APPROXIMATION OF THE MUMFORD-SHAH FUNCTIONAL
BY PHASE FIELDS OF BOUNDED VARIATION

SANDRO BELZ AND KRISTIAN BREDIES

1. Introduction and Main Results

In this paper we introduce a new phase-field approximation by \( \Gamma \)-convergence of the Mumford-Shah functional, which has been introduced by D. Mumford and J. Shah in [21] in the context of image segmentation. For a domain \( \Omega \subset \mathbb{R}^2 \) of a given 2-dimensional image \( g \in L^\infty(\Omega) \) they suggest the minimization of the functional

\[
\alpha \int_\Omega |\nabla u|^2 \, dx + \beta \int_\Omega |u - g|^2 \, dx + \gamma \mathcal{H}^1(\Gamma)
\]

with respect to \( u \in C^1(\Omega \setminus \Gamma) \) and \( \Gamma \subset \Omega \) closed. The solution \((u, \Gamma)\) of this minimization problem represents the approximation of the image \( g \) and the borders of the detected segments.

Until recently this model has been the subject of several theoretical and numerical investigations. It even became more important when the functional was introduced in [17] in the context of fracture mechanics. The difficulty in minimizing (1.1) certainly lies in the dependency of the closed set \( \Gamma \). It could be overcome by relaxing the functional to the space of special functions of bounded variations, where the set \( \Gamma \) is replaced by the discontinuity set \( S_u \) (see Section 2.3 for more details on these functions). Namely, one defines

\[
\mathcal{MS}(u) = \alpha \int_\Omega |\nabla u|^2 \, dx + \beta \int_\Omega |u - g|^2 \, dx + \gamma \mathcal{H}^1(S_u) .
\]

for \( u \in \text{SBV}(\Omega) \), the set of special functions of bounded variation.

For the relaxed functional \( \mathcal{MS} \) the existence of minimizers could be proved with the help of a compactness property in SBV(\( \Omega \)), which includes some semi-continuity properties and has been shown in [1–3]. Furthermore, it is well-known that for a minimizer \( u \in \text{SBV}(\Omega) \) of \( \mathcal{MS} \) the pair \((u, S_u)\) is a minimizer of (1.1). This is due to a regularity property for \( S_u \), proven by E. De Giorgi in [13], namely \( \mathcal{H}^1(S_u \setminus S_{\bar{u}}) = 0 \).

Because of the second integral term of \( \mathcal{MS} \), any minimizer of \( \mathcal{MS} \) is also in \( L^\infty(\Omega) \), which follows by a simple truncation argument. However, in the field of fracture mechanics this part of the functional is not present and theoretically the function \( u \), representing the displacement, can be arbitrarily large. For this reason generalized special functions of bounded variation have been introduced and in what follows we define the Mumford-Shah functional on the set of such functions denoted by GSBV(\( \Omega \)). In this way our results can also be adapted to cover a broader perspective. A precise definition of these functions is given in Section 2.3.

A very powerful tool in order to tackle the problem numerically is the variational approximation of the Mumford-Shah functional in terms of \( \Gamma \)-convergence. The basic idea is to find a sequence of functionals whose minimizers are easy to compute; the \( \Gamma \)-convergence then ensures that these minimizers converge to a solution of the
original problem (see Section 2.2). The probably most frequently used approach using this technique goes back to the work of L. Ambrosio and V.M. Tortorelli in [5,6]. In the latter they define the functionals \( \mathcal{AT}_\varepsilon: L^1(\Omega) \to \mathbb{R} \) for \( \varepsilon > 0 \) by

\[
(1.2) \quad \mathcal{AT}_\varepsilon(u,v) = \int_\Omega (v^2 + \eta_\varepsilon)|\nabla u|^2 \, dx + \frac{1}{4\varepsilon} \int_\Omega (1-v)^2 \, dx + \varepsilon \int_\Omega |\nabla v|^2 \, dx
\]

for \( u \in H^1(\Omega) \) and \( v \in H^1(\Omega; [0,1]) \) and \( \mathcal{AT}_\varepsilon = +\infty \) otherwise, and they prove that these functionals \( \Gamma \)-converge to the Mumford-Shah functional defined in the following form: \( \overline{\mathcal{MS}}: L^1(\Omega) \times L^1(\Omega) \) with

\[
\overline{\mathcal{MS}}(u,v) = \int_\Omega |\nabla u|^2 \, dx + \mathcal{H}^1(S_u)
\]

\( u \in \text{GSBV}(\Omega) \) and \( v = 1 \) a.e. on \( \Omega \), and \( \overline{\mathcal{MS}}(u,v) = +\infty \) otherwise. In this statement the second integral in (1.1) is omitted as the \( \Gamma \)-convergence is preserved under continuous perturbations. Furthermore, in the field of fracture mechanics the mentioned term is not present at all, so that results presented in this form are also applicable in other contexts beside image segmentation.

In the presented approximation \( v \) works as a phase field variable describing the jump set of \( u \). Following the intuition it seems reasonable that for small \( \varepsilon > 0 \) the function \( v \) is close to 0 where \( u \) is steep or jumps and that it is close to 1 where \( u \) is flat. In practice the weights of the different integral terms declare what is meant to be "steep" and "flat".

In this way the jump set is smoothed out by a \( H^1 \)-function. The presented variational approximations has turned out to be very useful for numerical computations. One usually exploits the appearing quadratic structure by using an alternating minimization, where one variable is fixed while minimizing the other and doing this repeatedly. We will also follow this approach in our numerical examples (see Section 4) where we also explain the method in more details.

From the work presented in [7] one can get a generalization by replacing \( \mathcal{AT}_\varepsilon \) above with

\[
(1.3) \quad \int_\Omega (v^2 + \eta_\varepsilon)|\nabla u|^2 \, dx + \frac{1}{2p\varepsilon} \int_\Omega (1-v)^p \, dx + \varepsilon^{p-1} \int_\Omega |\nabla v|^p \, dx
\]

for \( u \in H^1(\Omega) \) and \( v \in W^{1,p}(\Omega) \) with \( p > 1 \) and \( p' \) being the Hölder conjugate of \( p \).

In this paper we present a similar approximation of the Mumford-Shah functional, but this time allowing the phase field variable \( v \) to be in \( \text{BV}(\Omega) \), the set of functions of bounded variation. This represents in some sense the case in (1.3) with \( p = 1 \).

In order to formulate the main result we first state some necessary assumptions which are quite technical.

**Assumption 1.1.** Let \( \varepsilon_0 > 0 \). For each \( 0 < \varepsilon < \varepsilon_0 \) let

- \( W_\varepsilon: [0,1] \to [0,\infty) \) be continuous such that \( W_\varepsilon \to W \) in \( L^1([0,1]) \) as \( \varepsilon \to 0 \) for some \( W \in L^1([0,1]) \), with \( 1 \in \text{ess.sup} W \), and \( W_\varepsilon \leq \int_0^1 W(s) \, ds \) for a.a. \( t \in [0,1] \),
- \( \varphi_\varepsilon: W_\varepsilon([0,1]) \to \mathbb{R} \) be a convex function such that \( \varphi_\varepsilon(W_\varepsilon(1)) \to 0 \) and \( \varphi_\varepsilon(W_\varepsilon(0)) \to +\infty \) uniformly on \( [0,T] \) for all \( 0 < T < 1 \), i.e. for all \( C > 0 \) there exists \( 0 < \tilde{\varepsilon} < \varepsilon_0 \) such that \( \varphi_\varepsilon(W_\varepsilon(t)) > C \) for all \( t \in [0,T] \) and \( \varepsilon < \tilde{\varepsilon} \).
Corollary 1.3. Let \( \varepsilon > 0 \) be an increasing function such that \( \psi_\varepsilon(0) \to 0 \) as \( \varepsilon \to 0 \) and \( \varphi_\varepsilon^* \leq \psi_\varepsilon \) on \([0, \infty)\), where \( \varphi_\varepsilon^* \) denotes the convex conjugate of \( \varphi_\varepsilon \) (see Section 2.4).

[A4] \( \eta_\varepsilon > 0 \) such that \( \eta_\varepsilon \varphi_\varepsilon(W_\varepsilon(0)) \to 0 \) as \( \varepsilon \to \infty \).

Furthermore, assume that

[A5] \( f: [0, 1] \to [0, \infty) \) is a continuous, non-decreasing function with \( f(0) = 0 \) and \( f > 0 \) on \((0, 1]\).

We are now ready to state our main result.

**Theorem 1.2.** Let \( \Omega \subset \mathbb{R}^n \) be an open bounded set with Lipschitz boundary. For each \( \varepsilon > 0 \) let \( \eta_\varepsilon > 0 \) such that \( \frac{\eta_\varepsilon}{\varepsilon} \to 0 \) as \( \varepsilon \to 0 \), and define the functional \( F_\varepsilon: L^1(\Omega) \times L^1(\Omega) \to \mathbb{R} \) by

\[
F_\varepsilon(u, v) := \int_\Omega (f(v) + \eta_\varepsilon)|\nabla u|^2 + \varphi_\varepsilon(W_\varepsilon(v)) + \psi_\varepsilon(|\nabla v|) \, dx + c_\varepsilon(|D^1 v| + |D^2 v|)
\]

for all \( u \in H^1(\Omega), v \in BV(\Omega; [0, 1]) \) and \( F_\varepsilon(u, v) := +\infty \) otherwise.

Moreover, define \( F: L^1(\Omega) \times L^1(\Omega) \to \mathbb{R} \) by

\[
F(u, v) := \begin{cases} 
\int_\Omega (f(1)|\nabla u|^2 + 2c_\varepsilon \mathcal{H}^{n-1}(S_u)) & \text{for } u \in GSBV^2(\Omega), v = 1 \text{ a.e.} \\
+\infty & \text{otherwise}
\end{cases}
\]

with \( c_\varepsilon = \int_0^1 W(s) \, ds \). Then there holds \( F = \Gamma\text{-lim}_{\varepsilon \to 0} F_\varepsilon \).

As a function of bounded variation the phase field variable \( v \) may have jumps. This fact is exploit when constructing the recovery sequence in the proof of Proposition 3.9. The basic idea is to set \( v \) to zero in a small neighbourhood of \( S_u \) and to one everywhere else. Therefore, we expect to obtain sharper interfaces compared to the \( H^1 \)-phase field appearing in (1.2), which can be observed in the numerical examples in Section 4.

The following corollary represents a special case of the previous theorem.

**Corollary 1.3.** Let \( \Omega \subset \mathbb{R}^n \) be an open bounded set with Lipschitz boundary. For each \( \varepsilon > 0 \) define the functional

\[
F_\varepsilon(u, v) := \frac{\alpha}{2} \int_\Omega (v^2 + \eta_\varepsilon)|\nabla u|^2 \, dx + \frac{\gamma}{2\varepsilon} \int_\Omega (1 - v) \, dx + \frac{\gamma}{2} |Dv| \Omega
\]

if \( u \in H^1(\Omega), v \in BV(\Omega; [0, 1]) \) and \( F_\varepsilon(u, v) := +\infty \) otherwise. Moreover, define \( F: L^1(\Omega) \times L^1(\Omega) \to \mathbb{R} \) by

\[
F(u, v) := \begin{cases} 
\frac{\alpha}{2} \int_\Omega |\nabla u|^2 \, dx + \gamma \mathcal{H}^{n-1}(S_u) & \text{for } u \in GSBV^2(\Omega), v = 1 \text{ a.e.} \\
+\infty & \text{otherwise}
\end{cases}
\]

with \( c_\varepsilon = \int_0^1 W(s) \, ds \). Then there holds \( F = \Gamma\text{-lim}_{\varepsilon \to 0} F_\varepsilon \).

This is the approximation which is actually used for numerical computations in Section 4. The appearing total variation \(|Dv| \Omega\) can be discretized by a finite difference scheme.
2. Preliminaries and Notation

In this section, we collect the notation and the well-known results which are used in this paper.

With $B_R(x)$ we denote the Euclidean ball with radius $R > 0$ and center $x \in \mathbb{R}^n$. For a set $S \subset \mathbb{R}^n$ the $\rho$-neighborhood of $S$ is referred to as $B_\rho(S)$. We refer to $\mathbb{S}^{n-1}$ as the $n-1$-dimensional sphere in $\mathbb{R}^n$. At some places it will be convenient to use the short notation $a \vee b$ and $a \wedge b$ for $\max\{a,b\}$ and $\min\{a,b\}$, respectively.

2.1. Measure Theory. For any set $\Omega \subset \mathbb{R}^n$ we denote by $\mathcal{L}^n(\Omega)$ the $n$-dimensional Lebesgue measure and by $\mathcal{H}^k(\Omega)$ the $k$-dimensional Hausdorff measure. Instead of $\mathcal{H}^0$ we also use the symbol $\#$ as the counting measure. For a (signed, vector-valued) measure $\mu$ we write $|\mu|$ for its total variation.

2.2. $\Gamma$-convergence. For some sequence of functionals $(F_j)$ and a functional $F$ defined on some metric space $X$ we say that $F_j \Gamma$-converges to $F$ as $j \to \infty$ and write $\Gamma\lim_{j \to \infty} F_j = F$ if there holds the

**lim inf-inequality:** for all $u \in X$ and all sequences $(u_j)$ in $X$ with $u_j \to u$ there holds

$$F(u) \leq \liminf_{j \to \infty} F_j(u_j).$$

**lim sup-inequality:** for all $u \in X$ there exists a sequence $(u_j)$ in $X$ such that

$$u_j \to u$$

and

$$\limsup_{j \to \infty} F_j(u_j) \leq F(u).$$

One often defines

$$\Gamma\liminf_{j \to \infty} F_j(u) := \inf\{\liminf_{j \to \infty} F_j(u_j) : u_j \in X \text{ for all } j > 0, u_j \to u \text{ as } j \to \infty\},$$

$$\Gamma\limsup_{j \to \infty} F_j(u) := \inf\{\limsup_{j \to \infty} F_j(u_j) : u_j \in X \text{ for all } j > 0, u_j \to u \text{ as } j \to \infty\}.$$ 

Then the lim inf-inequality is equivalent to $F \leq \Gamma\liminf_{j \to \infty} F_j$ and the lim sup-inequality is equivalent to $\Gamma\limsup_{j \to \infty} F_j \leq F$. Note that $\Gamma\liminf_{j \to \infty} F_j$ as well as $\Gamma\limsup_{j \to \infty} F_j$ are lower semi-continuous.

If one has a family of functionals $(F_\varepsilon)$ for $\varepsilon \in I \subset \mathbb{R}$ the definition is adapted in the usual way, i.e. $F_\varepsilon \Gamma$-converges to $F$ as $\varepsilon \to a$ (for some $a \in I$) if $F_{\varepsilon_j}$ $\Gamma$-converges to $F$ for all sequences $(\varepsilon_j)$ in $I$ with $\varepsilon_j \to a$.

The most important property of $\Gamma$-convergent sequences is the convergence of minimizers to a minimizer of the limit functional, which is stated in the following proposition.

**Proposition 2.1.** Let $F_\varepsilon : X \to \mathbb{R} \cup \{\infty\}$ be a sequence of functionals $\Gamma$-converging to $F : X \to \mathbb{R} \cup \{\infty\}$, where $X$ is a metric space. Assume that $\inf_X F_\varepsilon = \inf_K F_\varepsilon$ for some compact set $K \subset X$. Then, there holds $\lim_{\varepsilon \to 0} \inf_X F_\varepsilon = \inf_X F$. Furthermore, for any sequence $x_\varepsilon$ in $X$ converging to $x \in X$ with $F_\varepsilon(x_\varepsilon) = \inf_X F_\varepsilon$ we have $F(x) = \inf_X F$.

If $F = \Gamma\lim_{j \to \infty} F_j$ and $u \in X$, a sequence $(u_j)$, for which (2.2) holds, is called a recovery sequence for $u$, and there clearly holds $\lim_{j \to \infty} F_j(u_j) = F(u)$. It is actually the case that a sequence of minimizers is a recovery sequence for the minimizer of the $\Gamma$-limit. For this reason knowing the recovery sequences provides lots of information about the structure of the limit behaviour of the functional sequence.
Furthermore, on an open interval the pointwise variation of $u$ points is defined in (2.4) coincide. Precisely, for $x \in \Omega$, their common value $V$ that $V(u, \Omega) = V(u, \Omega)$ for all $u \in BV(\Omega)$.

We call $\tilde{u}$ a precise representative of $u$ if $\tilde{u}(x) = u^+(x)$ for all $x \in \Omega \setminus S_u$ and $\tilde{u}(x) = \frac{1}{2}(u^+(x) + u^-(x))$ for all jump points $x \in S_u$. For functions of bounded variation on the real line we actually have that every point in $S_u$ is a jump point. Furthermore, on an open interval the pointwise variation of $\tilde{u}$ and the variation as defined in (2.4) coincide. Precisely, for $a < b$ and $u \in BV(a,b)$ there holds

$$V(u, (a,b)) = \sup \left\{ \sum_{i=1}^{N} |\tilde{u}(t_i) - \tilde{u}(t_{i-1})| : N \in \mathbb{N}, a < t_0 < \cdots < t_N < b \right\}. $$

For any $u \in BV(\Omega)$ one can split the measure $Du$ in the following way

$$Du = \nabla u \mathcal{L}^n + (u^+ - u^-) \cdot \nu_u \mathcal{H}^{n-1} \circ S_u + D^a u,$$

where the first term, which we will denote by $D^a u$, is the absolutely continuous part of $Du$ with respect to the Lebesgue measure. Therefore, with $\nabla u$ we denote
its density function. The second term represents the jump part of $u$, also referred to as $D^1u$, and $D^c u$ is the Cantor part.

There also holds a chain rule for the composition of a Lipschitz functions and some function of bounded variation (see [4, Theorem 3.99]). Precisely, for $\Omega$ being bounded and $f: \mathbb{R} \to \mathbb{R}$ being Lipschitz we get that $f \circ u \in BV(\Omega)$ and

$$D(f \circ u) = f'(u)\nabla u\mathcal{L}^n + (f(u^+) - f(u^-))\nu_u\mathcal{H}^{n-1}S_u + f'(u)D^c u.$$  

Note that $f'$ exists almost everywhere, which follows from Rademacher’s theorem.

The set of special functions of bounded variation, denoted by $SBV(\Omega)$, contains those functions of bounded variation whose cantor part is zero, i.e. we have $SBV(\Omega) = \{u \in BV(\Omega) : D^c u = 0\}$. The singular part of such functions is therefore only concentrated on the jump set.

A measurable function $u: \Omega \to \mathbb{R}$ is a generalized special function of bounded variation, where we write $u \in SBV(\Omega)$, if any truncation of $u$ is locally a special function of bounded variation, i.e. $u^M \in SBV_{loc}(\Omega)$ for all $M > 0$, with $u^M = (-M) \vee u \wedge M$. Note that for $u \in SBV(\Omega)$ we cannot define $\nabla u$ as above, because the distributional derivative does not need to be a measure on that space. However, $\nabla u^M$ is well defined for all $M > 0$ and converges pointwise a.e. for $M \to \infty$. Thus, we simply define $\nabla u(x) = \lim_{M \to \infty} \nabla u^M(x)$ for a.e. $x \in \Omega$. Furthermore, one can show that $S_u = \bigcup_{M>0} S_{u^M}$. These results and more details can be found in [4, Section 4.5] and the references therein.

Moreover, we will use the following two subspaces of $GSBV(\Omega)$ and $SBV(\Omega)$ defined for every $p > 0$ by

$$SBV^p(\Omega) = \{u \in SBV(\Omega) : \nabla u \in L^p(\Omega), \mathcal{H}^{n-1}(S_u) < \infty\}$$

$$GSBV^p(\Omega) = \{u \in GSBV(\Omega) : \nabla u \in L^p(\Omega), \mathcal{H}^{n-1}(S_u) < \infty\}.$$

A density result, which plays an important role in the proof of the limsup-inequality for our main assertion, is stated in the next theorem. It follows directly from [10, Theorem 3.1] and the following remarks therein.

**Theorem 2.2.** Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary, and take $u \in SBV^2(\Omega) \cap L^\infty(\Omega)$. Then, there exists a sequence $(w_j)$ in $SBV^2(\Omega) \cap L^\infty(\Omega)$ such that

1. $\overline{S_{w_j}}$ is a polyhedral set,
2. $\mathcal{H}^{n-1}(\overline{S_{w_j}} \setminus S_{w_j}) = 0$,
3. $w_j \in W^{1,\infty}(\Omega \setminus S_{w_j})$ for all $j \in \mathbb{N}$,
4. $w_j \to u$ in $L^1(\Omega)$ as $j \to \infty$,
5. $\nabla w_j \to \nabla u$ in $L^2(\Omega)$ as $j \to \infty$,
6. $\mathcal{H}^{n-1}(S_{w_j}) \to \mathcal{H}^{n-1}(S_u)$ as $j \to \infty$.

We now shortly introduce the concept of slicing, which is essential for the proof of the liminf-inequality. For that let $\Omega \subset \mathbb{R}^n$ be open and bounded and let $\xi \in \mathbb{S}^{n-1}$ be a unique normal vector. Then, we write $\Omega_\xi$ for the projection of $\Omega$ onto $\xi^\perp$, and we set

$${\Omega_\xi}_y := \{t \in \mathbb{R} : y + t\xi \subset \Omega\} \quad \text{for all } y \in \Omega_\xi.$$  

Furthermore, for any function $u \in L^1(\Omega)$ and for $\mathcal{L}^{n-1}$-a.a. $y \in \Omega_\xi$ we can define $u^\xi_y(t) := u(y + t\xi)$ for a.a. $t \in {\Omega_\xi}_y$.

One can show the following important results showing the connection between a function $u \in SBV(\Omega)$ and its sliced functions $u^\xi_y$. There are more general results

$$\text{SBV}^p(\Omega) = \{u \in SBV(\Omega) : \nabla u \in L^p(\Omega), \mathcal{H}^{n-1}(S_u) < \infty\}$$

$$\text{GSBV}^p(\Omega) = \{u \in GSBV(\Omega) : \nabla u \in L^p(\Omega), \mathcal{H}^{n-1}(S_u) < \infty\}.$$
for BV-functions, which are not needed in this context. The interested reader can find the details in [4, Section 3.11].

**Theorem 2.3.** Let \( u \in L^1(\Omega) \). Then \( u \in \text{SBV}(\Omega) \) if and only if for all \( \xi \in \mathbb{S}^{n-1} \) there holds \( u^\xi_t \in \text{SBV}(\Omega^\xi_t) \) for \( \mathcal{L}^{n-1} \)-a.a. \( y \in \Omega^\xi_t \) and

\[
\int_{\Omega^\xi_t} |Du^\xi_t|(\Omega^\xi_t) d\mathcal{L}^{n-1}(y) < \infty.
\]

Furthermore, if \( u \in \text{BV}(\Omega) \) there holds for all \( \xi \in \mathbb{S}^{n-1} \), for \( \mathcal{L}^{n-1} \)-a.a. \( y \in \Omega^\xi_t \) and for \( \text{a.a.} \ t \in \Omega^\xi_t \)

1. \( (u^\xi_t)'(t) = \langle \nabla u(y + t\xi), \xi \rangle \),
2. \( S_{u^\xi_t} = (S_u)^\xi_t \),
3. \( (u^\xi_t)^\pm(t) = u^\pm(y + t\xi) \),
4. \( |D^*u^\xi_t, \xi|(|\Omega^\xi_t|) = \int_{\Omega^\xi_t} |D^*u^\xi_t|(|\Omega^\xi_t|) d\mathcal{L}^{n-1}(y) \) for \( * = a, j, c \).

The following Corollary directly follows by a truncation argument.

**Corollary 2.4.** Let \( u \in L^1(\Omega) \). Then \( u \in \text{GSBV}(\Omega) \) if and only if for all \( \xi \in \mathbb{S}^{n-1} \) there holds \( u^\xi_t \in \text{SBV}(\Omega^\xi_t) \) for \( \mathcal{L}^{n-1} \)-a.e. \( y \in \Omega^\xi_t \) and

\[
\int_{\Omega^\xi_t} |D((-M) \lor u^\xi_t \land M)|(|\Omega^\xi_t|) d\mathcal{L}^{n-1}(y) < \infty \quad \text{for all } M > 0.
\]

### 2.4. Convex Functions

Especially, for the numerical part of this paper we also need some theory about convex functions. A good reference for this topic is [18] and [14]. In this context it is sufficient to consider functions defined on the real line. All the discussed issues can easily be adapted to a multi dimensional setting.

Let therefore \( I \subset \mathbb{R} \). The characteristic function over \( I \) is given by \( \chi_I = 0 \) on \( I \) and \( \chi_I = +\infty \) on \( \mathbb{R} \setminus I \). For any function \( f : I \to \mathbb{R} \), bounded from below by some affine function, \( f^* : \mathbb{R} \to \mathbb{R} \) denotes its convex conjugate, i.e.

\[
f^*(s) = \sup_{t \in \mathbb{R}} (ts - f(t)) \quad \text{for all } s \in \mathbb{R}
\]

where \( f \) is set to \(+\infty\) outside of \( I \). This definition directly yields Fenchel's inequality, which says

\[
ts \leq f(t) + f^*(s) \quad \text{for all } t, s \in \mathbb{R}.
\]

We remark that \( f^* \) is always convex and lower semi-continuous and the biconjugate \( f^{**} = (f^*)^* \) is the lower semi-continuous convex hull of \( f \). Furthermore, \( f \) is convex and lower semi-continuous if and only if \( f = f^{**} \).

### 3. Proof of Theorem 1.2

The proof of Theorem 1.2 follows the usual strategy that has been used for the classical Ambrosio-Tortorelli approximation and various generalization (see [5–7, 12, 19, 20]). Firstly, we show the lim inf-inequality on the real line (see Proposition 3.1). The generalization to the multi-dimensional case, stated in Proposition 3.2, is then shown by a slicing argument.

The lim sup-inequality is shown by the usual density result in SBV (see Proposition 2.2). Here, we exploit the fact that the phase field variable is allowed to have jumps, which enables the construction of a much simpler recovery sequence than in the traditional setting.
Proposition 3.1. In the setting of Theorem 1.2 with $\Omega \subset \mathbb{R}$ we redefine $F : L^1(\Omega) \times L^1(\Omega) \to \mathbb{R}$ by

$$F(u, v) := \begin{cases} \int_{\Omega} f(1) |u'|^2 \, dx + 2cW \# S_u & \text{for } u \in SBV^2(\Omega), v = 1 \text{ a.e.} \\ +\infty & \text{otherwise} \end{cases}$$

Then there holds $F \leq \Gamma\text{-lim inf}_{\varepsilon \to 0} F_\varepsilon$.

Proof. First of all, for each open set $I \subset \Omega$ we define the localized functionals

$$F_\varepsilon(u, v; I) := \int_I \left( f(v) + \eta \right) |u'|^2 + \varphi_\varepsilon(W_\varepsilon(v)) + \psi_\varepsilon(|v'|) \, dx$$

$$+ cW (|D^1v(I)| + |D^2v(I)|)$$

for all $u \in H^1(I)$ and $v \in BV(I; [0, 1])$, and $F_\varepsilon(u, v; I) := +\infty$ otherwise.

Now, let $(\varepsilon_j)$ be a sequence greater than zero with $\varepsilon_j \to 0$ as $j \to \infty$ and let $(u_j)$ and $(v_j)$ be sequences in $L^1(\Omega)$ such that $u_j \to u$ and $v_j \to v$ as $j \to \infty$. We can assume (up to a subsequence) that

$$\liminf_{j \to \infty} F_{\varepsilon_j}(u_j, v_j) = \lim_{j \to \infty} F_{\varepsilon_j}(u_j, v_j) < \infty.$$ 

Therefore, we must have $\int_I \varphi_\varepsilon(W_\varepsilon(v_j)) \, dx < \infty$ and due to the uniform convergence of $\varphi_\varepsilon(W_\varepsilon(\cdot))$ to $+\infty$ as $\varepsilon \to 0$ (see [A1]) we can assume that $v = 1$ a.e. on $\Omega$.

We first show that $\# S_u$ is finite and

$$(3.1) \quad 2cW \# S_u \leq \liminf_{j \to \infty} F_{\varepsilon_j}(u_j, v_j; B_\delta(S_u))$$

for all $\delta > 0$ sufficiently small.

For that let $y_0 \in S_u$, and let $\delta > 0$ sufficiently small such that $B_\delta(y_0) \subset \Omega$. Set $M := \liminf_{j \to \infty} \text{ess inf}_{B_{\delta/2}(y_0)} (f \circ v_j)$ and assume that $M > 0$. Furthermore, let $0 < \eta < M$ and choose $j_0 > 0$ such that up to subsequence there holds $M < \text{ess inf}_{B_{\delta/2}(y_0)} (f \circ v_j) + \eta$ for all $j > j_0$. Then there holds

$$\int_{y_0 - \frac{\delta}{2}}^{y_0 + \frac{\delta}{2}} |u'_j|^2 \, dx \leq \frac{1}{M - \eta} \int_{y_0 - \frac{\delta}{2}}^{y_0 + \frac{\delta}{2}} f(v_j)|u'_j|^2 \, dx \leq \frac{C}{M - \eta}$$

for all $j > j_0$ so that $u'_j$ converges weakly to $u'$ in $L^2(B_{\delta/2}(y_0))$ and consequently $y_0 \notin S_u$. Hence, we must have $M = 0$, and we can find a sequence $(y_j)$ such that $f(\tilde{v}_j(y_j)) \to 0$, $\tilde{v}_j$ being a precise representative of $v_j$ (see Section 2.3). The assumptions on $f$ in [A5] imply $\tilde{v}_j(y_j) \to 0$ as $j \to \infty$. Since $\tilde{v}_j \to 1$ a.e. we can, therefore, find $y^-, y^+ \in B_\delta(y_0)$ with $y^- < y_0 < y^+$ such that $\tilde{v}_j(y^-) \to 1$ as well as $\tilde{v}_j(y^+) \to 1$.

With this at hand we have, due to the $L^1$-convergence of $W_\varepsilon$ from [A1],

$$(3.2) \quad 2cW = \lim_{j \to \infty} \left[ \int_{\tilde{v}_j(y_j)}^{\tilde{v}_j(y_+)} W_{\varepsilon_j}(s) \, ds + \int_{\tilde{v}_j(y_j)}^{\tilde{v}_j(y_-)} W_{\varepsilon_j}(s) \, ds \right].$$

Defining

$$(3.3) \quad \Phi_\varepsilon(t) := \int_0^t W_\varepsilon(s) \, ds \quad \text{for all } t \in [0, 1], \varepsilon > 0$$

we get

$$\int_{\tilde{v}_j(y_j)}^{\tilde{v}_j(y_+)} W_{\varepsilon_j}(s) \, ds + \int_{\tilde{v}_j(y_j)}^{\tilde{v}_j(y_-)} W_{\varepsilon_j}(s) \, ds$$

and
and together with (2.5)

\[
\int_{\tilde{v}_j(y)}^{\tilde{v}_j(y^+)} W_{\varepsilon_j}(s) \, ds + \int_{\tilde{v}_j(y)}^{\tilde{v}_j(y^-)} W_{\varepsilon_j}(s) \, ds \leq |D(\Phi_{\varepsilon_j} \circ v_j)|(B_\delta(y_0)) + |D(\Phi_{\varepsilon_j} \circ v_j)|.(B_\delta(y_0)).
\]

Applying the chain rule (see (2.6)) and Fenchel’s inequality (see (2.7)) yields

\[
|D(\Phi_{\varepsilon_j} \circ v_j)|(B_\delta(y_0)) = \int_{y_0}^{y_0+\delta} W_{\varepsilon_j}(v_j) |v_j'| \, dx + \int_{J_{i\varepsilon}} \Phi_{\varepsilon_j}(v_j^+) - \Phi_{\varepsilon_j}(v_j^-) \, dH^0 + \int_{B_\delta(y_0)} \Phi'(\tilde{v}_j) \, d|D^c v_j| \leq \int_{y_0}^{y_0+\delta} \varphi_{\varepsilon_j}(W_{\varepsilon_j}(v_j)) + \varphi_{\varepsilon_j}^*(|v_j'|) \, dx + \int_{J_{i\varepsilon}} \int_{v_j^-}^{v_j^+} W_{\varepsilon_j}(s) \, ds \, dH^0 + \int_{B_\delta(y_0)} W_{\varepsilon_j}(\tilde{v}_j) \, d|D^c v_j| \leq F_{\varepsilon_j}(u_j, v_j; B_\delta(y_0)).
\]

In the last inequality we used $\varphi_{\varepsilon_j} \leq \psi_{\varepsilon_j}$ on $[0, \infty)$ from [A3] and $W_{\varepsilon_j} \leq cW$ from [A1]. By merging (3.2), (3.4) and (3.5) we deduce

\[
2cW \leq \liminf_{j \to \infty} F_{\varepsilon_j}(u_j, v_j; B_\delta(y_0)).
\]

For each $N \leq \#S_u$ we can repeat the preceding arguments for each element in a set $\{y_1, \ldots, y_N\} \subset S_u$ with $\delta > 0$ sufficiently small such that $B_\delta(y_k) \cap B_\delta(y_\ell) = \emptyset$ for $k \neq \ell$ in order to obtain

\[
2cW \cdot N \leq \sum_{k=1}^{N} \liminf_{j \to \infty} F_{\varepsilon_j}(u_j, v_j; B_\delta(y_k)) \leq \liminf_{j \to \infty} F_{\varepsilon_j}\left(u_j, v_j; \bigcup_{k=1}^{N} B_\delta(y_k)\right).
\]

By assumption the right hand side is finite independent of $\delta$; hence, there must hold $\#S_u < \infty$ and we deduce (3.1).

In the next step we show that for all $\delta > 0$

\[
\int_{\Omega \setminus B_\delta(S_u)} f(1)|u'|^2 \, dx \leq \liminf_{j \to \infty} F_{\varepsilon_j}(u_j, v_j; \Omega \setminus B_\delta(S_u)).
\]

Let $I := (a, b) \subset \Omega$ be an open interval such that $I \cap S_u = \emptyset$. For $k \in \mathbb{N}$ and $\ell \in \{1, \ldots, k\}$ we define the intervals

\[
I^k_\ell := \left(a + \frac{\ell - 1}{k}(b - a), a + \frac{\ell}{k}(b - a)\right)
\]

and we extract a subsequence of $(v_j)$ (not relabeled) such that $\lim_{j \to \infty} \text{ess inf}_{I^k_\ell} v_j$ exists for all $k$. Moreover, for $0 < z < 1$ we define the set

\[
T^k_z := \{ \ell \in \{1, \ldots, k\}: \lim_{j \to \infty} \text{ess inf}_{I^k_\ell} v_j \leq z\}.
\]

For any $\ell \in T^k_z$ there exists a sequence $(x_j)$ in $I^k_\ell$ and $y \in I^k_\ell$ such that

\[
\lim_{j \to \infty} \tilde{v}_j(x_j) = \lim_{j \to \infty} \text{ess inf}_{I^k_\ell} v_j \quad \text{and} \quad v_j(y) \to 1.
\]
With this at hand we can estimate precisely as in (3.4) and (3.5)
\[\int_z^1 W(s) \, ds \leq \lim_{j \to \infty} \int_{\varepsilon_j(x_j)}^{\varepsilon_j(y)} W_{\varepsilon_j}(s) \, ds \leq \liminf_{j \to \infty} F_{\varepsilon_j}(u_j, v_j; T^k_\ell) \leq C\]
for some $C > 0$ by assumption.

Repeating this argument for every $\ell \in T^k_\ell$ we get
\[\#(T^k_\ell) \int_z^1 W(s) \, ds \leq \liminf_{j \to \infty} F_{\varepsilon_j}(u_j, v_j; I) \leq C.\]

Note that, since $1 \in \operatorname{ess} \sup W$ from $[A1]$, there holds $\int_z^1 W(s) \, ds > 0$ for all $0 < z < 1$ and hence, $\#T^k_\ell$ is bounded independently of $k$. Because $\#T^k_\ell$ is non-decreasing with respect to $k$, for $k$ large enough we can pick $\ell^k_1 < \ell^k_2 < \cdots < \ell^k_N \in T^k_\ell$ with $N = \max_{k \in \mathbb{N}} \#(T^k_\ell)$, such that $\frac{\ell^k_i}{k}$ converges to some $y_i \in I$ as $k \to \infty$. Define $T_z = \{y_1, \ldots, y_N\}$; let $\delta > 0$, and choose $k > \frac{b-a}{2\delta}$ and $\ell \in T^k_\ell$. Then we have $T^k_\ell \subset B_\delta(T_z)$. Therefore,
\[
\liminf_{j \to \infty} f(z) \int_{I \setminus B_\delta(T_z)} |u_j'|^2 \, dx \leq \liminf_{j \to \infty} \int_{I} f(v_j) |u_j'|^2 \, dx
\leq \liminf_{j \to \infty} F_{\varepsilon_j}(u_j, v_j; I).
\]

Since $\delta > 0$ was chosen arbitrarily it is enough to integrate over $I \setminus T_z$ on the left hand side. Moreover, from $[A5]$ we have $f(z) > 1$, and thus, we obtain $u_j' \rightharpoonup u'$ in $L^2(I \setminus T_z)$ up to subsequence, and consequently $u \in H^1(I \setminus T_z)$. Since $I \cap S_u = \emptyset$ there even holds $u \in H^1(I)$. Letting $z \to 1$ and using the weakly lower semi-continuity of the norm as well as the continuity of $f$ from $[A5]$ we get
\[
\int_{I} f(1) |u'|^2 \, dx \leq \liminf_{j \to \infty} F_{\varepsilon_j}(u_j, v_j; I).
\]

Since $I \subset \Omega$ was chosen arbitrarily such that $I \cap S_u = \emptyset$ we conclude (3.6). Together with (3.1) we eventually obtain $F(u, v) \leq \liminf_{j \to \infty} F_{\varepsilon_j}(u_j, v_j)$. \hfill \Box

**Proposition 3.2.** In the setting of Theorem 1.2 there holds
\[F(u, v) \leq \Gamma^- \liminf_{\varepsilon \to 0} F_{\varepsilon}(u, v) \quad \text{for all } u, v \in L^1(\Omega).\]

**Proof.** For the proof we use the usual notation in the setting of slicing, introduced in Section 2.3. In what follows let $\xi \in S^{n-1}$ and $y \in \Omega_\xi$, and let $A \subset \Omega$ be open. We define the localized version of (1.4) by
\[
F_{\varepsilon}(u, v; A) := \int_A (f(v) + \eta_\varepsilon) |\nabla u|^2 + \varphi_\varepsilon(W_{\varepsilon}(v) + \psi_\varepsilon(|\nabla v|) \, dx
+ c_W(|D^1 v|(A) + |D^\varepsilon v|(A))
\]
if $u \in H^1(A), v \in \operatorname{BV}(A; [0, 1])$ and $F_{\varepsilon}(u, v; A) := +\infty$ otherwise. Furthermore, we define for $I \subset \mathbb{R}$ open
\[
F^W_{\varepsilon}(u, v; I) := \int_I (f(v) + \eta_\varepsilon) |u'|^2 + \varphi_\varepsilon(W_{\varepsilon}(v) + \psi_\varepsilon(|u'|) \, dx
+ c_W(|D^1 v|(I) + |D^\varepsilon v|(I))
\]
if \( u \in H^1(I), v \in BV(I; [0, 1]) \) and \( F^\xi_y(u, v; I) := +\infty \) otherwise. Now, we set for all \( u, v \in L^1(\Omega) \) and \( A \subset \Omega \)
\[
F^\xi_x(u, v; A) := \int_{A_1} F^\xi_y(u^\xi_y, v^\xi_y; A^\xi_y) d\mathcal{L}^{n-1}(y).
\]
Thus, we have by Fubini’s theorem and Theorem 2.3
\[
F^\xi_x(u, v; A) = \int_A \left( f(v) + \eta_r \right) |(\nabla u, \xi)|^2 + \varphi_r(W^e(v)) + \psi_r \left( |(\nabla v, \xi)| \right) dx
+ c_W |(D^j v, \xi)(A)| + c_W |(D^v, \xi)(A)|
\]
if \( |(Du, \xi)| \) is absolutely continuous with respect to \( \mathcal{L}^n \), and \( F^\xi_x = +\infty \) otherwise. There clearly holds
\[
(F^\xi_x(u, v; A) \leq F^\xi_x(u, v; I).
\]
From Proposition 3.1 we know that \( F^\xi_y(u, v; I) \leq \Gamma\text{-lim\,inf}_{\varepsilon \to 0} F^\xi_y(u, v; I) \) with
\[
F^\xi_y(u, v; I) := \left\{ \begin{array}{ll}
\int_I |f(1)| u^2 \, dx + 2C_W \# S_u & \text{for } u \in SBV^2(I), \nu = 1 \text{ a.e.}, \\
+\infty & \text{otherwise.}
\end{array} \right.
\]
Choosing
\[
F^\xi_x(u, v; A) := \int_{A_1} F^\xi_y(u^\xi_y, v^\xi_y; A^\xi_y) d\mathcal{L}^{n-1}(y),
\]
there holds for all sequences \( (u_j) \) and \( (v_j) \) with \( u_j \to u \) and \( v_j \to v \) in \( L^1(\Omega) \) as \( j \to \infty \)
\[
F^\xi_x(u, v; A) \leq \liminf_{j \to \infty} F^\xi_y(u_j^\xi_y, v_j^\xi_y; A_j^\xi_y) d\mathcal{L}^{n-1}(y).
\]
Fatou’s Lemma and (3.7), then yield
\[
F^\xi_x(u, v; A) \leq \Gamma\text{-lim\,inf}_{\varepsilon \to 0} F^\xi_x(u, v; A) \leq \Gamma\text{-lim\,inf}_{\varepsilon \to 0} F^\xi_x(u, v; A).
\]
Moreover, we get that \( F^\xi_x(u, v; A) \) is finite if and only if for all \( \xi \in \mathbb{S}^{n-1} \) and for a.a. \( y \in A_1 \) there holds \( v = 1 \) a.e. on \( A^\xi_y \) and \( u \in SBV^2(A^\xi_y) \) as well as
\[
\int_{A_1} \int_{A^\xi_y} f(1)|\nabla u^\xi_y|^2 \, dx + 2C_W \# S_u^\xi d\mathcal{L}^{n-1}(y) < \infty.
\]
Since there holds for every \( M > 0 \) and every \( u \in L^1(\Omega) \) with \( u^\xi_y \in SBV^2(A^\xi_y) \) for a.a. \( y \in A_1 \)
\[
\int_{A_1} |D((-M) \vee u^\xi_y \wedge M)|(A^\xi_y) d\mathcal{L}^{n-1}(y)
\leq \int_{A_1} C ||\nabla ((-M) \vee u^\xi_y \wedge M)||_{L^2(A^\xi_y)} + 2M \# S_u^\xi d\mathcal{L}^{n-1}(y)
\leq C \int_{A_1} \int_{A^\xi_y} f(1)|\nabla((-M) \vee u^\xi_y \wedge M)|^2 \, dx + 2C_W \# S_u^\xi d\mathcal{L}^{n-1}(y),
\]
we get by Corollary 2.4 that \( F^\xi_x(u, v; A) \) is finite only if \( u \in GSBV^2(A) \). Hence,
\[
F^\xi_x(u, v; A) = \int_A f(1)|\nabla(u, \xi)|^2 \, dx + 2C_W \int_{S_u} |(u, \xi)| d\mathcal{H}^{n-1}
\]
if \( u \in GSBV^2(A) \) and \( F^\xi_x(u, v; A) = +\infty \) otherwise.
Since \( A \) and \( \xi \) where chosen arbitrarily, [7, Theorem 1.16] and (3.8) imply
\[
F(u, v; A) = \int_A f(1) \sup_{\xi \in \mathbb{S}^{n-1}} |\langle \nabla u, \xi \rangle|^2 \, d\mathcal{L}^n + 2c_W \int_{S_h} \sup_{\xi \in \mathbb{S}^{n-1}} |\langle \nu, \xi \rangle| \mathcal{H}^{n-1}
\]
\[
\leq \Gamma\text{-lim inf}_{\varepsilon \to 0} F_{\varepsilon}(u, v; A) .
\]

The following proposition now shows the \( \text{lim sup} \)-inequality.

**Proposition 3.3.** In the setting of Theorem 1.2 there holds
\[
\Gamma\text{-lim sup}_{\varepsilon \to 0} F_{\varepsilon}(u, v) \leq F(u, v) \quad \text{for all } u, v \in L^1(\Omega) .
\]

**Proof.** If \( u \notin \text{GSBV}^2(\Omega) \) or \( v \neq 1 \) on some set with measure greater than zero the assertion is obvious. We first show that the result holds for \( w \in \text{SBV}^2(\Omega) \cap L^\infty(\Omega) \) for which (1)–(3) (replacing \( w \) by \( w \)) in Theorem 2.2 hold.

For this purpose choose for every \( \varepsilon > 0 \) some \( \delta_\varepsilon > 0 \) such that \( \frac{\eta}{\delta_\varepsilon} \to 0 \) as \( \varepsilon \to 0 \) but still \( \delta_\varepsilon \phi_\varepsilon(W_\varepsilon(0)) \to 0 \) as \( \varepsilon \to 0 \), for instance
\[
\delta_\varepsilon = \frac{\sqrt{\eta\varepsilon}}{\sqrt{\phi_\varepsilon(W_\varepsilon(0))}} .
\]

Take some cutoff function \( \phi : \mathbb{R} \to \mathbb{R} \) with \( \phi = 1 \) on \( B_1(0) \) and \( \phi = 0 \) on \( \Omega \setminus B_1(0) \), and define \( \tau(x) = d(x, S_w) \) for all \( x \in \Omega \). Then, we set \( \phi_\varepsilon(x) = \phi(\tau(x)/\delta_\varepsilon) \) for all \( x \in \Omega \), and we fix for every \( \varepsilon > 0 \) the function \( w_\varepsilon = (1 - \phi_\varepsilon)w \), for which holds \( w_\varepsilon \in H^3(\Omega) \), \( w_\varepsilon = w \) on \( \Omega \setminus B_{\delta_\varepsilon}(S_w) \) and \( w_\varepsilon \to w \) in \( L^1(\Omega) \) as \( \varepsilon \to 0 \). Furthermore we define
\[
v_\varepsilon = \begin{cases} 0 & \text{on } B_{\delta_\varepsilon}(S_w) \cap \Omega , \\ 1 & \text{elsewhere} . \end{cases}
\]

Since \( S_w \) is polyhedral there holds \( \mathcal{H}^{n-1}(\partial B_{\delta_\varepsilon}(S_w) \cap \Omega) < \infty \). Consequently, we have \( v_\varepsilon \in \text{BV}(\Omega; [0, 1]) \) for all \( \varepsilon > 0 \).

With this at hand, recalling [A5], we get
\[
(3.9) \quad F_{\varepsilon}(w_\varepsilon, v_\varepsilon) \leq \int_\Omega f(1)|\nabla w|^2 \, dx + \eta \varepsilon \int_\Omega |\nabla w_\varepsilon|^2 \, dx \\
+ \mathcal{L}^n(\Omega)(\phi_\varepsilon(W_\varepsilon(1)) + \psi_\varepsilon(0)) \\
+ \mathcal{L}^n(B_{\delta_\varepsilon}(S_w))\phi_\varepsilon(W_\varepsilon(0)) + \mathcal{H}^{n-1}(\partial B_{\delta_\varepsilon}(S_w))c_W .
\]

By the choice of \( w_\varepsilon \), the fact that \( \|w\|_{L^\infty(\Omega)} \leq M \) and that \( |\nabla \tau(x)| = 1 \) for a.e. on \( \Omega \) (see [16, Lemma 3.2.34]) we get on \( B_{\delta_\varepsilon}(S_w) \)
\[
|\nabla w_\varepsilon| \leq |w \nabla \phi_\varepsilon| + |(1 - \phi_\varepsilon)\nabla w| \leq \frac{M}{\delta_\varepsilon} \|\varphi\|_{L^\infty(\Omega)} + |\nabla w| ,
\]
which implies
\[
\eta \varepsilon \int_\Omega |\nabla w_\varepsilon|^2 \, dx \leq \eta \varepsilon \int_{\Omega \setminus B_{\delta_\varepsilon}(S_w)} |\nabla w|^2 \, dx \\
+ C\frac{\eta \varepsilon}{\delta_\varepsilon} \mathcal{L}^n(B_{\delta_\varepsilon}(S_w)) + 2\eta \varepsilon \int_{B_{\delta_\varepsilon}(S_w)} |\nabla w|^2 \, dx .
\]
with \( C = 2M^2\|\nabla \phi\|_{L^\infty(\Omega)}^2 \) independent of \( \varepsilon \). The first and the last term obviously converge to 0 as \( \varepsilon \to 0 \). For the second term we remark that for a polyhedral
set, the Hausdorff measure coincides with the Minkowski content (see, e.g., [16, Theorem 3.2.29]), so that

$$L^n(B_{\delta \epsilon}(S_w)) \to H^{n-1}(S_w) = H^{n-1}(S_w) < \infty \text{ as } \epsilon \to 0.$$  

As a consequence, recalling that \( \frac{\delta}{\epsilon} \to 0 \) we get

$$C_{\frac{\delta}{\epsilon}} L^n(B_{\delta \epsilon}(S_w)) \to 0 \text{ as } \epsilon \to 0.$$  

and therefore

$$\eta \int_\Omega |\nabla w_\epsilon|^2 dx \to 0 \text{ as } \epsilon \to 0.$$  

Additionally, (3.10) and \( \delta_\epsilon \varphi_\epsilon(W_\epsilon(0)) \to 0 \text{ as } \epsilon \to 0 \) implies

$$L^n(B_{\delta \epsilon}(S_w)) \varphi_\epsilon(W_\epsilon(0)) \to 0 \text{ as } \epsilon \to 0.$$  

Furthermore, there holds

$$H^{n-1}(\partial B_{\delta \epsilon}(S_w)) \to 2H^{n-1}(S_w) \text{ as } \epsilon \to 0,$$

which is again due to \( S_w \) being a polyhedral set.

Applying the previous three convergence statements in (3.9) together with the limit behaviour of \( \varphi_\epsilon(W_\epsilon(1)) \) and \( \psi_\epsilon(0) \) from [A2] and [A3], we get

$$\limsup_{\epsilon \to 0} F_\epsilon(w_\epsilon, v_\epsilon) \leq F(w, 1).$$  

Here 1 represents the function that maps to 1 a.e. on \( \Omega \).

If \( u \in GSBV^2(\Omega) \) we have for every \( M > 0 \) that \( u^M \in \text{SBV}^2(\Omega) \cap L^\infty(\Omega) \) with \( u^M := (-M) \vee u \wedge M \), and we can find a sequence \( (w_j) \) in \( \text{SBV}^2(\Omega) \cap L^\infty(\Omega) \) such that (1)–(6) in Theorem 2.2 (replacing \( u \) by \( u^M \)) holds. Together with the lower semi-continuity of \( \Gamma \)-lim sup \( F_\epsilon \) in \( L^1(\Omega) \times L^1(\Omega) \) and (3.11) we deduce

$$\Gamma \text{-lim sup } F_\epsilon(u^M, 1) \leq \liminf_{j \to \infty} \Gamma \text{-lim sup } F_\epsilon(w_j, 1) \leq \liminf_{j \to \infty} F(w_j, 1) = F(u^M, 1).$$  

Since \( \nabla u \in L^2(\Omega) \) we get by the dominated convergence theorem

$$\lim_{M \to \infty} \int_\Omega |\nabla u^M|^2 dx \to \int_\Omega |\nabla u|^2 dx.$$  

From \( S_u = \bigcup_{M > 0} S_{u, M} \) (see Section 2.3) we imply \( H^{n-1}(S_{u, M}) \leq H^{n-1}(S_u) \). Thus, using again the lower semi-continuity of \( \Gamma \)-lim sup \( F_\epsilon \) we conclude the proof letting \( M \to \infty \).

The proof of Theorem 1.2 is now a direct consequence of Proposition 3.2 and Proposition 3.3. It remains the proof of Corollary 1.3 which is basically an application of the just shown theorem.

**Proof of Corollary 1.3.** We define \( \tilde{F}_\epsilon := \frac{2}{\gamma} F_\epsilon \) and, choose the functions \( f, W_\epsilon, \varphi_\epsilon \) and \( \psi_\epsilon \) in the following way:

$$f(t) = \frac{\alpha}{\gamma} t^2, \quad W_\epsilon(t) = (1 - t) \tilde{\epsilon}, \quad \varphi_\epsilon(t) = \frac{1}{\epsilon} t^{1/\epsilon}, \quad \psi_\epsilon(s) = s$$  

for all \( t \in [0, 1], s \in [0, \infty) \) and \( 0 < \epsilon < 1 \). Note that in this setting we have

$$\varphi_\epsilon^*(s) = \begin{cases} (1 - \epsilon)(\epsilon s)^{1/\epsilon} & \text{for } s \in [0, \epsilon^{-2}], \\ s - \frac{1}{\epsilon} & \text{for } s > \epsilon^{-2}, \end{cases}$$  

where \( \varphi_\epsilon^* \) is the right inverse of \( \varphi_\epsilon \).
and hence, one can simply verify that Assumption 1.1 is fulfilled with $W = 1$, the constant one function.

From Theorem 1.2 we get that $\tilde{F}$ $\Gamma$-converges to

$$
\tilde{F}(u, v) := \begin{cases} 
\frac{\alpha}{\gamma} \int_{\Omega} |\nabla u|^2 \, dx + 2H^{n-1}(S_u) & \text{for } u \in \text{GSBV}^2(\Omega), v = 1 \text{ a.e.} \\
+\infty & \text{otherwise}
\end{cases}
$$

Since $\Gamma$-convergence is preserved under constant multiplication we get the result by multiplying $\tilde{F}$ and $\tilde{F}$ with $\frac{1}{\gamma}$.

**Remark 3.4.** We remark that the

### 4. Numerical Examples

As we have seen in the introduction for $\Omega \subset \mathbb{R}^n$ open, bounded and with Lipschitz boundary the Mumford-Shah functional in its complete form is given by

$$
(4.1) \quad E(u) = \begin{cases} 
\frac{\alpha}{2} \int_{\Omega} |\nabla u|^2 \, dx + \frac{\beta}{2} \int_{\Omega} |u - g|^2 \, dx + \gamma \int_{\Gamma}(S_u) & \text{for } u \in \text{SBV}^2(\Omega), \\
+\infty & \text{otherwise}
\end{cases}
$$

where $g \in L^\infty(\Omega)$ is the given image and $\alpha, \beta, \gamma > 0$ are parameters for weighting the appearing terms differently.

It is clear that we can add the additional term to $F$ and $\tilde{F}$ in Theorem 1.2 and the $\Gamma$-convergence still holds true, because $\Gamma$-convergence is stable under continuous perturbations. Hence, by Corollary 1.3 we approximately minimize $E$ by minimizing $E_\varepsilon : L^1(\Omega) \times L^1(\Omega)$ given by

$$
(4.2) \quad E_\varepsilon(u, v) := \frac{\alpha}{2} \int_{\Omega} (u^2 + \eta_\varepsilon)|\nabla u|^2 \, dx + \frac{\beta}{2} \int_{\Omega} |u - g|^2 \, dx + \gamma \int_{\Gamma}(1 - v) \, dx + \frac{\gamma}{2} |Dv|_1(\Omega)
$$

for $u \in H^1(\Omega), v \in \text{BV}(\Omega; [0, 1])$ and $E_\varepsilon = +\infty$ otherwise.

In order to discretize the functional $E_\varepsilon$ we consider a 2-dimensional image with its natural pixel grid with pixel length $h > 0$. If the picture is given by $M \times N$ pixels we have the discrete space $\Omega_h = [0, Mh] \times [0, Nh]$ and we identify the functions $u, g, v$ as elements in the Euclidean space $\mathbb{R}^{M \times N}$. Precisely, one sets $u = \sum_{i,j} u_{ij} \chi_{((i-1)h, jh) \times [(j-1)h, jh)]}$.

The discretization of the gradient is performed by finite differences. Hence, we set

$$
(\nabla u)_{ij} = ((\nabla u)^{(1)})_{ij}, (\nabla u)^{(2)}_{ij}) \quad \text{for } u \in \mathbb{R}^{M \times N}
$$

with

$$(\nabla_h u)^{(1)}_{ij} := \frac{u_{i+1,j} - u_{ij}}{h} \quad \text{for } i \in \{1, \ldots, M-1\}, j \in \{1, \ldots, N\},
$$

$$(\nabla_h u)^{(2)}_{ij} := 0 \quad \text{for } j \in \{1, \ldots, N\},
$$

$$(\nabla_h u)^{(1)}_{ij} := \frac{u_{i,j+1} - u_{ij}}{h} \quad \text{for } i \in \{1, \ldots, M\}, j \in \{1, \ldots, N-1\},
$$

$$(\nabla_h u)^{(2)}_{ij} := 0 \quad \text{for } i \in \{1, \ldots, M\}.
$$

We also replace $Dv$ in (4.2) by the discretized gradient. This is due to the fact that $\int |\nabla_h v| \, dx$ $\Gamma$-converges to the total variation of $v \in \mathbb{R}^{M \times N}$ as the grid size goes to $0$. Whereas, if one would just plug in $v = \sum_{ij} v_{ij} \chi_{((i-1)h, jh) \times [(j-1)h, jh])}$
one replaces $|Dv|$ by $\|\nabla_h v\|_h$ instead by the Euclidean norm $|\nabla_h v|$. However, the \( \Gamma \)-convergence result does not hold in this case (see REFERENCE).

Summing up we can define the discretized functional

\[ E^d_\varepsilon(u, v) := \sum_{i,j} \left[ \frac{\alpha}{2} (v^2_{ij} + \eta_{ij}) |(\nabla_h u)_{ij}|^2 + \frac{\beta}{2} |u_{ij} - g_{ij}|^2 + \frac{\gamma}{2\varepsilon} (1 - v_{ij}) + \frac{\gamma}{2\varepsilon} |(\nabla_h v)_{ij}| \right] \]

where $\| \cdot \|$ denotes the Euclidean Norm. Note that we neglected the factor $h^2$ in $E^d_\varepsilon$ since it does not change its minimum.

The usual method in order to minimize the functional (4.2) is an alternating scheme in which we fix one of the variables in order to minimize with respect to the other (REFERENCE). This method exploits the fact that the functional is strictly convex and lower semi-continuous in each variable separately. However, the computed solutions are not necessarily minimizers, not even local ones, but rather critical points of $E_\varepsilon$.

Once we fixed some initial value $v^0 \in BV(\Omega; [0, 1])$ we set for each $k \in \mathbb{N}$

\[ u^k = \arg \min \{ E_\varepsilon(u, v^{k-1}) : u \in H^1(\Omega) \}, \]
\[ v^k = \arg \min \{ E_\varepsilon(u^k, v) : v \in BV(\Omega; [0, 1]) \}, \]

or in the discrete setting for some $v^0 \in \mathbb{R}^{M \times N}$

\[ u^k = \arg \min \{ E^d_\varepsilon(u, v^{k-1}) : u \in \mathbb{R}^{M \times N} \}, \]
\[ v^k = \arg \min \{ E^d_\varepsilon(u^k, v) : v \in \mathbb{R}^{M \times N}, 0 \leq v_{ij} \leq 1 \}. \]

The minimization (4.6) can simply be solved by solving the corresponding Euler-Equation. That is

\[ (\alpha (D^{(1)})^T \operatorname{diag}(v^2 + \eta) D^{(1)} + \alpha (D^{(2)})^T \operatorname{diag}(v^2 + \eta) D^{(2)} + \beta \operatorname{Id}) u = \beta g \]

where $D^{(m)}$ is the matrix such that $D^{(m)} u = (\nabla_h u)^{(m)}$ for $m = 1, 2$.

We are going to tackle problem (4.7) by the corresponding primal-dual problem with the algorithm introduced by A. Chambolle and T. Pock in [9]. Therefore, we define for fixed $u \in \mathbb{R}^n$ the functionals

\[ F_u(v) = \frac{\alpha}{2} \sum_{i,j} (v^2_{ij} + \eta_{ij}) |(\nabla_h u)_{ij}|^2 + \frac{\gamma}{2\varepsilon} (1 - v_{ij}) + \chi_{[0,1]}(v_{ij}) \quad \text{and} \quad G(p) = \frac{\gamma}{2} \|p\|_1 \]

for all $v \in \mathbb{R}^{M \times N}, p \in \mathbb{R}^{2 \times M \times N}$, where

\[ \|p\|_1 = \sum_{i,j} |p_{ij}| = \sum_{i,j} \sqrt{|p_{ij}^{(1)}|^2 + |p_{ij}^{(2)}|^2}. \]

The minimizer in (4.7) is then equal to

\[ \arg \min \{ F_u(v) + G(\nabla_h v) : v \in \mathbb{R}^{M \times N} \}. \]

With $G^* = \chi(\| \cdot \|_\infty \leq \frac{\varepsilon}{2})$ the corresponding min-max problem of (4.9) is given by

\[ \min_{u \in \mathbb{R}^{M \times N}} \max_{q \in \mathbb{R}^{2 \times M \times N}} (\nabla_h v, q) + F_u(v) - \chi(\| \cdot \|_\infty \leq \frac{\varepsilon}{2}) (q) \]

where $\|q\|_\infty = \max_{i,j} |q_{ij}| = \max_{i,j} \sqrt{|q_{ij}^{(1)}|^2 + |q_{ij}^{(2)}|^2}$. We solve this problem by [9, Algorithm 1]. Namely, for $0 < \tau^2 < \frac{h^2}{\varepsilon}$ and $v^0 \in \mathbb{R}^{M \times N}$, $q^0 \in \mathbb{R}^{2 \times M \times N}$ as well as
\[ q^{k+1} = (\text{Id} + \tau \partial G^*)^{-1}(q^k + \tau \nabla_h v^k) \]
\[ v^{k+1} = (\text{Id} + \tau \partial F_w)^{-1}(v^k - \tau \nabla_T q^{k+1}) \]
\[ \bar{v}^{k+1} = 2q^{k+1} - v^k. \]

Here, for a convex lower semi-continuous function \( F \) and \( \tau > 0 \) the operator \((\text{Id} + \tau \partial F)^{-1}\) is the resolvent operator, i.e.
\[ (\text{Id} + \tau \partial F)^{-1}(p) := \arg \min_{x} \left\{ \frac{\|x - p\|^2}{2\tau} + F(x) \right\}. \]

Then, [9, Theorem 1] guarantees the convergence of \((v^k, q^k)\) to a solution of (4.10). Considering the resolvent operators, since \( G^* \) is just the indicator functions of a convex set, line (4.11) is the projection of \( q^k + \tau \nabla_h v^k \) onto \( \{\|\cdot\|_{\infty} \leq \frac{\gamma}{2}\} \) (cf. [9, Section 6.2]). Thus we simply get
\[ q^{k+1}_{ij} = \frac{\gamma}{2 \max \{1, |q^k_{ij}|\}} \text{ with } q = \frac{2}{\gamma}(q^k + \tau \nabla_h v^k). \]

The resolvent operator appearing in (4.12) requires the minimization of
\[ \tilde{F}_{u,w}(v) := \frac{|v - w|^2}{2\tau} + F_u(w) \text{ for some } w \in \mathbb{R}^{M \times N} \]
with respect to \( v \). The so-defined functional is clearly convex and coercive and consists of the sum of a continuous differentiable part given by
\[ \sum_{i,j} \frac{|v_{ij} - w_{ij}|^2}{2\tau} + \frac{\alpha}{2}(v^2_{ij} + \eta_e)(|\nabla_h u)_{ij}|^2) + \frac{\gamma}{2e}(1 - v_{ij}) \]
and the sum over the characteristic functions \( \chi_{[0,1]}(v_{ij}) \). If one of the \( v_{ij} \) is not contained in \([0,1] \), the subdifferential is clearly empty. Otherwise we have
\[ \partial v_{ij} \tilde{F}_{u,w}(v) = \frac{v_{ij} - w_{ij}}{\tau} + \alpha v_{ij}(|\nabla_h u)_{ij}|^2 - \frac{\gamma}{2e} + \partial \chi_{[0,1]}(v_{ij}). \]
Thus we only have to tackle a coordinate wise one-dimensional problem which is minimizing
\[ t \mapsto \frac{|t - w_{ij}|^2}{2\tau} + \frac{\alpha}{2} t^2(|\nabla_h u)_{ij}|^2 - \frac{\gamma}{2e} t \]
on \([0,1] \) for each \( i \) and \( j \). Precisely, with \( w = v^k - \tau \nabla_h q^{k+1} \) we simply have to compute for each \( i,j \)
\[ v^{k+1}_{ij} = 0 \lor \left( \frac{2e w_{ij} + \tau \gamma}{2e(1 + \alpha \tau (|\nabla_h u)_{ij}|^2)} \right) \land 1. \]

4.1. **Classical Ambrosio-Tortorelli.** In comparison with the introduced approximation of the Mumford-Shah functional we give a second example with the frequently used classical Ambrosio-Tortorelli functional given by
\[ \mathcal{AT}_\varepsilon(u,v) := \frac{\alpha}{2} \int_\Omega (v^2 + \eta_e)|\nabla_h u|^2 \, dx + \frac{\beta}{2} \int_\Omega |u - g|^2 \, dx + \gamma \int_\Omega \frac{1}{4\varepsilon}(1 - v)^2 + \varepsilon |\nabla_h v|^2 \, dx \]
for $u \in H^1(\Omega)$ and $v \in H^1(\Omega; [0,1])$. The discretization with finite differences as above then given by

$$
\mathcal{A}^d\varepsilon(u,v) := \sum_{ij} \alpha \frac{1}{2} (v_{ij}^2 + \eta \varepsilon) |(\nabla_h u)_{ij}|^2 + \frac{\beta}{2} |u_{ij} - g_{ij}|^2 + \frac{\gamma}{4\varepsilon} (1 - v_{ij})^2 + \gamma \varepsilon |(\nabla_h v)_{ij}|^2.
$$

The two minimization problems (4.4) and (4.5) can be simply solved by solving the Euler-Lagrange equation. That is solving (4.8) with respect to $u \in \mathbb{R}^{M \times N}$ and (4.15)

$$
\left( \text{diag}(|\nabla_h u|^2) + \frac{\gamma}{2\varepsilon} \text{Id} + 2\gamma \varepsilon \left((D^{(1)})^\top D^{(1)} + (D^{(2)})^\top D^{(2)}\right) \right) v = \frac{\gamma}{2\varepsilon}
$$

with respect to $v \in \mathbb{R}^{M \times N}$ subject to $0 \leq v_{ij} \leq 1$.

**References**


