# Higher integrability for the gradient of Mumford-Shah *almost*-minimizers

Sebastian Piontek Thomas Schmidt*

*Fachbereich Mathematik, Universität Hamburg,*
*Bundesstr. 55, 20146 Hamburg, Germany*

March 28, 2019

**Abstract**

We extend a recent higher-integrability result for the gradient of minimizers of the Mumford-Shah functional to a suitable class of almost-minimizers. The extension crucially depends on an $\mathrm{L}^\infty$ gradient estimate up to regular portions of the discontinuity set of an almost-minimizer.

# 1 Introduction

Consider a bounded open set $\Omega$ in $\mathbb{R}^n$, $n \geq 2$, a parameter $\lambda \in (0, \infty)$, and a scalar function $g \in \mathrm{L}^\infty(\Omega)$. The Mumford-Shah functional

$$\int_{\Omega\setminus K} |\nabla u|^2 \,\mathrm{d}x + \lambda \int_{\Omega\setminus K} |u - g|^2 \,\mathrm{d}x + \mathcal{H}^{n-1}(K) \tag{1.1}$$

(with the $(n-1)$-dimensional Hausdorff measure $\mathcal{H}^{n-1}$) is defined on pairs $(u, K)$, where $K$ is a closed subset of $\Omega$ and the scalar function $u \in \mathrm{L}^2(\Omega \setminus K)$ has a classical (or weak) derivative $\nabla u \in \mathrm{L}^2(\Omega \setminus K, \mathbb{R}^n)$ on $\Omega \setminus K$, but is allowed to be non-differentiable and discontinuous at points of $K$.

The functional in (1.1) with $n = 2$ has been originally introduced by Mumford–Shah [MS89] in connection with the segmentation of a noisy greyscale image, which can be thought of as a $[0, 1]$-valued $g$. The hope is then that an unconstrained minimizer $(u, K)$ of the functional consists of a denoised version $u$ of the image and, more crucially, of an edge set $K$ which segments the image into comparably homogeneous regions. In addition, the Mumford-Shah functional has also emerged

*Email address: thomas.schmidt@math.uni-hamburg.de.
URL: http://www.math.uni-hamburg.de/home/schmidt/.

into a basic object of theoretical interest, since the interplay between the volume term $\int_{\Omega\setminus K}|\nabla u|^2\,\mathrm{d}x$ and the surface term $\mathcal{H}^{n-1}(K)$ has turned out to be highly non-trivial. This has lead to the development of an elaborate analytical theory (see [DGCL89, Bon96, Dav96, AP97, AFP97, Leg99, AFP00, Rig00, MS01a, MS01b, Fus03, Dav05], for instance), but still some finer issues in the regularity theory of minimizers $(u,K)$ have remained unsolved. The most prominent such issue is certainly the Mumford-Shah conjecture on the precise nature of possible singularities of the edge set $K$ in dimension $n=2$. Here we focus on another such issue, vaguely related[1] to the conjecture, namely on $\mathrm{L}^p$ estimates for $\nabla u$, locally on $\Omega$ but up to $K$. In this regard De Lellis–Focardi [DLF13] (in dimension $n=2$) and subsequently De Phillipis–Figalli [DPF14] (in arbitrary dimension $n\geq 2$) have recently proven slight extra integrability of $\nabla u$. Precisely they proved the existence of some $\varepsilon>0$ such that $\nabla u\in\mathrm{L}^{2+\varepsilon}(\Omega'\setminus K,\mathbb{R}^n)$ holds for every open $\Omega'\Subset\Omega$; see also [Foc16] for a survey on these issues. Here we extend the extra integrability result from minimizers to a suitable class of almost-minimizers (see below and Section 2 for the relevant terminology for SBV functions):

**Theorem 1.1.** *If $u\in\mathrm{SBV}(\Omega)$ is an $(a,b)$-almost-minimizer of the Mumford-Shah functional in the sense of Definition* 3.1*, then there exists some $\varepsilon>0$ such that $\nabla u\in\mathrm{L}^{2+\varepsilon}(\Omega',\mathbb{R}^n)$ holds for every open $\Omega'\Subset\Omega$.*

This extension is in line with the overall Mumford-Shah regularity theory, in which nowadays almost all results are available for almost-minimizers [AP97, AFP97, AFP00, Dav05]. Indeed, the transition to almost-minimizers is technically very convenient also in other variational problems, since minimizers of related problems with additional coefficients, terms, or constraints (compare [Anz83] and [DGG00, Section 2]) are almost-minimizers. Particularly in the Mumford-Shah case, minimizers of the *full* Mumford-Shah functional (1.1) can be viewed as almost-minimizers of its variant

$$\int_{\Omega\setminus K}|\nabla u|^2\,\mathrm{d}x+\mathcal{H}^{n-1}(K)\,, \tag{1.2}$$

defined on the same pairs $(u,K)$ as before. With this connection (see Proposition 4.1 for a precise statement in our setting) at hand we can then limit all further considerations to the *reduced* Mumford-Shah functional (1.2) without zero-order term.

We emphasize that the treatment of almost-minimizers requires one essential deviation from [DPF14], which we now describe. Indeed the proof in [DPF14] draws on the fact that a minimizing $u$ for (1.2) is harmonic (or, in case of (1.1) satisfies a similar equation) on $\Omega\setminus K$ and gradient estimates can be deduced. In our case

[1]In fact, the relation is the following: By a result of Ambrosio–Fusco–Hutchinson [AFH03] (cf. [DLFR14]), local $\mathrm{L}^p$ integrability of $\nabla u$ up to $K$ with $p\in(2,4)$ implies that the singular set of $K$ has Hausdorff dimension $\leq n-\frac{p}{2}$. Thus, if one could obtain this integrability for $p$ arbitrarily close to 4, then one could conclude that the singular set has dimension $\leq n-2$. Moreover, in dimension $n=2$, if one could even establish that $\nabla u$ is locally in the Lorentz space $\mathrm{L}^{4,\infty}$, then a result of [DLF13] comes yet closer to the conjecture. It shows that the singular set is a locally finite subset of $\Omega$ and a classification of singularities à la Bonnet [Bon96] is possible.

instead, since $u$ just *almost*-minimizes the Dirichlet integral on $\Omega\setminus K$, we cannot rely on an equation, and an additional comparison between the almost-minimizer and a minimizing harmonic function renders necessary. In "interior" away-from-$K$ situations, gradient estimates for $u$ follow easily. However, the crucial point of the proof lies in the treatment of a sort-of "boundary" up-to-$K$ situation, and in this case the deduction of gradient estimates for $u$ also depends on flattening of (regular parts of) $K$ and reflection of $u$ across the flattened boundary. While the basic approach is still standard in boundary regularity issues, the details get somewhat technical, and we believe that this part of the proof deserves the careful account which we provide in Section 5. The other arguments in the proof stay closer to [DPF14] and are thus described much more concisely.

Finally, we close this introduction with a rough discussion of the setting and the technically feasible class of almost-minimizers for our result. To this end we first mention that we work in the natural framework introduced in [DGA88, Amb89, DGA89, Amb90] of SBV functions $u$ on $\Omega$, on which the reduced Mumford-Shah functional takes the form

$$\mathrm{MS}[u;A] = \int_A |\nabla u|^2 \,\mathrm{d}x + \mathcal{H}^{n-1}(\mathrm{S}_u \cap A) \tag{1.3}$$

(where $A$ is a measurable subset of $\Omega$, $\nabla u$ denotes the density of the absolutely continuous part of the gradient measure of $u$, and $\mathrm{S}_u$ stands for the approximate discontinuity set of $u$). Basically our almost-minimizers are then defined by conditions of the type

$$\mathrm{MS}[u;\mathrm{B}_r] \le \mathrm{MS}[u+\varphi;\mathrm{B}_r] + \Theta(r) \tag{1.4}$$

for all SBV functions $\varphi$ with support in a ball $\mathrm{B}_r \subset \Omega$ of radius $r > 0$, where $\Theta\colon [0,\infty) \to [0,\infty)$ is a certain fixed modulus. While most results on Mumford-Shah almost-minimizers apply under the hypotheses[2] $\Theta(r) \lesssim r^{n-1}$ or $\Theta(r) \lesssim r^{n-1+a}$ with arbitrarily small $a > 0$, this does not seem achievable for our result. In fact, the proof draws on auxiliary $\mathrm{C}^{1,\beta}$ estimates, and obtaining those even for almost-minimizers of the Dirichlet integral requires the stronger condition $\Theta(r) \lesssim r^{n+b}$ with $b > 0$. Yet again true minimizers of the *full* Mumford-Shah functional (1.1) only satisfy (1.4) with $\Theta(r) \lesssim r^n$ but not neccesarily with $\Theta(r) \lesssim r^{n+b}$. This dilemma drives us to introduce a suitable class of SBV almost-minimizers, which we call $(a,b)$-almost-minimizers, by generally requiring (1.4) with $\Theta(r) \lesssim r^{n-1+a}$, $a > 0$, but also imposing (1.4) as in [DGG00] with $\Theta(r) = \Theta_{u,\varphi}(r) \lesssim r^b \fint_{\mathrm{B}_r}(1+|\nabla u|^2+|\nabla\varphi|^2)\,\mathrm{d}x$ in specific situations[3]; see Definition 3.1. While the resulting notion may seem awkwardly technical, it *does* meet the basic requirements: It is wide enough to include all true minimizers of (1.1) (see Section 4), but also restrictive enough to enable gradient comparison estimates and to ultimately carry out the proof of the higher integrability result. Since furthermore the notion is very general, we believe

[2]Here $A(r) \lesssim B(r)$ means $A(r) \le CB(r)$ for $r \in [0,\infty)$ with some $r$-independent constant $C \in [0,\infty)$.

[3]Indeed the specific situations are those in which $\mathrm{B}_r$ does not intersect the singular part of $\overline{\mathrm{S}_u}$ and the variation $\varphi$ does not enlarge the discontinuity set.

that it indeed constitutes a technically adequate basis for results which depend on regularity of both $\nabla u$ and $\mathrm{S}_u$.

The results of this paper are partially contained in the first author's master thesis [Pio16], which has been directed by the second author.

# 2 Preliminaries

In this paper, we use $\mathbb{N}$ for the positive integers and $\mathbb{N}_0$ for the non-negative integers. We take $2 \le n \in \mathbb{N}$ and assume $\Omega \subset \mathbb{R}^n$ to be a bounded open set. For a measurable set $A \subset \mathbb{R}^n$ we write $\mathcal{L}^n(A)$ for the Lebesgue measure and $\mathcal{H}^k(A)$ for the $k$-dimensional Hausdorff measure. If $0 < \mathcal{L}^n(A) < \infty$, we use the notation

$$\fint_A u(y)\,\mathrm{d}y = \frac{1}{\mathcal{L}^n(A)} \int_A u(y)\,\mathrm{d}y$$

for the mean value integral of a function $u \in \mathrm{L}^1(A)$. If the domain of integration is a ball $\mathrm{B}_r(x)$ and the center is unambiguous, we shorten this to $(u)_r := \fint_{\mathrm{B}_r(x)} u(y)\,\mathrm{d}y$. We set $\omega_n := \mathcal{L}^n(\mathrm{B}_1(x))$. For $\varrho > 0$ we call $\mathcal{N}_\varrho(E) := \{x \in \mathbb{R}^n : \operatorname{dist}(x, E) < \varrho\}$ the $\varrho$-neighborhood of a set $E \subset \mathbb{R}^n$. We will express by $A \Subset \Omega$ that $\overline{A}$ is a compact set with $\overline{A} \subset \Omega$.

We briefly recall some notions related to BV functions and refer to [AFP00] for general information on this topic. For a function $u \in \mathrm{BV}(\Omega)$ we define the approximate discontinuity set $\mathrm{S}_u \subset \Omega$ by

$$x \notin \mathrm{S}_u \qquad \Longleftrightarrow \qquad \exists z \in \mathbb{R}, \text{ s.t. } \lim_{\varrho \to 0} \fint_{\mathrm{B}_\varrho(x)} |u(y) - z|\,\mathrm{d}y = 0.$$

The distributional derivative $Du$ of $u \in \mathrm{BV}(\Omega)$ can be decomposed into an absolutely continuous part $D^a u$ and a singular part $D^s u$ with respect to the Lebesgue measure $\mathcal{L}^n$. We have $D^a u = (\nabla u)\mathcal{L}^n$, where $\nabla u$ is the approximate differential of $u$. We call $u \in \mathrm{BV}(\Omega)$ a *special function of bounded variation*, if $D^s u$ is concentrated on $\mathrm{S}_u$, that is $|D^s u|(\Omega \setminus \mathrm{S}_u) = 0$, and write $\mathrm{SBV}(\Omega)$ for the corresponding function space. Under this assumption $D^s u$ is absolutely continuous with respect to the measure $\mathcal{H}^{n-1} \llcorner \mathrm{S}_u$, which denotes the restriction of $\mathcal{H}^{n-1}$ to $\mathrm{S}_u$. Because the derivative of a $\mathrm{W}^{1,1}$ function consists only of the absolutely continuous part, we have, for $u \in \mathrm{SBV}(\Omega)$, the useful equivalence

$$u \in W^{1,1}(\Omega) \iff \mathcal{H}^{n-1}(\mathrm{S}_u) = 0\,. \tag{2.1}$$

**Lemma 2.1.** *There exists a dimensional constant $C > 0$ such that, if $h$ is a harmonic function defined on a ball $\mathrm{B}_r \subset \mathbb{R}^n$ with radius $r > 0$ and $0 < \tau \le \frac{1}{2}$, it holds*

$$\fint_{\mathrm{B}_{r\tau}} |\nabla h - (\nabla h)_{r\tau}|^2\,\mathrm{d}x \le C\tau^2 \fint_{\mathrm{B}_r} |\nabla h - (\nabla h)_r|^2\,\mathrm{d}x. \tag{2.2}$$

*Proof.* This is a well-known estimate for harmonic functions. It can be derived from the bound $|V(x)-(V)_\varrho| \le \fint_{\mathrm{B}_\varrho} |V(x)-V(y)|\,\mathrm{d}y \le \varrho \sup_{\mathrm{B}_\varrho} |\nabla V|$ for $V \in \mathrm{C}^1(\mathrm{B}_\varrho, \mathbb{R}^n)$, $x \in \mathrm{B}_\varrho$ and an interior estimate for harmonic functions (see [AFP00, Lemma 7.44]). In short one has

$$\fint_{\mathrm{B}_{r\tau}} |\nabla h - (\nabla h)_{r\tau}|^2\,\mathrm{d}x \le (\tau r)^2 \sup_{x\in \mathrm{B}_{\tau r}} |\nabla^2 h|^2 \le C\tau^2 \fint_{\mathrm{B}_r} |\nabla h - (\nabla h)_r|^2\,\mathrm{d}x\,. \qquad \square$$

**Lemma 2.2.** *If $v \in \mathrm{L}^2(\mathrm{B}_r)$ on a ball $\mathrm{B}_r \subset \mathbb{R}^n$, then for all $y \in \mathbb{R}^n$ there holds*

$$\fint_{\mathrm{B}_r} |v-(v)_r|^2\,\mathrm{d}x \le \fint_{\mathrm{B}_r} |v-y|^2\,\mathrm{d}x.$$

*Proof.* This can be easily seen by calculating the minimum of the function $f(y) := \fint_{\mathrm{B}_r} |v-y|^2\,\mathrm{d}x$ in $\mathbb{R}^n$. $\square$

# 3 Definition and basic properties of almost-minimizers

Here we spell out our definition of almost-minimizers, which has already been motivated and explained in the introduction.

**Definition 3.1** (almost-minimizer)**.** *Let $a, b > 0$. For a function $u \in \mathrm{SBV}(\Omega)$ and a measurable set $A \subset \Omega$ we define the (reduced) Mumford-Shah functional* MS *by* (1.3)*. We say that a function $u \in \mathrm{SBV}(\Omega)$ that fulfills*

$$\mathrm{MS}[u; \mathrm{B_r}] < \infty \qquad \textit{for all balls } \mathrm{B}_r \Subset \Omega \tag{3.1}$$

*is an $(a, b)$-*almost-minimizer *on $\Omega$, if there exists a constant $C_m > 0$ such that the following conditions hold:*

1. *For all balls $\mathrm{B}_r \Subset \Omega$ and functions $v \in \mathrm{SBV}(\Omega)$ with $\{u \ne v\} \Subset \mathrm{B}_r$ we have*

$$\mathrm{MS}[u; \mathrm{B}_r] \le \mathrm{MS}[v; \mathrm{B}_r] + C_m r^{n-1+a}. \tag{3.2}$$

2. *For all balls $\mathrm{B}_r \Subset \Omega$, such that $\overline{\mathrm{S}}_u$ coincides in $\mathrm{B}_r$ with the rotated graph of an arbitrary function $f : \mathbb{R}^{n-1} \to \mathbb{R}$, and for all $\varphi \in \mathrm{SBV}(\mathrm{B}_r) \cap \mathrm{W}^{1,2}(\mathrm{B}_r \setminus \overline{\mathrm{S}}_u)$ with* $\operatorname{supp}\varphi \Subset \mathrm{B}_r$ *we have*

$$\mathrm{MS}[u; \mathrm{B}_r] \le \mathrm{MS}[u+\varphi; \mathrm{B}_r] + C_m r^b \int_{\mathrm{B}_r} \left(1 + |\nabla u|^2 + |\nabla \varphi|^2\right)\mathrm{d}x. \tag{3.3}$$

**Remark 3.2.** Consider a function $u \in \mathrm{SBV}(\Omega)$ that satisfies (3.1) and condition (3.2) for $a > 1$. Then (3.3) holds for $b = a - 1$ and $u$ is an $(a, b)$-almost-minimizer.

**Proposition 3.3** (energy upper bound and density lower bound)**.** *If $u \in \mathrm{SBV}(\Omega)$ is an $(a,b)$-almost-minimizer of* MS*, then there exists a $C_0 \geq 1$ and a radius $r_0 \leq 1$, such that for all balls $\mathrm{B}_r(x) \subset \Omega$ with $r \leq r_0$ we have*

$$\int_{\mathrm{B}_r(x)} |\nabla u|^2 \,\mathrm{d}y + \mathcal{H}^{n-1}(\mathrm{S}_u \cap \mathrm{B}_r(x)) \leq C_0 r^{n-1} \tag{3.4}$$

*and for all balls $\mathrm{B}_r(x) \subset \Omega$ with center $x \in \overline{\mathrm{S}}_u$ and $r \leq r_0$ we get*

$$C_0^{-1} r^{n-1} \leq \mathcal{H}^{n-1}(\mathrm{S}_u \cap \mathrm{B}_r(x)) \leq C_0 r^{n-1}. \tag{3.5}$$

*Proof.* This is a well-known fact about almost-minimizers in the sense of (3.2). See for example [AFP00, Lemma 7.19] and [AP97]. □

**Remark 3.4.** As a routine consequence of the density lower bound in (3.5) and a theorem on $k$-dimensional densities (see [AFP00, Theorem 2.56]), almost-minimizers in the sense of (3.2) satisfy

$$\mathcal{H}^{n-1}((\overline{\mathrm{S}}_u \setminus \mathrm{S}_u) \cap \Omega) = 0\,.$$

If $u$ also fulfills (3.3), we have $\mathcal{H}^{n-1}(\mathrm{S}_{u+\varphi}) \leq \mathcal{H}^{n-1}(\mathrm{S}_u)$ for $\varphi$ as in Definition 3.1, and (3.3) actually implies

$$\int_{\mathrm{B}_r} |\nabla u|^2 \,\mathrm{d}x \leq \int_{\mathrm{B}_r} |\nabla u + \nabla\varphi|^2 \,\mathrm{d}x + C_m r^b \int_{\mathrm{B}_r} \left(1 + |\nabla u|^2 + |\nabla\varphi|^2\right) \mathrm{d}x.$$

# 4 Minimizers of the full functional are almost-minimizers

In this section we show that minimizers of the full Mumford-Shah functional are almost-minimizers of the reduced functional. Actually, we consider the SBV version of (1.1), that is

$$\mathrm{MS}_g[u] := \int_\Omega |\nabla u|^2 \ \mathrm{d}x + \lambda \int_\Omega |u-g|^2 \ \mathrm{d}x + \mathcal{H}^{n-1}(\mathrm{S}_u) \qquad \text{for } u \in \mathrm{SBV}(\Omega) \tag{4.1}$$

with $g \in \mathrm{L}^\infty(\Omega), \lambda \in (0,\infty)$, and we recall that minimizers $u \in \mathrm{SBV}(\Omega)$ of $\mathrm{MS}_g$ correspond to minimizers $(u,K) = (u, \overline{\mathrm{S}}_u)$ of (1.1); see for instance [Fus03]. Since, in this sense, the minimization problems for $\mathrm{MS}_g$ and (1.1) are equivalent, it is enough to show that minimizers of $\mathrm{MS}_g$ are also almost-minimizers in our sense. A precise statement of this result reads:

**Proposition 4.1.** *If $u \in \mathrm{SBV}(\Omega)$ is a minimizer of $\mathrm{MS}_g$, then $u$ is an $(a,b)$-almost-minimizer of* MS *with $a = b = 1$.*

*Proof.* First we observe, that every minimizer $u$ of $\mathrm{MS}_g$ is bounded almost everywhere by $\|g\|_\infty$, i.e. $\|u\|_\infty \leq \|g\|_\infty$. This can be checked by comparing the minimizer

$u$ with the truncated function $\widetilde{u}(x) = \max\{-\|g\|_\infty, \min\{\|g\|_\infty, u(x)\}\}$. Notice that $\widetilde{u} \in \mathrm{SBV}(\Omega)$, $\mathrm{S}_w \subset \mathrm{S}_u$ and $|\nabla\widetilde{u}| \le |\nabla u|$ almost everywhere. Arguing by contradiction yields $u = \widetilde{u}$ almost everywhere and consequently $\|u\|_\infty \le \|g\|_\infty$. Moreover, comparing with $w \equiv 0 \in \mathrm{SBV}(\Omega)$, we get an upper bound $\mathrm{MS}_g[u] \le \mathrm{MS}_g[0] = \lambda\|g\|^2_{L^2(\Omega)}$ for $\mathrm{MS}_g[u]$.

**Step 1:** We show that $u$ satisfies condition (3.2). For this purpose let $\mathrm{B}_r \Subset \Omega$ and $v \in \mathrm{SBV}(\Omega)$ with $\{v \ne u\} \Subset \mathrm{B}_r$. Using a truncation we define $\widetilde{v}(x) = \max\{-\|g\|_\infty, \min\{\|g\|_\infty, v(x)\}\}$ and as before $\widetilde{v} \in \mathrm{SBV}(\Omega)$, $\mathrm{S}_{\widetilde{v}} \subset \mathrm{S}_v$ and $|\nabla\widetilde{v}| \le |\nabla v|$ almost everywhere. Observe that there is no cutoff outside $\mathrm{B}_r$, that is $\widetilde{v}(x) = v(x) = u(x)$ for $x \notin \mathrm{B}_r$, because $u$ is bounded by $\|g\|_\infty$ and $\{v \ne u\} \Subset \mathrm{B}_r$. We can estimate the difference between $g$ and the truncated function $\widetilde{v}$ by calculating

$$\int_{\mathrm{B}_r} |\widetilde{v} - g|^2 \,\mathrm{d}x \le \int_{\mathrm{B}_r} 2\left(|\widetilde{v}|^2 + |g|^2\right) \mathrm{d}x \le 4\omega_n\|g\|_\infty r^n.$$

Using this estimate and applying the minimality $\mathrm{MS}_g[u] \le \mathrm{MS}_g[\widetilde{v}]$ on the ball $\mathrm{B}_r$ we get

$$\begin{aligned}\mathrm{MS}[u; \mathrm{B}_r] &\le \int_{\mathrm{B}_r} |\nabla\widetilde{v}|^2 \,\mathrm{d}x + \mathcal{H}^{n-1}(\mathrm{S}_{\widetilde{v}} \cap \mathrm{B}_r) + \lambda\int_{\mathrm{B}_r} |\widetilde{v} - g|^2 \,\mathrm{d}x - \lambda\int_{\mathrm{B}_r} |u - g|^2 \,\mathrm{d}x \\ &\le \int_{\mathrm{B}_r} |\nabla v|^2 \,\mathrm{d}x + \mathcal{H}^{n-1}(\mathrm{S}_v \cap \mathrm{B}_r) + 4\omega_n\lambda\|g\|_\infty r^n \\ &\le \mathrm{MS}[v; \mathrm{B}_r] + Cr^{n-1+a}\end{aligned}$$

with $C = 4\omega_n\lambda\|g\|_\infty$ and $a = 1$. This proves the statement.

**Step 2:** We show that $u$ also satisfies condition (3.3). Consider a ball $\mathrm{B}_r \Subset \Omega$, such that $\overline{\mathrm{S}}_u \cap \mathrm{B}_r$ coincides with the rotated graph $\Gamma$ of a function $f : \mathbb{R}^{n-1} \to \mathbb{R}$. Assume for simplicity that $\mathrm{B}_r$ is centered at zero and the rotation is the identity. We first establish a Poincaré type inequality for $\widetilde{\varphi} \in \mathrm{C}^1(\mathrm{B}_r \setminus \overline{\mathrm{S}}_u)$ with $\operatorname{supp}\widetilde{\varphi} \Subset \mathrm{B}_r$. To this end, we abbreviate $D_r := \{x' \in \mathbb{R}^{n-1} : |x'| < r\}$,

$$\mathrm{B}^+ := \{(x', x_n) \in \mathrm{B}_r : x_n > f(x')\}, \qquad \mathrm{B}^- := \{(x', x_n) \in \mathrm{B}_r : x_n < f(x')\}.$$

Now we write $\widetilde{\varphi}(x', x_n) = \int_{-r}^{x_n} \partial_n\widetilde{\varphi}(x', t)\,\mathrm{d}t$ and use Fubini's theorem and Hölder's inequality in a standard way to get

$$\begin{aligned}\int_{\mathrm{B}^-} |\widetilde{\varphi}|^2 \,\mathrm{d}x &\le 2r\int_{D_r}\int_{-r}^{\min\{f(x'),r\}}\int_{-r}^{x_n} |\partial_n\widetilde{\varphi}(x', t)|^2 \,\mathrm{d}t\,\mathrm{d}x_n\,\mathrm{d}x' \\ &\le 4r^2\int_{D_r}\int_{-r}^{f(x')} |\partial_n\widetilde{\varphi}(x', t)|^2 \,\mathrm{d}t\,\mathrm{d}x' \le 4r^2\int_{\mathrm{B}^-} |\nabla\widetilde{\varphi}|^2 \,\mathrm{d}x.\end{aligned}$$

Clearly, the analogous inequality holds on $\mathrm{B}^+$ and thus on the whole ball $\mathrm{B}_r$. By approximation the estimate on $\mathrm{B}_r$ extends to $\varphi \in \mathrm{SBV}(\mathrm{B}_r) \cap \mathrm{W}^{1,2}(\mathrm{B}_r \setminus \overline{\mathrm{S}}_u)$ with $\operatorname{supp}\varphi \Subset \mathrm{B}_r$. Indeed, by the Meyers-Serrin theorem we find a sequence $\{\varphi_n\}_{n\in\mathbb{N}} \subset \mathrm{C}^1(\mathrm{B}_r \setminus \overline{\mathrm{S}}_u) \cap \mathrm{W}^{1,2}(\mathrm{B}_r \setminus \overline{\mathrm{S}}_u)$ which converges to $\varphi$ in the $\mathrm{W}^{1,2}$-norm. We have $\operatorname{supp}\varphi \subset \mathrm{B}_t$ for some $t < r$ and consider a cut-off function $\eta \in \mathrm{C}^1_{\mathrm{cpt}}(\mathrm{B}_r)$ with $\eta \equiv 1$ in $\mathrm{B}_t$.

Then $\{\eta\varphi_n\}_{n\in\mathbb{N}}$ also converges to $\varphi$ in $\mathrm{W}^{1,2}(\mathrm{B}_r \setminus \overline{\mathrm{S}}_u)$, and the previously derived Poincaré inequality holds for $\eta\varphi_n$ on $\mathrm{B}_r$. Passing to the limit, we conclude

$$\int_{\mathrm{B}_r} |\varphi|^2 \,\mathrm{d}x \le 4r^2 \int_{\mathrm{B}_r} |\nabla\varphi|^2 \,\mathrm{d}x\,. \tag{4.2}$$

Now we turn to our main concern. As in Step 1 we utilize $\|u\|_\infty \le \|g\|_\infty$ and the fact that $u$ is a minimizer for $\mathrm{MS}_g$. Combining this with Young's inequality in the form $2|\varphi| \le r + \frac{1}{r}|\varphi|^2$ and, in the last step, with the inequality (4.2), we infer

$$\begin{aligned}
\mathrm{MS}[u;\mathrm{B}_r] &\le \mathrm{MS}[u+\varphi;\mathrm{B}_r] + \lambda\int_{\mathrm{B}_r} |u-g+\varphi|^2\,\mathrm{d}x - \lambda\int_{\mathrm{B}_r}|u-g|^2\,\mathrm{d}x\\
&\le \mathrm{MS}[u+\varphi;\mathrm{B}_r] + 2\lambda\int_{\mathrm{B}_r}|u-g|\,|\varphi|\,\mathrm{d}x + \lambda\int_{\mathrm{B}_r}|\varphi|^2\,\mathrm{d}x\\
&\le \mathrm{MS}[u+\varphi;\mathrm{B}_r] + 2\lambda r\|g\|_\infty\mathcal{L}^n(\mathrm{B}_r) + \Big(2\lambda\|g\|_\infty\frac{1}{r}+\lambda\Big)\int_{\mathrm{B}_r}|\varphi|^2\,\mathrm{d}x\\
&\le \mathrm{MS}[u+\varphi;\mathrm{B}_r] + Cr^b\int_{\mathrm{B}_r}\Big(1+|\nabla u|^2+|\nabla\varphi|^2\Big)\,\mathrm{d}x,
\end{aligned}$$

with $b = 1$ and a constant $C$, that depends on $\|g\|_\infty$, $\lambda$, and the diameter of the bounded set $\Omega$. This completes the proof. □

# 5 $\mathrm{L}^\infty$ gradient estimates

In this section we derive $\mathrm{L}^\infty$ bounds for $\nabla u$, which are crucial for our purposes. We start with an "interior" case.

**Lemma 5.1** (gradient estimate away from $\mathrm{S}_u$)**.** *If $u \in \mathrm{SBV}(\Omega)$ is an $(a,b)$-almost-minimizer of* MS*, then, for every $\beta \in (0,1)$ with $\beta \le \frac{b}{2}$, there exist constants $C' \ge 1$ and $0 < r' \le 1$ with the following property. For every Lebesgue point $x_0 \in \Omega$ of $\nabla u$ and every $r < r'$ with $\mathrm{B}_r(x_0) \cap \overline{\mathrm{S}}_u = \emptyset$ and $\mathrm{B}_r(x_0) \Subset \Omega$, it holds*

$$|\nabla u(x_0)|^2 \;\le\; C' \fint_{\mathrm{B}_r} \Big(|\nabla u|^2 + r^{2\beta}\Big)\,\mathrm{d}x. \tag{5.1}$$

*Proof.* We will only deal with balls centered at $x_0$, and for this reason we simplify our notation by writing $\mathrm{B}_r$ instead of $\mathrm{B}_r(x_0)$. Notice that $r^b \le r^{2\beta}$ and thus $u$ satisfies condition (3.3) with $2\beta$ instead of $b$. Recall that $u \in \mathrm{W}^{1,2}(\mathrm{B}_r)$ because of $\overline{\mathrm{S}}_u \cap \mathrm{B}_r = \emptyset$ and (2.1).

**Step 1:** It is well known that there exists a unique harmonic function $h$ on $\mathrm{B}_r$ with $\varphi := h - u \in \mathrm{W}^{1,2}_0(\mathrm{B}_r)$. We now employ condition (3.3) with the test function $\varphi$ on balls $\mathrm{B}_s$ with $s > r$ and then send $s \searrow r$. In this way we obtain

$$\int_{\mathrm{B}_r}|\nabla u|^2\,\mathrm{d}x \le \int_{\mathrm{B}_r}|\nabla h|^2\,\mathrm{d}x + C_m r^{2\beta}\int_{\mathrm{B}_r}\Big(1+|\nabla u|^2+|\nabla u-\nabla h|^2\Big)\,\mathrm{d}x\,.$$

Taking into account the harmonicity of $h$ we infer

$$\begin{aligned}\fint_{B_r} |\nabla u - \nabla h|^2 \,\mathrm{d}x &= \fint_{B_r} |\nabla u|^2 \,\mathrm{d}x - \fint_{B_r} |\nabla h|^2 \,\mathrm{d}x + 2\fint_{B_r} \nabla h \cdot \nabla \varphi \,\mathrm{d}x \\ &\le C_m r^{2\beta} \fint_{B_r} \Big(1 + |\nabla u|^2 + |\nabla u - \nabla h|^2\Big) \,\mathrm{d}x \,.\end{aligned}$$

Choosing $r'$ small enough that $C_m {r'}^{2\beta} \le \frac{1}{2}$ and absorbing a term on the left-hand side, we arrive at

$$\fint_{B_r} |\nabla u - \nabla h|^2 \,\mathrm{d}x \le 2C_m r^{2\beta} \fint_{B_r} \Big(1 + |\nabla u|^2\Big) \,\mathrm{d}x \,.$$

**Step 2:** Here we deduce the basic excess decay. We work with a constant $C$, which varies from line to line, and with $0 < \tau \le \frac{1}{2}$. By Step 1, Lemma 2.1, and Lemma 2.2 we get

$$\begin{aligned}\fint_{B_{\tau r}} |\nabla u - (\nabla u)_{\tau r}|^2 \,\mathrm{d}x &\le \fint_{B_{\tau r}} |\nabla u - (\nabla h)_{\tau r}|^2 \,\mathrm{d}x \\ &\le C\bigg(\fint_{B_{\tau r}} |\nabla h - (\nabla h)_{\tau r}|^2 \,\mathrm{d}x + \fint_{B_{\tau r}} |\nabla u - \nabla h|^2 \,\mathrm{d}x\bigg) \\ &\le C\bigg(\tau^2 \fint_{B_r} |\nabla h - (\nabla h)_r|^2 \,\mathrm{d}x + \fint_{B_{\tau r}} |\nabla u - \nabla h|^2 \,\mathrm{d}x\bigg) \\ &\le \widehat{C}\tau^2 \fint_{B_r} |\nabla u - (\nabla u)_r|^2 \,\mathrm{d}x + \mathrm{C}\tau^{-n} \fint_{B_r} |\nabla u - \nabla h|^2 \,\mathrm{d}x \\ &\le \tau^{2\gamma} \fint_{B_r} |\nabla u - (\nabla u)_r|^2 \,\mathrm{d}x + C_* r^{2\beta} \fint_{B_r} \Big(1 + |\nabla u|^2\Big) \,\mathrm{d}x \,.\end{aligned}$$

Here, in the last step we fixed $\tau$ such that $\widehat{C}\tau^2 \le \tau^{2\gamma}$ for some $\gamma \in (\beta, 1)$. We emphasize that $\widehat{C}$, $\tau$, and $C_*$ depend only on the dimension $n$, on $C_m$, and on $b$.

**Step 3:** Next we iterate the estimate from Step 2. We introduce the abbreviation

$$E(\tau^i r) := \fint_{B_{\tau^i r}} |\nabla u - (\nabla u)_{\tau^i r}|^2 \,\mathrm{d}x \qquad \text{for } i \in \mathbb{N}_0$$

for the excess on $B_{\tau^i r}$ and set

$$M_i := \tau^{2i(\gamma-\beta)} + 2C_* \tau^{-2\beta} \sum_{\ell=0}^{i-1} \tau^{2\ell(\gamma-\beta)} \qquad \text{for } i \in \mathbb{N}_0 \,.$$

We can assume $C_* \ge 1$ and therefore $M_0 \le M_1 \le \cdots \le M_i$ for $i \in \mathbb{N}$. Because of $\tau \in (0,1)$ and $\gamma - \beta > 0$ it follows that $M_i$ are bounded for $i \to \infty$ and consequently we have $M := \sup_{i \in \mathbb{N}_0} M_i < \infty$.

We now prove by induction the hypothesis

$$E(\tau^i r) \le M_i \tau^{2i\beta}\Big(r^{2\beta} + \fint_{B_r} |\nabla u|^2 \,\mathrm{d}x\Big) \qquad \text{for every } r \in (0, r') \text{ and } i \in \mathbb{N}_0 \tag{5.2}$$

with $r' := \min\{[M\tau^{-n}(1+\frac{1}{1-\tau^\beta})^2]^{\frac{1}{-2\beta}}, (2C_m)^{-\beta}, 1\}$.

Base case: For $i = 0$, we have $M_0 = 1$, and by employing Lemma 2.2 it follows

$$E(r) = \fint_{B_r} |\nabla u - (\nabla u)_r|^2 \, dx \le \fint_{B_r} |\nabla u - 0|^2 \, dx + r^{2\beta}.$$

Inductive step: Assume that (5.2) holds for $i = 0, 1, \dots k$ with a $k \in \mathbb{N}_0$. We show that (5.2) is true for $i = k+1$. First of all, notice, that by replacing $r$ with $\tau^i r$ in Step 1 and Step 2 we immediately get the corresponding result

$$E(\tau^{k+1}r) \le \tau^{2\gamma} E(\tau^k r) + C_*(\tau^k r)^{2\beta}\Big(1 + \fint_{B_{\tau^k r}} |\nabla u|^2 \, dx\Big) \tag{5.3}$$

for $i = k+1$. Next we derive an estimate for the integral on the right-hand side of (5.3). Adding and subtracting mean values iteratively gives

$$\begin{aligned}\Big(\fint_{B_{\tau^k r}} |\nabla u|^2 \, dx\Big)^{\frac{1}{2}} &\le \Big(\fint_{B_{\tau^k r}} |\nabla u - (\nabla u)_{\tau^k r}|^2 \, dx\Big)^{\frac{1}{2}} + |(\nabla u)_r| + \sum_{i=0}^{k-1} |(\nabla u)_{\tau^{i+1} r} - (\nabla u)_{\tau^i r}| \\ &\le |(\nabla u)_r| + E(\tau^k r)^{\frac{1}{2}} + \sum_{i=0}^{k-1} \Big(\fint_{B_{\tau^{i+1} r}} |\nabla u - (\nabla u)_{\tau^i r}|^2 \, dx\Big)^{\frac{1}{2}} \\ &\le |(\nabla u)_r| + \tau^{-\frac{n}{2}} \sum_{i=0}^{k} E(\tau^i r)^{\frac{1}{2}}.\end{aligned}$$

Using the induction hypothesis and evaluating the geometric series, we get

$$\sum_{i=0}^{k} E(\tau^i r)^{\frac{1}{2}} \le \sum_{i=0}^{k} M_i^{\frac{1}{2}} \tau^{i\beta}\Big(r^{2\beta} + \fint_{B_r} |\nabla u|^2 \, dx\Big)^{\frac{1}{2}} \le \frac{M_k^{\frac{1}{2}}}{1-\tau^\beta}\Big(r^{2\beta} + \fint_{B_r} |\nabla u|^2 \, dx\Big)^{\frac{1}{2}}.$$

Setting $C_\tau := \tau^{-\frac{n}{2}}(1 + \frac{1}{1-\tau^\beta})$ we conclude

$$\fint_{B_{\tau^k r}} |\nabla u|^2 \, dx \le C_\tau^2 M_k \Big(r^{2\beta} + \fint_{B_r} |\nabla u|^2 \, dx\Big). \tag{5.4}$$

We use this estimate on the right-hand side of (5.3) and use the induction hypothesis once more to get

$$\begin{aligned}&E(\tau^{k+1} r) \\ &\le \tau^{2\gamma} E(\tau^k r) + C_* \tau^{2k\beta} r^{2\beta}\Big(1 + C_\tau^2 M_k \Big(r^{2\beta} + \fint_{B_r} |\nabla u|^2 \, dx\Big)\Big) \\ &\le \tau^{2\gamma} M_k \tau^{2k\beta}\Big(r^{2\beta} + \fint_{B_r} |\nabla u|^2 \, dx\Big) + C_* \tau^{2k\beta}\Big(r^{2\beta} + C_\tau^2 M_k r^{4\beta} + C_\tau^2 M_k r^{2\beta} \fint_{B_r} |\nabla u|^2 \, dx\Big) \\ &\le \tau^{2(k+1)\beta}\Big(C_* \tau^{-2\beta}(1 + C_\tau^2 M_k r^{2\beta}) + \tau^{2(\gamma-\beta)} M_k\Big)\Big(r^{2\beta} + \fint_{B_r} |\nabla u|^2 \, dx\Big).\end{aligned}$$

Due to the way we defined $r'$ and $C_\tau$ it follows

$$C_\tau^2 M_k r^{2\beta} \le C_\tau^2 M_k \Big(\frac{1}{C_\tau^2 M_k}\Big)^{\frac{2\beta}{2\beta}} = 1$$

and because of $M_{k+1} = \tau^{2(\gamma-\beta)} M_k + 2C_* \tau^{-2\beta}$ we conclude

$$\begin{aligned} E(\tau^{k+1} r) &\le \tau^{2(k+1)\beta}\Big(2C_*\tau^{-2\beta} + \tau^{2(\gamma-\beta)} M_k\Big)\Big(r^{2\beta} + \fint_{\mathrm{B}_r} |\nabla u|^2 \,\mathrm{d}x\Big) \\ &= \tau^{2(k+1)\beta} M_{k+1}\Big(r^{2\beta} + \fint_{\mathrm{B}_r} |\nabla u|^2 \,\mathrm{d}x\Big), \end{aligned}$$

and this proves the inductive step.

**Step 4:** We recall that, by assumption, $x_0$ is a Lebesgue point of $\nabla u$. Moreover, we record that Step 3 also gives (5.4) for all $k \in \mathbb{N}$. All in all, we thus conclude

$$|\nabla u(x_0)|^2 = \lim_{k\to\infty} |(\nabla u)_{\tau^k r}|^2 \le \limsup_{k\to\infty} \fint_{\mathrm{B}_{\tau^k r}} |\nabla u|^2 \,\mathrm{d}x \le C_\tau^2 M \left( r^{2\beta} + \fint_{\mathrm{B}_r} |\nabla u|^2 \,\mathrm{d}x \right).$$

This proves the lemma. □

Next we establish a similar $\mathrm{L}^\infty$ bound also in a sort-of "boundary" case.

**Lemma 5.2** (gradient estimate near regular points of $\mathrm{S}_u$)**.** *Let $u \in \mathrm{SBV}(\Omega)$ be an $(a,b)$-almost-minimizer of* MS. *Then, for all $\alpha, \beta \in (0,1)$ with $\beta \le \frac{1}{2}\min\{\alpha, b\}$, there exist constants $C'' \ge 1$ and $0 < r'' \le 1$ with the following property. If $\mathrm{B}_{2r}(x_0) \Subset \Omega$ is a ball with $x_0 \in \overline{\mathrm{S}}_u$ and $0 < r \le r''$ such that*

$$\overline{\mathrm{S}}_u \cap \mathrm{B}_{2r}(x_0) = [x_0 + \Gamma] \cap \mathrm{B}_{2r}(x_0)\,,$$

*where $\Gamma$ is the rotated graph of a $\mathrm{C}^{1,\alpha}$ function $f : \mathbb{R}^{n-1} \to \mathbb{R}$ with $f(0) = 0$ that fulfills*

$$\|\nabla f\|_\infty + r^\alpha [\nabla f]_{\mathrm{C}^{0,\alpha}} \le \frac{1}{10}\,, \tag{5.5}$$

*then we have*

$$\sup_{x \in \mathrm{B}_{\frac{r}{100}}(x_0)} |\nabla u(x)|^2 \le C'' \fint_{\mathrm{B}_{2r}(x_0)} \Big(|\nabla u|^2 + r^{2\beta}\Big) \,\mathrm{d}x. \tag{5.6}$$

*Proof.* We assume $x_0 = 0$, and we omit the center for balls around 0. Possibly reparametrizing the graph $\Gamma$ over a different hyperplane, we can also assume that $\Gamma$ is the graph of $f$ with $\nabla f(0) = 0$ without need for further rotation. We start by investigating a transformation $\Phi$ that maps $\overline{\mathrm{S}}_u \cap \mathrm{B}_r$ in the hyperplane $H := \mathbb{R}^{n-1} \times \{0\}$. For $x = (x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$ we define this $\mathrm{C}^{1,\alpha}$ diffeomorphism $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ by setting

$$\Phi(x', x_n) := (x', x_n - f(x')).$$

It is easy to check $\det \mathrm{D}\Phi(x) = 1$ and $\mathrm{D}\Phi(0) = \mathrm{I}_n$ (with the $n{\times}n$ unit matrix $\mathrm{I}_n$). Moreover, $\mathrm{D}\Phi$ inherits the Hölder property from $\nabla f$, in fact $\|\mathrm{D}\Phi(x) - \mathrm{D}\Phi(y)\| \le nL|x-y|^\alpha$ with operator norm $\|A\| := \max\limits_{|x|=1}|Ax|$ and Hölder constant $L$ of $\nabla f$.

For $x \in \mathrm{B}_r$, the assumption $\|\nabla f\|_\infty \le \frac{1}{10}$ immediately implies $|\Phi(x) - \Phi(0)| \le \frac{11}{10}r$. Taking into account that the inverse function $\Phi^{-1}(x', x_n) = (x', x_n + f(x'))$ has the same properties, for $V_r := \Phi(\mathrm{B}_r)$, we have

$$\mathrm{B}_{\frac{9}{10}r} \subset V_r \subset \mathrm{B}_{\frac{11}{10}r}.$$

After these preliminary considerations we focus our attention on the transformed function $w := u \circ \Phi^{-1}$ on $V_r$. We obtain $w \in \mathrm{SBV}(V_r)$ and $w \in \mathrm{W}^{1,2}(V_r \setminus \overline{\mathrm{S}}_w)$, because $\Phi$ is a $\mathrm{C}^1$ diffeomorphism. Furthermore, we have

$$\overline{\mathrm{S}}_w \cap V_r = \Phi(\overline{\mathrm{S}}_u \cap \mathrm{B}_r) = H \cap V_r$$

by definition of $\Phi$. Let $\widetilde{\varphi}$ be a function that fulfills $\widetilde{\varphi} \in \mathrm{W}^{1,2}(V_r \setminus \overline{\mathrm{S}}_w) \cap \mathrm{SBV}(V_r)$ and $\mathrm{supp}(\widetilde{\varphi}) \Subset V_r$. Thus $\varphi := \widetilde{\varphi} \circ \Phi$ is a comparison function for $u$, that is $\varphi \in \mathrm{W}^{1,2}(\mathrm{B}_r \setminus \overline{\mathrm{S}}_u) \cap \mathrm{SBV}(\mathrm{B}_r)$ and $\mathrm{supp}(\varphi) \Subset \mathrm{B}_r$. Now we check in which way condition (3.3) for $u$ transfers to $w$. To this end we compute

$$\begin{aligned}
&\int_{V_r \setminus \mathrm{S}_w} |\nabla w(y)|^2 \,\mathrm{d}y = \int_{\mathrm{B}_r \setminus \mathrm{S}_u} |\mathrm{D}\Phi^{-1}(\Phi(x)) \cdot \nabla u(x)|^2 \, |\det(\mathrm{D}\Phi)| \,\mathrm{d}x \\
&\le \int_{\mathrm{B}_r \setminus \mathrm{S}_u} \|\mathrm{D}\Phi^{-1}(\Phi(x)) - \mathrm{D}\Phi^{-1}(\Phi(0)) + \mathrm{I}_n\|^2 \, |\nabla u(x)|^2 \,\mathrm{d}x \\
&\le \int_{\mathrm{B}_r \setminus \mathrm{S}_u} \Big(nL|\Phi(x) - \Phi(0)|^\alpha + \|\mathrm{I}_n\|\Big)^2 |\nabla u(x)|^2 \,\mathrm{d}x \\
&\le \left[1 + 2nL\left(\frac{11}{10}r\right)^\alpha + n^2L^2\left(\frac{11}{10}r\right)^{2\alpha}\right] \int_{\mathrm{B}_r \setminus \mathrm{S}_u} |\nabla u(x)|^2 \,\mathrm{d}x \\
&\le \Big[1 + T_r\Big]\left[\int_{\mathrm{B}_r \setminus \mathrm{S}_u} |\nabla u(x) + \nabla \varphi(x)|^2 \,\mathrm{d}x + C_m r^b \int_{\mathrm{B}_r \setminus \mathrm{S}_u} \Big(1 + |\nabla u(x)|^2 + |\nabla \varphi(x)|^2\Big)\,\mathrm{d}x\right],
\end{aligned}$$

where we set $T_r := 2nL\left(\frac{11}{10}r\right)^\alpha + n^2L^2\left(\frac{11}{10}r\right)^{2\alpha}$. Using the inverse transformation, we get additional factors $T_r$ and we obtain

$$\begin{aligned}
&\int_{V_r \setminus \mathrm{S}_w} |\nabla w|^2 \,\mathrm{d}y \\
&\le \int_{V_r \setminus \mathrm{S}_w} |\nabla w + \nabla \widetilde{\varphi}|^2 \,\mathrm{d}y + \Big[2T_r + T_r^2 + (1+T_r)^2 C_m r^b\Big] \int_{V_r \setminus \mathrm{S}_w} \Big(1 + |\nabla w|^2 + |\nabla \widetilde{\varphi}|^2\Big)\,\mathrm{d}y \\
&\le \int_{V_r \setminus \mathrm{S}_w} |\nabla w + \nabla \widetilde{\varphi}|^2 \,\mathrm{d}y + \widetilde{C} r^{2\beta} \int_{V_r \setminus \mathrm{S}_w} \Big(1 + |\nabla w|^2 + |\nabla \widetilde{\varphi}|^2\Big)\,\mathrm{d}y,
\end{aligned}$$

where we used $2\beta = \min\{\alpha, b\}$ to eliminate the different powers of $r$. Next, we have to transfer this estimate to balls $\mathrm{B}_\varrho \subset \mathrm{B}_{\frac{9}{10}r} \subset V_r$ with center $\Phi(0) = 0$. To this purpose consider $\phi \in \mathrm{W}^{1,2}(\mathrm{B}_\varrho \setminus \overline{\mathrm{S}}_w) \cap \mathrm{SBV}(\mathrm{B}_\varrho)$ with $\mathrm{supp}(\phi) \Subset \mathrm{B}_\varrho$, so that $\nabla \phi = 0$ on $V_r \setminus \mathrm{B}_\varrho$. This yields

$$\int_{\mathrm{B}_\varrho} |\nabla w|^2 \,\mathrm{d}y \le \int_{\mathrm{B}_\varrho} |\nabla w + \nabla \phi|^2 \,\mathrm{d}y + \hat{C} r^{2\beta} \int_{\mathrm{B}_\varrho} |\nabla \phi|^2 \,\mathrm{d}y + \hat{C} r^{2\beta} \int_{\mathrm{B}_s} \Big(1 + |\nabla w|^2\Big)\,\mathrm{d}y, \tag{5.7}$$

with $s := \frac{11}{10}r$. Notice, that $w$ is well defined on $\mathrm{B}_s$ because of $\Phi^{-1}(\mathrm{B}_{\frac{11}{10}r}) \subset \mathrm{B}_{2r} \subset \Omega$.

To recap, we now have a SBV function $w$, that fulfills (5.7) on a ball $\mathrm{B}_\varrho$ and whose discontinuity set $\overline{\mathrm{S}}_u$ is equal to a hyperplane $H$. Therefore, the set $\overline{\mathrm{S}}_u$ divides the ball $\mathrm{B}_\varrho$ into two half-balls

$$\mathrm{B}_\varrho^\pm = \{(y', y_n) \in \mathrm{B}_\varrho : \pm y_n > 0\}$$

and $w$ is a $\mathrm{W}^{1,2}$ function on each of them. This makes it possible, to define two $\mathrm{W}^{1,2}$ functions $w_+$ and $w_-$ on the whole ball by even reflection:

$$w_\pm(y', y_n) := \begin{cases} w(y', y_n) & \text{for } \pm y_n > 0 \\ w(y', -y_n) & \text{for } \pm y_n < 0 \end{cases}$$

for $(y', y_n) \in \mathrm{B}_s$. From now on, we fix $\varrho := \frac{8}{10}r = \frac{8}{11}s$. Consequently, we are in a similar situation as in Lemma 5.1 and can closely follow the steps of its proof.

**Step 1:** Let $h_\pm \in \mathrm{W}^{1,2}(\mathrm{B}_\varrho)$ denote the harmonic functions such that $\varphi_\pm := h_\pm - w_\pm \in \mathrm{W}_0^{1,2}(\mathrm{B}_\varrho)$. We use these functions to define

$$h := h_+ \mathbb{1}_{\mathrm{B}_\varrho^+} + h_- \mathbb{1}_{\mathrm{B}_\varrho^-} \qquad \text{and} \qquad \varphi := \varphi_+ \mathbb{1}_{\mathrm{B}_\varrho^+} + \varphi_- \mathbb{1}_{\mathrm{B}_\varrho^-}.$$

Thus, we have $\varphi \in \mathrm{SBV}(\mathrm{B}_\varrho) \cap \mathrm{W}^{1,2}(\mathrm{B}_\varrho \setminus H)$, $\varphi = h - w$ with zero boundary values at $\partial \mathrm{B}_\varrho$ (see [AFP00, Theoreme 3.84 and Theorem 3.87]). This means, $\varphi$ is a valid comparison function for $w$ (at first in a slightly bigger ball $\mathrm{B}_{\varrho'}$ where $\varphi$ has compact support, then taking the limit $\varrho' \searrow \varrho$) and from (5.7) we infer

$$\begin{aligned} \int_{\mathrm{B}_\varrho} |\nabla \varphi|^2 \,\mathrm{d}y &= \int_{\mathrm{B}_\varrho} |\nabla w|^2 \,\mathrm{d}y - \int_{\mathrm{B}_\varrho} |\nabla h|^2 \,\mathrm{d}y + \int_{\mathrm{B}_\varrho} 2\nabla w \cdot \nabla h \,\mathrm{d}y \\ &\le \hat{C} r^{2\beta} \int_{\mathrm{B}_\varrho} |\nabla \varphi|^2 \,\mathrm{d}y + \hat{C} r^{2\beta} \int_{\mathrm{B}_s} \left(1 + |\nabla w|^2\right) \mathrm{d}y. \end{aligned}$$

For $r$ small enough, we can absorb one term to deduce

$$⨍_{\mathrm{B}_\varrho} |\nabla h_+ - \nabla w_+|^2 \,\mathrm{d}y + ⨍_{\mathrm{B}_\varrho} |\nabla h_- - \nabla w_-|^2 \,\mathrm{d}y \le \bar{C} r^{2\beta} ⨍_{\mathrm{B}_s} \left(1 + |\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y. \tag{5.8}$$

**Step 2:** For $i \in \mathbb{N}_0$ and $0 < \tau \le \frac{1}{2}$ we define the excess as

$$E(\tau^i \varrho) := ⨍_{\mathrm{B}_{\tau^i \varrho}} |\nabla w_+ - (\nabla w_+)_{\tau^i \varrho}|^2 \,\mathrm{d}y + ⨍_{\mathrm{B}_{\tau^i \varrho}} |\nabla w_- - (\nabla w_-)_{\tau^i \varrho}|^2 \,\mathrm{d}y.$$

Using (5.8), Lemma 2.1, Lemma 2.2 we follow the calculations of Step 2 in Lemma 5.1. In short this yields

$$\begin{aligned} E(\tau \varrho) &\le \widetilde{C} \left( \tau^2 E(\varrho) + \frac{2}{\tau^n} ⨍_{\mathrm{B}_\varrho} |\nabla w_+ - \nabla h_+|^2 \,\mathrm{d}y + \frac{2}{\tau^n} ⨍_{\mathrm{B}_\varrho} |\nabla w_- - \nabla h_-|^2 \,\mathrm{d}y \right) \\ &\le \tau^{2\gamma} E(\varrho) + C_* \varrho^{2\beta} ⨍_{\mathrm{B}_s} \left(1 + |\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y\,, \end{aligned}$$

where we have chosen $0 < \tau \leq \frac{1}{2}$ such that $\widetilde{C}\tau^2 \leq \tau^{2\gamma}$ for some $\gamma \in (\beta, 1)$.

**Step 3:** Again, the previous calculations can be easily transfered to smaller balls $\mathrm{B}_{\tau^i \varrho}$, so that we get

$$E(\tau^{i+1}\varrho) \leq \tau^{2\gamma} E(\tau^i \varrho) + C_*(\tau^i \varrho)^{2\beta} \fint_{\mathrm{B}_{\tau^i s}} \left(1 + |\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y \tag{5.9}$$

for $i \in \mathbb{N}_0$. Set

$$M_i := \tau^{2i(\gamma-\beta)} + 2C_* \tau^{-2\beta} \sum_{l=0}^{i-1} \tau^{2l(\gamma-\beta)} \qquad \text{for } i \in \mathbb{N}_0$$

and recall, that the sequence $M_i$ is non-decreasing and bounded by a value $M$.

The following induction is mostly similiar to the one in the proof of Lemma 5.1, but we examine both functions $w_+$ and $w_-$ at once, and the transformation between the radii $\varrho = \frac{8}{10} r = \frac{8}{11} s$ gives rise to some additional factors. Our goal is to prove for every $r \in (0, r'')$ and $i \in \mathbb{N}_0$ the hypothesis

$$E(\tau^i \varrho) \leq M_i \tau^{2i\beta} \left(\varrho^{2\beta} + 2^n \fint_{\mathrm{B}_s} \left(|\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y\right), \tag{5.10}$$

where $\varrho = \frac{8}{10} r$ and $r'' := \min\{[M\tau^{-2n}(1 + \frac{1}{1-\tau^\beta})^2]^{\frac{1}{-2\beta}}, (2\hat{C})^{-\beta}, 1\}$.

Base cases: For $i = 0$, we get $M_0 = 1$, and using Lemma 2.2 and $s/\varrho \leq 2$ we immediately have

$$E(\varrho) \leq \fint_{\mathrm{B}_\varrho} \left(|\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y \leq \varrho^{2\beta} + 2^n \fint_{\mathrm{B}_s} \left(|\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y\,.$$

We also treat the case $i = 1$ as a base case. In this case the claim follows from the case $i = 0$ and (5.9) as follows:

$$\begin{aligned}
&E(\tau\varrho) \leq \tau^{2\gamma} E(\varrho) + C_* \varrho^{2\beta} \fint_{\mathrm{B}_s} \left(1 + |\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y \\
&\leq \tau^{2\gamma}\varrho^{2\beta} + \tau^{2\gamma} 2^n \fint_{\mathrm{B}_s} \left(|\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y + C_* \varrho^{2\beta} \fint_{\mathrm{B}_s} \left(1 + |\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y \\
&\leq (\tau^{2\gamma} + 2C_*) \left(\varrho^{2\beta} + 2^n \fint_{\mathrm{B}_s} \left(|\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y\right) \\
&= M_1 \tau^{2\beta} \left(\varrho^{2\beta} + 2^n \fint_{\mathrm{B}_s} \left(|\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y\right).
\end{aligned}$$

Inductive step: Suppose (5.10) is true for $i = 0, 1, 2, \ldots, k$ with a positive integer $k$. Then we show that it holds for $i = k + 1$. The calculations are mostly the same as before, but we need to consider an integral with domain $\mathrm{B}_{\tau^{k-1}\varrho}$ (which makes sense since $k - 1 \geq 0$). We first use the estimate $\tau s \leq \frac{1}{2} \cdot \frac{11}{8} \varrho \leq \varrho$, then add and subtract mean values iteratively. In this way we arrive at

$$\begin{aligned}
\Big(\fint_{\mathrm{B}_{\tau^k s}} \left(|\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y\Big)^{\frac{1}{2}} &\leq \Big(\tau^{-n} \fint_{\mathrm{B}_{\tau^{k-1}\varrho}} \left(|\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y\Big)^{\frac{1}{2}} \\
&\leq \tau^{-n/2} \Big(\fint_{\mathrm{B}_\varrho} \left(|\nabla w_+|^2 + |\nabla w_-|^2\right) \mathrm{d}y\Big)^{\frac{1}{2}} + \tau^{-n} \sum_{i=0}^{k-1} E(\tau^i \varrho)^{\frac{1}{2}}.
\end{aligned}$$

Using the inductive hypothesis we get

$$\sum_{i=0}^{k-1} E(\tau^i\varrho)^{\frac{1}{2}} \le \sum_{i=0}^{k-1}\left(M_i\tau^{2i\beta}\Big(\varrho^{2\beta}+2^n\fint_{\mathrm{B}_s}\big(|\nabla w_+|^2+|\nabla w_-|^2\big)\,\mathrm{d}y\Big)\right)^{\frac{1}{2}}$$
$$\le \Big(\frac{1}{1-\tau^\beta}\Big)M_k^{\frac{1}{2}}\left(\varrho^{2\beta}+2^n\fint_{\mathrm{B}_s}\big(|\nabla w_+|^2+|\nabla w_-|^2\big)\,\mathrm{d}y\right)^{\frac{1}{2}}.$$

Setting $C_\tau := \tau^{-n}(1+\frac{1}{1-\tau^\beta})$ we conclude

$$\fint_{\mathrm{B}_{\tau^k s}}\big(|\nabla w_+|^2+|\nabla w_-|^2\big)\,\mathrm{d}y \le C_\tau^2 M_k\left(\varrho^{2\beta}+2^n\fint_{\mathrm{B}_s}\big(|\nabla w_+|^2+|\nabla w_-|^2\big)\,\mathrm{d}y\right).$$

From the estimate (5.9), the preceding inequality, and the hypothesis (5.10) for $i=k$, we infer

$$E(\tau^{k+1}\varrho) \le \tau^{2\gamma}E(\tau^k\varrho)+C_*(\tau^k\varrho)^{2\beta}\fint_{\mathrm{B}_{\tau^k s}}\big(1+|\nabla w_+|^2+|\nabla w_-|^2\big)\,\mathrm{d}y$$
$$\le \tau^{2\gamma}E(\tau^k\varrho)+C_*(\tau^k\varrho)^{2\beta}\left(1+C_\tau^2M_k\Big(\varrho^{2\beta}+2^n\fint_{\mathrm{B}_s}\big(|\nabla w_+|^2+|\nabla w_-|^2\big)\,\mathrm{d}y\Big)\right)$$
$$\le \tau^{2(k+1)\beta}\Big(C_*\tau^{-2\beta}(1+C_\tau^2M_k\varrho^{2\beta})+\tau^{2(\gamma-\beta)}M_k\Big)\Big(\varrho^{2\beta}+2^n\fint_{\mathrm{B}_s}\big(|\nabla w_+|^2+|\nabla w_-|^2\big)\,\mathrm{d}y\Big).$$

In view of the choice of $r''$ it follows $C_\tau^2M_k\varrho^{2\beta}\le 1$ and we conclude

$$E(\tau^{k+1}\varrho) \le \tau^{2(k+1)\beta}\Big(2C_*\tau^{-2\beta}+\tau^{2(\gamma-\beta)}M_k\Big)\Big(\varrho^{2\beta}+2^n\fint_{\mathrm{B}_s}\big(|\nabla w_+|^2+|\nabla w_-|^2\big)\,\mathrm{d}y\Big)$$
$$=\tau^{2(k+1)\beta}M_{k+1}\Big(\varrho^{2\beta}+2^n\fint_{\mathrm{B}_s}\big(|\nabla w_+|^2+|\nabla w_-|^2\big)\,\mathrm{d}y\Big).$$

**Step 4:** Notice, that the deduced estimates are also true for balls $\mathrm{B}_\varrho(z)$ with $z\in\mathrm{B}_{r/50}\cap H$, because we still have $\mathrm{B}_\varrho(z)\subset\mathrm{B}_{\frac{9}{10}r}$ and the hyperplane $H$ splits the ball in two half-balls.

We now want an estimate for the gradient $\nabla w$ for Lebesgue points $\hat{y}\in\mathrm{B}_{r/50}\setminus H$ which again fulfill $\mathrm{B}_\varrho(\hat{y})\subset\mathrm{B}_{\frac{9}{10}r}$. Let $z\in\mathrm{B}_{r/50}\cap H$ be such that $|\hat{y}-z|=\operatorname{dist}(\hat{y},H)$. Let $l\in\mathbb{N}_0$ be the power of $\tau$, such that $\mathrm{B}_{\tau^l\varrho}(\hat{y})\cap H\neq\emptyset$ and $\mathrm{B}_{\tau^{l+1}\varrho}(\hat{y})\cap H=\emptyset$. This means, $\mathrm{B}_{\tau^l\varrho}(\hat{y})$ is the last ball of the sequence $(\mathrm{B}_{\tau^i\varrho}(\hat{y}))_{i\in\mathbb{N}_0}$ that intersects $H$, thus we have $\operatorname{dist}(\hat{y},H)\le\tau^l\varrho$. The basic idea is to use an estimate like in Lemma 5.1 for the small balls around $\hat{y}$ that do not intersect $H$, and the remaining balls can be estimated by balls around $z\in H$ where we can use (5.10). Because of $\tau\le\frac{1}{2}$ we have $\mathrm{B}_{\tau^{l+1}\varrho}(\hat{y})\subset\mathrm{B}_{\tau^{l-1}\varrho}(z)$ and this yields the helpful estimate

$$\fint_{\mathrm{B}_{\tau^{l+1}\varrho}(\hat{y})}|\nabla w|\,\mathrm{d}y \le \frac{1}{\tau^{2n}}\fint_{\mathrm{B}_{\tau^{l-1}\varrho}(z)}(|\nabla w_+|+|\nabla w_-|)\,\mathrm{d}y. \tag{5.11}$$

(We can assume $l\ge1$, because the case $l=0$ can be directly calculated by comparison with the ball $\mathrm{B}_s$, so that $z\in H$ is not needed.) In order to control

$$|\nabla w(\hat{y})| = \lim_{k\to\infty}|(\nabla w)_{\mathrm{B}_{\tau^k\varrho}(\hat{y})}| \le |(\nabla w)_{\mathrm{B}_{\tau^{l+2}\varrho}(\hat{y})}|+\sum_{i=0}^{\infty}|(\nabla w)_{\mathrm{B}_{\tau^{i+l+3}\varrho}(\hat{y})}-(\nabla w)_{\mathrm{B}_{\tau^{i+l+2}\varrho}(\hat{y})}|, \tag{5.12}$$

we first treat the sum on the right side. Rewriting a single summand we can estimate it by

$$\tau^{-n} \fint_{B_{\tau^{i+l+2}\varrho}(\hat{y})} |\nabla w - (\nabla w)_{B_{\tau^{i+l+2}\varrho}(\hat{y})}| \,\mathrm{d}y \le \tau^{-\frac{n}{2}} [E(\tau^i(\tau^{l+2}\varrho); \hat{y})]^{\frac{1}{2}}.$$

Notice that $w \in W^{1,2}(B_{\tau^{l+1}\varrho}(\hat{y}))$, which allows us to follow the exact steps of Lemma 5.1 (using the transformed minimality condition (5.7)) to get an estimate for the excess. This results in

$$\sum_{i=0}^{\infty} [E(\tau^i(\tau^{l+2}\varrho); \hat{y})]^{\frac{1}{2}} \le \left( C \fint_{B_{\tau^{l+1}\varrho}(\hat{y})} \left( |\nabla w|^2 + r^{2\beta} \right) \mathrm{d}y \right)^{\frac{1}{2}}$$

with the obvious notation for the excess on balls centered at $\hat{y}$. Using (5.11) to pass over to $B_{\tau^{l-1}\varrho}(z)$ we can use a similar argument utilizing (5.10) to conclude

$$\begin{aligned}
\left( \fint_{B_{\tau^{l+1}\varrho}(\hat{y})} |\nabla w|^2 \,\mathrm{d}y \right)^{\frac{1}{2}} &\le \left( \frac{1}{\tau^{2n}} \fint_{B_{\tau^{l-1}\varrho}(z)} \left( |\nabla w_+|^2 + |\nabla w_-|^2 \right) \mathrm{d}y \right)^{\frac{1}{2}} \\
&\le \tau^{-\frac{3}{2}} [E(\tau^{l-2}\varrho); z]^{\frac{1}{2}} + \tau^{-n} \Big( |(\nabla w_+)_{B_{\tau^{l-2}\varrho}(z)}| + |(\nabla w_-)_{B_{\tau^{l-2}\varrho}(z)}| \Big) \\
&\le 2^{n+2} \tau^{-n} \left( \fint_{B_s} |\nabla w|^2 \,\mathrm{d}y \right)^{\frac{1}{2}} + 2\tau^{-\frac{3n}{2}} \sum_{i=0}^{l-2} [E(\tau^i \varrho; z)]^{\frac{1}{2}} \\
&\le 2^{n+2} \tau^{-n} \left( \fint_{B_s} |\nabla w|^2 \,\mathrm{d}y \right)^{\frac{1}{2}} + 4\tau^{-\frac{3n}{2}} \sum_{i=0}^{\infty} M^{\frac{1}{2}} \tau^{i\beta} \left( \varrho^{2\beta} + 2^n \fint_{B_s} |\nabla w|^2 \,\mathrm{d}y \right)^{\frac{1}{2}} \\
&\le \left( C \fint_{B_s} \left( |\nabla w|^2 + r^{2\beta} \right) \mathrm{d}y \right)^{\frac{1}{2}}.
\end{aligned}$$

We can use the same argument to estimate the remaining term $|(\nabla w)_{B_{\tau^{l+2}\varrho}(\hat{y})}|$ of equation (5.12), and in conclusion (5.12) yields

$$|\nabla w(\hat{y})| \le \left( C \fint_{B_s} \left( |\nabla w|^2 + r^{2\beta} \right) \mathrm{d}y \right)^{\frac{1}{2}}.$$

Transforming this estimate back to $u$ generates once again factors $T_r = 2nL\left(\frac{11}{10}r\right)^{\alpha} + n^2L^2\left(\frac{11}{10}r\right)^{2\alpha}$ which can be estimated by a dimensional constant because of the assumption $r^{\alpha}\|\nabla f\|_{C^{0,\alpha}} \le \frac{1}{10}$. For $\hat{x} \in B_{r/100}$ we clearly have $\hat{y} = \Phi(\hat{x}) \in B_{r/50}$ and it follows

$$\begin{aligned}
|\nabla u(\hat{x})|^2 &= |D\Phi(\hat{x}) \cdot \nabla w(\hat{y})|^2 \le (1 + T_{r/100}) |\nabla w(\hat{y})|^2 \\
&\le (1 + T_{r/100}) C \fint_{B_s} \left( |\nabla w|^2 + r^{2\beta} \right) \mathrm{d}y \\
&\le (1 + T_{r/100}) C (1 + T_{r/100}) \fint_{\Phi^{-1}(B_s)} \left( |\nabla u|^2 + r^{2\beta} \right) \mathrm{d}x \\
&\le C'' \fint_{B_{2r}} \left( |\nabla u|^2 + r^{2\beta} \right) \mathrm{d}x.
\end{aligned}$$

Finally, because almost every point is a Lebesgue point, we have

$$\sup_{x\in \mathrm{B}_{\frac{r}{100}}} |\nabla u(x)|^2 \le C'' \fint_{\mathrm{B}_{2r}} \left(|\nabla u|^2 + r^{2\beta}\right) \mathrm{d}x\,.$$

This completes the proof. □

In the next statement, we summarize the highly developed partial regularity theory for Mumford-Shah almost-minimizers with its culmination in the porosity of the set of singular points of $\mathrm{S}_u$, and we also incorporate — the decisive feature for our purposes — the gradient estimate near regular points of $\mathrm{S}_u$.

**Proposition 5.3.** *If $u \in \mathrm{SBV}(\Omega)$ is an $(a,b)$-almost-minimizer of* MS*, then there exist $r_0 > 0$, $\frac{1}{15} > \varepsilon > 0$, and $L_0 > 0$, such that for all $x \in \overline{\mathrm{S}}_u$ with $\mathrm{B}_r(x) \Subset \Omega$ and $r < r_0$, we find $y \in \mathrm{B}_{r/2}(x) \cap \overline{\mathrm{S}}_u$ such that*

$$\overline{\mathrm{S}}_u \cap \mathrm{B}_{2r/L_0}(y) = [y + \Gamma] \cap \mathrm{B}_{2r/L_0}(y),$$

*where $\Gamma$ is a rotated graph of a $\mathrm{C}^{1,\alpha}$ function $f : \mathbb{R}^{n-1} \to \mathbb{R}$ with $f(0) = 0$ and $\alpha = \frac{\min\{a,1/2\}}{2(n+1)}$. Furthermore, the $\mathrm{L}^\infty$-norm and the $\mathrm{C}^{0,\alpha}$-seminorm of $f$ can be bounded by*

$$\|\nabla f\|_\infty + r^\alpha [\nabla f]_{\mathrm{C}^{0,\alpha}} \le \varepsilon\,, \tag{5.13}$$

*and we have*

$$r \sup_{\mathrm{B}_{2r/L_0}(y)} |\nabla u|^2 \le L_0^2. \tag{5.14}$$

*Proof.* Taking into account that $u$ fulfills the requirements of the almost-minimizer concept used in [Dav05], we employ [Dav05, Corollary 75.15]. Thus, we find a $y \in \mathrm{B}_{r/2}(x) \cap \overline{\mathrm{S}}_u$ such that $\overline{\mathrm{S}}_u$ coincides in $\mathrm{B}_s(y)$ with a rotation of the graph of a $\mathrm{C}^1$ function $f$, for some $C \gg 1$ and $\frac{200}{C} r \le s < \frac{1}{6} r$. Indeed, the proof of this statement in [Dav05] basically follows the reasoning for true minimizers in [Dav96, Rig00]. It first establishes, for some $\varepsilon < \frac{1}{15}$, the smallness condition for the excess (cf. [Dav05, Theorem 75.2] and the remarks thereafter)

$$s^{1-n}\left(\int_{\mathrm{B}_{2s}(y)} |\nabla u(z)|^2 \,\mathrm{d}z + s^{-2} \inf_{A\in\mathcal{A}} \int_{\mathrm{S}_u \cap \mathrm{B}_{2s}(y)} \operatorname{dist}(z,A)^2 \,\mathrm{d}\mathcal{H}^{n-1}(z)\right) \le \varepsilon$$

(where $\mathcal{A}$ is the set of all affine hyperplanes in $\mathbb{R}^n$) and then applies the $\varepsilon$-regularity theorem [AFP97, Theorem 3.1]. From the latter theorem we read off that $f$ can even be taken $C^{1,\alpha}$ with the exponent $\alpha$ stated above, and by tracing the corresponding estimates in [AP97, Theorem 5.3, Remark 5.4, Lemma 6.1], [AFP97, Corollary 6.2] we arrive at (5.13); compare also [AFP00, Theorems 8.1, 8.2, 8.3] for the case $a \ge 1$. In view of (5.13) we can finally apply Lemma 5.2 on $\mathrm{B}_s(y)$ (with any admissible choice of $\beta$) to obtain

$$\begin{aligned}\sup_{\mathrm{B}_{2r/C}(y)} r|\nabla u|^2 \leq \sup_{\mathrm{B}_{\frac{s}{200}}(y)} r|\nabla u|^2 &\leq rC'' \fint_{\mathrm{B}_s(y)} \left(|\nabla u|^2 + s^{2\beta}\right) \mathrm{d}z \\ &\leq C'' \frac{r}{s\omega_n} s^{1-n} \int_{\mathrm{B}_{2s}(y)} |\nabla u|^2 \,\mathrm{d}z + C'' r s^{2\beta} \\ &\leq \frac{C''}{400\omega_n} C\varepsilon + C'' r_0^{2\beta+1} \leq L_0^2,\end{aligned}$$

for sufficiently large $L_0$. This shows (5.14), and the proof is complete. □

# 6 Proof of the higher integrability result

This section follows closely [DPF14]. However, while in [DPF14] gradient estimates follow easily from the fact that $u$ is harmonic in $\Omega \setminus K$ and solves a Neumann problem, this basic reasoning is longer available in our case of almost-minimizers. We thus rely, as a substitute, on the gradient estimates of Section 5 and also stay in SBV setting where $\overline{\mathrm{S}}_u$ replaces $K$. Still, since the reasoning remains close enough to [DPF14], we only provide a comparably brief rereading and refer to [DPF14] for full details.

As usual, let $u \in \mathrm{SBV}(\Omega)$ be an $(a, b)$-almost-minimizer of MS on an open bounded set $\Omega \subset \mathbb{R}^n$. We work on a fixed ball $\mathrm{B}_{2r_0}(x_0) \Subset \Omega$ with radius $r_0$ small enough that the requirements of Corollary 3.3, Lemma 5.1 and Proposition 5.3 are fulfilled. To simplify notation we omit the center for every ball centered at $x_0$.

In the following we examine the superlevel sets

$$A_h := \{x \in \mathrm{B}_{2r_0} \setminus \overline{\mathrm{S}}_u \;:\; |\nabla u(x)|^2 \geq M^{h+1}\}$$

for $M \gg 1$ and $h \in \mathbb{N}$.

With the help of the gradient estimates away from $\mathrm{S}_u$ we first establish the following lemma, which in turn plays a role in the full proof of Lemma 6.2.

**Lemma 6.1.** *There exists $M_0 > 0$, such that for $M \geq M_0$ and $r \leq r_0$ we have*

$$A_h \cap \mathrm{B}_{r-2M^{-h}} \subset \mathcal{N}_{M^{-h}}(\overline{\mathrm{S}}_u \cap \overline{\mathrm{B}}_r) \qquad \text{for every } h \in \mathbb{N}.$$

*Proof.* Using the notation of Corollary 3.3 and Lemma 5.1 we make the choice $M_0 := \max\{C'(C_0/\omega_n+2), 1/r_0, 1/r'\}$ and assume $M \geq M_0$. Let $h \in \mathbb{N}$, $x \in A_h \cap \mathrm{B}_{r-2M^{-h}} \neq \emptyset$ be a Lebesgue point of $\nabla u$, $d := \mathrm{dist}(x, \overline{\mathrm{S}}_u)$ and $z \in \overline{\mathrm{S}}_u$ such that $|x - z| = d$.

We assume $d > M^{-h}$ and argue by contradiction. It follows $\mathrm{B}_{M^{-h}}(x) \cap \overline{\mathrm{S}}_u = \emptyset$ and because of our choice of $M_0$, the requirements for the energy upper bound (3.4) and the gradient estimate (5.1) away from $\mathrm{S}_u$ are fulfilled on $\mathrm{B}_{M^{-h}}(x)$. Using these estimates and the definition of $A_h$ we conclude

$$M^{h+1} \le |\nabla u(x)|^2 \le C' \fint_{\mathrm{B}_{M^{-h}}(x)} \Big(|\nabla u|^2 + M^{-h2\beta}\Big)\, \mathrm{d}y$$
$$\le C'\Big(\frac{C_0}{\omega_n} M^h + M^{-h2\beta}\Big) \le C'\Big(\frac{C_0}{\omega_n} + 1\Big) M^h.$$

This gives $M \le C'(C_0/\omega_n + 1)$ which is impossible, since we took $M_0 \ge C'(C_0/\omega_n + 2)$. Therefore, we have $d \le M^{-h}$, which yields $z \in \overline{\mathrm{B}}_r$ and $x \in \mathcal{N}_{M^{-h}}(\overline{\mathrm{S}}_u \cap \overline{\mathrm{B}}_r)$. Taking into account that almost every point is a Lebesgue point, this proves the lemma. □

**Lemma 6.2.** *Assume that $\varepsilon$ and $L_0$ are as in Proposition* 5.3*. Then there exist $C_1, C_2, M_2 \ge 1$, $\alpha \in (0, \frac{1}{4})$ and sequences of radii $\{R_h\}_{h\in\mathbb{N}}$, $\{S_h\}_{h\in\mathbb{N}}$, such that for $M \ge M_2$ and every $h \in \mathbb{N}$ we have:*

1. *The radii fulfill*
   - $r_0 \ge R_h \ge S_h \ge R_{h+1} \ge r_0/2$,
   - $R_h - R_{h+1} \le M^{-\frac{h+1}{2}}$ *and* $S_h - R_{h+1} = 8M^{-(h+1)}$,
   - $\mathcal{H}^{n-1}\Big(\overline{\mathrm{S}}_u \cap (\overline{B}_{S_h} \setminus \overline{B}_{R_{h+1}})\Big) \le C_1 M^{-\frac{h+1}{2}}$.
2. *We can find suitable sets $K_h \subset (\overline{\mathrm{S}}_u \cap \overline{\mathrm{B}}_{S_h})$ which describe the "bad parts" of $\overline{\mathrm{S}}_u$ in such a way that the size of the superlevel sets $A_h$ can be estimated by*

$$|A_{h+2} \cap \mathrm{B}_{R_{h+2}}| \le C_2 M^{-(h+1)} \mathcal{H}^{n-1}(K_h) \tag{6.1}$$

*and the size of $K_h$ is bounded by*

$$\mathcal{H}^{n-1}(K_h) \le C_1 h M^{-2\alpha(h-1)}. \tag{6.2}$$

*Proof.* For the complete formulation of the lemma and its proof we refer to [DPF14, Lemma 3.3]. The only difference occurs in Step 3 of the proof given there, where we have to use Proposition 5.3 to find a family

$$\mathcal{F}_{h+1} := \{\mathrm{B}_{M^{-(h+1)}/L_0}(y_i)\}_{i\in I_h}$$

that fulfills

$$\sup_{\mathrm{B}_{2M^{-(h+1)}/L_0}(y_i)} |\nabla u|^2 \le L_0^2 M^{h+1} < M^{h+2}$$

for sufficiently large $M$. The remainder of the proof is analogous to [DPF14]. □

*Proof of Theorem* 1.1*.* Fix $M := M_2$. Combining (6.1) and (6.2) yields

$$|A_{h+2} \cap \mathrm{B}_{R_{h+2}}| \le C_1 C_2 h M^{-h(1+2\alpha)}$$

for all $h \geq 2$. Using the definition of $A_h$ and $R_h \geq r_0/2$ we get

$$|\{x \in \mathrm{B}_{r_0/2} \setminus \overline{\mathrm{S}}_u : |\nabla u|^2(x) \geq M^h\}| \leq C_1 C_2 M^{3(1+2\alpha)} h M^{-h(1+2\alpha)} \tag{6.3}$$

for all $h \geq 5$. Setting $\gamma = 1 + \alpha$ this implies

$$\begin{aligned} \int_{\mathrm{B}_{r_0/2}} |\nabla u|^{2\gamma} \, \mathrm{d}x &= \gamma \int_0^\infty t^{\gamma-1} |(\mathrm{B}_{r_0/2} \setminus \overline{\mathrm{S}}_u) \cap \{|\nabla u|^2 \geq t\}| \, \mathrm{d}t \\ &\leq C + \gamma M^\gamma \sum_{h=5}^\infty M^{h\gamma} |(\mathrm{B}_{r_0/2} \setminus \overline{\mathrm{S}}_u) \cap \{|\nabla u|^2 \geq M^h\}| \\ &\leq C + \widetilde{C} \sum_{h=5}^\infty h (M^{\gamma-1-2\alpha})^h < \infty. \end{aligned}$$

Since $\mathrm{B}_{r_0/2}$ denotes a ball with arbitrary center $x_0 \in \Omega$ (and sufficiently small radius), this implies $\nabla u \in \mathrm{L}^{2\gamma}(\Omega', \mathbb{R}^n)$ for every open $\Omega' \Subset \Omega$ and completes the proof of the theorem. □

# References


[AFH03] L. Ambrosio, N. Fusco, and J. E. Hutchinson. Higher integrability of the gradient and dimension of the singular set for minimisers of the Mumford-Shah functional. *Calc. Var. Partial Differ. Equ.*, 16(2):187–215, 2003.

[AFP97] L. Ambrosio, N. Fusco, and D. Pallara. Partial regularity of free discontinuity sets, II. *Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser.*, 24(1):39–62, 1997.

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems.* Oxford Science Publications. Clarendon Press, 2000.

[Amb89] L. Ambrosio. A compactness theorem for a new class of functions of bounded variation. *Boll. Unione Mat. Ital., VII. Ser., B*, 3(4):857–881, 1989.

[Amb90] L. Ambrosio. Existence theory for a new class of variational problems. *Arch. Ration. Mech. Anal.*, 111(4):291–322, 1990.

[Anz83] G. Anzellotti. On the $C^{1,\alpha}$-regularity of $\omega$-minima of quadratic functionals. *Boll. Unione Mat. Ital., VI. Ser., C, Anal. Funz. Appl.*, 2:195–212, 1983.

[AP97] L. Ambrosio and D. Pallara. Partial regularity of free discontinuity sets, I. *Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser.*, 24(1):1–38, 1997.

[Bon96] A. Bonnet. On the regularity of edges in image segmentation. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, 13(4):485–528, 1996.

[Dav96] G. David. $C^1$-arcs for minimizers of the Mumford-Shah functional. *SIAM J. Appl. Math.*, 56(3):783–888, 1996.

[Dav05] G. David. *Singular Sets of Minimizers for the Mumford-Shah Functional.* Progress in Mathematics. Birkhäuser, 2005.

[DGA88] E. De Giorgi and L. Ambrosio. Un nuovo tipo di funzionale del calcolo delle variazioni. *Atti Accad. Naz. Lincei, VIII. Ser., Rend., Cl. Sci. Fis. Mat. Nat.*, 82(2):199–210, 1988.

[DGA89] E. De Giorgi and L. Ambrosio. New functionals in calculus of variations. In *Nonsmooth Optimization and Related Topics* (edited by F.H. Clarke, V.F. Dem'yanov, and F. Giannessi), pages 49–59. Springer, Boston, 1989.

[DGCL89] E. De Giorgi, M. Carriero, and A. Leaci. Existence theorem for a minimum problem with free discontinuity set. *Arch. Ration. Mech. Anal.*, 108(3):195–218, 1989.

[DGG00] F. Duzaar, A. Gastel, and J. F. Grotowski. Partial regularity for almost minimizers of quasi-convex integrals. *SIAM J. Math. Anal.*, 32(3):665–687, 2000.

[DLF13] C. De Lellis and M. Focardi. Higher integrability of the gradient for minimizers of the 2d Mumford-Shah energy. *J. Math. Pures Appl. (9)*, 100(3):391–409, 2013.

[DLFR14] C. De Lellis, M. Focardi, and B. Ruffini. A note on the Hausdorff dimension of the singular set for minimizers of the Mumford-Shah energy. *Adv. Calc. Var.*, 7(4):539–545, 2014.

[DPF14] G. De Philippis and A. Figalli. Higher integrability for minimizers of the Mumford-Shah functional. *Arch. Ration. Mech. Anal.*, 213(2):491–502, 2014.

[Foc16] M. Focardi. Fine regularity results for Mumford-Shah minimizers: porosity, higher integrability and the Mumford-Shah conjecture. In *Free discontinuity problems* (edited by N. Fusco and A. Pratelli), pages 1–68. Edizioni della Normale, Scuola Normale Superiore, Pisa, 2016.

[Fus03] N. Fusco. An overview of the Mumford-Shah problem. *Milan J. Math.*, 71(1):95–119, 2003.

[Leg99] J. C. Leger. Flatness and finiteness in the Mumford-Shah problem. *J. Math. Pures Appl. (9)*, 78(4):431–459, 1999.

[MS89] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math.*, 42(5):577–685, 1989.

[MS01a] F. Maddalena and S. Solimini. Concentration and flatness properties of the singular set of bisected balls. *Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser.*, 30(3-4):623–659, 2001.

[MS01b] F. Maddalena and S. Solimini. Regularity properties of free discontinuity sets. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, 18(6):675–685, 2001.

[Pio16] S. Piontek. *Höhere Integrierbarkeit für Fast-Minimierer des Mumford-Shah-Funktionals*. Master Thesis. University Erlangen-Nuremberg, 2016.

[Rig00] S. Rigot. Big pieces of $C^{1,\alpha}$-graphs for minimizers of the Mumford-Shah functional. *Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser.*, 29(2):329–349, 2000.