Image reconstruction by Mumford-Shah regularization with a priori edge information

Thomas Sebastian Page

Image reconstruction by Mumford-Shah regularization with *a priori* edge information

von

Thomas Sebastian Page

Dissertation

zur Erlangung des Grades eines Doktors der Naturwissenschaften
– Dr. rer. nat. –

Vorgelegt im Fachbereich 3 (Mathematik & Informatik)
der Universität Bremen
26. Januar 2015

Datum des Promotionskolloquiums: 26. Februar 2015

Gutachter: Prof. Dr. Peter Maaß (Universität Bremen)
Prof. Dr. Ming Jiang (Peking University)

Acknowledgements

I thank Prof. Dr. Peter Maaß and Prof. Dr. Ming Jiang for their supervision, insight and valuable input throughout my graduate studies. In 2010 Prof. Dr. Ming Jiang introduced me to the interesting field of regularization with the Mumford-Shah functional and later to the problem of multimodal imaging. The combination of both was the starting point for this work. I owe him my gratitude for his help and thoughtful advice throughout the last years.

The work presented in this thesis was done to a large part during my stays at Peking University. I wholeheartedly thank Prof. Dr. Peter Maaß and Prof. Dr. Ming Jiang for giving me the opportunity and support to be able to work and live in Bremen and Beijing. I am grateful for the International Office of Peking University for their help when visiting.

During my time as a visiting student to Peking University I was partially supported by the National Basic Research Program of China (973 Program) (2011CB809105), the National Science Foundation of China (61121002, 60325101), the German Academic Exchange Service and the Graduate Center University of Bremen. I want to thank all the above organizations and people that helped me during the application processes.

I want to thank Dr. Robin Strehlow, Dr. Andreas Bartels, Patrick Dülk, Simon Grützner, Delf Lachmund and Simone Bökenheide for proof reading my manuscript. This work was greatly improved by their corrections and comments.

Most importantly, I would like to thank my family for their support.

Abstract

The Mumford-Shah functional has provided an important approach for image denoising and segmentation. Recently, it has been applied to image reconstruction in fields such as X-ray tomography and electric impedance tomography. In this thesis we study the applicability of the Mumford-Shah model to a setting, where *a priori* edge information is available and reliable. Such a situation occurs for example in biomedical imaging, where multimodal imaging systems have received a lot of interest.

The regularization terms in the Mumford-Shah functional force smoothness of the image within individual regions and simultaneously detect edges across which smoothing is prevented. We propose to divide the edge penalty into two parts depending on the *a priori* edge information. We investigate the proposed model for well-posedness and regularization properties under an assumption of pointwise boundedness of the underlying image.

Furthermore, we present two variational approximations that allow numerical implementations. For one we prove that it Γ-converges to a special case of our proposed model, the other we motivate heuristically. The resulting algorithm alternates between an image reconstruction and an image evaluation step. We illustrate the feasibility with two numerical examples.

Contents

Chapter 1

Introduction

In this thesis, we study a novel approach to image reconstruction from ill-posed operator equations, using *a priori* edge knowledge. The proposed model is based on the well-known Mumford-Shah regularization for image denoising and segmentation.

In this introduction we will first state the underlying motivation and then describe our proposed approach. Finally, we describe our contributions and give an outline of the thesis.

1.1 Multimodal image reconstruction

The underlying motivation to our work stems from multimodal image reconstruction in medical imaging. This field has gained considerable interest in recent years and is still rapidly developing Ehrhardt et al. [2014]; Kazantsev et al. [2014]; Leahy and Yan [1991]; Schweiger and Arridge [1999]; Townsend [2008]; Vauhkonen et al. [1998].

Over the last decades, a number of imaging modalities have emerged in the field of biomedical imaging. For example there are well established procedures, such as X-ray computed tomography (X-ray CT) or magnetic resonance imaging (MRI), that can visualize anatomical information with a high resolution and newer methods, such as diffuse optical tomography (DOT), positron emission tomography (PET) or electric impedance tomography (EIT), which are capable of visualizing chemical or biological processes, but have a poor resolution, see Bushberg and Boone [2011]; Intes [2008]. The natural observation that motivates this thesis is that images of the same object, but obtained from different modalities, possess similar complementary feature information.

A multimodal imaging approach treats the reconstruction from two or more imaging modalities as a combined task, rather than solving each imaging problem individually, see Intes [2008]. The terminology covers both innovations on the hardware side, where scanners have been developed that can acquire data from different modalities, either sequentially or simultaneously Cherry [2006], and innovations on the algorithmic side, where methods have been developed making use of the cross modal information in the inversion of the data, see for example Ehrhardt et al. [2014]; Somayajula et al. [2005]. Bi-modal examples are PET/MRI and PET/X-ray CT that are also in commercial use, where PET gives functional and MRI or X-ray CT yield anatomical information.

In applications, the measured data is almost always incomplete or inaccurate, due to limitations for example in the measurement geometry or accuracy. Furthermore, indirect imaging problems lead to inverse problems, that are typically ill-posed, see Engl et al. [1996]. In light of the ill-posedness and corrupted measured data it is therefore desirable to make use of all available information from different methods in a complementary manner to narrow the solution space.

Steps for a multimodal imaging approach are to identify the features that are shared across the considered modalities, model these features and design algorithms that can make use of the additional complementary information. For many applications, such as PET/X-ray CT and DOT/X-ray CT, the edges of the underlying images are correlated as different anatomical regions usually also present different functional information. We therefore want to use edges as the connecting feature, see also Kazantsev et al. [2014]; Leahy and Yan [1991]; Schweiger and Arridge [1999]. Other possible features are information theoretic similarities as in Somayajula et al. [2005].

In this work we consider a variational approach to bimodal image reconstruction. In the following section we present the model we base our approach on.

1.2 The Mumford-Shah model for image segmentation and denoising

In the following, we consider an image as a function on a two dimensional bounded domain $\Omega \subset \mathbb{R}^2$.

Computing a segmentation of an image is a key step in image processing. The segmentation problem for an image $g \in L^\infty(\Omega)$ can be defined as finding a decomposition

$$\Omega = R_1 \cup R_2 \cup \ldots \cup R_l \cup K$$

of Ω, where $R_i \subset \Omega$ are disjoint connected open subsets and K is the union of the boundaries of R_i in Ω. We consider a decomposition meaningful if

1. the image g varies smoothly and/or slowly within each R_i,
2. the image g varies discontinuously and/or rapidly across K between different R_i.

A segmentation is often used as a starting point for further analysis and thus plays a key role in image processing.

Inspired by the discrete Gibbs energy of Geman and Geman [1984], in Mumford and Shah [1989] a variational approach to the image segmentation problem was introduced. They proposed to minimize a functional (see equation (1.1)) with the aim to find a piecewise smooth approximation of an image and also detect its edges. Since its introduction the Mumford-Shah model has received a lot of attention, see for example Ambrosio et al. [2000]; David [2005]; Fusco [2003].

To a given noisy image $g \in L^\infty(\Omega)$, the ***Mumford-Shah functional*** is defined as

$$\mathrm{MS}(f, K) := \int_\Omega |f - g|^2 dx + \alpha \int_{\Omega \setminus K} |\nabla f|^2 dx + \beta \mathcal{H}^1(K) \tag{1.1}$$

for every closed subset $K \subset \Omega$, every $f \in W^{1,2}(\Omega \setminus K)$ and positive parameters α and β. The one dimensional Hausdorff measure $\mathcal{H}^1(K)$ measures a generalized length of K, see Section 2.2 for the definition. The last term forces K to be a 1-dimensional rectifiable curve for a minimizing pair (f, K).

A minimizer (f, K) of the Mumford-Shah functional (1.1) must balance three requirements each coming from one of the terms:

1. f must be a good approximation of g in the L_2 norm,
2. f must be smooth everywhere in Ω except at the edges K,
3. the edges K must be as "short" as possible.

The minimizer (f, K) can be understood as a simplification of the original picture g. Regions in $\Omega \setminus K$ are drawn smoothly and details are lost, but main objects are sharply marked through K. In this thesis we consider a pair (f, K) fulfilling the above described requirements as a "good segmentation", i.e. the smaller $\mathrm{MS}(f, K)$ the better.

The Mumford-Shah functional has a natural generalization for signals on N-dimensional domains $\Omega \subset \mathbb{R}^N$. Instead of penalizing the length of an edge, in dimension N the discontinuity set K is penalized by its $N-1$ dimensional Hausdorff measure $\mathcal{H}^{N-1}(K)$. As a result, for minimizers (f, K) the set K essentially is a $N-1$ dimensional subset of Ω (in the sense of Hausdorff dimension).

1.3 Regularizing with the Mumford-Shah functional

The Mumford-Shah functional and its variants have been applied to many imaging applications. Examples are electric impedance tomography Rondi and Santosa [2001], image inpainting Esedoglu and Shen [2002], image deblurring Bar et al. [2006], X-ray tomography Ramlau and Ring [2007], electron tomography Klann [2011] and SPECT Klann and Ramlau [2013].

Let $A : X \to Y$ be a forward operator of an imaging application mapping from its image space X to its data space Y (both spaces suitably defined). The

Mumford-Shah functional with corresponding least squares fidelity term then is

$$\mathrm{MS}(f, K) := \|A(f) - g\|_Y^2 + \alpha \int_{\Omega \setminus K} |\nabla f|^2 dx + \beta \mathcal{H}^{N-1}(K). \tag{1.2}$$

With a non-trivial operator A, the original existence and regularity theory cannot be applied directly. In fact, the functional (1.2) may not have a minimizer without extra constraint, as shown in Fornasier et al. [2011] for the image deblurring problem. Therefore it is necessary to introduce reasonable constraints on the image f, edge set K or forward operator A to establish existence results, see Rondi [2007, 2008a,b]; Rondi and Santosa [2001]. In Rondi [2008b], it was shown that for electric impedance tomography and for certain linear forward operators under a pointwise boundedness constraint on f, the Mumford-Shah functional yields a regularization on the image. In Jiang et al. [2014] it was shown that under the same pointwise boundedness constraint the Mumford-Shah functional yields a regularization for both the image and edge for an *a priori* parameter choice rule. When the image f is restricted to piecewise constants, existence and other regularization properties have also been established by Klann and Ramlau [2013]; Ramlau and Ring [2010]. In Fornasier et al. [2011] it was shown that under an additional regularity constraint on K, the Mumford-Shah functional yields a minimizer for linear forward operators, provided they are either compact and injective or first-order differential operators.

1.4 Image reconstruction with the Mumford-Shah model using *a priori* edge information

We now present the model we propose to bimodal imaging based on the Mumford-Shah model.

We study one specific case of feature based bimodal image reconstruction, where one modality is severely more ill-posed than the other. Examples for such a setting are PET/X-ray CT and DOT/X-ray CT , where DOT and PET are the more ill-posed problems.

Because of this asymmetrical setting we use the feature similarity only in one direction, which is from the less to the more ill-posed problem. We will presume that the less ill-posed problem is already solved and its edge set $K^0 \subset \Omega$ is available and reliable. Our approach is to use the *a priori* edge knowledge of K^0 in the reconstruction of the more ill-posed problem.

We propose to separate the edge penalty $\mathcal{H}^{N-1}(K)$ into two parts depending on K^0 leading to the ***Mumford-Shah type functional*** with *a priori* edge knowledge

$$MS(f,K)_{K^0} = \|A(f) - g\|_Y^2 + \alpha \int_{\Omega \setminus K} |\nabla f|^2 dx \\ + \beta \mathcal{H}^{N-1}(K \setminus K^0) + \gamma \mathcal{H}^{N-1}(K \cap K^0), \tag{1.3}$$

with parameters $0 \leq \gamma < \beta$ and $0 < \alpha$.

In the modified functional (1.3), edges K that coincide with *a priori* edges K^0 are penalized less (with factor γ), than edges not included in K^0 (with factor β). Thus, edges coinciding with K^0 are more likely to be reconstructed. In the early work Leahy and Yan [1991] a closely related approach based on the discrete model of Geman and Geman [1984] was proposed for coupled PET/MRI.

In Chapter 3 we will study the proposed functional with regard to existence of a minimizer, stability and parameter choice rules.

1.5 Variational approximation in the sense of Γ-convergence

In applying the Mumford-Shah regularization to practical applications, several issues arise. The primary difficulty comes from the edge part because it is not easy to represent in programming and to trace its updates. One solution is to use the level-set method Chan and Vese [2001]. Another approach is based on the Γ-convergence theory by approximating the edge set with smooth indicator functions Ambrosio and Tortorelli [1992].

In this work we study the latter approach for the above described Mumford-Shah functional with *a priori* edge information (1.3). The aim is to define a sequence of regular functionals, that on the one hand can easily be implemented, and on the other hand yield minimizers that approximate solutions of the original problem. A suitable notion of convergence for such variational approximations is Γ-convergence, see Braides [2002].

In Chapter 4 we follow the Γ-convergence approximation of the Mumford-Shah functional with elliptic functionals from Ambrosio and Tortorelli [1992]. In Ambrosio and Tortorelli [1992] a phase field approach, where the edges are approximated by smooth indicator functions $v \in W^{1,2}(\Omega), 0 \leq v \leq 1$, was proposed. There $v \approx 0$ indicates the presence and $v \approx 1$ the absence of an edge.

There are several advantages in choosing a phase field approach to model edges. Firstly, it gives a non-parametric global description of edges with which topological changes, such as introducing new edges, do not require extra effort. Secondly, the edges are available in a format that is easy to access and therefore can be used for other image processing tasks such as image registration, see Droske et al. [2009]. Thirdly, in contrast to sharp edge representations, the phase-field function only indicates the approximate position of edges in a blurry way. As a result, stronger and weaker edges can be distinguished with the indicator function. Naturally, these characteristics are disadvantageous for certain tasks, for example if a sharp segmentation into different regions is desired.

For the original Mumford-Shah penalty, the Γ-approximation of ***Ambrosio-Tortorelli*** is

$$\mathrm{AT}_{\varepsilon_n}(f, v) = \alpha \int\limits_{\Omega} (v^2 + k_{\varepsilon_n}) |\nabla f|^2 dx + \beta \int\limits_{\Omega} \varepsilon_n |\nabla v|^2 + \frac{(1-v)^2}{4\varepsilon_n} dx, \tag{1.4}$$

where $f, v \in W^{1,2}(\Omega), \varepsilon_n, k_{\varepsilon_n} \in \mathbb{R}^+$ and k_{ε_n} is of higher order than ε_n. As for the Mumford-Shah penalty, the first integral of (1.4) enforces smoothness whenever there is no edge, i.e. $v \approx 1$, and the second integral penalizes edges.

To derive a variational approximation for (1.3) we need to describe the *a priori* edges in a suitable way. We assume that for the *a priori* edge set K^0 we have a sequence of smooth functions $\{v_n^0\}$, for which $v_n^0(x) \to 0$ if $x \in K^0$ and otherwise $v_n^0(x) \to 1$. We will add further technical assumptions later, see Assumption 4.1.

In this work we study the modified penalty

$$\mathrm{AT}_{\varepsilon_n, v_n^0}(f, v) = \alpha \int_\Omega (v^2 + k_{\varepsilon_n})|\nabla f|^2 dx + \beta \int_\Omega \varepsilon_n |\nabla v|^2 + \frac{(v_n^0 - v)^2}{4\varepsilon_n} dx, \qquad (1.5)$$

where the last quadratic term was changed from $(1-v)^2$ to $(v_n^0 - v)^2$. In areas where no *a priori* edge information is available, that is locally $v_n^0 \approx 1$ for all n large enough, the penalties AT_{ε_n} and $AT_{\varepsilon_n, v_n^0}$ are approximately the same. In areas where there is *a priori* edge information, that is locally $v_n^0 \approx 0$ for all n large enough, the penalty guides v to take an edge at the location.

In Chapter 4 we will present sufficient conditions under which the proposed sequence of functionals (1.5) Γ-converges to the Mumford-Shah penalty including *a priori* edge information. Unfortunately, we can only establish the Γ-convergence under the condition $\gamma = 0$. Nevertheless, we propose a heuristic approximation which overcomes shortcomings of the above proposed penalty in numerical implementations.

1.6 Contribution and related work

There are several similar variational models to (1.3) in the literature. For example in the already mentioned work Leahy and Yan [1991] a finite difference model using the same idea of penalizing jumps differently according to some *a priori* known set is studied. Mathematically very close variational problems are studied in the context of fracture mechanics by Amar et al. [2010]; Babadjian and Giacomini [2013]; Dal Maso et al. [2005]; Giacomini and Ponsiglione [2006]. Motivated by crack evolution in anisotropic materials these works cover rather general edge penalty terms. The same compactness and lower semicontinuity properties in the aforementioned work are used in this thesis, for example to obtain existence of a solution. For this reason the main contribution of Chapter 3 are the results on the regularization properties of the Mumford-Shah penalty (also for the case $K^0 = \emptyset$). As shown by Fornasier et al. [2011] the topologies and assumptions have to be chosen with care to ensure that the regularization is well posed. Once a setting is fixed, the actual proofs for stability and the regularization properties follow established paths such as in Anzengruber and Ramlau [2010]; Klann and Ramlau [2013]; Ramlau and Ring [2010]. However, the proofs still needed to be verified in detail, for example to see what constraints on the parameters are necessary. Our work extends the regularization results of Jiang et al. [2014]; Rondi [2008b] and gives further justification of using the Mumford-Shah approach as a regularization.

There are many approximations of the original Mumford-Shah functional that allow numerical implementations. For example, there are finite difference approximations Chambolle [1995], finite element approximations Bourdin and Chambolle [2000], elliptic approximations Ambrosio and Tortorelli [1992] or non local approximations Braides and Dal Maso [1997]. See also Braides [2002]. In Chapter 4 we extend the classic approach of Ambrosio and Tortorelli [1992], which is still used as the reference algorithm for comparisons with new implementations of the Mumford-Shah functional. To our knowledge this kind of extension is new. The difference to other generalizations is that our approximations depend not only on a scalar parameter but also on a sequence of functions.

Our numerical results are meant as a first illustration of the proposed model rather than an exhaustive study. The resulting algorithm alternates between an image reconstruction and image evaluation step. Such a procedure can also be understood in the context of adaptive regularization methods with non constant regularization parameter, see for example Alexandrov et al. [2010]; Gilboa et al. [2006]; Grasmair [2009].

1.7 Outline

The thesis is structured as follows.

In Chapter 2 we state some notations and results needed in the later parts of the thesis.

In Chapter 3 we study the proposed model regarding existence of a minimizer, stability with respect to the data and regularization parameters and parameter choice rules.

In Chapter 4 we study the proposed approximation. We first investigate the *a priori* sequence $\{v_n^0\}$ and note the assumptions we impose. Then we establish convergence in one dimension for the case $\gamma = 0$ and lift the result to dimension $N \geq 2$ by standard arguments in the theory of Γ-convergence. Finally, we give a heuristic motivation for a second variational approximation.

In Chapter 5 we evaluate the approach for two inverse problems with simulated data. One is X-ray CT and one is 2 dimensional DOT. Both applications are covered by our theoretical results.

In Chapter 6 we draw a conclusion and provide an outlook for future work.

Chapter 2

Preliminaries

2.1 Inverse Problems

In this section a brief introduction on ill-posed inverse problems is provided. The reference for this section are Engl et al. [1996]; Louis [1989]; Rieder [2003].

Let X and Y be Banach spaces and $A : X \to Y$ a operator, possibly non-linear. Computing for a given $f \in X$ its effect under the operator $A(f) = g$ is called the *direct problem*. In many applications the opposite is desired, that is for a given observation $g \in Y$ the cause $f \in X$ resulting in g is wanted. Solving the operator equation

$$A(f) = g \tag{2.1}$$

for a g, is called the *inverse problem*. If the operator A is linear and bounded, it is called a *linear inverse problem*. Inverse problems arise in many branches of science and engineering, including computer vision, geophysics, medical imaging and nondestructive testing Rieder [2003]. In most applications the inverse problem is ill-posed, which leads to sever difficulties due to the inaccuracy in the model and the noise in the data.

Hadamard [1923] proposed the following definition for the well-posedness of an inverse problem.

Definition 2.1. *Let $A : X \to Y$ be a map between the topological spaces X and*

Y. The problem (A, X, Y) is called **well-posed**, *if the following properties hold.*

(i) *The equation $Af = g$ has a solution for every $g \in Y$.*

(ii) *The solution is uniquely determined by the data g.*

(iii) *The inverse map $A^{-1} : Y \to X$ is continuous, that is the solution f is continuously dependent on the data g.*

If one of the conditions does not hold, the problem is called **ill-posed**.

We denote by $R(A) := A(X)$ the range of A. Condition (iii) is usually the crucial one. In applications, the measurement typically carries some errors. Often instead of the exact data $g \in R(A)$ only a corrupted version $g^\delta \in Y$ is measured. A simple way to characterize the noise is by its noise level δ, with $\|g^\delta - g\|_Y \leq \delta$.

As g^δ is not necessarily in the range of A, more general solutions to the inverse problem have to be considered. A possible approach is to consider *least square solutions* $f \in X$ for which

$$f := \operatorname{argmin}_{f \in X} \|Af - g^\delta\|_Y^2. \tag{2.2}$$

Other criteria might also be used in 2.2, for example the Kullback-Leibler divergence. If least square solutions exist there might be infinite many and in order to uniquely determine a solution an additional selection is needed. Furthermore, the least squares approach does not resolve the crucial problem of condition (iii).

To stabilize the inversion often the minimization of *penalized least squares functionals*

$$F(f) = \|A(f) - g^\delta\|_{L^2}^2 + \alpha\Psi(f) \tag{2.3}$$

is considered, as in Tikhonov [1963]. In this approach, besides the fitting to the observed data g^δ enforced through the first term (called the fidelity term), *a priori* knowledge on the solution f can be introduced into the reconstruction through the penalty functional Ψ. Some choices of the regularization term are $\Psi(f) = \|\nabla f\|_{L^2}^2$, $\Psi(f) = \|f\|_{L^2}^2$ Tikhonov [1963] or $\Psi(f) = \|\nabla f\|_{L^1}$ (the total variation of f) Rudin et al. [1992]. Other choices of Ψ could be norms or semi-norms of f and/or ∇f or "norms" with respect to a certain basis or frame as in sparsity regularization Daubechies et al. [2004]. The parameter $\alpha > 0$ balances the influence of the data fitting term and the penalty term Ψ.

The choice of α has a big influence on the quality of the reconstruction. A too large α leads to a minimizer that might not approximate the data well enough, a too small α leads to a minimizer where the errors of the data might be carried over and amplified. Furthermore it has to be ensured that the bias introduced through the penalty is reasonable. That is the *a priori* knowledge is used to deal with the errors and noise in the data, but it can be ensured, with a suitable parameter choice rule, that for decreasing noise the reconstruction tends to the true solution.

In our work we consider non-linear operators and non-convex variational methods. Therefore we can not ask for uniqueness of the minimizer and hence also not for convergence in norm. Moreover we have multiple parameters. These requirements, together with the need of stability, lead to the following definition of a regularization we will use.

Definition 2.2. *Let $A^\dagger : R(A) \to X$ be an operator that maps $g \in R(A)$ to a solution $f^\dagger = A^\dagger(g)$ of (2.1). A family of operators $R^\alpha : Y \to 2^X$ for $\alpha \in \mathbb{R}^N$ is called a* **regularization** *for $A^\dagger$ if:*

(i) (Existence) *The set $R^\alpha(g)$ is non-empty for all $g \in Y$.*

(ii) (Stability) *For every $\alpha \in \mathbb{R}_+^N$, $\{g_n\}, g \in Y$, if $g_n \to g$, then there exists a sequence $\{f_n\} \in R^\alpha(g_n)$ that convergences, at least subsequentially, to a $f \in R^\alpha(g)$.*

(iii) (Convergence) *For $g^\delta \in Y$ such that $\|g^\delta - g\|_Y \le \delta$ and suitably chosen $\alpha = \alpha(\delta, g^\delta)$ with $\alpha(\delta, g^\delta) \to 0$ as $\delta \to 0$, there exists a sequence $\{f_\delta\} \in R^\alpha(g^\delta)$ that convergences, at least subsequentially, to a solution of (2.1) as $\delta \to 0$.*

2.2 Radon measures and the Hausdorff measure

In this section the theory of measures needed for the space of bounded variation is recalled. Moreover the Hausdorff measure is defined. The main reference is [Ambrosio et al., 2000, Chapter 1].

First the σ-algebra and measure spaces are defined.

Definition 2.3 ([Ambrosio et al., 2000, p.1 Definition 1.1]). *Let M be a nonempty set and let $\mathcal{A}$ be a collection of subsets of M.*

(i) *We say that $\mathcal{A}$ is a* **σ-algebra** *if $\emptyset \in \mathcal{A}$, $M/E_1 \in \mathcal{A}$ whenever $E_1 \in \mathcal{A}$ and for every sequence $(E_n)_n \subset \mathcal{A}$ the union $\bigcup_n E_n \in \mathcal{A}$.*

(ii) *If $\mathcal{A}$ is a σ-algebra in M, we call the pair $(M, \mathcal{A})$ a* **measure space**.

(iii) *The smallest σ-algebra containing all open subsets of M is called* **Borel-algebra** *and is denoted by $\mathcal{B}(M)$, when M is a topological space.*

For simplicity we restrict the following definitions to bounded domains $\Omega \subset \mathbb{R}^N$.

Definition 2.4 ([Ambrosio et al., 2000, p.2 Definition 1.4, p.19 Definition 1.40]). *Let $(\Omega, \mathcal{A})$ be a measure space.*

(i) *We say that $\mu : \mathcal{A} \to \mathbb{R}^m$ is a* **measure** *if $\mu(\emptyset) = 0$ and for any sequence $(E_n)_n$ of pairwise disjoint elements of $\mathcal{A}$*

$$\mu\Big(\bigcup_n E_n\Big) = \sum_n \mu(E_n).$$

(ii) *A set function defined on the relatively compact Borel subsets of Ω that is a measure on $(K, \mathcal{B}(K))$ for every compact set $K \subset \Omega$ is called a* **Radon measure** *on Ω.*
If the function is a measure on $(\Omega, \mathcal{B}(\Omega))$ it is called a **finite Radon measure**.

We state a special case of the Radon-Nykodim Theorem, which we will need to characterize functions of bounded variation. We denote the N dimensional Lebesgue measure by $\mathcal{L}^N$.

Theorem 2.5 ([Ambrosio et al., 2000, p.14 Theorem 1.28] Radon-Nykodim). *Let ν be a $\mathbb{R}^m$-valued finite Radon measure on Ω. Then there exists a unique pair of*

$\mathbb{R}^m$-valued measures ν^a, ν^s with

$$\nu = \nu^a + \nu^s,$$

where ν^a is absolutely continuous with respect to $\mathcal{L}^N$ and ν^s is singular with respect to $\mathcal{L}^N$. Moreover there is a unique function $f \in \big(L^1(\Omega)\big)^m$ so that $\nu^a = f\mathcal{L}^N$.

The Hausdorff measure below is a generalized volume for lower dimensional sets. We will use the Hausdorff measure to penalize the discontinuities in the Mumford-Shah functional.

Definition 2.6 ([Ambrosio et al., 2000, p.72 Definition 2.46])**.** *Let $k \in \mathbb{N}$ and $E \subset \mathbb{R}^N$. The* **k-dimensional Hausdorff measure** *of E is given by*

$$\mathcal{H}^k(E) = \lim_{\rho \downarrow 0} \mathcal{H}^k_\rho(E)$$

where, for $0 < \delta \leq \infty, \mathcal{H}^k_\rho$ is defined by

$$\mathcal{H}^k_\rho(E) := \frac{\omega_k}{2^k} \inf\{\sum_{i \in I} (\mathrm{diam}(E_i))^k : \mathrm{diam}(E_i) < \rho, E \subset \cup_{i \in I} E_i\}$$

for finite or countable covers $\{E_i\}_{i \in I}$, with the convention $\mathrm{diam}(\emptyset) = 0$. The constant ω_k is the Lebesgue measure of the unit ball in $\mathbb{R}^k$.

We will need the following lemma to lift our Γ-convergence result from one dimension to N dimensions in Chapter 4.

Lemma 2.7 ([Attouch et al., 2006, p.122 Lemma 4.22])**.** *Let μ be a non-negative $\mathbb{R}-$valued Radon measure, $\{f_n\}$ be a family of non-negative functions in $L^1_\mu(\Omega)$. Then*

$$\int_\Omega sup_n f_n d\mu = \sup \Big[\sum_{i \in I} \int_{A_i} f_i d\mu \Big],$$

where the supremum is taken over all finite families $\{A_i\}_{i \in I}$ of pairwise disjoint open subsets of Ω.

2.3 Sobolev spaces

In this section Sobolev spaces are introduced. The reference is [Ambrosio et al., 2000, Chapter 2]. Let $\Omega \subset \mathbb{R}^N$ be an open domain.

We denote the Banach space of real valued p-integrable functions as $L^p(\Omega, \mathbb{R})$ for $1 \le p < \infty$, i.e. $f \in L^p(\Omega, \mathbb{R})$ if and only if

$$\|f\|_{L^p(\Omega)} := \left(\int_\Omega |f|^p dx \right)^{\frac{1}{p}} < \infty.$$

The Banach space of essentially bounded functions is written as $L^\infty(\Omega, \mathbb{R})$ with norm

$$\|f\|_{L^\infty(\Omega)} := \inf_{N \subset \Omega \,:\, \mathcal{L}^n(N)=0} \big(\sup_{x \in \Omega \setminus N} |f(x)| \big).$$

In the following we will write $L^p(\Omega) := L^p(\Omega, \mathbb{R})$ for $1 \le p \le \infty$ and also omit $\mathbb{R}$ for the Sobolev spaces $W^{m,p}(\Omega)$ which we define in the following.

We first need the definition of weak derivatives.

Definition 2.8 ([Ambrosio et al., 2000, p.43 Definition 2.3 + p.45 Remark 2.10])**.** *Let α be a multiindex and $f \in L^1_{loc}(\Omega)$; if there is $g_\alpha \in L^1_{loc}$ such that*

$$\int_\Omega f \frac{\partial^{|\alpha|}}{\partial x^\alpha} \varphi dx = (-1)^{|\alpha|} \int_\Omega g_\alpha \varphi dx \qquad \forall \varphi \in C_0^\infty(\Omega),$$

then we say that g_α is the **weak α-th derivative** *of f. The α-th weak derivative if exists is unique and is denoted by $\frac{\partial^{|\alpha|}}{\partial x^\alpha} f$ or $\nabla^\alpha f$.*

We can now define the Sobolev spaces.

Definition 2.9 ([Ambrosio et al., 2000, p.43 Definition 2.4 and p.45 Remark 2.10])**.** *We say that $f \in W^{m,p}(\Omega)$ if $f \in L^p(\Omega)$ and for every $|\alpha| \le m$ all weak derivatives $\frac{\partial^{|\alpha|}}{\partial x^\alpha} f$ exist and belong to $L^p(\Omega)$.*

The Sobolev spaces $W^{m,p}(\Omega)$ for $1 \leq p < \infty$ endowed with the norm

$$\|f\|_{W^{m,p}} = \left(\|f\|_{L^p}^p + \sum_{i=1}^{m} \sum_{|\alpha|=i} \|\nabla^\alpha f\|_{L^p}^p \right)^{\frac{1}{p}}$$

and for $p = \infty$ with

$$\|f\|_{W^{m,\infty}} = \|f\|_{L^\infty} + \sum_{i=1}^{m} \sum_{|\alpha|=i} \|\nabla^\alpha f\|_{L^\infty}$$

are Banach spaces (for $p = 2$ Hilbert spaces).

We make a remark about the weak convergence in Sobolev spaces, which we will need later. First we note the weak convergence in $L^p(\Omega)$ and in $W^{m,p}(\Omega)$.

Remark 2.10 ([Alt, 1985, p.228 example 1 and 3]). *Let $1 \leq p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. If $p = 1$ set $q = \infty$.*

(i) *For $f_n, f \in L^p(\Omega)$:*

$$\begin{aligned} & f_n \rightharpoonup f \quad \text{weakly in } L^p(\Omega) \\ \Longleftrightarrow \quad & \int_\Omega f_n \, g \, dx \to \int_\Omega f \, g \, dx \quad \text{for all } g \in L^q(\Omega). \end{aligned}$$

(ii) *For $f_n, f \in W^{m,p}(\Omega)$:*

$$\begin{aligned} & f_n \rightharpoonup f \quad \text{weakly in } W^{m,p}(\Omega) \\ \Longleftrightarrow \quad & \frac{\partial^\alpha}{\partial x_\alpha} f_n \rightharpoonup \frac{\partial^\alpha}{\partial x_\alpha} f \quad \text{weakly in } L^p(\Omega) \quad \text{for all } |\alpha| \leq m. \end{aligned}$$

Now we show that for bounded domains weak convergence in $W^{m,p}(\Omega)$ for $1 < p < \infty$ implies weak convergence in $W^{m,1}(\Omega)$.

Remark 2.11. *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain and $f_n \rightharpoonup f$ weakly in $W^{m,p}(\Omega)$ with $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$.*
As the domain is bounded for every $|\alpha| \leq m$ the functions $\frac{\partial^{|\alpha|}}{\partial x^\alpha} f_n, \frac{\partial^{|\alpha|}}{\partial x^\alpha} f \in L^p(\Omega)$

are also in $L^1(\Omega)$*. Moreover* $L^\infty(\Omega) \subset L^q(\Omega)$ *and therefore for every* $g \in L^\infty(\Omega)$

$$\int_\Omega \frac{\partial^{|\alpha|}}{\partial x^\alpha} f_n \, g \, dx \to \int_\Omega \frac{\partial^{|\alpha|}}{\partial x^\alpha} f \, g \, dx.$$

2.4 Functions of bounded variation

In this section the space of functions of bounded variations $BV(\Omega, \mathbb{R})$ is introduced. It is the space that we use to model images for a relaxed version of the Mumford-Shah functional. The reference for this chapter is Ambrosio et al. [2000]. Let $\Omega \subset \mathbb{R}^N$ be open and bounded. We again write $BV(\Omega) := BV(\Omega, \mathbb{R})$

Definition 2.12 ([Ambrosio et al., 2000, p. 117 Definition 3.1] The space $BV(\Omega)$)**.** *Let* $f \in L^1(\Omega)$*; we say that* f *is a* **function of bounded variation in** Ω *if the distributional derivative of* f *is representable by a finite Radon measure in* Ω*, i.e. if*

$$\int_\Omega f \frac{\partial \varphi}{\partial x_i} dx = -\int_\Omega \varphi d D_i f \quad \forall \varphi \in C_0^\infty(\Omega), \quad i = 1, ..., n$$

for some $\mathbb{R}^N$*-valued measure* $Df = (D_1 f, ..., D_N f)$ *in* Ω*. The vector space of all functions of bounded variation in* Ω *is denoted by* $BV(\Omega)$ *.*
Moreover we define the **total variation** *of* Df *as*

$$|Df| := \sup \left\{ \sum_{n=0}^{\infty} |Df(E_n)| : E_n \in \mathcal{B}(\Omega) \text{ pairwise disjoint}, \Omega = \bigcup_n^\infty E_n \right\}$$

To explain the name of $BV(\Omega)$ the following definition of variation is needed.

Definition 2.13 ([Ambrosio et al., 2000, p. 119 Definition 3.4])**.** *Let* $f \in L^1_{loc}(\Omega)$*. The* **variation** $V(f, \Omega)$ *of* f *in* Ω *is defined by*

$$V(f, \Omega) := \sup \left\{ \int_\Omega f \operatorname{div}(\varphi) \, dx : \varphi \in C_0^1(\Omega)^n, \|\varphi\|_\infty \leq 1 \right\}$$

The space $BV(\Omega)$ contains exactly those functions $f \in L^1(\Omega)$ for which the variation $V(f, \Omega)$ is bounded.

Proposition 2.14 ([Ambrosio et al., 2000, p. 120 Proposition 3.6])**.** *Let $f \in L^1_{loc}(\Omega)$. Then, f belongs to $BV(\Omega)$ if and only if $V(f,\Omega) < \infty$. In addition for any $f \in BV(\Omega)$ it holds $V(f,\Omega) = |Df|(\Omega)$.*

The space $BV(\Omega)$ is a Banach space with the norm

$$\|f\|_{BV(\Omega)} := \|f\|_{L^1(\Omega)} + |Df|(\Omega).$$

For every $f \in W^{1,1}(\Omega)$ it is $\|f\|_{W^{1,1}(\Omega)} = \|f\|_{BV(\Omega)}$. The inclusion $W^{1,1}(\Omega) \subset BV(\Omega)$ is strict, as for example the Heaviside function $h : (-1,1) \to \mathbb{R}$

$$h(x) := \begin{cases} 1, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

is in $BV(-1,1) \setminus W^{1,1}(-1,1)$.

The notion of continuity and differentiability has to be generalized suitably for functions in $BV(\Omega)$.

Definition 2.15 ([Ambrosio et al., 2000, p. 160 Definition 3.63] Approximate limit)**.** *Let $f \in L^1_{loc}(\Omega)$; we say that f has an* **approximate limit at** *$x \in \Omega$ if there exists $z \in \mathbb{R}$ such that*

$$\lim_{\rho \downarrow 0} \frac{1}{\mathcal{L}^N(B_\rho(x))} \int_{B_\rho(x)} |f(y) - z| dy = 0. \tag{2.4}$$

The set $S_f \subset \Omega$ without an approximate limit is called the **approximate discontinuity set**. *For any $x \in \Omega \setminus S_f$ the value z, uniquely determined by (2.4), is called the approximate limit of f at x and is denoted by $\tilde{f}(x)$. A function f is called* **approximately continuous** *at x if $x \notin S_f$ and $\tilde{f}(x) = f(x)$.*

For a function in $BV(\Omega)$ the approximate discontinuity set is of zero Lebesgue measure $\mathcal{L}^N(S_f) = 0$.

Proposition 2.16 ([Ambrosio et al., 2000, p. 160 Proposition 3.64 (a)])**.** *Let $f \in L^1_{loc}(\Omega)$. Then S_f is a $\mathcal{L}^N$-negligible set and $\tilde{f}(x) : \Omega \setminus S_f \to \mathbb{R}$ coincides $\mathcal{L}^N$-almost everywhere in $\Omega \setminus S_f$ with f.*

In the same idea the approximate gradient is defined.

Definition 2.17 ([Ambrosio et al., 2000, p. 165 Definition 3.70] Approximate differentiability)**.** *Let $f \in L^1_{loc}(\Omega)$ and let $x \in \Omega \setminus S_f$; we say that f is* **approximately differentiable** *at x if there exists a $1 \times N$ matrix L such that*

$$\lim_{\rho \downarrow 0} \frac{1}{\mathcal{L}^N(B_\rho(x))} \int_{B_\rho(x)} \frac{|f(y) - \tilde{f}(x) - L(y-x)|}{\rho} dy = 0. \tag{2.5}$$

If f is approximately differentiable at x the matrix, uniquely determined by (2.5), is called the **approximate differential** *of f at x and is denoted by $\nabla f(x)$.*

For the definition of the approximate jump set the following notations for two halves of a ball $B_\rho(x)$ cut by a hyperplane is used. Let $\nu \in \mathbb{R}^N$ and $\langle \cdot, \cdot \rangle$ denote the euclidean inner product in $\mathbb{R}^N$. We define:

$$B^+_\rho(x,\nu) := \{y \in B_\rho(x) | \, \langle y - x, \nu \rangle > 0\},$$
$$B^-_\rho(x,\nu) := \{y \in B_\rho(x) | \, \langle y - x, \nu \rangle < 0.\}.$$

We write the unit sphere as S^{N-1}.

Definition 2.18 ([Ambrosio et al., 2000, p. 163 Definition 3.67] Approximate jump points)**.** *Let $f \in L^1_{loc}(\Omega)$ and $x \in \Omega$. We say that x is an* **approximate jump point** *of f if there exist $a, b \in \mathbb{R}$ and $\nu \in S^{N-1}$ such that $a \neq b$ and*

$$\lim_{\rho \downarrow 0} \frac{1}{\mathcal{L}^N(B^+_\rho(x,\nu))} \int_{B^+_\rho(x,\nu)} |f(y) - a| dy = 0, \tag{2.6}$$

and

$$\lim_{\rho \downarrow 0} \frac{1}{\mathcal{L}^N(B^-_\rho(x,\nu))} \int_{B^-_\rho(x,\nu)} |f(y) - b| dy = 0. \tag{2.7}$$

The triplet (a, b, ν), uniquely determined by equation (2.6) up to a permutation of (a, b) and a change of sign of ν, is denoted by $\big(f^+(x), f^-(x), \nu_f(x)\big)$. The set of approximate jump points is denoted by J_f.

In the following the two triplets (a, b, ν) and $(b, a, -\nu)$ are viewed as equivalent.

Using the Radon-Nykodim theorem 2.5 the distributional derivative Df can be split into an absolutely continuous part $D^a f$, which is integrable with respect to the Lebesgue measure, and a singular part $D^s f$

$$Df = D^a f + D^s f.$$

The singular part can be further divided into a jump part, restricted to the jump set $D^j f = D^s f|_{J_f}$, and a diffuse part $D^c f = D^s f - D^j f = D^s f|_{\Omega \setminus J_f}$. The part D^c is called Cantor part, as the Cantor-Vitali function is the most prominent example for a function with $Df = D^c f$.

The density of the absolutely continuous part D^a can be characterized more precisely.

Theorem 2.19 ([Ambrosio et al., 2000, p. 177 Theorem 3.83] Calderon Zygmund)**.** *Any function $f \in BV(\Omega)$ is approximately differentiable at $\mathcal{L}^N$-almost everywhere point of Ω. Moreover, the approximate differential ∇f is the density of the absolutely continuous part of Df with respect to $\mathcal{L}^N$.*

The jump part $D^j f$ can also be characterized.

Theorem 2.20 ([Ambrosio et al., 2000, p. 173 Theorem 3.78])**.** *Let $f \in BV(\Omega)$ then*

$$\mathcal{H}^{N-1}(S_f \setminus J_f) = 0 \tag{2.8}$$

and

$$D^j f = (f^+ - f^-)\nu_f \mathcal{H}^{n-1}|_{J_f}.$$

Because of (2.8) the jump set J_f and discontinuity set S_f are essentially the same and we will only use S_f in the following chapters.

It turns out that the space $BV(\Omega)$ is too big for minimizing Mumford-Shah type functionals. Instead the subspace of functions with $D^c = 0$ will be used.

Definition 2.21 ($SBV(\Omega)$)**.** *Let $f \in BV(\Omega)$; we say that f is a* **special function of bounded variation in** Ω *if*

$$D^c f = 0. \tag{2.9}$$

The vector space of all special functions of bounded variation in Ω is denoted by $SBV(\Omega)$.

In the following some properties are listed that will be useful in the later parts of the thesis. The most important property is the compactness and semicontinuity theorem in $SBV(\Omega)$.

First we states that the variation is reduced if a function is cut off at a constant value.

Proposition 2.22 ([De Giorgi et al., 1989, p.198 Remark. 2.2])**.** *Let $f \in BV(\Omega)$ and set $f_a := \min\{a, \max\{f, -a\}\}$ for $0 < a < \infty$. The following properties hold:*

$$
\begin{aligned}
|\nabla f_a| \leq |\nabla f| \quad & a.e. \ on \ \Omega, \\
\mathcal{H}^{N-1}(S_{f_a} \backslash S_f \cap \Omega) = 0&, \\
\int_\Omega |Df_a| dx \leq \int_\Omega |D_f| dx&.
\end{aligned}
$$

From the above proposition it directly follows $\int_\Omega |\nabla f_a| dx \leq \int_\Omega |\nabla f| dx$ and $\mathcal{H}^{N-1}(S_{f_a} \cap \Omega) \leq \mathcal{H}^{N-1}(S_f \cap \Omega)$.

The following propositions are about the relation between the spaces $W^{1,1}(\Omega)$ and $SBV(\Omega)$.

Proposition 2.23 ([De Giorgi et al., 1989, p.198 Lemma 2.3])**.** *Let $f \in L^\infty(\Omega) \cap L^1(\Omega)$. Let $K \subset \mathbb{R}^N$ be closed and assume*

$$f \in W^{1,1}(\Omega \setminus K) \quad and \quad \mathcal{H}^{N-1}(K \cap \Omega) < +\infty.$$

Then

$$f \in SBV(\Omega) \quad and \quad S_f \cap \Omega \subset K.$$

Also note that for a bound domain Ω if $\int_{\Omega \backslash K} |\nabla f|^2 dx < +\infty$ then also $\int_{\Omega \backslash K} |\nabla f| dx < +\infty$.

Proposition 2.24. *Let $f \in SBV(\Omega) \cap L^\infty(\Omega)$ and*

$$\int_\Omega |\nabla f|^p dx < +\infty,$$

then $f \in W^{1,p}(\Omega \setminus \overline{S_f})$.

Proof. For a function $f \in SBV(\Omega)$ the distributional derivative restricted to $\Omega \setminus \overline{S_f}$ is the absolutely continuous part, i.e. $Df|_{\Omega\setminus\overline{S_f}} = \nabla f \mathcal{L}^N$. The approximate differential $\nabla f \in L^1(\Omega)$ is therefore the weak derivative of f in $\Omega \setminus \overline{S_f}$. Together with the L^∞ bound of f and the boundedness of the domain Ω we have $f \in W^{1,1}(\Omega \setminus \overline{S_f})$. As the L^p-norm of the weak derivative is also bounded it follows $f \in W^{1,p}(\Omega \setminus \overline{S_f})$. □

The following compactness and semicontinuity theorem was proven in Ambrosio [1989]. We state a version from Attouch et al. [2006]. It will be the key tool to prove the existence of a solution to the Mumford-Shah functional and for the study of the regularization properties of the Mumford-Shah penalty.

Theorem 2.25 ([Attouch et al., 2006, p. 515 Theorem 13.4.3])**.** *Let $\{f_n\}$ be a sequence in $SBV(\Omega)$ satisfying for $p > 1$,*

$$\sup_{k\in\mathbb{N}} \left\{ \|f_k\|_\infty + \int_\Omega |\nabla f_k|^p dx + \mathcal{H}^{N-1}(S_{f_k}) \right\} < +\infty.$$

Then there exists a subsequence, still denoted as $\{f_n\}$ and a function f in $SBV(\Omega)$, such that

$$\begin{aligned} f_k &\to f && \textit{strongly in } L^1_{loc}(\Omega), \\ \nabla f_k &\rightharpoonup \nabla f && \textit{weakly in } L^p(\Omega), \\ \mathcal{H}^{N-1}(S_f) &\leq \liminf_{k\to\infty} \mathcal{H}^{N-1}(S_{f_k}). \end{aligned}$$

For the case including the *a priori* edge knowledge $K^0 \subset \Omega$ we will also need to extend the lower semicontinuity to the following penalty

$$K \mapsto \beta\mathcal{H}^{N-1}(K \setminus K^0) + \gamma\mathcal{H}^{N-1}(K \cap K^0)$$

with $0 \leq \gamma < \beta$ and K^0 fixed. See [Amar et al., 2010, Theorem 3.2] for a lower semicontinuity result that covers our penalty amongst other generalizations.

We state a different proof here.

Corollary 2.26. *Let K^0 be a compact subset of Ω with $\mathcal{H}^{N-1}(K^0) < \infty$ and $\{f_n\}$ be a sequence in $SBV(\Omega)$ satisfying for $p > 1$,*

$$\sup_{n\in\mathbb{N}} \left\{ \|f_n\|_\infty + \int_\Omega |\nabla f_n|^p dx + \mathcal{H}^{N-1}(S_{f_n}) \right\} < +\infty.$$

Then there exists a subsequence, still denoted as $\{f_n\}$ and a function f in $SBV(\Omega)$, such that for $\alpha > 0$ and $\beta > \gamma \geq 0$

$$\begin{aligned} \alpha \int_\Omega |\nabla f|^p dx + \beta\mathcal{H}^{N-1}(S_f \setminus K^0) + \gamma\mathcal{H}^{N-1}(S_f \cap K^0) \\ \leq \liminf_{n\to\infty} \left[\alpha \int_\Omega |\nabla f_n|^p dx + \beta\mathcal{H}^{N-1}(S_{f_n} \setminus K^0) + \gamma\mathcal{H}^{N-1}(S_{f_n} \cap K^0) \right]. \end{aligned} \tag{2.10}$$

Proof. Because of the lower semicontinuity Theorem 2.25 we directly have

$$\gamma\mathcal{H}^{N-1}(S_f) \leq \gamma \liminf_{n\to\infty} \mathcal{H}^{N-1}(S_{f_n}) \tag{2.11}$$

and

$$\int_\Omega |\nabla f|^p dx \leq \liminf_{n\to\infty} \int_\Omega |\nabla f_n|^p dx. \tag{2.12}$$

Furthermore if we define the new open domain $\Omega^* := \Omega \setminus K^0$ we can apply the theorem again to obtain

$$\begin{aligned} & \mathcal{H}^{N-1}(S_f \setminus K^0) \leq \liminf_{n\to\infty} \mathcal{H}^{N-1}(S_{f_n} \setminus K^0) \\ \iff \quad & 0 \leq \liminf_{n\to\infty} \left[\mathcal{H}^{N-1}(S_{f_n} \setminus K^0) - \mathcal{H}^{N-1}(S_f \setminus K^0) \right]. \end{aligned} \tag{2.13}$$

By rearranging (2.11) we can get

$$\begin{aligned} & \gamma\mathcal{H}^{N-1}(S_f \cap K^0) \\ & \leq \liminf_{n\to\infty} \left[\gamma \left(\mathcal{H}^{N-1}(S_{f_n} \setminus K^0) - \mathcal{H}^{N-1}(S_f \setminus K^0) \right) + \gamma\mathcal{H}^{N-1}(S_{f_n} \cap K^0) \right]. \end{aligned}$$

Using (2.13) together with $\beta > \gamma$ and $0 \leq \mathcal{H}^{N-1}(S_{f_n} \cap K^0)$ we can follow

$$\begin{aligned}&\gamma \mathcal{H}^{N-1}(S_f \cap K^0)\\&\leq \liminf_{n\to\infty} \left[\beta \left(\mathcal{H}^{N-1}(S_{f_n} \setminus K^0) - \mathcal{H}^{N-1}(S_f \setminus K^0)\right) + \gamma \mathcal{H}^{N-1}(S_{f_n} \cap K^0)\right].\end{aligned}$$

Bringing the constant $\beta \mathcal{H}^{N-1}(S_f \setminus K^0)$ back to the left side together with (2.12) leads to the claim. □

Functions in $SVB(\Omega)$ can also be characterized through their restrictions to one dimensional slices. The following notations are used to lift the variational approximation in Chapter 4 from one dimension to higher dimensions. We adapt the notation of [Attouch et al., 2006, p. 414]. Let S^{N-1} be the $N-1$ dimensional sphere. For every $\nu \in S^{N-1}$ we define

$$\begin{aligned}\pi_\nu &:= \{y \in \mathbb{R}^N : \langle y, \nu \rangle = 0\},\\ \Omega_y &:= \{t \in \mathbb{R} : y + t\nu \in \Omega\}, y \in \pi_\nu,\\ \Omega_\nu &:= \{y \in \pi_\nu : \Omega_y \neq \emptyset\}.\end{aligned}$$

Furthermore we define for all Borel functions $f : \Omega \to \mathbb{R}$ and y in Ω_ν the Borel

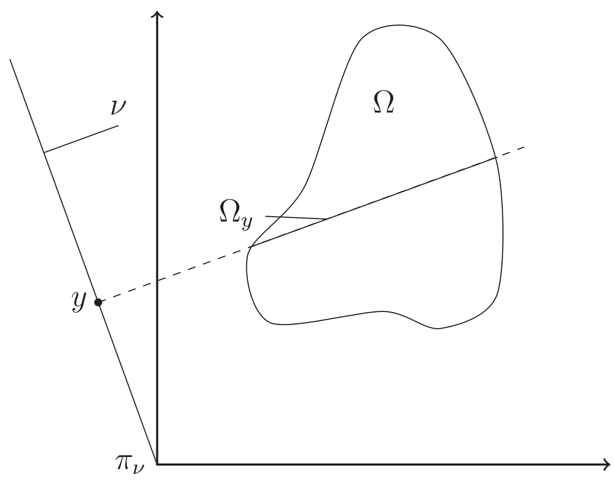


Figure 2.1: The domain Ω and a slice Ω_y for a fixed $\nu \in S^{N-1}$.

function f_y for all $t \in \Omega_y$ by $f_y(t) = f(y + t\nu)$.

Theorem 2.27 ([Attouch et al., 2006, p. 414 Theorem 10.5.2]). *Let f be a given function in $L^\infty(\Omega)$ such that for all $\nu \in S^{N-1}$*

(i) $f_y \in SBV(\Omega_y)$ *for* $\mathcal{H}^{N-1}$ *a.e. in* $y \in \Omega_\nu$,

(ii) $\int_{\Omega_\nu} \big(\int_{\Omega_y} |\nabla f_y| dt + \mathcal{H}^0(S_{f_y}) \big) d\mathcal{H}^{N-1} \leq \infty$.

Then f belongs to $SBV(\Omega)$. Conversely, if f belongs to $SBV(\Omega) \cap L^\infty(\Omega)$, conditions (i) and (ii) are satisfied for all $\nu \in S^{N-1}$. Moreover, for $\mathcal{H}^{N-1}$ a.e. in $y \in \Omega_\nu$

$$\langle \nabla f(y + t\nu), \nu \rangle = \nabla f_y(t)$$

and

$$\int_{\Omega_\nu} \mathcal{H}^0(S_{f_y}) d\mathcal{H}^{N-1} = \int_{S_f} |\langle \nu_f, \nu \rangle| d\mathcal{H}^{N-1}.$$

We conclude the section by giving an example of a BV function for which the jump set is not well behaved in the context of our segmentation problem.

Example 2.28 ([Chambolle, 2000, p. 34]). *Let $\Omega = B_1(0) \subset \mathbb{R}^2$ and*

$$f = \sum_{n=1}^{\infty} \frac{1}{2^n} \mathcal{X}_{B_{2^{-n}}(x_n) \cap \Omega},$$

where $(x_n)_n$ is the sequence of all points $\mathbb{Q}^2 \cap \Omega$ and $\mathcal{X}$ the characteristic function. For every $x \in \Omega$ it then is

$$\begin{aligned} f(x) &= \sum_{n: x \in B_{2^{-n}}(x_n)} \frac{1}{2^n} \leq \sum_{n=1}^{\infty} \frac{1}{2^n} \\ &= \frac{1}{1 - \frac{1}{2}} = 2. \end{aligned}$$

The Hausdorff measure of $\bigcup_n \partial B_{2^{-n}}(x_n) \cap \Omega$ *is bounded through*

$$\begin{aligned}\mathcal{H}^1\left(\bigcup_n \partial B_{2^{-n}}(x_n) \cap \Omega\right) &\leq \sum_{n=1}^{\infty} 2\pi \frac{1}{2^n} \\ &= 2\pi \frac{1}{1-\frac{1}{2}} = 4\pi.\end{aligned}$$

Every increase of the function f *is restricted to the singular jump set* $\bigcup_n \partial B_{2^{-n}}(x_n)$ *and so the absolutely continuous part of the distributional derivative and the Cantor part are zero. As the highest possible jump the function can take is* 2 *the function* f *is in* $SBV(\Omega)$ *with jump set* $\cup_n \partial B_{2^{-n}}(x_n) \cap \Omega$*.*

The jump set of f *is dense in* Ω

$$\overline{S_f} = \overline{\bigcup_n \partial B_{2^{-n}}(x_n) \cap \Omega} = \Omega.$$

That a behavior as in the example above can not occur for minimizers of the weak Mumford-Shah functional is a key step to prove the existence of a minimizer for the strong formulation.

2.5 Γ-convergence

In this section the theory of Γ-convergence is introduced.

Γ-convergence is a convergence designed for the approximation of variational problems. Loosely speaking, for a sequence of functionals $\{F_n\}$ that Γ-converge to a limit functional F, the minimizers of $\{F_n\}$ also approximate the minimizers of F. This allows the approximation of computational difficult problems by more feasible ones.

The main reference for this part is Braides [2002].

Definition 2.29 ([Braides, 2002, p.22 Definition 1.5])**.** *Let* (X, d) *be a metric space. We say that a sequence* $F_n : X \to [-\infty, +\infty]$ **Γ-converges** *in* X *to* $F : X \to [-\infty, +\infty]$ *if for all* $f \in X$*:*

(i) (lim inf inequality) *for every sequence* $\{f_n\}$ *converging to* f

$$F(f) \leq \liminf_n F_n(f_n); \tag{2.14}$$

(ii) (lim sup inequality / recovery sequence) *there exists a sequence* $\{f_n\}$ *converging to* f *such that*

$$\limsup_n F_n(f_n) \leq F(f). \tag{2.15}$$

The function F *is called the* **Γ-limit** *of* $\{F_n\}$ *and we write* $F = \Gamma\text{-}\lim_n F_n$.

The Γ-limit is uniquely determined if it exists. The definition of Γ-convergence can be extended to continuous parameters ε.

Definition 2.30. *We say that a sequence* $F_\varepsilon : X \to [-\infty, +\infty]$ Γ-converges *to* $F : X \to [-\infty, +\infty]$ *if for all sequences* $\{\varepsilon_n\}$ *converging to* 0 *we have* $\Gamma\text{-}\lim_n F_{\varepsilon_n} = F$.

The main advantage of Γ-convergence is characterized in the Fundamental Theorem of Γ-convergence.

Theorem 2.31 ([Braides, 2002, p.29 Theorem 1.21]Fundamental Theorem of Γ-convergence)**.** *Let* (X, d) *be a metric space, let* $\{F_n\}$ *be a sequence such that there exists a compact set* $K \subset X$, *so that for all* $n \in \mathbb{N}$

$$\inf_X F_n = \inf_K F_n.$$

If $\{F_n\}$ Γ*-convergences to* F, *then there exists a minimum of* F *and*

$$\min_X F = \lim_{n\to\infty} \inf_X F_n.$$

Moreover, if for a precompact sequence $\{x_n\}$*:* $\lim_{n\to\infty} F_n(x_n) = \lim_{n\to\infty} \inf_X F_n$, *then every limit point of* $\{x_n\}$ *is a minimum point of* F.

We list some properties of Γ-convergence which we will use in later chapters .

Proposition 2.32. *Let (X, d) be a metric space and let $F_n : X \to [-\infty, +\infty]$ Γ-converges in X to $F : X \to [-\infty, +\infty]$.*

(i) *The Γ-limit F is a lower semicontinuous function (see [Braides, 2002, p.32 Proposition 1.28]).*

(ii) *Every subsequence of $\{F_n\}$ Γ-convergences to the same Γ-limit F ([Braides, 2002, p.34 Proposition 1.37]).*

(iii) ***(Stability under continuous perturbation)***
If G is a continuous function, then $\{F_n + G\}$ Γ-converges to $(F + G)$ ([Braides, 2002, p.23 Remark 1.7]).

(iv) *Pointwise convergence does not lead to Γ-convergence.*
For example if and only if F is lower semicontinuous the constant sequence of functions $F_n = F$, $n = 1, 2, ..$ Γ-converges to F ([Braides, 2002, p.24 Remark 1.8]).

(v) *If for every $n \in \mathbb{N}$ F_n is convex, then also the Γ-limit F is convex ([Braides, 2002, p.38 Exercise 1.6]).*

2.6 σ-convergence

In this section the notion of σ-convergence is presented. The reference for this part is Dal Maso et al. [2005]. In our work we will characterize the convergence of edges in terms of σ-convergence.

First we give a definition of weak convergence in $SBV(\Omega)$ that was introduced in Dal Maso et al. [2005].

Definition 2.33 (Weak convergence in $SBV(\Omega)$ Dal Maso et al. [2005])**.** *We say a sequence $\{f_n\}$ converges weakly to f in $SBV(\Omega)$ if and only if f_n $(n = 1, \cdots)$ and $f \in SBV(\Omega) \cap L^\infty(\Omega)$, $f_n \to f$ a.e. in Ω, $\nabla f_n \rightharpoonup \nabla f$ in $L^1(\Omega)$, and both sequences $\{\|f_n\|_{L^\infty}\}$ and $\{\mathcal{H}^1(S_{f_n})\}$ are bounded.*

Note that this weak convergence is not meant in the usual Banach Space setting seeing as the dual of $BV(\Omega)$ is hard to characterize [Ambrosio et al., 2000, Remark 3.12 p. 124].

Given two sets A and $B \subset \mathbb{R}^N$, we write $A \tilde{\subset} B$ if $\mathcal{H}^{N-1}(A \backslash B) = 0$ and $A \tilde{=} B$ if $\mathcal{H}^{N-1}((A \backslash B) \cup (B \backslash A)) = 0$.

Definition 2.34 (Convergence of sets Dal Maso et al. [2005])**.** *A sequence of sets $\{E_n\}$ is said to σ-**converge** to E in Ω if E_n $(n = 1, \cdots)$ and $E \subset \Omega$, $\{\mathcal{H}^{N-1}(E_n)\}$ is bounded, and the following conditions are satisfied:*

- **(i)** *If $\{f_k\}$ converges weakly to f in $SBV(\Omega)$ and $S_{f_k} \tilde{\subset} E_{n_k}$ for some sequence $n_k \to \infty$, then $S_f \tilde{\subset} E$.*
- **(ii)** *There exist a function $f \in SBV(\Omega)$ and a sequence $\{f_n\}$ converging to f weakly in $SBV(\Omega)$ such that $S_f \tilde{=} E$ and $S_{f_n} \tilde{\subset} E_n$ for every n.*

From the second condition and the compactness and semicontinuity Theorem 2.25, it follows that if $\{E_n\}$ σ-converges to E, then

$$\mathcal{H}^{N-1}(E) \leq \liminf_{n \to \infty} \mathcal{H}^{N-1}(E_n). \tag{2.16}$$

The following compactness property was also proven in Dal Maso et al. [2005].

Theorem 2.35 (Compactness for σ-convergence Dal Maso et al. [2005])**.** *For every sequence $\{E_n\} \subset \Omega$, if the sequence of its Hausdorff measures $\{\mathcal{H}^{N-1}(E_n)\}$ is bounded, there is a σ-convergent subsequence of $\{E_n\}$.*

We will also frequently use the following relation.

Lemma 2.36. *For two sets $A, B \subset \mathbb{R}^N$ with $\mathcal{H}^{N-1}(A) \leq \infty$ and $\mathcal{H}^{N-1}(B) \leq \infty$, if $B \tilde{\subset} A$ and $\mathcal{H}^{N-1}(A) \leq \mathcal{H}^{N-1}(B)$, then $B \tilde{=} A$.*

Proof. Because $B \tilde{\subset} A$, by definition $\mathcal{H}^{N-1}(B \setminus A) = 0$. Rewriting the sets as $A = (A \setminus B) \cup (A \cap B)$ and $B = (B \setminus A) \cup (B \cap A)$, we can follow

$$\begin{aligned}
\mathcal{H}^{N-1}(A \setminus B) + \mathcal{H}^{N-1}(A \cap B) &= \mathcal{H}^{N-1}(A) \\
&\leq \mathcal{H}^{N-1}(B) \\
&= \mathcal{H}^{N-1}(B \setminus A) + \mathcal{H}^{N-1}(B \cap A) \\
&= \mathcal{H}^{N-1}(B \cap A).
\end{aligned}$$

Therefore $\mathcal{H}^{N-1}(A \setminus B) = 0$, which yields $\mathcal{H}^{N-1}((A\setminus B) \cup (B\setminus A)) = 0$. $\square$

Chapter 3

Regularizing with *a priori* structural information

In this chapter we study the approximation of solutions of ill-posed inverse problems given via operator equations

$$A(f) = g^\delta$$

through the solutions of variational problems

$$\min_{f,K} \left\{ \|A(f) - g^\delta\|^2_{L^2(\Theta)} + \alpha \int_{\Omega \setminus K} |\nabla f|^p dx + \beta \mathcal{H}^{N-1}(K \setminus K^0) + \gamma \mathcal{H}^{N-1}(K \cap K^0) \right\} \tag{3.1}$$

in the presence of *a priori* edge knowledge $K^0 \subset \Omega$. Here $\Omega \subset \mathbb{R}^N$, and $\Theta \subset \mathbb{R}^M$ for $N, M \in \mathbb{N}$ are bounded domains. We consider a continuous operator $A : L^2(\Omega) \to L^2(\Theta)$, scalar parameters $0 \leq \gamma < \beta$, $0 < \alpha$ and $p > 1$. We will discuss existence and regularity of minimizers, stability with respect to the data and parameters and present two parameter choice rules with which the approach yields a regularization.

Throughout this thesis we assume that

(i) K^0 is a compact subset of Ω,

(ii) $\mathcal{H}^{N-1}(K^0) < \infty$.

The *a priori* edge K^0 gives information where the true signal is likely to be discontinuous. This information can be obtained from a secondary modality that is less ill-posed, an application specific template or a reconstruction at a previous time point.

Let $\mathcal{K}$ be defined as the set of all closed subsets in Ω. For brevity we denote the Mumford-Shah type penalty term as

$$\Psi_{K^0,\alpha,\beta,\gamma}(f,K) := \alpha \int_{\Omega\setminus K} |\nabla f|^p dx + \beta\mathcal{H}^{N-1}(K \setminus K^0) + \gamma\mathcal{H}^{N-1}(K \cap K^0) \quad (3.2)$$

for $f \in W^{1,p}(\Omega\setminus K)$, $K \in \mathcal{K}$.

The objective functional then is

$$\mathrm{MS}_{g,K^0,\alpha,\beta,\gamma}(f,K) := \|A(f) - g\|^2_{L^2(\Theta)} + \Psi_{K^0,\alpha,\beta,\gamma}(f,K) \quad (3.3)$$

for $f \in W^{1,p}(\Omega\setminus K)$, $K \in \mathcal{K}$.

If the parameters α, β, γ or the data g are fixed and there is no chance of confusion, we will use shorter notations such as $\mathrm{MS}_{K^0}(f,K)$ and $\Psi_{K^0}(f,K)$.

The main difficulty in studying the Mumford-Shah functional lies in the different nature of f and K with their respective parts in the penalty term. The function f is defined on a N-dimensional domain and K is a singular set in N dimensions. To show the existence of a minimizer the direct approach would be to take a minimizing sequence $(f_n, K_n)_n \in W^{1,p}(\Omega\setminus K_n) \times \mathcal{K}$:

$$\lim_{n\to\infty} \mathrm{MS}(f_n, K_n) = \inf_{(f,K)\in W^{1,p}(\Omega\setminus K)\times\mathcal{K}} \mathrm{MS}(f,K)$$

and try to extract a subsequence, still denoted as $(f_n, K_n)_n$, converging to a pair (f, K) so that

$$\mathrm{MS}(f,K) \le \liminf_{n\to\infty} \mathrm{MS}(f_n, K_n)$$

by some lower semicontinuity property. The approach fails as the map

$$K \mapsto \mathcal{H}^{N-1}(K)$$

(and also the modified edge term) is not lower semicontinuous with respect to any topology that is weak enough to obtain a convergent subsequence simply from the boundedness of $\{\mathcal{H}^{N-1}(K_n)\}$, see David [2005].

Nevertheless in the original paper Mumford and Shah [1989] the existence of a minimizer was conjectured and a first proof was given shortly after in De Giorgi et al. [1989].

In this chapter we follow ideas of De Giorgi et al. [1989] by considering a relaxed version of the Mumford-Shah type functional on $SBV(\Omega)$. For $f \in SBV(\Omega)$ we define

$$\overline{\Psi}_{K^0,\alpha,\beta,\gamma}(f) := \alpha \int_\Omega |\nabla f|^p dx + \beta \mathcal{H}^{N-1}(S_f \setminus K^0) + \gamma \mathcal{H}^{N-1}(S_f \cap K^0), \qquad (3.4)$$

where ∇f is the density of the Lebesgue integrable part of Df and S_f is the jump set. Because of the p-integral over the gradient, for $f \in SBV(\Omega)$ it can also be $\overline{\Psi}_{K^0,\alpha,\beta,\gamma}(f) = +\infty$. We allow this as we are considering a minimization problem and thus only are interested in functions for which $\overline{\Psi}_{K^0,\alpha,\beta,\gamma}(f) < +\infty$.

The ***weak Mumford-Shah functional*** on $SBV(\Omega)$ is defined as

$$\overline{\text{MS}}_{g,K^0,\alpha,\beta,\gamma}(f) := \|A(f) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha,\beta,\gamma}(f). \qquad (3.5)$$

We again use the more compact notation $\overline{\text{MS}}_{K^0}(f)$ and $\overline{\Psi}_{K^0}(f)$ if α, β, γ or the data g are fixed and there is no chance of confusion. In contrast, the functional (3.3) will be called the ***strong Mumford-Shah functional***.

To show a desired property for the strong Mumford-Shah functional, the proofs follow the same strategy.

1. Prove the desired feature for minimizers of the weak version (3.5) using the compactness and semicontinuity theorem in $SBV(\Omega)$.

2. Follow by the regularity results of De Giorgi et al. [1989] that the desired property also holds for strong minimizers $(f, K) \in W^{1,p}(\Omega \backslash K) \times \mathcal{K}$.

The second step is necessary as it is not *a priori* clear that minimizers of the weak functional are well behaved in the context of image segmentation, see Example 2.28.

3.1 Existence

In this section results on the existence of a minimizer of $\mathrm{MS}_{g,K^0,\alpha,\beta,\gamma}(f,K)$ and $\overline{\mathrm{MS}}_{g,K^0,\alpha,\beta,\gamma}(f)$ are presented. A short overview on existence for the original Mumford-Shah functional can be found in Fusco [2003]. We will use the short notations $\mathrm{MS}_{K^0}(f,K)$ and $\overline{\mathrm{MS}}_{K^0}(f)$ for (3.3) and (3.5) respectively.

Fornasier et al. [2011] showed that restrictions are necessary to establish existence of a minimizer for Mumford-Shah type functionals with non trivial operators. We study our problem with a boundedness constraint, that is we consider functions f in the set

$$X_a^b(\Omega) = \{f \in L^\infty(\Omega) : \ a \leq f \leq b \ \text{ a.e. in } \Omega\} \tag{3.6}$$

where $-\infty < a < b < \infty$ are constants. This constraint was also used in Jiang et al. [2014]; Rondi [2008b]; Rondi and Santosa [2001].

By applying the direct method in the calculus of variations, the compactness and lower semicontinuity Theorem 2.25 with Corollary 2.26 yield the existence of a minimizer for the weak setting in $SBV(\Omega)$.

Lemma 3.1. *Let $p > 1$, $\alpha > 0$, $\beta > \gamma \geq 0$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ be fixed with $\mathcal{H}^{N-1}(K^0) < \infty$ and $A : L^2(\Omega) \to L^2(\Theta)$ be a continuous operator. For every $g \in L^\infty(\Theta)$, there exists at least one minimizer of the weak Mumford-Shah functional (3.5) in $SBV(\Omega) \cap X_a^b(\Omega)$.*

Proof. Let $\{f_n\}_n \in SBV(\Omega) \cap X_a^b(\Omega)$ be a minimizing sequence:

$$\lim_{n\to\infty} \overline{\mathrm{MS}}_{K^0}(f_n) = \inf_{u \in SBV} \overline{\mathrm{MS}}_{K^0}(u) < +\infty.$$

By the minimality of the sequence and the pointwise boundedness there exists a constant $C > 0$ so that for every $n \in \mathbb{N}$

$$\|f_n\|_{L^\infty} + \int_\Omega |\nabla f_n|^p dx + \mathcal{H}^{N-1}(S_{f_n}) \leq C.$$

From Theorem 2.35 and Corollary 2.26 it follows: There exists a function $f \in$

$SBV(\Omega) \cap X_a^b(\Omega)$ and a subsequence, still denoted with $(f_n)_n$, so that

$$f_n \to f \quad \text{in } L^1(\Omega),$$
$$\overline{\Psi}_{K^0}(f) \leq \liminf_{n\to\infty} \overline{\Psi}_{K^0}(f_n).$$

As $f, f_n \in X_a^b(\Omega)$ we also have by Fatou's lemma that $f_n \to f$ in $L^q(\Omega)$, for any $1 \leq q < \infty$.

By the continuity of the data fitting term with regards to $L^2(\Omega)$ convergence it follows

$$\begin{aligned}\overline{\mathrm{MS}}_{K^0}(f) &= \|A(f) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0}(f) \\ &\leq \liminf_{n\to\infty} \left[\|A(f_n) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0}(f_n)\right] = \inf_{u \in SBV} \overline{MS}_{K^0}(u).\end{aligned}$$

Therefore the function f is a minimizer of the weak Mumford-Shah functional in $SBV(\Omega) \cap X_a^b(\Omega)$. □

In general a special function of bounded variation f may have a complex jump set S_f. For example, there are functions $f \in SBV(\Omega)$ for which $\overline{S_f} = \Omega$, see Example 2.28. Therefore it is necessary to find conditions that ensure certain regularity of the jump set. In De Giorgi et al. [1989] it was proven that if the fidelity term decays with higher order than the penalty terms, then the jump set of minimizers of the weak Mumford-Shah functional is essentially closed, that is

$$\mathcal{H}^{N-1}(\overline{S_f} \backslash S_f \cap \Omega) = 0. \tag{3.7}$$

This is sufficient for weak minimizers $f \in SBV(\Omega)$ to induce a strong minimizer $(f, \overline{S_f}) \in W^{1,p}(\Omega \backslash \overline{S_f}) \times \mathcal{K}$. We summarize the above in the following definition and theorem.

Definition 3.2. *We say that the fidelity term* $f \mapsto \|Af - g\|^2_{L^2(\Theta)}$ decays with order $k > 0$ for pointwise bound functions *if for some constant* $C > 0$*, for every ball* $B_\rho \subset\subset \Omega$ *of radius* ρ*, and for every* $f, v \in SBV(\Omega) \cap X_a^b(\Omega)$ *with* supp $(f - v) \subset B_\rho$ *it is* $|\, \|A(f) - g\|^2_{L^2(\Theta)} - \|A(v) - g\|^2_{L^2(\Theta)} \,| \leq C\rho^k$.

Theorem 3.3. *Let* $p > 1$*,* $\alpha > 0$*,* $\beta > \gamma \geq 0$*,* $-\infty < a < b < \infty$*,* $K^0 \subset\subset \Omega$ *be fixed with* $\mathcal{H}^{N-1}(K^0) < \infty$ *and* $A : L^2(\Omega) \to L^2(\Theta)$ *be a continuous operator. If* *for some $\varepsilon > 0$ the fidelity term $f \mapsto \|A(f) - g\|^2_{L^2(\Theta)}$ decays with order $N - 1 + \varepsilon$ for pointwise bound functions, then there exists at least one minimizing pair of the Mumford-Shah functional (3.3) in $W^{1,p}(\Omega \backslash K) \cap X_a^b(\Omega) \times \mathcal{K}$.*

Proof. As Ω is bounded and $p > 1$, for every $(f, K) \in W^{1,p}(\Omega \backslash K) \cap X_a^b \times \mathcal{K}$ with $\mathrm{MS}_{K^0}(f, K) < +\infty$ it holds that $f \in SBV(\Omega) \cap X_a^b$ and $S_f \subset K$. Therefore we directly have

$$\begin{aligned} &\min\left\{\overline{\mathrm{MS}}_{K^0}(f) : f \in SBV(\Omega) \cap X_a^b\right\} \\ &\qquad \leq \inf\{\mathrm{MS}_{K^0}(f, K) : (f, K) \in W^{1,p}(\Omega \backslash K) \cap X_a^b \times \mathcal{K}\}. \end{aligned}$$

For a minimizer f^* of the weak Mumford-Shah functional from Theorem 3.8 we have that

$$\mathcal{H}^{N-1}(\overline{S_{f^*}} \backslash S_{f^*} \cap \Omega) = 0.$$

Therefore we can follow

$$\begin{aligned} &\overline{\mathrm{MS}}_{K^0}(f^*) \\ &= \|A(f^*) - g\|^2_{L^2(\Theta)} + \alpha \int_\Omega |\nabla f^*|^p dx + \beta \mathcal{H}^{N-1}(S_{f^*} \setminus K^0) + \gamma \mathcal{H}^{N-1}(S_{f^*} \setminus K^0) \\ &= \|A(f^*) - g\|^2_{L^2(\Theta)} + \alpha \int_{\Omega \backslash \overline{S_{f^*}}} |\nabla f^*|^p dx + \beta \mathcal{H}^{N-1}(\overline{S_{f^*}} \setminus K^0) + \gamma \mathcal{H}^{N-1}(\overline{S_{f^*}} \setminus K^0) \\ &= \mathrm{MS}_{K^0}(f^*, \overline{S_{f^*}}). \end{aligned}$$

Due to $\overline{\mathrm{MS}}_{K^0}(f^*) < \infty$ the weak minimizer f^* is contained in $W^{1,p}(\Omega \backslash \overline{S_f}) \cap X_a^b$ and therefore the last equality above is valid. As a result

$$\begin{aligned} &\min\left\{\overline{\mathrm{MS}}_{K^0}(f) : f \in SBV(\Omega) \cap X_a^b\right\} \\ &\qquad = \min\{\mathrm{MS}_{K^0}(f, K) : (f, K) \in W^{1,p}(\Omega \backslash K) \cap X_a^b \times \mathcal{K}\}, \end{aligned}$$

with $(f^*, \overline{S_{f^*}})$ being a minimizer of the strong Mumford-Shah functional. □

The above proof showed, that once the regularity of the jump set is ensured through Theorem 3.8, weak and strong minimizers are essentially the same. In

Proposition 3.10 we will give conditions on the operator A that ensure a decay of the least squares fidelity term with order $N-1+\varepsilon$ for pointwise bound functions.

Example 3.4 (Minimizers are not unique)**.** *Let* $\Omega = [-1, 1]$*,* $A = Id$*,* $\alpha = \beta = 1$ $\gamma = 0$*,* $K^0 = \emptyset$*,* $\lambda \in \mathbb{R}$*,* $p = 2$ *and* $h : \Omega \to \mathbb{R}$ *be the Heaviside function*

$$h(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

The corresponding one dimensional Mumford-Shah functional to the signal $g = \lambda h$ *is defined as*

$$MS(f, K) := \int\limits_{\Omega} (f - \lambda h)^2 \, dx + \int\limits_{\Omega \backslash K} |f'(x)|^2 dx + \mathcal{H}^0(K)$$

for $(f, K) \in W^{1,2}(\Omega \backslash K) \cap L^\infty(\Omega) \times \mathcal{K}$*. In one dimension the discontinuity set* K *is a finite set of points and* $\mathcal{H}^0$ *is the counting measure.*

There are only two possible candidates for a minimizer:

1. $(f_1, K_1) = (\lambda h, \{0\})$ *with*

$$MS(\lambda h, \{0\}) = \mathcal{H}^0(\{0\}) = 1,$$

2. $(f_2, K_2) = (f^*, \emptyset)$*, where* f^* *is the uniquely determined minimizer of*

$$G(f) = \int\limits_{\Omega} (\lambda h - f)^2 \, dx + \int\limits_{\Omega} |f'(x)|^2 dx$$

in $W^{1,2}([-1, 1])$*.*

By regularity results (see for example [David, 2005, p.17 Lemma 13]) the function f^* *is in* $C^1(\Omega)$ *and fulfills the linear equation*

$$\begin{aligned} f'' + f &= \lambda h \quad \text{on } [-1, 1] \\ f'(-1) &= f'(1) = 0 \end{aligned}$$

in a weak sense. We can find a fundamental solution f_0 so that, for any $\lambda > 0$, the solution can be represented as $f^* = \lambda f_0$. Computing $MS(f^*, \emptyset)$ therefore is,

$$MS(f^*, \emptyset) = MS(\lambda f_0, \emptyset) = \lambda^2 \int\limits_{\Omega} (h - f_0)^2 \, dx + \lambda^2 \int\limits_{\Omega} |f_0'(x)|^2 dx = \lambda^2 C,$$

for some $C > 0$ independent of λ. Therefore, if $\lambda > \frac{1}{\sqrt{C}}$, then $(f_1, K_1) = (\lambda h, \{0\})$ is the unique minimizer. If $\lambda < \frac{1}{\sqrt{C}}$, then $(f_2, K_2) = (f^*, \emptyset)$ is the unique minimizer. And if $\lambda = \frac{1}{\sqrt{C}}$, then there are exactly two minimizers (f_1, K_1) and (f_2, K_2).

The fact that minimizers are not unique matches the empiric experience, that in some cases even for human vision more than one segmentation can be good.

Definition 3.5. *For $g \in L^\infty(\Theta)$, $p > 1$, $\alpha > 0$, $\beta > \gamma \geq 0$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ with $\mathcal{H}^{N-1}(K^0) < \infty$ and A continuous from $L^2(\Omega)$ to $L^2(\Theta)$ we define $\overline{M}_{g,K^0,\alpha,\beta,\gamma}$ and $M_{g,K^0,\alpha,\beta,\gamma}$ as the set of minimizers of (3.5) in $SBV(\Omega) \cap X_a^b$ and (3.3) in $W^{1,p}(\Omega \backslash K) \cap X_a^b(\Omega) \times \mathcal{K}$ respectively.*

3.2 Regularity of K

In this section we will recall conditions for the essential closedness of the edge set following [Ambrosio et al., 2000, Chapter 7]. The essential closedness is necessary to conclude the existence proof of the previous section.

A central point in the regularity theory of the edges is that at small scales minimizers of the Mumford-Shah functional, for certain data fitting terms, behave similar to minimizers of

$$F(f) = \alpha \int_{\Omega} |\nabla f|^2 dx + \beta \mathcal{H}^{N-1}(S_f).$$

For this it is necessary that the data fitting term $f \mapsto \|A(f) - g\|^2_{L^2(\Theta)}$ becomes negligible at small scales, see Fusco [2003].

If we keep the *a priori* edge K^0 fixed, then the proofs in [Ambrosio et al., 2000, Chapter 7] can also be applied to our setting by considering

$$F_{K^0}(f) = \alpha \int_\Omega |\nabla f|^p dx + \beta \mathcal{H}^{N-1}(S_f \setminus K^0) + \gamma \mathcal{H}^{N-1}(S_f \cap K^0).$$

The main arguments of the proof use the scaling property of the penalty terms, that do not change, by having different but fixed weights $\beta > 0$ and $\gamma > 0$ on different parts of the domain. As $K^0 \subset\subset \Omega$ with $\mathcal{H}^{N-1}(K^0) < \infty$ the case $\gamma = 0$ can be rewritten as a classic Mumford-Shah functional on the new open domain $\Omega \setminus K^0$ and is therefore also covered by [Ambrosio et al., 2000, Chapter 7].

To give a more precise meaning to this the following definitions are needed.

Definition 3.6. *Let the function $f \in SBV_{loc}(\Omega)$ and the ball $\overline{B_\rho(x)} \subset \Omega$ be given. We call a function $v \in SBV_{loc}(\Omega)$ a* **competitor of** f **in** $B_\rho(x)$ *if f and v only differ inside of $B_\rho(x)$, i.e. $\overline{\operatorname{supp}(f - v)} \subset B_\rho(x)$.*

We define the functional

$$\mathcal{F}_{K^0}(f, U) := \alpha \int_U |\nabla f|^p dx + \beta \mathcal{H}^1\big((S_f \cap U) \setminus K^0\big) + \gamma \mathcal{H}^1\big((S_f \cap U) \cap K^0\big) \tag{3.8}$$

for every $f \in SBV(\Omega)$ and open set $U \subset \Omega$.
With this it is possible to define the notion of *quasiminimality*.

Definition 3.7 (see [Ambrosio et al., 2000, p. 339 Definition 7.2 and p. 350 7.17])**.** *We say that a function $f \in SBV_{loc}(\Omega)$ is a* **quasiminimizer** *of $\mathcal{F}_{K^0}$ in Ω if there exists constants $\omega, \varepsilon \geq 0$ such that for all balls $B_\rho(x) \subset\subset \Omega$ and all competitors $v \in SBV_{loc}(\Omega)$ of f in $B_\rho(x)$ it is*

$$\mathcal{F}_{K^0}(f, B_\rho(x)) \leq \mathcal{F}(v, B_\rho(x)) + \omega \rho^{N-1+\varepsilon}. \tag{3.9}$$

Theorem 3.8. *Let f be a function in $SBV(\Omega)$. If f is a quasiminimizer of $\mathcal{F}_{K^0}$ in Ω, then*

$$\mathcal{H}^{N-1}(\overline{S_f} \setminus S_f \cap \Omega) = 0. \tag{3.10}$$

Proof. The proof works the same as [Ambrosio et al., 2000, p. 351 Theorem 7.21] together with [Ambrosio et al., 2000, p. 78 Theorem 2.56]. □

See also Babadjian and Giacomini [2013] following [Ambrosio et al., 2000, Chap. 7] for the proof of existence and regularity of a minimizer to a similar variational problem.

The most important example of a quasiminimizer of $\mathcal{F}_{K^0}$ are minimizers of the weak Mumford-Shah functional.

Proposition 3.9. *Let the function f be a minimizer of the weak Mumford-Shah functional (3.5) in $SBV(\Omega) \cap X_a^b(\Omega)$. If for some $\varepsilon > 0$ the fidelity term $f \mapsto \|A(f) - g\|^2_{L^2(\Theta)}$ decays with order $N - 1 + \varepsilon$, then f is a quasiminimizer of $\mathcal{F}_{K^0}$.*

We will use the short notations $\overline{\Psi}_{K^0}$ for (3.4).

Proof. Let $f \in SBV(\Omega) \cap X_a^b(\Omega)$ be a minimizer of the weak Mumford-Shah functional and $v \in SBV(\Omega)$ be a competitor of f in the ball $\overline{B_\rho(x)} \subset \Omega$.

With the minimality of f it is

$$\|A(f) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0}(f) \leq \|A(v) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0}(v).$$

As v is also a competitor of f in $B_\rho(x)$ we follow

$$\begin{aligned}&\mathcal{F}_{K^0}(f, B_\rho(x))\\ &\leq \|A(v) - g\|^2_{L^2(\Theta)} - \|A(f) - g\|^2_{L^2(\Theta)} + \mathcal{F}_{K^0}(v, B_\rho(x)).\end{aligned}$$

The decay of the fidelity term then leads to the claim. □

This also concludes the existence proof of the previous section.

The following proposition is helpful to verify the decay property for least squares penalty terms.

Proposition 3.10. *Let $A : L^2(\Omega) \to L^2(\Theta)$ be the forward operator and $g \in L^\infty(\Theta)$ be the measured data. If there exist exponents $1 \leq q$ and $1 \leq \hat{q}, q' \leq \infty$*

with $\frac{1}{\hat{q}} + \frac{1}{q'} = 1$ *such that for all functions* $f, v \in SBV(\Omega) \cap X_a^b(\Omega)$ *it holds*

$$\|A(f) - A(v)\|_{L^{\hat{q}}(\Theta)} \leq L\|f - v\|_{L^q(\Omega)} \tag{3.11}$$

and

$$\|A(f) + A(v)\|_{L^{q'}(\Theta)} \leq C \tag{3.12}$$

for some constants L *and* $C > 0$*, then the fidelity term* $f \mapsto \|A(f) - g\|^2_{L^2(\Theta)}$ *decays with order* $\frac{N}{q}$ *for pointwise bound functions.*

Proof. We fix a ball $B_\rho \subset\subset \Omega$ of radius ρ. For any $f, v \in SBV(\Omega) \cap X_a^b(\Omega)$ with $\overline{\text{supp }(f - v)} \subset B_\rho$ it holds

$$\begin{aligned}\|A(f) - g\|^2_{L^2(\Theta)} - \|A(v) - g\|^2_{L^2(\Theta)} &= \int_\Theta (A(f) + A(v) - 2g)(A(f) - A(v)) \\ &\leq \|A(f) + A(v) - 2g\|_{L^{q'}(\Theta)}\|A(f) - A(v)\|_{L^{\hat{q}}(\Theta)} \\ &\leq (C + 2\|g\|_{L^{q'}(\Theta)})L\|f - v\|_{L^q(\Omega)}.\end{aligned}$$

As $\overline{\text{supp }(f - v)} \subset B_\rho$ and $f, v \in X_a^b(\Omega)$ we conclude

$$\|A(f) - g\|^2_{L^2(\Theta)} - \|A(v) - g\|^2_{L^2(\Theta)} \leq \tilde{C}\|f - v\|_{L^q(B_\rho)} \leq \tilde{\tilde{C}}\rho^{N/q}$$

for some $\tilde{C}, \tilde{\tilde{C}} > 0$. □

As a first example we show that image deblurring fits into our framework. In Chapter 4 we will show that the lemma above can also be applied to X-ray CT and 2 dimensional diffuse optical tomography.

Example 3.11. *A classical imaging task is image deconvolution, that is restoring an image from its blurred version, see Bertero and Boccacci [1998]. Given a blurring kernel* $\phi \in L^1(\Omega)$*, we define for* $1 \leq q \leq \infty$ *the forward operator* $A : L^q(\Omega) \to L^q(\Omega)$ *as*

$$Af(x) = (\phi * f)(x) := \int_\Omega f(y)\phi(x - y)dy.$$

The operator A is linear and bounded for $1 \leq q \leq \infty$. We have for f, $v \in SBV(\Omega) \cap X_a^b(\Omega)$

$$\|\phi * f - \phi * v\|_{L^1(\Omega)} \leq L\|f - v\|_{L^1(\Omega)}$$

and

$$\|\phi * f + \phi * v\|_{L^\infty(\Omega)} \leq C\|f + v\|_{L^\infty(\Omega)} \leq \tilde{C},$$

for some constants L, $\tilde{C}$, $C > 0$. By Proposition 3.10, with exponents $q = \hat{q} = 1$ and $q' = \infty$, the fidelity term decays with order N for pointwise bound functions.

3.3 Stability

In this section we prove stability with respect to the data and parameters.

Again we first consider the weak setting on $SBV(\Omega) \cap X_a^b(\Omega)$ for continuous operators $A: L^2(\Omega) \to L^2(\Theta)$. If additionally the data fitting term decays with order $N - 1 + \varepsilon$ for pointwise bound functions for some $\varepsilon > 0$, then the results also hold in the strong setting.

Lemma 3.12. *Let $p > 1$, $\alpha > 0$, $\beta > \gamma \geq 0$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ be fixed with $\mathcal{H}^{N-1}(K^0) < \infty$ and A be continuous from $L^2(\Omega)$ to $L^2(\Theta)$.*

Assume we have a converging sequence of data $g_n \overset{L^2}{\to} g$, $g_n \in L^\infty(\Theta)$ ($n = 1, \dots$). Then every sequence of minimizers $\{f_n\}$, with $f_n \in \overline{M}_{g_n,K^0,\alpha,\beta,\gamma}$, converges subsequentially to a minimizer $f^* \in \overline{M}_{g,K^0,\alpha,\beta,\gamma}$ in $L^1(\Omega)$. Moreover $\overline{MS}_{g,K^0,\alpha,\beta,\gamma}(f^*) = \lim_n \overline{MS}_{g_n,K^0,\alpha,\beta,\gamma}(f_n)$.

As the parameters α, β, γ are fixed we will use the shorter notations $\overline{\Psi}_{K^0}$ and $\overline{\text{MS}}_{g,K^0}$.

Proof. Because of the minimality of f_n, we have that for every $n \in \mathbb{N}$,

$$\begin{aligned}\overline{\Psi}_{K^0}(f_n) &\leq \|A(f_n) - g_n\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0}(f_n) \\ &\leq \|A(a) - g_n\|^2_{L^2(\Theta)},\end{aligned}$$

where a denotes the function of constant value a for which $\overline{\Psi}_{K^0}(a) = 0$. Since $g_n \stackrel{L^2}{\to} g$, there is some constant $C > 0$ such that

$$\sup_{n\in\mathbb{N}} \left\{ \|f_n\|_{L^\infty(\Omega)} + \int_\Omega |\nabla f_n|^p dx + \mathcal{H}^{N-1}(S_{f_n}) \right\} \leq \max\{|a|, |b|\} + C.$$

Corollary 2.26 and Theorem 2.35 yield a function $f^* \in SBV(\Omega) \cap X_a^b(\Omega)$ and a set $E \subset \Omega$, so that a subsequence of $\{f_n\}$, still denoted by $\{f_n\}$, weakly converges to f^* in SBV and $\{S_{f_n}\}$ σ-converges to E. Because Ω is a bounded domain and $p > 1$, the weak convergence $\nabla f_n \rightharpoonup \nabla f \quad$ in $L^p(\Omega)$ implies weak convergence in $L^1(\Omega)$. By the semicontinuity conclusion (2.10) of Corollary 2.26, it follows that

$$\overline{\Psi}_{K^0}(f^*) \leq \liminf_{n\to\infty} \left(\overline{\Psi}_{K^0}(f_n) \right). \tag{3.13}$$

Because f_n and $f^* \in X_a^b(\Omega)$, the sequence $\{f_n\}$ actually converges to f^* in $L^q(\Omega)$ for every $1 \leq q < \infty$, and consequently $A(f_n) \stackrel{L^2(\Theta)}{\to} A(f^*)$. Together we have

$$\begin{aligned} &\|A(f^*) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0}(f^*) \\ &\qquad \leq \lim_{n\to\infty} \|A(f_n) - g_n\|^2_{L^2(\Theta)} + \liminf_{n\to\infty} \left(\overline{\Psi}_{K^0}(f_n) \right). \end{aligned} \tag{3.14}$$

Let $v \in SBV(\Omega) \cap X_a^b(\Omega)$. Then comparing $\overline{\mathrm{MS}}_{g_n,K^0}(f_n)$ with $\overline{\mathrm{MS}}_{g_n,K^0}(v)$ yields the desired minimality of f^* :

$$\begin{aligned} \|A(f^*) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0}(f^*) &\leq \liminf_{n\to\infty} \left(\|A(v) - g_n\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0}(v) \right) \\ &= \|A(v) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0}(v). \end{aligned}$$

Setting $v = f^*$ leads to $\overline{\mathrm{MS}}_{g,K^0}(f^*) = \lim_n \overline{\mathrm{MS}}_{g_n,K^0}(f_n)$ for the subsequence $\{f_n\}$. The procedure can be repeated to obtain the convergence of function values for the entire sequence. □

Next we prove that for the classic Mumford-Shah regularization, that is $K^0 = \emptyset$, also the edges of minimizers converge for converging data.

Lemma 3.13. *Let the same assumptions as in Lemma 3.12 hold. Additionally let $K^0 = \emptyset$, then the edge sets $\{S_{f_n}\}$ σ-converge to S_{f^*}.*

Proof. We take the same converging subsequence $\{f_n\}$ with limit function $f^* \in SBV(\Omega) \cap X_a^b(\Omega)$ as in the proof of Lemma 3.12. Furthermore we have the set $E \subset \Omega$, for which $\{S_{f_n}\}$ σ-converges to E.

It remains to show that $E \tilde{=} S_{f^*}$. As $\{f_n\}$ weakly converges to f^* in $SBV(\Omega)$ it follows from Definition 2.34 (i) of σ-convergence that $S_{f^*} \tilde{\subset} E$. Just as in obtaining (3.14), by the lower semicontinuity (2.16) of the Hausdorff measure regarding σ-convergence, it follows that

$$\begin{aligned} \|A(f^*) - g\|^2_{L^2(\Theta)} + \alpha \int_\Omega |\nabla f^*|^p dx + \beta \mathcal{H}^{N-1}(E) \\ \leq \lim_{n\to\infty} \|Af_n - g_n\|^2_{L^2(\Theta)} + \liminf_{n\to\infty} \left(\alpha \int_\Omega |\nabla f_n|^p dx + \beta \mathcal{H}^{N-1}(S_{f_n}) \right) \\ \leq \|A(f^*) - g\|^2_{L^2(\Theta)} + \alpha \int_\Omega |\nabla f^*|^p dx + \beta \mathcal{H}^{N-1}(S_{f^*}). \end{aligned}$$

Therefore, $\mathcal{H}^{N-1}(E) \leq \mathcal{H}^{N-1}(S_{f^*})$, which together with $S_{f^*} \tilde{\subset} E$ yields $E \tilde{=} S_f^*$, by Lemma 2.36. □

Theorem 3.14 (Stability with respect to g)**.** *Let $p > 1$, $\alpha > 0$, $\beta > \gamma \geq 0$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ be fixed with $\mathcal{H}^{N-1}(K^0) < \infty$ and A be continuous from $L^2(\Omega)$ to $L^2(\Theta)$. Additionally let for some $\varepsilon > 0$ the fidelity term $f \mapsto \|A(f) - g\|^2_{L^2(\Theta)}$ decay with order $N - 1 + \varepsilon$ for pointwise bound functions.*

Assume we have a converging sequence of data $g_n \stackrel{L^2}{\to} g$, $g_n \in L^\infty(\Theta)$ ($n = 1, \dots$). Then for every sequence of minimizers $\{(f_n, K_n)\}$, with $(f_n, K_n) \in M_{g_n,K^0,\alpha,\beta,\gamma}$, there exists a pair $(f^*, K^*) \in M_{g,K^0,\alpha,\beta,\gamma}$, for which $\{f_n\}$ converges subsequentially to f^* in $L^1(\Omega)$.

Moreover $MS_{g,K^0,\alpha,\beta,\gamma}(f^*, K^*) = \lim_n MS_{g_n,K^0,\alpha,\beta,\gamma}(f_n, K_n)$.

Proof. For $n \in \mathbb{N}$, by the regularity of the edge sets K_n, it follows that f_n are in $SBV(\Omega) \cap X_a^b(\Omega)$ with $K_n \tilde{=} \overline{S_{f_n}}$. Furthermore we have $f_n \in \overline{M}_{g_n,K^0,\alpha,\beta,\gamma}$. Using Lemma 3.12 above there exists a function $f^* \in \overline{M}_{g,K^0,\alpha,\beta,\gamma}$ so that for a subsequence of $\{f_n\}$ it is $f_n \stackrel{L^1}{\to} f^*$ and $\overline{MS}_{g,K^0}(f^*) = \lim_n \overline{MS}_{g_n,K^0}(f_n)$. Again by regularity of the edge sets it follows that $(f^*, \overline{S_{f^*}})$ is a minimizer in $W^{1,p}(\Omega \backslash \overline{S_{f^*}}) \cap X_a^b(\Omega) \times \mathcal{K}$ to the data g for the strong Mumford-Shah functional (3.3) and $MS_{g,K^0}(f^*, \overline{S_{f^*}}) = \lim_n MS_{g_n,K^0}(f_n, \overline{S_{f_n}})$. We arrive at the claim by setting $K^* = \overline{S_{f^*}}$. □

Corollary 3.15 (Stability with respect to g)**.** *Let the same conditions as in Theorem 3.14 hold. Additionally let $K^0 = \emptyset$, then the edges $\{K_n\}$ σ-converge to K^*.*

Proof. We take the same converging subsequence $\{f_n\}$ with limit function $f^* \in SBV(\Omega) \cap X_a^b(\Omega)$ as above. Furthermore we have the same sequence $\{K_n\}$ and the set $K^* = \overline{S_{f^*}}$.

It remains to show that $K_n \xrightarrow{\sigma} K^*$. Again for $n \in \mathbb{N}$, by the regularity of the edge sets K_n, it follows that f_n are in $SBV(\Omega) \cap X_a^b(\Omega)$ with $K_n \tilde{=} \overline{S_{f_n}} \tilde{=} S_{f_n}$. By Lemma 3.13 we have that S_{f_n} σ-converge to S_{f^*}. As $K_n \tilde{=} S_{f_n}$ and $K^* \tilde{=} S_{f^*}$ it follows that K_n σ-converge to K^*. □

Now let us consider stability with respect to the parameters α, β and γ.

Lemma 3.16 (Stability with respect to (α, β, γ))**.** *Let $p > 1$, $g \in L^\infty(\Theta)$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ be fixed with $\mathcal{H}^{N-1}(K^0) < \infty$ and A be continuous from $L^2(\Omega)$ to $L^2(\Theta)$.*

Assume we have converging parameters $(\alpha_n, \beta_n, \gamma_n) \to (\alpha^*, \beta^*, \gamma^*)$ with $\alpha_n \geq \alpha_0 > 0$, $\beta_n \geq \beta_0 > 0$ and $\beta_n > \gamma_n \geq 0$. Then every sequence of minimizers $\{f_n\}$, with $f_n \in \overline{M}_{g,K^0,\alpha_n,\beta_n,\gamma_n} (n = 1, \cdots)$, converges subsequentially to a minimizer $f^* \in \overline{M}_{g,K^0,\alpha^*,\beta^*,\gamma^*}$ in $L^1(\Omega)$. Moreover $\overline{MS}_{g,K^0,\alpha^*,\beta^*,\gamma^*}(f^*) = \lim_n \overline{MS}_{g,K^0,\alpha_n,\beta_n,\gamma_n}(f_n)$.

Proof. Let $(\alpha_n, \beta_n, \gamma_n) \to (\alpha^*, \beta^*, \gamma^*)$ be a positive sequence as above and $\{f_n\}$ a sequence of respective minimizers, that is $f_n \in \overline{M}_{g,K^0,\alpha_n,\beta_n,\gamma_n}$.

Because of the minimality of f_n, we have that for every $n \in \mathbb{N}$,

$$\begin{aligned}\overline{\Psi}_{K^0,\alpha_n,\beta_n,\gamma_n}(f_n) &\leq \|A(f_n) - g\|_{L^2(\Theta)}^2 + \overline{\Psi}_{K^0,\alpha_n,\beta_n,\gamma_n}(f_n) \\ &\leq \|A(a) - g\|_{L^2(\Theta)}^2,\end{aligned}$$

where a denotes the function of constant value a for which $\overline{\Psi}_{K^0,\alpha_n,\beta_n,\gamma_n}(a) = 0$. Since α_n and β_n are bounded from below there is a constant $C > 0$ such that

$$\sup_{n\in\mathbb{N}} \left\{ \|f_n\|_{L^\infty(\Omega)} + \int_\Omega |\nabla f_n|^p dx + \mathcal{H}^{N-1}(S_{f_n}) \right\} \leq \max\{|a|, |b|\} + C. \tag{3.15}$$

Corollary 2.26 and Theorem 2.35 yield a function $f^* \in SBV(\Omega) \cap X_a^b(\Omega)$ and a set $E \subset \Omega$, so that a subsequence of $\{f_n\}$, still denoted by $\{f_n\}$, weakly converges to f^* in SBV and $\{S_{f_n}\}$ σ-converges to E.

For an arbitrary fixed function $v \in SBV \cap X_a^b(\Omega)$ we have

$$\begin{aligned}\overline{\mathrm{MS}}_{g,K^0,\alpha^*,\beta^*,\gamma^*}(f^*) &= \|A(f^*) - g\|^2 + \Psi_{K^0,\alpha^*,\beta^*,\gamma^*}(f^*) \\ &\leq \liminf_n \left(\|A(f_n) - g\|^2 + \Psi_{\alpha^*,\beta^*,\gamma^*,K^0}(f_n)\right) \\ &= \liminf_n \left(\|A(f_n) - g\|^2 + \Psi_{\alpha_n,\beta_n,\gamma_n,K^0}(f_n)\right) \\ &\leq \liminf_n \left(\|A(v) - g\|^2 + \Psi_{\alpha_n,\beta_n,\gamma_n,K^0}(v)\right) \\ &= \|A(v) - g\|^2 + \Psi_{K^0,\alpha^*,\beta^*,\gamma^*}(v).\end{aligned}$$

This shows that $f^* \in \overline{M}_{g,K^0,\alpha^*,\beta^*,\gamma^*}$.

By setting $v = f^*$ we get $\lim_n \overline{MS}_{g,K^0,\alpha_n,\beta_n,\gamma_n}(f_n) = \overline{MS}_{g,K^0,\alpha^*,\beta^*,\gamma^*}(f^*)$. □

Next we again consider the special case of classic Mumford-Shah regularization.

Lemma 3.17 (Stability with respect to (α, β, γ)). *Let the same assumptions as in Lemma 3.16 hold. Additionally let $K^0 = \emptyset$, then the edge sets $\{S_{f_n}\}$ σ-converge to S_{f^*}.*

Proof. We take the same converging subsequence $\{f_n\}$ with limit function $f^* \in SBV(\Omega) \cap X_a^b(\Omega)$ as above. Furthermore we have the set $E \subset \Omega$, for which $\{S_{f_n}\}$ σ-converges to E.

It remains to show that $E \tilde{=} S_{f^*}$. As $\{f_n\}$ weakly converges to f^* in $SBV(\Omega)$ it follows from Definition 2.34 (i) of σ-convergence that $S_{f^*} \tilde{\subset} E$. Just as in obtaining (3.14), by the lower semicontinuity (2.16) of the Hausdorff measure regarding σ-

convergence, it follows that

$$\begin{aligned}
&\|A(f^*)-g\|^2_{L^2(\Theta)}+\alpha^*\int_\Omega |\nabla f^*|^p dx+\beta^*\mathcal{H}^{N-1}(E)\\
&\qquad\le \liminf_{n\to\infty}\left(\|Af_n-g\|^2_{L^2(\Theta)}+\alpha_n\int_\Omega|\nabla f_n|^p dx+\beta_n\mathcal{H}^{N-1}(S_{f_n})\right)\\
&\qquad\le \liminf_{n\to\infty}\left(\|A(f^*)-g\|^2_{L^2(\Theta)}+\alpha_n\int_\Omega|\nabla f^*|^p dx+\beta_n\mathcal{H}^{N-1}(S_{f^*})\right)\\
&\qquad= \|A(f^*)-g\|^2_{L^2(\Theta)}+\alpha^*\int_\Omega|\nabla f^*|^p dx+\beta^*\mathcal{H}^{N-1}(S_{f^*}).
\end{aligned}$$

Therefore, $\mathcal{H}^{N-1}(E)\le\mathcal{H}^{N-1}(S_{f^*})$, which together with $S_{f^*}\tilde{\subset}E$ yields $E\tilde{=}S_{f^*}$, by Lemma 2.36. □

We again can formulate the strong version for the case when the fidelity term decays fast enough in small balls.

Theorem 3.18 (Stability with respect to (α,β,γ)). *Let $p>1$, $g\in L^\infty(\Theta)$, $-\infty<a<b<\infty$, $K^0\subset\subset\Omega$ be fixed with $\mathcal{H}^{N-1}(K^0)<\infty$ and A be continuous from $L^2(\Omega)$ to $L^2(\Theta)$. Additionally let for some $\varepsilon>0$ the fidelity term $f\mapsto\|A(f)-g\|^2_{L^2(\Theta)}$ decay with order $N-1+\varepsilon$ for pointwise bound functions.*

Assume we have converging parameters $(\alpha_n,\beta_n,\gamma_n)\to(\alpha^*,\beta^*,\gamma^*)$ with $\alpha_n\ge\alpha_0>0$, $\beta_n\ge\beta_0>0$ and $\beta_n>\gamma_n\ge0$. Then for every sequence of minimizers $\{(f_n,K_n)\}$, with $(f_n,K_n)\in M_{g,K^0,\alpha_n,\beta_n,\gamma_n}(n=1,\dots)$, there exists a pair $(f^*,K^*)\in M_{g,K^0,\alpha^*,\beta^*,\gamma^*}$ for which $\{f_n\}$ converges subsequentially to f^* in $L^1(\Omega)$. Moreover $MS_{g,K^0,\alpha^*,\beta^*,\gamma^*}(f^*,K^*)=\lim_n MS_{g,K^0,\alpha_n,\beta_n,\gamma_n}(f_n,K_n)$.

Proof. The proof is the same as for Theorem 3.14. □

Corollary 3.19 (Stability with respect to (α,β,γ)). *Let the same assumptions as in the above Theorem 3.18 hold. Additionally let $K^0=\emptyset$, then the edges $\{K_n\}$ σ-converge to K^*.*

Proof. The proof is the same as for Corollary 3.15. □

3.4 Monotonicity

In this section we will discuss monotonicity of the penalty and the fidelity term for parameters (α, β, γ) where for some constants $C_\beta, C_\gamma > 0$

$$\beta = C_\beta \alpha, \qquad \text{and} \qquad \gamma = C_\gamma \alpha.$$

Thus the problem is reduced to a single parameter setting. We will use the monotonicity in Section 3.6 for the discrepancy principle and only state it here for the weak setting on $SBV(\Omega)$. The results can be transferred to the strong setting as usual. Similar work has been done Anzengruber and Ramlau [2010] where the discrepancy principle was studied for inverse problems with non-linear operators.

Before we state our monotonicity results, we give an example where both decreasing and increasing the parameters α and β (but each to a different degree) yield an increase in the residual.

Example 3.20. *Let $\Omega = [-1, 1]$, $A = Id$, $\alpha > 0$, $\beta > 0$, $\gamma = 0$, $K^0 = \emptyset$, $p = 2$, $\lambda \in \mathbb{R}$ and $h : \Omega \to \mathbb{R}$ be the Heaviside function*

$$h(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

As before the one dimensional Mumford-Shah functional is

$$MS(f, K) := \int_{\Omega} (\lambda h - f)^2 \, dx + \alpha \int_{\Omega \setminus K} |f'(x)|^2 dx + \beta \mathcal{H}^0(K)$$

for $(f, K) \in W^{1,2}(\Omega \setminus K) \cap L^\infty \times \mathcal{K}$. Let λ be chosen such that $(\lambda h, \{0\})$ is the unique minimizer. This is possible, as shown in Example 3.4. As the first variable of the minimizer $(\lambda h, \{0\})$ equals the signal, the residual is zero.

We will now show that there are parameters $\alpha, \beta < 1$ and $\alpha, \beta > 1$ for which the residual is greater than zero.

For any pair of parameters $(\alpha, \beta) > 0$ there are only two possible candidates for a minimizer: $(f_1, K_1) = (\lambda h, \{0\})$ with $MS(\lambda h, \{0\}) = \beta$ or $(f_2, K_2) = (f^*, \emptyset)$

where f^* is the uniquely determined minimizer of

$$G(f) = \int_\Omega (\lambda h - f)^2 \, dx + \alpha \int_\Omega |f'(x)|^2 dx$$

in $W^{1,2}([-1,1])$. It is well known that for $\alpha \to 0$ it is $G(f^*) \to 0$.

We first consider $\alpha, \beta < 1$. Let us fix $\beta < 1$, we can than choose α so small, that $G(f^*) < \beta$. It follows that $(f^*, \emptyset)$ is the unique minimizer of $MS(f,K)$. As $f^* \in W^{1,2}([-1,1]) \subset C^0([-1,1])$ and λh is a step function, the residual is strictly greater than zero.

Now consider $\alpha, \beta > 1$. Let us fix $\alpha > 1$. We can then set $\beta := G(f^*) + 1$. It again follows that $(f^*, \emptyset)$ is the unique minimizer of $MS(f,K)$.

We will follow the presentation in Anzengruber and Ramlau [2010]. For $g \in L^\infty(\Theta)$, $\alpha > 0$, $C_\beta, C_\gamma > 0$ and $K^0 \subset \Omega$ we use the notation:

$$\begin{aligned} d(\alpha) &:= \{\|A(f) - g\|_{L^2(\Theta)} | f \in \overline{M}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}\}, \\ \psi(\alpha) &:= \{\overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f) | f \in \overline{M}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}\}\}, \\ J(\alpha) &:= \overline{\text{MS}}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}(f), \;\; f \in \overline{M}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}. \end{aligned} \tag{3.16}$$

As the minimizers are not unique the functions d and ψ are set-valued.

Lemma 3.21. *For every $C_\beta, C_\gamma > 0$ the maps $\alpha \mapsto d(\alpha)$ and $\alpha \mapsto J(\alpha)$ are* non-decreasing *and the map $\alpha \mapsto \psi(\alpha)$ is* non-increasing *in the sense that for $0 < \alpha_1 < \alpha_2$ we have*

$$J(\alpha_1) \le J(\alpha_2), \tag{3.17}$$

$$\sup d(\alpha_1) \le \inf d(\alpha_2), \tag{3.18}$$

$$\inf \psi(\alpha_1) \ge \sup \psi(\alpha_2). \tag{3.19}$$

Proof. Let $g \in L^\infty(\Theta)$, $0 < \alpha_1 < \alpha_2$, $f_1 \in \overline{M}_{g,\alpha_1,C_\beta\alpha_1,C_\gamma\alpha_1,K^0}$ and $f_2 \in \overline{M}_{g,\alpha_2,C_\beta\alpha_2,C_\gamma\alpha_2,K^0}$ be arbitrary but fixed. By the minimality of f_1 we have

$$\begin{aligned} J(\alpha_1) &= \|A(f_1) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha_1,C_\beta\alpha_1,C_\gamma\alpha_1}(f_1) \\ &\le \|A(f_2) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha_1,C_\beta\alpha_1,C_\gamma\alpha_1}(f_2). \end{aligned}$$

The first inequality then follows as $\alpha_1 < \alpha_2$ by

$$\begin{aligned} J(\alpha_1) &\leq \|A(f_2) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha_1,C_\beta\alpha_1,C_\gamma\alpha_1}(f_2) \\ &\leq \|A(f_2) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha_2,C_\beta\alpha_2,C_\gamma\alpha_2}(f_2) \\ &= J(\alpha_2). \end{aligned}$$

For the other two inequalities, we start with

$$\|A(f_1) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha_1,C_\beta\alpha_1,C_\gamma\alpha_1}(f_1) \leq \|A(f_2) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha_1,C_\beta\alpha_1,C_\gamma\alpha_1}(f_2)$$

which leads to

$$\frac{1}{\alpha_1}\big[\|A(f_1) - g\|^2 - \|A(f_2) - g\|^2\big] \leq \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f_2) - \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f_1).$$

And in the same way

$$\|A(f_2) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha_2,C_\beta\alpha_2,C_\gamma\alpha_2}(f_2) \leq \|A(f_1) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha_2,C_\beta\alpha_2,C_\gamma\alpha_1}(f_1)$$

leads to

$$\overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f_2) - \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f_1) \leq \frac{1}{\alpha_2}\big[\|A(f_1) - g\|^2 - \|A(f_2) - g\|^2\big].$$

We can combine these inequalities and obtain

$$\begin{aligned} \frac{1}{\alpha_1}(\|A(f_1) - g\|^2 - \|A(f_2) - g\|^2) &\leq \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f_2) - \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f_1) \\ &\leq \frac{1}{\alpha_2}(\|A(f_1) - g\|^2 - \|A(f_2) - g\|^2). \end{aligned}$$

As $0 < \frac{1}{\alpha_2} < \frac{1}{\alpha_1}$ we can follow that $\|A(f_1) - g\|^2 - \|A(f_2) - g\|^2 \leq 0$ and therefore

$$\|A(f_1) - g\|^2 \leq \|A(f_2) - g\|^2.$$

Furthermore we obtain

$$\overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f_2) \leq \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f_1).$$

As the functions f_1, f_2 are chosen arbitrarily we obtain the desired monotonicity. □

Lemma 3.22. *For every $C_\beta, C_\gamma > 0$ the discontinuity set*

$$A := \{\alpha > 0 | \inf d(\alpha) < \sup d(\alpha)\} \tag{3.20}$$

has at most countable many points. The same holds for ψ and the respective sets of discontinuity points coincide.

Proof. For every $\alpha \in A$ the set $d(\alpha)$ has at least two values and consequently the interval $(\inf d(\alpha), \sup d(\alpha))$ is not empty and contains a rational number. Due to the monotonicity Lemma 3.21, for different α_1 and α_2 in A the open intervals $(\inf d(\alpha_1), \sup d(\alpha_1))$ and $(\inf d(\alpha_2), \sup d(\alpha_2))$ are disjoint. As the rational numbers are countable, the set A has also countable many points at most.

As for two minimizers $f_1, f_2 \in \overline{M}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}$ it holds

$$\|A(f_2) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}(f_2) = \|A(f_1) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}(f_1),$$

it is clear that $d(\alpha)$ is set-valued if and only if $\psi(\alpha)$ is set-valued. □

3.5 Regularization with an *a priori* parameter choice

In this section we prove an *a priori* parameter choice rule with which our approach yields a regularization for the image and its edges.

As before we will first state the result for the weak version on $SBV(\Omega)$ and later lift it to the strong version by the regularity of the edge set.

Lemma 3.23. *Let $A : L^2(\Omega) \to L^2(\Theta)$ be continuous, $g \in L^\infty(\Theta)$, $p > 1$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ with $\mathcal{H}^{N-1}(K^0) < \infty$ and $f^\dagger \in SBV \cap X^b_a(\Omega)$ be such that $A(f^\dagger) = g$.*

Assume that the noisy data $g^\delta \in L^\infty(\Theta)$ with $\|g - g^\delta\|_{L^2(\Theta)} \le \delta$. Let the regularization parameters be chosen so that as $\delta \to 0$,

$$\alpha(\delta) \to 0, \quad \beta(\delta) \to 0, \quad \gamma(\delta) \to 0, \tag{3.21}$$

$$\frac{\delta^2}{\min\{\alpha(\delta), \beta(\delta), \gamma(\delta)\}} \to 0 \quad \textit{and} \quad \frac{\max\{\alpha(\delta), \beta(\delta), \gamma(\delta)\}}{\min\{\alpha(\delta), \beta(\delta), \gamma(\delta)\}} \to C, \tag{3.22}$$

for some $C > 0$ and $\alpha(\delta) > 0$, $\beta(\delta) > \gamma(\delta) > 0$. For any sequence $\delta_n \to 0$ let f^{δ_n} $(n = 1, \dots)$ be in $\overline{M}_{g,K^0,\alpha(\delta),\beta(\delta),\gamma(\delta)}$. Then

- **(i)** *there exists a function $f^* \in SBV(\Omega) \cap X_a^b(\Omega)$ and a convergent subsequence of $\{f^{\delta_n}\}$, still denoted as $\{f^{\delta_n}\}$, such that $\{f^{\delta_n}\}$ converges weakly to f^* in $SBV(\Omega)$;*
- **(ii)** *it holds $A(f^*) = g$;*
- **(iii)** *for every other solution $\phi \in SBV(\Omega) \cap X_a^b(\Omega)$ of the operator equation $A(f) = g$, it is*

$$\int_\Omega |\nabla f^*|^p dx + \mathcal{H}^{N-1}(S_{f^*}) \le C\Big(\int_\Omega |\nabla \phi|^p dx + \mathcal{H}^{N-1}(S_\phi)\Big). \tag{3.23}$$

Proof. We first prove (i). For every $n \in \mathbb{N}$, comparing the given true solution $f^\dagger$ with the minimizer f^{δ_n} yields

$$\begin{aligned} &\|A(f^{\delta_n}) - g^{\delta_n}\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha(\delta),\beta(\delta),\gamma(\delta),K^0}(f^{\delta_n}) \\ &\qquad \le \|A(f^\dagger) - g^{\delta_n}\|^2_{L^2(\Theta)} + \overline{\Psi}_{K^0,\alpha(\delta),\beta(\delta),\gamma(\delta)}(f^\dagger) \\ &\qquad \le \delta_n^2 + \overline{\Psi}_{K^0,\alpha(\delta),\beta(\delta),\gamma(\delta)}(f^\dagger). \end{aligned} \tag{3.24}$$

This leads to

$$\begin{aligned} &\int_\Omega |\nabla f^{\delta_n}|^p dx + \mathcal{H}^{N-1}(S_{f^{\delta_n}}) \\ &\le \frac{\delta_n^2}{\min\{\alpha(\delta_n), \beta(\delta_n), \gamma(\delta)\}} + \frac{\max\{\alpha(\delta_n), \beta(\delta_n), \gamma(\delta)\}}{\min\{\alpha(\delta_n), \beta(\delta_n), \gamma(\delta)\}}\Big(\int_\Omega |\nabla f^\dagger|^p dx + \mathcal{H}^{N-1}(S_{f^\dagger})\Big). \end{aligned} \tag{3.25}$$

As $\delta_n \to 0$, by definition of the parameters, the right hand side of (3.25) is bounded. Therefore, there exists a constant $\tilde{C} > 0$, so that:

$$\sup_{n\in\mathbb{N}} \left\{ \|f^{\delta_n}\|_{L^\infty(\Omega)} + \int_\Omega |\nabla f^{\delta_n}|^p dx + \mathcal{H}^{N-1}(S_{f^{\delta_n}}) \right\} \leq \max\{|a|, |b|\} + \tilde{C}. \tag{3.26}$$

Again Corollary 2.26 and 2.35 yield a function $f^* \in SBV(\Omega) \cap X_a^b(\Omega)$ and a set $E \subset \Omega$, so that there is a subsequence of $\{f^{\delta_n}\}$, still denoted by $\{f^{\delta_n}\}$, that weakly converges to f^* in SBV and $\{S_{f^{\delta_n}}\}$ σ-converges to E.

Next we prove (ii). Using (3.24) we have

$$\|A(f^{\delta_n}) - g^{\delta_n}\|^2_{L^2(\Theta)} \leq \delta_n^2 + \overline{\Psi}_{\alpha(\delta),\beta(\delta),\gamma(\delta),K^0}(f^\dagger). \tag{3.27}$$

Because $f_{f^{\delta_n}}$ and $f^* \in X_a^b(\Omega)$, the sequence $\{f_{f^{\delta_n}}\}$ actually converges to f^* in $L^q(\Omega)$ for every $1 \leq q < \infty$, and consequently $A(f_{f^{\delta_n}}) \xrightarrow{L^2} A(f^*)$. Therefore, due to the parameter choice rule for $n \to \infty$, it follows that

$$\|A(f^*) - g\|^2_{L^2(\Theta)} = \lim_{n\to\infty} \|A(f^{\delta_n}) - g^{\delta_n}\|^2_{L^2(\Theta)} = 0, \tag{3.28}$$

which proves the convergence to a solution of $A(f) = g$.

To prove (iii), denote that (3.25) holds for any ϕ in place of $f^\dagger$, provided $\phi \in SBV(\Omega) \cap X_a^b(\Omega)$ is a solution of $A(f) = g$. Using the lower semicontinuity Theorem 2.25 and (3.25), we obtain

$$\begin{aligned}
&\int_\Omega |\nabla f^*|^p dx + \mathcal{H}^{N-1}(S_{f^*}) && (3.29)\\
\leq &\liminf_{n\to\infty} \Big(\int_\Omega |\nabla f^{\delta_n}|^p dx + \mathcal{H}^{N-1}(S_{f^{\delta_n}}) \Big)\\
\leq &\liminf_{n\to\infty} \Big(\frac{\delta_n^2}{\min\{\alpha(\delta_n), \beta(\delta_n)\}} + \frac{\max\{\alpha(\delta_n), \beta(\delta_n)\}}{\min\{\alpha(\delta_n), \beta(\delta_n)\}} \Big(\int_\Omega |\nabla \phi|^p dx + \mathcal{H}^{N-1}(S_\phi) \Big) \Big)\\
= &C \Big(\int_\Omega |\nabla \phi|^p dx + \mathcal{H}^{N-1}(S_\phi) \Big). && (3.30)
\end{aligned}$$

□

Corollary 3.24. *Let the same assumptions as in the above Theorem 3.23 hold. If additionally $C = 1$, then the edge sets $\{S_{f^{\delta_n}}\}$ σ-converge to S_{f^*}.*

Proof. We take the same converging subsequence $\{f^{\delta_n}\}$ with limit function $f^* \in$

$SBV(\Omega) \cap X_a^b(\Omega)$ as as in the proof above. Furthermore we have the set $E \subset \Omega$, for which $\{S_{f^{\delta_n}}\}$ σ-converges to E.

It remains to show that $E \tilde{=} S_{f^*}$. Since $\{f^{\delta_n}\}$ weakly converges to f^* in $SBV(\Omega)$, it follows from Definition 2.34 (i) that $S_{f^*} \tilde{\subset} E$. As f^* is a solution of the operator equation, we can replace $f^\dagger$ by f^* in (3.25). Then by using the lower semicontinuity (2.16) and the same argument as from (3.29) to (3.30), we obtain

$$\int_\Omega |\nabla f^*|^p dx + \mathcal{H}^{N-1}(E) \leq \int_\Omega |\nabla f^*|^p dx + \mathcal{H}^{N-1}(S_{f^*}). \tag{3.31}$$

Therefore $\mathcal{H}^{N-1}(E) \leq \mathcal{H}^{N-1}(S_{f^*})$, which together with $S_{f^*} \tilde{\subset} E$ yields $E \tilde{=} S_{f^*}$, by Lemma 2.36. □

Theorem 3.25 (Regularization)**.** *Let the same notations and parameter choice rule from Lemma 3.23 hold with $C = 1$. Additionally suppose that for some $\varepsilon > 0$, the fidelty term $f \mapsto \|A(f) - g^\delta\|^2_{L^2(\Theta)}$ decays with order $N - 1 + \varepsilon$ for pointwise bound functions.*

For any sequence $\delta_n \to 0$ let $(f^{\delta_n}, K^{\delta_n})$ be in $M_{g^\delta, K^0, \alpha(\delta), \beta(\delta), \gamma(\delta)}$. Then

- **(i)** *there exists a pair $(f^*, K^*) \in W^{1,p}(\Omega \backslash K^*) \cap X_a^b(\Omega) \times \mathcal{K}$ and a convergent subsequence of $\{f^{\delta_n}, K^{\delta_n}\}$, still denoted as $\{f^{\delta_n}, K^{\delta_n}\}$, such that $f^{\delta_n} \stackrel{L^1}{\to} f^*$ and $\{K^{\delta_n}\}$ σ-converge to K^*;*
- **(ii)** *it holds $A(f^*) = g$;*
- **(iii)** *for every other solution $\phi \in W^{1,p}(\Omega \backslash K_\phi) \cap X_a^b(\Omega)$ of the operator equation, it is*

$$\int_{\Omega \backslash K^*} |\nabla f^*|^p dx + \mathcal{H}^{N-1}(K^*) \leq \int_{\Omega \backslash K_\phi} |\nabla \phi|^p dx + \mathcal{H}^{N-1}(K_\phi), \tag{3.32}$$

where K_ϕ is any suitable compact set so that $\phi \in W^{1,p}(\Omega \backslash K_\phi)$.

Proof. The proof is an application of Lemma 3.23 with the same arguments as in the proof of Theorem 3.14. □

Remark 3.26. *In Rondi [2008b] an a priori parameter choice rule $\alpha(\delta) = \beta(\delta) = \gamma(\delta) = c_1\delta^{c_2}$ was studied for the Mumford-Shah functional, where $c_1, c_2 > 0$. The fidelity terms considered in Rondi [2008b] are powers of distance functions other than least-squares functionals. Our result on the a priori parameter choice rule is an extension in several aspects. In Rondi [2008b] convergence is only obtained for the image f, in contrast we characterize when the edges converge in the sense of σ-convergence. Furthermore we consider the strong setting and our parameter choice is more general.*

3.6 Regularization with the discrepancy principle

In this section we prove that a parameter choice via Morozovs discrepancy principle yields a regularization for our approach under certain restrictions on the parameters. We again refer to Anzengruber and Ramlau [2010] for related work.

We will make use of the results in Section 3.21, and therefore restrict ourselves to parameters (α, β, γ) for which

$$\beta = C_\beta \alpha \qquad \text{and} \qquad \gamma = C_\gamma \alpha \tag{3.33}$$

for some constants $C_\gamma, C_\beta > 0$.

We use the following version of the discrepancy principle.

Definition 3.27 (MDP)**.** *Let $g \in L^\infty(\Theta)$, $p > 1$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ with $\mathcal{H}^{N-1}(K^0) < \infty$.*

Let $\tau_2 \geq \tau_1 > 1$. For $\delta > 0$ and $g^\delta \in L^2(\Theta)$ with $\|g - g^\delta\|_{L^2(\Theta)} \leq \delta$ we say that (α, β, γ) are chosen according to Morozov's discrepancy principle (MDP) if there exists $f^\delta \in \overline{M}_{g^\delta, K^0, \alpha, \beta, \gamma}$ such that

$$\tau_1 \delta \leq \|A(f^\delta) - g^\delta\|_{L^2(\Theta)} \leq \tau_2 \delta. \tag{3.34}$$

Morozov's discrepancy principle fails if it is not possible to find parameters such that (3.34) is true for at least one minimizer. This is the case, for example, when the uncorrupted data is obtained from a function $f^\dagger$ for which

$\overline{\Psi}_{K^0,\alpha,\beta,\gamma}(f^\dagger) = 0$. Because then for any (α, β, γ) we have

$$\min_{f \in SBV \cap X_a^b(\Omega)} \overline{\mathrm{MS}}_{g^\delta,K^0,\alpha,\beta,\gamma}(f) \leq \|A(f^\dagger) - g^\delta\|^2_{L^2(\Theta)} + 0 \leq \delta^2 < \tau_1 \delta^2. \tag{3.35}$$

For the weak Mumford-Shah penalty on $SBV(\Omega)$ constant functions are in the kernel.

We first study in which situation it is possible to choose (α, β, γ) according to MDP.

Lemma 3.28. *Let $A : L^2(\Omega) \to L^2(\Theta)$ be continuous, $g \in L^\infty(\Theta)$, $p > 1$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ with $\mathcal{H}^{N-1}(K^0) < \infty$.*

For fixed $C_\beta, C_\gamma > 0$, to each $\alpha > 0$ there exist $f_1, f_2 \in \overline{M}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}$ so that for $\alpha_n \uparrow \alpha$ and any sequence $\{f_n\} \in \overline{M}_{g,K^0,\alpha_n,C_\beta\alpha_n,C_\gamma\alpha_n}$ it holds

$$\begin{aligned}\lim_n \|A(f_n) - g\|_{L^2(\Theta)} &= \|A(f_1) - g\|_{L^2(\Theta)} \\ &= \min_{f \in \overline{M}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}} \|A(f) - g^\delta\|_{L^2(\Theta)}.\end{aligned}$$

In the same way for $\alpha_n \downarrow \alpha$ and any sequence $\{f_n\} \in \overline{M}_{g,K^0,\alpha_n,C_\beta\alpha_n,C_\gamma\alpha_n}$ it holds

$$\begin{aligned}\lim_n \|A(f_n) - g\|_{L^2(\Theta)} &= \|A(f_2) - g\|_{L^2(\Theta)} \\ &= \max_{f \in \overline{M}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}} \|A(f) - g^\delta\|_{L^2(\Theta)}.\end{aligned}$$

Proof. Let$\{\alpha_n\}$ be a positive strictly increasing sequence converging to α and $f_n \in \overline{M}_{g,K^0,\alpha_n,C_\beta\alpha_n,C_\gamma\alpha_n}$ be a corresponding sequence of arbitrary minimizers. Lemma 3.16 yields a $L^2(\Omega)$ convergent subsequence $\{f_n\}$ with limit $f_1 \in \overline{M}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}$. Using the monotonicity Lemma 3.21 of the fidelity term with respect to α we get for $\alpha_n \uparrow \alpha$

$$\begin{aligned}\|A(f_1) - g^\delta\|_{L^2(\Theta)} &= \lim_n \|A(f_n) - g^\delta\|_{L^2(\Theta)} \\ &\leq \inf_{f \in \overline{M}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}} \|A(f) - g\|_{L^2(\Theta)} \leq \|A(f_1) - g\|_{L^2(\Theta)}.\end{aligned}$$

We can repeat the reasoning for every subsequence of $\{f_n\}$ and obtain for the entire sequence

$$\lim_n \|A(f_n) - g\|_{L^2(\Theta)} = \min_{f \in \overline{M}_{g,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}} \|A(f) - g\|_{L^2(\Theta)}. \tag{3.36}$$

The claim for $\{\alpha_n\}$ strictly decreasing can be shown in exactly the same way. □

We now show that in many cases it is possible to choose α so that the data fitting term is either very large or vanishes.

Lemma 3.29. *Let $A : L^2(\Omega) \to L^2(\Theta)$ be continuous, $g \in L^\infty(\Theta)$, $p > 1$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ with $\mathcal{H}^{N-1}(K^0) < \infty$ and $f^\dagger \in SBV \cap X_a^b(\Omega)$ be such that $A(f^\dagger) = g$.*

Assume that the noisy data $g^\delta \in L^\infty(\Theta)$ with $\|g - g^\delta\|_{L^2(\Theta)} \leq \delta$ and that g^δ satisfies $\|A(C) - g^\delta\| > \tau_2\delta$ for all constant functions $C(x) = C$.

Then we can find parameters α_1, $\alpha_2 > 0$ and respective minimizers $f_1 \in \overline{M}_{g^\delta,K^0,\alpha_1,C_\beta\alpha_1,C_\gamma\alpha_1}$ and $f_2 \in \overline{M}_{g^\delta,K^0,\alpha_2,C_\beta\alpha_2,C_\gamma\alpha_2}$ such that

$$\|A(f_1) - g^\delta\|_{L^2(\Theta)} \leq \tau_1\delta \leq \tau_2\delta \leq \|A(f_2) - g^\delta\|_{L^2(\Theta)} \tag{3.37}$$

for $\tau_2 \geq \tau_1 > 1$.

Proof. First consider $\alpha_n \to 0$ and corresponding minimizers $f_n \in \overline{M}_{g^\delta,K^0,\alpha_n,C_\beta\alpha_n,C_\gamma\alpha_n}$. We have for each $n \in \mathbb{N}$

$$\|A(f_n) - g^\delta\|^2_{L^2(\Theta)} \leq \overline{\text{MS}}_{g^\delta,K^0,\alpha_n,C_\beta\alpha_n,C_\gamma\alpha_n}(f^\dagger) = \delta^2 + \alpha_n\overline{\Psi}_{K^0,1,C_\beta,,C_\gamma}(f^\dagger) \to \delta^2. \tag{3.38}$$

As $\tau_1 > 1$, for small enough α_n, we therefore have $f_n \in \overline{M}_{g^\delta,K^0,\alpha_n,C_\beta\alpha_n,C_\gamma\alpha_n}$ with $\|A(f_n) - g^\delta\|_{L^2(\Theta)} < \tau_1\delta$.

On the other hand assume that $\alpha_n \to \infty$, then for each $n \in \mathbb{N}$

$$\begin{aligned}\overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f_n) &\leq \frac{1}{\alpha_n}\overline{\text{MS}}_{g^\delta,K^0,\alpha_n,C_\beta\alpha_n,C_\gamma\alpha_n}(f_n) \\ &\leq \frac{1}{\alpha_n}\|A(a) - g^\delta\|^2_{L^2(\Theta)} \to 0.\end{aligned}$$

Therefore $\{\int_\Omega |\nabla f_n|^p dx + \mathcal{H}^{N-1}(S_{f_n})\} \to 0$, which together with the $L^\infty(\Omega)$ bound by Corollary 2.26 implies that $\{f_n\}$ has a convergent subsequence that converges to a function $f^* \in SBV(\Omega)$ with $\int_\Omega |\nabla f^*|^p dx + \mathcal{H}^{N-1}(S_{f^*}) = 0$, i.e. f^* is a constant function. With the assumption that for all constant functions $\|A(C) - g^\delta\|_{L^2(\Theta)} > \tau_2\delta$ this yields

$$\lim_n \|A(f_n) - g^\delta\|_{L^2(\Theta)} = \|A(f^*) - g^\delta\|_{L^2(\Theta)} > \tau_2\delta. \tag{3.39}$$

Therefore for large enough α_n, we have $f_n \in \overline{M}_{g^\delta,K^0,\alpha_n,C_\beta\alpha_n,C_\gamma\alpha_n}$ with $\|A(f_n) - g^\delta\|_{L^2(\Theta)} > \tau_2\delta$. □

We summarize the assumptions on the data.

Assumption 3.30. *Assume that for $\delta > 0$ and $\tau_2 \geq \tau_1 > 1$ the measured data $g^\delta \in L^2(\Theta)$ satisfies*

$$\|g - g^\delta\|_{L^2(\Theta)} \leq \delta < \tau_2\delta < \|A(C) - g^\delta\|_{L^2(\Theta)}, \tag{3.40}$$

for all constant functions $C(x) = C \in \mathbb{R}$. Moreover assume that there is no parameter $\alpha > 0$ with minimizers $f_1, f_2 \in \overline{M}_{g^\delta,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}$, such that

$$\|A(f_1) - g^\delta\|_{L^2(\Theta)} \leq \tau_1\delta \leq \tau_2\delta \leq \|A(f_2) - g^\delta\|_{L^2(\Theta)}. \tag{3.41}$$

Theorem 3.31. *Let $A : L^2(\Omega) \to L^2(\Theta)$ be continuous, $g \in L^\infty(\Theta)$, $p > 1$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ with $\mathcal{H}^{N-1}(K^0) < \infty$ and $f^\dagger \in SBV \cap X_a^b(\Omega)$ be such that $A(f^\dagger) = g$.*

If Assumption 3.30 is fulfilled then there exist parameters (α, β, γ) fulfilling Morozovs Discrepancy Principle (3.34).

Proof. Assume no (α, β, γ) exists that fulfills the MDP. Define the sets

$$S^1 = \{\alpha > 0 \mid \|A(f) - g^\delta\|_{L^2(\Theta)} < \tau_1\delta \quad \text{for some} \quad f \in \overline{M}_{g^\delta,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}\}, \tag{3.42}$$
$$S^2 = \{\alpha > 0 \mid \|A(f) - g^\delta\|_{L^2(\Theta)} > \tau_2\delta \quad \text{for some} \quad f \in \overline{M}_{g^\delta,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}\}. \tag{3.43}$$

Due to Lemma 3.29 we know the sets S^1 and S^2 are not empty. Note that for $\alpha \in S^1$ it must actually hold $\|A(f) - g^\delta\|_{L^2(\Theta)} < \tau_1\delta$ for all $f \in \overline{M}_{g^\delta,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}$ or else either the MDP would be fulfilled or Assumption 3.30 violated. In the same way we obtain for $\alpha \in S^2$ it holds $\|A(f) - g^\delta\|_{L^2(\Theta)} > \tau_2\delta$ for all $f \in \overline{M}_{g^\delta,K^0,\alpha,C_\beta\alpha,C_\gamma\alpha}$. Therefore we have $S^1 \cap S^2 = \emptyset$ and $S^1 \cup S^2 = \mathbb{R}^+$.

Let us define $\alpha* := \sup S^1$. Then it follows from the monotonicity and Lemma 3.29 that $0 < \alpha^* < \infty$ and therefore α^* is in S^1 or S^2. We treat the cases separately.

If $\alpha^* \in S^1$ then we can choose a strictly decreasing sequence $\alpha_n \to \alpha^*$ and $f_n \in \overline{M}_{g^\delta,K^0,\alpha_n,C_\beta\alpha_n,C_\gamma\alpha_n}$. Since all α_n belong to S^2, with Lemma 3.28 we get

$$\tau_2\delta \leq \lim_n \|A(f_n) - g^\delta\|_{L^2(\Theta)} = \sup_{f\in\overline{M}_{g^\delta,K^0,\alpha^*,C_\beta\alpha^*,C_\gamma\alpha^*}} \|A(f) - g^\delta\|_{L^2(\Theta)} < \tau_1\delta. \tag{3.44}$$

This is a contradiction to $\tau_1 \leq \tau_2$.

If $\alpha^* \in S^2$ we can choose a strictly increasing sequence $\alpha_n \to \alpha^*$ and argument in the same way. □

Now we are in the position to prove the main result of this section.

Lemma 3.32. *Let $A : L^2(\Omega) \to L^2(\Theta)$ be continuous, $g \in L^\infty(\Theta)$, $p > 1$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ with $\mathcal{H}^{N-1}(K^0) < \infty$ and $f^\dagger \in SBV \cap X_a^b(\Omega)$ be such that $A(f^\dagger) = g$. Assume that the noisy data $g^\delta \in L^\infty(\Theta)$ with $\|g - g^\delta\|_{L^2(\Theta)} \leq \delta$ and Assumptions 3.30 are met.*

For $C_\beta, C_\gamma > 0$ and any sequence $\delta_n \to 0$ let $(\alpha_n, \beta_n, \gamma_n) = (\alpha_n, C_\beta\alpha_n, C_\gamma\alpha_n)$ and $f^{\delta_n} \in \overline{M}_{g^\delta,K^0,\alpha_n,C_\beta\alpha_n,C_\gamma\alpha_n}$ be chosen according to MDP. Then

(i) *there exists a function $f^* \in SBV(\Omega) \cap X_a^b(\Omega)$ and a convergent subsequence of $\{f^{\delta_n}\}$, still denoted as $\{f^{\delta_n}\}$, such that $\{f^{\delta_n}\}$ converges weakly to f^* in $SBV(\Omega)$;*

(ii) *it holds $A(f^*) = g$;*

(iii) *for every other solution $\phi \in SBV(\Omega) \cap X_a^b(\Omega)$ of the operator equation $A(f) = g$, it is*

$$\overline{\Psi}_{1,C_\beta,C_\gamma,K^0}(f^*) \leq \overline{\Psi}_{1,C_\beta,C_\gamma,K^0}(\phi). \tag{3.45}$$

Proof. We first prove (i). For every $n \in \mathbb{N}$, using MDP and comparing the given true solution $f^\dagger$ with the minimizer $f^{\delta_n} \in \overline{M}_{g^\delta, K^0, \alpha_n, C_\beta\alpha_n, C_\gamma\alpha_n}$ we have

$$\begin{aligned} &\tau_1\delta^2 + \alpha_n \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^{\delta_n}) \\ &\quad \leq \|A(f^{\delta_n}) - g^{\delta_n}\|^2_{L^2(\Theta)} + \alpha_n \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^{\delta_n}) \\ &\quad \leq \delta_n^2 + \alpha_n \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^\dagger). \end{aligned}$$

This leads to

$$\begin{aligned} &0 \leq (\tau_1 - 1)\delta_n^2 \\ &\quad \leq \alpha_n \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^\dagger) - \alpha_n \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^{\delta_n}). \end{aligned}$$

And we can follow

$$0 \leq (\tau_1 - 1)\frac{\delta^2}{\alpha_n} \leq \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^\dagger) - \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^{\delta_n}). \tag{3.46}$$

As $0 \leq (\tau_1 - 1)\frac{\delta^2}{\alpha_n}$ for all $n \in \mathbb{N}$ it is

$$\overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^{\delta_n}) \leq \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^\dagger). \tag{3.47}$$

Therefore, there exists a constant $\tilde{C} > 0$, so that:

$$\sup_{n\in\mathbb{N}} \left\{ \|f^{\delta_n}\|_{L^\infty(\Omega)} + \int_\Omega |\nabla f^{\delta_n}|^p dx + \mathcal{H}^{N-1}(S_{f^{\delta_n}}) \right\} \leq \max\{|a|,|b|\} + \tilde{C}. \tag{3.48}$$

Again Corollary 2.26 and 2.35 yield a function $f^* \in SBV(\Omega) \cap X_a^b(\Omega)$ and a set $E \subset \Omega$, so that there is a subsequence of $\{f^{\delta_n}\}$, still denoted by $\{f^{\delta_n}\}$, that weakly converges to f^* in SBV and $\{S_{f^{\delta_n}}\}$ σ-converges to E.

Next we prove (ii). Using the MDP we have

$$\|A(f^{\delta_n}) - g^{\delta_n}\|^2_{L^2(\Theta)} \leq \tau_2^2\delta_n^2. \tag{3.49}$$

As in the proof of Lemma 3.12, we have $A(f^{\delta_n}) \xrightarrow{L^2} A(f^*)$. Therefore for $\delta_n \to 0$

it follows that

$$\|A(f^*) - g\|^2_{L^2(\Theta)} \leq \lim_{n\to\infty} (\|A(f^{\delta_n}) - g^{\delta_n}\|^2_{L^2(\Theta)} + \delta_n^2) = 0, \tag{3.50}$$

which proves that the limit f^* is a solution of $A(f) = g$.

To prove (iii), note that (3.47) holds for any ϕ in place of $f^\dagger$, provided $\phi \in SBV(\Omega) \cap X_a^b(\Omega)$ is a solution of the operator equation. Using Corollary 2.26, we obtain

$$\begin{aligned}\overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^*) \leq& \liminf_{n\to\infty} \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(f^{\delta_n}) \leq \liminf_{n\to\infty} \overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(\phi) \\ =&\overline{\Psi}_{K^0,1,C_\beta,C_\gamma}(\phi).\end{aligned} \tag{3.51}$$

□

For the classic Mumford-Shah regularization we can also get convergence of the edges.

Corollary 3.33. *Let the same assumptions as in the above Lemma 3.32 hold. If additionally $C_\beta = C_\gamma = 1$, then the edge sets $\{S_{f^{\delta_n}}\}$ σ-converge to S_{f^*}.*

Proof. We take the same converging subsequence $\{f^{\delta_n}\}$ with limit function $f^* \in SBV(\Omega) \cap X_a^b(\Omega)$ as in the proof above. Furthermore we have the set $E \subset \Omega$, for which $\{S_{f^{\delta_n}}\}$ σ-converges to E.

It remains to show that $E \tilde{=} S_{f^*}$. Since $\{f^{\delta_n}\}$ weakly converges to f^* in $SBV(\Omega)$, it follows from Definition 2.34 (i) that $S_{f^*} \tilde{\subset} E$. Since f^* is a solution of the operator equation, we can replace $f^\dagger$ by f^* in (3.47). By using the lower semicontinuity (2.16) and the same argument as for (3.51), we obtain

$$\int_\Omega |\nabla f^*|^p dx + \mathcal{H}^{N-1}(E) \leq \int_\Omega |\nabla f^*|^p dx + \mathcal{H}^{N-1}(S_{f^*}). \tag{3.52}$$

Therefore $\mathcal{H}^{N-1}(E) \leq \mathcal{H}^{N-1}(S_{f^*})$, which together with $S_{f^*} \tilde{\subset} E$ yields $E \tilde{=} S_{f^*}$ by Lemma 2.36. □

We conclude the chapter by stating the strong version of the regularization result for $C_\beta = C_\gamma = 1$.

Theorem 3.34. *Let $A : L^2(\Omega) \to L^2(\Theta)$ be continuous, $g \in L^\infty(\Theta)$, $p > 1$, $-\infty < a < b < \infty$, $K^0 \subset\subset \Omega$ with $\mathcal{H}^{N-1}(K^0) < \infty$ and $f^\dagger \in SBV \cap X_a^b(\Omega)$ be such that $A(f^\dagger) = g$. Assume that the noisy data $g^\delta \in L^\infty(\Theta)$ with $\|g - g^\delta\|_{L^2(\Theta)} \leq \delta$ and Assumptions 3.30 are met.*

Additionally suppose that for some $\varepsilon > 0$, the fidelity term $f \mapsto \|A(f) - g^\delta\|^2_{L^2(\Theta)}$ decays with order $N - 1 + \varepsilon$ for pointwise bound functions.

For any sequence $\delta_n \to 0$ let $(\alpha_n, \beta_n, \gamma_n) = (\alpha_n, \alpha_n, \alpha_n)$ and $(f^{\delta_n}, K^{\delta_n}) \in M_{g^\delta, K^0, \alpha_n, \alpha_n, \alpha_n}$ be chosen according to the MDP. Then

(i) *there exists a pair $(f^*, K^*) \in W^{1,p}(\Omega \backslash K^*) \cap X_a^b(\Omega) \times \mathcal{E}$ and a convergent subsequence of $\{f^{\delta_n}, K^{\delta_n}\}$, still denoted as $\{f^{\delta_n}, K^{\delta_n}\}$, such that $f^{\delta_n} \overset{L^1}{\to} f^*$ and $\{K^{\delta_n}\}$ σ-converge to K^*;*

(ii) *it holds $A(f^*) = g$;*

(iii) *for every other solution $\phi \in W^{1,p}(\Omega \backslash K_\phi) \cap X_a^b(\Omega)$ of the operator equation $A(f) = g$, we have that*

$$\int_{\Omega \backslash K^*} |\nabla f^*|^p dx + \mathcal{H}^{N-1}(K^*) \leq \int_{\Omega \backslash K_\phi} |\nabla \phi|^p dx + \mathcal{H}^{N-1}(K_\phi), \tag{3.53}$$

where K_ϕ is any suitable compact set so that $\phi \in W^{1,p}(\Omega \backslash K_\phi)$.

Proof. The proof works the same as for Theorem 3.14. □

Chapter 4

A variational approximation in the sense of Γ-convergence

In this chapter we will study variational approximations of the weak Mumford-Shah type functional

$$\overline{\mathrm{MS}}_{K^0}(f,v) = \\ \|A(f)-g\|_{L^2}^2 + \alpha\int_{\Omega\setminus K} |\nabla f|^2 dx + \beta\mathcal{H}^{N-1}(S_f \setminus K^0) \tag{4.1}$$

in the sense of Γ-convergence following the ideas of Ambrosio and Tortorelli [1992]. Compared to the functional (3.5) in the previous chapter, the functional (4.1) above corresponds to the special case where $\gamma = 0$, that is edges coinciding with the *a priori* edge are not penalized at all. We will also motivate a heuristic approximation for $\gamma \neq 0$ at the end of this chapter.

Sometimes it is possible to view a difficult problem, in our case the Mumford-Shah type functional, as the limit of a series of more computational feasible or more understandable problems. In such a scenario it is important to choose the correct notion of convergence. A suitable convergence for our problem is Γ-convergence. Its objective is the description of asymptotic behavior of families of minimum problems, which are often depending on some parameters.

For the approximation, instead of a set K we consider a sequences of smooth edge indicator functions $\{v_n\} \in W^{1,2}(\Omega), 0 \le v_n \le 1$, where $v_n \approx 0$ and $v_n \approx 1$

indicate an edge and no edge respectively. This smoothed setting is referred to as the phase field setting. It is also necessary to describe the *a priori* edge information K^0 in the phase field setting. We assume that for a given K^0 we have a sequence $\{v_n^0\}$ of edge indicator functions that fulfill certain conditions, which we will specify in the next section.

If the spaces are chosen appropriately, by Proposition 2.32, it suffices to derive an approximation of the penalty term. The data fitting term can then be regarded as a continuous perturbation in the sense of Proposition 2.32(ii).

Let us recall that $\Omega \subset \mathbb{R}^N$ is a bounded domain and K^0 is a compact subset of Ω with $\mathcal{H}^{N-1}(K^0) < \infty$. Without loss of generality, we assume that $\alpha = \beta = 1$. Furthermore, by $\{\varepsilon_n\}$ we denote a strictly positive sequence for which $\varepsilon_n \to 0$ as $n \to \infty$.

In this chapter we show that under certain assumptions on the *a priori* edge information $\{v_n^0\}$ and K^0 the sequence of functionals

$$\overline{\Psi}_{v_n^0,n}(f,v) = \begin{cases} \alpha \int_\Omega v^2 |\nabla f|^2 dx + \beta \int\limits_\Omega \varepsilon_n |\nabla v|^2 + \frac{(v_n^0 - v)^2}{4\varepsilon_n} dx, & \text{for } \begin{array}{l} f, v \in W^{1,2}(\Omega), \\ 0 \leq v \leq 1 \end{array} \\ +\infty, & \text{else} \end{cases} \tag{4.2}$$

approximates

$$\overline{\Psi}_{K^0}(f,v) = \begin{cases} \alpha \int_\Omega |\nabla f|^2 dx + \beta \mathcal{H}^{N-1}(S_f \setminus K^0), & v = 1, f \in SBV(\Omega) \\ +\infty, & \text{else,} \end{cases} \tag{4.3}$$

in the sense of Γ-convergence as $n \to \infty$ on $L^2(\Omega) \times L^2(\Omega)$. The difference of (4.2) to the usual Ambrosio-Tortorelli penalty is in the last term, where instead of a constant function 1 the *a priori* edges $\{v_n^0\}$ are used as a constraint.

For the case where $\gamma \neq 0$ we propose the heuristically motivated penalty

$$H_{v_n^0,n}(f,v) = \alpha \int_\Omega v^2|\nabla f|^2 dx + \beta \int\limits_\Omega \big(\varepsilon_n|\nabla v|^2 + \frac{(1-v)^2}{4\varepsilon_n}\big)(1+\gamma(v_n^0 - v)^2)dx \tag{4.4}$$

for $f,\ v \in W^{1,2}(\Omega), 0 \leq v \leq 1$ with $0 < \beta, \gamma, \alpha$. We do not have any convergence results in this case and therefore only give a brief motivation at the end of this chapter.

To apply the theory of Γ-convergence we need to show that for suitable topologies the lim inf-inequality and the lim sup-inequality from Definition 2.29 hold. We adapt the original proof from Ambrosio and Tortorelli [1992] as it is presented in Braides [2002]. First, we consider the approximation on an interval and then lift the result to the N dimensional case by standard techniques. The main contribution and difficulty is to establish suitable conditions on the *a priori* edge set and modify the proofs accordingly.

4.1 Representing K^0 in the phase field setting

In this section we note the assumptions we take on the sequence $\{v_n^0\} \in W^{1,N+1}(\Omega)$ that represents the *a priori* edge information $K^0 \subset \Omega$ in the phase field setting. We use these assumptions to describe the asymptotic behavior of the sequence $\{v_n^0\}$ near K^0. Furthermore, we give a second more natural set of assumptions and show why they unfortunately do not work for the proof in the next sections.

We begin by listing the assumptions we will use in this chapter.

Assumption 4.1 (Assumptions on $\{v_n^0\}$). *For a compact set $K^0 \subset \Omega \subset \mathbb{R}^N$ with $\mathcal{H}^{N-1}(K^0) < +\infty$ we assume that for every sequence $\{\varepsilon_n\}$ with $\varepsilon_n \to 0$ as $n \to \infty$ there exists a sequence $\{v_n^0\} \in W^{1,N+1}(\Omega), 0 \leq v_n^0 \leq 1$ such that*

(i) *$v_n^0(x) \to 0$ uniformly in K^0 as $n \to \infty$.*

(ii) *$v_n^0(x) \to 1$ if $x \notin K^0$ as $n \to \infty$.*

(iii) *For every subset $A \subset \Omega$ with $A \cap K^0 = \emptyset$ there exists a sequence $\{\eta_{A,n}\} \in \mathbb{R}$*

so that

$$1 - \eta_{A,n} \le v_n^0(x) \le 1 \qquad \text{for } x \in A \tag{4.5}$$

and

$$\frac{\eta_{A,n}}{\varepsilon_n} \to 0 \tag{4.6}$$

as $n \to \infty$.

(iv) *It holds*

$$\int_\Omega \varepsilon_n |\nabla v_n^0|^{N+1} dx \to 0 \tag{4.7}$$

as $n \to \infty$.

Assumption (i) and (ii) ensure that, in the limit, only real edges are indicated as such, assumption (iii) ensures that away from edges the sequence converges to 1 fast and assumption (iv) prohibits the valleys around edges to be too steep. Note that in the assumptions above for $W^{1,N+1}(\Omega)$ the Sobolev number $1 - \frac{N}{N+1}$ is strictly positive and therefore for every $n \in \mathbb{N}$ it holds $v_n^0 \in C^0(\Omega)$.

Remark 4.2. *As the domain* Ω *is bounded and* $N \ge 1$, *from Assumption 4.1* (iv) *it follows that*

$$\int_\Omega \varepsilon_n |\nabla v_n^0|^2 dx \to 0 \tag{4.8}$$

as $n \to \infty$.

We first characterize the behavior of the indicator functions $\{v_n^0\}$ for $n \to \infty$ in the neighborhood of edges. Throughout this chapter we will use the following estimate on an interval $[a, b]$: for $1 < p < \infty$ and $u \in W^{1,p}[a, b]$ it holds

$$\int_a^b |\nabla u|^p dx \ge |\frac{u(b) - u(a)}{b - a}|^p (b - a). \tag{4.9}$$

The estimate comes from the minimization of the p-Dirichlet integral

$$\min_{u \in W^{1,p}(\Omega)} \int_\Omega |\nabla u|^p dx$$
$$\text{subject to} \qquad (u-h) \in W_0^{1,p}(\Omega),$$

for some function $h : \Omega \to \mathbb{R}$. If the boundary data h is regular enough then a minimizer is attained and the associated Euler-Lagrange equation is

$$\operatorname{div}(|\nabla h|^{p-2} \nabla h) = 0,$$

see Lewis [1977]. The operator on the left hand side is called the p-Laplacian. In one dimension linear functions solve the p-Laplacian and thus also minimize the p-Dirichlet integral over an interval for a given set of boundary values.

Lemma 4.3. *Let $\Omega \subset \mathbb{R}^N$, $N \geq 2$, $K^0 \subset\subset \Omega$, $\mathcal{H}^{N-1}(K^0) < \infty$ and $\{\varepsilon_n\}$ be a sequence converging to 0. Furthermore, assume $\{v_n^0\}$ fulfills Assumption 4.1 for the given sequence $\{\varepsilon_n\}$ and K^0. Then for any sequence $\{x_n\}$ with $\limsup_n \frac{dist(x_n,K^0)}{\varepsilon_n} < \infty$, it holds $v_n^0(x_n) \to 0$ as $\varepsilon_n \to 0$.*

Proof. Without loss of generality we assume that in the following we are sufficiently far away from the boundary of Ω.

Step one: Let $\{x_n^0\}$ be an arbitrary sequence in K^0 and $\{\rho_n\}$ be a sequence of positive real numbers with $\limsup_n \frac{\rho_n}{\varepsilon_n} < \infty$. We first show that for almost every $x \in B_1(0)$ it holds $v_n^0(x_n^0 + \rho_n x) \to 0$ as $n \to \infty$. Let us fix a $c \in (0,1]$ and S^{N-1} be the $N-1$ dimensional sphere. Using polar coordinates we compute

$$\begin{aligned}\int_\Omega \varepsilon_n |\nabla v_n^0|^{N+1} dx &\geq \varepsilon_n \int_{B_{c\rho_n}(x_n^0)} |\nabla v_n^0(x)|^{N+1} dx \\ &= \varepsilon_n \int_{S^{N-1}} \int_0^{c\rho_n} |\nabla v_n^0(x_n^0 + t\cdot\nu)|^{N+1} t^{N-1} dt \, d\mathcal{H}^{N-1},\end{aligned}$$

where ν denotes an element in S^{N-1}. We estimate the one dimensional integral via the reverse Hölder inequality. Let $u(t) = |\nabla v_n^0(x_n^0 + t \cdot \nu)|^{N+1}$, $h(t) = t^{N-1}$

and $s \in (1, N)$, then

$$\int_0^{c\rho_n} u\ hdt = \|\ u\ h\ \|_{L^1([0,c\rho_n])}$$
$$\geq \|u\|_{L^{1/s}([0,c\rho_n])}\ \|h\|_{L^{-1/(s-1)}([0,c\rho_n])}.$$

The second norm can be computed as

$$\|h\|_{L^{-1/(s-1)}([0,c\rho_n])} = \Big(\int_0^{c\rho_n} (t^{N-1})^{-\frac{1}{s-1}} dt\Big)^{-(s-1)} = \Big(\frac{s-1}{s-N}(c\rho_n)^{(\frac{s-N}{s-1})}\Big)^{-(s-1)}$$
$$= C(c\rho_n)^{-s+N},$$

with $C = (\frac{s-N}{s-1})^{(s-1)}$. Now we estimate the p-Dirichlet integral with a linear function taking the same boundary values as v_n^0

$$\|u\|_{L^{1/s}([0,c\rho_n])} = \Big(\int_0^{c\rho_n} (|\nabla v_n^0(x_n^0 + t\cdot\nu)|^{N+1})^{\frac{1}{s}} dt\Big)^s$$
$$\geq \Big(\int_0^{c\rho_n} \Big|\frac{v_n^0(x_n^0) - v_n^0(x_n^0 + c\rho_n\nu)}{|c\rho_n|}\Big|^{\frac{N+1}{s}} dt\Big)^s$$
$$= \frac{(c\rho_n)^s}{(c\rho_n)^{N+1}}\ \big|v_n^0(x_n^0) - v_n^0(x_n^0 + c\rho_n\nu)\big|^{N+1}$$
$$= (c\rho_n)^{s-N-1}\ \big|v_n^0(x_n^0) - v_n^0(x_n^0 + c\rho_n\nu)\big|^{N+1}.$$

We summarize with

$$\int_\Omega \varepsilon_n|\nabla v_n^0|^{N+1}dx$$
$$\geq \varepsilon_n C(c\rho_n)^{-s+N}(c\rho_n)^{s-N-1}\int_{S^{N-1}} \big|v_n^0(x_n^0) - v_n^0(x_n^0 + c\rho_n\nu)\big|^{N+1} d\mathcal{H}^{N-1}$$
$$= C\frac{\varepsilon_n}{c\rho_n}\int_{S^{N-1}} \big|v_n^0(x_n^0) - v_n^0(x_n^0 + c\rho_n\nu)\big|^{N+1} d\mathcal{H}^{N-1}.$$

The integral $\int_\Omega \varepsilon_n|\nabla v_n^0|^{N+1}dx$ goes to zero via the Assumption 4.1 (iv). As $\liminf_n \frac{\varepsilon_n}{c\rho_n} > 0$ and $v_n^0 \to 0$ uniformly in K^0, for $\mathcal{H}^{N-1}$ - almost every $\nu \in S^{N-1}$ it holds $v_n^0(x_n + c\rho_n\nu) \to 0$. As $c \in (0, 1]$ was chosen arbitrarily, we can follow for $\mathcal{L}^N$ - almost everywhere $x \in B^1(0)$ it holds $v_n^0(x_n^0 + \rho_n x) \to 0$.

Step two: Now let us assume there is a sequence $\{x_n\} \in \Omega$ with

$\limsup_n \frac{dist(x_n,K^0)}{\varepsilon_n} < \infty$ and $\limsup_n v_n^0(x_n) > \hat{C} > 0$ as $n \to \infty$. Because of Assumption 4.1 (i) we can extract a subsequence such that $\{x_n\} \in \Omega \setminus K^0$. Because of $\limsup_n \frac{dist(x_n,K^0)}{\varepsilon_n} < \infty$ we can find a sequence $\{x_n^0\}$ in K^0 and $\{\rho_n\} \in \mathbb{R}_+$ with $\limsup_n \frac{\rho_n}{\varepsilon_n} < \infty$ such that

$$x_n \in \mathring{B}_{\rho_n}(x_n^0),$$

where $\mathring{B}_{\rho_n}(x_n^0)$ denotes the interior of $B_{\rho_n}(x_n^0)$. Due to step one for $\mathcal{L}^N$ - almost every $x \in B_1(0)$ it holds $v_n^0(x_n^0 + \rho_n x) \to 0$ as $n \to \infty$. Therefore we can fine a second sequence $\{\tau_n\} \in \mathbb{R}_+$ with

$$B_{\tau_n}(x_n) \subset \mathring{B}_{\rho_n}(x_n^0)$$

and for $\mathcal{H}^{N-1}$ - almost every $\nu \in S^{N-1}$ it holds $v_n^0(x_n + \tau_n \nu) \to 0$ as $n \to \infty$.

Step three: We estimate the gradient over $B_{\tau_n}(x_n)$ to obtain a contradiction. As above we have

$$\begin{aligned}
&\int_\Omega \varepsilon_n |\nabla v_n^0|^{N+1} dx \\
&\geq \int_{B_{\tau_n}(x_n)} \varepsilon_n |\nabla v_n^0|^{N+1} dx \\
&= \varepsilon_n \int_{S^{N-1}} \int_0^{\tau_n} |\nabla v_n^0(x_n + t\nu)|^{N+1} (t)^{N-1} dt \, d\mathcal{H}^{N-1} \\
&\geq C \frac{\varepsilon_n}{\tau_n} \int_{S^{N-1}} |v_n^0(x_n) - v_n^0(x_n + \tau_n \nu)|^{N+1} d\mathcal{H}^{N-1}.
\end{aligned}$$

We have $\liminf_n \frac{\varepsilon_n}{\tau_n} > 0$. Because of step one, for $\mathcal{H}^{N-1}$ - almost every $\nu \in S^{N-1}$ it holds $v(x_n + \tau_n \nu) \to 0$. For the sequence $\{x_n\}$ with $\limsup_n \frac{dist(x_n,K^0)}{\varepsilon_n} < \infty$ and $\limsup_n v_n^0(x_n) > \hat{C} > 0$ as $n \to \infty$ the last integral therefore does not go to zero. This contradicts Assumption 4.1 (iv). □

Lemma 4.3 also holds for dimension $N = 1$, but the proof does not require the transformation to polar coordinates and the reverse Hölder inequality. Instead the integral can be estimated directly with linear functions to obtain the claim.

To have a consistent model, it would be desirable to only assume that the *a priori* phase field functions were obtained as minimizers from a sequence of

Ambrosio-Tortorelli functions. In this regard we show a negative result in one dimension.

Assumption 4.4 (Desirable assumptions on $\{v_n^0\}$). *For a given set $K^0 \subset \Omega$ with $\mathcal{H}^{N-1}(K^0) < +\infty$ we assume that for every sequence $\{\varepsilon_n\}$ with $\varepsilon_n \to 0$ as $n \to \infty$ there exists a sequence $\{v_n^0\} \in W^{1,2}(\Omega) \cap C^0(\Omega), 0 \leq v_n^0 \leq 1$ such that*

(i) $v_n^0(x) \to 0$ *if* $x \in K^0$ *as* $n \to \infty$.

(ii) $v_n^0(x) \to 1$ *if* $x \notin K^0$ *as* $n \to \infty$.

(iii) *For every subset $A \subset \Omega$ it holds*

$$\int_A \varepsilon_n |\nabla v_n^0|^2 + \frac{(1-v_n^0)^2}{4\varepsilon_n} dx \to \mathcal{H}^{N-1}(A \cap K^0) \tag{4.10}$$

as $n \to \infty$.

We now introduce a one dimensional example to see that these assumptions are not sufficient. We divide the example into two parts.

Example 4.5 (No recovery sequence). *Let $\Omega = (-1,1)$, $K^0 = \{0\}$ and $\{\varepsilon_n\}$ be a sequence converging to 0. Furthermore, assume $\{v_n^0\}$ fulfills Assumption 4.4 for the given sequence $\{\varepsilon_n\}$. Then for every convergent sequence $\{x_n\}$ with limit x:*

(i) *If* $\lim_n \frac{x_n}{\varepsilon_n} \to 0$ *then* $v_n^0(x_n) \to 0$.

(ii) *If* $\lim_n \frac{\varepsilon_n}{x_n} \to 0$, *then the sequence $\{v_n^0\}$ converges to 1 on $[0, x_n]$ in the sense that*

$$\int_0^1 (1 - v_n^0(x_n \cdot \tau))^2 d\tau \to 0.$$

Proof. We first show (i). The case $x_n = 0$ is treated in the Assumptions 4.4 (i). Let $\{x_n\}$ be a positive sequence converging to 0 with $\lim_n \frac{x_n}{\varepsilon_n} \to 0$. Using Assumption 4.4 (iii), we follow that the sequence $\{\varepsilon_n \int_\Omega |(v_n^0)'|^2 dx\}$ is uniformly bound by some constant $C > 0$ and we can estimate for all $n \in \mathbb{N}$

$$C > \varepsilon_n \int_\Omega |(v_n^0)'|^2 dx \geq \varepsilon_n \int_0^{x_n} |(v_n^0)'|^2 dx$$
$$\geq \varepsilon_n \int_0^{x_n} |\frac{(v_n^0(x_n) - v_n^0(0))}{x_n}|^2 dx = \frac{\varepsilon_n}{x_n}(v_n^0(x_n) - v_n^0(0))^2.$$

As $\frac{\varepsilon_n}{x_n} \to \infty$ for $n \to \infty$, it is necessary that $(v_n^0(x_n) - v_n^0(0))^2 \to 0$. With $\lim_n v_n^0(0) = 0$, we can follow $\lim_n v_n^0(x_n) = 0$.

Now we turn to (ii). If $\{x_n\}$ does not converge to $K^0 = 0$, then due to Assumption 4.4 (ii) the claim holds. Let $\{x_n\}$ be a sequence converging to 0 with $\lim_n \frac{\varepsilon_n}{x_n} \to 0$. Using Assumption 4.4 (iii), we can compute for some constant $C > 0$ and every $n \in \mathbb{N}$

$$C > \int_\Omega \frac{(1 - v_n^0(x))^2}{4\varepsilon_n} dx \geq \int_0^{x_n} \frac{(1 - v_n^0(x))^2}{4\varepsilon_n} dx$$
$$= \frac{1}{4\varepsilon_n} \int_0^1 (1 - v_n^0(x_n \cdot \tau)))^2 x_n d\tau = \frac{x_n}{4\varepsilon_n} \int_0^1 (1 - v_n^0(x_n \cdot \tau)))^2 d\tau,$$

where we used a change of variable $x = \tau x_n$ and $dx = x_n d\tau$.

As $\lim_n \frac{\varepsilon_n}{x_n} \to 0$, for the last term to be uniformly bounded it is necessary that

$$\int_0^1 (1 - v_n^0(x_n \cdot \tau))^2 d\tau \to 0.$$

□

Now we will show that for this example, under the above Assumptions 4.4, no sequence $\{v_n\} \in W^{1,2}(-1,1)$ *can be found such that*

$$\lim_n \int_{-1}^1 \varepsilon_n |v_n'|^2 + \frac{(v_n - v_n^0)^2}{4\varepsilon_n} dx = 0.$$

To be able to find such a (recovery) sequence is necessary for the variational approximation we are pursuing.

Proof. Assume there exists a sequence $\{v_n\} \in W^{1,2}(-1,1)$ such that $v_n(0) \to c$ for some $c \geq 0$ and

$$\lim \int_{-1}^{1} \varepsilon_n |v_n'|^2 + \frac{(v_n - v_n^0)^2}{4\varepsilon_n} dx = 0. \tag{4.11}$$

For any sequence $\{x_n\}$ with $\lim \frac{x_n}{\varepsilon_n} < \infty$, as above we have

$$\varepsilon_n \int_0^{x_n} |v_n'|^2 dx \geq \varepsilon_n \int_0^{x_n} |\frac{(v_n(x_n) - v_n(0))}{x_n}|^2 dx = \frac{\varepsilon_n}{x_n}(v_n(x_n) - v_n(0))^2.$$

As $\int_\Omega \varepsilon_n |v_n'|^2 dx \to 0$ and $v_n(0) \to c$ we can follow that $v_n(x_n) \to c$.

Therefore, for any $C > 0$ it holds $v_n(t) \to c$ for all $t \in [0, C\varepsilon_n]$ and we can compute

$$\lim_n \frac{1}{4\varepsilon_n} \int_0^{C\varepsilon_n} (v_n(x) - 1)^2 dx = \lim_n \frac{C\varepsilon_n}{4\varepsilon_n} \int_0^1 (v_n(C\varepsilon_n t) - 1)^2 dt = \frac{C}{4}(1-c)^2.$$

Assume that $c \neq 1$, then as $C > 0$ can be chosen arbitrarily large we have

$$\frac{1}{4\varepsilon} \int_0^1 (v_n - 1)^2 dx \to \infty.$$

This together with $\limsup_n \frac{1}{4\varepsilon_n} \int_0^1 (v_n^0 - 1)^2 dx \leq \mathcal{H}^0(K^0)$ contradicts the assumption

$$\lim_n \int_{-1}^{1} \frac{(v_n - v_n^0)^2}{4\varepsilon_n} dx = 0.$$

Assume that $c = 1$, we would then have

$$0 = \lim_n \int_0^{C\varepsilon_n} \frac{(v - v_n^0)^2}{4\varepsilon_n} dx = \lim_n \int_0^{C\varepsilon_n} \frac{(1 - v_n^0)^2}{4\varepsilon_n} dx,$$

for any $C > 0$. This yields a contradiction to $\limsup_n \int_\Omega \varepsilon_n |(v_n^0)'|^2 dx \leq \mathcal{H}^{N-1}(K^0)$ together with $v_n^0(0) \to 0$. This can be computed by choosing $C > \frac{1}{\mathcal{H}^{N-1}(K^0)}$ and a sequences $\{x_n\}$ in $[0, C\varepsilon_n]$ for which $v_n^0(x_n) \to 1$ and then estimating the gradient

$$\begin{aligned}\int_0^{x_n} \varepsilon_n |(v_n^0)'|^2 dx &\geq \int_0^{x_n} \varepsilon_n \left|\frac{v_n^0(0) - v_n^0(x_n)}{x_n}\right|^2 dx \\ &= \frac{x_n \varepsilon_n}{(x_n)^2}(v_n^0(0) - v_n^0(x_n))^2 \\ &\geq \frac{1}{C}(v_n^0(0) - v_n^0(x_n))^2 \to \frac{1}{C}.\end{aligned}$$

□

This concludes our one dimensional example, showing that Assumption 4.4 are not sufficient.

4.2 The $\liminf$ inequality in one dimension

In this section we prove the $\liminf$ inequality

$$\overline{\Psi}_{K^0}(f, v) \leq \liminf_n \overline{\Psi}_{v_n^0, n}(f_n, v_n) \tag{4.12}$$

for every sequence $\{f_n, v_n\} \in W^{1,2}(\Omega) \times W^{1,2}(\Omega)$ converging to $(f, v) \in SBV(\Omega) \times L^2(\Omega)$ in $L^2(\Omega) \times L^2(\Omega)$ for $\varepsilon_n \to 0$ as $n \to \infty$.

Let us fix $\Omega = (-1, 1)$ and recall the definitions for the one dimensional case

$$\overline{\Psi}_{v_n^0, n}(f, v) = \alpha \int_\Omega v^2 |f'|^2 dx + \beta \int_\Omega \varepsilon_n |v'|^2 + \frac{(v_n^0 - v)^2}{4\varepsilon_n} dx. \tag{4.13}$$

for $f,\, v \in W^{1,2}(\Omega), 0 \leq v \leq 1$, and

$$\overline{\Psi}_{K^0}(f, v) = \alpha \int\limits_{\Omega} |f'|^2 dx + \beta \mathcal{H}^0(S_f \setminus K^0) \tag{4.14}$$

for $f \in SBV(\Omega)$, $v = 1$ *a.e.*, and else $\overline{\Psi}_{K^0}(f, v) = +\infty$. For restrictions to subsets $A \subset \Omega$, we use the notation

$$\overline{\Psi}_{v_n^0,n}(f, v, A) = \alpha \int\limits_{A} v^2 |f'|^2 dx + \beta \int\limits_{A} \varepsilon_n |v'|^2 + \frac{(v_n^0 - v)^2}{4\varepsilon_n} dx$$

and

$$\overline{\Psi}_{K^0}(f, v, A) = \alpha \int\limits_{A} |f'|^2 dx + \beta \mathcal{H}^0(S_f \setminus K^0 \cap A).$$

We will first prove the following lemma.

Lemma 4.6 (Oscillation Lemma). *Let $\varepsilon_n \to 0$ as $n \to \infty$ and $\{v_n^0\}$, K^0 fulfill Assumptions 4.1. Then for every sequence $\{v_n\} \in W^{1,2}(\Omega)$ with*

$$\sup_n \int\limits_{-1}^{1} \varepsilon_n |v_n'|^2 + \frac{(v_n^0 - v_n)^2}{4\varepsilon_n} dx < +\infty \tag{4.15}$$

there exists a convergent subsequence of $\{v_n\}$ with limit v that satisfies

$$v = 1 \quad a.e.$$

Also, for every $\eta > 0$ there exists a finite set $S^\eta \subset \Omega$ and a subsequence of $\{v_n\}$, still denoted as $\{v_n\}$, so that for each compact subset I of $\Omega \setminus (S^\eta \cup K^0)$ there exists an index $n_{\eta,I} \in \mathbb{N}$ such that for every $n \geq n_{\eta,I}$ it holds $1 - \eta \leq v_n \leq 1 + \eta$ on I.

In the proof we use the notation

$$G_n(v_n) := \int_{-1}^{1} \varepsilon_n |v_n'|^2 + \frac{(v_n^0 - v_n)^2}{4\varepsilon_n} dx. \tag{4.16}$$

Proof. Let $\{v_n\} \in W^{1,2}(\Omega)$ be a sequence with (4.15). Let us define $I_n^\eta := \{x \in (-1,1) : |v_n(x) - v_n^0(x)| > \eta\}$ and denote its measure as $|I_n^\eta|$. Then for any $\eta > 0$ it is

$$|I_n^\eta|(\frac{\eta^2}{4}) \leq \int_{I_n^\eta} \frac{\varepsilon_n(v_n(x) - v_n^0(x))^2}{\varepsilon_n 4} dx \leq \varepsilon_n G_n(v_n). \tag{4.17}$$

As $\sup_n G_n(v_n) < \infty$, for any $\eta > 0$ the measure $|I_n^\eta|$ tends to zero as $n \to \infty$. Also, by the triangle inequality we have

$$\begin{aligned} |\{x \in (-1,1) : |v_n(x) - 1| > \eta\}| &\leq |\{x \in (-1,1) : |v_n(x) - v_n^0(x)| + |v_n^0(x) - 1| > \eta\}| \\ &\leq |\{x \in (-1,1) : |v_n(x) - v_n^0(x)| > \frac{\eta}{2}\}| \\ &+ |\{x \in (-1,1) : |v_n^0(x) - 1| > \frac{\eta}{2}\}|. \end{aligned}$$

As $\{v_n^0\}$ converges to 1 a.e., the right hand side converges to zero as $n \to \infty$. Therefore, also $\{v_n\}$ converges to 1 in measure and therefore also $v_n \to 1$ a.e.

It remains to show that $\{v_n\}$ can be bound as described above. For every $N \in \mathbb{N}$ let us define N equidistant distributed points $x_N^i := -1 + i\frac{2}{N}$, $i = 0,..,N$. We show that the number of intervals $[x_N^i, x_N^{i+1}]$ on which the functions $\{v_n\}$ are not bound is independent of N, for large enough n and N.

We first fix N and i so that $[x_N^i, x_N^{i+1}] \cap K^0 = \emptyset$. Let s, t be such that $v_n(s) = min_{[x_N^i, x_N^{i+1}]} v_n(x)$ and $v_n(t) = max_{[x_N^i, x_N^{i+1}]} v_n(x)$. Furthermore, by Assumption 4.1 *(iii)* we have a sequence $\eta_n^0 \to 0$ that bounds the oscillations of v_n^0 through $|1 - v_n^0| \leq \eta_n^0$ on $[x_N^i, x_N^{i+1}]$. Note that the sequence $\{\eta_n^0\}$ is independent of i.

For simplicity assume without loss of generality $s < t$. With the inequality

$a^2 + b^2 \geq 2ab$ we have

$$\begin{aligned}\int_s^t \varepsilon_n |v_n'|^2 + \frac{(v_n^0 - v_n)^2}{4\varepsilon_n} dx &\geq \int_s^t |v_n'||v_n^0 - v_n| dx \geq \int_s^t |v_n'| \, \big||1 - v_n| - |1 - v_n^0|\big| dx \\ &\geq \left| \int_s^t |v_n'| \left(|1 - v_n| - |1 - v_n^0|\right) dx \right| \\ &\geq \int_s^t |v_n'| \left(|1 - v_n| - \eta_n^0\right) dx \\ &\geq \cdots\end{aligned}$$

Using the substitution $\tau = v(x)$ we then have

$$\cdots \geq \int_{v_n(s)}^{v_n(t)} \left(|1 - \tau| - \eta_n^0\right) d\tau.$$

Now we fix $0 < \eta$ and define

$$\begin{aligned}J_{n,N}^\eta :=& \{i \in \{0, ..., N\} : max_{[x_i, x_{i+1}]} v_n(x) - min_{[x_i, x_{i+1}]} v_n(x) \geq \eta\} \\ &\cap \{i \in \{0, ..., N\} : [x_i, x_{i+1}] \cap K^0 = \emptyset\}.\end{aligned}$$

For $i \in J_{n,N}^\eta$ again let s, t be such that $v_n(s) = min_{[x_N^i, x_N^{i+1}]} v_n(x)$ and $v_n(t) = max_{[x_N^i, x_N^{i+1}]} v_n(x)$. We again assume $s < t$. For sufficiently small η_n^0, that is large n, Lemma A 1 gives us a constant $C > 0$ such that

$$\begin{aligned}\int_{x_i}^{x_{i+1}} \varepsilon_n |v_n'|^2 + \frac{(v_n^0 - v_n)^2}{4\varepsilon_n} dx &\geq \int_{v_n(s)}^{v_n(t)} |1 - \tau| - \eta_n^0 d\tau \\ &\geq C.\end{aligned}$$

We can therefore follow that for every $0 < \eta$ for large n it holds

$$\# J_{n,N}^\eta \leq \frac{1}{C} G_n(v_n). \tag{4.18}$$

As $sup_n G_n < \infty$, the number of intervals on which v_n oscillates more than η is bound and the bound is independent of N.

Let us then define the set $S_{n,N}^\eta := \{x_i \in [-1, 1] : \; i \in J_{n,N}^\eta\}$. For any $0 < \eta \leq 1$ as $\# J_{n,N}^\eta$ is bounded so is $\# S_{n,N}^\eta$ with respect to both N and n. Therefore, letting

$N \to \infty$ and $n \to \infty$ we can find a convergent subsequence of $S^{\eta}_{n,N}$ with limit $S^{\eta} \subset \Omega$. We denote the according subsequence also as $\{v_n\}$.

For any open set $I \subset \Omega \setminus (S^{\eta} \cup K^0)$, for N large enough, we can find a cover $\cup_i [x_i, x_{i+1}]$ and an index n_0 such that for all $n \geq n_0$ it holds $I \subset \cup_i [x_i, x_{i+1}] \subset \Omega \setminus (S^{\eta}_{n,N} \cup K^0)$.

As each $[x_i, x_{i+1}] \cap S^{\eta}_{n,N} = \emptyset$ we have that

$$max_{[x_i,x_{i+1}]} v_n(x) - min_{[x_i,x_{i+1}]} v_n(x) \leq \eta.$$

Because $v_n \to 1$ in measure, for large enough n it therefore holds

$$1 - \eta \leq v_n(x) \leq 1 + \eta. \tag{4.19}$$

This is valid for the entire cover $\cup_i [x_i, x_{i+1}]$ and therefore also for $I \subset \Omega \setminus (S^{\eta} \cup K^0)$. □

Now we turn to the proof of the lim inf inequality in one dimension.

Lemma 4.7 (The lim inf inequality)**.** *Let $\varepsilon_n \to 0$ as $n \to \infty$ and $\{v_n^0\}$, K^0 fulfill Assumptions 4.1. For every sequence $\{(f_n, v_n)\}$ converging to (f, v) in $L^2(\Omega) \times L^2(\Omega)$ it then holds*

$$\overline{\Psi}_{K^0}(f, v) \leq \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n). \tag{4.20}$$

Proof. Let $\varepsilon_n \to 0$, $f_n \to f$ and $v_n \to v$ in $L^2(\Omega)$ as $n \to \infty$. Up to a subsequence, we can suppose that

$$\liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n) = \lim_{n\to\infty} \overline{\Psi}_{v_n^0,n}(f_n, v_n) = C < +\infty$$

and $v_n \to 1$ a.e. (else there is nothing to show).

We first show

$$\#(S_f \setminus K^0 \cap I) \leq \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, I) \tag{4.21}$$

for any open subset I of Ω. If $S_f \setminus K^0 = \emptyset$ there is nothing to show. Otherwise choose $\{t_1, .., t_N\} \subset S_f \setminus K^0$ and disjoint intervals $I_i = (a_i, b_i) \subset \Omega$ with $t_i \in I_i$

and $(a_i, b_i) \cap K^0 = \emptyset$. Let $t_i \in I_i' \subset\subset I_i$, and let $m_i = \liminf_n \left(\inf_{t \in I_i'} \; v_n(t)^2\right)$. If $m_i > 0$, then

$$\int_{I_i'} |f_n'|^2 \leq \frac{1}{m_i} \int_{I_i'} v_n(t)^2 |f_n'|^2 \leq \hat{C}$$

and because of the $L^2(\Omega)$ convergence of f_n also

$$\int_{I_i'} |f_n'|^2 + \int_{I_i'} |f_n|^2 \leq \hat{\hat{C}},$$

for some constants $\hat{C}, \hat{\hat{C}} > 0$. In this case f_n converges weakly to f in $W^{1,2}(I_i') \subset C^0(I_i')$, which would imply $(S_f \cap I_i') = \emptyset$. Therefore, it has to be $m_i = 0$ and there exists a sequence $\{s_n^i\} \in I_i'$ such that $v_n(s_n^i) \to 0$. Moreover, as v_n converges to 1 a.e., we can find $r_i, r_i' \in I_i$ so that $r_i < s_n^i < r_i'$ and $v_n(r_i) \to 1$, $v_n(r_i') \to 1$. Then we can estimate

$$\begin{aligned}
\liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, I_i) &\geq \liminf_n \int_{r_i}^{r_i'} \varepsilon_n |v_n'|^2 + \frac{(v_n^0 - v_n)^2}{4\varepsilon_n} dx \\
&\geq \liminf_n \int_{r_i}^{s_n^i} \varepsilon_n |v_n'|^2 + \frac{(v_n^0 - v_n)^2}{4\varepsilon_n} dx \\
&+ \liminf_n \int_{s_n^i}^{r_i'} \varepsilon_n |v_n'|^2 + \frac{(v_n^0 - v_n)^2}{4\varepsilon_n} dx.
\end{aligned}$$

Using the inequality $a^2 + b^2 \geq 2ab$ we get

$$\begin{aligned}
\liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, I_i) \geq \liminf_n &\int_{r_i}^{s_n^i} |v_n'||v_n^0 - v_n| dx \\
+ \liminf_n &\int_{s_n^i}^{r_i'} |v_n'||v_n^0 - v_n| dx.
\end{aligned} \tag{4.22}$$

The two terms on the right hand side of (4.22) are of the same kind and we only look at the second one

$$\liminf_n \int_{s_n^i}^{r_i'} |v_n'||v_n^0 - v_n|dx \geq \liminf_n \int_{s_n^i}^{r_i'} |v_n'| \left(|v_n - 1| - |v_n^0 - 1|\right) dx$$
$$\geq \cdots$$

As $(r_i, r_i') \cap K^0 = \emptyset$ we can use Assumption 4.1 (*iii*) of the *a priori* edge and obtain

$$\cdots \geq \liminf_n \int_{s_n^i}^{r_i'} |v_n'| \left(|v_n - 1| - \eta_n^0\right) dx.$$

Using the substitution $t = v_n(x)$ and $dt = v_n'(x)dx$ we get

$$\liminf_n \int_{s_n^i}^{r_i'} |v_n'| \left(|v_n - 1| - \eta_n^0\right) dx \geq \liminf_n \int_{v_n(s_n^i)}^{v_n(r_i')} |t-1|dt - \eta_n^0 \limsup_n \int_{v_n(s_n^i)}^{v_n(r_i')} dt$$
$$= \liminf_n \int_{v_n(s_n^i)}^{v_n(r_i')} |t-1|dt - \eta_n^0.$$

Letting the limits $n \to \infty$, $\eta_n^0 \to 0$, $v_n(r_i') \to 1$ and $v_n(s_n^i) \to 0$ pass, we arrive at

$$\liminf_n \int_{s_n^i}^{r_i'} |v_n'||v_n^0 - v_n|dx \geq \int_0^1 |t-1|dt = \frac{1}{2}.$$

This yields

$$\liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, I_i) \geq \frac{1}{2} + \frac{1}{2} = 1. \tag{4.23}$$

Together with the arbitrary choice of points $\{t_1, .., t_N\} \subset S_f \setminus K^0$ we arrive at

$$\#(S_f \setminus K^0) \leq \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, I) \tag{4.24}$$

if I is any open set with $(S_f \setminus K^0) \subset I$.

Now we show that for every open subset I with $I \cap (S_f \cup K^0) = \emptyset$ the limit

f is an element of $W^{1,2}(I)$ and

$$\int_I |f'|^2 dx \leq \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, I). \tag{4.25}$$

From Lemma 4.6 we know that there exists a set S (not the jump set S_f) so that for every open set $I \subset (-1,1) \setminus (S \cup K^0)$ and n large enough, up to a subsequence, it holds $1/2 \leq v_n \leq 3/2$ on I and thus

$$\frac{1}{2} \liminf_n \int_I |f_n'|^2 dx \leq \liminf_n \int_a^b v_n^2 |f_n'|^2 dx \leq C.$$

Therefore, f_n also converges weakly to f in $W^{1,2}(I)$. We estimate

$$\int_I |f'|^2 dx \leq \lim_n \int_I v_n^2 |f'|^2 dx \leq \liminf_n \int_I v_n^2 |f_n'|^2 dx \leq \liminf_n \int_a^b v_n^2 |f_n'|^2 dx. \tag{4.26}$$

In the first inequality we used Fatou's lemma for $v_n \to 1$. In the second, we used that the map $f \mapsto \int_I v_n^2 |f'|^2 dx$ is convex and lower semicontinuous with regards to the weak convergence $f_n \rightharpoonup f$.

Since f is in $W^{1,2}(I)$, the jump set S_f is a subset of $(S \cup K^0)$. As the points $(S \cup K^0) \setminus S_f$ are only finite, the inequality (4.26) can be extended to $f \in W^{1,2}(\Omega \setminus S_f)$. We therefore have for every I with $I \cap S_f = \emptyset$

$$\int_I |f'|^2 dx \leq \liminf_n \int_a^b v_n^2 |f_n'|^2 dx. \leq \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, I). \tag{4.27}$$

It remains to combine the two inequalities (4.24) and (4.27). For every $\tau > 0$ we define the sets $I_\tau^0 := \Omega \setminus (S_f + [-\tau, \tau])$ and $I_1 := (S_f \setminus K^0 + [-\tau, \tau]) \cap \Omega$. The two sets are disjoint but don't make up Ω entirely. Then using the two inequalities (4.24) and (4.27), for every $(f_n, v_n) \to (f, v)$ in $L^2(\Omega) \times L^2(\Omega)$ and $\tau > 0$ we have

$$\begin{aligned}\int_{I_\tau^0} |f'|^2 dx + \#(S_f \setminus K^0 \cap I_\tau^1) &\leq \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, I_\tau^0) + \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, I_\tau^1) \\ &\leq \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n).\end{aligned}$$

Letting $\tau \to 0$, we arrive at

$$\overline{\Psi}_{K^0}(f, v) \le \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n), \tag{4.28}$$

the desired inequality. □

4.3 The $\limsup$ inequality in one dimension

In this section we prove the existence of a recovery sequence $\{f_n, v_n\} \in W^{1,2}(\Omega) \times W^{1,2}(\Omega)$ such that

$$\limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n) \le \overline{\Psi}_{K^0}(f, v). \tag{4.29}$$

As usual, $\varepsilon_n \to 0$ for $n \to \infty$.

Lemma 4.8 (The $\limsup$ inequality)**.** *Let $\varepsilon_n \to 0$ as $n \to \infty$ and $\{v_n^0\}$, K^0 fulfill Assumptions 4.1. For every $(f, v) \in L^2(\Omega) \times L^2(\Omega)$ there exists a sequence $\{f_n, v_n\}$ converging to (f, v) in $L^2(\Omega) \times L^2(\Omega)$ with*

$$\limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n) \le \overline{\Psi}_{K^0}(f, v). \tag{4.30}$$

Proof. We can assume $v = 1$ a.e. or else there is nothing to show. We now construct such a recovery sequence. It suffices to consider the cases where $(a, b) = (-1, 1)$ and

(i) $f \in W^{1,2}(-1, 1)$.

(ii) $f \in W^{1,2}(-1, 1) \setminus \{0\}$, $S_f = \{0\}$ and $(a, b) \cap K^0 = \emptyset$.

(iii) $f \in W^{1,2}(-1, 1) \setminus \{0\}$, $S_f = K^0 = \{0\}$.

Other situations can be reduced to the three cases above by separation into small intervals and shifts. We will define the sequence $\{v_n\}$ such that on the boundary of each interval it is $v_n = v_n^0$, at least in the limit $n \to \infty$. Thus a continuous patching can be done without difficulty.

For (i) we can simply choose $f_n(x) = f(x)$ and $v_n = (1 - C_n)v_n^0$ with $C_n = \exp(\frac{-1}{2\sqrt{\varepsilon_n}})$. The factor $(1-C_n)$ is only chosen to make the patching more straightforward. We can then compute

$$\begin{aligned}\overline{\Psi}_{v_n^0,n}(f_n, v_n) &= \alpha \int_{-1}^{1} (v_n^0)^2|f'|^2 + \frac{\beta}{4\varepsilon_n}(C_n v_n^0)^2 + \beta\varepsilon_n|(1 - C_n)(v_n^0)'|^2 dx \\ &\leq \alpha \int_{-1}^{1} |f'|^2 dx + \frac{\beta C_n^2}{2\varepsilon_n} + \beta \int_{-1}^{1} \varepsilon_n|(v_n^0)'|^2 dx.\end{aligned}$$

The term $\frac{C_n^2}{2\varepsilon_n} = \frac{\exp(\frac{-1}{2\sqrt{\varepsilon_n}})^2}{2\varepsilon_n} = \frac{\exp(\frac{-1}{\sqrt{\varepsilon_n}})}{2\varepsilon_n}$ goes to zero as $\varepsilon_n \to 0$ because of the exponential decay. By Assumption 4.1 (iv) we also have $\int_{-1}^{1} \varepsilon_n|(v_n^0)'|^2 dx \to 0$, which yields

$$\limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n) \leq \overline{\Psi}_{K^0}(f, v). \tag{4.31}$$

We now construct a sequence for (ii)**.** Set $f_n \in W^{1,2}(-1, 1)$ with $f_n(x) = f(x)$ if $|x| \geq \varepsilon_n^2$ and $v_n = v_n^0 \cdot \phi_n$ where

$$\phi_n(x) = \begin{cases} 0, & |x| \leq \varepsilon_n^2 \\ 1 - \exp(\frac{\varepsilon_n^2 - |x|}{2\varepsilon_n}), & \varepsilon_n^2 < |x| < \varepsilon_n^2 + \sqrt{\varepsilon_n} \\ \\ 1 - \exp(\frac{-1}{2\sqrt{\varepsilon_n}}), & |x| \geq \varepsilon_n^2 + \sqrt{\varepsilon_n}. \end{cases} \tag{4.32}$$

We can then compute

$$\begin{aligned}\overline{\Psi}_{v_n^0,n}(f_n, v_n) &= \int_{-1}^{1} \left(\alpha v_n^2|f'|^2 + \frac{\beta}{4\varepsilon_n}(v_n^0 - v_n)^2 + \beta\varepsilon_n|v_n'|^2 \right) \\ &\leq \alpha \int_{-1}^{1} |f'|^2 dx + \beta \int_{-1}^{1} \frac{1}{4\varepsilon_n}(v_n^0 - v_n)^2 + \varepsilon_n|v_n'|^2 dx.\end{aligned}$$

We rewrite the last term to

$$
\begin{aligned}
&\int_{-1}^{1} \frac{1}{4\varepsilon_n}(v_n^0)^2(1-\phi_n)^2 + \varepsilon_n|(v_n^0\cdot\phi_n)'|^2 dx \\
&\le \int_{-1}^{1} \frac{1}{4\varepsilon_n}(1-\phi_n)^2 + \varepsilon_n|v_n^0\cdot\phi_n^{'}|^2 + \varepsilon_n|(v_n^0)^{'}\cdot\phi_n|^2 dx \\
&\le \int_{-1}^{1} \frac{1}{4\varepsilon_n}(1-\phi_n)^2 + \varepsilon_n|\phi_n^{'}|^2 + \varepsilon_n|(v_n^0)^{'}|^2 dx.
\end{aligned}
$$

By Assumption 4.1 (*iii*) on the *a priori* edge $\{v_n^0\}$ it is

$$
\int_{-1}^{1} \varepsilon_n|(v_n^0)^{'}|^2 dx \to 0. \tag{4.33}
$$

The remaining term is the original Ambrosio-Tortorelli approximation of the edge penalty. Using the definition of ϕ_n, we can compute,

$$
\begin{aligned}
&\int_{-1}^{1} \frac{1}{4\varepsilon_n}(1-\phi_n)^2 + \varepsilon_n|\phi_n^{'}|^2 dx \\
&= \frac{1}{2\varepsilon_n}\int_0^{\varepsilon_n^2} dx + 2\int_{\varepsilon_n^2}^{\varepsilon_n^2+\sqrt{\varepsilon_n}} \frac{1}{4\varepsilon_n}(1-\phi_n)^2 + \varepsilon_n|\phi_n^{'}|^2 dx + \frac{1}{2\varepsilon_n}(\exp(\frac{-1}{2\sqrt{\varepsilon_n}}))^2 \int_{\varepsilon_n^2+\sqrt{\varepsilon_n}}^{1} dx.
\end{aligned}
$$

The first and last integral converge to 0 as $\varepsilon_n \to 0$. Finally, the second term is

$$
\begin{aligned}
&2\int_{\varepsilon_n^2}^{\varepsilon_n^2+\sqrt{\varepsilon_n}} \frac{1}{4\varepsilon_n}(1-\phi_n)^2 + \varepsilon_n|\phi_n^{'}|^2 dx \\
&= 2\int_{\varepsilon_n^2}^{\varepsilon_n^2+\sqrt{\varepsilon_n}} \frac{1}{4\varepsilon_n}(\exp(\frac{\varepsilon_n^2-x}{2\varepsilon_n}))^2 + \varepsilon_n|\frac{1}{2\varepsilon_n}\exp(\frac{\varepsilon_n^2-x}{2\varepsilon_n})|^2 dx \\
&= \frac{1}{\varepsilon_n}\int_{\varepsilon_n^2}^{\varepsilon_n^2+\sqrt{\varepsilon_n}} \exp(\frac{\varepsilon_n^2-x}{\varepsilon_n})dx \\
&= \frac{1}{\varepsilon_n}\big[-\varepsilon_n\exp(\frac{\varepsilon_n^2-x}{\varepsilon_n})\big]_{\varepsilon_n^2}^{\varepsilon_n^2+\sqrt{\varepsilon_n}} = -\exp(\frac{-1}{\sqrt{\varepsilon_n}}) + 1.
\end{aligned}
\tag{4.34}
$$

We define

$$
O_n := \int_{-1}^{1} \varepsilon_n|(v_n^0)^{'}|^2 dx + \frac{1}{2\varepsilon_n}\int_0^{\varepsilon_n^2} dx + \frac{1}{2\varepsilon_n}\exp(\frac{-1}{\sqrt{\varepsilon_n}}) - \exp(\frac{-1}{\sqrt{\varepsilon_n}})
$$

for which it is $O_n \to 0$ as $n \to \infty$. Together we have

$$\overline{\Psi}_{v_n^0,n}(f_n, v_n) \leq \alpha \int_{-1}^{1} |f'|^2 dx + \beta\big(1 + O_n\big).$$

In the limit this is

$$\limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n) \leq \alpha \int_{-1}^{1} |f'|^2 dx + \beta \tag{4.35}$$

and thus the claim for (ii) is shown.

We now construct a sequence for (iii)**.** As in the above case, set $f_n \in W^{1,2}(-1,1)$ with $f_n(x) = f(x)$ if $|x| \geq \varepsilon_n^2$. We define the edge indicator as $v_n := v_n^0 \cdot \phi_n$, with $C_n := \exp(\frac{-1}{2\sqrt{\varepsilon_n}})$ and

$$\phi_n(x) = \begin{cases} 0, & |x| \leq \varepsilon_n^2 \\ (1 - C_n) \cdot \frac{|x| - \varepsilon_n^2}{\varepsilon_n}, & \varepsilon_n^2 < |x| < \varepsilon_n^2 + \varepsilon_n \\ \\ 1 - C_n, & |x| \geq \varepsilon_n^2 + \varepsilon_n. \end{cases} \tag{4.36}$$

In the same way as above, we arrive at

$$\begin{aligned} &\overline{\Psi}_{v_n^0,n}(f_n, v_n) = \int_{-1}^{1} \left(\alpha v_n^2 |f'|^2 + \frac{\beta}{4\varepsilon_n}(v_n^0 - v_n)^2 + \beta\varepsilon_n |v_n'|^2 \right) \\ &\leq \alpha \int_{-1}^{1} |f'|^2 dx + \beta \int_{-1}^{1} \frac{1}{4\varepsilon_n}(v_n^0)^2(1 - \phi_n)^2 + \varepsilon_n |v_n^0 \cdot \phi_n^{'}|^2 + \varepsilon_n |(v_n^0)^{'} \cdot \phi_n|^2 dx \\ &\leq \alpha \int_{-1}^{1} |f'|^2 dx + \beta \int_{-1}^{1} (v_n^0)^2 \big(\frac{1}{4\varepsilon_n}(1 - \phi_n)^2 + \varepsilon_n |\phi_n^{'}|^2 \big) dx + \beta \int_{-1}^{1} \varepsilon_n |(v_n^0)^{'}|^2 dx. \end{aligned}$$

We look at the following part of the second integral

$$\int_{\varepsilon_n^2}^{\varepsilon_n^2 + \varepsilon_n} (v_n^0)^2 \big(\frac{1}{4\varepsilon_n}(1 - \phi_n)^2 + \varepsilon_n |\phi_n'|^2 \big) dx, \tag{4.37}$$

the other parts of the edge integral can be estimated by

$$\hat{O}_n := \int_{-1}^{1} \varepsilon_n |(v_n^0)'|^2 dx + \frac{1}{2\varepsilon_n} \int_0^{\varepsilon_n^2} dx + \frac{1}{2\varepsilon_n} C_n^2.$$

Now we use that $v_n^0 \geq 0$ and v_n^0 is continuous. Then, by the mean value theorem for integration, there exists a $\xi_n \in [\varepsilon_n^2, \varepsilon_n^2 + \varepsilon_n]$ such that

$$\int_{\varepsilon_n^2}^{\varepsilon_n^2+\varepsilon_n} (v_n^0)^2 \big(\frac{1}{4\varepsilon_n}(1-\phi_n)^2 + \varepsilon_n |\phi_n'|^2\big) dx$$
$$= (v_n^0(\xi_n))^2 \int_{\varepsilon_n^2}^{\varepsilon_n^2+\varepsilon_n} \big(\frac{1}{4\varepsilon_n}(1-\phi_n)^2 + \varepsilon_n |\phi_n'|^2\big) dx.$$

Using the definition of $\phi_n(x) = (1 - C_n) \cdot \frac{x-\varepsilon_n^2}{\varepsilon_n}$ and straightforward, but lengthy, calculation we have

$$\int_{\varepsilon_n^2}^{\varepsilon_n^2+\varepsilon_n} \big(\frac{1}{4\varepsilon_n}(1-\phi_n)^2 + \varepsilon_n |\phi_n'|^2\big) dx = \frac{C_n^2 + C_n + 1}{12} + (1 - C_n)^2 \qquad (4.38)$$
$$=: \hat{C}_n.$$

For $\varepsilon_n \to 0$, all the terms in (4.38) are bounded and we have that $\limsup_n \hat{C}_n < \infty$. We summarize

$$\overline{\Psi}_{v_n^0,n}(f_n, v_n) \leq \alpha \int_{-1}^{1} |f'|^2 dx + \beta\big(v_n^0(\xi_n)^2 \hat{C}_n + \hat{O}_n\big).$$

By Lemma 4.3, if $\limsup_n \frac{dist(K^0,x_n)}{\varepsilon_n} < \infty$, then $v_n^0(x_n) \to 0$ as $\varepsilon_n \to 0$. As $\xi_n \in [\varepsilon_n^2, \varepsilon_n^2 + \varepsilon_n]$ we therefore have $v_0^n(\xi_n) \to 0$, and thus, letting the limit pass, we obtain

$$\limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n) \leq \alpha \int_{-1}^{1} |f'|^2 dx. \qquad (4.39)$$

□

4.4 The N-dimensional case

In the previous sections we established lim inf and lim sup inequalities for the variational approximation in one dimension. In this section we will lift our results to dimension $N \geq 2$.

For the lim inf equality this is done by a slicing technique following Braides [2002][p. 188ff] which is the standard approach for such kind of variational approximations. For the lim sup inequality we follow Attouch et al. [2006]. Other references are Ambrosio and Tortorelli [1990]; Attouch et al. [2006]; Braides [1998].

We first fix the notation. Let $\Omega \subset \mathbb{R}^N$, $N \geq 2$ be an open set. We write

$$\overline{\Psi}_{K^0}(f,v) = \alpha \int_{\Omega \setminus K} |\nabla f|^2 dx + \beta \mathcal{H}^{N-1}(S_f \setminus K^0) \tag{4.40}$$

for $f \in SBV(\Omega)$, $v = 1$ *a.e.*, and else $\overline{\Psi}_{K^0}(f,v) = +\infty$ and

$$\overline{\Psi}_{v_n^0,n}(f,v) = \alpha \int\limits_{\Omega} v^2 |\nabla f|^2 dx + \beta \int\limits_{\Omega} \varepsilon_n |\nabla v|^2 + \frac{(v_n^0 - v)^2}{4\varepsilon_n} dx \tag{4.41}$$

for $v, f \in W^{1,2}(\Omega)$, $0 \leq v \leq 1$.

4.4.1 The lim inf inequality through slicing

We now prove the lim inf inequality

$$\overline{\Psi}_{K^0}(f,v) \leq \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n)$$

for every $\{(f_n, v_n)\}$ converging to (f,v) in $L^2(\Omega) \times L^2(\Omega)$ for dimensions $N \geq 2$. We follow the steps as described in Braides [2002][p. 188ff]. Once the one dimensional convergence in Theorem 4.7 is established, the N-dimensional case follows by standard arguments. This can also be seen in the proof below, as we do not have to address the *a priori* edge in any step.

Lemma 4.9 (The lim inf inequality)**.** *Let $\varepsilon_n \to 0$ as $n \to \infty$ and $\{v_n^0\}$, K^0 fulfill Assumptions 4.1. For every sequence $\{(f_n, v_n)\}$ converging to (f,v) in*

$L^2(\Omega) \times L^2(\Omega)$ *it then holds*

$$\overline{\Psi}_{K^0}(f,v) \leq \liminf_n \overline{\Psi}_{v_n^0,n}(f_n, v_n). \tag{4.42}$$

We will use the following notations in the proof. Let S^{N-1} be the $N-1$ dimensional sphere. For every $\nu \in S^{N-1}$ we define

$$\begin{aligned}
\pi_\nu &:= \{y \in \mathbb{R}^N : \langle y, \nu \rangle = 0\},\\
\Omega_y &:= \{t \in \mathbb{R} : y + t\nu \in \Omega\}, y \in \pi_\nu,\\
\Omega_\nu &:= \{y \in \pi_\nu : \Omega_y \neq \emptyset\}.
\end{aligned}$$

Furthermore we define for $f : \Omega \to \mathbb{R}$ and y in Ω_ν the function f_y for all $t \in \Omega_y$ by $f_y(t) = f(y + t\nu)$.

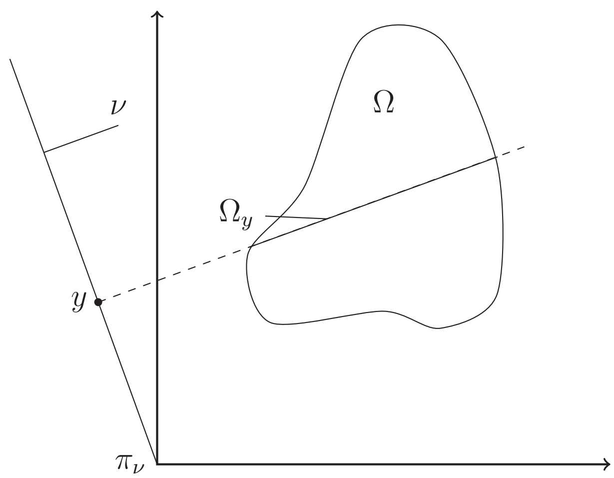


Figure 4.1: The domain Ω and a slice Ω_y for a fixed $\nu \in S^{N-1}$.

Proof. 1. *'Localize' the functional $\overline{\Psi}_{v_n^0,n}$ highlighting its dependence on the set of integration.*

For all open sets $A \subset \Omega$ we define

$$\overline{\Psi}_{v_n^0,n}(f,v,A) := \alpha \int_A v^2 |\nabla f|^2 dx + \beta \int_A \varepsilon_n |\nabla v|^2 + \frac{(v_n^0 - v)^2}{4\varepsilon_n} dx$$

for $v, f \in W^{1,2}(\Omega)$, $0 \le v \le 1$, and $\overline{\Psi}_{v_n^0,n}(f,v) = \infty$ else.

2. *We first fix $\nu \in S^{N-1}$ and define for all $y \in \pi_\nu$ one dimensional functionals:*

$$\overline{\Psi}^{\nu,y}_{v_n^0,n}(f,v,I) := \alpha \int\limits_I v^2|f'|^2 dx + \beta \int\limits_I \varepsilon_n |v'|^2 + \frac{(v_n^0 - v)^2}{4\varepsilon_n} dx$$

for $I \subset \mathbb{R}$ open and bounded, $f, v \in W^{1,2}(I), 0 \le v \le 1$ and $\overline{\Psi}^{\nu,y}_{v_n^0,n}(f,v,I) := \infty$ else (actually independent of y). By integrating over all y in π_ν we define the functional

$$\overline{\Psi}^{\nu}_{v_n^0,n}(f,v,A) := \int\limits_{\pi_\nu} \overline{\Psi}^{\nu,y}_{v_n^0,n}(f_y, v_y, A_y) d\mathcal{H}^{N-1}$$

for $A \subset \Omega$ open and bound, $f, v \in W^{1,2}(A), 0 \le v \le 1$. By Fubini's theorem this is

$$\overline{\Psi}^{\nu}_{v_n^0,n}(f,v,A) = \alpha \int\limits_A v^2 |\langle \nu, \nabla f\rangle|^2 dx + \beta \int\limits_A \varepsilon_n |\langle \nu, \nabla v\rangle|^2 + \frac{(v_n^0 - v)^2}{4\varepsilon_n} dx$$

if $\langle \nu, Df\rangle << \mathcal{L}^N$, $\langle \nu, Dv\rangle << \mathcal{L}^N$, $0 \le v \le 1$ and $\overline{\Psi}^{\nu}_{v_n^0,n}(f,v,A) = \infty$ otherwise.

3. *Compute the one dimensional* Γ-$\liminf_n \overline{\Psi}^{\nu,y}_{v_n^0,n}(f,v,I)$.

   By Theorem 4.7 we have

$$\overline{\Psi}^{\nu,y}_{K^0}(f,v,I) := \Gamma\text{-}\liminf_n \overline{\Psi}^{\nu,y}_{v_n^0,n}(f,v,I) = \alpha \int_I |f'|^2 dx + \beta \mathcal{H}^0(S_f \setminus K^0)$$

and we can define

$$\overline{\Psi}^{\nu}_{K^0}(f,v,A) := \int\limits_{\pi_\nu} \overline{\Psi}^{\nu,y}_{K^0}(f_y, v_y, A_y) d\mathcal{H}^{N-1}$$

for $A \subset \Omega$ open and bound, $f, v \in W^{1,2}(A), 0 \le v \le 1$.

Note that $\overline{\Psi}^{\nu}_{K^0}(f,v,A)$ is finite if and only if $v = 1$ a.e. in A, $f_y \in SBV(A_y)$

for $\mathcal{H}^{N-1}$ a.e. $y \in \pi_\nu$. If in addition $f \in L^\infty(\Omega)$ we have

$$\begin{aligned}&\int_{\pi_\nu} |Df_y|(\Omega_y) d\mathcal{H}^{N-1} \\ &\leq \int_{\pi_\nu} \Big[\int_{\Omega_y} \big(|f_y'|^2 + 2\|f\|_\infty \mathcal{H}^0(f_y)\big)\Big] d\mathcal{H}^{N-1} < \infty. \end{aligned} \tag{4.43}$$

4. *Apply Fatou's lemma.*

   For $(f_n, v_n) \to (f, v)$ in $L^2(\Omega) \times L^2(\Omega)$ we have

$$\begin{aligned}\liminf_n \overline{\Psi}_{v_n^0,n}(f,v,A) &\geq \liminf_n \overline{\Psi}^\nu_{v_n^0,n}(f,v,A) \\ &= \liminf_n \int_{\pi_\nu} \overline{\Psi}^{\nu,y}_{v_n^0,n}(f_y,v_y,A_y) d\mathcal{H}^{N-1} \\ &\geq \int_{\pi_k} \liminf_n \overline{\Psi}^{\nu,y}_{v_n^0,n}(f_y,v_y,A_y) d\mathcal{H}^{N-1} \\ &\geq \int_{\pi_k} \overline{\Psi}^{\nu,y}_{K^0}(f_y,v_y,A_y) d\mathcal{H}^{N-1} \\ &= \overline{\Psi}^\nu_{K^0}(f,v,A).\end{aligned}$$

   In the first inequality and first equality we used the definition of $\overline{\Psi}^\nu_{v_n^0,n}$. The second inequality follows from Fatou's lemma and the third from Theorem 4.7. The final equality is the definition of $\overline{\Psi}^\nu_{K^0}$. We can deduce that $\liminf_n \overline{\Psi}_{v_n^0,n}(f,v,A) \geq \overline{\Psi}^\nu_{K^0}(f,v,A)$ for all $\nu \in S^{N-1}$.

5. *Describe the domain of* Γ-$\liminf_n \overline{\Psi}_{v_n^0,n}(f,v,A)$.

   If f lies in $L^\infty(\Omega)$, by Equation (4.43) and Theorem 2.27 we deduce that Γ-$\liminf_n \overline{\Psi}_{v_n^0,n}(f,v,A)$ is finite if $f \in SBV(A)$ and $v = 1$ a.e.

6. *Obtain a direction dependent estimate.*

   If $f \in SBV(A)$ and $v = 1$ a.e. from Theorem 2.27 we have

$$\overline{\Psi}^\nu_{K^0}(f,v,A) = \alpha \int_A |\langle \nabla f, \nu\rangle|^2 dx + \beta \int_{A\cap(S_f\setminus K^0)} |\langle \nu_f, \nu\rangle| d\mathcal{H}^{N-1}$$

where ν_f is the normal on S_f as defined in Definition 2.18. We can therefore summarize

$$\begin{aligned}\liminf_n \overline{\Psi}_{v_n^0,n}(f,v,A) \geq \\ \sup_{\nu\in S^{N-1}} \Big(\alpha \int_A |\langle \nabla f, \nu\rangle|^2 dx + \beta \int_{A\cap(S_f\setminus K^0)} |\langle \nu_f, \nu\rangle| d\mathcal{H}^{N-1}\Big).\end{aligned} \tag{4.44}$$

7. *Optimize the lower estimate.*

We now apply Lemma 2.7 to optimize the lower estimate (4.44). Let $\{\nu_i\}$ be a dense sequence in S^{N-1}. We first define the measure $\mu_f = \mathcal{L}^N|_\Omega + \nu_f|_{S_f}$ and the functions

$$\phi_i(x)) = \begin{cases} |\langle \nabla f, \nu_i\rangle|^2, & x \notin S_f \setminus K^0 \\ |\langle \nu_f, \nu_i\rangle|, & x \in S_f \setminus K^0. \end{cases} \tag{4.45}$$

We then have

$$\begin{aligned}\int_\Omega \sup_i \phi_i(x)\, d\mu_f &= \int_{\Omega\setminus(S_f\setminus K^0)} \sup_i \phi_i(x)\, d\mu_f + \int_{\Omega\cap(S_f\setminus K^0)} \sup_i \phi_i(x)\, d\mu_f \\ &= \int_\Omega |\nabla f|^2\, dx + \int_{\Omega\cap(S_f\setminus K^0)} d\mathcal{H}^{N-1}.\end{aligned}$$

We can then apply Lemma 2.7 to conclude

$$\begin{aligned}\liminf_n \overline{\Psi}_{v_n^0,n}(f,v) &\geq \sup_{\{A_i\}} \Big[\sum_{i\in I} \liminf_n \overline{\Psi}_{v_n^0,n}(f,v,A_i)\Big] \\ &\geq \sup_{\{A_i\}} \Big[\sum_{i\in I} \alpha \int_{A_i} |\langle \nabla f, \nu_i\rangle|^2 dx + \beta \int_{A_i\cap(S_f\setminus K^0)} |\langle \nu_f, \nu_i\rangle| d\mathcal{H}^{N-1}\Big] \\ &= \alpha \int_\Omega |\nabla f|^2\, dx + \beta\mathcal{H}^{N-1}(S_f \setminus K^0) \\ &= \overline{\Psi}_{K^0}(f,v),\end{aligned} \tag{4.46}$$

where the supremum is taken over all finite families $\{A_i\}_{i\in I}$ of pairwise disjoint open subsets of Ω.

□

4.4.2 The $\limsup$ inequality through density

Finally we show the existence of a recovery sequence $\{(f_n, v_n)\}$ converging to (f, v) so that

$$\limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n) \leq \overline{\Psi}_{K^0}(f, v) \tag{4.47}$$

for every (f, v) in $L^2(\Omega) \times L^2(\Omega)$ for dimensions $N \geq 2$. We will reduce the N dimensional case to the one dimensional following the presentation in [Attouch et al., 2006, p. 492ff]. The arguments passing from the N dimensional case to the one dimensional are again independent of our *a priori* edge information $\{v_n^0\}$ and K^0.

For the proof, we need to assume certain regularity of the domain Ω. We assume that Ω satisfies the following "refection condition" on $\partial\Omega$: there exists an open neighborhood U of $\partial\Omega$ in $\mathbb{R}^N$ and an injective Lipschitz function $\phi : U \cap \Omega \to U \cap \overline{\Omega}$ such that ϕ^{-1} is Lipschitz.

Lemma 4.10 ($\limsup$ inequality)**.** *Let $\varepsilon_n \to 0$ as $n \to \infty$ and $\{v_n^0\}$, K^0 fulfill Assumptions 4.1 and Ω fulfill the reflection condition. For every (f, v) in $L^2(\Omega) \times L^2(\Omega)$ there exists a sequence $\{(f_n, v_n)\}$ converging to (f, v) for which*

$$\limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n) \leq \overline{\Psi}_{K^0}(f, v). \tag{4.48}$$

The idea of the proof is to modify $(f, 1)$ in a neighborhood of S_f to obtain, from the expression of $\overline{\Psi}_{v_n^0,n}(f_n, v_n)$, an equivalent of $\alpha \int_\Omega |\nabla f|^2 dx + \beta \mathcal{H}^{N-1}(S_f \setminus K^0)$. We will design the function in the same way as in the one dimensional case, only depending on the distance of a point x to the edge set. We write for any $x \in \mathbb{R}^N$ and any set $A \subset \mathbb{R}^N$

$$d(x, A) = \inf_{y \in A} \|x - y\|^2. \tag{4.49}$$

Proof. We first assume the following regularity condition on the jump set S_f: It holds $\mathcal{H}^{N-1}(\overline{S_f} \setminus\ S_f) = 0$ and for every compact set $A \subset \Omega$ it holds

$$\lim_{\rho \to 0} \frac{\mathcal{L}^N(\Omega \cap (S_f \setminus A)_\rho)}{2\rho} = \mathcal{H}^{N-1}(S_f \setminus A), \tag{4.50}$$

where $(S_f)_\rho$ is the tubular neighborhood $\{x \in \mathbb{R}^N : d(x, S_f) < \rho\}$ of order ρ around S_f.

Let us fix $a_n := \varepsilon_n^2 + \sqrt{\varepsilon_n}$, $b_n := \varepsilon_n^2 + \varepsilon_n$ and $c_n := \varepsilon_n^2 + 2\sqrt{\varepsilon_n}$ with corresponding sequence of tubular neighborhoods $(\overline{S_f})_{a_n}$, $(\overline{S_f})_{b_n}$, $(\overline{S_f})_{c_n}$. We assume that n is large enough, such that $b_n \leq a_n \leq c_n$. We again set $C_n = \exp(\frac{-1}{2\sqrt{\varepsilon_n}})$.

We will separate our domain in the following way:

$$\begin{aligned} \Omega = \Omega \setminus (\overline{S_f})_{a_n} \quad &\cup \quad (\overline{S_f} \setminus K^0)_{a_n} \quad \cup \quad (\overline{S_f} \cap K^0)_{a_n} \setminus (\overline{S_f} \setminus K^0)_{c_n} \\ &\cup \quad \Big((\overline{S_f} \cap K^0)_{a_n} \cap (\overline{S_f} \setminus K^0)_{c_n}\Big) \setminus (\overline{S_f} \setminus K^0)_{a_n}. \end{aligned}$$

The first three parts corresponds to the cases in dimension one, the last is needed to make sure that the sequence can be patched such that $v_n \in W^{1,2}(\Omega)$.

We set $f_n(x) = f(x)$ if $d(x, S_f) \geq a_n$ and extend each function f_n such that $f_n \in W^{1,2}(\Omega)$.

On $\Omega \setminus (\overline{S_f})_{a_n}$ and $(\overline{S_f} \setminus K^0)_{a_n}$ we set $v_n(x) = v_n^0(x) \cdot \phi_n(d(x, S_f))$ where $\phi_n : \mathbb{R} \to \mathbb{R}$ is defined as

$$\phi_n(t) := \begin{cases} 0, & t \leq \varepsilon_n^2 \\ 1 - \exp(\frac{\varepsilon_n^2 - t}{2\varepsilon_n}), & \varepsilon_n^2 < t < a_n \\ 1 - C_n, & t \geq a_n. \end{cases} \tag{4.51}$$

On $(\overline{S_f} \cap K^0)_{a_n} \setminus (\overline{S_f} \setminus K^0)_{c_n}$ we define $v_n(x) = v_n^0(x) \cdot \hat{\phi}_n(d(x, S_f))$ with

$$\hat{\phi}_n(t) = \begin{cases} 0, & t \leq \varepsilon_n^2 \\ (1 - C_n) \cdot \frac{t - \varepsilon_n^2}{\varepsilon_n}, & \varepsilon_n^2 < t < b_n \\ \\ 1 - C_n, & t \geq b_n. \end{cases} \tag{4.52}$$

On $\Big((\overline{S_f} \cap K^0)_{a_n} \cap (\overline{S_f} \setminus K^0)_{c_n}\Big) \setminus (\overline{S_f} \setminus K^0)_{a_n}$ we define $v_n(x) = v_n^0(x) \cdot \tilde{\phi}_n(d(x, S_f))$

with

$$\tilde{\phi}_n(t) = \frac{1}{c_n - a_n}\big((t-a_n)\hat{\phi}_n(t) + (c_n - t)\phi_n(t)\big). \tag{4.53}$$

Now we compute the integral on each domain.

On $\mathbf{\Omega \setminus (\overline{S_f})_{a_n}}$ we have

$$\begin{aligned}\overline{\Psi}_{v_n^0,n}(f_n, v_n, \Omega \setminus (\overline{S_f})_{a_n}) &\leq \int_{\Omega\setminus(\overline{S_f})_{a_n}} \left(\alpha (v_n^0)^2|\nabla f|^2 + \beta\frac{1}{4\varepsilon_n}(C_n)^2 + \beta\varepsilon_n|C_n\nabla v_n^0|^2\right)dx \\ &\leq \int_{\Omega\setminus(\overline{S_f})_{a_n}} \alpha|\nabla f|^2 dx + \beta\int_{\Omega\setminus(\overline{S_f})_{a_n}} \varepsilon_n|\nabla v_n^0|^2 + \beta\frac{1}{4\varepsilon_n}(C_n)^2 dx.\end{aligned}$$

We define

$$O_n^1 := \beta\int_{\Omega\setminus(\overline{S_f})_{a_n}} \varepsilon_n|\nabla v_n^0|^2 + \beta\frac{1}{4\varepsilon_n}(C_n)^2 dx.$$

As in the one dimensional case $O_n^1 \to 0$ as $n \to \infty$.

On $\mathbf{(\overline{S_f} \setminus K^0)_{a_n}}$ we compute

$$\begin{aligned}&\overline{\Psi}_{v_n^0,n}(f_n, v_n, (\overline{S_f}\setminus K^0)_{a_n}) = \int_{(\overline{S_f}\setminus K^0)_{a_n}} \left(\alpha v_n^2|\nabla f|^2 + \frac{\beta}{4\varepsilon}(v_n^0 - v_n)^2 + \beta\varepsilon|\nabla v_n|^2\right) \\ &\leq \alpha\int_{(\overline{S_f}\setminus K^0)_{a_n}} |\nabla f|^2 dx \\ &+ \beta\int_{(\overline{S_f}\setminus K^0)_{a_n}} \frac{1}{4\varepsilon}(1-\phi_n(d(x,S_f)))^2 + \varepsilon|\nabla(\phi_n(d(x,S_f))|^2 + \varepsilon|\nabla(v_n^0)|^2 dx\end{aligned}$$

and define

$$O_n^2 := \beta\int_{(\overline{S_f}\setminus K^0)_{a_n}} \varepsilon|\nabla(v_n^0)|^2 dx.$$

Besides O_n^2 the remaining integral over the edge part is the original Ambrosio-Tortorelli functional. We will address it in Lemma 4.11 below.

On $(\overline{\mathbf{S_f}} \cap \mathbf{K^0})_{\mathbf{a_n}} \setminus (\overline{\mathbf{S_f}} \setminus \mathbf{K^0})_{\mathbf{c_n}}$ we compute

$$\begin{aligned}&\overline{\Psi}_{v_n^0,n}(f_n, v_n, (\overline{S_f} \cap K^0)_{a_n} \setminus (\overline{S_f} \setminus K^0)_{a_n}) \\ &\leq \int_{(\overline{S_f} \cap K^0)_{a_n} \setminus (\overline{S_f} \setminus K^0)_{c_n}} \alpha |\nabla f|^2 dx + \int_{(\overline{S_f} \cap K^0)_{a_n}} \left(+ \frac{\beta}{4\varepsilon}(v_n^0 - v_n)^2 + \beta\varepsilon |\nabla v_n|^2 \right).\end{aligned}$$

For the edge we focus on the integration over $\hat{\phi}$, as the rest vanishes in the limit. By mean value theorem we have for some $\xi_n \in (\overline{S_f} \cap K^0)_{a_n}$

$$\begin{aligned}&\int_{(\overline{S_f} \cap K^0)_{a_n}} (v_n^0(x))^2 \big(\frac{1}{4\varepsilon}(1 - \hat{\phi}_n(d(x, S_f)))^2 + \varepsilon |\nabla (\hat{\phi}_n(d(x, S_f)))|^2 \big) dx \\ &= (v_n^0(\xi_n))^2 \int_{(\overline{S_f} \cap K^0)_{a_n}} \big(\frac{1}{4\varepsilon}(1 - \hat{\phi}_n(d(x, S_f)))^2 + \varepsilon |\nabla (\hat{\phi}_n(d(x, S_f)))|^2 \big) dx.\end{aligned}$$

As $\frac{dist(\xi_n, S_f)}{\varepsilon_n} < \infty$, from Lemma 4.3 , we have that $v_n^0(\xi_n) \to 0$. Moreover the remaining integral is uniformly bounded by Lemma 4.11 and we define

$$\begin{aligned}O_n^3 :=& \beta(v_n^0(\xi_n))^2 \int_{(\overline{S_f} \cap K^0)_{a_n}} \big(\frac{1}{4\varepsilon}(1 - \hat{\phi}_n(d(x, S_f)))^2 + \varepsilon |\nabla (\hat{\phi}_n(d(x, S_f)))|^2 \big) dx \\ &+ \beta \int_{(\overline{S_f} \cap K^0)_{a_n}} \varepsilon |\nabla (v_n^0)|^2 dx.\end{aligned}$$

On $\Big((\overline{\mathbf{S_f}} \cap \mathbf{K^0})_{\mathbf{a_n}} \cap (\overline{\mathbf{S_f}} \setminus \mathbf{K^0})_{\mathbf{c_n}}\Big) \setminus (\overline{\mathbf{S_f}} \setminus \mathbf{K^0})_{\mathbf{a_n}}$ we write

$$O_n^4 := \overline{\Psi}_{v_n^0,n}(f_n, v_n, \Big((\overline{S_f} \cap K^0)_{a_n} \cap (\overline{S_f} \setminus K^0)_{c_n}\Big) \setminus (\overline{S_f} \setminus K^0)_{a_n}).$$

We estimate the integral in Lemma A 3 in the appendix and have that $O_n^4 \to 0$ as $n \to \infty$.

We can then conclude

$$
\begin{aligned}
&\limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n) \\
&\leq \limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, \Omega \setminus (\overline{S_f})_{a_n}) \\
&+ \limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, (\overline{S_f} \setminus K^0)_{a_n}) \\
&+ \limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, (\overline{S_f} \cap K^0)_{a_n} \setminus (\overline{S_f} \setminus K^0)_{c_n}) \\
&+ \limsup_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, \Big((\overline{S_f} \cap K^0)_{a_n} \cap (\overline{S_f} \setminus K^0)_{c_n}\Big) \setminus (\overline{S_f} \setminus K^0)_{a_n}) \\
&\leq \alpha \limsup_n \int_{\Omega \setminus \overline{S_f}_{a_n}} |\nabla f|^2 dx + \alpha \limsup_n \int_{(\overline{S_f} \setminus K^0)_{a_n}} |\nabla f|^2 dx \\
&+ \beta \limsup_n \int_{(\overline{S_f} \setminus K^0)_{a_n}} \frac{1}{4\varepsilon}(1 - \phi_n(d(x, S_f)))^2 + \varepsilon |\nabla(\phi_n(d(x, S_f))|^2 dx \\
&+ \limsup_n \left(O_n^1 + O_n^2 + O_n^3 + O_n^4\right) \\
&= \alpha \int_\Omega |\nabla f|^2 dx + \beta \mathcal{H}^{N-1}(S_f \setminus K^0).
\end{aligned}
$$

The second step is to construct a sequence $f_n \to f$, where every f_n fulfills the regularity assumption (4.50) and that $\overline{\Psi}_{K^0}(f, 1) = \lim_n \overline{\Psi}_{K^0}(f_n, 1)$. The proof then follows from a diagonalization argument. We refer to [Attouch et al., 2006, p. 494 *Second step*] on how to construct such a sequence. In the mentioned proof it suffices to take $K = \overline{\Omega} \cap (\overline{S_f \setminus K^0})$. This step also makes use of the reflection condition . □

To conclude the proof above we need the following lemma following [Ambrosio and Tortorelli, 1990, Proposition 5.1]

Lemma 4.11. *Let $\varepsilon_n, a_n \to 0$ as $n \to \infty$ with $\frac{\varepsilon_n}{a_n} \to 0$ and $\{v_n^0\}$, K^0 fulfill Assumptions 4.1. If for $f \in SBV(\Omega)$ it holds*

$$
\lim_{\rho \to 0} \frac{\mathcal{L}(\Omega \cap (S_f \setminus K^0)_\rho)}{2\rho} = \mathcal{H}^{N-1}(S_f \setminus K^0), \tag{4.54}
$$

and we have a sequence of functions $\phi_n \in W^{1,2}([0,1])$, *for which*

$$\limsup_n \big(\frac{1}{4\varepsilon_n}\big(1-\phi_n(a_n)\big)^2 + \varepsilon_n|\phi_n(a_n)'|^2\big) < \infty, \tag{4.55}$$

then for any $A \subset \Omega$ *it is*

$$\begin{aligned}&\limsup_n \int_{(\overline{S_f}\setminus A)_{a_n}} \frac{1}{4\varepsilon_n}\big(1-\phi_n(d(x,S_f))\big)^2 + \varepsilon_n|\nabla\phi_n(d(x,S_f))|^2 dx \\ &\le \limsup_n \int_0^{a_n} \frac{1}{4\varepsilon_n}\big(1-\phi_n(t)\big)^2 + \varepsilon_n|\phi_n(t)'|^2 dt \cdot \mathcal{H}^{N-1}(\overline{S_f}\setminus A).\end{aligned}$$

Proof. Following [Ambrosio and Tortorelli, 1990, Proposition 5.1] we rewrite the integral

$$\begin{aligned}&\int_{(\overline{S_f}\setminus A)_{a_n}} \frac{1}{4\varepsilon}(1-\phi_n(d(x,S_f)))^2 + \varepsilon|\nabla(\phi_n(d(x,S_f)))|^2 dx \\ &= \int_0^{a_n} \Big(\frac{1}{4\varepsilon}(1-\phi_n(t))^2 + \varepsilon|\nabla(\phi_n(t))|^2\Big)\, \mathcal{H}^{N-1}[d(x,\overline{S_f}\setminus A)=t]\, dt.\end{aligned}$$

We define $h(t) = \mathcal{L}^N(d(x,\overline{S_f}\setminus A) < t)$, then by [Attouch et al., 2006, Corollary 4.2.3, p. 138] it is $h'(t) = \mathcal{H}^{N-1}[d(x,\overline{S_f}\setminus A) = t]$. Furthermore, we define $z_n(t) := \frac{1}{4\varepsilon_n}(1-\phi_n(t))^2 + \varepsilon_n|(\phi_n'(t))|^2$. We therefore can write

$$\begin{aligned}&\int_{(\overline{S_f}\setminus A)_{a_n}} \frac{1}{4\varepsilon}(1-\phi_n(d(x,S_f)))^2 + \varepsilon|\nabla(\phi_n(d(x,S_f)))|^2 dx \\ &= \int_0^{a_n} z_n(t)h'(t)dt.\end{aligned}$$

We can then use integration by parts and arrive at

$$\int_0^{a_n} z_n(t)h'(t)dt = [z(t)h(t)]_0^{a_n} - \int_0^{a_n} z'(t)h(t)dt.$$

The term $[z(t)h(t)]_0^{a_n}$ vanishes in the limit $n \to \infty$ because of the assumption (4.55) and (4.54). Also by (4.54) we have that for all $\eta > 0$ there exists a τ so that for all $t < \tau$ it is $h(t) \ge 2t(\mathcal{H}^{N-1}(\overline{S_f}\setminus A) - \eta)$. By this regularity and the definition of a_n we can find a sequence $\eta_n \to 0$ such that

$$-\int_0^{a_n} z_n'(t)h(t)dt \leq -2(\mathcal{H}^{N-1}(\overline{S_f} \setminus A) - \eta_n) \int_0^{a_n} z_n'(t)tdt.$$

A second integration by parts leads to

$$\int_0^{a_n} z'(t)\, tdt = [z(t)t]_0^{a_n} - \int_0^{a_n} z(t)dt.$$

The term $[z(t)t]_0^{a_n}$ vanishes in the limit $n \to \infty$ because of the assumption (4.55).

Together we have

$$\begin{aligned}
&\limsup_n \int_{(\overline{S_f} \setminus A)_{a_n}} \frac{1}{4\varepsilon}(1-\phi_n(x))^2 + \varepsilon|\nabla(\phi_n)|^2 dx \\
&\leq \limsup_n \left([z(t)h(t)]_0^{a_n} + (\mathcal{H}^{N-1}(\overline{S_f} \setminus A) - \eta_n)\Big(\int_0^{a_n} z(t)dt - [z(t)t]_0^{a_n} \Big) \right) \\
&= \int_0^{a_n} z(t)dt\, \mathcal{H}^{N-1}(\overline{S_f} \setminus A).
\end{aligned}$$

□

4.5 A heuristic approximation for $\gamma \neq 0$

In this section we introduce a second penalty term, that can be considered as an approximation of the case where $\gamma \neq 0$. The penalty for which we showed the convergence results in the sections above has numerical shortcomings. Most importantly, our experiments indicate that the *a priori* edge K^0 is always included in the detected edge, even if the data does not support an edge at a given point, see Figure 5.4. In some cases it is desirable to give the edge detector more flexibility, including *a priori* edges only when the measured data also supports this.

To this end we propose the penalty

$$H_{v_n^0,n}(f,v) = \alpha \int_\Omega v^2 |\nabla f|^2 dx + \beta \int_\Omega \big(\varepsilon_n |\nabla v|^2 + \frac{(1-v)^2}{4\varepsilon_n}\big)(1+\gamma(v_n^0 - v)^2)dx$$

for $f, v \in W^{1,2}(\Omega), 0 \leq v, \leq 1$ with $0 < \beta, \gamma, \alpha$.

We consider it a heuristic approximation of

$$\overline{\Psi}_{H_{K^0}}(f,v) = \alpha \int_{\Omega \setminus K} |\nabla f|^2 dx + \beta \mathcal{H}^{N-1}(S_f) + \gamma\beta\, \mathcal{H}^{N-1}(S_f \setminus K^0)$$

for parameters $0 < \beta, \gamma, \alpha$. The reasoning behind this is as follows. In the phase field setting, the integral

$$\int_{\Omega} \big(\varepsilon_n |\nabla v|^2 + \frac{(1-v)^2}{4\varepsilon_n}\big) dx$$

approximates $\mathcal{H}^{N-1}(S_f)$. In the case where v is close to the constant function 1 , i.e. there is no edge, this integral is small and the edge part in $H_{v_n^0,n}$ has a negligible contribution. Thus the factor $(1+\gamma(v_n^0 - v)^2)$ is not of big importance. In the case where v indicates an edge, that is $v(x) \approx 0$, the factor $(1+\gamma(v_n^0 - v)^2)$ decides how strongly it is penalized. If $v(x)$ is close to $v_n^0(x)$ (which is the case $x \in S_f \cap K^0$), then the edges are approximately penalizes by β. If $v(x)$ is close to 0 but $v_n^0(x)$ close to 1 (which is the case $x \in S_f \setminus K^0$), then the edges are approximately penalized by $\beta(1+\gamma)$. We show some reconstructions with this penalty in the next chapter.

Chapter 5

Applications

In this chapter we apply the variational approximations of

$$\min_{(f,K)} \left(\int_\Theta |A(f)-g|^2 dx + \alpha \int_{\Omega\setminus K} |\nabla f|^2 dx + \beta \mathcal{H}^{N-1}(K \setminus K^0) + \gamma \mathcal{H}^{N-1}(K \cap K^0) \right), \tag{5.1}$$

that we introduced in the previous chapter to two inverse problems. The first problem is 2D X-ray CT with parallel beam geometry. The second is the identification of the scattering coefficient in 2D diffuse optical tomography. Both inverse problems are covered by our theoretical setting and the examples are based on simulated data.

We compare the standard Mumford-Shah penalty (MS), the modified Mumford-Shah penalty (4.2) (MS_{K^0}), and the heuristic penalty (4.4) (H_{K^0}). In both examples we furthermore choose an additional regularization method not making use of the *a priori* knowledge. For X-ray CT we choose a smoothed TV penalty , see Rudin et al. [1992]. For 2D diffuse optical tomography we use a Landweber method as comparison, see Hanke et al. [1995].

The computations were done using Matlab (version R2013a). For X-ray CT we used the implementation of the Radon Transform and the adjoint operator written by Lutz Justen from the Software-Documentation of the Center for Industrial Mathematics, University of Bremen. For diffuse optical tomography we use the Toast package from Martin Schweiger and Simon Arridge Schweiger and Arridge

[2014] for the forward operator and the adjoint of its derivative.

As we only consider simulated data we can evaluate the quality of reconstruction f_{rec} by comparing it with the true image f_{true} via the peak signal to noise ratio

$$PSNR(f_{rec}, f_{true}) := 10 \log_{10} \left(\frac{\max |f_{true}|^2}{MSE(f_{true}, f_{rec})} \right), \tag{5.2}$$

where $MSE(f_{true}, f_{rec})$ is the mean square error and via the structural similarity index measure

$$SSIM(f_{rec}, f_{true}) := \frac{(2\mu_{f_{true}}\mu_{f_{rec}} + c_1)(2\sigma_{f_{true}f_{rec}} + c_2)}{(\mu^2_{f_{true}} + \mu^2_{f_{rec}} + c_1)(\sigma^2_{f_{true}} + \sigma^2_{f_{rec}} + c_2)}, \tag{5.3}$$

where $\mu_{f_{true}}$, $\mu_{f_{rec}}$ are averages, $\sigma^2_{f_{true}}$, $\sigma^2_{f_{rec}}$ are variances and $\sigma_{f_{true}f_{rec}}$ the covariance. The factors $c_1, c_2 > 0$ stabilize the division. The $SSIM$ returns values in $[-1, 1]$, where the maximum similarity $SSIM = 1$ is obtained only for identical images. For the $PSNR$ the larger the value the better. The $SSIM$ often gives a better indication of similarity between images than the $PSNR$. See Hore and Ziou [2010] for a comparison of the two quality measures.

Alternating minimization

Let $\overline{\Psi}_1(f, v) = \overline{\Psi}_{v^0_n,n}(f, v)$ as in (4.2), $\overline{\Psi}_2(f, v) = H_{v^0_n,n}(f, v)$ as in (4.4) and $\overline{\Psi}_3(f, v) = \mathrm{AT}_{\varepsilon_n}(f, v)$ be the original Ambrosio-Tortorelli penalty as in (1.4).

For $i = 1, 2, 3$ we solve the minimization problem

$$\min_{(f,v)} \left(\|A(f) - g\|^2_{L^2(\Theta)} + \overline{\Psi}_i(f, v) \right) \tag{5.4}$$

in an alternating manner described in Algorithm 1.

Data: g, parameters
Result: reconstruction, edge indicator
Initialization;
$v_0 = 1$;
for $j = 0$ **to** $NumberOfIterations$ **do**
  % Reconstruction:
  $f_j = \text{argmin}_f \left[\|A(f) - g\|^2_{L^2(\Theta)} + \alpha \int\limits_\Omega v_j^2 |\nabla f|^2 dx \right]$;
  % Edge detection:
  $v_{j+1} = \text{argmin}_v \overline{\Psi}_i(f_j, v)$;
end
reconstruction = f_j;
edge indicator = v_{j+1};

Algorithm 1: The alternate minimization to compute a reconstruction and its edge indicator function by solving the minimization problem 5.4 for $i = 1, 2, 3$. The three methods only differ in the edge detection step.

The algorithm alternates between a reconstruction step where f is updated and an evaluation of f in which v is updated. The obtained information of the second step is in form of the edge indicator function v, which is then used to update the regularization penalty for the reconstruction step. These steps are repeated several times. In each step we use a simple gradient descent method. We choose the descent direction as the negative gradient and the step size by backtracking line search with the Armijo-Goldstein stopping condition.

In the next sections we test the proposed method. For both applications we first state the mathematical model we are using and show that they are covered by our setting. For this it is necessary to prove that the fidelity decays of high enough order for pointwise bound functions.

5.1 Two dimensional X-ray CT

We first give a short review of a simple mathematical model for the X-ray computer tomography. The following is from [Natterer, 2001, Chapter II] and [Louis, 1989, Chapter 6].

In X-ray tomography, the image contrast comes from the X-ray absorption when X-ray beams pass through an object. The interaction of X-ray and the object could be a complex process. Nevertheless, the beam path in straight lines provides a good approximation for X-ray tomography in many cases.

A simple model is given by Beer's law. Let $I(x)$ be the intensity of an X-ray and $f(x)$ the X-ray attenuation coefficient at a point x. A X-ray passing a small distance Δx at x has an approximate relative intensity loss of

$$\frac{I(x+\Delta x) - I(x)}{I(x)} = f(x)\Delta x. \tag{5.5}$$

Let I_0 be the initial intensity of the X-ray and I_1 the intensity after passing through the object. As we assume that the beam travels in a straight line L, from (5.5) it follows

$$\frac{I_1}{I_0} = \exp^{-\int_L f(x)dx}$$

and taking the logarithm it is

$$-\ln\left(\frac{I_1}{I_0}\right) = \int_L f(x)dx.$$

We see that a measurement I_1 and the initial intensity I_0 give us a line integral of the X-ray attenuation coefficient f.

The operator mapping a function into the set of its line integrals in two dimensions is the ***Radon Transform***. The task in X-ray computer tomography is to invert this operator.

Let $\Omega \subset \mathbb{R}^2$ be an open bounded domain and f be the X-ray attenuation coefficient function on Ω. We can assume that the domain is the unit disk $\Omega := \{x \in \mathbb{R}^2 \mid |x| \leq 1\}$. The sphere in $\mathbb{R}^2$ is denoted as $S^1 := \{x \in \mathbb{R}^2 \mid |x| = 1\}$.

The Radon Transform R maps a function f into the set of its line integrals. Let $L(\sigma, \omega) \subset \Omega$ be a line in Ω, parameterized by a distance $\sigma \in [-1, 1]$ to the origin and a vector $\omega \in S^1$ perpendicular to the line. The Radon Transform is defined for every $f \in \{g \in C^\infty \mid x^k \frac{\partial^l}{\partial x^l} g$ is bounded for every $l, k \in \mathbb{N}_0\}$ as

$$Rf(L) := \int_L f(x)dx$$

for every line $L(\sigma, \omega) \subset \Omega$ or equivalently

$$Rf(\sigma, \omega) := \int_{\langle x, \omega \rangle = 0} f(\sigma\omega + x)dx.$$

We call $Z := [-1, 1] \times S^1$ the *Radon domain.* The following theorem states that the Radon Transform has a well defined extension from $L^2(\Omega)$ to $L^2(Z)$.

Theorem 5.1 ([Louis, 1989, p. 166. Theorem 6.1.1])**.** *The Radon Transform R has a continuous extension, still denoted by R, mapping from $L^2(\Omega)$ to $L^2(Z)$,*

$$R : L^2(\Omega) \to L^2(Z).$$

It follows that the Radon Transform has a Hilbert space adjoint.

Theorem 5.2 ([Louis, 1989, p.168 Theorem 6.1.4])**.** *Let $R : L^2(\Omega) \to L^2(Z)$ be the Radon Transform. Then $R^* : L^2(Z) \to L^2(\Omega)$ with*

$$R^*g(x) = \int_{S^1} g(\langle x, \omega \rangle, \omega)dx$$

is the adjoint operator of R. We call R^* the **backprojection***.*

The function R^*g maps a point $x \in \Omega$ to the integral over all line integrals, for which the line passes through x.

There exist inversion formulas for the Radon Transform (see [Natterer, 2001, p.18 Theorem 2.1]), but they are not feasible as they require complete and exact data. In practice inverting the Radon Transform from incomplete and noisy data is an ill-posed problem. Even very small errors in the data $g \in L^2(Z)$, which are not avoidable in practice, may lead to bad reconstructions.

In the following remark we discuss that the Mumford-Shah regularization can be applied to X-ray CT with a least squares fidelity term.

Remark 5.3. *It is well known that the Radon transform is bounded as* $R : L^1(B_1(0)) \to L^1([-1,1] \times S^1)$*, see Natterer [2001]. Furthermore it can easily be shown that* $R : L^\infty(B_1(0)) \to L^\infty([-1,1] \times S^1)$ *is also a bounded operator, see Page [2011]. Therefore, in the same way as for the image deblurring problem, Proposition 3.10 can be applied using the exponents* $q = \hat{q} = 1$ *and* $q' = \infty$*. As a result the least squares fidelity term with the Radon Transform decays with order* N *for pointwise bound functions.*

5.1.1 Numerical examples I: X-ray CT

In this section we present numerical examples illustrating the behavior of our proposed model compared to the standard Mumford-Shah regularization and TV regularization. The examples are created with synthetically generated data and are meant as a starting point to numerical investigations rather than an exhaustive study.

We compare our results with TV regularization, which is a variational method that is also able to reconstruct sharp edges. The minimization problem for TV regularization is

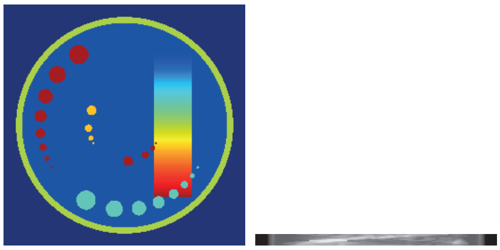
$$\min_f \int_\Theta |R(f) - g|^2 dL + \lambda \int_\Omega |\nabla f| dx, \tag{5.6}$$

for $f \in BV(\Omega)$ and $\lambda > 0$, see Rudin et al. [1992]. Compared to the smoothing term in the Mumford-Shah penalty, here the the L_1 norm instead of the L_2 norm of the gradient is measured. Unfortunately, the TV functional is not a smooth function of the image f and requires advanced convex optimization methods to be minimized Zhang et al. [2011]. We follow an alternative approach and replace the absolute value by a smoothed absolute value. The smoothed TV norm reads:

$$\Psi^h_{TV}(f) = \int_\Omega \sqrt{\|\nabla f\|^2 + h^2} dx, \tag{5.7}$$

with $h > 0$. When h tends to zero, the smoothed TV penalty becomes closer to the original one.

In the following examples we reconstruct the phantom illustrated in Figure 5.1.



$f^\dagger$ g^δ

Figure 5.1: The true image $f^\dagger$ (left) and the corresponding noisy projection data from 10 angles (right). The phantom is mostly piece-wise constant and has a linear slope on the right side.

Example 5.4 (MS and TV regularization)**.** *In the first example we compare the standard Mumford-Shah regularization without a priori edge knowledge to the smoothed TV regularization. The data is obtained from 10 views with* 4% *relative noise. The noise is additive Gaussian noise. Fig 5.2 shows the reconstructions for different weights on the smoothing penalty. For the Mumford-Shah reconstruction the edge weight is kept fixed. For TV the best reconstruction with regards to the peak-signal to noise ratio* ($PSNR$) *is in the third column, for the Mumford-Shah reconstruction it is in the second. The reconstructions from this sparsely sampled and noisy data have comparable quality. The different regions are more sharply divided for the Mumford-Shah reconstruction if the edges are detected. This can be seen for the circles in bottom or top left of the phantom. Depending on the regularization parameters the linear slope is reconstructed with or without a staircasing effect for the Mumford-Shah regularization.*

Example 5.5 (TV, MS, MS_{K^0} and H_{K^0} regularization)**.** *In this example we compare the smoothed TV regularization, standard Mumford-Shah regularization to the variational models with the Mumford-Shah priori using the a priori edge knowledge* $\overline{\Psi}_{MS_{K^0}}$ *and the heuristic penalty* $\overline{\Psi}_{H_{K^0}}$*. The data is again obtained from 10 views with* 4% *relative noise. Fig 5.3 shows the reconstructions, where the parameters were optimized with regard to the* $PSNR$ *of the reconstruction to the true image. The small circles at the bottom are reconstructed better when then a priori knowledge is used. This is expected, as it is additional correct side*

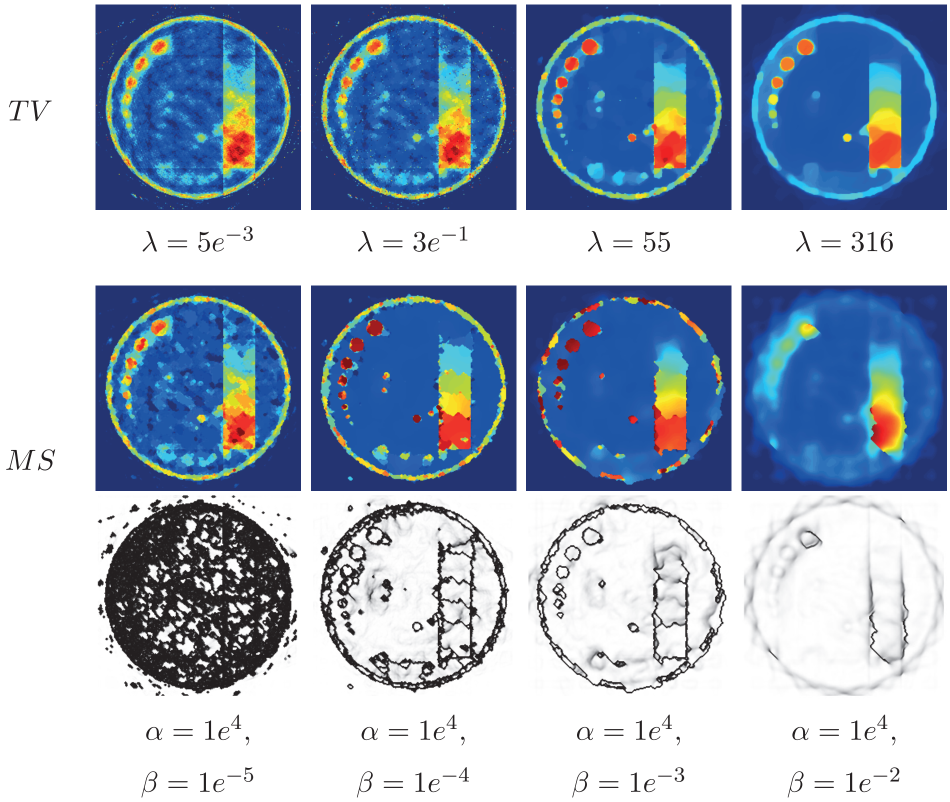
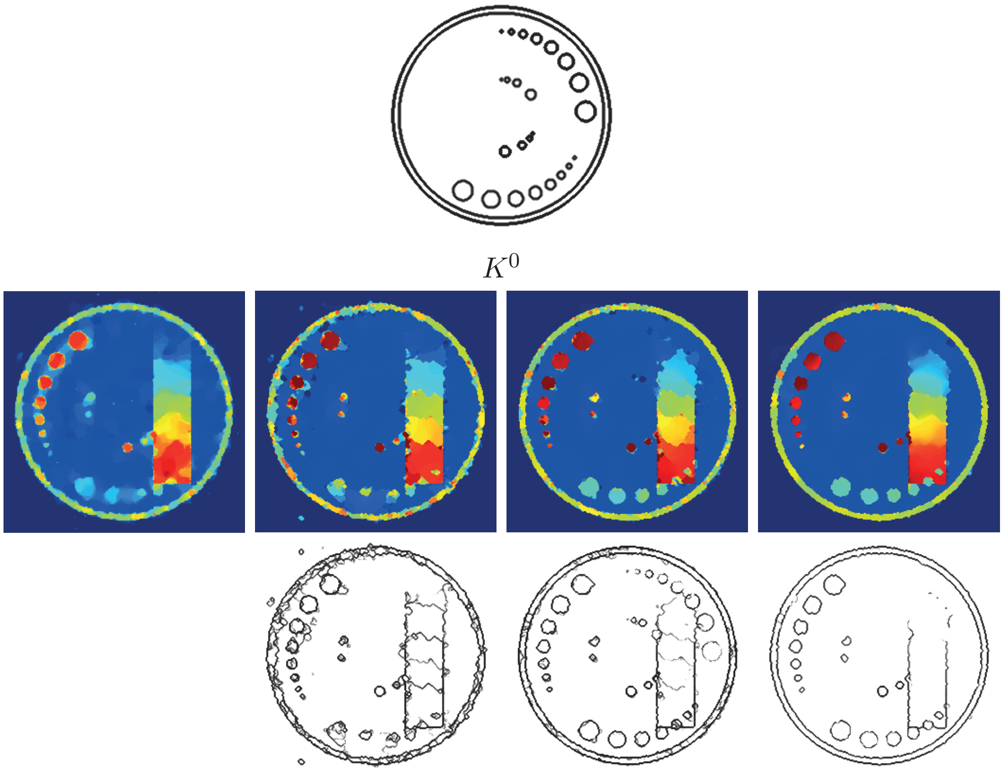


Figure 5.2: Reconstruction from 10 views with 4% relative noise. (top) TV reconstruction for different weights on the smoothing penalty, (middle) and (bottom) the Mumford-Shah reconstruction and edge set for different weights on the edge term and fixed weight on the smoothing parameter. For TV regularization the best reconstruction with regards to the peak-signal to noise ratio is in the third column with $PSNR = 11.51$ and $SSIM = 0.67$, for the Mumford-Shah reconstruction it is in the second with $PSNR = 10.89$ and $SSIM = 0.76$. The Mumford-Shah reconstructions also illustrate the non-convexity of the approach. Once an edge is smoothed away in the reconstruction step, it is lost and the reconstruction can not be guided back to it.

information. The false side information is detected if we use the $\overline{\Psi}_{MS_{K^0}}$ *penalty, for the heuristic penalty this is not the case. If the false edges are included, they yield artifacts in the reconstructed image.*

Example 5.6 (Varying parameters for $\Phi_{MS_{K^0}}$)**.** *In this example we illustrate the behavior of the penalty* $\overline{\Psi}_{MS_{K^0}}$*. We keep the smoothness parameter fixed and show reconstructions for different* β*. The data is again obtained from* 10 *views with* 4%



TV	MS	MS_{K^0}	H_{K^0}
$\lambda = 55$	$\alpha = 1.26e^3$,	$\alpha = 1e^4$,	$\alpha = 1e^4$,
	$\beta = 1e^{-3}$	$\beta = 1e^{-3}$	$\beta = 1e^{-4}$, $\gamma = 5$

Figure 5.3: Reconstruction from 10 views with 4% relative noise. Top: *a priori* edge information. Bottom rows: reconstruction and edge set from (left) TV regularization, standard Mumford-Shah regularization, Mumford-Shah regularization using *a priori* edge knowledge, and (right) the heuristic model. The areas where no *a priori* edge knowledge is available the Mumford-Shah reconstructions are similar, although slightly worse for the standard Mumford-Shah regularization. The correct side information improves the reconstruction as can be seen from the circles in the bottom of the images. The incorrect edge knowledge is only included in the third column, which results in some small artifacts in the reconstruction. For the TV reconstruction we have $PSNR = 11.51$ and $SSIM = 0.67$, for the MS reconstruction $PSNR = 10.89$ and $SSIM = 0.76$, for the MS_{K^0} reconstruction $PSNR = 12.7$ and $SSIM = 0.79$ and for the H_{K^0} reconstruction $PSNR = 12$ and $SSIM = 0.81$.

relative noise. Fig 5.4 shows the reconstructions.

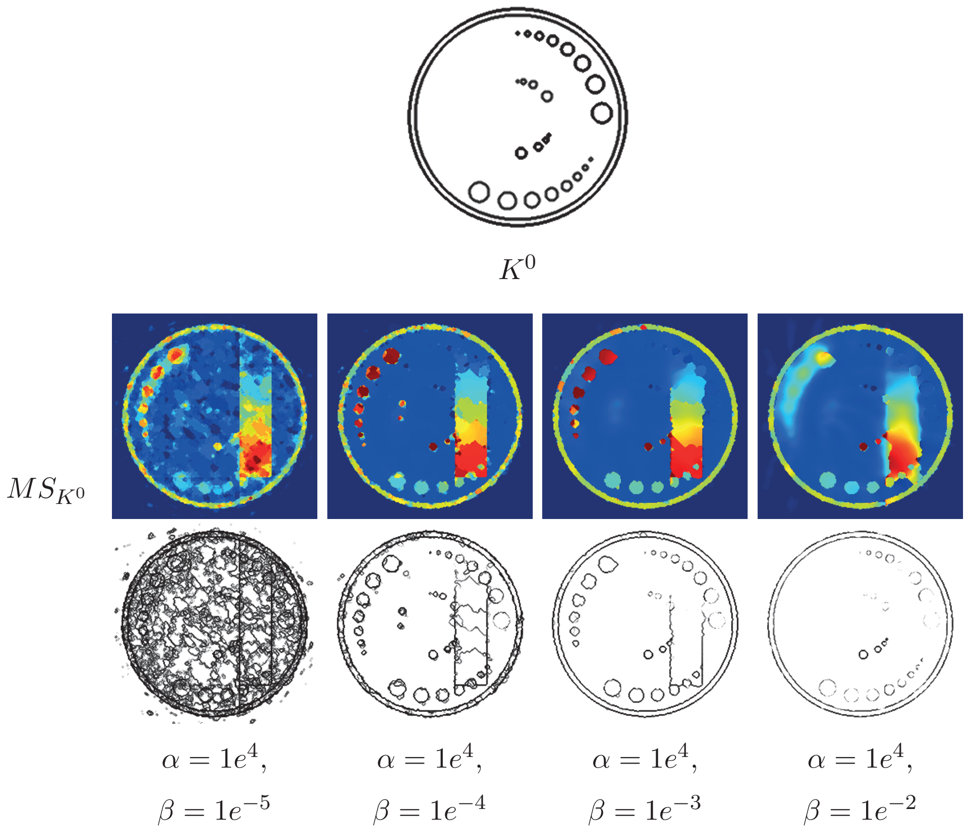


Figure 5.4: Reconstruction from 10 views with 4% relative noise using the penalty $\overline{\Psi}_{MS_{K^0}}$. The penalty on the smoothness term is kept fixed and the weight on the edge penalty increases from left to right. The *a priori* edge is almost always included in the detected edge set and yields strong artifacts. In regions with correct additional knowledge the reconstructions are considerably better than with the standard Mumford-Shah regularization.

Example 5.7 (Varying parameters for $\overline{\Psi}_{H_{K^0}}$)**.** *In this example we illustrate the behavior of the penalty* $\overline{\Psi}_{H_{K^0}}$*. We keep the smoothness parameter fixed, set* $\gamma = 5$ *and show reconstructions for different* β*. The data is again obtained from* 10 *views with* 4% *relative noise. Fig 5.5 shows the reconstructions. As can be seen, for increasing* β *the a priori edges are not detected anymore.*

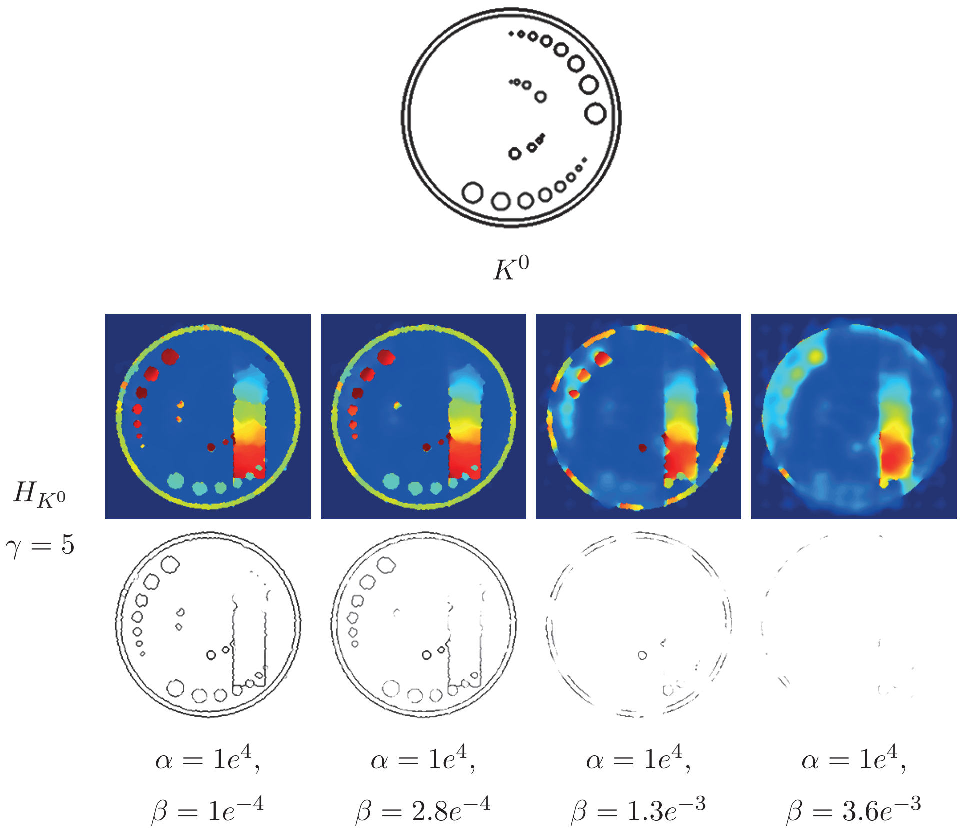


Figure 5.5: Reconstruction from 10 views with 4% relative noise using the penalty $\overline{\Psi}_{H_{K^0}}$.The weight on the smoothness penalty is kept fixed and the weight on the edge penalty increases from left to right. The *a priori* edge is only used when the data also supports an edge. In regions with correct additional knowledge the reconstructions are considerably better than with the standard Mumford-Shah regularization.

5.2 Two dimensional Diffuse Optical Tomography

We first give a description of the mathematical model we will consider for two dimensional diffuse optical tomography. For a detailed description see for example Arridge and Schotland [2009].

In steady-state diffuse optical tomography (DOT), the problem is to reconstruct the diffusion and absorption coefficients of an object using measurements of ingoing and corresponding outgoing near-infrared light passing through the object. The governing equation used for the diffuse light u in DOT is

$$-\mathrm{div}\,(D\nabla u) + \mu u = 0, \qquad \text{in } \Omega, \tag{5.8}$$

which is a diffuse approximation of the radiative transport equation. Here D is the diffusion coefficient and μ the absorption coefficient. We assume that D, $\mu \in L^\infty(\Omega)$.

We assume a single measurement is taken by shedding light into the object at its boundary and measuring the corresponding outgoing light at the boundary (or part of the boundary).

The incoming light $g_{\mathfrak{R}} \in L^2(\partial\Omega)$ can be modeled through a Robin boundary condition

$$u + 2D\nu \cdot \nabla u = g_{\mathfrak{R}}, \qquad \text{on } \partial\Omega, \tag{5.9}$$

where $\nu \in \mathbb{R}^N$ is the outer normal. The measurement $g_{\mathfrak{N}} \in L^2(\partial\Omega)$ is the negative Neumann boundary values of the solution u of (5.8)

$$g_{\mathfrak{N}} = -D\nu \cdot \nabla u, \tag{5.10}$$

either on the entire boundary, $\partial\Omega$, or part of it $\Gamma \subset \partial\Omega$, see Arridge [1999]; Arridge and Schotland [2009] for details.

For a pair D and μ under imaging with m-incoming light sources $g^i_{\mathfrak{R}} \in L^2(\partial\Omega)$, assume we have measured the corresponding Neumann data $g^i_{\mathfrak{N}} \in L^2(\partial\Omega)$ for $i = 1, 2, ..., m$. We assume the light sources $g^i_{\mathfrak{R}}$ are defined by the user and known. Therefore the measurements can be equivalently described as Dirichlet data

$$g_{\mathfrak{D}}^i = u^i_{|\partial\Omega} = g_{\mathfrak{R}}^i + 2g_{\mathfrak{N}}^i, \qquad \text{on } \partial\Omega. \tag{5.11}$$

Then we can define the forward operator $F : L^2(\Omega) \times L^2(\Omega) \to (L^2(\partial\Omega))^m$ that maps each pair of parameters (μ, D) to the Dirichlet data $F^i(\mu, D) = g_{\mathfrak{D}}^i \in L^2(\partial\Omega)$ of the solutions of (5.8) and (5.9) respectively, for $g_{\mathfrak{R}}^i$, $i = 1, \cdots, m$. Let $g_{\mathfrak{D}} = (g_{\mathfrak{D}}^1, ..., g_{\mathfrak{D}}^m) \in L^2(\partial\Omega)^m$ be a set of measured data. DOT is then to solve the operator equation

$$F(\mu, D) = g_{\mathfrak{D}}. \tag{5.12}$$

5.2.1 Decay of the least squares fidelity term

First we verify that the forward operator together with a least squares fidelity term fits into our theoretical framework, that is we need to verify that the fidelity term decays with order $N - 1 + \varepsilon$ for pointwise bound functions and $\varepsilon > 0$. As shown below we can only prove this for dimension 2, for higher dimensions this is still an open problem.

For the existence and uniqueness of the weak solutions of the boundary value problem (5.8), (5.9) we assume the following conditions.

Assumption 5.8.

1. *The function D is uniformly positive and bounded: there exist $a_{\mathfrak{D}}, b_{\mathfrak{D}} > 0$ such that $a_{\mathfrak{D}} \leq D \leq b_{\mathfrak{D}}$ on Ω.*

2. *The function μ is non-negative and bounded from above; i.e. there exists $b_\mu > 0$ such that $0 \leq \mu \leq b_\mu$.*

Moreover, in this section let Ω have at least Lipschitz boundary.

By the Lax-Milgram theorem and the Sobolev trace and embedding theorems Egger and Schlottbom [2010], for every $g_{\mathfrak{R}} \in L^2(\partial\Omega)$ there exists a unique weak solution $u \in W^{1,2}(\Omega)$ of (5.8) with the boundary values (5.9), that is for all $v \in W^{1,2}(\Omega)$:

$$\int_\Omega (D\nabla u \cdot \nabla v + \mu uv) \; dx + \frac{1}{2}\int_{\partial\Omega} uv \; dx = \frac{1}{2}\int_{\partial\Omega} g_{\mathfrak{R}} v \; dx. \tag{5.13}$$

We will need the following regularity theorem from Egger and Schlottbom [2010].

Theorem 5.9. *Let Assumption 5.8 hold. Then there exists a constant $p_0 > 2$ depending only on the domain and the bounds for the coefficients, such that the solution u of the variational problem (5.13) lies in $W^{1,p}(\Omega)$ whenever $g_{\mathfrak{R}} \in L^p(\partial\Omega)$ for some $\frac{p_0}{p_0-1} \leq p \leq p_0$. Moreover, there holds the a priori estimate*

$$\|u\|_{W^{1,p}(\Omega)} \leq C\|g_{\mathfrak{R}}\|_{L^{\hat{p}}(\partial\Omega)} \tag{5.14}$$

with a constant C that depends only on Ω and the bounds for the coefficients. If the domain Ω has a smooth boundary, and if $a_{\mathfrak{D}}/b_{\mathfrak{D}}$ approaches one, then the maximal p_0 such that the statement of the theorem holds, tends to infinity.

If $p > \frac{N}{N-1}$, then $\hat{p} = p/2$ for dimension $N = 2$ and $\hat{p} = 2p/3$ for dimension $N = 3$ respectively, see [Egger and Schlottbom, 2010, Remark 3.9]. As we will see later this restricts us to dimension $N = 2$.

In the following we consider the data fidelity term

$$(\mu, D) \mapsto \|F(\mu, D) - g_{\mathfrak{D}}\|^2_{(L^2(\partial\Omega))^m} \tag{5.15}$$

and show that it decays with order $N - 1 + \varepsilon$ for some $\varepsilon > 0$ for pointwise bound functions. In Rondi and Santosa [2001] a similar result is proven for electric impedance tomography. We follow their proof here. Assume that $\partial\Omega$ and the bounds $a_{\mathfrak{D}}, b_{\mathfrak{D}}$ are such that Theorem 5.9 yields a $p_0 > 4$. We now show that under these conditions, in dimension two, the forward operator F is Lipschitz continuous from $L^p(\Omega) \times L^p(\Omega) \to (L^2(\Omega))^m$ for $1 \leq p < \frac{N}{N-1}$. The required property for the fidelity term then follows from Proposition 3.10.

First we consider a single source $g^i_{\mathfrak{R}} \in L^2(\partial\Omega) \cap L^\infty(\Omega)$. Let (μ_0, D_0) and (μ_1, D_1) satisfy the conditions in Assumption 5.8 and $u_0, u_1 \in W^{1,2}(\Omega)$ be the respective weak solutions for (5.13) with the same incoming light $g^i_{\mathfrak{R}}$. Then we have for all $v \in W^{1,2}(\Omega)$

$$\int_\Omega (D_0 \nabla u_0 \cdot \nabla v + \mu_0 u_0 v) + \frac{1}{2}\int_{\partial\Omega} u_0 v = \int_\Omega (D_1 \nabla u_1 \cdot \nabla v + \mu_1 u_1 v) + \frac{1}{2}\int_{\partial\Omega} u_1 v.$$

Subtracting $\int_\Omega D_0 \nabla u_1 \nabla v$ and $\int_\Omega \mu_0 u_1 v$ from each side leads to

$$\int_\Omega (D_0 \nabla(u_0 - u_1) \cdot \nabla v + \mu_0(u_0 - u_1)v) + \frac{1}{2}\int_{\partial\Omega} (u_0 - u_1)v \\ = \int_\Omega ((D_1 - D_0)\nabla u_1 \cdot \nabla v + (\mu_1 - \mu_0)u_1 v)\,. \quad (5.16)$$

Let $w \in W^{1,2}(\Omega)$ be the solution to the following auxiliary boundary value problem

$$-\mathrm{div}\,(D_0 \nabla w) + \mu_0 w = 0, \quad \text{in } \Omega$$
$$w + 2D_0 \nu \cdot \nabla w = (u_0 - u_1), \quad \text{on } \partial\Omega.$$

Choosing $v = (u_0 - u_1)$ as a test function for the auxiliary problem and using equation (5.16), we obtain

$$\frac{1}{2}\int_{\partial\Omega} (u_0 - u_1)^2 = \int_\Omega (D_0 \nabla(u_0 - u_1) \cdot \nabla w + \mu_0(u_0 - u_1)w) + \frac{1}{2}\int_{\partial\Omega} (u_0 - u_1)w \\ = \int_\Omega ((D_1 - D_0)\nabla u_1 \cdot \nabla w + (\mu_1 - \mu_0)u_1 w)\,.$$

Via Hölders inequality we arrive at

$$\frac{1}{2}\int_{\partial\Omega} (u_0 - u_1)^2 \leq C \left(\|D_1 - D_0\|_{L^{p_1}(\Omega)} + \|\mu_1 - \mu_0\|_{L^{p_1}(\Omega)}\right) \|u_1\|_{W^{1,p_2}(\Omega)} \|w\|_{W^{1,p_3}(\Omega)}$$

with $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$.

For the decay property of the fidelity term we need not just $p_1 < \frac{N}{N-1}$ but also $\hat{p_3}$ from Theorem 5.9 to be smaller than or equal to 2. For $N = 2$ we can choose $p_2 = p_0 > 4$, $p_3 = 4$ and $p_1 = \frac{4p_0}{4p_0 - 4 - p_0} < 2$. From Theorem 5.9 we have

$$\|u_1\|_{W^{1,p_2}} \leq \|g_R^i\|_{L^{\frac{p_2}{2}}(\partial\Omega)} \quad (5.17)$$

and

$$\|w\|_{W^{1,p_3}} \leq \|u_0 - u_1\|_{L^2(\partial\Omega)}. \quad (5.18)$$

With these inequalities it follows for some constants $C > 0$ and $C_1 > 0$,

$$\begin{aligned} \|F^i(\mu_1, D_1) - F^i(\mu_0, D_0)\|_{L^2(\partial\Omega)} &= \|u_1 - u_0\|_{L^2(\partial\Omega)} \\ &\leq C\left(\|D_1 - D_0\|_{L^{p_1}} + \|\mu_1 - \mu_0\|_{L^{p_1}}\right)\|g_R^i\|_{L^{\frac{p_2}{2}}(\partial\Omega)} \\ &\leq C_1\|(\mu_1, D_1) - (\mu_0, D_0)\|_{L^{p_1}(\Omega)\times L^{p_1}(\Omega)}. \end{aligned} \tag{5.19}$$

In the same way it can be shown that for multiple light sources $(g_{\mathfrak{R}}^1, ..., g_{\mathfrak{R}}^m) \in L^2(\partial\Omega)^m \cap L^\infty(\Omega)^m$, it holds

$$\|F(\mu_1, D_1) - F(\mu_0, D_0)\|_{(L^2(\partial\Omega))^m} \leq L\|(\mu_1, D_1) - (\mu_0, D_0)\|_{L^{p_1}(\Omega)\times L^{p_1}(\Omega)} \tag{5.20}$$

where $L > 0$ is a constant and $1 \leq p_1 < \frac{N}{N-1}$. Proposition 3.10 can straightforwardly be extended to the case where the data is collected in $(L^2(\partial\Omega))^m$ and the reconstruction consists of multiple functions. If in Proposition 3.10 we choose $q' = \hat{q} = 2$ and $q = p_1$, then (5.20) yields the desired decay property of the fidelity term with order $\frac{N}{p_1} > N - 1$.

5.2.2 Numerical examples II: recovery of the absorption

In this section we consider a special case of DOT. We are interested in recovering the absorption coefficient μ, and will assume the diffusion coefficient D^* to be known in the following. This is a interesting but computational simpler example, see Egger [2010]. It is still a non-linear and highly ill-posed problem.

We keep the notation from the section above. For a fixed $D^* \in L^\infty(\Omega)$ and given light sources $g^i_{\mathfrak{R}}$, $i = 1, \cdots, m$, we introduce the new forward operator $G : L^2(\Omega) \to (L^2(\partial\Omega))^m$ as

$$G(\mu) = F(\mu, D^*). \tag{5.21}$$

The inverse problem then reduces to solve the operator equation

$$G(\mu) = g_{\mathfrak{D}}, \tag{5.22}$$

where $g_{\mathfrak{D}} = (g^1_{\mathfrak{D}}, ..., g^m_{\mathfrak{D}}) \in L^2(\partial\Omega)^m$ is the measured data.

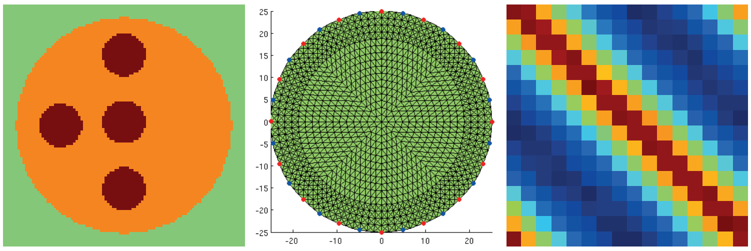
We compare our results with the iterative scheme

$$\mu_{j+1} = \mu_j + c_j G'(\mu_j)^*(g^\delta - G(\mu_j)), \tag{5.23}$$

where j is the iteration number, $G'(\mu_j)^*$ is the adjoint of the derivative of G at a given point μ_j and c_j is a step size, see Hanke et al. [1995] We choose the stopping index to maximize the $PSNR$ of the reconstruction to the true image.

In the following we investigate the reconstructions of a simple 100×100 phantom, which is shown on the left in Figure 5.6. The actual computational domain is only in the inner circle as depicted in the middle of Figure 5.6. We have 16 sources and 16 detectors equivalently distributed on the computational domain indicated by the red and blue points respectively. The data is shown on the (right), it is noiseless and was created on a much finer mesh to avoid the inverse crime.

Example 5.10 (MS and Landweber regularization)**.** *In the first example we compare the standard Mumford-Shah regularization without a priori edge knowledge to the the Landweber regularization for different parameters and stopping index.*



$\mu^\dagger$ Ω_n g

Figure 5.6: The true image (left), the actual computational domain for the reconstruction with light sources and detectors shown (middle) and the synthetic data (right). The domain corresponds to an object with 2 cm diameter.

Fig 5.7 shows the reconstructions for noise free data. Visually the Mumford-Shah reconstruction are sharper, but the ball in the center is lost, whereas it can still be guessed for the Landweber method. Also some artifacts appear at the boundaries of the computational domain.

Example 5.11 (LW, MS, MS_{K^0} and H_{K^0} regularization)**.** *In this example we compare Landweber regularization and standard Mumford-Shah regularization to the variational models using the a priori edge knowledge. The data is again noise free. Fig 5.8 shows the reconstructions, where the parameters were adjusted by hand. It can be seen that the a priori edge knowledge improves the reconstruction if the edges coincide with the ones from the true image. The method MS_{K^0} also introduces the false edges from the a priori edge set, the heuristic penalty does not do this. The false edges introduce small artifacts in the reconstruction.*

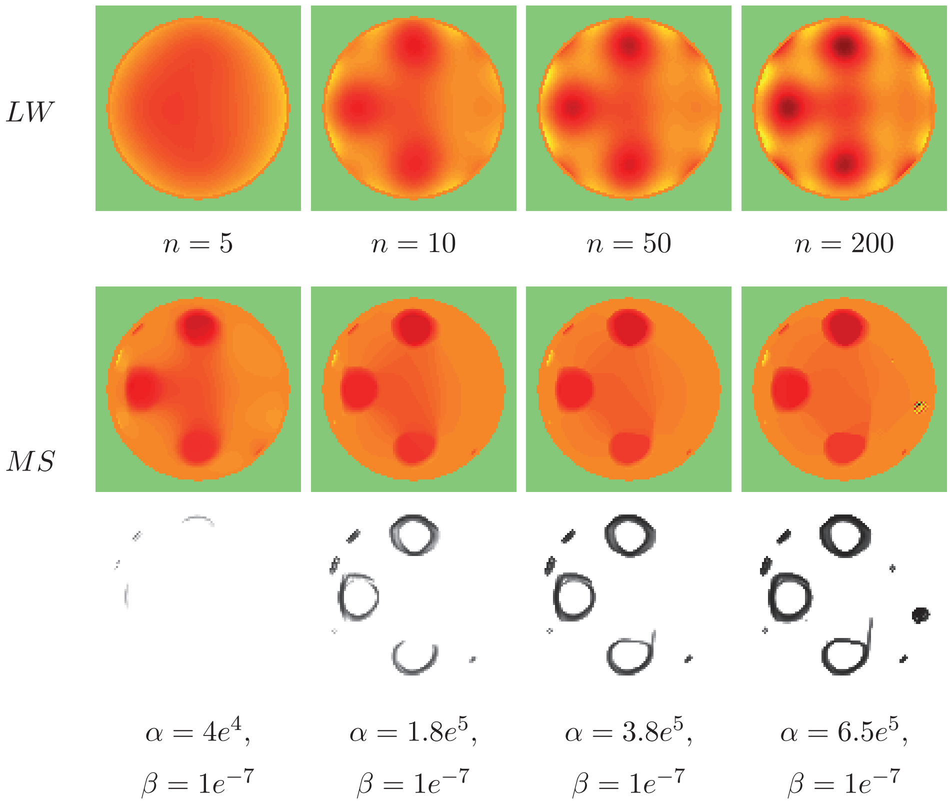


Figure 5.7: Reconstructions from noise free data. (top) Landweber reconstruction for different stopping index, (middle) and (bottom) the Mumford-Shah reconstruction and edge set for different weights on the smoothing term and a fixed weight on the edge term. Although the edges in the Mumford-Shah reconstruction are sharper, there are some clear artifacts at the edge of the computational domain.

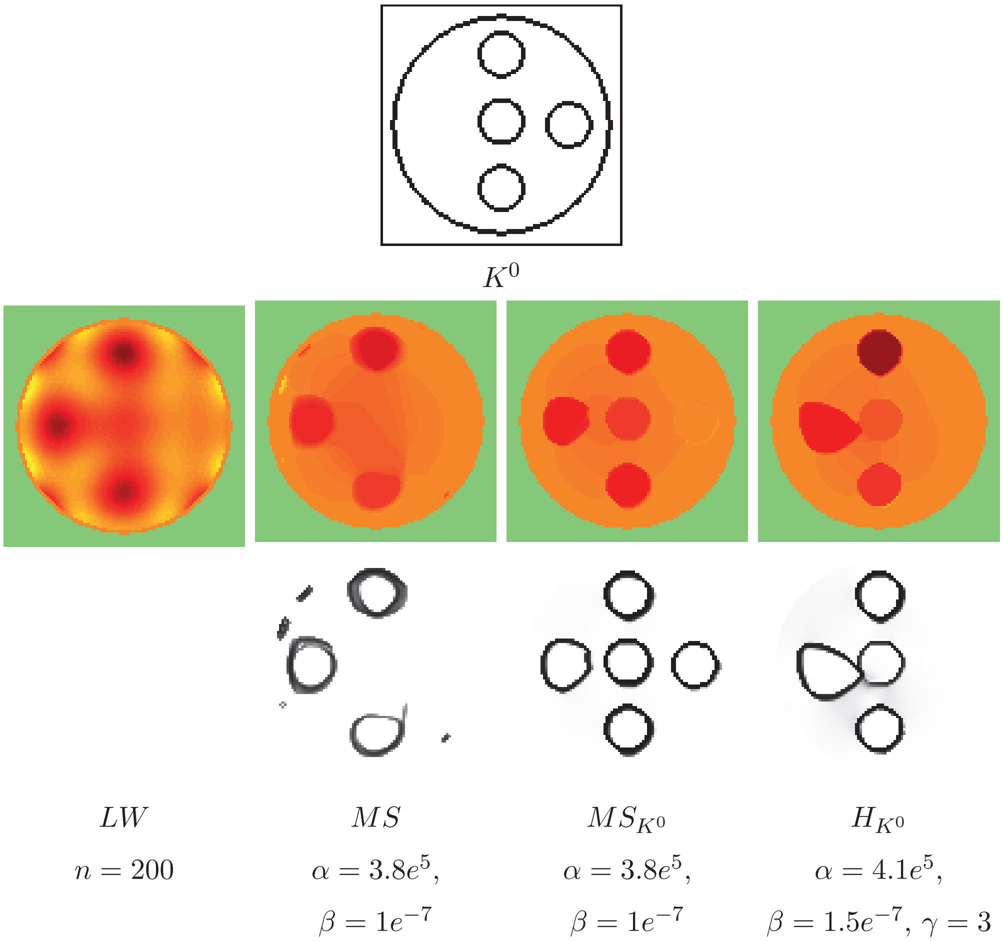


Figure 5.8: Reconstruction from noise free data. Top: *a priori* edge information. Bottom rows: reconstruction and edge set from (left) Landweber regularization, standard Mumford-Shah regularization, Mumford-Shah regularization using *a priori* edge knowledge, and (right) the heuristic model. In areas where no *a priori* edge knowledge is available the Mumford-Shah reconstructions are similar, although slightly worse for H_{K^0}. The correct side information improves the reconstruction as can be seen from the circle in the center. The incorrect edge knowledge is only included in the third column. For the LW reconstruction we have $PSNR = 25.75$ and $SSIM = 0.84$, for the MS reconstruction $PSNR = 25.80$ and $SSIM = 0.89$, for the MS_{K^0} reconstruction $PSNR = 28.04$ and $SSIM = 0.94$ and for the H_{K^0} reconstruction $PSNR = 26.31$ and $SSIM = 0.92$.

Chapter 6

Conclusions

In this thesis we study a variational approach based on the Mumford-Shah model for image reconstruction with *a priori* edge information. We presume that the *a priori* information was obtained beforehand, possibly from a secondary modality that is less ill-posed, an application specific template or a reconstruction at a previous time point. The difference of our approach to the usual Mumford-Shah regularization is the edge penalty term. We split it into two parts, such that the length of the edge is penalized less if it coincides with the *a priori* edge and penalized more if it does not. Similar extensions of the Mumford-Shah functional have been studied in the context of fracture mechanics, where the material has a different hardness depending on the place and orientation, see Babadjian and Giacomini [2013]; Giacomini and Ponsiglione [2006]. In those works the aim is mainly to investigate the well-posedness of the problem.

Following the classic proof of De Giorgi et al. [1989] we show that our proposed functional admits a minimizer under reasonable assumptions on the operator and underlying image. We assume that the *a priori* edge is fixed and that the underlying image is pointwise bound. Furthermore we show that the approach is stable and yields a regularization for the image and its edges with a general *a priori* parameter choice rule and also the discrepancy principle. The regularization results are the main contribution of Chapter 3. To our knowledge these results give the broadest justification so far to use the non-convex Mumford-Shah regularization for ill-posed inverse problems. For example, the fact that the discrepancy principle yields a regularization for the standard Mumford-Shah regularization (that is $K^0 = \emptyset$) is new. Naturally, there are still open questions, for example whether the parameter choice rules also allow convergence rates under additional assumptions.

When applying Mumford-Shah type regularization to practical applications several issues arise. The primary difficulty comes from the edge part because it is not easy to represent in programming. In Chapter 4 we prove a Γ-convergence result for the special case $\gamma = 0$. For the case $\gamma \neq 0$ we motivate a heuristic penalty that allows numerical implementation. Our approximations are in the phase field setting of Ambrosio-Tortorelli, where edges are described by blurry indicator functions. First we describe the *a priori* edge information in terms of an edge indicator function and note our set of assumptions. We then first prove the Γ-convergence in one dimension and lift the result to dimension $N \geq 2$ by standard arguments in the theory of Γ-convergence. Although the technique is well known, the recovery sequence needed some tedious computations. There are many Γ-convergence results for the Mumford-Shah functional or other free discontinuity problems, see Braides [2002]. To our knowledge this is the first extension of the phase field setting in this direction.

Finally, in Chapter 5 we evaluate our approach for the two inverse problems X-ray CT and DOT. Our numerical experiments indicate that our approach yields good reconstructions from incomplete data. A drawback of the model surely is the additional complexity and comparatively high number of parameters.

There are several avenues for future work. The edge penalty can be understood as being sparsity enforcing, in the sense that the discontinuity set has to be of zero Lebesgue measure. Although this kind of sparsity does not give rise to an efficient representation in a certain basis, it might be possible to exploit the sparsity in the computations, for example to choose the parameters, see Strehlow [2014]. From an applications point of view it is surely interesting to find a real application and see if the approach is feasible and the extra effort and complexity is justified. Looking at the resulting Algorithm 1 each of the two steps could be modified by itself. The second step can be understood as evaluating the current reconstruction with regards to some *a priori* expectation on the true image, in our case piece wise smoothness. Depending on this step, the regularization method in the first step is then adapted. Such kind of adaptive regularization methods with non constant regularization parameters have been studied before Alexandrov et al. [2010]; Gilboa et al. [2006]; Grasmair [2009]. Viewing each step by itself is more flexible, but on the other hand possibly does not have the mathematical justification as a Mumford-Shah type approach. Furthermore, other important image processing steps, such as image registration, could be incorporated into the model, see Droske et al. [2009].

Appendix A

Here we collect small proofs.

Lemma A 1. *Let $a, b \in \mathbb{R}$ with $b > a$ and define $h := b - a > 0$. Then there exist constants $C, c_0 > 0$, such that for all $c \leq c_0$ it holds*

$$\int_a^b |1 - x| - c \, dx \geq Ch. \tag{1}$$

Proof. Because of symmetry, the integral is smallest for $a = 1 - h/2$ and $b = 1 + h/2$. We can compute

$$\begin{aligned}\int_a^b |1 - x| - c \, dx &\geq 2 \int_{1-h/2}^{1} 1 - c - x dx \\ &= (1 - c)h - \big(1 - (1 - \frac{h}{2})^2\big) \\ &= \frac{h^2}{4} - ch.\end{aligned}$$

Now we can choose C and c_0 such that $0 < c_0 < \frac{h}{4} - C$. We then obtain because of $c \leq c_0 < \frac{h}{4} - C$

$$\int_a^b |1 - x| - c dx \geq \frac{h^2}{4} - ch > Ch.$$

∎

Lemma A 2. *Let $\{a_n\}$, $\{b_n\}$ be real sequences and $C > 1$. If $\liminf_n a_n \geq 0$ then*

$$\liminf_n (a_n + b_n) \leq \liminf_n (Ca_n + b_n). \tag{2}$$

Proof. Assume the claim is false. Then we have $\liminf_n (a_n + b_n) < \infty$ and there exists $\varepsilon > 0$ so that

$$\liminf_n (a_n + b_n) = \liminf_n (Ca_n + b_n) + \varepsilon.$$

Then for a sufficiently large index n_0 we can find a subsequence $\{m\} \in \mathbb{N}, m \geq n_0$ so that for each m

$$Ca_m + b_m + \frac{\varepsilon}{2} < a_m + b_m$$

which yields

$$a_m < -\frac{\varepsilon}{2(C-1)}.$$

This contradicts the assumptions $\liminf_n a_n \geq 0$. ∎

Lemma A 3. *Under the notations and assumptions of Lemma 4.10 we set*

$$\tilde{\phi}_n(t) = \frac{1}{c_n - a_n}\big((t - a_n)\hat{\phi}_n(t) + (c_n - t)\phi_n(t)\big). \tag{3}$$

It then is

$$\lim_n \overline{\Psi}_{v_n^0,n}(f_n, v_n, \big((\overline{S_f} \cap K^0)_{a_n} \cap (\overline{S_f} \setminus K^0)_{c_n}\big) \setminus (\overline{S_f} \setminus K^0)_{a_n}) = 0. \tag{4}$$

Note that from the proof of Lemma 4.8 we know that

$$\limsup_n \int_0^{c_n} \frac{1}{4\varepsilon_n}\big(1-\phi_n(t)\big)^2 + \varepsilon_n|(\phi_n(t)'|^2 dt < \infty$$

and

$$\limsup_n \int_0^{c_n} \frac{1}{4\varepsilon_n}\big(1-\hat{\phi}_n(t)\big)^2 + \varepsilon_n|(\hat{\phi}_n(t)'|^2 dt < \infty.$$

Proof. We first note that

$$\begin{aligned} A_n :=& \big((\overline{S_f} \cap K^0)_{a_n} \cap (\overline{S_f} \setminus K^0)_{c_n}\big) \setminus (\overline{S_f} \setminus K^0)_{a_n} \\ &\subset (K^0)_{a_n} \cap (\overline{S_f} \setminus K^0)_{c_n} \\ &\subset \big(\overline{S_f} \setminus K^0 \cap (K^0)_{a_n+c_n}\big)_{c_n}. \end{aligned}$$

As $\mathcal{H}^{N-1}(\overline{S_f} \setminus K^0 \cap (K^0)_{a_n+c_n}) \to 0$ as $n \to \infty$ the integral $\int_{A_n} v_n^2 |\nabla f_n|^2 \to 0$. By Lemma 4.11 it then suffices to show that

$$\limsup_n \int_0^{c_n} \frac{1}{4\varepsilon_n}\big(1-\tilde{\phi}_n(t)\big)^2 + \varepsilon_n|(\tilde{\phi}_n(t)'|^2 dt < \infty \tag{5}$$

to obtain the claim.

We will write out the integral step by step. First we look at

$$\begin{aligned}
\left(1-\tilde{\phi}_n(t)\right)^2 =& \frac{1}{(c_n-a_n)^2}\left(c_n-a_n-(t-a_n)\hat{\phi}_n(t)+(c_n-t)\phi_n(t)\right)^2 \\
=& \frac{1}{(c_n-a_n)^2}\left(2c_n+c_n(\phi_n(t)-1)+a_n(\hat{\phi}_n(t)-1)-t(\phi_n(t)+\hat{\phi}_n(t))\right)^2 \\
=& \frac{1}{(c_n-a_n)^2}\Big[\\
& c_n^2(1-\phi_n(t))^2+a_n^2(1-\hat{\phi}_n(t))^2 \\
& +(2c_n-t(\phi_n(t)+\hat{\phi}_n(t)))^2 \\
& +c_n(1-\phi_n(t))a_n(1-\hat{\phi}_n(t)) \\
& +c_n(\phi_n(t)-1)(2c_n-t(\phi_n(t)+\hat{\phi}_n(t))) \\
& +a_n(\hat{\phi}_n(t)-1)(2c_n-t(\phi_n(t)+\hat{\phi}_n(t))) \\
& \Big].
\end{aligned}$$

First note that $\limsup_n \frac{a_n}{c_n-a_n}$ and $\limsup_n \frac{c_n}{c_n-a_n}$ are bound. We already calculated the integrals over the terms $(1-\phi_n(t))^2$ and $(1-\hat{\phi}_n(t))^2$ in Section 4.3 and therefore know that the $\limsup_n$ is bound for those two parts. Then because for $t\in[0,c_n]$ it is $0\le t(\phi_n(t)+\hat{\phi}_n(t))\le 2c_n$ we have

$$\begin{aligned}
&\int_0^{c_n}\frac{1}{4\varepsilon_n}\frac{1}{(c_n-a_n)^2}(2c_n-t(\phi_n(t)+\hat{\phi}_n(t)))^2dx \\
&\le \int_0^{c_n}\frac{1}{\varepsilon_n}\frac{c_n^2}{(c_n-a_n)^2}dx
\end{aligned}$$

which is also bound as $n\to\infty$. The remaining three terms are products of terms for which we know that they are square integrable and bound for $n\to\infty$. For example using the Cauchy-Schwarz inequality we can estimate the 4-th term as

$$\int_0^{c_n} \frac{1}{4\varepsilon_n} c_n(1-\phi_n(t)) a_n (1-\hat{\phi}_n(t)) dx$$
$$\leq \int_0^{c_n} \frac{1}{4\varepsilon_n} c_n^2 (1-\phi_n(t))^2 dx \cdot \int_0^{c_n} \frac{1}{4\varepsilon_n} a_n^2 (1-\hat{\phi}_n(t))^2 dx.$$

As a result the integral $\int_0^{c_n} \frac{1}{4\varepsilon_n} \big(1-\tilde{\phi}_n(t)\big)^2 dx$ is bound for $n \to \infty$.

Now we turn to the integral over the gradient. We use the definition and sort the terms again

$$\begin{aligned}
|(\tilde{\phi}_n(t)'|^2 =& \frac{1}{(c_n-a_n)^2} \big(a_n \hat{\phi}_n'(t) - c_n \phi_n'(t) + (\phi_n(t) - \hat{\phi}_n(t)) + t(\phi_n'(t) - \hat{\phi}_n'(t))\big)^2 \\
\leq & \frac{1}{(c_n-a_n)^2} \big[\qquad (6) \\
& |a_n \hat{\phi}_n'(t)|^2 + |c_n \phi_n'(t)|^2 \\
& + \big| a_n \hat{\phi}_n'(t) c_n \phi_n'(t) \big| \\
& + \big| (\phi_n(t) - \hat{\phi}_n(t)) + t(\phi_n'(t) - \hat{\phi}_n'(t)) \big|^2 \\
& + \big| a_n \hat{\phi}_n'(t) \; \big((\phi_n(t) - \hat{\phi}_n(t)) + t(\phi_n'(t) - \hat{\phi}_n'(t))\big) \big| \\
& + \big| c_n \phi_n'(t) \; \big((\phi_n(t) - \hat{\phi}_n(t)) + t(\phi_n'(t) - \hat{\phi}_n'(t))\big) \big| \\
& \big].
\end{aligned}$$

We already calculated the integrals over the terms $|\phi_n'(t)|^2$ and $|\hat{\phi}_n'(t)|^2$ in Section 4.3 and therefore know that the $\limsup_n$ of these integrals are bound. We then look at

$$
\begin{aligned}
&\big((\phi_n(t) + \hat{\phi}_n(t)) + t(\phi_n'(t) + \hat{\phi}_n'(t))\big)^2 \\
&\leq \big(\phi_n(t) + \hat{\phi}_n(t)\big)^2 + \big|2\big(\phi_n(t) + \hat{\phi}_n(t)\big)t\big(\phi_n'(t) + \hat{\phi}_n'(t)\big)\big| + t^2\big(\phi_n'(t) + \hat{\phi}_n'(t)\big)^2
\end{aligned} \tag{7}
$$

We can estimate the last term by

$$t^2(\phi_n'(t) + \hat{\phi}_n'(t))^2 \leq c_n^2(\phi_n'(t) + \hat{\phi}_n'(t))^2$$

and as the integrals over $\phi_n'(t)^2$ and $\hat{\phi}_n'(t)^2$ are bound for $n \to \infty$ so is this. We can also estimate

$$\int_0^{c_n} \varepsilon_n \frac{1}{(c_n - a_n)^2}(\phi_n(t) + \hat{\phi}_n(t))^2 dt \leq 2\varepsilon_n c_n \frac{1}{(c_n - a_n)^2}$$

which by definition of c_n and a_n is also bound. The second term of (7) then is bound by Cauchy-Schwarz inequality. The remaining three terms of the gradient part (6) can all be bounded as above by the Cauchy-Schwarz inequality. □

References

T. Alexandrov, M. Becker, S. Deininger, G. Ernst, L. Wehder, M. Grasmair, F. von Eggeling, H. Thiele, and P. Maass. Spatial segmentation of imaging mass spectrometry data with edge-preserving image denoising and clustering. *Journal of proteome research*, 9(12):6535–6546, 2010. 9, 122

H. W. Alt. *Lineare Funktionalanalysis*, volume 2. Springer Berlin et al., 1985. 17

M. Amar, V. De Cicco, and N. Fusco. Lower semicontinuity results for free discontinuity energies. *Mathematical Models and Methods in Applied Sciences*, 20(05):707–730, 2010. 8, 24

L. Ambrosio. A compactness theorem for a new class of functions of bounded variation. *Unione Matematica Italiana. Bollettino. B. Serie VII*, 3(4):857–881, 1989. 23

L. Ambrosio and V. M. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Communications on Pure and Applied Mathematics*, 43(8):999–1036, 1990. 87, 96, 97

L. Ambrosio and V. M. Tortorelli. On the approximation of free discontinuity problems. *Unione Matematica Italiana. Bollettino. B. Serie VII*, 6(1):105–123, 1992. 6, 7, 9, 64, 66

L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford University Press, USA, 2000. ISBN 0198502451. 3, 13, 14, 15, 16, 18, 19, 20, 21, 30, 39, 40, 41

S. W. Anzengruber and R. Ramlau. Morozov's discrepancy principle for Tikhonov-type functionals with nonlinear operators. *Inverse Problems*, 26(2): 025001, 2010. 9, 49, 50, 56

S. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15(2): R41, 1999. 111

S. Arridge and J. Schotland. Optical tomography: forward and inverse problems. *Inverse Problems*, 25(12):123010, 2009. 111

H. Attouch, G. Buttazzo, and G. Michaille. *Variational analysis in Sobolev and BV spaces: applications to PDEs and optimization*. Society for Industrial Mathematics, 2006. ISBN 0898716004. 15, 23, 25, 26, 87, 92, 96, 97

J. Babadjian and A. Giacomini. Existence of strong solutions for quasi-static evolution in brittle fracture. Accepted in Annali della Scuola Normale Superiore di Pisa, Classe di Scienze., March 2013. URL `https://hal.archives-ouvertes.fr/hal-00797008`. 8, 41, 120

L. Bar, N. Sochen, and N. Kiryati. Semi-blind image restoration via Mumford-Shah regularization. *Image Processing, IEEE Transactions on*, 15(2):483–493, 2006. 4

M. Bertero and P. Boccacci. *Introduction to inverse problems in imaging*. IOP Publishing, Bristol, 1998. 42

B. Bourdin and A. Chambolle. Implementation of an adaptive finite-element approximation of the Mumford-Shah functional. *Numerische Mathematik*, 85 (4):609–646, 2000. 9

A. Braides. *Approximation of free-discontinuity problems*. Springer Verlag, 1998. ISBN 3540647716. 87

A. Braides. *Gamma-convergence for Beginners*. Oxford University Press, USA, 2002. ISBN 0198507844. 7, 9, 27, 28, 29, 66, 87, 121

A. Braides and G. Dal Maso. Non-local approximation of the Mumford-Shah functional. *Calculus of Variations and Partial Differential Equations*, 5(4): 293–322, 1997. 9

J. T. Bushberg and J. M. Boone. *The essential physics of medical imaging*. Lippincott Williams & Wilkins, 2011. 2

A. Chambolle. Image segmentation by variational methods: Mumford and Shah functional and the discrete approximations. *SIAM Journal on Applied Mathematics*, 55(3):827–863, 1995. 9

A. Chambolle. Inverse problems in image processing and image segmentation: some mathematical and numerical aspects. International Centre for Theoretical Physics, Trieste, Italy, 2000. URL `http://users.ictp.it/~pub_off/lectures/lns002/Chambolle/Chambolle.ps.gz`. 26

T. F. Chan and L. A. Vese. A level set algorithm for minimizing the Mumford-Shah functional in image processing. In *Variational and Level Set Methods in Computer Vision, 2001. Proceedings. IEEE Workshop on*, pages 161–168. IEEE, 2001. 6

S. R. Cherry. Multimodality in vivo imaging systems: twice the power or double the trouble? *Annu. Rev. Biomed. Eng.*, 8:35–62, 2006. 2

G. Dal Maso, G. A. Francfort, and R. Toader. Quasistatic crack growth in nonlinear elasticity. *Archive for rational mechanics and analysis*, 176(2):165–225, 2005. 8, 29, 30

I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on pure and applied mathematics*, 57(11):1413–1457, 2004. 13

G. David. *Singular sets of minimizers for the Mumford-Shah functional*. Birkhäuser, 2005. ISBN 376437182X. 3, 34, 38

E. De Giorgi, M. Carriero, and A. Leaci. Existence theorem for a minimum problem with free discontinuity set. *Archive for Rational Mechanics and Analysis*, 108(4):195–218, 1989. ISSN 0003-9527. 22, 34, 36, 120

M. Droske, W. Ring, and M. Rumpf. Mumford-Shah based registration: a comparison of a level set and a phase field approach. *Computing and visualization in science*, 12(3):101–114, 2009. 7, 122

H. Egger. On the convergence of modified Landweber iteration for nonlinear inverse problems. *Johann Radon Inst. Computat. Appl. Math., Tech. Rep. SFB-Rep. SFB-2010-017*, 2010. 116

H. Egger and M. Schlottbom. Analysis and regularization of problems in diffuse optical tomography. *SIAM J. Math. Anal.*, 42(5):1934–1948, 2010. 112, 113

M. Ehrhardt, K. Thielmans, L. Pizarro, D. Atkinson, S. Ourselin, B. Hutton, and S. Arridge. Joint reconstruction of PET-MRI by exploiting structural similarity. *Inverse Problems*, 2014. 1, 2

H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375. Springer, 1996. 2, 11

S. Esedoglu and J. Shen. Digital inpainting based on the Mumford-Shah-Euler image model. *European Journal of Applied Mathematics*, 13(4):353–370, 2002. 4

M. Fornasier, R. March, and F. Solombrino. Existence of minimizers of the Mumford-Shah functional with singular operators and unbounded data. *Annali di Matematica Pura ed Applicata*, pages 1–31, 2011. 5, 8, 35

N. Fusco. An overview of the Mumford-Shah problem. *Milan Journal of Mathematics*, 71(1):95–119, 2003. ISSN 1424-9286. 3, 35, 39

S. Geman and D. Geman. Stochastic relaxation. *Gibbs distributions, and the Bayesian*, 1984. 3, 6

A. Giacomini and M. Ponsiglione. A Γ-convergence approach to stability of unilateral minimality properties in fracture mechanics and applications. *Archive for rational mechanics and analysis*, 180(3):399–447, 2006. 8, 120

G. Gilboa, N. Sochen, and Y. Y. Zeevi. Variational denoising of partly textured images by spatially varying constraints. *Image Processing, IEEE Transactions on*, 15(8):2281–2289, 2006. 9, 122

M. Grasmair. Locally adaptive total variation regularization. In *Scale Space and Variational methods in computer Vision*, pages 331–342. Springer, 2009. 9, 122

J. Hadamard. *Lectures on Cauchy's problem in linear partial differential equations*. Yale university press, 1923. 11

M. Hanke, A. Neubauer, and O. Scherzer. A convergence analysis of the landweber iteration for nonlinear ill-posed problems. *Numerische Mathematik*, 72(1):21–37, 1995. 100, 116

A. Hore and D. Ziou. Image quality metrics: PSNR vs. SSIM. pages 2366–2369, 2010. 101

X. Intes. *Translational multimodality optical imaging*. Artech House, 2008. 2

M. Jiang, P. Maass, and T. Page. Regularizing properties of the Mumford-Shah functional for imaging applications. *Inverse Problems*, 30(3):035007, 2014. 5, 9, 35

D. Kazantsev, S. Ourselin, B. F. Hutton, K. J. Dobson, A. P. Kaestner, W. Lionheart, P. J. Withers, P. D. Lee, and S. R. Arridge. A novel technique to incorporate structural prior information into multi-modal tomographic reconstruction. *Inverse Problems*, 30(6):065004, 2014. 1, 2

E. Klann. A Mumford–Shah-like method for limited data tomography with an application to electron tomography. *SIAM Journal on Imaging Sciences*, 4(4): 1029 – 1048, 2011. 4

E. Klann and R. Ramlau. Regularization properties of Mumford-Shah-type functionals with perimeter and norm constraints for linear ill-posed problems. *SIAM Journal on Imaging Sciences*, 6(1):413 – 436, 2013. 4, 5, 9

R. Leahy and X. Yan. Incorporation of anatomical MR data for improved functional imaging with PET. In *Information Processing in Medical Imaging*, pages 105–120. Springer, 1991. 1, 2, 6, 8

J. L. Lewis. Capacitary functions in convex rings. *Archive for Rational Mechanics and Analysis*, 66(3):201–224, 1977. 68

A. K. Louis. *Inverse und schlecht gestellte Probleme*. Teubner, 1989. 11, 102, 104

D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math*, 42(5):577–685, 1989. 3, 34

F. Natterer. *The mathematics of computerized tomography*, volume 32. Society for Industrial Mathematics, 2001. 102, 104, 105

T. Page. Simultaneous reconstruction and segmentation with the Mumford-Shah functional for X-ray tomography. Diplomarbeit, University of Bremen, 2011. 105

R. Ramlau and W. Ring. A Mumford-Shah level-set approach for the inversion and segmentation of X-ray tomography data. *Journal of Computational Physics*, 221(2):539–557, 2007. ISSN 0021-9991. 4

R. Ramlau and W. Ring. Regularization of ill-posed Mumford–Shah models with perimeter penalization. *Inverse Problems*, 26(11):115001, 2010. 5, 9

A. Rieder. *Keine Probleme mit Inversen Problemen: Eine Einführung in ihre stabile Lösung*. Vieweg+ Teubner, 2003. 11

L. Rondi. A variational approach to the reconstruction of cracks by boundary measurements. *Journal de Mathématiques Pures et Appliqués*, 87(3):324–342, 2007. 5

L. Rondi. Reconstruction in the inverse crack problem by variational methods. *European Journal of Applied Mathematics*, 19(6):635–660, 2008a. 5

L. Rondi. On the regularization of the inverse conductivity problem with discontinuous conductivities. *Inverse Probl. Imaging*, 2(3):397–409, 2008b. 5, 9, 35, 56

L. Rondi and F. Santosa. Enhanced electrical impedance tomography via the Mumford-Shah functional. *ESAIM: Control, Optimisation and Calculus of Variations*, 6(1):517–538, 2001. 4, 5, 35, 113

L. T. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992. 12, 100, 105

M. Schweiger and S. Arridge. The Toast++ software suite for forward and inverse modeling in optical tomography. *Journal of Biomedical Optics*, 19(4):040801, 2014. doi: 10.1117/1.JBO.19.4.040801. URL `http://dx.doi.org/10.1117/1.JBO.19.4.040801`. 100

M. Schweiger and S. R. Arridge. Optical tomographic reconstruction in a complex head model using a priori region boundary information. *Physics in Medicine and Biology*, 44(11):2703, 1999. 1, 2

S. Somayajula, E. Asma, and R. M. Leahy. PET image reconstruction using anatomical information through mutual information based priors. In *Nuclear Science Symposium Conference Record, 2005 IEEE*, volume 5, pages 2722–2726. IEEE, 2005. 2

R. Strehlow. *Regularization of the inverse medium problem: on nonstandard methods for sparse reconstruction.* PhD thesis, Bremen, 2014. URL `http://elib.suub.uni-bremen.de/peid=D00104187`. 121

A. Tikhonov. Solution of incorrectly formulated problems and the regularization method. In *Soviet Math. Dokl.*, volume 5, page 1035, 1963. 12

D. W. Townsend. Multimodality imaging of structure and function. *Physics in medicine and biology*, 53(4):R1, 2008. 1

M. Vauhkonen, D. Vadasz, P. A. Karjalainen, E. Somersalo, and J. P. Kaipio. Tikhonov regularization and prior information in electrical impedance tomography. *Medical Imaging, IEEE Transactions on*, 17(2):285–293, 1998. 1

X. Zhang, M. Burger, and S. Osher. A unified primal-dual algorithm framework based on bregman iteration. *Journal of Scientific Computing*, 46(1):20–46, 2011. 105

Index