)(4\Gamma Q)}\iint_{4\Gamma Q\cap\mathfrak{S}^{r_2}(\eta\lambda)}g_u^{p/\vartheta_1}\,d\mu\,dt\right)^{\vartheta_1}+(\eta\lambda)^p\right],\end{aligned}$$

with $c\equiv c(\mathrm{n},N,L,\nu,L_1,p,\mathcal{Q},\Gamma)$. By the help of (4.21), choosing $\eta\equiv\eta(\mathrm{n},N,L,\nu,L_1,p,\mathcal{Q},\Gamma)$ in such a way that $c\eta^p\leq\frac{1}{2}$, we may absorb the expression $(\eta\lambda)^p$ into the left hand side of the inequality and therefore obtain

$$\fint\!\!\fint_{Q}(1+g_u)^p\,d\mu\,dt\leq\left[\frac{c}{(\mu\otimes\mathcal{L}^1)(4\Gamma Q)}\iint_{4\Gamma Q\cap\mathfrak{S}^{r_2}(\eta\lambda)}g_u^{p/\vartheta_1}\,d\mu\,dt\right]^{\vartheta_1},$$

This estimate holds for $\lambda>\lambda_1:=\frac{1}{\eta}=(2c)^{-1/p}$, where c denotes the constant appearing above in the choice of η. Splitting the bracket on the right hand side $[...]^{\vartheta_1}=[...][...]^{\vartheta_1-1}$ and using Hölder's inequality to get

$$\begin{aligned}\left[\frac{1}{(\mu\otimes\mathcal{L}^1)(4\Gamma Q)}\iint_{4\Gamma Q\cap\mathfrak{S}^{r_2}(\eta\lambda)}g_u^{p/\vartheta_1}\,d\mu\,dt\right]^{\vartheta_1-1} &\leq\left[\fint\!\!\fint_{4\Gamma Q}g_u^p\,d\mu\,dt\right]^{1-1/\vartheta_1}\\ &\leq c\lambda^{p(1-1/\vartheta_1)},\end{aligned}$$

we arrive at

$$\fint\!\!\fint_{Q}(1+g_u)^p\,d\mu\,dt\leq\frac{c}{(\mu\otimes\mathcal{L}^1)(4\Gamma Q)}\iint_{4\Gamma Q\cap\mathfrak{S}^{r_2}(\eta\lambda)}\lambda^{p(1-1/\vartheta_1)}g_u^{p/\vartheta_1}\,d\mu\,dt.$$

Finally, using the fact that we have '$<$' in (4.21) for the cylinder $Q^{(\lambda)}_{20\Gamma\varrho_{z_o}}(z_o)\equiv 20\Gamma Q$ and therefore getting

$$\fint\!\!\fint_{20\Gamma Q}(1+g_u)^p\,d\mu\,dt\leq\lambda^p\leq\fint\!\!\fint_{Q}(1+g_u)^p\,d\mu\,dt,$$

we conclude that

$$\iint_{20\Gamma Q}g_u^p\,d\mu\,dt\leq c\iint_{4\Gamma Q\cap\mathfrak{S}^{r_2}(\eta\lambda)}\lambda^{p(1-1/\vartheta_1)}g_u^{p/\vartheta_1}\,d\mu\,dt, \tag{4.23}$$

for a constant $c \equiv c(\mathrm{n}, N, L, \nu, L_1, p, \mathcal{Q}, \Gamma)$. Here, we have also used the doubling property of the measure to have $(\mu \otimes \mathcal{L}^1)(20\Gamma Q) \leq 2^{2\mathrm{n}} 5^{\mathrm{n}+2} (\mu \otimes \mathcal{L}^1)(4\Gamma Q)$.

Until now, we have proved the following: For scaling parameters $\lambda > \max\{\mathcal{B}\lambda_o, \lambda_1\}$ the levelset $\mathfrak{S}^{r_1}(\lambda)$ is covered by a family $\mathcal{F} = \{Q^{(\lambda)}_{4\Gamma\varrho_{z_o}}(z_o)\}$ of parabolic cylinders with centers $z_o \in \mathfrak{S}^{r_1}(\lambda)$, on which (4.23) holds. A standard covering theorem provides then a countable subfamily $\{Q^{(\lambda)}_{4\Gamma\varrho_{z_i}}(z_i)\}_{i=1}^{\infty} \subset \mathcal{F}$ of pairwise disjoint cylinders such that the 5-times enlarged cylinders $Q^{(\lambda)}_{20\Gamma\varrho_{z_i}}(z_i)$ still cover $\mathfrak{S}^{r_1}(\lambda)$, and moreover all cylinders $Q^{(\lambda)}_{20\Gamma\varrho_{z_i}}(z_i)$ are contained in $Q_{r_2}(\bar{z})$. As a consequence we have

$$
\begin{aligned}
\iint_{\mathfrak{S}^{r_1}(\lambda)} g_u^p \, d\mu \, dt &\leq \sum_{i=1}^{\infty} \iint_{Q^{(\lambda)}_{20\Gamma\varrho_{z_i}}} g_u^p \, d\mu \, dt \\
&\leq c \sum_{i=1}^{\infty} \iint_{Q^{(\lambda)}_{4\Gamma\varrho_{z_i}} \cap \mathfrak{S}^{r_2}(\eta\lambda)} \lambda^{p(1-1/\vartheta_1)} g_u^{p/\vartheta_1} \, d\mu \, dt \\
&\leq c \iint_{\mathfrak{S}^{r_2}(\eta\lambda)} \lambda^{p(1-1/\vartheta_1)} g_u^{p/\vartheta_1} \, d\mu \, dt.
\end{aligned}
$$

Now, since the previous estimate holds trivially on the set $\mathfrak{S}^{r_1}(\eta\lambda) \setminus \mathfrak{S}^{r_1}(\lambda)$, we conclude that it holds also if we replace on the left hand side the integral over the set $\mathfrak{S}^{r_1}(\lambda)$ by an integral over $\mathfrak{S}^{r_1}(\eta\lambda)$. In a next step, we replace the values $\eta\lambda$ by λ and we end up with the estimate

$$
\iint_{\mathfrak{S}^{r_1}(\lambda)} g_u^p \, d\mu \, dt \leq c \iint_{\mathfrak{S}^{r_2}(\lambda)} \lambda^{p(1-1/\vartheta_1)} g_u^{p/\vartheta_1} \, d\mu \, dt,
$$

which holds for every $\lambda > \max\{\mathcal{B}\eta\lambda_o, \eta\lambda_1\} = \max\{\mathcal{B}\eta\lambda_o, 1\} =: \lambda_2$.

Now, the result would follow by multiplying both sides of the preceding inequality by λ^{ε} and subsequently integrating the levelsets. However, as in the Euclidean case, we are faced with the technical difficulty that the obtained expressions could possibly not be finite. It is then a standard method to introduce truncations $|Du|_k$ to the level k of the gradient—and therefore here truncations $[g_u]_k$ of the upper gradient. For the convenience of the reader, we will shortly sketch the final steps of the proof, but we refer to [16] where the argument is carried out in detail in the Euclidean case. The proof in the metric case does not differ from the Euclidean one, since all higher integrability information is already contained in the previous estimate and the remaining arguments are purely measure theoretic ones, such as integral and convergence theorems. All these theorems hold also for integrals with respect to general Borel regular measures.

We introduce for $k>\lambda_2$ the truncated upper gradient and the corresponding levelset

$$[g_u]_k := \min\{g_u, k\} \quad \text{and} \quad \mathfrak{S}_k^r(\lambda) := \big\{z \in Q_r\colon [g_u]_k > \lambda\big\}.$$

With the notation $q:=p/\vartheta_1<p$, the previous estimate on the levelset implies

$$\iint_{\mathfrak{S}_k^{r_1}(\lambda)} [g_u]_k^{p-q} g_u^q \, d\mu\, dt \le c \iint_{\mathfrak{S}_k^{r_2}(\lambda)} \lambda^{p-q} g_u^q \, d\mu\, dt.$$

Now we multiply both sides of the inequality with $\lambda^{\varepsilon-1}$ for $\varepsilon\in(0,1)$ to be determined later. Integration of the resulting estimate with respect to λ over the interval (λ_2,∞) then provides

$$\int_{\lambda_2}^{\infty} \lambda^{\varepsilon-1} \iint_{\mathfrak{S}_k^{r_1}(\lambda)} [g_u]_k^{p-q} g_u^q \, d\mu\, dt \le c \int_{\lambda_2}^{\infty} \lambda^{\varepsilon-1} \iint_{\mathfrak{S}_k^{r_2}(\lambda)} \lambda^{p-q} g_u^q \, d\mu\, dt.$$

The integral on the left hand side can be calculated with the help of Fubini's theorem and we obtain

$$\int_{\lambda_2}^{\infty} \lambda^{\varepsilon-1} \iint_{\mathfrak{S}_k^{r_1}(\lambda)} [g_u]_k^{p-q} g_u^q \, d\mu\, dt = \frac{1}{\varepsilon} \iint_{\mathfrak{S}_k^{r_1}(\lambda_2)} \big[[g_u]_k^{p-q+\varepsilon} g_u^q - \lambda_2^{\varepsilon} [g_u]_k^{p-q} g_u^q\big] \, d\mu\, dt.$$

Again with Fubini's theorem we get for the right hand side the estimate

$$\int_{\lambda_2}^{\infty} \lambda^{\varepsilon-1} \iint_{\mathfrak{S}_k^{r_2}(\lambda)} \lambda^{p-q} g_u^q \, d\mu\, dt \le \frac{1}{p-q} \iint_{\mathfrak{S}_k^{r_2}(\lambda_2)} [g_u]_k^{p-q+\varepsilon} g_u^q \, d\mu\, dt,$$

and we therefore conclude

$$\begin{aligned} &\iint_{\mathfrak{S}_k^{r_1}(\lambda_2)} [g_u]_k^{p-q+\varepsilon} g_u^q \, d\mu\, dt \\ &\qquad \le \iint_{\mathfrak{S}_k^{r_1}(\lambda_2)} \lambda_2^{\varepsilon} [g_u]_k^{p-q} g_u^q \, d\mu\, dt + \frac{\varepsilon}{p-q} \iint_{\mathfrak{S}_k^{r_2}(\lambda_2)} [g_u]_k^{p-q+\varepsilon} g_u^q \, d\mu\, dt. \end{aligned}$$

Since on $Q_{r_1}(\bar z)\setminus \mathfrak{S}_k^{r_1}(\lambda_2)$ we have $[g_u]_k^{p-q+\varepsilon} g_u^q \le \lambda_2^{\varepsilon} [g_u]_k^{p-q} |g_u|^q$, we eventually arrive at

$$\begin{aligned} &\iint_{Q_{r_1}(\bar z)} [g_u]_k^{p-q+\varepsilon} g_u^q \, d\mu\, dt \\ &\qquad \le \frac{c\varepsilon}{p-q} \iint_{Q_{r_2}(\bar z)} [g_u]_k^{p-q+\varepsilon} g_u^q \, d\mu\, dt + \lambda_2^{\varepsilon} \iint_{Q_{2r}(\bar z)} [g_u]_k^{p-q} g_u^q \, d\mu\, dt, \end{aligned}$$

for a constant c which depends only on n, N, L, ν, p, L_1, Γ and $\mathcal{Q}$. At this point we choose $\varepsilon\le\varepsilon_o:=\frac{1}{2c}(q-p)=\frac{p(\vartheta_1-1)}{2c\vartheta_1}$, where c denotes the constant in the above estimate. Since $\vartheta_1\equiv\vartheta_1(\mathrm{n},p)$, we have that ε_o depends only on the structural parameters n, N, L, ν, L_1, p, Γ and $\mathcal{Q}$. On the other hand we recall that $\lambda_2^\varepsilon = \max\{\mathcal{B}\eta\lambda_o, 1\}$. Since $\mathcal{B}>1$, $\eta\leq 1$, $\varepsilon<1$ and $\lambda_o\geq 1$, we have in any case that $\lambda_2^\varepsilon \leq 1+(\mathcal{B}\eta\lambda_o)^\varepsilon \leq 2\mathcal{B}\lambda_o^\varepsilon$. Recalling now the definition of $\mathcal{B}$ in (4.19), we may apply the standard iteration Lemma 3.15 to conclude

$$\iint_{Q_r(\bar{z})} [g_u]_k^{p-q+\varepsilon} g_u^q \, d\mu \, dt \leq c\lambda_o^\varepsilon \iint_{Q_{2r}(\bar{z})} [g_u]_k^{p-q} g_u^q \, d\mu \, dt.$$

Finally we note that $[g_u]_k \leq g_u$ and $g_u \in L^p(Q_{2r}(\bar{z}))$, so the result follows by applying Fatou's lemma on the left hand side, the dominant convergence theorem on the right hand side and subsequently recalling the definition of λ_o in (4.16).

References

1. Björn, J., Poincaré inequalities for powers and products of admissible weights, *Ann. Acad. Sci. Fenn. Math.* **26** (2002), 175–188.
2. Björn, A. and Björn, J., *Nonlinear Potential Theory on Metric Spaces*, EMS Tracts in Mathematics, Eur. Math. Soc., Zurich, 2011.
3. Björn, A. and Marola, N., Moser iteration for (quasi)minimizers on metric spaces, *Manuscripta Math.* **121** (2006), 339–366.
4. Björn, J. and Shanmugalingam, N., Poincaré inequalities, uniform domains and extension properties for Newton–Sobolev functions in metric spaces, *J. Math. Anal. Appl.* **332** (2007), 190–208.
5. Björn, A., Björn, J. and Shanmugalingam, N., The Dirichlet problem for p-harmonic functions on metric spaces, *J. Reine Angew. Math.* **556** (2003), 173–203.
6. Bögelein, V., Higher integrability for weak solutions of higher order degenerate parabolic systems, *Ann. Acad. Sci. Fenn. Math.* **33** (2008), 387–412.
7. Bögelein, V. and Duzaar, F., Higher integrability for parabolic systems with non-standard growth and degenerate diffusions, *Publ. Mat.* **55** (2011), 201–250.
8. Bögelein, V. and Parviainen, M., Self-improving property of nonlinear higher order parabolic systems near the boundary, *NoDEA Nonlinear Differential Equations Appl.* **17** (2010), 21–54.
9. Cheeger, J., Differentiability of Lipschitz functions on metric spaces, *Geom. Funct. Anal.* **9** (1999), 428–517.
10. DiBenedetto, E., Harnack estimates in certain function classes, *Atti Sem. Mat. Fis. Univ. Modena* **37** (1989), 173–182.
11. DiBenedetto, E., Gianazza, U. and Vespri, V., Local clustering of the non-zero set of functions in $W^{1,1}(E)$, *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* (9) *Mat. Appl.* **917** (2006), 223–225.
12. Fornaro, S., Paronetto, F. and Vespri, V., Disuguaglianza di Harnack per equazioni paraboliche, *Quaderni Dip. Mt. Univ. Lecce*, 2/2008.
13. Franchi, B., Hajłasz, P. and Koskela, P., Definitions of Sobolev classes on metric spaces, *Ann. Inst. Fourier* (*Grenoble*) **49** (1999), 1903–1924.

14. Giusti, E., *Direct Methods in the Calculus of Variations*, World Scientific, Singapore, 2003.
15. Grigor'yan, A. A., The heat equation on noncompact Riemannian manifolds, *Mat. Sb.* **182** (1992), 55–87. Translation in *Math. USSR-Sb.* **72** (1992), 47–77.
16. Habermann, J., Global gradient estimates for vector-valued parabolic quasi-minimizers, *Nonlinear Anal.* **114** (2015), 43–73.
17. Hajłasz, P., Sobolev spaces on an arbitrary metric space, *Potential Anal.* **5** (1996), 403–415.
18. Hajłasz, P., Sobolev spaces on metric measure spaces. Heat kernels and analysis on manifolds, graphs and metric spaces, in *Contemp. Math.* **338**, pp. 173–218, Amer. Math. Soc., Providence, RI, 2003.
19. Hajłasz, P. and Koskela, P., Sobolev meets Poincaré, *C.R. Acad. Sci. Paris Ser. I Math.* **320** (1995), 1211–1215.
20. Hajłasz, P. and Koskela, P., Sobolev met Poincaré, *Mem. Amer. Math. Soc.* **145**, x+101 (2000).
21. Heinonen, J., *Lectures on Analysis on Metric Spaces*, Universitext, Springer, New York, 2001.
22. Heinonen, J. and Koskela, P., Quasiconformal maps in metric spaces with controlled geometry, *Acta Math.* **181** (1998), 1–61.
23. Heinonen, J. and Koskela, P., A note on Lipschitz functions, upper gradients, and the Poincaré inequality, *New Zealand J. Math.* **28** (1999), 37–42.
24. Heinonen, J., Koskela, P., Shanmugalingam, N. and Tyson, J., Sobolev classes of Banach space-valued functions and quasiconformal mappings, *J. Anal. Math.* **85** (2001), 87–139.
25. Keith, S. and Zhong, X., The Poincaré inequality is an open ended condition, *Ann. of Math.* (2) **167** (2008), 575–599.
26. Kilpeläinen, T., Kinnunen, J. and Martio, O., Sobolev spaces with zero boundary values on metric spaces, *Potential Anal.* **12** (2000), 233–241.
27. Kinnunen, J. and Lewis, J., Higher integrability for parabolic systems of p-Laplacian type, *Duke Math. J.* **102** (2000), 253–271.
28. Kinnunen, J. and Martio, O., The Sobolev capacity on metric spaces, *Ann. Acad. Sci. Fenn. Math.* **21** (1996), 367–382.
29. Kinnunen, J. and Shanmugalingam, N., Regularity of quasi-minimizers on metric spaces, *Manuscripta Math.* **105** (2001), 401–423.
30. Kinnunen, J., Marola, N., Miranda jr., M. and Paronetto, F., Parabolic variational problems and regularity in metric spaces, *Adv. Differential Equations* **17** (2012), 801–832.
31. Koskela, P., Shanmugalingam, N. and Tuominen, H., Removable sets for the Poincaré inequality on metric spaces, *Indiana Univ. Math. J.* **49** (2000), 333–352.
32. Kronz, M., Some function spaces and spaces of homogeneous type, *Manuscripta Math.* **106** (2001), 219–248.
33. Laakso, T., Ahlfors Q-regular spaces with arbitrary $Q>1$ admitting weak Poincaré inequality, *Geom. Funct. Anal.* **10** (2000), 111–123.
34. Mäkäläinen, T., Removable sets for Hölder continuous p-harmonic functions on metric measure spaces, *Ann. Acad. Sci. Fenn. Math.* **33** (2008), 605–624.

35. Marola, N. and Masson, M., On the Harnack inequality for parabolic minimizers in metric measure spaces, *Tohoku Math. J.* (2) **65** (2013), 569–589.
36. Masson, M. and Parviainen, M., Global higher integrability of parabolic quasi minimizers in metric spaces, *J. Anal. Math.* **126** (2014), 307–339.
37. Masson, M. and Siljander, J., Hölder regularity for parabolic De Giorgi classes in metric measure spaces, *Manuscripta Math.* **142** (2013), 187–214.
38. Masson, M., Miranda, M. Jr., Paronetto, F. and Parviainen, M., Local higher integrability for parabolic quasi minimizers in metric spaces, *Ric. Mat.* **62** (2013), 279–305.
39. Parviainen, M., Global higher integrability for parabolic quasiminimizers in nonsmooth domains, *Calc. Var. Partial Differential Equations* **31** (2008), 75–98.
40. Saloff-Coste, L., A note on Poincaré, Sobolev and Harnack inequalities, *Int. Math. Res. Not. IMRN* **2** (1992), 27–38.
41. Saloff-Coste, L., *Aspects of Sobolev-Type Inequalities*, London Mathematical Society Lecture Note Series **289**, Cambridge University Press, Cambridge, 2002.
42. Shanmugalingam, N., Newtonian spaces: an extension of Sobolev spaces to metric measure spaces, *Rev. Mat. Iberoam.* **16** (2000), 243–279.
43. Shanmugalingam, N., Harmonic functions on metric spaces, *Illinois J. Math.* **45** (2001), 1021–1050.
44. Wieser, W., Parabolic Q-minima and minimal solutions to variational flow, *Manuscripta Math.* **59** (1987), 63–107.

Jens Habermann
Department Mathematik
Universität Erlangen
Cauerstr. 11
DE-91058 Erlangen
Germany
habermann@math.fau.de

Received January 9, 2015
published online July 23, 2015