Inverse Problems and Imaging
Volume X, No. 0X, 20xx, X–XX
doi:10.3934/ipi.xx.xx.xx

POINT-WISE BEHAVIOR OF THE GEMAN–McCLURE AND THE HEBERT–LEAHY IMAGE RESTORATION MODELS

Petteri Harjulehto

Department of Mathematics and Statistics
FI-20014 University of Turku, Finland

Peter Hästö and Juha Tiirola

Department of Mathematical Sciences
FI-90014 University of Oulu, Finland

(Communicated by the associate editor name)

Abstract. We present new continuous variants of the Geman–McClure model and the Hebert–Leahy model for image restoration, where the energy is given by the nonconvex function $x \mapsto x^2/(1+x^2)$ or $x \mapsto \log(1+x^2)$, respectively. In addition to studying these models' Γ-convergence, we consider their point-wise behaviour when the scale of convolution tends to zero. In both cases the limit is the Mumford-Shah functional.

1. **Introduction.** A basic denoising problem encountered in image processing is obtaining an estimate u for an unknown image u_0 based on a corrupted observation f. In a variational approach, one seeks to solve this problem by minimizing an energy functional which typically consists of two parts, a fidelity term such as $\int |u-f|^2\,dx$ and a regularity term such as $\int |\nabla u|\,dx$. More generally, one can consider the class of problems:

$$\min_{u\in \mathrm{BV}} \int_\Omega \varphi(|Du|) + |u-f|^2\,dx,$$

where Du denotes the gradient of the BV function u; if $\frac{\varphi(t)}{t} \to \varphi_\infty$ as $t\to\infty$, then

$$\int_\Omega \varphi(|Du|)\,dx := \int_\Omega \varphi(|\nabla_a u|)\,dx + \varphi_\infty\,|D_s u|(\Omega),$$

where $\nabla_a u$ and $D_s u$ are the absolutely continuous and singular parts of Du. Note that the choice $\varphi(t)=t$ leads to the Total Variation (ROF) denoising model of Rudin, Osher and Fatemi [24] in 1992.

In their discussion about variational approaches to image restoration, Aubert and Kornprobst [5, Section 3.2.6] point out that non-convex potentials such as $\varphi(t) = \frac{t^2}{1+t^2}$ give better results in numerical tests than convex potentials such as $\varphi(t)=t$. In fact, this choice of φ leads to a discrete model which was first proposed by Geman and McClure [18] in 1985, many years before the ROF model. This model has been used, for instance, in tomography, road scene analysis, magnetic resonance imaging, geophysical imaging, volume reconstruction, target detection, functional MRI and image segmentation. Hebert and Leahy [20] presented in 1989 a somewhat

---

2010 *Mathematics Subject Classification.* Primary: 94A08 Secondary: 49J45, 49N45, 68U10).
Key words and phrases. Mumford-Shah, nonconvex, minimization problem, denoising, inverse problem, Gamma-convergence.

similar model whose potential is $t \mapsto \log(1+t^2)$; in a differential equation form this potential corresponds to the Perona–Malik model.

However, it was shown by Chipot, March, Rosati and Vergara Caffarelli [11] that the functional corresponding to this choice of φ,

$$\int_\Omega \frac{|Du|^2}{1+|Du|^2} + |u-f|^2\,dx, \tag{1}$$

does not have a minimizer in BV and that the energy minimum of the functional equals 0. The corresponding discrete model

$$\sum_i \frac{|D_h u_i|^2}{1+|D_h u_i|^2} + |u_i - f_i|^2, \tag{2}$$

where D_h is the discrete gradient with step size h, performs very well, see [8, 9, 19, 23]. Rosati [23] has shown that this discrete model Γ-convergences to a modified Mumford-Shah functional as the step size tends to zero. This corresponds to sampling a continuous function at individual points. Another way to get from the continuous to the discrete is to take the average of a gradient of the function in some pixel. (After smoothing one can equivalently use integrals or sums.) This route was followed by Braides and Dal Maso [7] and Chambolle and Dal Maso [10]; more on this below.

We also follow an averaging approach. To be more precise, let $G \in C^1(\mathbb{R}^n)$ be a non-negative, radial symmetric and decreasing function with $\operatorname{spt} G \subset B(0,1)$ and $\int G\,dx = 1$. Let $G_\sigma(x) := \frac{1}{\sigma^n} G(\frac{x}{\sigma})$ be the L^1-scaled version of G. For a (scalar- or vector valued) function $u \in L^1_{\rm loc}(\mathbb{R}^n)$, we denote

$$u_\sigma(x) := G_\sigma * u(x) = \int_{B(x,\sigma)} G_\sigma(x-y)u(y)\,dy.$$

If $u \in \mathrm{BV}(\mathbb{R}^n)$, then we denote

$$|Du|_\sigma(x) := \int_{B(x,\sigma)} G_\sigma(x-y)\,d|Du|(y)$$

and we write, for simplicity, $|Du|_\sigma^2 := (|Du|_\sigma)^2$.

Let $g : [0,\infty) \to [0,\infty)$ and $\ell : (0,\infty) \to (0,\infty)$ be functions. We define

$$\mathrm{F}_\sigma(u,U) := \int_U \frac{g\big(\ell(\sigma)|Du|_\sigma^2\big)}{\ell(\sigma)}\,dx, \quad u \in \mathrm{BV}(\Omega),$$

for $U \subset\subset \Omega$ and $\sigma > 0$ sufficiently small that the convolution be defined. Further we set

$$\mathrm{F}_\sigma(u) := \mathrm{F}_\sigma(u,\Omega_\sigma) \quad \text{where} \quad \Omega_\sigma := \{x \in \Omega : d(x,\partial\Omega) > \sigma\}.$$

We extend F_σ to GBV as a limit,

$$\mathrm{F}_\sigma(u) := \limsup_{\lambda\to\infty} \mathrm{F}_\sigma(u_\lambda), \quad u \in \mathrm{GBV}(\mathbb{R}^n),$$

and to $L^1(\mathbb{R}^n)$ by ∞. For definitions of GBV and other terms, we refer to Section 2.

Assumptions 1.1. We need the following assumptions on g and ℓ:

(G) g in increasing and concave, $g(0) = 0$ and $\lim_{x\to 0^+} \frac{g(x)}{x} = 1$;

(L) ℓ is increasing, $\lim_{x\to 0^+} \ell(x) = 0$ and $\lim_{x\to 0^+} \frac{\ell(x)}{x^2} = \infty$; and

(GL) $\lim_{x\to 0^+} \dfrac{g(c\,\ell(x)/x^2)}{\ell(x)/x} = 1$ and $\lim_{x\to 0^+} \dfrac{g(c\,\ell(x)/x^{2n})}{\ell(x)/x} \leqslant c_n$ for every fixed $c > 0$.

Note that (G) yields that $g(x) \leqslant x$, as the following deduction shows:

$$g(x) = \lim_{m\to\infty} g\left(m\frac{x}{m}\right) \leqslant \lim_{m\to\infty} mg\left(\frac{x}{m}\right) = \lim_{m\to\infty} x\frac{g(x/m)}{x/m} = x.$$

Since $g(0) = 0$, concavity implies that $g(ax) \leqslant ag(x)$ for every $a \geqslant 1$.

The Geman–McClure and Hebert–Leahy potentials satisfy the above conditions with a suitable scaling function ℓ. Thus our main examples are the functions:

$$g(x) := \frac{x}{1+x} \quad \text{with} \quad \ell(x) := x \tag{3}$$

and

$$g(x) := \log(1+x) \quad \text{with} \quad \ell(x) := x\log(1+\tfrac{1}{x}). \tag{4}$$

These correspond the the Geman–McClure model [18] and the Hebert–Leahy model [20], respectively. For simplicity, we denote

$$\hat{g}(\sigma, t) := \frac{g(\ell(\sigma)\,t^2)}{\ell(\sigma)}.$$

Note that $\hat{g}$ is increasing in t and that $\hat{g}(\sigma, t) \leqslant t^2$.

It turns out that the limit of this functional is intimately connected to the weak formulation of the Mumford–Shah (MS) functional. Let $\Omega \subset \mathbb{R}^n$ be a bounded open set and $f \in L^\infty(\Omega)$. The MS functional, in weak formulation and without fidelity term, is

$$\mathrm{MS}(u) := \int_\Omega |\nabla_a u|^2\,dx + 2\mathcal{H}^{n-1}(S_u),$$

for $u \in \mathrm{GSBV}(\Omega)$ and $\mathrm{MS}(u) := \infty$ for $u \in L^1(\Omega) \setminus \mathrm{GSBV}(\Omega)$. This functional was introduced by De Giorgi and Ambrosio [16] to solve the existence problem for the strong (C^1) functional proposed by Mumford and Shah in 1989 [22]. The lack of lower semicontinuity of $\mathcal{H}^{n-1}(S_u)$ means that it is difficult to deal with this functional and hence in many papers it has been approximated, in the sense of Γ-convergence, by more regular functionals, see for example [4, 14]. For overviews of results on the MS functional we refer to [3, 15]. The Geman–McClure model is rather easy to deal with numerically, whereas the opposite is true for the Mumford–Shah model. Therefore there is a practical incentive to establish connections between the models.

Following Braides, Chambolle and Dal Maso [7, 10], we study the Γ-convergence of our functional. We will need the following functional by Braides and Dal Maso [7]:

$$\mathrm{F}_\sigma^{BD}(u) := \int_\Omega \tfrac{1}{\sigma}\, g\left(\sigma \fint_{B(x,\sigma)\cap\Omega} |\nabla u(y)|^2\,dy\right)dx, \quad u \in H^1(\Omega);$$

we extend F_σ^{BD} to L^1 by ∞. Braides and Dal Maso have shown that F_σ^{BD} Γ-converges to MS with respect to the L^1-topology. In Section 5 we will use this to derive the Γ-converges of our functional in the case $g(t) = \frac{t}{1+t}$. Γ-convergence in the case $g(t) = \log(1+t)$ has been considered by Tiirola in a separate paper [25].

Theorem 1.2. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with Lipschitz boundary. If $g(t) = \frac{t}{1+t}$ then F_σ Γ-converges to* MS *with respect to the L^1 topology.*

The previous result holds also for the previously known approximating functionals. However, the next result indicates that the new approximation is more precise, in that we also obtain point-wise convergence. In this sense this result is the main innovation and motivation of this paper. It is proved in Sections 3 and 4.

Theorem 1.3. *Let Ω be a bounded open set. Suppose that $u \in \mathrm{SBV}^2(\Omega)$ and S_u is contained $\mathcal{H}^{n-1}$-a.e. in the union of a finite number of hyperplanes. Then*

$$\mathrm{MS}(u) = \lim_{\sigma\to 0^+} F_\sigma(u)$$

Note that each $u \in \mathrm{GSBV}^2(\Omega)$ can be estimated by functions u_i satisfying the conditions in the previous theorem such that $\mathrm{MS}(u_i) \to \mathrm{MS}(u)$. This result is by G. Cortesani [12, Corollary 3.11].

The functional $\lim_{\sigma\to 0} \mathrm{F}_\sigma^{BD}(u)$ is finite only for H^1-Sobolev functions and thus the previous theorem is not true for F_σ^{BD}. Therefore F_σ is point-wise better behaved than F_σ^{BD}. The same is true for the functional of Chambolle and Dal Maso [10].

Remark 1.4. When sampling a continuous function by averaging in these type of models, there are three operations to be performed: convolution, derivative, and raising to the power two. By varying the point at which the convolution operator is applied, we arrive at three possible orders: derivative-power-convolution, derivative-convolution-power, and convolution-derivative-power. These give rise to the functionals

$$\int \tfrac{1}{\sigma}\, g\Big(\sigma(|\nabla u|^2)_\sigma\Big)\,dx, \quad \int \tfrac{1}{\sigma}\, g\Big(\sigma(|\nabla u|_\sigma)^2\Big)\,dx, \quad \text{and} \quad \int \tfrac{1}{\sigma}\, g\Big(\sigma|\nabla(u_\sigma)|^2\Big)\,dx.$$

The argument of the first is finite in H^1, the second in BV, and the last in L^1. Braides and Dal Maso opt for the first (largest), we consider the second one. Most closely related to the discrete model would be the third. Unfortunately, that model is not well behaved analytically.

The model of Chambolle and Dal Maso [10] is of this type, but instead of the convolution u_σ they consider regularizations which are piece-wise affine on a triangular grid of controlled geometry. The extra structure allows them to handle the complications arising.

In [6], Bourdin and Chambolle study the numerical implementation of the model from [10]. In our case this remains for future work. We point out that the larger function space on which our functional is finite allows for many different approximation schemes. A natural starting point is with functions satisfying the condition of Theorem 1.3. On the theoretical side, it would be interesting to study whether the condition in the theorem can be removed.

Remark 1.5. Various indirect approaches have been proposed to deal with the Mumford–Shah functional, see [5, Section 4.2.4]. In the terminology of Aubert and Kornprobst, our approach falls under the heading “Approximation by introducing non-local terms”.

2. **Notation.** By c we denote a generic constant whose value can change between each appearance. We denote the Lebesgue n-measure of (a measurable set) $E \subset \mathbb{R}^n$ by $|E|$. The Hausdorff d-dimensional outer measure is denoted by $\mathcal{H}^d$.

We state some properties of BV-spaces that are needed. For further background and proofs see [3]. Let Ω be an open set of $\mathbb{R}^n$ and $u \in L^1(\Omega)$, set

$$\|Du\|(\Omega) := \sup\Big\{ \int_\Omega u \operatorname{div}\varphi\, dx : \varphi \in C_0^\infty(\Omega, \mathbb{R}^n), |\varphi| \leqslant 1 \Big\}$$

and define the space of functions of *bounded variation* as

$$\mathrm{BV}(\Omega) := \Big\{ u \in L^1(\Omega) : \|Du\|(\Omega) < \infty \Big\}.$$

We define S_u as the subset of Ω where the function u does not have approximate limit: $x \in \Omega \setminus S_u$ if and only if there exists $z \in \mathbb{R}$ such that

$$\lim_{r\to 0^+} \fint_{B(x,r)} |u(y) - z|\, dy = 0.$$

For $u \in \mathrm{BV}(\Omega)$, the set S_u is Borel, $|S_u| = 0$ and

$$S_u = N \cup \bigcup_{i\in\mathbb{N}} K_i,$$

where $\mathcal{H}^{n-1}(N) = 0$ and (K_i) is a sequence of compact sets, each one contained in a C^1-hypersurface.

The distributional gradient Du can be divided into an absolutely continuous part $\nabla_a u$ (with respect to Lebesgue measure) and a singular part $D_s u$. We say that $u \in \mathrm{BV}(\Omega)$ is a *special function of bounded variation* and denote $u \in \mathrm{SBV}(\Omega)$ if the singular part of the gradient $D_s u$ is concentrated on S_u, i.e. $\|D_s u\|(\Omega \setminus S_u) = 0$. The space SBV was introduced by De Giorgi and Ambrosio [16]. The spaces GBV and GSBV are defined to consist of functions u for which u_λ belongs to BV and SBV, respectively, for all $\lambda > 0$ where u_λ is the truncation of u at levels $-\lambda$ and λ. We write $u \in \mathrm{SBV}^2(\Omega)$ if $u \in \mathrm{SBV}(\Omega) \cap L^2(\Omega)$, $|\nabla_a u| \in L^2(\Omega)$ and $\mathcal{H}^{n-1}(S_u) < \infty$.

The concepts of Γ-convergence, introduced by De Giorgi, has been systematically studied in [13]. We present here only the definition. A family of functionals $\mathrm{F}_\sigma : X \to \overline{\mathbb{R}}$ is said to Γ-*converge* (in topology τ) to $F : X \to \overline{\mathbb{R}}$ if the following hold for every positive sequence (σ_i) converging to zero:

(a) for every $u \in X$ and every sequence $(u_i) \subset X$ τ-converging to u, we have

$$F(u) \leqslant \liminf_{\sigma_i \to 0^+} \mathrm{F}_{\sigma_i}(u_i);$$

(b) for every $u \in X$ there exists $(u_i) \subset X$ τ-converging to u such that

$$F(u) \geqslant \limsup_{\sigma_i \to 0^+} \mathrm{F}_{\sigma_i}(u_i).$$

The sequence in (b) is called the *recovery sequence*.

We conclude the introduction by showing that the Geman–McClure functions and the Hebert–Leahy functions satisfy the required conditions in the introduction.

Proposition 2.1. *The Geman–McClure functions* (3) *satisfy conditions (G), (L) and (GL).*

Proof. Condition (L) is clear. Since the second derivative of g is negative for every point in $[0, \infty)$, the function is concave. Now (G) is clear. We check the (GL):

$$\lim_{x\to 0^+} \frac{g(c\,\ell(x)/x^k)}{\ell(x)/x} = \lim_{x\to 0^+} \frac{g(c/x^{k-1})}{1} = 1$$

for every fixed $c > 0$ and $k > 1$. □

Proposition 2.2. *The Hebert–Leahy functions* (4) *satisfy conditions (G), (L) and (GL).*

Proof. Condition (L) is clear. Since the second derivative of g is negative for every point in $[0,\infty)$, the function is concave and (G) follows. For (GL) we note, with $k=2$ or $k=2n$, that

$$\frac{g(c\,\ell(x)/x^k)}{\ell(x)/x} = \frac{\log(1+c\,x^{1-k}\log(1+\frac{1}{x}))}{\log(1+\frac{1}{x})}$$
$$\sim \frac{\log(c\,x^{1-k}\log\frac{1}{x})}{\log\frac{1}{x}} = \frac{\log c+(k-1)\log\frac{1}{x}+\log\log\frac{1}{x}}{\log\frac{1}{x}} \to k-1$$

as $x\to 0$. The symbol $\sim$ means asymptotically equivalent. □

3. **Lower bound by the Hausdorff measure.** In this section we start with an estimate for the Hausdorff measure of the singular set S_u. This is done in several steps with increasingly more general sets S_u. We denote by $\mathbb{R}^{n-1}$ also the subset $\mathbb{R}^{n-1}\times\{0\}$ of $\mathbb{R}^n$.

Lemma 3.1. *If* $u\in \mathrm{BV}(\mathbb{R}^n)$, *then*

$$2\mathcal{H}^{n-1}\big(S_u\cap\mathbb{R}^{n-1}\big) \leqslant \liminf_{\sigma\to 0^+} \mathrm{F}_\sigma(u,\Omega^\sigma),$$

where $\Omega^\sigma := \mathbb{R}^{n-1}\times[-\sigma,\sigma]$.

Proof. By [3, Lemma 3.76] there exists a non-negative function $v\in L^1(\mathbb{R}^{n-1})$ representing the measure $|D_s u|$ restricted to $\mathbb{R}^{n-1}$ such that $|Du|\geqslant v\,d\mathcal{H}^{n-1}$. Then, by monotonicity,

$$\int_{\mathbb{R}^n} \hat g(\sigma, v_\sigma)\,dx \leqslant \int_{\mathbb{R}^n} \hat g(\sigma, |Du|_\sigma)\,dx,$$

where $v_\sigma := G_\sigma * (v\,\mathcal{H}^{n-1})$. Thus it suffices to show that

$$2\mathcal{H}^{n-1}\big(S_u\cap\mathbb{R}^{n-1}\big) \leqslant \liminf_{\sigma\to 0^+} \int_{\mathbb{R}^n} \hat g(\sigma, v_\sigma)\,dx. \tag{5}$$

For $x\in\mathbb{R}^{n-1}$ we denote by x_λ the point $(x,\lambda\sigma)\in\mathbb{R}^n$. Then, for $x\in\mathbb{R}^{n-1}$ and $\lambda\in(-1,1)$,

$$v_\sigma(x_\lambda) = \int_{\mathbb{R}^{n-1}\cap B(x,\sigma)} G_\sigma(x_\lambda - y)v(y)\,d\mathcal{H}^{n-1}(y).$$

For $\lambda\in(-1,1)$ and $y\in\mathbb{R}^{n-1}$, $G_\sigma(x_\lambda-y)=\frac{1}{\sigma}(G^\lambda)_\sigma(x-y)$, where $G^\lambda(x):=G(x,\lambda)$ is an $(n-1)$-dimensional section of G with $\mathbb{R}^{n-1}$, and with scaling $1/\sigma^{n-1}$. Now $\int G^\lambda\,d\mathcal{H}^{n-1}=c_\lambda>0$ and so

$$\sigma\, v_\sigma(x_\lambda)\to c_\lambda\, v(x) \tag{6}$$

as $\sigma\to 0^+$ at all $\mathcal{H}^{n-1}$-Lebesgue points of v.

Fix $r\in(0,1)$ and $\epsilon>0$ we define

$$E^\epsilon_\sigma := \big\{x\in S_u\cap\mathbb{R}^{n-1}\,\big|\,\forall\lambda\in(-r,r), \forall\sigma'\in(0,\sigma],\ \sigma'\,v_{\sigma'}(x,\lambda\sigma')\geqslant\epsilon\big\}.$$

Note that E_σ^ϵ is the intersection of the closed sets $\{\sigma' v_{\sigma'}(x, \lambda\sigma') \geqslant \epsilon\}$, so it is itself closed, in particular measurable. With this set, we find that

$$\begin{aligned}\int_{\mathbb{R}^n} \hat{g}(\sigma, v_\sigma)\, dx &\geqslant \int_{E_\sigma^\epsilon\times[-r\sigma,r\sigma]} \hat{g}(\sigma, v_\sigma)\, dx \\ &\geqslant \int_{E_\sigma^\epsilon\times[-r\sigma,r\sigma]} \frac{g(\ell(\sigma)/\sigma^2)}{\ell(\sigma)}\, dx \\ &= \frac{g(\ell(\sigma)/\sigma^2)}{\ell(\sigma)} \big|E_\sigma^\epsilon \times [-r\sigma, r\sigma]\big| = 2r\frac{g(\ell(\sigma)/\sigma^2)}{\ell(\sigma)/\sigma}\mathcal{H}^{n-1}(E_\sigma^\epsilon).\end{aligned}$$

Notice that the fraction in front of the Hausdorff measure tends to 1 as $\sigma \to 0^+$, by assumption (GL).

For $\lambda \in (-r, r)$, we have $c_\lambda > c > 0$. By (6), $\sigma\, v_\sigma(x, \lambda\sigma) \to c_\lambda v(x) > c\, v(x)$. Furthermore, $v(x) > 0$ in $S_u \cap \mathbb{R}^{n-1}$. Thus $E_\sigma^\epsilon \nearrow S_u \cap \mathbb{R}^{n-1}$ as $\sigma \to 0$ and $\epsilon \to 0$, up to a set of $\mathcal{H}^{n-1}$-measure zero. Hence

$$\begin{aligned}2\mathcal{H}^{n-1}(S_u \cap \mathbb{R}^{n-1}) = \lim_{\sigma\to 0^+} 2\mathcal{H}^{n-1}(E_\sigma^\epsilon) &= \lim_{\sigma\to 0^+} 2\frac{g(\ell(\sigma)/\sigma^2)}{\ell(\sigma)/\sigma}\mathcal{H}^{n-1}(E_\sigma^\epsilon) \\ &\leqslant \frac{1}{r}\liminf_{\sigma\to 0^+} \int_{\Omega^\sigma} \hat{g}(\sigma, v_\sigma)\, dx.\end{aligned}$$

Now (5) follows from this as $r \to 1$, which completes the proof. □

We define the push-forward $\Phi_\# Du$ of a measure as in [3, Definition 1.70].

Lemma 3.2. *Let $u \in \mathrm{BV}(\mathbb{R}^n)$ and let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ be an L-bilipschitz map, $L \geqslant 1$. Then, for open $\Omega \subset \mathbb{R}^n$,*

$$\mathrm{F}_\sigma(u \circ \Phi, \Omega) \leqslant c_L\, \mathrm{F}_\sigma(u, \Phi(\Omega)),$$

where $c_L \to 1^+$ as $L \to 1^+$.

Proof. Since Φ is bilipschitz, it is Lipschitz and proper. Hence by [3, Theorem 3.16], $|D(u \circ \Phi)| \leqslant L^{n-1}\, (\Phi^{-1})_\# |Du|$. Then, by [3, p. 32],

$$\begin{aligned}|D(u \circ \Phi)|_\sigma(x) &= \int_{\mathbb{R}^n} G_\sigma(x - y)\, d|D(u \circ \Phi)|(y) \\ &\leqslant L^{n-1} \int_{\mathbb{R}^n} G_\sigma(x - y)\, d(\Phi^{-1})_\# |Du|(y) \\ &= L^{n-1} \int_{\mathbb{R}^n} G_\sigma(x - \Phi^{-1}(z))\, d|Du|(z).\end{aligned}$$

Since Φ is bilipschitz and G_σ is radially decreasing and symmetric we further have

$$G_\sigma(x - \Phi^{-1}(z)) \leqslant G_\sigma\left(\frac{\Phi(x) - z}{L}\right) \leqslant c_L G_\sigma(\Phi(x) - z),$$

where $c_L := \sup_{r>0} G(r/L)/G(r) \geqslant 1$. Since G is continuous, $c_L \to 1^+$ as $L \to 1^+$. With this we may continue our previous estimate:

$$|D(u \circ \Phi)|_\sigma(x) \leqslant L^{n-1} c_L \int_{\mathbb{R}^n} G_\sigma(\Phi(x) - z)|Du|(z) = L^{n-1} c_L\, |Du|_\sigma(\Phi(x)).$$

Let us denote the constant in the inequality by d_L. Then we conclude that

$$\hat{g}\big(\sigma, |D(u \circ \Phi)|_\sigma(x)\big) \leqslant \hat{g}\big(\sigma, d_L\, |Du|_\sigma(\Phi(x))\big) \leqslant d_L^2 \hat{g}\big(\sigma, |Du|_\sigma(\Phi(x))\big),$$

where we used that $\hat{g}$ is increasing and the inequality $g(ax) \leqslant ag(x)$. The proof is concluded by a change of variables:

$$\int_\Omega \hat{g}(\sigma, |D(u\circ\Phi)|_\sigma(x))\,dx \leqslant d_L^2 \int_\Omega \hat{g}(\sigma, |Du|_\sigma(\Phi(x)))\,dx \leqslant L^n d_L^2 \int_{\Phi(\Omega)} \hat{g}(\sigma, |Du|_\sigma)\,dx. \qquad \square$$

Corollary 3.3. *If $u \in \mathrm{SBV}(\mathbb{R}^n)$ and $S \subset \mathbb{R}^n$ is a C^1-hypersurface, then*

$$2\mathcal{H}^{n-1}(S_u \cap S) \leqslant \liminf_{\sigma\to 0^+} \mathrm{F}_\sigma(u, \Sigma^\sigma),$$

where $\Sigma^\sigma := \{x \in \mathbb{R}^n \mid \mathrm{dist}(x,S) < 2\sigma\}$.

Proof. Every point $x_0 \in S$ has a neighborhood U such that there exists a C^1-isomorphism $\Phi : U \to \mathbb{R}^n$ under which $S\cap U$ maps into a hyperplane and $D\Phi(x_0) = I$. Since the mapping is C^1, we conclude that it is bilipschitz with constant L; the constant can be chosen arbitrarily close to 1 by making the neighbourhood U small. Hence $\mathcal{H}^{n-1}(S_u \cap U) \leqslant L^{n-1}\mathcal{H}^{n-1}(\Phi(S_u \cap U))$. By Lemma 3.2,

$$c_L \liminf_\sigma \mathrm{F}_\sigma(u, \Sigma^\sigma \cap U) \geqslant \liminf_\sigma \mathrm{F}_\sigma(u\circ\Phi^{-1}, \Phi(\Sigma^\sigma)\cap\Phi(U)).$$

Then it follows from Lemma 3.1 that

$$L^{n-1} c_L \liminf_\sigma \mathrm{F}_\sigma(u, \Sigma^\sigma \cap U) \geqslant 2\mathcal{H}^{n-1}(S_u \cap U).$$

Let $\epsilon > 0$. We can find a finite family of separated open sets U_k such that the above claim holds in each and

$$\sum_k \mathcal{H}^{n-1}(S_u \cap U_k) = \mathcal{H}^{n-1}\Big(S_u \cap \bigcup_k U_k\Big) \geqslant \mathcal{H}^{n-1}(S_u) - \epsilon.$$

Then

$$L^{n-1} c_L \liminf_\sigma \mathrm{F}_\sigma(u, \Sigma^\sigma) \geqslant \sum_k \liminf_\sigma \mathrm{F}_\sigma(u, \Sigma^\sigma \cap U_k) \geqslant 2\mathcal{H}^{n-1}(S_u) - 2\epsilon,$$

from which the claim follows as $\epsilon \to 0^+$ and $L \to 1^+$. $\square$

Next we leave the singular part and estimate the absolutely continuous part.

Lemma 3.4. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set. If $u \in \mathrm{BV}(\Omega)$, then*

$$\int_\Omega |\nabla_a u|^2\,dx \leqslant \liminf_{\sigma\to 0^+} \mathrm{F}_\sigma(u).$$

Proof. We assume the right hand side is finite since otherwise there is nothing to prove. Choose a decreasing sequence (σ_k) such that $\ell(\sigma_k) \leqslant 2^{-k}$ and $\lim_k \mathrm{F}_{\sigma_k}(u) < c < \infty$. Fix $\epsilon > 0$ and set

$$H_k := \big\{x \in \Omega \,\big|\, \ell(\sigma_k)|Du|^2_{\sigma_k} > \epsilon\big\} \quad \text{and} \quad G_j := \bigcup_{k\geqslant j} H_k.$$

Then we calculate

$$c > \mathrm{F}_{\sigma_k}(u) \geqslant \int_{H_k} \frac{g(\ell(\sigma_k)|Du|^2_{\sigma_k})}{\ell(\sigma_k)}\,dx \geqslant \int_{H_k} \frac{g(\epsilon)}{\ell(\sigma_k)}\,dx = \frac{g(\epsilon)}{\ell(\sigma_k)}\,|H_k|. \tag{7}$$

It follows that $|H_k| \leqslant c(\varepsilon)\,\ell(\sigma_k) \leqslant c(\varepsilon)2^{-k}$ and $|G_j| \leqslant c(\varepsilon)\sum_{k\geqslant j} 2^{-k} \leqslant c(\varepsilon)\,2^{-j}$.

Define $\Omega_j := \Omega \setminus G_j$. In Ω_j, $\ell(\sigma_k)|Du|^2_{\sigma_k} \leqslant \epsilon$ for every $k \geqslant j$. Since $g(t)/t \to 1$ as $t \to 0^+$ (by assumption (G)), we conclude that

$$g\big(\ell(\sigma_k)|Du|^2_{\sigma_k}\big) \geqslant c'(\epsilon)\ell(\sigma_k)|Du|^2_{\sigma_k},$$

where $c'(\epsilon) \to 1$ as $\epsilon \to 0^+$. Hence

$$\int_\Omega \frac{g(\ell(\sigma_k)|Du|^2_{\sigma_k})}{\ell(\sigma_k)}\,dx \geqslant c'(\epsilon)\int_{\Omega_j} |Du|^2_{\sigma_k}\,dx.$$

Since $\nabla_a u(x)\,dx$ and D_s are mutually singular measures [3, Proposition 3.92, p. 184], we obtain in every Borel set E that $\|Du\|(E)$ equals

$$\|\nabla_a u\,dx + D_s u\|(E) = \|\nabla_a u\,dx\|(E) + \|D_s u\|(E) = \int_E |\nabla_a u|\,dx + \|D_s u\|(E).$$

Thus

$$\begin{aligned}|Du|_\sigma(x) &= \int G_\sigma(x-y)\,d|Du|(y)\\ &= \int G_\sigma(x-y)|\nabla_a u(y)|\,dy + \int G_\sigma(x-y)\,d|D_s u|(y)\\ &= |\nabla_a u|_\sigma(x) + |D_s u|_\sigma(x).\end{aligned}$$

Since $\nabla_a u \in L^1(\Omega)$ we find that

$$\liminf_{\sigma\to 0} |Du|_\sigma = \liminf_{\sigma\to 0}\big[|\nabla_a u|_\sigma + |D_s u|_\sigma\big] \geqslant \liminf_{\sigma\to 0} |\nabla_a u|_\sigma = |\nabla_a u|$$

almost everywhere. Therefore the previous estimate implies that

$$\int_{\Omega_j} |\nabla_a u|^2\,dx \leqslant \liminf_{k\to\infty} \int_{\Omega_j} |Du|^2_{\sigma_k}\,dx \leqslant \frac{1}{c(\epsilon)} \liminf_{\sigma} \mathrm{F}_\sigma(u).$$

Hence $\nabla_a u \in L^2(\Omega_j)$ with norm independent of j. Then letting $j \to \infty$ yields that $\nabla_a u \in L^2(\cup_j \Omega_j)$. Since $|\Omega \setminus \cup_j \Omega_j| = 0$ and $\nabla_a u$ is absolutely continuous with respect to the Lebesgue measure, we conclude that $\nabla_a u \in L^2(\Omega)$. The estimate for the norm follows as $\epsilon \to 0$. □

Theorem 3.5. *Let $\Omega \subset \mathbb{R}^n$ be an open bounded set. If $u \in \mathrm{SBV}(\Omega)$, then*

$$\int_\Omega |\nabla_a u|^2\,dx + 2\mathcal{H}^{n-1}(S_u) \leqslant \liminf_{\sigma\to 0^+} \mathrm{F}_\sigma(u).$$

Proof. If we show that

$$\int_U |\nabla_a u|^2\,dx + 2\mathcal{H}^{n-1}(S_u \cap U) \leqslant \liminf_{\sigma\to 0^+} \mathrm{F}_\sigma(u, U)$$

for every $U \subset\subset \Omega$, then the claim follows by taking the supremum over U. Let $U \subset\subset V \subset\subset \Omega$; we define

$$u_{\text{ext}}(x) := \max\left\{0, 1 - \frac{\operatorname{dist}(x, V)}{\operatorname{dist}(\partial V, \partial\Omega)}\right\} u(x)$$

in Ω and $u_{\text{ext}} = 0$ in $\mathbb{R}^n \setminus \Omega$. Now $u_{\text{ext}} \in L^\infty(\mathbb{R}^n) \cap \mathrm{SBV}(\mathbb{R}^n)$ and $u|_V = u_{\text{ext}}|_V$. Hence for all sufficiently small σ, $\mathrm{F}_\sigma(u, U) = \mathrm{F}_\sigma(u_{\text{ext}}, U)$. Therefore it suffices to prove the claim of the theorem under the assumption $u \in \mathrm{SBV}(\mathbb{R}^n)$.

Let $K_i \subset U$ be a countable collection of compact sets, each contained in a C^1-hypersurface, such that $S_u \cap U = N \cup \bigcup K_i$ where $\mathcal{H}^{n-1}(N) = 0$. It suffices to show that

$$\int_U |\nabla_a u|^2\, dx + 2\mathcal{H}^{n-1}\bigg(\bigcup_{i=1}^m K_i\bigg) \leqslant \liminf_\sigma \mathrm{F}_\sigma(u),$$

since the claim then follows as we take the limit $m \to \infty$.

Let $K := \bigcup_{i=1}^m K_i$. Let V_1 be an open neighborhood of $K_1' := K_1$ such that $\mathcal{H}^{n-1}(V_1 \cap K) \leqslant \mathcal{H}^{n-1}(K_1) + \varepsilon/m$. Let V_2 be an open neighborhood of $K_2' := K_2 \setminus V_1$ such that $\mathcal{H}^{n-1}(V_2 \cap K) \leqslant \mathcal{H}^{n-1}(K_2 \setminus V_1) + \varepsilon/m$. We continue this process until we get to V_m, which is an open neighborhood of $K_m' := K_m \setminus (V_1 \cup \ldots \cup V_{m-1})$ such that $\mathcal{H}^{n-1}(V_m \cap K) \leqslant \mathcal{H}^{n-1}(K_m \setminus (V_1 \cup \ldots \cup V_{m-1})) + \varepsilon/m$. Now

$$\mathcal{H}^{n-1}\bigg(\bigcup_{i=1}^m K_i\bigg) \leqslant \mathcal{H}^{n-1}\bigg(\bigcup_{i=1}^m K_i'\bigg) + \epsilon,$$

and $\operatorname{dist}(K_i', K_j') > 0$ for distinct i and j. Set $\sigma_0 := \min_{i\neq j} \operatorname{dist}(K_i', K_j')$.

Let N_s^k be the s-neighborhood of K_k', $s > 0$. For $\sigma < \frac{1}{6}\sigma_0$, the sets $N_{3\sigma}^j$ and $N_{3\sigma}^k$ are disjoint ($j \neq k$), and when $\sigma < \frac{1}{3}\operatorname{dist}(\partial V, \partial U)$ they are contained in V. Thus Corollary 3.3 can be applied to the hypersurfaces individually, and it gives that

$$2\mathcal{H}^{n-1}(K_k') \leqslant c_\sigma\, \mathrm{F}_\sigma(u, N_{2\sigma}^k),$$

where $c_\sigma \to 1$ as $\sigma \to 0$. By Lemma 3.4 applied in the set $U \setminus \bigcup \overline{N_{4\sigma}^k}$, we obtain

$$\int_U |\nabla_a u|^2\, dx \leqslant c_\sigma\, \mathrm{F}_\sigma(u, U \setminus \cup N_{4\sigma}^k) + \int_{\cup N_{4\sigma}^k} |\nabla_a u|^2\, dx,$$

with c_σ converging to 1 as before. In view of this and the absolute continuity of the last integral,

$$\begin{aligned}\int_U |\nabla_a u|^2\, dx &+ 2\mathcal{H}^{n-1}\bigg(\bigcup_{k=1}^m K_k'\bigg)\\ &\leqslant \liminf_\sigma c_\sigma\bigg[\mathrm{F}_\sigma(u, U \setminus \cup N_{4\sigma}^k) + \sum_{k=1}^m \mathrm{F}_\sigma(u, N_{2\sigma}^k)\bigg].\end{aligned}$$

To complete the proof we observe that the sets $U \setminus \cup N_{4\sigma}^k$ and $N_{2\sigma}^k$ are disjoint, even when dilated by σ, so that

$$\mathrm{F}_\sigma(u, U \setminus \cup N_{4\sigma}^k) + \sum_k \mathrm{F}_\sigma(u, N_{2\sigma}^k) \leqslant \mathrm{F}_\sigma(u, V)$$

for $\sigma < \min\{\frac{1}{6}\sigma_0, \operatorname{dist}(\partial V, \partial U)\}$. Together with the previous estimate, this completes the proof. □

Next we generalize Theorem 3.5 for GSBV-function. We need the following lemma, which follows from the co-area formula [3, Theorem 3.40].

Lemma 3.6. *If $u \in \mathrm{BV}(\Omega)$ and $E \subset \Omega$ is Borel, then $|D(u_\lambda)|(E) \leqslant |Du|(E)$.*

Corollary 3.7. *If $u \in \mathrm{BV}(\Omega)$, then $G_\sigma * |Du_\lambda| \leqslant G_\sigma * |Du|$.*

Proof. Since G_σ is radially symmetric and decreasing, we have

$$G_\sigma * |Du|(x) = \int_0^\sigma h(r,\sigma)|Du|(B(x,r))\, dr,$$

where $h \geqslant 0$ is a suitably chosen weight function. This representation also holds for u_λ. Since $|Du_\lambda|(B(x,r)) \leqslant |Du|(B(x,r))$ by Lemma 3.6, the claim follows. □

We can now summarize the result for the lower bound.

Theorem 3.8. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set. If $u \in \mathrm{GSBV}(\Omega)$, then*

$$\mathrm{MS}(u) \leqslant \liminf_{\sigma\to 0^+} \mathrm{F}_\sigma(u).$$

Proof. By monotonicity of our functional (assumption (G)), it follows from Corollary 3.7 that $\mathrm{F}_\sigma(u_\lambda) \leqslant \mathrm{F}_\sigma(u)$ for every $\lambda > 0$. Theorem 3.5 yields that

$$\int_\Omega |\nabla_a u_\lambda|^2\,dx + 2\mathcal{H}^{n-1}(S_{u_\lambda}) \leqslant \liminf_{\sigma\to 0^+} \mathrm{F}_\sigma(u_\lambda) \leqslant \liminf_{\sigma\to 0^+} \mathrm{F}_\sigma(u).$$

When $\lambda \nearrow \infty$, $|\nabla_a u_\lambda| \nearrow |\nabla_a u|$ and $S_{u_\lambda} \nearrow S_u$. Hence, by monotone convergence,

$$\int_\Omega |\nabla_a u|^2\,dx + 2\mathcal{H}^{n-1}(S_u) \leqslant \liminf_{\sigma\to 0^+} \mathrm{F}_\sigma(u).$$

□

4. **Upper bound.** In this section we prove the upper bound of $\limsup F_\sigma$ by the Mumford–Shah functional. We start with a lemma. Recall that the Besov space $B^s_{p,p}(\Omega)$, $s \in (0,1)$, is defined as L^p-functions with

$$\int_\Omega\int_\Omega \frac{|u(x)-u(y)|^p}{|x-y|^{n+sp}}\,dx\,dy < \infty.$$

When $p=2$, $B^s_{2,2} = H^s$, the fractional Sobolev space. We denote by f_Q the average $\fint_Q f\,dx$ of f over the set Q and by $Q(z,\sigma)$ a cube with center z and side-length σ.

Lemma 4.1. *Let $\Omega \subset \mathbb{R}^n$ be an open set, $1 \leqslant p < \infty$ and $s \geqslant 0$. If $v \in B^s_{p,p}(\Omega)$, then*

$$\lim_{\sigma\to 0^+} \frac{1}{\sigma^s} v_{Q(x,\sigma)} = 0$$

for almost every $x \in \{v = 0\}$.

Proof. Suppose that $v \in B^s_{p,p}(\Omega)$ and set $V_\sigma(z) := \int_{Q(z,\sigma)} \frac{|v(x)-v(z)|^p}{|x-z|^{n+sp}}\,dx$ Then, by definition of the Besov space,

$$\int_\Omega V_\sigma(z)\,dz \leqslant \int_\Omega\int_\Omega \frac{|v(x)-v(z)|^p}{|x-z|^{n+sp}}\,dx\,dz < \infty.$$

Since the integrand is non-negative, we conclude that $V_\sigma \in L^1(\Omega)$.

Consider the sets $A := \{z \in \Omega \,|\, V_\sigma(z) < \infty \text{ for some } \sigma\}$ and $B := \{z \in \Omega \,|\, V_\sigma(z) = \infty \text{ for all } \sigma\}$. Since $V_\sigma \in L^1(\Omega)$, it follows that $|B| = 0$. Suppose that $z \in A$. Then for some cube

$$\int_{Q(z,\sigma)} \frac{|v(x)-v(z)|^p}{|x-z|^{n+sp}}\,dx < \infty.$$

Thus $x \mapsto \frac{|v(x)-v(z)|^p}{|x-z|^{n+sp}}$ is an integrable function, and so

$$V_\sigma(z) = \int_{Q(z,\sigma)} \frac{|v(x)-v(z)|^p}{|x-z|^{n+sp}}\,dx \to 0$$

when $\sigma \to 0^+$.

If $v(z) = 0$ and $z \in A$, then

$$\frac{1}{\sigma^s}|v_{Q(z,\sigma)}| \leqslant \frac{1}{\sigma^s}\fint_{Q(z,\sigma)} |v(x)|\,dx \leqslant \left(\frac{1}{\sigma^{sp}}\fint_{Q(z,\sigma)} |v(x)-v(z)|^p\,dx\right)^{1/p} \leqslant c\,V_\sigma(z)^{1/p}$$

since $|x-z|<2\sigma$ when $x\in Q(z,\sigma)$. The right hand side tends to zero, giving the claim. □

Lemma 4.2. *Suppose that $u\in \mathrm{SBV}^2(\Omega)$ in a rectangle Ω with $S_u\subset\mathbb{R}^{n-1}$. Then*

$$\limsup_{\sigma\to 0^+}\int_{U_\sigma}\hat g(\sigma,|D_s u|_\sigma)\,dx\leqslant 2\mathcal{H}^{n-1}(S_u),$$

where $U^\sigma:=(\mathbb{R}^{n-1}\times[-\sigma,\sigma])\cap\Omega_\sigma$.

Proof. As before, let $v:\mathbb{R}^{n-1}\to\mathbb{R}$ be the representative of the singular part of the derivative of u, $D_s u = v\,d\mathcal{H}^{n-1}$. Then $|v|=|u^+-u^-|$, where $u^\pm$ is the approximately continuous extension of u to $\mathbb{R}^{n-1}$ from the positive and negative half-spaces [3, (4.1)]. Since $S_u\cap\mathbb{R}^n_+=\emptyset$, it follows that $Du=\nabla_a u$ in $\Omega\cap\mathbb{R}^n_+$ and since $u\in\mathrm{SBV}^2(\Omega)$, it follows that $\nabla u\in L^2(\Omega\cap\mathbb{R}^n_+)$. As $\Omega\cap\mathbb{R}^n_+$ has Lipschitz boundary, we obtain by the Poincaré inequality that $u\in L^2(\Omega\cap\mathbb{R}^n_+)$ [21, Corollary on page 37], and thus $u\in H^1(\Omega\cap\mathbb{R}^n_+)$. Therefore $v^\pm$ are boundary-values of a Sobolev function, and hence in the trace space $H^{\frac12}(\mathbb{R}^{n-1})$ [1, Theorem 7.39]; the same holds for their difference v.

We have

$$G_\sigma*|D_s u|\leqslant\|G\|_{L^\infty}\frac{1}{\sigma^n}\chi_Q(\tfrac{\cdot}{\sigma})*|D_s u|$$

where $Q=[-1,1]^n$. Denote $x=(x',x_n)$ with $x'\in\mathbb{R}^{n-1}$ and $x_n\in\mathbb{R}$. Since $D_s u = v\,d\mathcal{H}^{n-1}$, the expression

$$\tfrac{1}{\sigma^n}\chi_Q(\tfrac{\cdot}{\sigma})*|D_s u|(x',x_n)=:\tfrac{1}{\sigma}V_\sigma$$

is independent of x_n, where V_σ denotes the $(n-1)$-dimensional average of $|v|$ over the cube $Q(x',\sigma)\cap\mathbb{R}^{n-1}$. Thus

$$\begin{aligned}\int_{U^\sigma}\hat g(\sigma,G_\sigma*|D_s u|)\,dx&\leqslant\int_{\Omega_\sigma\cap\mathbb{R}^{n-1}}\int_{-\sigma}^{\sigma}\hat g(\sigma,\tfrac{c}{\sigma}V_\sigma)\,dx_n\,d\mathcal{H}^{n-1}(x')\\&=2\int_{\Omega_\sigma\cap\mathbb{R}^{n-1}}\sigma\,\hat g(\sigma,\tfrac{c}{\sigma}V_\sigma)\,d\mathcal{H}^{n-1}\end{aligned}$$

By assumptions (G) and (L) we have the following two estimates:

$$\sigma\,\hat g(\sigma,\tfrac{c}{\sigma}V_\sigma)=\frac{g(c^2V_\sigma^2\ell(\sigma)/\sigma^2)}{\ell(\sigma)/\sigma}\leqslant\begin{cases}\frac{c^2}{\sigma}g(V_\sigma^2),\\ \max\left\{1,c^2V_\sigma^2\right\}\frac{g(\ell(\sigma)/\sigma^2)}{\ell(\sigma)/\sigma}.\end{cases}$$

By assumption (GL) $\frac{g(\ell(\sigma)/\sigma^2)}{\ell(\sigma)/\sigma}$ tends to 1 as $\sigma\to0^+$. Hence the latter upper bound can be estimated above by the function $c\,(Mv)^2+2$ which belongs to $L^1(\Omega\cap\mathbb{R}^{n-1})$ since $v\in L^2(\Omega\cap\mathbb{R}^{n-1})$ and the maximal operator is bounded. Hence this upper bound gives for $\sigma\,\hat g(\sigma,\frac{c}{\sigma}V_\sigma)$ a majorant to be used in the dominated convergence theorem.

Let us next look the point wise convergence. Since $|v|\in H^{\frac12}(\mathbb{R}^{n-1})$ we have that $\lim_{\sigma\to0^+}V_\sigma\in[0,\infty)$ for $\mathcal{H}^{n-1}$-almost all points. Assume first that $\lim_{\sigma\to0^+}V_\sigma=0$ and let us estimate the upper bound $\frac{c}{\sigma}g(V_\sigma^2)$. By (G), $\frac{c}{\sigma}g(V_\sigma^2)\leqslant\frac{c^2}{\sigma}V_\sigma^2$ and by Lemma 4.1, in dimension $n-1$ (with $s=\frac12$ and $p=2$), $\frac{c}{\sigma}V_\sigma^2$ tends to zero for $\mathcal{H}^{n-1}$-almost all points in the set $\{v=0\}$. Assume next that $\lim_{\sigma\to0^+}V_\sigma=c_1\in(0,\infty)$. Then there exists σ_0 such that $\frac12c_1\leqslant V_\sigma\leqslant2c_1$ for every $0<\sigma<\sigma_0$. Since $\hat g$ is increasing we obtain

$$\frac{g(\frac{1}{4}c_1^2c^2\ell(\sigma)/\sigma^2)}{\ell(\sigma)/\sigma} = \sigma\,\hat{g}(\sigma, \tfrac{c_1 c}{2\sigma}) \leqslant \sigma\,\hat{g}(\sigma, \tfrac{c}{\sigma}V_\sigma) \leqslant \sigma\,\hat{g}(\sigma, \tfrac{2c_1 c}{\sigma}) = \frac{g(4c_1^2c^2\,\ell(\sigma)/\sigma^2)}{\ell(\sigma)/\sigma}$$

for every $0<\sigma<\sigma_0$, and hence by assumption (GL) the term $\sigma\,\hat{g}(\sigma, \frac{c}{\sigma}V_\sigma)$ tends to 1 as $\sigma\to 0^+$. Hence it follows by dominated convergence that

$$\limsup_{\sigma\to 0^+}\int_{\Omega_\sigma\cap\mathbb{R}^{n-1}} \sigma\,\hat{g}(\sigma, \tfrac{c}{\sigma}V_\sigma)\,d\mathcal{H}^{n-1}(x) \leqslant \mathcal{H}^{n-1}\big(\mathbb{R}^{n-1}\cap\Omega\cap\{v\neq 0\}\big) = \mathcal{H}^{n-1}(S_u).$$

This gives the claim. □

Proof of Theorem 1.3. Let $\Sigma^\sigma := \{x\in\Omega_\sigma \,|\, d(x,S_u)<\sigma\}$. By the triangle inequality $|Du|_\sigma \leqslant |\nabla_a u|_\sigma + |D_s u|_\sigma$, and further $|Du|_\sigma = |\nabla_a u|_\sigma$ in $\Omega_\sigma\setminus\Sigma^\sigma$ since the support of G_σ does not intersect S_u in this case. Moreover,

$$|Du|_\sigma^2 \leqslant (|\nabla_a u|_\sigma + |D_s u|_\sigma)^2 \leqslant c_\epsilon|\nabla_a u|_\sigma^2 + (1+\epsilon)|D_s u|_\sigma^2$$

in Σ^σ. Since g is subadditive,

$$\begin{aligned} F_\sigma(u) &= \int_{\Omega_\sigma} \hat{g}(\sigma, |Du|_\sigma)\,dx = \int_{\Omega_\sigma} \frac{g(\ell(\sigma)|Du|_\sigma^2)}{\ell(\sigma)}\,dx \\ &\leqslant \int_{\Omega_\sigma} \hat{g}(\sigma, |\nabla_a u|_\sigma)\,dx + c_\epsilon\int_{\Sigma_\sigma} \hat{g}(\sigma, |\nabla_a u|_\sigma)\,dx + (1+\epsilon)\int_{\Sigma^\sigma} \hat{g}(\sigma, |D_s u|_\sigma)\,dx. \end{aligned}$$

Since $g(x)\leqslant x$ by assumption (G), we find that

$$\int_{\Omega_\sigma} \hat{g}(\sigma, |\nabla_a u|_\sigma)\,dx \leqslant \int_{\Omega_\sigma} |\nabla_a u|_\sigma^2\,dx \to \int_\Omega |\nabla_a u|^2\,dx$$

as $\sigma\to 0^+$. For the second integral we find that

$$c_\epsilon\int_{\Sigma_\sigma} \hat{g}(\sigma, |\nabla_a u|_\sigma)\,dx \leqslant c_\epsilon\int_{\Sigma_\sigma} |\nabla_a u|_\sigma^2\,dx \leqslant c_\epsilon\int_{\Sigma_{\sigma_0}} |\nabla_a u|_\sigma^2\,dx \to c_\epsilon\int_{\Sigma_{\sigma_0}} |\nabla_a u|^2\,dx,$$

when $\sigma\leqslant\sigma_0$. By absolute continuity the last integral tends to zero when $\sigma_0\to 0$. Then the factor $1+\epsilon$ is gotten rid of when $\epsilon\to 0$. Thus it remains only to bound the integral over Σ^σ by the Hausdorff measure.

By assumption the set S_u lies $\mathcal{H}^{n-1}$-a.e. in the finite union $\cup S_i$ of hyperplanes. Consider the components K of the hyperplane S_1 lying at distance at least δ from the other hyperplanes S_i and from $\partial\Omega$. In each of these we may apply Lemma 4.2 to conclude, for $\sigma<\frac{1}{2}\delta$, that

$$\limsup_{\sigma\to 0} \mathrm{F}_\sigma(u, K^\sigma) \leqslant 2\mathcal{H}^{n-1}(S_1\cap S_u);$$

here we denote by K^σ the points within distance σ of K.

In the complement $C := (S_1\setminus K)\cap\Omega_\sigma$ of the components we estimate

$$\limsup_{\sigma\to 0} \mathrm{F}_\sigma(u, C^\sigma) \leqslant \int_{C_\sigma} \hat{g}(\sigma, \tfrac{c}{\sigma^n})\,dx \leqslant \tfrac{c}{\sigma}|C^\sigma| \to c\,\mathcal{H}^{n-1}(C),$$

where we used the estimate $|Du|_\sigma \leqslant c\sigma^{-n}\|Du\|(\Omega)$ for the first inequality and assumption (GL) for the second. When $\delta\to 0$, C approaches the set consisting of a the finite number of $(n-2)$-planes; hence $\mathcal{H}^{n-1}(C)\to 0$. Thus we have estimated F_σ in the set S_1^σ. The same argument can be repeated for each S_i. This yields the desired estimate for Σ^σ by additivity of F_σ with respect to the set variable. □

5. $\Gamma-$**convergence.** In this section we prove the Γ-convergence to the Mumford–Shah functional. We restrict ourselves to the Geman–McClure functional, i.e. $g(x) = \frac{x}{1+x}$ and $\ell(x) = x$.

Lemma 5.1. *Let $\beta < 1$, $\sigma, \sigma' > 0$ and μ be a finite vector-valued measure. Then*

$$|G_\sigma * \mu(y)| \leqslant c_\beta \Big(\frac{\sigma' + \sigma}{\sigma}\Big)^n G_{\sigma'+\sigma} * |\mu|(x)$$

for all $y \in B(x, \beta\sigma')$, where $c_\beta \to 1$ as $\beta \to 0$.

Proof. Let $x \in \mathbb{R}^n$ and $y \in B(x, \beta\sigma')$. From the definition of G_σ it follows that

$$\sigma^n G_\sigma(z) = G(\tfrac{z}{\sigma}).$$

Let $z \in B(x, \sigma + \sigma')$. We consider two cases. If $z \in B(x, \beta\sigma')$, then

$$G(\tfrac{y-z}{\sigma}) \leqslant G(0) \leqslant \frac{G(0)}{G(\beta\sigma'/(\sigma+\sigma'))} G(\tfrac{x-z}{\sigma+\sigma'}).$$

If $z \notin B(x, \beta\sigma')$, then we denote $|x - z| =: r$ and estimate

$$G(\tfrac{y-z}{\sigma}) \leqslant G(\tfrac{r-\beta\sigma'}{\sigma}) \leqslant \frac{G((r-\beta\sigma')/\sigma)}{G(r/(\sigma+\sigma'))} G(\tfrac{x-z}{\sigma+\sigma'}).$$

Let

$$c_\beta := \sup_{r\in[\beta\sigma', \sigma+\sigma']} \frac{G((r-\beta\sigma')/\sigma)}{G(r/(\sigma+\sigma'))}.$$

If $r = \beta\sigma'$ above, then the value of the function is $\frac{G(0)}{G(\beta\sigma'/(\sigma+\sigma'))}$. At the upper bound it equals 0. Thus by continuity the supremum is finite for every $\beta < 1$. Furthermore, $c_\beta \to 1$ as $\beta \to 0$.

With this estimate we conclude the proof as follows:

$$\begin{aligned}\sigma^n |G_\sigma * \mu(y)| &= \Big| \int_{\mathbb{R}^n} G(\tfrac{y-z}{\sigma})\, d\mu(z) \Big| \\ &\leqslant c_\beta \int_{\mathbb{R}^n} G(\tfrac{x-z}{\sigma+\sigma'})\, d|\mu|(z) = c_\beta (\sigma+\sigma')^n G_{\sigma'+\sigma} * |\mu|(x). \qquad \square\end{aligned}$$

Next we prove the Γ-convergence result.

Proof of Theorem 1.2. We start the proof with the first condition of Γ-convergence. Let $u_i \to u$ in L^1. If $\liminf_{i\to\infty} \mathrm{F}_{\sigma_i}(u_i) = \infty$, then the claim holds. So we may assume that $\liminf_{i\to\infty} \mathrm{F}_{\sigma_i}(u_i) < \infty$. Hence $u_i \in \mathrm{GBV}$ for all sufficient large i (since $\mathrm{F}_{\sigma_i}(u_i) = \infty$ if $u_i \in L^1 \setminus \mathrm{GBV}$), and we may assume that $u_i \in \mathrm{GBV}$ for all i.

Assume first that $u_i \in \mathrm{BV}(\Omega)$ for all i. Let $\beta < 1$ and

$$v_i(x) := (u_i)_\sigma(x) = \int_\Omega G_\sigma(x-y) u_i(y)\, dy,$$

for $x \in \Omega_\sigma$ and $\sigma > 0$. Then for every $x \in \Omega_{\sigma+\sigma'}$ and $y \in B(x, \beta\sigma')$

$$|\nabla v_i(y)| = \big|G_\sigma * Du_i(y)\big| \leqslant c_\beta \Big(\frac{\sigma'+\sigma}{\sigma}\Big)^n \big(G_{\sigma'+\sigma} * |Du_i|(x)\big),$$

by Lemma 5.1. Let $U \subset\subset \Omega$ and suppose that $\sigma > 0$ is so small that $U \subset \Omega_\sigma$. Let $\alpha \in (0,1)$ and set $\sigma = (1-\alpha)\sigma_i$ and $\sigma' = \alpha\sigma_i$. Then the previous inequality implies that

$$|\nabla v_i(y)|^2 \leqslant c_\beta^2 \frac{1}{(1-\alpha)^{2n}} \big(G_{\sigma_i} * |Du_i|(x)\big)^2$$

for all $y \in B(x, \beta\alpha\sigma_i)$. Define $\tilde{v}_i := \frac{(1-\alpha)^n}{\sqrt{\beta\alpha}c_\beta} v_i$. Then we obtain that

$$\underbrace{\beta\alpha\sigma_i \fint_{B(x,\beta\alpha\sigma_i)} |\nabla \tilde{v}_i(y)|^2 \, dy}_{=:A} \leqslant \underbrace{\sigma_i \big(G_{\sigma_i} * |Du_i|(x)\big)^2}_{=:B}$$

for all $x \in U$. Since $t \mapsto \frac{t}{1+t}$ is increasing, it follows that $\frac{A}{1+A} \leqslant \frac{B}{1+B}$. Dividing by σ_i and integrating over $x \in U$, we find that

$$\beta\alpha \, \mathrm{F}^{BD}_{\beta\alpha\sigma_i} \left(\frac{(1-\alpha)^n}{\sqrt{\beta\alpha}c_\beta} (u_i)_{(1-\alpha)\sigma_i}, U \right) = \beta\alpha \, \mathrm{F}^{BD}_{\beta\alpha\sigma_i}(\tilde{v}_i, U) \leqslant \mathrm{F}_{\sigma_i}(u_i, U).$$

Since $\beta\alpha\sigma_i \to 0$ and $\frac{(1-\alpha)^n}{\sqrt{\beta\alpha}c_\beta} (u_i)_{(1-\alpha)\sigma_i} \to \frac{(1-\alpha)^n}{\sqrt{\beta\alpha}c_\beta} u$ in $L^1(U)$ as $\sigma_i \to 0$, it follows from the Γ-convergence of F^{BD}_σ, [7, Theorem 3.1], that

$$\beta\alpha \, \mathrm{MS} \left(\frac{(1-\alpha)^n}{\sqrt{\beta\alpha}c_\beta} u, U \right) \leqslant \liminf_{i\to\infty} \mathrm{F}_{\sigma_i}(u_i, U) \leqslant \liminf_{i\to\infty} \mathrm{F}_{\sigma_i}(u_i),$$

and thus $u|_U \in \mathrm{GSBV}^2(U)$. Since $S_u = S_{tu}$ for $t \neq 0$ constant, we conclude that

$$\begin{aligned} \liminf_{i\to\infty} \mathrm{F}_{\sigma_i}(u_i) &\geqslant \beta\alpha \int_U \left| \frac{(1-\alpha)^n}{\sqrt{\beta\alpha}c_\beta} \nabla_a u \right|^2 dx + \beta\alpha 2\mathcal{H}^{n-1}(S_u \cap U) \\ &= \frac{(1-\alpha)^{2n}}{c_\beta^2} \int_U |\nabla_a u|^2 dx + \beta\alpha 2\mathcal{H}^{n-1}(S_u \cap U) \end{aligned}$$

Then we take $\alpha \to 1$ and $\beta \to 1$ and obtain the lower bound $2\mathcal{H}^{n-1}(S_u \cap U)$. This inequality holds also if we restrict the right hand side to a neighborhood of S_u. If we let $\alpha \to 0$ and $\beta \to 0$, we obtain on the other hand the lower bound $\int_U |\nabla_a u|^2 \, dx$. Combining the part near S_u and the rest part as in the proof of Theorem 3.5, we obtain that

$$\mathrm{MS}(u, U) \leqslant \liminf_{i\to\infty} \mathrm{F}_{\sigma_i}(u_i).$$

Finally, we let $U \to \Omega$ and obtain the result in the case $u_i \in \mathrm{BV}(\Omega)$.

Assume then that $u_i \in \mathrm{GBV}(\Omega)$ for every i. Thus by the previous inequality we obtain

$$\mathrm{MS}(u_\lambda) \leqslant \liminf_{i\to\infty} \mathrm{F}_{\sigma_i}((u_i)_\lambda).$$

for every $\lambda > 0$. It follows from Corollary 3.7 that $\mathrm{F}_{\sigma_i}((u_i)_\lambda) \leqslant \mathrm{F}_{\sigma_i}(u_i)$. Thus

$$\int_\Omega |\nabla_a u_\lambda|^2 \, dx + 2\mathcal{H}^{n-1}(S_{u_\lambda}) \leqslant \liminf_{i\to\infty} \mathrm{F}_{\sigma_i}((u_\lambda)_i) \leqslant \liminf_{i\to\infty} \mathrm{F}_{\sigma_i}(u_i).$$

As $\lambda \nearrow \infty$, $|\nabla_a u_\lambda| \nearrow |\nabla_a u|$ and $S_{u_\lambda} \nearrow S_u$ and hence by monotone convergence

$$\int_\Omega |\nabla_a u|^2 \, dx + 2\mathcal{H}^{n-1}(S_u) \leqslant \liminf_{i\to\infty} \mathrm{F}_{\sigma_i}(u_i).$$

For the second condition of Γ-convergence, we note that Theorem 1.3 actually shows that $\limsup \mathrm{F}_\sigma(u) \leqslant \mathrm{MS}(u)$ for every $u \in L^1$ which satisfies the regularity condition. Furthermore, by [12, Theorem 3.10] we can choose a sequence u_i of functions satisfying the regularity condition of Theorem 1.3 such that $\mathrm{MS}(u_i) \to \mathrm{MS}(u)$. Since $\limsup \mathrm{F}_\sigma(u_i) \leqslant \mathrm{MS}(u_i)$, we can find for each i a value $\hat{\sigma}_i \in (0, \frac{1}{i})$ such that $\mathrm{F}_\sigma(u_i) \leqslant \mathrm{MS}(u_i) + \epsilon$ when $\sigma < \hat{\sigma}_i$. Now we define a new sequence v_i by setting $v_i = u_k$, where $k \in \mathbb{N}$ is the smallest number such that $\sigma_i \in (\hat{\sigma}_k, \hat{\sigma}_{k-1}]$. If no such k exists we set $v_i = u_1$. Since $\sigma_i \to 0$, only a finite number of u_1's will appear in (v_i). For all other terms we have $\mathrm{F}_\sigma(v_i) \leqslant \mathrm{MS}(v_i) + \epsilon$ by construction. For the same reason every other u_k appears at most a finite number of times, so we have $\mathrm{MS}(v_i) \to \mathrm{MS}(u)$ as well, which concludes the proof. □

Acknowledgement: *The authors thank Út V. Lê for his collaboration on this paper's predecessor, which never made it to publication. We also thank Luigi Ambrosio for pointing out a mistake in a preliminary version of this paper and the anonymous referees for corrections and valuable comments.*

REFERENCES

[1] R. Adams and J. Fournier, *Sobolev spaces, Second edition*, Pure and Applied Mathematics (Amsterdam), 140, Elsevier/Academic Press, Amsterdam, 2003.
[2] L. Ambrosio, Compactness for a special class of functions of bounded variation, *Boll. Un. Mat. Ital. B* **7** (1989), 857–881.
[3] L. Ambrosio, N. Fusco and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, Clarendon Press, Oxford University Press, New York, 2000.
[4] L. Ambrosio and V.M. Tortorelli, Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence, *Comm. Pure Appl. Math.* **43** (1990), no. 8, 999–1036.
[5] G. Aubert and P. Kornprobst, *Mathematical problems in image processing, Partial differential equations and the calculus of variations*, Second edition, Applied Mathematical Sciences, 147, Springer, New York, 2006.
[6] B. Bourdin and A. Chambolle, Implementation of an adaptive finite-element approximation of the Mumford-Shah functional, *Numerische Math.* **85** (2000), no. 4, 609–646.
[7] A. Braides and G. Dal Maso, Non-local approximation of the Mumford-Shah functional, *Calc. Var. Partial Differential Equations* **5** (1997), no. 4, 293–322.
[8] A. Chambolle, Image segmentation by variational methods: Mumford and Shah functional and the discrete approximations, *SIAM J. Appl. Math.* **55** (1995), no. 3, 827–863.
[9] A. Chambolle, Finite-differences discretizations of the Mumford-Shah functional, *M2AN Math. Model. Numer. Anal.* **33** (1999), no. 2, 261–288.
[10] A. Chambolle and G. Dal Maso, Discrete approximation of the Mumford-Shah functional in dimension two, *M2AN Math. Model. Numer. Anal.* **33** (1999), no. 4, 651–672.
[11] M. Chipot, R. March, M. Rosati and G. Vergara Caffarelli, Analysis of a nonconvex problem related to signal selective smoothing, *Math. Models Methods Appl. Sci.* **7** (1997), no. 3, 313–328.
[12] G. Cortesani, Strong approximation of GSBV functions by piecewise smooth functions, *Ann. Univ. Ferrara Sez. VII (N.S.)* **43** (1997), 27–49.
[13] G. Dal Maso, *An introduction to* Γ*-convergence*, Progress in Nonlinear Differential Equations and their Applications 8, Birkhuser, Boston, 1993.
[14] G. Dal Maso, J.-M. Morel and S. Solimini, A variational method in image segmentation: existence and approximation results. *Acta Math.* **168** (1992), no. 1-2, 89–151.
[15] G. David, Singular sets of minimizers for the Mumford-Shah functional, *Progress in Mathematics* **233**, Birkhuser Verlag, Basel, 2005.
[16] E. De Giorgi and L. Ambrosio, New functionals in the calculus of variations. (Italian) *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)* 82 (1988), no. 2, 199–210.
[17] L. Evans and R. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, 1992.
[18] S. Geman and D. McClure, Bayesian image analysis. An application to single photon emission tomography, *Proceedings of the American Statistical Association, Statistical Computing Section* (1985), 12–18.
[19] M. Gobbino, Finite difference approximation of the Mumford-Shah functional, *Comm. Pure Appl. Math.* **51** (1998), no. 2, 197–228.
[20] T. Hebert and R. Leahy, A generalised EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors, *IEEE Trans. Med. Imag.* **8** (1989), 194–202.
[21] V. Maz'ya nd S. Poborschi, Differentiable functions on bad domains, World Scientific, Singapore, 1997.
[22] D. Mumford and J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, *Comm. Pure Appl. Math.* **42** (1989), no. 5, 577–685.

[23] M. Rosati, Asymptotic behavior of a Geman and McClure discrete model, *Appl. Math. Optim.* **41** (2000), no. 1, 51–85.

[24] L. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena* **60** (1992), no. 1-4, 259–268.

[25] J. Tiirola, Gamma-convergence of certain modified Perona–Malik functionals, Preprint (2014), submitted.

Received xxxx 20xx; revised xxxx 20xx.