NUMERICAL APPROXIMATIONS OF THE MUMFORD-SHAH FUNCTIONAL FOR UNIT VECTOR FIELDS

JONAS HAEHNLE

Abstract. Two numerical approximation schemes for minimising the Mumford-Shah functional for unit vector fields are proposed, analysed, and compared. The first uses a projection strategy, the second a penalisation strategy to enforce the sphere constraint. Both schemes are then applied to the segmentation of colour images using the Chromaticity and Brightness colour model.

1. Introduction

For $\Omega \subset \mathbb{R}^d$, and $\gamma, \alpha, \lambda$ positive constants, we are interested in numerically minimising the following weak version of the Mumford-Shah energy functional:

$$G(u) := \frac{\gamma}{2} \int_{\Omega} |\nabla u|^2 \, dx + \alpha \mathcal{H}^{d-1}(S_u) + \frac{\lambda}{2} \int_{\Omega} |u - g|^2 \, dx,$$

with $u, g \in GSBV(\Omega, \mathbb{R}^m)$, and $|u|^2 = 1$ a.e. (see Section 2 for definitions). This is a prototype problem for studying interesting effects with applications in image processing (see e.g. [43, 44, 8, 10, 19, 50, 7]), and liquid crystal theory (see e.g. [39, 42, 21, 51, 1, 6, 16]).

We are sometimes going to refer to functional (1.1) as the “Mumford-Shah” functional. It is, in fact, a version (for sphere-valued functions) of a functional proposed by De Giorgi, Carriero, and Leaci in [27] (for scalar functions) as a weak formulation of the original functional proposed by Mumford and Shah in [43] for greyscale image segmentation,

$$E(u, K) := \frac{\gamma}{2} \int_{\Omega \setminus K} |\nabla u|^2 \, dx + \alpha \mathcal{H}^{d-1}(K) + \frac{\lambda}{2} \int_{\Omega} (u - g)^2 \, dx,$$

with $g \in L^2(\Omega)$, which is to be minimised for all closed sets $K \subset \Omega$, and functions $u \in H^1(\Omega \setminus K)$. It is shown in [27] that the two problems are essentially equivalent.

The goal of image segmentation is to partition images into meaningful regions, which is often done by finding the edges which bound these regions, and which are in our case identified with the set $K$. The first term in (1.2) ensures smoothness of $u$ outside of $K$, the second one ensures that there are not too many edges, and the last term ensures that the segmented image $u$ does not deviate too much from the original one $g$.

A more concrete motivation for studying functional (1.1), therefore is colour image segmentation in the Chromaticity and Brightness (CB) colour model, where the chromaticity (colour information) is represented by an $S^{m-1}$-valued function (usually $m = 3$) on the image domain $\Omega$. The brightness, represented by a function $b : \Omega \to [0, 1]$, can be separately treated just like a greyscale image. It has been proposed that this model is well-suited for colour image processing. Osher and Vese [44] studied $p$-harmonic flows to the sphere ($p \geq 1$, in particular $p \in \{1, 2\}$), and applied them to image chromaticity, for example; other sources include [19, 50, 7] and references therein.

The name free discontinuity problems was introduced by De Giorgi in [24] for variational problems like (1.2), which consist of minimising a functional with volume and surface terms, depending on a closed set $K$ and a function $u$ (usually smooth outside $K$). Other early sources include [26, 25]. Weak formulations like (1.1) allow to prove existence of solutions (see [27] for the scalar, and [17] for the sphere-valued case), but still require the computation of geometric properties of the unknown set of discontinuity boundaries.

Date: June 17, 2010.
Therefore, Ambrosio and Tortorelli introduced an elliptic approximation in [3, 4], whose vectorial version, if defined for sphere-valued functions, is to minimise

\[
AT_{\varepsilon}(u, s) := \frac{\gamma}{2} \int_\Omega (s^2 + k_\varepsilon) |\nabla u|^2 \, dx + \alpha \int_\Omega \left( \varepsilon |\nabla s|^2 + \frac{1}{4\varepsilon} (1 - s)^2 \right) \, dx \\
+ \frac{\lambda}{2} \int_\Omega |u - g|^2 \, dx
\]  

(1.3)

for \( u, g \in H^1(\Omega, S^{m-1}) \), \( s \in H^1(\Omega, [0, 1]) \), \( 0 < \varepsilon, k_\varepsilon \ll 1 \), and \( k_\varepsilon = o(\varepsilon) \). Here, \( s \) is a phase function approximating \( 1 - \chi_K \) by penalisation of phase transitions. Ambrosio and Tortorelli showed \( \Gamma \)-convergence of \( AT_{\varepsilon}(u, s) \) to \( G(u) \) in \( L^2 \) in the scalar case ([8]), as well as the \( S^{m-1} \)-valued case ([4]) for \( \varepsilon \to 0 \).

Bellettini and Coscia carried out a finite element approximation of the Mumford-Shah functional in the scalar case, based on this elliptic approximation in [8]. They showed that their approximation \( G_{\varepsilon,h} : V^h(\Omega) \times V^h(\Omega, [0, 1]) \to \mathbb{R} \) is \( \Gamma \)-convergent to \( G : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R} \) provided that the mesh size fulfils \( h = o(k_\varepsilon) \), and that \( S_n \) is piecewise \( C^2 \). Here, \( V^h(\Omega) \) is the continuous, piecewise affine finite element space. Using the approximation result in [28], Bourdin in [10] showed that \( S_n \) does not need to be assumed piecewise \( C^2 \); and he proposed an algorithm for actual computations — without providing a proof for its convergence, though. The problem here is that the two variables \( u \) and \( s \) appear strongly coupled in the energy and in the corresponding gradient flow.

As an alternative to the above phase-field approximation of the Mumford-Shah functional, Braides and Dal Maso proposed a non-local approximation approach in [13], on which Cortesani based a \( \Gamma \)-convergent, vector-valued finite element approximation in [22].

A different motivation for (1.3) comes from the theory of nematic liquid crystals. In order to overcome mathematical difficulties in showing existence and regularity of energy minimising static configurations in the Oseen-Frank model, Lin in [39] adapts Ericksen’s energy, which he simplifies to (see [39, equation (3.12)])

\[
\int_\Omega \frac{1}{2} \delta^2 |\nabla u|^2 + |\nabla s|^2 + W_0(s) \, dx
\]

with variable degree of orientation \( s \in [-1/2, 1] \) (in experiments, often \( s \geq 0 \)), and director \( n, |n| = 1 \) a.e. The strong similarities of this energy to functional (1.3) lets us hope that our analysis may be of use to this application, too.

The overall goal of the present work is to construct and analyse convergent discretisations for a prototype problem with several non-convexities; namely, we consider a non-convex functional (the Mumford-Shah functional) with a non-convex constraint (the sphere constraint), as an extension to existing work on convex functionals (in particular harmonic maps) with non-convex constraints, which have been intensely studied (see e.g. [1, 5, 6] and references therein). In particular, we deal with discretisations of the sphere constraint, which we account for using a projection and a penalisation strategy. The former turns out to deliver more convincing computational results, while the latter is analytically more satisfactory.

Below, we give a short overview over the two methods for the approximation of (1.1) that we shall present in Sections 3 - 7 of this paper, where we in particular discuss relevant stability properties of computed approximations, such as

- energy decay property for splitting schemes related to (1.3),
- the validity of a discrete or penalised sphere constraint for approximations of \( u \), and
- non-negativity and upper bounds for approximations of the phase field function \( s \).

1.1. Splitting & Projection Strategy. The problem of coupled variables is addressed through an iterative splitting strategy; i.e., in every step of the iteration the energy is first minimised with respect to the first variable while keeping the second variable fixed, and then minimised with respect to the second variable while keeping the first one fixed. A special projection idea as proposed by Alouges in [1] is used to enforce the sphere constraint. We propose a first-order finite element discretisation, which preserves the sphere constraint exactly at nodal points. The resulting discrete algorithm is simple, results in only linear equations to be solved in every step of the iteration, and every step is energy-decreasing (for acute triangulations). The algorithm converges weakly (up to subsequences) in \( H^1 \times H^1 \) to a tuple \( (u, s) \in H^1(\Omega, S^{m-1}) \times H^1(\Omega) \). For \( d = 2 \) we can show that \( s \) and iterates \( S_n \) fulfil \( S_n, s \in [-1, 1] \). However, we cannot show that \( (u, s) \) is a stationary point of the Ambrosio-Tortorelli energy for unit vector fields.
1.2. Penalisation & Splitting Strategy. This method again uses a splitting strategy, but the sphere constraint is now approximated by penalisation; i.e., we add a Ginzburg-Landau term \( \frac{1}{2\varepsilon^2} \int_\Omega (|u|^2 - 1)^2 \, dx \) (\( 0 < \varepsilon \ll 1 \)) to the energy \((1.3)\). We show that for proper scales of \( \delta_\varepsilon \) in terms of \( \varepsilon \), this does not affect \( \Gamma \)-convergence. Furthermore, we propose a first-order finite element algorithm based on this splitting and penalisation strategy. The resulting algorithm converges weakly (up to subsequences) in \( H^1 \times H^1 \) to a tuple \((u, s)\) in \( H^1(\Omega, \mathbb{R}^n) \times H^1(\Omega) \), without any mesh-constraint. For \( d = 2 \) we can also show that \( S_m, s \in [-1, 1] \). This allows to get strong convergence (up to subsequences) of iterates \( U_n \) in \( H^1 \), which in turn allows to pass to the limit and show that \((u, s)\) is a stationary point of the Ambrosio-Tortorelli-Ginzburg-Landau energy, and that \( s \geq 0 \). However, we now have to solve a nonlinear equation in every iteration.

In Section 6, comparative computational experiments for the “Penalisation & Splitting” and the “Splitting & Projection” methods are presented, which address in particular

1. the effect of perturbing the sphere constraint throughout minimisation, as well as proper scalings of regularisation and numerical parameters;
2. the accuracy of zero sets of \( s \) in the course of minimisation; and
3. comparative numerical studies to relate the CB and RGB models in colour image segmentation.

2. Preliminaries

We often use \( c \) and \( C \) as generic non-negative constants, capital letters for finite element functions and boldface for vectors or vector-valued functions. Given \( x, y \in \mathbb{R}^n \), \( \langle x, y \rangle \) or \( x \cdot y \) will denote their standard scalar product, and \( |x| \) the Euclidean norm of \( x \). For a set \( S \), \( |S| \) or \( L^d(S) \) denotes its Lebesgue measure of dimension \( d \), \( H^d(S) \) its Hausdorff measure. The \( L^2 \) scalar product and norm will be denoted by \( (\cdot, \cdot) \) and \( \| \cdot \| \), respectively, and \( S^{m-1} \) will be the unit sphere in \( \mathbb{R}^m \). For \( a, b \in \mathbb{R} \), let \( a \wedge b := \min\{a, b\} \), and \( a \vee b := \max\{a, b\} \). By \( A : B \) for \( A, B \in \mathbb{R}^{m \times m} \) we denote the dyadic product; i.e., \( A : B := \sum_{i,j=1}^m a_{ij} b_{ij} \) for \( A = (a_{ij}), B = (b_{ij}) \). Let \( |A| \) denote the Frobenius norm of \( A \); i.e., \( |A|^2 := \sum_{i,j=1}^m |a_{ij}|^2 \). For two vectors \( a \in \mathbb{R}^d, b \in \mathbb{R}^m \), let \( a \otimes b := M \) denote the matrix with entries \( m_{ij} := a_i b_j \).

2.1. Functions of Bounded Variation and \( \Gamma \)-Convergence. We summarise some definitions and results on functions of bounded variation and \( \Gamma \)-convergence. Sources are e.g. [2, 35, 30, 23, 11, 12, 18].

2.1.1. \( BV, SBV, \) and \( GSBV \) Functions. Let \( \Omega \subset \mathbb{R}^d \) be a bounded open set, \( u : \Omega \to \mathbb{R}^m \) a measurable function, \( S := \mathbb{R}^m \cup \{\infty\} \), and \( x \in \Omega \) be fixed. We call \( z \in S \) the approximate limit of \( u \) at \( x \), or \( z = \text{ap} - \lim_{y \to x} u(y) \), if for every neighbourhood \( U \) of \( z \in S \) we have

\[
\lim_{\varepsilon \to 0} \frac{1}{\varepsilon^d} |\{ y \in \Omega : |y - x| < \varepsilon, u(y) \notin U \}| = 0.
\]

If \( z \in \mathbb{R}^m \), we call \( x \) a Lebesgue point of \( u \), and we denote by \( S_u \) the complement of the set of Lebesgue points of \( u \) (approximate discontinuity set). Since \( |S_u| \) is known to be zero, \( u = \tilde{u} \) a.e. for

\[
\tilde{u}(x) := \text{ap} - \lim_{y \to x} u(y).
\]

Let \( x \in \Omega \setminus S_u \) such that \( \tilde{u}(x) \neq \infty \). If there exists \( L \in \mathbb{R}^{d \times m} \) such that

\[
\text{ap} - \lim_{y \to x} \frac{|u(y) - \tilde{u}(x) - L(y - x)|}{|y - x|} = 0,
\]

we call \( u \) approximately differentiable in \( x \), and \( \nabla u(x) := L \) the (uniquely determined) approximate gradient of \( u \) in \( x \). A function \( u \in L^1(\Omega, \mathbb{R}^m) \) is called a function of bounded variation in \( \Omega \), or \( u \in BV(\Omega, \mathbb{R}^m) \), if its distributional derivative \( Du \) is representable by a measure with finite total variation \( |Du| (\Omega) \); i.e., if

\[
\sum_{a=1}^m \int_\Omega u_a \text{div} (\varphi_a) \, dx = -\sum_{a=1}^m \sum_{i=1}^d \int_\Omega \varphi_a^i dD_a u^a \quad \forall \varphi \in C^1_c(\Omega, \mathbb{R}^{m \times d}),
\]

with \( Du \) an \( \mathbb{R}^{d \times m} \) valued matrix of measures \( D_i u^a \), and \( u = (u_1, \ldots, u_m) \). Defining

\[
\|u\|_{BV(\Omega, \mathbb{R}^m)} := \|u\|_{L^1(\Omega, \mathbb{R}^m)} + |Du|(\Omega),
\]

makes \( BV(\Omega, \mathbb{R}^m) \) a Banach space.
If \( \{ u_j \} \subset BV(\Omega, \mathbb{R}^m) \) with \( \sup_j \| u_j \|_{BV(\Omega, \mathbb{R}^m)} < +\infty \), then there exist a subsequence \( \{ u_{j_k} \} \) and a function \( u \in BV(\Omega, \mathbb{R}^m) \) such that \( u_{j_k} \rightharpoonup u \) in \( L^1(\Omega, \mathbb{R}^m) \), and \( Du_{j_k} \rightharpoonup Du \) weakly-* in the sense of measures.

Also, for \( u \in BV(\Omega, \mathbb{R}^m) \), \( S_u \) is countably \( H^{d-1} \)-rectifiable; i.e.,

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

where \( H^{d-1}(N) = 0 \), and each \( K_i \) is a compact subset of a \( C^1 \) manifold. So, for \( H^{d-1} \)-a.e. \( y \in S_u \) we can define an exterior unit normal \( \nu_u \) and outer and inner traces of \( u \) on \( S_u \) by

\[
\begin{align*}
    u^\pm(x) := & \text{ap} - \lim_{y \to x \atop y \in \pi^+(x, u(x))} u(y), \\
    \pi^\pm(x, \nu_u(x)) := & \{ y \in \mathbb{R}^d : \pm(y - x, \nu_u(x)) > 0 \}.
\end{align*}
\]

A point \( x \in \Omega \) is called a jump point of \( u \), \( x \in J_u \), if there exists \( \nu \in S^{d-1} \), such that

\[
\text{ap} - \lim_{y \to x \atop y \in \pi^-(x, \nu)} u(y) \neq \text{ap} - \lim_{y \to x \atop y \in \pi^+(x, \nu)} u(y).
\]

It is known that \( J_u \subseteq S_u \) and \( H^{d-1}(S_u \setminus J_u) = 0 \).

If we decompose \( Du \) into an absolutely continuous part \( D^a u \) and a singular part \( D^s u \), both with respect to the Lebesgue measure \( \mathcal{L}^d \), \( Du = D^a u + D^s u \), the density of \( D^a u \) with respect to \( \mathcal{L}^d \) coincides with the approximate gradient \( \nabla u \mathcal{L}^d \)-a.e. The restriction \( D^1_u \) of \( D^a u \) to \( S_u \) is called jump part of \( Du \), the restriction \( D^s_u \) of \( D^s u \) to \( \Omega \setminus S_u \) it called Cantor part. So,

\[
Du = D^a u + D^s u + D^c u.
\]

It is known that \( D^s u = (u^+ - u^-) \otimes \nu_u H^{d-1} | S_u \).

A function \( u \in BV(\Omega, \mathbb{R}^m) \) is called a special function of bounded variation in \( \Omega \), \( u \in SBV(\Omega, \mathbb{R}^m) \), if \( Du = 0 \). We call \( u \in BV(\Omega, \mathbb{R}^m) \) a generalised special function of bounded variation, \( u \in GSBV(\Omega, \mathbb{R}^m) \), if \( g(u) \in SBV(\Omega, \mathbb{R}^m) \) for every \( g \in C^1(\mathbb{R}^m) \) such that \( \nabla g \) has compact support. For \( 1 < p < +\infty \), let

\[
(G)SBV^p(\Omega, \mathbb{R}^m) := \{ u \in (G)SBV(\Omega, \mathbb{R}^m) : H^{d-1}(J_u) < +\infty, \nabla u \in L^p(\Omega, \mathbb{R}^{d \times m}) \}.
\]

We remark that \( W^{1,1}(\Omega, \mathbb{R}^m) \subseteq BV(\Omega, \mathbb{R}^m) \), that \( u \in SBV(\Omega, \mathbb{R}^m) \) implies \( u \in W^{1,1}(\Omega \setminus S_u, \mathbb{R}^m) \), and that \( SBV(\Omega, \mathbb{R}^m) \cap L^\infty(\Omega, \mathbb{R}^m) = GSBV(\Omega, \mathbb{R}^m) \cap L^\infty(\Omega, \mathbb{R}^m) \).

2.1.2. \( \Gamma \)-Convergence. Let \( X \) be a separable Banach space with a topology \( \tau \) and let \( F_\varepsilon : X \to \overline{\mathbb{R}} \) be a sequence of functionals. We say \( F_\varepsilon \) \( \Gamma \)-converges to \( F \) in the topology \( \tau \), or \( F = \Gamma - \lim \varepsilon \to 0 F_\varepsilon \), if the following two conditions hold:

1. For every \( x \in X \) and for every sequence \( \{ x_\varepsilon \} \subset X \) \( \tau \)-converging to \( x \in X \),

\[
F(x) \leq \inf_{\varepsilon \to 0} F_\varepsilon \left( x_\varepsilon \right).
\]

2. For every \( x \in X \) there exists a sequence \( \{ x_\varepsilon \} \subset X \) (recovery sequence) \( \tau \)-converging to \( x \in X \), such that

\[
F(x) \geq \limsup_{\varepsilon \to 0} F_\varepsilon \left( x_\varepsilon \right).
\]

**Lemma 2.1.** Let \( F_\varepsilon, F : X \to \overline{\mathbb{R}}, \) with \( \Gamma - \lim_{\varepsilon \to 0} F_\varepsilon = F \). Then

1. \( F \) is lower semicontinuous on \( X \).
2. \( F + G = \Gamma - \lim (F_\varepsilon + G) \) for all continuous \( G : X \to \overline{\mathbb{R}} \).
3. Let \( \{ u_\varepsilon \} \subset X \) be such that

\[
\lim_{\varepsilon \to 0^+} \left( F_\varepsilon (u_\varepsilon) - \inf_X F_\varepsilon \right) = 0,
\]

then every cluster point \( u \) of \( \{ u_\varepsilon \} \) minimises \( F \) over \( X \), and

\[
\lim_{\varepsilon \to 0^+} \inf_X F_\varepsilon = \min_X F = F(u).
\]

Here are some connections between \( \Gamma \)-convergence and pointwise convergence:

- If \( F_\varepsilon \) converges uniformly to \( F \), then \( F_\varepsilon \) \( \Gamma \)-converges to \( F \).
- If \( F_\varepsilon \) is decreasing and converges pointwise to \( F \), then \( F_\varepsilon \) \( \Gamma \)-converges to \( RF \), the lower semicontinuous envelope of \( F \).
3. Splitting & Projection Algorithm

Let \( \Omega \subset \mathbb{R}^d \) be a polyhedral Lipschitz domain, and \( \mathcal{T}_h \) be a quasi-uniform triangulation of \( \Omega \) with node set \( \mathcal{N} \) and maximal mesh size \( h > 0 \) (c.f. [14]). The space of globally continuous, piecewise affine finite element functions on \( \mathcal{T}_h \) is denoted by \( V_h(\Omega) \subseteq H^1(\Omega) \). The nodal basis functions are \( \{ \varphi_z : z \in \mathcal{N} \} \subseteq V_h(\Omega) \). Let \( V_h(\Omega, \mathbb{R}^m) \) be the finite element space of \( \mathbb{R}^m \)-valued mappings with basis functions \( \{ \varphi_z : z \in \mathcal{N}, 1 \leq i \leq m \} \), with \( \varphi_z^1 := (\varphi_z, 0, \ldots)^T \in V_h(\Omega, \mathbb{R}^m), \varphi_z^2 := (0, \varphi_z, 0, \ldots)^T \in V_h(\Omega, \mathbb{R}^m) \), and so forth. Let \( I_h(\cdot) : C^0(\overline{\Omega}) \to V_h(\Omega) \) be the Lagrange interpolation operator, and \( R_h(\cdot) : H^1(\Omega) \to V_h(\Omega) \) the Ritz projection, defined by 

\[
(\nabla (R_h(\varphi) - \varphi), \nabla V) + (R_h(\varphi) - \varphi, V) = 0 \quad \forall V \in V_h(\Omega),
\]

and \( r_h(\cdot) : L^2(\Omega) \to V_h(\Omega) \) the Clément operator [20] \((I_h(\cdot), R_h(\cdot), \text{and } r_h(\cdot) \text{ in the vector valued case}). The latter operator will be needed since it can be applied to non-continuous functions.

**Lemma 3.1.** The tuple \( (u, s) \in H^1(\Omega, S^{m-1}) \times H^1(\Omega, [0,1]) \) is a stationary point of \( AT_\varepsilon(\cdot, \cdot) \) if and only if

\[
\gamma \left((s^2 + k_\varepsilon) \nabla u, \nabla \varphi\right) = \lambda(g, \varphi)
\]

for all \( \varphi \in \overline{H^1(\Omega, \mathbb{R}^m)} \) such that \( \varphi(x) \in T_u(x)S^{m-1} \) (the tangent space of \( S^{m-1} \) at \( u(x) \)), and

\[
2\alpha \varepsilon (\nabla s, \nabla \varphi) + \left(\gamma |\nabla u|^2 + \frac{\alpha}{2\varepsilon}\right) s, \varphi = \left(\frac{\alpha}{2\varepsilon}\right) \varphi
\]

for all \( \varphi \in \overline{H^1(\Omega)} \cap L^\infty(\Omega) \).

**Proof.** Note \( u \cdot \varphi = 0 \) a.e. and derive the first variation of \( AT_\varepsilon(\cdot, \cdot) \) with respect to \( u \) and \( s \), respectively, c.f. [49] and [15, Proposition 1]. \( \square \)

The most natural approach to the discrete case would be to work with the original functional \( AT_\varepsilon(\cdot, \cdot) \). However, it is not clear how to get a uniform \( L^\infty \) bound on iterates \( S_n \) in this setting. We therefore introduce mass lumping into the last term: For \( G \in V_h(\Omega, \mathbb{R}^m) \), we define

\[
E_h(U, S) := \frac{\gamma}{2} \int_\Omega (S^2 + k_\varepsilon) |\nabla U|^2 \, dx + \frac{\lambda}{2} \int_\Omega |U - G|^2 \, dx
\]

\[
+ \alpha \int_\Omega \varepsilon |\nabla S|^2 + \frac{1}{4\varepsilon} I_h((1 - S)^2) \, dx,
\]

and

\[
\tilde{E}(U, S) := \frac{\gamma}{2} \int_\Omega (S^2 + k_\varepsilon) |\nabla U|^2 \, dx + \frac{\lambda}{2} \int_\Omega |U - G|^2 \, dx,
\]

with \( \gamma, \alpha, \varepsilon, k_\varepsilon \) fixed and positive, and \( \lambda \geq 0 \). We also assume \( d \leq 2 \), since so far, our arguments for the \( L^\infty \) bound on iterates \( S_n \) fail for higher dimensions (the rest of the analysis works for \( d \leq 3 \)), but we hope it will be possible to improve these results (and possibly remove lumping altogether).

Another solution would be to use mass lumping in all nonlinear terms involving \( S \); i.e., to use the functional

\[
\frac{\gamma}{2} \int_\Omega (I_h(S^2) + k_\varepsilon) |\nabla U|^2 \, dx + \frac{\lambda}{2} \int_\Omega |U - G|^2 \, dx
\]

\[
+ \alpha \int_\Omega \varepsilon |\nabla S|^2 + \frac{1}{4\varepsilon} I_h((1 - S)^2) \, dx.
\]

This introduces additional errors, but it still allows to get the necessary uniform \( H^1 \) bounds on iterates \((U_n, S_n)\), in addition to the \( L^\infty \) bound on \( S_n \), and it does not require \( d \leq 2 \); see [15] for details.

Functions \( V \in V_h(\Omega, \mathbb{R}^m) \) which satisfy the pointwise constraint \(|V| = 1 \) are necessarily constant. So it is more reasonable to work in the space

\[
H^1_h(T_h) := \left\{ V \in V_h(\Omega, \mathbb{R}^m) : V(z) \in S^{m-1} \forall z \in \mathcal{N} \right\}.
\]

We set

\[
K_n^h := \left\{ W \in V_h(\Omega, \mathbb{R}^m) : W(z) \cdot U_n(z) = 0 \forall z \in \mathcal{N} \right\},
\]

where \( U_n \in H^1_h(T_h) \) will be the iterates of the fully discrete algorithm.

The idea now is to find \( U \in K_n^h \) minimising \( \tilde{E}(\cdot, S) \) and then project to the sphere. This approach is based on [1] and [5] and replaces the nonlinear, non-convex constraint \( U \in H^1_h(T_h) \) by the linear one \( W(z) \cdot U_n(z) = 0 \forall z \in \mathcal{N} \), which in turn ensures that projection to the sphere does not increase the energy.
Algorithm 3.2. Let a quasi-uniform triangulation $T_h$ of $\Omega$, starting values $U_0, S_0$, and parameters $\varepsilon, k, \varrho > 0$ be given. For $n := 0, \ldots$

1. Minimise $\tilde{E}(U_n - W, S_n)$ for $W \in K^m_h$; i.e. solve

$$\gamma((S_n^2 + k_n) \nabla (U_n - W), \nabla V) - \lambda(W + G, V) = 0,$$

for all $V \in K^m_h$, and call the solution $W_n$.

2. If $\|W_n\|_{H^1(\Omega; \mathbb{R}^m)} \leq \varrho$ set $U_n := U_n, W := W_n, S := S_n$ and stop.

3. Set

$$U_{n+1} := \sum_{x \in \mathcal{N}} \frac{U_n(z) - W_n(z)}{|U_n(z) - W_n(z)|} \varphi_x.$$

4. Minimise $E_h(U_{n+1}, S)$ for all $S \in V_h(\Omega)$; i.e. solve

$$\gamma(S |\nabla U_{n+1}|^2; W) + \alpha \left(\frac{1}{2} \langle S - 1, W \rangle_h \right) = 0$$

for all $W \in V_h(\Omega)$, and call the solution $S_{n+1}$.

Here $(\varphi, \psi)_h := \int_T h \varphi \psi \, dx$ for $\varphi, \psi \in C(\Omega)$.

Definition 3.3. Let $T_h$ be a quasi-uniform triangulation of $\Omega$, and $s \in H^1(\Omega)$ be fixed. $T_h$ is said to satisfy an energy decreasing condition (ED) if

$$E_h(W, s) \leq E_h(V, s)$$

for all $V \in V_h(\Omega)$ fulfilling $|V(z)| \geq 1$ for $z \in \mathcal{N}$. Here $W \in V_h(\Omega, \mathbb{R}^m)$ is defined by

$$W := \sum_{x \in \mathcal{N}} \frac{V(z)}{|V(z)|} \varphi_x.$$

As demonstrated in [5, Lemma 3.2 & Remarks 3.3], for $d \leq 3$ (ED) is fulfilled if every angle in $T_h$ is $\leq \pi/2$ (i.e., if the triangulation is acute).

Lemma 3.4. Let $U \in V_h(\Omega, \mathbb{R}^m)$ be given, and $d \leq 2$. If $S \in V_h(\Omega)$ minimises $E_h(U, \cdot)$, then $-1 \leq S \leq 1$.

Proof. For $a \in \mathbb{R}$ define $\bar{a} := -1 \vee a \wedge 1$. Note that for this result it is crucial that we have piecewise affine finite element functions.

**Step 1:** If $a, b \in \mathbb{R}$, then $(\bar{a} + \bar{b})^2 \leq (a + b)^2$ and $(\bar{a} - \bar{b})^2 \leq (a - b)^2$.

A case differentiation gives

- $a, b \in [-1, 1]$ is trivial.
- $a, b > 1$ or $a, b < -1 \implies (\bar{a} + \bar{b})^2 = 2^2 \leq (a + b)^2$.
- $a > 1, b < -1 \implies (\bar{a} + \bar{b})^2 = 0 \leq (a + b)^2$.
- $a > 1, a < -1$ is symmetrical.
- $a \not\in [-1, 1], b \in [-1, 1] \implies 0 \leq |a + sign(ab)|b| \leq |a| + sign(ab)|b|,

$$\implies (\bar{a} + \bar{b})^2 = (1 + sign(ab)|b|)^2 \leq (|a| + sign(ab)|b|)^2 = (a + b)^2,$$

and $b \not\in [-1, 1], a \in [-1, 1]$ is symmetrical.

Therefore $(\bar{a} + \bar{b})^2 \leq (a + b)^2$, and $(\bar{a} - \bar{b})^2 \leq (a - b)^2$ follows by symmetry.

**Step 2:** We have $-1 \leq S \leq 1$.

In case $-1 \leq S \leq 1$ should not be true, we replace $S(x) = \sum_{x \in \mathcal{N}} S(z) \varphi_x(x)$ by

$$\bar{S}(x) := \sum_{x \in \mathcal{N}} (1 \vee S(z) \wedge 1) \varphi_x(x) = \bar{T}_h(-1 \vee S \wedge 1),$$

for which clearly $-1 \leq \bar{S} \leq 1$. We shall prove $E_h(U, \bar{S}) \leq E_h(U, S)$, by showing energy-decrease for every term involving $S$, on every triangle $T \in T_h$. Since $\nabla U$ is constant on every $T$, the terms we have to look at are $\int_T S^2 \, dx$, $\int_T |\nabla S|^2 \, dx$, and $\int_T \bar{T}_h((1 - S)^2) \, dx$. Let the values of $S$ at the nodal points of $T$ be $S_0, \ldots, S_d$, let $\bar{S}_0, \ldots, \bar{S}_d$ be the corresponding nodal basis functions, and $x := (x_1, \ldots, x_d)$. By a simple transformation argument, we can restrict ourselves to the standard simplex, which we shall still call $T$. Then

$$S(x)|_T = S_0 + \sum_{i=1}^d (S_i - S_0) x_i,$$
and
\[ \nabla S(x)|_T = (S_1 - S_0, \ldots, S_d - S_0), \]

For the first term, a calculation yields
\[ (3.5) \quad \int_T S^2 \, dx = \frac{2}{(d+2)!} \sum_{i=0}^d S_i \sum_{j=i}^d S_j. \]

If \( d = 1 \), then, by Step 1,
\[
\int_T S^2 \, dx = \frac{1}{3} \left( S_0^2 + S_0 S_1 + S_1^2 \right) \\
= \frac{1}{6} \left( (S_0 + S_1)^2 + S_0^2 + S_1^2 \right) \\
\leq \frac{1}{6} \left( (S_0 + S_1)^2 + S_0^2 + S_1^2 \right) \\
= \int_T S^2 \, dx.
\]

Similarly, if \( d = 2 \),
\[
\int_T S^2 \, dx = \frac{1}{12} \left( S_0^2 + S_1^2 + S_2^2 + S_0 S_1 + S_0 S_2 + S_1 S_2 \right) \\
= \frac{1}{24} \left( (S_0 + S_1)^2 + (S_0 + S_2)^2 + (S_1 + S_2)^2 \right) \\
\leq \frac{1}{24} \left( (S_0 + S_1)^2 + (S_0 + S_2)^2 + (S_1 + S_2)^2 \right) \\
= \int_T S^2 \, dx.
\]

Note: Both arguments break down for \( d \geq 3 \); in fact, counter-examples are easy to find, c.f. Remark 3.5.

The second term gives, by Step 1 and symmetry,
\[
\int_T |
abla S|^2 \, dx = \int_T (S_1 - S_0, \ldots, S_d - S_0)^2 \, dx \\
= \frac{1}{d} \left( (S_1 - S_0)^2 + \cdots + (S_d - S_0)^2 \right) \\
\leq \frac{1}{d} \left( (S_1 - S_0)^2 + \cdots + (S_d - S_0)^2 \right) \\
= \int_T |
abla S|^2.
\]

As for the last term, again by Step 1,
\[
\int_T \mathcal{I}_h ((1 - S)^2) \, dx = \sum_{i=1}^{d+1} (1 - S_i)^2 \int_T \varphi_i \, dx \\
\leq \sum_{i=1}^{d+1} (1 - S_i)^2 \int_T \varphi_i \, dx = \int_T \mathcal{I}_h ((1 - S)^2) \, dx.
\]

\[ \square \]

Remark 3.5. For \( d = 3 \), Step 2 in the above proof is wrong: Let \( S_0 := S_1 := S_2 := 1 \), and \( S_3 := -3/2 \).

Then, by (3.5),
\[ \int_T S^2 \, dx = \frac{1}{60} \sum_{i=0}^d S_i \sum_{j=i}^d S_j = \frac{1}{15}, \]

while
\[ \int_T S^2 \, dx = \frac{1}{60} \sum_{i=0}^d S_i \sum_{j=i}^d S_j = \frac{1}{16}. \]

We suspect that there exist dimension-dependent constants \( c_d \), at which one could crop \( |S| \), so that the energy is still decreasing (also replacing \((1 - s)^2\) by \((c_d - s)^2\) ).
Lemma 3.6. Let $T_h$ be a quasi-uniform triangulation of $\Omega$ satisfying (ED), $\rho > 0$ fixed, $S_0 \in V_h(\Omega)$, and $U_0 \in H^1_0(T_h)$. Then Algorithm 3.2 terminates within a finite number of iterations with output $(U, S) \in H^1_0(T_h) \times V_h(\Omega, [1, 1])$ and $W \in V_h(\Omega, \mathbb{R}^m)$ such that $\|\nabla W\| \leq \rho$, and $E_h(U, S) \leq E_h(U_0, S_0)$.

Proof. We proceed by induction. Suppose that for some $n \geq 0$ we have $(U_n, S_n) \in H^1_0(T_h) \times V_h(\Omega)$. The set $K^a_h$ is a subspace of $V_h(\Omega, \mathbb{R}^m)$. Therefore, by Lax-Milgram, there is a unique $W_n \in K^a_h$ such that (3.3) is fulfilled. Since $W_n(z) \cdot U_n(z) = 0$ and $|U_n(z)| = 1$, we have for $z \in N$ $|U_n(z) - W_n(z)|^2 = 1 + |W_n(z)|^2 \geq 1$.

Therefore, $U_{n+1}$ is well-defined and in $H^1_0(T_h)$. And since $0 \in K^a_h$ and $T_h$ fulfills (ED), we get $E_h(U_{n+1}, S_n) \leq E_h(U_n - W_n, S_n)$.

Step 4 of Algorithm 3.2 has a solution $S_{n+1}$ by convexity and coercivity of the functional. So

$E_h(U_{n+1}, S_{n+1}) \leq E_h(U_{n+1}, S_n) \leq E_h(U_n - W_n, S_n) \leq E_h(U_n, S_n)$.

In fact, $E_h(U_{n+1}, S_{n+1}) \leq E_h(U_n, W)$ for all $W \in V_h(\Omega)$. Therefore, by Lemma 3.4, we can assume $-1 \leq S_{n+1} \leq 1$. Furthermore,

$I := 2\tilde{E}(U_{n+1}, S_{n+1}) - 2\tilde{E}(U_n, S_n)$

\[ \leq 2\tilde{E}(U_n - W_n, S_n) - 2\tilde{E}(U_n, S_n) \]

\[ \leq \gamma \int_{\Omega} (S_n^2 + k_\varepsilon) \left( |\nabla U_n|^2 + |\nabla W_n|^2 - 2\nabla U_n : \nabla W_n \right) dx \]

\[ + \lambda \int_{\Omega} |U_n|^2 + |W_n|^2 + |G|^2 - 2G \cdot (U_n - W_n) + 2U_n \cdot W_n dx \]

\[ - \int_{\Omega} \gamma (S_n^2 + k_\varepsilon) |\nabla U_n|^2 + \lambda \left( |U_n|^2 + |G|^2 - 2G \cdot U_n \right) dx \]

\[ = \int_{\Omega} \gamma (S_n^2 + k_\varepsilon) |\nabla W_n|^2 - 2\nabla U_n : \nabla W_n + \lambda \left( |W_n|^2 + 2W_n \cdot (G - U_n) \right) dx. \]

Using equation (3.3) with $V := W_n$, we get

$I \leq -\int_{\Omega} \gamma (S_n^2 + k_\varepsilon) |\nabla W_n|^2 + \lambda |W_n|^2 dx$, \hfill (3.3)

whence

$0 \leq \frac{1}{2} \int_{\Omega} \gamma (S_n^2 + k_\varepsilon) |\nabla W_n|^2 + \lambda |W_n|^2 dx \leq \tilde{E}(U_n, S_n) - \tilde{E}(U_{n+1}, S_{n+1})$. \hfill (3.4)

Summing this from 0 to $N$ leads to

$\frac{1}{2} \sum_{n=0}^{N} \int_{\Omega} \gamma (S_n^2 + k_\varepsilon) |\nabla W_n|^2 + \lambda |W_n|^2 dx \leq \tilde{E}(U_0, S_0) - \tilde{E}(U_{N+1}, S_{N+1}) < +\infty$;

i.e., the series

$\frac{1}{2} \sum_{n=0}^{N} \int_{\Omega} \gamma (S_n^2 + k_\varepsilon) |\nabla W_n|^2 + \lambda |W_n|^2 dx$ is convergent. Therefore, $\|W_n\|_{H^1(\Omega, \mathbb{R}^m)} \leq \rho$ for $n$ large enough. $\square$

Theorem 3.7. Let $\{T_h\}$ be a sequence of quasi-uniform triangulations satisfying (ED) with maximal mesh size $h_l \to 0$ for $l \to +\infty$, $g_l \to 0$ for $l \to +\infty$, and $E_h(U_0, S_0) \leq C_0 < +\infty$ independently of $h_l$. Let $\{U_i, S_i\}$ be the output of Algorithm 3.2 (after termination) from input $(U^0_i, S^0_i, g_l)$. Then the sequence $\{U_i, S_i\}$ converges weakly in $H^1(\Omega, \mathbb{R}^m) \times H^1(\Omega)$ (up to subsequences, not relabelled) for $l \to +\infty$ to a point $(u, s) \in H^1(\Omega, S^{m-1}) \times H^1(\Omega, [-1, 1])$, with $AT_e(u, s) \leq \liminf_l AT_e(U_i, S_i) \leq \liminf_l AT_e(U^0_i, S^0_i)$.

Proof. By assumption and Lemma 3.6, we have

$E_h(U_i, S_i) \leq E_h(U^0_i, S^0_i) \leq C_0$,

and $-1 \leq S_i \leq 1$. This implies uniform boundedness of $H^1$-norms of iterates $U_i$ and $S_i$. Hence we can extract a subsequence that converges weakly in $H^1 \times H^1$ to some map $(u, s)$. Poincaré’s inequality (elementwise), $|U_i(z)| = 1$ for all $z \in N_{h_l}$, and $|U_i| \leq 1$ a.e. imply

$\left\|U_i^2 - 1\right\| \leq Ch_l \left\|2U_i^2 \nabla U_i\right\| \leq Ch_l$. 

8
So $U_t \rightarrow u$ a.e. leads to $|u| = 1$ a.e.

Since $H^1(\Omega)$ is a Hilbert space and $\{ \varphi \in H^1(\Omega) : 0 \leq \varphi \leq 1 \text{ a.e.} \} \subset H^1(\Omega)$ is a closed, convex set, it is weakly closed. Therefore, by the weak convergence in $H^1$ of $S_t \rightharpoonup s$, we get $-1 \leq s \leq 1$.

Finally, by weak lower semicontinuity of $AT_\varepsilon(\cdot, \cdot)$,

$$AT_\varepsilon(u, s) \leq \liminf_l AT_\varepsilon(U_l, S_l) \leq \liminf_l \left( E_{h_l}(U_l, S_l) + \epsilon \left( (1 - S_l)^2 \right) \right) \leq \liminf_l \left( E_{h_l}(U_l^0, S_l^0) + \epsilon \left( (1 - S_l^0)^2 \right) \right) \leq \liminf_l \left( AT_\varepsilon(U_l^0, S_l^0) + \epsilon \left( (1 - S_l^0)^2 \right) \right).$$

\[\square\]

**Remark 3.8.** We cannot prove that $(u, s)$ is a stationary point of $AT_\varepsilon(\cdot, \cdot)$. In particular, equation (3.4) in Step 4 of Algorithm 3.2 is

$$2\alpha \varepsilon (\nabla S, \nabla W) + \gamma \left( S |\nabla U_{n+1}|^2, W \right) + \frac{\alpha}{2\varepsilon} (S - 1, W)_h = 0$$

for all $W \in V_h(\Omega)$. Identifying limits on a term by term basis would require identifying the limit

$$\lim_{n \to +\infty} \left( |\nabla U_{n+1}|^2 S_n, W \right),$$

which so far we have to leave as an open problem.

What is missing for this identification of limits is strong convergence of $\nabla U_n$ in $L^2$. This is a fundamental shortcoming also observed in [1, 5] for the simpler case of harmonic maps to the sphere. In fact, we are not aware of any algorithm, even in the harmonic mapping case, that simultaneously gives strong convergence of $\nabla U_n$ in $L^2$ and assures the sphere constraint exactly.

However, the algorithm converges, decreases the energy, assures the sphere constraint exactly and delivers very convincing computational results (indeed, it is faster and delivers better results than the alternative algorithm described in the sequel, c.f. Section 6).

4. Γ-Convergence for Penalisation & Splitting

In order to resolve the problems with passing to the limit, we now use a penalisation approach instead of projection. This requires adding a term to the Ambrosio-Tortorelli energy, which penalises the sphere constraint. In this section, we show that this addition does not affect Γ-convergence to the Mumford-Shah functional, if the penalisation term is properly scaled.

Let $\Omega \subset \mathbb{R}^d$, $\gamma, \alpha, \lambda$ be fixed positive constants, $\varepsilon, \delta \varepsilon > 0$, $k_\varepsilon \geq 0$, $g \in L^{\infty}(\Omega, S^{m-1})$, and $G_{\varepsilon}, G : L^2(\Omega, \mathbb{R}) \times L^2(\Omega) \to [0, +\infty]$ be defined by

$$G_{\varepsilon}(u, s) := \begin{cases} \frac{\gamma}{2} \int_{\Omega} (s^2 + k_\varepsilon) |\nabla u|^2 dx + \frac{\lambda}{2} \int_{\Omega} |u - g|^2 dx & \text{if } u \in H^1(\Omega, \mathbb{R}), \\ + \alpha \int_{\Omega} \left( \varepsilon |\nabla s|^2 + \frac{(1 - s)^2}{4\varepsilon} \right) dx & s \in H^1(\Omega, [0, 1]), \\ + \frac{1}{4\varepsilon} \int_{\Omega} (|u|^2 - 1)^2 dx & \text{otherwise}, \end{cases}$$

and

$$G(u, s) := \begin{cases} \frac{\gamma}{2} \int_{\Omega} |\nabla u|^2 dx + \frac{\alpha}{4}\nu^d - 1(S_u) + \frac{\lambda}{2} \int_{\Omega} |u - g|^2 dx & \text{if } u \in GSBV(\Omega, S^{m-1}), \\ + \infty, & \text{and } s = 1 \text{ a.e.} \end{cases}$$
Theorem 4.1. If $\Omega \subset \mathbb{R}^d$ is open and bounded with Lipschitz boundary, $\delta_\varepsilon \to 0$, $k_\varepsilon = o(\varepsilon)$, and $k_\varepsilon = o(\delta_\varepsilon)$, then $G_\varepsilon \frac{1}{\varepsilon} \to G$ in $L^2(\Omega, \mathbb{R}^m) \times L^2(\Omega)$.

Moreover, there exists a solution \( \{u_\varepsilon, s_\varepsilon\} \) to the minimum problem

$$m_\varepsilon = \inf_{u \in H^1(\Omega, \mathbb{R}^m), \ s \in H^1(\Omega, [0,1])} G_\varepsilon(u, s)$$

with $\|u_\varepsilon\|_{L^\infty} \leq C$, and every cluster point of \( \{u_\varepsilon\} \) is a solution to the minimum problem

$$m = \inf_{u \in GSBV(\Omega, \mathbb{R}^{m-1})} G(u, 1),$$

and $m_\varepsilon \to m$ as $\varepsilon \to 0^+$.

For the liminf inequality we can apply the work of Focardi ([32, Lemma 3.3]). For the limsup inequality we use the same construction as Ambrosio and Tortorelli in [3], so it is enough to verify that the penalisation term we added vanishes for $\varepsilon \to 0^+$. This is explained in more detail below.

Proof. For notational convention, we first localise the functionals above, denoting by $G_\varepsilon(u, s, A)$ and $G(u, s, A)$ the same functionals with integration over $A \subseteq \Omega$ instead of $\Omega$, and $H^{d-1}(S_u)$ replaced by $H^{d-1}(S_u \cap A)$.

**Step 1: The Liminf Inequality.**

Let $\varepsilon \to 0^+$, and $(u_\varepsilon, s_\varepsilon) \to (u, s)$ in $L^2(\Omega, \mathbb{R}^m) \times L^2(\Omega)$. Up to subsequences, we can suppose that $(u_\varepsilon, s_\varepsilon) \to (u, s)$ a.e., and that $\lim_{\varepsilon \to 0^+} G_\varepsilon(u_\varepsilon, s_\varepsilon)$ exists and is finite. We can further assume $s = 1$ a.e., since otherwise $\int_\Omega (1 - s_\varepsilon)^2 \, dx \to 0$, and $G_\varepsilon(u_\varepsilon, s_\varepsilon) \to \infty$. Similarly, we get $|u|^2 = 1$ a.e.

We now have to show

$$\liminf_{\varepsilon \to 0^+} G_\varepsilon(u_\varepsilon, s_\varepsilon) \geq G(u, s).$$

Since it is clear that $\int_\Omega |u_\varepsilon - g|^2 \, dx \to \int_\Omega |u - g|^2 \, dx$, and that the penalisation term is non-negative, it is sufficient to prove that $u \in GSBV(\Omega, \mathbb{R}^m)$, and

$$\liminf_{\varepsilon \to 0^+} \int_\Omega (s_\varepsilon^2 + k_\varepsilon) |\nabla u_\varepsilon|^2 \, dx + 2 \int_\Omega \varepsilon |\nabla s_\varepsilon|^2 + \frac{(1 - s_\varepsilon)^2}{4\varepsilon} \, dx \geq \int_\Omega |\nabla u|^2 \, dx + 2H^{d-1}(S_u).$$

This was shown for a more general situation in [32, Lemma 3.3] (see also [33]).

**Step 2: The Limsup Inequality.**

It suffices to consider the case $u \in SBV(\Omega, \mathbb{R}^m) \cap L^\infty(\Omega, \mathbb{R}^m)$. We can also assume $\nabla u \in L^2(\Omega, \mathbb{R}^{d \times m})$, $|u|^2 = 1$ a.e., and (see [33, Theorem 2.7.14]) that $S_u$ is essentially closed in $\Omega$; i.e., $H^{d-1}(\Omega \cap (S_u \setminus S_u)) = 0$. Setting $d(x) := \text{dist}(x, S_u)$, we define the Minkowski content of $S_u$

$$\mathcal{M}^{d-1}(S_u) := \lim_{\delta \to 0^+} \mathcal{M}^{d-1}_\delta(S_u) := \lim_{\delta \to 0^+} \frac{|\{x \in \Omega : d(x) < \delta\}|}{2\delta}.$$ 

It is well-known that for $S_u$ essentially closed,

$$\lim_{\delta \to 0^+} \mathcal{M}^{d-1}_\delta(S_u) = H^{d-1}(S_u) \tag{4.1}$$

(see [31, Theorem 3.2.39]). So, there exists a sequence $w_\varepsilon \to 0^+$, such that

$$|\{x \in \Omega : d(x) < \delta\}| \leq 2\delta \left( H^{d-1}(S_u) + w_\varepsilon \right), \tag{4.2}$$

for every $\delta \geq 0$ small enough.

Given such functions $u$, and $s = 1$ a.e., we have to construct $\{u_\varepsilon, s_\varepsilon\}$ that converge in $L^2(\Omega, \mathbb{R}^m) \times L^2(\Omega)$ to $(u, s)$, such that

$$\limsup_{\varepsilon \to 0^+} G_\varepsilon(u_\varepsilon, s_\varepsilon) \leq G(u, s)$$

for any positive sequence $\varepsilon$ converging to zero.

It is natural to require $s_\varepsilon \equiv 0$ in some $\varepsilon$-dependent neighbourhood of $S_u$, $s_\varepsilon$ converging to 1 everywhere outside a larger neighbourhood of $S_u$, and smooth in between, as well as $u_\varepsilon \equiv u$ everywhere outside some neighbourhood of $S_u$. 


To this end, we use the same construction as in the paper [3] by Ambrosio and Tortorelli: Choose a positive sequence $b_\varepsilon$, such that $b_\varepsilon = o(\varepsilon)$, $b_\varepsilon = o(\delta_\varepsilon)$, and $k_\varepsilon = o(b_\varepsilon)$. For any $b > 0$, set $S_b := \{x \in \Omega : d(x) < b\}$. Thanks to (4.2), $|S_b| = O(b)$. For $t \geq b_\varepsilon$, let

$$\sigma_\varepsilon(t) := 1 - \exp\left(-\frac{t - b_\varepsilon}{2\varepsilon}\right),$$

$$\sigma'_\varepsilon(t) = \frac{1}{2\varepsilon} \exp\left(-\frac{t - b_\varepsilon}{2\varepsilon}\right).$$

We now set (c.f. Figure 1)

$$s_\varepsilon(x) := \begin{cases} 0 & \text{if } x \in S_{b_\varepsilon}, \\ \sigma_\varepsilon(d(x)) & \text{if } x \in S_{b_\varepsilon+2\varepsilon \ln \frac{1}{\varepsilon}} \setminus S_{b_\varepsilon}, \\ 1 - \varepsilon & \text{if } x \in \Omega \setminus S_{b_\varepsilon+2\varepsilon \ln \frac{1}{\varepsilon}}, \end{cases}$$

and

$$u_\varepsilon(x) := u(x) \min\left\{\frac{d(x)}{b_\varepsilon}, 1\right\}.$$ 

Note that $0 < 2\varepsilon \ln \frac{1}{\varepsilon} \to 0^+$, and $\varepsilon = o(2\varepsilon \ln \frac{1}{\varepsilon})$.

![Figure 1. Sketch of $s_\varepsilon(x)$ in the case $S_u = \{0\}$, and $d = 1.$](image)

By construction, $(u_\varepsilon, s_\varepsilon) \to (u, 1)$ in $L^2(\Omega, \mathbb{R}^m) \times L^2(\Omega)$, as $\varepsilon \to 0^+$. Therefore, we have for the term penalising the sphere constraint,

$$\frac{1}{4\delta_\varepsilon} \int_{\Omega} \left|u_\varepsilon\right|^2 \, dx \leq c \frac{|S_{b_\varepsilon}|}{\delta_\varepsilon} \leq c \frac{b_\varepsilon}{\delta_\varepsilon} \to 0.$$

So this term does not contribute to the lim sup. This calculation motivates why we cannot expect good experimental results for $\delta_\varepsilon$ too small (compared to $b_\varepsilon$, which in turn is between $\varepsilon$ and $k_\varepsilon$); i.e., we have to sacrifice something in terms of the sphere constraint, c.f. our experiments in Section 6.2.

The other terms are just like in the original paper [3].

**Step 3: Convergence of Minimisers.**

The functional $G_\varepsilon$ is coercive and lower semicontinuous in $L^2$. So for every $\varepsilon > 0$ there exists a minimising pair $(u_\varepsilon, s_\varepsilon)$ of $G_\varepsilon$. By a simple truncation argument, $\|u_\varepsilon\|_{L^\infty} \leq C$. As above, we can assume that $(u_\varepsilon, s_\varepsilon) \in SBV(\Omega, \mathbb{R}^m) \times SBV(\Omega) \cap L^\infty(\Omega, \mathbb{R}^m) \times L^\infty(\Omega)$. By the SBV Closure and Compactness Theorems [2, Theorems 4.7 and 4.8], there exists a subsequence $\{u_{\varepsilon_j}, s_{\varepsilon_j}\}$ converging to some $(u, 1)$ in $L^2(\Omega, \mathbb{R}^m) \times L^2(\Omega)$, with $u \in SBV(\Omega, \mathbb{R}^m)$. Thus, the stability of minimising sequences under $\Gamma$-convergence (Lemma 2.1(3)) concludes the proof.

5. **Penalisation & Splitting Algorithm**

Let $\Omega \subset \mathbb{R}^d$, be a polyhedral Lipschitz domain, and let $g : \Omega \to S^{m-1}$ be the chromaticity component of a given image. For $u, g \in H^1(\Omega, \mathbb{R}^m)$, $s \in H^1(\Omega, [0,1])$, and $0 < \varepsilon, k_\varepsilon, \delta_\varepsilon \ll 1$, we want to minimise
the following vector valued Ambrosio-Tortorelli-Ginzburg-Landau energy using a splitting strategy:

\[ G_\varepsilon(u, s) = \frac{\gamma}{2} \int_\Omega (s^2 + k_\varepsilon) |\nabla u|^2 \, dx + \frac{\lambda}{2} \int_\Omega |u - g|^2 \, dx \\
+ \alpha \int_\varepsilon |\nabla s|^2 + \frac{1}{4\varepsilon^2} (1 - s)^2 \, dx + \frac{1}{4\delta\varepsilon} \int_\Omega |(u^2 - 1)|^2 \, dx. \]  

(5.1)

In this section, we shall always assume \( \gamma, \alpha, \varepsilon, k_\varepsilon, \delta_\varepsilon \) to be fixed and positive, \( \lambda \geq 0 \), and \( d \leq 2 \) (the last assumption is again only used to show that iterates \( S_n \in [-1, 1] \), and that their weak limit \( s \in [0, 1] \)).

**Definition 5.1.** A tuple \((u, s) \in H^1(\Omega, \mathbb{R}^m) \times H^1(\Omega, [0, 1])\) is called a weak solution to the problem \( \inf G_\varepsilon \), if and only if

\[ \gamma \left((s^2 + k_\varepsilon) \nabla u, \nabla \varphi\right) + \lambda \left(u - g, \varphi\right) + \frac{1}{\delta\varepsilon} \left((|u|^2 - 1) u, \varphi\right) = 0 \]

for all \( \varphi \in H^1(\Omega, \mathbb{R}^m) \), and

\[ 2\alpha \varepsilon \left(\nabla s, \nabla \varphi\right) + \gamma \left(|\nabla u|^2 s, \varphi\right) + \frac{\alpha}{2\varepsilon} (s - 1, \varphi) = 0 \]

for all \( \varphi \in H^1(\Omega) \cap L^\infty(\Omega) \).

We use the same finite element setting as in Section 3, in particular, we shall always assume the triangulation \( T_h \) to be quasi-uniform. For \( U, G \in V_h(\Omega, \mathbb{R}^m) \) and \( S \in V_h(\Omega, [-1, 1]) \), let

\[ G_{\varepsilon, h}(U, S) = \frac{\gamma}{2} \int_\Omega (S^2 + k_\varepsilon) |\nabla U|^2 \, dx + \frac{\lambda}{2} \int_\Omega |U - G|^2 \, dx \\
+ \alpha \int_\varepsilon |\nabla S|^2 + \frac{1}{4\varepsilon^2} T_h((1 - S)^2) \, dx + \frac{1}{4\delta\varepsilon} \int_\Omega |(U^2 - 1)|^2 \, dx. \]  

(5.4)

In the algorithm below we use \( G := r_h(g) \in V_h(\Omega, \mathbb{R}^m) \), i.e., the Clément interpolation of \( g \). This allows the use of non-smooth \( g \). If \( g \in C^0(\Omega, \mathbb{R}^m) \), the Lagrange interpolation would do as well.

**Algorithm 5.2.** Let \( U_0, G \in V_h(\Omega, \mathbb{R}^m) \) and \( S_0 \in V_h(\Omega) \) be given. For \( n = 1, \ldots \) until convergence do

1. Compute \( U_n \in V_h(\Omega, \mathbb{R}^m) \) such that

\[ \gamma \left((S_{n-1}^2 + k_\varepsilon) \nabla U_n, \nabla W\right) + \lambda \left(U_n - G, W\right) + \frac{1}{\delta\varepsilon} \left((|U_n|^2 - 1) U_n, W\right) = 0 \]

for all \( W \in V_h(\Omega, \mathbb{R}^m) \).

2. Compute \( S_n \in V_h(\Omega) \) such that

\[ 2\alpha \varepsilon \left(\nabla S_n, \nabla W\right) + \gamma \left(S_n |\nabla U_n|^2, W\right) + \frac{\alpha}{2\varepsilon} (S_n - 1, W) = 0 \]

for all \( W \in V_h(\Omega) \).

We start with a discussion of relevant stability properties of iterates from Algorithm 5.2.

**Lemma 5.3.** Algorithm 5.2 decreases \( G_{\varepsilon, h} \) with respect to \( n \in \mathbb{N} \).

**Proof.** For any \( n \in \mathbb{N} \) fixed, Algorithm 5.2 ensures that

\[ G_{\varepsilon, h}(U_{n+1}, S_{n+1}) \leq G_{\varepsilon, h}(U_n, S_n) \leq G_{\varepsilon, h}(U_{n+1}, S_n). \]

\[ \square \]

The following existence and uniqueness result follows by standard coercivity and convexity arguments for \( G_{\varepsilon, h} \) (see e.g. [34, Section 8.4]). The fact \(-1 \leq S \leq 1\) follows from Lemma 3.4.

**Proposition 5.4.** There exists a function \( U \in V_h(\Omega, \mathbb{R}^m) \), such that equation (5.5) holds for all \( W \in V_h(\Omega, \mathbb{R}^m) \), and a unique function \( S \in V_h(\Omega, [-1, 1]) \), such that equation (5.6) holds for all \( W \in V_h(\Omega) \).

Main convergence properties of iterates from Algorithm 5.2 are given in the following

**Theorem 5.5.** Let \( \{T_h\} \) be a sequence of quasi-uniform triangulations with maximal mesh size \( h_l \to 0 \) for \( l \to +\infty \), and \( G_{\varepsilon, h_l}(U_0, S_0) \leq C_0 < +\infty \) independently of \( h_l \). Then the sequences \( \{U^l_m, S^l_m\}_{m,l} \), constructed by Algorithm 5.2 from inputs \( (U_0, S_0) \) have a (diagonal) subsequence called \( \{U_n, S_n\}_n \), such that \( U_n \) converges strongly in \( H^1(\Omega, \mathbb{R}^m) \), and \( S_n \) converges weakly in \( H^1(\Omega) \) to some \((u, s) \in H^1(\Omega, \mathbb{R}^m) \times H^1(\Omega, [0, 1])\), which is a weak solution as in Definition 5.1.
For identifying limits in the proof of Theorem 5.5, it will be crucial to prove strong $L^2$ convergence of $\nabla U_n$ to $\nabla u$, for which we use a strategy derived from [15, Proof of Theorem 2], where the authors show convergence of two adaptive, stationary finite element approximations for the minimisation of the unconfined Ambrosio-Tortorelli energy: In Step 2 we show that $u$ fulfils equation (5.2), then we use equations (5.2) and (5.5) and dominated convergence (c.f. Lemma 5.6, also derived from [15]) to show strong $L^2$ convergence of $\nabla U_n$ to $\nabla u$ in Step 3, and finally we use this to show that $s$ fulfils equation (5.3) in Step 4.

**Lemma 5.6.** Let $p_n, p \in H^1(\Omega) \cap L^\infty(\Omega)$, such that $\|p_n\|_{L^\infty(\Omega)} \leq C < +\infty$ a.e., independently of $n$, and $p_n \rightharpoonup p$ in $L^2(\Omega)$. Then

$$\lim_{n} \left( \|p_n - p\|, |\nabla \varphi|^2 \right) = 0 \quad \forall \varphi \in H^1(\Omega, \mathbb{R}^m).$$

**Proof.** See [15, Proof of Theorem 2]. \hfill \square

**Proof of Theorem 5.5.**

**Step 1:** For $m, l \to \infty$, there is a subsequence $\{U_n, S_n\}$, converging weakly in $H^1(\Omega, \mathbb{R}^m) \times H^1(\Omega)$ to some $(u, s) \in H^1(\Omega, \mathbb{R}^m) \times H^1(\Omega, [-1, 1])$.

For every $m, l \in \mathbb{N}$, Proposition 5.4 gives existence of $(U_m^l, S_m^l)$ and ensures that $-1 \leq S_m^l \leq 1$ a.e. By Lemma 5.3 and by assumption,

$$G_{\epsilon, h_{\lambda}}(U_m^l, S_m^l) \leq G_{\epsilon, h_{\lambda}}(U_0^l, S_0^l) \leq C_0,$$

independently of $l, m$. In particular, $G_{\epsilon, h_{\lambda}}(U_n^l, S_n^l) \leq C_0$. So, by the definition of $G_{\epsilon, h_{\lambda}}$, the $H^1$-norms of $U_n^l$ and $S_n^l$ are bounded independently of $n$. Therefore, since $H^1$ is a Hilbert space, there exist subsequences, called $\{U_n^l\}$ and $\{S_n^l\}$, which converge weakly in $H^1$ to some $(u, s) \in H^1(\Omega, \mathbb{R}^m) \times H^1(\Omega)$.

Finally, since $H^1(\Omega)$ is a Hilbert space and $\{\varphi \in H^1(\Omega) : -1 \leq \varphi \leq 1 \text{ a.e.} \} \subset H^1(\Omega)$ is a closed, convex set, it is weakly closed. Therefore, by the weak convergence in $H^1$ of $S_n \rightharpoonup s$, we get $-1 \leq s \leq 1$.

Below, we shall use the abbreviation $h$ for $h_n$.

**Step 2: $u$ solves equation (5.2).**

Let $\varphi \in C^\infty(\Omega, \mathbb{R}^m)$ be fixed, $n \in \mathbb{N}$, and $h > 0$. Consider

$$\gamma \left( (s^2 + k_{\epsilon}) \nabla u, \nabla \varphi \right) + \lambda (u - g, \varphi) + \frac{1}{\delta_{\epsilon}} ((|u|^2 - 1) u, \varphi) =: \gamma T_1 + \lambda T_2 + \frac{1}{\delta_{\epsilon}} T_3.$$

Since $H^1$ is compactly embedded in $L^p$ for $p < 6$, as long as the space dimension $d \leq 3$, we have $U_n \rightharpoonup u$ in $L^p(\Omega, \mathbb{R}^m)$ for $p < 6$.

We compute

$$T_1 = ((S_{n-1}^2 + k_{\epsilon}) \nabla U_n, \nabla \varphi) + ((S_n^2 - S_{n-1}^2) \nabla U_n, \nabla \varphi) + (s^2 + k_{\epsilon}) \nabla (u - U_n), \nabla \varphi) + ((S_n^2 - S_{n-1}^2 + k_{\epsilon}) \nabla U_n, \nabla \varphi - \nabla (U_n)).$$

Note that $\|\varphi - \nabla (U_n)\|_{L^2(\Omega, \mathbb{R}^m)} \leq ch^{2-r} \|\nabla \varphi\|_{L^2(\Omega, \mathbb{R}^m)}$ for $0 \leq r \leq 2$.

Since $-1 \leq S_{n-1}, s \leq 1$, we have $|S_{n-1}^2 - s^2| \leq C |S_{n-1} - s|^{1/2},$ whence, by Lemma 5.6,

$$T_{12} = ((s^2 - S_{n-1}^2) \nabla U_n, \nabla \varphi) \leq C (|s - S_{n-1}|, |\nabla \varphi|^2)^{1/2} \|\nabla U_n\|_{L^2(\Omega, \mathbb{R}^m)} \to 0 \text{ as } n \to +\infty.$$

Since $s \leq 1$, we know that $|s^2 + k_{\epsilon}| \nabla \varphi \in L^2(\Omega, \mathbb{R}^{d \times m}),$ so $T_{13} = ((\nabla (u - U_n), (s^2 + k_{\epsilon}) \nabla \varphi) \to 0,$ by weak convergence. And since $\|\varphi - \nabla S_n\|_{L^2(\Omega, \mathbb{R}^m)} \to 0$, using the bounds established in Step 1, the terms $T_{14}, T_2, T_3$ all clearly vanish.

Putting all of the above together, we have for $n \in \mathbb{N}$ and $h > 0$ fixed,

$$\gamma \left( (s^2 + k_{\epsilon}) \nabla u, \nabla \varphi \right) + \lambda (u - g, \varphi) + \frac{1}{\delta_{\epsilon}} ((|u|^2 - 1) u, \varphi) =: \gamma T_{14} + \lambda T_{21} + \frac{1}{\delta_{\epsilon}} T_{31} + T_n,$$

where $\gamma T_{21} + \lambda T_{31} + \frac{1}{\delta_{\epsilon}} T_{31} = 0$ by construction. Now, letting $n \to +\infty$ and $h \to 0$, we have $T_n \to 0$, as shown above. And by a density argument, the above is true for general $\varphi \in H^1(\Omega, \mathbb{R}^m)$.

**Step 3:** $\nabla U_n \rightharpoonup \nabla u$ strongly in $L^2(\Omega, \mathbb{R}^{m \times d})$, as $n \to +\infty$ and $h \to 0$. 13
Let \( n \in \mathbb{N} \) and \( h > 0 \). Then
\[
\gamma k \varepsilon \| \nabla (u - U_n) \|_{L^2, X}^2 \\
\leq \gamma \left( (S_{n-1}^2 + k \varepsilon) \nabla u, \nabla (R_h(u) - U_n) \right) \\
- \gamma \left( (S_{n-1}^2 + k \varepsilon) \nabla U_n, \nabla (R_h(u) - U_n) \right) \\
- \lambda (U_n - G, R_h(u) - U_n) - \frac{1}{\delta \varepsilon} \left( (|u|^2 - 1) U_n, R_h(u) - U_n \right) \\
+ \lambda (U_n - G, R_h(u) - U_n) + \frac{1}{\delta \varepsilon} \left( (|u|^2 - 1) U_n, R_h(u) - U_n \right) \\
\geq \left( (S_{n-1}^2 + k \varepsilon) \nabla (u - U_n), \nabla (u - R_h(u)) \right)
\]
\( =: T_1^n + \ldots + T_{16}^n. \)

By construction (equation (5.5) with \( W := R_h(u) - U_n \)), the expression \( T_2^n + T_3^n + T_4^n \) is zero.

\[
T_1^n = \gamma \left( (S_{n-1}^2 + s^2) \nabla u, \nabla (R_h(u) - U_n) \right) + \gamma \left( (s^2 + k \varepsilon) \nabla u, \nabla (R_h(u) - U_n) \right) \\
+ \lambda (u - g, R_h(u) - U_n) + \frac{1}{\delta \varepsilon} \left( (|u|^2 - 1) u, R_h(u) - U_n \right) \\
- \lambda (u - g, R_h(u) - U_n) - \frac{1}{\delta \varepsilon} \left( (|u|^2 - 1) u, R_h(u) - U_n \right)
\]
\( =: T_1^{n_1} + \ldots + T_{16}^{n_1}. \)

By Step 2, \( T_{12}^n + T_{15}^n + T_{14}^n = 0. \) Therefore
\[
\gamma k \varepsilon \| \nabla (u - U_n) \|_{L^2, X}^2 \leq T_1^{n_1} + T_{15}^n + T_{16}^n + T_5^n + T_6^n + T_7^n.
\]

All of the above is true for any \( n \in \mathbb{N} \). Now, consider the limit \( n \to +\infty \) and \( h \to 0 \). Note that, by a density-argument,
\[
\| R_h(u) - U_n \|_X \leq \| R_h(u) - u \|_X + \| u - U_n \|_X \xrightarrow{h \to 0} 0,
\]
for \( X = H^1 \) and, by embedding, \( X = L^p \) \((p < 6). \) Therefore we have, similarly to Step 2, that the terms \( T_3^n, T_4^n, T_7^n, T_{15}^n \), and \( T_{16}^n \) all vanish in the limit \( h \to 0 \) and \( n \to +\infty. \) Finally, \( T_{11}^n \) vanishes using Lemma 5.6, as in Step 2, and the \( H^1 \)-stability of the Ritz projection.

**Step 4:** \( s \) solves equation (5.3), and \( 0 \leq s \leq 1. \)

Let \( \varphi \in C^\infty(\Omega) \) be fixed, \( n \in \mathbb{N} \), and \( h > 0. \) Set
\[
2 \alpha \varepsilon (\nabla s, \nabla \varphi) + \gamma (|\nabla u|^2 s, \varphi) + \frac{\alpha}{2 \varepsilon} (s - 1, \varphi) := 2 \alpha \varepsilon T_1 + \gamma T_2 + \frac{\alpha}{2 \varepsilon} T_3.
\]

We have
\[
T_1 = (\nabla S_n, \nabla I_h(\varphi)) + (\nabla S_n, \nabla (\varphi - I_h(\varphi))) + (\nabla (s - S_n), \nabla \varphi)
\]
\( =: T_1^n + T_{12}^n + T_{13}^n, \)

with \( T_{12}^n, T_{13}^n \to 0 \) by the strong \( H^1 \) convergence of \( I_h(\cdot) \) and the weak \( H^1 \) convergence of \( S_n \), respectively, like in Step 2.

Also,
\[
T_2 = (|\nabla U_n|^2 S_n, I_h(\varphi)) + (|\nabla U_n|^2 S_n, \varphi - I_h(\varphi))
\]
\( =: T_{21}^n + T_{22}^n + T_{23}^n + T_{24}^n, \)

with \( T_{22}^n, T_{23}^n, T_{24}^n \to 0 \) by the properties of the Lagrange interpolation, Step 3, and Lemma 5.6, respectively.

Finally,
\[
T_3 = (S_n - 1, I_h(\varphi)) + (S_n - 1, I_h(\varphi)) - (S_n - 1, I_h(\varphi) - (S_n - 1, I_h(\varphi))
\]
\( =: T_{31}^n + \ldots + T_{35}^n, \)

with \( |T_{32}^n + T_{33}^n| \leq C h \| \nabla S_n \|_{L^\infty(\Omega)} \| I_h(\varphi) \|_{L^\infty(\Omega)} \to 0, \) and \( T_{34}^n, T_{35}^n \to 0 \) by the strong \( L^p \) convergence of \( I_h(\cdot) \) and \( S_n \), respectively.
So, putting all of the above together, we have for \( n \in \mathbb{N} \) and \( h > 0 \) fixed,
\[
2\alpha \varepsilon (\nabla s, \nabla \varphi) + \gamma (|\nabla u|^2 s, \varphi) + \frac{\alpha}{2\varepsilon} (s - 1, \varphi) =: 2\alpha \varepsilon T^n_{11} + \gamma T^n_{21} + \frac{\alpha}{2\varepsilon} T^n_{31} + T^n,
\]
where \( 2\alpha \varepsilon T^n_{11} + \gamma T^n_{21} + \frac{\alpha}{2\varepsilon} T^n_{31} = 0 \) by construction. Now, letting \( n \to +\infty \) and \( h \to 0 \), we get \( T^n \to 0 \), as shown above.

By a density argument, \( s \) solves equation (5.3) for \( \varphi \in H^1(\Omega) \cap L^\infty(\Omega) \). And since replacing \( s \) pointwise by \( 0 \leq s \leq 1 \) would only decrease every term of this energy, \( 0 \leq s \leq 1 \) follows.

\[\square\]

**Remark 5.7.** For \( d \leq 2 \), one can also get \( \nabla S_n \to \nabla s \) strongly in \( L^2(\Omega; \mathbb{R}^m) \), with an argument similar to Step 3, using the equations for \( S_n \) and \( s \) and a test function \( R_h(s) - S_n \). It breaks down for \( d \geq 3 \) because of the lack of \( L^\infty \)-stability of the Ritz projection.

6. Computational Studies

To implement Algorithm 5.2, we use a simple fixed-point strategy (with 3 iterations) for the Ginzburg-Landau term.

To process real images, we suggest to amend Ambrosio and Tortorelli’s energy to \( AT_z(u, v, s) : H^1(\Omega, S^{m-1}) \times H^1(\Omega) \times H^1(\Omega) \to [0, +\infty] \)
\[
AT_z(u, v, s) := \frac{\gamma}{2} \int_{\Omega} (s^2 + k_z) |\nabla u|^2 \, dx + \frac{\lambda}{2} \int_{\Omega} |u - g|^2 \, dx
+ \frac{\lambda_1}{2} \int_{\Omega} |v - b|^2 \, dx
+ \alpha \int_{\Omega} \varepsilon |\nabla s|^2 + \frac{1}{4\varepsilon} (1 - s)^2 \, dx,
\]
with \( \gamma, \gamma_1, \alpha, \lambda, \lambda_1 \) positive constants and \( b, v \in L^\infty(\Omega) \cap H^1(\Omega) \) the brightness component of the original and the processed image, respectively (normalised to lie in \([0,1]\)). So, we add a smoothing and a fidelity term for the brightness component in the second line of (6.1). The idea here is that the smoothing term for the chromaticity component forces \( |s| \) to be small whenever \( |\nabla u| \) is large, while the smoothing term for the brightness component does the same whenever \( |\nabla v| \) is large. So we expect \( |s \approx 0| \) to approximate the union of the essential jump sets of the chromaticity and the brightness component.

This necessitates the adaptation of the optimisation problem for \( s \) as well as the solution of a third optimisation problem, which we place between the two existing ones.

If we process an image with more noise in the chromaticity as in the brightness, as is usually the case with images from digital cameras, we can now choose to give more weight to the information of the brightness component through the joint edge set, as illustrated in Example 5.

6.1. Academic Images, Splitting & Projection. All arrows below are scaled in length to fit the plots. What we call \( h \) below is the length of the two shorter sides of the rectangular triangles in our triangulations; i.e., it is shorter than the actual diameter of the triangles (by a factor of \( \sqrt{2} \)).

**Example 1.** Let \( \Omega := (0,1)^2 \) and \( G \) as in the left plot in Figure 2. The right picture shows a section along \( x = 0.5 \), where the \( z \)-values of the two regions are the closest. We use a triangulation consisting of \( 2^{2+8} \) halved squares (along the direction \((1,1)) \); i.e., 131072 triangles, with 66049 nodes, and \( h = 2^{-8} \approx 4 \times 10^{-3} \). The initial values for \( U \) and \( S \) are \( U_0 \equiv G \) and \( S_0 \equiv 0.5 \), respectively. We choose \( \gamma = 1.2 \), \( \alpha = 0.5 \), \( \lambda = 2 \times 10^3 \), \( \varepsilon = 6 \times 10^{-4} \), and \( k_z = 10^{-6} \) (parameters chosen by experiment).

Figure 2 shows the initial values, Figure 3 the result after 10 iterations of our proposed algorithm. Figure 4 shows the detected edge set and Figure 5 the Ambrosio-Tortorelli energy over time.

The next example numerically studies blowup behaviour for the \( W^{1,\infty} \)-norm of iterates \( \{U_n, S_n\} \) in the absence of a fidelity term; i.e., \( \lambda = 0 \). This is motivated by blowup results for harmonic maps (to the sphere), see e.g. [46, 47, 48, 49, 37, 6]. In particular, it is known that for \( d = 2 \), singularities only appear for large initial energy. And any harmonic map (for general \( d \)) is smooth outside a set whose \((d-2)\)-dimensional Hausdorff measure is zero, see [45, 46, 38, 29, 9, 41, 40].

**Example 2.** Let \( \Omega \) be as above. We first use a triangulation consisting of \( 2^{2+r} \), \( r = 8 \) halved squares as above, and later use coarser ones \( \{r \in \{5, \ldots, 8\}\} \) for comparison. Let \( \gamma = 1 = \alpha, \lambda = 0, \varepsilon = h/6 \), and \( k_z = 10^{-6} \). We use two sets of initial data for \( U \) and \( S \), which are shown in Figures 6 and 9 (leftmost column). In both cases, \( U_0 \) is constantly \((0,0,1)\) in the periphery of the image, \((0,0,-1)\) at the centre,
and varying continuously inside a circle around the centre. In the first case, we choose $S = 0$ at the centre, $S = 1$ in the periphery, and smoothly varying in between; in the second case, we choose $S = 1$ at the centre, $S = 0$ in the periphery, and smoothly varying in between.

Figure 6 shows iterates $n \in \{0, 3, 5\}$ for $r = 8$ (crops in the case of $U_n$). Figure 7 shows the total energy for $r \in \{5, \ldots, 8\}$, while Figure 8 shows the $W^{1, \infty}$-norms of $U_n$ and $S_n$ for $r \in \{5, \ldots, 8\}$, which both show blowup behaviour. This time it is $U_n$ which appears one step ahead of $S_n$ with respect to blowup behaviour. Depending somewhat on $r$, the system matrices become close to singular after 6–7 iterations, so after this point, the results can no longer be expected to be reliable. The arrow at the centre of $U$ at this point still points down, while the rest of $U$ points up. The variable $S$, on the other hand, becomes 1 everywhere, except for the centre, where it stays 0. After breakdown, the arrows move erratically, but perfectly synchronised with one another.
The next example uses the same setting and the same initial data for $U$, but avoids blowup behaviour through a different choice of initial data for $S$.

**Example 3.** Except for the initial data for $S$ we use exactly the same setting as in Example 2. This time, we choose $S = 1$ at the centre, $S = 0$ in the periphery, and smoothly varying in between.

Figure 9 shows iterates $n \in \{0, 3, 6\}$ for $r = 8$ (crops in the case of $U_n$), Figure 10 shows the total energy for $r \in \{5, \ldots, 8\}$, while Figure 11 shows the $W^{1, \infty}$-norms of $U_n$ and $S_n$ for $r \in \{5, \ldots, 8\}$, which this time stay finite. The arrows of $U$ all point down at the end, while $S$ is 1 everywhere. After
6 iterations, the system matrices again become close to singular; in this case, however, iterates do not change dramatically after this point, if at all.

6.2. Academic Images, Penalisation & Splitting. The next example studies the same setting as Example 1, this time with Algorithm 5.2; i.e., the sphere constraint is enforced by penalisation instead of projection. Again, all arrows are scaled in length to fit the plots.

**Example 4.** The setting is as in Example 1. Parameters are $\gamma = 1.2$, $\alpha = 0.5$, $\lambda = 2 \times 10^3$, $\varepsilon = 10^{-3}$, $k_z = 10^{-6}$, and $\delta_z = 0.1$ (chosen by experiment).
Figure 11. Example 3: $W^{1,\infty}$-norm of $U$ and $S$ for $r \in \{5, \ldots, 8\}$.

Figure 12. Example 4: Edge set (left) and horizontal section through it ($y = 0.375$, right) after 10 iterations.

Figure 13. Example 4: min and max of $|U|$ (left), and Ambrosio-Tortorelli Energy (right) for 10 iterations.

The result $U$ after 10 iterations looks just like in Example 1 (Figure 3), so we omit the corresponding figures. The detected edge set after 10 iterations, however, is less exact, as shown in Figure 12. Figure 13 shows the global minimum and maximum of $|U|$ and the Ambrosio-Tortorelli energy over time.

For $\delta_\varepsilon$ between about $5 \times 10^{-3}$ and at least $10^2$, the results are qualitatively very similar to the ones in Example 1, but the detected edge set is less exact, and $|U|$ can be quite a bit shorter than 1. For $\delta_\varepsilon$ smaller than $5 \times 10^{-3}$ (which would be advantageous for the accuracy of $|U|$), the results break down, which is in accordance with our theoretical results.

Example 5. We try our algorithm on a small photograph (399 × 299 pixels), as shown in Figure 14. We choose \( \Omega := (0, 399/299) \times (0, 1) \), whence \( h = 1/298 \approx 3 \times 10^{-3} \), the pixels are used as nodes, each square of 4 pixels giving rise to two triangles. We further choose \( S_0 \equiv 1 \) and add two different kinds of noise to the image:

1. RGB noise: \( R = R_0 + 0.3 \times \text{randn} \), and \( G \) and \( B \) analogously, where \( \text{randn} \) are pseudo-random values drawn from the standard normal distribution. After this operation, we crop \( R, G, \) and \( B \) to lie in \([0, 1]\) (where \( R_0, G_0, B_0 \) were scaled to lie). This is shown in Figure 14.

2. CB noise, mainly in the chromaticity component: \( C = C_0 + 0.5 \times \text{randn} \times [1, 1, 1] \in \mathbb{S}^2 \), and \( B = B_0 + 0.01 \times \text{randn} \). After this operation, \( C \) is projected onto the sphere, and \( B \) is cropped to lie in \([0, 1]\). This is shown in Figure 18.

Our CB algorithm was in both cases compared to a channelwise RGB computation for the same image, with all channels sharing the same edge set. Parameters were chosen as follows (by experiment):

1. RGB computation: \( \alpha = 0.3, \beta = 10^{-2}, \gamma = 10^3, \varepsilon = 10^{-4}, \) and \( k_\varepsilon = 10^{-7} \).
   CB computation: \( \alpha = \alpha_1 = 0.5, \beta = 8 \times 10^{-3}, \gamma = \gamma_1 = 10^3, \varepsilon = 10^{-4}, \) and \( k_\varepsilon = 10^{-7} \).

2. RGB computation: \( \alpha = 0.5, \beta = 5 \times 10^{-3}, \gamma = 50, \varepsilon = 10^{-4}, \) and \( k_\varepsilon = 10^{-7} \).
   CB computation: \( \alpha = \alpha_1 = 0.3, \beta = 5 \times 10^{-2}, \gamma = 10^2, \gamma_1 = 5 \times 10^5, \varepsilon = 10^{-4}, \) and \( k_\varepsilon = 10^{-7} \).

Figure 14. Example 5.1: Original image (left) and image with RGB noise (right).

Figure 15. Example 5.1: Image after 10 iterations, RGB (left) and CB (right).

First, let us look at the computations with RGB noise: Figure 14 shows the noisy initial image, and Figure 15 the results after 10 iterations. Figure 16 shows the detected edge sets, and Figure 17 the expanded Ambrosio-Tortorelli energy over time. The energy terms labelled “...C” belong to the chromaticity component, those labelled “...B” to the brightness. The channelwise RGB algorithm has the advantage here.

Next, let us look at the image with CB noise: Figure 18 shows the noisy initial image, and Figure 19 the results after 10 iterations. Figure 20 shows the detected edge sets, and Figure 21 the expanded Ambrosio-Tortorelli energy over time. The CB algorithm has a very clear advantage here.
Figure 16. Example 5.1: Edge set, RGB (left) and CB (right).

Figure 17. Example 5.1: Expanded Ambrosio-Tortorelli Energy (10 iterations, y-logarithmic plots), RGB (left) and CB (right).

Figure 18. Example 5.2: Original image and image with CB noise (top), as well as noisy chromaticity (bottom left) and brightness (bottom right) components.
Figure 19. Example 5.2: Image after 10 iterations, RGB (left) and CB (right).

Figure 20. Example 5.2: Edge set, RGB (left) and CB (right).

Figure 21. Example 5.2: Expanded Ambrosio-Tortorelli Energy (10 iterations, $y$-logarithmic plots), RGB (left) and CB (right).

Acknowledgements
The research presented in this paper was done as part of the author’s PhD thesis [36] under the supervision of Andreas Prohl (U Tübingen). Further thanks go to Sören Bartels (U Bonn) for his help with the coding, and to Giovanni Bellettini (U Roma), Ludwig Gauckler (U Tübingen), Christoph Ortner (U Oxford), Reiner Schätzle (U Tübingen), and Markus Schmuck (MIT Cambridge, MA) for helpful discussions.

References


Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle 10, D-72076 Tübingen, Germany
E-mail address: haehnle@na.uni-tuebingen.de