Carnegie Mellon University

MELLON COLLEGE OF SCIENCE

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF Doctor of Philosophy

TITLE Variational and Partial Differential Equation Methods for Image Processing

PRESENTED BY Pan Liu

ACCEPTED BY THE DEPARTMENT OF Mathematical Sciences

Irene Fonseca	May 2017
MAJOR PROFESSOR	DATE
Thomas Bohman	May 2017
DEPARTMENT HEAD	DATE

APPROVED BY THE COLLEGE COUNCIL

Rebecca W. Doerge	May 2017
DEAN	DATE

VARIATIONAL AND PDE METHODS FOR IMAGE PROCESSING

PAN LIU

Center for Nonlinear Analysis, Department of Mathematics, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA, 15213

May 2, 2017

Abstract. In the first part (Chapter 2) of this thesis, a new fractional order seminorm, TGV^r, $r \in \mathbb{R}$, $r \geq 1$, is proposed in the one-dimensional setting, as a generalization of the standard TGV^k-seminorms, $k \in \mathbb{N}$. The fractional TGV^r-seminorms are shown to be intermediate between the standard TGV^k-seminorms of integer order. A bilevel training scheme is proposed, where under a box constraint a simultaneous optimization with respect to the parameter α and the order r is performed.

In the second part (Chapter 3) of this thesis, the Ambrosio-Tortorelli approximation scheme with weighted underlying metric is investigated. It is shown that it Γ-converges to a Mumford-Shah image segmentation functional depending on the weight $\omega\, dx$, where $\omega \in SBV(\Omega)$, and on its value ω^-.

Some new ideas about bilevel training scheme and future works are collected in Chapter 4-6.

Acknowledgements. I wish to express my deep gratitude to my advisor Professor Irene Fonseca for her guidance and encouragements. Professor Irene Fonseca has always been most supportive during my Ph.D. career and research along the way. Her enthusiasm for mathematics and dedication to research has significantly influenced my life.

Secondly, I am grateful to my other committee members Professor Noel Walkington, Professor Christopher Larsen, and Professor Hayden Schaeffer for their valuable comments and useful remarks.

I also want to thank the Department of Mathematical Sciences at Carnegie Mellon University for providing a friendly research environment. In particular, I want to thank Professor William Hrusa and Professor Giovanni Leoni for their many useful suggestions.

Special thanks to my parents for their unconditional support and love.

The research of author was partially funded by the National Science Foundation under Grant No. DMS - 1411646.

Contents

Chapter 1. Introduction 4
1.1. The image processing and bilevel training scheme 4
1.2. The fractional order total generalized variation 6
1.3. The weighted Ambrosio - Tortorelli approximation scheme 8
1.4. Some insights from finite resolution images 12
1.5. The comprehensive training scheme 15

Chapter 2. The fractional order total generalized variation 19
2.1. The theory of fractional sobolev spaces 19
2.2. The fractional order total generalized seminorm 21
2.3. The bilevel training scheme with respect to parameter and regularizer 30

Chapter 3. The weighted Ambrosio - Tortorelli approximation scheme 32
3.1. Definitions and preliminary results 32
3.2. The one dimensional case 36
3.3. The multi-dimensional case 48

Chapter 4. Some insights from finite resolution images 74
4.1. The finite resolution image and the unavoidable noise during acquisition 74
4.2. The total variation and some preliminary results 75
4.3. The quasi-convexity of bilevel training scheme 77
4.4. A direct search for a minimizer of error function 79
4.5. Another spatially dependent bilevel training scheme with respect to TV 85

Chapter 5. The comprehensive training scheme 93
5.1. Notations and basic assumptions 93
5.2. The regularizer training scheme 96
5.3. The comprehensive bilevel training scheme 100

Chapter 6. Work in progress and future projects 109
6.4. The arsenal of regularizer spaces 109
6.5. Γ-convergence with non-negative spatially dependent parameters 111

Appendix A. 112

Appendix B. 115
List of Symbols 134
References 135

Chapter 1. Introduction

1.1. **The image processing and bilevel training scheme.** The mathematical treatment of image processing is strongly hinged on variational methods, partial differential equation (PDE), and machine learning. Variational methods provide model-based approaches which are mathematically rigorous, yield stable solutions and error estimates. However, the underlying techniques have shortcomings in the adaptation to real data. Although machine learning provides data-based reconstruction approaches which are best fitted to the given data, it neither guarantees the reconstruction results as the variational method does, nor it offers insights into the structural properties of the image model. Hence, an unified approach that combines the advantages of a variational model with the data-based approach is needed, and many contributions toward this goal have been presented in recent articles (see [35, 54]). In particular, the bilevel training scheme is one of the most popular.

We start with a brief historical summary of the state of the art of the bilevel training scheme for model training. In machine learning, the bilevel training scheme is defined as a semi-supervised training scheme that optimally adapts itself to the given "perfect data". For example, in [23, 24, 37, 38, 73, 74]) authors consider the bilevel training scheme in the study of finite dimensional Markov random field models. In inverse problems, for instance in [49, 48], authors discussed the optimal inversion and experimental acquisition in the context of optimal model design. Recently, the bilevel training scheme framed, in the context of functional variational regularization models, has also entered the image processing community (see [35]), and our work will start from here.

The variational formulation of problems in image processing often has an underlying functional

$$\mathcal{I}(u) := \mathcal{F}\left(u - u_\eta, \Omega\right) + \mathcal{R}(u, \alpha, \Omega),$$

where u_η is a given corrupted (noised) data, $\Omega = I = (0,1)$ represents the domain of voice signal and $\Omega = Q := (0,1)\times(0,1)$ stands for the domain of a square image, $\mathcal{F}$ is the *fidelity term*, $\alpha > 0$ is the *tuning parameter*, and $\mathcal{R}$ is the *regularizer*.

Image denoising is a fundamental task in image processing, as it is always a necessary step prior to higher level image processing problems such as reconstruction and segmentation. For a fixed regularizer $\mathcal{R}$, the image denoising problem aims at computing an image

$$u_\alpha \in \arg\min\left\{ \|u - u_\eta\|^2_{L^2(Q)} + \mathcal{R}(u, \alpha, Q) : \ u \in X_{\mathcal{R}} \right\}, \tag{1.1}$$

where $\alpha := (\alpha_1, \alpha_2, \ldots, \alpha_N)$, the *tuning parameter*, is a given parameter with $\alpha_i > 0$ for all $1 \le i \le N$, and $N \in \mathbb{N}$ depends on $\mathcal{R}$. An example is given by the *ROF* model ([68]), in which the regularizer $\mathcal{R}(u, \alpha, Q) := \alpha \, |u|_{TV(Q)}$, where $|\cdot|_{TV(Q)}$ is the total variation, and the tuning parameter $\alpha \in \mathbb{R}^+$. That is, we are considering the following minimizing problem:

$$u_\alpha := \arg\min \left\{ \|u - u_\eta\|^2_{L^2(Q)} + \alpha \, |u|_{TV(Q)} : \; u \in BV(Q) \right\} \tag{1.2}$$

The quality of a reconstructed image u_α, generated in (1.1) or (1.2), highly depends α: choosing it too large may result in losing important fine details, and if it is too small then it may keep noise un-removed. Hence, the choice of the "optimal" tuning parameter α becomes an important task. In [35] the authors proposed a training scheme by using bilevel optimization. To be precise, we assume that we can decompose the corrupted image $u_\eta = u_c + \eta$ where u_c represents a noise-free clean image (the perfect data), and η encodes noise. A typical bilevel training scheme can be formulated as follows, using the ROF model as an example:

Level 1.

$$\tilde{\alpha} \in \arg\min \left\{ \|u_\alpha - u_c\|^2_{L^2} : \; \alpha > 0 \right\}, \tag{1.3}$$

Level 2.

$$u_\alpha := \arg\min \left\{ \|u - u_\eta\|^2_{L^2(Q)} + \alpha \, |u|_{TV(Q)} : \; u \in BV(Q) \right\}. \tag{1.4}$$

Roughly speaking, this training scheme searches $\alpha > 0$ such that the recovered image u_α, obtained from(1.4), best fits the given clean image u_c, measured in terms of the L^2-distance in (1.3). In [35] it has been proved that (1.3) admits a positive solution $\tilde{\alpha} > 0$ if $TV(u_\eta) > TV(u_c)$. Some other choices of regularizers have also been proposed in [35], for example, the second order total generalized variation TGV^2 which is defined as follows:

$$|u|_{TGV^2_{\alpha_0,\alpha_1}(I)} := \inf \left\{ \alpha_0 \left| u' - v_0 \right|_{\mathcal{M}_b(I)} + \alpha_1 \left| v_0' \right|_{\mathcal{M}_b(I)}, \; v_0 \in BV(I) \right\}, \tag{1.5}$$

and the corresponding bilevel training scheme becomes:

Level 1.

$$\tilde{\alpha} \in \arg\min \left\{ \|u_\alpha - u_c\|^2_{L^2} : \; \alpha = (\alpha_0, \alpha_1) > 0 \right\}, \tag{1.6}$$

Level 2.

$$u_{\alpha_0,\alpha_1} := \arg\min \left\{ \|u - u_\eta\|^2_{L^2(Q)} + |u|_{TGV^2_{\alpha_0,\alpha_1}(Q)} : \; u \in BV(Q) \right\}. \tag{1.7}$$

TGV^2, defined in (1.5), may yield geometric structures as compared with TV, therefore, we should expect that

$$\inf \left\{ \|u_\alpha - u_c\|^2_{L^2} : \; \alpha > 0 \right\} \neq \inf \left\{ \|u_{\alpha_0,\alpha_1} - u_c\|^2_{L^2} : \; \alpha_0 > 0, \alpha_1 > 0 \right\} \tag{1.8}$$

where u_α and u_{α_0,α_1} are defined in (1.4) and (1.7), respectively. It has been observed that for image data u_c, if u_c has large flat areas, then TV performs better than TGV^2. That is, the quantity on the left hand of (1.8) smaller than the quantity on the right hand side. However, if u_c has many smooth transitions and fine details, then TGV^2 performs better. Other situations are also observed in the case one dimension signal. In addition, the higher order seminorms TGV^k, $k \geq 2$, have rarely been analyzed, and hence their performance is largely unknown.

1.2. **The fractional order total generalized variation.** In the existing literature a regularizer is fixed a priori, and the biggest effort is concentrated on studying how to identify the best parameters. In the case of the TGV^k model, this amounts to set manually the value of k first, and then to determine the optimal α in (1.3) or (1.6). However, as we said before, there is no evidence suggesting that TGV^2 will always perform better than TV (or conversely). The main focus of Chapter 2 in this thesis is exactly to investigate how to optimally tune both the tuning parameter α and the order k of the TGV^k_α-seminorm, in order to achieve the best reconstructed image.

In Chapter 2 of this thesis we work in one dimension and generalize the bilevel training scheme introduced before so that it can not only do parameter training, but also it can determine the optimal order k of the regularizer TGV^k for image reconstruction. A straightforward modification of (1.6) would be to just insert the order of the regularizer inside the learning level 2 in (1.7). Namely,

Level 1.

$$(\tilde{\alpha}, \tilde{k}) \in \arg\min \left\{ \|u_{\alpha,k} - u_c\|^2_{L^2(I)} :\ \alpha > 0, k \in \mathbb{N} \right\},$$

Level 2.

$$u_{\alpha,k} := \arg\min \left\{ \|u - u_0\|^2_{L^2(I)} + |u|_{TGV^{k+1}_\alpha(I)} :\ u \in BV(I) \right\}.$$

Often, in order to show the existence of a solution of the training scheme and also for the numerical realization of the model, a *box constraint* is imposed (see, e.g. [12, 33]), i.e.,

$$(\alpha, k) \in [\mathrm{A}, 1/A]^{k+1} \times [1, \mathrm{R}], \tag{1.9}$$

where $0 < A < 1$ (called the index of box constraint) and $\mathrm{R} > 1$ are fixed real numbers. However, such constraint makes the above training scheme less interesting. To be precise, restricting the analysis to the case in which $k \in \mathbb{N}$ is an integer, the box constraint (1.9) would only allow k to take finitely many values, and hence the optimized order $\tilde{k}$ of regularizer would simply be determined by performing scheme (1.6) finitely many times, at each time with different values of k. In addition, finer texture effects, for which an "intermediate" reconstruction between the one provided by TGV^k and TGV^{k+1} for some $k \in \mathbb{N}$ would be needed, might be neglected in the optimization procedure.

Therefore, a main challenge in the setup of such a training scheme is to introduce a meaningful interpolation between the spaces TGV^{k+1} and TGV^k, to guarantee that the family of such spaces exhibits certain compactness and lower semicontinuity properties. For this purpose, we modify the definition of the TGV^k functionals by incorporating the theory of fractional Sobolev spaces, and we introduce the notion of fractional order TGV^{k+s} space (see Definition 2.11), where $k \in \mathbb{N}$ and $0 < s < 1$. For $k = 1$, our definition reads as follows:

$$|u|_{TGV^{1+s}(I)} := \inf \left\{ \left|u' - s v_0\right|_{\mathcal{M}_b(I)} + s(1-s) \left|v_0\right|_{W^{s,1+s(1-s)}(I)} : \right.$$
$$\left. v_0 \in W^{s,1+s(1-s)}(I),\ \int_I v_0(x)\,dx = 0 \right\},$$

and similarly, the TGV_α^{1+s} reads as follows, where $\alpha = (\alpha_0, \alpha_1) \in \mathbb{R}^+ \times \mathbb{R}^+$:

$$|u|_{TGV_\alpha^{1+s}(I)} := \inf\Big\{\alpha_0 \left|u' - sv_0\right|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s)\left|v_0\right|_{W^{s,1+s(1-s)}(I)} : \\ v_0 \in W^{s,1+s(1-s)}(I),\ \int_I v_0(x)\,dx = 0\Big\}.$$

In addition, for every $k \in \mathbb{N}$ and $s \in [0,1]$ we introduce the classes of functions with bounded infimal-convolution total variation seminorm

$$BGV^{k+s}(I) := \Big\{u \in L^1(I) :\ |u|_{TGV^{k+s}(I)} < +\infty\Big\}.$$

In the expressions above, $W^{s,1+s(1-s)}(I)$ is the *fractional Sobolev space* of order s and integrability $1+s(1-s)$. In Theorem 2.12 we show that the TGV^{1+s} seminorm is indeed intermediate between $TV(TGV^1)$ and TGV^2, i.e., we prove that,

$$\lim_{s\nearrow 1} |u|_{TGV_\alpha^{1+s}(I)} \geq |u|_{TGV_\alpha^2(I)} \text{ and } \lim_{s\searrow 0} |u|_{TGV_\alpha^{1+s}(I)} = |u|_{TV}.$$

Namely, for $s \nearrow 1$, the behavior of the TGV_α^{1+s}-seminorm is close to the one of the standard TGV_α^2-seminorm, whereas for $s \searrow 0$ it approaches the TV functional. We additionally prove (see Corollary 2.15) that analogous results hold for higher order TGV_α^{k+s}-seminorms.

The advantage in working with such interpolation spaces is twofold. First, TGV^{k+s} is expected to inherit the advantages of fractional order derivatives, which have been shown to reduce the staircasing and contrast effects in noise removal problems (see, e.g. [22]). Second, they allow us to introduce the following improved training scheme, which, under (1.9), simultaneously optimizes both with respect to the parameter α and to the order r of derivation.

Level 1.

$$(\tilde{\alpha}, \tilde{r}) \in \arg\min\Big\{\|u_{\alpha,r} - u_c\|^2_{L^2(I)},\ (\alpha, r) \in [\mathrm{A}, 1/A]^{\lfloor r\rfloor+1} \times [1, 2-A]\Big\}, \tag{1.10}$$

Level 2.

$$u_{\alpha,r} := \arg\min\Big\{\|u - u_0\|^2_{L^2} + TGV_\alpha^r(u) :\ u \in BGV_\alpha^r(I)\Big\}.$$

In the definition above, $\lfloor r\rfloor$ denotes the largest integer smaller than or equal to r. Note that, according to the test noise-free image u_c, the level 1 in our training scheme (1.10) directly indicates the higher order regularizer providing the best image reconstruction, as well as the associated corresponding optimal parameters.

The construction of TGV^r in two dimensions is work in progress. We will not include it in this thesis but instead refer to the upcoming paper [31].

1.3. **The weighted Ambrosio - Tortorelli approximation scheme.** Another drawback of the training scheme in (1.3) is that it uses a constant tuning parameter which provides an uniform regularization strength over the entire domain Q. It has been observed in [25] that an uniform regularization strength is undesirable when both fine details and large flat areas are present in an image, which is often the case in image denoising problems. Ideally, we should try to instruct a weak regularization strength in fine details area so that those details can be preserved, and to instruct a strong regularization strength should be used over large flat areas so that the noise can be removed.

To this purpose, in Chapter 3 we propose a spatially dependent training scheme with respect to the tuning parameter α. To write this precise, we introduce the following notation:

Notation 1.1. Recall that

$$0 < A << 1 \tag{1.11}$$

is a fixed constant.

1. For any $N \in \mathbb{N}$, we set

$$Q_N\left(i_N, j_N\right) := ((i_N - 1)/N, i_N/N) \times ((j_N - 1)/N, j_N/N)$$

for each $1 \leq i_N, j_N \leq N$,

$$\mathcal{Q}_N := \left\{Q_N(i_N, j_N),,\ 1 \leq i_N, j_N \leq N\right\},$$

and

$$\mathcal{Q}_A := \left\{L \in \mathcal{Q}_N,\ N \leq 1/A\right\}. \tag{1.12}$$

2. $\mathcal{L}$ denotes a collection of finitely many $L \in \mathcal{Q}_A$ such that

$$\mathcal{L} := \left\{L \in \mathcal{Q}_A :\ L \text{ are mutually disjoint},\ Q \subset \bigcup \overline{L}\right\}, \tag{1.13}$$

3. $\mathcal{V}_A$, called the *training ground*, is the collection of all possible $\mathcal{L}$. Note that we have

$$\#\left\{\mathcal{V}_A\right\} < +\infty.$$

We propose a new bilevel training scheme which uses the scheme (1.3) in each subdomain of Q, and searches for the best combination of different subdomains from which a recovered image $\tilde{u}$, which best fits u_c in certain sense, might be obtained. Moreover, in order to explore more regularizers, in Chapter 3 we use the Mumford-Shah image segmentation functional as the regularizer. The Mumford-Shah image segmentation functional is given by

$$\begin{gathered} MS(u) := \alpha \int_\Omega |\nabla u|^2 dx + \alpha \mathcal{H}^{N-1}(S_u) \\ \text{where } u \in SBV(\Omega),\ S_u \text{ stands for the jump set of } u, \end{gathered} \tag{1.14}$$

and was introduced in [67]. By minimizing the functional

$$\|u - u_\eta\|^2_{L^2(\Omega)} + MS(u, K)$$

one tries to find a "piecewise smooth" approximation of u_0. The existence of such minimizers can be proved by using compactness and lower semicontinuity theorems in $SBV(\Omega)$ (see [3]). Furthermore, regularity results in [32] assert that minimizers u satisfy $u \in C^1(\Omega \setminus \overline{S}_u)$ and $\mathcal{H}^{N-1}(\overline{S}_u \cap \Omega \setminus S_u) = 0$.

With the Mumford-Shah image segmentation functional as the regularizer, our new training scheme can be presented as follows:
Level 1.

$$\tilde{\mathcal{L}} \in \arg\min \left\{ \|u_c - u_{\mathcal{L}}\|^2_{L^2(Q)} :\ \mathcal{L} \in \mathcal{V}_A \right\} \tag{1.15}$$

Level 2.

$$\alpha_{\mathcal{L}}(x) := \alpha_L \text{ for } x \in L \in \mathcal{L}, \text{ and} \tag{1.16}$$
$$u_{\mathcal{L}} \in \arg\min \left\{ \int_Q \alpha_{\mathcal{L}} |\nabla u|^2 \, dx + \int_{S_u} \alpha_{\mathcal{L}}(x) d\mathcal{H}^{N-1} + \|u - u_\eta\|^2_{L^2(Q)} :\ u \in SBV(Q) \right\}$$

Level 3.

$$\begin{aligned} &\alpha_L \in \arg\min \left\{ \|u_\alpha - u_c\|^2_{L^2(L)} : \alpha \in [A, 1/A] \right\} \\ &u_\alpha \in \arg\min \left\{ \|u - u_\eta\|^2_{L^2(L)} + \alpha MS(u) :\ u \in SBV(L) \right\}. \end{aligned} \tag{1.17}$$

Scheme (1.15) allows us to perform the denoising procedure "pointwisely", and it is an improvement of the following training scheme
Level 1.

$$\tilde{\alpha} \in \arg\min \left\{ \|u_\alpha - u_c\|^2_{L^2} :\ \alpha > 0 \right\}, \tag{1.18}$$

Level 2.

$$u_\alpha \in \arg\min \left\{ \|u - u_\eta\|^2_{L^2(Q)} + \alpha MS(u, Q) :\ u \in SBV(Q) \right\},$$

which is the MS version of training scheme (1.3). Note that since $\{Q\} \in \mathcal{V}_A$, (1.15) must perform better than (1.18). We remark that the most important step is (1.16) for the following reasons:

1. (1.16) is the bridge connecting level 1 and level 2 in scheme (1.15);
2. since $\alpha_{\mathcal{L}}$ is defined by locally optimizing the parameter α_L, we expect $u_{\mathcal{L}}$ be "close" to u_c locally in L;
3. the last integrand in (1.16) keeps $u_{\mathcal{L}}$ close to u_η globally, hence we may expect $u_{\mathcal{L}}$ to have a good balance between local optimization and global optimization.

We may view (1.16) as a weighted version of (1.14) by changing the underlying metric from dx to $\alpha_{\mathcal{L}} dx$. By the construction of $\alpha_{\mathcal{L}}$ in (1.17), we know it is a piecewise constant function and, since $A > 0$ is positive, the discontinuity set of $\alpha_{\mathcal{L}}$ has finite $\mathcal{H}^{N-1}$ measure. However, $\alpha_{\mathcal{L}}$ is only defined $\mathcal{L}^N$-a.e., and hence the term

$$\int_{S_u} \alpha_{\mathcal{L}}(x) d\mathcal{H}^{N-1}$$

might be ill-defined.

In Chapter 3, we deal with the well-definedness of (1.16) by modifying $\alpha_{\mathcal{L}}$ accordingly, and by building a sequence of functionals which Γ-converges to (1.16). To be precise, we adopt the approximation scheme of Ambrosio and Tortorelli in [6] and change the underlying metric properly. In (1.14) Ambrosio and Tortorelli considered a sequence of functionals reminiscent of the Cahn-Hilliard approximation, and introduced a family of elliptic functionals

$$AT_\varepsilon(u,v) := \int_\Omega \alpha \left|\nabla u\right|^2 v^2 dx + \int_\Omega \alpha \left[\varepsilon \left|\nabla v\right|^2 + \frac{1}{4\varepsilon}(v-1)^2\right] dx + \int_\Omega (u-u_\eta)^2\, dx,$$

where $u \in W^{1,2}(\Omega)$, $(v-1) \in W_0^{1,2}(\Omega)$, and $u_\eta \in L^2(\Omega)$. The additional field v plays the role of controlling variable on the gradient of u. In [6] a rigorous argument has been made to show that $AT_\varepsilon \to MS$ in the sense of Γ-convergence ([7, 27]), where MS is defined in (1.14).

In view of (1.17), we fix a spatially dependent parameter, or a weight function, $\omega \in SBV(\Omega)$ such that

$$\omega \text{ is positive and bounded away from 0, and } \mathcal{H}^{N-1}(S_\omega) < +\infty. \tag{1.19}$$

Our new Mumford-Shah image segmentation functional with a spatially dependent parameter is defined as

$$MS_\omega(u) := \int_\Omega \left|\nabla u\right|^2 \omega\, dx + \int_{S_u} \omega^-\, d\mathcal{H}^{N-1},$$

and the Ambrosio - Tortorelli functionals with spatially dependent parameters are defined as

$$AT_{\omega,\varepsilon}(u,v) := \int_\Omega \left|\nabla u\right|^2 v^2 \omega\, dx + \int_\Omega \left[\varepsilon \left|\nabla v\right|^2 + \frac{1}{4\varepsilon}(v-1)^2\right] \omega\, dx.$$

(Note that $AT_{1,\varepsilon}(u,v)$ and $MS_1(u)$ are the Ambrosio-Tortorelli approximation scheme and Mumford-Shah functional studied in [6] with constant parameters, respectively). Moreover, since $A > 0$ is positive and $\alpha_L > A$ in (1.17) , it is not restricted to assume that

$$0 < A \le \operatorname{ess\,inf}\{\omega(x):\ x \in \Omega\} \le \operatorname{ess\,sup}\{\omega(x):\ x \in \Omega\} < +\infty, \tag{1.20}$$

where $A > 0$ is given in (1.11).

The main result of Chapter 3 is the following:

Theorem 1.2. *Let $\Omega \subset \mathbb{R}^N$ be open bounded with Lipschitz boundary, let $\omega \in SBV(\Omega)$ satisfy* (1.19) *and* (1.20), *and let $\mathcal{AT}_{\omega,\varepsilon}$: $L^1(\Omega) \times L^1(\Omega) \to [0,+\infty]$ be defined by*

$$\mathcal{AT}_{\omega,\varepsilon}(u,v) := \begin{cases} AT_{\omega,\varepsilon}(u,v) & \text{if } (u,v) \in W^{1,2}(\Omega) \times W^{1,2}(\Omega),\ 0 \le v \le 1 \text{ a.e.,} \\ +\infty & \text{otherwise.} \end{cases}$$

Then the functionals $\mathcal{AT}_{\omega,\varepsilon}$ Γ-converge, with respect to the $L^1 \times L^1$ topology, to the functional

$$\mathcal{MS}_\omega(u,v) := \begin{cases} MS_\omega(u) & \text{if } u \in GSBV(\Omega) \text{ and } v = 1 \text{ a.e.,} \\ +\infty & \text{otherwise.} \end{cases}$$

A direct inspection of the proof of Theorem 1.2 allows to also consider the case in which $\omega \in C(\Omega)$ and satisfies (1.20). To be precise:

Theorem 1.3. *Let $\Omega \subset \mathbb{R}^N$ be open bounded with Lipschitz boundary, let $\omega \in C(\overline{\Omega})$ satisfy (1.20). Then the functionals $\mathcal{AT}_{\omega,\varepsilon}$ Γ-converge, with respect to the $L^1 \times L^1$ topology, to the functional $\mathcal{MS}_\omega(u)$.*

We recall similar problems that have been studied for different types of weight functions ω (see, for example [9, 10, 43, 59, 76]). In particular, [9, 10] treated a uniformly continuous and strong A_∞ (defined in [70]) weight function on Modica-Mortola and Mumford-Shah-type functionals, respectively, and in [59] the authors considered a $C^{1,\beta}$-continuous weight function, with some other minor assumptions, in the one-dimensional Cahn-Hilliard model. Also, in [43] the author studies the family of energy functionals

$$\int_\Omega (v^2 + \eta_\varepsilon) f(x, u, \nabla u) dx + \int_\Omega \left[\frac{\varepsilon}{2} \varphi^2(\nabla v) + \frac{1}{2\varepsilon}(1-v)^2 \right] dx \tag{1.21}$$

where φ: $\mathbb{R}^N \to [0, +\infty)$ is a norm, and $(u, v) \in W^{1,2}(\Omega) \times W^{1,2}(\Omega)$. Note that in (1.21) $\varphi^2(\nabla v)$ is anisotropic and penalizing differently different orientations of the gradient, but spatially homogeneous, so it does not include the case where that term is replaced by $\omega(x) |\nabla v|^2$. Moreover, [76] addresses the family of energies

$$\int_\Omega |\nabla u|^2 v^2 \, dx + \int_\Omega \left[\varepsilon \varphi(x, \nabla v) + \frac{1}{4\varepsilon}(v-1)^2 \right] dx.$$

where φ is required to be continuous, and we cannot set $\varphi(x, \nabla v) = \omega(x) |\nabla v|^2$ with $\omega \in SBV(\Omega)$ as in our context. Lastly, we point out that in both [43, 76] the underlying measure of integration is the Lebesgue measure $\mathcal{L}^N$, while in our model the underlying measure is $\omega(x)\mathcal{L}^N$, which affects all terms.

The proof of Theorem 1.2 consists of two steps. The first step is to prove the "lim inf inequality" $\liminf_{\varepsilon \to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) \geq MS_\omega(u)$ for every sequence $u_\varepsilon \to u$, $v_\varepsilon \to v$. This is obtained in Section 3.2 in the case $N = 1$ by using most of the arguments proposed in [6], and extended to the case $N > 1$ by using a particular slicing argument (see Lemma 3.23). The second step is the construction of a recovery sequence ($u_\varepsilon \to u, v_\varepsilon \to 1$) such that the term

$$\int_\Omega \left[\varepsilon |\nabla v|^2 + \frac{1}{4\varepsilon}(v-1)^2 \right] \omega \, dx \tag{1.22}$$

only captures the information of ω^-, and this the main novelty of this chapter. We note that for small $\varepsilon > 0$, (1.22) only penalizes a ε-neighborhood around the jump point of u. By using fine properties of SBV functions (see Theorem 3.4), we are able to incorporate

u and v in our model such that (1.22) will only penalize along the direction $-\nu_{S_\omega}$ (see Notation 3.3). This will be carried out in Proposition 3.26.

1.4. **Some insights from finite resolution images.** In [35] the existence of a minimizer $\tilde{\alpha} > 0$ of the error function

$$\mathcal{E}(\alpha) := \|u_\alpha - u_c\|^2_{L^2(Q)}\,, \tag{1.23}$$

where u_α is obtained from (1.3) with TV as the regularizer, has been established. Still, some important properties like convexity and differentiability have not yet been addressed, and an efficient numerical scheme to locate $\tilde{\alpha}$ is in need (see, e.g., [60]). To develop such efficient scheme, we observe from numerical simulations that $\mathcal{E}(\cdot)$ is likely to be *strictly quasi-convex* [1] (see (1.26) for the definition), and if indeed it is, then a quasi-convex programming method can be inserted (see, e.g., [53]), which has been proven to be computationally efficient. Unfortunately, strict quasi-convexity of (1.23) is not easily established. Successful attempts have been made in settings in which the regularization term is linear and smooth, for example, using the $W^{1,2}$ Sobolev seminorm as the regularization term (see, e.g., [55]). But, to the author's knowledge, nothing of the kind has been investigated when regularization term is non-linear and non-smooth such as the total variation seminorm.

In Chapter 4 we take the first small step in this direction and only the one-dimensional case is investigated. Although it might relevant to the denoising of bar codes (see [77]), it is of marginal interest within the context of image reconstruction. However, extending a similar analysis to the two dimensional setting is quite challenge due to the lack of explicit expression for the minimizer u_α, which is an important ingredient in the analysis of the one dimensional case (see Theorem 4.5).

In this part of work we introduce a new way to represent the clean image and the noise, which is compatible with a discrete computer image data, in the domain $I := (0,1)$, and hence we may apply our PDEs and functional analysis tools on it. To be precise, we assume that an ideal clean image $u_c \in BV(I)$ can only be captured by a "super" camera which has infinite resolution, and we assume that a finite $N \in \mathbb{N}$ resolution level image captured by a real world digital camera is a piecewise constant function $u_{c,N}$, which is related to u_c via its averages

$$u_{c,N}(x) := \fint_{I_N(k)} u_c\,dy \text{ for } x \in I_N(k), \tag{1.24}$$

where $I_N(k) := ((k-1)/N, k/N)$, $1 \le k \le N$. We also introduce the family

$$\mathcal{I}_N := \{I_N(k),\ 1 \le k \le N\}\,.$$

[1] Note that this notation of quasi-convexity is not related in any way to the notation of quasiconvexity as introduced by Morrey (see [26, 47]), which is used in Section 6.4.2

and we use $u_{c,N}(I_N(k))$ to denote the value of $u_{c,N}(x)$ for $x \in I_N(k)$. Similarly, in two dimensions, we define

$$u_{c,N}(x) := \fint_{Q_N(i_N, j_N)} u_c \, dx \text{ for } x \in Q_N\left(i_N, j_N\right),$$

where $Q_N\left(i_N, j_N\right)$ is defined in Notation 1.1 item 1.

Then, we may write scheme (1.3) with respect to a finite resolution image $u_{c,N}$ as follows:

Level 1.

$$\tilde{\alpha} \in \arg\min \left\{ \|u_\alpha - u_{c,N}\|^2_{L^2(I)} : \ \alpha > 0 \right\},$$

Level 2.

$$u_\alpha := \arg\min \left\{ \|u - u_{\eta,N}\|^2_{L^2(I)} + \alpha \, |u|_{TV(I)} : \ u \in BV(I) \right\}.$$

We present our first main result of Chapter 4 in Theorem 1.4. To be precise, we define the *reconstruction operator* $\mathscr{L}$ by

$$\mathscr{L}(\alpha, v, I) := \arg\min \left\{ \frac{1}{2} \|u - v\|^2_{L^2(I)} + \alpha TV(u, I) \right\}, \tag{1.25}$$

for $v \in L^2(I)$ and $\alpha \in \mathbb{R}^+$.

In Theorem 1.4 we discuss the quasi-convexity of (1.23) with finite resolution data $u_{c,N}$ provided that $u_{c,N}$ is monotone. We say that a function $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ is *strictly quasi-convex* ([14], Section 3.4) if for all α_1, $\alpha_2 \in \mathbb{R}^+$ and $\lambda \in (0,1)$ we have

$$f(\lambda\alpha_1 + (1-\lambda)\alpha_2) < \max\left\{f(\alpha_1), f(\alpha_2)\right\}. \tag{1.26}$$

We show that under certain assumptions (see Assumption B.10) the following result holds (see Theorem 4.13):

Theorem 1.4. *Let $I := (0,1)$, let $u_c \in BV(I)$ be monotone, and let a resolution level $N \in \mathbb{N}$ be given. Then the error function*

$$\mathcal{E}_N(\alpha) := \frac{1}{2} \|\mathscr{L}(\alpha, u_{\eta,N}) - u_{c,N}\|^2_{L^2(I)}, \ \alpha \in \mathbb{R}^+, \tag{1.27}$$

is strictly quasi-convex provided that $u_{\eta,N}$ satisfies Assumption B.10.

We remark that the Assumption B.10 cannot be relaxed, and in Section 4.3.1 we present counterexamples to show that removing any of its conditions would result in losing the quasi-convexity of $\mathcal{E}_N(\alpha)$ nearby $\tilde{\alpha}_N$, which is the minimizer of (1.27).

However, Assumption B.10 is very restrictive and unlikely to be satisfied in concrete settings, and requiring u_c to be monotone renders Theorem 1.4 to be less interesting. Therefore, we need an alternative way to locate $\tilde{\alpha}_N$ without assuming Assumption B.10 nor the monotonicity of u_c. To overcome this drawback, we show in Chapter 4 that $\mathcal{E}_N$, defined in

(1.27), is piecewise convex with finite many pieces both for one dimension and two dimensions, and this is enough for our purpose. To do so, we define the *stopping time* $\alpha_s(v)$ of a function $v \in L^\infty(Q)$ via the following definition.

Definition 1.5. *Let $v \in L^\infty(Q)$ be given. We say that $\alpha_s(v) \in [0, +\infty)$ is the* stopping time *for v if*

$$\mathscr{L}(\alpha_s, v, Q) = \mathscr{L}(\alpha_s + \alpha, v, Q) =: C(v) \text{ and } \mathscr{L}(\alpha_s, v, Q) \neq \mathscr{L}(\alpha_s - \alpha, v, Q) \tag{1.28}$$

for all $\alpha > 0$, where $C(v)$ is a constant depends on v.

By its definition, if it exists then the stopping time is unique. In Section 4.4.1 we show that the stopping time $\alpha_s(u_{\eta,N})$ exists where $\alpha_s(u_{\eta,N})$ is defined in (4.3) with Q replaced by $I := (0,1)$. Next, in Proposition 4.16, using Theorem 4.5 repeatedly, we show that the *level N error function*

$$\mathcal{E}_N(\alpha) := \frac{1}{2}\int_I |\mathscr{L}(\alpha, u_{\eta,N}, I) - u_{c,N}|^2 \, dx \tag{1.29}$$

is continuous, and there exist finitely many $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_M = \alpha_s(u_{\eta,N}) < +\infty$ such that in each interval $[\alpha_i, \alpha_{i+1})$, $\mathcal{E}_N(\cdot)$ is convex and $\mathcal{E}_N'(\cdot)$ is linearly increasing. Hence, a direct search, which we detail in Section 4.4.2, of a minimizer α_m of (1.29) inside the finite interval $[0, \alpha_s(u_{\eta,N})]$ can be executed numerically and terminated within a finite time, although it may take a long CPU time.

The behavior of (1.29) in the two dimensional setting is also discussed in Chapter 4. We present a two dimensional version of Proposition 4.16 in Proposition 4.20. Although the statement of Proposition 4.20 is weaker compared to that of Proposition 4.16, due to the lack of the two dimensional version of Theorem 4.5, it is still sufficient to allow us to perform the same direct search to locate $\tilde{\alpha}$ within a finite time.

In Chapter 4 in addition to (1.15) we also introduce another spatially dependent training scheme. We recall Notation 1.1 and introduce our new spatially training scheme as follows:

Level 1.

$$\tilde{\mathcal{L}} \in \arg\min \left\{ \|u_c - u_{\mathcal{L}}\|^2_{L^2(Q)}, \;\; \mathcal{L} \in \mathcal{V}_A \right\} \tag{1.30}$$

Level 2.

$$\begin{aligned} &u_{\mathcal{L}}(x) := u_{\tilde{\alpha}_L}(x) \text{ for } x \in L \text{ and } L \in \mathcal{L}, \text{ where} \\ &u_{\tilde{\alpha}_L} := \arg\min \left\{ \|u - u_\eta\|^2_{L^2(L)} + \tilde{\alpha}\, |u|_{TV(L)} : \; u \in BV(L) \right\} \end{aligned} \tag{1.31}$$

Level 3. for any given $L \in \mathcal{Q}_A$, set

$$\begin{aligned} &\tilde{\alpha}_L \in \arg\min \left\{ \|u_\alpha - u_c\|^2_{L^2(L)} : \; \alpha > 0 \right\}, \text{ where} \\ &u_\alpha := \arg\min \left\{ \|u - u_\eta\|^2_{L^2(L)} + \alpha\, |u|_{TV(L)} : \; u \in BV(L) \right\} \end{aligned} \tag{1.32}$$

We prove in Chapter 4 that under a mild assumption on the noise η_N, the scheme (1.30) is able to fully recover the clean image u_c as the resolution level N goes to infinite and the box constraint index $A \to 0$. To be precise, let

$$\mathcal{P}_N(A) := \inf\left\{\|u_{c,N} - u_{\mathcal{L}}\|^2_{L^2(Q)},\ \ \mathcal{L} \in \mathcal{V}_A\right\}$$

and

$$\mathcal{P}(N) := \mathcal{P}_N(1/N) = \inf\left\{\|u_{c,N} - u_{\mathcal{L}}\|^2_{L^2(Q)},\ \ \mathcal{L} \in \mathcal{V}_{1/N}\right\}$$

where $u_{\mathcal{L}}$ is obtained by replacing u_η with $u_{\eta,N}$ in (1.31) and (1.32). In Theorem 4.24 we prove the following result:

Theorem 1.6. *Assume that the noise η_{N^2} has locally average 0, that is,*

$$⨍_L \eta_{N^2} = 0$$

for any $L \in \mathcal{Q}_{1/N}$ defined in Notation 1.1. Then

$$\lim_{N\to\infty} \mathcal{P}(N^2) = 0.$$

1.5. **The comprehensive training scheme.** Up to now we have introduced some new ideas to improve the original training scheme (1.3), including a training scheme with respect to regularizers in (1.10), and two spatially training schemes in (1.15) and (1.30) with a fixed regularizer. In Chapter 5 we summarize those ideas and generalize them even further.

The skeleton of the bilevel training scheme for image processing problem can be stated as follows:

Level 1.

Search for the best choice of α, $\mathcal{R}$, and $\mathcal{L}$ so that $\mathcal{A}(u_{\alpha,\mathcal{R},\mathcal{L}} - u_c, Q)$ attains its minimum

Level 2.

$$u_{\alpha,\mathcal{R},\mathcal{L}} \in \arg\min\left\{\text{certain functional of } u_\eta,\text{ defined by using } \mathcal{F},\ \mathcal{R},\ \alpha,\text{ and } \mathcal{L}\ \right\},$$

where the operator $\mathcal{A}$, called the *assessment operator*, assesses the quality of a denoised image. To be precise, we assume that the smaller is the value of $\mathcal{A}\left(u_c - u_{\alpha,\mathcal{R},\mathcal{L}}\right)$ then the higher is the quality of $u_{\alpha,\mathcal{R},\mathcal{L}}$. We also assume that the corrupted image u_η belongs to a Hilbert space Y, which is usually taken to be L^2 in image denoising and deblurring problems.

Recall that the training ground $\mathcal{V}_A$ defined in Notation 1.1 item 3 is only a finite collection (in discrete space), which renders training schemes (1.15) and (1.30) less interesting. While digitally acquired image data is discrete, the aim of high resolution image reconstruction and processing is always to compute an image that is close to the real, that is, infinite dimensional, and HD photography produces larger and larger images with a frequently increasing number of megapixels. Thus, we should aim for training schemes that accentuate and preserve qualitative properties in images independent of the resolution of the image itself.

To this purpose, we introduce a version of the training ground $\mathcal{V}_A$ in continuum space.

1. $\mathcal{H}_A$ is the family of rectangles such that

$\mathcal{H}_A := \{L \subset Q: \ L \text{ is an open rectangle with the shortest side length greater than or equal to } A\}$;

2. $\mathcal{L}$ is the collection of (finitely many) $L \in \mathcal{H}_A$ such that

$$\mathcal{L} := \left\{ L \in \mathcal{H}_A: \ L \text{ are mutually disjoint, } Q \subset \bigcup \overline{L} \right\}, \tag{1.33}$$

3. $\mathcal{V}_A$ denotes the collection of all possible $\mathcal{L}$, and we define, for any $\mathcal{L}$, $\mathcal{L}' \in \mathcal{V}_A$,

$$d_{\mathcal{V}_A}(\mathcal{L}, \mathcal{L}') := \max \left\{ \min \left\{ \|\chi_L - \chi_{L'}\|_{L^1(Q)}: \ L' \in \mathcal{L}' \right\}: \ L \in \mathcal{L} \right\}$$

where χ_L is the characteristic function over L.

Clearly, $\#\{\mathcal{V}_A\} = \infty$.

1.5.1. *The parameter training scheme* $(\mathcal{PT})$. The training scheme, stated in (1.3) and (1.6), with respect to the tuning parameter, can be generalized as follows:

Level 1.

$$\tilde{\alpha} \in \arg\min \left\{ \mathcal{A}\left(u_\alpha - u_c, Q\right): \ \alpha > 0 \right\}, \tag{1.34}$$

Level 2.

$$u_\alpha := \arg\min \left\{ \mathcal{F}\left(u - u_\eta, Q\right) + \mathcal{R}(u, \alpha, Q): \ u \in X_{\mathcal{R}} \right\},$$

where the regularizer $\mathcal{R}$ is given by

$$\mathcal{R}(u, \alpha, Q) := \sum_{i=1}^{N} \alpha_i \left| R_i u \right|_{\mathcal{M}(Q)}$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{N_{\mathcal{R}}}) \in (\mathbb{R}^+)^N$, R_i are linear operators, and the values R_iu are penalized in the Radon norm $|\cdot|_{\mathcal{M}}$. For example, we may take $\mathcal{R}(u, \alpha, Q)$ to be $\alpha TV(u)$ as in (1.4), with $N = 1$ and $\alpha \in \mathbb{R}^+$, or $TGV^2_{\alpha_1,\alpha_2}(u)$ with $N = 2$ and $\alpha = (\alpha_1, \alpha_2) \in (\mathbb{R}^+)^2$. Since the result of (1.34) is a parameter $\tilde{\alpha}$, we call this training scheme "**P**arameter **T**raining scheme", and hence the name $(\mathcal{PT})$. The scheme $(\mathcal{PT})$ has at least one solution $\tilde{\alpha}_{\mathcal{R}} \in (0, +\infty]^N$ provided some mild assumptions on u_c and η are satisfied, and we refer readers to [35] for details.

1.5.2. *The regularizer training scheme* $(\mathcal{RT})$. In Chapter 2 we introduced a bilevel training scheme which trains the optimal regularizer. In Chapter 5 we generalize such training scheme, named as "**R**egularizer **T**raining scheme", and hence $(\mathcal{RT})$, as follows:

Level 1.

$$\tilde{\gamma} \in \arg\min \left\{ \mathcal{A}\left(u_{c,N} - u_{\mathcal{R}[\gamma]}, Q\right): \ \gamma \in \Pi \right\} \tag{1.35}$$

Level 2.

$$u_{\mathcal{R}[\gamma]} := \arg\min \left\{ \mathcal{F}(u_{\eta,N} - u, Q) + \mathcal{R}[\gamma](u, Q),\ u \in X_{\mathcal{R}[\gamma]} \right\},$$

where the indexing set Π of the regularizer space $\mathscr{R}$ is defined as follows:

Definition 1.7 (The indexing set of $\mathscr{R}$). *Let $\Pi := \Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_{N_{\mathscr{R}}}$, where the* indexing dimension *is $N_{\mathscr{R}} \in \mathbb{N}$, and each Γ_i is a compact subset of $\mathbb{M}^{n_i \times k_i}$ (vector space of $n_i \times k_i$ real valued matrices, n_i, $m_i \in \mathbb{N}$). We say that a space (set) of regularizers $\mathscr{R}$ is indexed by Π if each $\mathcal{R} \in \mathscr{R}$ can be uniquely represented by an element $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_{N_{\mathscr{R}}}) \in \Pi$, and we use $\mathcal{R}[\gamma]$ to indicate that $\mathcal{R}$ is indexed by γ. Moreover, we endow $\mathscr{R}$ with the norm defined by*

$$d_{\mathscr{R}}(\mathcal{R}[\gamma], \mathcal{R}[\gamma']) := \sum_{i=1}^{N_{\mathscr{R}}} \left\| \gamma_i - \gamma_i' \right\|_{\Gamma_i}.$$

Definition 1.8. *Given $u_\eta \in Y$, we define the* reconstruction map *$\mathcal{S}$: $\mathscr{R} \to X_{\mathcal{R}}$ by*

$$\mathcal{S}_{u_\eta}(\mathcal{R}) := \arg\min \left\{ \mathcal{F}(u_\eta - u, Q) + \mathcal{R}(u, Q) :\ u \in X_{\mathcal{R}} \right\}.$$

Assumption 1.9 ($\mathcal{A}$-*l.s.c* with respect to $d_{\mathscr{R}}$)**.** *We say that the operator $\mathcal{S}_{u_\eta}(\mathcal{R})$ is $\mathcal{A}$-l.s.c. with respect to $d_{\mathscr{R}}$ if for every $\{\mathcal{R}[\gamma_n]\}_{n=1}^{\infty} \subset \mathscr{R}$ with $\lim_{n\to\infty} d_{\mathscr{R}}(\mathcal{R}[\gamma_n], \mathcal{R}[\gamma]) = 0$,*

$$\liminf_{n\to\infty} \mathcal{A}(\mathcal{S}_{u_\eta}(\mathcal{R}[\gamma_n]) - u_c) \geq \mathcal{A}(\mathcal{S}_{u_\eta}(\mathcal{R}[\gamma]) - u_c).$$

Theorem 1.10. *Let $\mathcal{A}$ be an assessment operator satisfying Assumption 5.3. If $\mathcal{S}_{u_\eta}$ is $\mathcal{A}$-l.s.c. with respect to $d_{\mathscr{R}}$, then problem* (5.5) *admits a solution $\tilde{\mathcal{R}} \in \mathscr{R}$.*

1.5.3. *Training scheme in regularizer and parameter spaces.* We recall the definition of the so-called *box constraint*.

Definition 1.11. *We say that a vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$ satisfies the box constraint if there exist two positive numbers $0 < A < 1$ such that $\alpha_i \in [A, 1/A]$ for $i = 1, 2, \ldots, n$.*

From now on, we use $\Pi_{\mathscr{R}}$ to denote the indexing set of the regularizer space $\mathscr{R}$ (see Definition 1.7).

Definition 1.12. *We say that a space $\mathscr{R}$ has* operator dimension *$n_{\mathscr{R}} \in \mathbb{N}$ if there exists a set of operators*

$$\left\{ R_i(\cdot,\cdot,\cdot) :\ Y \times Y^{n_{\mathscr{R}}} \times \Pi \to \mathbb{R}^+ \text{ for } i = 1, \ldots, n_{\mathscr{R}} \right\} \text{ with } R_i(tu, tv, \cdot) = tR_i(u, v, \cdot),\ t \in \mathbb{R}^+$$

such that each $\mathcal{R}[\gamma] \in \mathscr{R}$ can be represented by

$$\mathcal{R}[\gamma](u, Q) = \inf \left\{ R_1(u, v, \gamma) + R_2(v, \gamma) + \ldots + R_{n_{\mathcal{R}}}(v, \gamma) :\ v \in Y^{n_{\mathscr{R}}} \right\}. \tag{1.36}$$

We define a scaled version of $\mathcal{R}[\gamma] \in \mathscr{R}$ by adding a parameter $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{n_{\mathscr{R}}}) \in \mathbb{R}^{n_{\mathscr{R}}}$, which satisfies the box constraint in Definition 5.14, in the following sense:

$$\mathcal{R}[\alpha, \gamma](u) := \inf \left\{ R_1(\alpha_1 u, \alpha_1 v, \gamma) + R_2(\alpha_2 v, \gamma) + \ldots + R_{n_{\mathscr{R}}}(\alpha_{n_{\mathscr{R}}} v, \gamma) :\ v \in Y^{n_{\mathscr{R}}} \right\}, \tag{1.37}$$

and we let

$$d_{A,\mathscr{R}}(\mathcal{R}[\alpha,\gamma],\mathcal{R}[\alpha',\gamma']) := d_{\mathscr{R}}(\mathcal{R}[\gamma],\mathcal{R}[\gamma']) + \left|\alpha-\alpha'\right|.$$

We improve scheme $(\mathcal{RT})$ by inserting parameters in the way of (5.8), and hence $(\mathcal{RT})$ is now able to train the parameters and regularizers, simultaneously.

Level 1.

$$(\tilde{\alpha},\tilde{\gamma}) \in \arg\min\left\{\mathcal{A}\left(u_c - u_{\mathcal{R}[\alpha,\gamma]}, Q\right):\ \gamma\in\Pi_{\mathscr{R}},\ \alpha\in[A,1/A]^{n_{\mathscr{R}}}\right\} \tag{1.38}$$

Level 2.

$$u_{\mathcal{R}[\alpha,\gamma]} := \arg\min\left\{\mathcal{F}(u_\eta - u, Q) + \mathcal{R}[\alpha,\gamma](u,Q),\ u\in X_{\mathcal{R}[\gamma]}\right\}. \tag{1.39}$$

We improve Assumption 5.12 to accommodate the parameter spaces $[A,1/A]^{n_{\mathscr{R}}}$.

Assumption 1.13 (*A-$\mathcal{A}$-l.s.c.* with respect to $d_{A,\mathscr{R}}$)**.** *We say that the operator $\mathcal{S}(\mathcal{R})$ is A-$\mathcal{A}$-l.s.c. with respect to $d_{A,\mathscr{R}}$ if for every $\{(\alpha_n,\gamma_n)\}_{n=1}^{\infty}\subset[A,1/A]^{n_{\mathscr{R}}}\times\Pi$ with*

$$\lim_{n\to\infty} d_{A,\mathscr{R}}(\mathcal{R}[\alpha_n,\gamma_n],\mathcal{R}[\alpha,\gamma]) = 0,$$

we have

$$\liminf_{n\to\infty}\mathcal{A}(\mathcal{S}(\mathcal{R}[\alpha_n,\gamma_n]) - u_c) \geq \mathcal{A}(\mathcal{S}(\mathcal{R}[\alpha,\gamma]) - u_c).$$

1.5.4. *The comprehensive training scheme.* We propose a **C**omprehensive **T**raining scheme $(\mathcal{CT})$ taking into consideration several options:

Level 1.

$$\tilde{\mathcal{L}} \in \arg\min\left\{\mathcal{A}\left(u_c - \mathcal{P}(\mathcal{L})\right):\ \mathcal{L}\in\mathcal{V}_A\right\}, \tag{1.40}$$

Level 2.

$$\mathcal{P}(\mathcal{L}) \text{ is built upon the information of } \left\{(\tilde{\alpha}_L,\tilde{\gamma}_L)\right\}_L \text{ in each } L\in\mathcal{L},$$

Level 3.

$$\begin{aligned} \left\{(\tilde{\alpha}_L,\tilde{\gamma}_L)\right\}_L &\in \arg\min\left\{\mathcal{A}\left(u_c - u_{\mathcal{R}[\alpha,\gamma]}, L\right):\ \gamma\in\Pi,\ \alpha\in[A,1/A]^{n_{\mathscr{R}}}\right\}, \\ u_{\mathcal{R}[\alpha,\gamma]} &:= \arg\min\left\{\mathcal{F}(u_\eta - u, L) + \mathcal{R}[\alpha,\gamma](u,L),\ u\in X_{\mathcal{R}[\gamma]}\right\}. \end{aligned} \tag{1.41}$$

Here the operator $\mathcal{P}$: $\mathcal{V}_A \to Y$ acts as an assemble operator, using the local optimal re-construction information obtained in Level 3 within each subdomain L to construct a global re-constructed image $u_{\mathcal{L}}$, based on the partition domain $\mathcal{L}\in\mathcal{V}_A$.

Assumption 1.14. *We say that the operator $\mathcal{P}$: $\mathcal{V}_A\to Y$ is $\mathcal{A}$-l.s.c. with respect to $d_{\mathcal{V}_A}$ if for any sequence $\{\mathcal{L}_n\}_{n=1}^{\infty}\subset\mathcal{V}_A$ with $\lim_{n\to\infty} d_{\mathcal{V}_A}(\mathcal{L}_n,\mathcal{L})=0$,*

$$\liminf_{n\to\infty}\mathcal{A}(\mathcal{P}(\mathcal{L}_n) - u_c) \geq \mathcal{A}(\mathcal{P}(\mathcal{L}) - u_c).$$

Theorem 1.15. *If the assemble operator $\mathcal{P}$ is $\mathcal{A}$-l.s.c. with respect to $d_{\mathcal{V}_A}$, then problem* (5.24) *admints a solution $\tilde{\mathcal{L}}\in\mathcal{V}_A$.*

Two examples of assemble operator $\mathcal{P}(\mathcal{L})$ are presented in Chapter 5.

In the end, we would like to point out that, although the results obtained in this thesis are mainly motivated by problems from image processing, their applicability goes well beyond that and are related to problems involves parameter estimation. The box constraint we impose on (5.27), which is necessary to prove the existence of solutions, could be relaxed for certain regularizer spaces $\mathscr{R}$ based on observations made in [35]. However, to further open up the possibility to address more generalized regularizers, we proved our main result with box constraint, so that the scheme $(\mathcal{CT})$ is compatible with more regularizer spaces. Next steps from this work includes:

1. allow spatially dependent (weighted) tuning parameter $\omega \in BV$ to be 0 or $+\infty$ in subdomains;
2. new assessment operator $\mathcal{A}$. For example, the assessment operator optimized for edges enhancement and cancer detection;
3. construct new assemble operator $\mathcal{P}$. It is ideal to have a corresponding $\mathcal{P}$ for different assessment operators $\mathcal{A}$ optimized for different purposes;
4. design sophisticated numerical schemes to solve the optimal solution for scheme $(\mathcal{CT})$.

Chapter 2. **The fractional order total generalized variation**

In Chapter 2, we set $I := (0,1)$ is an unit interval.

2.1. **The theory of fractional sobolev spaces.** In what follows we will assume that $I = (0,1)$. We first recall a few results from the theory of Fractional Sobolev spaces. We refer to [36] for an introduction to the main results, and to [1, 56, 57, 65] and the references therein for a comprehensive treatment of the topic.

Definition 2.1 (Fractional Sobolev spaces)**.** *For $0 < s < 1$, $1 \le p < +\infty$, and $u \in L^p(I)$, we define the* Gagliardo seminorm *of u by*

$$|u|_{W^{s,p}(I)} := \left(\int_I \int_I \frac{|u(x)-u(y)|^p}{|x-y|^{1+sp}}\,dx\,dy\right)^{\frac{1}{p}}. \tag{2.1}$$

We say that $u \in W^{s,p}(I)$ if

$$\|u\|_{W^{s,p}(I)} := \|u\|_{L^p(I)} + |u|_{W^{s,p}(I)} < +\infty.$$

The following embedding results hold true ([36, Theorems 6.7, 6.10, and 8.2, and Corollary 7.2]).

Theorem 2.2 (Sobolev Embeddings - 1)**.** *Let $s \in (0,1)$ be given.*

1. *Let $p < 1/s$. Then there exists a positive constant $C = C(p,s)$ such that for every $u \in W^{s,p}(I)$ there holds*
$$\|u\|_{L^q(I)} \le C\|u\|_{W^{s,p}(I)} \tag{2.2}$$
for every $q \in [1, \frac{p}{1-sp}]$. If $q < \frac{p}{1-sp}$, then the embedding of $W^{s,1}(I)$ into $L^q(I)$ is also compact.

2. *Let* $p = 1/s$. *Then the embedding in* (2.2) *holds for every* $q \in [1, +\infty)$.
3. *Let* $p > 1/s$. *Then there exists a positive constant* $C = C(p, s)$ *such that for every* $u \in W^{s,p}(I)$ *we have*
$$\|u\|_{C^{0,\alpha}(I)} \leq C\|u\|_{W^{s,p}(I)},$$
with $\alpha := \frac{sp-1}{p}$.

The additional embedding result below is proved in [71, Corollary 19].

Theorem 2.3 (Sobolev Embeddings - 2)**.** *Let* $s \geq r$, $p \leq q$ *and* $s - 1/p \geq r - 1/q$, *with* $0 < r \leq s < 1$, *and* $1 \leq p \leq q \leq +\infty$. *Then*
$$W^{s,p}(I) \subset W^{r,q}(I)$$
and
$$|u|_{W^{r,q}(I)} \leq \frac{36}{rs} |u|_{W^{s,p}(I)}$$

The next inequality is a special case of [13, Theorem 1] and [64, Theorem 1].

Theorem 2.4 (Poincaré Inequality)**.** *Let* $p \geq 1$, *and let* $sp < 1$. *There exists a constant* $C > 0$ *such that*
$$\left\| u - \fint_I u(x)\, dx \right\|^p_{L^{\frac{p}{1-sp}}(I)} \leq \frac{Cs(1-s)}{(1-sp)^{p-1}} |u|^p_{W^{s,p}(I)}.$$

It is possible to construct a continuous extension operator from $W^{s,1}(I)$ to $W^{s,1}(\mathbb{R})$ (see, e.g., [36, Theorem 5.4]).

Theorem 2.5 (Extension Operator)**.** *Let* $s \in (0, 1)$, *and let* $1 \leq p < +\infty$. *Then* $W^{s,p}(I)$ *is continuously embedded in* $W^{s,p}(\mathbb{R})$, *namely there exists a constant* $C = C(p, s)$ *such that for every* $u \in W^{s,p}(I)$ *there exists* $\tilde{u} \in W^{s,p}(\mathbb{R})$ *satisfying* $\tilde{u}|_I = u$ *and*
$$\|\tilde{u}\|_{W^{s,p}(\mathbb{R})} \leq C \|u\|_{W^{s,p}(I)}.$$

The next two theorems ([75, Section 2.2.2, Remark 3, and Section 2.11.2]) yield an identification between fractional Sobolev spaces and Besov spaces in $\mathbb{R}$, and guarantee the reflexivity of Besov spaces $B^s_{p,q}$ for p, q finite.

Theorem 2.6 (Identification with Besov spaces)**.** *If* $1 \leq p < +\infty$ *and* $s \in \mathbb{R}^+ \setminus \mathbb{N}$, *then*
$$W^{s,p}(\mathbb{R}) = B^s_{p,p}(\mathbb{R})$$

Theorem 2.7 (Reflexivity of Besov spaces)**.** *Let* $-\infty < s < +\infty$, $1 \leq p < +\infty$ *and* $0 < q < +\infty$. *Then*
$$(B^s_{p,q}(\mathbb{R}))' = B^{-s}_{p',q'}(\mathbb{R}),$$
where $(B^s_{p,q}(\mathbb{R}))'$ *is the dual of the Besov space* $B^s_{p,q}(\mathbb{R})$, *and where* p' *and* q' *are the conjugate exponent of* p *and* q, *respectively.*

In view of Theorems 2.6 and 2.7 the following characterization holds true.

Corollary 2.8 (Reflexivity of Fractional Sobolev spaces). *Let $1 < p < +\infty$ and $s \in \mathbb{R}^+ \setminus \mathbb{N}$. Then the fractional Sobolev space $W^{s,p}(\mathbb{R})$ is reflexive.*

We conclude this section by recalling two theorems describing the limit behavior of the Gagliardo seminorm as $s \nearrow 1$ and $s \searrow 0$, respectively. The first result has been proved in [52, Theorem 3 and Remark 1], and [29, Theorem 1].

Theorem 2.9 (Asymptotic behavior as $s \nearrow 1$). *Let $u \in BV(I)$. Then*

$$\lim_{s \nearrow 1} (1-s) \left| u \right|_{W^{s,1}(I)} = \left| u' \right|_{\mathcal{M}_b(I)}.$$

Similarly, the asymptotic behavior of the Gagliardo seminorm has been characterized as $s \searrow 0$ in [64, Theorem 3].

Theorem 2.10 (Asymptotic behavior as $s \searrow 0$). *Let $u \in \cup_{0<s<1} W^{s,1}(\mathbb{R})$. Then,*

$$\lim_{s \searrow 0} s \left| u \right|_{W^{s,1}(\mathbb{R})} = 2 \left\| u \right\|_{L^1(\mathbb{R})}.$$

2.2. **The fractional order total generalized seminorm.** In this section we define the fractional order TGV^{1+s} seminorm, $0 < s < 1$, and prove some useful properties.

Definition 2.11 (The Fractional TGV Space). *Let $0 < s < 1$, $k \in \mathbb{N}$, and let $\alpha = (\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_k) \in \mathbb{R}_+^{k+1}$. For every $u \in L^1(I)$, we define its fractional TGV^{k+s} seminorm as follows.*

For $k = 1$ we set

$$\left| u \right|_{TGV_\alpha^{1+s}(I)} := \inf \Big\{ \alpha_0 \left| u' - s v_0 \right|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) \left| v_0 \right|_{W^{s,1+s(1-s)}(I)} : \\ v_0 \in W^{s,1+s(1-s)}(I), \int_I v_0(x)\, dx = 0 \Big\}.$$

For $k > 1$ we define

$$\left| u \right|_{TGV_\alpha^{k+s}(I)} := \inf \Big\{ \alpha_0 \left| u' - v_0 \right|_{\mathcal{M}_b(I)} + \alpha_1 \left| v_0' - v_1 \right|_{\mathcal{M}_b(I)} + \\ \cdots + \alpha_{k-1} \left| v_{k-2}' - s v_{k-1} \right|_{\mathcal{M}_b(I)} + \alpha_k s(1-s) \left| v_{k-1} \right|_{W^{s,1+s(1-s)}(I)} : \\ v_i \in BV(I) \text{ for } 0 \le i \le k-2,\ v_{k-1} \in W^{s,1+s(1-s)}(I), \int_I v_{k-1}(x)\, dx = 0 \Big\}.$$

For $0 < s < 1$, $k \in \mathbb{N}$, $\alpha = (\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_k) \in \mathbb{R}_+^{k+1}$, we say that $u \in BGV_\alpha^{k+s}(I)$ if

$$\left\| u \right\|_{BGV_\alpha^{k+s}(I)} := \left\| u \right\|_{L^1(I)} + \left| u \right|_{TGV_\alpha^{k+s}(I)} < +\infty,$$

and we write $u \in BGV^{k+s}(I)$ if there exists $\alpha \in \mathbb{R}_+^{k+1}$ such that $u \in BGV_\alpha^{k+s}(I)$. Note that if $u \in BGV_\alpha^{k+s}(I)$ for some $\alpha \in \mathbb{R}_+^{k+1}$, then $u \in BGV_\beta^{k+s}(I)$ for every $\beta \in \mathbb{R}_+^{k+1}$.

We observe that the TGV^{k+s} seminorm is actually "intermediate" between the TGV^k seminorm and the TGV^{k+1} seminorm. To be precise, we have the following identification.

Theorem 2.12. *For every $u \in BV(I)$, there holds*

$$\liminf_{s \nearrow 1} |u|_{TGV_\alpha^{1+s}(I)} \geq |u|_{TGV_\alpha^2(I)} \quad \text{and} \quad \lim_{s \searrow 0} |u|_{TGV_\alpha^{1+s}(I)} = |u'|_{\mathcal{M}_b(I)}.$$

Before proving the theorem we state and prove an intermediate result that will be crucial in determining the asymptotic behavior of the TGV^{1+s} seminorm as $s \nearrow 1$.

Proposition 2.13. *Let u be a Lipschitz function. Then*

$$\limsup_{s \nearrow 1} (1-s) \, |u|_{W^{s,1+s(1-s)}(I)} \leq |u'|_{\mathcal{M}_b(I)}.$$

Proof. We first observe that for $x, y \in I$ there holds

$$|x-y| \leq |x-y|^s.$$

Since u is Lipschitz, we have there exists a constant $L > 0$ such that for x, $y \in I$ there holds

$$|u(x) - u(y)| \leq L \, |x-y| \leq L \, |x-y|^s.$$

We observe that

$$\begin{aligned}
&|u|^{1+s(1-s)}_{W^{s,1+s(1-s)}(I)} \\
&= \int_I \int_I \frac{|u(x)-u(y)|^{1+s(1-s)}}{|x-y|^{1+s(1+s(1-s))}} dxdy = \int_I \int_I \frac{|u(x)-u(y)|^{s(1-s)} \, |u(x)-u(y)|}{|x-y|^{1+s(1+s(1-s))}} dxdy \\
&\leq L^{s(1-s)} \int_I \int_I \frac{|x-y|^{s^2(1-s)} \, |u(x)-u(y)|}{|x-y|^{1+s(1+s(1-s))}} dxdy = L^{s(1-s)} \int_I \int_I \frac{|u(x)-u(y)|}{|x-y|^{1+s}} dxdy \\
&= L^{s(1-s)} \, |u|_{W^{s,1}(I)}.
\end{aligned}$$

Therefore,

$$\limsup_{s \nearrow 1} (1-s) \, |u|_{W^{s,1+s(1-s)}(I)} \leq \limsup_{s \nearrow 1} (1-s) \, |u|^{\frac{1}{1+s(1-s)}}_{W^{s,1}(I)} \, L^{\frac{s(1-s)}{1+s(1-s)}}.$$

Thus, in view of Theorem 2.9 we conclude that

$$\limsup_{s \nearrow 1} (1-s) \, |u|_{W^{s,1+s(1-s)}(I)} \leq \limsup_{s \nearrow 1} (1-s) \, |u|^{\frac{1}{1+s(1-s)}}_{W^{s,1}(I)} \limsup_{s \nearrow 1} L^{\frac{s(1-s)}{1+s(1-s)}} = |u'|_{\mathcal{M}_b(I)}$$

as desired. □

A crucial ingredient in the proof of Theorem 2.12 is a compactness and lower semicontinuity result for maps with null averages and weighted $W^{s,1+s(1-s)}$-seminorm.

Proposition 2.14. *Let $\{s_n\} \subset (0,1)$ be such that $s_n \to \bar{s}$, with $\bar{s} \in (0,1]$. For every $n \in \mathbb{N}$ let $v_n \in W^{s_n,1+s_n(1-s_n)}(I)$ satisfy $\int_I v_n(x)\, dx = 0$, and*

$$\sup_{n \geq 1} s_n (1-s_n) \, |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} < +\infty. \tag{2.3}$$

Then, for $\bar{s} \in (0,1)$*, and up to the extraction of a (non-relabeled) subsequence, there exists* $\bar{v} \in W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)$ *such that*

$$v_n \to \bar{v} \quad \textit{strongly in } L^1(I),$$

and

$$\liminf_{n\to\infty} s_n(1-s_n)\,|v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} \geq \bar{s}(1-\bar{s})\,|\bar{v}|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)}\,. \tag{2.4}$$

The analogous statement holds for $\bar{s} = 1$*, by replacing* $W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)$ *with* $BV(I)$*, and* (2.4) *with*

$$\liminf_{n\to\infty} s_n(1-s_n)\,|v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} \geq \left|\bar{v}'\right|_{\mathcal{M}_b(I)}\,.$$

Proof. We first observe that for x, $y \in I$, and $s < t$, we have $|x-y|^{1+s} > |x-y|^{1+t}$. Hence, in view of (2.1) there holds

$$|u|_{W^{s,p}(I)} < |u|_{W^{t,p}(I)} \tag{2.5}$$

for $1 \leq p < +\infty$, and for every $u \in W^{t,p}(I)$.

Without loss of generality (and up to the extraction of a non-relabeled subsequence) we can assume that the sequences $\{s_n\}$ and $\{s_n(1-s_n)\}$ converge monotonically to $\bar{s}$ and $\bar{s}(1-\bar{s})$, respectively. Therefore, according to the value of $\bar{s}$ only 4 situations can arise:

Case 1: $0 < \bar{s} < \frac{1}{2}$: $s_n \searrow \bar{s}$ and $s_n(1-s_n) \searrow \bar{s}(1-\bar{s})$;
Case 2: $\frac{1}{2} \leq \bar{s} < 1$: $s_n \searrow \bar{s}$ and $s_n(1-s_n) \nearrow \bar{s}(1-\bar{s})$;
Case 3: $\frac{1}{2} < \bar{s} \leq 1$: $s_n \nearrow \bar{s}$ and $s_n(1-s_n) \searrow \bar{s}(1-\bar{s})$;
Case 4: $0 < \bar{s} \leq \frac{1}{2}$: $s_n \nearrow \bar{s}$ and $s_n(1-s_n) \nearrow \bar{s}(1-\bar{s})$.

We first consider Case 2. By (2.3) there exists a constant C such that

$$\sup_{n\geq 1} |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} \leq C.$$

We point out that the function $f : (0,1) \to \mathbb{R}$, defined as

$$f(x) := x - \frac{1}{1+x(1-x)} \quad \text{for every } x \in (0,1),$$

is strictly increasing on $(0,1)$. Thus, we can apply Theorem 2.3 with $s = s_n$, $r = \bar{s}$, $p = 1 + s_n(1-s_n)$, and $q = 1 + \bar{s}(1-\bar{s})$ and we obtain

$$|v_n|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)} \leq C\,|v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} \leq C. \tag{2.6}$$

Thus, by Theorem 2.4 and Corollary 2.8 there exists $\bar{v} \in W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)$ such that, up to the extraction of a (non-relabeled) subsequence, we have

$$v_n \rightharpoonup \bar{v} \quad \text{weakly in } W^{\bar{s},1+\bar{s}(1-\bar{s})}(I).$$

By the lower semicontinuity of the $W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)$ norm with respect to the weak convergence, and by (2.6) we deduce the inequality

$$\begin{aligned}\bar{s}(1-\bar{s})\,|\bar{v}|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)} &\le \liminf_{n\to+\infty} \bar{s}(1-\bar{s})\,|v_n|_{W^{s_n,1+s_n(1-s_n)}(I)}\\ &= \liminf_{n\to+\infty} s_n(1-s_n)\,|v_n|_{W^{s_n,1+s_n(1-s_n)}(I)}\,.\end{aligned}$$

In Case 1 we observe that the function $g:(0,1)\to\mathbb{R}$, defined as

$$g(x) := \frac{1}{1+x(1-x)} \quad \text{for every } x\in(0,1),$$

is strictly decreasing in $(0,\frac{1}{2}]$. Since $s_1 \ge s_n \ge \bar{s}$ for every n, there holds

$$\frac{1}{1+s_1(1-s_1)} \ge \frac{1}{1+\bar{s}(1-\bar{s})},$$

and by the properties of the functions f and g,

$$s_n - \frac{1}{1+s_n(1-s_n)} \ge \bar{s} - \frac{1}{1+\bar{s}(1-\bar{s})} \ge \bar{s} - \frac{1}{1+s_1(1-s_1)}.$$

By (2.3) there exists a constant C such that

$$\sup_{n\ge1} |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} \le C.$$

Choosing $s=s_n$, $r=\bar{s}$, $p=1+s_n(1-s_n)$, and $q=1+s_1(1-s_1)$ in Theorem 2.3 we have

$$|v_n|_{W^{\bar{s},1+s_1(1-s_1)}(I)} \le |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} \le C.$$

Thus, by Theorem 2.4 there exists a map $\bar{v}$ such that, up to the extraction of a (non-relabeled) subsequence, there holds

$$v_n \rightharpoonup \bar{v} \quad \text{weakly in } W^{\bar{s},1+s_1(1-s_1)}(I),$$

and by Theorem 2.2 also strongly in $L^1(I)$. In particular, Fatou's Lemma yields

$$|v|^{1+\bar{s}(1-\bar{s})}_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)} \le \liminf_{n_k\to+\infty} |v_{n_k}|^{1+s_{n_k}(1-s_{n_k})}_{W^{s_{n_k},1+s_{n_k}(1-s_{n_k})}(I)}\,,$$

which in turn implies the thesis.

We omit the proof of the result in Case 4, and in Case 3 for $\bar{s}<1$, as they follow from analogous arguments. Regarding Case 3 for $\bar{s}=1$, by (2.3) and (2.5) there exists a constant C such that

$$(1-s_n)\,|v_n|_{W^{s_n,1}(I)} \le C,$$

for every $n\in\mathbb{N}$. The thesis follows then by [52, Theorem 4]. □

We now prove Theorem 2.12.

Proof of Theorem 2.12. Let $\varepsilon > 0$ be given and $v_0 \in BV(I) \cap C^\infty(\mathbb{R})$ be such that

$$|u|_{TGV^2_\alpha(I)} \geq \alpha_0 \left|u' - v_0\right|_{\mathcal{M}_b(I)} + \alpha_1 |v_0'|_{\mathcal{M}_b(I)} - \varepsilon.$$

In view of Proposition 2.13 there holds

$$\begin{aligned}\limsup_{s\nearrow 1} |u|_{TGV^{1+s}_\alpha(I)} &\leq \limsup_{s \nearrow 1} \left\{ \alpha_0 \left| u' - sv_0 + s\int_I v_0(x)\,dx \right|_{\mathcal{M}_b(I)} \right. \\ &\qquad \left. + \alpha_1 s(1-s) \left|v_0\right|_{W^{s,1+s(1-s)}(I)} \right\} \\ &\leq |u|_{TGV^2_\alpha(I)} + \varepsilon + \left| \int_I v_0(x)\,dx \right|. \end{aligned} \tag{2.7}$$

For every $s \in (0,1)$, let $v_0^s \in W^{s,1+s(1-s)}(I)$ be such that $\int_I v_0^s(x)\,dx = 0$, and

$$\alpha_0 \left|u' - sv_0^s\right|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) \left|v_0^s\right|_{W^{s,1+s(1-s)}(I)} \leq |u|_{TGV^{1+s}_\alpha(I)} + (1-s). \tag{2.8}$$

In view of (2.7) and Proposition 2.14, there exists $v \in BV(I)$ such that, up to the extraction of a (non-relabeled) subsequence,

$$v_0^s \to v \quad \text{strongly in } L^1(I),$$

and

$$\lim_{s\nearrow 1} s(1-s) \left|v_0^s\right|_{W^{s,1+s(1-s)}(I)} \geq \left|v'\right|_{BV(I)}.$$

Passing to the limit in (2.8) we deduce the inequality

$$|u|_{TGV^2_\alpha(I)} \leq \alpha_0 \left|u' - v\right|_{\mathcal{M}_b(I)} + \alpha_1 \left|v'\right|_{BV(I)} \leq \liminf_{s\nearrow 1} |u|_{TGV^{1+s}_\alpha(I)},$$

which in turn implies the thesis.

To study the case $s \searrow 0$, we first observe that

$$\sup_{s \in (0,1)} |u|_{TGV^{1+s}_\alpha(I)} \leq |u'|_{BV(I)}. \tag{2.9}$$

Thus we only need to prove the opposite inequality. To this aim, for every $s \in (0,1)$ let $v_0^s \in W^{s,1+s(1-s)}(I)$ be such that $\int_I v_0^s(x)\,dx = 0$, and

$$\alpha_0 \left|u' - sv_0^s\right|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) \left|v_0^s\right|_{W^{s,1+s(1-s)}(I)} \leq |u|_{TGV^{1+s}_\alpha(I)} + s. \tag{2.10}$$

Since $s(1+s(1-s)) < 1$ for $s \in (0,1)$, by (2.9) and in view of Theorem 2.4 there holds

$$sv_0^s \to 0 \quad \text{strongly in } L^1(I).$$

Passing to the limit in (2.10) we deduce the inequality

$$\left|u'\right|_{\mathcal{M}_b(I)} \leq \limsup_{s\searrow 0} |u|_{TGV^{1+s}_\alpha(I)}.$$

The thesis follows owing to (2.9). $\square$

Corollary 2.15. *Let $k \geq 2$. For every $u \in BV(I)$, up to the extraction of a (non-relabeled) subsequence there holds*

$$\liminf_{s \nearrow 1} |u|_{TGV_\alpha^{k+s}(I)} \geq |u|_{TGV_\alpha^{k+1}(I)} \quad \text{and} \quad \lim_{s \searrow 0} |u|_{TGV_\alpha^{k+s}(I)} = |u|_{TGV_{\hat{\alpha}}^{k}(I)},$$

where $\hat{\alpha} := (\alpha_0, \ldots, \alpha_{k-1}) \in \mathbb{R}_+^k$.

Proof. The result follows by straightforward adaptations of the arguments in the proof of Theorem 2.12. □

We proceed by showing that the minimization problem in Definition 2.11 has a solution.

Proposition 2.16. *If the infimum in Definition 2.11 is finite, then it is attained.*

Proof. Let $k = 1$. Let $\alpha \in \mathbb{R}_+^2$, and let $u \in BGV_\alpha^{k+s}(I)$. We need to show that

$$|u|_{TGV_\alpha^{1+s}(I)} = \min \Big\{ \alpha_0 \left| u' - sv \right|_{\mathcal{M}_b(I)} + s(1-s)\alpha_1 \left| v \right|_{W^{s,1+s(1-s)}(I)} : \\ v \in W^{s,1+s(1-s)}(I), \int_I v(x)\,dx = 0 \Big\}. \quad (2.11)$$

We first observe that $u \in BV(I)$.

Indeed, let $\eta > 0$, and let $v \in W^{s,1+s(1-s)}(I)$ be such that $\int_I v(x)\,dx = 0$, and

$$\alpha_0 \left| u' - sv \right|_{\mathcal{M}_b(I)} + s(1-s)\alpha_1 \left| v \right|_{W^{s,1+s(1-s)}(I)} \leq |u|_{TGV_\alpha^{1+s}(I)} + \eta.$$

By Hölder inequality there holds

$$\begin{aligned} \alpha_0 \left| u' \right|_{\mathcal{M}_b(I)} &\leq \alpha_0 \left| u' - sv \right|_{\mathcal{M}_b(I)} + s\alpha_0 \|v\|_{L^1(I)} \\ &\leq \alpha_0 \left| u' - sv \right|_{\mathcal{M}_b(I)} + s\alpha_0 |v|_{W^{s,1+s(1-s)}(I)} + s\alpha_0 \|v\|_{L^{1+s(1-s)}(I)} \\ &\leq |u|_{TGV_\alpha^{1+s}(I)} + \eta + s(\alpha_0 - (1-s)\alpha_1) |v|_{W^{s,1+s(1-s)}(I)} + s\alpha_0 \|v\|_{L^{1+s(1-s)}(I)}. \end{aligned}$$

Let now $\{v_n\}_{n=1}^\infty \subset W^{s,1+s(1-s)}(I)$ be a minimizing sequence for (2.11). Since $s(1+s(1-s)) < 1$ for $s \in (0,1)$, by Theorem 2.2 there exists a constant C such that

$$\sup_{n \in \mathbb{N}} \|v_n\|_{W^{s,1+s(1-s)}(I)} \leq C.$$

Thus, by Corollary 2.8 there exists $\bar{v} \in W^{s,1+s(1-s)}(I)$ such that, up to the extraction of a (non-relabeled) subsequence, there holds

$$v_n \rightharpoonup \bar{v} \quad \text{weakly in } W^{s,1+s(1-s)}(I),$$

and hence by Theorem 2.2,

$$v_n \to \bar{v} \quad \text{strongly in } L^1(I).$$

The thesis follows now by the lower semicontinuity of the total variation and the $W^{s,1+s(1-s)}$ norm with respect to the L^1 convergence and the weak convergence in $W^{s,1+s(1-s)}(I)$, respectively.

For $k=2$, let $\{v_0^n\}_{n=1}^{\infty} \subset BV(I)$ and $\{v_1^n\}_{n=1}^{\infty} \subset W^{s,1+s(1-s)}(I)$ with $\int_I v_1^n(x)\,dx = 0$ for every $n \in \mathbb{N}$ be such that

$$\lim_{n\to+\infty} \left\{\alpha_0|u' - v_0^n|_{\mathcal{M}_b(I)} + \alpha_1|(v_0^n)' - sv_1^n|_{\mathcal{M}_b(I)} + \alpha_2 s(1-s)|v_1^n|_{W^{1,s(1-s)}(I)}\right\}$$
$$= TGV_\alpha^{2+s}(I).$$

Since $s(1+s(1-s)) < 1$ for $s \in (0,1)$, by Theorem 2.2 we obtain that $\{v_1^n\}_{n=1}^{\infty}$ is uniformly bounded in $W^{s,1+s(1-s)}(I)$. Therefore, $\{v_0^n\}_{n=1}^{\infty}$ is uniformly bounded in $BV(I)$, and there exist $v_0 \in BV(I)$ and $v_1 \in W^{s,1+s(1-s)}(I)$, with $\int_I v_1(x)\,dx = 0$, such that, up to the extraction of a (non-relabeled) subsequence,

$$v_0^n \rightharpoonup^* v_0 \quad \text{weakly* in } BV(I),$$

and

$$v_1^n \rightharpoonup v_1 \quad \text{weakly in } W^{s,1+s(1-s)}(I).$$

In particular, by Theorem 2.2,

$$v_1^n \to v_1 \quad \text{strongly in } L^1(I).$$

The minimality of v_0 and v_1 is a consequence of lower semicontinuity. The thesis for $k > 2$ follows by analogous arguments. □

We observe that the TGV^{k+s} seminorms are all equivalent to the total variation seminorm.

Lemma 2.17. *For every $k \geq 1$ and $0 < s < 1$, we have*

$$BV(I) \sim BGV^k(I) \sim BGV^{k+s}(I) \sim BGV^{k+1}(I).$$

Proof. We only show that

$$BV(I) \sim BGV^{1+s}(I) \sim BGV^2(I). \tag{2.12}$$

The proof of the inequality for $k > 1$ is analogous. In view of (2.9), to prove the first equivalence relation in (2.12) we only need to show that there exist a constant C and a multi-index $\alpha \in \mathbb{R}_+^2$ such that

$$\left|u'\right|_{\mathcal{M}_b(I)} \leq C\,|u|_{TGV_\alpha^{1+s}(I)}\,.$$

By Theorem 2.2 we have

$$\begin{aligned}\left|u'\right|_{\mathcal{M}_b(I)} &\leq \left|u' - sv_0\right|_{\mathcal{M}_b(I)} + s\,|v_0|_{L^1(I)}\\ &\leq \left|u' - sv_0\right|_{\mathcal{M}_b(I)} + Cs\,|v_0|_{W^{s,1+s(1-s)}(I)}\\ &= \left|u' - sv_0\right|_{\mathcal{M}_b(I)} + \frac{C}{(1-s)}s(1-s)\,|v_0|_{W^{s,1+s(1-s)}(I)}\end{aligned}$$

for every $v_0 \in W^{s,1+s(1-s)}(I)$. Thus

$$\left|u'\right|_{\mathcal{M}_b(I)} \leq C\,|u|_{TGV^{1+s}_{1,\frac{C}{(1-s)}}(I)}$$

for every $s \in (0,1)$. This completes the proof of the first equivalence in (2.12). Property (2.12) follows now by [17, Theorem 3.3]. □

We conclude this section with a proposition that will be crucial in establishing our new training scheme.

Proposition 2.18. *Let* $0 < A << 1$, $k \in \mathbb{N}$, $\{s_n\}_{n=1}^{\infty} \subset (0,1)$, *and* $\{\alpha^n\}_{n=1}^{\infty} \subset \mathbb{R}_+^{k+1}$, *with* $\alpha^n = (\alpha_0^n, \alpha_1^n, \dots, \alpha_k^n)$ *satisfying*

$$0 < \mathrm{A} < \inf\{\alpha_i^n,\ n \geq 1,\ 0 \leq i \leq k\} \leq \sup\{\alpha_i^n,\ n \geq 1,\ 0 \leq i \leq k\} < 1/A < +\infty. \quad (2.13)$$

Let $u_n \in BGV_{\alpha^n}^{k+s_n}(I)$ *be such that*

$$\sup_{n\in\mathbb{N}} \left\{ \|u_n\|_{BGV_{\alpha^n}^{k+s_n}(I)} \right\} < +\infty. \quad (2.14)$$

Then, up to the extraction of a (non-relabeled) subsequence, there exist $\bar{s} \in [0,1]$, $\alpha \in \mathbb{R}_+^{k+1}$ *and* $u \in BGV_{\alpha}^{k+\bar{s}}(I)$ *such that* $s_n \to \bar{s}$, $\alpha^n \to \alpha$, *and*

$$u_n \overset{*}{\rightharpoonup} u \text{ in } BV(I). \quad (2.15)$$

In addition, if $\bar{s} \in (0,1]$ *there holds*

$$|u|_{TGV_{\alpha}^{k+\bar{s}}(I)} \leq \liminf_{n\to+\infty} |u_n|_{TGV_{\alpha_n}^{k+s_n}(I)}. \quad (2.16)$$

If $\bar{s} = 0$ *we have*

$$|u|_{TGV_{\hat{\alpha}}^{k}(I)} \leq \liminf_{n\to+\infty} |u_n|_{TGV_{\alpha_n}^{k+s_n}(I)},$$

where $\hat{\alpha} \in \mathbb{R}_+^k$ *is the multi-index* $\hat{\alpha} := (\alpha_0, \dots, \alpha_{k-1})$, *and*

$$\alpha_0 \left|u'\right|_{\mathcal{M}_b(I)} \leq \liminf_{n\to+\infty} |u_n|_{TGV_{\alpha_n}^{k+s_n}(I)}, \quad (2.17)$$

for $k > 1$ *and* $k = 1$, *respectively.*

Proof. We prove the statement for $k = 1$. The proof of the result for $k > 1$ follows via straightforward modifications. For $k = 1$, we have $\{\alpha^n\}_{n=1}^{\infty} \subset \mathbb{R}^2$, and by (2.13) up to the extraction of a (non-relabeled) subsequence there holds

$$(\alpha_0^n, \alpha_1^n) \to (\alpha_0, \alpha_1), \text{ where } \mathrm{A} < \alpha_0,\ \alpha_1 < 1/A. \quad (2.18)$$

In addition, since $\{s_n\}_{n=1}^{\infty}$ is bounded, there exists $\bar{s} \in [0,1]$ such that $s_n \to \bar{s}$. By Proposition 2.16 we deduce that there exists $v_0^n \in W^{s_n, 1+s_n(1-s_n)}(I)$ such that

$$|u|_{TGV_{\alpha^n}^{1+s_n}(I)} = \alpha_0^n \left|u_n' - s_n v_0^n\right|_{\mathcal{M}_b(I)} + \alpha_1^n s_n(1-s_n) \left|v_0^n\right|_{W^{s_n,1+s_n(1-s_n)}(I)}.$$

Thus, by Theorem 2.4 and since $\int_I v_0^n(x)\,dx = 0$, we have (note that $s_n < 1$)

$$\begin{aligned} \left|u_n'\right|_{\mathcal{M}_b(I)} &\leq \left|u_n' - s_n v_0^n\right|_{\mathcal{M}_b(I)} + s_n \left|v_0^n\right|_{L^1(I)} \\ &\leq C\left\{ \left|u_n' - s_n v_0^n\right|_{\mathcal{M}_b(I)} + s_n(1-s_n)\left|v_0^n\right|_{W^{s_n,1+s_n(1-s_n)}(I)} \right\}. \end{aligned}$$

Hence, by (2.14) and (2.18),

$$\sup_{n\in\mathbb{N}} \left\{ \|u_n\|_{BV(I)} \right\} \leq C \sup_{n\in\mathbb{N}} \left\{ \|u_n\|_{BGV_{\alpha^n}^{1+s_n}(I)} \right\} < +\infty$$

which implies (2.15).

Note that for any $0 < s < 1$,

$$1 + s(1-s) - \frac{1}{1-s} = \frac{1}{1-s}(-s + s(1-s)^2) < 0.$$

Hence by applying again Theorem 2.4 we obtain

$$\|v_0^n\|_{L^{1+s_n(1-s_n)}(I)} \le \|v_0^n\|_{L^{\frac{1}{1-s_n}}(I)} \le C s_n(1-s_n)\, |v_0^n|_{W^{s_n,1+s_n(1-s_n)}(I)}\,,$$

which by (2.13) implies

$$\begin{aligned} s_n(1-s_n)\, \|v_n\|_{W^{s_n,1+s_n(1-s_n)}(I)} &\le \|v_0^n\|_{L^{1+s_n(1-s_n)}} + s_n(1-s_n)\, |v_0^n|_{W^{s_n,1+s_n(1-s_n)}(I)} \\ &\le C s_n(1-s_n)\, |v_0^n|_{W^{s_n,1+s_n(1-s_n)}(I)} \le C\, \|u_n\|_{BGV^{1+s_n}_{\alpha^n}(I)}\,. \end{aligned}$$

Therefore, by (2.14) we deduce the uniform bound

$$\sup_{n\ge 1} s_n(1-s_n)\, \|v_n\|_{W^{s_n,1+s_n(1-s_n)}(I)} < +\infty. \tag{2.19}$$

We subdivide the remaining part of the proof of Proposition 2.18 into 2 cases.

Case 1: $\bar{s} \in (0,1]$.

Assume first that $\bar{s} < 1$. By Proposition 2.14 there exists $v_0 \in W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)$ such that

$$v_0^n \to v_0 \quad \text{strongly in } L^1(I), \tag{2.20}$$

and

$$\liminf_{n\to\infty} s_n(1-s_n)\, |v_0^n|_{W^{s_n,1+s_n(1-s_n)}} \ge \bar{s}(1-\bar{s})\, |v_0|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)}\,. \tag{2.21}$$

By (2.20) and since $0 < \bar{s} < 1$, there holds

$$\begin{aligned} &\liminf_{n\to\infty} |u_n|_{TGV^{1+s_n}_{\alpha^n}(I)} \\ &\ge \liminf_{n\to\infty} \alpha_0^n \left|u_n' - s_n v_0^n\right|_{\mathcal{M}_b(I)} + \liminf_{n\to\infty} \alpha_1^n s_n(1-s_n)\, |v_0^n|_{W^{s_n,1+s_n(1-s_n)}(I)} \\ &\ge \alpha_0 \left|u' - \bar{s} v_0\right|_{\mathcal{M}_b(I)} + \alpha_1 \bar{s}(1-\bar{s})\, |v_0|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)} \ge |u|_{TGV^{1+\bar{s}}_{\alpha}(I)}\,, \end{aligned} \tag{2.22}$$

where in the last inequality we used the definition of the $TGV^{1+\bar{s}}_{\alpha}$-seminorm. In particular, $u \in BGV^{1+\bar{s}}_{\alpha}(I)$, and (2.16) is satisfied. The proof of the proposition for $\bar{s} = 1$ follows via the same argument, and by replacing $\bar{s}(1-\bar{s})\, |v_0|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)}$ in (2.21) and (2.22) with $|v_0'|_{\mathcal{M}_b(I)}$.

Case 2: $\bar{s} = 0$. In view of (2.19) and Theorem 2.4, up to the extraction of a (non-relabeled) subsequence we deduce that

$$s_n v_n \to 0 \quad \text{strongly in } L^1(I).$$

Hence, there holds

$$\liminf_{n\to\infty} |u_n|_{TGV^{1+s_n}_{\alpha^n}(I)}$$

$$\geq \liminf_{n\to\infty} \alpha_0^n \left|u_n' - s_n v_0^n\right|_{\mathcal{M}_b(I)} + \liminf_{n\to\infty} \alpha_1^n s_n(1-s_n)\left|v_0^n\right|_{W^{s_n,1+s_n(1-s_n)}(I)}$$
$$\geq \alpha_0 \left|u'\right|_{\mathcal{M}_b(I)},$$

which in turn implies (2.17). This completes the proof of the proposition. □

2.3. **The bilevel training scheme with respect to parameter and regularizer.** Let $r \geq 1$ be given and let $\lfloor r \rfloor$ denote the largest integer smaller than or equal to r. We propose the following training scheme $(\mathcal{R})$ which takes into account the order r of the regularizer and the parameter $\alpha \in \mathbb{R}_+^{\lfloor r \rfloor+1}$ simultaneously. We restrict our analysis to the case in which α and r satisfy the *box constraint*

$$(\alpha, r) \in [\mathrm{A}, 1/A]^{\lfloor r \rfloor+1} \times [K, K+1-A] \tag{2.23}$$

where $0 < \mathrm{A} << 1$ and $K \in \mathbb{N}$.

Our new training scheme $(\mathcal{R})$ is defined as follows:

Level 1.

$$(\bar{\alpha}, \bar{r}) := \arg\min \left\{ \int_I |u_{\alpha,r} - u_c|^2\, dx,\ (\alpha, r) \in [\mathrm{A}, 1/A]^{\lfloor r \rfloor+1} \times [K, K+S] \right\}, \tag{2.24}$$

Level 2.

$$u_{\alpha,r} := \mathop{\arg\min}_{u \in BGV_\alpha^r(I)} \left\{ \int_I |u - u_0|^2\, dx + |u|_{TGV_\alpha^r(I)} \right\}, \tag{2.25}$$

where $u_c \in L^1(I)$ represents a noise-free test picture, and $u_0 \in L^1(I)$ is the noisy image.

Note that we only allow the parameters α and the order r of regularizers to lie within a prescribed finite range. This is needed for the numerical realization of our model and also to force the optimal reconstructed image $u_{\bar{\alpha},\bar{r}}$ to remain inside our proposed space $BGV_{\bar{\alpha}}^{\bar{r}}(I)$ (see Proposition 2.18). In particular, if some of the components of α blow up to ∞, we might end up in the space $W^{r,1}(I)$, which is outside the purview of this chapter. We point out that no upper bound on R is required. Thus, despite the box constraint our analysis still incorporates a large class of image reconstruction regularizers, such as TV and TGV^2 (see, e.g., [34]).

Before we state the main theorem of this section, we prove a technical lemma that will guarantee the existence of a unique solution to (2.25).

Lemma 2.19. *For every $r \in [1, \mathrm{R}]$, and $\alpha \in \mathbb{R}_+^{\lfloor r \rfloor+1}$ there exists a unique $u_{\alpha,r} \in BGV_\alpha^r(I)$ solving the minimum problem* (2.25).

Proof. Let $\{u_n\}_{n=1}^\infty \subset BGV_\alpha^r(I)$ be a minimizing sequence for (2.25). By Lemma 2.17, $\{u_n\}_{n=1}^\infty$ is uniformly bounded in $BV(I)$. Thus there exists $\bar{u} \in BV(I)$ such that

$$u_n \rightharpoonup^* \bar{u} \quad \text{weakly* in } BV(I),$$

and hence also strongly in $L^2(I)$. The thesis follows then by Proposition 2.18 and by the strict convexity of the functional. $\square$

We are now in a position to prove existence and uniqueness of solutions to our training scheme.

Theorem 2.20. *Let $u_0, u_c \in BV(I)$ and $0 < A << 1$ be given. Under the box constraint* (2.23)*, the training scheme* ($\mathcal{R}$) *admits a unique solution $(\bar\alpha, \bar r) \in [\mathrm{A}, 1/A]^2 \times [1, 2-A]$ and provides an associated optimally reconstructed image $u_{\bar\alpha,\bar r} \in BGV^{\bar r}_{\bar\alpha}(I)$.*

Proof. Let $\{(\alpha_n, r_n)\}_{n=1}^{\infty}$ be a minimizing sequence for (2.24), with $(\alpha_n, r_n) \subset [\mathrm{A}, 1/A]^2 \times [1, 2-A]$ for every $n \in \mathbb{N}$. Let u_{α_n, r_n} be the unique solution to (2.25) provided by Lemma 2.19. By (2.25), there holds

$$|u_{\alpha_n,r_n}|_{BGV^{r_n}_{\alpha_n}(I)} \leq |u_0|_{TGV^{r_n}_{\alpha_n}(I)} \leq |u_0'|_{\mathcal{M}_b(I)}$$

for every $n \in \mathbb{N}$. There exists $\bar r \in [1, \mathrm{R}]$ such that, up to the extraction of a (non-relabeled) subsequence, there holds $r_n \to \bar r$. Note that for n big enough $r_n = 1 + s_n$, with $s_n \to \bar s$, and $\bar s \in [0, 1]$. We distinguish two cases.
Case 1. $\bar r \notin \mathbb{N}$. Then $\bar s \in (0, 1)$.
Case 2. $\bar r = 1$. Then $\bar s = 0$.
In both cases Proposition 2.18 yields the existence of a map $u_{\bar\alpha,\bar r} \in BGV^{\bar r}_{\bar\alpha}(I)$ such that

$$u_{\alpha_n,r_n} \overset{*}{\rightharpoonup} u_{\bar\alpha,\bar r} \quad \text{weakly* in } BV(I), \tag{2.26}$$

thus, in particular, strongly in $L^2(I)$. The existence of solutions follows then by lower semicontinuity, whereas the uniqueness is a direct consequence of the strict convexity of the L^2-error norm.

To be precise, we claim that

$$u_{\bar\alpha,\bar r} = \underset{u \in BGV^{\bar r}_{\bar\alpha}(I)}{\arg\min} \left\{ \int_I |u - u_0|^2 \, dx + |u|_{TGV^{\bar r}_{\bar\alpha}(I)} \right\}. \tag{2.27}$$

Define

$$a_r := \inf_{u \in BGV^{r}_{\alpha}(I)} \left\{ \int_I |u - u_0|^2 \, dx + |u|_{TGV^{r}_{\alpha}(I)} \right\}. \tag{2.28}$$

Assume (2.27) does not hold, i.e., there exists another $\tilde u$, as the minimizer of (2.27), such that

$$a_{\bar r} = \int_I |\tilde u - u_0|^2 \, dx + |\tilde u|_{TGV^{\bar r}_{\bar\alpha}(I)} < \int_I |u_{\bar\alpha,\bar r} - u_0|^2 \, dx + |u_{\bar\alpha,\bar r}|_{TGV^{\bar r}_{\bar\alpha}(I)} := a'.$$

In view of (2.26) we have

$$\liminf a_{r_n} \geq a'. \tag{2.29}$$

Assume $\bar r = 1$. Then we have $s_n \searrow \bar s = 0$. By Theorem 2.12 we may write

$$\lim_{n\to\infty} |\tilde u|_{TGV^{r_n}_{\alpha_n}(I)} = |\tilde u|_{TGV^{\bar r}_{\bar\alpha}(I)}$$

and hence

$$\lim_{n\to\infty}\int_I |\tilde{u}-u_0|^2\,dx + |\tilde{u}|_{TGV^{r_n}_{\alpha_n}(I)} = m_{\bar{r}} < a' \tag{2.30}$$

Now, in view of (2.29), for n large enough we have

$$a_{r_n} \geq a' - \frac{1}{4}(a' - a_{\bar{r}}),$$

and hence, push r even larger if need to, together with (2.30) we have

$$\int_I |\tilde{u}-u_0|^2\,dx + |\tilde{u}|_{TGV^{r_n}_{\alpha_n}(I)} \leq a_{\bar{r}} + \frac{1}{4}(a'-a_{\bar{r}}) < a' - \frac{1}{4}(a'-a_{\bar{r}}) \leq a_{r_n}$$

which contradicts to the definition of m_{r_n} in (2.28). The case $\bar{r} > 1$, i.e., $\bar{s} \in (0,1)$ can be proved in the same way as above. $\square$

Remark 2.21. Theorem 2.20 hold with the box constraint $[A, 1/A]^K \times [K, K+1-A]$, where $K \in \mathbb{N}$ and $0 < A << 1$. The proof of the result for $K > 1$ follows via straightforward modifications. In practice, we may take the box constraint $[A, 1/A]^K \times ([1, 1+S] \cup [2, 2+S])$, which covers a large class of image reconstruction regularizers, such as TV and TGV^2.

Chapter 3. The weighted Ambrosio - Tortorelli approximation scheme

3.1. **Definitions and preliminary results.** Throughout this part, $\Omega \subset \mathbb{R}^N$ is an open bounded set with Lipschitz boundary, and $I := (-1, 1)$.

Definition 3.1. *We say that $u \in BV(\Omega)$ is a special function of bounded variation, and we write $u \in SBV(\Omega)$, if the Cantor part of its derivative, $D^c u$, is zero, so that (see [4], (3.89))*

$$Du = D^a u + D^j u = \nabla u\,\mathcal{L}^N \lfloor \Omega + (u^+ - u^-)\nu_u \mathcal{H}^{N-1} \lfloor S_u. \tag{3.1}$$

Moreover, we say that
1. $u \in SBV^2(\Omega)$ *if* $u \in SBV(\Omega)$ *and* $\nabla u \in L^2(\Omega)$;
2. $u \in GSBV(\Omega)$ *if* $K \wedge u \vee -K \in SBV(\Omega)$ *for all* $K \in \mathbb{N}$.

Here we always identify $u \in SBV(\Omega)$ with its approximation representative $\bar{u}$, where

$$\bar{u}(x) := \frac{1}{2}\left[u^+(x) + u^-(x)\right],$$

with

$$u^+(x) := \inf\left\{t \in \mathbb{R} : \lim_{r\to 0} \frac{\mathcal{L}^N(B(x,r) \cap \{u > t\})}{r^N} = 0\right\},$$

and

$$u^-(x) := \sup\left\{t \in \mathbb{R} : \lim_{r\to 0} \frac{\mathcal{L}^N(B(x,r) \cap \{u < t\})}{r^N} = 0\right\}.$$

We note that $\bar{u}$ is Borel measurable (see [41], Lemma 1, page 210), and it can be shown that $\bar{u} = u$ $\mathcal{L}^N$-a.e. $x \in \Omega$, and that

$$(\bar{u})^+(x) = u^+(x) \text{ and } (\bar{u})^-(x) = u^-(x)$$

for $\mathcal{H}^{N-1}$-a.e. $x \in \Omega$ (see [41], Corollary 1, page 216). Furthermore, we have that

$$- < u^-(x) \leq u^+(x) < +\infty \tag{3.2}$$

for $\mathcal{H}^{N-1}$-a.e. $x \in \Omega$ (see [41], Theorem 2, page 211). The inequality (3.2) uniquely determines the sign of ν_u in (3.1).

Definition 3.2. *(The weight function) We say that ω: $\Omega \to (0, +\infty]$ belongs to $\mathcal{W}(\Omega)$ if $\omega \in L^1(\Omega)$ and has a positive lower bound, i.e., there exists $l > 0$ such that*

$$\operatorname{ess\,inf} \{\omega(x), x \in \Omega\} \geq l. \tag{3.3}$$

Without loss of generality, we take $l = 1$. Moreover, in this chapter we will only consider the cases in which ω is either a continuous function or a SBV function. If $\omega \in SBV$ then, in addition, we require that

$$\mathcal{H}^{N-1}(S_\omega) < \infty \text{ and } \mathcal{H}^{N-1}(\overline{S_\omega} \setminus S_\omega) = 0.$$

We next fix some notation which will be used throughout this chapter.

Notation 3.3. Let $\Gamma \subset \Omega$ be a $\mathcal{H}^{N-1}$-rectifiable set and $x \in \Gamma$ be given.

1. We denote by $\nu_\Gamma(x)$ a normal vector at x with respect to Γ, and $Q_{\nu_\Gamma}(x, r)$ is the cube centered at x with side length r and two faces normal to $\nu_\Gamma(x)$;
2. T_{x,ν_Γ} stands for the hyperplane normal to $\nu_\Gamma(x)$ and passing through x, and $\mathbb{P}_{x,\nu_\Gamma}$ stands for the projection operator from Γ onto T_{x,ν_Γ};
3. we define the hyperplane $T_{x,\nu_\Gamma}(t) := T_{x,\nu_\Gamma} + t\nu_\Gamma(x)$ for $t \in \mathbb{R}$;
4. we introduce the half-spaces

   $$H_{\nu_\Gamma}(x)^+ := \left\{y \in \mathbb{R}^N : \ \nu_\Gamma(x) \cdot (y - x) \geq 0\right\}$$

   and

   $$H_{\nu_\Gamma}(x)^- := \left\{y \in \mathbb{R}^N : \ \nu_\Gamma(x) \cdot (y - x) \leq 0\right\}.$$

   Moreover, we define the half-cubes

   $$Q^{\pm}_{\nu_\Gamma}(x, r) := Q_{\nu_\Gamma}(x, r) \cap H_{\nu_\Gamma}(x)^{\pm};$$

5. for given $\tau > 0$, we denote by $R_{\tau,\nu_\Gamma}(x, r)$ the part of $Q_{\nu_\Gamma}(x, r)$ which lies strictly between the two hyperplanes $T_{x,\nu_\Gamma}(-\tau r)$ and $T_{x,\nu_\Gamma}(\tau r)$;
6. we set $A_\delta := \{x \in \Omega : \operatorname{dist}(x, A) < \delta\}$ for every $A \subset \Omega$ and $\delta > 0$.

Theorem 3.4 ([41], Theorem 3, page 213)**.** *Assume that $u \in BV(\Omega)$. Then*

1. for $\mathcal{H}^{N-1}$-a.e. $x_0 \in \Omega \setminus S_u$,

$$\lim_{r \to 0} ⨍_{B(x_0,r)} |u(x) - \bar{u}(x_0)|^{\frac{N}{N-1}} \, dx = 0;$$

2. for $\mathcal{H}^{N-1}$-a.e. $x_0 \in S_u$,

$$\lim_{r \to 0} ⨍_{B(x_0,r) \cap H_{\nu_{S_u}}(x_0)^\pm} \left|u(x) - u^\pm(x_0)\right|^{\frac{N}{N-1}} \, dx = 0;$$

3. for $\mathcal{H}^{N-1}$ a.e. $x_0 \in S_u$

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon^{N-1}} \int_{S_u \cap Q_{\nu_{S_u}}(x_0,\varepsilon)} \left|u^+(x) - u^-(x)\right| d\mathcal{H}^{N-1}(x) = \left|u^+(x_0) - u^-(x_0)\right|.$$

Lemma 3.5 (Lemma 5.2 and Remark 5.3 in [16])**.** *If $u \in SBV^2(\Omega) \cap L^\infty(\Omega)$ then there exists a sequence $\{u_n\}_{n=1}^\infty \subset SBV^2(\Omega) \cap L^\infty(\Omega)$ such that the following hold:*

1. $\|u_n\|_{L^\infty} \le \|u\|_{L^\infty}$ *for all $n \in \mathbb{N}$;*
2. $u_n \to u$ *in $L^1(\Omega)$;*
3. $\nabla u_n \to \nabla u$ *strongly in $L^2(\Omega;\mathbb{R}^2)$;*
4. $\mathcal{H}^{N-1}(S_{u_n} \triangle S_u) \to 0$;
5. $\lim_{n\to\infty} \mathcal{H}^{N-1}\left(\overline{S_{u_n}} \setminus S_{u_n}\right) = 0$ *(see Remark 5.3 in [16]).*

Lemma 3.6. *Let $\omega \in SBV(I)$ be such that $\mathcal{H}^0(S_\omega) < \infty$. For every $x \in I$ the following statements hold:*

1. if $\{x_n\}_{n=1}^\infty$ and $\{y_n\}_{n=1}^\infty \subset I$ are such that $x_n < x < y_n$, $n \in \mathbb{N}$, and $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = x$, then

$$\liminf_{n\to\infty} \operatorname*{ess\,inf}_{y \in (x_n,y_n)} \omega(y) \ge \omega^-(x); \tag{3.4}$$

2.

$$\lim_{\substack{z_n \to x \\ \{z_n\}_{n=1}^\infty \subset H^{\pm}_{\nu_{S_\omega}(x)}}} \bar{\omega}(z_n) = \omega^\pm(x); \tag{3.5}$$

3.

$$\limsup_{d_{\mathcal{H}}(K_n,x)\to 0} \operatorname*{ess\,sup}_{\substack{z \in K_n \\ K_n \subset\subset H^{\pm}_{\nu_{S_\omega}(x)}}} \omega(z) = \omega^\pm(x), \tag{3.6}$$

where $K_n \subset\subset H^{\pm}_{\nu_{S_\omega}(x)}$ and $d_{\mathcal{H}}$ denotes the Hausdorff distance (see Definition A.1).

Proof. If $x \notin S_\omega$, then there exists $\delta > 0$ such that

$$S_\omega \cap (x-\delta, x+\delta) = \varnothing,$$

and so ω is absolutely continuous in $(x-\delta, x+\delta)$, and (3.4)-(3.6) are trivially satisfied with $\omega(x) = \omega^-(x)$ and with equality in place of the inequality in (3.4).

Let $x \in S_\omega$ and, without loss of generality, assume that $x = 0$, and let x_n, $y_n \to 0$ with $x_n < 0 < y_n$ for all $n \in \mathbb{N}$. Since $\mathcal{H}^0(S_\omega) < \infty$, choose $\bar{r} > 0$ such that

$$S_\omega \cap (0 - \bar{r}, 0 + \bar{r}) = 0.$$

As $\bar{\omega}$ is absolutely continuous in $(-\bar{r}, 0)$ and $(0, \bar{r})$, we may extend $\bar{\omega}$ uniquely to $x = 0$ from the left and the right (see Exercise 3.7 (1) in [58]) to define

$$\bar{\omega}(0^+) := \lim_{x \searrow 0^+} \bar{\omega}(x) \text{ and } \bar{\omega}(0^-) := \lim_{x \nearrow 0^-} \bar{\omega}(x). \tag{3.7}$$

Assume that (the case $\bar{\omega}(0^-) \ge \bar{\omega}(0^+)$ can be treated similarly)

$$\bar{\omega}(0^-) \le \bar{\omega}(0^+). \tag{3.8}$$

We first claim that

$$\liminf_{n\to\infty} \inf_{x\in(x_n,y_n)} \bar{\omega}(x) \geq \bar{\omega}(0^-). \tag{3.9}$$

Let $\varepsilon > 0$ be given. By (3.7) find $\bar{r} > \delta > 0$ small enough such that

$$\left|\bar{\omega}(x) - \bar{\omega}(0^-)\right| \leq \frac{1}{2}\varepsilon \text{ for all } x \in (-\delta, 0), \text{ and } \left|\bar{\omega}(x) - \bar{\omega}(0^+)\right| \leq \frac{1}{2}\varepsilon \text{ for all } x \in (0, \delta).$$

This, together with (3.8), yields

$$\bar{\omega}(x) \geq \bar{\omega}(0^-) - \frac{1}{2}\varepsilon,$$

for all $x \in (-\delta, \delta)$. Since $x_n \to 0$ and $y_n \to 0$, we may choose n large enough such that $(x_n, y_n) \subset (-\delta, \delta)$ and hence

$$\inf_{x\in(x_n,y_n)} \bar{\omega}(x) \geq \bar{\omega}(0^-) - \varepsilon.$$

Thus, (3.9) follows by the arbitrariness of $\varepsilon > 0$.

We next claim that

$$\bar{\omega}(0^\pm) = \omega^\pm(0). \tag{3.10}$$

By Theorem 3.4 part 2 and the fact that $\bar{\omega} = \omega$ $\mathcal{L}^1$-a.e., we have

$$\omega^-(0) = \lim_{r\to 0} \frac{1}{r} \int_{-r}^0 \omega(t)\, dt = \lim_{r\to 0} \frac{1}{r} \int_{-r}^0 \bar{\omega}(t)\, dt = \bar{\omega}(0^-),$$

where at the last equality we used the properties of absolutely continuous function and the definition of $\bar{\omega}(0^-)$. The equation $\bar{\omega}(0^+) = \omega^+(0)$ can be proved similarly.

Therefore

$$\liminf_{n\to\infty} \operatorname*{ess\,inf}_{x\in(x_n,y_n)} \omega(x) = \liminf_{n\to\infty} \inf_{x\in(x_n,y_n)} \bar{\omega}(x) \geq \bar{\omega}(0^-) = \omega^-(0),$$

which concludes (3.4), and (3.5) and (3.6) hold by (3.7) and (3.10). □

Lemma 3.7. *The space L^2_ω is a Hilbert space endowed with the inner product*

$$(u, v)_{L^2_\omega} := (u, v\,\omega)_{L^2} = \int u\, v\, \omega\, dx. \tag{3.11}$$

Proof. It is clear that (3.11) is an inner product. Also, $(u, u)_{L^2_\omega} = (u\sqrt{\omega}, u\sqrt{\omega})_{L^2} \geq 0$, and if $(u, u)_{L^2_\omega} = 0$ then by (3.3)

$$\int_\Omega u^2 \omega\, dx \geq \int_\Omega u^2 dx = 0,$$

and thus $u = 0$ *a.e.*

To see that L^2_ω is complete, and therefore a Hilbert space, let $\{u_n\}_{n=1}^\infty$ be a Cauchy sequence in L^2_ω and notice that $\{u_n\sqrt{\omega}\}_{n=1}^\infty$ is a Cauchy sequence in L^2. Hence, there is a function $v \in L^2$ such that $u_n\sqrt{\omega} \to v$ in L^2. Defining $u := v/\sqrt{\omega}$, we have that $u \in L^2_\omega$ and $u_n \to u$ in L^2_ω. □

Lemma 3.8. *Let* $\{u_n\}_{n=1}^{\infty} \subset W_{\omega}^{1,2}(\Omega)$ *be such that* $u_n \to u$ *in* L_{ω}^{1} *and*

$$\sup \int_{\Omega} |\nabla u_n|^2 \, \omega \, dx < \infty.$$

Then, for every measurable set $A \subset \Omega$

$$\liminf_{n\to\infty} \int_{A} |\nabla u_n|^2 \, \omega \, dx \geq \int_{A} |\nabla u|^2 \, \omega \, dx,$$

and $u \in W_{\omega}^{1,2}(\Omega)$.

Proof. By (3.3) we have that $\{\nabla u_n\}_{n=1}^{\infty}$ is uniformly bounded in $L^2(\Omega, \mathbb{R}^N)$ and $u_n \to u$ in $L^1(\Omega)$. Hence $\nabla u_n \rightharpoonup \nabla u$ in $L^2(\Omega; \mathbb{R}^N)$, and using standard lower semi-continuity of convex energies (see [45], Theorem 6.3.7), we conclude that

$$+\infty > \liminf_{n\to\infty} \int_{A} |\nabla u_n|^2 \, \omega \, dx \geq \int_{A} |\nabla u|^2 \, \omega \, dx,$$

for every measurable subset $A \subset \Omega$. In particular, with $A = \Omega$ and using the fact that $1 \leq \omega$ a.e., we deduce that $u \in W_{\omega}^{1,2}(\Omega)$. $\square$

Lemma 3.9. *Let* $u \in L_{\omega}^{1}(\Omega)$ *be such that*

$$\int_{\Omega} |\nabla u|^2 \, \omega \, dx + \int_{S_u} \omega \, d\mathcal{H}^{N-1} < +\infty. \tag{3.12}$$

Then $\mathcal{H}^{N-1}(S_u) < +\infty$ *and* $u \in GSBV_{\omega}(\Omega)$.

Proof. By (3.12) and (3.3)

$$\int_{\Omega} |\nabla u|^2 \; dx + \mathcal{H}^{N-1}(S_u) < +\infty,$$

and hence by [6] we have that $u \in GSBV(\Omega)$. To show that $u \in GSBV_{\omega}(\Omega)$, we only need to verify that

$$\int_{S_{u_K}} \left| u_K^+ - u_K^- \right| \omega \, d\mathcal{H}^{N-1} < +\infty$$

for every $K \in \mathbb{N}$ and with $u_K := K \wedge u \vee -K$. Indeed, by (3.12)

$$\int_{S_{u_K}} \left| u_K^+ - u_K^- \right| \omega \, d\mathcal{H}^{N-1} \leq 2K \int_{S_{u_K}} \omega \, d\mathcal{H}^{N-1} \leq 2K \int_{S_u} \omega \, d\mathcal{H}^{N-1} < +\infty.$$

$\square$

3.2. **The one dimensional case.**

3.2.1. *The case $\omega \in \mathcal{W}(I) \cap C(I)$.*

Let $\omega \in \mathcal{W}(I) \cap C(I)$ be given. Consider the functionals

$$AT_{\omega,\varepsilon}(u,v) := \int_I v^2 \left|u'\right|^2 \omega\, dx + \int_I \left[\frac{\varepsilon}{2}\left|v'\right|^2 + \frac{1}{2\varepsilon}(v-1)^2\right]\omega\, dx$$

for $(u,v) \in W^{1,2}_\omega(I) \times W^{1,2}(I)$, and let

$$MS_\omega(u) := \int_I \left|u'\right|^2 \omega\, dx + \sum_{x\in S_u} \omega(x)$$

be defined for $u \in GSBV_\omega(I)$ (Note that $AT_{1,\varepsilon}(u,v)$ and $MS_1(u)$ are, respectively, the non-weighted Ambrosio-Tortorelli approximation scheme and Mumford-Shah functional studied in [6]).

Theorem 3.10 (Γ-Convergence)**.** *Let $\mathcal{AT}_{\omega,\varepsilon}\colon L^1(I)\times L^1(I) \to [0,+\infty]$ be defined by*

$$\mathcal{AT}_{\omega,\varepsilon}(u,v) := \begin{cases} AT_{\omega,\varepsilon}(u,v) & \text{if } (u,v) \in W^{1,2}(I)\times W^{1,2}(I),\ 0\le v\le 1,\\ +\infty & \text{otherwise.}\end{cases}$$

Then the functionals $\mathcal{AT}_{\omega,\varepsilon}$ Γ-converge, with respect to the $L^1 \times L^1$ topology, to the functional

$$\mathcal{MS}_\omega(u,v) := \begin{cases} MS_\omega(u) & \text{if } u \in GSBV(I) \text{ and } v = 1 \text{ a.e.},\\ +\infty & \text{otherwise.}\end{cases}$$

We begin with an auxiliary proposition.

Proposition 3.11. *Let $\{v_\varepsilon\}_{\varepsilon>0} \subset W^{1,2}(I)$ be such that $0 \le v_\varepsilon \le 1$, $v_\varepsilon \to 1$ in $L^1(I)$ and pointwise a.e., and*

$$\limsup_{\varepsilon\to 0} \int_I \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right] dx < \infty.$$

Then for arbitrary $0<\eta<1$ there exists an open set $H_\eta \subset I$ satisfying:

1. *the set $I \setminus H_\eta$ is a collection of finitely many points in I;*
2. *for every set K compactly contained in H_η, we have $K \subset B_\varepsilon^\eta$ for $\varepsilon > 0$ small enough, where*

$$B_\varepsilon^\eta := \left\{x \in I : v_\varepsilon^2(x) \ge \eta\right\}. \tag{3.13}$$

Proposition 3.11 is adapted from [6], page 1020-1021 (see Lemma A.3).

Proposition 3.12. (Γ-lim inf) *For $u \in L^1_\omega(I)$, let*

$$MS^-_\omega(u) := \inf\Big\{\liminf_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) : \\ (u_\varepsilon, v_\varepsilon) \in W^{1,2}(I)\times W^{1,2}(I), u_\varepsilon \to u \text{ in } L^1,\ v_\varepsilon \to 1 \text{ in } L^1,\ 0 \le v_\varepsilon \le 1\Big\}.$$

We have

$$MS^-_\omega(u) \ge MS_\omega(u).$$

Proof. If $MS^-_\omega(u) = +\infty$ then there is nothing to prove. Assume that $M := MS^-_\omega(u) < \infty$. Choose u_ε and v_ε admissible for $MS^-_\omega(u)$ such that

$$\lim_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) = MS^-_\omega(u) < \infty,$$

and note that $v_\varepsilon \to 1$ in $L^1(I)$. Since $\inf_{x\in\Omega}\omega(x) \geq 1$, we have

$$\liminf_{\varepsilon\to 0} AT_{1,\varepsilon}(u_\varepsilon, v_\varepsilon) \leq \liminf_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) < +\infty,$$

and by [6] we obtain that

$$u \in GSBV(I) \text{ and } \mathcal{H}^0(S_u) < +\infty. \tag{3.14}$$

Let $\bar\varepsilon > 0$ be sufficiently small so that, for all $0 < \varepsilon < \bar\varepsilon$,

$$AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) \leq M + 1.$$

We claim, separately, that

$$\int_I |u'|^2\, \omega\, dx \leq \liminf_{\varepsilon\to 0} \int_I |u'_\varepsilon|^2\, v_\varepsilon^2\, \omega\, dx < +\infty, \tag{3.15}$$

and

$$\sum_{x\in S_u} \omega(x) \leq \liminf_{\varepsilon\to 0} \int_I \left[\frac{1}{2}\varepsilon\, |v'_\varepsilon|^2 + \frac{1}{2\varepsilon}(1-v_\varepsilon)^2\right] \omega\, dx < +\infty. \tag{3.16}$$

Note that (3.15), (3.16), and Lemma 3.9 will yield $u \in GSBV_\omega(I)$.

Up to the extraction of a (not relabeled) subsequence, we have $u_\varepsilon \to u$ and $v_\varepsilon \to 1$ *a.e.* in I with

$$\limsup_{\varepsilon\to 0} \int_I \left[\frac{1}{2}\varepsilon\, |v'_\varepsilon|^2 + \frac{1}{2\varepsilon}(1-v_\varepsilon)^2\right] dx \leq \limsup_{\varepsilon\to 0} \int_I \left[\frac{1}{2}\varepsilon\, |v'_\varepsilon|^2 + \frac{1}{2\varepsilon}(1-v_\varepsilon)^2\right] \omega\, dx < +\infty.$$

Therefore, up to the extraction of a (not relabeled) subsequence, we can apply Proposition 3.11 and deduce that, for a fixed $\eta \in (1/2, 1)$, there exists an open set H_η such that the set $I \setminus H_\eta$ contains only a finite number of points, and for every compact subset $K \subset\subset H_\eta$, K is contained in B^η_ε for $0 < \varepsilon < \varepsilon(K)$, where B^η_ε is defined in (3.13). We have

$$\begin{aligned} &\int_K |u'|^2\, \omega\, dx \leq \liminf_{\varepsilon\to 0} \int_K |u'_\varepsilon|^2\, \omega\, dx \\ &\leq \frac{1}{\eta} \liminf_{\varepsilon\to 0} \int_K v_\varepsilon^2\, |u'_\varepsilon|^2\, \omega\, dx \leq \frac{1}{\eta} \liminf_{\varepsilon\to 0} \int_I v_\varepsilon^2\, |u'_\varepsilon|^2\, \omega\, dx, \end{aligned} \tag{3.17}$$

where we used Lemma 3.8 in the first inequality. By letting $K \nearrow H_\eta$ on the left hand side of (3.17) first and then $\eta \nearrow 1$ on the right hand side, we proved that

$$\int_I |u'|^2\, \omega\, dx \leq \liminf_{\varepsilon\to 0} \int_I v_\varepsilon^2\, |u'_\varepsilon|^2\, \omega\, dx, \tag{3.18}$$

where we used the fact that $|I \setminus H_\eta| = 0$.

We claim that $S_u \subset I \setminus H_\eta$. Indeed, if there is $x_0 \in S_u \cap H_\eta$, since H_η is open there

exists an open interval I_0' containing x_0 and compactly contained in H_η such that for $0<\varepsilon<\varepsilon_0'$
$$\int_{I_0'} \left|u_\varepsilon'\right|^2 dx \le \int_{I_0'} \left|u_\varepsilon'\right|^2 \omega\, dx \le \frac{1}{\eta}\int_I v_\varepsilon^2 \left|u_\varepsilon'\right|^2 \omega\, dx \le 2(M+1).$$
Thus $u \in W^{1,2}(I_0')$, and hence is continuous at x_0, which contradicts the fact that $x_0 \in S_u$.

Let $t \in S_u$, and for simplicity assume that $t=0$. We claim that there exist $\left\{t_n^1\right\}_{n=1}^\infty$, $\left\{t_n^2\right\}_{n=1}^\infty$, and $\{s_n\}_{n=1}^\infty$ such that $-1<t_n^1<s_n<t_n^2<1$,
$$\lim_{n\to\infty} t_n^1 = \lim_{n\to\infty} t_n^2 = \lim_{n\to\infty} s_n = 0,$$
and, up to the extraction of a subsequence of $\{v_\varepsilon\}_{\varepsilon>0}$,
$$\lim_{n\to\infty} v_{\varepsilon(n)}(t_n^1) = \lim_{n\to\infty} v_{\varepsilon(n)}(t_n^2) = 1, \text{ and } \lim_{n\to\infty} v_{\varepsilon(n)}(s_n) = 0. \tag{3.19}$$
Because $I \setminus H_\eta$ is discrete and $0 \in I\setminus H_\eta$, we may choose $\delta_0>0$ small enough such that
$$(-2\delta_0, 2\delta_0) \cap (I\setminus H_\eta) = \{0\}.$$
We claim that
$$\limsup_{\delta\to 0^+} \limsup_{\varepsilon\to 0^+} \inf_{x\in I_\delta} v_\varepsilon(x) = 0, \tag{3.20}$$
where $I_\delta := (-\delta,\delta)$. Assume that
$$\limsup_{\delta\to 0^+} \limsup_{\varepsilon\to 0^+} \inf_{x\in I_\delta} v_\varepsilon(x) =: \alpha > 0.$$
Then there exists $0<\delta_\alpha<\delta_0$ such that
$$\limsup_{\varepsilon\to 0^+} \inf_{x\in I_{\delta_\alpha}} v_\varepsilon(x) \ge \frac{2}{3}\alpha > 0.$$
Up to the extraction of a subsequence of $\{v_\varepsilon\}_{\varepsilon>0}$, there exists $\varepsilon_0^{\delta_\alpha}>0$ such that
$$\inf_{x\in I_{\delta_\alpha}} v_\varepsilon(x) \ge \frac{1}{2}\alpha > 0,$$
for all $0<\varepsilon<\varepsilon_0^{\delta_\alpha}$, and we have
$$\begin{aligned}
&\int_{I_{\delta_\alpha}} \left|u'\right|^2 dx \le \int_{I_{\delta_\alpha}} \left|u'\right|^2 \omega\, dx \\
&\le \liminf_{\varepsilon\to 0} \int_{I_{\delta_\alpha}} \left|u_\varepsilon'\right|^2 \omega\, dx \le \liminf_{\varepsilon\to 0} \frac{2}{\alpha} \int_{I_{\delta_\alpha}} \left|u_\varepsilon'\right|^2 v_\varepsilon^2\, \omega\, dx \\
&\le \liminf_{\varepsilon\to 0} \frac{2}{\alpha} \int_I \left|u_\varepsilon'\right|^2 v_\varepsilon^2\, \omega\, dx < \frac{2}{\alpha}(M+1).
\end{aligned}$$
Hence $u \in W^{1,2}(I_{\delta_\alpha})$ and so u is continuous at $0 \in S_u$, and we reduce a contradiction. Therefore, in view of (3.20) we may find $\delta_n \to 0^+$, $\varepsilon(n)\to 0^+$, and $s_n \in (-\delta_n, \delta_n)$ such that
$$\lim_{n\to\infty} s_n = 0 \text{ and } \lim_{n\to\infty} v_{\varepsilon(n)}(s_n) = 0.$$

We claim that for all $\tau \in (0, 1/2)$,

$$\lim_{n\to\infty} \left[\inf_{x\in(s_n-\tau,s_n)} (1 - v_{\varepsilon(n)}(x)) + \inf_{y\in(s_n,s_n+\tau)} (1 - v_{\varepsilon(n)}(x)) \right] = 0. \tag{3.21}$$

To reach a contradiction, assume that there exists $\tau \in (0, 1/2)$ such that

$$\limsup_{n\to\infty} \left[\inf_{x\in(s_n-\tau,s_n)} (1 - v_{\varepsilon(n)}(x)) + \inf_{x\in(s_n,s_n+\tau)} (1 - v_{\varepsilon(n)}(x)) \right] =: \beta > 0.$$

Without loss of generality, suppose that

$$\limsup_{n\to\infty} \inf_{x\in(s_n-\tau,s_n)} (1 - v_{\varepsilon(n)}(x)) \geq \frac{1}{2}\beta > 0.$$

Then

$$\liminf_{n\to\infty} \sup_{x\in(s_n-\tau,s_n)} v_{\varepsilon(n)}(x) \leq 1 - \frac{1}{2}\beta,$$

which implies that

$$\sup_{x\in(s_{n_k}-\tau,s_{n_k})} v_{\varepsilon(n_k)}(x) \leq 1 - \frac{1}{3}\beta \tag{3.22}$$

for a subsequence $\{\varepsilon(n_k)\}_{k=1}^{\infty} \subset \{\varepsilon(n)\}_{n=1}^{\infty}$. However, (3.22) contradicts the fact that $v_{\varepsilon(n_k)}(x) \to 1$ a.e. since for k large enough so that $|s_{n_k}| < \tau/4$ it holds

$$(s_{n_k} - \tau, s_{n_k}) \supset \left(-\frac{3}{4}\tau, -\frac{\tau}{4} \right).$$

Therefore, in view of (3.21) we may find $t_m^1 \in (s_{n(m)} - 1/m, s_{n(m)})$ and $t_m^2 \in (s_{n(m)}, s_{n(m)} + 1/m)$ such that

$$\lim_{n\to\infty} t_m^1 = \lim_{n\to\infty} t_m^2 = 0 \text{ and } \lim_{n\to\infty} v_{\varepsilon(n(m))}(t_m^1) = \lim_{n\to\infty} v_{\varepsilon(n(m))}(t_m^2) = 1.$$

We next show that

$$\liminf_{m\to\infty} \int_{t_m^1}^{s_{n(m)}} \left[\frac{1}{2}\varepsilon(n(m)) \left| (v_{\varepsilon(n(m))})' \right|^2 + \frac{1}{2\varepsilon(n(m))} (1 - v_{\varepsilon(n(m))})^2 \right] dx \geq \frac{1}{2}.$$

Indeed, we have

$$\begin{aligned}
&\liminf_{m\to\infty} \int_{t_m^1}^{s_{n(m)}} \left[\frac{1}{2}\varepsilon(n(m)) \left| (v_{\varepsilon(n)})' \right|^2 + \frac{1}{2\varepsilon(n(m))} (1 - v_{\varepsilon(n)})^2 \right] dx \\
&\geq \liminf_{m\to\infty} \int_{t_m^1}^{s_{n(m)}} (1 - v_{\varepsilon(n)}) \left| v'_{\varepsilon(n)} \right| dx \geq \liminf_{m\to\infty} \left| \int_{t_m^1}^{s_{n(m)}} (1 - v_{\varepsilon(n)}) v'_{\varepsilon(n)} dx \right| \\
&= \liminf_{m\to\infty} \frac{1}{2} \left| \int_{t_m^1}^{s_{n(m)}} \frac{d}{dt} (1 - v_{\varepsilon(n)})^2 dx \right| \\
&= \frac{1}{2} \lim_{n\to\infty} \left[(1 - v_{\varepsilon(n(m))}(s_{n(m)}))^2 - (1 - v_{\varepsilon(n(m))}(t_m^1))^2 \right] = \frac{1}{2},
\end{aligned}$$

where we used (3.19). Similarly, we obtain

$$\liminf_{m\to\infty}\int_{s_{n(m)}}^{t_m^2}\left[\frac{1}{2}\varepsilon(n(m))\left|(v_{\varepsilon(n)})'\right|^2+\frac{1}{2\varepsilon(n)}(1-v_{\varepsilon(n)})^2\right]dx\geq\frac{1}{2}.$$

We observe that, since ω is positive,

$$\begin{aligned}&\int_{t_m^1}^{t_m^2}\left[\frac{1}{2}\varepsilon(n)\left|v'_{\varepsilon(n)}\right|^2+\frac{1}{2\varepsilon(n)}(1-v_{\varepsilon(n)})^2\right]\omega(x)\,dx\\&\geq\left(\inf_{r\in(t_m^1,t_m^2)}\omega(r)\right)\cdot\left\{\int_{t_m^1}^{s_{n(m)}}\left[\frac{1}{2}\varepsilon(n)\left|v'_{\varepsilon(n)}\right|^2+\frac{1}{2\varepsilon(n)}(1-v_{\varepsilon(n)})^2\right]dx\right.\\&\quad\left.+\int_{s_{n(m)}}^{t_m^2}\left[\frac{1}{2}\varepsilon(n)\left|(v_{\varepsilon(n)})'\right|^2+\frac{1}{2\varepsilon(n)}(1-v_{\varepsilon(n)})^2\right]dx\right\},\end{aligned}\tag{3.23}$$

and so

$$\begin{aligned}&\liminf_{m\to\infty}\int_{t_m^1}^{t_m^2}\left[\frac{1}{2}\varepsilon(n)\left|v'_{\varepsilon(n)}\right|^2+\frac{1}{2\varepsilon(n)}(1-v_{\varepsilon(n)})^2\right]\omega(x)\,dx\\&\geq\left(\liminf_{m\to\infty}\inf_{r\in(t_m^1,t_m^2)}\omega(r)\right)\liminf_{n\to\infty}\left\{\int_{t_m^1}^{s_{n(m)}}\left[\frac{1}{2\varepsilon(n)}(1-v_{\varepsilon(n)})^2+\frac{\varepsilon}{2}\left|(v_{\varepsilon(n)})'\right|^2\right]dx\right.\\&\quad\left.+\int_{s_{n(m)}}^{t_m^2}\left[\frac{1}{2}\varepsilon(n)\left|v'_{\varepsilon(n)}\right|^2+\frac{1}{2\varepsilon(n)}(1-v_{\varepsilon(n)})^2\right]dx\right\}\\&\geq\left(\frac{1}{2}+\frac{1}{2}\right)\omega(0)=\omega(0),\end{aligned}$$

where we used the fact that ω is continuous at 0.

Finally, since $S_u\subset I\setminus H_\eta$, by (3.14) we have that S_u is a finite collection of points, and we may repeat the above argument for all $t\in S_u$ by partitioning I into non-overlaping intervals where there is at most one point of S_u, to deduce that

$$\liminf_{\varepsilon\to0}\int_I\left[\frac{1}{2}\varepsilon\left|v'_\varepsilon\right|^2+\frac{1}{2\varepsilon}(1-v_\varepsilon)^2\right]\omega(x)\,dx\geq\sum_{x\in S_u}\omega(x).\tag{3.24}$$

In view of (3.18) and (3.24), we conclude that

$$\liminf_{\varepsilon\to0}AT_{\omega,\varepsilon}(u_\varepsilon,v_\varepsilon)\geq MS_\omega(u).$$

□

Proposition 3.13. (Γ-lim sup) *For $u\in L^1(I)\cap L^\infty(I)$, let*

$$\begin{aligned}MS^+_\omega(u):=\inf\Big\{&\limsup_{\varepsilon\to0}AT_{\omega,\varepsilon}(u_\varepsilon,v_\varepsilon):\\&(u_\varepsilon,v_\varepsilon)\in W^{1,2}(I)\times W^{1,2}(I),u_\varepsilon\to u\text{ in }L^1,\,v_\varepsilon\to1\text{ in }L^1,\,0\leq v_\varepsilon\leq1\Big\}.\end{aligned}$$

We have

$$MS_\omega^+(u) \leq MS_\omega(u). \tag{3.25}$$

Proof. Without loss of generality, assume that $MS_\omega(u) < \infty$. Then by Lemma 3.9 we have $u \in GSBV_\omega(I)$ and $\mathcal{H}^0(S_u) < \infty$. To prove (3.25), we show that there exist $\{u_\varepsilon\}_{\varepsilon>0} \subset W_\omega^{1,2}(I)$ and $\{v_\varepsilon\}_{\varepsilon>0} \subset W^{1,2}(I)$ such that $u_\varepsilon \to u$ in L_ω^1, $v_\varepsilon \to 1$ in L^1, $0 \leq v_\varepsilon \leq 1$, and

$$\limsup_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) \leq MS_\omega(u). \tag{3.26}$$

<u>Step 1</u>: Assume that $S_u = \{0\}$.

Fix $\eta > 0$, and let $T > 0$ and $v_0 \in W^{1,2}(0,T)$ be such that

$$0 \leq v_0 \leq 1 \quad \text{and} \quad \int_0^T \left[(1-v_0)^2 + \left|v_0'\right|^2\right] dx \leq 1+\eta, \tag{3.27}$$

with $v_0(0) = 0$ and $v_0(T) = 1$.

For $\xi_\varepsilon = o(\varepsilon)$ we define

$$v_\varepsilon(x) := \begin{cases} 0 & \text{if } |x| \leq \xi_\varepsilon, \\ v_0\left(\frac{|x|-\xi_\varepsilon}{\varepsilon}\right) & \text{if } \xi_\varepsilon < |x| < \xi_\varepsilon + \varepsilon T, \\ 1 & \text{if } |x| \geq \xi_\varepsilon + \varepsilon T. \end{cases} \tag{3.28}$$

Since $\|v_\varepsilon\|_{L^\infty(I)} \leq 1$, by Lebesgue Dominated Convergence Theorem we have $v_\varepsilon \to 1$ in L^1. Let

$$u_\varepsilon(x) := \begin{cases} u(x) & \text{if } |x| \geq \frac{1}{2}\xi_\varepsilon, \\ \text{affine from } u\left(-\frac{1}{2}\xi_\varepsilon\right) \text{ to } u\left(\frac{1}{2}\xi_\varepsilon\right) & \text{if } |x| < \frac{1}{2}\xi_\varepsilon. \end{cases} \tag{3.29}$$

and we observe that (recall in assumption we have $u \in L^\infty(I)$)

$$\|u_\varepsilon\|_{L^\infty(I)} \leq \|u\|_{L^\infty(I)},$$

and

$$\int_I \|u\|_{L^\infty(I)} \,\omega\, dx < \infty.$$

Therefore, by Lebesgue Dominated Convergence Theorem we deduce that $u_\varepsilon \to u$ in L_ω^1. Moreover, by (3.28) and (3.29) we observe that

$$v_\varepsilon^2 \left|u_\varepsilon'\right|^2 = \begin{cases} v_\varepsilon^2 \left|u'\right|^2 & \text{if } x \geq |\xi_\varepsilon|, \\ 0 & \text{if } x < |\xi_\varepsilon|, \end{cases}$$

and so $v_\varepsilon^2 \left|u_\varepsilon'\right|^2 \leq \left|u'\right|^2$. Since $MS_\omega(u) < \infty$ we have $u' \in L_\omega^2(I)$, by Lebesgue Dominated Convergence Theorem we obtain

$$\lim_{\varepsilon\to 0} \int_I v_\varepsilon^2 \left|u_\varepsilon'\right|^2 \omega\, dx = \int_I \left|u'\right|^2 \omega\, dx.$$

Next, since ω is positive we have

$$
\begin{aligned}
&\int_I \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right]\omega(x)\,dx \\
=&\int_{-\xi_\varepsilon-\varepsilon T}^{-\xi_\varepsilon} \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right]\omega(x)\,dx \\
&+\int_{\xi_\varepsilon}^{\xi_\varepsilon+\varepsilon T} \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right]\omega(x)\,dx + \frac{1}{2\varepsilon}\int_{-\xi_\varepsilon}^{\xi_\varepsilon}\omega(x)dx \\
\leq&\left(\sup_{t\in(-\xi_\varepsilon-\varepsilon T,\xi_\varepsilon+\varepsilon T)}\omega(t)\right)\cdot\left\{\int_{-\xi_\varepsilon-\varepsilon T}^{-\xi_\varepsilon} \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right]\,dx\right. \\
&\left.+\int_{\xi_\varepsilon}^{\xi_\varepsilon+\varepsilon T} \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right]\,dx\right\} + \frac{\xi_\varepsilon}{\varepsilon}\left\|\omega\right\|_{L^\infty}.
\end{aligned}
$$

We obtain

$$
\begin{aligned}
&\limsup_{\varepsilon\to 0}\int_I \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right]\omega(x)\,dx \\
\leq&\limsup_{\varepsilon\to 0}\left(\sup_{t\in(-\xi_\varepsilon-\varepsilon T,\xi_\varepsilon+\varepsilon T)}\omega(t)\right)\cdot \\
&\limsup_{\varepsilon\to 0}\left\{\int_{-\xi_\varepsilon-\varepsilon T}^{-\xi_\varepsilon} \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right]\,dx + \int_{\xi_\varepsilon}^{\xi_\varepsilon+\varepsilon T} \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right]\,dx\right\} \\
\leq&\,\omega(0)(1+\eta),
\end{aligned}
$$

where we used (3.27).

We conclude that

$$\limsup_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) \leq \int_I \left|u'\right|^2\omega\,dx + \omega(0)(1+\eta),$$

and (3.26) follows by the arbitrariness of η.

Step 2: In the general case in which S_u is finite, we obtain u_ε by repeating the construction in Step 1 (see (3.29)) in small non-overlapping intervals centered at each point in S_u. To obtain v_ε, we repeat the construction (3.28) in those intervals and extend by 1 in the complement of the union of those intervals. Hence, by Step 1 we have

$$\limsup_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) \leq \int_I \left|u'\right|^2\omega\,dx + (1+\eta)\sum_{x\in S_u}\omega(x),$$

and again (3.26) follows by letting $\eta \to 0^+$. □

Proof of Theorem 3.10. The lim inf inequality follows from Proposition 3.12. For the lim sup inequality, we note that for any given $u \in GSBV_\omega$ such that $MS_\omega(u) < +\infty$, by Lebesgue

Monotone Convergence Theorem we have that

$$MS_\omega(u) = \lim_{K\to\infty} MS_\omega(K \wedge u \vee -K),$$

and hence a diagonal argument together with Proposition 3.13 conclude the proof. $\square$

3.2.2. *The Case $\omega \in \mathcal{W}(I) \cap SBV(I)$.*

Consider the functionals

$$AT_{\omega,\varepsilon}(u,v) := \int_I |u'|^2 v^2 \omega\, dx + \int_I \left[\frac{\varepsilon}{2}|v'|^2 + \frac{1}{2\varepsilon}(v-1)^2\right]\omega\, dx$$

for $(u,v) \in W^{1,2}(I) \times W^{1,2}(I)$, and for $u \in GSBV_\omega(I)$ let

$$MS_\omega(u) := \int_I |u'|^2 \omega\, dx + \sum_{x\in S_u} \omega^-(x).$$

We note that if $\omega \in \mathcal{W}(I) \cap SBV(I)$ and ω is continuous in a neighborhood of S_u, for $u \in GSBV_\omega(I)$, then

$$\sum_{x\in S_u} \omega^-(x) = \sum_{x\in S_u} \omega(x)$$

and Theorem 3.10 still holds.

Here we study the case in which ω is no longer continuous on a neighborhood of S_u. We recall that $\omega \in SBV(I)$ implies that $\omega \in L^\infty(I)$ and by definition of $\omega \in \mathcal{W}(I)$, we have $\mathcal{H}^0(S_\omega) < \infty$. Also, we note that ω^- is defined $\mathcal{H}^0$-a.e, hence everywhere in I.

Theorem 3.14. *Let $\mathcal{MS}_\varepsilon\colon L^1(I) \times L^1(I) \to [0,+\infty]$ be defined by*

$$\mathcal{AT}_{\omega,\varepsilon}(u,v) := \begin{cases} AT_{\omega,\varepsilon}(u,v) & \text{if } (u,v) \in W^{1,2}(I)\times W^{1,2}(I),\ 0\le v\le 1,\\ +\infty & \text{otherwise.}\end{cases}$$

Then the functionals $\mathcal{AT}_{\omega,\varepsilon}$ Γ-converge, with respect to the $L^1 \times L^1$ topology, to the functional

$$\mathcal{MS}_\omega(u,v) := \begin{cases} MS_\omega(u) & \text{if } u \in GSBV(I) \text{ and } v = 1 \text{ a.e.},\\ +\infty & \text{otherwise.}\end{cases}$$

The proof of Theorem 3.14 will be split into two propositions.

Proposition 3.15. (Γ-lim inf) *For $u \in L^1(I)$, let*

$$MS_\omega^-(u) := \inf\Big\{\liminf_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon,v_\varepsilon) : \\ (u_\varepsilon,v_\varepsilon) \in W^{1,2}(I)\times W^{1,2}(I),\ u_\varepsilon \to u \text{ in } L^1, v_\varepsilon \to 1 \text{ in } L^1,\ 0 \le v_\varepsilon \le 1\Big\}.$$

We have

$$MS_\omega^-(u) \ge MS_\omega(u)$$

Proof. Without lose of generality, assume that $MS^-_\omega(u) < +\infty$. We use the same arguments of the proof of Proposition 3.12 until (3.23). In particular, (3.14) and (3.15) still hold, that is

$$\mathcal{H}^0(S_u) < +\infty \text{ and } \int_I |u'|^2 \omega\, dx \le \liminf_{\varepsilon\to 0} \int_I |u'_\varepsilon|^2 v_\varepsilon^2 \omega\, dx.$$

Invoking (3.23), we have

$$\begin{aligned}
&\liminf_{m\to\infty} \int_{t^1_m}^{t^2_m} \left[\frac{1}{2}\varepsilon(n)\left|v'_{\varepsilon(n)}\right|^2 + \frac{1}{2\varepsilon(n)}(1-v_{\varepsilon(n)})^2\right]\omega(x)\,dx \\
&\ge \left(\liminf_{m\to\infty} \operatorname*{ess\,inf}_{r\in(t^1_m,t^2_m)} \omega(r)\right)\cdot \liminf_{n\to\infty}\left\{\int_{t^1_m}^{s_{n(m)}} \left[\frac{1}{2}\varepsilon(n)\left|(v_{\varepsilon(n)})'\right|^2 + \frac{1}{2\varepsilon(n)}(1-v_{\varepsilon(n)})^2\right]dx\right. \\
&\quad \left. + \int_{s_{n(m)}}^{t^2_m} \left[\frac{1}{2}\varepsilon(n)\left|(v_{\varepsilon(n)})'\right|^2 + \frac{1}{2\varepsilon(n)}(1-v_{\varepsilon(n)})^2\right]dx\right\} \\
&\ge \omega^-(0)\left(\frac{1}{2}+\frac{1}{2}\right) = \omega^-(0),
\end{aligned}$$

where the last step is justified by (3.4).

Since S_u is finite, we may repeat the above argument for all $t \in S_u$ by partitioning I into finitely many non-overlapping intervals where there is at most one point of S_u, to conclude that

$$\liminf_{\varepsilon\to 0} \int_I \left[\frac{1}{2}\varepsilon\left|v'_\varepsilon\right|^2 + \frac{1}{2\varepsilon}(1-v_\varepsilon)^2\right]\omega(x)\,dx \ge \sum_{x\in S_u} \omega^-(x),$$

as desired. □

The construction of the recovery sequence uses a reflection argument nearby points of $S_\omega \cap S_u$.

Proposition 3.16. (Γ-lim sup) *For $u \in L^1(I)\cap L^\infty(I)$, let*

$$\begin{aligned}
MS^+_\omega(u) := \inf \Big\{ &\limsup_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) : \\
&(u_\varepsilon, v_\varepsilon) \in W^{1,2}(I)\times W^{1,2}(I), u_\varepsilon \to u \text{ in } L^1,\ v_\varepsilon \to 1 \text{ in } L^1,\ 0\le v_\varepsilon \le 1\Big\}.
\end{aligned}$$

We have

$$MS^+_\omega(u) \le MS_\omega(u). \tag{3.30}$$

Proof. To prove (3.30), we only need to explicitly construct a sequence $\{(u_\varepsilon, v_\varepsilon)\}_{\varepsilon>0} \subset W^{1,2}(I)\times W^{1,2}(I)$ such that $u_\varepsilon \to u$ in L^1, $v_\varepsilon \to 1$ in L^1, $0 \le v_\varepsilon \le 1$, and

$$\limsup_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) \le MS_\omega(u). \tag{3.31}$$

Step 1: Assume that $\{0\} = S_u \subset S_\omega$.

Recall that we always identify ω with its approximation representative $\bar{\omega}$, and by (3.5) we may assume that (the converse situation may be dealt with similarly)

$$\lim_{t \nearrow 0^-} \omega(t) = \omega^-(0) \text{ and } \lim_{t \searrow 0^+} \omega(t) = \omega^+(0).$$

Fix $\eta > 0$. For $\varepsilon > 0$ small enough, and with $\xi_\varepsilon = o(\varepsilon)$, as in (3.27), (3.28) let

$$\tilde{v}_\varepsilon(x) := \begin{cases} 0 & \text{if } |x| \leq \xi_\varepsilon \\ v_0\left(\frac{|x|-\xi_\varepsilon}{\varepsilon}\right) & \text{if } \xi_\varepsilon < |x| < \xi_\varepsilon + \varepsilon T \\ 1 & \text{if } |x| \geq \xi_\varepsilon + \varepsilon T, \end{cases}$$

and define

$$v_\varepsilon(x) := \tilde{v}_\varepsilon(x + 2\xi_\varepsilon + \varepsilon T).$$

Note that from (3.28) $v_\varepsilon \to 1$ a.e., and since $0 \leq v_\varepsilon \leq 1$, by Lebesgue Dominated Convergence Theorem we have $v_\varepsilon \to v$ in L^1. We also note that

$$\frac{\varepsilon}{2}\left|v_\varepsilon'(x)\right|^2 + \frac{1}{2\varepsilon}(1 - v_\varepsilon(x))^2 = 0 \tag{3.32}$$

if $x \in (-1, -3\xi_\varepsilon - 2\varepsilon T) \cup (-\xi_\varepsilon, 1)$, and if $x \in (-3\xi_\varepsilon - \varepsilon T, -\xi_\varepsilon - \varepsilon T)$ then

$$v_\varepsilon(x) = 0. \tag{3.33}$$

Set

$$\tilde{u}_\varepsilon(x) := \begin{cases} u(x) & \text{if } x \in (-1, -2\xi_\varepsilon - \varepsilon T) \cup (0, 1), \\ u(-x) & \text{if } x \in [-2\xi_\varepsilon - \varepsilon T, 0]. \end{cases}$$

Observe that $\tilde{u}_\varepsilon(x)$ is continuous at 0 since $\tilde{u}_\varepsilon^+(0) = \tilde{u}_\varepsilon^-(0) = u^+(0)$ by the definition of $\tilde{u}_\varepsilon(x)$, and $\tilde{u}_\varepsilon$ may only jump at $t = -2\xi_\varepsilon - \varepsilon T$ but not at $t = 0$ where u jumps.

We define the recovery sequence

$$u_\varepsilon(x) :=$$

$$\begin{cases} \tilde{u}_\varepsilon(x) & \text{if } x \in I \setminus [-2.5\xi_\varepsilon - \varepsilon T, -1.5\xi_\varepsilon - \varepsilon T], \\ \text{affine from } \tilde{u}_\varepsilon(-2.5\xi_\varepsilon - \varepsilon T) \text{ to } \tilde{u}_\varepsilon(-1.5\xi_\varepsilon - \varepsilon T) & \text{if } x \in [-2.5\xi_\varepsilon - \varepsilon T, -1.5\xi_\varepsilon - \varepsilon T]. \end{cases}$$

We claim that

$$\lim_{\varepsilon \to 0} \int_I |u_\varepsilon - u|\, \omega\, dx = 0 \tag{3.34}$$

and

$$\limsup_{\varepsilon \to 0} \int_I \left|u_\varepsilon'\right|^2 v_\varepsilon^2\, \omega\, dx \leq \int_I \left|u'\right|^2 \omega\, dx. \tag{3.35}$$

To show (3.34), we observe that

$$\lim_{\varepsilon\to 0}\int_I |u_\varepsilon - u|\,\omega\,dx \le \lim_{\varepsilon\to 0}\int_{-2.5\xi_\varepsilon-\varepsilon T}^{0} |u_\varepsilon - u|\,\omega\,dx \le \lim_{\varepsilon\to 0} 2\,\|u\|_{L^\infty}\,\|\omega\|_{L^\infty}\,(2.5\xi_\varepsilon+\varepsilon T) = 0.$$

We next prove (3.35). By (3.32) we have

$$\int_I \left|u_\varepsilon'\right|^2 v_\varepsilon^2\,\omega\,dx \le \int_I \left|u'\right|^2\omega\,dx + \|\omega\|_{L^\infty}\int_{-\xi_\varepsilon-\varepsilon T}^{0}\left|u'(-x)\right|^2 dx,$$

and so

$$\limsup_{\varepsilon\to 0}\int_I \left|u_\varepsilon'\right|^2 v_\varepsilon^2\,\omega\,dx \le \int_I \left|u'\right|^2\omega\,dx,$$

since $u' \in L^2(I)$, and we conclude that $u' \in L^2(I)$.

On the other hand, by (3.32) and (3.33),

$$\begin{aligned}
&\int_I \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon-1)^2\right]\omega(x)\,dx\\
&= \int_{-3\xi_\varepsilon-2\varepsilon T}^{-\xi_\varepsilon}\left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon-1)^2\right]\omega(x)\,dx\\
&\le \left(\operatorname*{ess\,sup}_{t\in(-3\xi_\varepsilon-2\varepsilon T,-\xi_\varepsilon)}\omega(t)\right)\int_{-3\xi_\varepsilon-2\varepsilon T}^{-\xi_\varepsilon}\left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon-1)^2\right]dx\\
&= \left(\operatorname*{ess\,sup}_{t\in(-3\xi_\varepsilon-2\varepsilon T,-\xi_\varepsilon)}\omega(t)\right)\int_{-\xi_\varepsilon-\varepsilon T}^{\xi_\varepsilon+\varepsilon T}\left[\frac{\varepsilon}{2}\left|\tilde v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(\tilde v_\varepsilon-1)^2\right]dx.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\limsup_{\varepsilon\to 0}\int_I \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon-1)^2\right]\omega(x)\,dx\\
&\le \limsup_{\varepsilon\to 0}\left(\operatorname*{ess\,sup}_{t\in(-3\xi_\varepsilon-2\varepsilon T,-\xi_\varepsilon)}\omega(t)\right)\left\{\limsup_{\varepsilon\to 0}\int_{-\xi_\varepsilon-\varepsilon T}^{\xi_\varepsilon+\varepsilon T}\left[\frac{\varepsilon}{2}\left|\tilde v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(\tilde v_\varepsilon-1)^2\right]dx\right\}\\
&\le \omega^-(0)(1+\eta),
\end{aligned}$$

where at the last inequality we used the definition of $\tilde v_\varepsilon$, (3.27), and (3.5).

We conclude that

$$\limsup_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) \le \int_I \left|u'\right|^2\omega\,dx + \omega^-(0)(1+\eta),$$

and (3.31) follows due to the arbitrariness of η.

Step 2: In the general case, we recall that S_u is finite. We may obtain u_ε and v_ε by repeating the construction in Step 1 in small non-overlapping intervals centered at every point of $S_u \cap S_\omega$, and by repeating the construction in Step 1 in Lemma 3.13 in those non-overlaping intervals centered at points of $S_u \setminus S_\omega$. Hence, we have

$$\limsup_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) \le \int_I \left|u'\right|^2 \omega\, dx + (1+\eta) \sum_{x\in S_u} \omega^-(x),$$

and (3.31) follows due to the arbitrariness of η. □

Proof of Theorem 3.14. The proof follows that of Theorem 3.10, using Proposition 3.15 and Proposition 3.16, in place of Proposition 3.12 and 3.13, respectively. □

3.3. **The multi-dimensional case.**

3.3.1. *One-dimensional restrictions and slicing properties.*

Let $\mathcal{S}^{N-1}$ be the unit sphere in $\mathbb{R}^N$ and let $\nu \in \mathcal{S}^{N-1}$ be a fixed direction. We set

$$\begin{cases} \Pi_\nu := \left\{x \in \mathbb{R}^N : \langle x, \nu\rangle = 0\right\}, \ \Omega_\nu := \left\{x \in \Pi_\nu : \Omega_{x,\nu} \neq \emptyset\right\}, \\ \Omega^1_{x,\nu} := \left\{t \in \mathbb{R} : x + t\nu \in \Omega\right\} \quad \text{for } x \in \Pi_\nu, \\ \Omega_{x,\nu} := \left\{y = x + t\nu : \ t \in \mathbb{R}\right\} \cap \Omega, \\ u_{x,\nu}(t) := u(x+t\nu), \ x \in \Omega_\nu, \ t \in \Omega^1_{x,\nu}. \end{cases} \tag{3.36}$$

Set $x = (x', x_N) \in \mathbb{R}^N$, where $x' \in \mathbb{R}^{N-1}$ denotes the first $N-1$ component of $x \in \mathbb{R}^N$, and given ϕ: $\mathbb{R}^{N-1} \to \mathbb{R}$ and $G \subset \mathbb{R}^{N-1}$, we define the graph of ϕ over G as

$$F(\phi; G) := \left\{(x', x_N) \in \mathbb{R}^N : x' \in G, \, x_N = l(x')\right\}.$$

If ϕ is Lipschitz, then we call $F(\phi; G)$ a Lipschitz - $(N-1)$ - graph.

Theorem 3.17 ([6], Theorem 3.3)**.** *Let $\nu \in \mathcal{S}^{N-1}$ be given, and assume that $u \in W^{1,2}(\Omega)$. Then, for $\mathcal{H}^{N-1}$-a.e. $x \in \Omega_\nu$, $u_{x,\nu}$ belongs to $W^{1,2}(\Omega_{x,\nu})$ and*

$$u'_{x,\nu}(t) = \langle \nabla u(x+t\nu), \nu\rangle .$$

Lemma 3.18. *Let $\omega \in \mathcal{W}(\Omega)$ and ϕ: $\mathbb{R}^{N-1} \to \mathbb{R}$ be a Lipschitz function. Then for every Lebesgure measurable set $G \subset \mathbb{R}^{N-1}$ such that $F(\phi;\, G) \subset \Omega$, we have*

$$\int_{F(\phi:\, G)} \omega^- d\mathcal{H}^{N-1} = \int_G \omega^-(z, \phi(z)) \sqrt{1 + |\nabla\phi(z)|^2} dz.$$

Proof. Since $\mathcal{H}^{N-1}(S_\omega) < \infty$, we have that $\mu := \mathcal{H}^{N-1}\lfloor S_\omega$ is a nonnegative radon measure. Moreover, as $\omega \in L^\infty(\Omega)$ we deduce that $\omega^- \in L^1(\Omega, \mu)$, and the result now follows from by Remark 8.3 in [51], where the Jacobian of $F(\phi : G)$ is shown to be equal to $\sqrt{1 + |\nabla\phi(z)|^2}$ in Theorem 9.1 in [51]. □

Lemma 3.19. *Let $\omega \in \mathcal{W}(\Omega)$ and $u \in W^{1,p}_\omega(\Omega)$, for $p \in [1, \infty)$, be given. If $\nu \in \mathcal{S}^{N-1}$ and $v \in W^{1,p}(\Omega)$ is nonnegative, then*

$$\int_\Omega |\nabla u|^p \, v^p \, \omega \, dx \ge \int_{\Omega_\nu} \int_{\Omega^1_{x,\nu}} \left|u'_{x,\nu}(t)\right|^p v^p_{x,\nu}(t) \, \omega_{x,\nu}(t) \, dt dx.$$

Proof. Since $\operatorname{ess\,inf}_\Omega \omega \geq 1$, we have $W^{1,p}_\omega(\Omega) \subset W^{1,p}(\Omega)$. Given $\nu \in \mathcal{S}^{N-1}$ and a nonnegative function $v \in W^{1,p}(\Omega)$, by Fubini's Theorem and Theorem 3.17 we have

$$\begin{aligned}\int_\Omega |\nabla u|^p\, v^p\, \omega\, dx &= \int_{\Omega_\nu} \int_{\Omega^1_{x,\nu}} |\nabla u|^p\, v^p\, \omega\, dt\, d\mathcal{H}^{N-1}(x) \\ &\geq \int_{\Omega_\nu} \int_{\Omega^1_{x,\nu}} |\langle \nabla u(x+tv), \nu\rangle|^p\, v^p_{x,\nu}(t)\, \omega_{x,\nu}(t)\, dt d\mathcal{H}^{N-1}(x) \\ &= \int_{\Omega_\nu} \int_{\Omega^1_{x,\nu}} \left|u'_{x,\nu}(t)\right|^p v^p_{x,\nu}(t)\, \omega_{x,\nu}(t)\, dt d\mathcal{H}^{N-1}(x),\end{aligned}$$

where we used the fact that

$$\left|u'_{x,\nu}(t)\right| = |\langle \nabla u(x+t\nu), \nu\rangle| \leq |\nabla u(x+t\nu)|$$

$\mathcal{H}^{N-1}$-a.e. $x \in \Omega_\nu$. □

Proposition 3.20. *Let $\nu \in \mathcal{S}^{N-1}$ be a fixed direction, $\Gamma \subset \mathbb{R}^N$ be such that $\mathcal{H}^{N-1}(\Gamma) < \infty$, and $\mathbb{P}_\nu$: $\mathbb{R}^N \to \Pi_\nu$ be a projection operator, where by* (3.36) *$\Pi_\nu \subset \mathbb{R}^N$ is a hyperplane in $\mathbb{R}^{N-1}$. Then*

$$\mathcal{H}^{N-1}(\mathbb{P}_\nu(\Gamma)) \leq \mathcal{H}^{N-1}(\Gamma), \tag{3.37}$$

and for $\mathcal{H}^{N-1}$-a.e. $x \in \Pi_\nu$,

$$\mathcal{H}^0(\Omega_{x,\nu} \cap \Gamma) < +\infty. \tag{3.38}$$

Proof. Note that (3.37) follows immediately from Theorem 7.5 in [63] since $\mathbb{P}_\nu$ is a Lipschitz map with Lipschitz constant less or equal to one. To show (3.38), we apply co-area formula (see [4], Theorem 2.93) with $\mathbb{P}_\nu$ and again since $\mathbb{P}_\nu$ is a Lipschitz map with Lipschitz constant less or equal to one, we are done. □

Set $x = (x', x_N) \in \mathbb{R}^N$, where

$$x' \in \mathbb{R}^{N-1} \text{ denotes the first } N-1 \text{ component of } x \in \mathbb{R}^N,$$

and given u: $\mathbb{R}^{N-1} \to \mathbb{R}$ and $G \subset \mathbb{R}^{N-1}$, we define the graph of u over G as

$$F(u; G) := \left\{(x', x_N) \in \mathbb{R}^N : x' \in G,\, x_N = u(x')\right\}.$$

If u is Lipschitz, then we call $F(u; G)$ a Lipschitz -$(N-1)$-graph.

Lemma 3.21. *Let $\Gamma \subset \mathbb{R}^N$ be a $\mathcal{H}^{N-1}$-rectifiable set, and let $\mathbb{P}_{x,\nu_\Gamma}$: $\mathbb{R}^N \to T_{x,\nu_\Gamma}$ be a projection operator for $x \in \Gamma$. Then*

$$\lim_{r\to 0} \frac{\mathcal{H}^{N-1}(\mathbb{P}_{x_0,\nu_\Gamma}(\Gamma \cap Q_{\nu_\Gamma}(x_0, r)))}{r^{N-1}} = 1 \tag{3.39}$$

for $\mathcal{H}^{N-1}$-a.e. $x_0 \in \Gamma$.

Proof. By Proposition 3.20 we have

$$\limsup_{r\to 0} \frac{\mathcal{H}^{N-1}(\mathbb{P}_{x_0,\nu_\Gamma}(\Gamma \cap Q_{\nu_\Gamma}(x_0, r)))}{r^{N-1}} \leq \limsup_{r\to 0} \frac{\mathcal{H}^{N-1}(\Gamma \cap Q_{\nu_\Gamma}(x_0, r))}{r^{N-1}} = 1 \tag{3.40}$$

for a.e. $x_0 \in \Gamma$. By Theorem 2.76 in [4] we may write

$$\Gamma = \Gamma_0 \cup \bigcup_{i=1}^{\infty} \Gamma_i$$

as a disjoint union with $\mathcal{H}^{N-1}(\Gamma_0) = 0$, $\Gamma_i = (N_i, l_i(N_i))$ where $l_i : \mathbb{R}^{N-1} \to \mathbb{R}$ is of class C^1 and $N_i \subset \mathbb{R}^{N-1}$.

Let $x_0 \in \Gamma_{i_0}$ for some $i_0 \in \mathbb{N}$ and, without loss of generality, let $(-\nabla l_{i_0}(x_0'), 1) = \nu_\Gamma(x_0)$, with x_0 a point of density one in Γ_0 (see Exercise 10.6 in [51]). Up to a rotation and a translation, we may assume that $\nabla l_{i_0}(x_0') = (0, 0, \dots, 0) \in \mathbb{R}^{N-1}$, $x_0 = (0, 0, \dots, 0)$, and $\mathbb{P}_{x_0,\nu_\Gamma}$: $\Gamma_{i_0} \to \mathbb{R}^{N-1} \times \{0\}$. Therefore, for $r > 0$ small enough,

$$\Gamma_{i_0} \cap Q_{\nu_\Gamma}(x_0, r) = (\mathbb{P}_{x_0,\nu_\Gamma} (\Gamma_{i_0} \cap Q_{\nu_\Gamma}(x_0, r)), \, l_{i_0}((\mathbb{P}_{x_0,\nu_\Gamma} (\Gamma_{i_0} \cap Q_{\nu_\Gamma}(x_0, r)))')),$$

and by Theorem 9.1 in [63] we obtain that,

$$\mathcal{H}^{N-1}(\Gamma_{i_0} \cap Q_{\nu_\Gamma}(x_0, r)) = \int_{\mathbb{P}_{x_0,\nu_\Gamma}(\Gamma_{i_0} \cap Q_{\nu_\Gamma}(x_0,r))} \sqrt{1 + |\nabla l_{i_0}(x')|^2} d\mathcal{H}^{N-1}(x').$$

Since l_{i_0} is of class C^1 and $\nabla l_{i_0}(x_0) = 0$, for $\varepsilon > 0$ choose $r_\varepsilon > 0$ such that $|\nabla l_{i_0}(x)| < \varepsilon$ for all $0 < r < r_\varepsilon$. Therefore, we have that

$$\begin{aligned}
\mathcal{H}^{N-1}(\mathbb{P}_{x_0,\nu_\Gamma} (\Gamma \cap Q_{\nu_\Gamma}(x_0, r))) &\geq \mathcal{H}^{N-1}(\mathbb{P}_{x_0,\nu_\Gamma} (\Gamma_{i_0} \cap Q_{\nu_\Gamma}(x_0, r))) \\
&\geq \frac{1}{\sqrt{1+\varepsilon^2}} \int_{\mathbb{P}_{x_0,\nu_\Gamma}(\Gamma_{i_0} \cap Q_{\nu_\Gamma}(x_0,r))} \sqrt{1 + |\nabla l_{i_0}(x')|^2} dx' \\
&= \frac{1}{\sqrt{1+\varepsilon^2}} \mathcal{H}^{N-1}(\Gamma_{i_0} \cap Q_{\nu_\Gamma}(x_0, r)).
\end{aligned}$$

We obtain

$$\begin{aligned}
&\liminf_{r\to 0} \frac{\mathcal{H}^{N-1}(\mathbb{P}_{x_0,\nu_\Gamma} (\Gamma \cap Q_{\nu_\Gamma}(x_0, r)))}{r^{N-1}} \\
&\geq \liminf_{r\to 0} \frac{1}{\sqrt{1+\varepsilon^2}} \frac{\mathcal{H}^{N-1}(\Gamma_{i_0} \cap Q_{\nu_\Gamma}(x_0, r))}{r^{N-1}} = \frac{1}{\sqrt{1+\varepsilon^2}}.
\end{aligned}$$

By the arbitrariness of $\varepsilon > 0$, we deduce that

$$\liminf_{r\to 0} \frac{\mathcal{H}^{N-1}(\mathbb{P}_{x_0,\nu_\Gamma} (\Gamma \cap Q_{\nu_\Gamma}(x_0, r)))}{r^{N-1}} \geq 1,$$

and, in view of (3.40), we conclude that

$$\lim_{r\to 0} \frac{\mathcal{H}^{N-1}(\mathbb{P}_{x_0,\nu_\Gamma} (\Gamma \cap Q_{\nu_\Gamma}(x_0, r)))}{r^{N-1}} = 1.$$

□

Lemma 3.22. *Let $Q := (-1,1)^N$ and let $\Gamma \subset Q$ be a $\mathcal{H}^{N-1}$-rectifiable set such that $\mathcal{H}^{N-1}(\Gamma) < \infty$ and*

$$\mathcal{H}^0(\Gamma \cap (\{x'\} \times (-1,1))) \geq 1 \tag{3.41}$$

for $\mathcal{H}^{N-1}$-a.e. $x' \in (-1,1)^{N-1}$. Then there exists a $\mathcal{H}^{N-1}$-measurable subset $\Gamma' \subset \Gamma$ such that

$$\mathcal{H}^0(\Gamma' \cap (\{x'\} \times (-1,1))) = 1. \tag{3.42}$$

for $\mathcal{H}^{N-1}$-a.e. $x' \in (-1,1)^{N-1}$.

Proof. By Lemma 3.20 we have

$$\mathcal{H}^0(\Gamma' \cap (\{x'\} \times (-1,1))) < +\infty$$

for $\mathcal{H}^{N-1}$-a.e. $x' \in (-1,1)^{N-1}$. Thus, for $\mathcal{H}^{N-1}$-a.e. $x' \in (-1,1)^{N-1}$, the set

$$\Gamma_{x'} := \Gamma \cap (\{x'\} \times (-1,1))$$

is a finite collection of singletons, hence closed, and by (3.41) is non-empty. Applying Corollary 1.1 in [39], page 237, we obtain a $\mathcal{H}^{N-1}$ measurable subset $\Gamma' \subset \Gamma$ which satisfies (3.42). □

Lemma 3.23. *Let $\tau > 0$ and $\eta > 0$ be given. Let $u \in SBV(\Omega)$ and assume that $\mathcal{H}^{N-1}(S_u) < \infty$. The following statements hold:*

1. *there exist a set $S \subset S_u$ with $\mathcal{H}^{N-1}(S_u \backslash S) < \eta$, and a countable collection $\mathcal{Q}$ of mutually disjoint open cubes centered on elements of S_u such that*

$$\bigcup_{Q \in \mathcal{Q}} Q \subset \Omega,$$

and

$$\mathcal{H}^{N-1}\left(S \setminus \bigcup_{Q \in \mathcal{Q}} Q\right) = 0;$$

2. *for every $Q \in \mathcal{Q}$ there exists a direction vector $\nu_Q \in \mathcal{S}^{N-1}$ such that*

$$\mathcal{H}^0(S \cap Q_{x,\nu_Q}) = 1,$$

for $\mathcal{H}^{N-1}$ a.e. $x \in Q \cap S$;

3. *for every $Q \in \mathcal{Q}$, $S \cap Q$ is contained in a Lipschitz $(N-1)$- graph Γ_Q with Lipschitz constant less than τ.*

Proof. Let $\tau, \eta > 0$ be given. By Theorem 2.76 in [4], there exist countably many Lipschitz $(N-1)$- graphs $\Gamma_i \subset \mathbb{R}^N$ such that (up to a rotation and a translation)

$$\Gamma_i = \{(x', x_N) : \ x' \in N_i, \ x_N = l_i(x')\}$$

with $N_i \subset \mathbb{R}^{N-1}$, l_i: $\mathbb{R}^{N-1} \to \mathbb{R}$ of class C^1, $|\nabla l_i| < \tau$ for all $i \in \mathbb{N}$, and

$$\mathcal{H}^{N-1}\left(S_u \setminus \bigcup_{i=1}^{\infty} \Gamma_i\right) = 0. \tag{3.43}$$

Without lose of generality, we assume that

$$\mathcal{H}^{N-1}(\Gamma_i \cap \Gamma_{i'}) = 0 \text{ if } i \neq i' \in \mathbb{N}, \text{ and } \mathcal{H}^{N-1}(\Gamma_i) > 0. \tag{3.44}$$

We denote by $\mathcal{P}$ the collection of Lipschitz $(N-1)$-graphs Γ_i in (3.43)-(3.44). By (3.44), for $\mathcal{H}^{N-1}$- a.e. $x \in S_u$ there exists only one $\Gamma \in \mathcal{P}$ such that $x \in \Gamma$, and we denote such Γ by Γ_x and we write

$$\Gamma_x = \left\{ (y', y_N) : \; y' \in N_x \subset \mathbb{R}^{N-1}, \; y_N = l_x(y') \right\}.$$

For simplicity of notation, in what follows we will abbreviate $\nu_{\Gamma_x}(x) = \nu_{S_u}(x)$ by $\nu(x)$, $Q_{\nu_{S_u}}(x,r)$ by $Q(x,r)$, and $T_{x,\nu_{S_u}}$ by T_x.

We also note that $\mathcal{H}^{N-1}(\Gamma \cap S_u) < \mathcal{H}^{N-1}(S_u) < \infty$ for each $\Gamma \in \mathcal{P}$. Then $\mathcal{H}^{N-1}$- a.e. x has density 1 in $\Gamma_x \cap S_u$ (see Theorem 2.63 in [4]). Denote by S_1 the set of points such that S_u has density 1 at x and

$$\lim_{r \to 0} \frac{\mathcal{H}^{N-1}(S_u \cap \Gamma_x \cap Q_{\nu_{\Gamma_x}}(x,r))}{r^{N-1}} = 1. \tag{3.45}$$

Then $\mathcal{H}^{N-1}(S_u \setminus S_1) = 0$.

Define

$$f_r(x) := \frac{\mathcal{H}^{N-1}(S_1 \cap Q(x,r))}{r^{N-1}}.$$

Since $f_r(x) \to 1$ as $r \to 0^+$ for $x \in S_1$, by Egoroff's Theorem there exists a set $S_2 \subset S_1$ such that $\mathcal{H}^{N-1}(S_1 \setminus S_2) < \eta/4$ and $f_r \to 1$ uniformly on S_2. Find $r_1 > 0$ such that

$$\frac{\mathcal{H}^{N-1}(S_1 \cap Q(x,r))}{r^{N-1}} \geq \frac{1}{2},$$

i.e.,

$$\mathcal{H}^{N-1}(S_1 \cap Q(x,r)) \geq \frac{1}{2} r^{N-1} \tag{3.46}$$

for all $0 < r < r_1$ and $x \in S_2$. Since $S_2 \subset S_1$, S_2 is also $\mathcal{H}^{N-1}$-rectifiable and so $\mathcal{H}^{N-1}$ a.e. $x \in S_2$ has density one. Without loss of generality, we assume that every point in S_2 has density one and satisfies (3.39) in Lemma 3.21.

Let $x_0 \in S_2$ be given and recall (3.36). We define

$$\begin{aligned} T_b(x_0,r) &:= \left\{ x \in Q(x_0,r) \cap T_{x_0} : \; \mathcal{H}^0([Q(x_0,r)]_{x,\nu(x_0)} \cap S_2) \geq 2 \right\}, \\ T_g(x_0,r) &:= \left\{ x \in Q(x_0,r) \cap T_{x_0} : \; \mathcal{H}^0([Q(x_0,r)]_{x,\nu(x_0)} \cap S_2) = 1 \right\}, \\ S_b(x_0,r) &:= \bigcup_{x \in T_b(x_0,r)} \left(S_2 \cap [Q(x_0,r)]_{x,\nu(x_0)} \right), \\ S_g(x_0,r) &:= \bigcup_{x \in T_g(x_0,r)} \left(S_2 \cap [Q(x_0,r)]_{x,\nu(x_0)} \right). \end{aligned} \tag{3.47}$$

Note that

$$T_b(x_0,r)\cap T_g(x_0,r)=\emptyset \text{ and } S_b(x_0,r)\cap S_g(x_0,r)=\emptyset, \tag{3.48}$$

and by Proposition 3.20 we have

$$\mathcal{H}^{N-1}(S_g(x_0,r))\geq \mathcal{H}^{N-1}(T_g(x_0,r)). \tag{3.49}$$

We claim that

$$\mathcal{H}^{N-1}(S_b(x_0,r))\geq 2\mathcal{H}^{N-1}(T_b(x_0,r)). \tag{3.50}$$

By Lemma 3.22 there exists a measurable selection $S_b^1\subset S_b(x_0,r)$ such that

$$\mathcal{H}^{N-1}(S_b^1(x_0,r)\cap [Q(x_0,r)]_{x,\nu(x_0)})=1$$

for $\mathcal{H}^{N-1}$-a.e. $x\in T_b(x_0,r)$. We define

$$S_b^2(x_0,r):=S_b(x_0,r)\setminus S_b^1(x_0,r).$$

By the definition of $S_b(x_0,r)$ in (3.47), we have

$$\mathcal{H}^0([Q(x_0,r)]_{x,\nu(x_0)}\cap S_b^1(x_0,r))\geq 1 \text{ and } \mathcal{H}^0([Q(x_0,r)]_{x,\nu(x_0)}\cap S_b^2(x_0,r))\geq 1$$

for all $x\in T_b(x_0,r)$. We observe that

$$\mathcal{H}^{N-1}(S_b(x_0,r))=\mathcal{H}^{N-1}(S_b^1(x_0,r))+\mathcal{H}^{N-1}(S_b^2(x_0,r))\geq 2\mathcal{H}^{N-1}(T_b(x_0,r))$$

by Proposition 3.20 and we deduce (3.50).

We next show that

$$\lim_{r\to 0}\frac{\mathcal{H}^{N-1}(S_b(x_0,r))}{r^{N-1}}=0. \tag{3.51}$$

Indeed, since T_{x_0} is the tangent hyperplane to S_2 at x_0,

$$T_b(x_0,r)\cup T_g(x_0,r)=\mathbb{P}_{x_0,\nu_{S_u}}(S_2\cap Q(x_0,r)),$$

and by Lemma 3.21 it follows that

$$\lim_{r\to 0}\frac{\mathcal{H}^{N-1}(T_b(x_0,r)\cup T_g(x_0,r))}{r^{N-1}}=1. \tag{3.52}$$

On the other hand, in view of (3.48), (3.49), and (3.50), we deduce that

$$\begin{aligned}\mathcal{H}^{N-1}(S_b(x_0,r)\cup S_g(x_0,r))&=\mathcal{H}^{N-1}(S_b(x_0,r))+\mathcal{H}^{N-1}(S_g(x_0,r))\\&\geq 2\mathcal{H}^{N-1}(T_b(x_0,r))+\mathcal{H}^{N-1}(T_g(x_0,r)).\end{aligned}$$

That is,

$$\begin{aligned}\mathcal{H}^{N-1}(T_b(x_0,r))&\leq \mathcal{H}^{N-1}(S_b(x_0,r)\cup S_g(x_0,r))-\left[\mathcal{H}^{N-1}(T_b(x_0,r))+\mathcal{H}^{N-1}(T_g(x_0,r))\right]\\&=\mathcal{H}^{N-1}(S_b(x_0,r)\cup S_g(x_0,r))-\mathcal{H}^{N-1}(T_b(x_0,r)\cup T_g(x_0,r)).\end{aligned} \tag{3.53}$$

Since $x_0\in S_2$ has density 1, we have

$$\lim_{r\to 0}\frac{\mathcal{H}^{N-1}(S_b(x_0,r)\cup S_g(x_0,r))}{r^{N-1}}=\lim_{r\to 0}\frac{\mathcal{H}^{N-1}(S_2\cap Q(x_0,r))}{r^{N-1}}=1. \tag{3.54}$$

In view of (3.52), (3.53), and (3.54), we conclude that

$$\limsup_{r\to 0}\frac{\mathcal{H}^{N-1}(T_b(x_0,r))}{r^{N-1}}$$
$$\le \lim_{r\to 0}\frac{\mathcal{H}^{N-1}(S_b(x_0,r)\cup S_g(x_0,r))}{r^{N-1}} - \lim_{r\to 0}\frac{\mathcal{H}^{N-1}(T_b(x_0,r)\cup T_g(x_0,r))}{r^{N-1}} = 0,$$

which implies that

$$\lim_{r\to 0}\frac{\mathcal{H}^{N-1}(T_b(x_0,r))}{r^{N-1}} = 0.$$

This, together with (3.48) and (3.52), yields

$$\lim_{r\to 0}\frac{\mathcal{H}^{N-1}(T_g(x_0,r))}{r^{N-1}} = 1,$$

and so by (3.49) we have

$$\liminf_{r\to 0}\frac{\mathcal{H}^{N-1}(S_g(x_0,r))}{r^{N-1}} \ge \lim_{r\to 0}\frac{\mathcal{H}^{N-1}(T_g(x_0,r))}{r^{N-1}} = 1,$$

while by (3.54)

$$\limsup_{r\to 0}\frac{\mathcal{H}^{N-1}(S_g(x_0,r))}{r^{N-1}} \le \lim_{r\to 0}\frac{\mathcal{H}^{N-1}(S_b(x_0,r)\cup S_g(x_0,r))}{r^{N-1}} = 1,$$

and we conclude that

$$\lim_{r\to 0}\frac{\mathcal{H}^{N-1}(S_g(x_0,r))}{r^{N-1}} = 1.$$

Now, also in view of (3.48) and (3.54), we deduce (3.51).

We define, for $x\in S_2$,

$$g_r(x) := \frac{\mathcal{H}^{N-1}(S_b(x,r))}{r^{N-1}}.$$

By (3.51) we have $\lim_{r\to 0} g_r(x) = 0$ for all $x\in S_2$, therefore by Egoroff's Theorem there exists a set $S_3\subset S_2$ such that

$$\mathcal{H}^{N-1}(S_2\setminus S_3) < \frac{\eta}{4}$$

and $g_r\to 0$ uniformly on S_3. Choose $0<r_2<r_1$ such that

$$\frac{\mathcal{H}^{N-1}(S_b(x,r))}{r^{N-1}} < \frac{\eta}{16}\frac{1}{\mathcal{H}^{N-1}(S_u)} \tag{3.55}$$

for all $x\in S_3$ and $0<r<r_2$. We claim that, for $x\in S_3$ and the corresponding $\Gamma_x\in\mathcal{P}$,

$$\lim_{r\to 0}\frac{\mathcal{H}^{N-1}\left(S_g(x,r)\setminus[S_u\cap\Gamma_x\cap Q(x,r)]\right)}{r^{N-1}} = 0. \tag{3.56}$$

Suppose that

$$0 < \limsup_{r\to 0}\frac{\mathcal{H}^{N-1}\left(S_g(x,r)\setminus[S_u\cap\Gamma_x\cap Q(x,r)]\right)}{r^{N-1}} =: \delta.$$

By (3.45), and the fact that $\Gamma_x \subset S_u$, we have that

$$\begin{aligned}1 &= \lim_{r\to 0} \frac{\mathcal{H}^{N-1}(S_u \cap Q(x,r))}{r^{N-1}} \\ &= \lim_{r\to 0} \frac{\mathcal{H}^{N-1}\left(([S_u \cap Q(x,r)] \setminus [S_u \cap \Gamma_x \cap Q(x,r)]) \cup [S_u \cap \Gamma_x \cap Q(x,r)]\right)}{r^{N-1}} \\ &\geq \limsup_{r\to 0} \frac{\mathcal{H}^{N-1}\left([S_g(x,r)] \setminus [S_u \cap \Gamma_x \cap Q(x,r)]\right)}{r^{N-1}} \\ &\quad + \lim_{r\to 0} \frac{\mathcal{H}^{N-1}[S_u \cap \Gamma_x \cap Q(x,r)]}{r^{N-1}} \\ &= \delta + 1 > 1,\end{aligned}$$

which is a contradiction.

We define, for $x \in S_3$,

$$h_r(x) := \frac{\mathcal{H}^{N-1}\left(S_g(x,r) \setminus [S_u \cap \Gamma_x \cap Q(x,r)]\right)}{r^{N-1}}.$$

By (3.56) $\lim_{r\to 0} h_r(x) = 0$ for all $x \in S_3$, therefore by Egoroff's Theorem there exists a set of $S_4 \subset S_3$ such that

$$\mathcal{H}^{N-1}(S_3 \setminus S_4) < \frac{\eta}{4},$$

and $h_r \to 0$ uniformly on S_4. Choose $0 < r_3 < r_2$ such that

$$\frac{\mathcal{H}^{N-1}\left(S_g(x,r) \setminus [S_u \cap \Gamma_x \cap Q(x,r)]\right)}{r^{N-1}} < \frac{\eta}{16} \frac{1}{\mathcal{H}^{N-1}(S_u)} \tag{3.57}$$

for all $x \in S_4$ and $0 < r < r_3$, and let

$$\mathcal{Q}' := \{Q(x,r) : \ x \in S_4,\ 0 < r < r_3\}.$$

By Besicovitch's Covering Theorem we may extract a countable collection $\mathcal{Q}$ of mutually disjoint cubes from $\mathcal{Q}'$ such that

$$\bigcup_{Q\in\mathcal{Q}} Q \subset \Omega \text{ and } \mathcal{H}^{N-1}\left(S_4 \setminus \left(\bigcup_{Q\in\mathcal{Q}} Q\right)\right) = 0.$$

Define

$$S := S_4 \setminus \left[\left(\bigcup_{Q\in\mathcal{Q}} S_b(x_Q, r_Q)\right) \cup \left(\bigcup_{Q\in\mathcal{Q}} \left[S_g(x_Q, r_Q) \setminus \left(S_u \cap \Gamma_{x_Q} \cap Q\right)\right]\right)\right], \tag{3.58}$$

where x_Q is the center of cube Q and r_Q is the side length of Q. Note that the set S satisfies properties 2 and 3 in the statement of Lemma 3.23. Finally, we show that

$$\mathcal{H}^{N-1}(S_u \setminus S) < \eta.$$

Indeed, in view of (3.55) and (3.57), and using the fact that the cubes $Q \in \mathcal{Q}$ are mutually disjoint, we have

$$\mathcal{H}^{N-1}\left(\bigcup_{Q\in\mathcal{Q}} S_b(x_Q, r_Q)\right) = \sum_{Q\in\mathcal{Q}} \mathcal{H}^{N-1}(S_b(x_Q, r_Q)) \leq \frac{\eta}{16\mathcal{H}^{N-1}(S_u)} \sum_{Q\in\mathcal{Q}} r_Q^{N-1}, \tag{3.59}$$

and

$$\begin{aligned}&\mathcal{H}^{N-1}\left(\bigcup_{Q\in\mathcal{Q}} \left[S_g(x_Q, r_Q) \setminus \left(S_u \cap \Gamma_{x_Q} \cap Q\right)\right]\right) \\ &= \sum_{Q\in\mathcal{Q}} \mathcal{H}^{N-1}(S_g(x_Q, r_Q) \setminus \left(S_u \cap \Gamma_{x_Q} \cap Q\right)) \leq \frac{\eta}{16\mathcal{H}^{N-1}(S_u)} \sum_{Q\in\mathcal{Q}} r_Q^{N-1}.\end{aligned} \tag{3.60}$$

By (3.46) we obtain

$$\sum_{Q\in\mathcal{Q}} \frac{1}{2} r_Q^{N-1} \leq \sum_{Q\in\mathcal{Q}} \mathcal{H}^{N-1}(S_1 \cap Q) = \mathcal{H}^{N-1}\left(\bigcup_{Q\in\mathcal{Q}} S_u \cap Q\right) \leq \mathcal{H}^{N-1}(S_u). \tag{3.61}$$

Using (3.59), (3.60), and (3.61), we deduce that

$$\mathcal{H}^{N-1}\left(\bigcup_{Q\in\mathcal{Q}} S_b(x_Q, r_Q)\right) \leq \frac{\eta}{8},$$

and

$$\mathcal{H}^{N-1}\left(\bigcup_{Q\in\mathcal{Q}} \left[S_g(x_Q, r_Q) \setminus \left(S_u \cap \Gamma_{x_Q} \cap Q\right)\right]\right) \leq \frac{\eta}{8},$$

and so by (3.58) we get

$$\mathcal{H}^{N-1}(S_4 \setminus S) \leq \frac{\eta}{4}.$$

Since $S \subset S_4 \subset S_3 \subset S_2 \subset S_1 \subset S_u$, we conclude that

$$\begin{aligned}&\mathcal{H}^{N-1}(S_u \setminus S) \\ \leq &\mathcal{H}^{N-1}(S_u \setminus S_1) + \mathcal{H}^{N-1}(S_1 \setminus S_2) + \mathcal{H}^{N-1}(S_2 \setminus S_3) + \mathcal{H}^{N-1}(S_3 \setminus S_4) + \mathcal{H}^{N-1}(S_4 \setminus S) \\ \leq &\frac{\eta}{4} + \frac{\eta}{4} + \frac{\eta}{4} + \frac{\eta}{4} = \eta\end{aligned}$$

as desired. □

3.3.2. Γ-lim inf *Inequality: The Multi-Dimensional Case.* In this section we prove the Γ-lim inf inequality which is stated in the following proposition.

Proposition 3.24. (Γ-lim inf) *For $\omega \in \mathcal{W}(\Omega)$ and $u \in L^1(\Omega)$, let*

$$MS_\omega^-(u) := \inf\left\{\liminf_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) : \right.$$

$$(u_\varepsilon, v_\varepsilon) \in W^{1,2}(\Omega) \times W^{1,2}(\Omega),\ u_\varepsilon \to u \ \mathit{in}\ L^1, v_\varepsilon \to 1 \ \mathit{in}\ L^1,\ 0 \le v_\varepsilon \le 1 \ \mathit{a.e.}\big\}.$$

We have $MS^-_\omega(u) \ge MS_\omega(u)$.

To prove Proposition 3.24 we reduce the statement to the case $N = 1$ by a special slicing argument (Lemma 3.23), and we use the result from case $N = 1$.

Proof of Proposition 3.24. Without lose of generality we may assume that $M := MS^-_\omega(u) < \infty$. Let $\{(u_\varepsilon, v_\varepsilon)\}_{\varepsilon>0} \subset W^{1,2}(\Omega) \times W^{1,2}(\Omega)$ be such that $u_\varepsilon \to u$ in L^1, $v_\varepsilon \to 1$ in L^1, and $\lim_{\varepsilon\to0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) = MS^-_\omega(u)$. Since $\inf_{x\in\Omega} \omega(x) \ge 1$, we have

$$\liminf_{\varepsilon\to0} AT_{1,\varepsilon}(u_\varepsilon, v_\varepsilon) \le \liminf_{\varepsilon\to0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) < \infty,$$

and by [6] we deduce that $u \in GSBV(\Omega)$ and $\mathcal{H}^{N-1}(S_u) < \infty$. We prove separately that

$$\liminf_{\varepsilon\to0} \int_\Omega |\nabla u_\varepsilon|^2 v_\varepsilon\, \omega\, dx \ge \int_\Omega |\nabla u|^2 \omega\, dx, \tag{3.62}$$

and

$$\liminf_{\varepsilon\to0} \int_\Omega \left(\varepsilon\, |\nabla v_\varepsilon|^2 + \frac{1}{4\varepsilon}(1 - v_\varepsilon)^2\right) \omega\, dx \ge \int_{S_u} \omega^- \, d\mathcal{H}^{N-1}. \tag{3.63}$$

Let A be an open subset of Ω. Fix $\nu \in \mathcal{S}^{N-1}$, and define $A_{x,\nu}$, $A^1_{x,\nu}$, and A_ν as in (3.36). For $K \in \mathbb{R}^+$, set $u_K := K \wedge u \vee -K$, and we observe, by Fubini's Theorem, Fatou's Lemma, Theorem 3.17, equation (3.15), and Theorem 2.3 in [6], that

$$\begin{aligned} &\liminf_{\varepsilon\to0} \int_A |\nabla u_\varepsilon|^2 v_\varepsilon^2\, \omega\, dx \\ &\ge \int_{A_\nu} \liminf_{\varepsilon\to0} \int_{A^1_{x,\nu}} \left|(u_\varepsilon)'_{x,\nu}\right|^2 (v_\varepsilon)^2_{x,\nu}\, \omega_{x,\nu}\, dt\, d\mathcal{H}^{N-1}(x) \\ &\ge \int_{A_\nu} \int_{A^1_{x,\nu}} \left|(u_K)'_{x,\nu}\right|^2 \omega_{x,\nu}\, dt\, d\mathcal{H}^{N-1}(x) \ge \int_A |\langle \nabla u_K(x), \nu\rangle|^2 \omega\, dx. \end{aligned} \tag{3.64}$$

Letting $K \to \infty$ and using Lebesgue Monotone Convergence Theorem we have

$$\liminf_{\varepsilon\to0} \int_A |\nabla u_\varepsilon|^2 v_\varepsilon^2\, \omega\, dx \ge \int_A |\langle \nabla u(x), \nu\rangle|^2 \omega\, dx. \tag{3.65}$$

Let $\phi_n(x) := |\langle \nabla u(x), \nu_n\rangle|^2 \omega$ for $\mathcal{L}^N$-a.e. $x \in \Omega$, where $\{\nu_n\}_{n=1}^\infty$ is a dense subset of $\mathcal{S}^{N-1}$, and let $\mu(A) := \liminf_{\varepsilon\to0} \int_A |\nabla u_\varepsilon|^2 v_\varepsilon^2\, \omega\, dx$. Then μ is a positive function, super-additivity on open sets A, B, with disjoint closures, and hence by Lemma 15.2 in [15], together with (3.65), we conclude (3.62). Now we prove (3.63). By Fubini's Theorem, Fatou's Lemma, (3.38), and (3.16), and using a similar calculation as in (3.64) we have

$$\liminf_{\varepsilon\to0} \int_A \left(\varepsilon\, |\nabla v_\varepsilon|^2 + \frac{1}{4\varepsilon}(1 - v_\varepsilon)^2\right) \omega\, dx \ge \int_{A_\nu} \left[\sum_{t\in S_{u_{x,\nu}}\cap A^1_{x,\nu}} \omega^-_{x,\nu}(t)\right] d\mathcal{H}^{N-1}(x). \tag{3.66}$$

Next, given arbitrary $\tau > 0$ and $\eta > 0$ we choose a set $S \subset S_u$ and a collection $\mathcal{Q}$ of mutually disjoint cubes according to Lemma 3.23 with respect to S_u. Fix one such cube $Q_{\nu_S}(x_0, r_0) \in \mathcal{Q}$. By Lemma 3.23 we have, up to a rotation and a translation,

$$\Gamma_{x_0} = \left\{(y', l_{x_0}(y')) : \ y \in T_{x_0,\nu_S} \cap Q_{\nu_S}(x_0, r_0)\right\} \text{ and } \|\nabla l_{x_0}\|_{L^\infty} < \tau.$$

In (3.66) set $A = Q_{\nu_S}(x_0, r_0)$ and $\nu = \nu_S(x_0)$ and, using the same notation as in the proof of Lemma 3.23, we obtain

$$\begin{aligned} &\int_{\left[Q_{\nu_S}(x_0,r_0)\right]_{\nu_S(x_0)}} \left(\sum_{t \in S_{u_{x,\nu_S(x_0)}} \cap \left[Q_{\nu_S}(x_0,r_0)\right]_{x,\nu_S(x_0)}} \omega^-_{x,\nu_S(x_0)}(t) \right) d\mathcal{H}^{N-1}(x) \\ &\qquad \geq \int_{T_g(x_0,r_0)} \omega^-(x)\, d\mathcal{H}^{N-1}(x) = \int_{T_g(x_0,r_0)} \omega^-(x', l_{x_0}(x')) d\mathcal{L}^{N-1}(x'). \end{aligned} \tag{3.67}$$

where in the first inequality we used the fact that $\omega^-_{x,\nu}(t) = \omega^-(x + t\nu)$ (see Remark 3.109 in [4]). Next, in view of Lemma 3.18, we have that

$$\begin{aligned} \int_{Q_{\nu_S}(x_0,r_0)\cap S} \omega^-\, d\mathcal{H}^{N-1} &= \int_{T_{x_0,\nu_S}\cap Q_{\nu_S}(x_0,r_0)} \omega^-(x', l_{x_0}(x'))\sqrt{1 + |\nabla l_{x_0}(x')|^2} dx' \\ &\leq \sqrt{1+\tau^2} \int_{T_{x_0,\nu_S}\cap Q_{\nu_S}(x_0,r_0)} \omega^-(x', l_{x_0}(x')) dx', \end{aligned}$$

which, together with (3.66) and (3.67), yields

$$\begin{aligned} &\liminf_{\varepsilon\to 0} \int_\Omega \left[\varepsilon\, |\nabla v_\varepsilon|^2 + \frac{1}{4\varepsilon}(1 - v_\varepsilon)^2\right] \omega\, dx \\ &\quad \geq \liminf_{\varepsilon\to 0} \int_{\cup_{Q\in\mathcal{Q}} Q} \left[\varepsilon\, |\nabla v_\varepsilon|^2 + \frac{1}{4\varepsilon}(1 - v_\varepsilon)^2\right] \omega\, dx \\ &\quad \geq \frac{1}{\sqrt{1+\tau^2}} \sum_{Q\in\mathcal{Q}} \int_{S\cap Q} \omega^-\, d\mathcal{H}^{N-1} \geq \frac{1}{\sqrt{1+\tau^2}} \left(\int_{S_u} \omega^-\, d\mathcal{H}^{N-1} - \|\omega\|_{L^\infty}\, \eta \right), \end{aligned} \tag{3.68}$$

and (3.63) follows from the arbitrariness of η and τ, and the fact that η and τ are independent. □

Remark 3.25. The assumption $\omega \in L^\infty$ can be removed by applying (3.68) to $\omega_K := K \wedge \omega \vee -K$ and letting $K \nearrow \infty$ and using Lebesgue Monotone Convergence Theorem.

3.3.3. *The* Γ-$\limsup$ *Inequality.* This section is devoted to the proof of the Γ-$\limsup$ inequality and the proof of Theorem 1.2. The main task is to prove the following proposition.

Proposition 3.26. (Γ-$\limsup$) *For* $\omega \in \mathcal{W}(\Omega)$ *and* $u \in L^1(\Omega) \cap L^\infty(\Omega)$, *let*

$$\begin{aligned} MS^+_\omega(u) := \inf \Big\{ &\limsup_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) : \\ &(u_\varepsilon, v_\varepsilon) \in W^{1,2}(\Omega) \times W^{1,2}(\Omega), u_\varepsilon \to u \text{ in } L^1,\ v_\varepsilon \to 1 \text{ in } L^1,\ 0 \leq v_\varepsilon \leq 1 \Big\}. \end{aligned}$$

We have $MS_\omega^+(u) \leq MS_\omega(u)$.

Proposition 3.26 will be proved in several propositions. To get start, we recall $Q_{\nu_{S_\omega}}(x_0, r)$ and $T_{x_0,\nu_{S_\omega}}(l)$ from Notation 3.3 1 and 2, and define $I(t_0, t) := (t_0 - t, t_0 + t) \subset \mathbb{R}$ for $t_0 \in \mathbb{R}$ and $t \in \mathbb{R}^+$.

Proposition 3.27. *Let $\omega \in \mathcal{W}(\Omega)$ and $\tau \in (0, 1/4)$ be given. Then for $\mathcal{H}^{N-1}$ a.e. $x_0 \in S_\omega$, a point of density one, there exists $r_0 := r_0(x_0) > 0$ such that for each $0 < r < r_0$ there exist $t_0 \in (2.5\tau r, 3.5\tau r)$ and $0 < t_{0,r} = t_{0,r}(t_0, \tau, x_0, r) < t_0$ such that $I(t_0, t_{0,r}) \subset (2\tau r, 4\tau r)$ and*

$$\begin{aligned} &\sup_{0<t\leq t_{0,r}} \frac{1}{|I(t_0,t)|} \int_{I(t_0,t)} \int_{Q_{\nu_{S_\omega}}(x_0,r)\cap T_{x_0,\nu_{S_\omega}}(-l)} \omega(x) d\mathcal{H}^{N-1}(x) dl \\ &\qquad \leq \int_{S_\omega \cap Q_{\nu_{S_\omega}}(x_0,r)} \omega^-(x)\, d\mathcal{H}^{N-1} + O(\tau) r^{N-1}. \end{aligned} \tag{3.69}$$

Proof. For simplicity of notation, in what follows we abbreviate $Q_{\nu_{S_\omega}}(x_0, r)$ as $Q(x_0, r)$, $T_{x_0,\nu_{S_\omega}}$ as T_{x_0}, and $T_{x_0,\nu_{S_\omega}}(l)$ as $T_{x_0}(l)$. Since $\mathcal{H}^{N-1}(S_\omega) < \infty$, and so $\mu := \mathcal{H}^{N-1}\lfloor S_\omega$ is a nonnegative Radon measure, and since $\omega^- \in L^1(\Omega, \mu)$, it follows that for $\mathcal{H}^{N-1}$ a.e. $x_0 \in S_\omega$

$$\lim_{r\to 0} \fint_{Q(x_0,r)\cap S_\omega} \left|\omega^-(x) - \omega^-(x_0)\right| d\mathcal{H}^{N-1}(x) = 0, \tag{3.70}$$

Choose one such $x_0 \in S_\omega$, also a point of density 1 of S_ω, and let $\tau > 0$ be given. Select $r_1 > 0$ such that for all $0 < r < r_1$,

$$\frac{1}{1+\tau^2} \leq \frac{\mathcal{H}^{N-1}(S_\omega \cap Q(x_0, r))}{r^{N-1}} \leq 1 + \tau^2. \tag{3.71}$$

Let $0 < r_2 < r_1$ be such that, in view of (3.70),

$$\int_{Q(x_0,r)\cap S_\omega} \left|\omega^-(x) - \omega^-(x_0)\right| d\mathcal{H}^{N-1} \leq \tau^2 r^{N-1} \tag{3.72}$$

for all $0 < r < r_2$, and observe that, in view of (3.71),

$$\omega^-(x_0)\mathcal{H}^{N-1}\left[Q(x_0, r) \cap T_{x_0}(-t_0)\right] \leq (1 + \tau^2)\, \omega^-(x_0)\mathcal{H}^{N-1}\left[Q(x_0, r) \cap S_\omega\right]. \tag{3.73}$$

By Theorem 3.4 we may choose $0 < r_3 < r_2$ such that, for all $0 < r < r_3$,

$$\fint_{Q^-(x_0,r)} \left|\omega(x) - \omega^-(x_0)\right| dx \leq \tau^2, \tag{3.74}$$

and so, since $3.5\tau r < r$, we have

$$\begin{aligned} &\int_{2.5\tau r}^{3.5\tau r} \int_{Q^-(x_0,r)\cap T_{x_0}(-t)} \left|\omega(x) - \omega^-(x_0)\right| d\mathcal{H}^{N-1}(x) dt \\ &\qquad \leq \int_{Q^-(x_0,r)} \left|\omega(x) - \omega^-(x_0)\right| dx \leq \frac{1}{2}\tau^2 r^N. \end{aligned}$$

Therefore, there exists a set $A \subset (2.5\tau r, 3.5\tau r)$ with positive one dimension Lebesgure measure such that for every $t \in A$,

$$\int_{Q^-(x_0,r)\cap T_{x_0}(-t)} \left|\omega(x) - \omega^-(x_0)\right| d\mathcal{H}^{N-1}(x) \le \frac{1}{2}\frac{\tau^2 r^N}{\tau r} \le \tau r^{N-1}, \tag{3.75}$$

and we choose $t_0 \in A$ a Lebesgue point so that

$$\begin{aligned}&\lim_{t\to 0}\frac{1}{|I(t_0,t)|}\int_{I(t_0,t)}\int_{Q^-(x_0,r)\cap T_{x_0}(-l)}\omega(x)d\mathcal{H}^{N-1}(x)dl\\ &\quad=\int_{Q^-(x_0,r)\cap T_{x_0}(-t_0)}\omega(x)d\mathcal{H}^{N-1}(x).\end{aligned}$$

Hence, there exists $t_{0,r} > 0$, depending on t_0, τ, r, and x_0, such that $I(t_0, t_{0,r}) \subset (2.5\tau r, 3.5\tau r)$ and

$$\begin{aligned}&\sup_{0<t\le t_{0,r}}\frac{1}{|I(t_0,t)|}\int_{I(t_0,t)}\int_{Q^-(x_0,r)\cap T_{x_0}(-l)}\omega(x)d\mathcal{H}^{N-1}(x)dl\\ &\quad\le\int_{Q^-(x_0,r)\cap T_{x_0}(-t_0)}\omega(x)d\mathcal{H}^{N-1}+\tau r^{N-1}.\end{aligned} \tag{3.76}$$

In view of (3.76), (3.75), (3.73), and (3.72), in this order, we have that for every $0 < r < r_3$ there exist $t_0 \in (2.5\tau r, 3.5\tau r)$ and $0 < t_{0,r} < t_0$ such that

$$\begin{aligned}&\sup_{0<t\le t_{0,r}}\frac{1}{|I(t_0,t)|}\int_{I(t_0,t)}\int_{Q^-(x_0,r)\cap T_{x_0}(-l)}\omega(x)d\mathcal{H}^{N-1}(x)dl\\ &\quad\le\int_{Q^-(x_0,r)\cap T_{x_0}(-t_0)}\omega(x)d\mathcal{H}^{N-1}+\tau r^{N-1}\\ &\quad\le\int_{Q^-(x_0,r)\cap T_{x_0}(-t_0)}\left|\omega(x)-\omega^-(x_0)\right|d\mathcal{H}^{N-1}\\ &\qquad+\omega^-(x_0)\mathcal{H}^{N-1}\left[Q^-(x_0,r)\cap T_{x_0}(-t_0)\right]+\tau r^{N-1}\\ &\quad\le O(\tau)r^{N-1}+(1+\tau^2)\,\omega^-(x_0)\mathcal{H}^{N-1}\left[Q(x_0,r)\cap S_\omega\right]\\ &\quad\le O(\tau)r^{N-1}+(1+\tau^2)\int_{Q(x_0,r)\cap S_\omega}\omega^-(x)d\mathcal{H}^{N-1}.\end{aligned}$$

Since $\omega \in L^\infty(\Omega)$, we have $\omega^- \in L^\infty(\Omega)$ and thus, invoking (3.71),

$$\tau^2\int_{Q(x_0,r)\cap S_\omega}\omega^-(x)d\mathcal{H}^{N-1} \le O(\tau)\left\|\omega\right\|_{L^\infty}\mathcal{H}^{N-1}\left[Q(x_0,r)\cap S_\omega\right] \le O(\tau)r^{N-1},$$

and we conclude (3.69). □

Proposition 3.28. *Let $\omega \in \mathcal{W}(\Omega)$ and $\tau \in (0, 1/4)$ be given. There exist a set $S \subset S_\omega$ and a countable family of disjoint cubes $\mathcal{F} = \left\{Q_{\nu_{S_\omega}}(x_n, r_n)\right\}_{n=1}^\infty$ with $r_n \le \tau$, for all $n \in \mathbb{N}$, such that the following hold:*

1. *$\mathcal{H}^{N-1}(S_\omega \setminus S) < \tau$ and $S \subset \bigcup_{n=1}^\infty Q_{\nu_{S_\omega}}(x_n, r_n) \subset \Omega$;*

2. $(1+\tau^2)^{-1}r^{N-1} \le \mathcal{H}^{N-1}\left(S \cap Q_{\nu_{S_\omega}}(x_n, r)\right) \le (1+\tau^2)r^{N-1}$ *for all* $0 < r < r_n$;
3. $S \cap Q_{\nu_{S_\omega}}(x_n, r_n) \subset R_{\tau/2,\nu_{S_\omega}}(x_n, r_n)$ *(recall* $R_{\tau/2,\nu_{S_\omega}}$ *from Notation 3.3);*
4. *if* $0 < \kappa < 1$ *then for every* $n \in \mathbb{N}$ *there exist* $t_n^\kappa \in (2.5\tau\kappa r_n, 3.5\tau\kappa r_n)$ *and* $0 < t^\kappa_{x_n,r_n} < |t_n^\kappa|$*, depending on* τ*,* x_n*, and* κr_n*, such that*

$$\begin{aligned}&\sup_{0<t\le t^\kappa_{x_n,r_n}} \frac{1}{|I(t_n^\kappa, t)|}\int_{I(t_n^\kappa,t)}\int_{Q^-_{\nu_{S_\omega}}(x_n,\kappa r_n)\cap T_{x_n,\nu_{S_\omega}}(-l)} \omega^-(x)d\mathcal{H}^{N-1}dl \\ &\le \int_{S_\omega\cap Q_{\nu_{S_\omega}}(x_n,\kappa r_n)} \omega^-(x)d\mathcal{H}^{N-1} + (1+\omega^-(x_n))O(\tau)(\kappa r_n)^{N-1}.\end{aligned} \tag{3.77}$$

Proof. Let $\tau \in (0,1/4)$ and $\kappa \in (0,1)$ be given. Since $\mathcal{H}^{N-1}(S_\omega) < \infty$, there exists $S_1 \subset S_\omega$ such that $\mathcal{H}^{N-1}(S_\omega \setminus S_1) < \tau/3$, S_1 is compact and contained in a finite union of $(N-1)$-Lipschitz graphs Γ_i, $i = 1,\dots,M$, with Lipschitz constants less than $\tau/(2\sqrt{N})$. Moreover, since $\mathcal{H}^{N-1}$ a.e. $x \in S_1$ a point of density one, by Egorov's Theorem, we may find $S_2 \subset S_1$ compact such that $\mathcal{H}^{N-1}(S_1 \setminus S_2) < \tau/3$ and there exists $r_1 > 0$ such that $(1+\tau^2)^{-1}r^{N-1} \le \mathcal{H}^{N-1}\left(S_1 \cap Q_{\nu_{S_\omega}}(x,r)\right) \le (1+\tau^2)r^{N-1}$ for all $0 < r < r_1$ and $x \in S_2$.

Let $L_i := S_2 \cap \Gamma_i$ and without lose of generality we assume that L_i are mutually disjoint. Let $L'_i \subset L_i$ be such that $\mathcal{H}^{N-1}(L_i \setminus L'_i) < \tau/\left(3\cdot 2^i\right)$ and $d_{ij} := \operatorname{dist}(L'_i, L'_j) > 0$ for $i \ne j$. We observe that

$$\mathcal{H}^{N-1}\left(S_2 \setminus \bigcup_{i=1}^M L'_i\right) < \frac{\tau}{3} \text{ and } d := \min_{i\ne j}\{d_{ij}\} > 0.$$

Define $S := \bigcup_{i=1}^M L'_i$. Note that there exists $0 < r_2 < \min\left\{\tau^2, d/2, r_1\right\}$ such that for every $0 < r < r_2$ and every $x, y \in S$ with $|x-y| < \sqrt{N}r$ we have $S \cap Q_{\nu_{S_\omega}}(x,r) \subset R_{\tau/2,\nu_{S_\omega}}(x,r)$, where we are using the notation introduced in Notation 3.3. Next, for $\mathcal{H}^{N-1}$-a.e. $x \in S$ we may find $r_2(x) > 0$ such that $Q_{\nu_{S_\omega}}(x, r_2(x)) \subset \Omega$ and $\kappa r_2(x) \le r_0(x)$ where $r_0(x)$ is determined in Proposition 3.27. Let $\bar r_0(x) := \min\{r_1, r_2(x)\}$. The collection $\mathcal{F}' := \left\{Q_{\nu_{S_\omega}}(x,r) :\ x \in S,\ r < \bar r_0(x)\right\}$ is a fine cover for S, and so by Besicovitch's Covering Theorem we may obtain a countable sub-collection $\mathcal{F} \subset \mathcal{F}'$ with pairwise disjoint cubes such that $S \subset \bigcup_{Q_{\nu_{S_\omega}}(x_n,r_n)\in\mathcal{F}} Q_{\nu_{S_\omega}}(x_n,r_n) \subset \Omega$.

Finally, for each $Q_{\nu_{S_\omega}}(x_n,r_n) \in \mathcal{F}$ we apply Proposition 3.27 to obtain $t_n^\kappa \in (2.5\tau\kappa r_n, 3.5\tau\kappa r_n)$ and $t^\kappa_{x_n,r_n} > 0$, depending on $t^\kappa_{x_n}$, τ, κr_n, and x_n, such that (3.77) hold. We complete this proof by observing that

$$\mathcal{H}^{N-1}(S_\omega \setminus S) \le \mathcal{H}^{N-1}(S_\omega \setminus S_1) + \mathcal{H}^{N-1}(S_1 \setminus S_2) + \mathcal{H}^{N-1}(S_2 \setminus S) \le \tau.$$

□

Proposition 3.29. *Let* $\omega \in \mathcal{W}(\Omega)$ *and* $\tau \in (0,1/4)$ *be given. There exist a set* $S \subset S_\omega$ *and a countable family of disjoint cubes* $\mathcal{F} = \left\{Q_{\nu_{S_\omega}}(x_n,r_n)\right\}_{n=1}^\infty$*, with* $r_n < \tau$*, such that the following hold:*

1. $\mathcal{H}^{N-1}(S_\omega \setminus S) < \tau$ *and* $S \subset \bigcup_{n=1}^\infty Q_{\nu_{S_\omega}}(x_n,r_n) \subset \Omega$;

2. $\mathrm{dist}(Q_{\nu_{S_\omega}}(x_n, r_n), Q_{\nu_{S_\omega}}(x_{n'}, r_{n'})) > 0$ *for* $n \neq n'$;
3. $S \cap Q_{\nu_{S_\omega}}(x_n, r_n) \subset R_{\tau/2,\nu_{S_\omega}}(x_n, r_n)$;
4. $(1+\tau^2)^{-1} r_n^{N-1} \leq \mathcal{H}^{N-1}\left(S \cap Q_{\nu_{S_\omega}}(x_n, r_n)\right) \leq (1+\tau^2) r_n^{N-1}$;
5. $\sum_{n=1}^\infty r_n^{N-1} \leq 4\mathcal{H}^{N-1}(S_\omega)$;
6. *for each* $n \in \mathbb{N}$, *there exists* $t_n \in (2.5\tau r_n, 3.5\tau r_n)$ *and* $0 < t_{x_n,r_n} < t_n$, *depending on* τ, r_n, *and* x_n, *such that* $T_{x_n,\nu_{S_\omega}}(-t_n \pm t_{x_n,r_n}) \cap Q_{\nu_{S_\omega}}(x_n, r_n) \subset Q^-_{\nu_{S_\omega}}(x_n, r_n) \setminus R_{\tau/2,\nu_{S_\omega}}(x_n, r_n)$ *and, where we recall* $I(t_n, t) := (-t_n - t, -t_n + t)$,

$$\begin{aligned} \sup_{0<t\leq t_{x_n,r_n}} \frac{1}{|I(t_n,t)|} \int_{I(t_n,t)} \int_{Q_{\nu_{S_\omega}}(x_n,r_n)\cap T_{x_n,\nu_{S_\omega}}(-l)} \omega^-(x) d\mathcal{H}^{N-1} dl \\ \leq \int_{S\cap Q_{\nu_{S_\omega}}(x_n,r_n)} \omega^-\, d\mathcal{H}^{N-1} + O(\tau) r_n^{N-1}. \end{aligned} \tag{3.78}$$

Proof. By Proposition 3.28, we obtain a countable collection $\left\{Q_{\nu_{S_\omega}}(x_n, r_n')\right\}_{n=1}^\infty$ and a set $S' \subset S_\omega$ such that $\mathcal{H}^{N-1}(S_\omega \setminus S') < \frac{\tau}{2}$, $S' \subset \bigcup_{n=1}^\infty Q_{\nu_{S_\omega}}(x_n, r_n')$, $S' \cap Q_{\nu_{S_\omega}}(x_n, r_n') \subset R_{\tau/2,\nu_{S_\omega}}(x_n, r_n')$, and $(1+\tau^2)^{-1} r_n^{N-1} \leq \mathcal{H}^{N-1}\left(S \cap Q_{\nu_{S_\omega}}(x_n, r)\right) \leq (1+\tau^2) r^{N-1}$ for all $0 < r < r_n'$. Find $0 < \kappa < 1$ such that

$$\mathcal{H}^{N-1}\left(S' \setminus \bigcup_{n=1}^\infty Q_{\nu_{S_\omega}}(x_n, \kappa r_n')\right) < \frac{\tau}{2}, \text{ and let } S := S' \cap \left(\bigcup_{n=1}^\infty Q_{\nu_{S_\omega}}(x_n, \kappa r_n')\right).$$

Then $S \subset \bigcup_{n=1}^\infty Q_{\nu_{S_\omega}}(x_n, \kappa r_n')$ and $\mathcal{H}^{N-1}(S_\omega \setminus S) \leq \mathcal{H}^{N-1}(S_\omega \setminus S') + \mathcal{H}^{N-1}(S' \setminus S) \leq \tau$. Note that S satisfies Proposition 3.29 (1), (3), (4), and (5), and the collection $\left\{Q_{\nu_{S_\omega}}(x_n, \kappa r_n')\right\}_{n=1}^\infty$ satisfies Proposition 3.29 (2). Next, we apply Proposition 3.28 (4) with such $\kappa > 0$ to find t_n^κ, $t^\kappa_{x_n,r_n'}$ such that (3.77) holds. It suffices to set $r_n := \kappa r_n'$, $t_n := t_n^\kappa$, and $t_{x_n,r_n} := t^\kappa_{x_n,r_n'}$. □

The next lemma provides an approximation of $u \in SBV^2(\Omega) \cap L^\infty(\Omega)$ with functions u_n whose jump sets are more regular than that of u.

Lemma 3.30. *Let* $u \in SBV^2(\Omega) \cap L^\infty(\Omega)$ *with* $\mathcal{H}^{N-1}(S_u) < +\infty$, *and let* $\omega \in \mathcal{W}(\Omega)$ *be given. Then there exists a sequence* $\{u_n\}_{n=1}^\infty \subset SBV^2(\Omega) \cap L^\infty(\Omega)$ *such that the following hold:*

1. $\|u_n\|_{L^\infty} \leq \|u\|_{L^\infty}$;
2. $\mathcal{H}^{N-1}(\overline{S_{u_n}}) < +\infty$ *for each* $n \in \mathbb{N}$ *and*

$$\lim_{n\to\infty} \mathcal{H}^{N-1}(\overline{S_{u_n}} \setminus S_{u_n}) = 0;$$

3.

$$\lim_{n\to\infty}\left[\int_\Omega |\nabla u_n|^2\, \omega\, dx + \int_{S_{u_n}} \omega^- d\mathcal{H}^{N-1}\right] = \int_\Omega |\nabla u|^2\, \omega\, dx + \int_{S_u} \omega^- d\mathcal{H}^{N-1}. \tag{3.79}$$

Proof. We apply Lemma 3.5 to obtain a sequence $\{u_n\}_{n=1}^{\infty}$ such that Lemma 3.5 (2 - 5) hold. Since $\mathcal{H}^{N-1}(S_u) < +\infty$ by assumption and in view of Lemma 3.5 (4) and (5), we have $\mathcal{H}^{N-1}(\overline{S_{u_n}}) < +\infty$ for each $n \in \mathbb{N}$.
We write

$$\int_{S_{u_n}} \omega^- d\mathcal{H}^{N-1} = \int_{S_{u_n}\setminus S_u} \omega^- d\mathcal{H}^{N-1} + \int_{S_u} \omega^- d\mathcal{H}^{N-1} - \int_{S_u\setminus S_{u_n}} \omega^- d\mathcal{H}^{N-1}.$$

This, together with Lemma 3.5 (4) and the fact that $\|\omega\|_{L^\infty} < +\infty$, yields

$$\lim_{n\to\infty} \int_{S_{u_n}} \omega^- d\mathcal{H}^{N-1} = \int_{S_u} \omega^- d\mathcal{H}^{N-1}.$$

Moreover, by Lemma 3.5 (3), and again in view of $\|\omega\|_{L^\infty} < +\infty$, we conclude (3.79). □

The next proposition is key to proving Proposition 3.26. In Proposition 3.31 we construct a recovery sequence which converges to a function $u \in SBV^2(\Omega) \cap L^\infty(\Omega)$ with a regular enough jump set, such as those approximating function obtained in Lemma 3.30.

We summarize here the main ideas: We will modify most of S_u by replacing it with $(N-1)$ polyhedral sets located in the $-\nu_{S_\omega}$ side of S_ω, while ensuring that the L^1-norm of u and the L^2-norm of ∇u do not change much. This will be done by using a reflection argument around suitable hyperplanes (see (3.90)). We will cover the rest of S_u by using a finite collection of cubes, and change the value of u to 0 in those cubes (see (3.87)). Hence, in this way we replace the jump set of S_u by a finite union of $(N-1)$ - polyhedral sets.

Proposition 3.31. *Let $\omega \in \mathcal{W}(\Omega)$ and let $u \in SBV^2(\Omega) \cap L^\infty$ with $\mathcal{H}^{N-1}(\overline{S_u}) < +\infty$. Then there exists a sequence $\{(u_\varepsilon, v_\varepsilon)\}_{\varepsilon>0} \subset W^{1,2}(\Omega) \times W^{1,2}(\Omega)$ such that $u_\varepsilon \to u$ in L^1, $v_\varepsilon \to 1$ in L^1, $0 \le v_\varepsilon \le 1$ a.e., and*

$$\limsup_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_\varepsilon, v_\varepsilon) \le MS_\omega(u) + O(\mathcal{H}^{N-1}(\overline{S_u} \setminus S_u)).$$

Proof. Without lose of generality we assume that $MS_\omega(u) < +\infty$.

Step 1: Assume that $\mathcal{H}^{N-1}(S_\omega \,\Delta\, S_u) = 0$. Fix $\tau \in (0, 1/4)$. Applying Proposition 3.29 to ω we obtain a set S_τ, a collection $\mathcal{F}_\tau = \left\{Q_{\nu_{S_\omega}}(x_n, r_n)\right\}_{n=1}^{\infty}$, and corresponding $t_n \in (2.5\tau r_n, 3.5\tau r_n)$ and t_{x_n,r_n} for which (3.78) holds. Extract a finite collection $\mathcal{T}_\tau = \left\{Q_{\nu_{S_\omega}}(x_n, r_n)\right\}_{n=1}^{M_\tau}$ from $\mathcal{F}_\tau$ with $M_\tau > 0$ large enough such that

$$\mathcal{H}^{N-1}\left[S_\tau \setminus \bigcup_{n=1}^{M_\tau} Q_{\nu_{S_\omega}}(x_n, r_n)\right] < \tau,$$

we define $F_\tau := S_\tau \cap \left[\bigcup_{n=1}^{M_\tau} Q_{\nu_{S_\omega}}(x_n, r_n)\right]$, and note that

$$\mathcal{H}^{N-1}(S_u \setminus F_\tau) \le \mathcal{H}^{N-1}(S_u \setminus S_\tau) + \mathcal{H}^{N-1}(S_\tau \setminus F_\tau) < 2\tau. \tag{3.80}$$

Let $\tilde{\Omega}$ be an open bounded set such that $\Omega \subset\subset \tilde{\Omega}$ and, since $\partial\Omega$ is Lipschitz, using a reflection argument as in Lemma 7.1 in [**?**] we can extend u and ω to $\tilde{u} \in SBV^2(\tilde{\Omega}) \cap L^\infty(\tilde{\Omega})$

and $\tilde{\omega} \in \mathcal{W}(\tilde{\Omega})$, respectively, in a way that $\mathcal{H}^{N-1}(S_{\tilde{u}} \cap \partial\Omega) = \mathcal{H}^{N-1}(S_{\tilde{\omega}} \cap \partial\Omega) = 0$. Taking $\mathrm{dist}(\partial\Omega, \partial\tilde{\Omega})$ small enough, it is not restrictive to assume that

$$\mathcal{H}^{N-1}\left((\overline{S_{\tilde{u}}} \setminus S_{\tilde{u}}) \cap (\tilde{\Omega} \setminus \Omega)\right) \leq \mathcal{H}^{N-1}(\overline{S_u} \setminus S_u) \tag{3.81}$$

and

$$\int_{\tilde{\Omega}\setminus\Omega} \left(|\nabla \tilde{u}|^2 + |\tilde{u}|^2\right) dx + \mathcal{H}^{N-1}(S_{\tilde{u}} \setminus S_u) \leq O(\tau). \tag{3.82}$$

Define $A_\tau := \overline{S_{\tilde{u}}} \setminus F_\tau$ and recall $R_{\tau/2,\nu_{S_\omega}}(x_n, r_n)$ from Notation 3.3 (5). We show that the set $A_\tau \setminus \bigcup_{n=1}^{M_\tau} R_{\tau/2,\nu_{S_\omega}}(x_n, r_n)$ can be covered by a finite collection of cubes such that the sum of the $\mathcal{H}^{N-1}$ measure of the boundary of those cubes is at most $O(\tau)$. We first note that, in view of (3.80), (3.81), and (3.82),

$$\begin{aligned}
&\mathcal{H}^{N-1}(A_\tau)\\
&\leq \mathcal{H}^{N-1}(\overline{S_{\tilde{u}}} \setminus S_{\tilde{u}}) + \mathcal{H}^{N-1}(S_{\tilde{u}} \setminus F_\tau)\\
&\leq \mathcal{H}^{N-1}\left((\overline{S_{\tilde{u}}} \setminus S_{\tilde{u}}) \cap (\tilde{\Omega} \setminus \Omega)\right) + \mathcal{H}^{N-1}\left((\overline{S_{\tilde{u}}} \setminus S_{\tilde{u}}) \setminus (\tilde{\Omega} \setminus \Omega)\right) + \mathcal{H}^{N-1}(S_{\tilde{u}} \setminus F_\tau)\\
&\leq \mathcal{H}^{N-1}(\overline{S_u} \setminus S_u) + \mathcal{H}^{N-1}\left((\overline{S_{\tilde{u}}} \setminus S_{\tilde{u}}) \cap \Omega\right) + \mathcal{H}^{N-1}(S_{\tilde{u}} \setminus S_u) + \mathcal{H}^{N-1}(S_u \setminus F_\tau)\\
&\leq \mathcal{H}^{N-1}(\overline{S_u} \setminus S_u) + \mathcal{H}^{N-1}\left((\overline{S_{\tilde{u}}} \setminus S_{\tilde{u}}) \cap \Omega\right) + \mathcal{H}^{N-1}\left(S_{\tilde{u}} \cap (\tilde{\Omega} \setminus \Omega)\right) + 2\tau\\
&\leq 2\mathcal{H}^{N-1}(\overline{S_u} \setminus S_u) + O(\tau).
\end{aligned} \tag{3.83}$$

Let a'_τ denote the minimum distance between all cubes in $\mathcal{T}_\tau$ and

$$a_\tau := \frac{1}{4} \min\left\{\mathrm{dist}(\partial\Omega, \partial\tilde{\Omega}),\ a'_\tau,\ \tau,\ r_n,\ 1 \leq n \leq M_\tau\right\}. \tag{3.84}$$

In view of (3.83), and using the definition of Hausdorff measure, there exists a countable collection of balls $\{B(y_m, d_m/2)\}_{m=1}^{\infty}$, with center $y_m \in A_\tau$ and diameter $d_m > 0$, such that $A_\tau \subset \bigcup_{m\in\mathbb{N}} B(y_m, d_m/2)$, $\max_{m\in\mathbb{N}} \{d_m\} \leq \tau^2 a_\tau/2\sqrt{N}$, and

$$\sum_{m=1}^{\infty} \alpha(N-1) \left(\frac{d_m}{2}\right)^{N-1} \leq \mathcal{H}^{N-1}(A_\tau) + \tau \leq 2\mathcal{H}^{N-1}(\overline{S_u} \setminus S_u) + O(\tau)$$

where $\alpha(N-1)$ is a constant dependenting only on the dimension N (see [41], page 60), and in the last inequality we used (3.83).

We note that for each $m \in \mathbb{N}$, since $d_m \leq \tau^2 a_\tau/2\sqrt{N}$ and a_τ is at most one quarter of the minimum distance between all cubes in $\mathcal{T}_\tau$, there exists at most one index $n_m \in \{1, \ldots, M_\tau\}$ such that $Q_{\nu_{S_\omega}}(x_{n_m}, r_{n_m}) \in \mathcal{T}_\tau$ and

$$\overline{Q_{\nu_{S_\omega}}(x_{n_m}, r_{n_m})} \cap B(y_m, 2d_m) \neq \emptyset.$$

(It is of course possible that for some $m \in \mathbb{N}$ no such n_m exist). For every $B(y_m, d_m/2)$, $m \in \mathbb{N}$, define

$$Q_m := \begin{cases} Q_{\nu_{S_\omega(x_{n_m})}}(y_m, d_m/2) & \text{if there exists such } n_m \in \{1, \dots, M_\tau\} \\ Q_\nu(y_m, d_m) & \text{otherwise,} \end{cases}$$

where $\nu \in \mathcal{S}^{N-1}$ is an arbitrary direction (recall the notation from (3.36)). We have $B(y_m, d_m/2) \subset Q_m$ and

$$\sum_{m=1}^{\infty} \mathcal{H}^{N-1}(\partial Q_m) \le 2^N d_m^{N-1} \le 2^N (O(\tau) + \mathcal{H}^{N-1}(\overline{S_u} \setminus S_u)).$$

In view of Proposition 3.29 3,

$$F_\tau \cap Q_{\nu_{S_\omega}}(x_n, r_n) \subset R_{\tau/2, \nu_{S_\omega}}(x_n, r_n),$$

which, together with the fact that $F_\tau \subset \bigcup_{n=1}^{M_\tau} Q_{\nu_{S_\omega}}(x_n, r_n)$, implies that

$$A_\tau \setminus \bigcup_{n=1}^{M_\tau} R_{\tau/2, \nu_{S_\omega}}(x_n, r_n) = \overline{S_u} \setminus \bigcup_{n=1}^{M_\tau} R_{\tau/2, \nu_{S_\omega}}(x_n, r_n), \tag{3.85}$$

and hence, in view of (3.85) $A_\tau \setminus \bigcup_{n=1}^{M_\tau} R_{\tau/2, \nu_{S_\omega}}(x_n, r_n)$ is compact, as well as

$$A_\tau \cap \overline{\Omega} \setminus \bigcup_{n=1}^{M_\tau} R_{\tau/2, \nu_{S_\omega}}(x_n, r_n).$$

Therefore, we may extract a finite collection $\{Q_m\}_{m=1}^{Y_\tau}$, where $Y_\tau \in \mathbb{N}$, such that

$$A_\tau \setminus \bigcup_{n=1}^{M_\tau} R_{\tau/2, \nu_{S_\omega}}(x_n, r_n) \subset \bigcup_{m=1}^{Y_\tau} Q_m,$$

and further a $\{Q_m\}_{m=1}^{Y'_\tau} \subset \{Q_m\}_{m=1}^{Y_\tau}$, $Y'_\tau \le Y_\tau$, such that

$$A_\tau \cap \overline{\Omega} \setminus \bigcup_{n=1}^{M_\tau} R_{\tau/2, \nu_{S_\omega}}(x_n, r_n) \subset \bigcup_{m=1}^{Y'_\tau} Q_m. \tag{3.86}$$

We define $\bar{u}$ as follows:

$$\bar{u}(x) := \begin{cases} 0 & \text{if } x \in Q_m,\ 1 \le m \le Y_\tau, \\ \tilde{u}(x) & \text{otherwise.} \end{cases} \tag{3.87}$$

We remark that

$$\sum_{m=1}^{Y_\tau} \mathcal{H}^{N-1}(\partial Q_m) \le 2^N (2\tau + \mathcal{H}^{N-1}(\overline{S_u} \setminus S_u)) \text{ and } \sum_{m=1}^{Y_\tau} \mathcal{L}^N(Q_m) \le O(\tau). \tag{3.88}$$

Next, let U_n be the part of $Q_{\nu_{S_\omega}}(x_n, r_n)$ which lies between $T_{x_n,\nu_{S_\omega}}(\pm t_n)$, U_n^+ be the part above $T_{x_n,\nu_{S_\omega}}(t_n)$, and U_n^- be the part below $T_{x_n,\nu_{S_\omega}}(-t_n)$. We observe that, for each $1 \le n \le M_\tau$ fixed, since $r_n < \tau$ in Proposition 3.29,

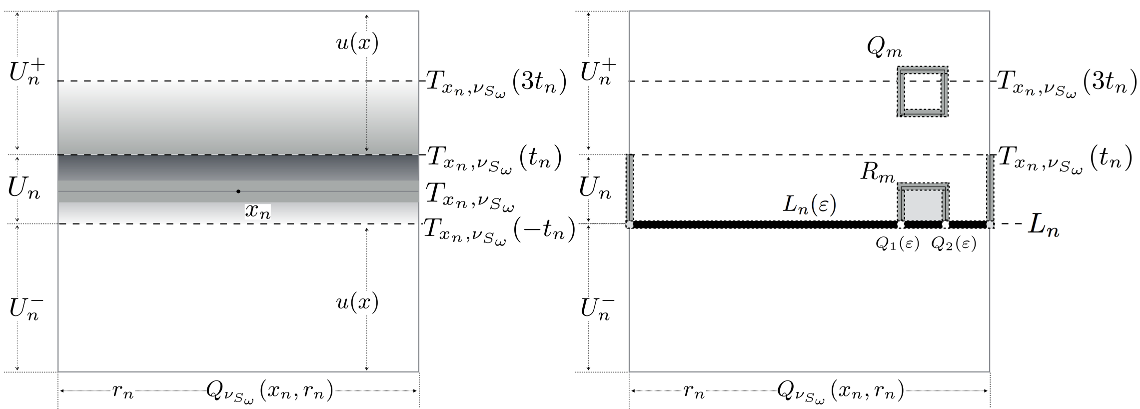
$$\mathcal{L}^N(U_n) = r_n^{N-1} \cdot 2t_n \le 7\tau r_n \cdot r_n^{N-1} \le 7\tau^2 r_n^{N-1}. \tag{3.89}$$

We define $\bar{u}_\tau$ as follows (see Figure 1a in page 66):

$$\bar{u}_\tau(x) := \begin{cases} \bar{u}\left(x + 2\mathrm{dist}(x, T_{x_n,\nu_{S_\omega}}(t_n))\nu_{S_\omega}(x_n)\right) & \text{if } x \in U_n, \\ \bar{u}(x) & \text{otherwise,} \end{cases} \tag{3.90}$$

and, for $1 \le m \le Y_\tau$ (see Figure 1b), where $R_{mn} := Q_m \cap U_n$ and $R_{mn}^+ := Q_m \cap U_n^+$,

$$R_m := \begin{cases} R_{mn}^+ - (d_m/2 - \mathrm{dist}(x_m, T_{x_n,\nu_{S_\omega}}(t_n)))\nu_{S_\omega}(x_n) & \text{if } \mathcal{L}^N(R_{mn}) > \mathcal{L}^N(R_{mn}^+) > 0, \\ R_{mn}^+ - (d_m/2 + \mathrm{dist}(x_m, T_{x_n,\nu_{S_\omega}}(t_n)))\nu_{S_\omega}(x_n) & \text{if } \mathcal{L}^N(R_{mn}^+) > \mathcal{L}^N(R_{mn}) > 0, \\ R_{mn}^+ - 2(\mathrm{dist}(x_m, T_{x_n,\nu_{S_\omega}}(t_n)))\nu_{S_\omega}(x_n) & \text{if } \mathcal{L}^N(R_{mn}^+) > \mathcal{L}^N(R_{mn}) = 0, \\ \emptyset & \text{otherwise.} \end{cases}$$



(A) Reflection construction in (3.90) (B) Recovery area

FIGURE 1. (A) depicts the construction in (3.90), with $\bar{u}$ in the light gray region, marked with color fill fainting toward above, reflected into the dark gray region, marked with color fill fainting toward below; the rectangle $R_{\tau/2,\nu_{S_\omega}}(x_n, r_n)$ is represented by the solid gray region in the middle of $Q_{\nu_{S_\omega}}(x_n, r_n)$.

In (B), the recovery region $(P_\tau)_\varepsilon$ is plotted by (dark, light) gray and black strips with dashed border. The region $L_n(\varepsilon)$ in which the core of the construction is undertaken is plotted as a dark strip. The small cubes $Q_1(\varepsilon)$ and $Q_2(\varepsilon)$ are plotted by light white cubes on $(L_n)_\varepsilon$.

We observe that

$$\mathcal{L}^N\left(\{x\in\Omega,\,\bar{u}(x)\neq\bar{u}_\tau(x)\}\right)=\mathcal{L}^N\left(\bigcup_{n=1}^{M_\tau}U_n\right)\leq\sum_{n=1}^{M_\tau}\mathcal{L}^N(U_n)\leq 7\tau^2\sum_{n=1}^{M_\tau}r_n^{N-1}\leq O(\tau),$$

where in the second last inequality we used (3.89), and in the last inequality we used Proposition 3.29 (5). We note that:

1. $\bar{u}_\tau$ is a reflection of $\bar{u}$ within the set with measure less than $O(\tau)$;
2. $\mathcal{L}^N(\{\bar{u}\neq u\})\leq\sum_{m=1}^{Y_\tau}\mathcal{L}^N(Q_m)\leq O(\tau)$;
3. $\bar{u}\in SBV^2(\Omega)\cap L^\infty(\Omega)$.

We conclude that

$$\lim_{\tau\to 0}\int_\Omega|\bar{u}_\tau-u|\,dx=0\text{ and }\lim_{\tau\to 0}\int_\Omega|\nabla\bar{u}_\tau-\nabla u|^2\,dx=0.\tag{3.91}$$

For simplicity of notation, in the rest of the proof of this proposition we shall abbreviate $Q_{\nu_{S_\omega}}(x_n,r_n)$ by Q_n, $T_{x_n,\nu_{S_u}}$ by T_{x_n}, and $T_{x_n,\nu_{S_u}}(-t_n)$ by $T_{x_n}(-t_n)$. Note that the jump set of $\bar{u}_\tau$ is contained in

$$P_\tau:=\bigcup_{n=1}^{M_\tau}[T_{x_n}(-t_n)\cap Q_n]\cup\bigcup_{n=1}^{M_\tau}\partial Q_n\cap\overline{U_n}\cup\bigcup_{m=1}^{Y_\tau}\partial Q_m\cup\bigcup_{m=1}^{Y_\tau}\partial R_m,$$

and note also that $S_{\bar{u}_\tau}\subset P_\tau$, with P_τ a union of finitely many $(N-1)$ - polyhedral sets. We also observe that, with $\mathrm{cl}(\cdot)$ denoting the closure of a set,

$$\begin{aligned}
&\mathcal{H}^{N-1}\left[\mathrm{cl}\left(\left(\bigcup_{n=1}^{M_\tau}\partial Q_n\cap\overline{U_n}\right)\cup\left(\bigcup_{m=1}^{Y_\tau}\partial Q_m\right)\cup\left(\bigcup_{m=1}^{Y_\tau}\partial R_m\right)\right)\right]\\
&\qquad\leq\sum_{n=1}^{M_\tau}\mathcal{H}^{N-1}(\partial Q_n\cap\overline{U_n})+\sum_{m=1}^{Y_\tau}\mathcal{H}^{N-1}(\partial Q_m)+\sum_{m=1}^{Y_\tau}\mathcal{H}^{N-1}(\partial R_m)\\
&\qquad\leq 2\tau+C\tau\sum_{n=1}^{\infty}r_n^{N-1}\tau+2\mathcal{H}^{N-1}(\overline{S_{\tilde{u}}}\setminus S_{\tilde{u}})\leq O(\tau)+4\mathcal{H}^{N-1}(\overline{S_u}\setminus S_u)<+\infty,
\end{aligned}\tag{3.92}$$

where we used Proposition 3.29 (5), (3.80), (3.81), (3.88), and the assumption that $\mathcal{H}^{N-1}(\overline{S_u})<+\infty$.

Recall a_τ from (3.84), and let $\varepsilon>0$ be such that

$$\varepsilon^2+\sqrt{\varepsilon}<<\min\{\,a_\tau,\,t_{x_n,r_n}\text{ for }1\leq n\leq M_\tau\}\,.\tag{3.93}$$

By Proposition 3.29 (6) we have

$$\varepsilon^2+\sqrt{\varepsilon}<t_{x_n,r_n}<t_n<\frac{1}{4}\tau r_n<r_n.$$

We set $u_{\tau,\varepsilon}:=(1-\varphi_\varepsilon)\bar{u}_\tau$, where φ_ε is such that $\varphi_\varepsilon\in C_c^\infty(\Omega;[0,1])$, $\varphi_\varepsilon\equiv 1$ on $(\overline{S_{\bar{u}_\tau}})_{\varepsilon^2/4}$, and $\varphi_\varepsilon\equiv 0$ in $\Omega\setminus(\overline{S_{\bar{u}_\tau}})_{\varepsilon^2/2}$. By (3.92) we have $\mathcal{H}^{N-1}(\overline{S_{\bar{u}_\tau}})<+\infty$ and hence $\{u_{\tau,\varepsilon}\}_{\varepsilon>0}\subset$

$W^{1,2}(\Omega)$. Moreover, using Lebesgue Dominated Convergence Theorem and (3.91), we conclude that $u_{\tau,\varepsilon} \to u$ in $L^1(\Omega)$.

Consider the sequence $\{v_{\tau,\varepsilon}\}_{\varepsilon>0} \subset W^{1,2}(\Omega)$ given by $v_{\tau,\varepsilon}(x) := \tilde{v}_\varepsilon(d_\tau(x))$ where $d_\tau(x) := \operatorname{dist}(x, P_\tau)$, and $\tilde{v}_\varepsilon \in W^{1,2}_{\mathrm{loc}}(\mathbb{R})$ is defined by

$$\tilde{v}_\varepsilon(t) := \begin{cases} 0 & \text{if } t \leq \varepsilon^2, \\ -e^{-\frac{1}{2}\frac{t-\varepsilon^2}{\varepsilon}} + 1 & \text{if } \varepsilon^2 \leq t \leq \sqrt{\varepsilon} + \varepsilon^2, \\ 1 - e^{-\frac{1}{2\sqrt{\varepsilon}}} & \text{if } t > \sqrt{\varepsilon} + \varepsilon^2. \end{cases} \tag{3.94}$$

An explicit computation shows that

$$\tilde{v}_\varepsilon'(t) = \frac{1}{2\varepsilon}(1 - \tilde{v}_\varepsilon(t)) \tag{3.95}$$

for $\varepsilon^2 \leq t \leq \sqrt{\varepsilon} + \varepsilon^2$, and we remark that

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} e^{-\frac{1}{2\sqrt{\varepsilon}}} = 0, \tag{3.96}$$

and

$$-\frac{d}{dt}\left(\frac{1}{2}\left(1 - \tilde{v}_\varepsilon(t)\right)^2\right) = \left(1 - \tilde{v}_\varepsilon(t)\right)\tilde{v}_\varepsilon'(t) \geq 0. \tag{3.97}$$

Moreover, since $S_{u_\tau} \subset P_\tau$, by (3.91) we conclude that

$$\int_\Omega |\nabla u_{\tau,\varepsilon}|^2 v_{\tau,\varepsilon}^2 \, \omega \, dx \leq \int_\Omega |\nabla \bar{u}_\tau|^2 \, \omega \, dx \leq \int_\Omega |\nabla u|^2 \, \omega \, dx + O(\tau). \tag{3.98}$$

Define $L_n := T_{x_n}(-t_n) \cap Q_n$ and $L_n(\varepsilon) := (T_{x_n}(-t_n) \cap Q_n)_\varepsilon$ and, without loss of generality, assume that there exists only one R_m such that $\mathcal{H}^{N-1}(T_{x_n}(-t_n) \cap Q_n \cap R_m) > 0$ (recall $Y_\tau < +\infty$). We claim that

$$\begin{aligned} &\int_{L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \left[\varepsilon |\nabla v_{\tau,\varepsilon}|^2 + \frac{1}{4\varepsilon}(1 - v_{\tau,\varepsilon})^2\right] \omega \, dx \\ &\leq \int_{L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \left[\varepsilon |\nabla v_{L_n,\varepsilon}|^2 + \frac{1}{4\varepsilon}(1 - v_{L_n,\varepsilon})^2\right] \omega \, dx + O(\varepsilon) r_n^{N-1} \end{aligned} \tag{3.99}$$

where $v_{L_n,\varepsilon}(x) := \tilde{v}_\varepsilon(\operatorname{dist}(x, L_n))$. Indeed, let $\{y_0, y_1\} := T_{x_n}(-t) \cap Q_n \cap \partial R_m$ and observe that, where we abbreviate $\left[\varepsilon \left|\nabla v_{(\cdot),\varepsilon}\right|^2 + \frac{1}{4\varepsilon}(1 - v_{(\cdot),\varepsilon})^2\right]$ by $\varphi_{(\cdot),\varepsilon}$, and $Q_{\nu_{S_\omega(x_n)}}(y_i, \varepsilon^2 + \sqrt{\varepsilon})$ by $Q_i(\varepsilon)$, $i = 1, 2$ (see Figure 1b where $Q_i(\varepsilon)$ is represented by a white cube),

$$\varphi_{\tau,\varepsilon}(x) = \varphi_{L_n,\varepsilon}(x) \text{ for } x \in L_n(\varepsilon^2 + \sqrt{\varepsilon}) \cap Q_n \setminus (Q_1(\varepsilon) \cup Q_2(\varepsilon)),$$

and

$$\begin{aligned} &\int_{L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \varphi_{\tau,\varepsilon} \, \omega \, dx \\ &= \int_{L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n \setminus (Q_1(\varepsilon)\cup Q_2(\varepsilon))} \varphi_{\tau,\varepsilon} \, \omega \, dx + \int_{Q_n \cap (Q_1(\varepsilon)\cup Q_2(\varepsilon))} \varphi_{\tau,\varepsilon} \, \omega \, dx \end{aligned}$$

$$\leq \int_{L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \varphi_{L_n,\varepsilon}\,\omega\,dx + \int_{Q_n\cap(Q_1(\varepsilon)\cup Q_2(\varepsilon))} \varphi_{\tau,\varepsilon}\,\omega\,dx.$$

Invoking Proposition 5.1 in [6] we have, for $\varepsilon_0 > 0$ fixed,

$$\begin{aligned}
\limsup_{\varepsilon\to 0} \int_{Q_n\cap(Q_1(\varepsilon)\cup Q_2(\varepsilon))} \varphi_{\tau,\varepsilon}\,\omega\,dx &\leq \limsup_{\varepsilon\to 0} \int_{Q_n\cap(Q_1(\varepsilon_0)\cup Q_2(\varepsilon_0))} \varphi_{\tau,\varepsilon}\,\omega\,dx \\
&\leq \|\omega\|_{L^\infty} \limsup_{\varepsilon\to 0} \int_{Q_n\cap(Q_1(\varepsilon_0)\cup Q_2(\varepsilon_0))} \varphi_{\tau,\varepsilon}\,dx \\
&\leq \|\omega\|_{L^\infty}\, \mathcal{H}^{N-1}(\overline{(Q_1(\varepsilon_0)\cup Q_2(\varepsilon_0))\cap(L_n\cup\partial R_m)}) \leq 2\,\|\omega\|_{L^\infty}\, 2^N \varepsilon_0^{N-1},
\end{aligned}$$

and, by letting $\varepsilon_0 \searrow 0$ on the right hand side, we conclude (3.99).

Next, for each $1 \leq n \leq M_\tau$ fixed, we observe that, using Fubini's Theorem,

$$\begin{aligned}
&\int_{L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \left[\varepsilon\,|\nabla v_{L_n,\varepsilon}|^2 + \frac{1}{4\varepsilon}(1-v_{L_n,\varepsilon})^2\right]\omega\,dx \\
&= \int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}} \left(\left[\varepsilon\left|\tilde{v}_\varepsilon'(l)\right|^2 + \frac{1}{4\varepsilon}(1-\tilde{v}_\varepsilon(l))^2\right]\right)\cdot \\
&\quad\cdot\left(\int_{\{d_\tau(y)=l\}\cap L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \omega(y)\,d\mathcal{H}^{N-1}(y)\right) dl + \frac{1}{4\varepsilon}\int_{L_n(\varepsilon^2)} \omega(x)dx \qquad (3.100)\\
&= -\int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}} \frac{1}{2\varepsilon}\frac{d}{dl}\left[(1-\tilde{v}_\varepsilon(l))^2\right] \int_{\{d_\tau(y)\leq l\}\cap L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \omega(y)\,dy\,dl \\
&\quad + \mathcal{A}^n_\omega(\varepsilon^2+\sqrt{\varepsilon}) - \mathcal{A}^n_\omega(\varepsilon^2) + \frac{1}{4\varepsilon}\int_{L_n(\varepsilon^2)} \omega(x)dx,
\end{aligned}$$

where

$$\mathcal{A}^n_\omega(t) := \frac{1}{2\varepsilon}(1-\tilde{v}_\varepsilon(t))^2 \int_{\{d_\tau(x)\leq t\}\cap L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \omega(y)\,dy.$$

Since $\omega \in L^\infty(\Omega)$ we have

$$\mathcal{A}^n_\omega(\varepsilon^2+\sqrt{\varepsilon}) \leq \frac{1}{2\varepsilon} e^{-\frac{1}{2\sqrt{\varepsilon}}}\,\|\omega\|_{L^\infty}\left[2(\varepsilon^2+\sqrt{\varepsilon})[r_n + (\varepsilon^2+\sqrt{\varepsilon})]^{N-1}\right] \leq O(\varepsilon) r_n^{N-1}, \qquad (3.101)$$

and

$$\begin{aligned}
&-\int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}} \frac{1}{2\varepsilon}\frac{d}{dl}\left[(1-\tilde{v}_\varepsilon(l))^2\right] \int_{\{d_\tau(y)\leq l\}\cap L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \omega(y)\,dy\,dl \\
&\qquad = \int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}} 2l\left(-\frac{1}{2\varepsilon}\frac{d}{dl}\left[(1-\tilde{v}_\varepsilon(l))^2\right]\right)\left[\frac{1}{2l}\int_{\{d_\tau(y)\leq l\}\cap L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \omega(x)\,dx\right] dl.
\end{aligned}$$

Recalling the notation from Proposition 3.29 and the fact that $\omega^-(x_n) \le \|\omega\|_{L^\infty}$, we have for $l \in (\varepsilon^2, \varepsilon^2 + \sqrt{\varepsilon})$

$$\begin{aligned}
&\frac{1}{2l}\int_{\{d_\tau(y)\le l\}\cap L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \omega(x)\,dx \\
&\quad\le \sup_{0<t\le\varepsilon^2+\sqrt{\varepsilon}}\left(\frac{1}{|I(t_n,t)|}\int_{I(t_n,t)}\int_{Q_n(x_n,r_n)\cap T_{x_n}(-s)}\omega(x)d\mathcal{H}^{N-1}ds\right)\\
&\quad\le \int_{S_\tau\cap Q_n(x_n,r_n)}\omega^-(x)\,d\mathcal{H}^{N-1} + O(\tau)r_n^{N-1},
\end{aligned}$$

where, in view of (3.93), we used (3.78) in the last inequality. Therefore, by (3.97) we obtain

$$\begin{aligned}
&-\int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}}\frac{1}{2\varepsilon}\frac{d}{dl}\left[(1-\tilde{v}_\varepsilon(l))^2\right]\int_{\{d_\tau(y)\le l\}\cap L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n}\omega(x)\,dx\,dl\\
&\le 2\left(\int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}}-\frac{1}{2\varepsilon}\frac{d}{dl}\left[(1-\tilde{v}_\varepsilon(l))^2\right]l\,dl\right)\left(\int_{S_\tau\cap Q_n(x_n,r_n)}\omega^-(x)\,d\mathcal{H}^{N-1}+O(\tau)r_n^{N-1}\right).
\end{aligned}\tag{3.102}$$

An integration by parts, and using (3.95), yields

$$\begin{aligned}
&\int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}}-\frac{1}{2\varepsilon}\frac{d}{dl}\left[(1-\tilde{v}_\varepsilon(l))^2\right]l\,dl\\
&=\int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}}\frac{1}{2\varepsilon}(1-\tilde{v}_\varepsilon(l))^2dl-\frac{1}{2\varepsilon}(\varepsilon^2+\sqrt{\varepsilon})(1-\tilde{v}_\varepsilon(\varepsilon^2+\sqrt{\varepsilon}))^2+\frac{\varepsilon^2}{2\varepsilon}(1-\tilde{v}_\varepsilon(\varepsilon^2))^2\\
&\le\int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}}2\varepsilon\left|\tilde{v}_\varepsilon'(l)\right|^2dl+\frac{\varepsilon^2}{2\varepsilon}(1-\tilde{v}_\varepsilon(\varepsilon^2))^2\\
&=\int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}}-\frac{1}{2}\frac{d}{dl}\left(e^{-\frac{l-\varepsilon^2}{\varepsilon}}\right)dl+\frac{\varepsilon}{2}(1-\tilde{v}_\varepsilon(\varepsilon^2))^2\\
&=\frac{1}{2}\left(1-e^{-\frac{1}{\sqrt{\varepsilon}}}\right)+\frac{\varepsilon}{2}(1-\tilde{v}_\varepsilon(\varepsilon^2))^2\le\frac{1}{2}+\frac{1}{2}\varepsilon,
\end{aligned}$$

which, together with (3.102) and Proposition 3.29 (4), gives

$$\begin{aligned}
&-\int_{\varepsilon^2}^{\varepsilon^2+\sqrt{\varepsilon}}\frac{1}{2\varepsilon}\frac{d}{dl}\left[(1-\tilde{v}_\varepsilon(l))^2\right]\int_{\{d_\tau(y)\le l\}\cap L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n}\omega^-(x)\,dx\,dl\\
&\qquad\le\int_{S_\tau\cap Q(x_n,r_n)}\omega(x)\,d\mathcal{H}^{N-1}+O(\tau)r_n^{N-1}\\
&\qquad\qquad+\varepsilon\,\|\omega\|_{L^\infty}\,\mathcal{H}^{N-1}(S_\tau\cap Q(x_n,r_n))+\varepsilon O(\tau)r_n^{N-1}\\
&\qquad\le\int_{S_\tau\cap Q(x_n,r_n)}\omega^-(x)\,d\mathcal{H}^{N-1}+O(\tau)r_n^{N-1}+O(\varepsilon)O(\tau)r_n^{N-1}.
\end{aligned}\tag{3.103}$$

Hence, in view of (3.99), (3.100), (3.101), (3.103), and since $A_\omega(\varepsilon^2) \geq 0$, we obtain that

$$\begin{aligned}
&\int_{L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n} \left[\varepsilon\left|\nabla v_{\tau,\varepsilon}\right|^2 + \frac{1}{4\varepsilon}(1-v_{\tau,\varepsilon})^2\right]\omega\, dx \\
&\qquad\qquad \leq \int_{S_\tau\cap Q(x_n,r_n)} \omega^-(x) d\mathcal{H}^{N-1} + O(\tau) r_n^{N-1} + O(\varepsilon)O(\tau) r_n^{N-1} + O(\varepsilon) r_n^{N-1},
\end{aligned}$$

which implies that

$$\begin{aligned}
&\int_{\bigcup_{n=1}^{M_\tau}\left(L_n(\varepsilon^2+\sqrt{\varepsilon})\cap Q_n\right)} \left[\varepsilon\left|\nabla v_{\tau,\varepsilon}\right|^2 + \frac{1}{4\varepsilon}(1-v_{\tau,\varepsilon})^2\right]\omega\, dx \\
&\quad \leq \int_{S_\tau} \omega^-(x) d\mathcal{H}^{N-1} + (O(\tau) + O(\varepsilon)O(\tau) + O(\varepsilon)) \sum_{n=1}^{M_\tau} r_n^{N-1} \qquad (3.104) \\
&\quad \leq \int_{S_u} \omega^-(x) d\mathcal{H}^{N-1} + (O(\tau) + O(\varepsilon)O(\tau) + O(\varepsilon)) \left(4\mathcal{H}^{N-1}(S_u) + \|\omega\|_{L^\infty(\Omega)}\right).
\end{aligned}$$

Next we define $L_0 := \left(\bigcup_{m=1}^{Y'_\tau} \partial Q_m \cup \bigcup_{m=1}^{Y'_\tau} \partial R_m\right) \cup \left(\bigcup_{n=1}^{M_\tau} \partial Q_n \cap \overline{U}_n\right)$, where we recall Y'_τ from (3.86), L_0 is a finite union of $(N-1)$ - polyhedral sets, with $L_0(\varepsilon) := (L_0)_\varepsilon \subset \tilde{\Omega}$. Then, invoking Proposition 5.1 in [6] and the calculations within, we conclude that

$$\begin{aligned}
&\limsup_{\varepsilon\to 0} \int_{L_0(\varepsilon^2+\sqrt{\varepsilon})} \left[\varepsilon\left|\nabla v_{\tau,\varepsilon}\right|^2 + \frac{1}{4\varepsilon}(1-v_{\tau,\varepsilon})^2\right]\omega\, dx \\
&\qquad \leq \|\omega\|_{L^\infty} \limsup_{\varepsilon\to 0} \int_{L_0(\varepsilon^2+\sqrt{\varepsilon})} \left[\varepsilon\left|\nabla v_{\tau,\varepsilon}\right|^2 + \frac{1}{4\varepsilon}(1-v_{\tau,\varepsilon})^2\right] dx \\
&\qquad \leq \|\omega\|_{L^\infty} \limsup_{\varepsilon\to 0} \frac{1}{2\varepsilon} \left|\left\{x\in\Omega : \operatorname{dist}(x,\overline{L_0}) < \varepsilon\right\}\right| \\
&\qquad = \|\omega\|_{L^\infty} \mathcal{H}^{N-1}\left(\overline{L_0}\right) \leq \|\omega\|_{L^\infty} \mathcal{H}^{N-1}\left(\left(\overline{S_u}\setminus S_u\right)\cap\overline{\Omega}\right) + O(\tau),
\end{aligned}$$

where in the last inequality we used (3.92) and in the last equality we used Theorem 3.2.39 in [42]. Hence, we have

$$\begin{aligned}
&\int_{L_0(\varepsilon^2+\sqrt{\varepsilon})} \left[\varepsilon\left|\nabla v_{\tau,\varepsilon}\right|^2 + \frac{1}{4\varepsilon}(1-v_{\tau,\varepsilon})^2\right]\omega\, dx \\
&\qquad \leq \|\omega\|_{L^\infty}\left[\mathcal{H}^{N-1}\left(\left(\overline{S_u}\setminus S_u\right)\cap\overline{\Omega}\right) + O(\tau) + O(\varepsilon)\right] \qquad (3.105) \\
&\qquad = \|\omega\|_{L^\infty}\left[\mathcal{H}^{N-1}\left(\left(\overline{S_u}\setminus S_u\right)\right) + O(\tau) + O(\varepsilon)\right].
\end{aligned}$$

Furthermore, by (3.94) and (3.96) we have that

$$\begin{aligned}
&\int_{\Omega\setminus(P_\tau)_{\varepsilon^2+\sqrt{\varepsilon}}} \left[\varepsilon\left|\nabla v_{\tau,\varepsilon}\right|^2 + \frac{1}{4\varepsilon}(1-v_{\tau,\varepsilon})^2\right]\omega\, dx \\
&\qquad \leq \frac{1}{4\varepsilon} e^{-\frac{1}{\sqrt{\varepsilon}}} \|\omega\|_{L^\infty} \mathcal{L}^N(\Omega) \leq O(\varepsilon). \qquad (3.106)
\end{aligned}$$

We note that, for each $1 \leq n \leq M_\tau$ fixed,

$$\left[\left(\partial Q_n \cap \overline{U_n}\right) \cup (T_{x_n}(-t_n) \cap Q_n)\right]_\varepsilon = \left(\partial Q_n \cap \overline{U_n}\right)_\varepsilon \cup \left[(T_{x_n}(-t_n) \cap Q_n)_\varepsilon \cap Q_n\right],$$

and in view of Proposition 3.29 (5), (3.104), (3.105), and (3.106), we have that

$$\begin{aligned}&\int_\Omega \left[\varepsilon\left|\nabla v_{\tau,\varepsilon}\right|^2 + \frac{1}{4\varepsilon}(1-v_{\tau,\varepsilon})^2\right]\omega\,dx \\ &\leq \int_{S_u} \omega^-(x)\,d\mathcal{H}^{N-1} + O(\varepsilon) + O(\tau) + O(\varepsilon)O(\tau) + \|\omega\|_{L^\infty}\,\mathcal{H}^{N-1}\left(\left(\overline{S_u}\setminus S_u\right)\right).\end{aligned} \tag{3.107}$$

Recall that $S_{\bar{u}_\tau} \subset P_\tau$. Hence, also by (3.98) and (3.107), for each $\tau > 0$ we may choose $\varepsilon(\tau)$ such that

$$\int_\Omega \left|\nabla u_{\tau,\varepsilon(\tau)}\right|^2 v^2_{\tau,\varepsilon(\tau)}\,\omega\,dx \leq \int_\Omega |\nabla u|^2\,\omega\,dx + O(\tau),$$

and

$$\begin{aligned}&\int_\Omega \left[\varepsilon\left|\nabla v_{\tau,\varepsilon(\tau)}\right|^2 + \frac{1}{4\varepsilon}(1-v_{\tau,\varepsilon(\tau)})^2\right]\omega\,dx \\ &\quad\leq \int_{S_u} \omega^-(x)d\mathcal{H}^{N-1} + O(\tau) + \|\omega\|_{L^\infty}\,\mathcal{H}^{N-1}\left(\left(\overline{S_u}\setminus S_u\right)\right).\end{aligned}$$

It suffices to define the recovery sequence $\{(u_\tau, v_\tau)\}_{\tau>0}$ by $u_\tau := u_{\tau,\varepsilon(\tau)}$ and $v_\tau := v_{\tau,\varepsilon(\tau)}$.

Step 2: In the general case in which $\mathcal{H}^{N-1}(S_\omega\,\Delta\,S_u) > 0$, we may apply the same construction in Step 1 to S_u, since it suffices to notice that $\omega^-(x) = \omega(x)$ if $x \in S_u \setminus S_\omega$. □

Proof of Proposition 3.26. If $u \in L^1(\Omega)\cap L^\infty(\Omega)$ is such that $MS_\omega(u) = +\infty$, then there is nothing to prove. Suppose that $MS_\omega(u) < +\infty$. Then $u \in SBV^2(\Omega)\cap L^\infty(\Omega)$ and $\mathcal{H}^{N-1}(S_u) < +\infty$. We apply Lemma 3.30 to obtain a sequence $\{u_n\}_{n=1}^\infty \subset SBV^2(\Omega)\cap L^\infty(\Omega)$ such that Lemma 3.30 (1), (2), and (3) hold. Then,

$$\lim_{n\to\infty} \mathcal{H}^{N-1}(\overline{S_{u_n}}\setminus S_{u_n}) = 0, \tag{3.108}$$

and

$$\lim_{n\to\infty}\left(\int_\Omega |\nabla u_n|^2\,\omega\,dx + \int_{S_{u_n}} \omega^- d\mathcal{H}^{N-1}\right) = \int_\Omega |\nabla u|^2\,\omega\,dx + \int_{S_u} \omega^- d\mathcal{H}^{N-1}. \tag{3.109}$$

By Proposition 3.31, for every $n \in \mathbb{N}$ we may construct a sequence $\{u_{n,\varepsilon}, v_{n,\varepsilon}\}_{\varepsilon>0} \subset W^{1,2}(\Omega)\times W^{1,2}(\Omega)$ such that $u_{n,\varepsilon}\to u_n$, $v_{n,\varepsilon}\to 1$, $0 \leq v_{n,\varepsilon} \leq 1$, and

$$\limsup_{\varepsilon\to 0} AT_{\omega,\varepsilon}(u_{n,\varepsilon}, v_{n,\varepsilon}) \leq MS_\omega(u_n) + \|\omega\|_{L^\infty}\,O(\mathcal{H}^{N-1}(\overline{S_{u_n}}\setminus S_{u_n})).$$

A diagonal argument, together with (3.108) and (3.109), concludes the proof. □

Proof of Theorem 1.2. The lim inf inequality follows from Proposition 3.24. On the other hand, for any given $u \in GSBV(\Omega)$ such that $MS_\omega(u) < \infty$, we have, by the Lebesgue Monotone Convergence Theorem,

$$MS_\omega(u) = \lim_{K\to\infty} MS_\omega(K \wedge u \vee -K),$$

and a diagonal argument, together with Proposition 3.26, concludes the proof. □

A direct inspection of the proof of Theorem 1.2 shows that the properties of SBV functions, more specifically, of the distributional derivative of ω, are only needed to ensure that Theorem 3.4 holds. Indeed, if $\omega \in C(\overline{\Omega})$, then for all $x_0 \in \Gamma$, where Γ is any given $\mathcal{H}^{N-1}$ -rectifiable set,

$$\lim_{r\to 0} \fint_{B(x_0,r)\cap H_{\nu_\Gamma}(x_0)^\pm} |u(x) - u(x_0)|^{\frac{N}{N-1}}\, dx = 0,$$

and

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon^{N-1}} \int_{\Gamma\cap Q_{\nu_\Gamma}(x_0,\varepsilon)} |u(x) - u(x_0)|\, d\mathcal{H}^{N-1}(x) = 0.$$

Proof of Theorem 1.3. Here we only highlight the main modifications needed in the proofs of Proposition 3.24 and Proposition 3.26.

Indeed, to prove Proposition 3.24 with $\omega \in C(\overline{\Omega})$ we only need the following modification:

1. In (3.23) we have

$$\liminf_{n\to\infty} \int_{t_n^1}^{t_n^2} \left[\varepsilon_n \left|\left(v_{\varepsilon(n)}\right)'\right|^2 + \frac{1}{4\varepsilon(n)}(1 - v_{\varepsilon(n)})^2\right] \omega(x)\, dx \geq \omega(0),$$

   because for $\omega \in C(I)$,

$$\liminf_{n\to\infty} \operatorname*{ess\,inf}_{r\in(t_n^1,t_n^2)} \omega(r) = \liminf_{n\to\infty} \inf_{r\in(t_n^1,t_n^2)} \omega(r) = \omega(0).$$

2. The proof of Lemma 3.18 with $\omega \in C(\overline{\Omega})$ can be obtained directly from Theorem 9.1 in [51].

By adapting to above modifications, the version of Proposition 3.24 with $\omega \in C(\overline{\Omega})$ can be now obtained following the argument provided in Section 3.2 *mutatis mutandis*, taking (1) and (2) into consideration.

Regarding the proof of Proposition 3.26 with $\omega \in C(\overline{\Omega})$, the only modification we need to make is (3.69). We must show that for any $u \in SBV(\Omega)$ such that $MS_\omega(u) < +\infty$, the following holds:

$$\begin{aligned} \sup_{0<t\leq t_{0,r}} \frac{1}{|I(t_0,t)|} \int_{I(t_0,t)} \int_{Q^-_{\nu_{S_u}}(x_0,r)\cap T_{x_0,\nu_{S_u}}(-l)} \omega(x)\, d\mathcal{H}^{N-1}(x)\, dl \\ \leq \int_{S_u\cap Q^-_{\nu_{S_u}}(x_0,r)} \omega(x)\, d\mathcal{H}^{N-1} + O(\tau) r^{N-1}. \end{aligned}$$

This can be obtained using the proof of Proposition 3.27 with (3.74) is replaced by

$$\fint_{Q^-(x_0,r)} |\omega(x) - \omega(x_0)|\, dx \leq \tau^2.$$

We conclude Theorem 1.3 by using the same arguments as in the proof of Theorem 1.2, taking the above modifications into consideration. □

Chapter 4. Some insights from finite resolution images

In Chapter 4 we collect some results which were investigated in the early stage of my research in image processing.

4.1. **The finite resolution image and the unavoidable noise during acquisition.** As we stated in the Introduction, in one dimension a finite $N \in \mathbb{N}$ resolution level image captured by a real world digital camera is a piecewise constant function $u_{c,N}$ which is related to u_c via its averages

$$u_{c,N}(x) := \fint_{I_N(k)} u_c\, dx \text{ for } x \in I_N(k),$$

where $I_N(k) := ((k-1)/N, k/N)$, for $1 \leq k \leq N$, and where we set

$$\mathcal{I}_N := \{I_N(k),\ 0 \leq k \leq N\}.$$

Definition 4.1. *We say that a piecewise constant function is an* image with resolution level N *if it is constant in each* $I_N(k) \in \mathcal{I}_N$*.*

The principal sources of noise in digital images are introduced during acquisition, for example, the sensor noise caused by poor illumination, high temperature, and circuity of a scanner. Other possible sources could be digital error during the transmission, and the unavoidable shot noise of an photon detector. The noise is only generated during the acquiring of the image, i.e., it is only added to $u_{c,N}$; and each time we acquire an image, we produce a different noise η_N. Therefore, we propose to use a piecewise constant function η_N over $\mathcal{I}_N$ to represent the noise at the resolution level $N \in \mathbb{N}$, and we write

$$u_{\eta,N} := u_{c,N} + \eta_N.$$

That is, when a image is taken with resolution $N \in \mathbb{N}$, although we only wish to observe $u_{c,N}$, the noise η_N is an unavoidable by-product, and hence the corrupted image $u_{\eta,N}$ is produced.

Since $u_{\eta,N}$ represents an image data, we may assume (after rescaling) that

$$\|u_{\eta,N}\|_{L^\infty} \leq 1.$$

When $N \to \infty$, $u_{c,N} \to u_c$ in L^2, but since η_N is randomly generated, although for a fixed N, $u_{c,N}$ is fixed, η_N would vary. As it often assumed in the literatures in image reconstruction papers (see, e.g., [21]), we also assume that

$$\int_Q \eta_N \, dx = 0. \tag{4.1}$$

Moreover, we use $\eta_N(I_N(k))$ to denote the value of $\eta_N(x)$ for $x \in I_N(k)$.

4.2. **The total variation and some preliminary results.** We start by introducing notations that will be used in the sequel.

Notation 4.2. Recall that $I := (0,1) \subset \mathbb{R}$ and $M \in \mathbb{N}$ is a positive integer.

1. we say a function w is a *piecewise constant function with M pieces* if there exist M intervals $I_M(j) := (x_j, x_{j+1})$, where $0 = x_1 < \cdots < x_j < \cdots < x_M = 1$, such that w is a constant in each $I_M(j)$. Moreover, we use $w(I_M(k))$ to denote the value of $w(x)$ for $x \in I_M(j)$, $1 \le j \le M$;
2. given a piecewise constant function ω with M pieces, we say that $I_M(j)$, $1 < j < M$, is a *step region* of w if

$$w(I_M(j-1)) \le w(I_M(j)) \le w(I_M(j+1)) \text{ or } w(I_M(j-1)) \ge w(I_M(j)) \ge w(I_M(j+1));$$

   and $(I_M(j))$ is a *high extreme region* of ω if

$$w(I_M(j)) > \max\{w(I_M(j-1)), w(I_M(j+1))\}$$

   and $(I_M(j))$ is a *low extreme region* of ω if

$$w(I_M(j)) < \min\{w(I_M(j-1)), w(I_M(j+1))\}. \tag{4.2}$$

3. we say $I_M(1)$ is a *high (low) boundary regions* of ω if $w(I_M(1)) > (<) w(I_M(2))$, and $I_M(M)$ is a *high (low) boundary regions* of ω if $w(I_M(M)) > (<) w(I_M(M-1))$, respectively.
4. we use $\mathcal{C}_E(w)$ to denote the collection of extreme regions, $\mathcal{C}_B(w)$ the collection of boundary regions, and $\mathcal{C}_I(w)$ the collection of step regions.

Note: By (1.24), $u_{c,N}$ is indeed a piecewise constant function with N pieces.

Definition 4.3. *We say a piecewise constant function is an* image with resolution level N *if it is a constant in each $I_N(k) \in \mathcal{I}_N$ ($Q_N(i,j)$ in two dimensions).*

Recall the reconstruction operator $\mathcal{L}$ from (1.25).

Definition 4.4. *Let $v \in L^\infty(I)$ be given. We say that $\alpha_s(v) \in [0, +\infty)$ is the* stopping time *for v if*

$$\mathscr{L}(\alpha_s, v) = \mathscr{L}(\alpha_s + \alpha, v) =: C(v) \textit{ and } \mathscr{L}(\alpha_s, v) \ne \mathscr{L}(\alpha_s - \alpha, v) \tag{4.3}$$

for all $\alpha > 0$, where $C(v)$ is a constant depends on v.

Theorem 4.5 ([72], Theorem 2)**.** *Suppose that the function u_0 is piecewise constant with M pieces, and let α be small enough. Then the unique solution $u_\alpha := \mathcal{L}(\alpha, u_0)$ is also piecewise constant with the same number of pieces of u_0, and we have*

$$\begin{aligned} u_\alpha(I_M(j)) &= u_0(I_M(j)) \mp \frac{2}{|I_M(j)|}\alpha, \;\; \text{if } I_M(j) \text{ is a high (low) extremum region,} \\ u_\alpha(I_M(j)) &= u_0(I_M(j)), \;\; \text{if } I_M(j) \text{ is a step region,} \\ u_\alpha(I_M(j)) &= u_0(I_M(j)) \mp \frac{1}{|I_M(j)|}\alpha, \;\; \text{if } I_M(j) \text{ is a high (low) boundary region.} \end{aligned}$$

Moreover, for α is large enough, the function u_α is a constant.

Notation 4.6. Let $v \in BV(I)$ be given.

1. we denote by
$$(v)_I := \fint_I v(x)\,dx,$$
i.e., the average of v over I;
2. we denote by J_v the jump set of v and for $x_0 \in J_v$,
$$v(x_0^-) := \lim_{x \nearrow x_0} v(x) \text{ and } v(x_0^+) := \lim_{x \searrow x_0} v(x);$$

Lemma 4.7 ([60], Lemma 3.1 and Lemma 4.1)**.** *Let w be a piecewise constant function with M pieces where $M > 1$ large is a positive integer, then there exists a positive integer $M' \leq M$ and*

$$0 = \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{M'} < +\infty \tag{4.4}$$

such that

1. *$\mathscr{L}(\alpha_i, w)$ has at least one more constant piece than $\mathscr{L}(\alpha_{i+1}, w)$ for $i = 0, 1, \ldots, M' - 1$;*
2. *$\mathscr{L}(\alpha_i + \alpha, w)$ has the same number of constant pieces of $\mathscr{L}(\alpha_i, w)$, for any $0 \leq \alpha < \alpha_{i+1} - \alpha_i$ where $0 \leq i \leq M' - 1$;*
3. *$\mathscr{L}(\alpha, w) = (w)_I$ for all $\alpha \geq \alpha_{M'}$.*

Moreover, the function t: $[0, +\infty) \to [0, +\infty)$ defined as

$$t(\alpha) := \|\mathscr{L}(\alpha, w)\|^2_{L^2(I)}$$

is continuous, and in each interval $[\alpha_j, \alpha_{j+1})$, t' is linearly increasing and t is convex.

Proposition 4.8 ([60], Proposition 3.2)**.** *For any given corrupted image $u_{\eta,N}$ and clean image $u_{c,N}$, there exists an integer $N' < N$ and*

$$0 = \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{N'} = \alpha_s(u_{\eta,N}) < +\infty \tag{4.5}$$

such that item 1, 2, and 3 of Lemma 4.15 holds. Moreover, in each interval (α_i, α_{i+1}) $\mathcal{E}_N(\cdot)$ is convex and $\mathcal{E}'_N(\cdot)$ is linearly increasing, where $\mathcal{E}_N$ is defined in (1.27).

Theorem 4.9 ([69], Theorem 10.10)**.** *Let $v \in L^\infty(I)$ and $\alpha_1, \alpha_2 \in \mathbb{R}^+$ be given. Then the semigroup property*

$$\mathscr{L}(\alpha_1 + \alpha_2, v) = \mathscr{L}(\alpha_2, \mathscr{L}(\alpha_1, v)) = \mathscr{L}(\alpha_1, \mathscr{L}(\alpha_2, v)) \tag{4.6}$$

holds for the one dimensional scalar total variation problem.

Theorem 4.10 ([20], Theorem 3.4)**.** *Let $v \in BV(I)$ be given. Then*

$$J_{\mathscr{L}(\alpha,v)} \subset J_v$$

for any $\alpha > 0$, where J_v denotes the jump set of v. Moreover, the same result holds if we replace Q by I.

Remark 4.11. It follows from Theorem 4.9 and Theorem 4.10 that

$$J_{\mathscr{L}(\alpha_2,v,I)} \subset J_{\mathscr{L}(\alpha_1,v,I)}$$

for any $\alpha_1 \leq \alpha_2$. Indeed, by Theorem 4.9

$$\mathscr{L}(\alpha_2, v, I) = \mathscr{L}(\alpha_2 - \alpha_1 + \alpha_1, v, I) = \mathscr{L}(\alpha_2 - \alpha_1, \mathscr{L}(\alpha_1, v, I), I),$$

and hence by Theorem 4.10 with $\alpha := \alpha_2 - \alpha_1$, we obtain the result.

Proposition 4.12 ([11], Theorem 3)**.** *Let $v \in L^2(Q)$ be given. Then $\mathscr{L}(\cdot, v, Q) \in C([0, +\infty); L^2(Q))$. The same result holds for one dimension case, i.e., $\mathscr{L}(\cdot, v, I) \in C([0, +\infty); L^2(I))$.*

4.3. **The quasi-convexity of bilevel training scheme.** Theorem 4.5 allows us analytically predict the effects of TV regularization applied to any piecewise constant function in $\mathbb{R}$. Together with Theorem 4.9, we can completely obtain the explicit solution of $\mathscr{L}(\alpha, v, I)$ for all $\alpha \in \mathbb{R}^+$. By using such result, we may prove the following theorem:

Theorem 4.13. *Let $u_c \in BV(I)$ be monotone and let $N \in \mathbb{N}$ be given. Then the error function*

$$\mathcal{E}_N(\alpha) := \frac{1}{2} \left\| \mathscr{L}(\alpha, u_{\eta,N}) - u_{c,N} \right\|^2_{L^2(I)}, \ \alpha \in \mathbb{R}^+ \tag{4.7}$$

is strictly quasi-convex under Assumption 4.14.

The Assumption 4.14 is stated as follows:

Assumption 4.14. *Let $u_c \in BV(I)$ be monotone and $N \in \mathbb{N}$ be givne. Here $1 \leq k \leq N$.*

1. The observed noise changes sign consecutively, that is,

$$\eta_N(I_N(k))\eta_N(I_N(k+1)) \leq 0;$$

2. $u_{\eta,N}$ is oscillating at least at the half rate of $u_{c,N}$. That is, we require that

$$|u_{\eta,N}(I_N(k)) - u_{\eta,N}(I_N(k+1))| \geq \frac{1}{2} |u_{c,N}(I_N(k)) - u_{c,N}(I_N(k+1))| ; \tag{4.8}$$

3. if $u_{\eta,N}$ changes the sign of of jump of $u_{c,N}$, that is, if

$$(u_{\eta,N}(I_N(k)) - u_{\eta,N}(I_N(k+1))) (u_{c,N}(I_N(k)) - u_{c,N}(I_N(k+1))) \leq 0,$$

we require that

$$\eta_N(I_N(k)) = -\eta_N(I_N(k+1)). \tag{4.9}$$

4. we assume that

$$u_{\eta,N}(I_N(1)) > (<) u_{c,N}(I_N(1))$$

if $u_{c,N}(I_N(1)) > (<) u_{c,N}(I_N(2))$*, and*

$$u_{\eta,N}(I_N(N)) > (<) u_{c,N}(I_N(N))$$

if $u_{c,N}(I_N(N)) > (<) u_{c,N}(I_N(N-1))$*. Lastly, we assume that*

$$|\eta(I_N(1))| = |\eta(I_N(N))| \geq \frac{1}{2} \max \{ |\eta(I_N(k_I))| \,,\ k_I \in \mathcal{C}_I(u_{c,N}) \} . \tag{4.10}$$

As we can see, Assumption 4.14 is very restrictive and unlikely to be satisfied in the concrete setting, and requiring u_c to be monotone renders Theorem 4.13 to be less interesting. Also, since the proof of Theorem 4.13 is purely technical, we leave it in Appendix B and move on to construct counterexamples directly.

4.3.1. *Counterexamples.* We first show that removing (4.10) results in losing quasi-convexity, although the perturbation is relatively small. An explicit example is provided in Figure 2 below, but here let us draw some theoretical analysis first. Let $u_{c,N}$ be monotone increasing, and we assume that

$$|\eta_N(I_N(N))| > |\eta_N(I_N(1))| \geq \max \{ |\eta_N(I_N(k))| \,,\ 1 < k < N \} ,$$

i.e., (4.10) is no longer satisfied. Moreover, to simplify our computation, we assume that $\eta_N(I_N(2)) = 0$.

By using the same argument of the proof of Theorem 4.13, at $\alpha_d := |\eta_N(I_N(1))| / N$, we have $\mathcal{E}'_{1,N}(\alpha_d) = 0$, $I_N(k) \in \mathcal{I}(\alpha_d)$, $1 < k < N$, but

$$\mathcal{E}'_{N,N}(\alpha_d) = N \left(|\eta_N(I_N(1))| - |\eta_N(I_N(N))| \right) < 0.$$

On the other hand, for any $\alpha > 0$ small such that

$$0 < \alpha < \alpha_1 := \frac{1}{N} \left(\mathscr{L}(\alpha_d, u_{\eta,N})(I_N(1)) - u_{c,N}(I_N(1)) \right) ,$$

we have

$$\mathcal{E}'_{1,N}(\alpha_d + \alpha) = N\alpha > 0.$$

Moreover, according to (B.31) and (B.32), and $\eta_N(I_N(2)) = 0$, we have that

$$\mathcal{E}'_{1,N}((\alpha_d + \alpha_1)^-) > \mathcal{E}'^{+}_{1,N}((\alpha_d + \alpha_1)^+) + \mathcal{E}'^{+}_{2,N}((\alpha_d + \alpha_1)^+). \tag{4.11}$$

Then by choosing $\eta_N(I_N(N))$ properly and using (4.11), the following inequalities could hold:

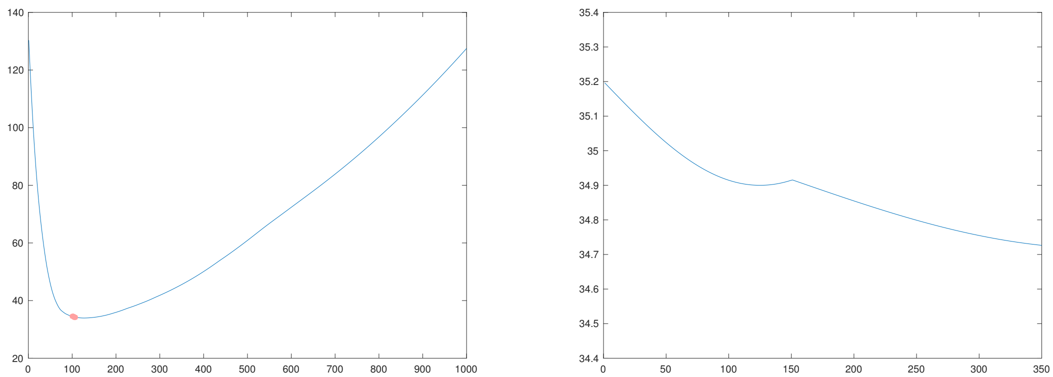
$$\begin{aligned} \mathcal{E}'_N((\alpha_d + \alpha_1)^-) &= \mathcal{E}'_{N,N}(\alpha_d + \alpha_1) + \mathcal{E}'_{1,N}((\alpha_d + \alpha_1)^-) \\ &= N \left[\eta_N(I_N(N)) - (\alpha_d + \alpha_1) \right] + \mathcal{E}'_{1,N}((\alpha_d + \alpha_1)^-) > 0, \end{aligned} \tag{4.12}$$

but

$$\mathcal{E}'_N((\alpha_d + \alpha_1)^+) = \mathcal{E}'_{N,N}(\alpha_d + \alpha_1) + \mathcal{E}'^{+}_{1,N}((\alpha_d + \alpha_1)^+) + \mathcal{E}'^{+}_{2,N}((\alpha_d + \alpha_1)^+) \tag{4.13}$$

$$=N\left[\eta_N(I_N(N)) - (\alpha_d + \alpha_1)\right] + \mathcal{E}'^{+}_{1,N}((\alpha_d+\alpha_1)^+) + \mathcal{E}'^{+}_{2,N}((\alpha_d+\alpha_1)^+) < 0,$$

and hence we lose quasi-convexity. We refer to Figure 2 for an explicit construction. In Figure 2a we see $\mathcal{E}_N(\alpha)$ is almost quasi-convex, but Figure 2b shows a small perturbation around the red point in Figure 2a, as we zoom in sufficient enough.



(A) $\mathcal{E}_N(\alpha)$ looks almost quasi-convex (B) a small perturbation around red point.

FIGURE 2. $N = 100$. $u_{c,N}(I_N(i)) = (i-1)/N$, $1 \le i \le N$. $\eta_N(I_N(1)) = -1/100$, $\eta_N(I_N(100)) = 1.8/100$, $\eta_N(I_N(3)) = 0.5/100$, $\eta_N(I_N(4)) = -0.1/100$, $\eta_N(I_N(5)) = 0.5/100$, $\eta_N(I_N(6)) = -0.1/100$, and $\eta_N(I_N(k)) = 0$ for all other intervals.

Moreover, removing either (4.9) or (4.8) will both result in losing (B.34) and hence some perturbation would happen shortly after α_m. By using the similar idea of (4.12) and (4.13), we may build the counterexamples as shown in Figure 3a to Figure 3b.

4.4. **A direct search for a minimizer of error function.**

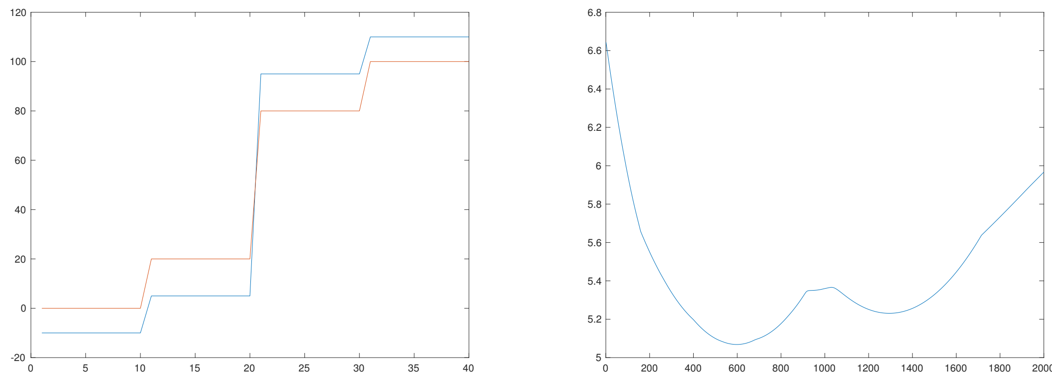
4.4.1. *The one dimensional case.* In Section 4.4.1 we will abbreviate $\mathscr{L}(\alpha, v, I)$ as $\mathscr{L}(\alpha, v)$ and $TV(v, I)$ as $TV(v)$.

Lemma 4.15. *Let w be a piecewise constant function with M pieces where $M > 1$ large is a positive integer, then there exists a positive integer $M' \le M$ and*

$$0 = \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{M'} < +\infty \tag{4.14}$$

such that

1. *$\mathscr{L}(\alpha_i, w)$ has at least one more constant piece than $\mathscr{L}(\alpha_{i+1}, w)$ for $i = 0, 1, \dots, M'-1$;*
2. *$\mathscr{L}(\alpha_i + \alpha, w)$ has the same number of constant pieces of $\mathscr{L}(\alpha_i, w)$, for any $0 \le \alpha < \alpha_{i+1} - \alpha_i$ where $0 \le i \le M'-1$;*
3. *$\mathscr{L}(\alpha, w) =: C(v)$ for all $\alpha \ge \alpha_{M'}$, where $C(v)$ is a constant depends on v.*



(A) violating of Assumption B.8

(B) $\mathcal{E}_N(\alpha)$ is not quasi-convex

FIGURE 3. $N = 4$. $u_{c,4}(I_4(1)) = 0$, $u_{c,4}(I_4(2)) = 0.2$, $u_{c,4}(I_4(3)) = 0.8$, and $u_{c,4}(I_4(4)) = 1$. $\eta_4(I_4(1)) = -0.1$, $\eta_4(I_4(2)) = -0.18$, $\eta_4(I_4(3)) = 0.18$, and $\eta_4(I_4(4)) = 0.1$.

Moreover, the function t*:* $[0, +\infty) \to [0, +\infty)$ *defined as*

$$t(\alpha) := \|\mathscr{L}(\alpha, w)\|^2_{L^2(I)}$$

is continuous, and in each interval $[\alpha_j, \alpha_{j+1})$*,* t' *is linearly increasing and* t *is convex.*

Proof. According to Theorem 4.5, for each $1 < j < M$ and $\alpha > 0$ small enough, we have

$$(\mathscr{L}(\alpha, w))(I_M(j)) = w(I_M(j)) \mp \frac{2}{|I_M(j)|}\alpha, \text{ if } I_M(j) \text{ is a high (low) extremum region of } w,$$

$$(\mathscr{L}(\alpha, w))(I_M(j)) = w(I_M(j)) \mp \frac{1}{|I_M(j)|}\alpha, \text{ if } I_M(j) \text{ is a high (low) boundary region of } w.$$

Therefore, we have

$$\|\mathscr{L}(\alpha, w)\|^2_{L^2(I_M(j))} = \left\| w(I_M(j)) \mp \frac{2}{|I_M(j)|}\alpha \right\|^2_{L^2(I_M(j))}$$

provided that $I_M(j)$ is a high (low) extremum region of ω. We obtain

$$\frac{1}{2}\frac{d}{d\alpha}\|\mathscr{L}(\alpha, w)\|^2_{L^2(I_M(j))} = 2\left(\frac{2}{|I_M(j)|}\alpha \mp w(I_M(j))\right),$$

which is continuous and linearly increasing in α, and

$$\frac{1}{2}\frac{d^2}{d\alpha^2}\left(\|\mathscr{L}(\alpha, w)\|^2_{L^2(I_M(j))}\right) = \frac{4}{|I_M(j)|}$$

which is strictly positive. A similar result holds if $I_M(j)$ is a boundary region. Moreover, since

$$t(\alpha) := \|\mathscr{L}(\alpha, w)\|^2_{L^2(I)} = \sum_{j=1}^{M} \|\mathscr{L}(\alpha, w)\|^2_{L^2(I_M(j))},$$

which is a finite summation of $\|\mathscr{L}(\alpha, w)\|^2_{L^2(I_M(j))}$, we conclude that $t'(\alpha)$ is continuous increasing and $t''(\alpha) > 0$ for $\alpha > 0$ small.

We claim that there exists an unique $\alpha_1 > 0$ such that for all $\alpha \in (0, \alpha_1)$

$$\mathscr{L}(\alpha_1 - \alpha, w) \text{ has } M \text{ pieces, but } \mathscr{L}(\beta, w) \text{ have at most } M-1 \text{ pieces,} \tag{4.15}$$

for all $\beta \geq \alpha_1$.

We first show the uniqueness. Assume there exist distinct α_1 and $\alpha_1' > 0$ such that (4.15) holds for both α_1 and α_1'. Without lose of generality we assume that $\alpha_1 < \alpha_1'$. Let $\alpha_1'' > 0$ be such that $\alpha_1 < \alpha_1'' < \alpha_1'$. Then, on the one hand, by (4.15) and Remark 4.11 we have

$$\mathscr{L}(\alpha_1'', w) = \mathscr{L}(\alpha_1 + (\alpha_1'' - \alpha_1), w) = \mathscr{L}(\alpha_1'' - \alpha_1, \mathscr{L}(\alpha_1, w))$$

has at most $M-1$ pieces, on the anther hand we have, again by (4.15), that $\mathscr{L}(\alpha_1' - (\alpha_1' - \alpha_1''), w)$ has M pieces since $\alpha_1' - \alpha_1'' > 0$, and we have a contradiction.

We define the set

$$\mathcal{A} := \left\{\alpha' > 0,\ \mathscr{L}(\alpha', w) \text{ has at most } M-1 \text{ pieces}\right\}$$

and we claim that

$$\beta := \inf_{\alpha > 0} \{\alpha \in \mathcal{A}\} \tag{4.16}$$

has the properties required by (4.15). First, we have that $\beta < +\infty$ since by Theorem 4.5 there exists $\alpha' > 0$ large enough such that $\mathscr{L}(\alpha, w)$ is a constant, i.e., it has only one constant piece, and hence $\mathcal{A} \neq \varnothing$. Next, let $\{\alpha_n\}_{n=1}^{\infty} \subset \mathcal{A}$ be such that $\alpha_n \searrow \beta$. We have $\mathscr{L}(\beta, w) = \lim_{n\to\infty} \mathscr{L}(\alpha_n, w)$ by Proposition, 4.12 and hence $\mathscr{L}(\beta, w)$ has at most $M-1$ pieces. Finally, we claim that $\mathscr{L}(\beta - \alpha, w)$ has M constant pieces for any $\alpha > 0$. If not, then there would be $\alpha'' > 0$ such that $\mathscr{L}(\beta - \alpha'', w)$ has at most $M-1$ constant pieces, but this contradicts (4.16).

We have shown that the function t has the required properties for $0 \leq \alpha < \alpha_1$ where α_1 is obtain via (4.16) ($\alpha_1 := \beta$), and α_1 satisfies items 1 and 2 in Lemma 4.15. Next, by (4.6) we may write, for $\alpha \geq \alpha_1$, that

$$\mathscr{L}(\alpha, w) = \mathscr{L}(\alpha_1 + \alpha - \alpha_1, w) = \mathscr{L}(\alpha - \alpha_1, w_1)$$

where $w_1 := \mathscr{L}(\alpha_1, w)$ is a piecewise constant function with M_1 pieces and $M_1 \leq M-1$. We can repeat the above argument to obtain α_2' such that $w_2 := \mathscr{L}(\alpha_2', \omega_1)$ has at most M_2 constant pieces where $M_2 \leq M_1 - 1$, and we define $\alpha_2 := \alpha_1 + \alpha_2'$. A recursive argument will lead to $w_{M'}$ a constant for M' sufficiently large. Since w only has M pieces, $M' \in \mathbb{N}$ is finite and $\alpha_{M'} < +\infty$ and so we obtain (4.14) as desired. Finally, since $w_{M'} := \mathscr{L}(\alpha_{M'}, w)$ has only one piece, $w_{M'}(x) =: C$ for all $x \in I$ and C is a constant. We conclude that for all $\alpha > \alpha_{M'}$

$$\mathscr{L}(\alpha, w) = \mathscr{L}(\alpha - \alpha_{M'}, w_{M'}) = w_{M'}.$$

□

Proposition 4.16. *For any given corrupted image $u_{\eta,N}$ and clean image $u_{c,N}$, there exists an integer $N' < N$ and*

$$0 = \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{N'} = \alpha_s(u_{\eta,N}) < +\infty \tag{4.17}$$

such that item 1, 2, and 3 of Lemma 4.15 holds. Moreover, in each interval (α_i, α_{i+1}) $\mathcal{E}_N(\cdot)$ is convex and $\mathcal{E}_N'(\cdot)$ is linearly increasing, where $\mathcal{E}_N$ is defined in (1.29).

Proof. Since $u_{c,N}$ is a fixed piecewise constant function, we may apply Lemma 4.15 to $u_{\eta,N}$ to obtain (4.17), and that $\mathcal{E}_N(\cdot)$ is convex and $\mathcal{E}_N'(\cdot)$ linearly increasing within each interval (α_i, α_{i+1}). Moreover, we conclude that $\alpha_{N'} = \alpha_s(u_{\eta,N})$ by applying items 1, 2, and 3 in Lemma 4.15 with $i = N'$. □

4.4.2. *The direct search for a minimizer α_m of level N error function.* Proposition 4.16 allows us to perform a direct search to find a minimizer α_m of (1.29). Indeed, recall that in each interval $[\alpha_i, \alpha_{i+1})$, $\mathcal{E}_N(\cdot)$ is convex and $\mathcal{E}_N'(\cdot)$ is linearly increasing. Hence, we may apply *Newton descent* (see, e.g., [8]) algorithm to locate the unique local minimizer $\alpha_{i,m}$ for $\mathcal{E}_N(\alpha)$ in $[\alpha_i, \alpha_{i+1}]$, and repeat over all intervals provided by (4.17). Since there are only finitely many intervals $[\alpha_i, \alpha_{i+1})$, we can locate all possible local minimizers $\alpha_{i,m}$ within a finite time. Finally, the finite stopping time $\alpha_s(u_{\eta,N})$ provides a natural stopping criterion for our searching algorithm. That is, we terminate our searching progress once we reach the point when $\mathcal{E}_N(\cdot)$ is a constant. After we terminate our searching progress, we only need to find the smallest local minimizer $\alpha_{i,m}$ and that is our α_m as desired. Lastly, if there is a tie, i.e., two local minimizer $\alpha_{i,m} < \alpha_{i',m}$ such that both gave the smallest value of $\mathcal{E}_N(\cdot)$, we choose $\alpha_{i,m}$ as our minimizer α_m and ignore $\alpha_{i'm}$.

4.4.3. *The two dimensional case.* In this section we present a two dimensional (weaker) version of Lemma 4.15 and Proposition 4.16 in Proposition 4.20. In particular, items 1 and 2 in Lemma 4.15 will be absent due to the lack of a two dimensional version of Theorem 4.5. We remark that so far we only have a weaker version of Theorem 4.5 in two dimensions and we refer readers to our follow up work [62].

We start by recalling the following theorem in [19].

Theorem 4.17 ([19], Theorem 4 and 5)**.** *Let $v \in L^\infty(Q)$ be given and let ∂TV denote the subgradient of the TV seminorm. Considering the* gradient flow *defined as*

$$\begin{cases} -\partial_t \mathcal{G}(t,v) \in \partial TV(\mathcal{G}(t,v)), \\ \mathcal{G}(0,v) := v. \end{cases} \tag{4.18}$$

Then following hold:

1. *the solution $\mathcal{G}(t,v)$ is uniquely defined;*
2. *the solution $\mathcal{G}(t,v)$ satisfies $\mathcal{G}(t,v) = \mathscr{L}(\alpha,v)$ for $t=\alpha$;*
3. *there exist finitely many $0 = t_0 < t_1 < t_2 < \cdots < t_K \le \infty$ such that the solution of (4.18) is given by*

$$\mathcal{G}(t,v) = \mathcal{G}(t_i,v) - (t-t_i)SG(t_{i+1})$$

for $t \in [t_i, t_{i+1})$, where $SG(t_{i+1}) \in \partial TV(\mathcal{G}(t_i,v))$.

We now prove the following two dimensional "semi-group" property.

Proposition 4.18. *Let $v \in L^\infty(Q)$ and let $0 < \alpha_1 < \alpha_2 < +\infty$ be given. Then*

$$\mathscr{L}(\alpha_2,v) = \mathscr{L}(\alpha_2-\alpha_1,\mathscr{L}(\alpha_1,v)). \tag{4.19}$$

Proof. Let $v_1 := \mathscr{L}(\alpha_1,v)$, and define a new gradient flow by

$$-\partial_t\mathcal{G}^1(t,v_1) \in \partial TV(\mathcal{G}^1(t,v_1)),\ \mathcal{G}^1(0,v_1) := v_1,$$

and we have $\mathcal{G}^1(t,v_1)$ is uniquely defined. By Theorem 4.17 we have that

$$\mathcal{G}^1(\alpha_2-\alpha_1,v_1) = \mathscr{L}(\alpha_2-\alpha_1,\mathscr{L}(\alpha_1,v)),$$

and

$$\mathcal{G}(\alpha_2,v) = \mathscr{L}(\alpha_2,v).$$

Moreover, by the property of gradient flow, we have

$$\mathcal{G}(\alpha_2,v) = \mathcal{G}^1(\alpha_2-\alpha_1,v_1),$$

and hence (4.19) hold. $\square$

We recall that the stopping time α_s was defined in Definition 4.4 and

$$\mathscr{L}(\alpha,v,Q) := \operatorname*{arg\,min}_{u\in SBV(Q)}\left\{\frac{1}{2}\int_Q |u-v|^2\,dx + \alpha TV(u,Q)\right\}. \tag{4.20}$$

Lemma 4.19. *Let $v \in L^\infty(Q)$ be given. Then $\alpha_s(v) < +\infty$ and $\mathscr{L}(\alpha_s(v),v)$ is a constant.*

Proof. We note that the null space of total variation seminorm

$$\mathcal{N}(TV) = \left\{v \in L^1(Q),\ TV(v) = 0\right\}, \tag{4.21}$$

is the space of constant function (see, e.g., [4]), and hence a linear subspace of $L^1(Q)$.

By Proposition 2.1 in [21], the optimality condition of (4.20), with v in place of u_η, is

$$\frac{1}{\alpha}\left(\mathscr{L}(\alpha,v) - v\right) \in \partial TV(\mathscr{L}(\alpha,v)).$$

Let P_{TV} denote the projection operator onto $\mathcal{N}(TV)$. Hence $P_{TV}(v)$ is a constant by (4.21). We claim that

$$\frac{1}{\alpha}\left(v - P_{TV}(v)\right) \in \partial TV(0) \tag{4.22}$$

for $\alpha > 0$ large enough. Indeed, since $\partial TV(0)$ has nonempty relative interior in $\mathcal{N}(TV)$ (see, e.g., [66]), we have that (4.22) holds for $\alpha > 0$ sufficient large since $v \in L^\infty(Q)$ and $P_{TV}(v)$ is a constant. Let $\alpha_S > 0$ be large enough such that (4.22) hold. Then we have

$$\frac{1}{\alpha_S}\left(v - P_{TV}(v)\right) \in \partial TV(0) = \partial TV(P_{TV}(v))$$

where in the last inequality we used again the fact that $P_{TV}(v)$ is a constant. That is, we have

$$\frac{1}{\alpha_S}\left(v - P_{TV}(v)\right) \in \partial TV(P_{TV}(v)),$$

and hence $P_{TV}(v)$ is a solution of (4.20). Since the minimizer of (4.20) is unique, we conclude that

$$P_{TV}(v) = \mathscr{L}(\alpha_S, v) \tag{4.23}$$

and thus $\mathscr{L}(\alpha_S, v)$ is a constant.

Define

$$\alpha_s := \inf\left\{\alpha > 0,\ \mathscr{L}(\alpha, v) = P_{TV}(v)\right\}.$$

Let $\{\alpha_n\}_{n=1}^\infty \subset \{\alpha > 0,\ \mathscr{L}(\alpha, v) = P_{TV}(v)\}$ and $\alpha_n \searrow \alpha_s$. We claim that α_s is indeed the stopping time of v. First, α_s is unique by its definition, and α_s is finite since there exists at least one $\alpha_S < +\infty$ such that (4.23) hold. Next, by Proposition 4.12 we have

$$\mathscr{L}(\alpha_s, v) = \lim_{n\to\infty} \mathscr{L}(\alpha_n, v) = P_{TV}(v).$$

Therefore, for all $\alpha > 0$, we have

$$\mathscr{L}(\alpha_s + \alpha, v) = \mathscr{L}(\alpha, \mathscr{L}(\alpha_s, v)) = \mathscr{L}(\alpha, P_{TV}(v)) = P_{TV}(v),$$

where in the first equality we used Proposition 4.18. This concludes the proof. □

Proposition 4.20. *For any given corrupted image $u_{\eta,N}$ and clean image $u_{c,N}$, there exists an integer $N' \in \mathbb{N}$ and*

$$0 = \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{N'} = \alpha_s(u_{\eta,N}) < +\infty$$

such that, in each interval (α_i, α_{i+1}), $\mathcal{E}_N(\cdot)$ is convex and $\mathcal{E}'_N(\cdot)$ is linearly increasing, where

$$\mathcal{E}_N(\alpha) := \frac{1}{2}\int_Q \left|\mathscr{L}(\alpha, u_{\eta,N}) - u_{c,N}\right|^2 dx. \tag{4.24}$$

Proof. Applying Theorem 4.17 to $u_{\eta,N}$, we obtain finitely many

$$0 := \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{N'} \le +\infty$$

such that

$$\mathscr{L}(\alpha, u_{\eta,N}) = \mathscr{L}(\alpha_i, u_{\eta,N}) - (\alpha - \alpha_i) SG(\alpha_{i+1}) \tag{4.25}$$

for $\alpha \in (\alpha_i, \alpha_{i+1})$, where $SG(\alpha_{i+1}) \in \partial TV(\mathscr{L}(\alpha_i, v))$. By Lemma 4.19 we have $\mathscr{L}(\alpha_s(u_{\eta,N}), u_{\eta,N})$ is a constant and hence $SG(\alpha_s(u_{\eta,N})) = 0$. Therefore, invoking (4.25) we deduce that $\alpha_{N'} \le \alpha_s(u_{\eta,N}) < +\infty$ and $\mathscr{L}(\alpha, u_{\eta,N}) = \mathscr{L}(\alpha_s(u_{\eta,N}), u_{\eta,N})$ for all $\alpha \ge \alpha_s(u_{\eta,N})$. Moreover, by (4.25) and the fact that $u_{c,N}$ is a fixed function, we conclude that in each interval (α_i, α_{i+1}), $\mathcal{E}_N(\cdot)$ is convex and $\mathcal{E}'_N(\cdot)$ is linearly increasing, as desired. □

4.5. **Another spatially dependent bilevel training scheme with respect to TV.** One significant drawback of TV denoising is the staircasing effect, and many attempts have been made to avoid such effect by, for example, introducing a higher level of derivative [28, 18], or by introducing a spatially dependent denoising parameter $\alpha(x)$ (see, e.g., [50]). In this section we present a new training scheme which is adapted from the bilevel training scheme (1.3).

Before we introduce our new training scheme, we prove a useful lemma.

Lemma 4.21. *Let $v \in L^\infty(Q)$ be given. Then*

$$\mathscr{L}(\alpha, v) =: u_\alpha \to (v)_Q := \fint_Q v\,dx \ \text{a.e..}$$

Proof. Recalling the definition of $\mathscr{L}(\alpha, v)$ from (4.20) and using $(v)_Q$ as test function, we have

$$\int_Q |u_\alpha - v|^2\,dx + \alpha TV(u_\alpha) \le \int_Q |(v)_Q - v_0|^2\,dx < +\infty.$$

Hence, $\{u_\alpha\}_{\alpha>0}$ is bounded in L^2, and (up to a not relabeled subsequence) there exists a $u_\infty \in L^2$ such that $u_\alpha \rightharpoonup u_\infty$ in L^2 as $\alpha \to \infty$. In turn, $TV(u_\alpha)$ is bounded, i.e., $\{u_\alpha\}_{\alpha>0}$ is bounded in BV. Hence $u_\infty \in BV$ and

$$TV(u_\infty) \le \liminf_{\alpha\to\infty} TV(u_\alpha) \le \lim_{\alpha\to\infty} \frac{1}{\alpha}\int_Q |(v)_Q - v_0|^2\,dx = 0,$$

which implies that $u_\infty =: c$ is a constant. Invoking the compactness embedding in BV space, we have $u_\alpha \to c$ in L^1, and we have (up to a not relabeled subsequence) $u_\alpha \to c$ a.e.. Moreover, by Fatou's Lemma,

$$\int_Q |v - c|^2\,dx \le \int_Q \left|v - (v)_Q\right|^2 dx. \tag{4.26}$$

Note that

$$\frac{d}{d\lambda}\int_Q |v-\lambda|^2\,dx = 2\int_Q (v-\lambda)\,dx,$$

and hence the left hand side of (4.26) reaches the minimum value at $\lambda = (v)_Q$. We conclude that $c = (v)_Q$, and the proof is completed. □

Remark 4.22. Combining the results from Lemma 4.21 and Lemma 4.19, we deduce that for $\alpha > \alpha_s(v)$, $\mathscr{L}(\alpha, v) = (v)_Q$, which is in agreement with Theorem 4.5.

4.5.1. *A spatially dependent construction.* Let $N \in \mathbb{N}$, $u_{c,N}$, and η_N be given. For $K \in \mathbb{N}$, $Q_K \subset \mathbb{R}^2$ denotes a cube with its faces normal to the orthonormal basis of $\mathbb{R}^2$, and with side-length greater than or equal to $1/K$. $\mathcal{L}_K$ will be a collection of finitely many Q_K such that

$$\mathcal{L}_K := \left\{Q_K \subset Q : \ Q_K \text{ are mutually disjoint, } Q \subset \bigcup \overline{Q_K}\right\}, \tag{4.27}$$

and $\mathcal{V}_K$ denotes the collection of all possible $\mathcal{L}_K$. For $K = 0$ we set $Q_0 := Q$, hence $\mathcal{L}_0 = \{Q\}$. We define our improved training scheme $(\mathcal{P})$ in resolution level N as:

Level 1.

$$u_{\mathcal{P},N} := \arg\min\left\{\int_\Omega |u_{c,N} - u_{\mathcal{L}_K}|^2\,dx,\ K \geq 0,\ \mathcal{L}_K \in \mathcal{V}_K\right\} \tag{4.28}$$

Level 2.

$$u_{\mathcal{L}_K}(x) := \mathscr{L}(\alpha_{Q_K}, u_{\eta,N}, Q_K) \text{ for } x \in Q_K \text{ and } Q_K \in \mathcal{L}_K, \tag{4.29}$$

$$\text{where } \ \alpha_{Q_K} := \mathop{\arg\min}_{\alpha>0} \int_{Q_K} |\mathscr{L}(\alpha, u_{\eta,N}, Q_K) - u_{c,N}|^2\,dx.$$

The training scheme $(\mathcal{P})$ performs the training scheme $(\mathcal{B})$ in each subdomain and combines it all together to achieve an improved global result. Let

$$\mathcal{P}_N(K) := \inf_{\mathcal{L}_K\in\mathcal{V}_K}\left\{\int_\Omega |u_{c,N} - u_{\mathcal{L}_K}|^2\,dx\right\}$$

where $u_{\mathcal{L}_K}$ is defined in (4.29), and

$$\mathcal{P}(N) := \int_Q |u_{c,N} - u_{\mathcal{P},N}|^2\,dx$$

where $u_{\mathcal{P},N}$ is obtained from (4.28). Since $\mathcal{V}_K \subset \mathcal{V}_{K+1}$, we have $\mathcal{P}_N(K) \geq \mathcal{P}_N(K+1)$ and hence

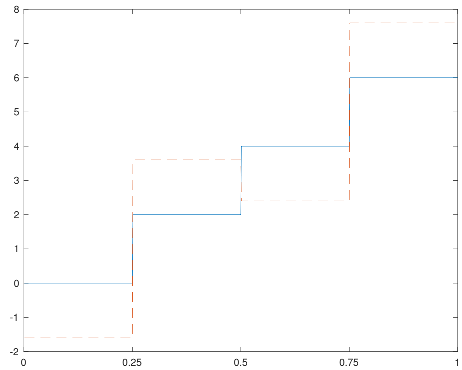
$$\lim_{K\to\infty} \mathcal{P}_N(K) \text{ exists}$$

and is equal to $\inf_{K\in\mathbb{N}_0} \mathcal{P}_N(K)$. Note that when $K = 0$, $\mathcal{P}_N(0) = \mathcal{E}_N(\alpha_m)$ where $\mathcal{E}_N(\cdot)$ is defined in (4.24) and α_m is the minimizer. That is, the improved scheme $(\mathcal{P})$ does make an improvement since $\mathcal{P}(N) \leq \mathcal{P}_N(0) = \mathcal{E}_N(\alpha_m)$.

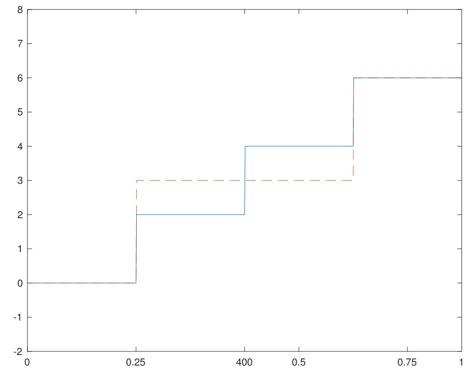
The assumption that $u_{\eta,N}$ is a piecewise constant function attaining finitely many values yields a natural stop criterion of scheme $(\mathcal{P})$ and prevents us from letting $K \to \infty$. Indeed, since $u_{\eta,N}$ is constant in each $Q_N \in \mathcal{Q}_N$ where $\mathcal{Q}_N$ is defined in (1.12), searching in cubes Q_K such that $K > N$ would not benefit us anymore since $\mathscr{L}(\alpha, v, Q_K) = v$ for any $\alpha \geq 0$ if v is constant in Q_K.

4.5.2. *The staircasing effect.* In this section we first illustrate with a simple example how $(\mathcal{P})$ avoids the staircasing effect. Figure 4a shows the given corrupted image $u_{\eta,N}$ and the clean image $u_{c,N}$, with $N = 4$. Scheme $(\mathcal{B})$ results in $\mathscr{L}(\alpha_m, u_{\eta,N}, I)(I_4(2)) = \mathscr{L}(\alpha_m, u_{\eta,N}, I)(I_4(3))$ and hence the staircasing effect occurs, as Figure 4b indicates. Scheme $(\mathcal{P})$ operates in the subintervals $I' := (0, 0.5)$ and $I'' := (0.5, 1)$ separately, and hence $\mathscr{L}(\alpha, u_{\eta,N}, I')(I_4(2))$ and $\mathscr{L}(\alpha, u_{\eta,N}, I'')(I_4(3))$ are able to break up the staircase produced in Figure 4b and go across each other, as shown in Figure 4c, and finally achieve a better result, as Figure 4d indicates. Moreover, as shown in the end of this chapter for the two dimensional case, where Figure 7 and 8 represents the clean image $u_{c,N}$ and corrupted image $u_{\eta,N}$, respectively. We see in Figure 10, the reconstructed image by scheme $(\mathcal{P})$ results in smaller error value, mitigated staircasing effect (upper right corner), and sharped edge (around the middle area), compare with the reconstructed image by scheme

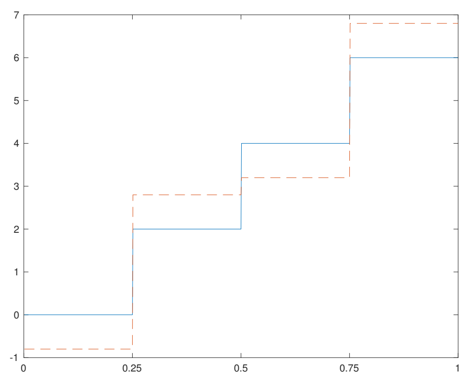
$(\mathcal{B})$ in Figure 9.



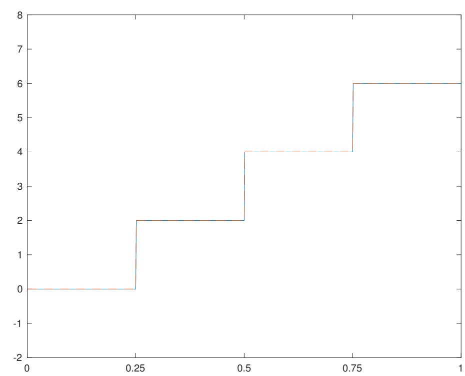
(A) $u_{\eta,N}$ and $u_{c,N}$



(B) $\mathscr{L}(\alpha_m, u_{\eta,N}, I)$ produces staircase



(C) Scheme (4.28) avoids staircasing



(D) $u_{\mathcal{P},N}$ and $u_{c,N}$ overlap, a perfect recovery

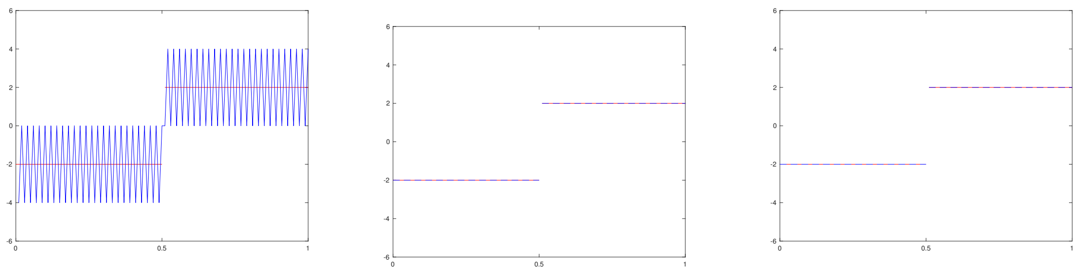
FIGURE 4. $I_4(1) = (0, 0.25)$, $I_4(2) = (0.25, 0.5)$, $I_4(4) = (0.5, 0.75)$, $I_4(4) = (0.75, 1)$

We remark that the ability to create a new jump point in $\mathscr{L}(\alpha, u_{\eta,N})$, as shown in Figure 4c, is key to avoid the staircasing effect. In [50], the authors proposed a method to avoid the staircasing effect by letting $\alpha = 0$ in certain points and hence at those points new jump points could be created in $\mathscr{L}(\alpha, u_{\eta,N})$. In Section 5.2 in [50] they showed that if η_N has average 0 in each subinterval I_i, where $I = \bigcup_{i=1}^{M} I_i$, and if $u_{c,N}$ is constant in each I_i, then they can achieve a perfect recovery (See Figure 5a to 5c). We remark that our scheme $(\mathcal{P})$ can produce the same perfect recovery result by choosing K large enough such that

$\{I_1, \ldots, I_M\} \subset \mathcal{L}_K$. Indeed, invoking Lemma 4.21 we have that, for $\alpha > 0$ large enough,

$$\mathscr{L}(\alpha, u_{\eta,N}, I_i) = \fint_{I_i} u_{\eta,N}\, dx = \fint_{I_i} (u_{c,N} + \eta_N)\, dx = \fint_{I_i} u_{c,N}\, dx = u_{c,N}(I_i)$$

for any $1 \leq i \leq M$, where in the last two equalities we used the assumptions that η_N has average 0 in I_i and that $u_{c,N}$ is constant in I_i.



(A) η_N has average 0 in I_1 and I_2

(B) perfect recovery by [50]

(C) perfect recovery by scheme (4.28)

FIGURE 5. $M = 2$. $I_1 = (0, 0.5)$, $I_2 = (0.5, 1)$

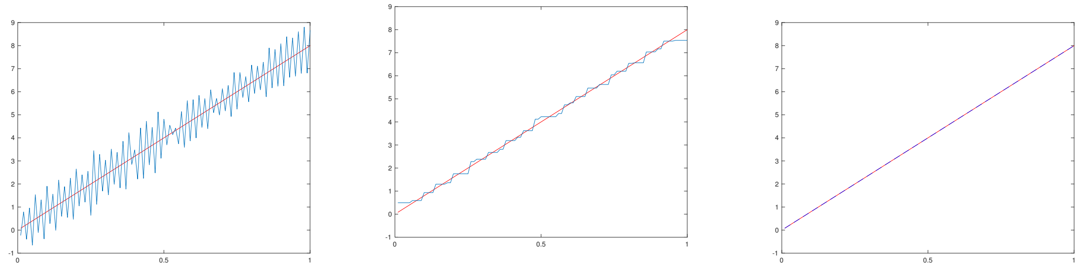
Finally, we remark that scheme $(\mathcal{P})$ can deal with more generalized situations which cannot be dealt by the method proposed in [50]. For example, in Figure 6a, $u_c(x) := x$ and hence $u_{c,N}(I_N(i)) = i/N$ for $x \in I_N(i)$, $1 \leq i \leq N$ (recall $I_N(i)$ from (1.24)). We define $\eta_N(2i-1) = -\eta_N(2i)$, $1 \leq i \leq N/2$. That is, η_N does not have average 0 in each subinterval $I_N(i)$ and so Proposition 5.5 in [50] can not be applied. However, scheme $(\mathcal{P})$ can still provide a perfect recovery result, as shown in Figure 4d, by choosing K large enough such that $\{I_{2i-1} \cup I_{2i},\ 1 \leq i \leq N/2\} \subset \mathcal{L}_K$. Moreover, we observe that scheme $(\mathcal{B})$ produces, again, the staircasing effect, as shown in Fig 6b.

4.5.3. *Approximation of the clean image.* In the last section of this chapter, we show that, under mild assumptions on the noise η_N, the scheme (4.28) can produce a perfect recovery result for an arbitrary clean image u_c, as the resolution level N goes to ∞.

We recall a useful corollary for Lusin's Theorem.

Corollary 4.23 ([40], Corollary 1, page 16. Also see [44], 7.10)**.** *Let μ be a Borel regular measure on $\mathbb{R}^N$ and let $f\colon \mathbb{R}^N \to \mathbb{R}^M$ be μ-measurable and bounded. Assume $A \subset \mathbb{R}^N$ is μ-measurable and $\mu(A) < +\infty$. Fix $\varepsilon > 0$. Then there exists a continuous function $\bar{f}\colon \mathbb{R}^N \to \mathbb{R}^M$ such that $\|\bar{f}\|_{L^\infty} \leq \|f\|_{L^\infty}$ and $\mu\{x \in A : \bar{f}(x) \neq f(x)\} < \varepsilon$.*

The main theorem of this section is as follows.



(A) $u_{\eta,N}$ and $u_{c,N}$ (B) $\mathscr{L}(\alpha_m, u_{\eta,N})$ by scheme (1.3) (C) $u_{\mathcal{P},N}$ by scheme (4.28)

FIGURE 6. $N = 100$. The noise η_N is designed such that $\eta_N(i) = -\eta_N(i+1)$. Note that in Figure 6b, scheme (1.3) produces staircasing; in Figure 6c, scheme (4.28) produces an almost perfect recovery

FIGURE 7. Clean image $u_{c,N}$

Theorem 4.24. *Assume that the noise η_{K^2} has locally average 0, that is*

$$\fint_{Q_K} \eta_{K^2} = 0 \tag{4.30}$$

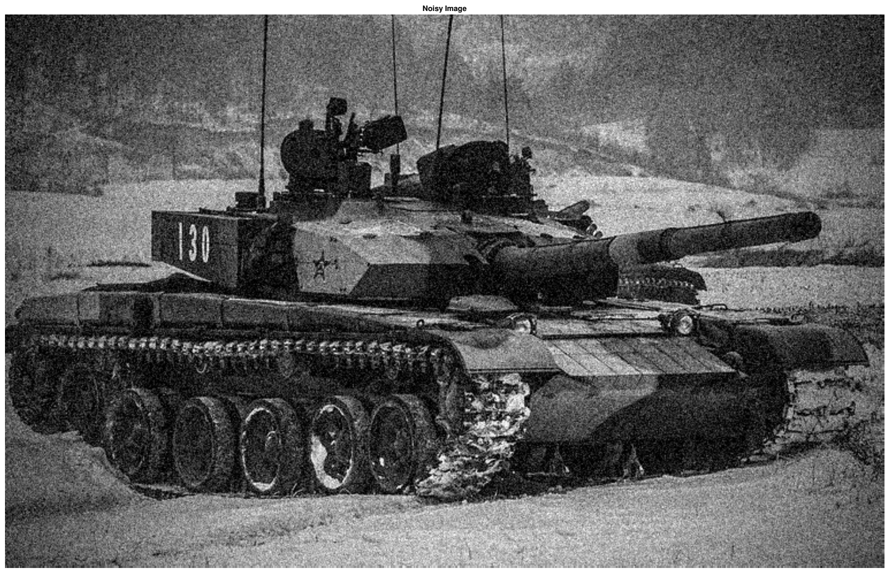


FIGURE 8. Corrupted image $u_{\eta,N}$, where the artificial noise is added by using a Gaussian noise distribution

for any $Q_K \in \mathcal{Q}_K$ *and all* $k \in \mathbb{N}$. *Then*

$$\lim_{K\to\infty} \mathcal{P}(K^2) = 0.$$

Proof. Let $K \in \mathbb{N}$ be fixed. Note that $\mathcal{Q}_K \in \mathcal{V}_K$. Then, according to (4.29) and invoking Lemma 4.21, for each $Q_K \in \mathcal{Q}_K$ we have

$$\begin{aligned}&\left\|\mathscr{L}(\alpha_{Q_K}, u_{\eta,K^2}, Q_K) - u_{c,K^2}\right\|^2_{L^2(Q_K)}\\ &\leq \left\|\fint_{Q_K} u_{\eta,K^2} dx - u_{c,K^2}\right\|^2_{L^2(Q_K)} = \left\|\fint_{Q_K} u_{c,K^2}\, dx - u_{c,K^2}\right\|^2_{L^2(Q_K)},\end{aligned}$$

where in the last equality we used (4.30).

Hence, we have

$$\begin{aligned}\mathcal{P}(K^2) &\leq \sum_{Q_K\in\mathcal{Q}_K} \left\|\mathscr{L}(\alpha_{Q_K}, u_{\eta,K^2}, Q_K) - u_{c,K^2}\right\|^2_{L^2(Q_K)}\\ &\leq \sum_{Q_K\in\mathcal{Q}_K} \left\|\fint_{Q_K} u_{c,K^2}(x)dx - u_{c,K^2}\right\|^2_{L^2(Q_K)}.\end{aligned}$$

Denoised Image sigle alpha

FIGURE 9. The reconstructed image by scheme $(\mathcal{B})$. The training error is 931.667. Note that the staircasing effect is observed, upper left and right corner.

We claim that

$$\lim_{K\to\infty}\sum_{Q_K\in\mathcal{Q}_K}\left\|u_{c,K^2}-u_c\right\|^2_{L^2(Q_K)}=\lim_{K\to\infty}\left\|u_{c,K^2}-u_c\right\|^2_{L^2(Q)}=0. \tag{4.31}$$

It is clear that (4.31) holds if u_c is continuous and using Lebesgue Dominated Convergence Theorem. We prove that (4.31) still hold if $u_c\in L^\infty(Q)$. For simplicity, assume that $\|u_c\|_{L^\infty(Q)}\leq 1$. Fix $\varepsilon>0$. By Corollary 4.23 there exists a compact set $W\subset\subset Q$ and a continuous function v such that $v\lfloor W=u_c\lfloor W$, $\|v\|_{L^\infty}\leq\|u_c\|_{L^\infty}$, and $\mathcal{L}^2(W)\geq\mathcal{L}^2(Q)-\varepsilon=1-\varepsilon$, where $\mathcal{L}^2$ stands for the two dimensional Lebesgure measure. Then we immediately have

$$\int_Q|v-u_c|^2\,dx<\varepsilon. \tag{4.32}$$

Let v_{K^2} be defined similarly to u_{c,K^2} and we observe that $v_{K^2}\to v$ in $L^2(Q)$. That is,

$$\lim_{K\to\infty}\|v_{K^2}-v\|^2_{L^2(Q)}=0. \tag{4.33}$$

FIGURE 10. The reconstructed image by scheme $(\mathcal{P})$. The training error is 900.325. Note that the staircasing effect is reduced, and edges are sharper

We obtain

$$\begin{aligned}\int_\Omega \left|v_{K^2} - u_{c,K^2}\right| dx &= \sum_{1\le i,j\le K^2} \int_{Q_{K^2}(i,j)} \left| K^2 \int_{Q_{K^2}(i,j)} (v-u_c)\, dx \right| dy \\ &\le \sum_{1\le i,j\le K^2} K^2 \int_{Q_{K^2}(i,j)} \int_{Q_{K^2}(i,j)} |v-u_c| \; dxdy \\ &= \sum_{1\le i,j\le K^2} \int_{Q_{K^2}(i,j)} |v-u_c| \; dx = \|v-u_c\|_{L^1(Q)} \le 2\varepsilon.\end{aligned}$$

Since $\left|v_{K^2} - u_{c,K^2}\right| \le 2$ uniformly in W we deduce that

$$\int_\Omega \left|v_{K^2} - u_{c,K^2}\right|^2 dx \le 2\int_\Omega \left|v_{K^2} - u_{c,K^2}\right| dx \le 4\varepsilon. \tag{4.34}$$

Hence, for $K \in \mathbb{N}$ large enough, and in view of (4.34), (4.33), and (4.32), in this order, we observe that

$$\begin{aligned}\left\|u_{c,K^2} - u_c\right\|^2_{L^2(Q)} &= \left\|u_{c,K^2} - v_{K^2} + v_{K^2} - v + v - u_c\right\|^2_{L^2(Q)} \\ &\le 3\left\|u_{c,K^2} - v_{K^2}\right\|^2_{L^2(Q_K)} + 3\left\|v_{K^2} - v\right\|^2_{L^2(Q_K)} + 3\left\|v-u_c\right\|^2_{L^2(Q_K)}\end{aligned}$$

$$\leq 12\varepsilon + 3\varepsilon + 3\varepsilon = 18\varepsilon$$

and (4.31) is verified.

Similarly, we could show that (note below we have $u_{c,K}$, but in (4.31) we have u_{c,K^2})

$$\lim_{K\to\infty} \sum_{Q_K} \|u_{c,K} - u_c\|^2_{L^2(Q_K)} = 0. \tag{4.35}$$

Note that

$$\fint_{Q_K} u_{c,K^2}(y)dy = u_{c,K}(x) \text{ for } x \in Q_K.$$

Then, in view of (4.31) and (4.35),

$$\begin{aligned}
&\sum_{Q_K\in\mathcal{Q}_K} \left\| \fint_{Q_K} u_{c,K^2}(x)dx - u_{c,K^2} \right\|^2_{L^2(Q_K)} \\
&= \sum_{Q_K\in\mathcal{Q}_K} \left\| u_{c,K} - u_c + u_c - u_{c,K^2} \right\|^2_{L^2(Q_K)} \\
&\leq \sum_{Q_K\in\mathcal{Q}_K} \|u_{c,K} - u_c\|^2_{L^2(Q_K)} + \sum_{Q_K\in\mathcal{Q}_K} \left\|u_{c,K^2} - u_c\right\|^2_{L^2(Q_K)} \to 0
\end{aligned}$$

as $K \to \infty$.

Therefore, we deduce that

$$\lim_{K\to\infty} \mathcal{P}(K^2) \leq \lim_{K\to\infty} \sum_{Q_K\in\mathcal{Q}_K} \left\| \fint_{Q_K} u_{c,K^2}(x)dx - u_{c,K^2} \right\|^2_{L^2(Q_K)} = 0,$$

and the proof is concluded. □

Remark 4.25. The noise η_{K^2} in Theorem 4.24, which has locally zero average, can be produced by using the *compound camera* which is the leading technology in robotic vision. Roughly speaking, the compound camera captures a corrupted image u_{η,K^2} with resolution K^2 by capturing with K^2 number of small cameras, each has resolution level K and captures a part of u_c in the subdomain Q_K, and these put together yield u_{η,K^2}. It is usually assumed that each individual camera produces noise with zero average (see, e.g., [21]), which implies that the nose η_{K^2} has average zero in each Q_K as required.

Chapter 5. The comprehensive training scheme

5.1. **Notations and basic assumptions.** Let $Q := (0,1) \times (0,1)$ be the unit square. This will be the domain of our image data. Generally we may take $Q \subset \mathbb{R}^2$ to be an open bounded domain with Lipschitz boundary, although such generalization would not be useful in image processing problems. The corrupted image u_η and the associated clean image u_c are assumed to lie in a Banach space Y, which is usually taken to be L^2 or L^1 in image denoising and deblurring problem.

We next describe the basic assumptions on the *assessment operators* $\mathcal{A}$, *fidelity operator* $\mathcal{F}$, and *regularizer* $\mathcal{R}$, and we will provide some examples to illustrate these assumptions.

Definition 5.1. *We say that an operator $\mathcal{R}$: $Y \to \mathbb{R}^+$ is a regularizer if it satisfies the following conditions:*

1. *$\mathcal{R}$ is convex;*
2. *the set*

$$X_{\mathcal{R}} := \{u \in Y : \ \mathcal{R}(u) < +\infty\}$$

equipped with the norm

$$\|u\|_{X_{\mathcal{R}}} := \|u\|_Y + \mathcal{R}(u).$$

is a normed subspace of Y.

Definition 5.2. *We say an operator $\mathcal{F}$: $Y \to [0, +\infty]$ is a fidelity operator if it is proper and strictly convex.*

Definition 5.3. *We say that an operator $\mathcal{A}$: $Y \to [0, +\infty]$ is an assessment operator if it is both continuous and weakly lower semicontinuous in Y.*

Assumption 5.4 (imaging-ready operator)**.** *Let $\{u_n\}_{n=1}^{\infty} \subset Y$ be such that*

$$\sup\{\mathcal{F}(u_n - u_\eta) + \mathcal{R}(u_n) : \ n \geq 1\} < +\infty. \tag{5.1}$$

We say that the operators $\mathcal{F}$ and $\mathcal{R}$ are satisfy Assumption imaging-ready *if, up to the extraction of a (non-relabeled) subsequence, there exists $\tilde{u} \in X_{\mathcal{R}}$ such that*

$$u_n \rightharpoonup \tilde{u} \text{ in } Y \text{ and } \liminf_{n\to\infty} \mathcal{R}(u_n) \geq \mathcal{R}(\tilde{u}).$$

Remark 5.5. In the context of existing literature, the fidelity operator $\mathcal{F}$ is usually assumed to be coercive in the sense that

$$\mathcal{F}(v) \to +\infty \text{ as } \|v\|_Y \to \infty. \tag{5.2}$$

We remark that (5.2), together with Definition 5.1 and Definition 5.2, implies Assumption 5.4 provided that the Banach space Y is assumed to be reflexive.

To sketch the proof, note that for any sequence $\{u_n\}_{n=1}^{\infty} \subset Y$ such that (5.1) holds, by (5.2) we have

$$\sup\left\{\|u_n - u_\eta\|_Y\right\} < +\infty.$$

Since Y is reflexive, by *Banach - Alaoglu* theorem there exists $\tilde{u} \in Y$ such that, up to the extraction of a subsequence, $u_n \rightharpoonup \tilde{u}$ in Y. Finally, since $\mathcal{R}$ is convex, we have $\mathcal{R}$ is *s.w.l.s.c.* and hence

$$+\infty > \liminf_{n\to\infty} \mathcal{R}(u_n, Q) \geq \mathcal{R}(\tilde{u}, Q),$$

and $\tilde{u} \in X_{\mathcal{R}}$.

Lemma 5.6. *Let $\mathcal{R}$ and $\mathcal{F}$ satisfy Definition 5.1 and Definition 5.2, respectively, and 5.4. Then the minimizing problem*

$$u_{\mathcal{R}} \in \arg\min \{\mathcal{F}(u_\eta - u, Q) + \mathcal{R}(u, Q) : \ u \in X_{\mathcal{R}}\}$$

has a unique solution.

Proof. Let

$$m := \inf \{\mathcal{F}(u_\eta - u, Q) + \mathcal{R}(u, Q) : \ u \in X_{\mathcal{R}}\} . \tag{5.3}$$

Since $\mathcal{F}$ and $\mathcal{R}$ are non-negative, we have $m \geq 0$. Let $\{u_n\}_{n=1}^{\infty} \subset X_{\mathcal{R}}$ be a minimizing sequence. Then, for n large enough, we have

$$\mathcal{F}(u_n - u_\eta, Q) + \mathcal{R}(u_n, Q) \leq m + 1.$$

In view of Assumption 5.4, up to a subsequence, we have there exists $\tilde{u} \in X_{\mathcal{R}}$ such that

$$u_n \rightharpoonup \tilde{u} \text{ in } Y \text{ and } \liminf_{n\to\infty} \mathcal{R}(u_n) \geq \mathcal{R}(\tilde{u}).$$

Since $\mathcal{F}$ is strictly convex, and hence weakly *l.s.c.*, we have

$$m = \liminf_{n\to\infty} \mathcal{F}(u_n - u_\eta, Q) + \mathcal{R}(u_n, Q) \geq \mathcal{F}(\tilde{u} - u_\eta, Q) + \mathcal{R}(\tilde{u}, Q),$$

which, together with (5.3), implies that

$$\mathcal{F}(\tilde{u} - u_\eta, Q) + \mathcal{R}(\tilde{u}, Q) = m.$$

Finally, since $\mathcal{R}(\cdot)$ is convex and $\mathcal{F}(\cdot)$ is strongly convex, we conclude that $\tilde{u}$ is unique and we set $u_{\mathcal{R}} := \tilde{u}$. □

We present a few examples to illustrate the abstract framework above.

Example 5.7 (Squared L^2 assessment and fidelity operator & (non)-smooth regularizer)**.** Let $Y = L^2$ with

$$\mathcal{A}(\cdot, Q) = \mathcal{F}(\cdot, Q) = \|\cdot\|^2_{L^2(Q)} ,$$

we recover the standard L^2-squared fidelity and assessment operators (the role of $\mathcal{A}$ will be explained in Section 1.5).

An example of a smooth regularizer is given by $\mathcal{R}(\cdot, Q) = |\cdot|_{W^{1,2}(Q)}$. In this case, $X_{\mathcal{R}} = W^{1,2}(Q)$ and we have

$$\mathcal{F}(u - u_\eta) + \mathcal{R}(u) = \|u - u_\eta\|^2_{L^2} + |u|_{W^{1,2}} . \tag{5.4}$$

It is clear that $\mathcal{F}$ and $\mathcal{R}$ in (5.4) satisfy Assumption 5.4.

An example of a non-smooth regularizer is given by $\mathcal{R}(\cdot, Q) := |\cdot|_{TV(Q)}$, the total variation, and we define $X_{\mathcal{R}} = BV(Q) \cap L^2(Q)$. Since the domain Q has Lipschitz boundary, and the dimension is either 1 (signal) or 2 (image), we have that the operators $\mathcal{F}$ and $\mathcal{R}$ in this example satisfy Assumption 5.4.

5.2. **The regularizer training scheme.** For the convenience of the reader, we re-state the definition of *indexing set* introduced in Chapter 1.

Notation 5.8. We use $\mathbb{M}^{n\times k}$ to denote the vector space of $n \times k$ real valued matrices, where n, $k \in \mathbb{N}$

Definition 5.9 (The indexing set of $\mathscr{R}$)**.** *Let $\Pi := \Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_{N_{\mathscr{R}}}$, where the* indexing dimension *is $N_{\mathscr{R}} \in \mathbb{N}$, and each Γ_i is a compact subset of $\mathbb{M}^{n_i\times k_i}$. We say that a space (set) of regularizers $\mathscr{R}$ is indexed by Π if each $\mathcal{R} \in \mathscr{R}$ can be uniquely represented (see Example 5.11 below) by an element $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_{N_{\mathscr{R}}}) \in \Pi$, and we use $\mathcal{R}[\gamma]$ to indicate that $\mathcal{R}$ is indexed by γ. Moreover, we endow $\mathscr{R}$ with the norm defined by*

$$d_{\mathscr{R}}(\mathcal{R}[\gamma], \mathcal{R}[\gamma']) := \sum_{i=1}^{N_{\mathscr{R}}} \left\|\gamma_i - \gamma_i'\right\|_{\Gamma_i}.$$

We introduce the following **R**egularizer **T**raining scheme ($\mathcal{RT}$).
Level 1.

$$\tilde{\gamma} := \arg\min \left\{\mathcal{A}\left(u_c - u_{\mathcal{R}[\gamma]}, Q\right) :\ \gamma \in \Pi\right\} \tag{5.5}$$

Level 2.

$$u_{\mathcal{R}[\gamma]} := \arg\min \left\{\mathcal{F}(u_\eta - u, Q) + \mathcal{R}[\gamma](u, Q),\ u \in X_{\mathcal{R}[\gamma]}\right\}.$$

Definition 5.10 (The reconstruction map)**.** *Given $u_\eta \in Y$, we define the* reconstruction map *$\mathcal{S}$: $\mathscr{R} \to X_{\mathcal{R}}$ by*

$$\mathcal{S}_{u_\eta}(\mathcal{R}) := \arg\min \left\{\mathcal{F}(u_\eta - u, Q) + \mathcal{R}(u, Q) :\ u \in X_{\mathcal{R}}\right\}.$$

Note that by Lemma 5.6, the operator $\mathcal{S}_{u_\eta}$ is well defined. Moreover, for simplicity of notation, we abbreviate $\mathcal{S}_{u_\eta}(\mathcal{R})$ as $\mathcal{S}(\mathcal{R})$ in the rest of this chapter.

Example 5.11. We present two examples to illustrate Definition 5.9.

1. Let $\Pi := \Gamma_1 \times \Gamma_2$ where $\Gamma_1 := [1, 2]$ and $\Gamma_2 := [2, 4]$. Then we may define a space $\mathscr{R}$ based on the indexing set Π via

$$\mathscr{R} := \left\{W^{s,p} :\ s \in \Gamma_1,\ p \in \Gamma_2\right\},$$

   and we have $\mathcal{R}[(s, p)] = W^{s,p}$ and

$$d_{\mathscr{R}}(\mathcal{R}[(s, p)], \mathcal{R}[(s', r')]) := \left|s - s'\right| + \left|p - p'\right|.$$

2. Let $\Pi := \Gamma_1 = [1, S]$ where $S \in (1, +\infty)$. The fractional order total generalized variation TGV^s ([31]) is defined as follows:

$$\begin{aligned} TGV^s(u) = \min \Big\{ &|\nabla u - \mathrm{mod}(s, 1)v_0|_{\mathcal{M}_b} + \mathrm{mod}(s, 1)\left|\mathcal{E}^{\mathrm{mod}(s,1)} v_0 - \mathrm{mod}(s, 2)v_1\right|_{\mathcal{M}_b} + \cdots \\ &+ \mathrm{mod}(s, l)\left|\mathcal{E}^{\mathrm{mod}(s,l)} v_{k-2} - \mathrm{mod}(s, l + 1)v_l\right|_{\mathcal{M}_b} + \\ &\cdots + \mathrm{mod}(s, \lfloor S\rfloor - 1)\left|\mathcal{E}^{\mathrm{mod}(s,\lfloor S\rfloor-1)} v_{k-2} - \mathrm{mod}(s, \lfloor S\rfloor)v_{\lfloor S\rfloor-1}\right|_{\mathcal{M}_b} + \end{aligned}$$

$$+\operatorname{mod}(s,\lfloor S\rfloor)\left|\mathcal{E}^{\operatorname{mod}(s,\lfloor S\rfloor)}v_{\lfloor S\rfloor}\right|_{\mathcal{M}_b} : v_l \in BV(Q,\operatorname{Sym}^l(\mathbb{R}^2)),\ l=0,\dots,\lfloor S\rfloor-1\Big\},$$

where $\operatorname{mod}(s,N) := 0 \vee (s-N) \wedge 1$. We introduce a regularizer space $\mathscr{R}$ via

$$\mathscr{R} := \{TGV^s :\ s \in \Gamma_1\},$$

and we have $\mathcal{R}[s] = TGV^s$ and

$$d_{\mathscr{R}}(\mathcal{R}[s],\mathcal{R}[s']) := |s-s'|.$$

Assumption 5.12 ($\mathcal{A}$-*l.s.c* with respect to $d_{\mathscr{R}}$)**.** *We say that the operator* $\mathcal{S}(\mathcal{R})$ *is* $\mathcal{A}$-*l.s.c. with respect to* $d_{\mathscr{R}}$ *if for every* $\{\mathcal{R}[\gamma_n]\}_{n=1}^{\infty} \subset \mathscr{R}$ *with* $\lim_{n\to\infty} d_{\mathscr{R}}(\mathcal{R}[\gamma_n],\mathcal{R}[\gamma]) = 0$,

$$\liminf_{n\to\infty} \mathcal{A}(\mathcal{S}(\mathcal{R}[\gamma_n]) - u_c) \geq \mathcal{A}(\mathcal{S}(\mathcal{R}[\gamma]) - u_c).$$

Theorem 5.13. *Let* $\mathcal{A}$ *be an assessment operator satisfying Assumption 5.3 and let* Y *be a Hilbert space. If* $\mathcal{S}$ *is* $\mathcal{A}$-*l.s.c. with respect to* $d_{\mathscr{R}}$*, then problem* (5.5) *admits a solution* $\mathcal{R}[\tilde{r}] \in \mathscr{R}$*.*

Proof. Since the assessment operator is non-negative, we may extract a minimizing sequence $\{\mathcal{R}[\gamma_n]\}_{n=1}^{\infty} \subset \mathscr{R}$ such that

$$\lim_{n\to\infty} \mathcal{A}\left(u_{c,N} - \mathcal{S}(\mathcal{R}[\gamma_n]),Q\right) = \inf\left\{\mathcal{A}\left(u_{c,N} - \mathcal{S}(\mathcal{R}[\gamma]),Q\right) :\ \mathcal{R}[\gamma] \in \mathscr{R}\right\} =: m \geq 0. \quad (5.6)$$

We claim that there exists a regularizer $\mathcal{R}[\tilde{r}] \in \mathscr{R}$ such that

$$\mathcal{A}(u_c - \mathcal{S}(\mathcal{R}[\tilde{r}])) = m.$$

Recall from Definition 5.9 the indexing dimension $N_{\mathscr{R}} \in \mathbb{N}$. Then we may write γ_n in (5.6) by

$$\gamma_n = (\gamma_{1,n},\gamma_{2,n},\dots,\gamma_{N_{\mathscr{R}},n}),\ \{\gamma_{i,n}\}_{n=1}^{\infty} \subset \Gamma_i \text{ for } i=1,\dots,N_{\mathscr{R}}.$$

Since each Γ_i is closed and compact, by a diagonal argument we find $\tilde{\gamma} = (\tilde{\gamma}_1,\tilde{\gamma}_2,\dots,\tilde{\gamma}_{N_{\mathscr{R}}}) \in \Pi$ such that, up to a subsequence,

$$\tilde{\gamma}_i = \lim_{n\to\infty} \gamma_{i,n}$$

for each $i = 1,\dots,N_{\mathscr{R}}$. Now, in view of Assumption 5.12 we have

$$m = \liminf_{n\to\infty} \mathcal{A}(\mathcal{S}(\mathcal{R}[\gamma_n]) - u_c) \geq \mathcal{A}(\mathcal{S}(\mathcal{R}[\tilde{\gamma}]) - u_c).$$

Since $\tilde{\gamma} \in \Pi$ and hence $\mathcal{R}[\tilde{\gamma}] \in \mathscr{R}$, we have

$$\mathcal{A}(\mathcal{S}(\mathcal{R}[\tilde{\gamma}]) - u_c) = m$$

as desired. ◻

5.2.1. *Training scheme in regularizer and parameter spaces.* In order to add parameters into the regularizer, we specific a more detailed structure of the set $\mathscr{R}$. Recall from Definition 5.9 that we require a set $\mathscr{R}$ to be indexed by a set Π, a product of compact subset of $\mathbb{M}$, so that we can define the *distance* between two regularizers. Here, in addition, we define the *dimension* of $\mathscr{R}$ so that we may specify the structure of $\mathscr{R}$. First, we give the definition of *box constraint.*

Definition 5.14. *We say that a vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{R}^n$ satisfies the* box constraint *if there exist a positive number $0 < A << 1$ such that $\alpha_i \in [A, 1/A]$ for $i = 1, 2, \dots, n$.*

Definition 5.15. *We say that a space $\mathscr{R}$ has* operator dimension $n_{\mathscr{R}} \in \mathbb{N}$ *if there exists a set of operators*

$$\left\{R_i(\cdot,\cdot,\cdot) :\ Y \times Y^{n_{\mathscr{R}}} \times \Pi \to \mathbb{R}^+ \text{ for } i = 1, \dots, n_{\mathscr{R}}\right\} \text{ with } R_i(tu, tv, \cdot) = tR_i(u, v, \cdot),\ t \in \mathbb{R}^+,$$

such that each $\mathcal{R}[\gamma] \in \mathscr{R}$ can be represented by

$$\mathcal{R}[\gamma](u, Q) = \inf\left\{R_1(u, v, \gamma) + R_2(u, v, \gamma) + \ldots + R_{n_{\mathcal{R}}}(u, v, \gamma) :\ v \in Y^{n_{\mathscr{R}}}\right\}. \tag{5.7}$$

We define a scaled version of $\mathcal{R}[\gamma] \in \mathscr{R}$ by adding a parameter $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{n_{\mathscr{R}}}) \in \mathbb{R}^{n_{\mathscr{R}}}$, which satisfies the box constraint in Definition 5.14, in the following sense:

$$\begin{aligned} &\mathcal{R}[\alpha, \gamma](u) \\ &:= \inf\left\{R_1(\alpha_1 u, \alpha_1 v, \gamma) + R_2(\alpha_2 u, \alpha_2 v, \gamma) + \ldots + R_{n_{\mathscr{R}}}(\alpha_{n_{\mathscr{R}}} u, \alpha_{n_{\mathscr{R}}} v, \gamma) :\ v \in Y^{n_{\mathscr{R}}}\right\}, \end{aligned} \tag{5.8}$$

and we set

$$d_{A,\mathscr{R}}(\mathcal{R}[\alpha, \gamma], \mathcal{R}[\alpha', \gamma']) := d_{\mathscr{R}}(\mathcal{R}[\gamma], \mathcal{R}[\gamma']) + \left|\alpha - \alpha'\right|.$$

Lemma 5.16. *If α satisfies the box constraint, then $X_{\mathcal{R}[\alpha,\gamma]} = X_{\mathcal{R}[\gamma]}$.*

Proof. We only show that

$$X_{\mathcal{R}[\alpha,\gamma]} \subset X_{\mathcal{R}[\gamma]}$$

The proof in the other direction is analogous.

Let $u \in X_{\mathcal{R}[\alpha,\gamma]}$ be given. Then by Definition 5.15 we have $u \in Y$ and

$$\begin{aligned} &\mathcal{R}[\alpha, \gamma](u) \\ &= \inf\left\{R_1(\alpha_1 u, \alpha_1 v, \gamma) + R_2(\alpha_2 u, \alpha_2 v, \gamma) + \ldots + R_{n_{\mathscr{R}}}(\alpha_{n_{\mathscr{R}}} u, \alpha_{n_{\mathscr{R}}} v, \gamma) :\ v \in Y^{n_{\mathscr{R}}}\right\} < +\infty, \end{aligned}$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{n_{\mathscr{R}}}) \in \mathbb{R}^{n_{\mathscr{R}}}$. Let $v_0 \in Y^{n_{\mathscr{R}}}$ be such that

$$R_1(\alpha_1 u, \alpha_1 v_0, \gamma) + R_2(\alpha_2 u, \alpha_2 v_0, \gamma) + \ldots + R_{n_{\mathscr{R}}}(\alpha_{n_{\mathscr{R}}} u, \alpha_{n_{\mathscr{R}}} v_0, \gamma) \le \mathcal{R}[\alpha, \gamma] + 1. \tag{5.9}$$

Since α satisfies the box constraint, i.e., $A \le \alpha_i$, $i = 1, \dots, n_{\mathscr{R}}$, we have

$$\begin{aligned} &R_1(u, v_0, \gamma) + R_2(u, v_0, \gamma) + \ldots + R_{n_{\mathscr{R}}}(u, v_0, \gamma) \\ &= \frac{1}{A} R_1(Au, Av_0, \gamma) + \frac{1}{A} R_2(Au, Av_0, \gamma) + \ldots + \frac{1}{A} R_{n_{\mathscr{R}}}(Au, Av_0, \gamma) \end{aligned}$$

$$\leq \frac{1}{A}\left[R_1(\alpha_1 u, \alpha_1 v_0, \gamma) + R_2(\alpha_2 u, \alpha_2 v_0, \gamma) + \ldots + R_{n_{\mathscr{R}}}(\alpha_{n_{\mathscr{R}}} u, \alpha_{n_{\mathscr{R}}} v_0, \gamma)\right]$$
$$\leq \frac{1}{A}\mathcal{R}[\alpha, \gamma] + \frac{1}{A} < +\infty.$$

Thus, in view of (5.9), we have

$$\mathcal{R}[\gamma](u) \leq R_1(u, v_0, \gamma) + R_2(u, v_0, \gamma) + \ldots + R_{n_{\mathscr{R}}}(u, v_0, \gamma) < +\infty,$$

and hence $u \in X_{\mathcal{R}[\gamma]}$. □

Example 5.17. Recall the fractional order total variation TV^s from [78].

1. We define a space of regularizers $\mathscr{R}$ by using the indexing set $\Pi := \Gamma_1 = [1/2, 1]$ and setting
$$\mathscr{R} := \{TV^s : \ s \in [0,1]\},$$
where TV^s can be written in the form of (5.7), to be precise,
$$TV^s = \inf\left\{TV^s(u) : \ v \in L^2\right\}.$$
That is, the operator R_1 in (5.7) is defined to be independent of the auxiliary variation v. Then, we have
$$\mathcal{R}\left[\alpha, s\right](u) = \alpha TV^s, \text{ where } \alpha \in [A, 1/A].$$

2. Recalling Example 5.11, in terms of (5.8) we have
$$\begin{aligned} TGV^s[\alpha](u) = \min\Bigg\{ & \alpha_0 \left|\nabla u - \mathrm{mod}(s,1)v_0\right|_{\mathcal{M}_b} + \alpha_1 \,\mathrm{mod}(s,1)\left|\mathcal{E}^{\mathrm{mod}(s,1)}v_0 - \mathrm{mod}(s,2)v_1\right|_{\mathcal{M}_b} \\ & \cdots + \alpha_2 \,\mathrm{mod}(s,l)\left|\mathcal{E}^{\mathrm{mod}(s,l)}v_{k-2} - \mathrm{mod}(s,l+1)v_l\right|_{\mathcal{M}_b} + \\ & \cdots + \alpha_l \,\mathrm{mod}(s,\lfloor S\rfloor - 1)\left|\mathcal{E}^{\mathrm{mod}(s,\lfloor S\rfloor-1)}v_{k-2} - \mathrm{mod}(s,\lfloor S\rfloor)v_{\lfloor S\rfloor-1}\right|_{\mathcal{M}_b} + \\ & + \alpha_{\lfloor S\rfloor}\,\mathrm{mod}(s,\lfloor S\rfloor)\left|\mathcal{E}^{\mathrm{mod}(s,\lfloor S\rfloor)}v_{\lfloor S\rfloor}\right|_{\mathcal{M}_b} : v_l \in BV(Q, \mathrm{Sym}^l(\mathbb{R}^2)),\ l = 0,\ldots,\lfloor S\rfloor - 1\Bigg\}, \end{aligned}$$
where $\bar{\alpha} = (\alpha_0, \ldots, \alpha_k)$ satisfies the box constraint. In [31] we proved that the regularizer space $\mathscr{R}$ constructed using $TGV_{\bar{\alpha}}^{k+s}$ satisfies Assumption 5.18.

We improve scheme $(\mathcal{RT})$ by inserting parameters as in (5.8), so that $(\mathcal{RT})$ is now able to train the parameters and regularizers, simultaneously. The scheme $(\mathcal{R})$ can be viewed as the generalization of the scheme defined (2.24).

Level 1.

$$(\tilde{\alpha}, \tilde{\gamma}) \in \arg\min\left\{\mathcal{A}\left(u_c - u_{\mathcal{R}[\alpha,\gamma]}, Q\right) : \ \gamma \in \Pi_{\mathscr{R}},\ \alpha \in [A, 1/A]^{n_{\mathscr{R}}}\right\}, \tag{5.10}$$

Level 2.

$$u_{\mathcal{R}[\alpha,\gamma]} := \arg\min\left\{\mathcal{F}(u_\eta - u, Q) + \mathcal{R}[\alpha,\gamma](u, Q),\ u \in X_{\mathcal{R}[\gamma]}\right\}. \tag{5.11}$$

We improve Assumption 5.12 to accommodate the parameter spaces $[A, 1/A]^{n_{\mathscr{R}}}$.

Assumption 5.18 (A-$\mathcal{A}$-$l.s.c.$ with respect to $d_{A,\mathscr{R}}$)**.** *We say that the operator $\mathcal{S}(\mathcal{R})$ is A-$\mathcal{A}$-l.s.c. with respect to $d_{A,\mathscr{R}}$ if for every $\{(\alpha_n,\gamma_n)\}_{n=1}^{\infty} \subset [A,1/A]^{n_{\mathscr{R}}} \times \Pi$ with* $\lim_{n\to\infty} d_{A,\mathscr{R}}(\mathcal{R}[\alpha_n,\gamma_n],\mathcal{R}[\alpha,\gamma]) = 0$,

$$\liminf_{n\to\infty} \mathcal{A}(\mathcal{S}(\mathcal{R}[\alpha_n,\gamma_n]) - u_c) \geq \mathcal{A}(\mathcal{S}(\mathcal{R}[\alpha,\gamma]) - u_c).$$

Remark 5.19. In most cases Assumption 5.18 on $\mathcal{R}[\alpha,\gamma]$ is redundant once Assumption 5.12 holds on $\mathcal{R}[\gamma]$ since α satisfies the box constraint defined in Definition 5.14.

Theorem 5.20. *Suppose that Assumption 5.18 holds. Then problem* (5.10) *admits a solution $(\tilde{\alpha},\tilde{\gamma})$.*

Proof. The proof can be carried out by using an argument similar to that adopted in the proof of Theorem 5.13 and Assumption 5.18. □

Corollary 5.21. *Recall the notations from* (5.10) *and* (5.11)*. The set*

$$\{(\alpha,\gamma)\}_{\text{opt}} := \arg\min \left\{ \mathcal{A}\left(u_{c,N} - u_{\mathcal{R}[\alpha,\gamma]}, Q\right) :\ \gamma \in \Pi,\ \alpha \in [A,1/A]^{n_{\mathscr{R}}} \right\}$$

is closed.

Proof. If $\#\{(\alpha,\gamma)\}_{\text{opt}} < +\infty$, we have nothing to prove. If not, in view of the box constraint we have that for any sequence $\{(\alpha_n,\gamma_n)\}_{n=1}^{\infty} \subset \{(\alpha,\gamma)\}_{\text{opt}}$ there exist $\tilde{\alpha}$ and $\tilde{\gamma}$ such that, up to a subsequence,

$$\alpha_n \to \tilde{\alpha} \text{ and } \gamma_n \to \tilde{\gamma}.$$

We claim that $(\tilde{\alpha},\tilde{\gamma}) \in \{(\alpha,\gamma)\}_{\text{opt}}$. Since $\{(\alpha_n,\gamma_n)\}_{n=1}^{\infty} \subset \{(\alpha,\gamma)\}_{\text{opt}}$, we have

$$\mathcal{A}(\mathcal{S}(\mathcal{R}[\alpha_n,\gamma_n]) - u_c) = m := \inf \left\{ \mathcal{A}\left(u_{c,N} - u_{\mathcal{R}[\alpha,\gamma]}, Q\right) :\ \gamma \in \Pi,\ \alpha \in [A,1/A]^{n_{\mathscr{R}}} \right\},$$

and in view of Assumption 5.18, we have

$$m \geq \liminf_{n\to\infty} \mathcal{A}(\mathcal{S}(\mathcal{R}[\alpha_n,\gamma_n]) - u_c) \geq \mathcal{A}(\mathcal{S}(\mathcal{R}[\tilde{\alpha},\tilde{\gamma}]) - u_c). \tag{5.12}$$

Since $(\tilde{\alpha},\tilde{\mathcal{R}}) \in [A,1/A]^{n_{\mathscr{R}}} \times \mathscr{R}$, we have

$$\mathcal{A}(\mathcal{S}(\mathcal{R}[\tilde{\alpha},\tilde{\gamma}]) - u_c) \geq m,$$

and together with (5.12), this conclude the proof. □

5.3. **The comprehensive bilevel training scheme.**

5.3.1. *The construction.* We first review the following notations from (1.33) (recall that $0 < A < 1$ from Definition 5.14):

1. $\mathcal{H}_A$ is the collection of rectangles such that

$\mathcal{H}_A := \{L \subset Q :\ L$ is an open rectangle with the shortest side-length greater than or equal to $A\}$;

2. $\mathcal{L}$ stands for a collection of finitely many $L \in \mathcal{H}_A$ such that

$$\mathcal{L} := \left\{ L \in \mathcal{H}_A :\ L \text{ are mutually disjoint},\ Q \subset \bigcup \overline{L} \right\}, \tag{5.13}$$

3. $\mathcal{V}_A$ is the collection of all possible $\mathcal{L}$, and we define, for any $\mathcal{L}$, $\mathcal{L}' \in \mathcal{V}_A$,
$$d_{\mathcal{V}_A}(\mathcal{L},\mathcal{L}') := \max\left\{\min\left\{d_{\mathcal{H}_A}(L,L') : \ L' \in \mathcal{L}'\right\} : \ L \in \mathcal{L}\right\},$$
and
$$d_{\mathcal{H}_A}(L,L') := \|\chi_L - \chi_{L'}\|_{L^1(Q)} \tag{5.14}$$
where χ_L is the characteristic function over L.

Next, we introduce a comprehensive training scheme $(\mathcal{CT})$ as follows:
Level 1.
$$\tilde{\mathcal{L}} \in \arg\min\left\{\mathcal{A}\left(u_{c,N} - \mathcal{P}(\mathcal{L})\right) : \ \mathcal{L} \in \mathcal{V}_A\right\}, \tag{5.15}$$
Level 2.
$$\mathcal{P}(\mathcal{L}) \text{ is built upon the information of } \left\{(\tilde{\alpha}_L, \tilde{\gamma}_L)\right\}_L \text{ in each } L \in \mathcal{L},$$
Level 3.
$$\left\{(\tilde{\alpha}_L, \tilde{\gamma}_L)\right\}_L := \arg\min\left\{\mathcal{A}\left(u_{c,N} - u_{\mathcal{R}[\alpha,\gamma]}, L\right) : \ \gamma \in \Pi, \ \alpha \in [A, 1/A]^{n_{\mathscr{R}}}\right\}, \tag{5.16}$$
$$u_{\mathcal{R}[\alpha,\gamma]} := \arg\min\left\{\mathcal{F}(u_{\eta,N} - u, L) + \mathcal{R}[\alpha,\gamma](u,L), \ u \in X_{\mathcal{R}[\gamma]}\right\}.$$
Here the operator $\mathcal{P}$: $\mathcal{V}_A \to Y$ acts as an assemble operator, using the local optimal re-construction information obtained in Level 3 within each subdomain L to construct a global re-constructed image $u_{\mathcal{L}}$, based on the partition domain $\mathcal{L} \in \mathcal{V}_A$.

The delicate part of the training scheme $(\mathcal{CT})$ is the construction of an assemble operator $\mathcal{P}$. We will provide two constructions in Section 5.3.2. Here we first give a sufficient condition for an assemble operator so that the scheme $(\mathcal{CT})$ admits a solution.

Assumption 5.22. *We say that the operator $\mathcal{P}$: $\mathcal{V}_A \to Y$ is $\mathcal{A}$-l.s.c. with respect to $d_{\mathcal{V}_A}$ if for any sequence $\{\mathcal{L}_n\}_{n=1}^{\infty} \subset \mathcal{V}_A$ with $\lim_{n\to\infty} d_{\mathcal{V}_A}(\mathcal{L}_n,\mathcal{L}) = 0$,*
$$\liminf_{n\to\infty} \mathcal{A}(\mathcal{P}(\mathcal{L}_n) - u_c) \geq \mathcal{A}(\mathcal{P}(\mathcal{L}) - u_c).$$

Theorem 5.23. *If the assemble operator $\mathcal{P}$ is $\mathcal{A}$-l.s.c. with respect to $d_{\mathcal{V}_A}$, then problem* (5.24) *admits a solution $\tilde{\mathcal{L}} \in \mathcal{V}_A$.*

To prove Theorem 5.23, we first establish two compactness results in the space $\mathcal{H}_A$ and on training ground $\mathcal{V}_A$.

Lemma 5.24. *Let a sequence of $\{L_n\}_{n=1}^{\infty} \subset \mathcal{H}_A$ be given. Then, up to a subsequence, there exists $\tilde{L} \in \mathcal{H}_A$ such that*
$$\lim_{n\to\infty} d_{\mathcal{H}_A}(L_n, \tilde{L}) = 0.$$

Proof. Select L_n from $\mathcal{L}_n$ for each $n \in \mathbb{N}$ and define $\chi_n := I_{L_n}$, the characteristic function of L_n. Since $\{L_n\}_{n=1}^{\infty} \subset \mathcal{H}_A$, we have
$$\|\chi_n\|_{BV(Q)} \leq |L_n| + |\chi_n|_{TV} \leq |Q| + 4 \leq 5,$$
and hence
$$\sup\left\{\|\chi_n\|_{BV(Q)} : \ n \in \mathbb{N}\right\} < +\infty.$$

Therefore, by weak*-compactness in BV, up to a subsequence (not relabeled), there exists $\chi \in BV$ such that

$$\chi_n \overset{*}{\rightharpoonup} \chi \text{ in } BV. \tag{5.17}$$

We claim that χ is the characteristic function for a set $S \subset Q$ such that $S \in \mathcal{H}_A$. Indeed, since $\chi_n \to \chi$ in L^1 strong, we have χ is a characteristic function of certain set S, and we only need to prove that $S \in \mathcal{H}_A$.

Next, let the four vertices of L_n be $a_n = (a_{1,n}, a_{2,n})$, $b_n = (b_{1,n}, b_{2,n})$, $c_n = (c_{1,n}, c_{2,n})$, $d_n = (d_{1,n}, d_{2,n}) \in [0,1] \times [0,1]$. Upon a further extract subsequence, there exist $\tilde{a} = (\tilde{a}_1, \tilde{a}_2)$, $\tilde{b} = (\tilde{b}_1, \tilde{b}_2)$, $\tilde{c} = (\tilde{c}_1, \tilde{c}_2)$, and $\tilde{d} = (\tilde{d}_1, \tilde{d}_2) \in [0,1] \times [0,1]$ such that

$$a_n \to \tilde{a},\ b_n \to \tilde{b},\ c_n \to \tilde{c},\ \text{and } d_n \to \tilde{d}. \tag{5.18}$$

We claim that S is a rectangle with vertices $\tilde{a}$, $\tilde{b}$, $\tilde{c}$, and $\tilde{d}$. Indeed, let χ' be the characteristic function of rectangle with vertices $\tilde{a}$, $\tilde{b}$, $\tilde{c}$, and $\tilde{d}$, we show that $\chi_n \to \chi'$ in L^1. We observe that

$$\begin{aligned}
&\lim_{n\to\infty} \int_Q |\chi_n - \chi'|\, dx \\
&\le \limsup_{n\to\infty} \Big(|a_{1,n} - \tilde{a}_1|\, |a_{2,n} - b_{2,n}| + \left|b_{2,n} - \tilde{b}_2\right| |b_{1,n} - c_{1,n}| \\
&\quad + |a_{2,n} - \tilde{a}_2| \left|\tilde{a}_1 - \tilde{d}_1\right| + |c_{1,n} - \tilde{c}_1| \left|\tilde{a}_2 - \tilde{b}_2\right| \Big) \\
&\le \limsup_{n\to\infty} |a_{1,n} - \tilde{a}_1| + \limsup_{n\to\infty} \left|b_{2,n} - \tilde{b}_2\right| + \limsup_{n\to\infty} |a_{2,n} - \tilde{a}_2| + \limsup_{n\to\infty} |c_{1,n} - \tilde{c}_1| = 0.
\end{aligned}$$

Hence, we have $\chi_n \to \chi'$ in L^1 which forces $\chi' = \chi$ since by (5.17) we have $\chi_n \to \chi$ in L^1, too. Therefore, we have S is a rectangle with vertices $\tilde{a}$, $\tilde{b}$, $\tilde{c}$, and $\tilde{d}$, and by (5.18) we conclude that $S \in \mathcal{H}_A$ and we set $\tilde{L} := S$. □

Lemma 5.25. *Let a sequence of $\{\mathcal{L}_n\}_{n=1}^{\infty} \subset \mathcal{V}_A$ be given. Then, up to a subsequence, there exists $\tilde{\mathcal{L}} \in \mathcal{V}_A$ such that*

$$\lim_{n\to\infty} d_{\mathcal{V}_A}(\mathcal{L}_n, \tilde{\mathcal{L}}) = 0.$$

Proof. Select L_n from $\mathcal{L}_n$ for each $n \in \mathbb{N}$. By Lemma 5.24, up to a subsequence, there exists $S_1 \in \mathcal{H}_A$ such that

$$\lim_{n\to\infty} d_{\mathcal{H}_A}(L_n, S_1) = 0.$$

Define

$$\mathcal{L}_n^1 := \{L \in \mathcal{L}_n :\ L \cap S_1 = \varnothing\} \text{ for each } n \in \mathbb{N}. \tag{5.19}$$

Repeating the argument above with $\mathcal{L}_n^1$, we may obtain a rectangle $S_2 \in \mathcal{H}_A$, and, in view of (5.19), we have $S_2 \cap S_1 = \varnothing$. We next define

$$\mathcal{L}_n^2 := \left\{L \in \mathcal{L}_n^1 :\ L \cap S_2 = \varnothing\right\} \text{ for each } n \in \mathbb{N}.$$

Recursively, we obtain S_1, S_2, ... Since $|L| \geq A^2$ for arbitrary $L \in \mathcal{H}_A$, we have

$$M := \sup \{ \# \{ \mathcal{L}_n \} : \ n \in \mathbb{N} \} \leq \frac{1}{A^2} < +\infty.$$

Therefore, the above argument can only be repeated finitely many times, and we obtain a set

$$\mathcal{S} := \{ S_1, S_2, \ldots, S_T \} \tag{5.20}$$

where $T \leq M$ and each $S_i \in \mathcal{H}_A$. We finally claim that $\mathcal{S} \in \mathcal{V}_A$. To do so, we only need to prove that

$$\bigcup_{i=1}^{T} \bar{S}_i \supset Q.$$

Suppose not, i.e.,

$$\left| Q \setminus \bigcup_{i=1}^{T} S_i \right| > 0.$$

Since S_i are all rectangles, there exists a rectangle

$$L' \subset Q \setminus \bigcup_{i=1}^{T} S_i \tag{5.21}$$

(Note that L' might be small and $L' \notin \mathcal{H}_A$). In view of (5.13), for each $n \in \mathbb{N}$ there exists $L'_n \in \mathcal{L}_n$ such that

$$\left| L' \cap L'_n \right| \geq \frac{1}{M} \left| L' \right|.$$

Let L'' be the limit of L'_n, up to a subsequence, in the sense of (5.17). We have

$$\left| L' \cap L'' \right| \geq \frac{1}{M} \left| L' \right|. \tag{5.22}$$

Hence, using $\{L'_n\}_{n=1}^{\infty}$ in step one (5.19) above, we have $L'' \in \mathcal{S}$ where $\mathcal{S}$ is defined in (5.20). That is, in view of (5.22),

$$\left| L' \cap \left(Q \setminus \bigcup_{i=1}^{T} S_i \right) \right| \leq \left| L' \setminus L'' \right| < \left| L' \right|. \tag{5.23}$$

However, (5.21) implies that

$$\left| L' \cap \left(Q \setminus \bigcup_{i=1}^{T} S_i \right) \right| = \left| L' \right|,$$

which contradicts to (5.23). $\square$

We are now ready to proof Theorem 5.23

Proof of Theorem 5.23. Let $\{\mathcal{L}_n\}_{n=1}^{\infty} \subset \mathcal{V}_A$ be a minimizing sequence such that $\mathcal{A}(u_c - \mathcal{P}(\mathcal{L}_n)) \to m$, where

$$m := \inf \left\{ \mathcal{A}\left(u_{c,N} - \mathcal{P}(\mathcal{L})\right) : \ \mathcal{L} \in \mathcal{V}_A \right\}.$$

Applying Lemma 5.25 to $\{\mathcal{L}_n\}_{n=1}^{\infty}$, there exists $\tilde{\mathcal{L}} \in \mathcal{V}_A$ such that

$$\lim_{n\to\infty} d_{\mathcal{V}_A}(\mathcal{L}_n, \tilde{\mathcal{L}}) = 0.$$

In view of Assumption 5.22, we have

$$m \geq \liminf_{n\to\infty} \mathcal{A}(\mathcal{P}(\mathcal{L}_n) - u_c) \geq \mathcal{A}(\mathcal{P}(\tilde{\mathcal{L}}) - u_c).$$

Since $\tilde{\mathcal{L}} \in \mathcal{V}_A$, we obtain

$$\mathcal{A}(\mathcal{P}(\tilde{\mathcal{L}}) - u_c) \geq m,$$

and this concludes the proof. □

5.3.2. *Construction of assemble operators.* In view of Theorem 5.20 and Theorem 5.23 we observe that the existence of a solution to the training scheme $(\mathcal{RT})$ and $(\mathcal{CT})$ depends on the *l.s.c.* of the solution operators, such as $\mathcal{S}(\alpha, \gamma)$ in $(\mathcal{RT})$ and $\mathcal{P}(\mathcal{L})$ in $(\mathcal{CT})$.

It is usually not easy to prove that $\mathcal{S}(\alpha, \gamma)$ and $\mathcal{P}(\mathcal{L})$ satisfying the *l.s.c.* properties, as required in Assumption 5.18 and Assumption 5.22, respectively. In Chapter 2 we showed that to prove TGV^{1+s} satisfies Assumption 5.18, even in one dimension, a serious amount of knowledge in PDEs and calculation of variations are required. However, the construction of the assemble operator $\mathcal{P}(\mathcal{L})$, which is closer to a data-based approach, need to use knowledge other than that used in the analysis of regularizers, for example, deep learning techniques introduced in Machine Learning, but this is beyond the scope of this thesis. Here, we will only present two elementary level constructions.

5.3.3. *Assemble operator directly from a local optimization result.* This construction is the generalization of the training scheme (1.30).

Notation 5.26. Let a partition domain $\mathcal{L} \in \mathcal{V}_A$ be given.

1. Fix $L \in \mathcal{L}$. We denote by $E(L)$ the space of functions u, defined only in L, such that $\|u\|_{Y(L)} < +\infty$.
2. We use $\mathcal{E}(\mathcal{L})$ to denote the collection of functions

$$\mathcal{E}(\mathcal{L}) := \left\{ u : \ u\lfloor_L \in E(L) \text{ for each } L \in \mathcal{L} \right\}.$$

The training scheme $(\mathcal{CT})$ will read as follows:
Level 1.

$$\tilde{\mathcal{L}} \in \arg\min \left\{ \mathcal{A}\left(u_c - \mathcal{P}(\mathcal{L})\right) : \ \mathcal{L} \in \mathcal{V}_A \right\} \tag{5.24}$$

Level 2.

$$\mathcal{P}(\mathcal{L}) := \arg\min \left\{ \mathcal{A}(u - u_c, Q) : \ u \in \mathcal{E}(\mathcal{L}) \right\} \tag{5.25}$$

$$\text{where } \mathcal{E}(\mathcal{L}) \text{ is built by using } E(L) := \left\{ \mathcal{S}(\tilde{\alpha}_L, \tilde{\gamma}_L) : \ (\tilde{\alpha}_L, \tilde{\gamma}_L) \in \{(\tilde{\alpha}_L, \tilde{\gamma}_L)\}_L \right\}, \tag{5.26}$$

Level 3.

$$\{(\tilde{\alpha}_L, \tilde{\gamma}_L)\}_L := \arg\min \left\{ \mathcal{A}\left(u_c - u_{\mathcal{R}[\alpha,\gamma]}, L\right) : \ \gamma \in \Pi, \ \alpha \in [A, 1/A]^{n_{\mathscr{R}}} \right\} \tag{5.27}$$
$$u_{\mathcal{R}[\alpha,\gamma]} := \arg\min \left\{ \mathcal{F}(u_\eta - u, L) + \mathcal{R}[\alpha,\gamma](u, L), \ u \in X_{\mathcal{R}[\gamma]} \right\}.$$

Lemma 5.27. *The collection $\mathcal{E}(\mathcal{L})$ defined in* (5.26) *is closed.*

Proof. This is a direct consequence of Corollary 5.21. Moreover, we have $\mathcal{P}(\mathcal{L}) \in \mathcal{E}(\mathcal{L})$. □

To show that $\mathcal{P}(\mathcal{L})$ defined in (5.25) satisfies Assumption 5.22, the following assumption on $\mathscr{R}_\Pi$ is needed.

Assumption 5.28. *Let a regularizer space $\mathscr{R}_\Pi$, a fidelity operator $\mathcal{F}$, and v, $\tilde{v} \in Y$ be given.*

1. *Let $\{(\alpha_n, \gamma_n)\}_{n=1}^\infty \subset [A, 1/A]^{n_{\mathscr{R}}} \times \Pi$ and $\{v_n\}_{n=1}^\infty \subset Y$ be such that $(\alpha_n, \gamma_n) \to (\tilde{\alpha}, \tilde{\gamma})$ and $v_n \to \tilde{v}$ in Y, and define*
$$u_n := \arg\min \left\{ \mathcal{F}(u - v_n, Q) + \mathcal{R}[\alpha_n, \gamma_n](u, Q) : \ u \in X_{\mathcal{R}[\gamma_n]} \right\}. \tag{5.28}$$
Then, up to a subsequence, there exists $\tilde{u} \in Y$ such that $u_n \to \tilde{u}$ in Y and
$$\tilde{u} = \arg\min \left\{ \mathcal{F}(u - \tilde{v}, Q) + \mathcal{R}[\tilde{\alpha}, \tilde{\gamma}](u, Q) : \ u \in X_{\mathcal{R}[\tilde{\gamma}]} \right\}. \tag{5.29}$$
2. *Let a sequence $\{L_n\}_{n=1}^\infty \subset \mathcal{H}_A$ and $L \in \mathcal{H}_A$ be such that $L_n \to L$ in the sense of* (5.14)*, and define*
$$w_n := \arg\min \left\{ \mathcal{F}(u - \tilde{v}, L_n) + \mathcal{R}[\alpha_n, \gamma_n](u, L_n) : \ u \in X_{\mathcal{R}[\gamma_n]} \right\}.$$
Moreover, let $\{\varphi_n\}_{n=1}^\infty$ be a sequence continuously differentiable bijective maps from L into L_n. Then, we have
$$\lim_{n\to\infty} \mathcal{A}(w_n - v, L_n) = \lim_{n\to\infty} \mathcal{A}(\tilde{w}_n - v, L),$$
where
$$\tilde{w}_n := \arg\min \left\{ \mathcal{F}(u - \tilde{v} \circ \varphi_n, L) + \mathcal{R}[\tilde{\alpha}, \tilde{\gamma}](u, L) : \ u \in X_{\mathcal{R}[\tilde{\gamma}]} \right\}.$$

Lemma 5.29. *Let two sequences $\{w_n\}_{n=1}^\infty$ and $\{v_n\}_{n=1}^\infty \subset Y$ be given such that $w_n \to \tilde{w}$ and $v_n \to \tilde{v}$ in Y. Let (α_n, γ_n) be a solution of training scheme* (5.10) *in which we set $u_c := w_n$ and $u_\eta := v_n$. Then there exists $(\tilde{\alpha}, \tilde{\gamma})$, a solution of* (5.10) *with $u_c = \tilde{w}$ and $u_\eta = \tilde{v}$, such that $(\alpha_n, \gamma_n) \to (\tilde{\alpha}, \tilde{\gamma})$.*

Proof. In view of the box constraint and the compactness of Π, there exists $(\tilde{\alpha}, \tilde{\mathcal{R}})$ such that, up to a subsequence,
$$(\alpha_n, \gamma_n) \to (\tilde{\alpha}, \tilde{\gamma}).$$
We claim that $(\tilde{\alpha}, \tilde{\gamma})$ is a solution of (5.10) with $u_c = \tilde{w}$ and $u_\eta = \tilde{v}$.

Let u_n and $\tilde{u}$ be defined by (5.28) and (5.29), respectively. By Assumption 5.28, we have that, up to a subsequence, $u_n \to \tilde{u}$ in Y and, by Assumption 5.3,
$$\mathcal{A}(u_n - w_n, Q) \to \mathcal{A}(\tilde{u} - \tilde{w}, Q). \tag{5.30}$$

Suppose that $(\tilde{\alpha}, \tilde{\gamma})$ is not a solution of (5.10) with $u_c = \tilde{w}$ and $u_\eta = \tilde{v}$. That is, there exist (α', γ') and $\varepsilon > 0$ such that

$$\mathcal{A}(\tilde{u} - \tilde{w}, Q) > \mathcal{A}(u' - \tilde{w}, Q) + \varepsilon. \tag{5.31}$$

Let u_n' and u' be defined by letting $(\alpha_n, \gamma_n) = (\alpha', \gamma')$ and $(\tilde{\alpha}, \tilde{\gamma}) = (\alpha', \gamma')$ in (5.28) and (5.29), respectively. Then, in view of Assumption 5.28 item 1 again we have, up to a subsequence, $u_n' \to u'$ and

$$\mathcal{A}(u_n' - w_n, Q) \to \mathcal{A}(u' - \tilde{w}, Q),$$

which implies that for n large enough

$$\mathcal{A}(u_n' - w_n, Q) \le \mathcal{A}(u' - \tilde{w}, Q) + \frac{1}{4}\varepsilon \le \mathcal{A}(\tilde{u} - \tilde{w}, Q) - \frac{3}{4}\varepsilon \le \mathcal{A}(u_n - w_n, Q) - \frac{1}{2}\varepsilon.$$

where at the second inequality we invoked (5.31) and at the last inequality we used (5.30). Finally, we conclude that

$$\mathcal{A}(u_n' - w_n, Q) < \mathcal{A}(u_n - w_n, Q)$$

which contradicts the definition of (α_n, γ_n). $\square$

Theorem 5.30. *The assemble operator $\mathcal{P}(\mathcal{L})$ defined in* (5.25) *satisfies Assumption 5.22.*

Proof. Let a sequence $\{\mathcal{L}_n\}_{n=1}^{\infty} \subset \mathcal{V}_A$ be such that

$$\lim_{n\to\infty} d_{\mathcal{V}_A}(\mathcal{L}_n, \mathcal{L}) = 0, \tag{5.32}$$

where $\mathcal{L} \in \mathcal{V}_A$. Let $\mathcal{P}(\mathcal{L}_n)$ be defined as in (5.25). By Lemma 5.27, for each $L_n \in \mathcal{L}_n$ we have there exists $(\alpha_{n,L_n}, \gamma_{n,L_n})$ such that

$$\mathcal{P}(\mathcal{L}_n)(x) = \mathcal{S}(\alpha_{n,L_n}, \gamma_{n,L_n})(x), \text{ if } x \in L_n.$$

In view of (5.32), for arbitrary $L \in \mathcal{L}$ there exists a sequence $\{L_n\}_{n=1}^{\infty}$, $L_n \in \mathcal{L}_n$ for each $n \in \mathbb{N}$, such that $L_n \to L$. Since $\{(\alpha_{n,L_n}, \gamma_{n,L_n})\}_{n=1}^{\infty}$ satisfies the box constraint, up to a subsequence, we have

$$(\alpha_{n,L_n}, \gamma_{n,L_n}) \to (\tilde{\alpha}_L, \tilde{\gamma}_L). \tag{5.33}$$

Now, for a fixed $L_n \in \{L_n\}_{n=1}^{\infty}$ we may find a continuously differentiable bijective map φ_n of L onto L_n, such that φ_n^{-1} is also continuously differentiable. Let

$$u_{c,n}(x) := u_c(\varphi_n(x)) \text{ and } u_{\eta,n}(x) := u_\eta(\varphi_n(x)) \text{ for } x \in L.$$

We have

$$\lim_{n\to\infty} \|u_{c,n} - u_c\|_{Y(L)} = \lim_{n\to\infty} \|u_{\eta,n} - u_\eta\|_{Y(L)} = 0.$$

Define

$$u_n := \arg\min \left\{ \mathcal{F}\left(u - u_{\eta,n}, L\right) + \mathcal{R}[\alpha_{n,L_n}, \gamma_{n,L_n}](u, L) : \; u \in X_{\mathcal{R}[\gamma_{n,L_n}]} \right\},$$

and

$$u_L := \arg\min \left\{ \mathcal{F}\left(u - u_\eta, L\right) + \mathcal{R}[\tilde{\alpha}_L, \tilde{\gamma}_L](u, L) : \; u \in X_{\mathcal{R}[\tilde{\gamma}_L]} \right\}. \tag{5.34}$$

By Assumption 5.28, up to a subsequence, $u_n \to u_L$ in Y, and so

$$\lim_{n\to\infty} \mathcal{A}(u_n - u_c, L) = \mathcal{A}(u_L - u_c, L). \tag{5.35}$$

Moreover, since $\{\varphi_n\}_{n=1}^{\infty}$ is a sequence of continuously differentiable bijective maps, again by Assumption 5.28, we have

$$\lim_{n\to\infty} \mathcal{A}(\mathcal{P}(\mathcal{L}_n) - u_c, L_n) = \lim_{n\to\infty} \mathcal{A}(u_n - u_c, L),$$

which, together with (5.35), implies that

$$\lim_{n\to\infty} \mathcal{A}(\mathcal{P}(\mathcal{L}_n) - u_c, L_n) = \mathcal{A}(u_L - u_c, L). \tag{5.36}$$

Next, in view of (5.33), (5.34), and Lemma 5.29, we have $u_L \in E(L)$ for each $L \in \mathcal{L}$, where $E(L)$ is defined in (5.26). Hence, the function

$$\tilde{u}(x) := u_L(x) \text{ for } x \in L \in \mathcal{L}$$

belongs to $\mathcal{E}(\mathcal{L})$. Therefore, we conclude that

$$\mathcal{A}(\tilde{u} - u_c, Q) \geq \inf\left\{\mathcal{A}(u - u_c, Q) :\ u \in \mathcal{E}(\mathcal{L})\right\}. \tag{5.37}$$

Finally, since $\sup \#\{\mathcal{L}_n\} < +\infty$, we obtain

$$\lim_{n\to\infty} \mathcal{A}(\mathcal{P}(\mathcal{L}_n) - u_c, Q) = \sum_{L_n \in \mathcal{L}_n} \lim_{n\to\infty} \mathcal{A}(\mathcal{P}(\mathcal{L}_n) - u_c, L_n) = \sum_{L \in \mathcal{L}} \mathcal{A}(u_L - u_c, L) = \mathcal{A}(\tilde{u} - u_c, Q),$$

where on the second equality we invoked (5.36). Together with (5.37), we conclude that

$$\liminf_{n\to\infty} \mathcal{A}\left(\mathcal{P}(\mathcal{L}_n) - u_c, Q\right) \geq \inf\left\{\mathcal{A}(u - u_c, Q) :\ u \in \mathcal{E}(\mathcal{L})\right\} = \mathcal{A}(\mathcal{P}(\mathcal{L}) - u_c, Q)$$

as desired. □

5.3.4. *Assemble operator with spatially dependent tuning parameter.* The operator $\mathcal{P}(\mathcal{L})$ defined in (5.25) has one natural drawback: the construction is too local. To be precise, for two adjacent rectangles L_1 and L_2, the construction of $\mathcal{P}(\mathcal{L})$ in L_1 is independent of $\mathcal{P}(\mathcal{L})$ in L_2, what may cause overfitting and edging problems, especially when the constant A goes smaller.

In this section we propose another assemble operator which provides a good balance between local and global optimization. We remark that this new $\mathcal{P}(\mathcal{L})$ only works with a fixed regularizer. That is, we need to fix a regularizer $\mathcal{R}[\gamma] \in \mathscr{R}$ at the beginning. For simplicity of the notation, in Section 5.3.4 we will abbreviate $\mathcal{R}[\alpha, \gamma]$ by $\mathcal{R}[\alpha]$ since γ is fixed and we will not train with respect to γ.

Notation 5.31. Let a partition domain $\mathcal{L} \in \mathcal{V}_A$ be given.

1. Fix $L \in \mathcal{L}$. We denote by A_L a collection of positive vectors $\alpha \in [A, 1/A]^{n_{\mathscr{R}}}$, and $A_{\mathcal{L}}$ a collection of A_L i.e., $A_{\mathcal{L}} := \{A_L :\ L \in \mathcal{L}\}$;

2. We denote by $W_{A_{\mathcal{L}}}$ a collection of weighted (spatially dependent) parameters based on the collection $A_{\mathcal{L}}$. We say $\omega_{\mathcal{L}} \in W_{A_{\mathcal{L}}}$ if $\omega_{\mathcal{L}}(x)$ is constant in each $L \in \mathcal{L}$, and $\omega_{\mathcal{L}}(x) \in A_L$ if $x \in L$. Moreover, we use $W_{\mathcal{V}_A}$ to denote
$$W_{\mathcal{V}_A} := \bigcup_{\mathcal{L} \in \mathcal{V}_A} W_{A_{\mathcal{L}}}.$$
3. We say $\mathcal{W}_{A_{\mathcal{L}}}$ is a collection of weighted reconstructed images
$$\mathcal{W}_{A_{\mathcal{L}}} := \{u_\omega : \ \omega \in W_{A_{\mathcal{L}}}\},$$
where
$$u_\omega := \arg\min \{\mathcal{F}(u - u_\eta, Q) + \mathcal{R}[\omega](u, Q) : \ u \in X_{\mathcal{R}}\} \tag{5.38}$$

The following training scheme is generalized from scheme (1.15).

Level 1.
$$\tilde{\mathcal{L}} \in \arg\min \{\mathcal{A}(u_c - \mathcal{P}(\mathcal{L}), Q) : \ \mathcal{L} \in \mathcal{V}_A\},$$

Level 2.
$$\mathcal{P}(\mathcal{L}) := \arg\min \{\mathcal{A}(u - u_c, Q) : \ u \in \mathcal{W}_{A_{\mathcal{L}}}\} \tag{5.39}$$
$$\text{where } A_{\mathcal{L}} := \{A_L : \ L \in \mathcal{L}\} \text{ and } A_L \text{ is defined in (5.40),}$$

Level 3.
$$A_L := \arg\min \{\mathcal{A}(u_c - u_\alpha, L) : \alpha \in [A, 1/A]^{n_{\mathscr{R}}}\} \tag{5.40}$$
$$u_\alpha := \arg\min \{\mathcal{F}(u_\eta - u, L) + \mathcal{R}[\alpha](u, L), \ u \in X_{\mathcal{R}}\}.$$

Lemma 5.32. *The set $W_{\mathcal{V}_A}$ is a closed set under the L^2 norm.*

Proof. By Corollary 5.21 we have that the set A_L defined in (5.40) is closed, and hence the collection $W_{A_{\mathcal{L}}}$ is closed under the L^2 norm.

Let $\{\omega_n\}_{n=1}^\infty \subset W_{\mathcal{V}_A}$ be given. By definition of ω_n, there exists sequence $\{\mathcal{L}_n\}_{n=1}^\infty \subset \mathcal{V}_A$ such that $\omega_n \in W_{A_{\mathcal{L}_n}}$ for each $n \in \mathbb{N}$. By Lemma 5.25 there exists $\tilde{\mathcal{L}} \in \mathcal{V}_A$ such that, up to a subsequence, $\mathcal{L}_n \to \tilde{\mathcal{L}}$. Fixing an arbitrary $\tilde{L} \in \tilde{\mathcal{L}}$, we may extract a sequence $\{L_n\}_{n=1}^\infty$ from $\{\mathcal{L}_n\}_{n=1}^\infty$ such that $\chi_{L_n} \to \chi_{\tilde{L}}$. Since ω_n is constant on L_n, we have $\omega_n \lfloor_{L_n} \to a_{\tilde{L}}$ where $a_{\tilde{L}} \in [A, 1/A]^{n_{\mathscr{R}}}$, and by Lemma 5.29 we have $a_{\tilde{L}} \in A_{\tilde{L}}$.

In the end, since there are only finitely many $\tilde{L}$ inside $\tilde{\mathcal{L}}$, we may repeat the above argument only finitely many times and conclude that $\omega_n \to \tilde{\omega}$ in L^2, where $\tilde{\omega}(x) := a_{\tilde{L}}$ for $x \in \tilde{L} \in \tilde{\mathcal{L}}$. □

Assumption 5.33. *Let $\{\omega_n\}_{n=1}^\infty$ be a sequence of piecewise constant function such that $\omega_n \to \omega$ in L^2. Then $u_{\omega_n} \to u_\omega$ in Y, where u_{ω_n} and u_ω are defined in* (5.38)*.*

Theorem 5.34. *The assemble operator $\mathcal{P}(\mathcal{L})$ defined in* (5.39) *satisfies Assumption 5.22.*

Proof. Let $\{\mathcal{L}_n\}_{n=1}^{\infty}$ and $\mathcal{L}$ be such that

$$d_{\mathcal{V}_A}(\mathcal{L}_n, \mathcal{L}) = 0.$$

By Lemma 5.32 there exist $\omega_n \in W_{A_{\mathcal{L}_n}}$ such that

$$\mathcal{P}(\mathcal{L}_n) = \arg\min \{\mathcal{F}(u - u_\eta, Q) + \mathcal{R}[\omega_n](u, Q) : \ u \in X_{\mathcal{R}}\}.$$

Since $\{\omega_n\}_{n=1}^{\infty}$ satisfies the box-constraint, and finitely piece-wise constant, there exists ω, a finitely piecewise constant function, such that $\omega_n \to \omega$ in L^2, and by Lemma 5.32 we have

$$\omega \in W_{A_{\mathcal{L}}}. \tag{5.41}$$

Next, in view of Assumption 5.33 we obtain $\mathcal{P}(\mathcal{L}_n) \to \tilde{u}$, where

$$\tilde{u} = \arg\min \{\mathcal{F}(u - u_\eta, Q) + \mathcal{R}[\omega](u, Q) : \ u \in X_{\mathcal{R}}\}.$$

Since $\omega \in W_{A_{\mathcal{L}}}$, we have $\tilde{u} \in \mathcal{W}_{A_{\mathcal{L}}}$, an hence, in view of Assumption 5.4, (5.41), and (5.39), we deduce that

$$\liminf \mathcal{A}(u_c - \mathcal{P}(\mathcal{L}_n)) \geq \mathcal{A}(u_c - \tilde{u}) \geq \mathcal{A}(u_c - \mathcal{P}(\mathcal{L}))$$

as desired. □

Chapter 6. Work in progress and future projects

In Chapter 5, we studied necessary conditions for regularizer spaces and assemble operators, such that the scheme $(\mathcal{CT})$, admit a solution. As a next step of my work in this direction, I will derive more meaningful constructions of regularizer spaces and assemble operators.

6.4. **The arsenal of regularizer spaces.** Concerning regularizer spaces, one example has been provided in Chapter 2, but only the one-dimensional case was investigated, and this may be of marginal interest within the context of image reconstruction.

The two dimensional setting of fractional order generalized total variation is proposed below and is being currently undertaken in [31].

6.4.1. *The TGV^r regularizers with Riemann-Liouville (R-L) fractional derivative.* For $x \in (0,1)$, $k \in \mathbb{N}_0$, and $0 < s < 1$, we define the order s left-sided Riemann-Liouville derivative by:

$$d_{[0,x]}^{k+s} u(x) = \frac{1}{\Gamma(k+1-s)} \left(\frac{d}{dx}\right)^{k+1} \int_0^x \frac{u(y)}{(x-y)^s} dy$$

where $\Gamma(s) := \int_0^\infty e^{-t} t^{s-1} dt$. For example, $k = 0$, then $d_{[0,x]}^s$ is defined a.e. and we can use the fractional order derivative in the same manner of usual integer order derivative. That is, we may define $\nabla^s u$ by

$$\nabla^s u := \left(d_{[0,x]}^s u, d_{[0,y]}^s u\right).$$

We introduce TGV^{k+s}, $k \in \mathbb{N}$ and $0 < s < 1$, as

$$TGV^{k+s}(u) = \min \Big\{ |\nabla u - v_0|_{\mathcal{M}_b} + |\mathcal{E} v_0 - v_1|_{\mathcal{M}_b} + \cdots + |\mathcal{E} v_{k-2} - s v_{k-1}|_{\mathcal{M}_b} + s\, |\mathcal{E}^s v_{k-1}|_{\mathcal{M}_b} :$$

$$v_l \in BV(Q, \mathrm{Sym}^l(\mathbb{R}^2)),\ l = 0, \dots, k-1\Big\}$$

where $\mathcal{E}^s$ denotes the fractional order symmetric derivative. Hence, we can define a regularizer space $\mathscr{R}$ by $\mathscr{R} := \{TGV^r :\ r \in [1, R]\}$, where $R > 1$.

The proof that $\mathscr{R}$ satisfies Assumption 5.12 is undertaking in [31].

6.4.2. *$\mathscr{A}$-$\mathscr{B}$ Morrey-quasiconvex regularizer [30].* A more generalized regularizer space can by constructed by using the $\mathscr{A}$-$\mathscr{B}$ (Morrey)-quasiconvex operator theory (see [26, 47]).[2] To be precise, let $0 \le s \le 1$ be given and let $\mathscr{A}$ be a differential operator of the form

$$\mathscr{A} u := \sum_{i,j=1,2} A^{ij} \frac{\partial^{1+s}}{\partial x_i \partial x_j} u \quad \text{for every } u \in L^1_{\mathrm{loc}}(Q; \mathbb{R}^3),$$

where ∂^{1+s} is the fractional order derivative, and $A^{ij} \in \mathbb{M}^{3\times 3}$ (see Notation 5.8), $i, j = 1, 2$. Let $\mathscr{B}$ be a first order differential operator such that

$$\mathscr{B} u := \sum_{k=1,2} B^k \frac{\partial}{\partial x_k} u \quad \text{for every } u \in L^1_{\mathrm{loc}}(Q),$$

where $B^k \in \mathbb{R}^3$ for $k = 1, 2$. We define the regularizer (seminorm)

$$\begin{aligned} &ABQ^{1+s}_{A,B}(u) \\ &:= \inf\left\{ \|\nabla u - v\|_{\mathcal{M}(Q;\mathbb{R}^3)} + \int_Q \mathcal{Q}_{\mathscr{A}} f(\mathscr{B} v(x))\, dx : v \in L^1(Q; \mathbb{R}^3),\ \mathscr{B} v \in L^1(Q; \mathbb{R}^3) \right\}, \end{aligned} \tag{6.1}$$

where $\mathcal{Q}_{\mathscr{A}} f$ represents the $\mathscr{A}$-quasiconvex envelope of f, in the sense of Definition 3.2 in [47], with $f : \mathbb{M}^{3\times 3} \to [0, +\infty)$ is Lipschitz continuous, and there exists $C > 0$ such that $C^{-1}|\xi| \le f(\xi) \le C|\xi|$ for every $\xi \in \mathbb{M}^{3\times 3}$ and same $C > 0$.

The proof that $\mathscr{R}$, composed by $ABQ^{1+s}_{A,B}$ defined in (6.1), satisfies Assumption 5.12 is undertaken in [30].

6.4.3. *Generalized Mumford-Shah functional.* For $k \in \mathbb{N}$, we set

$$\begin{aligned} &MS^k_{\tilde{\alpha},\tilde{\beta}}(u) := \\ &\quad \inf\Bigg\{ \alpha_0 \int_\Omega |\nabla u - v_0|^2\, dx + \alpha_1 \int_\Omega |\mathcal{E} v_0 - v_1|^2\, dx + \cdots + \alpha_{k-1} \int_\Omega |\mathcal{E} v_{k-2} - v_{k-1}|^2\, dx \\ &\quad + \alpha_k \int_\Omega |\mathcal{E} v_{k-1}|^2\, dx + \beta_0 \mathcal{H}^{N-1}(S_u) + \beta_1 \mathcal{H}^{N-1}(S_{v_0^1} \cup S_{v_0^2}) + \beta_{l+1} \mathcal{H}^{N-1}\left(\bigcup_{i=1}^{C_l} S_{v_l^i} \right) + \\ &\quad\quad \cdots + \beta_k \mathcal{H}^{N-1}\left(\bigcup_{i=1}^{C_k} S_{v_{k-1}^i} \right) :\ v_l \in GSBV(\Omega, \mathrm{Sym}^l(\mathbb{R})),\ l = 0, \dots, k-1 \Bigg\}, \end{aligned} \tag{6.2}$$

[2] The notation of quasiconvexity is unrelated to the concept of quasi-convexity introduced in (1.26)

where $\tilde{\alpha} := (\alpha_0, \ldots, \alpha_k)$, $\tilde{\beta} := (\beta_0, \ldots, \beta_k) \in \mathbb{R}^{k+1}$, C_l denotes the number of components in v_l, $\mathcal{E}$ the symmetric derivative, and Sym^k the space of symmetric tensors of order k. Moreover, an approximation scheme of (6.2) based on the framework of the Ambrosio - Tortorelli functionals, as well as the relevant numerical scheme, are undertaken in [61].

6.5. **Γ-convergence with non-negative spatially dependent parameters.** It has been observed that, at least in one dimension we may reduce the staircasing effect by allowing the weight function ω to be 0 in certain subdomains (see [50]). It is then interesting to investigate precisely whether the same holds in two dimensions, and to study the corresponding Γ-convergence problem.

We introduce some notations and definitions in order to precisely state the problem.

Notation 6.1. Let $\Omega \subset \mathbb{R}^N$ be an open bounded Lipschitz domain, and let $\omega \in SBV(\Omega)$ be a non-negative function. Let $S \subset \Omega$ be given.

1. We say that $S \in \mathcal{R}(\Omega)$ if $\bar{S}$ (the closure of S) is $\mathcal{H}^{N-1}$-rectifiable and $\mathcal{H}^{N-1}(\bar{S} \setminus S) = 0$ (Note that if $\bar{S}$ is $\mathcal{H}^{N-1}$-rectifiable then S is $\mathcal{H}^{N-1}$-rectifiable (See [4], Proposition 2.76)).
2. We set $P^t(\omega) := \{x \in \Omega : \ \omega(x) > t\}$, for $t > 0$, and
$$P^\infty(\omega) := \bigcap_{t>0} P^t(\omega) \text{ and } P^0(\omega) := \bigcap_{t>0} \left(\Omega \setminus P^t(\omega)\right).$$
3. We set $S_\delta := \{x \in \Omega : \ \mathrm{dist}(x, S) < \delta\}$ for $A \subset \Omega$ and $\delta > 0$.

We allow ω to be 0 in certain subdomains as follows:

Definition 6.2. *Let ω: $\Omega \to [0, +\infty]$ belong to $SBV(\Omega)$.*

1. *We say that $\omega \in \mathcal{P}(\Omega)$ if $\mathcal{H}^{N-1}(S_\omega) < +\infty$ and $P^0(\omega) \cup P^\infty(\Omega) \in \mathcal{R}(\Omega)$.*
2. *We say that $\omega \in \mathcal{P}_r(\Omega)$ if $\omega \in \mathcal{P}(\Omega)$ and*
$$\lim_{\delta \to 0} \left[\int_{\partial((P^\infty(\omega))_\delta)} \omega \, d\mathcal{H}^{N-1} + \int_{\partial((P^0(\omega))_\delta)} \omega \, d\mathcal{H}^{N-1} \right] = 0.$$
3. *We say that $\omega \in \mathcal{P}_b(\Omega)$ if $\omega \in \mathcal{P}(\Omega)$ and satisfies* (3.3).

We also define the function spaces SBV_ω and $GSBV_\omega$.

Definition 6.3. *Let $\omega \in \mathcal{P}(\Omega)$ be given. We say that $u \in SBV_\omega(\Omega)$ if $u \in L^1(\Omega)$, $u \in SBV(\Omega \setminus (P^0(\omega))_\delta)$ for every $\delta > 0$, and*
$$\int_\Omega |\nabla u|^2 \, \omega \, dx + \int_{S_u^0} \left|u^+ - u^-\right| \omega \, d\mathcal{H}^{N-1} < +\infty,$$
where the jump set S_u^0 of $u \in SBV_\omega(\Omega)$, with a vanishing parameter ω, is defined by
$$S_u^0 := \left(\bigcup_{\delta > 0} S_u^\delta \right) \cup P^0(\omega)$$
where S_u^δ denotes the jump set of u in $SBV(\Omega \setminus (P^0(\omega))_\delta)$. Moreover, we say that $u \in GSBV_\omega(\Omega)$ if $K \wedge u \vee -K \in SBV_\omega(\Omega)$ for all $K \in \mathbb{N}$.

The proof of following theorem is undertaken in [46].

Theorem 6.4. *Let $\Omega \subset \mathbb{R}^N$ be an open bounded Lipschitz domain, let $\omega \in \mathcal{P}_r(\Omega)$, and for $k \in \mathbb{N}$, $\varepsilon > 0$, let $\mathcal{AT}^k_{\omega,\varepsilon}\colon L^1(\Omega) \times L^1(\Omega) \to [0, +\infty]$ be defined by*

$$\mathcal{AT}^k_{\omega,\varepsilon}(u,v) := \begin{cases} AT^k_{\omega,\varepsilon}(u,v) & \text{if } (u,v) \in W^{1,2}(\Omega) \times W^{1,2}(\Omega),\ 0 \le v \le 1, \\ +\infty & \text{otherwise.} \end{cases}$$

Then the functionals $\mathcal{AT}^k_{\omega,\varepsilon}$ Γ-converge, with respect to the $L^1 \times L^1$ topology, to the functional

$$\mathcal{MS}_\omega(u,v) := \begin{cases} MS_\omega(u) & \text{if } u \in GSBV_\omega(\Omega) \text{ and } v = 1 \text{ a.e.,} \\ +\infty & \text{otherwise.} \end{cases}$$

We remark that the techniques we developed here can be adapted to other functional models. For example,

1. the weighted Cahn-Hilliard model defined as

$$CH_{\omega,\varepsilon}(u) := \int_I \left[\varepsilon\, |\nabla u(x)|^2 + \frac{1}{\varepsilon} W(u) \right] \omega\, dx,$$

   for $u \in W^{1,2}(\Omega)$ and with a double well potential function $W\colon \mathbb{R} \to [0, +\infty)$ such that $\{W = 0\} = \{0, 1\}$ with the Γ-limit $CH_\omega(u) := c_W P_\omega(u)$ defined for $u = \chi_E \in BV_\omega(\Omega)$, where

$$c_W := 2 \int_0^1 \sqrt{W(s)}\, ds \text{ and } P_\omega(u) := \int_{S_u} \omega^- d\mathcal{H}^{N-1};$$

2. the weighted version of functionals involving the L^1-norm of the gradient [2].

$$G_\omega(u) := \int_\Omega |\nabla u|\, \omega\, dx + \int_{S_u} \frac{1}{2} \left[\omega(x^-) \right] g(\left|u^+ - u^-\right|) d\mathcal{H}^{N-1},$$

   where $u \in SBV_\omega(\Omega)$, $v = 1$ *a.e.*;

$$G_{\omega,\varepsilon}(u,v) := \int_\Omega \varphi(v)\, |\nabla u|\, \omega\, dx + \int_\Omega \left[\frac{1}{\varepsilon} W(v) + \varepsilon\, |\nabla v|^2 \right] \omega\, dx, \quad u \in W^{1,2}_\omega(\Omega),\ v \in W^{1,2}(\Omega).$$

Appendix A.

Definition A.1 ([5], Definition 4.4.9)**.** *Let $\mathcal{X}$ be a metric space. We denote by $\mathcal{C}_\mathcal{X}$ the family of all nonempty closed subsets of X. Then*

$$d_\mathcal{H}(C, D) := \min \{1, h(C, D)\}, \quad C, D \in \mathcal{C}_\mathcal{X},$$

where

$$h(C, D) := \inf \{\delta \in [0, +\infty] : C \subset D_\delta \text{ and } D \subset C_\delta\},$$

is a metric on $\mathcal{C}_\mathcal{X}$, and is called the Hausdorff *distance between the set C and D (see Notation 3.3 for definition of D_δ and C_δ).*

Consider $\mathcal{X}$ to be the interval $(0, 1)$ with the Euclidian distance. We remark that for two intervals $[a_1, b_1]$ and $[a_2, b_2]$ in $(0, 1)$,

$$d_{\mathcal{H}}([a_1, b_1],\, [a_2, b_2]) = \min\{1,\, \max\{|a_1 - a_2|\,, |b_1 - b_2|\}\}\,. \tag{A.1}$$

Indeed, the δ-neighborhood of $[a_1, b_1]$ is $[a_1 - \delta, b_1 + \delta]$, and contains $[a_2, b_2]$ if and only if

$$\delta \geq \max\{a_1 - a_2, b_2 - b_1\}\,.$$

Similarly, the δ-neighborhood of $[a_2, b_2]$ contains $[a_1, b_1]$ if and only if

$$\delta \geq \max\{a_2 - a_1, b_1 - b_2\}\,,$$

and we conclude (A.1).

Lemma A.2. *Let $I_n := [a_n, b_n] \subset (-1, 1)$. Then, up to the extraction of a subsequence,*

$$I_n \overset{\mathcal{H}}{\to} I_\infty \subset (-1, 1),$$

where I_∞ is connected and closed in $(-1, 1)$, and

$$\mathcal{L}^1(I_\infty) = \lim_{n\to\infty} \mathcal{L}^1(I_n).$$

Moreover, for arbitrary $K \subset\subset I_\infty$, K must be contained in I_n for n large enough.

Proof. Because $I_n \subset (-1, 1)$, we have that $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ are bounded and so, up to the extraction of a subsequence, there exist

$$a_\infty := \lim_{n\to\infty} a_n \text{ and } b_\infty := \lim_{n\to\infty} b_n, \tag{A.2}$$

where $-1 \leq a_\infty \leq b_\infty \leq 1$. We define $I_\infty := [a_\infty, b_\infty]$ if $-1 < a_\infty \leq b_\infty < 1$, $I_\infty := (-1, b_\infty]$ if $a_\infty = -1$, and $I_\infty := [a_\infty, 1)$ if $b_\infty = 1$. Hence I_∞ is connected and closed in $(-1, 1)$ (in the case in which $a_\infty = b_\infty = -1$, or $a_\infty = b_\infty = 1$, we have $I_\infty = \varnothing$ and it is still closed in $(-1, 1)$).

Therefore

$$\lim_{n\to\infty} d_{\mathcal{H}}(I_n,\, I_\infty) = \lim_{n\to\infty} \max\{|a_n - a_\infty|\,, |b_n - b_\infty|\} = 0,$$

and we have for $I_\infty \neq \varnothing$,

$$\mathcal{L}^1(I_\infty) = b_\infty - a_\infty = \lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty} \mathcal{L}^1(I_n),$$

as desired.

Next, if $K \subset\subset I_\infty$ then $K \subset (\alpha, \beta)$ for some α, β such that $a_\infty < \alpha < \beta < b_\infty$. By (A.2) choose N large enough such that for all $n \geq N$,

$$a_n < \alpha < \beta < b_n,$$

so that $K \subset I_n$ for all $n \geq N$. $\square$

Lemma A.3. *Let $\{v_\varepsilon\}_{\varepsilon>0} \subset W^{1,2}(I)$ be such that $0 \leq v_\varepsilon \leq 1$, $v_\varepsilon \to 1$ in $L^1(I)$ and pointwise a.e., and*

$$\limsup_{\varepsilon\to 0} \int_I \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right] dx < \infty. \tag{A.3}$$

Then for arbitrary $0 < \eta < 1$ there exists an open set $H_\eta \subset I$ satisfying the following properties:

1. the set $I \setminus H_\eta$ is a collection of finitely many points in I;
2. for every set K compactly contained in H_η, we have $K \subset B_\varepsilon^\eta$ for $\varepsilon > 0$ small enough, where

$$B_\varepsilon^\eta := \left\{x \in I : v_\varepsilon^2(x) \geq \eta\right\}.$$

Proof. Choose a constant $M > 0$ such that

$$M \geq \limsup_{\varepsilon\to 0} \int_I \left[\frac{\varepsilon}{2}\left|v_\varepsilon'\right|^2 + \frac{1}{2\varepsilon}(v_\varepsilon - 1)^2\right] dx \geq \limsup_{\varepsilon\to 0} \int_I \left|v_\varepsilon'\right| |1 - v_\varepsilon|\, dx = \limsup_{\varepsilon\to 0} \frac{1}{2}\int_I \left|c_\varepsilon'\right| dx,$$

where $c_\varepsilon(x) := (1 - v_\varepsilon(x))^2$. Note that by (A.3), $c_\varepsilon \to 0$ in $L^1(I)$. Fix σ, δ with

$$0 < \sigma < \delta < 1.$$

By the co-area formula, for $0 < \varepsilon < \varepsilon_0$ with ε_0 sufficiently small, we have

$$2M + 1 \geq \int_I \left|c_\varepsilon'(x)\right| dx = \int_{-\infty}^{\infty} \mathcal{H}^0(\{x : c_\varepsilon(x) = t\})\, dt \geq \int_\sigma^\delta \mathcal{H}^0(\{x : c_\varepsilon(x) = t\})\, dt.$$

Hence, for each $\varepsilon > 0$ there exists $\delta_\varepsilon \in (\sigma, \delta)$ such that

$$\frac{2M+1}{\delta - \sigma} \geq \mathcal{H}^0(\{x : c_\varepsilon(x) = \delta_\varepsilon\}). \tag{A.4}$$

Define, for a fixed $r > 0$,

$$A_\varepsilon^r := \{x \in I : c_\varepsilon(x) \leq r\}.$$

Since $v_\varepsilon \in W^{1,2}(I)$, v_ε is continuous and so is c_ε, therefore $A_\varepsilon^{\delta_\varepsilon}$ is closed and has at most $(2M+1)/(\delta-\sigma)+1$ connected components because of (A.4) and in view of the continuity of c_ε. Note that the number $(2M+1)/(\delta - \sigma)$ does not depend on $\varepsilon > 0$.

For $\varepsilon \in (0, \varepsilon_0)$ and $k \in \mathbb{N}$ depending only on $\delta - \sigma$ and M, we have

1. $A_\varepsilon^{\delta_\varepsilon} = \bigcup_{i=1}^k I_\varepsilon^i$, where each I_ε^i is a closed interval or $\varnothing$;
2. for all $i < j$, $\max\left\{x : x \in I_\varepsilon^i\right\} < \min\left\{x : x \in I_\varepsilon^j\right\}$.

By Lemma A.2, up to the extraction of a subsequence, for each $i \in \{1, 2, \ldots, k\}$ let I_0^i be the *Hausdorff* limit of the I_ε^i as $\varepsilon \to 0$, i.e., $I_\varepsilon^i \xrightarrow{\mathcal{H}} I_0^i$, with I_0^i is connected and closed in I, and for all $i < j$, $\max I_0^i \leq \min I_0^j$.

Set

$$T_\delta := \bigcup_{i=1}^k (I_0^i)^\circ \text{ and } T_{\delta,\varepsilon} := \bigcup_{i=1}^k (I_\varepsilon^i)^\circ, \tag{A.5}$$

where by $(\cdot)^\circ$ we denote the interior of a set. Since

$$I \setminus A_\varepsilon^{\delta_\varepsilon} \subset \{x \in I : \ c_\varepsilon(x) \geq \sigma\}$$

and $c_\varepsilon \to 0$ in $L^1(I)$, by Chebyshev's inequality we have

$$\lim_{\varepsilon\to 0} \mathcal{L}^1(T_{\delta,\varepsilon}) = \lim_{\varepsilon\to 0} \mathcal{L}^1(A_\varepsilon) = 2.$$

Moreover, since $T_{\delta,\varepsilon} \xrightarrow{\mathcal{H}} T_\delta$, by Lemma A.2 we have

$$\mathcal{L}^1(T_\delta) = \sum_{i=1}^k \mathcal{L}^1(I_0^i)^\circ = \sum_{i=1}^k \lim_{\varepsilon\to 0} \mathcal{L}^1(I_\varepsilon^i)^\circ = \lim_{\varepsilon\to 0} \sum_{i=1}^k \mathcal{L}^1(I_\varepsilon^i)^\circ = \lim_{\varepsilon\to 0} \mathcal{L}^1(T_{\delta,\varepsilon}) = 2.$$

Thus $|I \setminus T_\delta| = 0$. Moreover, since T_δ has at most k connected components, $I \setminus T_\delta$ is a finite collection of points in I.

Next, let $K \subset\subset T_\delta$ be a compact subset. We claim that K must be contained in $A_\varepsilon^{\delta_\varepsilon}$ for $\varepsilon > 0$ small enough. Recall I_0^i and I_ε^i from (A.5). Define $K_i := K \cap (I_0^i)^\circ$ for $i = 1, \dots, k$. Then $K_i \subset\subset (I_o^i)^\circ$ for each i, and so by Lemma A.2 there exists $\varepsilon_i > 0$ such that for all $0 < \varepsilon < \varepsilon_i$, $K_i \subset I_\varepsilon^i$. Define

$$\varepsilon' := \min_{i\in\{1,\dots,k\}} \{\varepsilon_i\}.$$

For $0 < \varepsilon < \varepsilon'$ we have $K_i \subset I_\varepsilon^i$, and so

$$K = \bigcup_{i=1}^k K_i \subset \bigcup_{i=1}^k I_\varepsilon^i = A_\varepsilon^{\delta_\varepsilon}.$$

Finally, given $\eta \in (0,1)$, set $\delta := \left(1 - \sqrt{\eta}\right)^2$ with $H_\eta := T_{(1-\sqrt{\eta})^2}$ and $B_\varepsilon^\eta := A_\varepsilon^{(1-\sqrt{\eta})^2}$, and properties 1 and 2 are satisfied. □

Appendix B.

B.1. *The forward and backward properties of $\mathscr{L}$.* Let operator $\mathscr{L}$ be defined as in (1.25) and $v \in L^2$ be given. We define the *forward error* by

$$\mathscr{F}(\alpha) := \|v - \mathscr{L}(v,\alpha)\|^2_{L^2(I)}, \tag{B.1}$$

and the *backward error* by

$$\mathscr{B}(\alpha) := \frac{1}{2} \|\mathscr{L}(v,\alpha) - (v)_I\|^2_{L^2(I)} \tag{B.2}$$

where

$$(v)_I := \fint_I v\,dx.$$

In Proposition B.1 and B.2 we establish some basic properties of (B.1) and (B.2).

Proposition B.1. *Let $v \in L^\infty(I)$ be given. Then the forward error $\mathscr{F}(\cdot)$ is non-decreasing.*

Proof. Let $0 \leq \alpha_1 < \alpha_2$ be given. We observe that

$$\frac{1}{2}\left\|\mathscr{L}(\alpha_1, v) - v\right\|^2_{L^2(I)} + \alpha_1 TV\left(\mathscr{L}(\alpha_1, v)\right) \leq \frac{1}{2}\left\|\mathscr{L}(\alpha_2, v) - v\right\|^2_{L^2(I)} + \alpha_1 TV\left(\mathscr{L}(\alpha_2, v)\right) \tag{B.3}$$

and

$$\frac{1}{2}\left\|\mathscr{L}(\alpha_2, v) - v\right\|^2_{L^2(I)} + \alpha_2 TV\left(\mathscr{L}(\alpha_2, v)\right) \leq \frac{1}{2}\left\|\mathscr{L}(\alpha_1, v) - v\right\|^2_{L^2(I)} + \alpha_2 TV\left(\mathscr{L}(\alpha_1, v)\right).$$

Adding up the previous two inequalities yields

$$\alpha_1 TV\left(\mathscr{L}(\alpha_1, v)\right) + \alpha_2 TV\left(\mathscr{L}(\alpha_2, v)\right) \leq \alpha_1 TV\left(\mathscr{L}(\alpha_2, v)\right) + \alpha_2 TV\left(\mathscr{L}(\alpha_1, v)\right),$$

that is,

$$(\alpha_2 - \alpha_1)\, TV\left(\mathscr{L}(\alpha_2, v)\right) \leq (\alpha_2 - \alpha_1)\, TV\left(\mathscr{L}(\alpha_1, v)\right)$$

which implies that

$$TV\left(\mathscr{L}(\alpha_2, v)\right) \leq TV\left(\mathscr{L}(\alpha_1, v)\right). \tag{B.4}$$

Hence, in the view of (B.3) and (B.4) we have

$$\begin{aligned}
&\frac{1}{2}\left\|\mathscr{L}(\alpha_1, v) - v\right\|^2_{L^2(I)} + \alpha_1 TV\left(\mathscr{L}(\alpha_1, v)\right) \\
&\leq \frac{1}{2}\left\|\mathscr{L}(\alpha_2, v) - v\right\|^2_{L^2(I)} + \alpha_1 TV\left(\mathscr{L}(\alpha_2, v)\right) \\
&\leq \frac{1}{2}\left\|\mathscr{L}(\alpha_2, v) - v\right\|^2_{L^2(I)} + \alpha_1 TV\left(\mathscr{L}(\alpha_1, v)\right)
\end{aligned}$$

which implies that $f(\alpha)$ is non-decreasing as desired. □

Proposition B.2. *Let $w \in BV(I)$ be a piecewise constant function with M pieces as defined in Notation 4.6. Then the backward error $\mathscr{B}(\alpha)$, defined in* (B.2)*, is continuous, piecewise convex, and strict decreasing to* 0.

Proof. According to Lemma 4.15 there exist $M' \leq M$ positive numbers

$$0 = \alpha_0 < \alpha_1 < \cdots \alpha_{M-1} < \alpha_M < +\infty \tag{B.5}$$

such that items 1-3 in Lemma 4.15 hold. Without loss of generality, we assume that

$$(w)_I = 0. \tag{B.6}$$

We claim that for arbitrary $0 \leq i < M'$, $i \in \mathbb{N}$, we have that $\mathscr{B}'(\alpha) < 0$ for all $\alpha \in [\alpha_i, \alpha_{i+1})$. We first deal with $0 \leq \alpha < \alpha_1$. Let

$$\mathcal{A} := \left\{I_M(j),\ I_M(j) \in \mathcal{C}_E(w) \cup \mathcal{C}_B(w),\ 1 \leq j \leq M\right\}.$$

Case 1: if $I_M(j)$ is a low (high) extreme (boundary) region and $w(I_M(j)) < (>)0$ for all $I_M(j) \in \mathcal{A}$, we are done. Indeed, we observe that, for each region $I_M(j) \in \mathcal{A}$ and in view of (B.6),

$$\int_{I_M(j)} \left|\mathscr{L}(\alpha, w) - (w)_I\right|^2 dx = \int_{I_M(j)} \left|\mathscr{L}(\alpha, w)(x)\right|^2 dx$$

$$= \begin{cases} |I_M(j)| \left(|w(I_M(j))| - \frac{2\alpha}{|I_M(j)|} \right)^2, & \text{if } I_M(j) \in \mathcal{C}_E(w), \\ |I_M(j)| \left(|w(I_M(j))| - \frac{\alpha}{|I_M(j)|} \right)^2, & \text{if } I_M(j) \in \mathcal{C}_B(w). \end{cases}$$

and hence

$$\frac{1}{2}\frac{d}{d\alpha}\left(\int_{I_M(j)} |\mathscr{L}(\alpha, w) - (w)_I|^2 \, dx \right) = \begin{cases} 2\left(\frac{2\alpha}{|I_M(j)|} - |w(I_M(j))| \right) < 0, & \text{if } I_M(j) \in \mathcal{C}_E(w), \\ \left(\frac{\alpha}{|I_M(j)|} - |w(I_M(j))| \right) < 0, & \text{if } I_M(j) \in \mathcal{C}_B(w), \end{cases}$$

as long as $2\alpha < |I_M(j)| \, |w(I_M(j))|$ if $I_M(j) \in \mathcal{C}_E(w)$, or $\alpha < |I_M(j)| \, |w(I_M(j))|$ if $I_M(j) \in \mathcal{C}_B(w)$. Therefore, we have $\mathscr{B}'(\alpha) < 0$ if $0 \le \alpha < \alpha'$, where

$$\alpha' := \min \left\{ \frac{1}{2} |I_M(j)| \, |w(I_M(j))| \, , \left| I_M(j') \right| \left| w(I_M(j')) \right| , \; I_M(j) \in \mathcal{A} \cap \mathcal{C}_E(w), \, I_M(j') \in \mathcal{C}_B(w) \right\}. \tag{B.7}$$

The case $\alpha \ge \alpha'$ will be dealt with later.

Case 2: there exists $j_0 \in A$ such that $I_M(j_0)$ is a low extreme region but $w(I_M(j_0)) \ge 0$ (the case $I_M(j_0)$ is a high extreme region but $w(I_M(j_0)) \le 0$ could be dealt in a similar way). Then we have

$$\int_{I_M(j_0)} |\mathscr{L}(\alpha, w)(x)|^2 \, dx = |I_M(j_0)| \left(w(I_M(j_0)) + \frac{2\alpha}{|I_M(j_0)|} \right)^2,$$

and hence

$$\frac{1}{2}\frac{d}{d\alpha}\left(\int_{I_M(j_0)} |\mathscr{L}(\alpha, w)|^2 \, dx \right) = 2\left(\frac{2\alpha}{|I_M(j_0)|} + w(I_M(j_0)) \right) > 0, \tag{B.8}$$

which might cause $\mathscr{B}'(\alpha) > 0$.

However, in view of (4.2), if $I_M(j_0)$ is a low extreme region, there must exist two indexes $1 \le j_0', j_0'' \le M$ such that

$$0 \le w(I_M(j_0)) < \min \left\{ w(I_M(j_0')), w(I_M(j_0'')) \right\}, \tag{B.9}$$

and hence one of the following three situations must hold:

1. $I_M(j_0')$ and $I_M(j_0'')$ are two high extreme regions, or
2. $I_M(j_0')$ is a high extreme region and $I_M(j_0'')$ is a high boundary region (that is, $j_0'' = 1$ or M), or
3. $j_0' = 1$ and $j_0'' = M$, i.e., two high boundary regions.

(if not, $I_M(j_0)$ would not be a low extreme region).

We treat situation 1 first. We first assume that $j_0' = j_0 + 1$, $j_0'' = j_0 - 1$, and

$$(|I_M(j_0)| + |I_M(j_0 + 1)|) \, |w(I_M(j_0)) - w(I_M(j_0 + 1))|$$

$$\leq (|I_M(j_0)| + |I_M(j_0-1)|)\,|w(I_M(j_0)) - w(I_M(j_0-1))|\,. \quad \text{(B.10)}$$

We note that, in view of (B.9),

$$w(I_M(j_0)) + \frac{2\alpha}{|I_M(j_0)|} < w(I_M\,(j_0+1)) - \frac{2\alpha}{|I_M\,(j_0+1)|}, \quad \text{(B.11)}$$

for $\alpha < \bar{\alpha}$, where $\bar{\alpha}$ is defined that that

$$w(I_M(j_0)) + \frac{2\tilde{\alpha}}{|I_M(j_0)|} = w(I_M\,(j_0+1)) - \frac{2\tilde{\alpha}}{|I_M\,(j_0+1)|}.$$

By (B.10) we have that

$$\begin{aligned}&\frac{1}{2}\frac{d}{d\alpha}\int_{I_M(j_0+1)} |\mathscr{L}(\alpha,w)|^2\,dx\\ &= \frac{1}{2}\frac{d}{d\alpha}\left[|I_M\,(j_0+1)|\left(w(I_M\,(j_0+1)) - \frac{2\alpha}{|I_M\,(j_0+1)|}\right)^2\right]\\ &= 2\left(\frac{2\alpha}{|I_M\,(j_0+1)|} - w\,(I_M\,(j_0+1))\right)\end{aligned}$$

for all $\alpha < \bar{\alpha}$. Hence, in the view of (B.8) and (B.11), we have that

$$\begin{aligned}&\frac{1}{2}\frac{d}{d\alpha}\left(\int_{I_M(j_0)} |\mathscr{L}(\alpha,w)|^2\,dx\right) + \frac{1}{2}\frac{d}{d\alpha}\left(\int_{I_M(j_0+1)} |\mathscr{L}(\alpha,w)|^2\,dx\right) \quad \text{(B.12)}\\ &= 2\left(\frac{2\alpha}{|I_M(j_0)|} + w(I_M(j_0))\right) + 2\left(\frac{2\alpha}{|I_M\,(j_0+1)|} - w\,(I_M\,(j_0+1))\right)\\ &= 2\left[\left(w(I_M(j_0)) + \frac{2\alpha}{|I_M(j_0)|}\right) - \left(w(I_M\,(j_0+1)) - \frac{2\alpha}{|I_M\,(j_0+1)|}\right)\right] < 0\end{aligned}$$

for all $\alpha < \bar{\alpha}$. Note that at $\alpha = \bar{\alpha}$, we have

$$\mathscr{L}(\bar{\alpha},w)(I_M(j_0)) = \mathscr{L}(\bar{\alpha},w)(I_M(j_0+1)),$$

which means that $\mathscr{L}(\bar{\alpha},w)$ has at most $M-1$ constant pieces. Therefore, we have $\bar{\alpha} \geq \alpha_1$, where α_1 is obtained in (B.5). If (B.10) does not hold, we use the region $I_M(j_0-1)$ instead of $I_M(j_0+1)$ in (B.11) all the way to (B.12), and obtain the same result.

For the case in which $I_M(j_0')$ is not adjacent to $I_M(j_0)$, we may obtain a chain such that

$$w(I_M(j_0)) < w(I_M(j_0+1)) < \cdots < w(I_M(j_0'-1)) < w(I_M(j_0')), \quad \text{(B.13)}$$

and we again have $\alpha_1 < \bar{\alpha}$ where $\bar{\alpha}$ is defined in (B.11) (actually, α_1 in this case would be much smaller than $\bar{\alpha}$ since the value of $\mathscr{L}(\alpha,w)(I_M(j_0))$ will reach the value of $w(I_M(j_0+1))$ early than $w(I_M(j_0')))$. Moreover, for j such that $j_0+1 \leq j \leq j_0'-1$, we have $I_M(j)$ are step regions and we don't need to worry about them.

Note that in the above argument, we only used the strength of $I_M(j_0')$ but not yet $I_M(j_0'')$. Moreover, we can deal with situation 2 similarly by choosing j_0' to be the extreme region but not a boundary region since, according to Theorem 4.5, a boundary region only moves with half speed compare with extreme region.

Now we deal with situation 3. First we assume that $j_0' = 1$, $j_0 = 2$, and $j_0'' = 3$. That is, $M = 3$. We also assume that

$$(|I_3(1)| + |I_3(2)|/2)(w(I_3(1)) - w(I_3(2))) \leq (|I_3(3)| + |I_3(2)|/2)(w(I_3(3)) - w(I_3(2))). \tag{B.14}$$

According to Theorem 4.5, we have

$$\frac{1}{2}\frac{d}{d\alpha}\int_{I_3(1)} (\mathscr{L}(\alpha,w))^2\,dx = \frac{d}{d\alpha}\left[|I_3(1)|\left(w(I_3(1)) - \frac{\alpha}{|I_3(1)|}\right)^2\right] = \left(\frac{\alpha}{|I_3(1)|} - w(I_3(1))\right),$$

and

$$\frac{1}{2}\frac{d}{d\alpha}\int_{I_3(3)} (\mathscr{L}(\alpha,w))^2\,dx = \frac{d}{d\alpha}\left[|I_3(3)|\left(w(I_3(3)) - \frac{\alpha}{|I_3(3)|}\right)^2\right] = \left(\frac{\alpha}{|I_3(3)|} - w(I_3(3))\right).$$

Let $\bar{\alpha}'$ be such that

$$w(I_3(1)) - \frac{\bar{\alpha}'}{|I_3(1)|} = w(I_3(2)) + \frac{2\bar{\alpha}'}{|I_3(2)|}, \tag{B.15}$$

and we observe that for $0 \leq \alpha < \bar{\alpha}$,

$$\begin{aligned}
&\frac{1}{2}\frac{d}{d\alpha}\int_{I_3(1)} (\mathscr{L}(\alpha,w))^2\,dx + \frac{1}{2}\frac{d}{d\alpha}\int_{I_3(2)} (\mathscr{L}(\alpha,w))^2\,dx + \frac{1}{2}\frac{d}{d\alpha}\int_{I_3(3)} (\mathscr{L}(\alpha,w))^2\,dx \\
&= \left(\frac{\alpha}{|I_3(1)|} - w(I_3(1))\right) + \left(\frac{\alpha}{|I_3(3)|} - w(I_3(3))\right) + 2\left(w(I_3(2)) + \frac{2\alpha}{|I_3(2)|}\right) \\
&\leq 2\left(\frac{\alpha}{|I_3(1)|} - w(I_3(1))\right) + 2\left(w(I_3(2)) + \frac{2\alpha}{|I_3(2)|}\right) < 0,
\end{aligned}$$

where on the first inequality we used (B.14), and (B.15) in the last inequality. In the general case that $M > 0$, we may obtain a chain as in (B.13), and the same result holds since step regions do not count. Moreover, we have $\bar{\alpha}' \geq \alpha_1$ according to (B.15).

In the general case, where there is a collection

$$\mathcal{S} = \{I_M(j) \in \mathcal{A},\ I_M(j) \text{ is low extreme region and } w(I_M(j)) > 0,\ 1 \leq j \leq m\}$$

such that $\#(\mathcal{S}) > 1$, there must exists a collection $\mathcal{H}$ of regions $I_M(j)$ such that $\#(\mathcal{H}) \geq \#(\mathcal{S}) + 1$ and for each $I_M(j) \in \mathcal{S}$, there exist $I_M(j')$ and $I_M(j'') \in H$ such that one of situations above is satisfied. Therefore, since $\mathscr{B}(\alpha)$ is a finite summation over each region $I_M(j)$, we conclude that $\mathscr{B}'(\alpha) < 0$ for $0 \leq \alpha < \alpha_1$ and we finish Case 2.

Now we deal with what we left below (B.7). Let

$$\alpha' := \frac{1}{2}\min\{|I_M(j)|\,|w(I_M(j))|,\ j \in A\},$$

and define $w' = \mathscr{L}(\alpha', w)$. Then we may treat$\mathscr{L}(\alpha, w)$ for $\alpha > \alpha'$ by looking at $\mathscr{L}(\alpha - \alpha', w')$ and applying Case 2 above.

Hence, we have shown that $\mathscr{B}'(\alpha) < 0$ for all $0 \le \alpha < \alpha_1$. To show that $\mathscr{B}'(\alpha) < 0$ for $\alpha_1 \le \alpha < \alpha_2$, we set $w_1 := \mathscr{L}(\alpha_1, w)$ and apply the same argument above to $0 \le \alpha < \alpha_2 - \alpha_1$ on w_1 to obtain that

$$\frac{d}{d\alpha}\left(\frac{1}{2}\left\|\mathscr{L}(w_1,\alpha)\right\|^2_{L^2(I)}\right) < 0,\ 0 \le \alpha < \alpha_2 - \alpha_1$$

and this yield

$$\mathscr{B}'(\alpha) = \frac{d}{d\alpha}\left(\frac{1}{2}\left\|\mathscr{L}(w,\alpha)\right\|^2_{L^2(I)}\right) < 0,\ \alpha_1 \le \alpha < \alpha_2.$$

Note that $(w_1)_I = 0$ since, if not,

$$\lim_{\alpha\to\infty} \mathscr{L}(\alpha, w) = \lim_{\alpha\to\infty} \mathscr{L}(\alpha - \alpha_1, w_1) = (w_1)_I \ne 0,$$

contradicting (B.6).

Since $\mathscr{B}(\alpha)$ is continuous according to Lemma 4.15, we conclude that $\mathscr{B}(\alpha)$ is strictly decreasing for $0 \le \alpha < \alpha_M$ and $\mathscr{B}(\alpha_M) = 0$ since $(w)_I = \mathscr{L}(\alpha_M, w)$, by Lemma 4.15 again. □

Proposition B.3. *Let $v \in BV(I)$ be a monotone function and recall v_N from Notation 4.6. Then the following statements hold:*

1. $(v_N)_I = (v)_I$*;*

2.

$$\alpha_s(v_N) = \frac{1}{2}\int_I |v_N(x) - (v_N)_I|\ dx; \tag{B.16}$$

3.

$$\int_I |v_N(x) - \mathscr{L}(\alpha, v_N)(x)|\, dx = 2\alpha,\ \textit{for } 0 \le \alpha \le \frac{1}{2}\int_I |v_N(x) - (v_N)_I|\, dx. \tag{B.17}$$

Proof. Without loss of generality, we assume that v is monotone increasing. Let $N > 0$ be fixed and define v_N as in (1.24). Then, by definition of v_N we have $(v_N)_I = (v)_I$.

We now prove (B.16) and (B.17). First we assume that there exists a region $I_N(m_N)$, $1 \le m_N \le N$, such that

$$(v_N)_I = v_N(I_N(m_N)). \tag{B.18}$$

Define the subintervals (see Figure 11a)

$$I_l := \bigcup_{1\le k<m_N} I_N(k) \text{ and } I_h := \bigcup_{m_N<k\le N} I_N(k),$$

and we first focus on the subinterval I_l.

Since v_N is monotone increasing, we have

$$I_N(k) \in \mathcal{C}_S(v_N),\ 1 < k < m_N,$$

and for

$$0 \leq N\alpha < v_N(I_N(2)) - v_N(I_N(1)),$$

we obtain, according to Theorem 4.5,

$$\begin{aligned} &\mathscr{L}(\alpha, v_N)(I_N(1)) = v(I_N(1)) + N\alpha, \text{ and} \\ &\mathscr{L}(\alpha, v_N)(I_N(k)) = v(I_N(k)),\ 1 < k \leq m_N. \end{aligned} \tag{B.19}$$

There,

$$\alpha_1 := (v_N(I_N(2)) - v_N(I_N(1)))/N, \tag{B.20}$$

and

$$\mathscr{L}(\alpha_1, v_N)(I_N(1)) = v_N(I_N(2)).$$

Moreover, from (B.19) we have, for $0 \leq \alpha \leq \alpha_1$,

$$\int_{I_l} |v_N(x) - \mathscr{L}(\alpha, v_N)(x)|\, dx = \int_{I_l} (\mathscr{L}(\alpha, v_N)(x) - v_N(x))\, dx = \frac{1}{N}(N\alpha) = \alpha. \tag{B.21}$$

Next, set $v_N^1 := \mathscr{L}(\alpha_1, v_N)$ and observe that, for

$$0 \leq \alpha < \alpha_2' := \frac{N}{2}\left(v_N(I_N(3)) - v_N(I_N(2))\right),$$

we have

$$\begin{aligned} &\mathscr{L}(\alpha, v_N^1)(I_N(1)) = \mathscr{L}(\alpha, v_N^1)(I_N(2)) = v(I_N(2)) + \frac{N}{2}, \text{ and} \\ &\mathscr{L}(\alpha, v_N^1)(I_N(k)) = v(I_N(k)),\ 2 < k \leq m_N. \end{aligned} \tag{B.22}$$

Hence for $\alpha = \alpha_2'$ we obtain

$$\mathscr{L}(\alpha_2', v_N')(I_N(1)) = \mathscr{L}(\alpha_2', v_N')(I_N(2)) = v_N(I_N(3)),$$

and for

$$\alpha_2 := \alpha_1 + \alpha_2' = \frac{1}{N}(v_N(I_N(2)) - v_N(I_N(1))) + \frac{2}{N}\left(v_N(I_N(3)) - v_N(I_N(2))\right), \tag{B.23}$$

we have that

$$\mathscr{L}(\alpha_2, v_N)(I_N(1)) = \mathscr{L}(\alpha_2, v_N)(I_N(2)) = v_N(I_N(3)).$$

Moreover, we observe that for $\alpha_1 \leq \alpha \leq \alpha_2$,

$$\begin{aligned} &\int_{I_l} (\mathscr{L}(\alpha, v_N) - v_N)\, dx = \int_{I_l} (\mathscr{L}(\alpha_1, v_N) - v_N)\, dx + \int_{I_l} (\mathscr{L}(\alpha, v_N) - \mathscr{L}(\alpha_1, v_N))\, dx \\ &= \int_{I_N(1)} (v_N(I_N(2)) - v_N(I_N(1))) + \int_{I_N(1)\cup I_N(2)} (\mathscr{L}(\alpha - \alpha_1, \mathscr{L}(\alpha_1, v_N)) - \mathscr{L}(\alpha_1, v_N))\, dx \end{aligned}$$

$$=\alpha_1 + \frac{2}{N}\frac{N}{2}(\alpha - \alpha_1) = \alpha,$$

where on the last equality we used (B.21) and (B.22).

Similarly, for

$$\alpha_3' = \frac{3}{N}\left(v_N(I_N(4)) - v_N(I_N(3))\right),$$

we have that

$$\mathscr{L}(\alpha_2+\alpha_3', v_N)(I_N(1)) = \mathscr{L}(\alpha_2+\alpha_3', v_N)(I_N(2)) = \mathscr{L}(\alpha_2+\alpha_3', v_N).(I_N(3)) = v_N(I_N(4))$$

Recursively, we obtain

$$\alpha'_{m_N-1} = \frac{m_N - 1}{N}\left(v_N(I_N(m_N)) - v_N(I_N(m_N-1))\right), \tag{B.24}$$

and at $\alpha = \alpha_{m_N-1}$, where

$$\begin{aligned}\alpha_{m_N-1} := \sum_{k=1}^{m_N-1} \alpha_k' &= \sum_{k=1}^{m_N-1} \frac{k}{N}\left(v_N(I_N(k+1)) - v_N(I_N(k))\right) \qquad \text{(B.25)}\\ &= \sum_{k=1}^{m_N-1} \frac{1}{N}\left(v_N(m_N) - v_N(I_N(k))\right),\end{aligned}$$

it holds

$$\begin{aligned}&\mathscr{L}(\alpha_{m_N-1}, v_N)(I_N(1)) = \mathscr{L}(\alpha_{m_N-1}, v_N)(I_N(2)) = \\ &\cdots = \mathscr{L}(\alpha_{m_N-1}, v_N)(I_N(m_N-1)) = \mathscr{L}(\alpha_{m_N-1}, v_N)(I_N(m_N)) = v_N(I_N(m_N)).\end{aligned} \tag{B.26}$$

Moreover, by using a similar computation as above we deduce that

$$\int_{I_l} \left(\mathscr{L}(\alpha, v_N) - v_N\right) dx = \alpha \text{ for } 0 \le \alpha \le \alpha_{m_N-1} \tag{B.27}$$

and

$$\int_{I_l} \left(\mathscr{L}(\alpha_{m_N-1}, v_N)(x) - v_N(x)\right) dx = \int_{I_l} \left((v_N)_I - v_N(x)\right) dx.$$

Next, we claim that

$$\begin{aligned}&\mathscr{L}(\alpha_{m_N-1}, v_N)(I_N(N)) = \mathscr{L}(\alpha_{m_N-1}, v_N)(I_N(N-1)) = \\ &\qquad \cdots = \mathscr{L}(\alpha_{m_N-1}, v_N)(I_N(m_N+1)) = \mathscr{L}(\alpha_{m_N-1}, v_N)(I_N(m_N)) = v_N(I_N(m_N))\end{aligned}$$

and

$$\int_{I_h} \left(v_N - \mathscr{L}(\alpha, v_N)\right) = \alpha. \tag{B.28}$$

Indeed, in order to reach a contradiction we assume that $\mathscr{L}(\alpha_{m_N-1}, v_N)(I_N(m_N)) > v_N(I_N(m_N))$. In view of (B.26), we have

$$(\mathscr{L}(\alpha_{m_N-1}, v_N))_I > \mathscr{L}(\alpha_{m_N-1}, v_N)(I_N(m_N)) = v_N(I_N(m_N))$$

and hence by (B.18), Proposition 4.12, and Theorem 4.9, we obtain

$$\begin{aligned}(v_N)_I &= v_N(I_N(m_N)) = \lim_{\alpha\to\infty} \mathscr{L}(\alpha, v_N)\\ &= \lim_{\alpha\to\infty} \mathscr{L}(\alpha, \mathscr{L}(\alpha_{m_N-1}, v_N)) = (\mathscr{L}(\alpha_{m_N-1}, v_N))_I > v_N(I_N(m_N)) = (v_N)_I, \end{aligned} \tag{B.29}$$

a contradiction. Moreover, using the same argument as in (B.19) and (B.21), we may deduce (B.28).

That is, we have that $\mathscr{L}(\alpha_{m_N-1}, v_N) = v_N(I_N(m_N))$ is a constant, and we conclude that

$$\alpha_s(v_N) = \int_{I_l} (v_N - (v_N)_I)\,dx = \frac{1}{2}\int_I |v_N - (v_N)_I|\,dx, \tag{B.30}$$

where $\alpha_s(v_N) := \alpha_{m_N-1}$. The behavior of $\mathscr{L}(\alpha, v_N)$ in subinterval the I_h exactly mirrors its behavior in I_l. That is, $\mathscr{L}(\alpha, v_N)$ behaves "symmetrically" with respect to the average value $(v)_I$. See Figure 11a to Figure 11c for an illustration. Moreover, we remark that the set $\{\alpha_1, \dots, \alpha_{m_N-1}\}$ are only a subset of (4.14) in Lemma 4.15. We refer to Remark B.5 for details.
For the general case that there is no index m_N such that (B.18) holds, but an index m_N such that

$$v_N(I_N(m_N)) < (v_N)_I < v_N(I_N(m_N+1)),$$

we only need one more step to obtain (B.30).

Indeed, using the same argument until (B.24), we have that there exist $\alpha_l = \alpha_h > 0$ such that (see Figure 11d)

$$\mathscr{L}(\alpha_l, v_N)(I_N(1)) = \cdots = \mathscr{L}(\alpha_l, v_N)(I_N(m_N)) < (v_N)_I$$

and

$$\mathscr{L}(\alpha_h, v_N)(I_N(N)) = \cdots = \mathscr{L}(\alpha_h, v_N)(I_N(m_N+1)) > (v_N)_I.$$

Also, we have

$$\mathscr{L}(\alpha_l+\alpha, v_N)(I_N(1)) = \cdots = \mathscr{L}(\alpha_l+\alpha, v_N)(I_N(m_N)) = v_N(I_N(m_N)) + \frac{N}{m_N}\alpha$$

for

$$0 \le \alpha \le \alpha'_{m_N} := \frac{m_N}{N}\left((v_N)_I - v_N(I_N(m_N))\right).$$

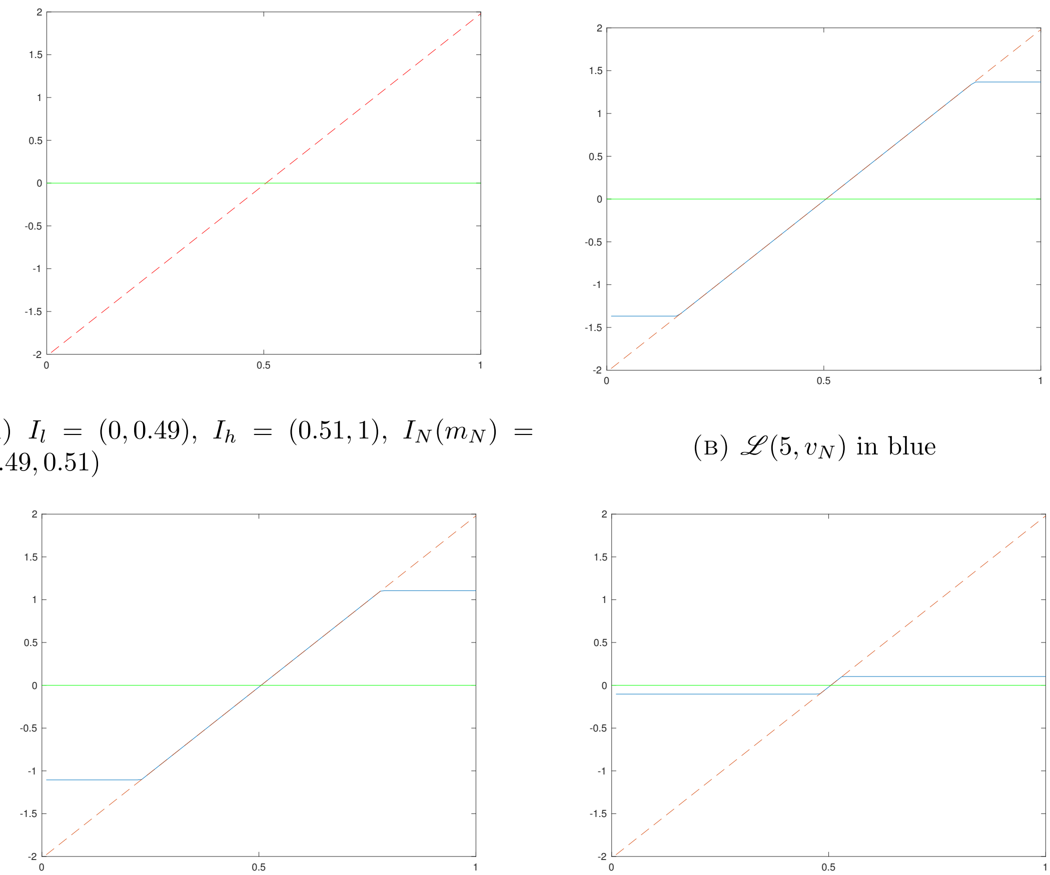
Hence,

$$\mathscr{L}(\alpha_{m_N}, v_N)(I_N(k)) = (v_N)_I \text{ for } 1 \le k \le m_N,$$

where $\alpha_{m_N} := \alpha_l + \alpha'_{m_N}$, and again we have (B.30) with $\alpha_s(v_N) := \alpha_{m_N}$.

Note that we also obtain

$$\mathscr{L}(\alpha_{m_N}, v_N)(I_N(k)) = (v_N)_I \text{ for } m_N+1 \le k \le N,$$



(A) $I_l = (0, 0.49)$, $I_h = (0.51, 1)$, $I_N(m_N) = (0.49, 0.51)$

(B) $\mathscr{L}(5, v_N)$ in blue

(C) $\mathscr{L}(10, v_N)$ in blue

(D) $\mathscr{L}(50, v_N)$ in blue

FIGURE 11. $N = 100$. v_N is plotted in red dashed line, and the constant $(v)_I$ is plotted in green. Moreover, in Figure 11b to 11d, α equals to the area of the triangles formed by the blue line and the red line (left lower corner and right upper corner)

since $\mathscr{L}(\alpha_s(v_N), v_N) = (v_N)_I$ and use the same contradiction argument as in (B.29).

Finally, in view of (B.27) and (B.28), we observe that

$$\begin{aligned}\alpha &= \int_{I_l} (\mathscr{L}(\alpha, v_N) - v_N)\, dx = \int_{I_h} (v_N - \mathscr{L}(\alpha, v_N))\, dx \\ &= \frac{1}{2}\int_I |v_N - \mathscr{L}(\alpha, v_N)|\,, \text{ for } 0 \le \alpha < \frac{1}{2}\int_I |v_N - (v_N)_I|\, dx,\end{aligned}$$

and this conclude the proof. □

Remark B.4. By (B.16), and invoking Lebesgue Dominated convergence theorem, we have that

$$\lim_{N\to\infty}\alpha_s(v_N)=\lim_{N\to\infty}\frac{1}{2}\int_I|v_N-(v_N)_I|\,dx=\frac{1}{2}\int_I|(v)_I-v(t)|\,dt,$$

and thus $\alpha_s(v_N)$ is bounded and convergence.

Remark B.5. Since v_N is a piecewise constant function with N pieces, we may apply Lemma 4.15 to get a chain

$$0=\alpha_0<\alpha_1<\cdots<\alpha_{N'}<+\infty$$

such that items 1-3 in Lemma 4.15 are satisfied, where $N'\le N$. We use

$$0=\alpha_0<\alpha_1^l<\alpha_2^l<\cdots<\alpha_{m_N-1}^l<+\infty$$

to denote those α's we found in (B.20), (B.23), and (B.25). Then,

$$\left\{\alpha_i^l,\ 1\le i\le m_N-1\right\}\subset\left\{\alpha_i,\ 1\le i\le N'\right\}$$

since at each α_i^l we have $\mathscr{L}(\alpha_i^l,v_N)(I_N(i))=v_N(I_N(i+1))$, and hence $\mathscr{L}(\alpha,v_N)$ loses one piece in I_l. Moreover, we may repeat the argument from (B.18) to (B.25) on the subinterval I_h and obtain

$$0<\alpha_N^h<\alpha_{N-1}^l<\cdots<\alpha_{m_N+1}^h<+\infty$$

such that $\mathscr{L}(\alpha_k^h,v_N)(I_N(k))=v_N(I_N(k-1))$ at each α_k^h, and hence $\mathscr{L}(\alpha,v_N)$ loses one piece in I_h, and this

$$\left\{\alpha_k^h,\ N\ge k\ge m_N+1\right\}\subset\left\{\alpha_i,\ 1\le i\le N'\right\}.$$

Moreover, we have

$$\left\{\alpha_i^l,\ 1\le i\le m_N-1\right\}\cup\left\{\alpha_k^h,\ N\ge k\ge m_N+1\right\}=\left\{\alpha_i,\ 1\le i\le N'\right\}.$$

Proposition B.6. *Let $v\in BV(I)$ be a monotone function and define*

$$\mathcal{E}_{v_N}(\alpha):=\frac{1}{2}\int_I|\mathscr{L}(\alpha,v_N)-v_N|^2\,dx.$$

Then $\mathcal{E}'_{v_N}(0)=0$, and $\mathcal{E}'_{v_N}(\alpha)$ is piecewise linear and increasing in each linear piece. Moreover, if $\tilde\alpha\in J_{\mathcal{E}'_{v_N}(\alpha)}$, then

$$\mathcal{E}'_{v_N}(\tilde\alpha^-)>\mathcal{E}'_{v_N}(\tilde\alpha^+).$$

Proof. Following the same argument as in Proposition B.3, we have for

$$0\le\alpha<\alpha_1:=\frac{1}{N}\left(v_N(I_N(2))-v_N(I_N(1))\right),$$

that

$$\mathscr{L}(\alpha,v_N)=v_N(I_N(1))+N\alpha\text{ and }\mathcal{E}_{v_N}(\alpha)=\frac{1}{2}\int_I|\mathscr{L}(\alpha,u_c)-v_N|^2\,dx=\frac{1}{2}\frac{1}{N}(N\alpha)^2.$$

Therefore, for $0 \le \alpha < \alpha_1$, we have $\mathcal{E}'_{v_N}(\alpha) = N\alpha$, and hence

$$\mathcal{E}'_{v_N}\left(\alpha_1^-\right) = \lim_{\alpha \nearrow \alpha_1} \mathcal{E}'_{v_N}(\alpha) = \left(v_N(I_N(2)) - v_N(I_N(1))\right). \tag{B.31}$$

However, at $\alpha = \alpha_1$, $\mathscr{L}(\alpha_1, v_N)(I_N(1)) = v_N(I_N(2))$, and thus for

$$\alpha_1 \le \alpha < \alpha_2 := \alpha_1 + \frac{2}{N}\left(v_N(I_N(3)) - v_N(I_N(2))\right)$$

we obtain

$$\mathcal{E}_{v_N}(\alpha) = \frac{1}{2}\left[\frac{1}{N}\left(\frac{N}{2}(\alpha-\alpha_1) + v_N(I_N(2)) - v_N(I_N(1))\right)^2 + \frac{1}{N}\left(\frac{N}{2}(\alpha-\alpha_1)\right)^2\right],$$

therefore

$$\mathcal{E}'_{v_N}(\alpha) = \frac{N}{2}(\alpha-\alpha_1) + \frac{1}{2}\left(v_N(I_N(2)) - v_N(I_N(1))\right).$$

We deduce that

$$\mathcal{E}'_{v_N}\left(\alpha_1^+\right) = \lim_{\alpha \searrow \alpha_1} \mathcal{E}'_{v_N}(\alpha) = \frac{1}{2}\left(v_N(I_N(2)) - v_N(I_N(1))\right) \tag{B.32}$$

and

$$\mathcal{E}'_{v_N}\left(\alpha_2^-\right) = \lim_{\alpha \nearrow \alpha_2} \mathcal{E}'_{v_N}(\alpha) = \frac{1}{2}\left(v_N(I_N(2)) - v_N(I_N(1))\right) + \left(v_N(I_N(3)) - v_N(I_N(2))\right).$$

Following a similar computation, we have

$$\begin{aligned}
\mathcal{E}'_{v_N}\left(\alpha_2^-\right) &= \frac{1}{2}\left(v_N(I_N(2)) - v_N(I_N(1))\right) + \left(v_N(I_N(3)) - v_N(I_N(2))\right),\\
\mathcal{E}'_{v_N}\left(\alpha_2^+\right) &= \frac{1}{3}\left(v_N(I_N(3)) - v_N(I_N(1))\right) + \frac{1}{3}\left(v_N(I_N(3)) - v_N(I_N(2))\right),
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{E}'_{v_N}\left(\alpha_3^-\right)\\
&= \frac{1}{3}\left(v_N(I_N(3)) - v_N(I_N(1))\right) + \frac{1}{3}\left(v_N(I_N(3)) - v_N(I_N(2))\right) + \left(v_N(I_N(4)) - v_N(I_N(3))\right),\\
&\mathcal{E}'_{v_N}\left(\alpha_3^+\right)\\
&= \frac{1}{4}\left(v_N(I_N(4)) - v_N(I_N(1))\right) + \frac{1}{4}\left(v_N(I_N(4)) - v_N(I_N(2))\right) + \frac{1}{4}\left(v_N(I_N(4)) - v_N(I_N(3))\right),
\end{aligned}$$

$$\vdots$$

and

$$\mathcal{E}'_{v_N}\left(\alpha_l^-\right) = \frac{1}{l}\left[\sum_{k=1}^{l-1} v_N(I_N(l)) - v_N(I_N(k))\right] + v_N(I_N(l+1)) - v_N(I_N(l)),$$

$$\mathcal{E}'_{v_N}\left(\alpha_l^+\right) = \frac{1}{l+1}\sum_{k=1}^{l}\left(v_N(I_N(l+1)) - v_N(I_N(k))\right),$$

for any l such that $1 \leq l < m_N$. That is, we conclude that

$$\left|J_{\mathcal{E}'_{v_N}(\alpha)}\right| \leq N$$

and for each $\tilde{\alpha} \in J_{\mathcal{E}'_{v_N}(\alpha)}$,

$$\mathcal{E}'_{v_N}\left(\tilde{\alpha}^-\right) > \mathcal{E}'_{v_N}\left(\tilde{\alpha}^+\right).$$

□

B.2. *The uniqueness of solution of bilevel training scheme.* In this section we prove Theorem 4.13. To do so, we introduce several definitions, assumptions, and propositions first.

Definition B.7. *For a given clean image $u_c \in BV$, we define the* deformation error *by*

$$\mathcal{E}_{u_c}(\alpha) := \frac{1}{2}\left\|\mathscr{L}(\alpha, u_c) - u_c\right\|^2_{L^2(I)},$$

and the level N deformation error *by*

$$\mathcal{E}_{\mathscr{C}_N}(\alpha) := \frac{1}{2}\left\|\mathscr{L}(\alpha, u_{c,N}) - u_{c,N}\right\|^2_{L^2(I)}.$$

Moreover, we define the denoising error *by*

$$\mathcal{E}_{\eta_N}(\alpha) := \frac{1}{2}\left\|\mathscr{L}(\alpha, \eta_N) - \fint_I \eta_N\right\|^2_{L^2(I)}. \tag{B.33}$$

Note that by Proposition B.1 we have that $\mathcal{E}_{u_c}(\alpha)$ and $\mathcal{E}_{\mathscr{C},N}(\alpha)$ are monotone increasing and, according to Proposition B.3, we have that $\mathcal{E}_{u_c}(\alpha)$ ($\mathcal{E}_{\mathscr{C},N}(\alpha)$) is a constant for all $\alpha \geq \alpha_s(u_c)$, with

$$\mathcal{E}_{u_c}(\alpha_s(u_c)) = \frac{1}{2}\left\|\fint_I u_c\,dx - u_c\right\|^2_{L^2(I)}.$$

Moreover, we have $\mathcal{E}_{\eta_N}(\alpha)$ is strictly decreasing to 0 by Proposition B.2 with $w = \eta_N$.

Assumption B.8. *Let $u_c \in BV(I)$ and $N \in \mathbb{N}$ be given. We say that an image $u_{d,N}$ with resolution N is an* acceptable compressed deformation *of $u_{c,N}$ if the following conditions are satisfied:*

1. *if $I_N(k)$ is a high (low) boundary (extreme) region of $u_{c,N}$, then it is also a high (low) boundary (extreme) region of $u_{d,N}$ and*

$$u_{c,N}(I_N(k)) \geq (\leq) u_{d,N}(I_N(k));$$

2. *if $I_N(k)$ is a step region of $u_{c,N}$, it is also a step region of $u_{d,N}$ and we have*

$$\begin{aligned}\frac{1}{2}\left(u_{c,N}(I_N(k-1)) + u_{c,N}(I_N(k))\right) &\leq u_{d,N}(I_N(k)) \\ &\leq \frac{1}{2}\left(u_{c,N}(I_N(k)) + u_{c,N}(I_N(k+1))\right).\end{aligned} \tag{B.34}$$

Proposition B.9. *Let $u_{c,N}$ be monotone and $u_{d,N}$ satisfying Assumption B.8. Then*

$$\mathcal{E}_{u_{d,N}}(\alpha) := \frac{1}{2}\left\|\mathscr{L}(\alpha, u_{d,N}) - u_{c,N}\right\|_{L^2}^2$$

is non-decreasing as $\alpha \to \infty$.

Proof. Without loss of generality, we assume that $u_{c,N}$ is monotone increasing. That is,

$$u_{c,N}(I_N(1)) < u_{c,N}(I_N(2)) < \cdots u_{c,N}(I_N(N-1)) < u_{c,N}(I_N(N)),$$

and hence by Assumption B.8 we have

$$\begin{aligned} u_{c,N}(I_N(1)) &\le u_{d,N}(I_N(1)) \le u_{d,N}(I_N(2)) \le \cdots \\ &\le u_{d,N}(I_N(N-1)) \le u_{d,N}(I_N(N)) \le u_{c,N}(I_N(N)). \end{aligned}$$

Therefore, in the same spirit of the argument used in the proof of Proposition B.3, we have for

$$0 \le \alpha < \alpha_1 := \frac{1}{N}\left[u_{d,N}(I_N(2)) - u_{d,N}(I_N(1))\right],$$

that

$$\mathscr{L}(\alpha, u_{d,N})(I_N(1)) = \mathscr{D}_N(I_N(1)) + N\alpha,$$

and

$$\begin{aligned} \mathcal{E}_{u_{d,N}}(\alpha) &= \frac{1}{2}\left\|\mathscr{L}(\alpha, \mathscr{D}_N) - u_{c,N}\right\|_{L^2(I)}^2 \\ &= \frac{1}{2}\frac{1}{|I_N(1)|}\left[\mathscr{D}_N(I_N(1)) + N\alpha - u_{c,N}(I_N(1))\right]^2; \end{aligned} \tag{B.35}$$

and for

$$\alpha_1 \le \alpha < \alpha_2 := \frac{2}{N}\left[u_{d,N}(I_N(3)) - u_{d,N}(I_N(2))\right] + \alpha_1,$$

that

$$\mathscr{L}(\alpha, u_{d,N})(I_N(1)) = \mathscr{L}(\alpha, u_{d,N})(I_N(2)) = \mathscr{D}_N(I_N(2)) + \frac{N}{2}(\alpha - \alpha_1),$$

and

$$\begin{aligned} \mathcal{E}_{u_{d,N}}(\alpha) &= \frac{1}{2}\frac{1}{N}\left[\mathscr{D}_N(I_N(2)) + \frac{N}{2}(\alpha - \alpha_1) - u_{c,N}(I_N(1))\right]^2 \\ &+ \frac{1}{2}\frac{1}{N}\left[\mathscr{D}_N(I_N(2)) + \frac{N}{2}(\alpha - \alpha_1) - u_{c,N}(I_N(2))\right]^2. \end{aligned} \tag{B.36}$$

We note that (B.35) is increasing for $\alpha \in [0, \alpha_1)$ since $\mathscr{D}_N(I_N(1)) \ge u_{c,N}(I_N(1))$ by assumption. We discuss the monotonicity of (B.36) in two cases:

<u>Case 1:</u> if

$$u_{c,N}(I_N(1)) \le u_{d,N}(I_N(1)) \le u_{c,N}(I_N(2)) \le u_{d,N}(I_N(2)),$$

then both (B.35) and (B.36) are increasing, and so $\mathcal{E}_{\mathscr{D}_N}(\alpha)$ is increasing for $0 \le \alpha < \alpha_2$.

Case 2: if

$$u_{c,N}(I_N(1)) \le u_{d,N}(I_N(1)) \le u_{d,N}(I_N(2)) \le u_{c,N}(I_N(2)),$$

then the second term in (B.36) is decreasing, and hence $\mathcal{E}_{\mathscr{D}_N}(\alpha)$ might decrease. However, we show that with condition (B.34) this will not happen. Indeed, in view of (B.34) we have

$$u_{d,N}(I_N(2)) - u_{c,N}(I_N(2)) \ge u_{c,N}(I_N(1)) - u_{d,N}(I_N(2)),$$

and thus, by (B.36), that for $\alpha_1 \le \alpha < \alpha_2$,

$$\begin{aligned}
&\mathcal{E}'_{u_{d,N}}(\alpha)\\
&= \frac{1}{2}\left[u_{d,N}(I_N(2)) - u_{c,N}(I_N(1))\right] + \frac{1}{2}\left[u_{d,N}(I_N(2)) - u_{c,N}(I_N(2))\right] + \frac{N}{2}(\alpha - \alpha_1)\\
&\ge \frac{1}{2}\left[u_{d,N}(I_N(2)) - u_{c,N}(I_N(1)) + u_{c,N}(I_N(1)) - u_{d,N}(I_N(2))\right] + \frac{N}{2}(\alpha - \alpha_1)\\
&\ge \frac{N}{2}(\alpha - \alpha_1) \ge 0.
\end{aligned}$$

For $\alpha_2 \le \alpha < \alpha_3$, $\alpha_3 \le \alpha < \alpha_4, \ldots$, we may prove that $\mathcal{E}'_{u_{d,N}}(\alpha)$ is non-decreasing by adopting the same computation, and this concludes the proof. □

So far we have not assume any relation between the clean image $u_{c,N}$ and the noise η_N. However, Theorem 4.13 does not hold for arbitrary η_N, and the following assumptions have to be enforced.

Assumption B.10. *Let $u_c \in BV(I)$ be monotone and let $N \in \mathbb{N}$ be given. Here $1 \le k \le N$.*

1. *The observed noise changes sign consecutively, that is,*

$$\eta_N(I_N(k))\eta_N(I_N(k+1)) \le 0; \tag{B.37}$$

2. *$u_{\eta,N}$ is oscillating at least at half rate of $u_{c,N}$, that is, we require that*

$$|u_{\eta,N}(I_N(k)) - u_{\eta,N}(I_N(k+1))| \ge \frac{1}{2}\,|u_{c,N}(I_N(k)) - u_{c,N}(I_N(k+1))|\,; \tag{B.38}$$

3. *if $u_{\eta,N}$ changes the sign of jump of $u_{c,N}$, that is, if*

$$\left(u_{\eta,N}(I_N(k)) - u_{\eta,N}(I_N(k+1))\right)\left(u_{c,N}(I_N(k)) - u_{c,N}(I_N(k+1))\right) \le 0,$$

then we require that

$$\eta_N(I_N(k)) = -\eta_N(I_N(k+1)). \tag{B.39}$$

4. *we assume that*

$$u_{\eta,N}(I_N(1)) > (<) u_{c,N}(I_N(1))$$

if $u_{c,N}(I_N(1)) > (<) u_{c,N}(I_N(2))$, and

$$u_{\eta,N}(I_N(N)) > (<) u_{c,N}(I_N(N))$$

if $u_{c,N}(I_N(N)) > (<) u_{c,N}(I_N(N-1))$. *Lastly, we assume that*

$$|\eta(I_N(1))| = |\eta(I_N(N))| \geq \frac{1}{2} \max \{|\eta(I_N(k_I))|,\ k_I \in \mathcal{C}_I(u_{c,N})\}. \tag{B.40}$$

Proof of Theorem 4.13. We denote the local error by

$$\mathcal{E}_{k,N}(\alpha) := \frac{1}{2} \|\mathscr{L}(\alpha, u_{\eta,N}) - u_{c,N}\|^2_{L^2(I_N(k))} = \frac{1}{2N} |\mathscr{L}(\alpha, u_{\eta,N})(I_N(k)) - u_{c,N}(I_N(k))|^2,$$

where $1 \leq k \leq N$. We write that

$$\mathcal{E}'_N(\alpha) = \sum_{k=1}^{N} \mathcal{E}'_{k,N}(\alpha) = \sum_{I_N(k)\in\mathcal{P}(\alpha)} \mathcal{E}'_{k,N}(\alpha) + \sum_{I_N(k)\in\mathcal{N}(\alpha)} \mathcal{E}'_{k,N}(\alpha) + \sum_{I_N(k)\in\mathcal{I}(\alpha)} \mathcal{E}'_{k,N}(\alpha),$$

where, for each $\alpha > 0$, we say that a region $I_N(k)$ is positive active of $\mathcal{E}_N(\alpha)$ if $\mathcal{E}'_{k,N}(\alpha) > 0$, negative active if $\mathcal{E}'_{k,N}(\alpha) < 0$, and inactive if $\mathcal{E}'_{k,N}(\alpha^+) = 0$. Moreover, we use $\mathcal{N}(\alpha)$, $\mathcal{P}(\alpha)$, and $\mathcal{I}(\alpha)$ to denote the collections of such regions for each $\alpha > 0$.

We will prove that

$$\alpha_m = \frac{1}{N} |\eta_N(1)| \tag{B.41}$$

is the desired minimizer of (4.7) by showing that

$$\mathcal{P}(\alpha) = \varnothing \text{ for all } \alpha \leq \alpha_m,\ \mathcal{I}(\alpha_m) = \{I_N(k),\ 1 \leq k \leq N\}, \text{ and } \mathcal{N}(\alpha) = \varnothing \text{ for all } \alpha \geq \alpha_m.$$

Items 1 and 3 in Assumption B.10 implying that if $I_N(k) \in \mathcal{P}(0) \cup \mathcal{N}(0)$, then it is an extreme region of η_N. We claim that

$$\mathcal{E}'_{k,N}(0) = -|\eta_N(I_N(k))| < 0.$$

Indeed, without loss of generality, we assume that $u_{c,N}$ is monotone increasing. Then $I_N(1) \in \mathcal{P}(0) \cup \mathcal{N}(0)$ is a low boundary region of $u_{\eta,N}$ by item 4 in Assumption B.10, and hence for α small enough

$$\mathcal{E}'_{1,N}(\alpha) = \frac{d}{d\alpha}\frac{1}{2}\frac{1}{N} |u_{\eta,N}(I_N(1)) - N\alpha - u_{c,N}(I_N(1))|^2 = \frac{d}{d\alpha}\frac{1}{2}\frac{1}{N} |\eta_N(I_N(1)) - N\alpha|^2, \tag{B.42}$$

that is,

$$\mathcal{E}'_{1,N}(0) = -|\eta_N(I_N(1))| < 0.$$

The case in which $I_N(N)$ is a high boundary region of $u_{c,N}$ and hence a high boundary of $u_{\eta,N}$ can be dealt similarly. Now we assume that $I_N(k) \in \mathcal{C}_I(u_{c,N})$ is a step region, that is, we have

$$u_{c,N}(I_N(k-1)) \leq u_{c,N}(I_N(k)) \leq u_{c,N}(I_N(k+1)).$$

We claim that if $I_N(k)$ is a high extreme region of $u_{\eta,N}$, then $u_{\eta,N}(I_N(k)) > u_{c,N}(I_N(k))$. Suppose not. We have

$$u_{c,N}(I_N(k)) > u_{\eta,N}(I_N(k)) > \max \{u_{\eta,N}(I_N(k-1)), u_{\eta,N}(I_N(k+1))\}. \tag{B.43}$$

However, $u_{c,N}(I_N(k)) > u_{\eta,N}(I_N(k))$ implies that $\eta_N(I_N(k)) < 0$, and thus by (B.37) we have $\eta_N(I_N(k+1)) > 0$. Therefore, we have that

$$u_{\eta,N}(I_N(k{+}1)) = u_{c,N}(I_N(k{+}1)){+}\eta_N(I_N(k{+}1)) > u_{c,N}(I_N(k{+}1)) \geq u_{c,N}(I_N(k)) > u_{\eta,N}(I_N(k)),$$

which contradicts to (B.43). Hence, by using the same computation as in (B.42), we deduce that

$$\mathcal{E}'_{k,N}(0) = -\left|\eta_N(I_N(k))\right| < 0. \tag{B.44}$$

The case in which $I_N(k)$ is a low extreme region of $u_{\eta,N}$, and hence $u_{\eta,N}(I_N(k)) < u_{c,N}(I_N(k))$ and (B.44) holds, can be proved similarly.

Therefore, at $\alpha = 0$, we have $\mathcal{P}(0) = \varnothing$ and

$$\mathcal{E}'_N(0) = -\sum_{I_N(k)\in\mathcal{N}(0)} \left|\eta_N(I_N(k))\right| < 0 \text{ and } \mathcal{E}'_N(0) \geq \mathcal{E}'_{\eta_N}(0) = -\sum_{k=1}^{N} \left|\eta_N(I_N(k))\right|,$$

where $\mathcal{E}_{\eta_N}$ is defined in (B.33) and $(\eta_N)_I = 0$ by (4.1).

We next claim that $\mathcal{P}(\alpha) = \varnothing$ if $\alpha < \alpha_m$ where α_m is defined in (B.41). Assume that $u_{c,N}(I_N(k+1))$ and $u_{c,N}(I_N(k+2))$ are two step regions of $u_{c,N}$, that is,

$$u_{c,N}(I_N(k)) \leq u_{c,N}(I_N(k+1)) \leq u_{c,N}(I_N(k+2)) \leq u_{c,N}(I_N(k+3)).$$

Without loss of generality, by (B.37) we assume that

$$\eta_N(I_N(k)) \leq 0,\ \eta_N(I_N(k+1)) \geq 0,\ \eta_N(I_N(k+2)) \leq 0, \text{ and } \eta_N(I_N(k+3)) \geq 0. \tag{B.45}$$

Two situations could hold within interval $I_N(k+1)$ and $I_N(k+2)$:

<u>Case 1:</u> if $u_{\eta,N}(I_N(k+1)) < u_{\eta,N}(I_N(k+2))$, then we have, by (B.45), that

$$u_{\eta,N}(I_N(k)) < u_{c,N}(I_N(k+1)) < u_{\eta,N}(I_N(k+1)) < u_{\eta,N}(I_N(k+2)) < u_{\eta,N}(I_N(k+3)),$$

and hence $u_{\eta,N}(I_N(k+1))$ and $u_{\eta,N}(I_N(k+2))$ are steps regions, which implies that

$$\mathcal{E}'_{k+1,N}(0) = 0 > \mathcal{E}'_{k+1,\eta_N}(0) \text{ and } \mathcal{E}'_{k+2,N}(0) = 0 > \mathcal{E}'_{k+2,\eta_N}(0).$$

That is, $I_N(k+1)$ and $I_N(k+1) \in \mathcal{I}(0)$ and the noise $\eta_N(I_N(k+1))$ and $\eta_N(I_N(k+2))$ remain un-removed. Moreover, according to the argument in the Proposition B.3, the region $I_N(k+1)$ remains in-active until $\alpha > 0$ large enough such that

$$\mathscr{L}(\alpha, u_{\eta,N})(I_N(1)) = \mathscr{L}(\alpha, u_{\eta,N})(I_N(k)) = u_{c,N}(I_N(k)).$$

However, according to (B.40) and (B.41), we have

$$\mathscr{L}(\alpha_m, u_{\eta,N})(I_N(1)) = u_{c,N}(I_N(1)) < u_{c,N}(I_N(k)),$$

and hence we have

$$I_N(k+1) \text{ and } I_N(k+1) \in \mathcal{I}(\alpha) \text{ for } \alpha \leq \alpha_m. \tag{B.46}$$

Case 2: If $u_{\eta,N}(I_N(k+1)) \geq u_{\eta,N}(I_N(k+2))$, then we have by (B.45) that

$$\begin{aligned}&u_{\eta,N}(I_N(k)) < u_{\eta,N}(I_N(k+1)),\\ &u_{\eta,N}(I_N(k+1)) \geq u_{\eta,N}(I_N(k+2)), \text{ and}\\ &u_{\eta,N}(I_N(k+2)) < u_{\eta,N}(I_N(k+3)),\end{aligned}$$

and (B.39) implies that

$$\eta_N(I_N(k+1)) = -\eta_N(I_N(k+2)). \tag{B.47}$$

Hence, $u_{\eta,N}(I_N(k+1))$ and $u_{\eta,N}(I_N(k+2))$ are two extreme regions and, according to Theorem 4.5, we have

$$\mathscr{L}(\alpha, u_{\eta,N})(I_N(k+1)) = u_{c,N}(I_N(k+1)) + \eta_N(I_N(k+1)) - 2N\alpha$$

and

$$\mathscr{L}(\alpha, u_{\eta,N})(I_N(k+2)) = u_{c,N}(I_N(k+2)) + \eta_N(I_N(k+2)) + 2N\alpha.$$

Thus, by (B.47), at $\alpha = \tilde{\alpha}_{k+1}$, where $\tilde{\alpha}_{k+1}$ is defined by

$$2N\tilde{\alpha}_{k+1} := \eta_N(I_N(k+1)) - \frac{1}{2}\left(u_{c,N}(I_N(k+2)) - u_{c,N}(I_N(k+1))\right) < \left|\eta_N(I_N(k+1))\right|, \tag{B.48}$$

we have

$$\mathscr{L}(\tilde{\alpha}_{k+1}, u_{\eta,N})(I_N(k+1)) = \mathscr{L}(\tilde{\alpha}_{k+1}, u_{\eta,N})(I_N(k+2)) \tag{B.49}$$

and both $I_N(k+1)$ and $I_N(k+2)$ are step regions of $\mathscr{L}(\tilde{\alpha}_{k+1}, u_{\eta,N})$.

We observe that (B.49) not only causes a staircasing effect but also, together with (B.48), leaves part of $\eta_N(I_N(k+1))$ and $\eta_N(I_N(k+2))$ un-removed. Moreover, we have

$$\mathcal{E}'_{k+1,N}(\alpha) = \mathcal{E}'_{k+1,\eta_N}(\alpha) < 0 \text{ and } \mathcal{E}'_{k+2,N}(\alpha) = \mathcal{E}'_{k+2,\eta_N}(\alpha) < 0 \text{ for } \alpha \leq (\tilde{\alpha}_{k+1})^-,$$

but

$$\mathcal{E}'_{k+1,N}((\tilde{\alpha}_{k+1})^+) = 0 > \mathcal{E}'_{k+1,\eta_N}(\tilde{\alpha}_{k+1}) \text{ and } \mathcal{E}'_{k+2,N}((\tilde{\alpha}_{k+1})^+) = 0 > \mathcal{E}'_{k+2,\eta_N}(\tilde{\alpha}_{k+1}).$$

That is, $I_N(k+1)$ and $I_N(k+2) \in \mathcal{N}(\alpha)$ for $\alpha < \tilde{\alpha}_{k+1}$, and $I_N(k+1)$ and $I_N(k+2) \in \mathcal{I}(\alpha)$ for $\alpha = \tilde{\alpha}_{k+1}$, and according to (B.40), we have $\tilde{\alpha}_{k+1} \leq \alpha_m$.

We next claim that if $\tilde{\alpha}_{k+1} < \alpha_m$, then $I_N(k+1)$ and $I_N(k+2) \in \mathcal{I}(\alpha)$ for $\tilde{\alpha}_{k+1} < \alpha \leq \alpha_m$. Indeed, let ${u_{\eta,N}}^1 := \mathscr{L}(\tilde{\alpha}_{k+1}, u_{\eta,N})$. One can check that $\eta_N^1 := {u_{\eta,N}}^1 - u_{c,N}$ satisfies all assumptions in Assumption B.10. Then, since $I_N(k+1)$ and $I_N(k+2)$ are two step regions of ${u_{\eta,N}}^1$ by (B.49), we may apply Case 1 above to ${u_{\eta,N}}^1$ and obtain that $I_N(k+1)$ and $I_N(k+2)$ are inactive for at least $0 < \alpha < \alpha_m - \tilde{\alpha}_{k+1}$ because

$$\mathscr{L}(\tilde{\alpha}_{k+1}, u_{\eta,N})(I_N(1)) = u_{c,N}(I_N(1)) - (|\eta_N(I_N(1))| - N\tilde{\alpha}_{k+1}).$$

Hence, we have

$$I_N(k+1) \text{ and } I_N(k+2) \in \mathcal{I}(\alpha) \text{ for all } \tilde{\alpha}_{k+1} < \alpha \leq \alpha_m. \tag{B.50}$$

Therefore, we conclude that

$$\sum_{I_N(k)\in\mathcal{C}_I(u_{c,N})} \mathcal{E}'_{k,N}(\alpha) < 0 \text{ for } 0 \le \alpha < \tilde{\alpha} := \max\{\tilde{\alpha}_k,\ I_N(k) \in \mathcal{C}_I(u_{c,N})\} \le \alpha_m, \quad \text{(B.51)}$$

and, in the view of (B.46) and (B.50), we have that

$$\sum_{I_N(k)\in\mathcal{C}_I(u_{c,N})} \mathcal{E}'_{k,N}(\alpha) = 0 \text{ for } \tilde{\alpha} \le \alpha \le \alpha_m. \quad \text{(B.52)}$$

In the end, since $|\eta_N(I_N(1))| = |\eta_N(I_N(N))|$, then we have, by item 4 in Assumption B.10, that

$$\mathcal{E}'_{1,N}(\alpha) + \mathcal{E}'_{N,N}(\alpha) < 0 \text{ for } \alpha < \alpha_m$$

and

$$\mathcal{E}'_{1,N}(\alpha_m) + \mathcal{E}'_{N,N}(\alpha_m) = 0.$$

Therefore, and taking into consideration of (B.51) and (B.52), we have that

$$0 > \mathcal{E}'_N(\alpha) = \sum_{k\in\mathcal{N}(\alpha)} \mathcal{E}'_{k,N}(\alpha) + \sum_{k\in\mathcal{I}(\alpha)} \mathcal{E}'_{k,N}(\alpha) \ge \mathcal{E}'_{\eta_N}(\alpha)$$

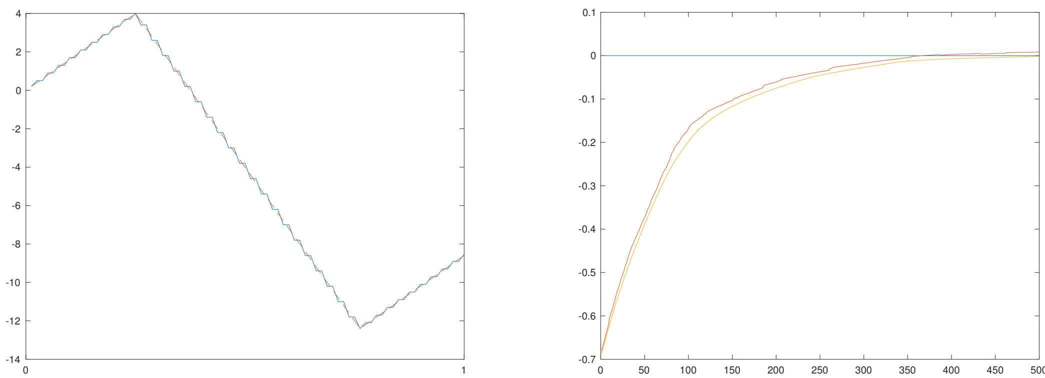
for $0 \le \alpha < \alpha_m$, and $\mathcal{E}'_N(\alpha_m) = 0$.

We next claim that $\mathscr{L}(u_{\eta,N}, \alpha_m)$ is an acceptable deformation of $u_{c,N}$. Item 1 in Assumption B.8 is satisfied because of (B.40), and (B.34) is satisfied because of (B.38) and (B.39). Hence, we may apply Proposition B.9, and deduce that

$$\mathcal{E}'_N(\alpha) > 0 \text{ for } \alpha_m \le \alpha < \alpha_s(u_{\eta,N}).$$

This, together with the fact that $\mathscr{L}(\alpha, u_{\eta,N})$ is a constant for $\alpha > \alpha_s(u_{\eta,N})$, concludes the argument. □

Remark B.11. We cannot reach full-recovery if u_c is not a constant. As (B.49) holds, the damage is permanent, since once two pieces of $\mathscr{L}(\alpha, u_{\eta,N})$ merged together, they can never be separated again.



(A) $\mathscr{L}(\alpha_m, u_{\eta,N})$ in blue, $u_{c,N}$ in red.

(B) $\mathcal{E}'_N(\alpha)$ in red and $\mathcal{E}'_{\eta_N}(\alpha)$ in orange

FIGURE 12. Heavily staircasing effect is observed in Figure 12a, even if with an ideal noise. In Figure 12b, $\mathcal{E}'_N(\alpha) \geq \mathcal{E}_{\eta_N}(\alpha)$ indicts that not all noise are removed

List of Symbols

AT_ε: Ambrosio - Tortorelli approximation. 6
$AT_{\omega,\varepsilon}$: Weighted Ambrosio - Tortorelli approximation. 7

u_c: Clean data (signal in one dimension, image in two dimension). 3
u_η: Corrupted image data (signal in one dimension, image in two dimension). 2

$\mathcal{F}$: Fidelity term in image processing. 2
$u_{c,N}$: Level N finite resolution clean image. 8
$u_{\eta,N}$: Level N finite resolution corrupted image. 9

$\mathcal{E}_N$: Level N error function. 9

MS: Mumford-Shah image segmentation functional. 5
MS_ω: Weighted Mumford-Shah image segmentation functional. 7

$\mathscr{L}$: Reconstruction operator. 9
$\mathcal{R}$: Regularization term (regularizer) in image processing. 2

SBV: Space of special bounded variation. 5

TGV^{1+s}: Fractional order total generalized variation. 4

References

[1] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[2] R. Alicandro, A. Braides, and J. Shah. Free-discontinuity problems via functionals involving the L^1-norm of the gradient and their approximations. *Interfaces Free Bound.*, 1(1):17–37, 1999.

[3] L. Ambrosio. Variational problems in SBV and image segmentation. *Acta Appl. Math.*, 17(1):1–40, 1989.

[4] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[5] L. Ambrosio and P. Tilli. *Topics on analysis in metric spaces*, volume 25 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2004.

[6] L. Ambrosio and V. M. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Comm. Pure Appl. Math.*, 43(8):999–1036, 1990.

[7] H. Attouch. *Variational convergence for functions and operators*. Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[8] M. Avriel. *Nonlinear programming: analysis and methods*. Courier Corporation, 2003.

[9] A. Baldi and B. Franchi. A Γ-convergence result for doubling metric measures and associated perimeters. *Calc. Var. Partial Differential Equations*, 16(3):283–298, 2003.

[10] A. Baldi and B. Franchi. Mumford-Shah-type functionals associated with doubling metric measures. *Proc. Roy. Soc. Edinburgh Sect. A*, 135(1):1–23, 2005.

[11] G. Bellettini, V. Caselles, and M. Novaga. The total variation flow in $\mathbb{R}^N$. *J. Differential Equations*, 184(2):475–525, 2002.

[12] M. Bergounioux. Optimal control of problems governed by abstract elliptic variational inequalities with state constraints. *SIAM J. Control Optim.*, 36(1):273–289 (electronic), 1998.

[13] J. Bourgain, H. m. Brezis, and P. Mironescu. Limiting embedding theorems for $W^{s,p}$ when $s \uparrow 1$ and applications. *J. Anal. Math.*, 87:77–101, 2002. Dedicated to the memory of Thomas H. Wolff.

[14] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.

[15] A. Braides. *Γ-convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[16] A. Braides and V. Chiadò Piat. Integral representation results for functionals defined on $\mathrm{SBV}(\Omega; \mathbf{R}^m)$. *J. Math. Pures Appl. (9)*, 75(6):595–626, 1996.

[17] K. Bredies and T. Valkonen. Inverse problems with second-order total generalized variation constraints. *Proceedings of SampTA 2011 - 9th International Conference on Sampling Theory and Applications, Singapore*, 2011.

[18] K. Bredies and T. Valkonen. Inverse problems with second-order total generalized variation constraints. *Proceedings of SampTA*, 201, 2011.

[19] M. Burger, G. Gilboa, M. Moeller, L. Eckardt, and D. Cremers. Spectral Decompositions using One-Homogeneous Functionals. *ArXiv e-prints*, Jan. 2016.

[20] V. Caselles, A. Chambolle, and M. Novaga. The discontinuity set of solutions of the TV denoising problem and some extensions. *Multiscale Model. Simul.*, 6(3):879–894, 2007.

[21] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76(2):167–188, 1997.

[22] D. Chen, S. Sun, C. Zhang, Y. Chen, and D. Xue. Fractional-order TV-L^2 model for image denoising. *Central European Journal of Physics*, 11(10):1414–1422, 2013.

[23] Y. Chen, T. Pock, R. Ranftl, and H. Bischof. Revisiting loss-specific training of filter-based mrfs for image restoration. In *Pattern Recognition*, pages 271–281. Springer, 2013.

[24] Y. Chen, R. Ranftl, and T. Pock. Insights into analysis operator learning: From patch-based sparse models to higher order mrfs. *IEEE Transactions on Image Processing*, 23(3):1060–1072, March 2014.

[25] C. Chung, J. C. De los Reyes, and C. B. Schoenlieb. Learning optimal spatially-dependent regularization parameters in total variation image restoration. *ArXiv e-prints*, Mar. 2016.

[26] B. Dacorogna. *Weak continuity and weak lower semicontinuity of nonlinear functionals*, volume 922 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1982.

[27] G. Dal Maso. *An introduction to* Γ*-convergence*, volume 8 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Boston, Inc., Boston, MA, 1993.

[28] G. Dal Maso, I. Fonseca, G. Leoni, and M. Morini. A higher order model for image restoration: the one-dimensional case. *SIAM J. Math. Anal.*, 40(6):2351–2391, 2009.

[29] J. Dávila. On an open question about functions of bounded variation. *Calc. Var. Partial Differential Equations*, 15(4):519–527, 2002.

[30] E. Davoli, I. Fonseca, and P. Liu. Image reconstruction and $\mathscr{A}$-$\mathscr{B}$ quasiconvexity. *In preparation.*

[31] E. Davoli and P. Liu. Optimal parameters and regularizers for image reconstruction problems - two dimensional case. *In preparation.*

[32] E. De Giorgi, M. Carriero, and A. Leaci. Existence theorem for a minimum problem with free discontinuity set. *Arch. Rational Mech. Anal.*, 108(3):195–218, 1989.

[33] J. C. De los Reyes and C.-B. Schönlieb. Image denoising: learning the noise model via nonsmooth PDE-constrained optimization. *Inverse Probl. Imaging*, 7(4):1183–1214, 2013.

[34] J. C. De los Reyes, C.-B. Schönlieb, and T. Valkonen. Bilevel parameter learning for higher-order total variation regularisation models. *Journal of Mathematical Imaging and Vision*, pages 1–25, 2016.

[35] J. C. De Los Reyes, C.-B. Schönlieb, and T. Valkonen. The structure of optimal parameters for image restoration problems. *J. Math. Anal. Appl.*, 434(1):464–500, 2016.

[36] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012.

[37] J. Domke. Generic methods for optimization-based modeling. In *AISTATS*, volume 22, pages 318–326, 2012.

[38] J. Domke. Learning graphical model parameters with approximate marginal inference. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(10):2454–2467, Oct 2013.

[39] I. Ekeland and R. Temam. *Convex analysis and variational problems.* SIAM, 1976.

[40] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[41] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions.* CRC press, 2015.

[42] H. Federer. *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[43] M. Focardi. On the variational approximation of free-discontinuity problems in the vectorial case. *Math. Models Methods Appl. Sci.*, 11(4):663–684, 2001.

[44] G. B. Folland. *Real analysis: modern techniques and their applications.* John Wiley & Sons, 2013.

[45] I. Fonseca and G. Leoni. *Modern methods in the calculus of variations: sobolev spaces.* Springer Monographs in Mathematics. Unpublished yet, 2015.

[46] I. Fonseca and P. Liu. Higher order Ambrosio-Tortorelli scheme with non-negative spatially dependent parameters. *In preparation.*

[47] I. Fonseca and S. Müller. $\mathscr{A}$-quasiconvexity, lower semicontinuity, and Young measures. *SIAM J. Math. Anal.*, 30(6):1355–1390 (electronic), 1999.

[48] E. Haber, L. Horesh, and L. Tenorio. Numerical methods for the design of large-scale nonlinear discrete ill-posed inverse problems. *Inverse Problems*, 26(2):025002, 14, 2010.

[49] E. Haber and L. Tenorio. Learning regularization functionals—a supervised training approach. *Inverse Problems*, 19(3):611–626, 2003.

[50] M. Hintermüller, K. Papafitsoros, and C. N. Rautenberg. Analytical aspects of spatially adapted total variation regularisation. *ArXiv e-prints*, Sept. 2016.

[51] A. Iosevich. *Sets of finite perimeter and geometric variational problems. An introduction to geometric measure theory*, volume 53. 2016.

[52] H. B. J. Bourgain and P. Mironescu. Another look at sobolev spaces. *Optimal Control and Partial Differential Equations, J. L. Menaldi, E. Rofman, A. Sulem (Eds.), a volume in honor of A. Bensoussan's 60th birthday, IOS Press, Amsterdam*, pages 439–455, 2001.

[53] K. C. Kiwiel. Convergence and efficiency of subgradient methods for quasiconvex minimization. *Math. Program.*, 90(1, Ser. A):1–25, 2001.

[54] K. Kunisch and T. Pock. A bilevel optimization approach for parameter learning in variational models. *SIAM J. Imaging Sci.*, 6(2):938–983, 2013.

[55] K. Kunisch and T. Pock. A bilevel optimization approach for parameter learning in variational models. *SIAM J. Imaging Sci.*, 6(2):938–983, 2013.

[56] N. S. Landkof. *Foundations of modern potential theory.* Springer-Verlag, New York-Heidelberg, 1972. Translated from the Russian by A. P. Doohovskoy, Die Grundlehren der mathematischen Wissenschaften, Band 180.

[57] G. Leoni. *A first course in Sobolev spaces*, volume 105 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2009.

[58] G. Leoni. *A first course in Sobolev spaces*, volume 105 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2009.

[59] G. Leoni and R. Murray. Second-order Γ-limit for the Cahn-Hilliard functional. *Arch. Ration. Mech. Anal.*, 219(3):1383–1451, 2016.

[60] P. Liu. The comprehensive bi-level learning scheme for image processing. *To be submitted.*

[61] P. Liu. The generalized Mumford-Shah functionals for image processing. *In preparation.*

[62] P. Liu. The quasi-convexity of bi-level learning scheme. *In preparation.*

[63] P. Mattila. *Geometry of sets and measures in Euclidean spaces*, volume 44 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.

[64] V. Maz′ya and T. Shaposhnikova. On the Bourgain, Brezis, and Mironescu theorem concerning limiting embeddings of fractional Sobolev spaces. *J. Funct. Anal.*, 195(2):230–238, 2002.

[65] V. G. Maz'ya and T. O. Shaposhnikova. *Theory of Sobolev multipliers*, volume 337 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 2009. With applications to differential and integral operators.

[66] Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures*, volume 22. American Mathematical Soc., 2001.

[67] D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 17, pages 137–154. San Francisco, 1985.

[68] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992. Experimental mathematics: computational issues in nonlinear science (Los Alamos, NM, 1991).

[69] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences.* Springer, New York, 2009.

[70] S. Semmes. Bi-Lipschitz mappings and strong A_∞ weights. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 18(2):211–248, 1993.

[71] J. Simon. Sobolev, Besov and Nikol′skiĭ fractional spaces: imbeddings and comparisons for vector valued spaces on an interval. *Ann. Mat. Pura Appl. (4)*, 157:117–148, 1990.

[72] D. M. Strong, T. F. Chan, et al. Exact solutions to total variation regularization problems. In *UCLA CAM Report*. Citeseer, 1996.

[73] M. F. Tappen. Utilizing variational optimization to learn markov random fields. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2007.

[74] M. F. Tappen, C. Liu, E. H. Adelson, and W. T. Freeman. Learning gaussian conditional random fields for low-level vision. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2007.

[75] H. Triebel. *Theory of function spaces*. Modern Birkhäuser Classics. Birkhäuser/Springer Basel AG, Basel, 2010. Reprint of 1983 edition, Also published in 1983 by Birkhäuser Verlag.

[76] D. Vicente. An anisotropic Mumford–Shah model. *J. Math. Anal. Appl.*, 447(1):181–205, 2017.

[77] T. Wittman. Lost in the supermarket: Decoding blurry barcodes. *SIAM News*, 37(7):16, 2004.

[78] J. Zhang and K. Chen. A total fractional-order variation model for image restoration with nonhomogeneous boundary conditions and its numerical solution. *SIAM J. Imaging Sci.*, 8(4):2487–2518, 2015.