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’ $\Gamma$-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:

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

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 $\varphi$ 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 \( \varphi \)

\[
\int_\Omega \frac{|Du|^2}{1 + |Du|^2} + |u - f|^2 \, dx,
\]
does not have a minimizer in BV and that the energy minimum of the functional equals 0. The corresponding discrete model

\[
\sum_i \frac{|D_hu_i|^2}{1 + |D_hu_i|^2} + |u_i - f_i|^2,
\]
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 \( \Gamma \)-converges 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 \( \text{spt } G \subset B(0, 1) \) 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 \( u \in L^1_{\text{loc}}(\mathbb{R}^n) \), we denote

\[
u_\sigma(x) := G_\sigma \ast u(x) = \int_{B(x, \sigma)} G_\sigma(x - y) u(y) \, dy.
\]

If \( u \in \text{BV}(\Omega) \), then we denote

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

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

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

\[
F_\sigma(u, U) := \int_U \frac{g(\ell(\sigma)|Du|^2_\sigma)}{\ell(\sigma)} \, dx, \quad u \in \text{BV}(\Omega),
\]

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

\[
F_\sigma(u) := F_\sigma(u, \Omega_\sigma) \quad \text{where} \quad \Omega_\sigma := \{ x \in \Omega : d(x, \partial \Omega) > \sigma \}.
\]

We extend \( F_\sigma \) to GBV as a limit,

\[
F_\sigma(u) := \limsup_{\lambda \to \infty} F_\sigma(u_\lambda), \quad u \in \text{GBV}(\mathbb{R}^n),
\]

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

**Assumptions 1.1.** We need the following assumptions on \( g \) and \( \ell \):

- (G) \( g \) in increasing and concave, \( g(0) = 0 \) and \( \lim_{x \to 0^+} \frac{g(x)}{x} = 1 \);
- (L) \( \ell \) is increasing, \( \lim_{x \to 0^+} \frac{\ell(x)}{x} = 0 \) and \( \lim_{x \to 0^+} \frac{\ell(x)}{x^2} = \infty \); and
Point-wise behavior of the Geman–McClure and the Hebert–Leahy models

Point-wise behavior of the Geman–McClure and the Hebert–Leahy models

\[ \lim_{x \to 0^+} \frac{g(c \ell(x)/x^2)}{\ell(x)/x} = 1 \] and \[ \lim_{x \to 0^+} \frac{g(c \ell(x)/x^{2n})}{\ell(x)/x} \leq c_n \] for every fixed \( c > 0 \).

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

\[ g(x) = \lim_{m \to \infty} g \left( \frac{m x}{m} \right) \leq \lim_{m \to \infty} mg \left( \frac{x}{m} \right) = \lim_{m \to \infty} x \frac{g(x/m)}{x/m} = x. \]

Since \( g(0) = 0 \), concavity implies that \( g(ax) \leq ag(x) \) for every \( a \geq 1 \).

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

1. \( g(x) := \frac{x}{1+x} \) with \( \ell(x) := x \)
2. \( g(x) := \log(1 + x) \) with \( \ell(x) := x \log(1 + \frac{1}{x}) \).

These correspond to the Geman–McClure model [18] and the Hebert–Leahy model [20], respectively.

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

\[ F_{BD}^{\sigma}(u) := \int_{\Omega} \frac{1}{2} g \left( \int_{B(x,\sigma) \cap \Omega} |\nabla u(y)|^2 \, dy \right) \, dx, \quad u \in H^1(\Omega); \]

we extend \( F_{BD}^{\sigma} \) to \( L^1 \) by \( \infty \). Braides and Dal Maso have shown that \( F_{BD}^{\sigma} \) \( \Gamma \)-converges to MS with respect to the \( L^1 \)-topology. In Section 5 we will use this to derive the \( \Gamma \)-convergences of our functional in the case \( g(t) = \frac{t}{1+t} \). \( \Gamma \)-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 \( F_{BD}^{\sigma} \) \( \Gamma \)-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 \( \Omega \) be a bounded open set. Suppose that \( u \in \text{SBV}^2(\Omega) \) and \( S_u \) is contained \( \mathcal{H}^{n-1} \)-a.e. in the union of a finite number of hyperplanes. Then

\[
\text{MS}(u) = \lim_{\sigma \to 0^+} F_\sigma(u)
\]

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

The functional \( \lim_{\sigma \to 0} F_\sigma^{BD}(u) \) is finite only for \( H^1 \)-Sobolev functions and thus the previous theorem is not true for \( F_\sigma^{BD} \). Therefore \( F_\sigma \) is point-wise better behaved than \( F_\sigma^{BD} \). 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

\[
\begin{align*}
\int \frac{1}{\sigma} g(\sigma(|\nabla u|^2) \sigma) \, dx, & \quad \int \frac{1}{\sigma} g(\sigma(|\nabla u|^2) \sigma^2) \, dx, & \quad \int \frac{1}{\sigma} g(\sigma(|\nabla (u_{\sigma})|^2) \sigma) \, dx.
\end{align*}
\]

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 \( u_{\sigma} \) 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

$$\|Du\|_\Omega := \sup \left\{ \int_\Omega u \text{div} \varphi \, dx : \varphi \in C_0^\infty(\Omega, \mathbb{R}^n), |\varphi| \leq 1 \right\}$$

and define the space of functions of bounded variation as

$$\text{BV}(\Omega) := \left\{ u \in L^1(\Omega) : \|Du\|_\Omega < \infty \right\}.$$

We define $S_u$ as the subset of $\Omega$ 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

$$\lim_{r \to 0^+} \frac{1}{|B(x,r)|} \int_{B(x,r)} |u(y) - z| \, dy = 0.$$

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

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

where $H^{n-1}(N) = 0$ 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 $u \in \text{BV}(\Omega)$ is a special function of bounded variation and denote $u \in \text{SBV}(\Omega)$ 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 $u_\lambda$ belongs to $\text{BV}$ and $\text{SBV}$, respectively, for all $\lambda > 0$ where $u_\lambda$ is the truncation of $u$ at levels $-\lambda$ and $\lambda$. We write $u \in \text{SBV}^2(\Omega)$ if $u \in \text{SBV}(\Omega) \cap L^2(\Omega)$, $|\nabla_a u| \in L^2(\Omega)$ and $H^{n-1}(S_u) < \infty$.

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 $F_\sigma : X \to \mathbb{R}$ is said to Γ-converge (in topology $\tau$) to $F : X \to \mathbb{R}$ if the following hold for every positive sequence $(\sigma_i)$ converging to zero:

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

$$F(u) \leq \lim \inf_{\sigma_i \to 0^+} F_{\sigma_i}(u_i);$$

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

$$F(u) \geq \lim \sup_{\sigma_i \to 0^+} F_{\sigma_i}(u_i).$$

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 $k > 1$. 

**Inverse Problems and Imaging**

**Volume X, No. X (20xx), X–XX**
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

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

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

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 $u \in BV(\mathbb{R}^n)$, then

$$
2\mathcal{H}^{n-1}(S_u \cap \mathbb{R}^{n-1}) \leq \liminf_{\sigma \to 0^+} F_{\sigma}(u, \Omega^\sigma),
$$

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 $|D_u u|$ restricted to $\mathbb{R}^{n-1}$ such that $|D_u u| \geq v \, d\mathcal{H}^{n-1}$. Then, by monotonicity,

$$
\int_{\mathbb{R}^n} \tilde{g}(\sigma, v_\sigma(x)) \, dx \leq \int_{\mathbb{R}^n} \tilde{g}(\sigma, |D_u u|_\sigma) \, dx,
$$

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

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

For $x \in \mathbb{R}^{n-1}$ we denote by $x_\lambda$ 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 $y \in \mathbb{R}^{n-1}$, $G_\sigma(x_\lambda - y) = \frac{1}{c_\lambda}(G^\lambda)_\sigma(x - y)$, 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

$$
\sigma v_\sigma(x_\lambda) \to c_\lambda v(x)
$$

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

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

$$
E_{\sigma}^r := \{x \in S_u \cap \mathbb{R}^{n-1} | \forall \lambda \in (-r, r), \forall \sigma' \in (0, \sigma], \sigma' v_{\sigma'}(x, \lambda \sigma') \geq \epsilon\}.
$$
Note that $E'_\sigma$ is the intersection of the closed sets $\{\sigma' v_\sigma'(x,\lambda \sigma') \geq \epsilon\}$, so it is itself closed, in particular measurable. With this set, we find that

$$
\int_{\mathbb{R}^n} \hat{g}(\sigma, v_\sigma) \, dx \geq \int_{E'_\sigma \times [-r_\sigma, r_\sigma]} \hat{g}(\sigma, v_\sigma) \, dx
$$

$$
\geq \int_{E'_\sigma \times [-r_\sigma, r_\sigma]} \frac{g(\ell(\sigma)/\sigma^2)}{\ell(\sigma)} \, dx
$$

$$
= \frac{g(\ell(\sigma)/\sigma^2)}{\ell(\sigma)} |E'_\sigma \times [-r_\sigma, r_\sigma]| = 2r c g(\ell(\sigma)/\sigma^2) H^{n-1}(E'_\sigma).
$$

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 $E'_\sigma \cap S_u \cap \mathbb{R}^{n-1}$ as $\sigma \to 0$ and $\epsilon \to 0$, up to a set of $H^{n-1}$-measure zero. Hence

$$
2 H^{n-1}(S_u \cap \mathbb{R}^{n-1}) = \lim_{\sigma \to 0^+} 2 H^{n-1}(E'_\sigma) = \lim_{\sigma \to 0^+} 2 \frac{g(\ell(\sigma)/\sigma^2)}{\ell(\sigma)/\sigma} H^{n-1}(E'_\sigma)
$$

$$
\leq \frac{1}{r} \liminf_{\sigma \to 0^+} \int_{\mathbb{R}^n} \hat{g}(\sigma, v_\sigma) \, dx.
$$

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

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

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

$$
F_\sigma(u \circ \Phi, \Omega) \leq c_L F_\sigma(u, \Phi(\Omega)),
$$

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

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

$$
|D(u \circ \Phi)|_\sigma(x) = \int_{\mathbb{R}^n} G_\sigma(x - y) \, d|D(u \circ \Phi)|(y)
$$

$$
\leq L^{n-1} \int_{\mathbb{R}^n} G_\sigma(x - y) \, d(\Phi^{-1})\# |Du|(y)
$$

$$
= L^{n-1} \int_{\mathbb{R}^n} G_\sigma(x - \Phi^{-1}(z)) \, d|Du|(z).
$$

Since $\Phi$ is bilipschitz and $G_\sigma$ is radially decreasing and symmetric we further have

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

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

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

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

$$
\hat{g}(\sigma, |D(u \circ \Phi)|_\sigma(x)) \leq \hat{g}(\sigma, d_L |Du|_\sigma(\Phi(x))) \leq d_L^2 \hat{g}(\sigma, |Du|_\sigma(\Phi(x))),
$$
Lemma 3.4. If \( \sigma > \infty \), then we calculate \( c < \infty \).

Proof. Choose a decreasing sequence \( \sigma_k \) such that \( \ell(\sigma_k) \leq 2^{-k} \) and \( \lim_k F_{\sigma_k}(u) < c < \infty \). Fix \( \epsilon > 0 \) and set

\[
H_k := \{ x \in \Omega \mid (\ell(\sigma_k)|Du|_{\sigma_k}^2 > \epsilon) \}
\]

and \( G_j := \bigcup_{k \geq j} H_k \).

Then we calculate

\[
(7) \quad c > F_{\sigma_k}(u) \geq \int_{H_k} g(\ell(\sigma_k)|Du|_{\sigma_k}^2) \, dx \geq \int_{H_k} g(\epsilon) \ell(\sigma_k) \, dx = \frac{g(\epsilon)}{\ell(\sigma_k)} |H_k|.
\]

It follows that \( |H_k| \leq c(\epsilon) \ell(\sigma_k) \leq c(\epsilon)2^{-k} \) and \( |G_j| \leq c(\epsilon) \sum_{k \geq j} 2^{-k} \leq c(\epsilon) 2^{-j} \).

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

\[
2\mathcal{H}^{n-1}(S_u \cap S) \leq \liminf_{\sigma \to 0^+} F_{\sigma}(u, \Sigma^\sigma),
\]

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

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 \( \mathcal{H}^{n-1}(S_u \cap U) \leq L^{-1}\mathcal{H}^{n-1}(\Phi(S_u \cap U)) \).

By Lemma 3.2,

\[
c_L \liminf_{\sigma \to 0^+} F_{\sigma}(u, \Sigma^\sigma \cap U) \geq \liminf_{\sigma \to 0^+} F_{\sigma}(u, \Sigma^\sigma \cap \Phi(U)).
\]

Then it follows from Lemma 3.1 that

\[
L^{-1}c_L \liminf_{\sigma \to 0^+} F_{\sigma}(u, \Sigma^\sigma \cap U) \geq 2\mathcal{H}^{n-1}(S_u \cap U).
\]

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

\[
\sum_k \mathcal{H}^{n-1}(S_u \cap U_k) = \mathcal{H}^{n-1}\left( S_u \cap \bigcup_k U_k \right) \geq \mathcal{H}^{n-1}(S_u) - \epsilon.
\]

Then

\[
L^{-1}c_L \liminf_{\sigma \to 0^+} F_{\sigma}(u, \Sigma^\sigma) \geq \sum_k \liminf_{\sigma \to 0^+} F_{\sigma}(u, \Sigma^\sigma \cap U_k) \geq 2\mathcal{H}^{n-1}(S_u) - 2\epsilon,
\]

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

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 \( u \in BV(\Omega) \), then

\[
\int_{\Omega} |\nabla u|^2 \, dx \leq \liminf_{\sigma \to 0^+} F_{\sigma}(u).
\]

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

\[
H_k := \{ x \in \Omega \mid \ell(\sigma_k)|Du|_{\sigma_k}^2 > \epsilon \}
\]

and \( G_j := \bigcup_{k \geq j} H_k \).

Then we calculate

\[
\int_{H_k} g(\ell(\sigma_k)|Du|_{\sigma_k}^2) \, dx \geq \int_{H_k} g(\epsilon) \ell(\sigma_k) \, dx = \frac{g(\epsilon)}{\ell(\sigma_k)} |H_k|.
\]

It follows that \( |H_k| \leq c(\epsilon) \ell(\sigma_k) \leq c(\epsilon)2^{-k} \) and \( |G_j| \leq c(\epsilon) \sum_{k \geq j} 2^{-k} \leq c(\epsilon) 2^{-j} \).
Define $\Omega_j := \Omega \setminus G_j$. In $\Omega_j$, $\ell(\sigma_k)|Du|_{\sigma_k}^2 \leq \epsilon$ for every $k \geq j$. Since $g(t)/t \to 1$ as $t \to 0^+$ (by assumption (G)), we conclude that
\[
g(\ell(\sigma_k)|Du|_{\sigma_k}^2) \geq c'(\epsilon)\ell(\sigma_k)|Du|_{\sigma_k}^2,
\]
where $c'(\epsilon) \to 1$ as $\epsilon \to 0^+$. Hence
\[
\int_{\Omega} g(\ell(\sigma_k)|Du|_{\sigma_k}^2) \ell(\sigma_k) \, dx \geq c'(\epsilon) \int_{\Omega_j} |Du|_{\sigma_k}^2 \, dx.
\]
Since $\nabla_a u(x) \, dx$ and $D_\sigma$ are mutually singular measures \cite[Proposition 3.92, p. 184]{3}, we obtain in every Borel set $E$ that $\|Du\|(E)$ equals
\[
\|\nabla_a u \, dx + D_\sigma u\|(E) = \|\nabla_a u\|\,(E) + \|D_\sigma u\|\,(E) = \int_E |\nabla_a u| \, dx + \|D_\sigma u\|\,(E).
\]
Thus
\[
|Du|_\sigma(x) = \int G_\sigma(x - y) \, d|Du|(y)
= \int G_\sigma(x - y)|\nabla_a u(y)| \, dy + \int G_\sigma(x - y) \, d|D_\sigma u|(y)
= |\nabla_a u|_\sigma(x) + |D_\sigma u|_\sigma(x).
\]
Since $\nabla_a u \in L^1(\Omega)$ we find that
\[
\liminf_{\sigma \to 0} |Du|_\sigma = \liminf_{\sigma \to 0} \left[ |\nabla_a u|_\sigma + |D_\sigma u|_\sigma \right] \geq \liminf_{\sigma \to 0} |\nabla_a u|_\sigma = |\nabla_a u|
\]
almost everywhere. Therefore the previous estimate implies that
\[
\int_{\Omega_j} |\nabla_a u|^2 \, dx \leq \liminf_{k \to \infty} \int_{\Omega_j} |Du|^2_{\sigma_k} \, dx \leq \frac{1}{c'(\epsilon)} \liminf_{\sigma \to 0^+} F_\sigma(u).
\]
Hence $\nabla_a u \in L^2(\Omega_j)$ with norm independent of $j$. Then letting $j \to \infty$ yields that $\nabla_a u \in L^2(\bigcup_j \Omega_j)$. Since $|\Omega \setminus \bigcup_j \Omega_j| = 0$ 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 $\epsilon \to 0$.

\textbf{Theorem 3.5.} Let $\Omega \subset \mathbb{R}^n$ be an open bounded set. If $u \in SBV(\Omega)$, then
\[
\int_{\Omega} |\nabla_a u|^2 \, dx + 2H^{n-1}(S_u) \leq \liminf_{\sigma \to 0^+} F_\sigma(u).
\]
\textbf{Proof.} If we show that
\[
\int_U |\nabla_a u|^2 \, dx + 2H^{n-1}(S_u \cap U) \leq \liminf_{\sigma \to 0^+} F_\sigma(u, U)
\]
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
\[
u_{\text{ext}}(x) := \max\left\{0, 1 - \frac{\text{dist}(x, V)}{\text{dist}(\partial V, \partial \Omega)}\right\} u(x)
\]
in $\Omega$ and $\nu_{\text{ext}} = 0$ in $\mathbb{R}^n \setminus \Omega$. Now $\nu_{\text{ext}} \in L^\infty(\mathbb{R}^n) \cap SBV(\mathbb{R}^n)$ and $u|_V = \nu_{\text{ext}}|_V$. Hence for all sufficiently small $\sigma$, $F_\sigma(u, U) = F_\sigma(\nu_{\text{ext}}, U)$. Therefore it suffices to prove the claim of the theorem under the assumption $u \in SBV(\mathbb{R}^n)$. 

\textit{Inverse Problems and Imaging} Volume X, No. X (20xx), X–XX
Let $K_i \subset U$ be a countable collection of compact sets, each contained in a $C^1$-hypersurface, such that $S_u \cap U = N \cup \bigcup_i K_i$ where $\mathcal{H}^{n-1}(N) = 0$. It suffices to show that

$$\int_U |\nabla u|^2 \, dx + 2\mathcal{H}^{n-1}\left( \bigcup_{i=1}^m K_i \right) \leq \liminf_{\sigma} F_\sigma(u),$$

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 $\mathcal{H}^{n-1}(V_1 \cap K) \leq \mathcal{H}^{n-1}(K_1) + \varepsilon/m$. Let $V_2$ be an open neighborhood of $K_2 := K_2 \setminus V_1$ such that $\mathcal{H}^{n-1}(V_2 \cap K) \leq \mathcal{H}^{n-1}(K_2 \setminus V_1) + \varepsilon/m$. 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 $\mathcal{H}^{n-1}(V_m \cap K) \leq \mathcal{H}^{n-1}(K_m \setminus (V_1 \cup \ldots \cup V_{m-1})) + \varepsilon/m$. Now

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

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

Let $N^k$ be the $\sigma$-neighborhood of $K_i^k$, $\sigma > 0$. For $\sigma < \frac{1}{6}\sigma_0$, the sets $N_j^{k_0}$ and $N_j^{k_0}$ are disjoint ($j \neq k$), and when $\sigma < \frac{1}{3}\text{dist}(\partial V, \partial U)$ they are contained in $V$. Thus Corollary 3.3 can be applied to the hypersurfaces individually, and it gives that

$$2\mathcal{H}^{n-1}(K_i^k) \leq c_\sigma F_\sigma(u, N^k),$$

where $c_\sigma \to 1$ as $\sigma \to 0$. By Lemma 3.4 applied in the set $U \setminus \bigcup N^k_\sigma$, we obtain

$$\int_U |\nabla u|^2 \, dx \leq c_\sigma F_\sigma(u, U \setminus \bigcup N^k_\sigma) + \int_{U \setminus \bigcup N^k_\sigma} |\nabla u|^2 \, dx,$$

with $c_\sigma$ converging to 1 as before. In view of this and the absolute continuity of the last integral,

$$\int_U |\nabla u|^2 \, dx + 2\mathcal{H}^{n-1}\left( \bigcup_{k=1}^m K_k' \right) \leq \liminf_{\sigma} c_\sigma \left[ F_\sigma(u, U \setminus \bigcup N^k_\sigma) + \sum_{k=1}^m F_\sigma(u, N^k_\sigma) \right].$$

To complete the proof we observe that the sets $U \setminus \bigcup N^k_\sigma$ and $N^k_\sigma$ are disjoint, even when dilated by $\sigma$, so that

$$F_\sigma(u, U \setminus \bigcup N^k_\sigma) + \sum_k F_\sigma(u, N^k_\sigma) \leq F_\sigma(u, V)$$

for $\sigma < \min\left\{ \frac{1}{3}\sigma_0, \text{dist}(\partial V, \partial U) \right\}$. Together with the previous estimate, this completes the proof. \hfill \Box

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 $u \in \text{BV}(\Omega)$ and $E \subset \Omega$ is Borel, then $|D(u_\lambda)|(E) \leq |Du|(E)$.

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

**Proof.** Since $G_\sigma$ is radially symmetric and decreasing, we have

$$G_\sigma \ast |Du|(x) = \int_0^\sigma h(r, \sigma)|Du|(B(x, r)) \, dr,$$
where \( h \geq 0 \) is a suitably chosen weight function. This representation also holds for \( u_\lambda \). Since \( |Du_\lambda|(B(x, r)) \leq |Du|(B(x, r)) \) by Lemma 3.6, the claim follows. \( \square \)

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 \( u \in \text{GSBV}(\Omega) \), then
\[
\text{MS}(u) \leq \liminf_{\sigma \to 0^+} F_\sigma(u).
\]

**Proof.** By monotonicity of our functional (assumption (G)), it follows from Corollary 3.7 that \( F_\lambda(u_\lambda) \leq F_\lambda(u) \) for every \( \lambda > 0 \). Theorem 3.5 yields that
\[
\int_{\Omega} |\nabla u_\lambda|^2 \, dx + 2H^{n-1}(S_{u_\lambda}) \leq \liminf_{\sigma \to 0^+} F_\sigma(u_\lambda) \leq \liminf_{\sigma \to 0^+} F_\sigma(u).
\]

When \( \lambda \nearrow \infty \), \( |\nabla u_\lambda| \nearrow |\nabla u| \) and \( S_{u_\lambda} \nearrow S_u \). Hence, by monotone convergence,
\[
\int_{\Omega} |\nabla u|^2 \, dx + 2H^{n-1}(S_u) \leq \liminf_{\sigma \to 0^+} F_\sigma(u). \quad \square
\]

4. Upper bound. In this section we prove the upper bound of \( \text{limsup} F_\sigma \) 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 \( \frac{1}{|Q|} \int_Q f \, dx \) of \( f \) over the set \( Q \) and by \( Q(z, \sigma) \) a cube with center \( z \) and side-length \( \sigma \).

**Lemma 4.1.** Let \( \Omega \subset \mathbb{R}^n \) be an open set, \( 1 \leq p < \infty \) and \( s \geq 0 \). 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,
\[
\int_{\Omega} V_\sigma(z) \, dz \leq \int_{\Omega} \int_{\Omega} \frac{|u(x) - v(z)|^p}{|x-z|^{n+sp}} \, dx \, dz < \infty.
\]

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

Consider the sets \( A := \{ z \in \Omega \mid V_\sigma(z) < \infty \text{ for some } \sigma \} \) and \( B := \{ z \in \Omega \mid V_\sigma(z) = \infty \text{ for all } \sigma \} \). 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
\[
\frac{1}{\sigma^s} |v_{Q(z, \sigma)}| \leq \frac{1}{\sigma^s} \int_{Q(z, \sigma)} |v(x)| \, dx \leq \left( \frac{1}{\sigma^{sp}} \int_{Q(z, \sigma)} |v(x) - v(z)|^p \, dx \right)^{1/p} \leq cV_\sigma(z)^{1/p}
\]
since \(|x - z| < 2\sigma\) when \(x \in Q(z, \sigma)\). The right hand side tends to zero, giving the claim. \(\square\)

**Lemma 4.2.** Suppose that \(u \in SBV^2(\Omega)\) in a rectangle \(\Omega\) with \(S_u \subset \mathbb{R}^{n-1}\). Then
\[
\limsup_{\sigma \to 0^+} \int_{U^\sigma} \hat{g}(\sigma, |D_s u|_\sigma) \, dx \leq 2\mathcal{H}^{n-1}(S_u),
\]
where \(U^\sigma := (\mathbb{R}^n \times [-\sigma, \sigma]) \cap \Omega_\sigma\).

**Proof.** As before, let \(v : \mathbb{R}^{n-1} \to \mathbb{R}\) be the representative of the singular part of \(u\), \(D_s u = v d\mathcal{H}^{n-1}\). Then \(|v| = |u^+ - u^-|\), where \(v^\pm\) is the approximately continuous extension of \(u\) to \(\mathbb{R}^n\) from the positive and negative half-spaces [3, (4.1)]. Since \(S_u \cap \mathbb{R}^n_+ = \emptyset\), it follows that \(D_u = \nabla u u \in \Omega \cap \mathbb{R}^n_+\) and since \(u \in SBV^2(\Omega)\), 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 \(H^\frac{1}{2}(\mathbb{R}^{n-1})\) [1, Theorem 7.39]; the same holds for their difference \(v\).

We have
\[G_\sigma * |D_s u| \leq ||G||_{L^\infty} \frac{1}{\sigma^n} \chi_Q(\sigma) * |D_s u|\]
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
\[
\frac{1}{\sigma^n} \chi_Q(\sigma) * |D_s u|(x', x_n) =: \frac{1}{\sigma} V_\sigma
\]
is independent of \(x_n\), where \(V_\sigma\) denotes the \((n-1)\)-dimensional average of \(|v|\) over the cube \(Q(x', \sigma) \cap \mathbb{R}^{n-1}\). Thus
\[
\int_{U^\sigma} \frac{\hat{g}(\sigma, G_\sigma * |D_s u|)}{\sigma} d\mathcal{H}^{n-1}(x') = 2 \int_{\Omega_\sigma \cap \mathbb{R}^{n-1}} \sigma \frac{\hat{g}(\sigma, \frac{1}{\sigma} V_\sigma)}{\sigma} d\mathcal{H}^{n-1}(x')
\]
By assumptions (G) and (L) we have the following two estimates:
\[
\sigma \hat{g}(\sigma, \frac{1}{\sigma} V_\sigma) = \frac{g(c^2 V_\sigma^2 \ell(\sigma)/\sigma^2)}{\ell(\sigma)/\sigma} \leq \begin{cases} \frac{\hat{g}^2(V_\sigma^2)}{\ell(\sigma)/\sigma}, & \min \{1, c^2 V_\sigma^2\} \frac{g(\ell(\sigma)/\sigma)}{\ell(\sigma)/\sigma}. \end{cases}
\]
By assumption (GL) \(\hat{g}(\ell(\sigma)/\sigma)\) tends to 1 as \(\sigma \to 0^+\). 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 \(\sigma \hat{g}(\sigma, \frac{1}{\sigma} V_\sigma)\) a majorant to be used in the dominated convergence theorem.

Let us next look the point wise convergence. Since \(|v| \in H^\frac{1}{2}(\mathbb{R}^{n-1})\) we have that \(\lim_{\sigma \to 0^+} V_\sigma = 0\) for \(\mathcal{H}^{n-1}\)-almost all points. Assume first that \(\lim_{\sigma \to 0^+} V_\sigma = 0\) and let us estimate the upper bound \(\frac{\hat{g}^2(V_\sigma^2)}{\sigma V_\sigma^2}\). By (G), \(\frac{\hat{g}^2(V_\sigma^2)}{\sigma V_\sigma^2} \leq \frac{c^2 V_\sigma^2}{\sigma V_\sigma^2}\) and by Lemma 4.1, in dimension \(n-1\) (with \(s = \frac{1}{2}\) and \(p = 2\)), \(\frac{c^2 V_\sigma^2}{\sigma V_\sigma^2}\) tends to zero for \(\mathcal{H}^{n-1}\)-almost all points in the set \(\{v = 0\}\). Assume next that \(\lim_{\sigma \to 0^+} V_\sigma = c_1 \in (0, \infty)\). Then there exists \(\sigma_0\) such that \(\frac{1}{2} c_1 \leq V_\sigma \leq 2c_1\) for every \(0 < \sigma < \sigma_0\). Since \(\hat{g}\) is
increasing we obtain
\[
\frac{g}{\ell} \left( \frac{c^2 \ell^2}{\sigma} \right) \leq \frac{\sigma \hat{g}(\sigma, \frac{x}{\sigma})}{\ell(\sigma)} \leq \frac{\sigma \hat{g}(\sigma, 2 \ell/\sigma)}{\ell(\sigma)} = \frac{g \left( \frac{4c^2 \ell}{\sigma} \right)}{\ell(\sigma)}
\]
for every \(0 < \sigma < \sigma_0\), and hence by assumption (GL) the term \(\hat{g}(\sigma, \frac{x}{\sigma})\) tends to 1 as \(\sigma \to 0^+\). Hence it follows by dominated convergence that

\[
\limsup_{\sigma \to 0^+} \int_{\Omega^0 \cap \mathbb{R}^{n-1}} \sigma \hat{g}(\sigma, \frac{x}{\sigma}) \, d\mathcal{H}^{n-1}(x) \leq \mathcal{H}^{n-1}(\mathbb{R}^{n-1} \cap \Omega \cap \{v \neq 0\}) = \mathcal{H}^{n-1}(S_u).
\]

This gives the claim.

**Proof of Theorem 1.3.** Let \(\Sigma^\sigma := \{x \in \Omega^\sigma \mid d(x, S_u) < \sigma\}\). By the triangle inequality \(|Du|_\sigma \leq |\nabla u|_\sigma + |D_s u|_\sigma\), and further \(|Du|_\sigma = |\nabla u|_\sigma\) in \(\Omega^\sigma \setminus \Sigma^\sigma\) since the support of \(G_{\sigma}\) does not intersect \(S_u\) in this case. Moreover,

\[
|Du|^2_\sigma \leq (|\nabla u|_\sigma + |D_s u|_\sigma)^2 \leq c_\epsilon |\nabla u|_\sigma^2 + (1 + \epsilon) |D_s u|_\sigma^2
\]
in \(\Sigma^\sigma\). Since \(g\) is subadditive,

\[
F_\sigma(u) = \int_{\Omega^\sigma} \hat{g}(\sigma, |Du|_\sigma) \, dx \leq \int_{\Omega^\sigma} \hat{g}(\sigma, |\nabla u|_\sigma) \, dx + c_\epsilon \int_{\Omega^\sigma \setminus \Sigma^\sigma} \hat{g}(\sigma, |\nabla u|_\sigma) \, dx + (1 + \epsilon) \int_{\Sigma^\sigma} \hat{g}(\sigma, |D_s u|_\sigma) \, dx.
\]

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

\[
\int_{\Omega^\sigma} \hat{g}(\sigma, |\nabla u|_\sigma) \, dx \leq \int_{\Omega^\sigma} |\nabla u|_\sigma^2 \, dx \to \int_{\Omega} |\nabla u|_\sigma^2 \, dx
\]
as \(\sigma \to 0^+\). For the second integral we find that

\[
c_\epsilon \int_{\Omega^\sigma} \hat{g}(\sigma, |\nabla u|_\sigma) \, dx \leq c_\epsilon \int_{\Omega^\sigma} |\nabla u|_\sigma^2 \, dx \leq c_\epsilon \int_{\Sigma^\sigma} |\nabla u|_\sigma^2 \, dx \to c_\epsilon \int_{\Omega^\sigma} |\nabla u|_\sigma^2 \, dx,
\]

when \(\sigma \leq \sigma_0\). 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 \(\Sigma^\sigma\) by the Hausdorff measure.

By assumption the set \(S_u\) lies \(\mathcal{H}^{n-1}\)-a.e. in the finite union \(\bigcup S_i\) of hyperplanes.

Consider the components \(K\) of the hyperplane \(S_1\) lying at distance at least \(\delta\) 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

\[
\limsup_{\sigma \to 0} F_\sigma(u, K^\sigma) \leq 2\mathcal{H}^{n-1}(S_1 \cap S_u);
\]

here we denote by \(K^\sigma\) the points within distance \(\sigma\) of \(K\).

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

\[
\limsup_{\sigma \to 0} F_\sigma(u, C^\sigma) \leq \int_{C^\sigma} \hat{g}(\sigma, \frac{x}{\sigma}) \, dx \leq \frac{c}{\sigma} |C^\sigma| \to c\mathcal{H}^{n-1}(C),
\]

where we used the estimate \(|Du|_\sigma \leq c\sigma^{-n} \|Du\|_\Omega\) 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 \(F_\sigma\) in the set \(S_1^\sigma\). The same argument can be repeated for each \(S_i\). This yields the desired estimate for \(\Sigma^\sigma\) by additivity of \(F_\sigma\) with respect to the set variable. \(\square\)
5. \(\Gamma\)-convergence. In this section we prove the \(\Gamma\)-convergence to the Mumford–Shah functional. We restrict ourselves to the Geman–McClure functional, i.e. \(g(x) = \frac{x^2}{1 + x^2}\) and \(\ell(x) = x\).

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

\[
|G_\sigma \ast \mu(y)| \leq c_\beta \left( \frac{\sigma' + \sigma}{\sigma} \right)^n G_{\sigma'+\sigma} \ast |\mu|(x)
\]

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 \(G_\sigma\) it follows that

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

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

\[
G\left(\frac{y-z}{\sigma}\right) \leq G(0) \leq \frac{G(0)}{G(\beta \sigma'/(\sigma + \sigma'))} G\left(\frac{y-z}{\beta \sigma'}\right).
\]

If \(z \not\in B(x, \beta \sigma')\), then we denote \(|x-z| =: r\) and estimate

\[
G\left(\frac{y-z}{\sigma}\right) \leq G\left(\frac{r-\beta \sigma'}/\sigma\right) \leq \frac{G\left((r-\beta \sigma')/\sigma\right)}{G\left(r/(\sigma + \sigma')\right)} G\left(\frac{z-r}{\sigma + \sigma'}\right).
\]

Let

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

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:

\[
\sigma^n |G_\sigma \ast \mu(y)| = \left| \int_{\mathbb{R}^n} G\left(\frac{y-z}{\sigma}\right) \, d\mu(z) \right| 
\leq c_\beta \int_{\mathbb{R}^n} G\left(\frac{z-r}{\sigma + \sigma'}\right) \, |\mu|(z) = c_\beta (\sigma + \sigma')^n G_{\sigma'+\sigma} \ast |\mu|(x).
\]

Next we prove the \(\Gamma\)-convergence result.

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

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

\[
v_i(x) := (u_i)_{\alpha}(x) = \int_{\Omega} G_\sigma(x-y)u_i(y) \, dy,
\]

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

\[
|\nabla v_i(y)| = |G_\sigma \ast Du_i(y)| \leq c_\beta \left( \frac{\sigma' + \sigma}{\sigma} \right)^n (G_{\sigma'+\sigma} \ast |Du_i|(x)),
\]

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 \(\sigma = (1-\alpha)\sigma_1\) and \(\sigma' = \alpha \sigma_1\). Then the previous inequality implies that

\[
|\nabla v_i(y)|^2 \leq c_\beta^2 \frac{1}{(1-\alpha)^2n} (G_{\sigma_1} \ast |Du_i|(x))^2.
\]
for all $y \in B(x, \beta \alpha \sigma_i)$. Define $\tilde{v}_i := \frac{(1-\alpha)^n}{\sqrt{\beta \alpha c_\beta}} u_i$. Then we obtain that
\[
\beta \alpha \sigma_i \int_{B(x, \beta \alpha \sigma_i)} |\nabla \tilde{v}_i(y)|^2 \, dy \leq \sigma_i (G_{\sigma_i} * |Du_i|(x))^2 \bigg|_B
\]
for all $x \in U$. Since $t \mapsto \frac{t}{1+t}$ is increasing, it follows that $\frac{A}{1+A} \leq \frac{B}{1+B}$. Dividing by $\sigma_i$ and integrating over $x \in U$, we find that
\[
\beta \alpha F_{\beta \alpha \sigma_i}^{BD} \left( \frac{(1-\alpha)^n}{\sqrt{\beta \alpha c_\beta}} (u_i)_{(1-\alpha)\sigma_i}, U \right) = \beta \alpha F_{\beta \alpha \sigma_i}^{BD}(\tilde{v}_i, U) \leq F_{\sigma_i}(u_i, U).
\]
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 $\Gamma$-convergence of $F_{\sigma_i}^{BD}$, [7, Theorem 3.1], that
\[
\beta \alpha M_S \left( \frac{(1-\alpha)^n}{\sqrt{\beta \alpha c_\beta}} u, U \right) \leq \liminf_{i \to \infty} F_{\sigma_i}(u_i, U) = \liminf_{i \to \infty} F_{\sigma_i}(u_i),
\]
and thus $u|_U \in GSBV^2(U)$. Since $S_u = S_{u_t}$ for $t \neq 0$ constant, we conclude that
\[
\liminf_{i \to \infty} F_{\sigma_i}(u_i) \geq \beta \alpha \int_U \left( \frac{(1-\alpha)^n}{\sqrt{\beta \alpha c_\beta}} \nabla_a u \right)^2 \, dx + \beta \alpha 2\mathcal{H}^{n-1}(S_u \cap U)
\]
\[
= \frac{(1-\alpha)^n}{c_\beta^n} \int_U |\nabla_a u|^2 \, dx + \beta \alpha 2\mathcal{H}^{n-1}(S_u \cap U)
\]
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
\[
MS(u_i, U) \leq \liminf_{i \to \infty} F_{\sigma_i}(u_i).
\]
Finally, we let $U \to \Omega$ and obtain the result in the case $u_i \in BV(\Omega)$.

Assume then that $u_i \in GBV(\Omega)$ for every $i$. Thus by the previous inequality we obtain
\[
MS(u, \lambda) \leq \liminf_{i \to \infty} F_{\sigma_i}((u_i)\lambda).
\]
for every $\lambda > 0$. It follows from Corollary 3.7 that $F_{\sigma_i}((u_i)\lambda) \leq F_{\sigma_i}(u_i)$. Thus
\[
\int_{\Omega} |\nabla_a u\lambda|^2 \, dx + 2\mathcal{H}^{n-1}(S_u\lambda) \leq \liminf_{i \to \infty} F_{\sigma_i}((u_i)\lambda) \leq \liminf_{i \to \infty} F_{\sigma_i}(u_i).
\]
As $\lambda \to \infty$, $|\nabla_a u\lambda| \to |\nabla_a u|$ and $S_u\lambda \to S_u$ and hence by monotone convergence
\[
\int_{\Omega} |\nabla_a u|^2 \, dx + 2\mathcal{H}^{n-1}(S_u) \leq \liminf_{i \to \infty} F_{\sigma_i}(u_i).
\]
For the second condition of $\Gamma$-convergence, we note that Theorem 1.3 actually shows that $\limsup F_{\sigma}(u) \leq MS(u)$ 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 $MS(u_i) \to MS(u)$. Since $\limsup F_{\sigma}(u_i) \leq MS(u_i)$, we can find for each $i$ a value $\delta_i \in (0, \frac{1}{2})$ such that $F_{\sigma}(u_i) \leq MS(u_i) + \epsilon$ when $\sigma < \delta_i$. 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 $\sigma_i \in (\delta_k, \delta_{k-1}]$. If no such $k$ exists we set $v_i = u_1$. Since $\sigma_i \to 0$, only a finite number of $u_i$’s will appear in $(v_i)$. For all other terms we have $F_{\sigma}(v_i) \leq MS(v_i) + \epsilon$ by construction. For the
same reason every other $u_k$ appears at most a finite number of times, so we have $\text{MS}(v_i) \rightarrow \text{MS}(u)$ 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


Received xxxx 20xx; revised xxxx 20xx.