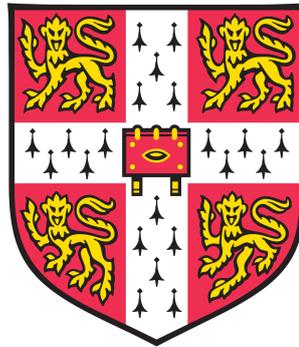


First-order gradient regularisation methods for image restoration

Reconstruction of tomographic images with thin structures
and denoising piecewise affine images



Evangelos Papoutsellis

Jesus College

DAMTP, Center for Mathematical Sciences

University of Cambridge

A thesis submitted for the degree of

Doctor of Philosophy

August 2015

Abstract

The focus of this thesis is variational image restoration techniques that involve novel non-smooth first-order gradient regularisers: Total Variation (TV) regularisation in image and data space for reconstruction of thin structures from PET data and regularisers given by an infimal-convolution of TV and L^p seminorms for denoising images with piecewise affine structures.

In the first part of this thesis, we present a novel variational model for PET reconstruction. During a PET scan, we encounter two different spaces: the sinogram space that consists of all the PET data collected from the detectors and the image space where the reconstruction of the unknown density is finally obtained. Unlike most of the state of the art reconstruction methods in which an appropriate regulariser is designed in the image space only, we introduce a new variational method incorporating regularisation in image and sinogram space. In particular, the corresponding minimisation problem is formed by a total variational regularisation on both the sinogram and the image and with a suitable weighted L^2 fidelity term, which serves as an approximation to the Poisson noise model for PET. We establish the well-posedness of this new model for functions of Bounded Variation (BV) and perform an error analysis through the notion of the Bregman distance. We examine analytically how TV regularisation on the sinogram affects the reconstructed image especially the boundaries of objects in the image. This analysis motivates the use of a combined regularisation principally for reconstructing images with thin structures.

In the second part of this thesis we propose a first-order regulariser that is a combination of the total variation and L^p seminorms with $1 < p \leq \infty$. A well-posedness analysis is presented and a detailed study of the one dimensional model is performed by computing exact solutions for simple functions such as the step function and a piecewise affine function, for the regulariser with $p = 2$ and $p = \infty$. We derive necessary and sufficient conditions for a pair in $BV \times L^p$ to be a solution for our proposed model and determine the structure of solutions dependent on the value of p . In the case $p = 2$, we show that the regulariser is equivalent to the Huber-type variant of total variation regularisation. Moreover, there is a certain class of one dimensional data functions for which the regularised solutions are equivalent to high-order regularisers such as the state of the art total generalised variation (TGV) model.

The key assets of our regulariser are the elimination of the staircasing effect - a well-known disadvantage of total variation regularisation - the capability of obtaining piecewise affine structures for $p = \infty$ and qualitatively comparable results to TGV. In addition, our first-order TVL^p regulariser is capable of preserving spike-like structures that TGV is forced to smooth. The numerical solution of the proposed first-order model is in general computationally more efficient compared to high-order approaches.

Keywords: Total variation regularisation, Radon transform, Positron Emission Tomography reconstruction, Sinogram space regularisation, Bregman distance, Split Bregman algorithm, Thin structures, Staircasing effect, First-order regularisers, Infimal convolution, Image decomposition, TVL^p regularisers, Total generalised variation.

Dedicated to my beloved grandmother Evangelia (1915-2004)

Acknowledgements

I never thought that writing the acknowledgements section in my thesis would be more difficult than any other section. In my opinion, a proper acknowledgement is to gather all the people below and organise a decent *Greek-feast acknowledgement* with lot of *ouzo* and *papalina* in Agiasos. However, since it is practically impossible, I would like to devote the following lines to thank each and everyone for his or her help and support in any level during the last five years, when I first moved to Cambridge to continue my mathematical studies.

First of all, I would like to express my gratitude to my Master and Ph.D supervisor Carola-Bibiane Schönlieb for giving me the opportunity to work on a very interesting and challenging topic in one of the most beautiful Universities. For being an excellent supervisor, for her continuous support and encouragement and for being the first person who introduced me to the field of mathematical imaging after my completion of my Master. Thank you for believing in me and for motivating me to be a better mathematician.

My friends and colleagues Luca Calatroni and Konstantinos Papafitsoros for very helpful and fruitful mathematical discussions for the last four years. Luca for making me laugh with his amazing imitations and Konstantinos for sharing his passion about turtles. Also, I would like to thank all the members of the Cambridge Image Analysis (CIA) group and especially Jan Lellmann for his numerical analysis expertise, informing the CIA group about CVX and Tuomo Valkonen for his excellent TGV knowledge.

My officemates Amandeep Kaur and Pranav Singh who created a beautiful working environment and for helping me learn Punjabi and Hindi from time to time!!!

I am extremely thankful and indebted to Martin Burger and his imaging group in University of Münster for making my three research visits a beautiful place to work outside Cambridge. For helping me with very interesting ideas and discussions on both my two research papers and for organising nice football games during my visits. I am also grateful to the other Martin (Benning), for welcoming me in the mathematics department, helping me understand the concept of Bregman distance and always demonstrating his passion about

football and BVB. Special thanks also to Jahn Müller for sharing his PET expertise, helping me with my first publication and organising a very impressive visit to the European Institute for Molecular imaging (EIMI) in Münster.

Many thanks to all my friends from the Hellenic Basketball society Dionyssios Mantzavinos, Vassilis Taralis and Antonis Tsiflis who create a very strong and competitive basketball team and accomplished to be the champions of the Cambridge Basketball League for two years in a row.

I consider it an honour to meet Reverend Nikolaos Vernezos, a very kind person with amazing strength and spirituality who tried every Sunday to bring together all the Christians in Cambridge and for organising proper Greek feasts, especially during Easter.

I also acknowledge the financial support provided from Jesus College of University of Cambridge, the Engineering and Physical Sciences Research Council (EPSRC) and the King Abdullah University of Science and Technology.

This thesis would have remained a dream had it not been for my parents Efstratios and Mirsini and by brothers Stavros and Christos. For always believe in me and for your unconditional support.

Above all, I would like to thank my life partner Antonina Pliatsika, for her love, support and patience for the last five years especially when we were away from each other. I could not accomplish my goal without your beautiful *joker* smile.

Statement of Originality

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. Specifically, I certify that:

Chapter 1 is an introduction of the thesis and provides the reader with the basic concepts of mathematical image processing. It also contains a summary of my own research work.

Chapter 2 is a mathematical summary of some known results in measure theory, functions of bounded variation and convex analysis. It is not my personal work.

Chapter 3 introduces a novel approach related to PET reconstruction and can be found in [BMPS14]. It is a result of a collaboration with my supervisor Carola Schönlieb and her prior work in [BFS11], Martin Burger and Jahn Müller from University of Münster, Germany. All the proofs and numerical implementation were written by me. The Radon matrix and the realistic PET phantoms used in the numerical implementations were provided by Jahn Müller and the European Institute for Molecular imaging (EIMI) in Münster.

Chapter 4 introduces a novel first-order regulariser based on the total variation and L^p norm and it can be found in [BPPS15a] and [BPPS15b]. It has been done in collaboration with my supervisor Carola Schönlieb, Konstantinos Papafitsoros from University of Cambridge and Martin Burger from University of Münster. All the proofs and numerical implementation were written by me except Propositions 4.3.1 and 4.4.10 where they were proved by Martin Burger and Konstantinos Papafitsoros respectively. The numerical implementation that involves the TGV regulariser is due to Konstantinos Papafitsoros.

Contents

List of Notation	13
1 Introduction	19
1.1 Digital image processing	20
1.2 Inverse Problems: Variational approach	22
1.3 Contribution	25
2 Mathematical Preliminaries	33
2.1 Measure Theory	34
2.2 Lebesgue and Sobolev spaces	38
2.2.1 L^p spaces	38
2.2.2 $W^{k,p}$ spaces	40
2.3 Functions of bounded variation	42
2.4 Convex Analysis	44
2.4.1 Fenchel-Rockafellar duality	49
3 Total variation regularisation in image and sinogram space for PET reconstruction	51
3.1 Introduction	51
3.2 Positron Emission Tomography	53
3.3 Radon transform	57
3.4 State of the art image reconstruction methods	63
3.5 TV regularisation on image and sinogram	65
3.5.1 BV continuity of the Radon transform	66
3.5.2 Well-posedness of the model	68
3.5.3 Error analysis using the Bregman distance	73
3.5.4 An explicit example of TV regularisation on the sinogram	76
3.6 Numerical Implementation	81
3.6.1 Discrete Setting	81
3.6.2 Split Bregman Algorithm	83
3.7 Numerical Results	86

3.7.1	Image reconstruction from corrupted simulated and real PET data	88
3.7.2	Scale space of sinogram regularisation	92
3.7.3	Thin Structure Reconstruction	97
4	Infimal convolution regularisation functionals of BV and L^p spaces	105
4.1	Introduction	105
4.2	Total variation and L^p regularisation	109
4.3	The p -homogeneous analogue and relation to Huber-TV	114
4.4	The one dimensional case	119
4.4.1	Optimality conditions	120
4.4.2	Structure of the solutions	125
4.4.3	Exact solutions	130
4.5	An image decomposition approach	140
4.5.1	Characterisation of minimisers	142
4.6	Numerical Experiments	147
4.6.1	Split Bregman for TVL^p-L^2	149
4.6.2	One dimensional results	152
4.6.3	Two dimensional results	156
4.6.4	The $p = \infty$ case	162
4.6.5	Open problems	168
5	Conclusions	175
A	Exact solutions of TVL^∞: Piecewise affine function	177
	Bibliography	183

List of Notation

Against each entry is the page at which the notation is introduced.

Spaces of continuous and differentiable functions:

$\mathcal{C}^k(X, \mathbb{R}^\ell)$	33	The space of k -times continuously differentiable \mathbb{R}^ℓ -valued functions in X .
$\mathcal{C}_c^k(X, \mathbb{R}^\ell)$	33	The space of k -times continuously differentiable \mathbb{R}^ℓ -valued functions with compact support in X .
$\mathcal{C}^\infty(X, \mathbb{R}^\ell)$	33	The space of infinitely many times continuously differentiable \mathbb{R}^ℓ -valued functions in X .
$\mathcal{C}_c^\infty(X, \mathbb{R}^\ell)$	33	The space of infinitely many times continuously differentiable \mathbb{R}^ℓ -valued functions with compact support in X .
$\mathcal{D}(X, \mathbb{R}^\ell)$	33	Another notation of $\mathcal{C}_c^\infty(X, \mathbb{R}^\ell)$, known as test functions.
$\mathcal{C}_c(X, \mathbb{R}^\ell)$	33	The space of continuous \mathbb{R}^ℓ -valued functions with compact support in X .
$\mathcal{C}_0(X, \mathbb{R}^\ell)$	33	The completion of $\mathcal{C}_c(X, \mathbb{R}^\ell)$ under the supremum norm $\ u\ _\infty = \sup_{x \in X} u(x) $.
$\mathcal{C}_c^k(X)$	33	The space of k -times continuously differentiable real valued functions with compact support in X .
$\mathcal{C}_c^\infty(X)$	33	The space of infinitely many times continuously differentiable real valued functions with compact support in X .
$\mathcal{C}_c(X)$	33	The space of continuous real valued functions with compact support in X .
$\mathcal{C}_0(X)$	33	The completion of $\mathcal{C}_c(X)$ under the supremum norm $\ u\ _\infty = \sup_{x \in X} u(x) $.

Measure theory:

\mathfrak{C}	34	σ -algebra.
$\mathcal{B}(X)$	34	The Borel σ -algebra of X .
$ \mu $	34	The total variation measure of μ .
$\nu \ll \mu$	35	ν is absolutely continuous with respect to μ .
$\nu \perp \mu$	35	ν and μ are mutually singular.
ν^a	35	The absolutely continuous part of measure ν .

ν^s	35	The singular part of measure ν .
$\text{sgn}(\mu)$	36	The sign of measure μ defined as $\frac{\mu}{ \mu }$.
$\mathcal{M}(X, \mathbb{R}^\ell)$	36	The set of \mathbb{R}^ℓ -valued finite Radon measures on X .
$\mathcal{M}_{loc}(X, \mathbb{R}^\ell)$	36	The set of \mathbb{R}^ℓ -valued Radon measures on X .
$\mathcal{M}(X)$	36	The set of real valued finite Radon measures on X .
$\mathcal{M}_{loc}(X)$	36	The set of real valued Radon measures on X .
$\mu_n \xrightarrow{w^*} \mu$	37	Weak* convergence for measures i.e.,

$$\lim_{n \rightarrow \infty} \int_X u d\mu_n = \int_X u d\mu.$$

Space of Lebesgue integrable functions:

\mathcal{L}^d	38	The Lebesgue measure on \mathbb{R}^d with $d \geq 1$.
$L^p(X, \mathbb{R}^\ell; \mu)$	38	The space of \mathbb{R}^ℓ -valued, μ -measurable functions such that $\ u\ _{L^p(X, \mathbb{R}^\ell; \mu)} = (\int_X u ^p d\mu)^{\frac{1}{p}} < \infty$ for $1 \leq p < \infty$ and μ a positive measure.
$L^\infty(X, \mathbb{R}^\ell; \mu)$	38	The space of \mathbb{R}^ℓ -valued, μ -essentially bounded measurable functions such that $\ u\ _{L^\infty(X, \mathbb{R}^\ell; \mu)} = \text{ess sup}_{u \in X} u < \infty$ and μ a positive measure.
$L^p(X, \mathbb{R}^\ell)$	38	The space of \mathbb{R}^ℓ -valued, Lebesgue measurable functions with $1 \leq p \leq \infty$.
$L^p(X)$	38	The space of real valued, Lebesgue measurable functions with $1 \leq p \leq \infty$.
$L^1_+(X)$	68	The space of positive valued, Lebesgue measurable functions.

Sobolev spaces:

Ω	40	An open subset of \mathbb{R}^d .
u_Ω	41	The mean value of u in Ω , i.e., $u_\Omega = \frac{1}{\mathcal{L}^d(\Omega)} \int_\Omega u dx$.
$D^a u$	41	The distributional (weak) a -th derivative of the function u .
$W^{k,p}(\Omega)$	41	The space of $u \in L^p(\Omega)$ such that the distributional derivative $Du \in L^p(\Omega)$ for $0 \leq a \leq k$.
$H^k(\Omega)$	41	The Sobolev space $W^{k,2}(\Omega)$.
$H^k_0(\Omega)$	41	The completion of $\mathcal{C}^\infty_c(\Omega)$ under the $\ \cdot\ _{W^{k,2}(\Omega)}$ norm.

Functions of bounded variation:

$BV(\Omega, \mathbb{R}^\ell)$	42	The space of \mathbb{R}^ℓ -valued functions of bounded variation on Ω .
$BV(\Omega)$	42	The space of real valued functions of bounded variation on Ω .

$D^\alpha u$	43	The absolutely continuous part of the Radon measure Du with respect to Lebesgue measure, equivalently denoted as ∇u or u' in one dimension.
$D^s u$	43	The singular part of the Radon measure Du with respect to Lebesgue measure.
δ_x	43	The Dirac atomic measure concentrated on $\{x\}$.
$\text{TV}(u)$	43	The total variation of $u \in L^1(\Omega, \mathbb{R}^\ell)$, equivalently denoted as $ Du (\Omega)$ when $u \in \text{BV}(\Omega, \mathbb{R}^\ell)$.
$u_n \xrightarrow{w^*} u$	43	Weak* convergence in BV spaces i.e., $u_n \rightarrow u$ in $L^1(\Omega, \mathbb{R}^\ell)$ and $Du_n \xrightarrow{w^*} Du$ in Ω .

Convex analysis:

$\text{dom}F$	45	The effective domain of a functional $F : X \rightarrow \overline{\mathbb{R}}$ i.e., $\text{dom}F = \{u \in X : F(u) < \infty\}$.
F^*	45	The convex conjugate or Legendre-Fenchel transformation of F .
F^{**}	45	The convex biconjugate of F .
$F'_G(u)$	46	The Gâteaux derivative of F at u .
$F'_F(u)$	46	The Fréchet derivative of F at u .
$\partial F(u)$	47	The subdifferential of F at u .
K^*	49	The adjoint operator $K^* : Y^* \rightarrow X^*$ of the bounded linear functional $K : X \rightarrow Y$, where X, Y are Banach spaces.
$\sigma_C(u)$	45	The support function of u in $C \subset X$.

Miscellaneous:

\mathbb{R}^+	33	The set of positive real numbers.
$\overline{\mathbb{R}}$	33	The set of extended real numbers, $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$.
\mathcal{X}_Ω	77	The characteristic function of the set Ω , i.e., $\mathcal{X}_\Omega(x) = \begin{cases} 1 & \text{if } x \in \Omega, \\ 0 & \text{if } x \notin \Omega. \end{cases}$
\mathcal{I}_Ω	45	The indicator function of the set Ω , i.e., $\mathcal{I}_\Omega(x) = \begin{cases} 0 & \text{if } x \in \Omega, \\ \infty & \text{if } x \notin \Omega. \end{cases}$
\mathcal{S}^{d-1}	57	The unit sphere in \mathbb{R}^d .
$\mathcal{R}u(\theta, s)$	57	The Radon transform of $u \in L^1(\mathbb{R}^d)$ in θ direction and s distance from the origin, i.e., defined as

$$\mathcal{R}u(\theta, s) = \int_{\mathbb{R}^d} \delta(s - x \cdot \theta) u(x) dx,$$

where δ is the one dimensional Dirac function.

Σ^d	57	The domain of the Radon transform, i.e., $\Sigma^d = \left\{ (\theta, s) : \theta \in \mathcal{S}^{d-1}, s \in \mathbb{R} \right\},$ a cylinder of dimension d .
Σ^2	59	The domain of the Radon transform for a two dimensional function u with compact support known as sinogram space, i.e., $\Sigma^2 = \{(\phi, s) : s \in [-r, r], \phi \in [0, \pi)\},$ for some $r > 0$.
$\mathcal{R}^*g(x)$	61	The adjoint operator of the Radon transform \mathcal{R} for $g \in L^1(\Sigma^d)$ known as backprojection operator defined as $\mathcal{R}^*g(x) = \int_{\mathcal{S}^{d-1}} g(\theta, x \cdot \theta) d\theta.$
$\text{diam}(K)$	60	The diameter of set K , i.e., $\text{diam}(K) = \sup \{ \rho(x, y) : x, y \in K \},$ under a metric ρ .
EM	65	Abbreviation of the Expectation-Maximisation algorithm.
$D_J^\xi(u, v)$	74	The Bregman distance of a convex functional $J : X \rightarrow \overline{\mathbb{R}}$ for a fixed subgradient ξ .
$\mathcal{A}(u(\tilde{r}))(x)$	80	The Abel transform of u defined as $2 \int_x^\infty \frac{\tilde{r}u(\tilde{r})}{\sqrt{\tilde{r}^2 - x^2}} d\tilde{r}.$
$\text{TVL}_{\alpha, \beta}^p$	105	Total variation and L^p regulariser defined as $\text{TVL}_{\alpha, \beta}^p(u) := \inf_{w \in L^p(\Omega)} \alpha \ Du - w\ _{\mathcal{M}} + \beta \ w\ _{L^p(\Omega)}.$
$\text{BV}^2(\Omega)$	106	The space of functions of bounded Hessian on Ω , i.e., all the functions $u \in W^{1,1}(\Omega)$ such that $\nabla u \in \text{BV}(\Omega, \mathbb{R}^d)$.
$\text{TV}^2(u)$	106	The second order total variation of u , equal to $\ D^2u\ _{\mathcal{M}}$ for a function $u \in \text{BV}^2(\Omega)$.
$\text{ICTV}_{\alpha, \beta}$	106	The TV-TV ² infimal convolution of u , i.e., $\text{ICTV}_{\alpha, \beta}(u) := \min_{v \in \text{BV}^2(\Omega)} \alpha \text{TV}(u - v) + \beta \text{TV}^2(v).$
$\text{BD}(\Omega)$	107	The space of functions of bounded deformation on Ω , i.e., all the functions $u \in L^1(\Omega, \mathbb{R}^d)$ such that the distributional symmetrised gradient of u , $\mathcal{E}u$ is a $\mathbb{R}^{d \times d}$ -valued finite Radon measure.

$\text{TGV}_{\alpha,\beta}^2$ 107 The second order total generalised variation defined as

$$\text{TGV}_{\alpha,\beta}^2(u) := \min_{w \in \text{BD}(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|\mathcal{E}w\|_{\mathcal{M}}.$$

$\mu \lfloor A$ 127 The restriction of the measure μ to the set A .

Basics:

\mathbb{N} - The set of natural numbers.

$\|\cdot\|_p$ - The discrete norm in ℓ^p space for $1 \leq p \leq \infty$.

B_X - The unit ball of space X .

Δu - The Laplacian of u i.e., $\Delta u = \sum_{i=1}^d \frac{\partial^2 u_i}{\partial x_i^2}$.

Chapter 1

Introduction

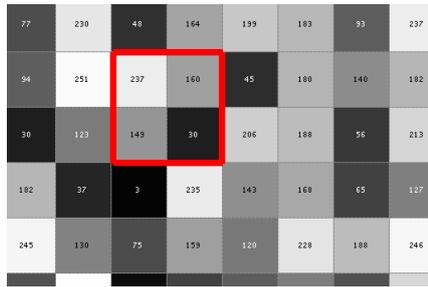
The basic goal of my Ph.D thesis is to explore novel variational methods applied in image processing. I am especially interested in *total variation* (TV) based regularisation techniques for general image restoration as well as for image reconstruction in *positron emission tomography* (PET).

The disciplines of image processing, image analysis and image restoration are terms that are naturally related to computing with images. In fact, for more than three decades the aforementioned fields were occupied mostly by computer scientists and electrical engineers without any significant interest by mathematicians. However, in the late 90's many mathematicians developed an increasing interest in these fields, especially due to Rudin, Osher and Fatemi and their research paper *Nonlinear total variation based noise removal algorithms* in [ROF92, 1992]. This was not restricted to image denoising applications, but also to image segmentation, see for instance the paper of Mumford and Shah in [MS89, 1989] or to a theoretical analysis of imaging models as in [AV94, 1994]. In addition, Rudin et al. introduced the notion of the total variational regularisation which in practice is considered as the integral of the absolute value of the gradient. The main outcome of this pioneering work was the introduction of *the space of bounded variation* as the appropriate space to explore images. It is considered by many researchers the *holy grail* of mathematical imaging and a significant breakthrough towards image reconstruction.

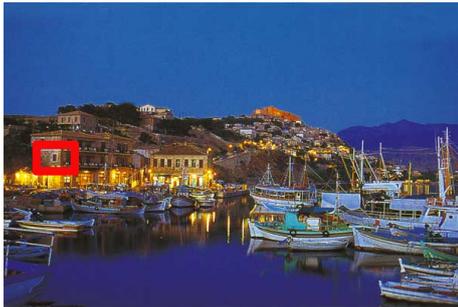
It is a standard process to examine images from two different perspectives. One is from the computer science framework where an image is regarded as a collection of numbers, for instance a matrix with fixed columns and rows. Here, we say that an image *lives* in the *discrete* world. The other is as an abstract mathematical entity, for instance a function that assigns for every point in the *real* world a certain value. In the following, we discuss these concepts of digital and mathematical image processing and lay emphasis on properties such as *edges or jump discontinuities*.

1.1 Digital image processing

Everyday, we encounter a plethora of digital images captured in different situations: a simple photo from our camera or our smartphone to climatological (weather prediction), astronomical (Pluto images) and medical images (tomography) are only some of them. One way to understand digital images is as a natural projection of the real world to the digital world, for example our computer screen. *Digital or discrete* images, as they are commonly referred to, are matrices of a fixed grid size $n \times m$ that contain squared pictures elements, i.e., *pixels* that are arranged in columns and rows.



(a) Chessboard grey image



(b) Coloured image of Molyvos port in Lesvos



(c) RGB pixel values of selected region

Figure 1.1: Grey and coloured digital images with their pixel values for selected regions.

For example, in a greyscale image every picture element has a specific value ranged from 0 to 255. It is normally interpreted as a black (0) and white (255) image but also include many shades of grey, see a *chessboard* kind of image with their pixel values in Figure 1.1a. Another common convention that is also used in this thesis is to rescale the pixel values to $[0, 1]$ range. Extending to coloured images (RGB images), we assign a value from the red, green and blue channels to each pixel as in Figures 1.1b, 1.1c.

Although, a digital image lies in the *discrete world*, a useful convention for modelling image reconstruction models mathematically is to consider images as scalar or vector valued functions from a *continuous space* $\Omega \subset \mathbb{R}^2$ i.e.,

$$u : \Omega \rightarrow \mathbb{R} \quad \text{or} \quad u : \Omega \rightarrow \mathbb{R}^3. \quad (1.1)$$



Figure 1.2: Andromeda Galaxy: Captured with the NASA/ESA Hubble Space Telescope and is a cropped version of the full image with 1.5 billion pixels. One would need more than 600 HD television screens to display the whole image. <http://www.spacetelescope.org/images/heic1502a/>

An image-function defined in the continuous setting is easier to analyse for its properties by using all the necessary tools that are available from our mathematical toolbox e.g. *functional* analysis, *convex* analysis and *measure* theory. It is an admissible and fair approximation to interpret analytically real world images, since one may think that a 21st century digital image can actually be represented by billions and even more pixels as it appears in Figure 1.2.

Having an image as a function, one may ask the following question: what is the most appropriate space to describe images? For example, an image can be easily considered as an L^2 , L^∞ or even a $W^{1,2}$ function, see Section 2.2. However, we need to take into account the most basic property that the majority of the images have, known as the *edge* of the image. In computer vision terminology, it is identified as a set of continuous pixel positions where an abrupt change of intensity values occur, see for instance the red regions and their pixels values in Figure 1.1. In the context of mathematical image processing, an image-function has a *jump discontinuity* in this point and hence the derivative in this point is not defined. Therefore, it is essential to find a suitable space whose elements can describe this kind of behaviour.

Let us consider the following useful example of a function defined in $\Omega = [-1, 1]$ for $\varepsilon > 0$ as:

$$u_\varepsilon(x) = \begin{cases} -1, & -1 \leq x < -\varepsilon, \\ \frac{x}{\varepsilon}, & -\varepsilon \leq x \leq \varepsilon, \\ 1, & \varepsilon < x \leq 1, \end{cases} \quad \text{with} \quad \nabla u_\varepsilon(x) = \begin{cases} \frac{1}{\varepsilon}, & -\varepsilon \leq x \leq \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

This function has a piecewise constant part and a relatively small linear part in $[-\varepsilon, \varepsilon]$

internal, see Figure 1.3. Clearly $u_\varepsilon(x) \in W^{1,p}$ for $p \geq 1$. The pointwise limit of u_ε is

$$u_0(x) = \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0, \end{cases}$$

which is not differentiable and is not an element of $W^{1,p}(\Omega)$ for any $p > 1$, cf. [Eva10]. We usually refer to u_0 as a *step function* having a jump discontinuity at $x = 0$. In fact,

$$\|\nabla u_\varepsilon\|_{L^p} = \frac{2}{\varepsilon^{p-1}} \xrightarrow{\varepsilon \rightarrow 0} \infty$$

for $p \in (1, \infty)$ and therefore Sobolev spaces cannot model edges or discontinuities.

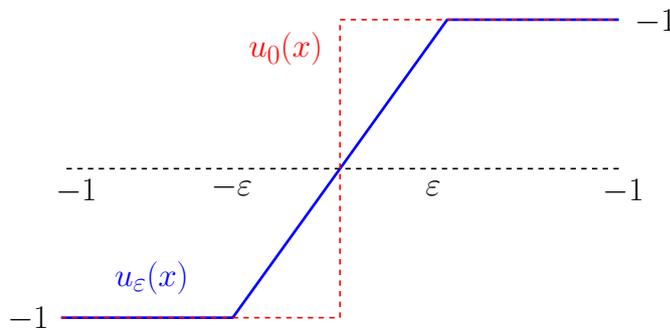


Figure 1.3: The step function u_0 approximated by a $W^{1,1}$ function u_ε .

On the other hand, for $p = 1$ we observe a different result. The space of functions of bounded variation denoted by $BV(\Omega)$, where the gradient is identified by a measure and not a function, is suitable to model edges, see Section 2.3. Indeed, as we will encounter later $W^{1,1}(\Omega) \subset BV(\Omega)$ and in particular $\|\nabla u_\varepsilon\|_{L^1} = 2$. The *variation* from the -1 to 1 value is bounded and equal to 2 and (u_ε) converge to $u_0 \in BV(\Omega) \setminus W^{1,1}(\Omega)$ in some topology. Therefore, it is sufficient to consider image-function as elements of the BV space, a space that was initially introduced in image processing by Rudin et. al having an enormous advantage on preserving the edges of images.

The main goal of this thesis is to propose novel variational imaging models, where the BV space is our guideline towards imaging models with new capabilities for image reconstruction.

1.2 Inverse Problems: Variational approach

The image processing problems that we encounter in this dissertation can be written as linear *inverse problems*. Suppose that we have acquired a data f from an image acquisition device. Usually, due to possible defects of the imaging system the data f contains several

undesirable artifacts such as noise and blur. In some situations e.g. in tomography, the data f constitutes a transformation of the image for instance a Fourier or Radon transform. Our goal is to find an image u that is visually close to the underlying real image without these kind of artifacts. Equivalently, we solve the following *inverse problem* with respect to u

$$f = Ru + \eta \quad (1.2)$$

where R is a linear operator also known as *forward operator*. In practice, finding u is a quite difficult task since R is mostly not invertible and there is an additional degradation to our data, creating an *ill-posed* problem in the sense of Hadamard. To be more precise, let us assume that f satisfies the equation above where η is the noisy component that follows a well-known distribution e.g. additive *Gaussian* noise. Certainly, there are several applications that use different types of noise, such as Poisson or impulse noise. In particular, Poisson noise or *photon noise* will be thoroughly examined in Chapter 3, since it is related to positron emission tomography. The operator R can be the identity operator for image denoising tasks, a convolution (blurring) operator for image deblurring or the Radon transform for tomography related problems. For a Gaussian distribution η , one way of rephrasing the problem is as a *least-squares* problem seeking u which solves

$$\min_{u \in X} \frac{1}{2} \int_{\Omega} |f - Ru|^2 dx, \quad (1.3)$$

analogously finding u which minimises the noise η in the L^2 sense. The first-order optimality condition of this problem is $R^*Ru = R^*f$ which is an ill-posed problem in general. The matrix R^*R is not necessarily invertible and even when it is invertible, its inverse is difficult to compute numerically causing numerical instability. The existence or even the uniqueness of the solutions are not guaranteed. One way to overcome this problem is to apply a common procedure known as *regularisation* and add some a priori information to the model. Now, our model can be formed with the following variational expression:

$$\min_{u \in X} \frac{1}{2} \int_{\Omega} |f - Ru|^2 dx + \Psi(u) \quad (1.4)$$

and consists of the least squares term, which in general could be any distance function $H(Ru, f)$, namely the *data fidelity* or *fitting term* that measures the distance between the data f and the reconstruction u under the forward operator R . The $\Psi(u)$ term is called *regulariser* and is responsible to provide additional information or regularity to the final solution u . These two terms are typically weighted by a positive parameter, a *regularising parameter* that balances the strength between the fidelity and the regulariser. The Banach space X , where the minimisation takes place, is chosen for the above problem to be well-defined and usually depends on the Ψ term. Finally, a successful or unsuccessful reconstruction or solution of the above problem depends on the choice of these two terms

as well as the choice of the regularising parameter. As we will see in the main chapters of this thesis, the type of the fidelity term depends on the noise distribution that our model demands. For instance, we choose the squared L^2 norm i.e., $H(Ru, f) = \|Ru - f\|_{L^2(\Omega)}^2$ for an image corrupted with additive Gaussian noise and for the case of impulse noise the L^1 norm is the appropriate choice for our fidelity. Furthermore, for the Poisson (photon) noise we use the *Kullback-Leibler* divergence defined as $H(Ru, f) = \int_{\Omega} Ru - f \log Ru \, d\sigma$ or the weighted- L^2 norm $H(Ru, f) = \int_{\Omega} \frac{(f - Ru)^2}{f} \, d\sigma$. There are many research papers that analyse these types of noise and we refer the reader to some of them such as [BLZ08, Nik04, CE05, LCA07, BBS⁺09, BLZ08, BBS⁺09, CDIRS14, CCDIR⁺15].

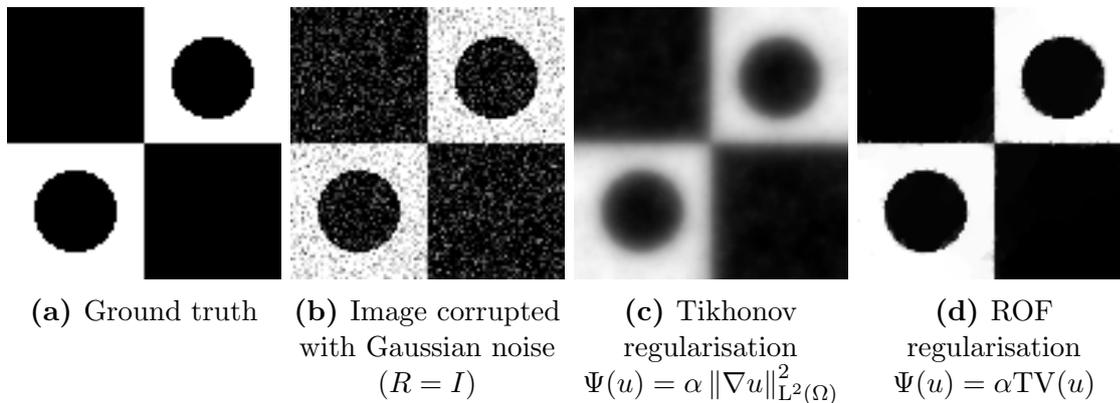


Figure 1.4: Illustration of how different regularisers behave towards preservation of edges.

While the choice of the fidelity is enforced by the particular type of noise, the choice of the regulariser is a personal selection aimed to preserve significant properties of the solution. For instance, if one chooses a regulariser as $\Psi(u) = \|\nabla u\|_{L^2(\Omega)}^2$ and solve the so-called *Tikhonov regularisation* minimisation problem, [TA77],

$$\min_{u \in W^{1,2}(\Omega)} \frac{1}{2} \|Ru - f\|_{L^2(\Omega)}^2 + \alpha \|\nabla u\|_{L^2(\Omega)}^2, \quad (1.5)$$

then will observe that although the noise is eliminated, edges are lost since we have already demonstrated that $W^{1,2}$ is not a suitable space to preserve them, see Figure 1.4. On the other hand, as it was suggested in [ROF92], if one selects the total variation of a function $u \in L^1(\Omega)$ i.e.,

$$\Psi(u) = \text{TV}(u) = \sup \left\{ \int_{\Omega} u \operatorname{div} \phi : \phi \in C_c^1(\Omega), \|\phi\|_{\infty} \leq 1 \right\} \quad (1.6)$$

then reconstruction under a proper choice of α , will preserve accurately all the edges in the image. In this case, $X = \text{BV}(\Omega)$ and for the denoising task we use the identity operator,

i.e., $R = I$ under Gaussian noise where we recover

$$\min_{u \in \text{BV}(\Omega)} \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2 + \alpha \text{TV}(u) \quad (1.7)$$

the well-known Rudin, Osher and Fatemi (ROF) model.

1.3 Contribution

This thesis consists of two main contributions as a result of two research papers [BMPS14, BPPS15a] produced during my Ph.D.

Chapter 3: Total variation regularisation in image and sinogram space for PET reconstruction

We focus on reconstructing images related to positron emission tomography (PET). We propose a combined TV regularisation in the image and sinogram space. For a given PET data f , it is expressed as the following variational minimisation problem:

$$\operatorname{argmin}_{u \in \text{BV}(\mathbb{R}^2)} \left\{ \alpha |Du|(\mathbb{R}^2) + \beta |D(\mathcal{R}u)|(\Sigma^2) + \frac{1}{2} \int_{\Sigma^2} \frac{(f - \mathcal{R}u)^2}{f} d\sigma \right\}. \quad (1.8)$$

The first two terms act as total variation penalisation on the image space \mathbb{R}^2 and the sinogram space Σ^2 weighted by positive parameters α, β . The fidelity term is a weighted-L² norm, an approximation to the Kullback-Leibler divergence that is suitable for denoising images corrupted with Poisson noise, see also [SBMB09, LCA07]. We prove the well-posedness of the proposed model and continuity properties of the Radon transform \mathcal{R} in the $\text{BV}(\Sigma^2)$ for images with compact support. The theoretical analysis also explores stability issues regarding a small perturbation on the sinogram data as well as a quantitative error analysis through the Bregman distance. In order to have a deeper understanding on how sinogram regularisation affects the final object reconstruction, we choose to examine first a regularisation only on the sinogram space i.e., when $\alpha = 0$. We provide an analytical computation for a sinogram image obtained from the characteristic function of a ball with radius r , i.e., $u = \mathcal{X}_{B_r}$ and determine how the β parameter affects the final reconstruction, presented in Figure 1.5. We use the Abel transform in order to recover an analytical expression of u which is

$$u(\tilde{r}) = \frac{\delta}{\pi \sqrt{r^2 - \tilde{r}^2}},$$

where δ is the corresponding loss of contrast in the image derived from TV regularisation on the sinogram space.

We observe that as we increase β the regularised solution starts to lose details from the inner part of the circle concentrating on the boundaries of the object. This motivates the

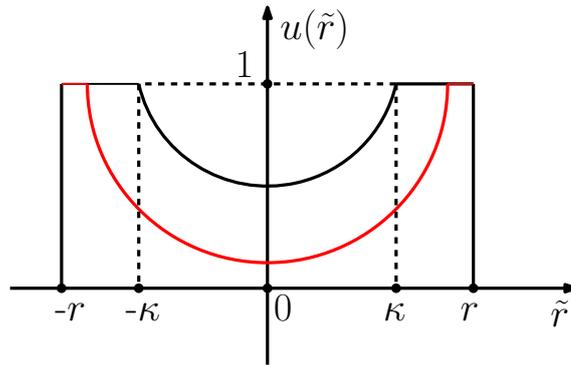


Figure 1.5: TV regularisation on the sinogram of a ball with radius r . We set $\alpha = 0$ in (1.8). The black and the red curve constitute the regularised solution expressed by the Abel inversion formula for a smaller and larger value of β respectively. The larger β the more the solution concentrates around the boundaries of the ball.

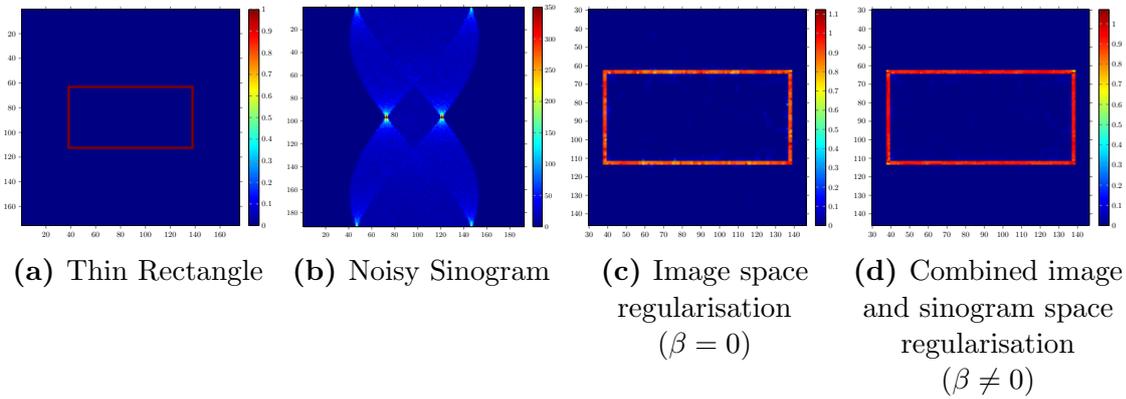


Figure 1.6: Thin rectangle: Illustration of the combined image and sinogram space regularisation.

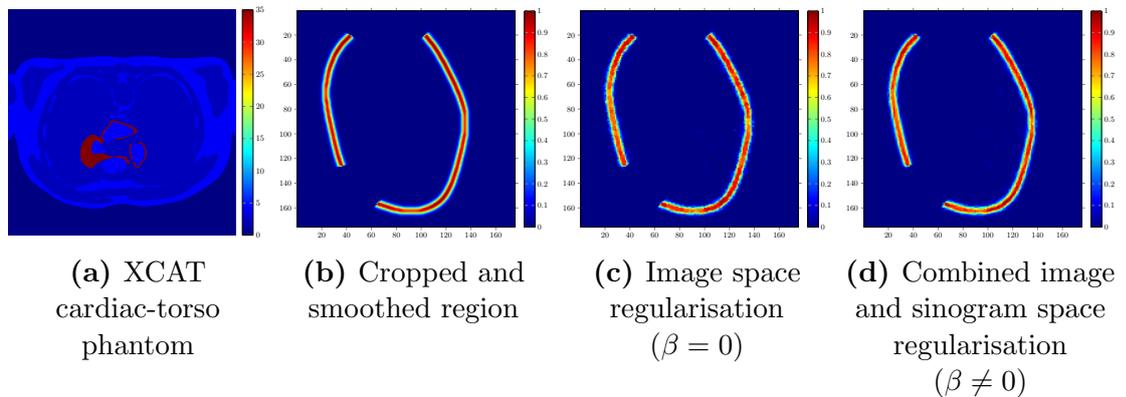


Figure 1.7: XCAT: Illustration of the combined image and sinogram space regularisation.

examination of (1.8) for reconstructing structures as in Figures 1.6, 1.7 where we see that the extra penalisation on the sinogram space has a significant improvement compared to TV regularisation in image space only, i.e., $\beta = 0$.

Chapter 4: Infimal convolution regularisation functionals of BV and L^p spaces

We begin with a literature review on high-order TV based variational methods such as the infimal convolution (ICTV) regularisation proposed in [CL97] and the second order total generalised variation (TGV²) introduced in [BKP10]. The main purpose of high-order regularisation is to reduce and eliminate a well known artifact that first-order regularisation promotes, namely the *staircasing* effect. For instance, a ROF regularised solution (1.7) behaves as a piecewise constant approximation to the noisy data f due to the TV term with a small L^2 norm. This is illustrated in Figure 1.8, where the staircasing term is clearly justified. Notice that in two dimensional images this can be interpreted as blocky-like artifacts, see for example Figures 1.11c and 1.12c.

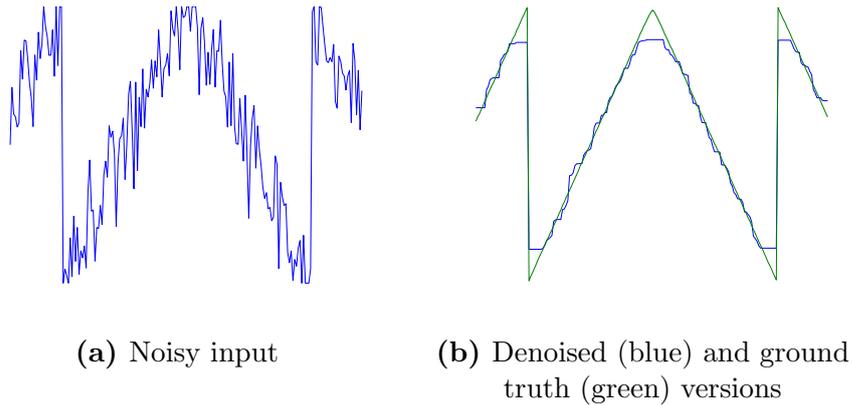


Figure 1.8: Illustration of the staircasing effect for the ROF model.

However, as we discuss in Chapter 4 not only high-order regularisers are capable of eliminating the staircasing effect. We propose a family of first-order infimal convolution regularisation between BV and L^p spaces with $1 < p \leq \infty$ defined as

$$\mathrm{TVL}_{\alpha,\beta}^p(u) := \inf_{w \in L^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)}$$

which are mainly used for image denoising tasks under Gaussian noise. We discuss the analytical properties of

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \mathrm{TVL}_{\alpha,\beta}^p(u)$$

and focus on the structure of solutions that are produced for different cases of f and

p and different range of regularising parameters α and β . For example, if $\beta \geq \alpha|\Omega|^{\frac{1}{q}}$, where q if the conjugate exponent of p , we have an equivalent ROF solution and for the case of $p = 2$ our variational model coincides with the Huber-TV regularisation. In order to understand the structure of the corresponding solutions, we begin our analysis with the one dimensional version of our model and derive necessary and sufficient conditions (*optimality* conditions) that the optimal $(w, u) \in L^p \times \text{BV}$ pair must satisfy. We recover also the optimality conditions for the Huber-TV regularisation approach.

We present the analytical solutions for f being a step function as well as for a piecewise affine function f in order to illustrate how the geometry of the solution differentiates in terms of p . In addition, we realise that for finite values of p a piecewise smooth solution is enforced whereas for $p = \infty$ we obtain piecewise affine structures which is an innovative result regarding first-order methods, see Figure 1.9. Let us emphasise the fact that our proposed regulariser can preserve the exact geometry of the initial data compared to a high-order regulariser, e.g. TGV², see Figure 1.10.

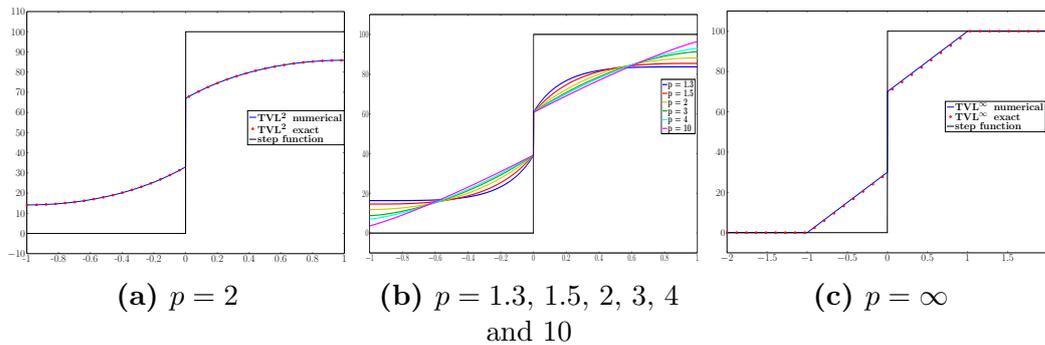


Figure 1.9: Step function: Structures of solutions for different values of $p \in (1, \infty]$.

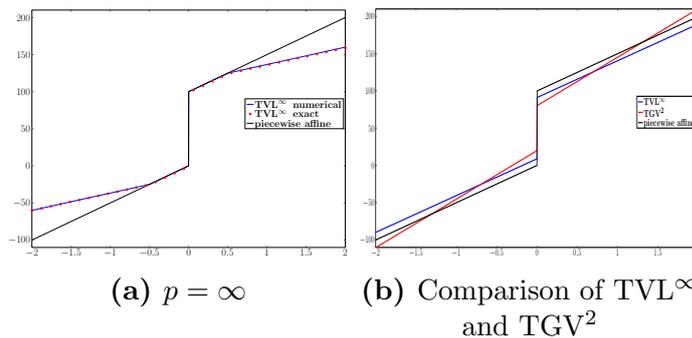


Figure 1.10: Piecewise affine function: Reconstruction for $p = \infty$ and comparison with TGV². Geometry of the data is preserved with the proposed model.

We proceed with the two dimensional numerical experiments where we clearly demonstrate that one can achieve equivalent visual results to high-order regularisers without any

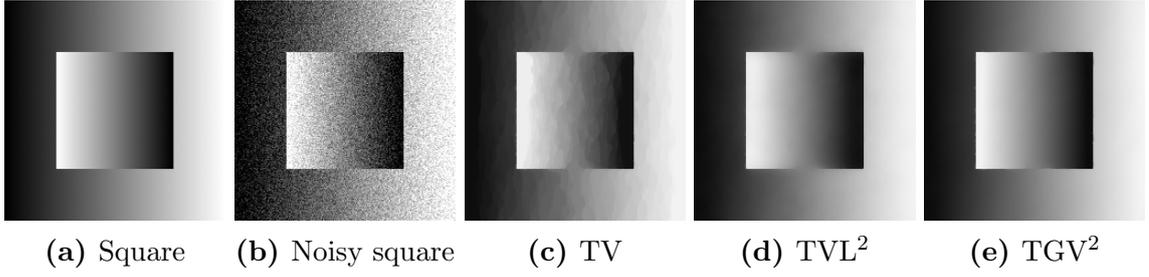


Figure 1.11: Square: TV, TVL² and TGV² reconstructions. Staircasing can be eliminated with first-order methods.

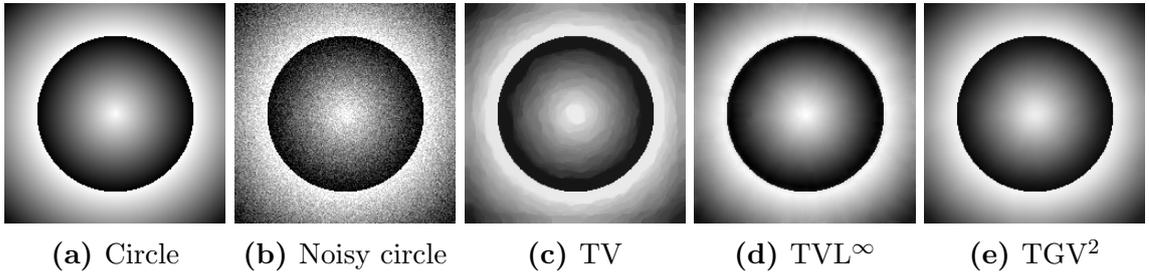


Figure 1.12: Circle: TV, TVL[∞] and TGV² reconstructions. Spike-like structures around the origin are better preserved with the proposed method.

staircasing just by using first-order regularisers and finite values of p , see Figure 1.11. Furthermore, we demonstrate that our regulariser is more suitable than TGV² in order to recover *spike-like* structures in the image, see Figure 1.12. However, there is a significant disadvantage for the structure of the solution u when we deal with a scalar β . For the $p = \infty$, we prove that $|w| = \|w\|_\infty$ which is equivalent to say that the gradient of the solution behaves as piecewise constant function. This is quite restrictive and introduces similar staircasing artifacts but with different geometry e.g. *affine staircasing*, see Figure 1.13d on the background of the parrot.

One remedy to this problem is to introduce the *weighted-TV*L[∞] version where β is now a matrix with a priori information on the gradient of a *smoothed* image. It is written as

$$\min_{\substack{u \in \text{BV}(\Omega) \\ w \in \text{L}^\infty(\Omega)}} \frac{1}{2} \|f - u\|_2^2 + \alpha \|\nabla u - w\|_1 + \|\beta \circ w\|_\infty$$

where \circ denotes the Hadamard product and

$$\beta := \frac{c}{|\nabla f_\sigma| + \varepsilon}$$

with f_σ be a smoothed version of the noisy image and ε a small positive constant. Ideally, we would like β to imitate the gradient of the ground truth image meaning that is zero

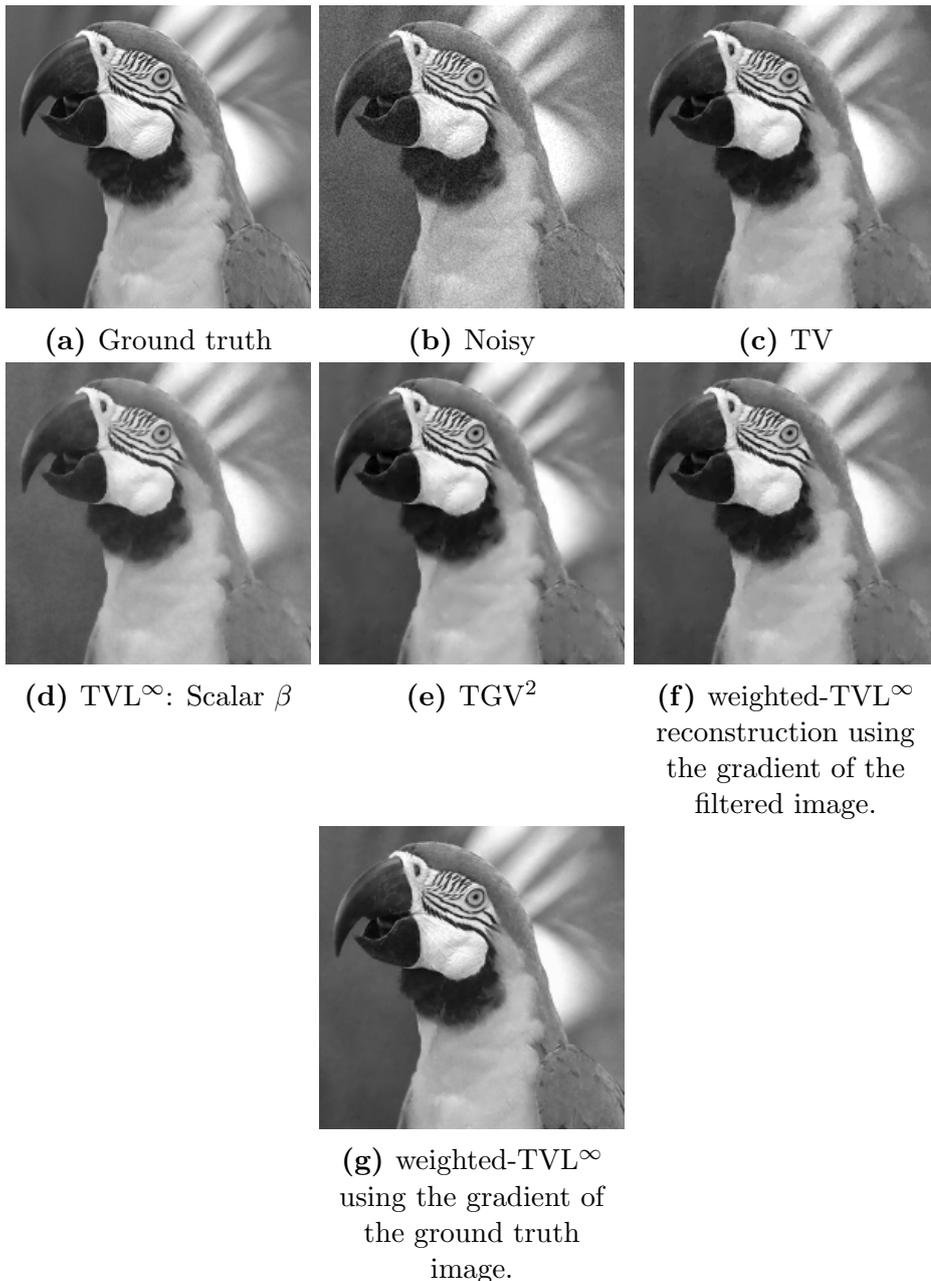


Figure 1.13: Parrot: Illustration of TVL^∞ reconstructions compared with TGV^2 . Scalar β introduces affine staircasing artifacts whereas matrix β computed via the gradient of the filtered image is visually close to TGV^2 . It still remains an open question on how we can accurately choose the matrix β and obtain almost perfect reconstructions.

near the edges and w will vanish while preserving the edges. Moreover, a suitable choice of the positive constant c provides us with an approximation to the correct slope for other regions in the image. This approach provides enough freedom to the TVL^∞ regularised solution and achieve comparable results with high-order regularisers also for real-world

images. It still remains an open question on how we can select the matrix β properly, since computing β based on the gradient of the ground truth image, an image we do not have a priori, results to almost perfect quality reconstructions, see Figure 1.13g.

Chapter 2

Mathematical Preliminaries

The scope of this chapter is twofold. Firstly, we would like to provide all the necessary mathematical background used in this thesis. We focus mainly on aspects of *measure theory*, *functional* and *convex* analysis. The reader is advised to focus specifically to the notions of Radon measures and functions of bounded variation, on duality theory and subdifferential calculus and finally recall some of the fundamental properties and embedding theorems for Lebesgue and Sobolev spaces. Secondly, we would like to help the reader familiarise with the notation that is introduced and repeatedly used throughout this thesis.

We usually denote by X an open subset of \mathbb{R}^d with $d \geq 1$. We denote by \mathbb{R} (resp. \mathbb{R}^+) the space of real numbers (resp. the space of real positive numbers) and \mathbb{R}^d the Euclidean d -dimensional space. The extended version of the real line is denoted by $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. We recall some standard definitions on the space of continuous functions. The space of k -times continuously differentiable \mathbb{R}^ℓ -valued functions is denoted by $\mathcal{C}^k(X, \mathbb{R}^\ell)$ and if $k = \infty$ we call this the space of *smooth* functions denoted by $\mathcal{C}^\infty(X, \mathbb{R}^\ell)$. If $k = 0$, we mean the space of continuous functions and write $\mathcal{C}(X, \mathbb{R}^\ell)$. If $\ell = 1$, we follow the notation $\mathcal{C}^k(X)$ and $\mathcal{C}^\infty(X)$ instead of $\mathcal{C}^k(X, \mathbb{R})$ and $\mathcal{C}^\infty(X, \mathbb{R})$. Moreover, we define the space of k -times continuously differentiable \mathbb{R}^ℓ -functions with compact support on X denoted by $\mathcal{C}_c^k(X, \mathbb{R}^\ell)$. If $k = \infty$, we write $\mathcal{C}_c^\infty(X, \mathbb{R}^\ell)$ as the smooth \mathbb{R}^ℓ -valued functions with compact support which are sometimes found in literature as the *test functions* denoted by $\mathcal{D}(X, \mathbb{R}^\ell)$. Analogously, if $\ell = 1$ we write $\mathcal{C}_c(X)$, $\mathcal{C}_c^k(X)$ or $\mathcal{C}_c^\infty(X)$ instead of $\mathcal{C}_c(X, \mathbb{R})$, $\mathcal{C}_c^k(X, \mathbb{R})$ and $\mathcal{C}_c^\infty(X, \mathbb{R})$ respectively. The closure of $\mathcal{C}_c(X, \mathbb{R}^\ell)$ under the supremum norm $\|u\|_\infty := \sup_{x \in X} |u(x)|$ describe the \mathbb{R}^ℓ -valued functions that vanish at infinity denoted by $\mathcal{C}_0^k(X, \mathbb{R}^\ell)$ and if $\ell = 1$ we simply write $\mathcal{C}_0^k(X)$.

As a general notation rule for this thesis, when there is no possibility of confusion and there is enough information from the context, we usually omit the range of functions.

2.1 Measure Theory

This section covers the basic notions of measure theory and the key ingredient is the *Radon measure*. We mainly follow [AFP00], [Rud87] although the reader is referred to classical textbooks such as [EG92] and [FL07]. In order to have a smooth transition to the theory presented below, we begin with the standard definitions of the σ -algebra and the measure.

Definition 2.1.1 (σ -algebra). Let X be a nonempty set and \mathfrak{C} be a collection of subsets of X . We call \mathfrak{C} a σ -algebra if it satisfies the following properties :

- (i) $X \in \mathfrak{C}$
- (ii) if $A \in \mathfrak{C}$, then $X \setminus A \in \mathfrak{C}$,
- (iii) if $A_n \in \mathfrak{C}$ for every $n \in \mathbb{N}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{C}$

We call the pair (X, \mathfrak{C}) a *measure space*. If X is a topological space, we define the smallest σ -algebra containing all open subsets of X by $\mathcal{B}(X)$ i.e., the *Borel σ -algebra* generated by open subsets of X .

Definition 2.1.2 (Measures). Let X be a nonempty set and \mathfrak{C} be a σ -algebra with $\ell \in \mathbb{N}$. We say that a map $\mu : \mathfrak{C} \rightarrow \mathbb{R}^{\ell}$ is a *measure* if

$$\mu(\emptyset) = 0 \quad \text{and} \quad \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

for any sequence A_n of pairwise disjoint elements of \mathfrak{C} . We call μ a *real measure* if $\ell = 1$ and a *vector measure* if $\ell > 1$ with $\mu = (\mu_1, \dots, \mu_{\ell})$ where $\mu_i : \mathfrak{C} \rightarrow \mathbb{R}$ for every $i = 1, \dots, \ell$.

Notice, that Definition 2.1.2 applies also to $\mu : \mathfrak{C} \rightarrow [0, \infty]$ or $\mu : \mathfrak{C} \rightarrow \overline{\mathbb{R}}$ and in these cases μ is called *positive measure* and *signed measure* respectively. If there exists a sequence $A_n \in \mathfrak{C}$ such that $\bigcup_{n=1}^{\infty} A_n = X$ and $\mu(A_n) < \infty$ for every n , then μ is called σ -finite. If $\mu(X) < \infty$, the measure μ is called *finite*.

Definition 2.1.3 (Total variation measure). Let μ be a measure on (X, \mathfrak{C}) measure space. We define its total variation measure $|\mu| : \mathfrak{C} \rightarrow \mathbb{R}^+$ as

$$|\mu|(A) := \sup \left\{ \sum_{i=1}^{\infty} |\mu(A_n)| : A_n \in \mathfrak{C}, \text{ pairwise disjoint with } \bigcup_{n=1}^{\infty} A_n = A \right\}.$$

The difference between positive and real or vector measures is that in the latter case, it can be shown that any real or vector measure must be finite. Notice that this is not true for signed measures. This is a direct consequence of the fact that the total variation

of measure μ is a finite positive measure. We proceed with two essential properties of measures that are revealed in the following definition.

Definition 2.1.4. (i) Let μ be a positive and ν a real or vector measure on measure space (X, \mathfrak{C}) . We say that ν is *absolutely continuous* with respect to μ and write $\nu \ll \mu$ if for every $A \in \mathfrak{C}$

$$\mu(A) = 0 \Rightarrow |\nu|(A) = 0.$$

(ii) Let μ, ν positive measures (resp. real or vector valued) we say that they are *mutually singular* and write $\nu \perp \mu$ if there exists $A \in \mathfrak{C}$ such that $\mu(A) = 0$ (resp. $|\mu|(A) = 0$) and $\nu(X \setminus A) = 0$ (resp. $|\nu|(X \setminus A) = 0$).

Remark 2.1.5. An important result regarding the notion of mutually singular measures is the following: if μ_1 and μ_2 are two positive mutually singular measures then

$$|\mu_1 + \mu_2| = |\mu_1| + |\mu_2|.$$

Let $\mu := \mu_1 + \mu_2$, then obviously

$$|\mu_1 + \mu_2|(X) \leq |\mu_1|(X) + |\mu_2|(X).$$

For the other direction, since $\mu_1 \perp \mu_2$ there exist two disjoint sets $A_1, A_2 \in \mathfrak{C}$ with $A_1 \cup A_2 = X$ such that $\mu_1(E) = \mu_1(A_1 \cap E) = \mu(A_1 \cap E)$ and $\mu_2(E) = \mu_2(A_2 \cap E) = \mu(A_2 \cap E)$ for any $E \in \mathfrak{C}$. If $\{E_i\}$ and $\{F_i\}$ are two partitions of X , i.e., pairwise disjoint whose union form X and

$$\sum_{i=1}^{\infty} |\mu_1(E_i)| + \sum_{i=1}^{\infty} |\mu_2(F_i)| = \sum_{i=1}^{\infty} |\mu(E_i \cap A_1)| + \sum_{i=1}^{\infty} |\mu(F_i \cap A_2)| \leq |\mu|(X),$$

since $\{E_i \cap A_1\}_{i=1}^{\infty}$ and $\{F_i \cap A_2\}_{i=1}^{\infty}$ are a partition of X . Taking the supremum over the corresponding partitions, we conclude that $|\mu_1|(X) + |\mu_2|(X) \leq |\mu|(X) = |\mu_1 + \mu_2|(X)$.

Let $f \in L^1(X, \mathbb{R}^\ell; \mu)$ (see Section 2.2) and μ a positive measure on the measure space (X, \mathfrak{C}) . We define an \mathbb{R}^ℓ -valued measure and its corresponding total variation as

$$f\mu(A) = \int_A f d\mu, \quad |f\mu|(A) = \int_A |f| d\mu, \quad \forall A \in \mathfrak{C}.$$

Clearly, $f\mu \ll \mu$ and for every measurable function g , we write

$$\int_X g d(f\mu) = \int_X gf d\mu.$$

The next fundamental theorem of measure theory allows to decompose a measure into an absolutely continuous and singular part.

Theorem 2.1.6 (Lebesgue decomposition). *Let (X, \mathfrak{C}) a measure space and μ, ν be two measures as in Definition 2.1.4(i) with a σ -finite μ . Then, there exists a unique pair of \mathbb{R}^ℓ -valued measures ν^a and ν^s such that*

$$\nu = \nu^a + \nu^s, \quad \nu^a \ll \mu \quad \text{and} \quad \nu^s \perp \mu$$

Furthermore, there exists a function $f \in L^1(X, \mathbb{R}^\ell; \mu)$ such that $\nu^a = f\mu$. The function f is called the density of ν with respect to μ and is denoted by $\frac{\nu}{\mu}$. It is often found in literature as Radon-Nikodým derivative of ν with respect to μ and is denoted by $\frac{d\nu}{d\mu}$.

It is immediate that $\mu \ll |\mu|$ for a real or vector measure μ . Hence, by the previous theorem we obtain the following decomposition.

Theorem 2.1.7 (Polar decomposition). *Let μ be an \mathbb{R}^ℓ -valued measure on (X, \mathfrak{C}) measure space, then there exists a unique $\mathcal{S}^{\ell-1}$ -valued function $f \in L^1(X, \mathbb{R}^\ell; |\mu|)$ such that $\mu = f|\mu|$. We naturally denote f as $\text{sgn}(\mu)$ and $|f| = 1$, $|\mu|$ -almost everywhere.*

For the rest of this section we assume that X is a locally compact and separable metric space abbreviated as *l.c.s* and $\mathfrak{C} = \mathcal{B}(X)$.

Definition 2.1.8 (Radon measures). Let $(X, \mathcal{B}(X))$ a measure space, where X is a l.c.s metric space.

- (i) A positive measure on $(X, \mathcal{B}(X))$ is called *Borel* measure. Furthermore, if it is finite on the compact sets then it is called *positive Radon measure*.
- (ii) A set function defined on relatively compact Borel subsets of X that is a measure on $(K, \mathcal{B}(K))$ for every compact set $K \subset X$ is called a *real or vector Radon measure* on X .

If $\mu : \mathcal{B}(X) \rightarrow \mathbb{R}^\ell$ is a measure according to Definition 2.1.2, then we say that μ is a *finite Radon measure*. The space of \mathbb{R}^ℓ -valued Radon measures is denoted by $\mathcal{M}_{loc}(X, \mathbb{R}^\ell)$ and the finite \mathbb{R}^ℓ -valued Radon measures as $\mathcal{M}(X, \mathbb{R}^\ell)$. We let $\mathcal{M}_{loc}(X)$ and $\mathcal{M}(X)$ respectively for $\ell = 1$. We identify the space $\mathcal{M}(X, \mathbb{R}^\ell)$ as a Banach space under the norm $\|\mu\|_{\mathcal{M}(X, \mathbb{R}^\ell)} := |\mu|(X)$. The following theorem introduces another formula of computing the total variation measure.

Theorem 2.1.9. *Let X be an l.c.s. metric space and $\mu \in \mathcal{M}(X, \mathbb{R}^\ell)$. Then, for every open $A \subset X$, the total variation measure is:*

$$|\mu|(A) = \sup \left\{ \sum_{i=1}^{\ell} \int_X u_i d\mu_i : u \in \mathcal{C}_c(A, \mathbb{R}^\ell), \|u\|_\infty \leq 1 \right\}.$$

We continue with a standard result which indicates that $\mathcal{M}(X, \mathbb{R}^\ell)$ can be identified as the dual space of $(\mathcal{C}_0(X, \mathbb{R}^\ell), \|\cdot\|_\infty)$ under the pairing $\langle u, \mu \rangle := \sum_{i=1}^{\ell} \int_X u_i d\mu_i$.

Theorem 2.1.10 (Riesz representation theorem). *Let X be a l.c.s. metric space and L a bounded linear functional on $(\mathcal{C}_0(X, \mathbb{R}^\ell), \|\cdot\|_\infty)$. Then there exists a unique $\mu \in \mathcal{M}(X, \mathbb{R}^\ell)$ such that*

$$L(u) = \sum_{i=1}^{\ell} \int_X u_i d\mu_i, \quad \forall u \in \mathcal{C}_0(X, \mathbb{R}^\ell)$$

i.e., a bounded linear functional is characterised by a Radon measure μ and

$$\|L\| = |\mu|(X),$$

where $\|\cdot\|$ denotes the operator norm in the dual space $\mathcal{C}_0(X, \mathbb{R}^\ell)$.

Notice that the above theorem is clearly valid when $L : \mathcal{C}_c(X, \mathbb{R}^\ell) \rightarrow \mathbb{R}$ is a bounded linear functional and in this case we can identify the dual space of $\mathcal{C}_c(X, \mathbb{R}^\ell)$ as $\mathcal{M}_{loc}(X, \mathbb{R}^\ell)$. In addition, an equivalent expression of the Radon norm derived from Theorem 2.1.10 is

$$\|\mu\|_{\mathcal{M}} = \sup \left\{ \langle u, \mu \rangle : u \in \mathcal{C}_0(X, \mathbb{R}^\ell), \|u\|_\infty \leq 1 \right\}.$$

Using the density of the *test functions*, i.e., $u \in \mathcal{C}_c^\infty(X, \mathbb{R}^\ell)$ dense in $\mathcal{C}_0(X, \mathbb{R}^\ell)$ under the supremum norm, we have that μ can be regarded as a distribution, i.e., a bounded and linear functional $L : \mathcal{C}_c^\infty(X, \mathbb{R}^\ell) \rightarrow \mathbb{R}$ provided that the supremum of $\langle u, \mu \rangle$ tested against all $\mathcal{C}_c^\infty(X, \mathbb{R}^\ell)$ with $\|u\|_\infty \leq 1$ is finite. This result will be useful in Section 4.4.1. We conclude this section with the notion of convergence properties with respect to Radon measures that will be useful when we deal with functions of bounded variation, see Section 2.3.

Definition 2.1.11 (Weak*-convergence). Let $\mu, \mu_n \in \mathcal{M}(X, \mathbb{R}^\ell)$ (resp. $\mathcal{M}_{loc}(X, \mathbb{R}^\ell)$). We say that μ_n (resp. locally) weakly* converges to μ if

$$\lim_{n \rightarrow \infty} \int_X u d\mu_n = \int_X u d\mu, \quad \text{where} \quad \int_X u d\mu = \left(\int_X u d\mu_1, \dots, \int_X u d\mu_\ell \right)$$

for every $u \in \mathcal{C}_0(X)$ (resp. $u \in \mathcal{C}_c(X)$).

Theorem 2.1.12 (Weak*-compactness for measures). *Let $\mu_n \in \mathcal{M}(X, \mathbb{R}^\ell)$, a sequence of finite Radon measures with $\sup \{|\mu_n|(X) : n \in \mathbb{N}\} < \infty$, then there is a weakly* converging subsequence. In addition, the map $\mu \rightarrow |\mu|(X)$ is lower semicontinuous with respect to the weak* convergence i.e., $|\mu|(X) \leq \liminf_{n \rightarrow \infty} |\mu_n|(X)$, as a supremum of weakly* continuous functionals.*

2.2 Lebesgue and Sobolev spaces

This section presents a brief summary of Lebesgue spaces denoted by L^p and Sobolev spaces denoted by $W^{1,p}$ that will be used in the following chapters of this thesis. We primarily focus on inclusion and convergence properties for L^p spaces and main embedding theorems for $W^{1,p}$ spaces. We refer the reader to some classical textbooks related to both kind of spaces [Bre11], [AF03], [Eva10], [Leo09] and [LL01].

2.2.1 L^p spaces

In the following, we let X be a l.c.s metric space and μ a positive measure.

Definition 2.2.1. Let (X, \mathfrak{C}, μ) be a measure space with $1 \leq p \leq \infty$. Then, we define

(i) $L^p(X, \mathbb{R}^\ell; \mu) := \left\{ u : X \rightarrow \mathbb{R}^\ell : u \text{ is } \mu\text{-measurable and } \|u\|_{L^p(X, \mathbb{R}^\ell; \mu)} < \infty \right\}$, where

$$\|u\|_{L^p(X, \mathbb{R}^\ell; \mu)} := \left(\int_X |u|^p d\mu \right)^{\frac{1}{p}}, \quad \text{for } 1 \leq p < \infty.$$

(ii) $L^\infty(X, \mathbb{R}^\ell; \mu) := \left\{ u : X \rightarrow \mathbb{R}^\ell : u \text{ is } \mu\text{-measurable and } \|u\|_{L^\infty(X, \mathbb{R}^\ell; \mu)} < \infty \right\}$, where

$$\|u\|_{L^\infty(X, \mathbb{R}^\ell; \mu)} = \operatorname{ess\,sup}_{u \in X} |u| := \inf \{ C > 0 : \mu \{ x \in X : |u(x)| > C \} = 0 \}$$

i.e., the *essential supremum* of $|u|$.

It is known that the $L^p(X, \mathbb{R}^\ell; \mu)$ are Banach spaces under the norm $\|\cdot\|_{L^p(X, \mathbb{R}^\ell; \mu)}$ for $1 \leq p \leq \infty$. Notice that we assign in $L^p(X, \mathbb{R}^\ell; \mu)$ spaces an *equivalence relation* which transforms the initial seminorm $\|\cdot\|_{L^p(X, \mathbb{R}^\ell; \mu)}$ into a norm. For simplicity of notation we use the following rule: if μ is the Lebesgue measure \mathcal{L}^d , where X is now a subset of \mathbb{R}^d , we use the notation $L^p(X, \mathbb{R}^\ell)$ instead of $L^p(X, \mathbb{R}^\ell; \mathcal{L}^d)$ with $d \geq 1$. Moreover, if it is clear from the context and there is no possibility of confusion about the range of functions and the corresponding measure μ , we simply write $L^p(X)$. The following two propositions are based on Hölder's inequality and provide a relation between different L^p spaces and their norms.

Proposition 2.2.2 (L^p inclusions). *If $\mu(X) < \infty$, $1 \leq p_1 < p_2 \leq \infty$ and $u \in L^{p_2}(X, \mathbb{R}^\ell; \mu)$.*

Then

$$\|u\|_{L^{p_1}(X, \mathbb{R}^\ell; \mu)} \leq \|u\|_{L^{p_2}(X, \mathbb{R}^\ell; \mu)} \cdot \begin{cases} |\mu(X)|^{\frac{1}{p_1} - \frac{1}{p_2}}, & p_2 < \infty \\ |\mu(X)|^{\frac{1}{p_1}}, & p_2 = \infty \end{cases}$$

and $u \in L^{p_1}(X, \mathbb{R}^\ell; \mu)$, i.e., $L^{p_2}(X, \mathbb{R}^\ell; \mu) \hookrightarrow L^{p_1}(X, \mathbb{R}^\ell; \mu)$ continuously.

Let q be the Hölder conjugate exponent of p i.e., p, q satisfy $\frac{1}{p} + \frac{1}{q} = 1$ and

$$q := \begin{cases} \frac{p}{p-1}, & 1 < p < \infty, \\ \infty, & p = 1, \\ 1, & p = \infty. \end{cases}$$

The following proposition can be found in [Leo09, Appendix B.7].

Proposition 2.2.3. *Let (X, \mathfrak{C}, μ) be a measure space, μ is σ -finite and q the Hölder exponent of p with $1 \leq p \leq \infty$. Then, for any measurable function $u : X \rightarrow \mathbb{R}^\ell$,*

$$\|u\|_{L^p(X, \mathbb{R}^\ell; \mu)} = \sup \left\{ \int_X |uv| d\mu : v \in L^q(X, \mathbb{R}^\ell; \mu), \|v\|_{L^q(X, \mathbb{R}^\ell; \mu)} \leq 1 \right\}.$$

We now state the Riesz representation theorem in L^p spaces.

Theorem 2.2.4 (Riesz representation in L^p). *Let (X, \mathfrak{C}, μ) be a measure space with $1 \leq p < \infty$ and L a bounded linear operator in $L^p(X, \mathbb{R}^\ell; \mu)$. Then, there exists a unique $v \in L^q(X, \mathbb{R}^\ell; \mu)$ such that*

$$L(u) = \int_X uv d\mu, \quad \forall u \in L^p(X, \mathbb{R}^\ell; \mu)$$

and $\|L\|_{(L^p(X, \mathbb{R}^\ell; \mu))^*} = \|v\|_{L^q(X, \mathbb{R}^\ell; \mu)}$. If $p = 1$, we require that μ is σ -finite.

We can identify the dual of $L^p(X, \mathbb{R}^\ell; \mu)$ by $L^q(X, \mathbb{R}^\ell; \mu)$ if $1 < p < \infty$ and the dual of $L^1(X, \mathbb{R}^\ell; \mu)$ by $L^\infty(X, \mathbb{R}^\ell; \mu)$. Note, that for $1 < p < \infty$ the L^p spaces are reflexive. For sake of completeness, let us mention that the dual of $L^\infty(X, \mathbb{R}^\ell; \mu)$, can be identified by the space of all *bounded finitely additive measures absolutely continuous with respect to μ* and is denoted by $\text{ba}(X, \mathfrak{C}; \mu)$. Specifically, if $L : L^\infty(X, \mathbb{R}^\ell; \mu) \rightarrow \mathbb{R}$ is a bounded, linear functional then there is a unique $\lambda \in \text{ba}(X, \mathfrak{C}; \mu)$ i.e.,

$$\begin{aligned} \lambda(A_1 \cup A_2) &= \lambda(A_1) + \lambda(A_2), \quad \forall A_1, A_2 \in \mathfrak{C} \text{ with } A_1 \cap A_2 = \emptyset, \\ \lambda(A) &= 0, \text{ if } \mu(A) = 0, \end{aligned}$$

such that

$$L(u) = \int_X u d\lambda, \quad \forall u \in L^\infty(X, \mathbb{R}^\ell; \mu).$$

The space $\text{ba}(X, \mathfrak{C}; \mu)$ is a Banach space endowed with the total variation of λ . For more details, we refer the reader to [Yos80, Chapter 4] and [FL07, Chapter 2]. Note that in general $L^1(X, \mathbb{R}^\ell; \mu) \subsetneq (L^\infty(X, \mathbb{R}^\ell; \mu))^*$. We conclude this section with some standard results related to weak convergence in L^p spaces.

Definition 2.2.5 (Weak convergence in L^p). Let (X, \mathfrak{C}, μ) be a measure space, $(u_n) \in L^p(X, \mathbb{R}^\ell; \mu)$ and q be the Hölder exponent of p .

- (i) If $1 \leq p < \infty$, we say that u_n converges weakly to $u \in L^p(X, \mathbb{R}^\ell; \mu)$, denoted by $u_n \rightharpoonup u$, if

$$\int_X u_n v \, d\mu \rightarrow \int_X u v \, d\mu, \quad \forall v \in L^q(X, \mathbb{R}^\ell; \mu)$$

- (ii) If $p = \infty$, we say that u_n converges weakly* to $u \in L^\infty(X, \mathbb{R}^\ell; \mu)$, denoted by $u_n \xrightarrow{w^*} u$, if

$$\int_X u_n v \, d\mu \rightarrow \int_X u v \, d\mu, \quad \forall v \in L^1(X, \mathbb{R}^\ell; \mu)$$

- (iii) If $u_n \rightharpoonup u$ (resp. $u_n \xrightarrow{w^*} u$) then (u_n) is bounded in $L^p(X, \mathbb{R}^d; \mu)$ and

$$\|u\|_{L^p(X, \mathbb{R}^\ell; \mu)} \leq \liminf_{n \rightarrow \infty} \|u_n\|_{L^p(X, \mathbb{R}^d, \mu)}.$$

Theorem 2.2.6 (Weak compactness in L^p). Let $(X, \mathfrak{C}; \mu)$ be a measure space with $1 < p \leq \infty$ and (u_n) bounded in $L^p(X, \mathbb{R}^\ell; \mu)$. Then, there exists a subsequence (u_{n_k}) , $u \in L^p(X, \mathbb{R}^\ell; \mu)$ such that $u_n \rightharpoonup u$ in $L^p(X, \mathbb{R}^\ell; \mu)$ for $1 < p \leq \infty$ or $u_n \xrightarrow{w^*} u$ in $L^\infty(X, \mathbb{R}^\ell; \mu)$.

Finally, we know that the $L^1(X, \mathbb{R}^\ell; \mu)$ space enjoys no compactness properties. However, as we will see later for functions of bounded variation, it is convenient to identify $L^1(X, \mathbb{R}^\ell; \mu)$ as a subspace of $\mathcal{M}_{loc}(X, \mathbb{R}^\ell)$. Consider $u \in L^1(X, \mathbb{R}^\ell)$, then the mapping $\phi \in \mathcal{C}_c(X, \mathbb{R}^\ell) \mapsto \int_X u \phi \, dx$ is continuous and linear functional on $\mathcal{C}_c(X, \mathbb{R}^\ell)$ denoted by $L(u)$, i.e.,

$$\langle Lu, \phi \rangle = \int_X u \phi \, dx, \quad \forall \phi \in \mathcal{C}_c(X, \mathbb{R}^\ell).$$

In particular, one can prove that L is an isometry, since

$$\|Lu\|_{\mathcal{M}(X, \mathbb{R}^\ell)} = \sup_{\substack{\phi \in \mathcal{C}_c(X, \mathbb{R}^\ell) \\ \|\phi\|_\infty \leq 1}} \int_X u \phi \, dx = \|u\|_{L^1(X, \mathbb{R}^\ell)}$$

and without loss of generality, we write $\|u\|_{\mathcal{M}(X, \mathbb{R}^\ell)} = \|u\|_{L^1(X, \mathbb{R}^\ell)}$.

2.2.2 $W^{k,p}$ spaces

Let $\Omega \subset \mathbb{R}^d$, $d \geq 1$ be an open set with $1 \leq p \leq \infty$. If $a = (a_1, \dots, a_d)$ is an d -tuple of nonnegative integers a_j , we call a a *multi-index* with degree $|a| = \sum_{j=1}^d a_j$. For a differential operator $D = \frac{\partial}{\partial x}$ we write $D^a \phi = \frac{\partial^{a_1}}{\partial x_1^{a_1}} \cdots \frac{\partial^{a_d}}{\partial x_d^{a_d}} \phi$. The corresponding integrals are now with respect to the Lebesgue measure \mathcal{L}^d .

Definition 2.2.7. Let k a nonnegative integer. The Sobolev space $W^{k,p}(\Omega)$ consists of functions $u : \Omega \rightarrow \mathbb{R}$ such that $u \in L^p(\Omega)$ and the distributional derivative denoted by $D^a u$ belongs to $L^p(\Omega)$ in the *weak* sense for $0 \leq |a| \leq k$ i.e.,

$$\int_{\Omega} u D^a \phi \, dx = (-1)^{|a|} \int_{\Omega} \phi v \, dx, \quad \forall \phi \in C_c^\infty(\Omega).$$

We call v the *weak a -th derivative of u* and we write $v = D^a u$. Analogously, we define the \mathbb{R}^ℓ -valued Sobolev spaces $W^{k,p}(\Omega, \mathbb{R}^\ell)$.

We recall that $W^{k,p}(\Omega)$ is a Banach space endowed with the norm

$$\begin{aligned} \|u\|_{W^{1,p}(\Omega)} &= \sum_{0 \leq |a| \leq k} \|D^a u\|_{L^p(\Omega)} \quad \text{for } 1 \leq p < \infty, \\ \|u\|_{W^{1,\infty}(\Omega)} &= \max_{0 \leq |a| \leq k} \|D^a u\|_{L^\infty(\Omega)} \quad \text{for } p = \infty. \end{aligned}$$

It is also reflexive when $1 < p < \infty$ and separable when $1 \leq p < \infty$. For $p = 2$, we denote $W^{k,2}(\Omega)$ as $H^k(\Omega)$, which are Hilbert spaces. We write $H^0(\Omega) = L^2(\Omega)$. Moreover, we identify the closure of $C_c^\infty(\Omega)$ under the $W^{k,p}$ norm as $W_0^{k,p}(\Omega)$, namely functions that vanish on the boundary of Ω , denoted by $\partial\Omega$, in the weak sense. Analogously, we write $H_0^k(\Omega)$ instead of $W_0^{k,2}(\Omega)$. Next, we state two of the most fundamental properties for Sobolev spaces when $k = 1$. We define the *Sobolev conjugate* of p with $1 \leq p \leq d$,

$$p^* := \begin{cases} \frac{dp}{d-p} & , \text{ if } p < d \\ \infty, & \text{ if } p = d \end{cases}$$

and the mean value of u in Ω , i.e.,

$$u_\Omega := \frac{1}{\mathcal{L}^d(\Omega)} \int_{\Omega} u \, dx.$$

Theorem 2.2.8 (Rellich-Kondrachov). *Let $\Omega \subset \mathbb{R}^d$ bounded with Lipschitz boundary. Then, we have the following compact injections:*

$$\begin{aligned} W^{1,p}(\Omega) &\hookrightarrow L^q(\Omega), \quad \forall q \in [1, p^*), \, p < d \\ W^{1,p}(\Omega) &\hookrightarrow L^q(\Omega), \quad \forall q \in [p, \infty), \, p = d \\ W^{1,p}(\Omega) &\hookrightarrow C(\bar{\Omega}), \quad p > d \end{aligned}$$

In particular $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ with compact injection for all p and d . The same embeddings are valid also for the $W_0^{1,p}(\Omega)$ space.

Theorem 2.2.9. *Let $\Omega \subset \mathbb{R}^d$ bounded open set with Lipschitz boundary.*

(i) Poincaré inequality: If $1 \leq p < \infty$ then there exists a constant $C = C(\Omega, p)$ such that

$$\|u\|_{L^p(\Omega)} \leq C \|\nabla u\|_{L^p(\Omega)}, \quad \forall u \in W_0^{1,p}(\Omega)$$

(ii) Poincaré-Wirtinger inequality: If $1 \leq p \leq \infty$ then there exist a constant $C = C(\Omega, p)$ such that

$$\|u - u_\Omega\|_{L^p(\Omega)} \leq C \|\nabla u\|_{L^p(\Omega)}, \quad \forall u \in W^{1,p}(\Omega)$$

Finally, we recall the notion of weak convergence in $W^{1,p}$ spaces similar to the L^p spaces analogue.

Definition 2.2.10. Let $\Omega \subset \mathbb{R}^d$ with a Lipschitz boundary and $1 \leq p \leq \infty$. If $u_n, u \in W^{1,p}(\Omega)$, we say that u_n converges weakly (weakly* if $p = \infty$) if $u_n \rightarrow u$ strongly in L^p and $\nabla u_n \rightharpoonup \nabla u$ in $L^p(\Omega)$ ($\nabla u_n \xrightarrow{w^*} \nabla u$ in $L^\infty(\Omega)$).

2.3 Functions of bounded variation

The space of functions of bounded variation is an essential tool in mathematical imaging and in particular in this thesis. This space is connected to the notion of *edges* that we encounter in image processing. In a digital image, pixels at which image brightness changes sharply are identified by edges. In the continuum setting, this results in the need to find a suitable function space that allows *jump discontinuities*. Here, we provide a concrete summary of this space and focus on important properties and theorems that will be used in the forthcoming chapters. We recall that Ω denotes an open subset in \mathbb{R}^d .

Definition 2.3.1. A function $u \in L^1(\Omega, \mathbb{R}^\ell)$ is said to be a function of bounded variation and write $u \in \text{BV}(\Omega, \mathbb{R}^\ell)$ if its distributional derivative Du can be represented by a $\mathbb{R}^{\ell \times d}$ -valued finite Radon measure i.e., $Du \in \mathcal{M}(\Omega, \mathbb{R}^{\ell \times d})$ and for $u = (u^1, \dots, u^\ell)$

$$\int_\Omega u^a \frac{\partial \phi}{\partial x_i} dx = - \int_\Omega \phi dD_i u^a, \quad \forall \phi \in \mathcal{C}_c^\infty(\Omega, \mathbb{R}^\ell), \quad i = 1, \dots, d, \quad a = 1, \dots, \ell$$

and Du is an $\ell \times d$ matrix of measures $D_i u^a$.

As usual, if it is clear from the context what the range of the functions is, we simply write $u \in \text{BV}(\Omega)$. One of the main differences to the Sobolev spaces $W^{1,p}$ is that we require the distributional (weak) derivative to be a measure rather than an L^p function. It is straightforward from the definition that $W^{1,1}(\Omega, \mathbb{R}^\ell) \subsetneq \text{BV}(\Omega, \mathbb{R}^\ell)$. Indeed, if $u \in W^{1,1}(\Omega, \mathbb{R}^\ell)$ one can define $Du := \nabla u \mathcal{L}^d$ which is a Radon measure absolutely continuous with respect to \mathcal{L}^d and we can write

$$\int_\Omega \phi dDu = \int_\Omega \phi \nabla u d\mathcal{L}^d.$$

Also for the strict inclusion, we consider the *Heaviside* function $\mathcal{X}_{(0,\infty)}$ whose distributional derivative is the *Dirac* measure $\delta_{\{0\}}$ and is singular to \mathcal{L}^d .

According to the Lebesgue decomposition theorem Du can be decomposed to an absolutely continuous part denoted by $D^a u$ and a singular part $D^s u$ with respect to \mathcal{L}^d measure and write

$$Du = D^a u + D^s u.$$

We use the ∇u notation for the absolutely continuous part $D^a u = \nabla u$ and when $d = \ell = 1$ we simply write u' as ∇u . We define for $u \in L^1(\Omega, \mathbb{R}^\ell)$ the *total variation* of u , denoted by $\text{TV}(u)$ as

$$\text{TV}(u) = \sup \left\{ \sum_{a=1}^{\ell} \int_{\Omega} u^a \text{div} \phi^a dx : \phi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^{\ell \times d}), \|\phi\|_{\infty} \leq 1 \right\}.$$

One can show that $u \in \text{BV}(\Omega, \mathbb{R}^\ell)$ if and only if $\text{TV}(u) < \infty$ and in that case $\text{TV}(u) = |Du|(\Omega)$. If $u \in W^{1,1}(\Omega, \mathbb{R}^\ell)$ then $\text{TV}(u) = |Du|(\Omega) = \int_{\Omega} |\nabla u| dx$. Similarly to Theorem 2.1.12, the total variation is a lower semicontinuous functional in BV with respect to the strong L^1 topology i.e., if $u_n \rightarrow u$ in $L^1(\Omega, \mathbb{R}^\ell)$ then

$$|Du|(\Omega) \leq \liminf_{n \rightarrow \infty} |Du_n|(\Omega).$$

The space $\text{BV}(\Omega, \mathbb{R}^\ell)$ endowed with the norm

$$\|u\|_{\text{BV}(\Omega, \mathbb{R}^\ell)} = \|u\|_{L^1(\Omega, \mathbb{R}^\ell)} + |Du|(\Omega)$$

is a Banach space, namely an extension to the $\|\cdot\|_{W^{1,1}(\Omega)}$. However, this topology is very strong and too restrictive in most of the applications related to image processing. Therefore, the notion of the weak* convergence is introduced in BV spaces according also to Definition 2.1.11.

Definition 2.3.2 (Weak* convergence in BV). Let $u, u_n \in \text{BV}(\Omega, \mathbb{R}^\ell)$ we say that u_n weakly* converges to u , if u_n converges to u strongly in $L^1(\Omega, \mathbb{R}^\ell)$ and Du_n weakly* converges to Du in Ω i.e.,

$$\lim_{n \rightarrow \infty} \int_{\Omega} \phi dDu_n = \int_{\Omega} \phi dDu, \quad \forall \phi \in \mathcal{C}_0(\Omega)$$

Theorem 2.3.3 (Compactness in BV). Let Ω be an open bounded domain in \mathbb{R}^d with Lipschitz boundary and $(u_n)_n \in \text{BV}(\Omega, \mathbb{R}^\ell)$ such that

$$\sup_{n \in \mathbb{N}} \|u_n\|_{\text{BV}(\Omega, \mathbb{R}^\ell)} < \infty.$$

Then, there exists a subsequence $(u_{n_k})_{k \in \mathbb{N}}$ that converges to some $u \in \text{BV}(\Omega, \mathbb{R}^\ell)$ weakly* in $\text{BV}(\Omega, \mathbb{R}^\ell)$.

Naturally, the theorem above takes into account the previous definition and provides a compactness property for the BV spaces. In fact, it is a good substitute in order to obtain a compactness argument on $W^{1,1}$ space due to its lack of reflexivity. In addition, the compactness property of the BV space is a crucial factor in order to guarantee existence of solutions for our forthcoming variational problems. Finally, we present some important embeddings of BV with respect to L^p spaces and the Poincaré-Wirtinger inequality. Recall, that $1^* := \frac{d}{d-1}$ if $d > 1$ and $1^* = \infty$ if $d = 1$.

Theorem 2.3.4. *Let Ω be an open bounded domain in \mathbb{R}^d with Lipschitz boundary.*

(i) *Then $\text{BV}(\Omega) \hookrightarrow L^{1^*}(\Omega)$ with continuous injection. If $1 \leq p < 1^*$, then $\text{BV}(\Omega) \hookrightarrow L^p(\Omega)$ with a compact injection.*

(ii) *If Ω is connected, then there exists a constant $C = C(\Omega)$ such that*

$$\|u - u_\Omega\|_{L^p(\Omega)} \leq C|Du|(\Omega), \quad \forall u \in \text{BV}(\Omega), \quad 1 \leq p \leq 1^*.$$

2.4 Convex Analysis

This section summarises the basic convex analysis tools used in this thesis. Basically, we are interested in the convex conjugate and the subdifferential of a functional as well as the Fenchel-Rockafellar duality theorem. For a detailed review on convex analysis we refer to [ET99], [BC11] and [Roc70].

In the following, we assume $(X, \|\cdot\|_X)$ to be a Banach space. We start with some standard definitions on convex sets and convex functionals.

Definition 2.4.1. Let $(X, \|\cdot\|_X)$ be a Banach space, a subset C of X and a functional $F : X \rightarrow \overline{\mathbb{R}}$.

(i) The set C is called convex if $\lambda u + (1 - \lambda)v \in C$ for all $\lambda \in [0, 1]$ and $u, v \in C$.

(ii) The functional F is called convex (resp. strictly convex) if

$$F(\lambda u + (1 - \lambda)v) \leq \lambda F(u) + (1 - \lambda)F(v),$$

for all $\lambda \in [0, 1]$ and for all $u, v \in X$ (resp. for all $\lambda \in (0, 1)$, $u \neq v$ with a strict inequality).

(iii) The *effective domain* of F is the set

$$\text{dom}F = \{u \in X : F(u) < \infty\}.$$

A functional F is called *proper* if $\text{dom}F \neq \emptyset$, i.e., F is not identically equal to $+\infty$.

We proceed with the definition of the *convex conjugate* or also know as *Legendre–Fenchel transformation*. Let X^* be the dual space of X under the duality pairing $\langle \cdot, \cdot \rangle_{X, X^*}$.

Definition 2.4.2 (Convex conjugate). Let X be a Banach space and X^* its dual. For a functional $F : X \rightarrow \overline{\mathbb{R}}$, the convex conjugate of F denoted by $F^* : X^* \rightarrow \overline{\mathbb{R}}$ is

$$F^*(u^*) = \sup_{u \in X} \langle u, u^* \rangle - F(u) \text{ for all } u^* \in X^*.$$

Obviously, we can consider taking the supremum over all $u \in \text{dom}F$ which is translated as the pointwise supremum of continuous and affine functions $\langle u, \cdot \rangle - F(u)$. Hence, F^* is lower semicontinuous and convex. Furthermore, we may repeat the conjugation process, leading to the *biconjugate* $F^{**} : X \rightarrow \overline{\mathbb{R}}$ defined analogously as

$$F^{**}(u) = \sup_{u^* \in X^*} \langle u, u^* \rangle - F^*(u^*), \text{ for all } u \in X.$$

The next fundamental theorem allows us to determine under which conditions on the functional F , the biconjugate F^{**} is indeed equal to F . The proof can be found in [Bre11].

Theorem 2.4.3 (Fenchel-Moreau). *Let $F : X \rightarrow \overline{\mathbb{R}}$ be a convex, lower semicontinuous and proper functional. Then $F^{**} = F$.*

Let us demonstrate the importance of this theorem with few useful examples:

Example 2.4.4. (i) *Let C a nonempty subset of X . We set the indicator function of C as*

$$\mathcal{I}_C(u) = \begin{cases} 0, & u \in C \\ +\infty, & u \notin C. \end{cases}$$

Clearly, \mathcal{I}_C is convex and lower semicontinuous if and only C is convex and closed respectively. Now, we obtain $\mathcal{I}_C^(u^*) = \sup_{u \in C} \langle u, u^* \rangle$ where the right-hand side is commonly referred to as the support function of C and denoted by $\sigma_C(u^*)$. Additionally, if we assume that C is a closed and convex subset of X then by Theorem 2.4.3 we obtain that $\mathcal{I}_C^{**} = \mathcal{I}_C$.*

(ii) *If $F(u) = \|u\|_X$, then by the Cauchy-Schwarz inequality one has*

$$F^*(u^*) \leq \sup_{u \in X} \|u\|_X (\|u^*\|_{X^*} - 1).$$

Then, if we consider separately the cases where $\|u\|_{X^} \leq 1$ and $\|u\|_{X^*} > 1$ we conclude that $F^* = I_{B_{X^*}}$, i.e., the indicator function of the unit ball of the dual*

space X^* . Similarly, for the biconjugate of F , we obtain that

$$F^{**}(u) = \sup_{\substack{u^* \in X^* \\ \|u\|_{X^*} \leq 1}} \langle u, u^* \rangle = \|u\|_X = F(u).$$

(iii) If $1 < p < \infty$ and $F(u) = \frac{1}{p} \|u\|_{L^p(X)}^p$, then by Young's inequality we have that

$$F^*(u^*) = \frac{1}{q} \|u^*\|_{L^q(X)}^q,$$

where q is the Hölder exponent of p .

Let us discuss now the *subdifferential* of a functional. The concept of subdifferentiability constitutes a generalisation of differentiability in Banach spaces regarding for example the Gâteaux and Fréchet derivatives.

Definition 2.4.5 (Gâteaux-Fréchet). Let X, Y be two Banach spaces. A map $F : X \rightarrow Y$ is said to be *Gâteaux differentiable* at $u \in X$ if there exists a bounded linear operator $T_u : X \rightarrow Y$ such that $\forall h \in X$,

$$\lim_{t \rightarrow 0} \frac{F(u + th) - F(u)}{t} = T_u h.$$

If the limit exists, we say that F has a *directional derivative at u in the direction h* . The operator T_u is said to be the Gâteaux derivative of F at u denoted by $F'_G(u)$. If F is Gâteaux differentiable at every point $u \in X$, then we say that F is Gâteaux differentiable.

If the limit exists *uniformly* in h on every bounded set in X , we say that F is Fréchet differentiable at $u \in X$ and T_u is the Fréchet derivative of F at u denoted by $F'_F(u)$. Equivalently, if we set $z = ht$ then if $t \rightarrow 0$ we have that $z \rightarrow 0$ and F is Fréchet differentiable at u if for all $z \in X$, $F(u + z) - F(u) - T_u(z) = o(\|z\|_X)$, i.e.,

$$\frac{\|F(u + z) - F(u) - T_u(z)\|_Y}{\|z\|_X} \rightarrow 0, \text{ as } \|z\|_X \rightarrow 0.$$

Definition 2.4.6 (Subdifferential). Let $(X, \|\cdot\|_X)$ be a Banach space and $F : X \rightarrow \overline{\mathbb{R}}$. We say that F is subdifferentiable at $u \in X$ if there exists $u^* \in X^*$ such that

$$F(v) \geq F(u) + \langle u^*, v - u \rangle, \quad \forall v \in X.$$

We call u^* the subgradient of F at $u \in X$. The collection of all subgradients is called the subdifferential of F at u denoted by $\partial F(u)$ i.e.,

$$\partial F(u) = \{u^* \in X^* : F(v) \geq F(u) + \langle u^*, v - u \rangle, \forall v \in X\}.$$

Combining the two previous definitions, we can prove that if F is Fréchet differentiable then it also Gâteaux differentiable. Moreover, if F is Gâteaux differentiable at $u \in X$ then $\partial F(u) = \{F'_G(u)\}$. In general, the subdifferential can be multivalued. A standard example is $F : \mathbb{R} \rightarrow \mathbb{R}^+$ with $F(u) = |u|$. Then, the subdifferential is

$$\partial F(u) = \text{sgn}(u) = \begin{cases} 1, & \text{if } u > 0, \\ [-1, 1], & \text{if } u = 0, \\ -1, & \text{if } u < 0. \end{cases}$$

Next, we recall some useful properties for the subdifferential. Most of the proofs can be found in [ET99].

Proposition 2.4.7. *Let $(X, \|\cdot\|_X)$ be a Banach space and $F : X \rightarrow \overline{\mathbb{R}}$ with $F^* : X^* \rightarrow \overline{\mathbb{R}}$ its convex conjugate.*

(i) $u^* \in \partial F(u) \Leftrightarrow F(u) + F^*(u^*) = \langle u, u^* \rangle.$

(ii) *If F is a proper, lower semicontinuous and convex functional then*

$$u^* \in \partial F(u) \Leftrightarrow u \in \partial F^*(u^*).$$

(iii) *Let $F_1 : X \rightarrow \overline{\mathbb{R}}$, $F_2 : Y \rightarrow \overline{\mathbb{R}}$ be two proper, lower semicontinuous and convex functionals with $\mathcal{R} : X \rightarrow Y$ be a linear operator. If $\mathcal{R}(\text{dom}F_1) \cap \text{dom}F_2 \neq \emptyset$ with F_2 continuous at $\mathcal{R}u \in Y$ then*

$$\partial(F_1(u) + F_2(\mathcal{R}u)) = \partial F_1(u) + \partial F_2(\mathcal{R}u), \quad \partial F_2(\mathcal{R}u) = \partial(F_2 \circ \mathcal{R})(u) = \mathcal{R}^* \partial F_2(\mathcal{R}u),$$

where $\mathcal{R}^* : Y^* \rightarrow X^*$ is the adjoint operator of \mathcal{R} .

Throughout this thesis, we mainly consider functionals that are convex, non-differentiable and homogeneous. A functional F is said to be (positively) homogeneous of degree p if for $\alpha \in \mathbb{R}$ then $F(\alpha u) = |\alpha|^p F(u)$. We can obtain a useful characterisation of the subdifferential when we are dealing with one-homogenous functionals.

Proposition 2.4.8 (One-homogenous). *Let $F : X \rightarrow \overline{\mathbb{R}}$ be a proper, lower semicontinuous, convex and one-homogeneous functional. Then, $F = \sigma_C$ with $C = \partial F(0)$ and*

$$\partial F(u) = \{u^* \in X^* : \langle u, u^* \rangle = F(u), \langle u^*, v \rangle \leq F(v), \forall v \in X\}.$$

Proof. Let $u^* \in C$, then $F^*(u^*) = \sup_{u \in X} \langle u, u^* \rangle - F(u) \leq 0 = \langle 0, u^* \rangle - F(0) \leq F^*(u^*)$ i.e., $F^*(u^*) = 0$. On the other hand, there exists $z \in X$ such that $\langle z, u^* \rangle - F(z) > 0$. Then,

$\forall \lambda > 0$ we obtain $F^*(u^*) \geq \langle \lambda z, u^* \rangle - F(\lambda z) = \lambda(\langle z, u^* \rangle - F(z))$ which can be arbitrarily large i.e., $F^*(u^*) = +\infty$. Hence, we obtain that $F^*(u^*) = I_C(u^*)$ and since F^* is also convex and lower semicontinuous, we have that C is a closed, convex set in X^* . Finally, by Theorem 2.4.3 and Proposition 2.4.7 (i) we conclude that $F = F^{**} = (I_C)^* = \sigma_C$, $F(u) = \langle u, u^* \rangle$ with $u^* \in \partial F(0)$, see also Example 2.4.4 (i), and

$$\partial F(u) = \{u^* \in X^* : \langle u, u^* \rangle = F(u), \langle u^*, v \rangle \leq F(v), \forall v \in X\}.$$

□

An important consequence is in the case of total variation norm, where one can obtain the following characterisation:

$$\partial \|Du\|_{\mathcal{M}} = \{\operatorname{div} \phi, \|\phi\|_{\infty} \leq 1, \phi|_{\partial\Omega} \cdot n = 0, \langle \operatorname{div} \phi, u \rangle = \|Du\|_{\mathcal{M}}\}$$

Indeed, see also [BO13],[CCC⁺10], by Proposition 2.4.8, we have that

$$\partial \|Du\|_{\mathcal{M}} = \left\{ \xi \in \operatorname{BV}(\Omega)^* : \langle \xi, u \rangle_{\operatorname{BV}, \operatorname{BV}^*} = \|Du\|_{\mathcal{M}}, \langle \xi, v \rangle_{\operatorname{BV}, \operatorname{BV}^*} \leq \|Dv\|_{\mathcal{M}}, \forall v \in \operatorname{BV}(\Omega) \right\}.$$

Then,

$$\|\xi\|_{\operatorname{BV}(\Omega)^*} = \sup_{\substack{v \in \operatorname{BV}(\Omega) \\ \|Dv\|_{\mathcal{M}} \leq 1}} \langle \xi, v \rangle \leq \sup_{\substack{v \in \operatorname{BV}(\Omega) \\ \|Dv\|_{\mathcal{M}} \leq 1}} \|Dv\|_{\mathcal{M}} = 1$$

and the subdifferential can now be written as

$$\partial \|Du\|_{\mathcal{M}} = \left\{ \xi \in \operatorname{BV}(\Omega)^* : \|\xi\|_{\operatorname{BV}(\Omega)^*} \leq 1, \langle \xi, u \rangle_{\operatorname{BV}, \operatorname{BV}^*} = \|Du\|_{\mathcal{M}} \right\}.$$

Now let us define the dual of BV, which is also discussed in Section 4.5.1. It can be identified by the space $G(\Omega)$, where it was initially introduced first in \mathbb{R}^2 by Y. Meyer in [Mey01] and generalised for $\Omega \subset \mathbb{R}^2$ in [AA05], i.e.,

$$G(\Omega) = \left\{ \xi \in L^2(\Omega, \mathbb{R}^\ell) : \xi = \operatorname{div} \phi, \phi \in L^\infty(\Omega, \mathbb{R}^\ell), \phi|_{\partial\Omega} \cdot n = 0 \right\}$$

under the so-called G^* or Meyer norm

$$\|\xi\|_* := \inf \{ \|\phi(x)\|_{\infty} \mid \operatorname{div} \phi = \xi \}$$

with $\int_{\Omega} \xi \, dx = 0$. The existence of such ϕ is guaranteed by [BB03, Theorem 3] and therefore we obtain the required expression.

2.4.1 Fenchel-Rockafellar duality

We discuss the Fenchel-Rockafellar duality related to variational problems that we encounter in this thesis. Our goal is to associate to a minimisation problem (\mathcal{P}), referred to as the *primal problem*, a maximisation problem (\mathcal{P}^*), referred to as the *dual problem* and examine the relationship between these two problems, especially the connection between their solutions.

Let X, Y two Banach spaces with $F_1 : X \rightarrow \overline{\mathbb{R}}, F_2 : Y \rightarrow \overline{\mathbb{R}}$ be two proper, convex, lower semicontinuous functionals and a bounded linear operator $K : X \rightarrow Y$. We consider a minimisation problem (primal problem) of the following form i.e.,

$$\inf_{u \in X} F_1(u) + F_2(Ku).$$

The corresponding dual problem is a maximisation problem defined as

$$\sup_{u^* \in Y^*} -F_1^*(K^*(u^*)) - F_2^*(-u^*),$$

where $K^* : Y^* \rightarrow X^*$ is the adjoint of K . We denote the infimum of (\mathcal{P}) problem as $\inf \mathcal{P}$ and analogously the supremum of (\mathcal{P}^*) as $\sup \mathcal{P}^*$. In general, we have that $\sup \mathcal{P}^* \leq \inf \mathcal{P}$ and both of them are finite. We say that there is no *duality gap* if they are equal and the theorem below, see [ET99, Chapter III] provides a characterisation between their solutions.

Theorem 2.4.9. *Assume that the primal and dual problems have solutions and $\sup \mathcal{P}^* = \inf \mathcal{P}$. Then, u and u^* are solutions of (\mathcal{P}) and (\mathcal{P}^*) respectively if and only if satisfy the following optimality conditions*

$$\begin{aligned} K^*(u^*) &\in \partial F_1(u), \\ -u^* &\in \partial F_2(Ku). \end{aligned}$$

In order to guarantee that $\sup \mathcal{P}^* = \inf \mathcal{P}$, we need to make the following assumption: that there exists $u_0 \in X$ such that both $F_1(u_0)$ and $F_2(Ku_0)$ are finite and F_2 is continuous at Ku_0 . However, Attouch and Brezis in [AB86] proved the same result with a weaker and more geometrical assumption, see also [BC11, Chapter 15].

Theorem 2.4.10 (Attouch-Brezis). *Let X, Y be two Banach spaces, $K : X \rightarrow Y$ a bounded linear operator and $F_1 : X \rightarrow \overline{\mathbb{R}}, F_2 : Y \rightarrow \overline{\mathbb{R}}$ be two proper, convex, lower semicontinuous functionals. If*

$$\bigcup_{\lambda \geq 0} \lambda (\text{dom}(F_2) - K(\text{dom}F_1))$$

is a closed subspace of Y , then

$$\inf_{u \in X} F_1(u) + F_2(Ku) + \min_{u^* \in Y^*} F_1^*(K^*(u^*)) + F_2^*(-u^*) = 0,$$

the dual problem has at least one solution and there is no duality gap.

Chapter 3

Total variation regularisation in image and sinogram space for PET reconstruction

3.1 Introduction

In this chapter, we introduce our novel approach in image reconstruction for Positron Emission Tomography abbreviated as PET. It is based on the publication [BMPS14]. Generally speaking, *emission tomography* is a specific branch of medical imaging that employs radioactive materials to depict biochemical and physiological processes such as metabolic activities, blood flow in the heart, detection of tumors, and to identify brain regions influenced by drugs. It can be understood as the combination of two basic principles: (1) Emission: imaging through the use of a *radioactive tracer* or *marker*. The idea of the tracer principle was first introduced by the Nobel prize winner in chemistry George de Hevesy in the early 1900's. His research on how plants absorb radioactive tracers was the foundation of the current state of medical techniques such as PET. (2) Tomography: projection imaging or drawing a cross-sectional slices of the region of interest. This idea allows to obtain images that cannot be physically acquired, except by making a planar cut through the body and then looking at the exposed surface. Hence, the importance of tomography is to visualise characteristics of the human body without the need for surgery leading to a significant breakthrough of non-invasive medical techniques.

In general, the radioactive tracers emit gamma rays which are *photons* with high energy. The emitted photons are recorded by appropriate detectors, placed around the body and have geometrical information about the origin of every photon. Through these measurements, which are in the form of projections, our scope is to apply mathematical techniques in order to reconstruct accurately an image that represents the region of interest, namely where the radiotracer started to emit photons and how it is distributed in the

body.

Our proposed model can be summarised by the following minimisation problem: given noisy measurements $g \in L^2(\Sigma^2)$, we reconstruct $u \in \text{BV}(\mathbb{R}^2)$ by solving

$$\operatorname{argmin}_{u \in \text{BV}(\mathbb{R}^2)} \left\{ \alpha |Du|(\mathbb{R}^2) + \beta |D(\mathcal{R}u)|(\Sigma^2) + \frac{1}{2} \int_{\Sigma^2} \frac{(g - \mathcal{R}u)^2}{g} d\sigma \right\}. \quad (3.1)$$

The terms $|Du|(\mathbb{R}^2)$ and $|D(\mathcal{R}u)|(\Sigma^2)$ are the total variation penalties on the image u and on the measurement space $\Sigma^2 = \{(\theta, s) : \theta \in \mathcal{S}^1, s \in \mathbb{R}\}$ of $\mathcal{R}u$. The space Σ^2 is commonly referred as the *sinogram* space and contains the values of $\mathcal{R}u(\theta, s)$, where \mathcal{R} denotes the Radon transform, $\mathcal{R} : L^1(\mathbb{R}^2) \rightarrow L^1(\Sigma^2)$ with

$$\mathcal{R}u(\theta, s) = \int_{\mathbb{R}^2} \delta(s - x \cdot \theta) u(x) dx. \quad (3.2)$$

The fidelity term represents a weighted L^2 difference between $\mathcal{R}u$ and the noisy data g , which is a suitable choice due to its the statistics of the noise. In fact, as we describe in Section 3.2, the measurements g during a PET scan are corrupted by *Poisson* noise and the weighted L^2 norm constitutes a standard approximation of the *Kullback-Leibler divergence* $\int_{\Sigma^2} (\mathcal{R}u - g) \log \mathcal{R}u d\sigma$, see [Saw11, Chapter 4]. The positive parameters α and β are tuned appropriately in order to balance the strength of the regularisation on the image and on the sinogram space that we require in the final solution u .

PET reconstruction using total variation regularisation is not new. Most of the state of the art methods, see Section 3.4 are focused on regularising only the image space. However, the novelty of our model is that we add an additional regulariser on the sinogram space which enforces a priori information on the total variation of $\mathcal{R}u$. This approach can be very useful when we are dealing with measurements corrupted with high levels of noise and images that are governed by thin and elongated structures, see Section 3.7.3.

Organisation of the chapter

We first begin with a proper introduction to emission tomography and the clinical stages during a PET scan, see Section 3.2.

In Section 3.3, we summarise all the basic properties of the Radon transform and the *sinogram* space. We focus on its continuity in L^p spaces for functions with compact support. We also recall the projection-slice theorem, which provides a connection between the Radon and Fourier transforms and finally the filtered backprojection, a technique that reconstructs an image from its noisy projections.

We proceed in Section 3.4 with a general review of the state of the art reconstruction methods such as the expectation-maximisation (EM) algorithm and its extension via the TV regularisation, namely the EM-TV approach introduced in [BBS⁺09],[BLZ08]. In Section 3.5, we begin the analysis of the proposed model (3.1). One of the crucial step

towards the well-posedness of the model is the BV continuity of the Radon transform of functions with compact support. Existence, uniqueness and stability results are provided as well as an error analysis using the *Bregman* distance.

In order to understand how regularisation of the sinogram space behaves and how the reconstruction reflects in the image space, we compute analytically the TV regularisation applied only on the sinogram of the characteristic function of ball with radius r i.e., $u = \chi_{B_r}$. In particular, we derive the optimality conditions of (3.1) when $\beta = 0$ and obtain the corresponding regularised solution u , see Section 3.5.4.

Finally, in Sections 3.6 and 3.7, we proceed with the discretised problem of (3.1) and perform our numerical experiments. We use the *split Bregman* algorithm, see Section 3.6.2, to solve our minimisation problem and apply our method to both simulated phantoms and real PET data. We demonstrate numerically the relation between regularising only on the sinogram space and its corresponding image reconstruction which validates the theoretical results of Section 3.5.4. We focus on how geometrical features of the reconstructing image u , especially on the boundaries of the image, are affected in terms of the choice of the parameter β . We conclude this chapter with numerical simulations of PET data with thin and elongated structures that clearly demonstrate the advantage of our method.

3.2 Positron Emission Tomography

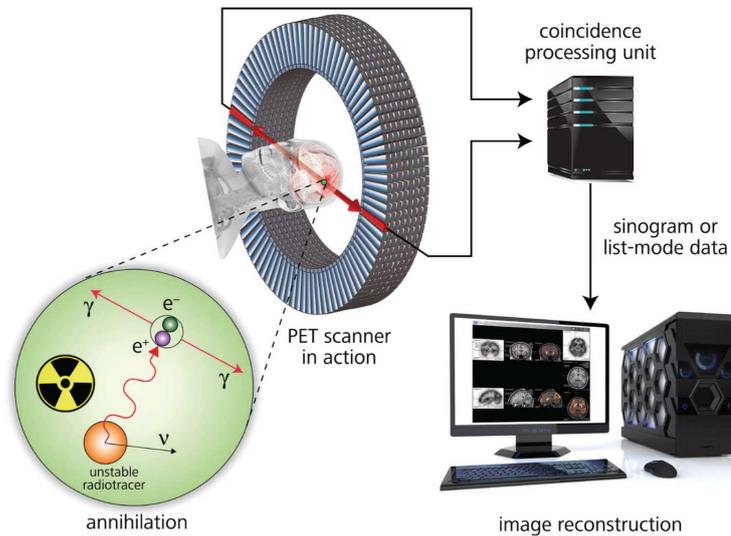


Figure 3.1: Stages of PET scan, Image courtesy of *Wikipedia*.

In this section, we describe the clinical stages during a PET scan as presented in Figure 3.1. For a thorough analysis of the physics and instrumentation aspects of PET imaging as well as reconstruction methods we refer the reader to [Phe06] and [WA04].

The clinical Positron Emission tomography is divided into four stages:

- (i) A patient is injected (usually into blood circulation) with a radioactive tracer or radiopharmaceutical. One of the most important radiopharmaceutical and widely used in brain imaging is the ^{18}F -FDG. This element consists of two parts: the tracer compound, FDG which stands for *fluorodeoxyglucose* an analogue of glucose that interacts with the body and a positron emitter, ^{18}F a *fluorine* radioisotope. In recent years, ^{18}F -FDG has also become an important asset for cancer imaging since malignant tumors cells can produce high levels of glucose metabolism in comparison with normal cells.

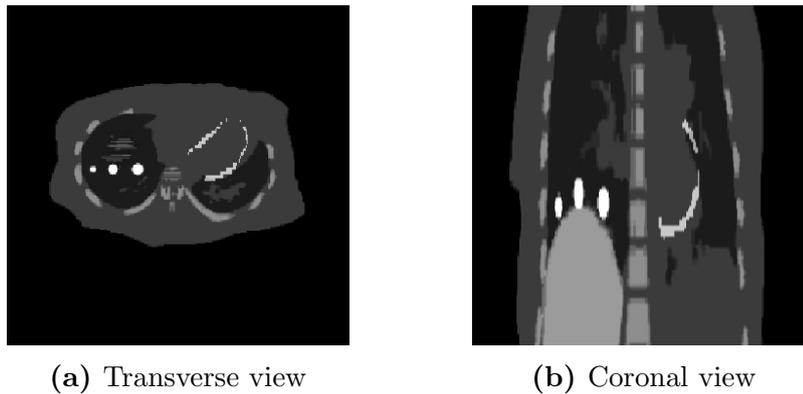


Figure 3.2: Realistic PET data for the chest simulated in [TPT⁺13]. Six different tumour lesions appear in transverse and coronal views.

- (ii) After the administration of the radiopharmaceutical, there is a waiting period that depends on the nature of the tracer. For example, for the ^{18}F -FDG is usually one hour. During this period, ^{18}F -FDG, as a glucose analogue, is taken up by high-glucose-using cells such as brain, kidney, and cancer cells. In practice, cells in the tissue which are more active have a higher metabolism, i.e., need more energy, and hence will absorb more tracer isotope than cells which are less active. Thus, tumors cells are highlighted in emission tomography images as bright regions against the relatively dark background of surrounding normal tissues, see Figure 3.2.
- (iii) The radioactive isotope starts to emit positron from its nucleus as a product of the radioactive decay. Then, the emitted positron travels in tissue for a very short distance (less than 1mm) losing enough of its kinetic energy until it meets with an electron. At this point, the positron and the electron annihilate and emit two gamma photons carrying an energy equal to 511Kev and travel in almost opposite directions.
- (iv) As the gamma rays travel through the patient's body in opposite directions, they are detected and recorded by an imaging hardware, as it is shown in Figure 3.1. It consists by a large number of detectors e.g. block detectors ordered in a ring

geometry around the body. Once the photons are recorded from these detectors, they can indicate the annihilation point which occurred somewhere along this *line*. This line from which a pair of detectors can detect coincidences usually is referred as *line of response* denoted by LOR. A collection of a large number of these LORs in different angles can help us reconstruct a complete cross-sectional image of the object.

From a physical point of view, the change of intensity I of the emitted photons travelling through the body can be described according to Lambert-Beer law:

$$\frac{dI}{I} = - \int_{L'(x)} \mu(z) dz \Rightarrow I = I_0 e^{-\int_{L'(x)} \mu(z) dz} \quad (3.3)$$

where $\mu(z)$ is the *attenuation-coefficient* distribution of the body known as μ -map, and $L'(x)$ is the distance which extends from the annihilation point x through the detector, see Figure 3.3. The attenuation-coefficient distribution is the tendency of an object to absorb or scatter rays of a given energy. For example, a bone has a much higher attenuation coefficient than from a soft tissue. It is usually compared with the attenuation coefficient of the water and is quoted in terms of the *Hounsfield unit*, see [Eps07, Chapter 3]. Analogously, for the spatial distribution of the radiotracer $u(x)$, we have that the measurements recorded from the detectors are mathematically described by

$$g = \int_L u(x) e^{-\int_{L'(x)} \mu(z) dz} dx \quad (3.4)$$

which is called the *attenuated Radon transform* of u and is denoted by $\mathcal{R}_\mu u$. In this case, the distribution $u(x)$ is weighted by an exponential term that represents the attenuation coefficient of the body, where L denotes all the object points along the line connecting the detectors. A detailed mathematical analysis of the Radon transform will be provided in Section 3.3. In most modern PET devices, the distribution of the attenuation coefficient $\mu(z)$ is already computed via a second imaging unit, like CT (computed tomography) or MRI (magnetic resonance imaging). For example, there is the PET/CT imaging device that combines the functionality of both PET and CT, see [WA04, Chapters 10,11] for more details. In our analysis, we assume that $\mu(z) \equiv 0$ and (3.4) becomes

$$g = \int_L u(x) dx. \quad (3.5)$$

Our ultimate goal is to find the distribution $u(x)$ given the measurements g . This is a complicated task on its own and can be even more difficult, if one takes into account that during a PET scan, these measurements are corrupted with noise. Individual photon detections can be treated as independent events that follow a random distribution. It is usually referred as *photon noise* or *Poisson noise* due to the fact that the number of

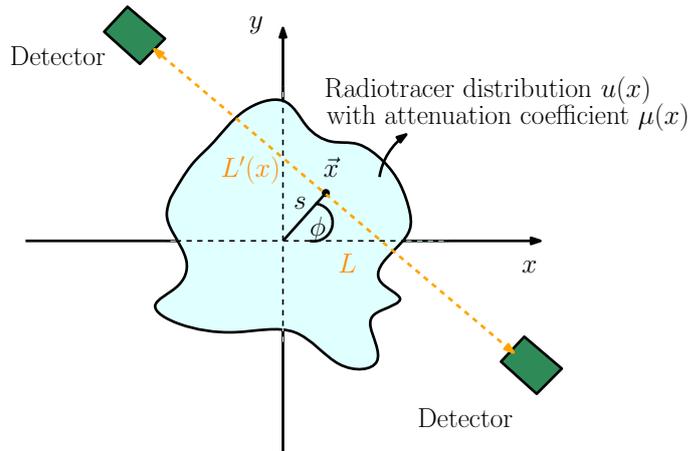


Figure 3.3: PET scan geometry: The measurements g of a radiotracer distribution $u(x)$ with attenuation coefficient $\mu(x)$ are given by (3.4) where $L'(x)$ is a line segment of L for all object points measuring the distance from the annihilation point to the detector.

events recorded by the detectors in any fixed interval of time obeys the well-known Poisson probability distribution. In detail, let X be a discrete random variable assigned to the number of photons measured by a given sensor/detector over a time interval t . Then, the discrete probability distribution is

$$P(X = k; \lambda) = \frac{e^{-\lambda} \lambda^k}{k!} \quad (3.6)$$

where λ is the expected number of photons per unit time interval. Since, both the

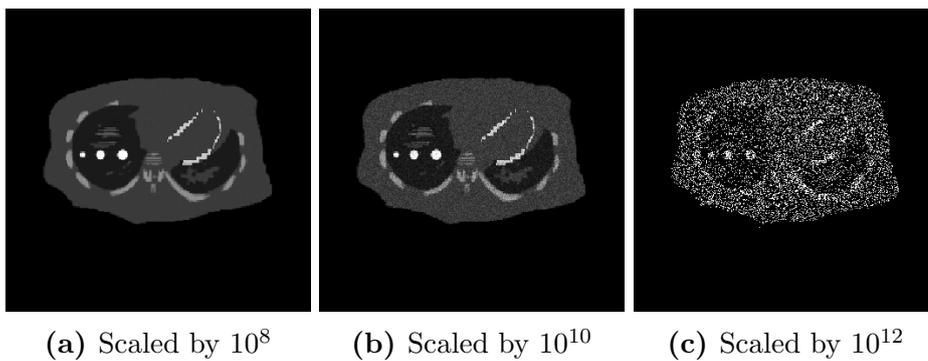


Figure 3.4: Poisson noise with different scaling factors of Figure 3.2, meaning different photon counts see also section 3.7.

expectation and the variance of Poisson distribution are equal to λ , this indicates that Poisson noise is signal dependent meaning that the noise intensity perturbing the image pixel is proportional to the clean image. In fact the standard deviation grows with the square root of the clean image. This can be justified in Figure 3.4, where the image quality

deteriorates seriously when the photon count decreases. We use the MATLAB's *imnoise* routine with different scaling factors in order to simulate the number of photon counts, see Section 3.7 for more details. The Poisson distribution will be used again in Section 3.4 in order to obtain a variational approach via the Bayesian framework.

3.3 Radon transform

In 1917, an Austrian Mathematician named Johann Radon solved the following problem: *Could we determine an object provided that the total density of every line through the object is known?* In [Rad17] or in the first English translation in [Dea83, Appendix], Radon introduced an operator measuring the density of a function u along lines and provide an analytical solution, known as *the Radon inversion formula*, of recovering the function u by its density. In this section, we briefly summarise all the basic properties of the Radon transform such as its continuity, its connection with the Fourier transform and the filtered back-projection. We mainly follow [NW01] and [Nat01]. Although, we refer the reader to other standard textbooks such as [Mar14], [Eps07], [Dea83].



Figure 3.5: Johann Radon (1887-1956)

Definition 3.3.1 (Radon transform). Let $u \in L^1(\mathbb{R}^d)$ and $H(\theta, s) = \{x \in \mathbb{R}^d : x \cdot \theta = s\}$ be the hyperplane perpendicular to $\theta \in \mathcal{S}^{d-1}$ with distance $s \in \mathbb{R}$ from the origin. The d -dimensional Radon transform of u is defined as

$$\mathcal{R}u(\theta, s) = \int_{H(\theta, s)} u(x) dx. \quad (3.7)$$

Equivalently, we can write

$$\mathcal{R}u(\theta, s) = \int_{y \in \theta^\perp} u(s\theta + y) dy \quad (3.8)$$

integrating along the orthogonal subspace $\theta^\perp = \{x \in \mathbb{R}^d : x \cdot \theta = 0\}$. The domain of the Radon transform is $\Sigma^d = \{(\theta, s) : \theta \in \mathcal{S}^{d-1}, s \in \mathbb{R}\}$ a cylinder of dimension d and is often referred as *projection space* or *Radon space* or *sinogram space*.

One common expression regarding the Radon transform, see [Nat01] for instance, is

by using the Dirac δ_0 measure such that

$$\int_{\mathbb{R}^d} u(x) d\delta_0(x) = u(0).$$

Let δ^k be an exponential approximation of δ then

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} \delta^k(x) u(x) dx = u(0).$$

Hence, we have the following

$$\begin{aligned} \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} u(x) \delta^k(s - x \cdot \theta) dx &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}} \int_{\theta^\perp} u(t\theta + y) \delta^k(s - t) dy dt \\ &= \int_{\theta^\perp} u(s\theta + y) dy, \end{aligned}$$

where in a sense we can write

$$\mathcal{R}u(\theta, s) = \int_{\mathbb{R}^d} u(x) \delta(s - x \cdot \theta) dx. \quad (3.9)$$

Clearly, the Radon transform is a well-defined linear operator for $u \in L^1(\mathbb{R}^d)$, and in fact is L^1 continuous, since by a simple change of variables and Fubini's theorem we obtain

$$\int_{\mathbb{R}} \mathcal{R}u(\theta, s) ds = \int_{\mathbb{R}} \int_{y \in \theta^\perp} u(s\theta + y) dy ds = \int_{\mathbb{R}^d} u(x) dx$$

and

$$\|\mathcal{R}u\|_{L^1(\Sigma^d)} \leq |\mathcal{S}^{d-1}| \|u\|_{L^1(\mathbb{R}^d)}. \quad (3.10)$$

However, this is not true in general. It can be shown, see [Bor98], that for $p \geq \frac{d}{d-1} > 1$ the function

$$u(x) = (2 + |x|)^{-\frac{d}{p}} \ln \frac{1}{2 + |x|}$$

belongs to $L^p(\mathbb{R}^d)$, since $\|u(x)\|_{L^p} \leq \frac{\ln^{1-p}(2)}{p-1}$, but $\mathcal{R}u(\theta, s) \rightarrow \infty$.

Let us now examine the Radon transform in a two dimensional setting. For $d = 2$ equation (3.7) can be simplified to

$$\mathcal{R}u(\theta, s) = \int_{-\infty}^{\infty} u(s \cos \phi - t \sin \phi, s \sin \phi + t \cos \phi) dt. \quad (3.11)$$

Indeed, let a point (x, y) and line L passing through it with distance s from the origin being perpendicular to $\theta = (\cos \phi, \sin \phi)$, see Figure 3.6. Then, a line passing through the origin with angle ϕ from the x -axis intersects the line L at the point $(s \cos \phi, s \sin \phi) = \vec{r}_0$. Hence, for every $t \in \mathbb{R}$ we have that $\vec{r} = \vec{r}_0 + t\theta^\perp$ i.e.,

$$\begin{aligned} x &= s \cos \phi - t \sin \phi \\ y &= s \sin \phi + t \cos \phi. \end{aligned} \quad (3.12)$$

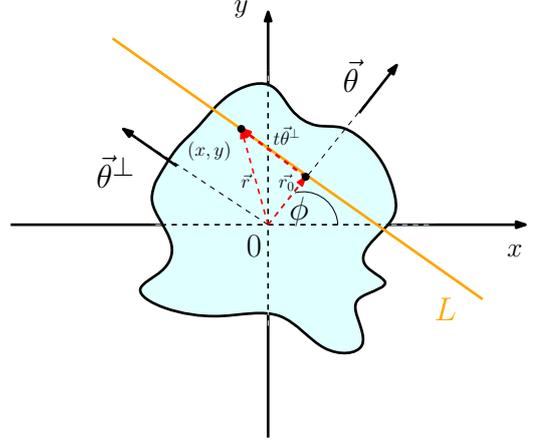


Figure 3.6: Local coordinates of PET in \mathbb{R}^2

Moreover, the space $\Sigma^2 = S^1 \times \mathbb{R}$ can be identified as a half-cylinder with unit radius whose surface contains all the measurements $\mathcal{R}u(\theta, s)$ with $\theta = (\cos \phi, \sin \phi) \in S^1$. From (3.9), we observe that $\mathcal{R}u(\theta, s) = \mathcal{R}u(-\theta, -s)$. In addition, if for example $\mathcal{R}u(\theta, s)$ is known for $-\infty < s < \infty$, then the only values of ϕ we need are in the range of $\phi \in [0, \pi)$, since the points $(-s, \phi)$ and $(s, \phi + \pi)$ produce the same value, see Figure 3.7a. Likewise, we can also have $s \in [0, \infty)$ and $\phi \in [0, 2\pi)$. However, in practice the distance s of a line L from the origin has always finite value and cannot extend to infinity, hence we usually write $\Sigma^2 = \{(\phi, s) : s \in [-r, r], \phi \in [0, \pi)\}$ for some $r > 0$. Now, if we unroll this cylinder, the resulting surface is a rectangular grid containing the values of the line integral (3.11) for every distance s and angle ϕ . This unrolled space is commonly referred as the *sinogram* space. This is due to the fact that a single point in the (x, y) plane has a *sinusoidal* form in the Radon space. Indeed, we are interested on lines that passing through this point for different angles ϕ . These lines can be expressed using (3.12), i.e., $s = x_0 \cos \phi + y_0 \sin \phi$. Hence, for the polar coordinates $(\sqrt{x_0^2 + y_0^2}, \phi_0)$, we have that

$$\begin{aligned} s &= x_0 \cos \phi + y_0 \sin \phi = \sqrt{x_0^2 + y_0^2} \cos \phi_0 \cos \phi + \sqrt{x_0^2 + y_0^2} \sin \phi_0 \sin \phi \\ &= \sqrt{x_0^2 + y_0^2} \cos(\phi - \phi_0). \end{aligned} \quad (3.13)$$

Roughly, a single a point in the (x, y) plane, e.g. a bright spot, is interpreted as a bright sinusoidal curve over a grey background, see Figure 3.7.

Let us further lay emphasis on the boundedness of the object that we aim to reconstruct. In medical applications, we mainly deal with functions with compact support. For instance, if $u(x, y) = 0$ for $x^2 + y^2 > r^2$ then by (3.9) $\mathcal{R}u(\theta, s) = 0, \forall |s| > r$. In fact, under these assumptions one can guarantee the L^p continuity of the Radon transform. It has been shown for $p = 2$ in [SSW77, Section 12]. The proof for general p is based on a similar

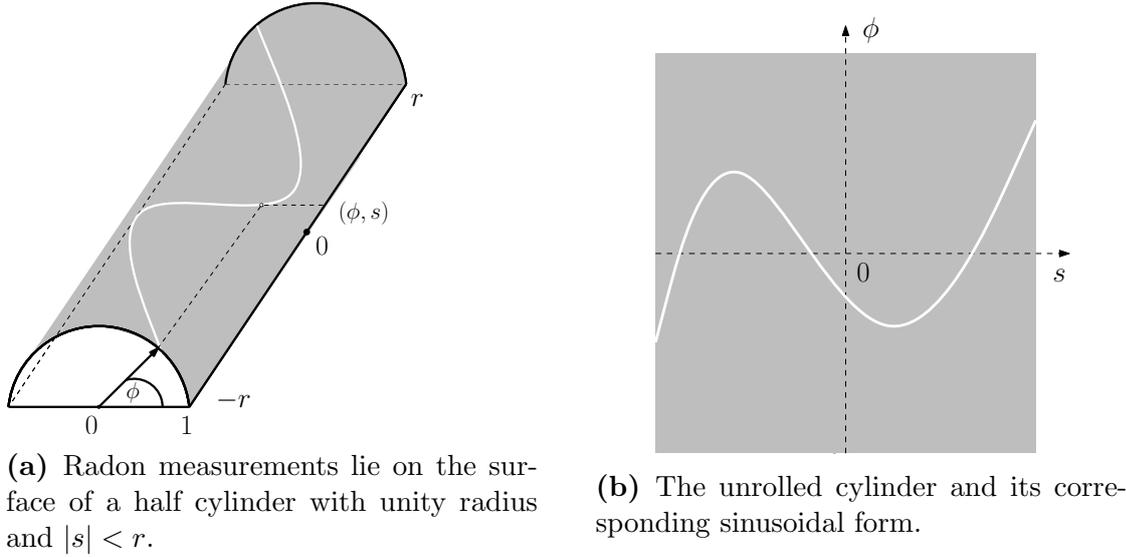


Figure 3.7: The sinogram space $\Sigma^2 = \{(\phi, s) : s \in [-r, r], \phi \in [0, \pi]\}$ of a function which is 0 everywhere except at (x_0, y_0) , see also (3.13).

proof for the k -plane transforms, an extension to Radon transform, see [Mar14, Chapter 3] for more details. In Section 3.5.1, we will prove a similar continuity result regarding locally supported functions of bounded variation.

Proposition 3.3.2. *Let $u \in L^p_{loc}(\mathbb{R}^d)$ i.e., an L^p function vanishing outside a compact set K with $1 \leq p \leq \infty$ and q its Hölder exponent as in (2.2.1). Then, the Radon transform is continuous in $L^p(\Sigma^d)$ and*

$$\|\mathcal{R}u\|_{L^p(\Sigma^d)} := \left(\int_{\Sigma^d} |\mathcal{R}u(\theta, s)|^p ds d\theta \right)^{\frac{1}{p}} \leq |C \mathcal{S}^{d-1}|^{\frac{1}{p}} \text{diam}(K)^{\frac{d-1}{q}} \|u\|_{L^p(\mathbb{R}^d)}.$$

Proof. Let K be a compact subset of \mathbb{R}^d such that $u(x) = 0$ for every $x \in \mathbb{R}^d \setminus K$. We compute the following using Hölder inequality and writing $u = u \cdot \mathcal{X}_K$

$$|\mathcal{R}u(\theta, s)| = \left| \int_{y \in \theta^\perp} u(s\theta + y) dy \right| \leq \left(\int_{y \in \theta^\perp} |u(s\theta + y)|^p dy \right)^{\frac{1}{p}} |\theta^\perp \cap K|^{\frac{1}{q}}.$$

For any point $x \in \theta^\perp \cap K$ we choose a ball centered at x with radius $\text{diam}(K)$. Hence $\theta^\perp \cap K \subset B(x, \text{diam}(K))$ and

$$|\mathcal{R}u(\theta, s)| \leq C \left(\int_{y \in \theta^\perp} |u(s\theta + y)|^p dy \right)^{\frac{1}{p}} \text{diam}(K)^{\frac{d-1}{q}}$$

where C is a constant related to the volume of the ball and depends only on d . Similarly

to (3.10), using a simple change of variables and Fubini's theorem, we obtain

$$\|\mathcal{R}u\|_{L^p(\mathbb{R})}^p \leq C^p \text{diam}(K)^{\frac{p}{q}(d-1)} \|u\|_{L^p(\mathbb{R}^d)}^p$$

and finally integrating over $\theta \in \mathcal{S}^{d-1}$ we conclude that

$$\|\mathcal{R}u\|_{L^p(\Sigma^d)} \leq C |\mathcal{S}^{d-1}|^{\frac{1}{p}} \text{diam}(K)^{\frac{d-1}{q}} \|u\|_{L^p(\mathbb{R}^d)}.$$

□

Another useful result of the Radon transform is its relation with the Fourier transform \mathcal{F} , known as *Projection-Slice* theorem.

Theorem 3.3.3 (Projection-Slice). *Let $u \in L^1(\mathbb{R}^d)$, then for $\theta \in \mathcal{S}^{d-1}$ and $\xi \in \mathbb{R}$*

$$\mathcal{F}_1(\mathcal{R}u)(\theta, \xi) = (2\pi)^{\frac{d-1}{2}} \mathcal{F}_d u(\xi\theta),$$

where on the left-hand side is the one dimensional Fourier transform and on the right-hand side in the Fourier transform in \mathbb{R}^d , $d > 1$.

An important consequence of the previous theorem is the uniqueness of the Radon transform where the proof is based on the uniqueness of Fourier transform.

Corollary 3.3.4. *If $u_1, u_2 \in L^1(\mathbb{R}^d)$ with $\mathcal{R}u_1(\theta, s) = \mathcal{R}u_2(\theta, s)$ for all $(\theta, s) \in \Sigma^d$ then $u_1(x) = u_2(x)$, $\forall x \in \mathbb{R}^d$.*

Finally, let us briefly describe how to reconstruct our initial unknown function u , once we have obtained all the Radon measurements, namely a function $g \in \Sigma^d$. A reasonable idea is to take the average value of all the line integrals over all lines passing through a point $x \in \mathbb{R}^d$. We assign a value to the point $x \in \mathbb{R}^d$ by *averaging* our projected measurements g *back* to the initial space.

Definition 3.3.5 (Backprojection). Let $g \in L^1(\Sigma^d)$, we define the backprojection operator \mathcal{R}^* as

$$\mathcal{R}^*g(x) = \int_{\mathcal{S}^{d-1}} g(\theta, x \cdot \theta) d\theta. \quad (3.14)$$

It is the adjoint operator of \mathcal{R} and for every $u \in L^1(\mathbb{R}^d)$ and $g \in L^1(\Sigma^d)$ we have that

$$\langle \mathcal{R}u, g \rangle_{L^1(\Sigma^d)} = \langle u, \mathcal{R}^*g \rangle_{L^1(\mathbb{R}^d)}.$$

However, the backprojection operation is not a good approximation of the function u , as it is clearly indicated by the proposition below, see [Nat01, Chapter 1] for its proof.

Proposition 3.3.6. *Let $u \in L^1(\mathbb{R}^d)$, then*

$$\mathcal{R}^*(\mathcal{R}u(\theta, s))(x) = |S^{d-2}| \left(\frac{1}{|x|} * u(x) \right)$$

i.e., a smoothed approximation of u that expands radially from the origin.

In order to overcome this defect, we employ first a filtering on the projected measurements and then apply the backprojection operation. This process is widely known as the *filtered-backprojection* abbreviated as FBP. Let $g = \mathcal{R}u$, the idea is to find a suitable filter V approximating the Dirac δ function and determine v from $V = \mathcal{R}^*v$. Then

$$u \cong V * u = \mathcal{R}^*v * u = \mathcal{R}^*(v * \mathcal{R}u) = \mathcal{R}^*(v * g),$$

where the following property on convolution was used: for $u \in L^1(\mathbb{R}^d)$ and $g \in L^1(\Sigma^d)$

$$(\mathcal{R}^*g) * u = \mathcal{R}^*(g * \mathcal{R}u).$$

The process of finding a suitable filter V and then determine v is described in [NW01, Chapter 5]. One of the most common filter that we also use in our analysis is the *Ram-Lak* filter, see Section 3.7.2. In Figure 3.8, we compare our reconstructions of the characteristic function of a square using the backprojection operation and the filtered backprojection with a Ram-Lak filter. Clearly, there is a radial smoothing expanding from the center of the square when we backproject without using any filter, as it is indicated by Proposition 3.3.6. On the other hand, the filtered backprojection can avoid this artifact and the reconstruction is visually close to the initial image. Filtered backprojection is considered as the most important algorithm in tomographic imaging and is still widely used in clinical applications.

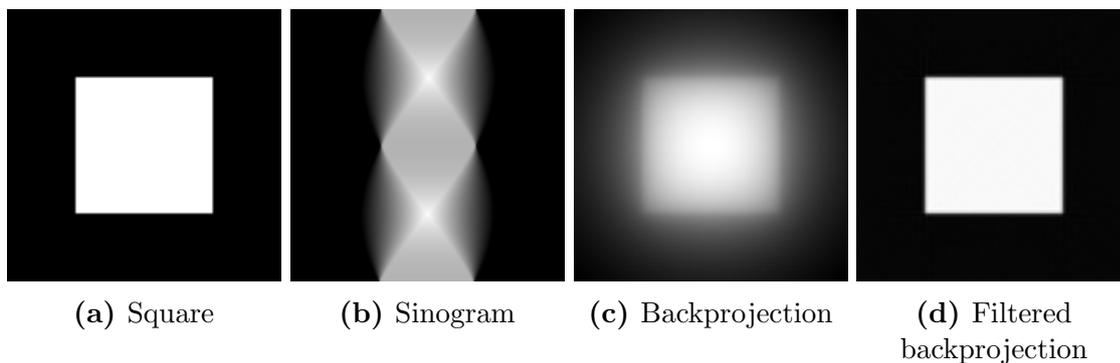


Figure 3.8: Backprojection and filtered backprojection operations on a characteristic function of a square.

3.4 State of the art image reconstruction methods

In (3.1), we reconstruct an image from PET measurements g by *smoothing* both in measurement and image space. Our approach is inspired by an alternating regularisation procedure for PET reconstruction initially introduced by Barbano et. al in [BFS11]. Given possible under sampled and noisy PET measurements $g \in \Sigma^2$, an image u is reconstructed by solving in the discrete setting

$$\min_{\substack{(h,u): h \in \Sigma^2 \\ u = \mathcal{R}^{-1}h}} \alpha |Dh|(\Sigma^2) + \beta \|u\|_1 + \frac{\lambda}{2} \|g - h\|_2^2, \quad (3.15)$$

where $|Dh|(\Sigma^2)$ is the discrete total variation of h in the sinogram space, \mathcal{R}^{-1} is the inverse Radon transform approximated by the FBP and α, β, λ are positive weighting parameters. The image u is forced to be sparse by an ℓ_1 penalty. The main focus of [BFS11] is to study the effect of total variation regularisation on the sinogram, rather than the image as usually done in variational PET reconstruction [SBW⁺08, SBMB09, BBS⁺09, BSB10]. Therefore, in their numerical experiments the effect of the image regularisation is kept low by choosing an appropriate weighting $\alpha \gg \beta$. In [BFS11], it is proved that (3.15) has a unique solution. Moreover, the authors show the effect that the total variation regularisation of the sinogram h has on the reconstructed image u by a computational experiment on a simulated data set. The main idea of adding a total variation prior on the projection space originated in the works of Thirion in [Thi91] and Prince *et al.* in [PW90]. In [Thi91], the author introduces a geometrical based approach in order to detect inner and outer boundaries of an object. Since every point on the sinogram corresponds to a line in the image space with fixed angle and distance from the origin, he proposes to follow continuously *edge* points on the sinogram that will reflect to either inner or outer edge points on the image. In [PW90], Prince and Willsky focus on reconstructing tomographic images by using a Markov random field prior on the sinogram, particularly in the presence of data with a low signal-to-noise ratio (SNR) and limited angle or sparse-angle measurement configurations. Their approach leads to the computation of a smoothed sinogram from which the image u is reconstructed using the filtered backprojection.

Besides adding an additional prior to the sinogram the novelty of our method is that it combines two of the main philosophies encountered in PET reconstruction: direct methods and iterative / variational methods.

- (i) In a direct method, the goal is to perform an one step calculation in order to recover a reconstruction from the PET measurements. Typically, direct methods neglect most of the important features of PET such as noise distribution. For example, FBP is a direct method as was described in Section 3.3. Another direct method presented in

[BFK09], [MFI06] is to compute analytically the inverse Radon transform defined as

$$u(x, y) = \mathcal{R}^{-1}(\mathcal{R}u) := \frac{1}{4i\pi^2} (\partial_x - i\partial_y) \int_0^{2\pi} e^{i\theta} H(\mathcal{R}u)(y \cos \phi - x \sin \phi) d\phi$$

where $H(\cdot)$ denotes the Hilbert transform. They approximate the PET datum $g = \mathcal{R}u$ by cubic splines, whereas the Hilbert transform is approximated by quadrature formulas, see also [KGSF11, KGSF15].

- (ii) On the other hand, iterative methods such as *expectation-maximisation* algorithm (EM), introduced in [DLR] and [SV82] or variational methods such as EM-TV, see [BBS⁺09] are based on the standard Bayesian modelling approach in inverse problems. Prior knowledge in terms of regularity is expressed for the image function u rather than the measurements g . In these cases, the possibility to include statistical noise models is a important asset of iterative and variational methods.

Now, we discuss the Bayesian approach in iterative methods for PET reconstruction. Assume that we have P detectors and denote by X_i , $i = 1, \dots, P$ the random variables assigned to each detector that are responsible to photon counting. Assume also that X_i are pairwise independent and identically distributed. Since, every random variable corresponds to a different detector, the latter assumption is indeed acceptable. Every X_i random variable is Poisson distributed with parameter $(\mathcal{R}u)_i$ with a detected value g_i . Then using (3.6), the measurements $g = \mathcal{R}u$ have conditional probability

$$P(g = g_i | u; (\mathcal{R}u)_i) = \prod_{i=1}^P \frac{(\mathcal{R}u)_i^{g_i}}{g_i!} e^{-(\mathcal{R}u)_i}. \quad (3.16)$$

In order to decide among all possible images u , which is the best estimate of the true object we employ the *maximum-likelihood* (ML) criterion. Therefore, we seek a solution for which the measured data g would have the greatest likelihood $P(g|u)$ i.e., $u \in \underset{u}{\operatorname{argmax}} P(g|u)$. Generally speaking, using information only from the noise distribution generates problems that are ill-posed and additional information on the solution has to be imposed. We can obtain this information using the Bayes' Law and choose an estimate u that maximises the posterior probability density function $P(u|g)$. Hence, we write

$$u \in \underset{u}{\operatorname{argmax}} P(u|g), \quad \text{where} \quad P(u|g) = \frac{P(g|u)P(u)}{P(g)}. \quad (3.17)$$

Since the denominator does not depend on u , it suffices to maximise $P(g|u)P(u)$. Hence, we can enforce to our model a prior probability of the image u , see [CS05, LCA07]. The most frequently used prior densities are the Gibbs priors, i.e.,

$$P(u) = e^{-\alpha J(u)} \quad (3.18)$$

where $\alpha > 0$ is a regularisation parameter and $J(u)$ is a convex energy functional. Instead of maximising $P(g|u)P(u)$, we minimise $-\log(P(g|u)P(u))$. We seek a minimiser of the following problem

$$\operatorname{argmin}_{u \geq 0} \left\{ \sum_{i=1}^P ((\mathcal{R}u)_i - g_i \log(\mathcal{R}u)_i) + \alpha J(u) \right\}, \quad (3.19)$$

where the first term is the so-called *Kullback-Leibler* divergence of $\mathcal{R}u$ and g . This often serves as a motivation to consider the continuous variational problem

$$\operatorname{argmin}_{u \geq 0} \left\{ \int_{\Sigma^2} \mathcal{R}u - g \log(\mathcal{R}u) d\sigma + \alpha J(u) \right\}. \quad (3.20)$$

In the case where $J \equiv 0$, the first optimality condition in (3.20) yields the following iterative scheme, known as EM algorithm

$$u^{k+1} = \frac{u^k}{\mathcal{R}^* \mathbb{1}} \mathcal{R}^* \left(\frac{g}{\mathcal{R}u_k} \right). \quad (3.21)$$

Additionally imposing prior information on the solution, e.g. that the solutions have small total variation with $J(u) = |Du|$, leads to an extension of the EM algorithm, e.g. the EM-TV algorithm. The reader is also referred to [Mül13] and [MBS⁺11] where an extension to the EM-TV algorithm is applied to PET data using the Bregman distance, see Section 3.5.3. Finally, in order to reduce the computational cost of the Kullback-Leibler fidelity, it is useful to replace it with a second order Taylor approximation, see [Saw11, Section 4.5.2] for more details, which results to a weighted-L² squared version i.e.,

$$\int_{\Sigma^2} \frac{(g - \mathcal{R}u)^2}{g} d\sigma.$$

3.5 TV regularisation on image and sinogram

In this section, we will discuss the well-posedness of our minimisation problem. We typically considered image functions u that are defined on an open and bounded domain $\Omega \subset \mathbb{R}^2$, including sufficiently large balls around zero. Our proposed variational model is

$$\operatorname{argmin}_{\substack{u \in \text{BV}(\Omega) \\ u \geq 0 \text{ a.e. in } \Omega}} \left\{ F(u) = \alpha |Du|(\Omega) + \beta |D(\mathcal{R}u)|(\Sigma^2) + \frac{1}{2} \int_{\Sigma^2} \frac{(g - \mathcal{R}u)^2}{g} d\sigma \right\}. \quad (3.22)$$

Prior information for both the image and sinogram spaces is enforced and weighted by positive parameters α and β . Note, that the TV regularisation on the sinogram in (3.22) has a different effect on the reconstructed image compared to regularising in image space ($\beta = 0$) only. Of course, regularisation of the image u enforces a certain regularisation of

the sinogram $\mathcal{R}u$ as well. However, because of the nonlinear character of the total variation regularisation, TV regularisation on the sinogram is not equivalent to a regularisation on the image and vice-versa. In (3.22), the two types of TV regularisation impose different structures in the subgradients of the two terms. This is also emphasised in Section 3.5.3, where the *source condition* (3.41) and the subgradient elements ξ_1, ξ_2 are described.

We start with some observations that are crucial ingredients of the well-posedness analysis for (3.22). In order not to divide by zero in the weighted L^2 norm in (3.22), we first assume that there exists constant $c_1 > 0$ such that

$$0 < c_1 \leq g(\theta, s) \leq \|g\|_{L^\infty(\Sigma)}. \quad (3.23)$$

The constraint (3.23) is not significantly restrictive in most medical experiments. Since u is assumed to be nonnegative, this basically can be achieved if the lines in the Radon transform are confined to those intersecting the support of u , at least in a discretised setting. Moreover, to justify the definition of $F(u)$ in (3.22) over the admissible set $\{u \in \text{BV}(\Omega), u \geq 0 \text{ a.e. in } \Omega\}$ in Theorem 3.5.1 we show that the Radon transform of u is again a function of bounded variation and $\mathcal{R}u \in \text{BV}(\Sigma^2)$ is well defined. To do so it is important to assume that the object we wish to recover is compactly supported. Hence, we may assume that $\text{supp } u \subset B_r \subset \Omega$, where B_r is the ball with radius r centered at the origin. Consequently, (3.9) implies that $\mathcal{R}u(\theta, s) = 0$, when $s \notin (-r, r)$ and the projection space becomes:

$$\Sigma^2 = \{(\theta, s) : \theta = (\cos \phi, \sin \phi) \in \mathcal{S}^1, -r \leq s \leq r, 0 \leq \phi < \pi\}. \quad (3.24)$$

If it is not stated otherwise, we will always assume that the reconstructed image is compactly supported. Notice that, we may allow either negative values on the s variable and not consider the Radon transform for $\phi = \pi$ or allow that $s \geq 0$ and $0 \leq \phi < 2\pi$, see Section 3.3.

3.5.1 BV continuity of the Radon transform

Our first result deals with the continuity of the Radon transform as a mapping operator in the space of functions with bounded variation. A similar result is proved by M. Bergounioux and E. Trélat in [BT10] in three dimensional case for bounded and axially symmetric objects. Under these assumptions, they obtain a nice expression of the Radon transform which can be identified as an Abel transform, see Section 3.5.4 for more details on the Abel transform. They also provide a similar continuity result to Proposition 3.3.2 in [ABT08] for this specific expression of the Radon transform.

In what follows we do not need this symmetry assumption, but prove that the Radon transform is BV continuous for compactly supported function u in two dimensions.

Theorem 3.5.1. *Let $u \in \text{BV}(\Omega)$ and the ball B_r with radius r be its compact support, then $\mathcal{R}u \in \text{BV}(\Sigma^2)$ and the Radon transform is BV continuous on the subspace of functions supported in B_r .*

Proof. Since the Radon transform is L^1 continuous, see (3.10), in order to prove that is BV continuous with respect to $\|\cdot\|_{\text{BV}} = \|\cdot\|_{L^1} + |D(\cdot)|$ norm, it suffices to prove that the total variation of $\mathcal{R}u$ over Σ^2 is finite and bounded by the total variation of u , i.e.,

$$|D(\mathcal{R}u)|(\Sigma^2) = \sup \left\{ \int_{\Omega} \mathcal{R}u(\theta, s) \text{div}h(\theta, s) d\theta ds : h \in \mathcal{C}_c^1(\Sigma^2, \mathbb{R}^2), \|h\|_{\infty} \leq 1 \right\} < \infty.$$

Using the local coordinates (3.12), see also Figure 3.6 with $t, s \in [-r, r]$, we compute that

$$\begin{aligned} \int_{\Sigma^2} \mathcal{R}u(\theta, s) \text{div}h(\theta, s) d\theta ds &= \int_0^{\pi} \int_{-r}^r \int_{-r}^r u(s \cos \phi - t \sin \phi, s \sin \phi + t \cos \phi) \\ &\quad \text{div}h(\phi, s) dt ds d\phi \\ &= \int_0^{\pi} \int_{-r}^r \int_{-r}^r u(x, y) \left[\nabla h_1 \cdot \vec{\alpha} + \nabla h_2 \cdot \vec{\theta} \right] dx dy d\theta \end{aligned}$$

where $\vec{\alpha} = (-y, x)$, $\vec{\theta} = (\cos \phi, \sin \phi)$. Set $\vec{G}(x, y) = (G_1(x, y), G_2(x, y))$ with

$$\begin{aligned} G_1(x, y) &:= \int_0^{\pi} -yh_1(\phi, x \cos \phi + y \sin \phi) + h_2(\phi, x \cos \phi + y \sin \phi) \cos \phi d\phi \\ G_2(x, y) &:= \int_0^{\pi} xh_1(\phi, x \cos \phi + y \sin \phi) + h_2(\phi, x \cos \phi + y \sin \phi) \sin \phi d\phi \end{aligned}$$

then,

$$\text{div}G(x, y) = \frac{\partial G_1(x, y)}{\partial x} + \frac{\partial G_2(x, y)}{\partial y} = \int_0^{\pi} \left(\nabla h_1 \cdot \vec{\alpha} + \nabla h_2 \cdot \vec{\theta} \right) d\phi$$

and finally by Fubini's theorem

$$\int_0^{\pi} \int_{-r}^r \int_{-r}^r u(x, y) \left[\nabla h_1 \cdot \vec{\alpha} + \nabla h_2 \cdot \vec{\theta} \right] dx dy d\theta = \int_{-r}^r \int_{-r}^r u(x, y) \text{div}G(x, y) dx dy.$$

The function $G \in \mathcal{C}^1(\Omega, \mathbb{R}^2)$ and if we consider $G \cdot \mathcal{X}_{B_r}$, then $G \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2)$. Moreover,

$$|G_1(x, y)| \leq \int_0^{\pi} |y||h_1| + |h_2| \leq \pi \|h\|_{\infty} (1 + |y|) \leq \pi \|h\|_{\infty} (1 + r) = C$$

and similarly $|G_2(x, y)| \leq C$. Finally, let

$$A := \int_{-r}^r \int_{-r}^r u(x, y) \text{div}G(x, y) dx dy = C \int_{-r}^r \int_{-r}^r u(x, y) \text{div}\left(\frac{G(x, y)}{C}\right) dx dy$$

and

$$B := \int_{\Sigma^2} \mathcal{R}u(\theta, s) \text{div}h(\theta, s) ds d\theta.$$

Then, taking the supremum over all $G \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2)$ with $\|G/C\|_\infty \leq 1$, we have that $B = C \cdot |Du|(\Omega)$. Similarly, for all $h \in \mathcal{C}_c^1(\Sigma^2, \mathbb{R}^2)$ with $\|h\|_\infty \leq 1$, we conclude that

$$|D(\mathcal{R}u)|(\Sigma^2) \leq \pi(1+r)|Du|(\Omega) < \infty.$$

Therefore, $\mathcal{R}u \in \text{BV}(\Sigma^2)$ and the variation coincides with the total variation $|D(\mathcal{R}u)|(\Sigma^2)$. We deduce from (3.10) that

$$\|\mathcal{R}u\|_{\text{BV}(\Sigma^2)} \leq \pi(1+r)\|u\|_{\text{BV}(\Omega)}.$$

□

3.5.2 Well-posedness of the model

Next, we show existence and uniqueness of a minimiser for the problem (3.22). The following proof uses standard tools of variational calculus and mainly the existence over the BV space is based on the Poincaré-Wirtinger inequality, see for example [AV94] and [Ves01].

Theorem 3.5.2. *Let $\alpha > 0$, $\beta \geq 0$ and $g \in L^\infty(\Sigma^2)$ which satisfies (3.23). Then, the functional $F(u)$ in (3.22) is lower semicontinuous and strictly convex and the minimisation problem (3.22) attains a unique solution $u \in \text{BV}(\Omega) \cap L_+^1(\Omega)$.*

Proof. Let $(u_n)_n \in \text{BV}(\Omega)$ be a minimising sequence of nonnegative functions, then there exists a constant $C_1 > 0$ such that

$$F(u_n) = \alpha|Du_n|(\Omega) + \beta|D(\mathcal{R}u_n)|(\Sigma^2) + \frac{1}{2} \int_{\Sigma^2} \frac{(g - \mathcal{R}u_n)^2}{g} d\sigma < C_1. \quad (3.25)$$

Let $\bar{u}_n = \frac{1}{|\Omega|} \int_{\Omega} u_n dx$, then by the Poincaré-Wirtinger inequality, see Theorem 2.3.4, we can find a constant $C_2 > 0$ such that

$$\|u_n - \bar{u}_n\|_{L^2(\Omega)} \leq C_2|Du_n|(\Omega) \quad (3.26)$$

and

$$\|u_n\|_{L^2(\Omega)} \leq C_2|Du_n|(\Omega) + \|\bar{u}_n\|_{L^2(\Omega)} = C_2|Du_n|(\Omega) + \left| \int_{\Omega} u_n dx \right|. \quad (3.27)$$

We set $v_n = u_n - \bar{u}_n$ and by (3.23)

$$C_1 > \int_{\Sigma} \frac{(g - \mathcal{R}u_n)^2}{g} \geq \frac{1}{\|g\|_{L^\infty(\Sigma^2)}} \|g - \mathcal{R}u_n\|_{L^2(\Sigma^2)}^2$$

we proceed to the following calculations

$$\begin{aligned}
 \|\mathcal{R}\bar{u}_n\|_{L^2(\Sigma^2)} &= \|\mathcal{R}\bar{u}_n + g - g\|_{L^2(\Sigma^2)} = \|\mathcal{R}u_n - \mathcal{R}v_n + g - g\|_{L^2(\Sigma^2)} \\
 &\leq C_1^{\frac{1}{2}} \|g\|_{L^\infty(\Sigma^2)}^{\frac{1}{2}} + \|\mathcal{R}v_n\|_{L^2(\Sigma^2)} + \|g\|_{L^2(\Sigma^2)} \\
 &\leq C_1^{\frac{1}{2}} \|g\|_{L^\infty(\Sigma^2)}^{\frac{1}{2}} + C_3 \|g\|_{L^\infty(\Sigma^2)} + \|\mathcal{R}\| \|v_n\|_{L^2(\Omega)} \leq C_4
 \end{aligned}$$

where the constant C_3 depends on the sinogram space Σ^2 and the last term is bounded due to (3.26) and Proposition 3.3.2. Since,

$$\left| \frac{1}{|\Omega|} \int_{\Omega} u_n dx \right| \cdot \|\mathcal{R}(\mathcal{X}_\Omega)\|_{L^2(\Sigma^2)} = \|\mathcal{R}\bar{u}_n\|_{L^2(\Sigma^2)} \quad (3.28)$$

without loss of generality, we may assume that the image domain Ω is a unit square, then $\mathcal{R}(\mathcal{X}_\Omega) \neq 0$, see [Pou10, Chapter 8] and conclude that $|\int_{\Omega} u_n dx|$ is uniformly bounded. Hence, by (3.27), (u_n) is $L^2(\Omega)$ bounded and since $\Omega \subset \mathbb{R}^2$ is a bounded and open set, it is also $L^1(\Omega)$ bounded. Moreover, since the Radon transform is L^2 continuous for functions with compact support, see Proposition (3.3.2), using the L^1 continuity, we can summarise the following results so far: The sequence $(u_n)_{n \in \mathbb{N}}$ is bounded in $L^1(\Omega)$ and $|Du_n|(\Omega) < \infty$ i.e., is $BV(\Omega)$ bounded, hence we obtain a subsequence $(u_{n_k})_{k \in \mathbb{N}}$, $u \in BV(\Omega)$ such that u_{n_k} converges weakly* to u . Also, u_{n_k} converges weakly to u in $L^2(\Omega)$. Then,

$$\begin{aligned}
 \mathcal{R}u_{n_k} &\rightarrow \mathcal{R}u \text{ in } L^1(\Sigma^2) \\
 \mathcal{R}u_{n_k} &\rightharpoonup \mathcal{R}u \text{ in } L^2(\Sigma^2).
 \end{aligned} \quad (3.29)$$

Hence, for $h \in C_c^1(\Sigma^2, \mathbb{R}^2)$ we obtain

$$\liminf_{k \rightarrow \infty} |D(\mathcal{R}u_{n_k})|(\Sigma^2) \geq \liminf_{k \rightarrow \infty} \int_{\Sigma^2} \mathcal{R}u_{n_k} \operatorname{div} h \, d\sigma \geq \int_{\Sigma^2} \mathcal{R}u \operatorname{div} h \, d\sigma$$

and taking the supremum over all $\|h\|_{L^\infty(\Sigma^2)} \leq 1$ we conclude that

$$|D(\mathcal{R}u)|(\Sigma^2) \leq \liminf_{k \rightarrow \infty} |D(\mathcal{R}u_{n_k})|(\Sigma^2).$$

The weak lower semicontinuity of the L^2 norm and the lower semicontinuity of total variation semi-norm, see Section 2.3 for both the image and the projection space imply that

$$F(u) \leq \liminf_{k \rightarrow \infty} F(u_{n_k}).$$

Finally, in order to prove uniqueness let $0 \leq u_1, u_2 \in BV(\Omega)$ be two minimisers. If $\mathcal{R}u_1 \neq \mathcal{R}u_2$, then the strict convexity of the weighted L^2 fidelity term together with the

convexity of the total variation of $\mathcal{R}u$ imply that:

$$F\left(\frac{u_1 + u_2}{2}\right) < \frac{F(u_1)}{2} + \frac{F(u_2)}{2} = \inf_{\substack{u \in \text{BV}(\Omega) \\ u \geq 0 \text{ a.e.}}} F(u)$$

which is a contradiction. Hence, $\mathcal{R}u_1 = \mathcal{R}u_2$ and using Corollary 3.3.4, we conclude that $u_1 = u_2$. \square

Further, we discuss the stability of problem (3.22) in terms of a small perturbation on the data. Following [AV94], we consider a perturbation on the projection space i.e.,

$$g_n = g + \tau_n \text{ with } \|\tau_n\|_{L^2(\Sigma^2)} \rightarrow 0 \quad (3.30)$$

and define the corresponding minimisation problem on the perturbed functionals:

$$\underset{\substack{u \in \text{BV}(\Omega) \\ u \geq 0 \text{ a.e.}}}{\text{argmin}} \left\{ F^n(u) = \alpha |Du|(\Omega) + \beta |D(\mathcal{R}u)|(\Sigma^2) + \frac{1}{2} \int_{\Sigma^2} \frac{(g_n - \mathcal{R}u)^2}{g_n} d\sigma \right\}. \quad (3.31)$$

For (3.31) to be well-defined, we assume an L^∞ bound on τ_n such that g_n is still positive. More precisely we assume that

$$0 < c_1 \leq g_n(\theta, s) \leq \|g\|_{L^\infty(\Sigma^2)} + \varepsilon, \text{ for all } n \geq 1, \quad (3.32)$$

which is the same as assuming that the perturbations τ_n are bounded from above by a small enough constant. Then, by the previous theorem we have that both F^n and F are lower semicontinuous, strictly convex with unique minimisers u_n and u^* respectively. In a sense, we will prove that for a small change on our data g , our solution's behaviour does not change significantly. Before, we proceed with the stability analysis we need to ensure that the functional is indeed BV-coercive. That is coercive with respect to the bounded variation norm $\|u\|_{\text{BV}(\Omega)} = \|u\|_{L^1(\Omega)} + |Du|(\Omega)$, rather than the total variation seminorm only.

Lemma 3.5.3. *Let $g \in L^\infty(\Sigma^2)$ a strictly positive and bounded function, then the functional F in (3.22) is BV coercive i.e., there exists a constant $C > 0$ such that*

$$F(u) \geq C \|u\|_{\text{BV}(\Omega)}. \quad (3.33)$$

Proof. Let $u \geq 0$ a.e with $u \in \text{BV}(\Omega)$ and consider $v = u - \bar{u}$. Then, using again Poincaré-Wirtinger inequality for $1 \leq p \leq 2$, we obtain that

$$\|v\|_{L^p(\Omega)} \leq C_1 |Dv|(\Omega)$$

and the corresponding estimate for the BV norm holds:

$$\begin{aligned} \|u\|_{\text{BV}(\Omega)} &= \|u\|_{L^1(\Omega)} + |Du|(\Omega) \leq \|u - \bar{u}\|_{L^1(\Omega)} + \|\bar{u}\|_{L^1(\Omega)} + |Dv|(\Omega) \\ &\leq \|\bar{u}\|_{L^1(\Omega)} + (C_1 + 1)|Dv|(\Omega). \end{aligned} \quad (3.34)$$

Note that in the above calculations we have used the fact that $|Du|(\Omega) = |Dv|(\Omega)$. Moreover, see (3.28) we know that there exists a constant $C_2 > 0$ such that

$$\|\mathcal{R}\bar{u}\|_{L^2(\Sigma^2)} = C_2 \|\bar{u}\|_{L^1(\Omega)}$$

since $\mathcal{R}(\mathcal{X}_\Omega) \neq 0$. Hence, we can derive the following bound:

$$\begin{aligned} F(u) &= \alpha|Du|(\Omega) + \beta|D(\mathcal{R}u)|(\Sigma^2) + \frac{1}{2} \int_{\Sigma^2} \frac{(g - \mathcal{R}u)^2}{g} d\sigma \geq \alpha|Dv|(\Omega) + \frac{1}{2\|g\|_{L^\infty(\Sigma^2)}} \|g - \mathcal{R}u\|_{L^2(\Sigma^2)}^2 \\ &\geq \alpha|Dv|(\Omega) + \frac{1}{2\|g\|_{L^\infty(\Sigma^2)}} \|g - \mathcal{R}v - \mathcal{R}\bar{u}\|_{L^2(\Sigma^2)}^2 \\ &\geq \alpha|Dv|(\Omega) + \frac{1}{2\|g\|_{L^\infty(\Sigma^2)}} \left(\|g - \mathcal{R}v\|_{L^2(\Sigma^2)} - \|\mathcal{R}\bar{u}\|_{L^2(\Sigma^2)} \right)^2 \\ &\geq \alpha|Dv|(\Omega) + \frac{1}{2\|g\|_{L^\infty(\Sigma^2)}} \|\mathcal{R}\bar{u}\|_{L^2(\Sigma^2)} \left(\|\mathcal{R}\bar{u}\|_{L^2(\Sigma^2)} - 2\|g - \mathcal{R}v\|_{L^2(\Sigma^2)} \right) \\ &\geq \alpha|Dv|(\Omega) + \frac{C_2 \|\bar{u}\|_{L^1(\Omega)}}{2\|g\|_{L^\infty(\Sigma^2)}} \left(C_2 \|\bar{u}\|_{L^1(\Omega)} - 2 \left(\|\mathcal{R}\| C_1 |Dv|(\Omega) + \|g\|_{L^2(\Sigma^2)} \right) \right) \end{aligned} \quad (3.35)$$

Setting

$$A = C_2 \|\bar{u}\|_{L^1(\Omega)} - 2 \left(C_1 \|\mathcal{R}\| |Dv|(\Omega) + \|g\|_{L^2(\Sigma^2)} \right)$$

we consider 2 cases:

(a) If $A \geq 1$, then using (3.35), one can prove that

$$\frac{C_1 + 1}{\alpha} F(u) \geq (C_1 + 1)|Dv| \quad \text{and} \quad \frac{2\|g\|_{L^2(\Sigma^2)}}{C_2} F(u) \geq \|\bar{u}\|_{L^1(\Omega)}$$

and hence by (3.34)

$$F(u) \left(\frac{C_1 + 1}{\alpha} + \frac{2\|g\|_{L^2(\Sigma^2)}}{C_2} \right) \geq \|u\|_{\text{BV}(\Omega)}. \quad (3.36)$$

(b) If $A \leq 1$, then

$$\|\bar{u}\|_{L^1(\Omega)} \leq \frac{1 + 2 \left(\|\mathcal{R}\| C_1 |Dv|(\Omega) + \|g\|_{L^2(\Sigma^2)} \right)}{C_2}$$

and using equation (3.34) we derive that:

$$\|u\|_{BV(\Omega)} - \frac{1 + 2\|g\|_{L^2(\Sigma^2)}}{C_2} \leq \left(\frac{2C_1\|\mathcal{R}\|}{C_2} + C_1 + 1\right)|Dv| \leq \frac{K}{\alpha}F(u) \quad (3.37)$$

where $K = \frac{2C_1\|\mathcal{R}\|}{C_2} + C_1 + 1$. From equations (3.36), (3.37), we have that the functional F is BV coercive. \square

Moreover, we can prove that given constants $C > 0$ and $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that

$$|F^n(u) - F(u)| \leq \varepsilon \text{ for } n \geq n_0 \text{ and } \|u\|_{BV(\Omega)} \leq C. \quad (3.38)$$

Indeed, after few computations using (3.30) and the constraints on g and g_n we conclude that

$$\begin{aligned} \frac{1}{2} \left| \int_{\Sigma^2} \frac{(g_n - \mathcal{R}u)^2}{g_n} - \frac{(g - \mathcal{R}u)^2}{g} - \frac{(g_n - g)^2}{g_n} d\sigma \right| &\leq \frac{1}{2c_1^2} \left| \int_{\Sigma^2} g^2 \tau_n - \tau_n (\mathcal{R}u)^2 d\sigma \right| \\ &\leq \frac{1}{2c_1^2} \|\tau_n\|_{L^2(\Sigma)} \|g^2 - (\mathcal{R}u)^2\|_{L^2(\Sigma^2)}. \end{aligned}$$

The continuity of Radon transform for L^2 functions with compact support, see Proposition 3.3.2 and that $BV \hookrightarrow L^2$ continuously, imply that we can find an appropriate constant such that (3.38) is valid since the third term on the left-hand side vanishes for large n . With these preparations we can prove the following weak stability result for minimisers of (3.22).

Theorem 3.5.4. *Let $0 < u_n, u^* \in BV(\Omega)$ be the minimisers of the functionals F^n and F defined in (3.31) and (3.22) respectively. Then*

$$u_n \rightharpoonup u^* \text{ in } L^2(\Omega). \quad (3.39)$$

Proof. Observe that $F^n(u_n) \leq F^n(u^*)$ and using (3.38) we have that

$$\liminf_{n \rightarrow \infty} F^n(u_n) \leq \limsup_{n \rightarrow \infty} F^n(u_n) \leq F(u^*) < \infty$$

Lemma 3.5.3 implies that $(u_n)_{n \in \mathbb{N}}$ is BV bounded. Assume now that (3.39) is not true, then there exists a subsequence u_{n_k} which converges weakly to some $u \neq u^*$ in $L^2(\Omega)$. Hence, by the lower semicontinuity of F , equivalently weak lower semicontinuity since F is convex,

$$\begin{aligned} F(u) &\leq \liminf_{n \rightarrow \infty} F(u_{n_k}) \\ &= \liminf_{k \rightarrow \infty} (F(u_{n_k}) - F^{n_k}(u_{n_k})) + \liminf_{k \rightarrow \infty} F^{n_k}(u_{n_k}) \leq F(u^*) \end{aligned}$$

which is a contradiction to the uniqueness of minimiser of F . \square

3.5.3 Error analysis using the Bregman distance

In the following we discuss a similar approach as presented in [BO04] for deriving an error estimate for our model (3.22) in terms of the Bregman distance. Let us note that what follows holds for a more general minimisation problem

$$\operatorname{argmin}_{u \in X} \left\{ F(u) = \alpha J(u) + \beta J(\mathcal{R}u) + \frac{1}{2} \int_{\Sigma^2} \frac{(g - \mathcal{R}u)^2}{g} d\sigma \right\}, \quad (3.40)$$

where $J : X \rightarrow \mathbb{R}$ is a convex functional and X is a Banach space such that $\mathcal{R} : X \rightarrow L^2(\Sigma^2) \cap X$ is a bounded operator. Before we proceed with proving an error estimate for (3.40), we first restate the terminologies found in [BO04] of a *minimising solution*, the *source-condition* and the notion of *Bregman distance* for a convex functional.

Definition 3.5.5. An element $\tilde{u} \in X$ is called a minimising solution of $\mathcal{R}u = g$ with respect to the functional $J : X \rightarrow \mathbb{R}$ if:

- (i) $\mathcal{R}\tilde{u} = g$,
- (ii) $J(\tilde{u}) \leq J(v)$, $\forall v \in X$, $\mathcal{R}v = g$.

Based on the definition of the subdifferential of J at u , see Definition 2.4.6, we consider the following condition known as the *source condition* for an element \tilde{u}

$$\exists \tilde{w} \in L^2(\Sigma^2) \text{ such that } \mathcal{R}^* \tilde{w} \in \partial J(\tilde{u}). \quad (3.41)$$

One of the importance of the source condition is that the elements satisfying the source condition (3.41) are exactly the minimisers of the (3.40) for $\beta = 0$, see [BO04, Prop. 1]. Next, we recall the Bregman distance for a convex functional J together with some of its basic properties as it was first introduced in [Bre67].

Definition 3.5.6. Let X be a Banach space and $J : X \rightarrow \overline{\mathbb{R}}$ a convex and proper functional with non empty subdifferential ∂J . The Bregman distance is defined as

$$D_J(u, v) = \{J(u) - J(v) - \langle \xi, u - v \rangle \mid \xi \in \partial J(v)\}.$$

Moreover, the Bregman distance for a specify subgradient $\xi \in \partial J(v) \subset X^*$, $v \in X$ is defined as $D_J^\xi : X \times X \rightarrow \mathbb{R}^+$ with

$$D_J^\xi(u, v) = J(u) - J(v) - \langle \xi, u - v \rangle, \quad (3.42)$$

where $\langle \cdot, \cdot \rangle_{X, X^*}$ is the standard duality product between X and its dual X^* .

Let us mention few properties on the Bregman distance. For instance, if J is Gâteaux differentiable then the subdifferential of J contains a unique element ξ and in that case $D_J^\xi(u, v) = 0$, see Definition 2.4.5. On the other hand, if J is not Gâteaux differentiable then the subdifferential is multivalued and each element $d \in D_J^\xi(u, v)$ represents a distance between u and v . One can immediately observe that the Bregman distance is not a metric in the usual sense, since it is not in general symmetric or the triangular inequality is not true. Moreover, for not strictly convex functionals, it is possible to have $0 \in D_J^\xi(u, v)$ for $u \neq v$ and also there is no guarantee that $D_J^\xi(u, v)$ is nonempty since the subdifferential can be nonempty. However, it produces a measure of how close are the elements $u, v \in X$, see Figure 3.9. More precisely, it is a difference between the value of the tangent at v evaluated at u and the value of J at u .

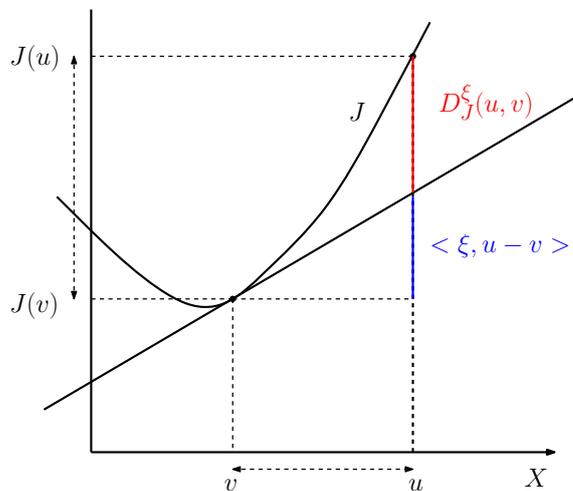


Figure 3.9: Bregman distance $D_J^\xi(u, v)$.

Let us mention some basic properties on the Bregman distance and their proofs are based in the definitions of Bregman distance and the subdifferential of functional J .

- (1) $D_J^\xi(u, u) = 0$, $\forall u \in X$.
- (2) $D_J^\xi(u, v) \geq 0$, since $\xi \in \partial J(v)$.
- (3) If J is a strictly convex functional then $D_J^\xi(u, v) = 0 \Leftrightarrow u = v$.
- (4) If $w \in X$ lies on the line segment between u, v then $D_J^\xi(u, v) \geq D_J^\xi(w, v)$.

Bregman distance is strongly related to the source condition (3.41) through the following observation. With a specific subgradient ξ satisfying the source condition we have $D_J^\xi(u, \tilde{u}) = J(u) - J(\tilde{u}) - \langle \xi, u - \tilde{u} \rangle_{X, X^*} = J(u) - J(\tilde{u}) - \langle \tilde{w}, \mathcal{R}(u - \tilde{u}) \rangle_{X, X^*}$, hence it provides a relation between the error in the regularisation functionals and the output error

$\mathcal{R}u - \mathcal{R}\tilde{u}$. Now, we can derive an estimate for the difference of a minimising solution \tilde{u} in Definition 3.5.5 and a regularised solution \hat{u} of (3.40).

Let $\alpha > 0, \beta \geq 0$ and the data g fulfil (3.23). Then, for a minimiser \hat{u} of (3.40) and the exact solution \tilde{u} satisfying $\mathcal{R}\tilde{u} = f$ with a fixed noise bound $\|g - f\|_{L^2(\Sigma^2)} \leq \delta$ from the exact data f , we have

$$\begin{aligned} \alpha J(\hat{u}) + \beta J(\mathcal{R}\hat{u}) + \frac{\|g - \mathcal{R}\hat{u}\|_{L^2(\Sigma^2)}^2}{2\|g\|_{L^\infty(\Sigma^2)}} &\leq \alpha J(\tilde{u}) + \beta J(\mathcal{R}\tilde{u}) + \frac{\delta^2}{2c_1} \Leftrightarrow \\ \alpha D_J^{\xi_1}(\hat{u}, \tilde{u}) + \alpha \langle \xi_1, \hat{u} - \tilde{u} \rangle + \beta D_J^{\xi_2}(\mathcal{R}\hat{u}, f) + \beta \langle \xi_2, \mathcal{R}\hat{u} - f \rangle + \frac{\|g - \mathcal{R}\hat{u}\|_{L^2(\Sigma^2)}^2}{2\|g\|_{L^\infty(\Sigma^2)}} &\leq \frac{\delta^2}{2c_1} \end{aligned}$$

where, we have used the corresponding Bregman distances related to the functional J and their subgradients ξ_1, ξ_2 regarding the image and the sinogram regularisation. We also require that

$$\partial(J(u) + J(\mathcal{R}u)) = \partial J(u) + \partial(J(\mathcal{R}u)) \quad (3.43)$$

holds, subject to the assumption that the related *effective domains* have a common point, that is

$$\text{dom}J(u) \cap \text{dom}J(\mathcal{R}u) \neq \emptyset \text{ for some } u \in X \quad (3.44)$$

see Proposition 2.4.7. In our case, this is valid due to Theorem 3.5.1. Let

- (i) $\xi_1 \in \partial J(\tilde{u})$
- (ii) $\xi_2 \in \partial(J(\mathcal{R}\tilde{u})) = \mathcal{R}^*(\partial J(\mathcal{R}\tilde{u})) \in \mathcal{R}^*w_2$

Moreover, assume that the source condition (3.41) is satisfied with respect to J , that is

$$\exists \xi_1 \in \partial J(\tilde{u}) \text{ s.t. } \xi_1 = \mathcal{R}^*w_1, w_1 \in L^2(\Sigma^2)$$

Then, by generalised Young's inequality, that is for every $\varepsilon > 0$ we have

$$ab \leq \frac{a^2}{2\varepsilon} + \frac{\varepsilon b^2}{2},$$

we conclude that

$$\begin{aligned} \alpha D_J^{\xi_1}(\hat{u}, \tilde{u}) + \alpha \langle \xi_1, \hat{u} - \tilde{u} \rangle + \beta D_J^{\xi_2}(\mathcal{R}\hat{u}, f) + \beta \langle \xi_2, \mathcal{R}\hat{u} - f \rangle + \frac{\|g - \mathcal{R}\hat{u}\|_{L^2(\Sigma^2)}^2}{2\|g\|_{L^\infty(\Sigma^2)}} &\leq \frac{\delta^2}{2c_1} \Leftrightarrow \\ \alpha D_J^{\xi_1}(\hat{u}, \tilde{u}) + \beta D_J^{\xi_2}(\mathcal{R}\hat{u}, f) + \langle \alpha w_1 + \beta \xi_2, \mathcal{R}\hat{u} - f \rangle + \frac{\|g - \mathcal{R}\hat{u}\|_{L^2(\Sigma^2)}^2}{2\|g\|_{L^\infty(\Sigma^2)}} &\leq \frac{\delta^2}{2c_1} \Leftrightarrow \\ \alpha D_J^{\xi_1}(\hat{u}, \tilde{u}) + \beta D_J^{\xi_2}(\mathcal{R}\hat{u}, f) + \frac{\|g - \mathcal{R}\hat{u}\|_{L^2(\Sigma^2)}^2}{2\|g\|_{L^\infty(\Sigma^2)}} &\leq \frac{\delta^2}{2c_1} + \langle \alpha w_1 + \beta \xi_2, f - \mathcal{R}\hat{u} + g - g \rangle \Leftrightarrow \end{aligned}$$

$$\begin{aligned} \alpha D_J^{\xi_1}(\hat{u}, \tilde{u}) + \beta D_J^{\xi_2}(\mathcal{R}\hat{u}, f) + \frac{\|g - \mathcal{R}\hat{u}\|_{L^2(\Sigma^2)}^2}{2\|g\|_{L^\infty(\Sigma^2)}} &\leq \frac{\delta^2}{2c_1} + \frac{\|\alpha w_1 + \beta \xi_2\|_{L^2(\Sigma^2)}^2}{2\varepsilon} + \frac{\varepsilon}{2} \|g - \mathcal{R}\hat{u}\|_{L^2(\Sigma^2)}^2 \\ &\quad + \frac{\|\alpha w_1 + \beta \xi_2\|_{L^2(\Sigma^2)}^2}{2\varepsilon} + \frac{\varepsilon \|f - g\|_{L^2(\Sigma^2)}^2}{2} \Leftrightarrow \\ \alpha D_J^{\xi_1}(\hat{u}, \tilde{u}) + \beta D_J^{\xi_2}(\mathcal{R}\hat{u}, f) + \frac{\|g - \mathcal{R}\hat{u}\|_{L^2(\Sigma^2)}^2}{2\|g\|_{L^\infty(\Sigma^2)}} &\leq \frac{\delta^2}{2c_1} + \frac{\|\alpha w_1 + \beta \xi_2\|_{L^2(\Sigma^2)}^2}{\varepsilon} + \frac{\varepsilon}{2} \|g - \mathcal{R}\hat{u}\|_{L^2(\Sigma^2)}^2 + \frac{\varepsilon \delta^2}{2} \end{aligned}$$

Hence, for $\varepsilon = \|g\|_{L^\infty(\Sigma)}^{-1} > 0$ we have

$$D_J^{\xi_1}(\hat{u}, \tilde{u}) + \frac{\beta}{\alpha} D_J^{\xi_2}(\mathcal{R}\hat{u}, f) \leq \frac{\tilde{c}_1 \delta^2}{\alpha} + \alpha \|g\|_{L^\infty(\Sigma^2)} \left\| w_1 + \frac{\beta}{\alpha} \mathcal{R}^* w_2 \right\|_{L^2(\Sigma^2)}^2$$

and prove the following theorem:

Theorem 3.5.7. *Let $\delta > 0$ be the noise bound related to the exact data f and the noise data g . Moreover let (3.43) holds. If \hat{u} is a minimiser of (3.40) and \tilde{u} the exact solution of $\mathcal{R}\tilde{u} = f$ which satisfies the source condition (3.41), then for $\alpha > 0, \beta \geq 0$ we have the following estimate:*

$$D_J^{\xi_1}(\hat{u}, \tilde{u}) + \frac{\beta}{\alpha} D_J^{\xi_2}(\mathcal{R}\hat{u}, f) \leq \frac{\tilde{c}_1 \delta^2}{\alpha} + \alpha \|g\|_{L^\infty(\Sigma^2)} \left\| w_1 + \frac{\beta}{\alpha} \mathcal{R}^* w_2 \right\|_{L^2(\Sigma^2)}^2 \quad (3.45)$$

where $\tilde{c}_1 = \frac{c_1 + \|g\|_{L^\infty(\Sigma^2)}}{2c_1 \|g\|_{L^\infty(\Sigma^2)}}$.

This result describes the connection between a regularised solution and a minimising solution related to the regularised functional J . Finally, the estimate above contains the case of $\beta = 0$ where it has been proved in [BO04, Theorems 1,2].

3.5.4 An explicit example of TV regularisation on the sinogram

Before we proceed with the numerical simulations of (3.22), we discuss how a regularised solution in the projection space behaves in terms of an appropriate positive parameter β . In particular, we derive an explicit solution of the weighted ROF minimisation problem on the sinogram, that is

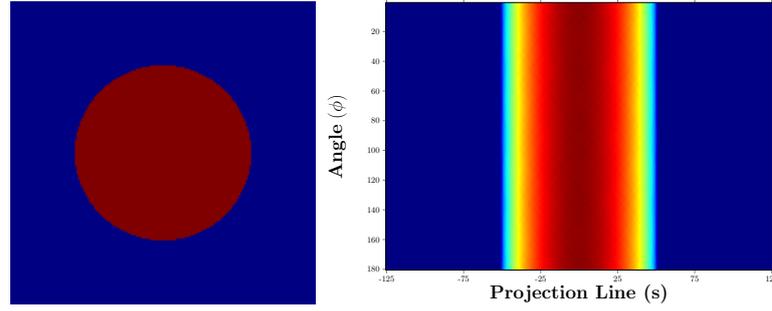
$$\operatorname{argmin}_{\substack{v \in \text{BV}(\Sigma^2) \\ v \geq 0 \text{ a.e}}} \left\{ \beta |Dv|(\Sigma^2) + \frac{1}{2} \int_{\Sigma^2} \frac{(g-v)^2}{g} d\sigma \right\}. \quad (3.46)$$

We examine the sinogram of the characteristic function u of a ball with radius r centered at the origin. Therefore, in this case it is easy to compute the corresponding

Radon transform, where there is no angle dependence due to its radial symmetry. Let

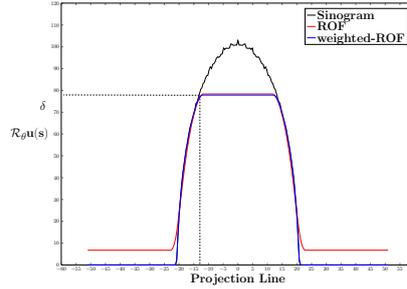
$$u(x, y) = \begin{cases} 1, & \text{if } x^2 + y^2 \leq r \\ 0, & \text{otherwise} \end{cases} \quad (3.47)$$

$$\text{then, } g(\theta, s) = \mathcal{R}_\theta u(s) = \begin{cases} 2\sqrt{r^2 - s^2}, & \text{for } |s| < r \\ 0, & \text{otherwise.} \end{cases} \quad (3.48)$$



(a) $u = \mathcal{X}_{B_r}$

(b) Sinogram g



(c) 45° angle profiles: The loss of contrast from the maximum value $2r$ is δ

Figure 3.10: The characteristic function of a ball with $r = 50.5$ and its sinogram given in (3.48). The total variation reconstructions in 45° angle profiles with the weighted L^2 and the standard L^2 fidelities.

In Figure 3.10, we present $u = \mathcal{X}_{B_r}$ with radius $r = 50.5$ and the corresponding sinogram g . We also put emphasis on the weighted L^2 fidelity compared to the classical ROF regularisation on the sinogram space, namely

$$\operatorname{argmin}_{v \in \text{BV}(\Sigma^2)} \beta |Dv|(\Sigma^2) + \frac{1}{2} \|g - v\|_{L^2(\Sigma^2)}^2.$$

We observe that the weight g on (3.46) allows on the solution v to respect the geometry

of the initial sinogram and for example vanish outside the support $(-r, r)$. We compare the two types of total variation reconstructions where the β 's used, have been selected appropriately in order to produce the same loss of contrast. Since, there is no angle dependence on the sinogram, it suffices to present for demonstration reasons only one angle profile of the sinogram.

Now, we proceed by finding the optimality conditions for (3.46). These conditions allow us to identify formally the structure of the solution. We first make the following ansatz for a solution of (3.46) which is

$$v(s) = \begin{cases} \delta = g(\kappa), & \text{for } |s| \leq \kappa, \\ g(s), & \text{for } \kappa < |s| < r, \\ 0, & \text{otherwise.} \end{cases} \quad (3.49)$$

The solution v is zero outside the interval $(-r, r)$, equal to the data in some subinterval and constant δ in $(-\kappa, \kappa)$ for $\kappa > 0$ representing the loss of contrast from the maximum value of the sinogram. Notice that from (3.48), the maximum value of the sinogram is obtained at $s = 0$ and is $2r$.

The sinogram $g \in \mathcal{C}(-r, r)$ and a solution v of (3.46) is in $\mathcal{C}(-r, r)$ and hence also in $W^{1,1}(-r, r)$ [CCN07]. Therefore, $|Dv|(\Sigma) = \int_{\Sigma^2} |\nabla v| dx$. Due to the lack of dependence on ϕ , we do not consider in the following calculations the integral with respect to the angle $\phi \in [0, \pi)$ and use only the symmetry on the sinogram space. Then, if we plug-in (3.49) in (3.46), we obtain

$$\begin{aligned} \operatorname{argmin}_{v \geq 0 \text{ a.e}} \left\{ \beta \|\nabla v\|_{L^1(\Sigma^2)} + \frac{1}{2} \int_{\Sigma^2} \frac{(g-v)^2}{g} d\sigma \right\} &= \operatorname{argmin}_{v \geq 0 \text{ a.e}} \left\{ \beta \|\nabla v\|_{L^1([-r,r])} + \int_0^\kappa \frac{(g-v)^2}{g} ds \right\} = \\ \operatorname{argmin}_{\kappa} \left\{ 4\beta\sqrt{r^2 - \kappa^2} + \int_0^\kappa \left(2\sqrt{r^2 - s^2} + \frac{\delta^2}{2\sqrt{r^2 - s^2}} - 2\delta \right) ds \right\} &= \\ \operatorname{argmin}_{\kappa} \left\{ 4\beta\sqrt{r^2 - \kappa^2} + \kappa\sqrt{r^2 - \kappa^2} + r^2 \arcsin\left(\frac{\kappa}{r}\right) + 2(r^2 - \kappa^2) \arcsin\left(\frac{\kappa}{r}\right) - 4\kappa\sqrt{r^2 - \kappa^2} \right\} \end{aligned}$$

which can be simplified to

$$\operatorname{argmin}_{\kappa} \left\{ (4\beta - 3\kappa)\sqrt{r^2 - \kappa^2} + (3r^2 - 2\kappa^2) \arcsin\left(\frac{\kappa}{r}\right) \right\}. \quad (3.50)$$

Since, it is impossible to compute an analytical expression for κ , we solve numerically (3.50) under the constraint $0 < |\kappa| < r$. Then, we obtain a value for κ which can be substituted in (3.49) and find the corresponding loss of contrast δ for our solution after the regularisation. We solve (3.50) with MATLAB's built-in routine *fminbnd* in $\kappa \in [0, r)$. In Figure 3.11, we present how the β parameter relates to the constant height value δ of the computed regularised solution. Clearly, for small values of β , there is no significant effect of the total variation regularisation but as we increase β , we have that δ decreases

to 0, while κ tends to r .

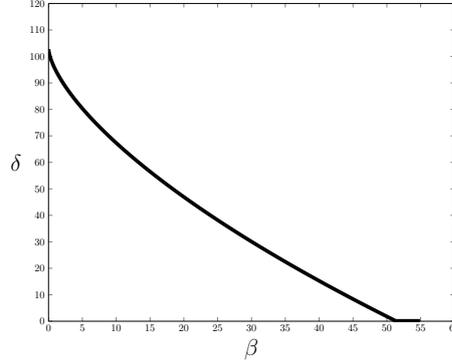


Figure 3.11: The relation between the regularisation parameter β and δ in (3.49), computed using (3.50) for the example in Figure 3.10. The parameter β varies from 0.001 to 55 with step size 0.1.

Before we apply the inverse Radon transform on (3.49) and find the corresponding solution in the image space, we need to verify its optimality. The following theorem ensures that the candidate solution (3.49) for the problem (3.46) is indeed optimal.

Theorem 3.5.8. *The unique solution of the minimisation problem (3.46) is defined by (3.49).*

Proof. The optimality condition on (3.46) implies that:

$$\beta q + \frac{v - g}{g} = 0, \quad q \in \partial|Du|(\Omega). \quad (3.51)$$

Recall that the subdifferential of the total variation, see Section 2.4, can be expressed as

$$\partial|Du|(\Sigma) = \left\{ \operatorname{div} p : p \in L_0^\infty(\Sigma^2), \|p\|_\infty \leq 1, \langle \operatorname{div} p, v \rangle = |Du|(\Sigma^2) \right\} \quad (3.52)$$

where $L_0^\infty(\Sigma^2)$ denotes the space of all functions in $L^\infty(\Sigma^2)$ that vanish at the boundary. Therefore, in our case (3.51) becomes

$$\beta p'(s) + \frac{v(s) - g(s)}{g(s)} = 0, \quad \text{for } s \in \Sigma^2 = (-r, r) \times [0, \pi) \quad (3.53)$$

with $-1 \leq p(s) \leq 1$ and $\int_{\Sigma^2} p'(s)v(s) d\sigma = \int_{\Sigma^2} |v'(s)| d\sigma$. Again, we can neglect the integral over all the angles. Hence, If v is either increasing or decreasing on an interval $I \subset [-r, r]$, then through integration by parts one obtains $-p(s)v'(s) = |v'(s)|$ which immediately implies that $p' = 0$ and $v = g$ on I . However, when $v \neq g$ on an interval $J \subset [-r, r]$, then $p' \neq 0$ which is true only if $v'(s) = 0$ on J , i.e., v is constant. Overall, the solution v is either constant or coincides with the initial data g . \square

In order to compute the regularised image that corresponds to a solution of (3.46) we first note that the rotational symmetry of the object in image space allows us to simplify the Radon transform and its inverse. In this case the Radon transform coincides with the so-called *Abel transform*, cf. [Pou10, Chapter 8]. More precisely, if u is a radial function and $u(x, y) = f(r = \sqrt{x^2 + y^2})$, then the Radon transform can be written as

$$\mathcal{R}_\theta u(s) = 2 \int_s^\infty \frac{f(r)r}{\sqrt{r^2 - s^2}} dr \quad (3.54)$$

where there is no angle dependence. Using (3.54), we can recover analytically the solution u for a regularised sinogram (3.49). Indeed, the Abel transform and the inverse Abel transform in this case are

$$\mathcal{A}(u(\tilde{r}))(x) = 2 \int_x^\infty \frac{\tilde{r}u(\tilde{r})}{\sqrt{\tilde{r}^2 - x^2}} d\tilde{r}, \quad (3.55)$$

$$u(\tilde{r}) = -\frac{1}{\tilde{r}\pi} \frac{d}{d\tilde{r}} \int_{\tilde{r}}^\infty \frac{x\mathcal{A}(u(\tilde{r}))(x)}{\sqrt{x^2 - \tilde{r}^2}} dx. \quad (3.56)$$

Setting $u(\tilde{r}) = 1$ and replacing the upper limit of the integral in (3.55) by \tilde{r} , the expression in (3.55) matches the expression for the Radon transform in (3.48). Finally, we plug-in (3.49) in (3.56) and focus on the constant part of the sinogram for $-\kappa \leq \tilde{r} \leq \kappa$,

$$u(\tilde{r}) = -\frac{1}{\tilde{r}\pi} \frac{d}{d\tilde{r}} \int_{\tilde{r}}^r \frac{x\delta}{\sqrt{x^2 - \tilde{r}^2}} dx = \frac{\delta}{\pi\sqrt{r^2 - \tilde{r}^2}}. \quad (3.57)$$

We observe that the reconstructed image is affected by the initial loss of contrast δ of the sinogram regularisation in (3.49) and depends radially on \tilde{r} . In Figure 3.12, we present the regularised solution u given by (3.57), which can be also interpreted as any angle profile for the two dimensional sinogram, see for instance Figure 3.18. We present the

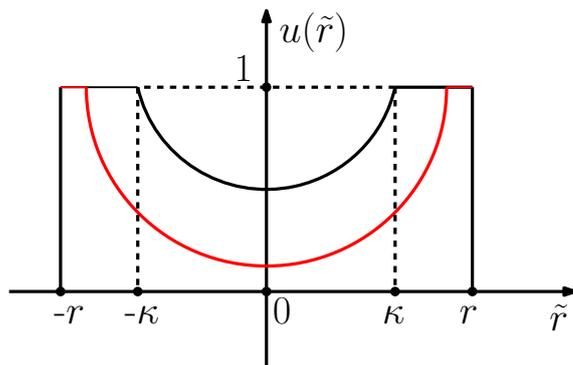


Figure 3.12: The solution $u(\tilde{r})$ (solid line) given in (3.57) inside the interval $[-r, r]$ and 0 outside. The black and the red curve constitute the regularised solution for a smaller and larger value of β , respectively. The larger β the more the solution concentrates around the boundaries of the ball.

reconstruction for two values of β : the initial black curve expands radially around 0 and as we increase β (red curve), we have that $\delta \rightarrow 0$ and $\kappa \rightarrow r$. The solution after regularising the sinogram starts losing details from the interior of the function and for extreme values of β , it approximates the outer boundaries of the object. This theoretical analysis will be an important motivation on using our approach to sinograms governed by thin structures. The reader is advised to compare these results with the numerical ones obtained in Section 3.7.2.

3.6 Numerical Implementation

In this section, we discuss the numerical solution of the minimisation problem (3.22). Since, we are dealing with regularisation on two spaces, the image and the sinogram space, we employ the split Bregman technique. It was introduced in [GO09] and its efficiency is based on splitting the initial minimisation problem to several other that are computationally easy to solve. In our case, we can separate the problem into two space related subproblems – one in image space and one in sinogram space – that are solved iteratively in an alternating fashion. In order to present the numerical solution we start with formulating (3.22) in a discrete setting.

3.6.1 Discrete Setting

Let $(u_{i,j})$, $i = 1, \dots, n$, $j = 1, \dots, m$ be the discretised image defined on a rectangular grid of size $n \times m$, $n, m > 0$. Let $(v_{i,j})$, $i = 1, \dots, k$, $j = 1, \dots, \ell$ the discretisation for an element in the sinogram space $\Sigma^2 = [0, \pi] \times [-r, r]$, where k denotes the number of angles and ℓ the number of lines. The values $u_{i,j}$ and $v_{i,j}$ are defined on two-dimensional grids. They are rearranged into one-dimensional vectors $u \in \mathbb{R}^{nm}$ and $v \in \mathbb{R}^{k\ell}$ by appending the columns of the array to each other, starting from the leftmost. Then, the discrete gradient for $u \in \mathbb{R}^{n \times m}$ is a matrix $\nabla \in \mathbb{R}^{nm \times 2nm}$ which is the standard forward difference approximation of the gradient in the continuum. More precisely, applying the discrete gradient to u gives $\nabla u = ((\nabla u)_1, (\nabla u)_2)$ with Neumann/mirror boundary conditions

$$(\nabla u)_1(i, j) = \begin{cases} u(i, j+1) - u(i, j), & \text{if } 1 \leq i \leq n, 1 \leq j < m, \\ 0, & \text{if } 1 \leq i \leq n, j = m, \end{cases}$$

$$(\nabla u)_2(i, j) = \begin{cases} u(i+1, j) - u(i, j), & \text{if } 1 \leq i < n, 1 \leq j \leq m, \\ 0, & \text{if } i = n, 1 \leq j \leq m. \end{cases}$$

The discrete divergence is defined as its adjoint, see for instance [Cha04], and is given by

$$\text{div} : (\mathbb{R}^{n \times m})^2 \rightarrow \mathbb{R}^{nm} \text{ with } \langle \text{div} z, u \rangle = - \langle z, \nabla u \rangle$$

where $u \in \mathbb{R}^{n \times m}$, $z = (z^1, z^2) \in (\mathbb{R}^{n \times m})^2$ and

$$\operatorname{div} z = \begin{cases} z^1(i, j) - z^1(i-1, j), & \text{if } 1 < i < n, 1 \leq j \leq m, \\ z^1(i, j), & \text{if } i = 1, 1 \leq j \leq m \\ -z^1(i-1, j), & \text{if } i = n, 1 \leq j \leq m \end{cases}$$

$$+ \begin{cases} z^2(i, j) - z^2(i, j-1), & \text{if } 1 \leq i \leq n, 1 < j < m, \\ z^2(i, j), & \text{if } j = 1, 1 \leq i \leq n \\ -z^2(i, j-1), & \text{if } j = m, 1 \leq i \leq n \end{cases}$$

Analogously, we define the discrete gradient and discrete divergence for $v \in \mathbb{R}^{k\ell}$ in the projection space. Further, to approximate the Radon transform \mathcal{R} we introduce the discrete Radon transform as a mapping $\mathcal{R} : \mathbb{R}^{nm} \rightarrow \mathbb{R}^{k\ell}$ and its inverse $\mathcal{R}^{-1} : \mathbb{R}^{k\ell} \rightarrow \mathbb{R}^{nm}$. In the numerical implementation the discrete Radon transform is represented by a sparse matrix $\mathcal{R} \in \mathbb{R}^{k\ell \times nm}$ which acts on a column vector $u \in \mathbb{R}^{nm}$ to obtain a sinogram image. Let $x(\theta_{\hat{i}}, s_{\hat{j}})$, $\hat{i} = 1, \dots, k$, $\hat{j} = 1, \dots, \ell$, be the line formed by $\theta_{\hat{i}}, s_{\hat{j}}$ and for $i = 1, \dots, n$ and $j = 1, \dots, m$ define

$$\psi_{i,j}(\theta_{\hat{i}}, s_{\hat{j}}) = \begin{cases} 1, & \text{where the line } x(\theta_{\hat{i}}, s_{\hat{j}}) \text{ goes through the pixel } (i, j) \\ 0, & \text{otherwise.} \end{cases} \quad (3.58)$$

Using this notation and the linearity of the Radon transform, we define the discrete Radon transform as

$$\mathcal{R}u(\theta_{\hat{i}}, s_{\hat{j}}) = \sum_{i=1}^n \sum_{j=1}^m u_{i,j} \mathcal{R}\psi_{i,j}(\theta_{\hat{i}}, s_{\hat{j}}) \quad (3.59)$$

where $\mathcal{R}\psi_{i,j}(\theta_{\hat{i}}, s_{\hat{j}})$ is equal to the length of the intersection of the projection line with the pixel (i, j) .

With these discrete quantities, we can now define the discrete functional F by

$$\min_{u \in \mathbb{R}^{n \times m}} \left\{ F(u) = \alpha \|\nabla u\|_1 + \beta \|\nabla(\mathcal{R}u)\|_1 + \frac{1}{2} \sum_{k,\ell} \frac{(g - \mathcal{R}u)^2}{g} \right\} \quad (3.60)$$

where the following discrete ℓ^p norms for $I = \{(i, j) : i = 1, \dots, n \text{ and } j = 1, \dots, m\}$ and $J = \{(i, j) : i = 1, \dots, k \text{ and } j = 1, \dots, \ell\}$ are

$$\|u\|_p = \left(\sum_{(i,j) \in I} |u_{i,j}|^p \right)^{1/p}, \quad \|v\|_p = \left(\sum_{(i,j) \in J} |v_{i,j}|^p \right)^{1/p},$$

$$\|\nabla u\|_p = \left(\sum_{(i,j) \in I} (|((\nabla u)_1)_{i,j}|^2 + |((\nabla u)_2)_{i,j}|^2)^{p/2} \right)^{1/p},$$

$$\|\nabla v\|_p = \left(\sum_{(i,j) \in J} (|((\nabla v)_1)_{i,j}|^2 + |((\nabla v)_2)_{i,j}|^2)^{p/2} \right)^{1/p}.$$

3.6.2 Split Bregman Algorithm

To solve the problem defined in (3.60), we employ the Bregman iteration [OBG⁺05] combined with a splitting technique. The resulting algorithm is called Split Bregman and is proposed in [GO09] to solve efficiently total variation and ℓ_1 regularised image problems. The concept of this splitting procedure is to replace a complex and costly minimisation problem by a sequence of simple and cheaply to solve minimisation problems and to set up an iteration in which they are solved alternately. Note, that the Split Bregman method can be equivalently phrased in terms of an *Augmented Lagrangian* method and *Douglas-Rachford* splitting, cf. [Set09, EZC10, Set11].

Suppose, we have to solve a constrained minimization problem:

$$\min_u E(u) \text{ such that } Au = b \quad (3.61)$$

where A is a linear operator, $E(u)$ is a convex functional and b is a vector. We transform (3.61) into an unconstrained minimisation problem

$$\min_u E(u) + \frac{\lambda}{2} \|Au - b\|_2^2 \quad (3.62)$$

where for λ sufficiently large the problem coincides with (3.61). Instead of solving (3.62), the authors in [OBG⁺05], introduce the concept of Bregman distance and propose to solve

$$\begin{aligned} u^{k+1} &= \operatorname{argmin}_u D_E^{\xi^k}(u, u^k) + \frac{\lambda}{2} \|Au - b\|_2^2 \\ &= \operatorname{argmin}_u E(u) - \langle \xi^k, u \rangle + \frac{\lambda}{2} \|Au - b\|_2^2 \\ \xi^{k+1} &= \xi^k - \lambda A^\top (Au^{k+1} - b). \end{aligned} \quad (3.63)$$

However, the unpleasant minimisation (3.63) can be reduced to the following as it is proved in [YOGD08, Theorem. 3.1]:

$$u^{k+1} = \min_u E(u) + \frac{\lambda}{2} \|Au - b^k\|_2^2 \quad (3.64)$$

$$b^{k+1} = b^k + b - Au^k \quad (3.65)$$

adding up the error in the fidelity term in every iteration. The splitting technique applied

on total variation minimisation is based on separating the L^1 and L^2 components appeared in the problem. As we will observe, the split Bregman algorithm, provide us with two subproblems solved alternately in the projection and image space.

We follow [GO09] to adapt the split Bregman algorithm to (3.60). To do so, we consider

$$\min_{\{u: u \geq 0 \text{ a.e.}\}} \alpha \|\nabla u\|_1 + \beta \|\nabla(\mathcal{R}u)\|_1 + \frac{1}{2} \sum_{k,\ell} \frac{(g - \mathcal{R}u)^2}{g}. \quad (3.66)$$

We start by replacing (3.66) with an equivalent constrained minimisation problem for two unknowns, the image $u \in \mathbb{R}^{n \times m}$ and the sinogram $v \in \mathbb{R}^{k \times \ell}$, related to each other by $v = \mathcal{R}u$. This results to

$$\min_{\{(u,v): u \geq 0 \text{ a.e.}\}} \alpha \|\nabla u\|_1 + \beta \|\nabla v\|_1 + \frac{1}{2} \sum_{k,\ell} \frac{(g - v)^2}{g} \text{ s.t } v = \mathcal{R}u. \quad (3.67)$$

For computational efficiency reasons, we introduce three additional variables

$$z = \nabla u, w = \nabla v, u = \tilde{u} \text{ and } v = \mathcal{R}u \quad (3.68)$$

and rephrase (3.66) again into

$$\min_{\{(u,\tilde{u},v,z,w): \tilde{u} \geq 0 \text{ a.e., satisfying (3.68)}\}} \alpha \|z\|_1 + \beta \|w\|_1 + \frac{1}{2} \sum_{k,\ell} \frac{(g - v)^2}{g}. \quad (3.69)$$

Then, we could iteratively solve the constrained minimisation problem (3.69) by Bregman iteration. For $b_1^0 \in \mathbb{R}^{k \times \ell}$, $b_2^0 \in (\mathbb{R}^{k \times \ell})^2$, $b_3^0 \in (\mathbb{R}^{n \times m})^2$, $b_4^0 \in \mathbb{R}^{n \times m}$, we iteratively solve for $k = 0, 1, \dots$

$$\begin{aligned} \operatorname{argmin}_{u,\tilde{u},v,z,w} \left\{ \alpha \|z\|_1 + \beta \|w\|_1 + \sum_{k,\ell} \frac{(g - v)^2}{g} + \mathcal{X}_{(\tilde{u} > 0)} + \frac{\lambda_1}{2} \|b_1^k + \mathcal{R}u - v\|_2^2 + \frac{\lambda_2}{2} \|b_2^k + \nabla v - w\|_2^2 \right. \\ \left. + \frac{\lambda_3}{2} \|b_3^k + \nabla u - z\|_2^2 + \frac{\lambda_4}{2} \|b_4^k + u - \tilde{u}\|_2^2 \right\}, \\ b_1^{k+1} = b_1^k + \mathcal{R}u^{k+1} - v^{k+1}, \quad b_2^{k+1} = b_2^k + \nabla v^{k+1} - w^{k+1}, \\ b_3^{k+1} = b_3^k + \nabla u^{k+1} - z^{k+1}, \quad b_4^{k+1} = b_4^k + u^{k+1} - \tilde{u}^{k+1}, \end{aligned}$$

with Lagrange multipliers $(\lambda_i)_{i=1}^4 > 0$, $b_1^k \in \mathbb{R}^{k \times \ell}$, $b_2^k \in (\mathbb{R}^{k \times \ell})^2$, $b_3^k \in (\mathbb{R}^{n \times m})^2$, $b_4^k \in (\mathbb{R}^{n \times m})$ and $\mathcal{X}_{(\tilde{u} > 0)}$ being the characteristic function for the positivity constraint on \tilde{u} . To progress, in each iteration above we would need to solve a minimisation problem with respect to all u, \tilde{u}, v, z, w at the same time which is numerically a very complicated task. Instead, we use the split Bregman idea and in each iteration solve a sequence of decoupled problems with respect to u, \tilde{u}, v, z and w respectively. This procedure leads to five minimisation problems that have to be solved sequentially in each iteration. The split

Bregman algorithm for (3.60) is summarised below. For $k = 0, 1, \dots$ we iteratively solve:

Split Bregman algorithm for (3.60)

$$v^{k+1} = \operatorname{argmin}_v \left\{ \frac{1}{2} \sum \frac{(g-v)^2}{g} + \frac{\lambda_1}{2} \|b_1^k + \mathcal{R}u^k - v\|_2^2 + \frac{\lambda_2}{2} \|b_2^k + \nabla v - w^k\|_2^2 \right\} \quad (3.70)$$

$$u^{k+1} = \operatorname{argmin}_u \left\{ \frac{\lambda_1}{2} \|b_1^k + \mathcal{R}u - v^{k+1}\|_2^2 + \frac{\lambda_3}{2} \|b_3^k + \nabla u - z^k\|_2^2 + \frac{\lambda_4}{2} \|b_4^k + u - \tilde{u}^k\|_2^2 \right\} \quad (3.71)$$

$$\tilde{u}^{k+1} = \operatorname{argmin}_{\tilde{u}} \left\{ \mathcal{X}_{(\tilde{u}>0)} + \frac{\lambda_4}{2} \|b_4^k + u^{k+1} - \tilde{u}\|_2^2 \right\} \quad (3.72)$$

$$z^{k+1} = \operatorname{argmin}_z \left\{ \alpha \|z\|_1 + \frac{\lambda_3}{2} \|b_3^k + \nabla u^{k+1} - z\|_2^2 \right\} \quad (3.73)$$

$$w^{k+1} = \operatorname{argmin}_w \left\{ \beta \|w\|_1 + \frac{\lambda_2}{2} \|b_2^k + \nabla v^{k+1} - w\|_2^2 \right\} \quad (3.74)$$

$$b_1^{k+1} = b_1^k + \mathcal{R}u^{k+1} - v^{k+1} \quad (3.75)$$

$$b_2^{k+1} = b_2^k + \nabla v^{k+1} - w^{k+1} \quad (3.76)$$

$$b_3^{k+1} = b_3^k + \nabla u^{k+1} - z^{k+1} \quad (3.77)$$

$$b_4^{k+1} = b_4^k + u^{k+1} - \tilde{u}^{k+1} \quad (3.78)$$

Every subproblem either has an explicit solution or involves the solution of a linear system of equations that can be efficiently solved with an iterative method such as conjugate gradient. We iterate until

$$\frac{\|\tilde{u}^{K+1} - \tilde{u}^K\|_2}{\|\tilde{u}^{K+1}\|_2} < 10^{-4}$$

and take v^{K+1} as the regularised sinogram and \tilde{u}^{K+1} as the reconstructed image. Let us go into more detail for the solution of each minimisation problem.

Solution of (3.70): To solve (3.70), we derive the corresponding Euler-Lagrange equation for v and obtain a linear system of equations with $k \cdot \ell$ unknowns $v_{i,j}$, $i = 1, \dots, k$, $j = 1, \dots, \ell$ which reads

$$(3.70) \Rightarrow ((1 + \lambda_1)g - \lambda_2 g \operatorname{div} \cdot \nabla)v = g + \lambda_1 g (b_1^k + \mathcal{R}u^k) + \lambda_2 g \operatorname{div}(b_2^k - w^k). \quad (3.79)$$

The system (3.79) is solved by a conjugate gradient method.

Solution of (3.71): The Euler-Lagrange equation of (3.71) with respect to u is

$$(3.71) \Rightarrow (\lambda_1 \mathcal{R}^* \mathcal{R} - \lambda_3 \operatorname{div} \cdot \nabla + \lambda_4) u = \lambda_1 \mathcal{R}^* (v^{k+1} - b_1^k) + \lambda_3 \operatorname{div} (b_3^k - z^k) - \lambda_4 (b_4 - \tilde{u}^k), \quad (3.80)$$

where \mathcal{R}^* is the adjoint of \mathcal{R} , i.e., that is the discrete backprojection. As before, the system (3.80) is solved by a conjugate gradient method.

Solution of (3.72): The solution of (3.72) is given by

$$\tilde{u}^{k+1} = \max\{b_4^{k+1} + u^{k+1}, 0\}.$$

Solution of (3.73) and (3.74): Finally, the solution of the minimisation problems (3.73), (3.74) can be obtained exactly through *soft shrinkage*, see for instance [WYYZ08]. That is,

$$z^{k+1} = \mathcal{S}_{\frac{\alpha}{\lambda_3}}(b_3^k + \nabla u^{k+1}) := \max\left(\left\|b_3^k + \nabla u^{k+1}\right\|_2 - \frac{\alpha}{\lambda_3}, 0\right) \frac{b_3^k + \nabla u^{k+1}}{\|b_3^k + \nabla u^{k+1}\|_2}, \quad (3.81)$$

$$w^{k+1} = \mathcal{S}_{\frac{\beta}{\lambda_2}}(b_2^k + \nabla v^{k+1}) := \max\left(\left\|b_2^k + \nabla v^{k+1}\right\|_2 - \frac{\beta}{\lambda_2}, 0\right) \frac{b_2^k + \nabla v^{k+1}}{\|b_2^k + \nabla v^{k+1}\|_2}. \quad (3.82)$$

3.7 Numerical Results

In this section, we present our numerical results on both artificial and real PET data. The Radon matrix that we described in (3.59) is fixed and produces sinograms of size 192×192 , that is the sinogram is given in 192 projection lines, 192° degrees with 1° degree incrementation and the corresponding reconstructed image is of size 175×175 pixels. We corrupt the sinograms with Poisson noise of different levels. In order to create noisy images corrupted by Poisson noise, we apply the MATLAB routine *imnoise (sinogram, poisson)*. MATLAB's *imnoise* function acts in the following way: for an image of double precision, the input pixel values are interpreted as means of a Poisson distribution scaled by a factor of 10^{-12} . For example, if an input pixel has the value $5.5 * 10^{-12}$ then the corresponding output pixel will be generated from a Poisson distribution with mean of 5.5 and afterwards scaled back to its original range by 10^{12} . The factor 10^{12} is fixed to represent the maximal number of detectable photons. Our simulated sinograms are in $[0, 1]$ intensity and in order to create different noise levels, we have to rescale the initial sinogram with a suitable factor before applying *imnoise* and then scale it back with the same factor, i.e.,

$$\text{Noisy Sinogram} = \text{scale} * \text{imnoise} \left(\frac{\text{sinogram}}{\text{scale}}, \text{poisson} \right).$$

To simulate realistic sinograms with higher noise level, we use 10^{13} as a scaling factor, see for example Figure 3.13. The real data, in Figure 3.16, was obtained from the hardware

phantom “Wilhelm”, a self-built phantom modelled of the human body. Beside the activity in the heart a small source is placed in the phantom to simulate a lesion, see Section 3.7.1 for more information.

Before presenting our results we give some numerical details on how equations (3.70)-(3.74) are solved and how the parameters α, β and $(\lambda_i)_{i=1}^4$ are chosen. Both linear systems (3.79) and (3.80) are solved using MATLAB’s built-in function *cg* which performs a conjugate gradient method. As a stopping criterium we either stop after at most 200 iterations or if the relative residual is smaller than 10^{-3} . As it is observed in [GO09], it seems optimal to apply only a few steps of an iterative solver for both subproblems (3.79) and (3.80) since the error in the split Bregman algorithm is updated in every iteration. Typically, in the case of $\beta > 0$, the computational times for the linear system of the sinogram is approximately 0.83 seconds per iteration and the corresponding linear system in the image is solved approximately in 1.26 seconds per iteration. This computational time difference can be justified by the fact that the corresponding matrix for the sinogram subproblem is sparse, whereas the matrix of the image subproblem is dense since $\mathcal{R}^*\mathcal{R}$ is used. Even though we use $\mathcal{R}^*\mathcal{R}$ in order to project in every iteration our regularised sinogram, we do not observe any radial smoothing in the final reconstruction, as it is described in Proposition 3.3.6. This is due to the alternating procedure that our algorithm imposes and the total variation regularisation is sufficient to eliminate this artifact within the iterations.

The Lagrange multipliers $(\lambda_i)_{i=1}^4$ in equations (3.70)-(3.74) are selected following [P. 12] in order to optimise the convergence speed. They are fixed as $\lambda_1 = 0.001$, $\lambda_2 = 1$, and $\lambda_3 = \lambda_4 = 100$. Furthermore, note that these parameters have been tuned appropriately in order to obtain a relatively small condition number for both system matrices

$$\begin{aligned} A_{img} &= \lambda_1 \mathcal{R}^* \mathcal{R} - \lambda_3 \operatorname{div} \cdot \nabla + \lambda_4 I, \\ A_{sin} &= (1 + \lambda_1)g - \lambda_2 g \operatorname{div} \cdot \nabla, \end{aligned}$$

appear in (3.80) and (3.79) respectively. Finally, we observe that after 150 split Bregman iterations, there are no significant changes in the reconstructed image and therefore we choose a stopping criteria of either at most $K = 400$ iterations or we stop at iteration K where for the first time we have

$$\frac{\|\tilde{u}^{K+1} - \tilde{u}^K\|_2}{\|\tilde{u}^{K+1}\|_2} < 10^{-4}$$

and \tilde{u}^{K+1} is the final regularised image. To evaluate the quality of reconstructed images we choose the *Signal-to-Noise Ratio* (SNR) as a quality measure. The SNR is defined as

$$\text{SNR} = 20 \log \left(\frac{\|u\|_2}{\|u - \tilde{u}\|_2} \right) \quad (3.83)$$

where u and \tilde{u} denote the ground truth and the reconstructed image, respectively. In what follows, we first evaluate the proposed reconstruction technique (3.22) against pure total variation regularisation on the image ($\beta = 0$) for a synthetic image of two circles and for different noise levels, as well as for a real data set acquired for the Wilhelm phantom. Then, we numerically analyse the scale space properties of pure sinogram regularisation, that is for $\alpha = 0$, which will be a motivation for the final section where we discuss the merit of the proposed reconstruction method for PET data that encodes thin image structures, see Section 3.7.3.

3.7.1 Image reconstruction from corrupted simulated and real PET data

Our numerical demonstration begin with a simple toy-phantom example as it is shown in Figure 3.13. It is a simulated phantom of two discs with different radii and the corresponding noiseless and noisy sinograms corrupted with low and high level Poisson noise as described above.

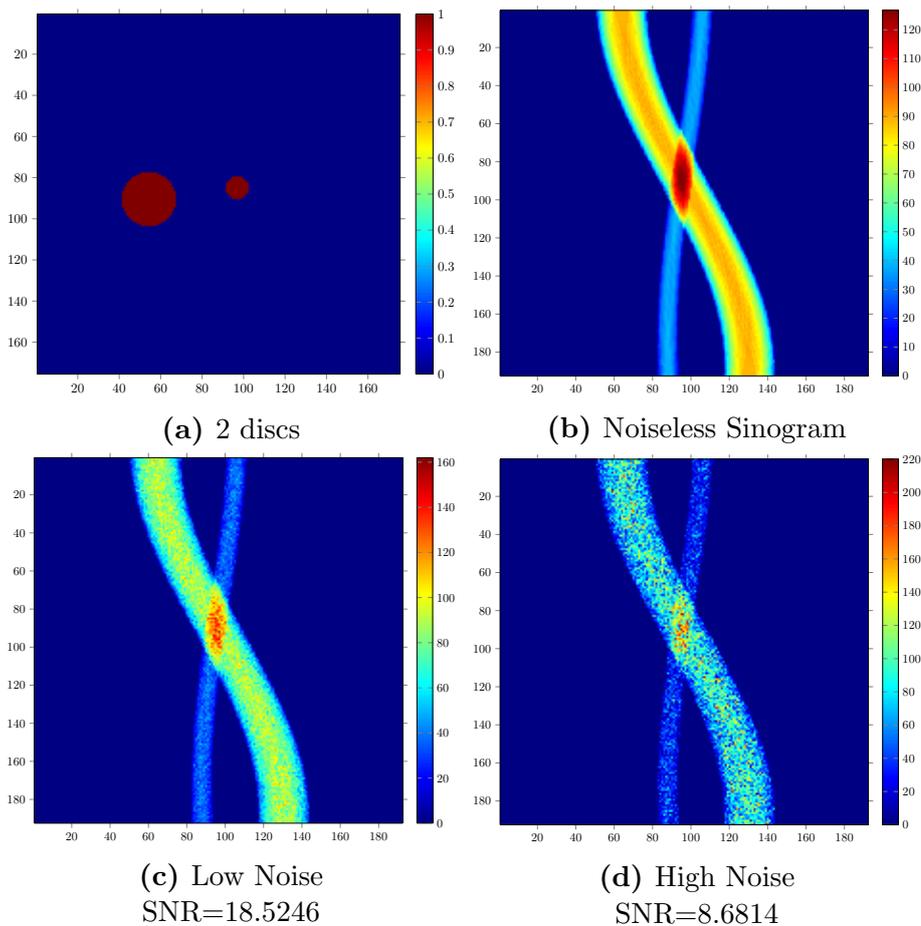


Figure 3.13: The phantom image includes two discs of radii $r_1 = 26$ and $r_2 = 11$ pixels. Its sinogram has 192 angles and 192 rays with low and high noise.

First, we evaluate the proposed algorithm for reconstructing an image from the sinogram corrupted by low level Poisson noise with $\text{SNR}=18.5246$, see Figure 3.13c. The proposed reconstruction algorithm with joint total variation regularisation on image and sinogram (that is $\alpha, \beta > 0$) is compared with the algorithm that uses pure total variation regularisation on the image (that is $\alpha > 0$ and $\beta = 0$). Both reconstruction strategies are tested for a range of parameters α, β and in each case the reconstruction which has the highest SNR value is found. For $\beta = 0$ we compute the reconstructed image for $\alpha = 3, 4, 5, 6, 7$. The optimal reconstructed image in terms of the best $\text{SNR}=25.8589$ is obtained for $\alpha = 6$, see Figure 3.14a. Then, we test the proposed reconstruction method applying total variation regularisation on both the image and the sinogram using the same range of $\alpha = 3, 4, 5, 6, 7$ and $\beta = 0.001, 0.005, 0.01, 0.05$. Here, the optimal reconstruction was obtained for $\alpha = 6$ and $\beta = 0.001$ with $\text{SNR}=25.3127$, see Figure 3.14b. In Table 3.1, a full list of tested parameters and SNRs for the corresponding reconstructed images is given. The results do not indicate a significant difference between the algorithm with and without total variation regularisation on the sinogram, both visually and also in terms of the SNR. Indeed, in the low noise case additional total variation regularisation on the sinogram produces even slightly worse results in terms of SNR than using no regularisation on the sinogram at all.

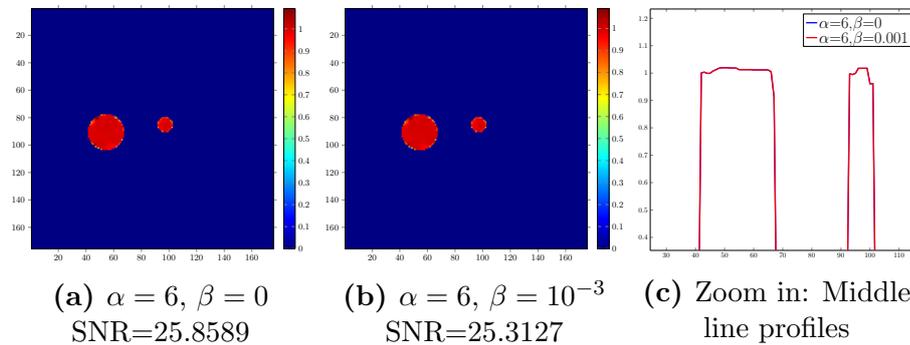


Figure 3.14: Low level noise: Optimal reconstruction results of the two discs image with sinogram shown in Figure 3.13c with and without sinogram regularisation and a comparison of the line profiles for the two results.

The TV regularisation on the sinogram gains importance in the reconstruction algorithm when the noise in the corruption of the sinogram is increased. The sinogram with high level noise is shown in Figure 3.13d and has $\text{SNR}=8.6814$. We tested the proposed method for $\alpha = 250, 275, 300, 325, 350$ and $\beta = 0, 0.001, 0.01, 0.05, 0.1$. The results are reported in Table 3.2.

The highest SNR is obtained when $\alpha = 250$ and $\beta = 0.001$, cf. Figure 3.15b. Although, it is hard to distinguish any difference between the cases of β , we observe that the extra penalisation on the sinogram produces better results in terms of the SNR value. The

		β					
		0	0.001	0.005	0.01	0.05	0.1
α	3	24.0819	22.0172	22.4415	22.8894	23.2414	21.6533
	4	25.3682	24.0926	24.2951	24.4801	23.6303	21.9382
	5	25.7867	25.0829	25.0779	25.0469	23.9432	22.0367
	6	25.8589	25.3127	24.7787	25.0602	24.0095	22.1034
	7	25.7436	24.8499	24.8278	25.0148	23.9662	22.2289

Table 3.1: Low level noise for simulated example in Figure 3.13: SNRs of reconstructed images for different combinations of α and β values.

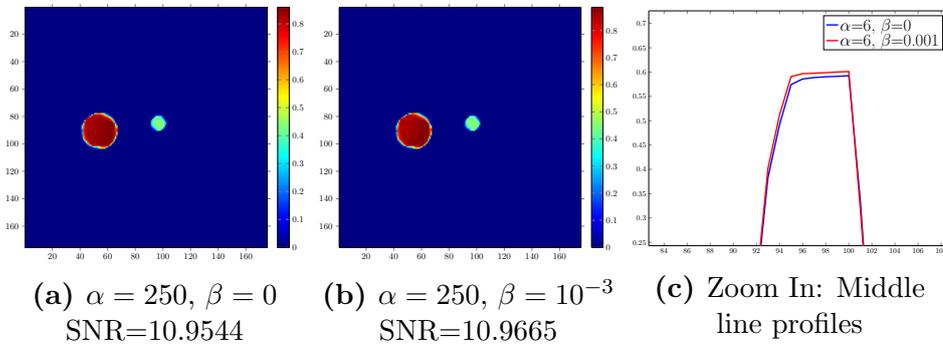


Figure 3.15: High level noise: Optimal reconstruction results of the two discs image with sinogram shown in Figure 3.13d with and without sinogram regularisation and a comparison of the line profiles for the two results.

		β					
		0	0.001	0.005	0.01	0.05	0.1
α	250	10.9544	10.9665	10.9557	10.9464	10.8531	10.8058
	275	10.9502	10.9599	10.9501	10.9381	10.8595	10.8013
	300	10.9425	10.9543	10.9415	10.9257	10.8267	10.7777
	325	10.9167	10.9551	10.9434	10.9283	10.8101	10.7293
	350	10.8784	10.9289	10.9165	10.9014	10.7946	10.7104

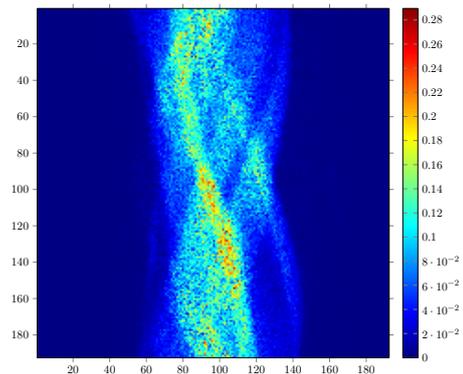
Table 3.2: High level noise for simulated example in Figure 3.13: SNRs of reconstructed images for different combinations of α and β values.

increase in SNR for $\beta > 0$ can be seen when comparing the middle line profiles of the reconstructed images with and without sinogram regularisation in Figure 3.15c.

As a second example for our evaluation of the algorithm for PET reconstruction, we consider real PET data obtained from scanning a self-built phantom of a human breast with a small source which simulates a lesion, compare Figure 3.16a. The data has been acquired with a Siemens Biograph Sensation 16 PET/CT scanner (Siemens Medical Solutions) located at the University Hospital in Münster. From the acquired 3D PET data, we use only one sinogram slice. The 2D sinogram dimension is 192×192 with a pixel size of 3.375mm^2 . The size of the reconstructed image is 175×175 , covering a field of view of 590.625mm in diameter. The 2D slice of the noisy sinogram which has been used in our computations is shown in Figure 3.16b. Reconstructions obtained from the proposed algorithm, with and without sinogram regularisation, are shown in Figure 3.17. The additional regularisation of the sinogram seems to allow for *smoother* image structures (such as the boundary of the red lesion) and results in a slight reduction of the staircasing effect of total variation regularisation.



(a) Top: Phantom "Wilhelm", consisting of a plastic torso and inserts for the lungs, heart and liver. Bottom: Phantom reconstruction with combined PET-MRI. Data courtesy of the European Institute for Molecular Imaging (EIMI), Münster.



(b) Noisy sinogram (2D slice)

Figure 3.16: Real PET data.

In the following two sections we will aim to improve our understanding of this new sinogram regularisation, taking the analytic solution of Section 3.5.4 as a starting point. A thorough numerical discussion of this example in Section 3.7.2 leads us to Section 3.7.3 where the benefits of total variation regularisation of the sinogram for the reconstruction of thin objects are discussed.

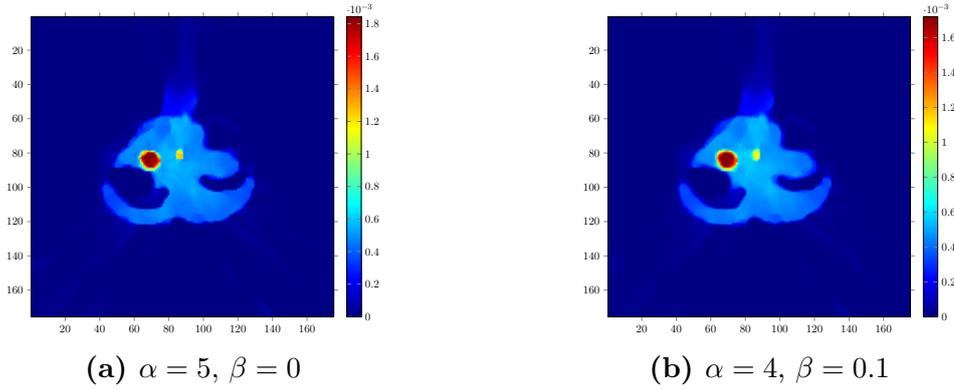


Figure 3.17: Real Data: Best reconstructions in terms of visual comparison for the noisy slice in Figure 3.16b.

3.7.2 Scale space of sinogram regularisation

Following up on the computations in Section 3.5.4, we now discuss how the regularisation on the sinogram affects the backprojected image. Let us recall that every point (θ, s) on the sinogram corresponds to a line $s = x \cos \phi + y \sin \phi$ that passes through a point (x, y) on the image, with a distance s from the origin and normal to the direction $\theta = (\cos \phi, \sin \phi)$. Moreover, every point on an edge in the sinogram corresponds to a line in the object space which is tangent to the boundary of the object. To further understand how sinogram regularisation acts, we consider the effect of the regularisation when reconstructing an image from simulated noise-free Radon data. To this end, we set $\alpha = 0$, regularise the noise-free sinogram with different values of β , and apply FBP to the regularised sinogram to obtain the corresponding reconstructed image. We call the set of reconstructed images from regularised sinograms with varying β regularisation, *the scale space* of total variation regularisation of the sinogram.

We consider the discrete variational model of (3.22) for $\alpha = 0$ which results in the following weighted total variation denoising problem for the sinogram g

$$\operatorname{argmin}_{v \geq 0 \text{ a.e.}} \beta \|\nabla v\|_1 + \frac{1}{2} \sum_{k,\ell} \frac{(g - v)^2}{g} \quad (3.84)$$

Similar to Section 3.6.2, we solve (3.84) by a split Bregman technique, introducing two additional variables $w = \nabla v$ and $\tilde{v} = v$. Let $b_1^0 \in (\mathbb{R}^{k \times \ell})^2$ and $b_2^0 \in \mathbb{R}^{k \times \ell}$, we iteratively solve for $k = 0, 1, \dots$

$$v^{k+1} = \operatorname{argmin}_v \frac{\lambda_1}{2} \|b_1^k + \nabla v - w^k\|_2^2 + \frac{\lambda_2}{2} \|b_2^k + v - \tilde{v}^k\|_2^2, \quad (3.85)$$

$$\tilde{v}^{k+1} = \operatorname{argmin}_{\tilde{v} \geq 0} \frac{1}{2} \sum \frac{(g - \tilde{v})^2}{g} + \frac{\lambda_2}{2} \|b_2^k + v^{k+1} - \tilde{v}\|_2^2, \quad (3.86)$$

$$w^{k+1} = \underset{w}{\operatorname{argmin}} \beta \|w\|_1 + \frac{\lambda_1}{2} \|b_1^k + \nabla v^{k+1} - w\|_2^2, \quad (3.87)$$

$$b_1^{k+1} = b_1^k + \nabla w^{k+1} - v^{k+1}, \quad (3.88)$$

$$b_2^{k+1} = b_2^k + v^{k+1} - \tilde{v}^{k+1}. \quad (3.89)$$

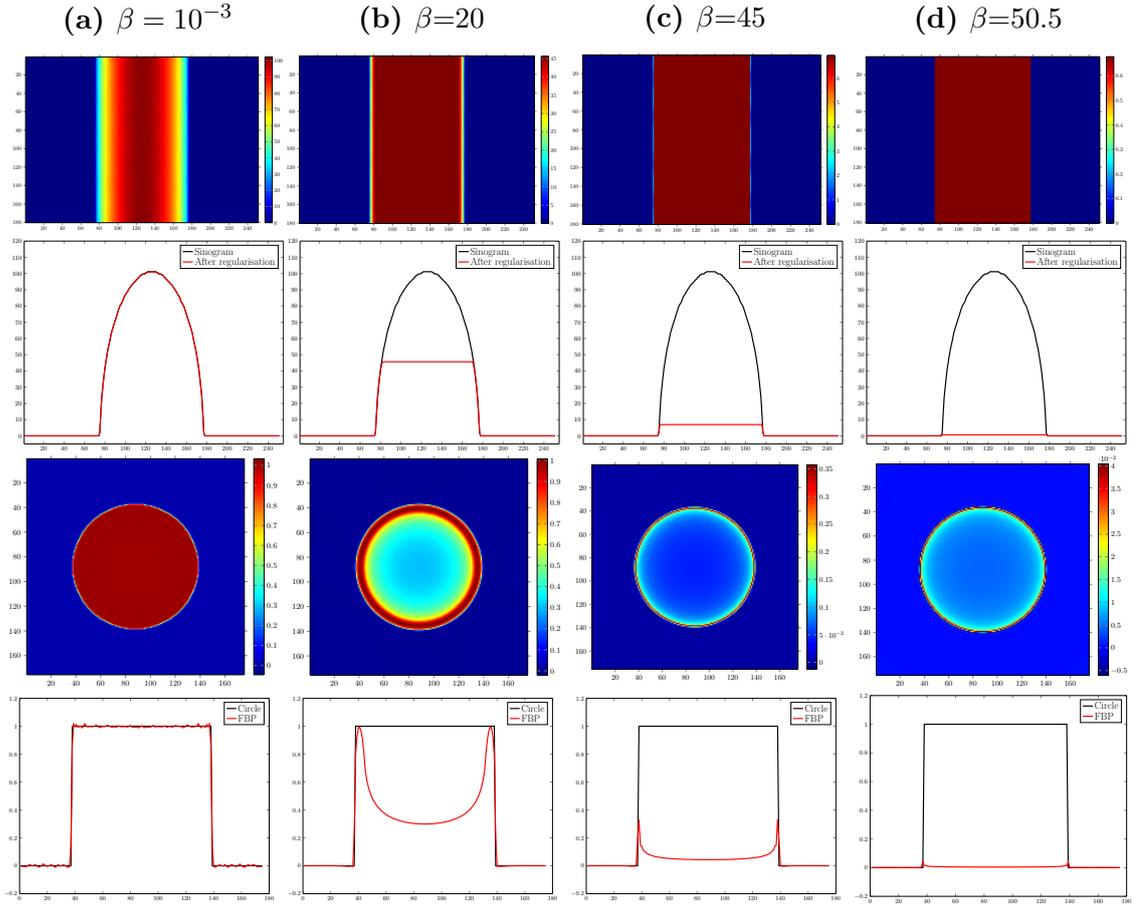


Figure 3.18: Sinogram regularisation with different values of β (first row) and the corresponding filtered backprojected images (third row). The second row represents a 45° comparison of the original sinogram and the sinogram after regularisation. The numerical highest value of the sinogram is 102.8. The fourth row represents the middle line profiles of the reconstructed images in the third row compared with the original one.

Moreover, since we do not apply any positivity constraint on the image as it is done in the full algorithm used in Section 3.7.1, we might observe small negative values in the reconstructed images presented in the following. The optimality condition of (3.85) leads to a linear system, which is solved quite efficiently using the conjugate gradient method. The subproblem (3.86) is an element-wise division and the subproblem (3.87) can be computed explicitly by the shrinkage operator in a similar way as in (3.82). Note, that we enforce a positivity constraint on the \tilde{v} variable, which is our final solution. We

fix $\lambda_1 = \lambda_2 = 1$ and solve alternately the subproblem (3.85), (3.86) and (3.87), until we reached our stopping criteria:

$$\frac{\|\tilde{v}^{k+1} - \tilde{v}^k\|_2}{\|\tilde{v}^{k+1}\|_2} < 10^{-5}.$$

First, we consider image functions with radial symmetry such as in Section 3.5.4 equation (3.49). Figure 3.18 shows the numerically computed regularised sinograms and corresponding images for an original image of a ball with radius $r = 50.5$. Here, we use MATLAB's built-in function *iradon* with a *Ram-Lak* filter and *spline* interpolation to compute the filtered backprojection of the regularised sinogram. Moreover, Table 3.3, demonstrates the correspondence of the numerical solution with the analytic solution obtained in Section 3.5.4 for three balls of radii $r = 15.5, 30.5$ and 50.5 . The values δ^{an} and δ^{num} denote the analytic and numerical loss of contrast δ , respectively, in the expression of the regularised solution in (3.49). As predicted from the analysis in Section 3.5.4, we observe that as β converges to the radius r , the regularised image emphasises the boundary of the ball, see for instance the third row in Figure 3.18.

r = 15.5	β	10^{-3}	0.1	1	5	10	15	15.5
	δ^{an}	30.94	29.88	25.84	16.09	7.59	0.64	0.084
	δ^{num}	31.32	29.76	25.71	15.96	7.37	0.67	0.37
r = 30.5	β	10^{-3}	1	10	15	20	25	30.5
	δ^{an}	60.93	59.6	31.35	22.37	14.46	7.27	0.09
	δ^{num}	61.98	54.58	31.42	22.47	14.55	7.34	0.65
r = 50.5	β	10^{-3}	1	10	20	30	45	50.5
	δ^{an}	100.92	93.33	65.74	45.46	28.71	7.16	0.12
	δ^{num}	101.83	93.26	65.75	45.41	28.82	7.24	0.68

Table 3.3: Comparison of analytic and numerical computations of sinogram regularisation for three test images of characteristic functions of circles with radii $r = 15.5, 30.5$ and 50.5 . The parameters δ^{an} and δ^{num} denote the analytic and numerical δ , respectively, in the expression of the regularised solution in (3.49). Compare also Figure 3.18 for regularised reconstructions for the circle with radius $r = 50.5$.

Going beyond radial symmetry, we consider three additional examples where the sinogram depends on the direction θ . First, we simply consider the image that we use in the previous section in Figure 3.13 without adding noise to its sinogram. The effect of β regularisation in this case is presented in Figure 3.19. We see that as we increase β we lose details in the image, starting again from the inner structure of the discs, while enhancing

the boundaries of the objects. Here, the connection of the choice of β with the radius of every circle is clearly visible. More precisely, for $\beta < r_2$ the boundary of the smaller circle is enhanced and for $r_2 < \beta < r_1$ the small circle is lost and the boundary of the larger circle is enhanced. Again, due to the FBP we observe small negative values in the reconstructed images.

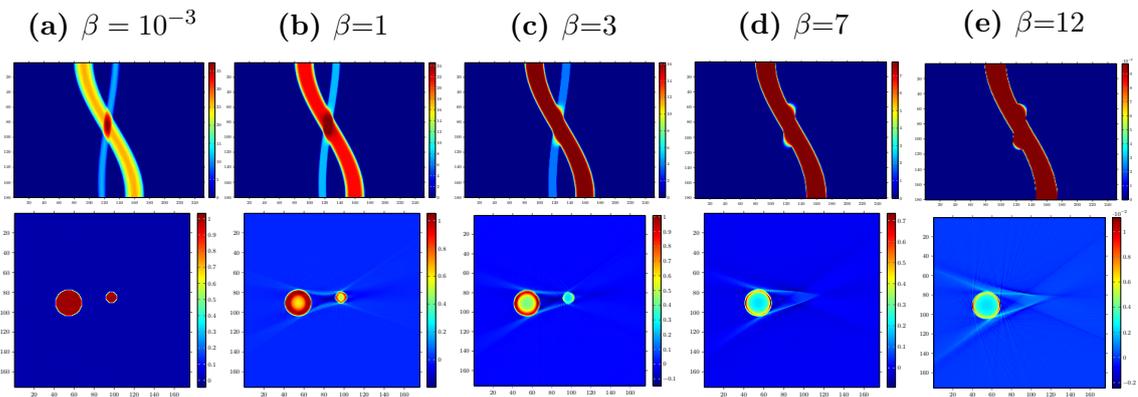


Figure 3.19: Sinogram regularisation with different values of β and the corresponding filtered backprojected images using MATLAB's *iradon* built-in function. The radii for the discs are $r_1 = 13$ and $r_2 = 5.5$.

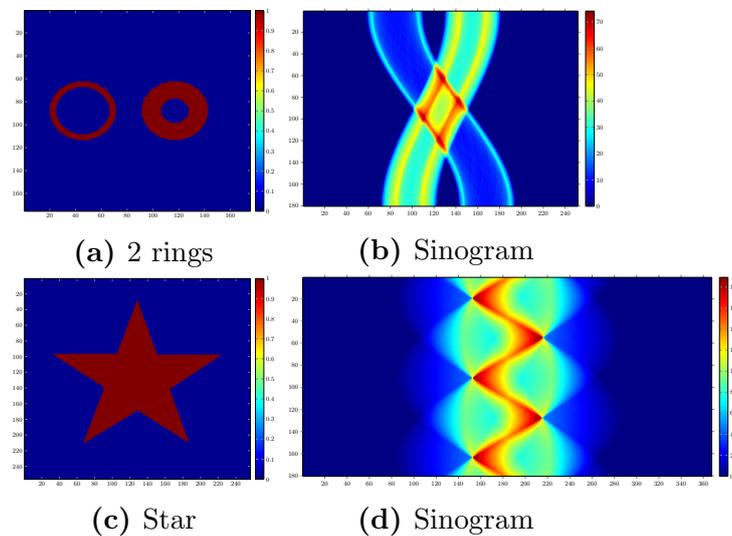


Figure 3.20: 2 rings with different annulus regions and its sinogram. Star-shaped image of 5 points and its sinogram.

In Figure 3.20, we present two more test images. The first one is an image of two rings with the same outer radius but with different annulus regions, compare Figure 3.20a. A similar scale-space analysis as for the previous examples is carried out in Figure 3.21. Additionally to the enhancement of the outer boundaries of the two rings we see that for

increasing β regularisation the reconstructed image approaches the convex hull of the two rings. This is even more apparent for the last example of a star-shaped object in Figure 3.20c. See Figure 3.22 in particular.

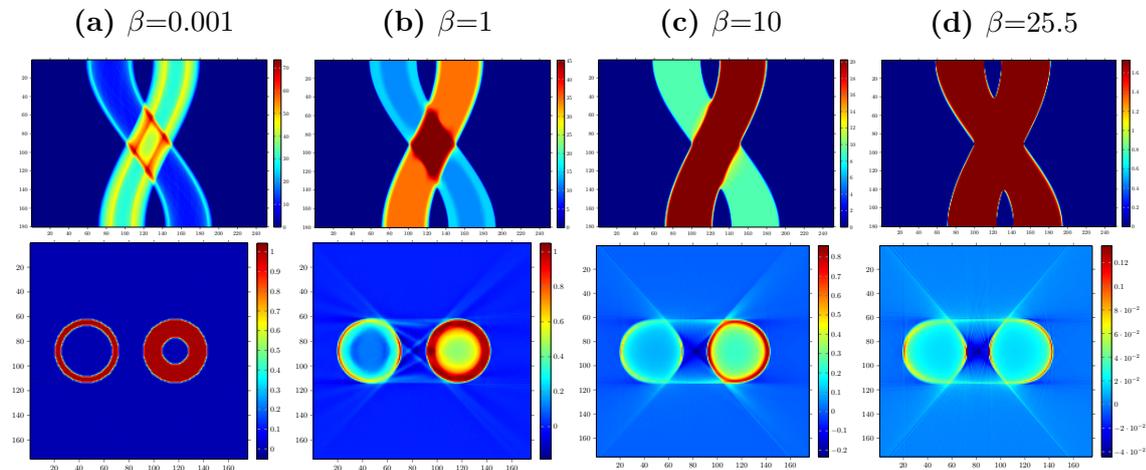


Figure 3.21: 2 rings with different annulus regions: The outer radius for both rings is $r=25.5$ and the inner radii are $r_1 = 21$ and $r_2 = 11$. In Figures (a)-(d), we present the sinogram regularisation for increasing values of β with the corresponding filtered backprojected using MATLAB's *iradon* built-in function.

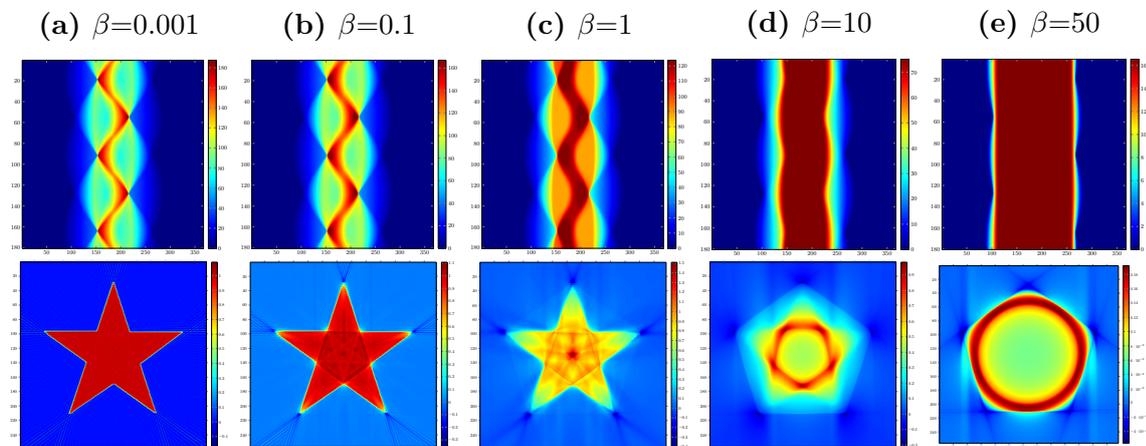


Figure 3.22: Star-shaped image of 5 corners: In Figures (a)-(e), we present the sinogram regularisation for increasing values of β with the corresponding filtered backprojected image using MATLAB's built-in function *iradon*.

The conclusion of this section is at the same time the motivation for the next section. Analysing the effect of total variation regularisation on the sinogram by considering its scale space and its effect on the reconstructed image we have seen in Figures 3.18–3.22 the potential use of this method is for the enhancement and detection of object boundaries.

As we will see in the next section, this effect can be exploited for enhancing thin structures in images obtained from Radon measurements.

3.7.3 Thin Structure Reconstruction

In what follows, we discuss how total variation regularisation of the sinogram can improve the quality of the reconstruction in comparison with pure total variation regularisation of the image in the presence of *thin* structures. Our first example is a thin rectangular frame in Figure 3.23. Similarly as in Section 3.7.1, we start by finding an optimal value of α with $\beta = 0$, in terms of SNR. Then, we select a range of α values close to this optimal one and we allow strictly positive values for β . The noise that is added on the sinogram, is generated by MATLAB's *imnoise* routine, with a 10^{12} scaling factor, see the beginning of Section 3.7 for more explanation. The test image that is shown in Figure 3.23 has 50 pixels

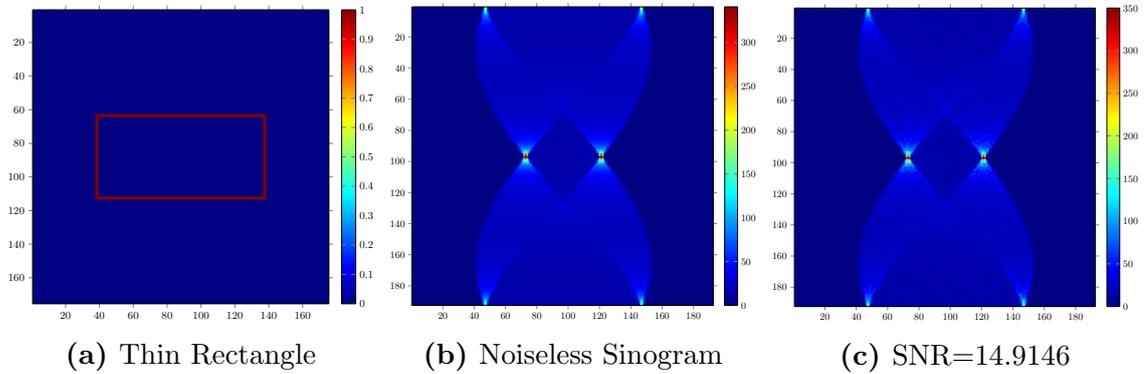


Figure 3.23: A thin rectangle of 50 pixels width, 100 pixels length and 2 pixels length on the boundaries. The corresponding noiseless and noisy sinograms with 10^{12} scaling factor in *imnoise*.

width and 100 pixels length and the rectangular frame has a width of 2 pixels. In Figure 3.24, we first present some of the results obtained with pure total variation regularisation on the image, that is when $\beta = 0$. As we increase the α parameter, we observe that the best SNR corresponds to $\alpha=5$ with SNR=19.9764. That is because for small values of α we observe that the large-scale structure of the object is still intact, with the cost that noise is still present in the reconstructed image, see Figures 3.24a–3.24c. However, with higher values of α noise is further eliminated but at the expense of a significant loss of contrast and some unpleasant artifacts along the boundaries of the frame, see Figure 3.24d–3.24f.

If we switch on total variation regularisation on the sinogram, that is taking $\beta > 0$, we obtain results which are greatly improved both in terms of the SNR of the reconstructed images but also – visually – in terms of finding the right balance of eliminating the noise and accurately preserving the thin structures, see Table 3.4 and Figure 3.25. This observation

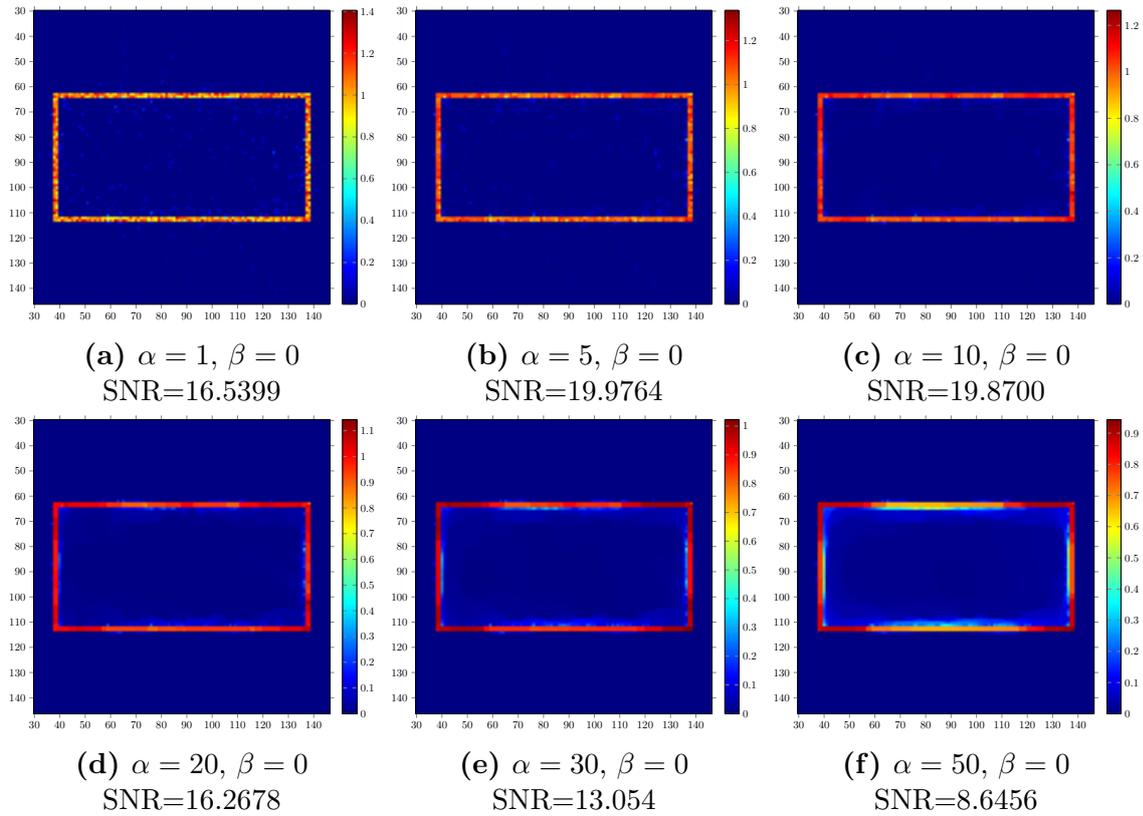


Figure 3.24: Thin Rectangle: Reconstruction without total variation regularisation on the sinogram and different parameters of α .

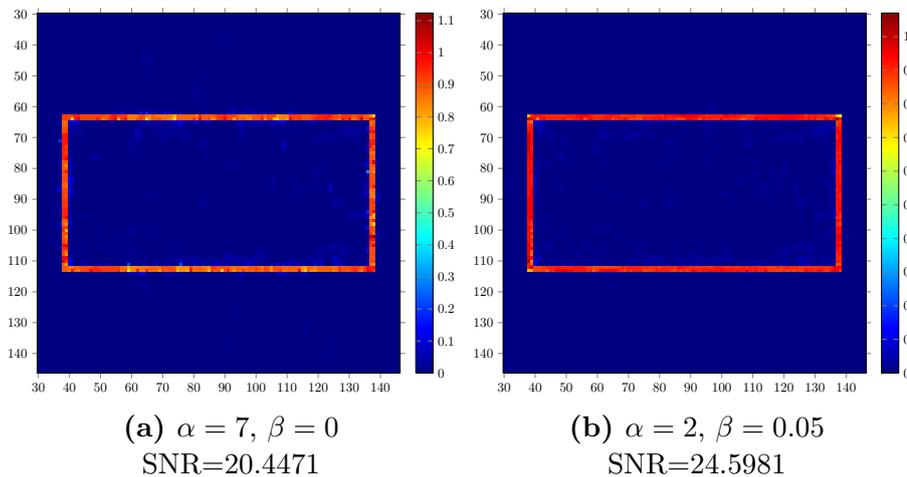


Figure 3.25: Thin Rectangle: Best reconstructions with and without total variation regularisation on the sinogram as reported in Table 3.4.

is confirmed by a second example of an image of two thin straight lines which cross,

		β				
		0	0.005	0.01	0.05	0.1
α	2	17.6798	18.0078	19.7238	24.5981	24.2978
	3	18.6444	18.9855	20.6460	23.9028	24.2647
	4	19.4269	19.7539	21.5305	23.9178	23.2860
	5	19.9764	20.2979	21.7962	23.6466	22.8525
	6	20.2583	20.5771	21.9057	23.2213	22.3440
	7	20.4471	20.8665	21.8372	22.7554	21.8147
	8	20.3511	20.3276	20.9859	22.2391	21.2477

Table 3.4: Thin Rectangle: SNR with $\beta \geq 0$.

compare Figure 3.26. The width of the thin lines is 3 pixels. The length of the horizontal line is 121 pixels and of the vertical line is 100 pixels. The noise, added on the sinogram, is generated with the same scaling factor of 10^{12} as before. Again, we observe that for positive values of β , we obtain much better reconstructions with almost all noise eliminated while keeping the boundaries of the thin structures intact, see Figure 3.27. We present our numerical results for both cases of β , according to best SNR values that are produced for specific ranges of α and β , see Table 3.5.

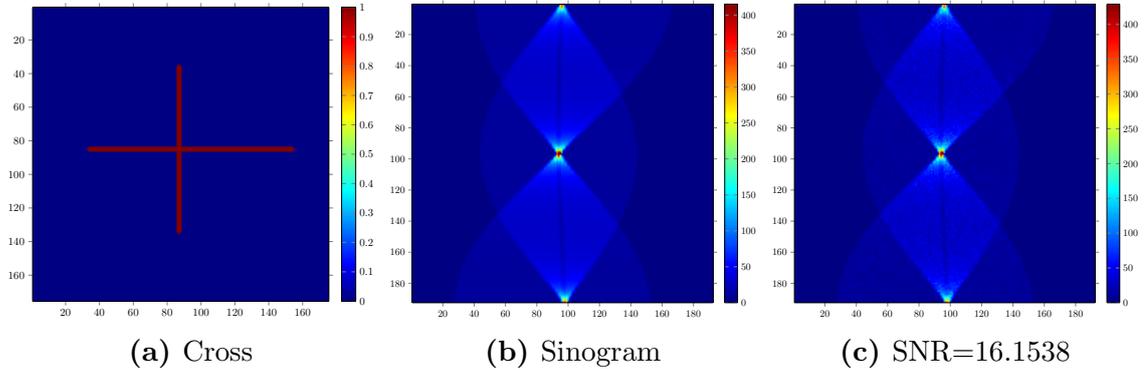


Figure 3.26: Cross image: Two thin crossing lines and its noiseless and noisy sinograms respectively. The width of the thin lines is 3 pixels. The length of the horizontal line is 121 pixels and of the vertical line is 100 pixels.

The previous toy-phantoms were used in order to demonstrate the advantages of our proposed model where thin formations are present. Now, we proceed to a more realistic PET phantom, which visualises the activity of the human heart. The XCAT phantom is a 3D phantom. For our purpose we used only one z-slice through the centre of the

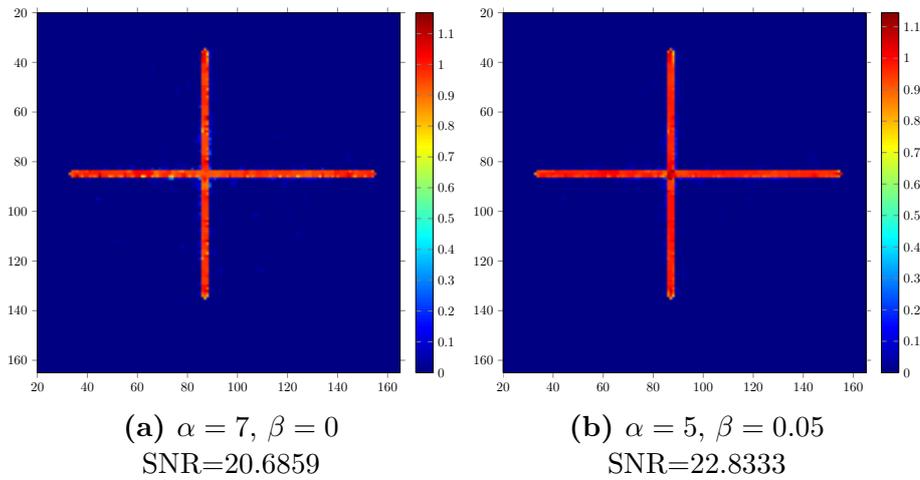


Figure 3.27: Cross image: Best reconstructions with and without total variation regularisation on the sinogram as reported in Table 3.5.

		β				
		0	0.005	0.01	0.05	0.1
α	5	20.2146	20.6041	21.1127	22.8333	21.5244
	7	20.6859	20.9067	21.2490	22.1166	20.8109
	10	20.3563	20.2762	20.4687	20.6954	19.5800
	13	19.3596	19.1946	19.2891	19.2428	18.3060
	15	18.5511	18.3874	18.4518	18.3379	17.5022

Table 3.5: Cross: SNR with $\beta \geq 0$.

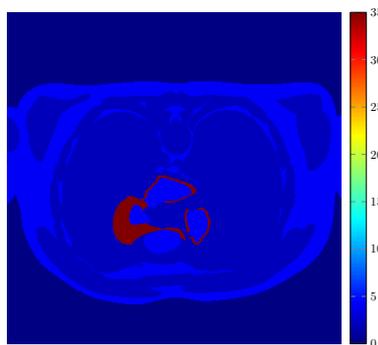


Figure 3.28: XCAT cardiac-torso phantom.

phantom which represents a transverse plane view of the human body, see Figure 3.28. In particular, we can see the activity of the heart through the myocardium (the muscle

surrounding the heart) in red. We focus on regions where thin structures are observed, see Figures 3.29a-3.29b and add the usual level of Poisson noise to their corresponding sinograms, see Figures 3.29c-3.29d.

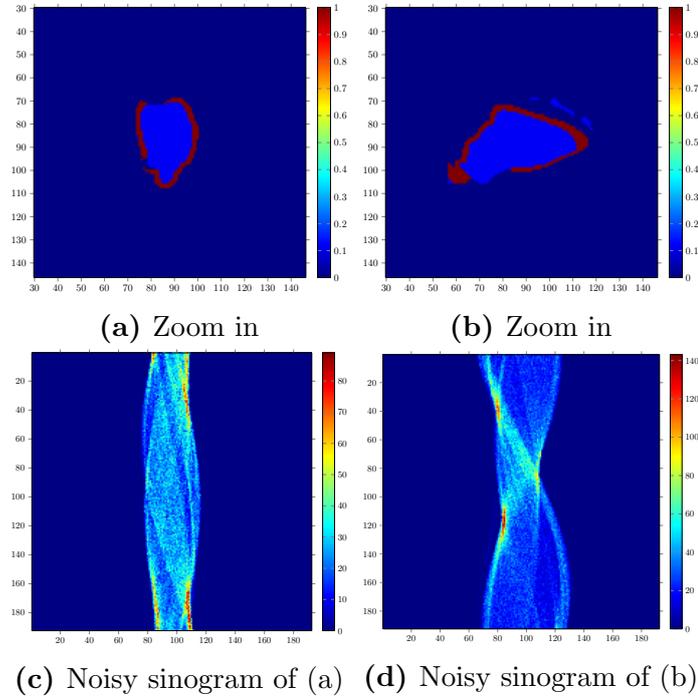


Figure 3.29: Selected regions of the XCAT phantom with the corresponding noisy sinograms.

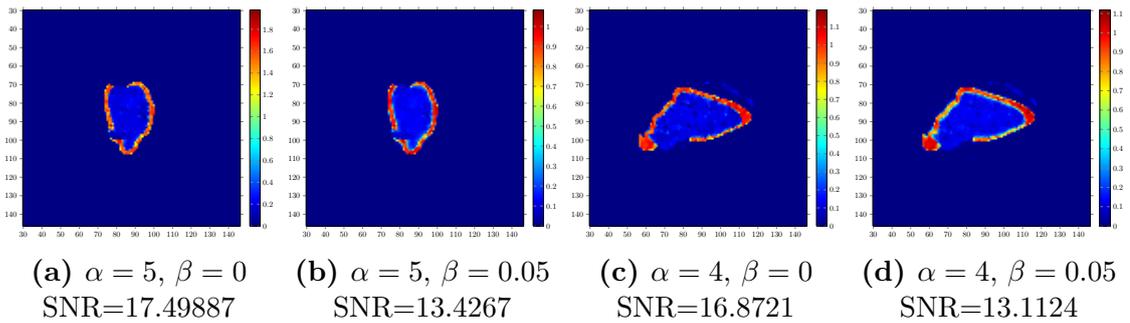


Figure 3.30: Reconstructions with and without total variation regularisation on the details of the XCAT sinogram in Figure 3.29.

In Figures 3.30a-3.30d, we present our best reconstructions for these two different data-regions in terms of the SNR values for both cases of sinogram regularisation. We observe that the best reconstructions are achieved when there is no regularisation on the sinogram. That is because for increasing values of β a smoothing on the originally blocky

boundaries is enforced and hence the SNR value is reduced. Indeed, as we show in the following experiments this is only true if the initial data that we start our experiments with is of low resolution and the thin structures have *blocky* instead of smooth boundaries.

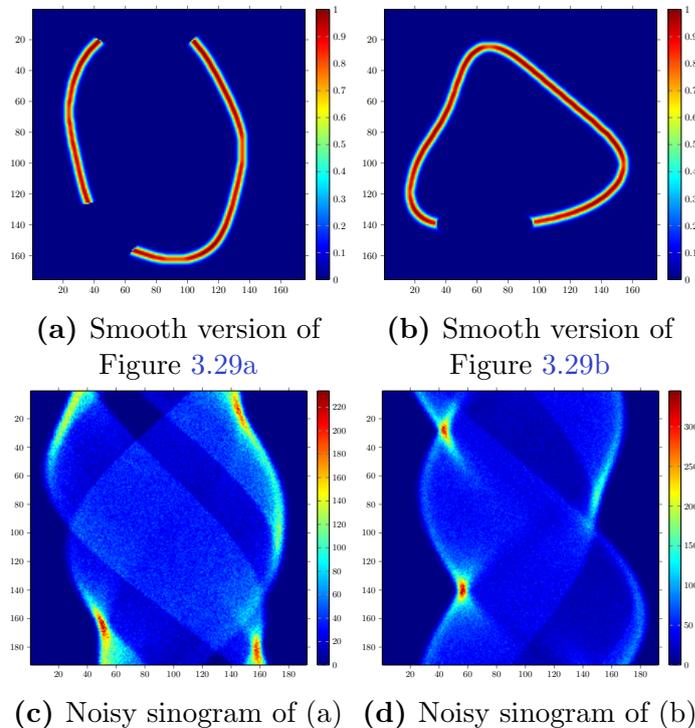


Figure 3.31: High resolution XCAT: smooth versions of Figures 3.29a-3.29b and their noisy sinograms.

We observe a serious improvement to our reconstructions when our initial phantom data is of higher resolution, see Figure 3.31. We change our experiment to the consideration of a high resolution version of the XCAT phantom with thin structures as in Figures 3.29a-3.29b but with medically more realistic smooth boundaries. As it is expected, regularising only on the image space creates a rather unpleasant staircasing effect along the boundaries which is clearly eliminated when we combine the regularisation on both spaces, see Figure 3.32. Indeed, a significant increase of the SNR when turning on the TV regularisation on the sinogram ($\beta > 0$) can be observed.

Overall, our proposed variational model clearly indicates that the additional penalisation on the sinogram space has a powerful impact on PET reconstruction. The analysis on the scale space total variation provides us with an interpretation on how our algorithm can be beneficial and to what extent. Although, the combined regularisation is computationally slower, we obtain better quality images with smoother boundaries when we deal with images with thin and elongated structures. Finally, we would like to point out

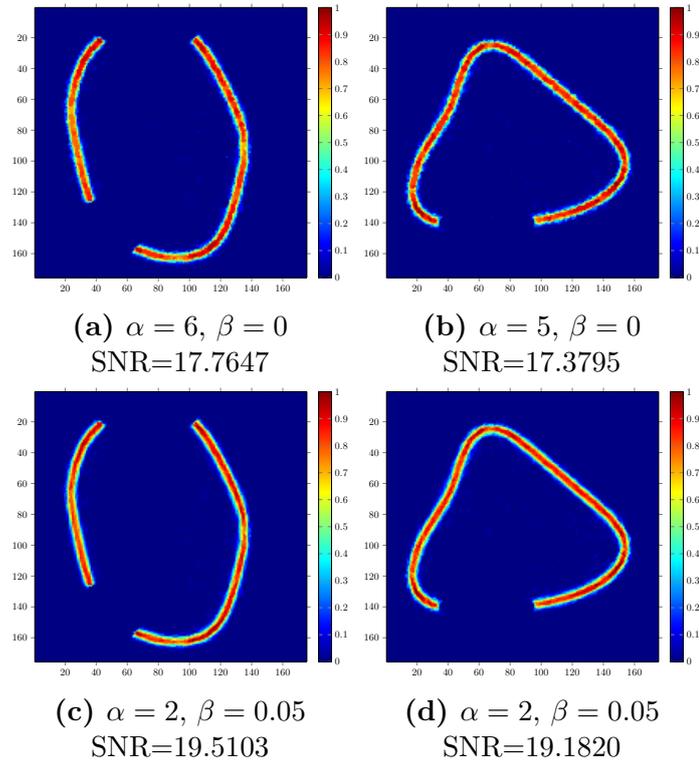


Figure 3.32: Reconstructions with and without total variation regularisation on the sinogram. Reduced staircasing along the boundaries is achieved when sinogram regularisation is active, resulting in a significant improvement of the SNR.

that our method could also be beneficial to applications related to computed tomography reconstruction.

Chapter 4

Infimal convolution regularisation functionals of BV and L^p spaces

4.1 Introduction

In this chapter we introduce a family of a novel TV – L^p infimal convolution regularisation functionals with applications to image processing. It is based on the publications [BPPS15a] and [BPPS15b]. So far, we have encountered the total variation (TV) regularisation functional in the context of medical imaging and in particular in PET reconstruction. Here, we propose a regulariser that is based on an “infimal-convolution process”, a combination of the TV seminorm and the L^p norm, namely $BV(\Omega)$ and $L^p(\Omega)$ spaces. Our regulariser can be expressed by the following minimisation problem:

$$\text{TVL}_{\alpha,\beta}^p(u) := \inf_{w \in L^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)}. \quad (4.1)$$

The $\|\cdot\|_{\mathcal{M}}$ denotes the Radon norm of a measure and the positive parameters α and β are tuned appropriately for every $p \in (1, \infty]$ in order to balance the strength between these two terms. The functional (4.1) is certainly based on TV seminorm with an additional L^p norm which is responsible of obtaining different type of structures depending on the value of p . As we will see in the forthcoming analysis, p regulates the *smoothness* of the solution u for the minimisation problem

$$\min_u \frac{1}{2} \|f - Tu\|_{L^s(\Omega)}^s + \Psi(u), \quad (4.2)$$

where $\Psi(u)$ is the regulariser in (4.1) and $s \geq 1$. We assume that the data f , defined on a domain $\Omega \subset \mathbb{R}^d$ with $d = 1, 2$, has been corrupted through a bounded linear operator T and additive (random) noise. Different values of s can be considered for the first term of (4.2), the *fidelity term*. For example, models incorporating a L^2 fidelity term (resp. L^1) are suitable for the restoration of images corrupted by additive Gaussian noise (resp.

impulse noise), when T is the identity operator. Certainly, as we have observed in the previous chapter, other types of noise can also be considered, e.g. Poisson noise and in this case the form of the fidelity term is adjusted accordingly, e.g. Kullback-Leibler divergence or weighted- L^2 norm, see (3.20) and (3.1) respectively.

One should notice that the regulariser (4.1) is not a formal expression of an infimal-convolution action between two convex functionals. However, an equivalent way to reformulate (4.1) and as a result (4.2) is by using an infimal-convolution of the $\|D(\cdot)\|_{\mathcal{M}}$ and $\|D(\cdot)\|_{L^p(\Omega)}$ seminorms with $1 < p \leq \infty$. Hence, given noisy data f , the solution u is obtained by solving

$$\min_{\substack{u=u_1+u_2 \\ u_1 \in \text{BV}(\Omega) \\ u_2 \in W^{1,p}(\Omega)}} \alpha \|Du_1\|_{\mathcal{M}} + \beta \|Du_2\|_{L^p(\Omega)} + \frac{1}{s} \|f - T(u_1 + u_2)\|_{L^s(\Omega)}^s, \quad s \geq 1. \quad (4.3)$$

In a sense, we decompose our image into two structural components. One that promotes piecewise constant structures and one that promotes structures specified by the value of p , particularly a combination of BV and Sobolev spaces $W^{1,p}$. Later, we will prove that (4.3) and (4.2) with $\Psi(u) = \text{TVL}_{\alpha,\beta}^p$ are equivalent in the one dimensional setting.

Infimal convolution approach as in (4.3) is not new. It was first introduced in [CL97] by Chambolle and Lions in the context of image denoising. They proposed to use an infimal convolution of functionals with first and second derivatives in order to reduce the staircasing effect that the total variation regularisation creates. Explicitly, they propose the following minimisation problem:

$$\min_{\substack{u=u_1+u_2 \\ u_1 \in \text{BV}(\Omega) \\ u_2 \in \text{BV}^2(\Omega)}} \alpha \|Du_1\|_{\mathcal{M}} + \beta \|D^2u_2\|_{\mathcal{M}} + \frac{1}{2} \|f - u_1 - u_2\|_{L^2(\Omega)}^2 \quad (4.4)$$

where $\text{BV}^2(\Omega) = \{u \in W^{1,1}(\Omega) : \nabla u \in \text{BV}(\Omega)\}$ is the space of functions of *bounded Hessian* [Dem85]. Note that by the definition of the total variation, see Section 2.3, we can write $\|Du_1\|_{\mathcal{M}} = \text{TV}(u_1)$ and $\|D^2u_2\|_{\mathcal{M}} = \text{TV}^2(u_2) = \text{TV}(\nabla u_2)$ and hence the first two terms can be expressed as the ICTV (Infimal Convolution TV) regulariser defined below:

$$\begin{aligned} \text{ICTV}_{\alpha,\beta}(u) &:= \min_{v \in \text{BV}^2(\Omega)} \alpha \text{TV}(u - v) + \beta \text{TV}^2(v) \\ &= \min_{\substack{w \in \text{BV}(\Omega) \\ w = \nabla v}} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|Dw\|_{\mathcal{M}}. \end{aligned} \quad (4.5)$$

Similarly, the positive parameters α and β are responsible to obtain a decomposition into piecewise constant component from the first term and a *piecewise affine* component from the second term. Compared to our regulariser (4.1), the reader is advised to focus on

three differences with (4.5). We minimise over the L^p space without any constraint on w , which is a function and not a measure and is penalised by the L^p norm compared to the Radon norm. Moreover, our case involves only first-order variants of the total variation whereas the ICTV includes also second order derivatives. In the following sections, these key factors give sufficient capabilities to our first-order and TV-based model to reduce and eliminate the staircasing effect.

We should mention that there exist many interesting and important approaches in the context of high-order regularisers and in particular to image decomposition. For example, in [CEP10] and [CEP07], the authors introduce a fourth-order model applied to denoising and texture extraction. They replaced $\|D^2u_2\|_{\mathcal{M}}$ with $\|\Delta u_2\|_{L^2}^2$ in (4.4) and either use an L^2 -squared fidelity for denoising tasks or an H^{-1} -squared fidelity which is suitable for treating texture components in an image, see [OSV03], [LV08] and [Sch09]. The main idea is to decompose a given image into structural-geometric and texture-oscillating components. The reader is also referred to other structure-texture decomposition models as in [AABFC05], [AGCO], [LV05] and [GLML07]. From another point of view, regularisers that takes into account only second order derivatives, i.e., $\Psi(u) = \text{TV}^2(u)$ and minimise over $\text{BV}^2(\Omega)$ space, constitutes a significant improvement towards image restoration and specifically staircasing elimination. For example, in [Sch98], [BP10] and [HS06a] a concrete variational analysis is presented for (4.2) in the context of image denoising. Moreover, the authors in [PS08] studied a high-order extension of the k -th derivative of total variation and provided a characterisation of minimisers and analytical solutions for specific one dimensional data.

The most successful regulariser of this kind is the second order total generalised variation (TGV) introduced by Bredies et al. [BKP10]. In practice, one can understand TGV exactly as (4.5) by neglecting the constraint on w . Its definition reads

$$\text{TGV}_{\alpha,\beta}^2(u) := \min_{w \in \text{BD}(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|\mathcal{E}w\|_{\mathcal{M}}. \quad (4.6)$$

The minimisation of $\text{ICTV}_{\alpha,\beta}$ is over functions that are gradients whereas the minimisation of $\text{TGV}_{\alpha,\beta}^2$ is over the wider space $\text{BD}(\Omega)$. It is the space of functions of *bounded deformation*, i.e., the space of all $w \in L^1(\Omega)$, whose distributional symmetrized derivative $\mathcal{E}w = \frac{1}{2}(\nabla w + \nabla w^\top)$ is a finite Radon measure, see [TS80]. As a result, the total generalised regularisation when $\Psi(u) = \text{TGV}_{\alpha,\beta}^2(u)$ in (4.2) has the ability to find the appropriate balance between piecewise constant and affine structures and provide a staircasing-free reconstruction. In general, TGV and ICTV are not equivalent except in the one dimensional setting, see [Mül13, Chapter 3]. Also, a recent paper on a *modified* infimal convolution approach indicates that similar results to TGV can be obtained and in certain cases the two methods are equivalent, see [SST11].

One of the main disadvantage of TGV is its computational cost that second-order

derivatives create. For instance, the authors in [PS13] propose a regulariser $\Psi(u) = \text{TV}(u) + \text{TV}^2(u)$ that can achieve similar results to TGV but with less computational cost. In general, variational models that contain high-order derivatives can have significant impact on the computational time involved to acquire the desired solution. Even though, there are many efficient algorithms, as the *split Bregman algorithm* or the *primal-dual algorithm* of Chambolle-Pock cf [CP11], that can be applied in order to solve minimisation problems as (4.2), it is still much slower than a standard first-order methods. Our main contribution of this chapter is to introduce a family of first-order regularisers (4.1) that are capable of obtaining results comparable to high-order regularisers such as TGV. It is an extension of total variation regularisation by incorporating the capabilities of L^p norms for image processing purposes. To the best of our knowledge, the use of L^p norms for $p > 1$ has been exploited in different contexts, see for instance the p -Laplacian and an L^∞ decomposition in [Kui07] and [EV07] respectively.

Organisation of the chapter

In Section 4.2, we discuss the well-posedness of our regularisers (4.1) for $p \in (1, \infty]$ and provide an equivalent expression based on standard duality arguments. We also examine its relation to the total variation seminorm and prove existence and uniqueness results for (4.2) when $\Psi(u) = \text{TVL}_{\alpha,\beta}^p(u)$.

We continue with the analysis of the p -homogenous analogue of (4.1), namely we replace $\|w\|_{L^p(\Omega)}$ by $\frac{1}{p} \|w\|_{L^p(\Omega)}^p$ for finite values of p . An important outcome of this analysis is that for $p = 2$, the 2-homogenous analogue coincide with the well-known Huber-TV regularisation, see Section 4.3.

For a better understanding of our proposed regulariser, we provide a thorough analysis in the one dimensional setting in Section 4.4. There, we derive necessary and sufficient conditions for a pair $(w, u) \in L^p \times \text{BV}$ to be a solution of (4.2) for every value of $p \in (1, \infty]$ and provide information about the structure of the solutions. We compute analytically the exact solutions on representative one dimensional data such as the step function and a piecewise affine function.

In Section 4.5, we formulate an image decomposition approach based on an infimal convolution of BV and $W^{1,p}$ spaces. A characterisation of minimisers for the decomposition approach is also presented in order to determine the nature of each components, i.e., piecewise constant and p -smooth components.

Our numerical experiments begin in Section 4.6, where we propose a split Bregman algorithm in order to solve efficiently (4.2) for $p \in (1, \infty]$. We compare the analytical and numerical one dimensional results for the exact solutions of the step and the piecewise affine function. Moreover, we concentrate on the affine behaviour of our solution for increasing values of p and yet for $p = \infty$. We demonstrate the capability of our first-order variational model not only to eliminate the staircasing and achieve equivalent result to

TGV but also provide even better reconstructions when spike-like structures appear in the images. Finally, we introduce a weighted version of the (4.1) regulariser where β is a space-dependent matrix that imitates the gradient of the reconstructed image. Here, we pose some open questions on how a suitable selection of the matrix β can be achieved in order to obtain reconstructions that are far better than TGV and visually close to the ground truth.

4.2 Total variation and L^p regularisation

We begin with the analysis of the TVL^p functional as well as some of its main properties. Let $\Omega \subset \mathbb{R}^d$ with $d \geq 1$, $\alpha, \beta > 0$ and $1 < p \leq \infty$, we define $\text{TVL}_{\alpha, \beta}^p : L^1(\Omega) \rightarrow \overline{\mathbb{R}}$ as follows:

$$\text{TVL}_{\alpha, \beta}^p(u) := \min_{w \in L^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)}. \quad (4.7)$$

First, we ensure that the above minimisation problem is indeed well-defined and the minimum is attained.

Proposition 4.2.1. *Let $u \in \text{BV}(\Omega)$ with $1 < p \leq \infty$ and $\alpha, \beta > 0$. Then (4.7) is well defined.*

Proof. By the Lebesgue decomposition of the measure Du , we have that $Du = D^a u + D^s u$. Using the Remark 2.1.5, we obtain the following

$$\begin{aligned} w \in \operatorname{argmin}_{\tilde{w} \in L^p(\Omega)} \alpha \|Du - \tilde{w}\|_{\mathcal{M}} + \beta \|\tilde{w}\|_{L^p(\Omega)} &\Leftrightarrow \\ w \in \operatorname{argmin}_{\tilde{w} \in L^p(\Omega)} \alpha \|D^s u + D^a u - \tilde{w}\|_{\mathcal{M}} + \beta \|\tilde{w}\|_{L^p(\Omega)} &\Leftrightarrow \\ w \in \operatorname{argmin}_{\tilde{w} \in L^p(\Omega)} \alpha (\|D^s u\|_{\mathcal{M}} + \|D^a u - \tilde{w}\|_{\mathcal{M}}) + \beta \|\tilde{w}\|_{L^p(\Omega)} &\Leftrightarrow \\ w \in \operatorname{argmin}_{\tilde{w} \in L^p(\Omega)} \alpha \|\nabla u - \tilde{w}\|_{L^1(\Omega)} + \beta \|\tilde{w}\|_{L^p(\Omega)}, &\quad (4.8) \end{aligned}$$

where we have used that the measures are mutual singular with respect to the Lebesgue measure $(\nabla u - w)\mathcal{L}^d$ and consider the minimisation over \tilde{w} . Thus, it suffices to prove well-posedness for (4.8) which is an L^1 - L^p minimisation problem. Consider a minimising sequence $(\tilde{w}_n)_{n \in \mathbb{N}} \in L^p(\Omega)$ and a constant $C > 0$ such that $0 \leq F(\tilde{w}_n) \leq C$, where

$$F(\tilde{w}_n) := \alpha \|\nabla u - \tilde{w}_n\|_{L^1(\Omega)} + \beta \|\tilde{w}_n\|_{L^p(\Omega)}.$$

Since \tilde{w}_n is bounded in $L^p(\Omega)$, there exists a subsequence $(\tilde{w}_{n_k})_{k \in \mathbb{N}}$ and $w \in L^p(\Omega)$ such that $\tilde{w}_{n_k} \rightharpoonup w$ in $L^p(\Omega)$ for $1 < p < \infty$ (or $\tilde{w}_{n_k} \xrightarrow{*} w$ in $L^\infty(\Omega)$). It remains to show that

$$\|\nabla u - w\|_{L^1(\Omega)} \leq \liminf_{k \rightarrow \infty} \|\nabla u - \tilde{w}_{n_k}\|_{L^1(\Omega)}.$$

If $1 < p < \infty$ then by the dual expression of the L^1 norm, one has that

$$\|\nabla u - \tilde{w}_{n_k}\|_{L^1(\Omega)} \geq \int_{\Omega} (\nabla u - \tilde{w}_{n_k})v \, dx, \quad \forall v \in L^\infty(\Omega) \subset L^q(\Omega) \quad \text{with} \quad \|v\|_\infty \leq 1.$$

Then, we have

$$\liminf_{k \rightarrow \infty} \|\nabla u - \tilde{w}_{n_k}\|_{L^1(\Omega)} \geq \liminf_{k \rightarrow \infty} \int_{\Omega} (\nabla u - \tilde{w}_{n_k})v \, dx = \int_{\Omega} (\nabla u - w)v \, dx,$$

for all $v \in L^\infty(\Omega) \subset L^q(\Omega)$ with $\|v\|_\infty \leq 1$. Taking the supremum over all $\|v\|_\infty \leq 1$, we conclude that

$$\|\nabla u - w\|_{L^1(\Omega)} \leq \liminf_{k \rightarrow \infty} \|\nabla u - \tilde{w}_{n_k}\|_{L^1(\Omega)}.$$

The case $p = \infty$ is proved using similar steps. Hence,

$$\inf F \leq F(w) \leq \liminf_{k \rightarrow \infty} F(\tilde{w}_{n_k}) = \inf F,$$

and thus w is a solution to the problem (4.7). \square

It is often useful to identify a *primal* problem, as (4.7) with its *dual* formulation. For example, the dual formulation of the $\text{ICTV}_{\alpha,\beta}$ regulariser in (4.5) is

$$\text{ICTV}_{\alpha,\beta}(u) = \sup_{\substack{\phi_1 \in \mathcal{C}_0^\infty(\Omega, \mathbb{R}^d) \\ \phi_2 \in \mathcal{C}_0^\infty(\Omega, \text{Sym}^2(\mathbb{R}^d)) \\ \|\phi\|_\infty \leq \alpha, \|\phi_2\|_\infty \leq \beta \\ \text{div}^2(\phi_2) = \text{div}(\phi_1)}} \int_{\Omega} u \text{div}^2(\phi_2) \, dx$$

and for (4.6) is the same but with a different connection for ϕ_1 and ϕ_2 ,

$$\text{TGV}_{\alpha,\beta}^2(u) = \sup_{\substack{\phi_1 \in \mathcal{C}_0^\infty(\Omega, \mathbb{R}^d) \\ \phi_2 \in \mathcal{C}_0^\infty(\Omega, \text{Sym}^2(\mathbb{R}^d)) \\ \|\phi\|_\infty \leq \alpha, \|\phi_2\|_\infty \leq \beta \\ \text{div}(\phi_2) = \phi_1}} \int_{\Omega} u \text{div}^2(\phi_2) \, dx$$

see [BBBM13], [BKP10]. $\text{Sym}^2(\mathbb{R}^d)$ denotes the space of symmetric tensors of order 2. In our case, we have that

$$\text{TVL}_{\alpha,\beta}^p(u) = \sup \left\{ \int_{\Omega} u \text{div} \phi \, dx : \phi \in \mathcal{C}_c^1(\Omega), \|\phi\|_\infty \leq \alpha, \|\phi\|_{L^q(\Omega)} \leq \beta \right\}, \quad (4.9)$$

which constitutes another verification of our first-order method. The equivalence of (4.9) and (4.7) is shown below.

Proposition 4.2.2. *Let $u \in \text{BV}(\Omega)$, then*

$$\min_{w \in L^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)} = \sup \left\{ \int_{\Omega} u \operatorname{div} \phi \, dx : \phi \in \mathcal{C}_c^1(\Omega), \|\phi\|_{\infty} \leq \alpha, \|\phi\|_{L^q(\Omega)} \leq \beta \right\}$$

for $1 < p \leq \infty$.

Proof. First notice that in (4.9), we can replace $\mathcal{C}_c^1(\Omega)$ by $\mathcal{C}_0^1(\Omega)$, since $\overline{\mathcal{C}_c^1(\Omega)} = \mathcal{C}_0^1(\Omega)$ with the closure taken with respect to the \mathcal{C}^1 norm. We define

$$\begin{aligned} X &= \mathcal{C}_0^1(\Omega), \\ F_1 : X &\rightarrow \overline{\mathbb{R}}, \text{ with } F_1(\phi) = \mathcal{I}_{\{\|\cdot\|_{L^q(\Omega)} \leq \beta\}}(\phi), \\ F_2 : X &\rightarrow \overline{\mathbb{R}}, \text{ with } F_2(\phi) = \mathcal{I}_{\{\|\cdot\|_{\infty} \leq \alpha\}}(\phi) - \int_{\Omega} u \operatorname{div} \phi \, dx. \end{aligned}$$

Then, we can rewrite (4.9) as

$$\text{TVL}_{\alpha, \beta}^p(u) = - \inf_{\substack{\phi \in X \\ \|\phi\|_{\infty} \leq \alpha \\ \|\phi\|_{L^q(\Omega)} \leq \beta}} \left\{ - \int_{\Omega} u \operatorname{div} \phi \, dx \right\} = - \inf_{\phi \in X} F_1(\phi) + F_2(\phi).$$

The Fenchel–Rockafellar duality theory, see Section 2.4.1, allows to establish a relation between the primal problem

$$- \inf_{\phi \in X} F_1(\phi) + F_2(\phi),$$

and its dual

$$\min_{w \in X^*} F_1^*(-w) + F_2^*(w).$$

Here F_1^* and F_2^* denote the convex conjugate of F_1 and F_2 respectively. In order to obtain such a connection, we follow [AB86] where it suffices to show that

$$\bigcup_{\lambda \geq 0} \lambda (\operatorname{dom} F_2 - \operatorname{dom} F_1)$$

is a closed vector space. Indeed, we have that

$$\bigcup_{\lambda \geq 0} \lambda (\operatorname{dom} F_2 - \operatorname{dom} F_1) \subset X$$

and for every $\phi \in X$, we can write $\phi = \lambda(\lambda^{-1}\phi - 0)$ with $\|\lambda^{-1}\phi\|_{\infty} \leq \alpha$ and $0 \in \operatorname{dom} F_1$. Hence, $\bigcup_{\lambda \geq 0} \lambda (\operatorname{dom} F_2 - \operatorname{dom} F_1) = X$ is a closed vector space and there is no duality gap i.e.,

$$\inf_{\phi \in X} \{F_1(\phi) + F_2(\phi)\} + \min_{w \in X^*} \{F_1^*(-w) + F_2^*(w)\} = 0.$$

Finally, we compute that

$$F_1^*(-w) = \sup_{\substack{\phi \in \mathcal{C}_0^1(\Omega) \\ \|\phi\|_{L^q(\Omega)} \leq \beta}} \langle \phi, w \rangle = \sup_{\substack{\phi \in L^q(\Omega) \\ \|\phi\|_{L^q(\Omega)} \leq \beta}} \langle \phi, w \rangle = \beta \|w\|_{L^p(\Omega)}$$

and similarly,

$$F_2^*(w) = \sup_{\substack{\phi \in \mathcal{C}_0^1(\Omega) \\ \|\phi\|_\infty \leq \alpha}} \langle w, \phi \rangle + \langle u, \operatorname{div} \phi \rangle = \sup_{\substack{\phi \in \mathcal{C}_0^1(\Omega) \\ \|\phi\|_\infty \leq \alpha}} \langle -Du + w, \phi \rangle = \alpha \|Du - w\|_{\mathcal{M}}.$$

Thus the desired equality is proven. \square

Remark 4.2.3. *The dual formulation of $\operatorname{TVL}_{\alpha,\beta}^p : L^1(\Omega) \rightarrow \overline{\mathbb{R}}$ is useful since one can easily derive that $\operatorname{TVL}_{\alpha,\beta}^p$ is lower semicontinuous with respect to the strong L^1 topology as a pointwise supremum of continuous functions.*

We proceed with a relation of the $\operatorname{TVL}_{\alpha,\beta}^p$ and the TV regulariser. The following lemma shows that the $\operatorname{TVL}_{\alpha,\beta}^p$ functional is Lipschitz equivalent to the total variation seminorm.

Lemma 4.2.4. *Let $u \in L^1(\Omega)$ and $1 < p \leq \infty$. Then $\operatorname{TVL}_{\alpha,\beta}^p(u) < \infty$ if and only if $u \in \operatorname{BV}(\Omega)$ and there exist constants $0 < C_1, C_2 < \infty$ such that*

$$C_2 \|Du\|_{\mathcal{M}} \leq \operatorname{TVL}_{\alpha,\beta}^p(u) \leq C_1 \|Du\|_{\mathcal{M}}. \quad (4.10)$$

Proof. Let $u \in \operatorname{BV}(\Omega)$, using (4.7) we have that

$$\operatorname{TVL}_{\alpha,\beta}^p(u) \leq \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)},$$

for every $w \in L^p(\Omega)$. Setting $w = 0$ and $C_1 = \alpha$, we obtain

$$\operatorname{TVL}_{\alpha,\beta}^p(u) \leq C_1 \|Du\|_{\mathcal{M}}.$$

For the other direction, we have that for any $w \in L^p(\Omega) \subset L^1(\Omega)$ by the triangle inequality and from the L^p inclusion, see Proposition 2.2.2, we have

$$\begin{aligned} \|Du\|_{\mathcal{M}} &\leq \|Du - w\|_{\mathcal{M}} + \|w\|_{L^1(\Omega)} \leq \|Du - w\|_{\mathcal{M}} + |\Omega|^{\frac{1}{q}} \|w\|_{L^p(\Omega)} \\ &\leq C \left(\|Du - w\|_{\mathcal{M}} + \|w\|_{L^p(\Omega)} \right), \end{aligned}$$

with $C = \max(1, |\Omega|^{\frac{1}{q}})$. Then, for $C_2 = \frac{1}{C}$ with $\tilde{C} = \max(\frac{1}{\alpha}, \frac{1}{\beta})$ we obtain

$$C_2 \|Du\|_{\mathcal{M}} \leq \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)}$$

which yields the left-hand side inequality by taking the minimum over $w \in L^p$. \square

Having shown the basic properties of the $\text{TVL}_{\alpha,\beta}^p$ functional, we can use it as a regulariser for variational imaging problems, by minimising

$$\min_{u \in \text{BV}(\Omega)} \frac{1}{s} \|f - Tu\|_{L^s(\Omega)}^s + \text{TVL}_{\alpha,\beta}^p(u), \quad 1 \leq s < \infty, \quad (4.11)$$

where $T : L^s(\Omega) \rightarrow L^s(\Omega)$ is a bounded, linear operator and $f \in L^s(\Omega)$. We conclude our analysis with existence and uniqueness results for the minimisation problem (4.11).

Theorem 4.2.5. *Let $1 < p \leq \infty$ and $f \in L^s(\Omega)$. If $T(\mathcal{X}_\Omega) \neq 0$ then there exists a solution $u \in L^s(\Omega) \cap \text{BV}(\Omega)$ for the problem (4.11). If $s > 1$ and T is injective then the solution is unique.*

Proof. The proof is a straightforward application of the direct method of calculus of variations, see for instance [Ves01] or the Theorem 3.5.2 in the previous chapter. For a minimising sequence $u_n \in L^s(\Omega) \cap \text{BV}(\Omega)$, using Lemma 4.2.4 and that T does not annihilate constants, we can prove that (u_n) is BV bounded. Hence, $u_{n_k} \xrightarrow{*} u$ and the lower semicontinuity of $\text{TVL}_{\alpha,\beta}^p$ ensures that u is a solution. Finally, if $s > 1$, then the corresponding energy functional is strictly convex and the uniqueness is valid if T is injective. \square

Since we are mainly interested in studying the regularising properties of $\text{TVL}_{\alpha,\beta}^p$, from now on we focus on the case where $s = 2$ and T is the identity function, i.e., denoising task where rigorous analysis can be carried out. We thus define the following problem

$$\min_{u \in \text{BV}(\Omega)} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \text{TVL}_{\alpha,\beta}^p(u),$$

or equivalently

$$\min_{\substack{u \in \text{BV}(\Omega) \\ w \in L^p(\Omega)}} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)}. \quad (\mathcal{P})$$

Remark 4.2.6. *In general the minimisation problem (\mathcal{P}) is not well-defined for $p = 1$. For instance, when $\beta \geq \alpha$ by Proposition 4.4.6, we have that $\text{TVL}_{\alpha,\beta}^p(u)$ is equivalent to $\alpha \|Du\|_{\mathcal{M}}$ and the problem is well-posed, see [AV94, Ves01]. However, when $\alpha > \beta$ using Remark 2.1.5, we have that*

$$\text{TVL}_{\alpha,\beta}^1(u) = \min_{w \in L^1(\Omega)} \alpha \|\nabla u - w\|_{L^1(\Omega)} + \beta \|w\|_{L^1},$$

where the optimal solution is $w = \nabla u$. Therefore, the minimisation becomes

$$\inf_{u \in \text{BV}(\Omega)} \alpha \|D^s u\|_{\mathcal{M}} + \beta \|\nabla u\|_{L^1(\Omega)} + \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2, \text{ or equivalent}$$

$$\inf_{u \in \text{BV}(\Omega)} \left\{ F(u) := (\alpha - \beta) \|D^s u\|_{\mathcal{M}} + \beta \|Du\|_{\mathcal{M}} + \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 \right\}.$$

We claim that if $u \in \text{BV}(\Omega)$ is a solution of the above minimisation problem, then $D^s u = 0$. Assume by contradiction that $D^s u \neq 0$, we can find $(u_n)_n \in W^{1,1}(\Omega)$, see [ABM14, Section 10.1], such that

$$u_n \rightarrow u \text{ in } L^2(\Omega), \quad \|Du_n\|_{\mathcal{M}} \rightarrow \|Du\|_{\mathcal{M}}.$$

Then, we have that $\liminf_{n \rightarrow \infty} F(u_n) < F(u)$, which is a contradiction since u is a minimiser. Hence, if u is a solution of the above problem then $D^s u = 0$ and it is also a solution of

$$\inf_{u \in W^{1,1}(\Omega)} \beta \|Du\|_{\mathcal{M}} + \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2,$$

which is not a well posed problem in general.

4.3 The p -homogeneous analogue and relation to Huber-TV

In this section, we focus on the $p < \infty$ case and consider the p -homogeneous analogue of (\mathcal{P}) , where we replace $\|w\|_{L^p(\Omega)}$ with $\frac{1}{p} \|w\|_{L^p(\Omega)}^p$, that is

$$\min_{\substack{u \in \text{BV}(\Omega) \\ w \in L^p(\Omega)}} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta}{p} \|w\|_{L^p(\Omega)}^p, \quad 1 < p < \infty. \quad (\mathcal{P}_{p\text{-hom}})$$

Basically, we would like to understand how the variational problem changes with respect to the one homogeneous model. The reason for the introduction of $(\mathcal{P}_{p\text{-hom}})$ is that, in certain cases, it is technically easier to derive exact solutions for $(\mathcal{P}_{p\text{-hom}})$ rather than for (\mathcal{P}) straightforwardly, see Section 4.4.3. We show in Proposition 4.3.2 that there is a strong connection between the models (\mathcal{P}) and $(\mathcal{P}_{p\text{-hom}})$. Here, we can guarantee the uniqueness of the optimal w^* , since

$$w^* = \operatorname{argmin}_{w \in L^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta}{p} \|w\|_{L^p(\Omega)}^p$$

and thus w^* is unique as a minimiser of a strictly convex functional with $1 < p < \infty$. Hence, compared to (\mathcal{P}) , an optimal solution pair of $(\mathcal{P}_{p\text{-hom}})$ is unique. The next Proposition says that, unless f is a constant function then the optimal w in $(\mathcal{P}_{p\text{-hom}})$ cannot be 0 but nonetheless converges to 0 as $\beta \rightarrow \infty$. In essence, this means that one cannot obtain TV-type solutions with the p -homogeneous model.

Proposition 4.3.1. *Let $1 < p < \infty$, $f \in L^2(\Omega)$ and let (w^*, u^*) be an optimal solution pair of the p -homogeneous problem $(\mathcal{P}_{p\text{-hom}})$. Then $w^* = 0$ if and only if f is a constant function. For general data f , we have that $w^* \rightarrow 0$ in $L^p(\Omega)$ for $\beta \rightarrow \infty$.*

Proof. It follows immediately that if f is constant then $(0, f)$ is the optimal pair for $(\mathcal{P}_{p\text{-hom}})$. Suppose that (w^*, u^*) solve $(\mathcal{P}_{p\text{-hom}})$. Notice that in this case we also have

$$u^* = \operatorname{argmin}_{u \in \operatorname{BV}(\Omega)} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du - w^*\|_{\mathcal{M}}. \quad (4.12)$$

Suppose now that $w^* = 0$. Then (4.12) becomes

$$u^* = \operatorname{argmin}_{u \in \operatorname{BV}(\Omega)} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du\|_{\mathcal{M}}. \quad (4.13)$$

Furthermore, since $(0, u^*)$ solve $(\mathcal{P}_{p\text{-hom}})$, then for every $h \in C_c^\infty(\Omega)$ and $\varepsilon > 0$, the pair $(\varepsilon \nabla h, u^* + \varepsilon h) \in L^p(\Omega) \times \operatorname{BV}(\Omega)$ is suboptimal for $(\mathcal{P}_{p\text{-hom}})$, i.e.,

$$\frac{1}{2} \|f - u^*\|_{L^2(\Omega)}^2 + \alpha \|Du^*\|_{\mathcal{M}} \leq \frac{1}{2} \|f - u^* - \varepsilon h\|_{L^2(\Omega)}^2 + \alpha \|D(u^* + \varepsilon h) - \varepsilon \nabla h\|_{\mathcal{M}} + \frac{\beta}{p} \|\varepsilon \nabla h\|_{L^p(\Omega)}^p,$$

from which we take

$$\begin{aligned} \frac{1}{2} \|f - u^*\|_{L^2(\Omega)}^2 &\leq \frac{1}{2} \|f - u^* - \varepsilon h\|_{L^2(\Omega)}^2 + \frac{\beta}{p} \|\varepsilon \nabla h\|_{L^p(\Omega)}^p \\ 0 &\leq \frac{\varepsilon^2}{2} \|h\|_{L^2(\Omega)}^2 - \varepsilon \int_{\Omega} (f - u^*)h \, dx + \frac{\beta \varepsilon^p}{p} \|\nabla h\|_{L^p(\Omega)}^p. \end{aligned}$$

By dividing the last inequality by ε and taking the limit $\varepsilon \rightarrow 0$ we have that $\int_{\Omega} (f - u^*)h \, dx \leq 0$. By considering the analogous perturbations $u^* - \varepsilon h$, we obtain similarly that $\int_{\Omega} (f - u^*)h \, dx \geq 0$ and thus

$$\int_{\Omega} (f - u^*)h \, dx = 0 \quad \forall h \in C_c^\infty(\Omega).$$

Hence $u^* = f$ and by taking the optimality condition of (4.13) we get that $0 \in \partial \|D(\cdot)\|_{\mathcal{M}}(f)$, which implies that $Df = 0$, i.e., f is a constant function. Finally, for the last part of the proposition, simply observe that for every $u \in \operatorname{BV}(\Omega)$ and $w \in L^p(\Omega)$ we have that

$$\frac{1}{2} \|f - u^*\|_{L^2(\Omega)}^2 + \alpha \|Du^* - w^*\|_{\mathcal{M}} + \frac{\beta}{p} \|w^*\|_{L^p(\Omega)}^p \leq \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta}{p} \|w\|_{L^p(\Omega)}^p$$

and setting $u = w = 0$, we obtain

$$\frac{1}{p} \|w^*\|_{L^p(\Omega)}^p \leq \frac{1}{2\beta} \|f\|_{L^2(\Omega)}^2,$$

and thus $\|w^*\|_{L^p(\Omega)}^p \rightarrow 0$ when $\beta \rightarrow \infty$. \square

We can further establish a connection, up to a rescaling of the β 's parameters, between the 1-homogeneous (\mathcal{P}) and the p -homogeneous model $(\mathcal{P}_{p\text{-hom}})$.

Proposition 4.3.2. *Let $1 < p < \infty$ and $f \in L^2(\Omega)$ not a constant. A pair (w^*, u^*) is a solution of $(\mathcal{P}_{p\text{-hom}})$ with parameters $(\alpha, \beta_{p\text{-hom}})$ if and only if it is also a solution of (\mathcal{P}) with parameters $(\alpha, \beta_{1\text{-hom}})$ where $\beta_{1\text{-hom}} = \beta_{p\text{-hom}} \|w^*\|_{L^p(\Omega)}^{p-1}$.*

Proof. Since f is not a constant by the previous proposition we have that $w^* \neq 0$. Note that for an arbitrary function $u \in \text{BV}(\Omega)$:

$$\begin{aligned} w^* \in \operatorname{argmin}_{w \in L^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta_{p\text{-hom}}}{p} \|w\|_{L^p(\Omega)}^p &\Leftrightarrow \\ 0 \in \alpha \partial \|Du - \cdot\|_{\mathcal{M}}(w^*) + \beta_{p\text{-hom}} |w^*|^{p-2} w^* &\Leftrightarrow \\ 0 \in \alpha \partial \|Du - \cdot\|_{\mathcal{M}}(w^*) + \frac{\beta_{1\text{-hom}}}{\|w^*\|_{L^p(\Omega)}^{p-1}} |w^*|^{p-2} w^* &\Leftrightarrow \\ w^* \in \operatorname{argmin}_{w \in L^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta_{1\text{-hom}} \|w\|_{L^p(\Omega)} &\cdot \end{aligned}$$

This means that w^* is an admissible solution for both problems (\mathcal{P}) and $(\mathcal{P}_{p\text{-hom}})$, with the corresponding set of parameters $(\alpha, \beta_{1\text{-hom}})$ and $(\alpha, \beta_{p\text{-hom}})$ respectively. The fact that the same holds for u^* as well, is due to

$$u^* \in \operatorname{argmin}_{u \in \text{BV}(\Omega)} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du - w^*\|_{\mathcal{M}},$$

which holds for both problems. \square

If we examine a particular case of $p = 2$, it turns out that problem $(\mathcal{P}_{p\text{-hom}})$ is essentially equivalent to the widely used *Huber* regularisation, [Hub64, Section 4]. We recall that in the Huber total variation regularisation we approximate the gradient to behave quadratically near 0, i.e., near flat regions and linearly otherwise, see Figure 4.1. This approximation of the gradient is very useful in order to reduce the staircasing that the total variation creates, see for instance [PCBC10],[AHH06]. Let $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ with



Figure 4.1: Huber function

and the Huber-TV minimisation problem is written as:

$$\operatorname{argmin}_{u \in \text{BV}(\Omega)} \int_{\Omega} \varphi(\nabla u) \, dx + \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2. \quad (4.14)$$

Proposition 4.3.3. Consider the functional $F : \text{BV}(\Omega) \rightarrow \mathbb{R}$ as

$$F(u) = \min_{w \in L^2(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta}{2} \|w\|_{L^2(\Omega)}^2, \quad (4.15)$$

which is the 2-homogeneous analogue of $\text{TVL}_{\alpha, \beta}^2$. Then

$$F(u) = \int_{\Omega} \varphi(\nabla u) \, dx + \alpha |D^s u|(\Omega),$$

where $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ with

$$\varphi(x) = \begin{cases} \alpha|x| - \frac{\alpha^2}{2\beta}, & |x| \geq \frac{\alpha}{\beta}, \\ \frac{\beta}{2}|x|^2, & |x| \leq \frac{\alpha}{\beta}. \end{cases}$$

Proof. Using the Remark 2.1.5, we have

$$\begin{aligned} F(u) &= \min_{w \in L^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta}{2} \|w\|_{L^2(\Omega)}^2 \\ &= \alpha |D^s u|(\Omega) + \alpha \min_{w \in L^2(\Omega)} \int_{\Omega} |\nabla u - w| + \frac{\beta}{2\alpha} |w|^2 \, dx. \end{aligned}$$

So we focus on the minimisation problem

$$\min_{w \in L^2(\Omega)} \int_{\Omega} |\nabla u - w| + \frac{\beta}{2\alpha} |w|^2 \, dx. \quad (4.16)$$

Bearing in mind that (as it can easily be checked) for $c \in \mathbb{R}^d$ and $\gamma > 0$ that

$$\operatorname{argmin}_{y \in \mathbb{R}^d} |c - y| + \frac{\gamma}{2} |y|^2 = \begin{cases} \frac{1}{\gamma} \frac{c}{|c|} & \text{if } |c| \geq \frac{1}{\gamma}, \\ c & \text{if } |c| < \frac{1}{\gamma}, \end{cases}$$

and

$$\min_{y \in \mathbb{R}^d} |c - y| + \frac{\gamma}{2} |y|^2 = \begin{cases} |c| - \frac{1}{2\gamma} & \text{if } |c| \geq \frac{1}{\gamma}, \\ \frac{\gamma}{2} |c|^2 & \text{if } |c| < \frac{1}{\gamma}. \end{cases}$$

It is straightforwardly verified setting $\gamma = \beta/\alpha$ that the function

$$w^* = \frac{\alpha}{\beta} \frac{\nabla u}{|\nabla u|} \mathcal{X}_{\{|\nabla u| \geq \frac{\alpha}{\beta}\}} + \nabla u \mathcal{X}_{\{|\nabla u| < \frac{\alpha}{\beta}\}},$$

belongs to $L^\infty(\Omega) \subset L^2(\Omega)$ and solves (4.16) with optimal value equal to $\frac{1}{\alpha} \int_{\Omega} \varphi(\nabla u) dx$. \square

The result above can also be generalised for general $1 < p < \infty$ in $(\mathcal{P}_{p\text{-hom}})$ which is equivalent to a generalised Huber total variation regularisation, see also [HS06b].

Proposition 4.3.4. *Let $1 < p < \infty$ and consider the functional $\text{TVL}_{\alpha,\beta}^{p\text{-hom}} : \text{BV}(\Omega) \rightarrow \mathbb{R}$ with*

$$\text{TVL}_{\alpha,\beta}^{p\text{-hom}}(u) = \min_{w \in L^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta}{p} \|w\|_{L^p(\Omega)}^p. \quad (4.17)$$

Then

$$\text{TVL}_{\alpha,\beta}^{p\text{-hom}}(u) = \int_{\Omega} \varphi_p(\nabla u) dx + \alpha |D^s u|(\Omega),$$

where $\varphi_p : \mathbb{R}^d \rightarrow \mathbb{R}$ with

$$\varphi_p(x) = \begin{cases} \alpha |x| - \left(1 - \frac{1}{p}\right) \frac{\alpha}{\lambda^{\frac{1}{p-1}}}, & |x| \geq \frac{1}{\lambda^{\frac{1}{p-1}}}, \\ \frac{\beta}{p} |x|^p, & |x| \leq \frac{1}{\lambda^{\frac{1}{p-1}}}, \end{cases} \quad \text{where } \lambda := \frac{\beta}{\alpha}.$$

Proof. Similarly to the previous proposition, we have

$$\text{TVL}_{\alpha,\beta}^{p\text{-hom}}(u) = \alpha |D^s u|(\Omega) + \alpha \min_{w \in L^p(\Omega)} \int_{\Omega} |\nabla u - w| + \frac{\beta}{p\alpha} |w|^p dx$$

and it suffices to examine the problem

$$\min_{w \in L^p(\Omega)} \int_{\Omega} |\nabla u - w| + \frac{\beta}{p\alpha} |w|^p dx. \quad (4.18)$$

Then for $c \in \mathbb{R}^d$, $\lambda > 0$,

$$\operatorname{argmin}_{y \in \mathbb{R}^d} |c - y| + \frac{\lambda}{p} |y|^p = \begin{cases} \frac{1}{\lambda^{\frac{1}{p-1}}} \frac{c}{|c|} & \text{if } |c| \geq \frac{1}{\lambda^{\frac{1}{p-1}}}, \\ c & \text{if } |c| < \frac{1}{\lambda^{\frac{1}{p-1}}}, \end{cases}$$

$$\min_{y \in \mathbb{R}^d} |c - y| + \frac{\lambda}{p} |y|^p = \begin{cases} |c| - \left(1 - \frac{1}{p}\right) \frac{1}{\lambda^{\frac{1}{p-1}}} & \text{if } |c| \geq \frac{1}{\lambda^{\frac{1}{p-1}}}, \\ \frac{\lambda}{p} |c|^p & \text{if } |c| < \frac{1}{\lambda^{\frac{1}{p-1}}}, \end{cases}$$

and by setting $\lambda = \beta/\alpha$ the function

$$w^* = \lambda^{-\frac{1}{p-1}} \frac{\nabla u}{|\nabla u|} \mathcal{X}_{\{|\nabla u| \geq \lambda^{-\frac{1}{p-1}}\}} + \nabla u \mathcal{X}_{\{|\nabla u| < \lambda^{-\frac{1}{p-1}}\}},$$

belongs to $L^\infty(\Omega) \subset L^p(\Omega)$ and solves (4.18) with optimal value equal to $\frac{1}{\alpha} \int_{\Omega} \varphi_p(\nabla u) dx$. \square

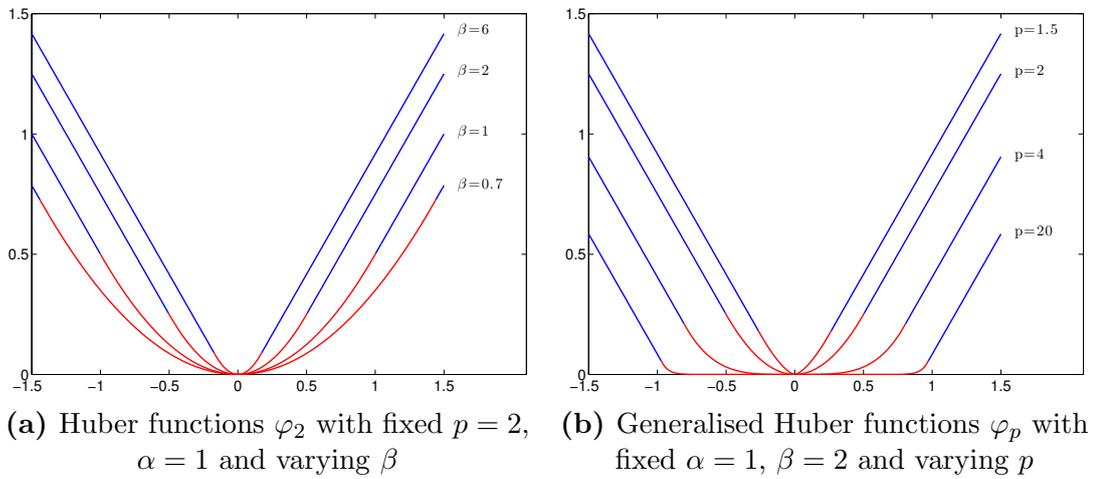


Figure 4.2: Illustration of the forms of the Huber type functions φ_p of Proposition 4.3.4. Their linear and p -power parts are plotted with blue and red colour respectively

For the reader's convenience, in Figure 4.2 we have plotted some of the functions φ_p in order to illustrate how their form changes when their parameters vary. Note for instance in Figure 4.2a how ϕ_2 is converging to an absolute type function when β is getting large, i.e., approaching a total variation regularisation. This can also be seen from Proposition 4.3.1 where the optimal variable w is converging to 0 when $\beta \rightarrow \infty$. On the other hand when p is getting large, Figure 4.2b, small gradients are essentially not penalised at all, allowing the gradient to be almost constant, equal to its maximum value, leading to piecewise affine structures.

4.4 The one dimensional case

In order to get more insight into the structure of solutions of the problem (\mathcal{P}) , in this section we study its one dimensional version. For this Section $\Omega \subset \mathbb{R}$ is an open and bounded

interval, i.e., $\Omega = (a, b)$. Similar analysis has been presented in [BKV12] and [PB15] where the one dimensional L^1 -TGV² and L^2 -TGV² problems are studied respectively.

4.4.1 Optimality conditions

The main outcome of this section is the derivation of the optimality conditions for the one dimensional problem (\mathcal{P}) . These conditions can provide practical information about the structure of the solutions for every value of α and β and for every $p \in (1, \infty]$. We begin our analysis by defining the predual problem (\mathcal{P}') , proving existence and uniqueness for its solutions. Then, we follow similar steps as in Proposition 4.2.2 using the Fenchel–Rockafellar duality theory in order to find a connection between their solution pairs, see Section 2.4.1.

We define the predual problem (\mathcal{P}')

$$-\inf \left\{ \int_{\Omega} f \phi' dx + \frac{1}{2} \int_{\Omega} (\phi')^2 dx : \phi \in H_0^1(\Omega), \|\phi\|_{L^q(\Omega)} \leq \beta, \|\phi\|_{\infty} \leq \alpha \right\}, \quad (\mathcal{P}')$$

where the conjugate exponent of $p \in (1, \infty]$ is

$$q = \begin{cases} \frac{p}{p-1}, & \text{if } 1 < p < \infty, \\ 1, & \text{if } p = \infty. \end{cases} \quad (4.19)$$

Proposition 4.4.1. *For $f \in L^2(\Omega)$, the predual problem (\mathcal{P}') admits a unique solution in $H_0^1(\Omega)$.*

Proof. Let $\mathcal{A} := \left\{ \phi \in H_0^1(\Omega) : \|\phi\|_{L^q(\Omega)} \leq \beta, \|\phi\|_{\infty} \leq \alpha \right\}$ be the admissible set of (\mathcal{P}') . Consider a minimising sequence $(\phi_n)_{n \in \mathbb{N}} \in \mathcal{A}$, i.e., $\phi_n \in H_0^1(\Omega)$ with $\|\phi_n\|_{L^q(\Omega)} \leq \beta$ and $\|\phi_n\|_{\infty} \leq \alpha$. Then there exists $M > 0$ such that

$$\int_{\Omega} f \phi_n' dx + \frac{1}{2} \int_{\Omega} (\phi_n')^2 dx \leq M.$$

Clearly, $(\phi_n')_{n \in \mathbb{N}}$ is bounded in $L^2(\Omega)$ and also $\|\phi_n\|_{L^2(\Omega)} \leq C \|\phi_n\|_{\infty} \leq C\alpha$, hence it is bounded in $H_0^1(\Omega)$. By the reflexivity of $H_0^1(\Omega)$ there exists a weakly convergent subsequence $(\phi_{n_k})_{k \in \mathbb{N}}$ and $\phi \in H_0^1(\Omega)$ such that $\phi_{n_k} \rightharpoonup \phi$ in $H_0^1(\Omega)$. Since \mathcal{A} is convex and closed, it is also weakly closed by the Mazur theorem, see [Bre11], and thus $\phi \in \mathcal{A}$. Finally, the minimising energy functional is weakly lower semicontinuous and ϕ is a unique solution for the predual problem due to its strict convexity. \square

Observe now that we can also write down the predual problem (\mathcal{P}') using the following

equivalent formulation:

$$- \inf_{(\phi, \xi) \in X} F_1(\phi, \xi) + F_2(K(\phi, \xi)), \quad (4.20)$$

where $X = H_0^1(\Omega) \times H_0^1(\Omega)$, $Y = H_0^1(\Omega) \times L^2(\Omega)$ and

$$\begin{aligned} K : X &\rightarrow Y, \quad K(\phi, \xi) = (\xi - \phi, \xi'), \\ F_1 : X &\rightarrow \overline{\mathbb{R}}, \quad \text{with } F_1(\phi, \xi) = \mathcal{I}_{\{\|\cdot\|_{L^q(\Omega)} \leq \beta\}}(\phi) + \mathcal{I}_{\{\|\cdot\|_\infty \leq \alpha\}}(\xi), \\ F_2 : Y &\rightarrow \overline{\mathbb{R}}, \quad \text{with } F_2(\phi, \psi) = \mathcal{I}_{\{0\}}(\phi) + \int_{\Omega} f\psi \, dx + \frac{1}{2} \int_{\Omega} \psi^2 \, dx. \end{aligned} \quad (4.21)$$

We denote the infimum in (\mathcal{P}') as $\inf \mathcal{P}^*$. Then, it is immediate that

$$- \inf \mathcal{P}^* = - \inf_{(\phi, \xi) \in X} F_1(\phi, \xi) + F_2(K(\phi, \xi)).$$

The dual problem of (4.20) is defined as

$$\min_{(w, u) \in Y^*} F_1^*(-K^*(w, u)) + F_2^*(w, u), \quad (4.22)$$

where K^* denotes the adjoint of K . Let (σ, τ) be elements of $H_0^1(\Omega)^* \times H_0^1(\Omega)^*$. Note that elements on the dual space of $H_0^1(\Omega)$ can be considered as distributions of 0 order, since we restrict $\sigma \in (H_0^1(\Omega))^*$ to $C_c^\infty(\Omega)$ the dense subspace of $H_0^1(\Omega)$, see [ABM14, Section 5.2]. For the convex conjugate of F_1 , we write

$$F_1^*(\sigma, \tau) = \sup_{\substack{(\phi, \xi) \in X \\ \|\phi\|_{L^q(\Omega)} \leq \beta \\ \|\xi\|_\infty \leq \alpha}} \langle \sigma, \phi \rangle + \langle \tau, \xi \rangle = \beta \sup_{\substack{\phi \in H_0^1(\Omega) \\ \|\phi\|_{L^q(\Omega)} \leq 1}} \langle \sigma, \phi \rangle + \alpha \sup_{\substack{\xi \in H_0^1(\Omega) \\ \|\xi\|_\infty \leq 1}} \langle \tau, \xi \rangle. \quad (4.23)$$

In the minimisation (4.22), both terms are finite and by standard density arguments we have that

$$F_1^*(\sigma, \tau) = \beta \sup_{\substack{\phi \in C_c^\infty(\Omega) \\ \|\phi\|_{L^q(\Omega)} \leq 1}} \langle \sigma, \phi \rangle + \alpha \sup_{\substack{\xi \in C_c^\infty(\Omega) \\ \|\xi\|_\infty \leq 1}} \langle \tau, \xi \rangle = \beta \|\sigma\|_{L^p(\Omega)} + \alpha \|\tau\|_{\mathcal{M}}, \quad (4.24)$$

where we have used Theorems 2.2.4 and 2.1.10. Moreover, let $K^* : Y^* \rightarrow X^*$ with

$$\begin{aligned} \langle -K^*(w, u), (\phi, \xi) \rangle &= - \langle (w, u), K(\phi, \xi) \rangle = - \langle (w, u), (\xi - \phi, \xi') \rangle \\ &= - \langle w, \xi \rangle + \langle w, \phi \rangle - \langle u, \xi' \rangle = \langle Du - w, \xi \rangle + \langle w, \phi \rangle. \end{aligned}$$

Hence, we obtain

$$F_1^*(-K^*(w, u)) = \beta \|w\|_{L^p(\Omega)} + \alpha \|Du - w\|_{\mathcal{M}}, \quad (4.25)$$

and

$$\begin{aligned} F_2^*(w, u) &= \sup_{\substack{(\phi, \psi) \in Y \\ \phi=0}} \langle w, \phi \rangle + \langle u, \psi \rangle - \langle f, \psi \rangle - \frac{1}{2} \int_{\Omega} \psi^2 \\ &= \sup_{\psi \in L^2(\Omega)} \langle u - f, \psi \rangle - \frac{1}{2} \int_{\Omega} \psi^2 dx \\ &:= \left(\frac{1}{2} \|\cdot\|_{L^2(\Omega)}^2 \right)^* (u - f) = \frac{1}{2} \|u - f\|_{L^2(\Omega)}^2. \end{aligned} \quad (4.26)$$

Therefore, we have proved the following:

Proposition 4.4.2. *The dual problem of (\mathcal{P}') is equivalent to the primal problem (\mathcal{P}) in the sense that $(w, u) \in Y^*$ is a solution of the dual of (\mathcal{P}') if and only if $(w, u) \in L^p(\Omega) \times \text{BV}(\Omega)$ is a solution of (\mathcal{P}) .*

It remains to verify that we have no duality gap between the two minimisation problems (\mathcal{P}) and (\mathcal{P}') . Similar to Proposition 4.2.2 we use the geometric condition that is presented in [AB86], see Theorem 2.4.10. The proof of the following proposition follows the proof of the corresponding proposition in [BKV12]. For the sake of completeness, we slightly modify it for our case.

Proposition 4.4.3. *Let F_1, F_2, K be defined as in (4.21). Then*

$$Y = \bigcup_{\lambda \geq 0} \lambda(\text{dom}F_2 - K(\text{dom}F_1)) \quad (4.27)$$

is a closed vector space and

$$\min_{(\phi, \xi) \in X} F_1((\phi, \xi)) + F_2(K(\phi, \xi)) + \min_{(w, u) \in Y^*} F_1^*(-K^*(w, u)) + F_2^*((w, u)) = 0. \quad (4.28)$$

Proof. Let $(v, \psi) \in Y$ and define $\psi_0(x) = c_1$, where $c_1 = \frac{1}{|\Omega|} \int_{\Omega} \psi(x) dx$. This condition is useful in order to find elements that lie on the right-hand side of (4.27). For instance, let $\xi(x) = \int_a^x (\psi_0 - \psi)(y) dy$. Since by construction, $\xi' = \psi_0 - \psi \in L^2(\Omega)$ with $\xi(a) = \xi(b) = 0$, we have that $\xi \in H_0^1(\Omega)$. Furthermore, let $\phi = -v + \xi \in H_0^1(\Omega)$ and $(\phi, \xi) \in X$ with

$$(v, \psi) = (\xi - \phi, \psi_0 - \xi') = (0, \psi_0) - (\xi - \phi, \xi') = (0, \psi_0) - K(\phi, \xi).$$

Choosing appropriately $\lambda > 0$ such that $\|\lambda^{-1}\phi\|_{L^q(\Omega)} \leq \beta$, $\|\lambda^{-1}\xi\|_{\infty} \leq \alpha$, we can write

$$(v, \psi) = \lambda \left((0, \lambda^{-1}\psi_0) - K(\lambda^{-1}\phi, \lambda^{-1}\xi) \right),$$

with $\text{dom}F_2 = \{0\} \times L^2(\Omega)$ and $\text{dom}F_1 = \{(\phi, \xi) : \|\phi\|_{L^q(\Omega)} \leq \beta, \|\xi\|_\infty \leq \alpha\}$. Since $(v, \psi) \in Y$ were chosen arbitrarily, (4.27) holds. \square

Since there is no duality gap, we can find a relationship between the solutions of (\mathcal{P}') and (\mathcal{P}) via the optimality conditions, see [ET99, Prop. 4.1 (III)]. The following definition is useful when one wants to characterise optimality conditions based on the subdifferential calculus. It was introduced in [BKV12] as a generalisation of the sign function.

Definition 4.4.4. Let $\mu \in \mathcal{M}(\Omega)$. We define the set-valued sign, $\text{Sgn}(\mu)$ as

$$\text{Sgn}(\mu) = \{v \in L^\infty(\Omega) \cap L^\infty(\Omega, |\mu|) : \|v\|_\infty \leq 1, v = \text{sgn}(\mu), |\mu| - a.e.\}. \quad (4.29)$$

It is proved in [BKV12] that if $\mu \in \mathcal{M}(\Omega)$ then

$$\partial \|\cdot\|_{\mathcal{M}}(\mu) \cap C_0(\Omega) = \text{Sgn}(\mu) \cap C_0(\Omega). \quad (4.30)$$

Theorem 4.4.5 (Optimality conditions). *Let $p \in (1, \infty]$ and $f \in L^2(\Omega)$. A pair $(w, u) \in L^p(\Omega) \times \text{BV}(\Omega)$ is a solution of (\mathcal{P}) if and only if there exists a function $\phi \in H_0^1(\Omega)$ such that*

$$\begin{aligned} \phi' &= u - f, \\ \phi &\in \alpha \text{Sgn}(Du - w), \end{aligned} \quad (4.31)$$

and

$$\left\{ \begin{array}{l} \phi \in \left\{ \tilde{\phi} \in L^q(\Omega) : \|\tilde{\phi}\|_{L^q(\Omega)} \leq \beta \right\}, \quad \text{if } w = 0, 1 < p \leq \infty \\ \phi = \beta \frac{|w|^{p-2}w}{\|w\|_{L^p(\Omega)}^{p-1}}, \quad \text{if } w \neq 0, 1 < p < \infty \\ \phi \in \left\{ \tilde{\phi} \in L^1(\Omega) : \langle \tilde{\phi}, w \rangle = \beta \|w\|_\infty, \|\tilde{\phi}\|_{L^1(\Omega)} \leq \beta \right\}, \quad \text{if } w \neq 0, p = \infty. \end{array} \right. \quad (4.32)$$

Proof. Since there is not duality gap, the optimality conditions read, see Theorem 2.4.9 :

$$(\phi, \xi) \in \partial F_1^*(-K^*(w, u)), \quad (4.33)$$

$$K(\phi, \xi) \in \partial F_2^*(w, u), \quad (4.34)$$

for every (ϕ, ξ) and (w, u) solutions of (\mathcal{P}') and (\mathcal{P}) respectively. Note that in dimension one we have $H_0^1(\Omega) \subseteq \mathcal{C}_0(\Omega)$. Hence, for every $(\sigma, \tau) \in X^*$, we have the following:

$$\begin{aligned} F_1^*(\sigma, \tau) &\geq F_1^*(-K^*(w, u)) + \langle (\sigma, \tau) + K^*(w, u), (\phi, \xi) \rangle \Leftrightarrow \\ \alpha \|\tau\|_{\mathcal{M}} + \beta \|\sigma\|_{L^p(\Omega)} &\geq \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)} + \langle (\sigma, \tau) - (w, Du - w), (\phi, \xi) \rangle \Leftrightarrow \\ \alpha \|\tau\|_{\mathcal{M}} + \beta \|\sigma\|_{L^p(\Omega)} &\geq \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)} + \langle \sigma - w, \phi \rangle + \langle \tau - (Du - w), \xi \rangle \Leftrightarrow \end{aligned}$$

$$\begin{cases} \alpha \|\tau\|_{\mathcal{M}} \geq \alpha \|Du - w\|_{\mathcal{M}} + \langle \tau - (Du - w), \xi \rangle, \forall \tau \in H_0^1(\Omega)^*, \\ \beta \|\sigma\|_{L^p(\Omega)} \geq \beta \|w\|_{L^p(\Omega)} + \langle \sigma - w, \phi \rangle, \forall \sigma \in H_0^1(\Omega)^* \end{cases} \Leftrightarrow \\
 \begin{cases} \alpha \|\tau\|_{\mathcal{M}} \geq \alpha \|Du - w\|_{\mathcal{M}} + \langle \tau - (Du - w), \xi \rangle, \forall \tau \in \mathcal{M}(\Omega), \\ \beta \|\sigma\|_{L^p(\Omega)} \geq \beta \|w\|_{L^p(\Omega)} + \langle \sigma - w, \phi \rangle, \forall \sigma \in L^p(\Omega) \end{cases} \Leftrightarrow \\
 \begin{cases} \xi \in \alpha \partial \|\cdot\|_{\mathcal{M}}(Du - w), \\ \phi \in \beta \partial \|\cdot\|_{L^p(\Omega)}(w), \end{cases}$$

and using (4.30) we can simplify the last expressions to:

$$\xi \in \alpha \text{Sgn}(Du - w).$$

For the (4.32) let us consider the $p < \infty$ and $p = \infty$ separately. If $1 < p < \infty$ the homogeneity of the L^p norm implies that its subdifferential, see Proposition 2.4.8, can be written as:

$$\partial \|\cdot\|_{L^p(\Omega)}(w) = \left\{ z \in (L^p(\Omega))^* = L^q(\Omega) : \langle z, w \rangle = \|w\|_{L^p(\Omega)}, \langle z, \sigma \rangle \leq \|\sigma\|_{L^p(\Omega)}, \forall \sigma \in L^p(\Omega) \right\}.$$

Then, if $w = 0$, the above expression reduces to $\|\sigma\|_{L^p(\Omega)} \geq \langle z, \sigma \rangle, \forall \sigma \in L^p(\Omega)$ which is valid for any $z \in L^q(\Omega)$ with $\|z\|_{L^q(\Omega)} \leq 1$, i.e., the unit ball of $L^q(\Omega)$. If $w \neq 0$ then the subdifferential reduces to the Gâteaux derivative of the L^p norm, i.e., $\partial \|\cdot\|_{L^p(\Omega)}(w) = \frac{|w|^{p-2}w}{\|w\|_{L^p(\Omega)}^{p-1}}$. For the $p = \infty$ case, the subdifferential is written analogously

$$\partial \|\cdot\|_{L^p(\Omega)}(w) = \left\{ z \in (L^\infty(\Omega))^* : \langle z, w \rangle = \|w\|_\infty, \|\sigma\|_\infty \geq \langle z, \sigma \rangle, \forall \sigma \in L^\infty(\Omega) \right\}.$$

The dual space of $L^\infty(\Omega)$ consists of all bounded and finitely additive signed measures, see [Yos80, Chapter IV] for more details. However, since $\phi \in H_0^1(\Omega) \subset L^1(\Omega)$ we do not take into account all the elements of $(L^\infty(\Omega))^*$, rather than the intersection with $L^1(\Omega)$, where $L^1(\Omega) \subset (L^\infty(\Omega))^*$. Hence, we have that

$$\partial \|w\|_\infty = \left\{ z \in L^1(\Omega) : \langle z, w \rangle = \|w\|_\infty, \|\sigma\|_\infty \geq \langle z, \sigma \rangle, \forall \sigma \in L^\infty(\Omega) \right\}$$

where for $w = 0$, it reduces to the unit ball on $L^1(\Omega)$ i.e., $\|z\|_{L^1(\Omega)} \leq 1$. Finally, from (4.34) we have for every $(\hat{w}, \hat{u}) \in Y^*$

$$\begin{aligned} F_2^*(\hat{w}, \hat{u}) &\geq F_2^*(w, u) + \langle K(\phi, \xi), ((\hat{w}, \hat{u}) - (w, u)) \rangle \Leftrightarrow \\ &\frac{1}{2} \int_{\Omega} (f - \hat{u})^2 dx \geq \frac{1}{2} \int_{\Omega} (f - u)^2 dx + \langle (\xi - \phi, \xi'), (\hat{w} - w, \hat{u} - u) \rangle \Leftrightarrow \\ &\frac{1}{2} \int_{\Omega} (f - \hat{u})^2 dx \geq \frac{1}{2} \int_{\Omega} (f - u)^2 dx + \langle \xi - \phi, \hat{w} - w \rangle + \langle \xi', \hat{u} - u \rangle \Leftrightarrow \end{aligned}$$

$$\begin{cases} \langle \xi - \phi, \hat{w} - w \rangle \leq 0, \forall \hat{w} \in (\mathbf{H}_0^1(\Omega))^*, \\ \frac{1}{2} \int_{\Omega} (f - \hat{u})^2 \geq \frac{1}{2} \int_{\Omega} (f - u)^2 + \langle \xi', u - \hat{u} \rangle, \forall \hat{u} \in (\mathbf{L}^2(\Omega))^* \end{cases} \Leftrightarrow \begin{cases} \xi = \phi, \\ \xi' \in \partial \left(\frac{1}{2} \|f - \cdot\|_2^2 \right) (u) = u - f. \end{cases}$$

Combining all the above results, we obtain the optimality conditions (4.31) and (4.32). \square

One can immediately recover, by letting $w = 0$, the optimality conditions for the (ROF) minimisation problem with regularising parameter α i.e.,

$$\min_{u \in \text{BV}(\Omega)} \frac{1}{2} \|f - u\|_{\mathbf{L}^2(\Omega)}^2 + \alpha \|Du\|_{\mathcal{M}}, \quad (\text{ROF})$$

see [Rin00]. On the other hand when $w \neq 0$, the additional condition (4.32) depends on the value of p and as we will see later it allows a certain degree of smoothness in the solution u . Finally, in Table 4.1, we compare our optimality conditions with respect to the TV and $\text{TGV}_{\alpha,\beta}^2$ regularisers for variational problems as in (4.2). The reader can again check the difference between first and second order regularisers. The optimality conditions of $(\mathbf{L}^2 - \text{TGV}_{\alpha,\beta}^2)$ variational problem can be found in [PB15].

Optimality conditions		
$\mathbf{L}^2 - \alpha \text{TV}$ (ROF)	$\mathbf{L}^2 - \text{TVL}_{\alpha,\beta}^p$	$\mathbf{L}^2 - \text{TGV}_{\alpha,\beta}^2$
$u \in \text{BV}(\Omega), \phi \in \mathbf{H}_0^1(\Omega)$	$u \in \text{BV}(\Omega), w \in \mathbf{L}^p(\Omega), \phi \in \mathbf{H}_0^1(\Omega)$	$u, w \in \text{BV}(\Omega), \phi \in \mathbf{H}_0^2(\Omega)$
$\phi' = u - f,$ $\phi \in \alpha \text{Sgn}(Du).$	$\phi' = u - f,$ $\phi \in \alpha \text{Sgn}(Du - w),$ $\ \phi\ _{\mathbf{L}^q(\Omega)} \leq \beta,$ if $w = 0, 1 < p \leq \infty,$ $\phi = \beta \frac{ w ^{p-2} w}{\ w\ _{\mathbf{L}^p(\Omega)}^{(p-1)}},$ if $w \neq 0, 1 < p < \infty,$ $\langle \phi, w \rangle = \beta \ w\ _{\infty}, \ \phi\ _{\mathbf{L}^1(\Omega)} \leq \beta,$ if $w \neq 0, p = \infty.$	$\phi'' = u - f,$ $\phi' \in \alpha \text{Sgn}(Du - w),$ $\phi \in \beta \text{Sgn}(Dw).$

Table 4.1: The optimality conditions for the one dimensional $\mathbf{L}^2 - \text{TV}$, $\mathbf{L}^2 - \text{TVL}_{\alpha,\beta}^p$ and $\mathbf{L}^2 - \text{TGV}_{\alpha,\beta}^2$ denoising problems.

4.4.2 Structure of the solutions

We are now ready to determine the structure of the solutions of (\mathcal{P}) for every value of $p \in (1, \infty]$. It is reasonable to discuss first the cases where the solution u of (\mathcal{P}) is a solution of a corresponding ROF minimisation problem i.e., when $w = 0$.

Proposition 4.4.6 (ROF-solutions). *Let q as in (4.19). If*

$$\frac{\beta}{\alpha} \geq |\Omega|^{\frac{1}{q}}, \quad (4.35)$$

then the solution u of (\mathcal{P}) coincides with the solution of (ROF) and $w = 0$. It is also true for $p = 1$ and $q = \infty$.

Proof. Let (w^*, u^*) be a solution pair for (\mathcal{P}) , then for every $(w, u) \in L^p(\Omega) \times \text{BV}(\Omega)$,

$$\frac{1}{2} \|f - u^*\|_{L^2(\Omega)}^2 + \alpha \|Du^* - w^*\|_{\mathcal{M}} + \beta \|w^*\|_{L^p(\Omega)} \leq \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)}.$$

Setting $w = 0$, we get

$$\frac{1}{2} \|f - u^*\|_{L^2(\Omega)}^2 + \alpha \|Du^* - w^*\|_{\mathcal{M}} + \beta \|w^*\|_{L^p(\Omega)} \leq \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du\|_{\mathcal{M}}. \quad (4.36)$$

Since $w \in L^p(\Omega)$ with $p \in (1, \infty]$, we have that

$$\|w^*\|_{\mathcal{M}} = \|w^*\|_{L^1(\Omega)} \leq |\Omega|^{1-\frac{1}{p}} \|w^*\|_{L^p(\Omega)},$$

and using the condition (4.35) we have

$$\begin{aligned} \alpha \|Du^*\|_{\mathcal{M}} &\leq \alpha \|Du^* - w^*\|_{\mathcal{M}} + \alpha \|w^*\|_{\mathcal{M}} \\ &\leq \alpha \|Du^* - w^*\|_{\mathcal{M}} + \beta |\Omega|^{\frac{1}{p}-1} \|w^*\|_{\mathcal{M}} \\ &\leq \alpha \|Du^* - w^*\|_{\mathcal{M}} + \beta \|w^*\|_{L^p(\Omega)}. \end{aligned} \quad (4.37)$$

From (4.36) and (4.37), we conclude that for every $u \in \text{BV}(\Omega)$,

$$\frac{1}{2} \|f - u^*\|_{L^2(\Omega)}^2 + \alpha \|Du^*\|_{\mathcal{M}} \leq \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du\|_{\mathcal{M}},$$

i.e., u^* solves

$$\min_{u \in \text{BV}(\Omega)} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du\|_{\mathcal{M}} = \min_{\substack{u \in \text{BV}(\Omega) \\ w=0}} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)}.$$

For $p = 1$ and $q = \infty$, the proof is obvious. \square

In fact, what we have essentially proved above is that if (4.35) holds then

$$\text{TVL}_{\alpha, \beta}^p(u) = \alpha \|Du\|_{\mathcal{M}}, \quad \forall u \in \text{BV}(\Omega).$$

Notice also that when (4.35) holds then we have that $w = 0$ is an admissible solution but in general we cannot prove that this solution is unique, unless we treat the p -homogeneous analogue. The condition (4.35) is valid for any dimension $d \geq 1$. It provides a rough threshold for obtaining ROF-type solutions in terms of the regularising parameters α, β and the image domain Ω . However, the condition is not sharp, since as we will see in the following sections, we can obtain a sharper estimate when a specific data of f is examined.

The following proposition is in the spirit of [BKV12, Proposition 4.3] and gives more insight into the structure of solutions of (\mathcal{P}) . It is formulated rigorously via the use of *good representatives* of BV functions, see [AFP00, Theorem 3.28]. The proof remains the same in our case hence it can be omitted.

Proposition 4.4.7. *Let $f \in \text{BV}(\Omega)$ and suppose that $(w, u) \in L^p(\Omega) \times \text{BV}(\Omega)$ is a solution pair for (\mathcal{P}) with $p \in (1, \infty]$. Suppose that $u > f$ (or $u < f$) on an open interval $I \subset \Omega$ then $(Du - w)|_I = 0$ i.e., $u' = w$ on I and $|D^s u|(I) = 0$.*

An acceptable solution of (ROF) regularisation is also the constant solution. Therefore, we consider the case where the solution is constant in Ω , which in fact coincides with the mean value \tilde{f} of the data f :

$$\tilde{f} := \operatorname{argmin}_{u \text{ constant}} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 = \frac{1}{|\Omega|} \int_{\Omega} f \, dx. \quad (4.38)$$

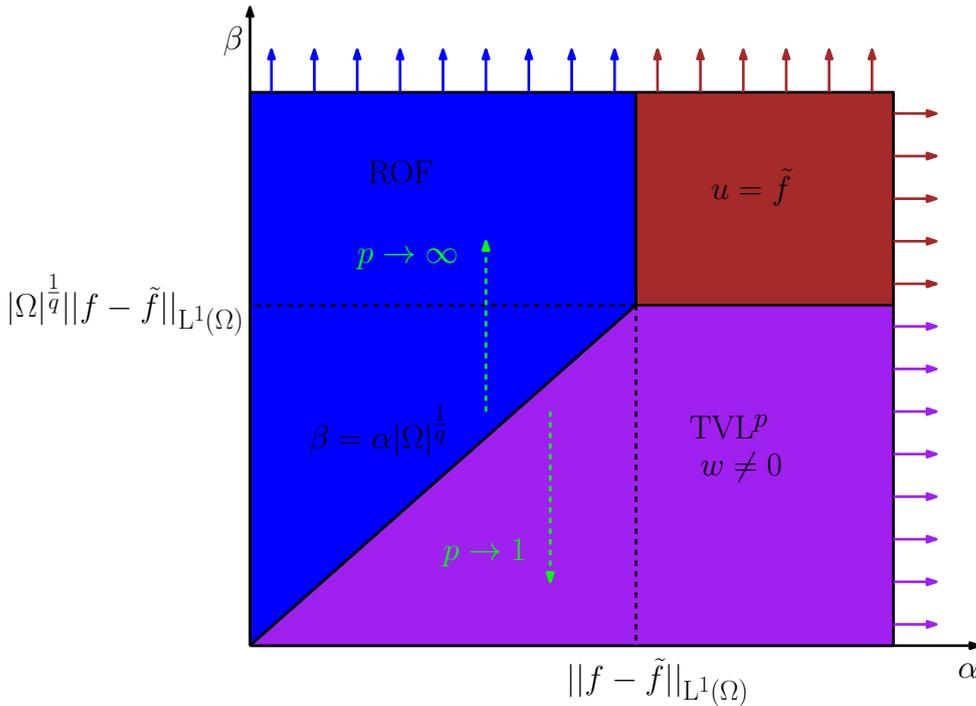


Figure 4.3: Characterisation of solutions of (\mathcal{P}) for any data f : The blue/brown areas correspond to the ROF type solutions ($w = 0$) and the purple area corresponds to the TVL^p solutions ($w \neq 0$) for $1 < p \leq \infty$. We note that the blue/purple areas are potentially larger/smaller as the conditions we have derived are not sharp.

Proposition 4.4.8 (Mean value solution). *If the following conditions hold*

$$\begin{aligned} \alpha &\geq \|f - \tilde{f}\|_{L^1(\Omega)}, \\ \beta &\geq |\Omega|^{\frac{1}{q}} \|f - \tilde{f}\|_{L^1(\Omega)}, \end{aligned} \quad (4.39)$$

then the solution of (\mathcal{P}) is constant and it is equal to \tilde{f} .

Proof. Clearly, if u is a constant solution of (\mathcal{P}) , then $Du = 0$ and from (4.10) we get $\text{TVL}_{\alpha,\beta}^p(u) = 0$ which implies (4.38) and $u = \tilde{f}$. In order to have $u = \tilde{f}$, from the optimality conditions (4.31) and (4.32), it suffices to find a function $\phi \in H_0^1(\Omega)$ and

$$\phi' = \tilde{f} - f, \quad \|\phi\|_\infty \leq \alpha, \quad \|\phi\|_{L^q(\Omega)} \leq \beta.$$

Let $\phi(x) = \int_a^x (\tilde{f} - f(s)) ds$, then obviously $\phi(a) = \phi(b) = 0$ and

$$|\phi(x)| \leq \int_a^x |\tilde{f} - f(s)| ds \leq \|f - \tilde{f}\|_{L^1(\Omega)} < \infty.$$

Therefore, $\|\phi\|_\infty \leq \|f - \tilde{f}\|_{L^1(\Omega)}$. Also, since $L^\infty(\Omega) \subset L^q(\Omega)$ we obtain

$$\|\phi\|_{L^q(\Omega)} \leq |\Omega|^{\frac{1}{q}} \|\phi\|_\infty \leq |\Omega|^{\frac{1}{q}} \|f - \tilde{f}\|_{L^1(\Omega)}.$$

Hence, it suffices to choose α and β as in (4.39). \square

In Figure 4.3, we summarise our results so far. There, we have partitioned the set $\{\alpha > 0, \beta > 0\}$ into different areas that correspond to different types of solutions of the problem (\mathcal{P}) . The brown area, arising from thresholds (4.39), corresponds to the choices of α and β that produce constant solutions while the blue area corresponds to ROF type solutions, according to threshold (4.35). Therefore, we can determine the area where the non-trivial solutions are obtained i.e., $w \neq 0$, see purple region. Note that since the conditions (4.35) and (4.39) are not sharp the blue and the purple areas are potentially larger or smaller respectively as it is shown in Figure 4.3.

Let us now, examine the non-trivial solutions of (\mathcal{P}) ($w \neq 0$) for the $p < \infty$ and $p = \infty$ cases. We begin with the $p < \infty$ case where the solution is obtained by a non-homogenous p -Laplace differential equation, cf. [Kui07].

Proposition 4.4.9 (TVL p -solutions). *Let $f \in \text{BV}(\Omega)$ and suppose that $(w, u) \in L^p(\Omega) \times \text{BV}(\Omega)$ is a solution pair for (\mathcal{P}) with $p \in (1, \infty)$ and $w \neq 0$. Suppose that $u > f$ (or $u < f$) on an open interval $I \subset \Omega$ then the solution u of (\mathcal{P}) is obtained by*

$$-C(|u'(x)|^{p-2}u'(x))' + u(x) = f(x), \quad \forall x \in I \quad \text{where} \quad C = \frac{\beta}{\|w\|_{L^p(\Omega)}^{p-1}}. \quad (4.40)$$

Proof. Since $1 < p < \infty$, $w \neq 0$ using Proposition 4.4.7 and the second optimality condition of (4.32), we have that

$$\phi = \beta \frac{|u'|^{p-2}u'}{\|w\|_{L^p(\Omega)}^{p-1}}.$$

Hence, by (4.31) we obtain (4.40) where $C = \frac{\beta}{\|w\|_{L^p(\Omega)}^{p-1}}$. \square

One very important property is revealed when $p = \infty$. In fact, it is the first time that first-order total variation based methods can indeed produce piecewise affine solutions.

Proposition 4.4.10. *(Affine structures) If $(w, u) \in L^\infty(\Omega) \times \text{BV}(\Omega)$ a solution pair of (\mathcal{P}) and assume that either $u > f$ or $u < f$ in some open interval $I \subset \Omega$. Then $|w(x)| = \|w\|_\infty$ a.e in I and u is a piecewise affine function.*

Proof. Let $U \subset I$ of positive measure such that $|w(x)| < \|w\|_\infty$ for every $x \in U$. Then, we have the following:

$$\begin{aligned} \int_{\Omega} \phi(x)w(x) dx &\leq \int_{\Omega} |\phi(x)||w(x)| dx = \int_{\Omega \setminus U} |\phi(x)||w(x)| dx + \int_U |\phi(x)||w(x)| dx \\ &< \|w\|_\infty \int_{\Omega \setminus U} |\phi(x)| dx + \|w\|_\infty \int_U |\phi(x)| dx = \|w\|_\infty \|\phi\|_{L^1(\Omega)} \leq \beta \|w\|_\infty \end{aligned}$$

Hence, we reach a contradiction by the third optimality condition of (4.32) and $|w(x)| = \|w\|_\infty$ a.e in I . Moreover by Proposition (4.4.7) we obtain that $u'(x) = \pm \|w\|_\infty$, meaning that the slope of the solution is constant and u is a piecewise affine function in I . \square

Let us mention, one particular case where our model coincides with $L^2 - \text{TGV}_{\alpha,\beta}^2$ minimisation problem. This indicates that one can achieve, for a specific class of one dimensional data, the same results with a high-order order model by using only first-order regularisers. Notice that in the one dimensional setting (4.6) coincides with (4.5) and the $\text{BD}(\Omega)$ space is replaced by $\text{BV}(\Omega)$.

Proposition 4.4.11 ($\text{TGV}_{\alpha,\beta}^2 \sim \text{TVL}_{\alpha,\beta}^\infty$). *Let $(u, w) \in \text{BV}(\Omega) \times \text{BV}(\Omega)$ is a solution pair of (4.2) with $\Psi(u)$ is given by (4.6) and has the properties that u' is monotone, odd and $u \neq f$ a.e. Then, u is a solution of (\mathcal{P}) with $\tilde{\beta} = 2\beta$.*

Proof. Since $Du = D^s u + D^a u$, u' is monotone, odd function then by Proposition 4.4.7 $\|Dw\|_{\mathcal{M}} = 2\|w\|_\infty$, with $u' = w$. Let $\mathcal{A} \subset \text{BV}(\Omega)$ the set of functions with the previous property, then using Remark 2.1.5

$$\begin{aligned} u \in \underset{u \in \mathcal{A}}{\text{argmin}} \left\{ \alpha \|Du - w\|_{\mathcal{M}} + \beta \|Dw\|_{\mathcal{M}} + \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 \right\} &\Leftrightarrow \\ u \in \underset{u \in \mathcal{A}}{\text{argmin}} \left\{ \alpha \|D^s u\|_{\mathcal{M}} + \beta \|Du'\|_{\mathcal{M}} + \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 \right\} &\Leftrightarrow \\ u \in \underset{u \in \mathcal{A}}{\text{argmin}} \left\{ \alpha \|D^s u\|_{\mathcal{M}} + 2\beta \|u'\|_\infty + \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 \right\} &\Leftrightarrow \\ u \in \underset{u \in \mathcal{A}}{\text{argmin}} \left\{ \alpha \|Du - w\|_{\mathcal{M}} + 2\beta \|w\|_\infty + \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 \right\} &\end{aligned}$$

i.e., u is a minimiser of (\mathcal{P}) with $\tilde{\beta} = 2\beta$. Equivalently, the one dimensional functionals $\text{TGV}_{\alpha,\beta}^2$ and $\text{TVL}_{\alpha,2\beta}^\infty$ coincide. \square

4.4.3 Exact solutions

In what follows, we compute explicit solutions of the (\mathcal{P}) model. We examine representative one dimensional data such as the step function and a piecewise affine function with one discontinuity. The step function in $\Omega = (-L, L)$, $L > 0$ is defined as:

$$f(x) = \begin{cases} 0, & \text{if } x \in (-L, 0], \\ h, & \text{if } x \in (0, L). \end{cases} \quad (4.41)$$

It reasonable to start with the ROF-type solutions that are obtained for $1 < p \leq \infty$ when $w = 0$.

ROF-type solutions (step function)

We are initially interested in a solution that respects the discontinuity at $x = 0$ and is piecewise constant. From the optimality conditions (4.31)–(4.32), it suffices to find a function $\phi \in H_0^1(\Omega)$ such that

$$\phi(-L) = \phi(L) = 0, \quad \|\phi\|_\infty \leq \alpha, \quad \phi(0) = \alpha, \quad (4.42)$$

which is piecewise affine. It is easy to see that by setting $\phi(x) = \frac{\alpha}{L}(L - |x|)$, the conditions (4.42) are satisfied and the solution u is piecewise constant. The first condition of (4.32) implies that $\|\phi\|_{L^q(\Omega)} \leq \beta \Leftrightarrow \frac{\beta}{\alpha} \geq \left(\frac{2L}{q+1}\right)^{\frac{1}{q}}$ and provides a necessary and sufficient condition that need to be fulfilled in order for u to be piecewise constant, that is to say

$$u(x) = \begin{cases} \frac{\alpha}{L}, & \text{if } x \in (-L, 0], \\ h - \frac{\alpha}{L}, & \text{if } x \in (0, L), \end{cases} \Leftrightarrow \frac{\beta}{\alpha} \geq \left(\frac{2L}{q+1}\right)^{\frac{1}{q}}. \quad (4.43)$$

For the constant solution, i.e., when $u = \tilde{f}$, the mean value of f , we define $\phi(x) = \frac{h}{2}(L - |x|)$ and in that case conditions (4.42) are satisfied with $\|\phi\|_\infty \leq \alpha \Leftrightarrow \alpha \geq \frac{hL}{2}$ and $\|\phi\|_{L^q(\Omega)} \leq \beta \Leftrightarrow \beta \geq \frac{h}{2} \left(\frac{2L^{q+1}}{q+1}\right)^{\frac{1}{q}}$. This implies that

$$u = \tilde{f} = \frac{h}{2} \Leftrightarrow \alpha \geq \frac{hL}{2} \quad \text{and} \quad \beta \geq \frac{h}{2} \left(\frac{2L^{q+1}}{q+1}\right)^{\frac{1}{q}}. \quad (4.44)$$

Using now (4.43)–(4.44) we can draw the exact regions in the quadrant of $\{\alpha > 0, \beta > 0\}$ that correspond to these two types of solutions, see the left graph in Figure 4.5 for the special case $p = 2$. Notice that in these regions $w = 0$ and the estimates are valid for any $p \in (1, \infty]$. The above estimates are sharp compared to those that we obtained in (4.35) and (4.39), where we compute that $\beta \geq (2L)^{\frac{1}{q}}\alpha$ and $\alpha \geq hL$, $\beta \geq h(2L^{q+1})^{\frac{1}{q}}$.

TVL²-type solutions (step function)

For simplicity, we examine here only the case $p = 2$ with $w \neq 0$ in Ω . However, we refer the reader to Section 4.6.2 where we compute numerically solutions for other finite values of p . Using Proposition 4.4.9, we observe that the solution is given by the following second order differential equation:

$$-Cu''(x) + u(x) = f(x), \quad \text{subject to } C = \frac{\beta}{\|w\|_{L^2(\Omega)}}. \quad (4.45)$$

Even though we can tell that the solution of (4.45) has an exponential form, the fact that the constraint on C depends on the solution w , creates a difficult computation to recover u analytically. In order to overcome this obstacle, we consider the one dimensional version of the 2-homogeneous analogue of (\mathcal{P}) that was introduced in Section 4.3:

$$\min_{\substack{u \in \text{BV}(\Omega) \\ w \in L^2(\Omega)}} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta_{2\text{-hom}}}{2} \|w\|_{L^2(\Omega)}^2. \quad (4.46)$$

Similarly to Section 4.4.1, one can derive the optimality conditions for (4.46). A pair $(w, u) \in L^p(\Omega) \times \text{BV}(\Omega)$ is a solution of (4.46) if and only if there exists a function $\phi \in H_0^1(\Omega)$ such that

$$\begin{aligned} \phi' &= u - f, \\ \phi &\in \alpha \text{Sgn}(Du - w), \\ \phi &= \beta_{2\text{-hom}} w. \end{aligned} \quad (4.47)$$

Equivalently, these are the optimality conditions for the one dimensional analogue of the Huber-TV minimisation problem (4.14). Therefore, if we want to recover analytically the solutions of (\mathcal{P}) for $p = 2$ and determine the purple region in Figure 4.5 it suffices to solve the equivalent model (4.46), where we know from Proposition 4.3.1 that $w \neq 0$. We may restrict our computations only on $I = (-L, 0] \subset \Omega$ and due to symmetry the solution in $I = (0, L)$ is given by $u(x) + u(-x) = h$. From the optimality conditions (4.47), we acquire that

$$-u''(x) + ku(x) = 0, \quad \text{where } k^2 = \frac{1}{\beta_{2\text{-hom}}} \quad \text{and } x \in I = (-L, 0] \subset \Omega.$$

Then, clearly $u(x) = c_1 e^{kx} + c_2 e^{-kx}$ with $\phi(x) = \frac{c_1}{k} e^{kx} - \frac{c_2}{k} e^{-kx} + c_3$ for all $x \in (-L, 0]$. Initially, we examine solutions that are continuous and due to symmetry must have the value $\frac{h}{2}$ at the $x = 0$, i.e., $u(0) = \frac{h}{2}$. Since $\phi \in H_0^1(-L, L)$, we have $\phi(-L) = 0$ and also $u'(-L) = 0$ by the last condition of (4.47). Finally, we require that $\phi(0) < \alpha$. After some

computations, we conclude that

$$u(x) = \begin{cases} c_1 e^{kx} + c_2 e^{-kx}, & \text{if } x \in (-L, 0], \\ h - c_1 e^{-kx} - c_2 e^{kx}, & \text{if } x \in (0, L), \end{cases} \Leftrightarrow \frac{\tanh(kL)}{k} < \frac{2\alpha}{h}, \quad (4.48)$$

where $c_1 = c_2 e^{2kL}$, $c_2 = \frac{h}{2(e^{2kL} + 1)}$ and $c_3 = 0$. On the other hand, in order to get solutions that preserve the discontinuity at $x = 0$, we require the following:

$$\begin{aligned} \phi(-L) &= 0, & u'(-L) &= 0, \\ u(0) &< \frac{h}{2}, & \phi(0) &= \alpha. \end{aligned}$$

Then, we find

$$u(x) = \begin{cases} c_1 e^{kx} + c_2 e^{-kx}, & \text{if } x \in (-L, 0], \\ h - c_1 e^{-kx} - c_2 e^{kx}, & \text{if } x \in (0, L), \end{cases} \Leftrightarrow \frac{\tanh(kL)}{k} > \frac{2\alpha}{h}, \quad (4.49)$$

where $c_1 = c_2 e^{2kL}$, $c_2 = \frac{\alpha k}{e^{2kL} - 1}$. Note that the conditions for α and β_{2-hom} in (4.48) and (4.49) are supplementary and thus only these type of solutions can occur, see the quadrant of $\{\alpha > 0, \beta_{2-hom} > 0\}$ as it presented in Figure 4.4. Letting $g(\beta_{2-hom}) = \sqrt{\beta_{2-hom}} \tanh\left(\frac{L}{\sqrt{\beta_{2-hom}}}\right)$, if $g(\beta_{2-hom}) < \frac{2\alpha}{h}$ then the solution is of the form (4.48), see the blue region in Figure 4.4.

On the other hand, in the complementary green region we obtain the solution (4.49). For extreme cases where $\beta_{2-hom} \rightarrow \infty$, i.e., $k \rightarrow 0$ we obtain $\frac{\tanh(kL)}{k} \rightarrow L$, which means that there is an asymptote of g at $\alpha = \frac{hL}{2}$. Although, we know the form of the inverse function of the hyperbolic tangent, we cannot compute analytically the inverse f^{-1} . However, we can obtain an approximation using a Taylor expansion which leads to

$$\sqrt{\beta_{2-hom}} \tanh\left(\frac{L}{\sqrt{\beta_{2-hom}}}\right) = L - \frac{L^3}{3\beta_{2-hom}} + \mathcal{O}\left(\frac{1}{\beta_{2-hom}^2}\right) = \frac{2\alpha}{h} \Leftrightarrow \beta_{2-hom} = \frac{hL^3}{3(hL - 2\alpha)},$$

where $\alpha > 0$ and $\alpha \neq \frac{hL}{2}$. Finally, we need to examine is what kind of solutions are produced in the limiting case $\beta_{2-hom} \rightarrow \infty$. Letting $\beta_{2-hom} \rightarrow \infty$ in (4.48), we have that $c_1, c_2 \rightarrow \frac{h}{4}$ and $u(x) \rightarrow \frac{h}{2}$ for every $x \in \Omega$, which in fact is the mean value solution \tilde{f} . For the discontinuous solutions, we have that $c_1, c_2 \rightarrow \frac{\alpha}{2L}$ and

$$u(x) \rightarrow \begin{cases} \frac{\alpha}{L}, & \text{if } x \in (-L, 0], \\ h - \frac{\alpha}{L}, & \text{if } x \in (0, L), \end{cases}$$

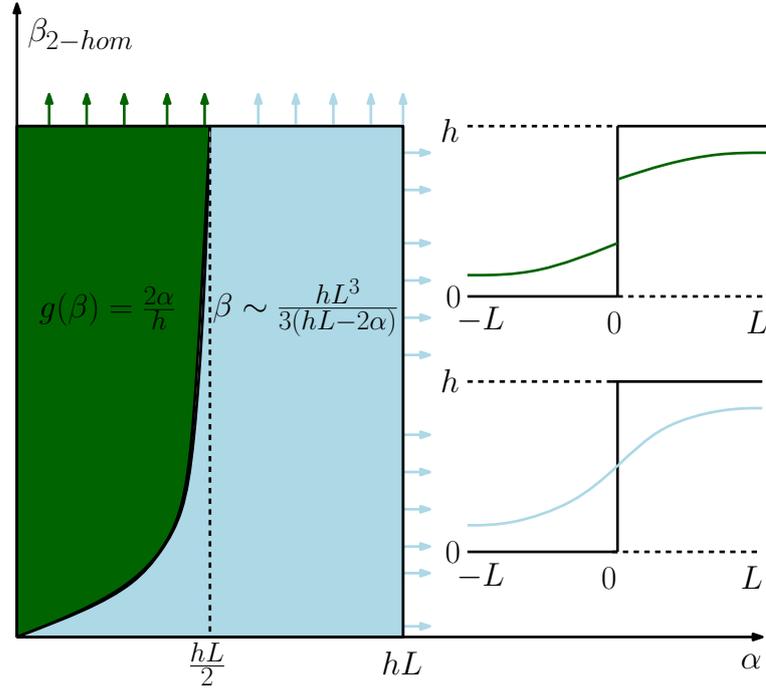


Figure 4.4: Characterisation of solutions of (4.46) for data f being a step function. The green region corresponds to solutions that preserve the discontinuity at $x = 0$, (4.49), while the lightblue region corresponds to continuous solutions, (4.48), both having an exponential form.

i.e., we converge to the solution (4.43). We also have that

$$w(x) = kc_2 \begin{cases} e^{2kL+kx} - e^{-kx}, & \text{if } x \in (-L, 0], \\ e^{2kL-kx} - e^{kx}, & \text{if } x \in (0, L], \end{cases}$$

with $\|w\|_{L^2(\Omega)} = c_2\sqrt{2k}e^{kL}(\sinh(2kL) - 2kL)^{\frac{1}{2}}$ and c_2 is expressed either by (4.48) or (4.49). Then, in both cases we have $w \rightarrow 0$ as $k \rightarrow 0$. Observe that the product of $\beta_{2-hom} \|w\|_{L^2(\Omega)}$ is bounded as $\beta_{2-hom} \rightarrow \infty$ for both types of solutions and in fact corresponds to the bounds found in (4.43) and (4.44) for $p = 2$. Indeed, since

$$\frac{(\sinh(2kL) - 2kL)^{\frac{1}{2}}}{k^{\frac{3}{2}}} \rightarrow 2\sqrt{\frac{L^3}{3}}, \quad \text{as } k \rightarrow 0,$$

then if $\alpha > \frac{hL}{2}$ with $c_2 = \frac{h}{4}$ and $\beta_{2-hom} \rightarrow \infty$

$$\beta_{2-hom} \|w\|_{L^2(\Omega)} = \frac{\sqrt{2}c_2 e^{kL}}{k^{\frac{3}{2}}} (\sinh(2kL) - 2kL)^{\frac{1}{2}} \rightarrow \frac{h}{2} \sqrt{\frac{2L^3}{3}} = \beta_{1-hom},$$

while if $\alpha \leq \frac{hL}{2}$ with $c_2 = \frac{\alpha}{2L}$ and $\beta_{2-hom} \rightarrow \infty$

$$\beta_{2-hom} \|w\|_{L^2(\Omega)} = \frac{\sqrt{2}c_2 e^{kL}}{k^{\frac{3}{2}}} (\sinh(2kL) - 2kL)^{\frac{1}{2}} \rightarrow \alpha \sqrt{\frac{2L}{3}} = \beta_{1-hom}.$$

The last result is yet another verification of Proposition 4.3.2 and it shows that there is an one to one correspondence, namely $\beta_{2-hom} \|w\|_{L^2(\Omega)} \leftrightarrow \beta_{1-hom}$ and the purple region of Figure 4.5 is characterised by the solutions obtained in Figure 4.4. Equivalently, we have derived the exact solutions for the step function under the Huber-TV regularisation.

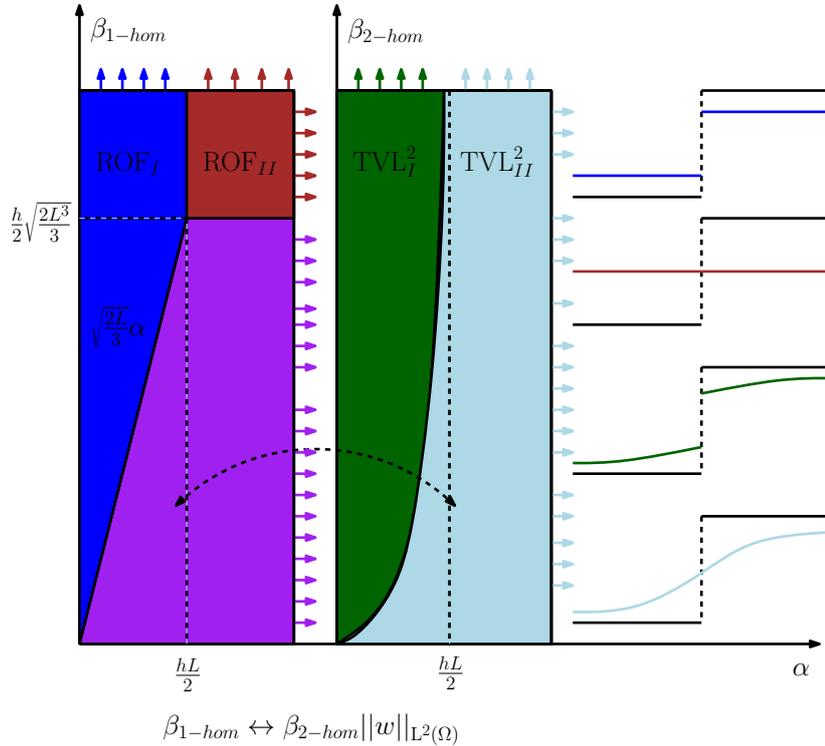


Figure 4.5: Characterisation of solutions of (\mathcal{P}) for $p = 2$ for data f being a step function. The type of solutions in the purple region of the left graph are exactly the solutions obtained for the 2-homogenous problem (4.46), on the right graph.

TVL $^\infty$ type solutions (step function)

Let us now derive the analytical solutions of (\mathcal{P}) for the step function (4.41) when $p = \infty$. The ROF-type solutions (4.43), (4.44) are still valid for $p = \infty$ with $q = 1$. For $w \neq 0$, the analysis presented below justifies Proposition 4.4.10, namely we recover affine solutions.

First, we consider a piecewise affine solution u preserving the discontinuity at $x = 0$ and $u = f$ on $(-L, -l) \cup (l, L)$ i.e., near the boundaries. Indeed, by the optimality condition (4.31), it suffices to find a quadratic continuous function ϕ , increasing (decreasing) on $(-l, 0) = \{x \in (-L, L) : u > f\}$ ($(0, l) = \{x \in (-L, L) : u < f\}$) with $\phi(0) = \alpha$ which van-

ishes near the boundaries. We also use Proposition 4.4.7 in order to define $w \in L^\infty(-L, L)$. Therefore, we define

$$\phi(x) = \begin{cases} 0, & x \in (-L, -l], \\ \frac{\alpha}{l^2}(l - |x|)^2, & x \in (-l, l], \\ 0, & x \in (l, L), \end{cases}, \quad w(x) = \begin{cases} 0, & x \in (-L, -l], \\ \frac{2\alpha}{l^2}, & x \in (-l, l], \\ 0, & x \in (l, L). \end{cases}$$

The optimality condition (4.32) implies that $\langle \phi, w \rangle = \beta \|w\|_\infty \Leftrightarrow l = \frac{3\beta}{2\alpha}$ and since $l < L$ we have that $\frac{\beta}{\alpha} < \frac{2}{3}L$. Finally, in order to preserve the discontinuity, we require that $u(0) < \frac{h}{2} \Leftrightarrow \alpha < \frac{hl}{4} \Leftrightarrow \beta > \frac{8\alpha^2}{3h}$ and conclude that the solution, see orange region in Figure 4.6, is

$$u(x) = \begin{cases} 0, & x \in (-L, -l], \\ \frac{2\alpha}{l^2}(x+l), & x \in (-l, 0], \\ h + \frac{2\alpha}{l^2}(x-l), & x \in (0, l], \\ h, & x \in (l, L], \end{cases} \Leftrightarrow \frac{\beta}{\alpha} < \frac{2}{3}L, \quad \beta > \frac{8\alpha^2}{3h} \quad (\text{TVL}_I^\infty : \text{step})$$

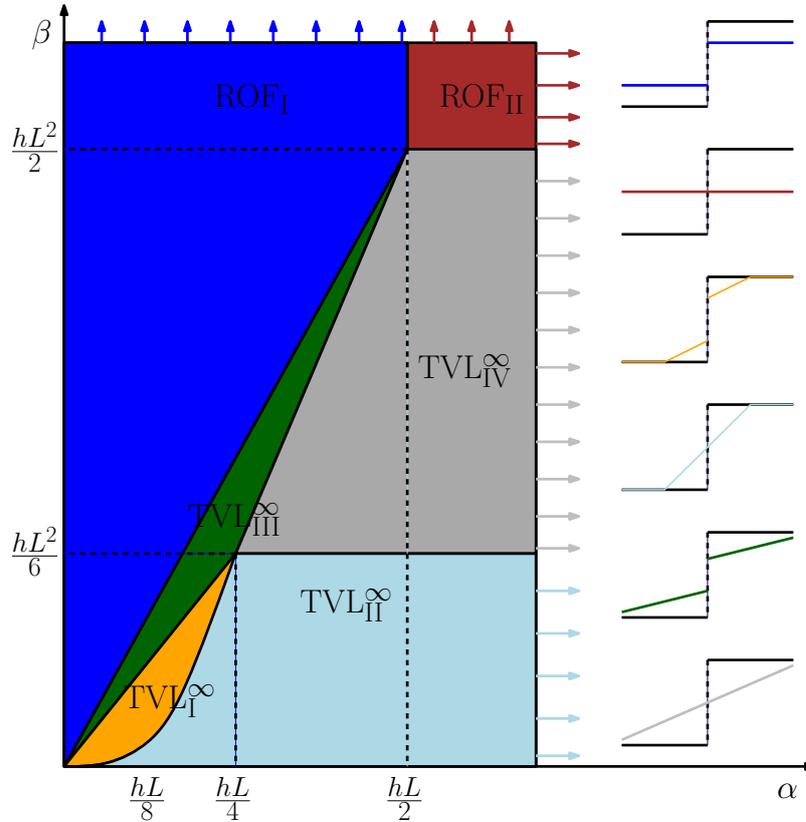


Figure 4.6: Step function: The solutions graph of (P) for TVL^∞ .

On the other hand for a continuous solution u that $u(0) = \frac{h}{2}$ and $u = f$ on $(-L, -k) \cup (k, L)$ i.e., near the boundaries we define,

$$\phi(x) = \begin{cases} 0, & x \in (-L, -k], \\ \frac{h}{4k}(k - |x|)^2, & x \in (-k, k], \\ 0, & x \in (k, L), \end{cases}, \quad w(x) = \begin{cases} 0, & x \in (-L, -k], \\ \frac{h}{2k}, & x \in (-k, k], \\ 0, & x \in (k, L). \end{cases}$$

Also, $\langle \phi, w \rangle = \beta \|w\|_\infty \Leftrightarrow k = \sqrt{\frac{6\beta}{h}}$. Moreover, $k < L \Leftrightarrow \beta < \frac{hL^2}{6}$ and $\phi(0) < \alpha \Leftrightarrow \beta < \frac{8\alpha^2}{3h}$. Hence the solution, see lightblue region in Figure 4.6, is

$$u(x) = \begin{cases} 0, & x \in (-L, -k], \\ \frac{h}{2k}(x + k), & x \in (-k, k], \\ h, & x \in (k, L], \end{cases} \Leftrightarrow \beta < \frac{hL^2}{6}, \quad \beta < \frac{8\alpha^2}{3h}. \quad (\text{TVL}_{\text{II}}^\infty : \text{step})$$

We now examine solutions for which $u = f$ is not allowed and are either preserve the discontinuity at $x = 0$ or they do not. For instance, let u be a piecewise affine function defined on intervals $I_1 = (-L, 0) = \{x \in (-L, L) : u > f\}$ and $I_2 = (0, L) = \{x \in (-L, L) : u < f\}$. Then, the optimality conditions can be expressed as

$$\begin{aligned} \phi(0) &= \alpha, & (\text{preserve discontinuity}) \\ \phi'(0) &< \frac{h}{2}, & (u(0) < \frac{h}{2}) \\ \phi'(-L) &> 0, & (\text{or } \phi'(L) < 0, \text{ since either } u > f \text{ or } u < f) \\ \phi''(-L) &> 0, & (u \text{ is strictly monotone, } w > 0). \end{aligned} \quad (4.50)$$

We define $\phi(x) = c_1x^2 - c_2|x| + \alpha$ and $w(x) = 2c_1$ for $x \in (-L, L)$ and since $\phi \in H_0^1(-L, L)$

$$\left. \begin{aligned} \phi(-L) = 0 &\Leftrightarrow c_1L^2 - c_2L + \alpha = 0 \\ \langle \phi, w \rangle = \beta \|w\|_\infty &\Leftrightarrow 2c_1\frac{L^3}{3} - c_2L^2 + 2\alpha L = \beta \end{aligned} \right\} \Rightarrow \begin{cases} c_1 = \frac{3(\alpha L - \beta)}{L^3}, \\ c_2 = \frac{4\alpha L - 3\beta}{L^2}. \end{cases}$$

Finally, by (4.50) we have the following:

$$\begin{aligned} \phi'(0) < \frac{h}{2} &\Leftrightarrow c_2 < \frac{h}{2} \Leftrightarrow \beta > \frac{4}{3}L\alpha - \frac{hL^2}{6}, \\ \phi'(-L) > 0 &\Leftrightarrow c_2 > 2c_1L \Leftrightarrow \beta > \frac{2L}{3}\alpha, \\ \phi''(-L) > 0 &\Leftrightarrow c_1 > 0 \Leftrightarrow \beta < L\alpha, \end{aligned} \quad (4.51)$$

and the solution, see green region in Figure 4.6, is

$$u(x) = \begin{cases} \frac{6(\alpha L - \beta)}{L^3}x + \frac{4\alpha L - 3\beta}{L^2}, & x \in (-L, 0] \\ h + \frac{6(\alpha L - \beta)}{L^3}x - \frac{4\alpha L - 3\beta}{L^2}, & x \in (0, L) \end{cases} \Leftrightarrow \begin{cases} \beta > \frac{4}{3}L\alpha - \frac{hL^2}{6} \\ \beta > \frac{2L}{3}\alpha, \beta < L\alpha. \end{cases} \quad (\text{TVL}_{\text{III}}^{\infty} : \text{step})$$

Similarly, for a continuous and affine solution u , we define $\phi(x) = c_3x^2 - \frac{h}{2}|x| + c_4$ with $w(x) = 2c_3$ for $x \in (-L, L)$ under the conditions:

$$\begin{aligned} \phi(0) &< \alpha, u \text{ is continuous} \\ \phi'(0) &= \frac{h}{2}, u(0) = \frac{h}{2} \\ \phi'(-L) &> 0, u > f \\ \phi''(-L) &> 0, u \text{ is increasing.} \end{aligned} \quad (4.52)$$

Hence, we compute

$$\left. \begin{aligned} \phi(-L) = 0 &\Leftrightarrow c_3L^2 - \frac{h}{2}L + c_4 = 0 \\ \langle \phi, w \rangle = \beta \|w\|_{\infty} &\Leftrightarrow 2c_3\frac{L^3}{3} - \frac{h}{2}L^2 + 2c_4L = \beta \end{aligned} \right\} \Rightarrow \begin{cases} c_3 = \frac{3(hL^2 - 2\beta)}{8L^3} \\ c_4 = \frac{hL^2 + 6\beta}{8L} \end{cases}$$

and the solution, see grey region in Figure 4.6, is

$$u(x) = \frac{3(hL^2 - 2\beta)}{4L^3}x + \frac{h}{2}, x \in (-L, L) \Leftrightarrow \begin{cases} \beta < \frac{4}{3}L\alpha - \frac{hL^2}{6}, \\ \beta > \frac{hL^2}{6}, \beta < \frac{hL^2}{2}. \end{cases} \quad (\text{TVL}_{\text{IV}}^{\infty} : \text{step})$$

TVL[∞] type solutions (piecewise affine function)

Here, we choose to examine the exact solutions for a piecewise affine function defined as

$$g(x) = \begin{cases} \lambda x, & , x \in (-L, 0], \\ \lambda x + h, & , x \in (0, L). \end{cases} \quad (4.53)$$

We demonstrate the capability of our regulariser to obtain piecewise affine solutions. In the following, we have summarised our computed solutions for every combination of $\alpha, \beta > 0$ using similar analysis with the one described above. We conclude with a total of eight types of solutions, three of them are ROF-type solutions, i.e., $w = 0$ and the rest are the corresponding TVL[∞] solutions, see also Figure 4.7. All the analytical computations regarding the solutions below can be found in Appendix A.

We begin with the ROF-type solutions. We consider a solution that is piecewise

constant near the boundaries and $u = f$ in $I_1 = (-l_1, 0) \cup (0, l_1)$, where $l_1 > 0$,

$$u(x) = \begin{cases} -\frac{\alpha}{l_1-L} - \frac{\lambda}{2}(l_1 + L), & x \in (-L, l_1), & 0 \leq \alpha \leq \frac{\lambda L^2}{2}, \\ \lambda x, & x \in (l_1, 0), \\ \lambda x + h, & x \in (0, l_1), \\ h + \frac{\alpha}{l_1-L} + \frac{\lambda}{2}(l_1 + L), & x \in (l_1, L), & l_1 = L - \sqrt{\frac{2\alpha}{\lambda}}. \end{cases} \Leftrightarrow \beta \geq 2\alpha L - \frac{2\alpha}{3}\sqrt{\frac{2\alpha}{\lambda}},$$

(ROF_I : affine)

We proceed with a piecewise constant solution where $u \neq f$ for all $x \in (-L, L)$,

$$u(x) = \begin{cases} \frac{\alpha}{L} - \frac{\lambda L}{2}, & x \in (-L, 0), \\ h - \frac{\alpha}{L} + \frac{\lambda L}{2}, & x \in (0, L), \end{cases} \Leftrightarrow \begin{cases} \frac{hL}{2} \leq \alpha < \frac{hL + \lambda L^2}{2}, \\ \beta \geq \alpha L + \frac{\lambda L^3}{6}. \end{cases} \quad (\text{ROF}_{\text{II}} : \text{affine})$$

Finally, we obtain the mean value solution $u = \tilde{f} = \frac{1}{2L} \int_{-L}^L g(x) dx$:

$$u(x) = \frac{h}{2}, \quad x \in (-L, L) \Leftrightarrow \alpha > \frac{hL + \lambda L^2}{2}, \quad \beta \geq \frac{hL^2}{2} + \frac{2\lambda L^3}{3}. \quad (\text{ROF}_{\text{III}} : \text{affine})$$

For $w \neq 0$, we have the following analytic solutions:

$$u(x) = \begin{cases} \lambda x, & x \in (-L, l_2), \\ \left(\frac{2\alpha}{l_2^2} + \lambda\right)x + \frac{2\alpha}{l_2}, & x \in (-l_2, 0), \\ \left(\frac{2\alpha}{l_2^2} + \lambda\right)x - \frac{2\alpha}{l_2} + h, & x \in (0, l_2), \\ \lambda x + h, & x \in (l_2, L), \end{cases} \Leftrightarrow \begin{cases} \beta > \frac{8\alpha^2}{3h}, \beta \leq \frac{2\alpha}{3}L, \\ l_2 = \frac{3\beta}{2\alpha}. \end{cases}$$

(TVL_I[∞] : affine)

$$u(x) = \begin{cases} \lambda x, & x \in (-L, l_3), \\ \left(\frac{h}{2l_3} + \lambda\right)x + \frac{h}{2}, & x \in (-l_3, l_3), \\ \lambda x + h, & x \in (l_3, L), \end{cases} \Leftrightarrow \begin{cases} \beta \leq \frac{hL^2}{6}, \beta < \frac{8\alpha^2}{3h}, \\ l_3 = \sqrt{\frac{6\beta}{h}}. \end{cases} \quad (\text{TVL}_{\text{II}}^{\infty} : \text{affine})$$

$$u(x) = \begin{cases} \left(\frac{6(\alpha L - \beta)}{L^3} + \lambda\right)x + \frac{4\alpha L - 3\beta}{L^2}, & x \in (-L, 0), \\ \left(\frac{6(\alpha L - \beta)}{L^3} + \lambda\right)x + h - \frac{4\alpha L - 3\beta}{L^2}, & x \in (0, L), \end{cases} \Leftrightarrow \begin{cases} \beta > \frac{4\alpha}{3}L - \frac{hL^2}{6}, \\ \beta > \frac{2}{3}\alpha L, \beta < \frac{4}{3}\alpha L, \\ \beta < \alpha L + \frac{\lambda L^3}{6}. \end{cases}$$

(TVL_{III}[∞] : affine)

$$u(x) = \begin{cases} \left(-\frac{2\alpha}{(L-l_4)^2} + \lambda\right)x - \frac{2\alpha l_4}{(L-l_4)^2}, & x \in (-L, -l_4), & \beta < 2\alpha L - \frac{2\alpha}{3}\sqrt{\frac{2\alpha}{\lambda}}, \\ \lambda x, & x \in (-l_4, 0), & \\ \lambda x + h, & x \in (0, l_4), & \Leftrightarrow \beta \geq \frac{4\alpha}{3}L, \\ \left(-\frac{2\alpha}{(L-l_4)^2} + \lambda\right)x + \frac{2\alpha l_4}{(L-l_4)^2} + h, & x \in (l_4, 0), & l_4 = \frac{3\beta}{2\alpha} - 2L. \end{cases} \quad (\text{TVL}_{\text{IV}}^\infty : \text{affine})$$

$$u(x) = \left(\frac{3(\frac{hL^2}{2} - \beta)}{2L^3} + \lambda\right)x + \frac{h}{2} \Leftrightarrow \begin{cases} \beta < \frac{4\alpha}{3}L - \frac{hL^2}{6}, & \beta > \frac{hL^2}{6} \\ \beta < \frac{hL^2}{2} + \frac{2\lambda L^3}{3}. \end{cases} \quad (\text{TVL}_{\text{V}}^\infty : \text{affine})$$

For demonstration reasons, the Figure 4.7 is formed when $\lambda = \frac{h}{L}$ and the colored regions are defined appropriately.

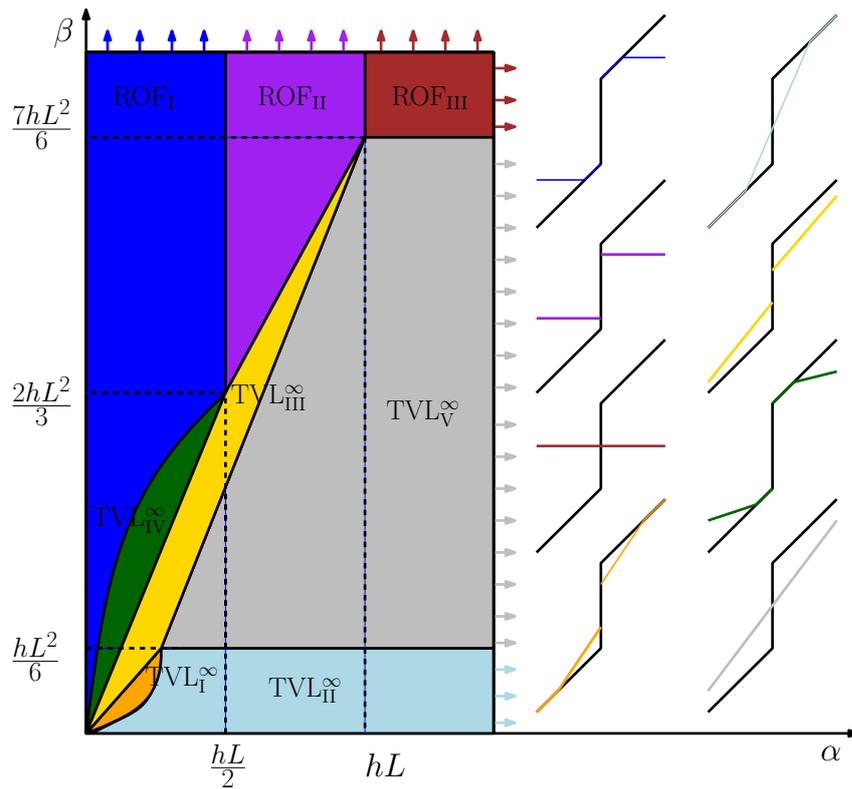


Figure 4.7: Piecewise affine function: The solutions graph of (\mathcal{P}) for TVL^∞ with $\lambda = \frac{h}{L}$.

4.5 An image decomposition approach

Here, we present another formulation for the problem (\mathcal{P}) by decomposing an image into a BV part (piecewise constant) and a part that belongs to $W^{1,p}(\Omega)$ (smooth). We also provide a characterisation of minimisers based on the analysis of Yves Meyer for the (ROF) model, cf. [Mey01]. Let $1 < p \leq \infty$ and $\Omega \subset \mathbb{R}^d$ and consider the following minimisation problem:

$$\min_{\substack{u \in \text{BV}(\Omega) \\ v \in W^{1,p}(\Omega)}} L(u, v) := \frac{1}{2} \|f - u - v\|_{L^2(\Omega)}^2 + \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)}. \quad (4.54)$$

This is in fact an infimal convolution operation between the total variation seminorm and the L^p norm of the gradient. In this way, we can decompose our image into two geometric components. The second term captures the piecewise constant structures in the image, as in the (ROF), whereas the third term captures the smoothness that depends on the value of p . In the one dimensional setting, we can prove that the problems (\mathcal{P}) and (4.54) are equivalent.

Proposition 4.5.1. *Let $\Omega = (a, b) \subset \mathbb{R}$, then a pair $(v^*, u^*) \in W^{1,p}(\Omega) \times \text{BV}(\Omega)$ is a solution of (4.54) if and only if $(\nabla v^*, u^* + v^*) \in L^p(\Omega) \times \text{BV}(\Omega)$ is a solution of (\mathcal{P}) .*

Proof. Let $\bar{u} = u + v$ then, we have the following

$$\begin{aligned} (v^*, u^*) &\in \underset{\substack{u \in \text{BV}(\Omega) \\ v \in W^{1,p}(\Omega)}}{\text{argmin}} \frac{1}{2} \|f - u - v\|_{L^2(\Omega)}^2 + \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} \Leftrightarrow \\ (v^*, u^*) &\in \underset{\substack{u \in \text{BV}(\Omega) \\ v \in W^{1,p}(\Omega)}}{\text{argmin}} \frac{1}{2} \|f - u - v\|_{L^2(\Omega)}^2 + \alpha \sup_{\substack{\phi \in C_c^\infty(\Omega) \\ \|\phi\|_\infty \leq 1}} \{\langle u, \phi' \rangle\} + \beta \|\nabla v\|_{L^p(\Omega)} \Leftrightarrow \\ (v^*, \bar{u}^*) &\in \underset{\substack{\bar{u} \in \text{BV}(\Omega) \\ v \in W^{1,p}(\Omega)}}{\text{argmin}} \frac{1}{2} \|f - \bar{u}\|_{L^2(\Omega)}^2 + \alpha \sup_{\substack{\phi \in C_c^\infty(\Omega) \\ \|\phi\|_\infty \leq 1}} \{\langle \bar{u}, \phi' \rangle + \langle \nabla v, \phi \rangle\} + \beta \|\nabla v\|_{L^p(\Omega)} \Leftrightarrow \\ (w^*, \bar{u}^*) &\in \underset{\substack{\bar{u} \in \text{BV}(\Omega) \\ w = \nabla v \\ v \in W^{1,p}(\Omega)}}{\text{argmin}} \frac{1}{2} \|f - \bar{u}\|_{L^2(\Omega)}^2 + \alpha \|D\bar{u} - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)}. \end{aligned}$$

However, we can eliminate the last constraint since

$$\{w \in L^p(\Omega) : \exists v \in W^{1,p}(\Omega), w = \nabla v\} = L^p(\Omega). \quad (4.55)$$

Indeed, let $w \in L^p(\Omega) \subset L^1(\Omega)$ for $p \in (1, \infty]$ and define $v(x) = \int_a^x w(s) ds$ for $x \in \Omega \subset \mathbb{R}$. Clearly, $v' = w$ a.e and by Hölder's inequality

$$|v(x)|^p = \left| \int_a^x w(s) ds \right|^p \leq (x-a)^{p-1} \int_a^x |w(s)|^p ds < C < \infty$$

and $v \in W^{1,p}(\Omega)$ for $p \in (1, \infty)$. Finally, for the case $p = \infty$, let $C > 0$ be a constant such that $|w(x)| \leq C$ a.e. on Ω . In that case we have $|v(x)| \leq \int_a^x |w(s)| ds \leq C|\Omega| < \infty$, i.e., $v \in L^\infty(\Omega)$ and hence $v \in W^{1,\infty}(\Omega)$ since $v' = w$, see also [Bre11, Chapter 8]. Therefore,

$$(w^*, \bar{u}^*) \in \operatorname{argmin}_{\substack{\bar{u} \in \operatorname{BV}(\Omega) \\ w \in L^p(\Omega)}} \frac{1}{2} \|f - \bar{u}\|_{L^2(\Omega)}^2 + \alpha \|D\bar{u} - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)},$$

where $\bar{u}^* = u^* + v^*$ and $w^* = \nabla v^*$. \square

Even though for $d = 1$ it is true that every L^p function can be written as a gradient, this is not true for higher dimensions. In fact, as we show in the following sections, the constraint (4.55) is quite restrictive and for example the staircasing effect cannot always be eliminated in the denoising process, see for instance Figure 4.25. The existence of minimisers of (4.54) is shown following again the same techniques as in Proposition 4.2.1. Moreover, one can immediately show if $(v^*, u^*) \in W^{1,p}(\Omega) \times \operatorname{BV}(\Omega)$ is a minimiser of (4.54) then

$$u^* = \operatorname{argmin}_{u \in \operatorname{BV}(\Omega)} \left\{ \frac{1}{2} \|f - u - v^*\|_{L^2(\Omega)}^2 + \alpha \|Du\|_{\mathcal{M}} \right\} \quad (4.56)$$

$$v^* = \operatorname{argmin}_{v \in W^{1,p}(\Omega)} \left\{ \frac{1}{2} \|f - u^* - v\|_{L^2(\Omega)}^2 + \beta \|\nabla v\|_{L^p(\Omega)} \right\}. \quad (4.57)$$

Hence by the optimality conditions of (4.56) and (4.57), we obtain that:

$$0 \in \alpha \partial \|\cdot\|_{\mathcal{M}}(u^*) + u^* + v^* - f, \quad (4.58)$$

$$0 \in \beta \partial \|\cdot\|_{L^p(\Omega)}(v^*) + u^* + v^* - f. \quad (4.59)$$

Using (4.59) and the characterisation of the subdifferential of total variation, see Section 2.4,

$$\partial \|\cdot\|_{\mathcal{M}}(u^*) = \left\{ \operatorname{div} \phi : \|\phi\|_{L^\infty(\Omega)} \leq 1, \phi \cdot n|_{\partial\Omega} = 0, \langle \operatorname{div} \phi, u \rangle_{L^2(\Omega)} = \|u^*\|_{\mathcal{M}} \right\}, \quad (4.60)$$

see also [CCC⁺10] and [BO13], we derive that

$$\int_{\Omega} f = \int_{\Omega} u^* + v^*. \quad (4.61)$$

Let $(u_1, v_1), (u_2, v_2) \in \operatorname{BV}(\Omega) \times W^{1,p}(\Omega)$ be two minimisers of (4.54). Since the energy functional is convex one has that for every $\lambda \in [0, 1]$

$$L(\lambda(u_1, v_1) + (1 - \lambda)(u_2, v_2)) \leq \lambda L(u_1, v_1) + (1 - \lambda)L(u_2, v_2),$$

which is in fact an equality i.e.,

$$L(\lambda(u_1, v_1) + (1 - \lambda)(u_2, v_2)) = \lambda L(u_1, v_1) + (1 - \lambda)L(u_2, v_2). \quad (4.62)$$

Moreover, if $u_1 + v_1 \neq u_2 + v_2$ due to the strict convexity of $\frac{1}{2} \|f - \cdot\|_{L^2(\Omega)}^2$ we reach a contradiction of (4.62), hence the sum is unique. In addition to this, we also have

$$\begin{aligned} & \alpha \|D(\lambda u_1 + (1 - \lambda)u_2)\|_{\mathcal{M}} + \beta \|\nabla(\lambda v_1 + (1 - \lambda)v_2)\|_{L^p(\Omega)} \\ &= \alpha(\lambda \|Du_1\|_{\mathcal{M}} + (1 - \lambda) \|Du_2\|_{\mathcal{M}}) + \beta(\lambda \|\nabla v_1\|_{L^p(\Omega)} + (1 - \lambda) \|\nabla v_2\|_{L^p(\Omega)}). \end{aligned} \quad (4.63)$$

If we assume that

$$\|\nabla(\lambda v_1 + (1 - \lambda)v_2)\|_{L^p(\Omega)} < \lambda \|\nabla v_1\|_{L^p(\Omega)} + (1 - \lambda) \|\nabla v_2\|_{L^p(\Omega)},$$

then we contradict the equality on (4.63). Hence, the Minkowski inequality becomes an equality which is equivalent to the existence of $\mu > 0$ such that $\nabla v_2 = \mu \nabla v_1$. In other words, we have proved the following proposition that was also shown in [KV09] in a similar context:

Proposition 4.5.2. *Let $(u_1, v_1), (u_2, v_2)$ be two minimisers of (4.54). Then*

$$u_1 + v_1 = u_2 + v_2, \quad (4.64)$$

$$\exists \mu > 0 \text{ such that } \nabla v_2 = \mu \nabla v_1. \quad (4.65)$$

4.5.1 Characterisation of minimisers

In this section, we provide a characterisation of minimisers for (4.54). Namely, we would like to identify the cases when one obtains $(u, 0)$, $(0, v)$ as minimisers and certainly when an optimal decomposition occur, i.e, u represents the piecewise constant component of the image and v represents a component whose structure depends on p . Based on Meyer analysis, who introduced the so-called *texture* norm in order to provide the decomposition for the (ROF) model, we define a similar *texture norm* for our proposed model (4.54) and provide necessary and sufficient conditions for *certain* types of solutions depending on the texture norm and the regularising parameters α, β . Related characterisations of minimisers can be found in [KV09], [LV08], [LV05], [OSV03].

In [Mey01, Theorem 3], Meyer proved the following theorem:

Theorem 4.5.3. *Let $f \in L^2(\mathbb{R}^2)$ and $(u, v) \in \text{BV}(\mathbb{R}^2) \times L^2(\mathbb{R}^2)$ then:*

- (i) *If $\|f\|_* \leq \alpha$, then the image f is seen as a texture and $f = 0 + f$ in the (ROF) decomposition.*

(ii) If $\|f\|_* > \alpha$, then the (ROF) decomposition $f = u+v$ is characterised by the following conditions

$$\|v\|_* = \alpha \text{ and } \int_{\Omega} u(x)v(x)dx = \alpha \|Du\|_{\mathcal{M}}.$$

Let us recall the definition $\|\cdot\|_*$ and the corresponding $G(\mathbb{R}^2)$ space. For further details, the reader is referred to [Mey01] and [Had05].

Definition 4.5.4. The space $G(\mathbb{R}^2)$ is a Banach space of distributions f which can be decomposed as

$$f(x) = \partial g_1(x) + \partial g_2(x)$$

with $g_1, g_2 \in L^\infty(\mathbb{R}^2)$. It is endowed with the following norm,

$$\|f\|_* := \inf \left\{ \| |g(x)| \|_{L^\infty(\mathbb{R}^2)} \mid \nabla \cdot g = f \right\}$$

where $g = (g_1, g_2)$ and $|g(x)| = \sqrt{g_1(x)^2 + g_2(x)^2}$.

Roughly speaking, the space $G(\mathbb{R}^2)$ could be seen as the dual space of $BV(\mathbb{R}^2)$. In fact, $G(\mathbb{R}^2)$ is the dual space \mathcal{BV}^* of \mathcal{BV} where $\mathcal{BV} = \overline{BV}^{\mathcal{S}(\mathbb{R}^2)}$ i.e, the closure of BV with the Schwartz norm. Hence, by a duality argument we observe that if $f \in \mathcal{BV}(\mathbb{R}^2)$ and $g \in L^2(\mathbb{R}^2)$ then,

$$\|g\|_* = \sup_{\substack{f \in \mathcal{BV} \\ \|Df\|_{\mathcal{M}} \leq 1}} \left\{ \int_{\mathbb{R}^2} f(x)g(x)dx \right\} \quad (4.66)$$

and the following inequality is valid:

$$\int_{\mathbb{R}^2} f(x)g(x)dx \leq \|Df\|_{\mathcal{M}} \|g\|_*, \forall f \in \mathcal{BV}, \forall g \in L^2(\mathbb{R}^2) \quad (4.67)$$

The above relation (4.67) can be generalised for functions $f \in BV(\mathbb{R}^2)$. The proof is straightforward using approximations with mollifiers, see [Mey01, Lemma 3]. Notice, that the Definition 4.5.4 applies for cases where $\Omega = \mathbb{R}^2$. Later, it was generalised also to $\Omega \subset \mathbb{R}^2$ being a bounded domain. For unbounded domains a constant non-zero function has unbounded G-norm and a normalisation condition is necessary as

$$\int_{\Omega} f(x)dx = 0. \quad (4.68)$$

In fact the authors in [AA05, Proposition 2.1] prove that

$$G(\Omega) = \left\{ f \in L^2(\Omega) : \int_{\Omega} f(x) dx = 0 \right\}$$

and the G-norm becomes

$$\|f\|_* := \inf \left\{ \|g(x)\|_{L^\infty(\Omega)} \mid \nabla \cdot g = f, g \cdot \vec{n} = 0 \text{ on } \partial\Omega \right\} \quad (4.69)$$

where \vec{n} is the unit normal vector on the boundary.

We are now ready to begin with the characterisation of minimisers for trivial cases, solutions which one of the components are zero, i.e., $(u^*, 0)$ or $(0, v^*)$. We observe from equations (4.56), (4.57) that if $(u^*, 0)$ or $(0, v^*)$ are minimisers of (4.54), then either u^* behaves as a ROF-type solution or v^* is a $L^2 - W^{1,p}$ solution:

$$v^* = \operatorname{argmin}_{v \in W^{1,p}(\Omega)} \left\{ \frac{1}{2} \|f - v\|_{L^2(\Omega)}^2 + \beta \|\nabla v\|_{L^p(\Omega)} \right\}.$$

Proposition 4.5.5. (i) If $\|f - u\|_* \leq \alpha$ and $\frac{\beta}{\alpha} \geq |\Omega|^{1/q}$, then $(u, 0)$ is a minimiser.

(ii) If $\|f - v\|_* \leq \beta |\Omega|^{\frac{1}{q}} \leq \alpha$, then $(0, v)$ is a minimiser.

Proof. (i) Since $\|f - u\|_* \leq \alpha$, then for every $h \in \operatorname{BV}(\Omega)$, we have that

$$\int_{\Omega} (f - u)h \, dx \leq \alpha \|Dh\|_{\mathcal{M}}.$$

The latter is also true for every $h \in W^{1,1}(\Omega) \subset \operatorname{BV}(\Omega)$. Hence,

$$\int_{\Omega} (f - u)h \, dx \leq \alpha \|\nabla h\|_{L^1(\Omega)}$$

and since $\Omega \subset \mathbb{R}^2$ bounded, we conclude that

$$\int_{\Omega} (f - u)h \, dx \leq \alpha \|\nabla h\|_1 \leq \alpha |\Omega|^{\frac{1}{q}} \|\nabla h\|_p \leq \beta \|\nabla v\|_p, \quad \forall h \in W^{1,p}(\Omega) \subset W^{1,1}(\Omega).$$

Finally, for every $h \in W^{1,p}(\Omega)$ one has that

$$\begin{aligned} \alpha \|Du\|_{\mathcal{M}} + \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 &\leq \alpha \|Du\|_{\mathcal{M}}(\Omega) + \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \frac{1}{2} \|h\|_{L^2(\Omega)}^2 - \langle f - u, h \rangle + \langle f - u, h \rangle \\ &= \alpha \|Du\|_{\mathcal{M}} + \frac{1}{2} \|f - u - h\|_{L^2(\Omega)}^2 + \langle f - u, h \rangle \\ &\leq \alpha \|Du\|_{\mathcal{M}} + \frac{1}{2} \|f - u - h\|_{L^2(\Omega)}^2 + \beta \|\nabla h\|_{L^p(\Omega)} \end{aligned}$$

i.e., $(u, 0)$ is a minimiser for (4.54).

(ii) Similarly, for every $h \in \operatorname{BV}(\Omega)$ we have that

$$\int_{\Omega} (f - v)h \, dx \leq \beta |\Omega|^{\frac{1}{q}} \|Dh\|_{\mathcal{M}} \leq \alpha \|Dh\|_{\mathcal{M}}.$$

Hence,

$$\begin{aligned}
 \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|f - v\|_{L^2(\Omega)}^2 &\leq \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|f - v\|_{L^2(\Omega)}^2 + \frac{1}{2} \|h\|_{L^2(\Omega)}^2 \\
 &\quad - \langle f - v, h \rangle + \langle f - v, h \rangle \\
 &= \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|f - v - h\|_{L^2(\Omega)}^2 + \langle f - v, h \rangle \\
 &\leq \beta \|\nabla v\|_{L^p(\Omega)} + \alpha \|Dh\|_{\mathcal{M}} + \frac{1}{2} \|f - v - h\|_{L^2(\Omega)}^2,
 \end{aligned}$$

i.e., $(0, v)$ is a minimiser for (4.54). □

The previous proposition implies that for $\frac{\beta}{\alpha} \geq |\Omega|^{\frac{1}{q}}$, we have the decomposition $f = u + 0 + (f - u)$ where $u \in \text{BV}(\Omega)$ and the residual part $r := f - u \in L^2(\Omega)$ is controlled by α . On the other hand, if $\frac{\beta}{\alpha} \leq |\Omega|^{-\frac{1}{q}}$, we have that $f = 0 + v + (f - v)$ where $v \in W^{1,p}(\Omega)$ for $p \in (1, \infty]$ and the residual part $r := f - v \in L^2(\Omega)$ is controlled by β . We observe that the first condition for the (4.54) model coincides with (4.35) for the (\mathcal{P}) , meaning that the u component behaves as an ROF solution whereas the v component is 0. For the other inequality, where $\beta \leq \alpha|\Omega|^{\frac{1}{q}}$ the u component is 0 and the v component is a solution of $L^2 - W^{1,p}$ problem, which in fact is continuous. This is also verified by the fact that in the right corner regions (lightblue regions) of the quadrants in Figures 4.4, 4.6 and 4.7 the corresponding solutions are indeed continuous. Compared to the analysis performed in Section 4.4.2 in the one dimensional setting, the estimates derived above provide us with a quantitative analysis also in the two dimensional setting.

Our next step is to examine the cases where non-trivial solutions are obtained. For this task, we need to define a similar *texture* norm as the $\|\cdot\|_*$ norm, dual to $\text{BV} + W^{1,p}$.

Definition 4.5.6. Let $g \in L^2(\Omega)$ and $\alpha, \beta > 0$ and $1 < p \leq \infty$. We define

$$\|g\|_{(\alpha, \beta)} = \sup_{\mathcal{G}} \frac{|\langle g, u + v \rangle|}{\alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)}} \quad (4.70)$$

with $\langle \cdot, \cdot \rangle$ be the L^2 inner product and

$$\mathcal{G} = \{(u, v) \in \text{BV} \times W^{1,p} : \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} \neq 0.\} \quad (4.71)$$

The following useful lemma is also proved in [LV08] in a similar context.

Lemma 4.5.7. If $f \in L^2(\Omega)$ with $\|f\|_{(\alpha, \beta)} < \infty$ then $\int_{\Omega} f = 0$.

Proof. Let $(u, v) \in \mathcal{G}$. Then, for any constant $\lambda \in \mathbb{R} \setminus \{0\}$, $(u, v + \lambda) \in \mathcal{G}$ and

$$\frac{|\langle f, u + v + \lambda \rangle|}{\alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla(v + \lambda)\|_{L^p(\Omega)}} = \frac{|\langle f, u + v \rangle + \langle f, \lambda \rangle|}{\alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)}} \leq \|f\|_{(\alpha, \beta)} < \infty.$$

The last condition implies that there exists a constant $\Lambda > 0$ such that $|\langle f, \lambda \rangle| < \Lambda < \infty$ for all $\lambda \in \mathbb{R} \setminus \{0\}$ and therefore $|\langle f, \mathcal{X}_\Omega \rangle|$ must be zero i.e., $\int_\Omega f = 0$ \square

Now, we are ready to provide a characterisation of minimisers via the *texture* norm defined above.

Theorem 4.5.8. *Let $(u, v) \in \mathcal{G}$ be an optimal decomposition $BV \times W^{1,p}$ of (4.54) with $f \in L^2(\Omega)$ and the residual part is $r := f - u - v \in L^2(\Omega)$. Then,*

$$\|f\|_{(\alpha,\beta)} \leq 1 \Leftrightarrow u = v = 0, r = f. \quad (4.72)$$

If $\|f\|_{(\alpha,\beta)} > 1$, then $(u, v) \in \mathcal{G}$ is characterised by

$$\|r\|_{(\alpha,\beta)} = 1 \text{ and } |\langle f, u + v \rangle| = \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)}. \quad (4.73)$$

Proof. Suppose $(0, 0)$ is a minimiser, then for every pair $(u, v) \in BV \times W^{1,p}$ we have

$$\begin{aligned} \frac{1}{2} \|f\|_{L^2(\Omega)}^2 &\leq \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|f - u - v\|_{L^2(\Omega)}^2 \\ \langle f, u + v \rangle &\leq \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|u + v\|_{L^2(\Omega)}^2. \end{aligned}$$

Replacing u by εu and v by εv with $\varepsilon \rightarrow 0^\pm$, we obtain

$$|\langle f, u + v \rangle| \leq \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} \Rightarrow \|f\|_{(\alpha,\beta)} \leq 1.$$

For the converse, we have that $|\langle f, u + v \rangle| \leq \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)}$ for $(u, v) \in \mathcal{G}$ and

$$\begin{aligned} \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|f - u - v\|_{L^2(\Omega)}^2 &= \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|f\|_{L^2(\Omega)}^2 \\ &\quad + \frac{1}{2} \|u + v\|_{L^2(\Omega)}^2 - \langle f, u + v \rangle \\ &\geq \frac{1}{2} \|f\|_{L^2(\Omega)}^2 + \frac{1}{2} \|u + v\|_{L^2(\Omega)}^2 \geq \frac{1}{2} \|f\|_{L^2(\Omega)}^2. \end{aligned}$$

Hence, the optimal decomposition in this case is $u = v = 0$ with $r = f$. For the condition (4.73), let $(h_1, h_2) \in BV \times W^{1,p}$ and $\varepsilon \in \mathbb{R}$ then,

$$\alpha \|D(u + \varepsilon h_1)\|_{\mathcal{M}} + \beta \|\nabla(v + \varepsilon h_2)\|_{L^p(\Omega)} + \frac{1}{2} \|r - \varepsilon(h_1 + h_2)\|_{L^2(\Omega)}^2 \geq \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|r\|_{L^2(\Omega)}^2$$

and can be simplified by the triangle inequality on the left hand side as

$$\begin{aligned} \alpha \|Du\|_{\mathcal{M}} + \alpha|\varepsilon| \|Dh_1\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} + \beta|\varepsilon| \|\nabla h_2\|_{L^p(\Omega)} + \frac{1}{2} \|r - \varepsilon(h_1 + h_2)\|_{L^2(\Omega)}^2 &\geq \\ \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|r\|_{L^2(\Omega)}^2 &\Leftrightarrow \end{aligned}$$

$$\begin{aligned} \alpha|\varepsilon| \|Dh_1\|_{\mathcal{M}} + \beta|\varepsilon| \|\nabla h_2\|_{L^p(\Omega)} + \frac{1}{2} \|r - \varepsilon(h_1 + h_2)\|_2^2 &\geq \frac{1}{2} \|r\|_{L^2(\Omega)}^2 \Leftrightarrow \\ \alpha|\varepsilon| \|Dh_1\|_{\mathcal{M}} + \beta|\varepsilon| \|\nabla v\|_{L^p(\Omega)} + \frac{\varepsilon^2}{2} \|h_1 + h_2\|_{L^2(\Omega)}^2 &\geq \varepsilon \langle r, h_1 + h_2 \rangle. \end{aligned}$$

Dividing with $\varepsilon > 0$ and letting $\varepsilon \rightarrow 0$, one has

$$\langle r, h_1 + h_2 \rangle \leq \alpha \|Dh_1\|_{\mathcal{M}} + \beta \|\nabla h_2\|_{L^p(\Omega)},$$

i.e., $\|r\|_{(\alpha,\beta)} \leq 1$. If we replace $(h_1, h_2) = (u, v) \in \text{BV} \times \text{W}^{1,p}$ and $\varepsilon \in (-1, 1)$ then

$$\alpha\varepsilon \|Du\|_{\mathcal{M}} + \beta\varepsilon \|\nabla v\|_{L^p(\Omega)} + \frac{\varepsilon^2}{2} \|u + v\|_{L^2(\Omega)}^2 \geq \varepsilon \langle r, u + v \rangle.$$

Again considering the cases with $\varepsilon \rightarrow 0^\pm$, we conclude that

$$|\langle r, u + v \rangle| = \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)},$$

with $\|r\|_{(\alpha,\beta)} = 1$.

Finally, if (4.73) is true then for every $(h_1, h_2) \in \text{BV} \times \text{W}^{1,p}$ and $\varepsilon \in \mathbb{R}$

$$\begin{aligned} &\alpha \|D(u + \varepsilon h_1)\|_{\mathcal{M}} + \beta \|\nabla(v + \varepsilon h_2)\|_{L^p(\Omega)} + \frac{1}{2} \|r - \varepsilon(h_1 + h_2)\|_{L^2(\Omega)}^2 \\ &\geq \langle r, u + v + \varepsilon(h_1 + h_2) \rangle + \frac{1}{2} \|r\|_{L^2(\Omega)}^2 + \frac{\varepsilon^2}{2} \|h_1 + h_2\|_{L^2(\Omega)}^2 - \varepsilon \langle r, h_1 + h_2 \rangle \\ &= \langle r, u + v \rangle + \frac{1}{2} \|r\|_{L^2(\Omega)}^2 + \frac{\varepsilon^2}{2} \|h_1 + h_2\|_{L^2(\Omega)}^2 \\ &= \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|r\|_{L^2(\Omega)}^2 + \frac{\varepsilon^2}{2} \|h_1 + h_2\|_{L^2(\Omega)}^2 \\ &\geq \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{L^p(\Omega)} + \frac{1}{2} \|r\|_{L^2(\Omega)}^2 \end{aligned}$$

i.e., $(u, v) \in \mathcal{G}$ is an optimal decomposition of f . □

Notice that condition (4.72) is valid if and only if $\int_{\Omega} f = 0$. Otherwise, by Lemma 4.5.7 we have that $\|f\|_{(\alpha,\beta)} = \infty$.

4.6 Numerical Experiments

Throughout this section, we present our numerical simulations in one and two dimensional setting for the problem (P). We start with the one dimensional case where we verify numerically the analytical solutions obtained in Section 4.4.3 for $p = 2$ and $p = \infty$. The type of structures that are promoted for different values of p is also examined. Moreover, we perform a structural decomposition using the equivalent model (4.54) for complex one dimensional data. Finally, we proceed to the two dimensional experiments where we

focus on image denoising tasks and in particular on the reduction and elimination of the staircasing effect.

The discretised version of problem (\mathcal{P}) is defined as

$$\min_{u \in \mathbb{R}^{n \times m}} \frac{1}{2} \|f - u\|_2^2 + \text{TVL}_{\alpha, \beta}^p(u). \quad (4.74)$$

Here $\text{TVL}_{\alpha, \beta}^p : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$ with $1 < p \leq \infty$ is

$$\text{TVL}_{\alpha, \beta}^p(u) = \operatorname{argmin}_{w \in (\mathbb{R}^{n \times m})^2} \alpha \|\nabla u - w\|_1 + \beta \|w\|_p,$$

where for $x \in \mathbb{R}^{n \times m}$, we set $\|x\|_p = \left(\sum_{i,j=1}^{n,m} |x(i,j)|^p \right)^{\frac{1}{p}}$ and for $x = (x_1, x_2) \in (\mathbb{R}^{n \times m})^2$ we set

$$\|x\|_p = \left(\sum_{i,j=1}^{n,m} \left(\sqrt{(x_1(i,j))^2 + (x_2(i,j))^2} \right)^p \right)^{\frac{1}{p}}.$$

We denote by $\nabla u = ((\nabla u)_1, (\nabla u)_2)$ the discretised gradient with forward differences and zero Neumann boundary conditions defined as

$$\begin{aligned} (\nabla u)_1(i,j) &= \begin{cases} \frac{u(i,j+1) - u(i,j)}{t} & \text{if } 1 \leq i \leq n, 1 \leq j < m, \\ 0 & \text{if } 1 \leq i \leq n, j = m, \end{cases} \\ (\nabla u)_2(i,j) &= \begin{cases} \frac{u(i+1,j) - u(i,j)}{t} & \text{if } 1 \leq i < n, 1 \leq j \leq m, \\ 0 & \text{if } i = n, 1 \leq j \leq m, \end{cases} \end{aligned}$$

where t denotes the step size. The discrete version of the divergence operator is defined as the adjoint of ∇ , i.e., for every $w = (w_1, w_2) \in (\mathbb{R}^{n \times m})^2$ and $u \in \mathbb{R}^{n \times m}$, we have that $\langle -\operatorname{div} w, u \rangle = \langle w, \nabla u \rangle$ with

$$\begin{aligned} \operatorname{div} w &= \begin{cases} \frac{w_1(i,j) - w_2(i-1,j)}{t} & \text{if } 1 < i < n, 1 \leq j \leq m, \\ \frac{w_1(i,j)}{t} & \text{if } i = 1, 1 \leq j \leq m, \\ -\frac{w_1(i-1,j)}{t} & \text{if } i = m, 1 \leq j \leq m, \end{cases} \\ &+ \begin{cases} \frac{w_2(i,j) - w_1(i,j-1)}{t} & \text{if } 1 < j < m, 1 \leq i \leq n, \\ \frac{w_2(i,j)}{t} & \text{if } j = 1, 1 \leq i \leq n, \\ -\frac{w_2(i,j-1)}{t} & \text{if } j = m, 1 \leq i \leq n. \end{cases} \end{aligned}$$

We solve the minimisation problem (4.74) in two ways. The first one is by using the CVX optimisation package under MOSEK solver (e.g. interior point methods), see [GB14, Mos08]. This method is efficient for small-medium scale optimisation problems

and thus it is a suitable choice in order to replicate one dimensional solutions. On the other hand, we prefer to solve large scale two dimensional versions of (4.74) with the split Bregman method [GO09] which has been widely used for the fast solution of non-smooth minimisation problems.

4.6.1 Split Bregman for TVL^p-L²

We describe how we adapt the split Bregman algorithm to our discrete model (4.74). We have already encountered the split Bregman algorithm in the previous chapter. Letting $z = \nabla u - w$, the corresponding unconstrained problem becomes

$$\min_{\substack{u \in \mathbb{R}^{n \times m} \\ w \in (\mathbb{R}^{n \times m})^2 \\ z \in (\mathbb{R}^{n \times m})^2}} \frac{1}{2} \|f - u\|_2^2 + \alpha \|z\|_1 + \beta \|w\|_p, \quad \text{such that } z = \nabla u - w.$$

Replacing the constraint by a Lagrange multiplier λ , we obtain the following unconstrained formulation:

$$\min_{\substack{u \in \mathbb{R}^{n \times m} \\ w \in (\mathbb{R}^{n \times m})^2 \\ z \in (\mathbb{R}^{n \times m})^2}} \frac{1}{2} \|f - u\|_2^2 + \alpha \|z\|_1 + \beta \|w\|_p + \frac{\lambda}{2} \|z - \nabla u + w\|_2^2. \quad (4.75)$$

The Bregman iteration, see [OBG⁺05], that corresponds to the minimisation (4.75) leads to the following two step algorithm:

$$\begin{aligned} (u^{k+1}, z^{k+1}, w^{k+1}) &= \underset{u, z, w}{\operatorname{argmin}} \frac{1}{2} \|f - u\|_2^2 + \alpha \|z\|_1 + \beta \|w\|_p + \frac{\lambda}{2} \|b^k - z + \nabla u - w\|_2^2, \\ b^{k+1} &= b^k + z^{k+1} - \nabla u^{k+1} - w^{k+1} \end{aligned} \quad (4.76)$$

Since solving (4.76) at once is a difficult task, we employ a splitting technique and minimise alternately for u, z and w . This yields the split Bregman iteration for our method:

Split Bregman algorithm for (4.74)

$$u^{k+1} = \operatorname{argmin}_{u \in \mathbb{R}^{n \times m}} \frac{1}{2} \|f - u\|_2^2 + \frac{\lambda}{2} \|b^k + z^k - \nabla u + w^k\|_2^2, \quad (4.77)$$

$$z^{k+1} = \operatorname{argmin}_{z \in (\mathbb{R}^{n \times m})^2} \alpha \|z\|_1 + \frac{\lambda}{2} \|b^k + z - \nabla u^{k+1} + w^k\|_2^2, \quad (4.78)$$

$$w^{k+1} = \operatorname{argmin}_{w \in (\mathbb{R}^{n \times m})^2} \beta \|w\|_p + \frac{\lambda}{2} \|b^k + z^{k+1} - \nabla u^{k+1} + w\|_2^2, \quad (4.79)$$

$$b^{k+1} = b^k + z^{k+1} - \nabla u^{k+1} - w^{k+1}. \quad (4.80)$$

Next, we discuss how we solve each of the subproblems (4.77)–(4.79) for the $p < \infty$ and $p = \infty$ cases.

Solution of (4.77): The first-order optimality condition of (4.77) results into the following linear system:

$$\underbrace{(I - \lambda \Delta)}_A u = \underbrace{f - \lambda \operatorname{div}(b^k + z^k - w^k)}_c. \quad (4.81)$$

Here A is a sparse, symmetric, positive definite and strictly diagonal dominant matrix, thus we can easily solve (4.81) with an iterative solver such as conjugate gradient or Gauss–Seidel method. However, due to the zero Neumann boundary conditions, the matrix A can be efficiently diagonalised by the two dimensional discrete cosine transform,

$$A = W_{nm}^\top D W_{nm}, \quad (4.82)$$

where here W_{nm} is the discrete cosine matrix and $D = \operatorname{diag}(\mu_1, \dots, \mu_{n^*m})$ is the diagonal matrix of the eigenvalues of A . In this case, A has a particular structure of a block symmetric *Toeplitz-plus-Hankel* matrix with *Toeplitz-plus-Hankel* blocks and one can obtain the solution of (4.77) by three operations involving the discrete cosine transform [Han10] as follows: Firstly, we calculate the eigenvalues of A by multiplying (4.82) with $e_1 = (1, 0, \dots, 0)^\top$ from both sides and using the fact that $W_{nm}^\top W_{nm} = W_{nm} W_{nm}^\top = I_{nm}$, we get

$$D_{i,i} = \frac{[W_{nm} A e_1]_i}{[W_{nm} e_1]_i}, \quad i = 1, 2, \dots, nm. \quad (4.83)$$

Then, the solution of (4.77) is computed exactly by

$$u = W_{nm}^\top D^{-1} W_{nm} c. \quad (4.84)$$

Solution of (4.78): The solution of the subproblem (4.78) is obtained in a closed form

via the following *shrinkage operator*, see also [GO09, WYYZ08]. Indeed, for $i = 1, 2$ we have

$$z_i^{k+1} = \text{shrink}_{\frac{\alpha}{\lambda}} \left(\underbrace{b_i^k - \nabla_i u^{k+1} + w_i^k}_{g_i} \right) := \max \left(\|g\|_2 - \frac{\alpha}{\lambda}, 0 \right) \frac{g_i}{\|g\|_2}. \quad (4.85)$$

Solution of (4.79): Finally, we discuss the solution of the subproblem (4.79). In the spirit of [VO96], we solve (4.79) by a fixed point iteration scheme, where we consider finite values of p . Letting $\kappa = \frac{\beta}{\lambda}$ and $\eta = -b^k - z^{k+1} + \nabla u^{k+1}$, the first-order optimality condition of (4.79) becomes

$$\kappa \frac{|w|^{p-2} w}{\|w\|_p^{p-1}} + w - \eta = 0. \quad (4.86)$$

For given w^k , we obtain w^{k+1} by the following fixed point iteration

$$w_i^{k+1} = \frac{\eta_i \|w^k\|_p^{p-1}}{\kappa |w^k|^{p-2} + \|w^k\|_p^{p-1}}, \quad (4.87)$$

under the convention that $0/0 = 0$. We can also consider solving the p -homogenous analogue ($\mathcal{P}_{p\text{-hom}}$), where for certain values of p , e.g. $p = 2$, we can solve exactly (4.87), since in that case $w_i^{k+1} = \frac{\eta_i}{\kappa+1}$. However, we observe numerically that there is no significant computational difference between these two methods.

On the other hand, for $p = \infty$ equation (4.79) becomes

$$\underset{w \in (\mathbb{R}^n \times m)^2}{\text{argmin}} \kappa \|w\|_\infty + \frac{1}{2} \|w - \eta\|_2^2, \quad (4.88)$$

which is the *proximity operator* associated to $\kappa \|\cdot\|_\infty$ and is denoted by $\text{prox}_{\kappa \|\cdot\|_\infty}(\eta)$. Then, since $(\kappa \|\cdot\|_\infty)^*(w) = \mathcal{I}_{\{\|\cdot\|_1 \leq \kappa\}}(w)$ and by the proximity identity, see for instance [CW05, CP08],

$$\eta = \text{prox}_{\kappa \|\cdot\|_\infty}(\eta) + \text{prox}_{\mathcal{I}_{\{\|\cdot\|_1 \leq \kappa\}}}(\eta) \quad (4.89)$$

it suffices to compute the *Euclidean* projection onto L^1 ball i.e.,

$$\begin{aligned} \text{prox}_{\mathcal{I}_{\{\|\cdot\|_1 \leq \kappa\}}}(\eta) &= \underset{v}{\text{argmin}} \frac{1}{2} \|v - \eta\|_2^2 + \mathcal{I}_{\{\|\cdot\|_1 \leq \kappa\}}(\eta) \\ &= \underset{\|v\|_1 \leq \kappa}{\text{argmin}} \frac{1}{2} \|v - \eta\|_2^2 := \pi_{\{\|\cdot\|_1 \leq \kappa\}}(\eta), \end{aligned} \quad (4.90)$$

and the solution of (4.88) is now $w = \eta - \pi_{\{\|\cdot\|_1 \leq \kappa\}}(\eta)$. A detailed analysis for an efficient algorithm solving (4.90) is proposed in [DSYT08] and for the vector-valued cases we follow [FP11]. Let us finally mention that since we do not solve exactly all the subproblems (4.77)–(4.79), we do not have a convergence proof for the split Bregman iteration. However in practice, the algorithm converges faster to the right solutions after comparing them with

the corresponding solutions obtained with the CVX package.

4.6.2 One dimensional results

Here, we set $m = 1$ and thus $u \in \mathbb{R}^{n \times 1}$, $w \in (\mathbb{R}^{n \times 1})^2$. Initially, we compare our numerical solutions with the analytical ones, obtained in Section 4.4.3 for the step function for $p = 2$, $p = \infty$ and the piecewise affine function for $p = \infty$. We also demonstrate that for a particular datum the solution of (\mathcal{P}) coincides with the one dimensional second order total generalised variation, see Proposition 4.4.11. Finally, we continue with numerical experiments related to the image decomposition approach in Section 4.5. There, we concentrate on the geometrical structure decomposition of one dimensional signal into a piecewise constant and structural components depending on the value of p .

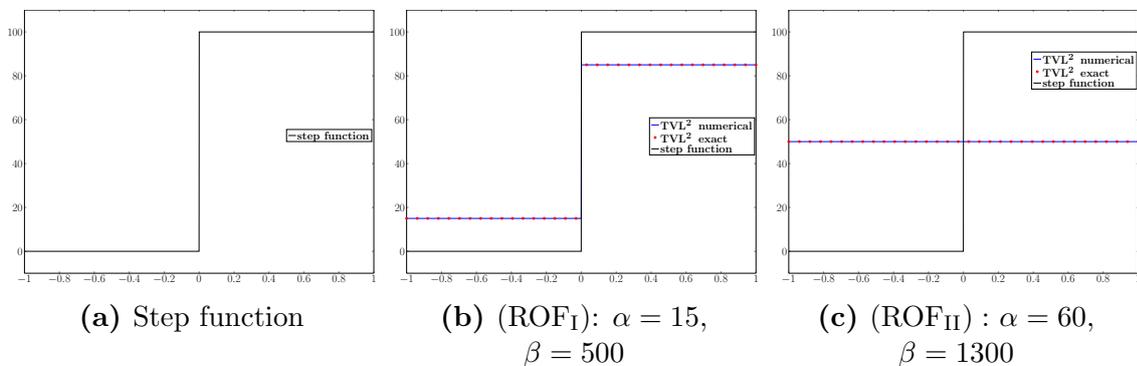


Figure 4.8: Comparison between numerical solutions of (\mathcal{P}) and the corresponding analytical solutions obtained in Section 4.4.3. The parameters α and β are chosen so that conditions (4.43) and (4.44) are satisfied.

The step function defined in (4.41) is discretised into 2000 points where $h = 100$, $L = 1$ and $\Omega = [-1, 1]$. We first examine the ROF-type solutions for $p = 2$, where the parameters α and β are selected according to the conditions (4.43) and (4.44), see Figure 4.8. There, we see that the analytical solutions coincide with the numerical ones. Obviously, we can obtain identical results if one selects $p = \infty$ with an appropriate choice of α and β .

Now, we proceed by computing the non-ROF solutions. The numerical solutions are obtained using the 2-homogeneous analogue of (4.46), since we have proved that the 1-homogeneous and p -homogeneous problems are equivalent modulo an appropriate rescaling of the parameter β , see Proposition 4.3.2. In fact, as it is described in Figure 4.5, in order to obtain solutions from the purple region, it suffices to seek solutions for the 2-homogeneous (4.46). Notice also that these solutions are exactly the solutions obtained solving a Huber-TV problem, see Proposition 4.3.3. The analytical solutions are given in (4.48) and (4.49) and are compared with the numerical ones in Figure 4.9, where we observe that they coincide. We also verify the equivalence between the 1-homogeneous and 2-homogeneous

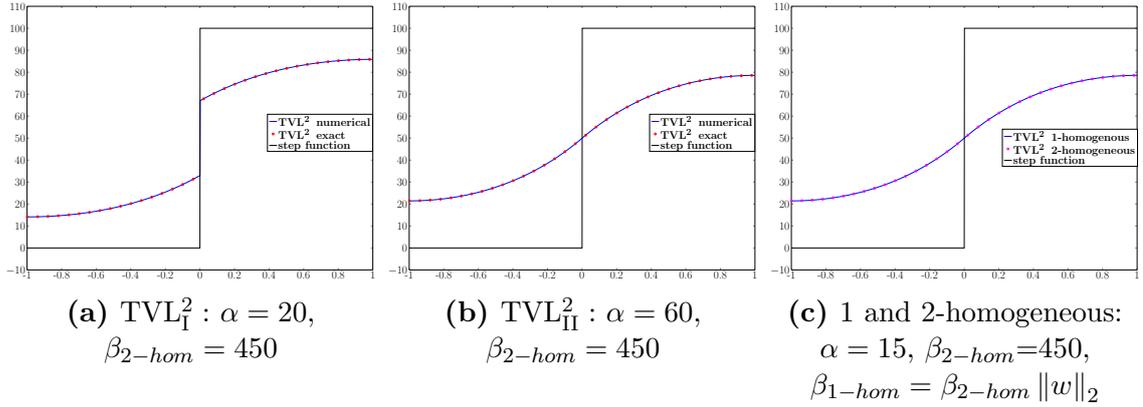


Figure 4.9: Comparison between numerical and analytical solutions obtained in Section 4.4.3 for the step function, by solving the 2-homogeneous problem (4.46). The parameters α and β are chosen so that conditions (4.48) and (4.49) are satisfied. The last plot indicates the equivalence between the 2-homogeneous (4.46) and the 1-homogeneous problem (\mathcal{P}).

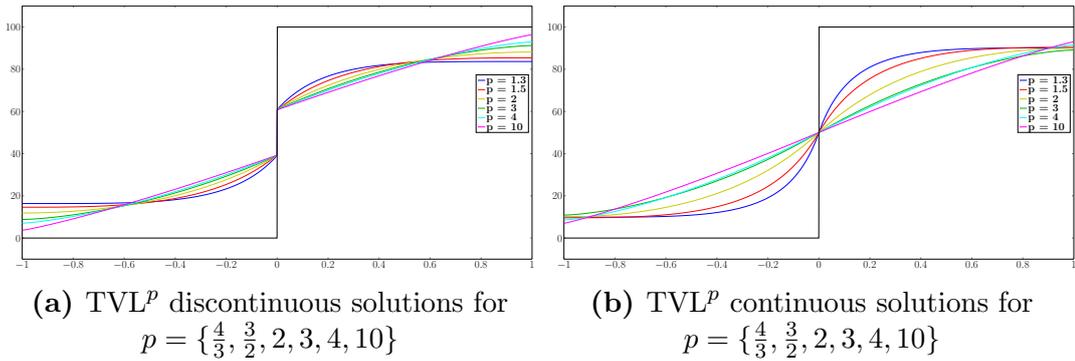


Figure 4.10: Step function: The structure of solutions for (\mathcal{P}) with different values of p .

problems where α is fixed and β is obtained from Proposition 4.3.2, see Figure 4.9c.

We continue our experiments for general finite values of p focusing on the geometric behaviour of the solutions as p increases. In order to compare the solutions for $p \in (1, \infty)$, we fix the parameter α and choose appropriate values of β and p . We choose α and β so that they belong to the purple region in Figure 4.5, i.e., $\beta < (\frac{2L}{q+1})^{\frac{1}{q}} \alpha$ and $\beta < \frac{h}{2} (\frac{2L^{q+1}}{q+1})^{\frac{1}{q}}$, hence non-ROF solutions are obtained. We set $p = \{\frac{4}{3}, \frac{3}{2}, 2, 3, 4, 10\}$ and for the solutions that preserve the discontinuity we select $\beta = \{72, 140, 430, 1350, 2400, 6800\}$ with fixed $\alpha = 20$ (observe that $\beta < (\frac{2L}{q+1})^{\frac{1}{q}} \alpha$ is valid in any case), see Figure 4.10a. For the continuous cases, we set $\alpha = 60$ and $\beta = \{50, 110, 430, 1700, 3000, 9500\}$ (again the conditions $\alpha \geq \frac{hL}{2}$ and $\beta < \frac{h}{2} (\frac{2L^{q+1}}{q+1})^{\frac{1}{q}}$ hold), see Figure 4.10b. Notice that for $p = \frac{4}{3}$, the solution has a similar behaviour to $p = 2$, but with a steeper gradient at the discontinuity point. Moreover, the solution becomes almost constant near the boundary of Ω . On the other hand, as we increase p , the slope of the solution near the discontinuity point reduces and it becomes

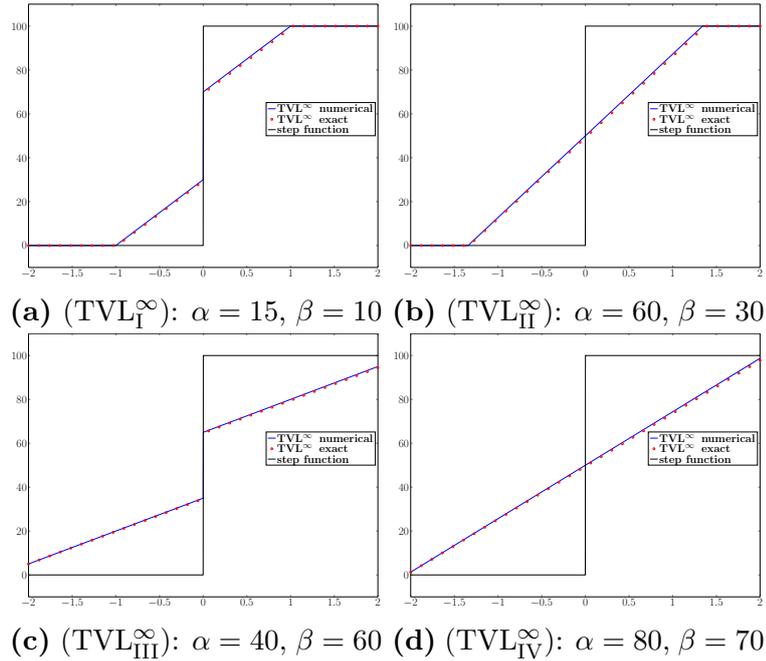


Figure 4.11: Step function: Comparison between numerical solutions of (\mathcal{P}) with $p = \infty$ and the corresponding analytical solutions obtained in Section 4.4.3. The parameters α and β satisfy the corresponding conditions for each solution, see $(\text{TVL}_I^\infty : \text{step})$ - $(\text{TVL}_{IV}^\infty : \text{step})$.

almost linear with a relative small constant part near the boundary.

The *almost* linear structure of the solutions that appears for large p motivates us to examine the case of $p = \infty$, firstly for the step function and secondly for a piecewise affine function defined in (4.53) where $\lambda = 50$ and the data is again discretised in 2000 points. For demonstration purposes, here $\Omega = [-2, 2]$. In Figures 4.11a-4.11d, we present our numerical solutions compared to the analytical ones for the step function with $p = \infty$. The parameters are selected according to $(\text{TVL}_I^\infty : \text{step})$ - $(\text{TVL}_{IV}^\infty : \text{step})$. Similarly, in Figures 4.12a-4.12h we present all the eight type of solutions regarding the piecewise affine function. In addition to this, we analyse a specific type of solution which appears in Figure 4.7. In the yellow region, namely the $(\text{TVL}_{III}^\infty : \text{affine})$ solution, if we select $\beta = \alpha L$ then we can achieve a solution with the same slope as our initial data. If we compare this result with a high order regulariser as $\text{TGV}_{\alpha,\beta}^2$, we observe that although in both cases we preserve the discontinuity, $\text{TGV}_{\alpha,\beta}^2$ cannot respect the geometry of the data. In fact, none of the acceptable solutions obtained using $\text{TGV}_{\alpha,\beta}^2$ for a piecewise affine function, see [PB15, Section 5.2] can achieve a similar result. In Figure 4.13, we present our results using the $\text{TGV}_{\alpha,\beta}^2$ and $\text{TVL}_{\alpha,\beta}^\infty$ regularisers. Furthermore, TVL^∞ is equivalent to TGV^2 on certain class of one dimensional data. In order to achieve this kind of result we use a symmetric input data as in Figure 4.14. It is a combination of two piecewise affine functions with two discontinuities creating a spike at the origin. The parameters

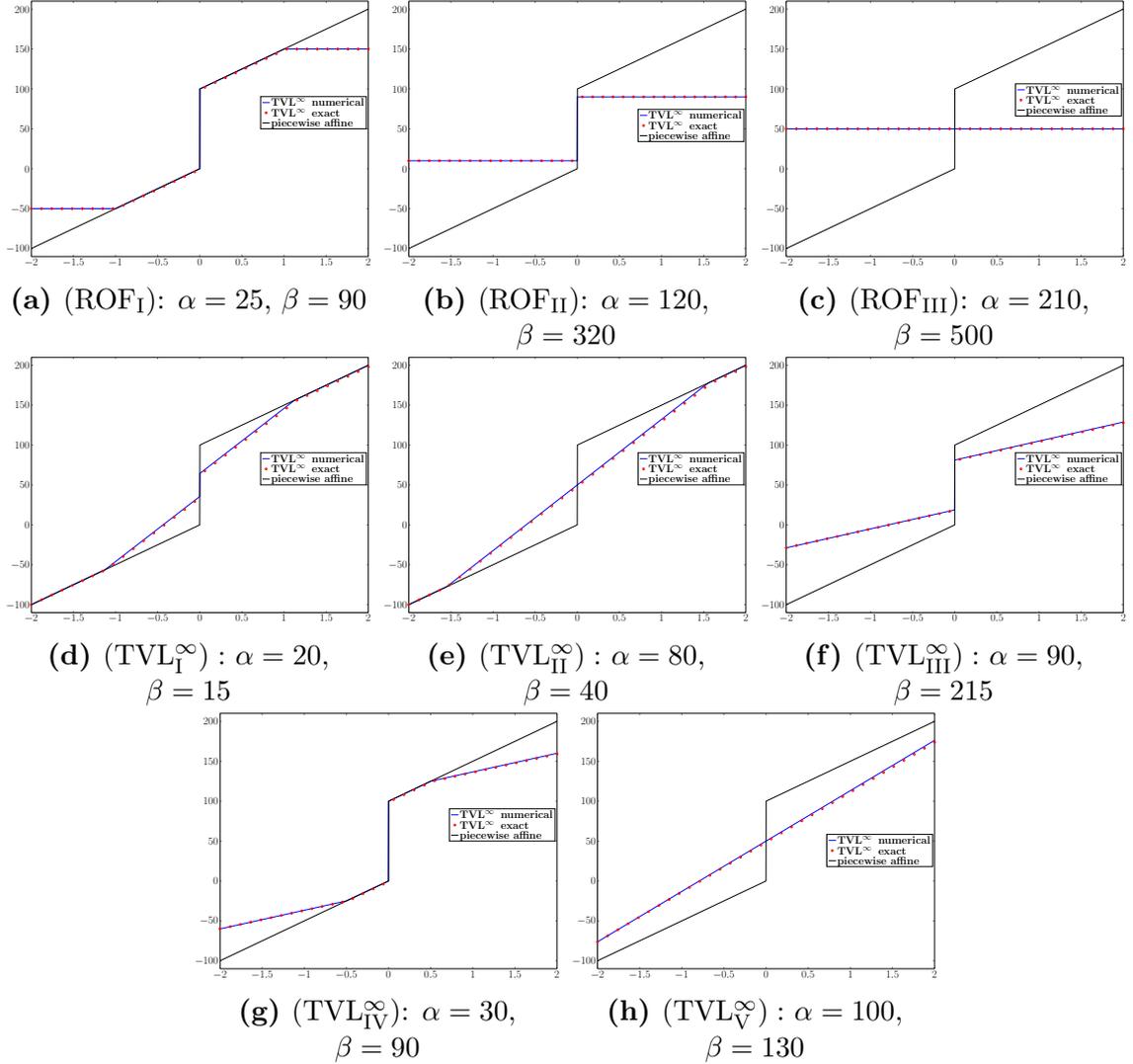


Figure 4.12: Piecewise affine function: Comparison between numerical solutions of (\mathcal{P}) with $p = \infty$ and the corresponding analytical solutions obtained in Section 4.4.3. The parameters α and β satisfy the corresponding conditions for each solution see (ROF_I : affine)-(TVL_V[∞] : affine).

are selected according to Proposition 4.4.11.

In the last part of this section, we discuss the image decomposition approach presented in Section 4.5. We treat a more complicated one dimensional noiseless signal with piecewise constant, affine and exponential components and solve the discretised version of (4.54) using CVX under MOSEK. We verify numerically the equivalence between (4.54) and (\mathcal{P}) for $p = 2$, i.e., $(\nabla v, u + v)$ corresponds to (w, \bar{u}) where (v, u) and (w, \bar{u}) are the solutions of (4.54) and (\mathcal{P}) respectively, see Figure 4.15. The parameters α and β for Figure 4.15a are selected appropriately in order to have an optimal decomposition into piecewise

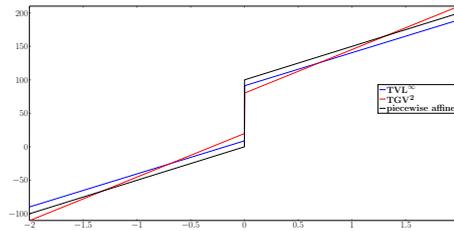


Figure 4.13: Numerical comparison of $\text{TGV}_{\alpha,\beta}^2$ and $\text{TVL}_{\alpha,\beta}^\infty$ regularisers for the piecewise affine function. The parameters α and β are chosen such that $\|f - u_{\text{TGV}^2}\|_2 = \|f - u_{\text{TVL}^\infty}\|_2$.

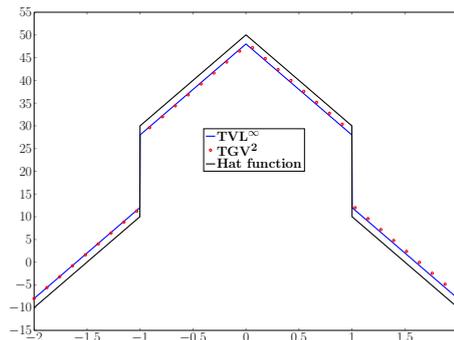


Figure 4.14: Numerical comparison of $\text{TGV}_{\alpha,\beta}^2$ and $\text{TVL}_{\alpha,\beta}^\infty$ regularisers for the symmetric hat function. The parameters are $\alpha = 2$ and $\beta_{\text{TGV}^2} = 2$ and $\beta_{\text{TVL}^\infty} = 4$.

constant and exponential components. We also compare the decomposed parts u, v for two extreme values of p , i.e., $p = \frac{4}{3}$ and $p = \infty$. In order to have a reasonable comparison on the corresponding solutions, here the parameters α, β are selected such that the residual $\|f - u - v\|_2$ is the same for both values of p . As we observe, the v decomposition with $p = \frac{4}{3}$ promotes some *flatness* on the solution, meaning that it converges to an almost piecewise constant solution, see Figure 4.16a. On the other hand for $p = \infty$, more spikes are introduced in the solution $u + v$, see Figure 4.16b.

4.6.3 Two dimensional results

Let us now consider images instead of signals, where $u \in \mathbb{R}^{n \times m}$, $w \in (\mathbb{R}^{n \times m})^2$ with $m > 1$ and Ω denotes a rectangular/square image domain. We focus on image denoising tasks and on eliminating the staircasing effect for different values of p . We may again use CVX under MOSEK but with some serious drawbacks in terms of computational efficiency. In fact, for large image data, CVX fails to obtain results for high values of p due to out of memory errors.

We start with the image in Figure 4.17, i.e., a square with piecewise affine structures. The image size is 200×200 pixels at a $[0, 1]$ intensity range. The noisy image, Figure 4.17b, is a corrupted version of the original image with Gaussian noise of zero mean and variance $\sigma = 0.01$.

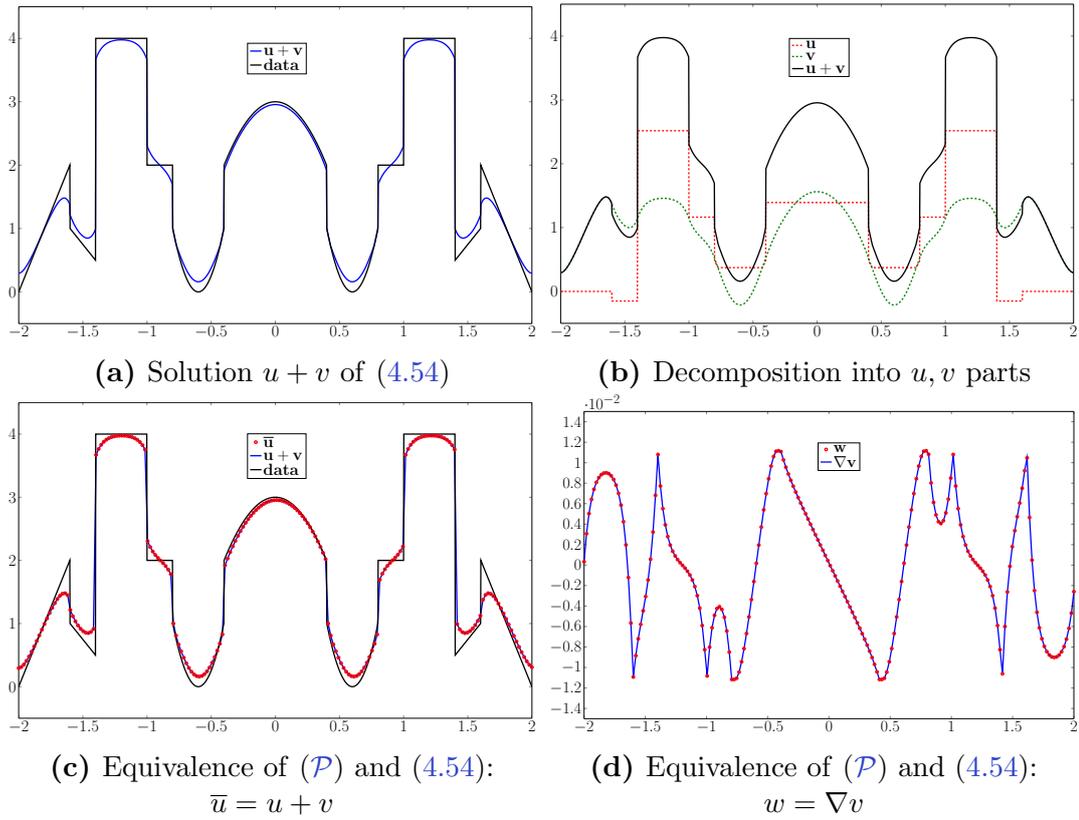


Figure 4.15: Numerical results on the image decomposition approach (4.54) for $p = 2$, see Section 4.5.

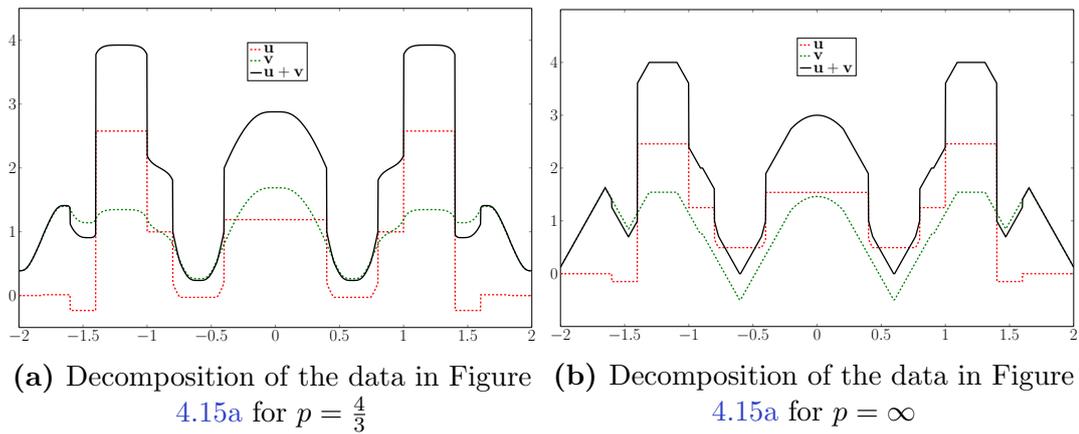


Figure 4.16: Decomposition of the data in Figure 4.15a into u, v parts for $p = \frac{4}{3}$ and $p = \infty$. The value $p = \frac{4}{3}$ produces a v component with flat structures while $p = \infty$ produces a component with affine structures. In both cases we have $\|f - u - v\|_2 = 6.667$.

In Figure 4.18, we present the best reconstructions results in terms of two quality measures, the *Peak Signal to Noise Ratio* (PSNR) and the *Structural Similarity Index*

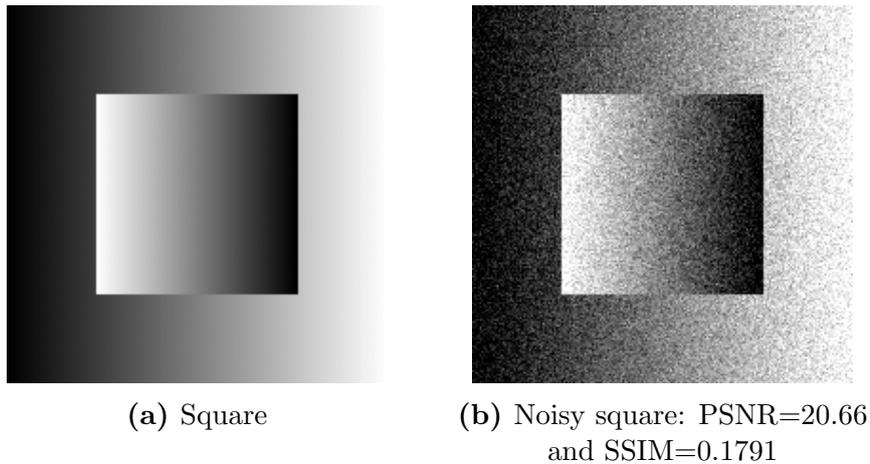


Figure 4.17: Square with piecewise affine structures and its noisy version with $\sigma = 0.01$.

(SSIM), see [WBSS04] for the definition of the latter. In each case, the values of α and β are selected appropriately for the optimal PSNR and SSIM. Our stopping criterion for the proposed algorithm in Section 4.6.1 is the relative residual error becoming less than 10^{-6} i.e.,

$$\frac{\|u^{k+1} - u^k\|_2}{\|u^{k+1}\|_2} \leq 10^{-6}. \quad (4.91)$$

We also fix a maximum number of iterations for the finite and large values of p , e.g. $p = 7$, which is at 5000 iterations. This is in order to ensure that the fixed point iteration we perform on the w subproblem (4.87) converges to the correct solution. We observe by comparing with the CVX results that for $p < 4$ the stopping rule in (4.91) is sufficient to obtain the correct solution.

Finally, for computational efficiency, we fix $\lambda = 10\alpha$ when $1 < p < 4$ and $\lambda = 1000\alpha$ when $4 \leq p < \infty$. This is an empirical rule in order to reduce the computational cost when we perform the fixed point iteration for large but finite values of p . For $p = \infty$, we fix again $\lambda = 10\alpha$. The best reconstructions in terms of the PSNR have no visual difference among $p = \frac{3}{2}, 2$ and 3 and the staircasing is obvious, see Figures 4.18a–4.18c. This is one more indication that the PSNR – which is based on the squares of the difference between the ground truth and the reconstruction – does not correspond to the optimal visual results. However, the best reconstructions in terms of SSIM are visually better. They exhibit significantly reduced staircasing for $p = \frac{3}{2}$ and $p = 3$ and is essentially absent in the case of $p = 2$, see Figures 4.18d–4.18f.

We can also get a total staircasing elimination by setting higher values for the parameters α and β , as we show in Figure 4.19. There, one observes that on one hand as we increase p , almost affine structures are promoted – see the middle row profiles (blue) in Figure 4.19, where we compare them with the middle row profiles of the ground truth

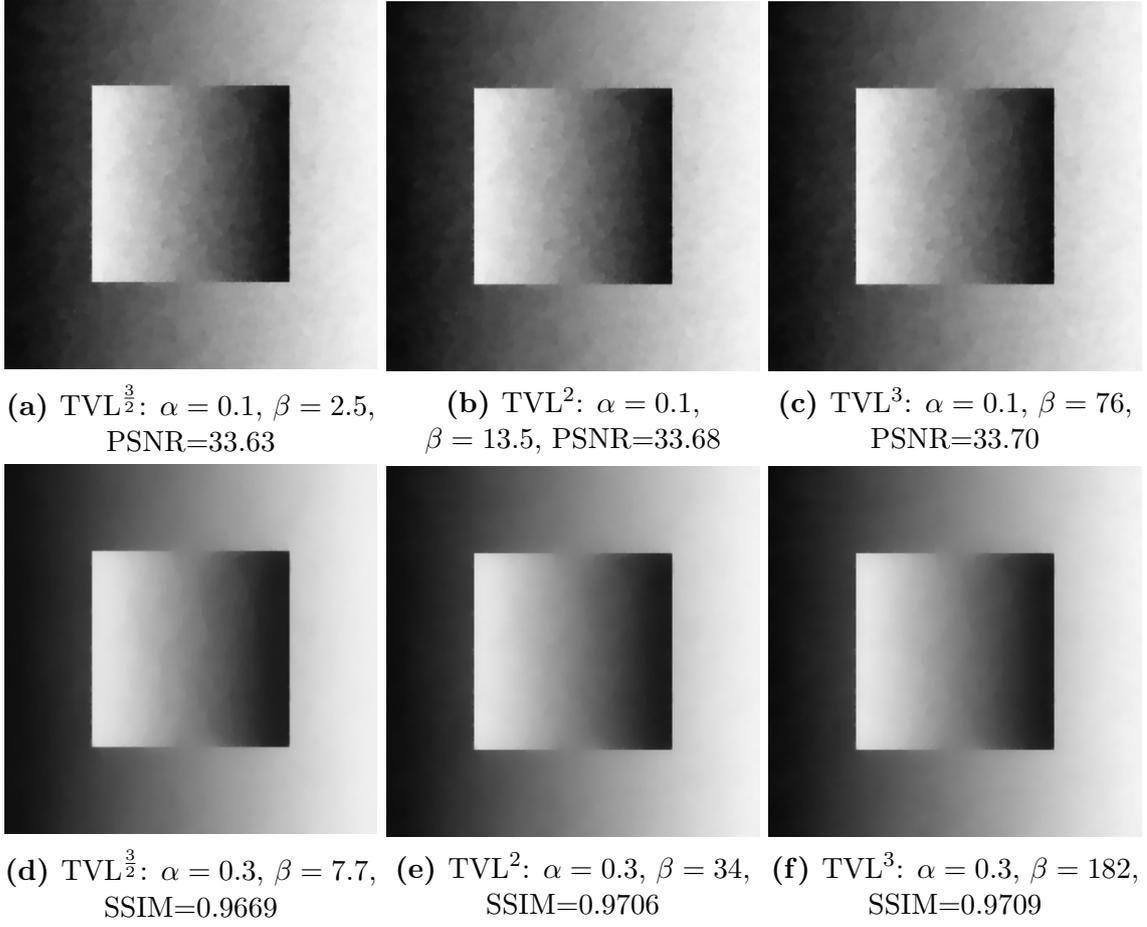


Figure 4.18: Best reconstructions in terms of PSNR and SSIM for $p = \frac{3}{2}, 2, 3$.

(green) – and on the other hand these choices of α, β produce a serious loss of contrast that however can be easily treated via the *Bregman iteration*. Contrast enhancement via Bregman iteration was introduced in [OBG⁺05], see also [BBBM13] for an application to higher-order models. It involves solving a modified version of the minimisation problem (\mathcal{P}). Setting $u^0 = f$, for $k = 1, 2, \dots$, we solve iteratively

$$\begin{aligned}
 u^{k+1} &= \underset{\substack{u \in \mathbb{R}^{n \times m} \\ w \in (\mathbb{R}^{n \times m})^2}}{\operatorname{argmin}} \frac{1}{2} \left\| f + \tilde{v}^k - u \right\|_2^2 + \alpha \|\nabla u - w\|_1 + \beta \|w\|_p, \\
 \tilde{v}^{k+1} &= \tilde{v}^k + f - u^{k+1}.
 \end{aligned} \tag{4.92}$$

Instead of solving (4.74) once for fixed α and β , we solve a sequence of similar problems adding back a noisy residual in each iteration which results to a contrast improvement. For stopping criteria, e.g. *discrepancy principle*, regarding the Bregman iteration we refer to [OBG⁺05]. In Figure 4.20 we present our best Bregmanised results in terms of SSIM. There, we notice that Bregman iteration leads to a significant contrast improvement, in

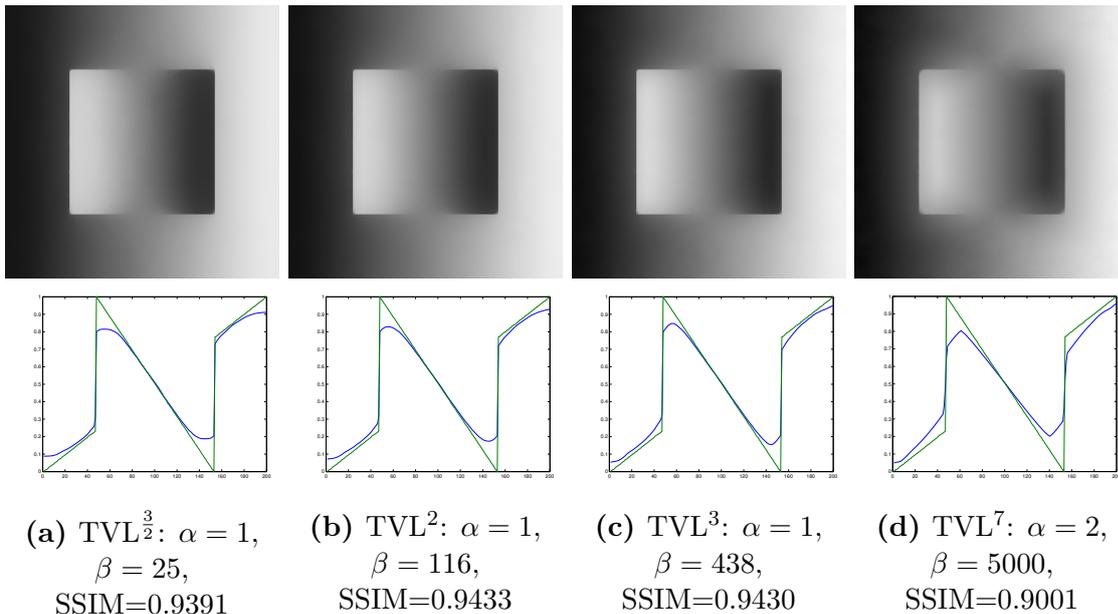


Figure 4.19: Staircasing elimination for $p = \frac{3}{2}, 2, 3$ and 7 . High values of p promotes almost affine structures as we observe from the middle-row comparison of the solution u (blue) and the ground truth (green).

comparison to the results of Figure 4.19. Notice that the Bregmanised TVL^2 (first-order) reconstruction is indeed visually close to the Bregmanised TGV^2 (4.6) (second-order), compare Figures 4.20e and 4.20f. We use the Chambolle–Pock primal-dual method, cf. [CP11], to solve the corresponding second-order TGV^2 and second-order Bregmanised TGV^2 .

We continue our experimental analysis with a radially symmetric image, see Figure 4.21. In Figure 4.22, we demonstrate that we can achieve staircasing-free reconstructions for $p = \frac{3}{2}, 2, 3$ and 7 . Let us emphasise the fact that when we increase p , we obtain results that preserve the spike in the centre of the circle, see Figure 4.22d.

The loss of contrast can be treated again using the Bregman iteration (4.92). The best results of the latter in terms of SSIM are presented in Figure 4.23, for $p = 2, 4$ and 7 and they are also compared with the corresponding Bregmanised TV and TGV^2 . We observe that we can obtain reconstructions that are visually close to the TGV^2 ones and in fact notice that for $p = 7$, the spike on the centre of the circle is better reconstructed compared to TGV^2 . This behaviour is clear when we visualise the surface plots of the corresponding reconstructions, see Figure 4.24. As we observe both in Figures 4.23f and 4.24h, TGV^2 regulariser tends to *smooth* out the spike-like structures appear in the origin of the circle and behaves similar to the TVL^2 regulariser.

We conclude with numerical results for the image decomposition approach of Section 4.5 which we solve again using the split Bregman algorithm. Recall that in two dimensions,

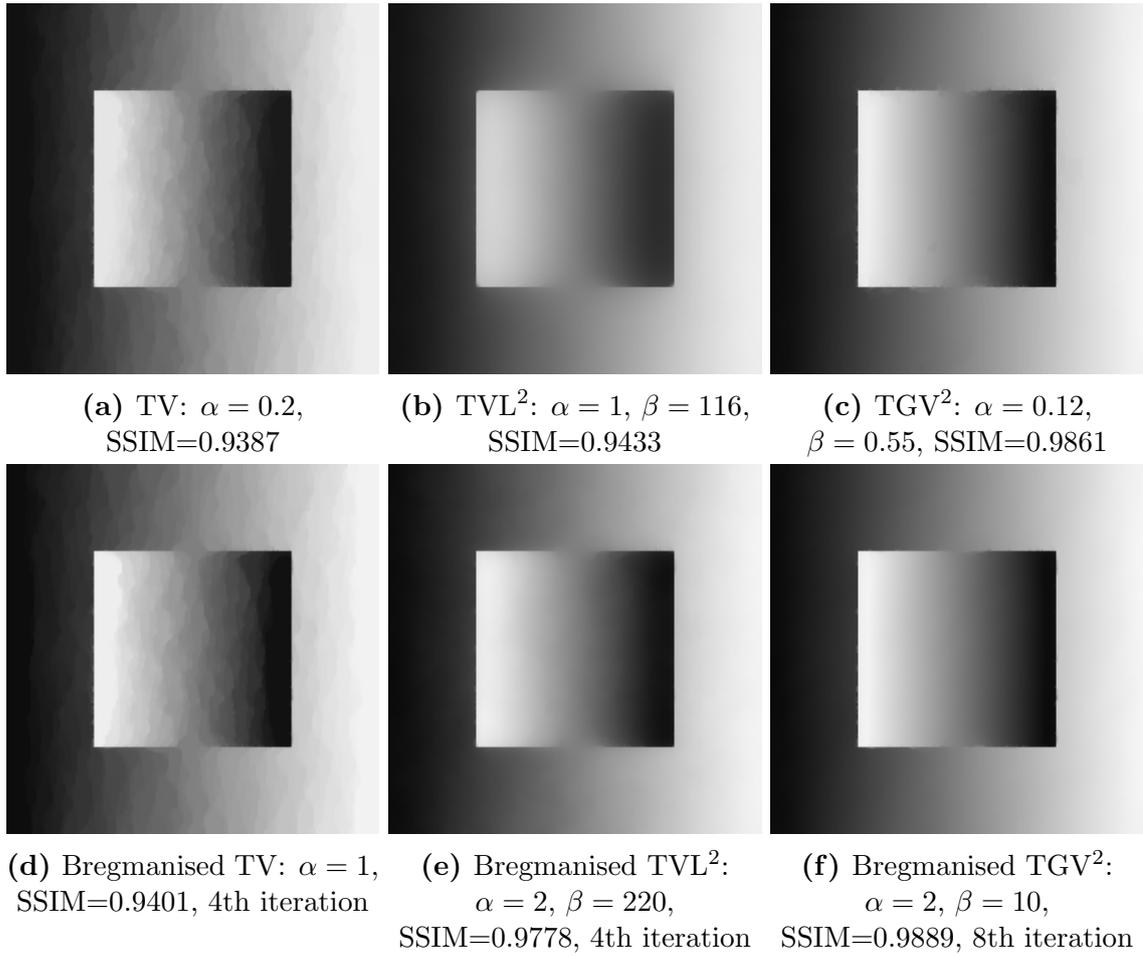


Figure 4.20: First Row: Best reconstruction in terms of SSIM for TV, TVL^2 and TGV^2 .
Second Row: Best reconstruction in terms of SSIM for Bregmanised TV, TVL^2 and TGV^2 .

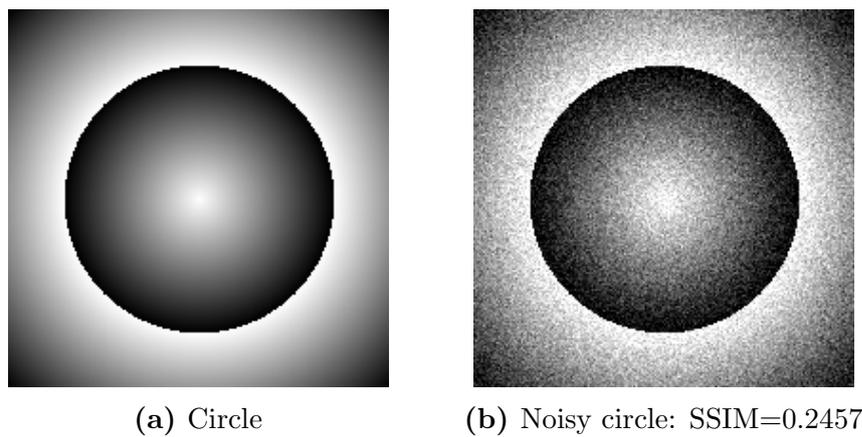


Figure 4.21: Image with symmetric radial structures and its noisy version with $\sigma = 0.01$.

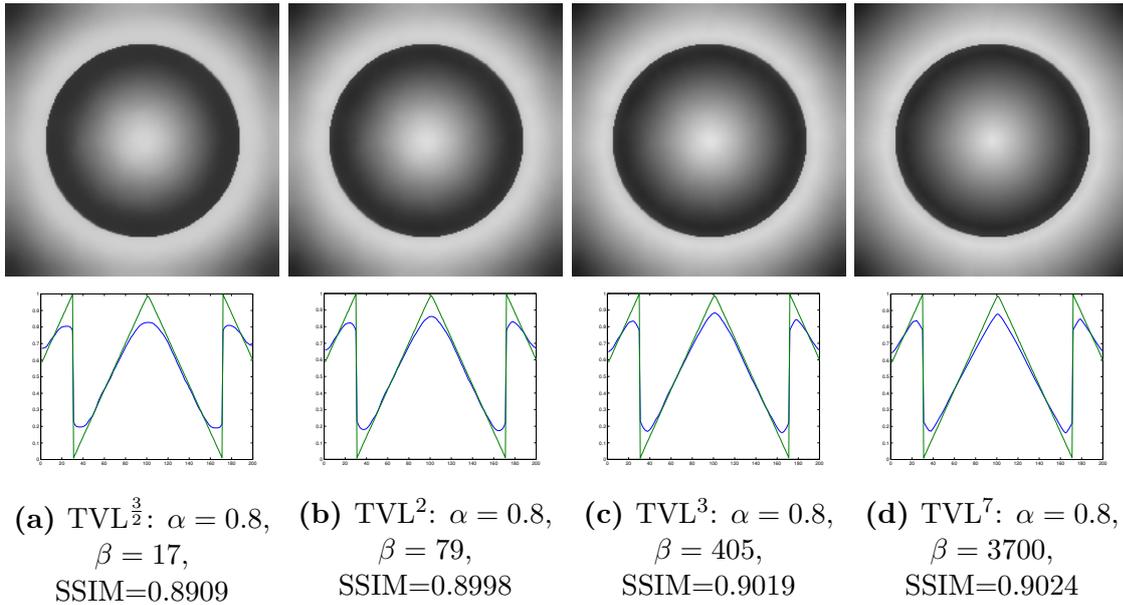


Figure 4.22: Staircasing elimination for $p = \frac{3}{2}, 2, 3$ and 7 . Better preservation of spike-like structures for large values of p . Middle-row comparison of the regularised solution u (blue) and the ground truth (green).

the solutions of (4.54) will not necessarily be the same with the ones of (\mathcal{P}). In fact, we observe that (4.54) cannot always eliminate the staircasing, see Figure 4.25. Even though, we can easily eliminate the staircasing both in the square and in the circle by applying TVL^p regularisation, Figures 4.25b and 4.25d, we cannot obtain equally satisfactory results by solving (4.54). While using the latter we can get rid of the staircasing in the circle, Figure 4.25c, this is not possible for the square, Figure 4.25a, where we observe – after extensive experimentation – that no values of α and β lead to a staircasing elimination. This is analogous to the difference between the TGV^2 and the ICTV regularisers, see also [Mül13, Chapter 4]. However, as we mentioned before, the strength of the formulation (4.54) lies on its ability to efficiently decompose an image into piecewise constant and smooth parts. We show that in Figure 4.26, for the image in Figure 4.25c.

4.6.4 The $p = \infty$ case

Now, we present our results using the TVL^∞ as our regulariser. We follow the same procedure as we did in the finite p case and choose high values of α and β in order to remove the staircasing. For instance, in the square example with fix $\alpha = 5$ and we gradually increase the values of β until we obtain an ROF solution, i.e., $|w| = 0$, see Figure 4.27. We realise two main properties for the reconstructed images. First, we can achieve staircasing-free reconstructions when $\beta = 4, 6, 10 \cdot 10^4$, see also the middle row profiles in Figure 4.27. Moreover, in every reconstruction we observe not only a serious loss

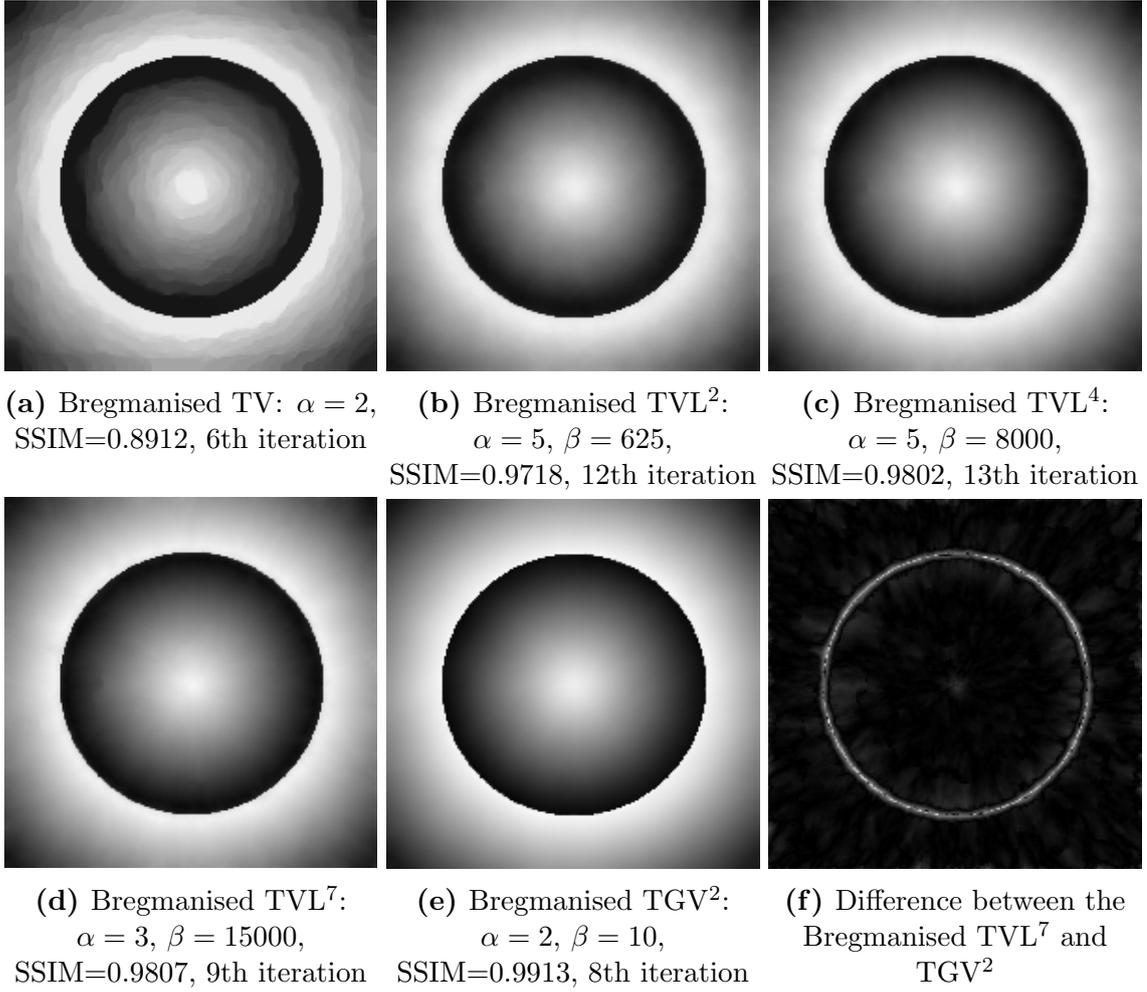


Figure 4.23: Best reconstruction in terms of SSIM for Bregmanised TV, TVL², TVL⁴, TVL⁷ and TGV². The absolute difference between TVL⁷ TGV², see also Figure 4.24.

of contrast but also a significant loss of geometrical information of the square. The reason for this kind of behaviour for the TVL[∞] regulariser has already mentioned in Section 4.4.2. We proved that in the one dimensional setting, see Proposition 4.4.10, a regularised solution u of (\mathcal{P}) behaves as a piecewise affine function with the same gradient, namely $\pm \|w\|_{\infty}$. Although, it is difficult to derive the same conclusion for the two dimensional setting, we expect the same behaviour that is $|w| = \|w\|_{\infty}$ in regions where $u \neq f$. This can be verified numerically as we demonstrate in Figure 4.28b, where we present the image of $|w| = \sqrt{w_1^2 + w_2^2}$. We observe that w is constant a.e. with $|w| = \|w\|_{\infty} = 0.0054$ and the slope of the solution u is the same in regions where $u \neq f$.

Let us focus on the loss of contrast. As before, in order to recover the loss of contrast we employ the Bregmanised version of our model and the regularised solution is shown in Figure 4.28c. One can observe that we recover a perfect reconstruction in the inner

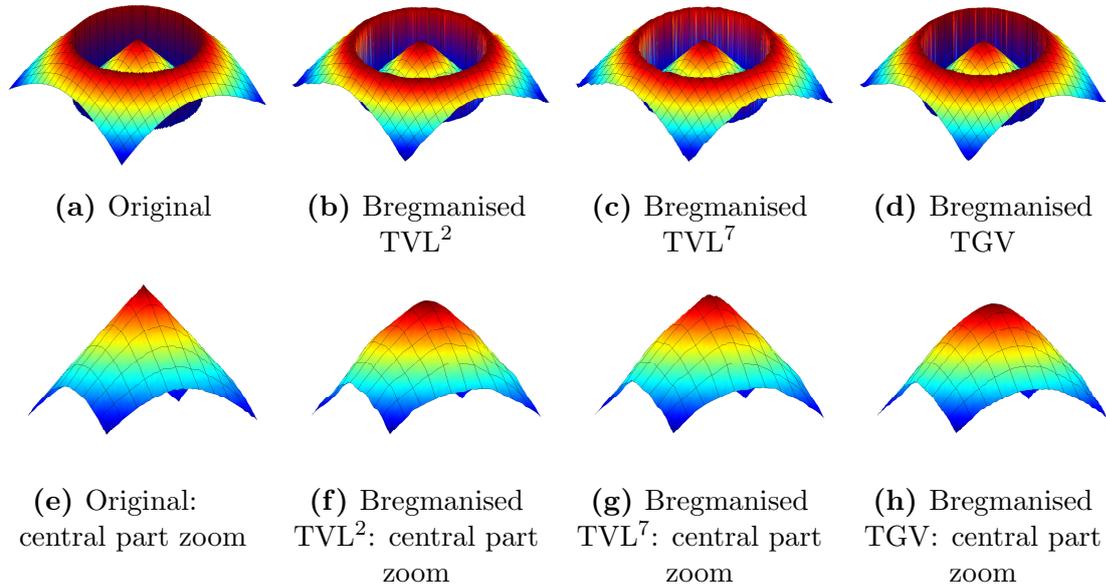


Figure 4.24: Surface plots of the images in Figure 4.23. Notice how high values of p , here for $p = 7$, can preserve the sharp spike in the middle of the image.

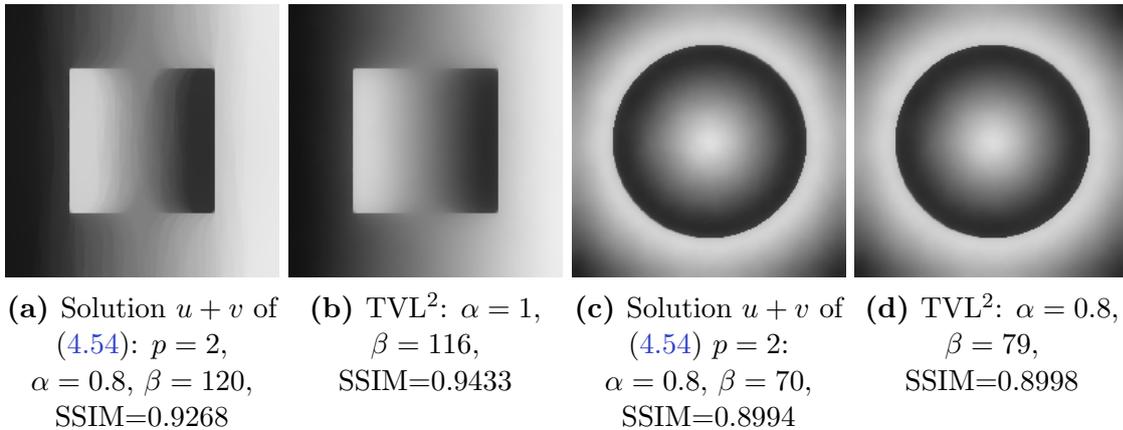


Figure 4.25: Comparison between the model (4.54) for $p = 2$ and TVL^2 : Staircasing cannot be always eliminated with (4.54).

square however we introduce a kind of staircasing in the background region. This is not the standard staircasing that we have encountered so far, see for instance Figure 4.27d. It acts as an affine approximation of the noisy data and we refer to this behaviour with the term *affine staircasing*, see the middle row profiles in Figure 4.28d that correspond to the background region.

For the sake of eliminating the affine staircasing in the background area of the inner square, we need to examine its structure. The square image is constructed with two

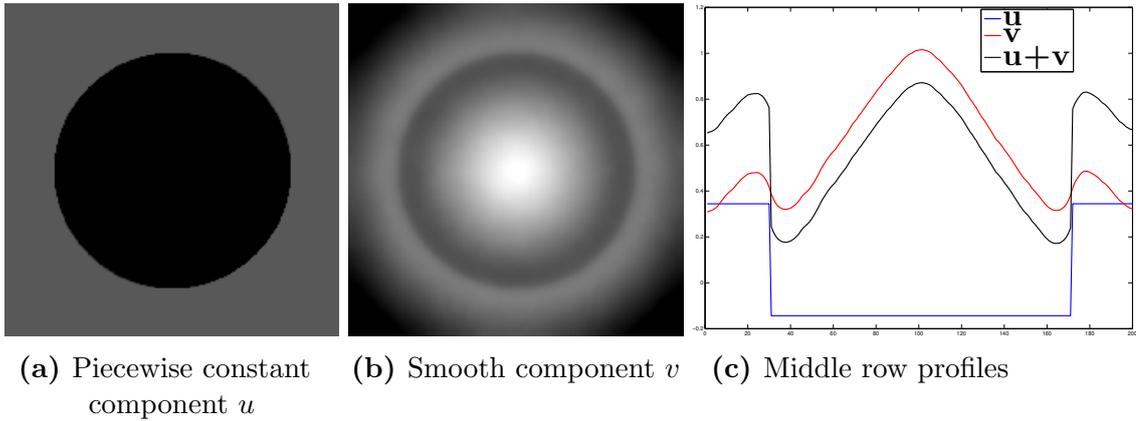


Figure 4.26: Geometric decomposition of the image in Figure 4.25c into a piecewise constant and smooth component, u and v respectively, by solving (4.54) for $p = 2$.

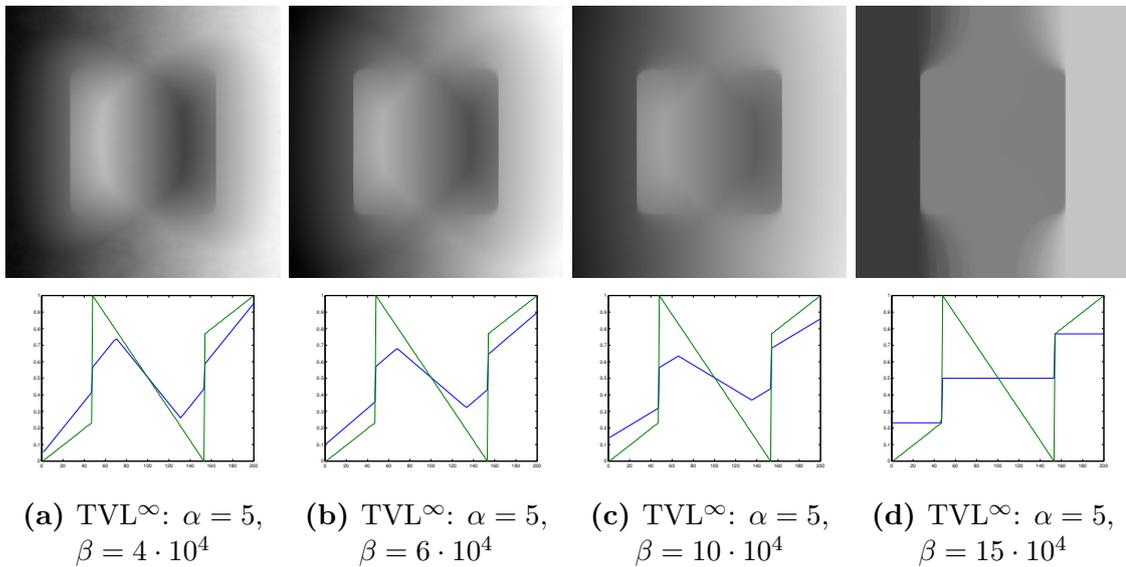


Figure 4.27: TVL^∞ reconstructions with fixed $\alpha = 5$ and different values of β until we reach a ROF solution. Staircasing can be eliminated with a serious loss of contrast and geometrical information. Middle row profiles (blue) compared with ground truth (green).

different slopes one in the inner square and one in the background region, see the green middle row profile in Figure 4.28d. Notice that the slope in the inner square is $s_{in} = \frac{0.2314}{47} = 0.0049$ whereas the slope in the background region is $s_{out} = \frac{1}{100} = 0.01$. As we illustrate in Figure 4.27, we know that the regularising parameter β is responsible to the slope of the regularised solution. In Figures 4.29a, 4.29b, we perform the following experiment. We select two different values of β , meaning that we enforce two different slopes in the solution u of (\mathcal{P}) . For a fixed $\alpha = 0.3$ and proper choices of β we can eliminate the staircasing only in one of these regions. This is equivalent to say that we

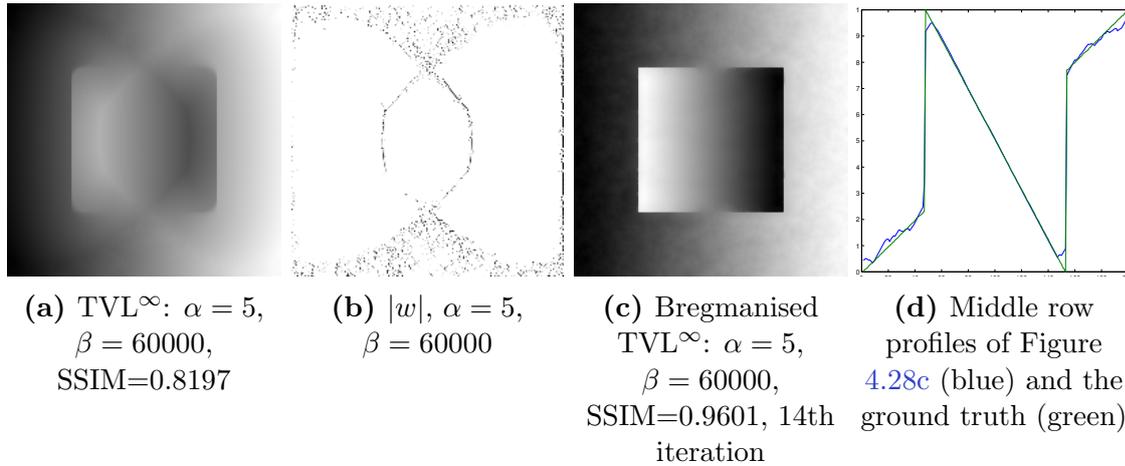


Figure 4.28: The TVL $^\infty$ reconstruction u and its $|w| = \sqrt{w_x^2 + w_y^2}$ solution of (\mathcal{P}) . Bregmanised TVL $^\infty$ reconstruction and the middle row profile comparison (blue) with the ground truth (green).

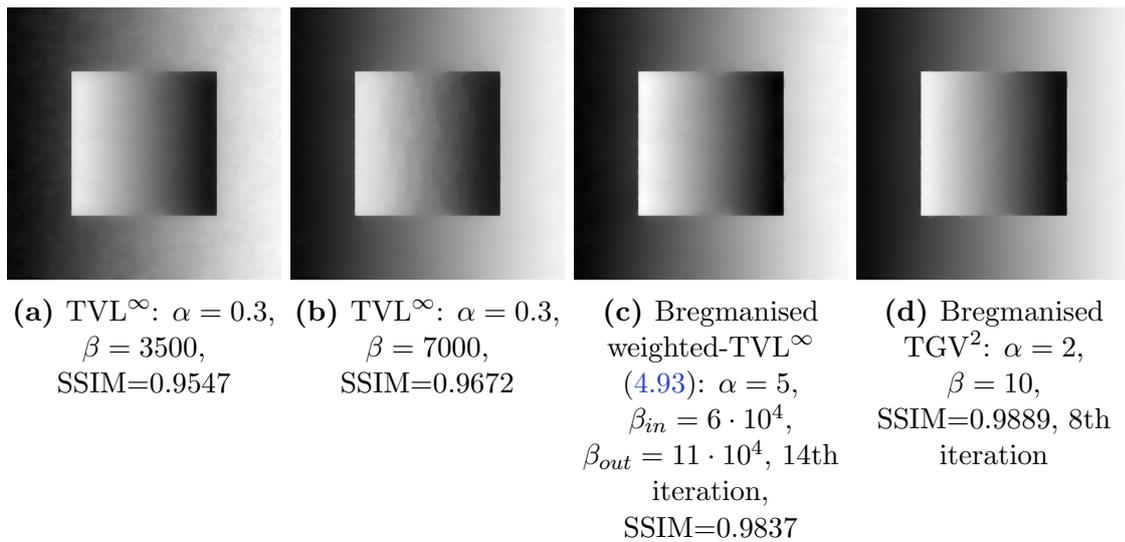


Figure 4.29: TVL $^\infty$ reconstructions with fixed α and different scalar values of β . Bregmanised weighted-TVL $^\infty$ reconstructions with a space dependent parameter β . Bregmanised TGV 2 reconstruction. A suitable choice of $\beta \in \mathbb{R}^{n \times m}$ can achieve similar results to high-order regularisers.

approximate the correct slope of the ground truth for the inner square when $\beta = 3500$ and for the background when $\beta = 7000$ and in these regions $u \neq f$. Notice, that the values of β behave inversely proportional to the slopes s_{in} and s_{out} . That is, we choose $\beta = 3500$ in order to eliminate the staircasing for the inner square which is the half of the value β we use so as to eliminate the staircasing in the background. This behaviour

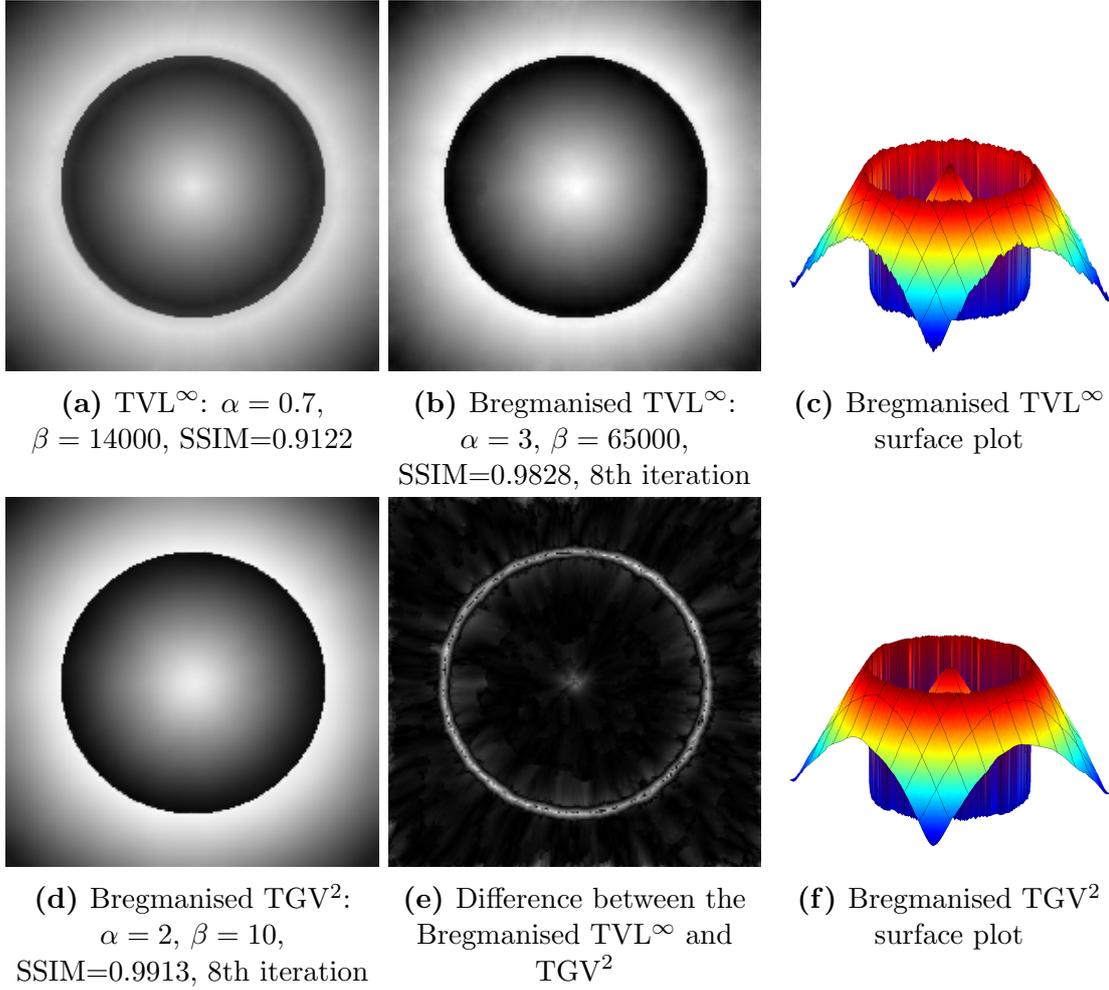


Figure 4.30: TVL^∞ , Bregmanised TVL^∞ and Bregmanised TGV^2 reconstructions for the circle test image. The surface plots indicate that spike-like structures are better preserved with the TVL^∞ regulariser. The reader is suggested to compare the absolute difference of the Bregmanised TVL^∞ and TGV^2 with the Figure 4.23f.

motivates us to treat β as a space dependent parameter that has two different values in these regions. Hence, $\beta \in \mathbb{R}^{n \times m}$ is a matrix multiplied componentwise with $|w|$ and we obtain a weighted- TVL^∞ version of (4.74), that is

$$\min_{u \in \mathbb{R}^{n \times m}} \frac{1}{2} \|f - u\|_2^2 + \alpha \|\nabla u - w\|_1 + \|\beta \circ w\|_\infty, \quad (4.93)$$

where \circ denotes the Hadamard product. Analogously, we consider the Bregmanised approach of the weighted- TVL^∞ where we choose $\beta_{in} = 6 \cdot 10^4$ and $\beta_{out} = 11 \cdot 10^4$ the values of the matrix β for the inner square and the background region respectively. As we observe in Figures 4.29c–4.29d, we obtain a reconstruction that is similar to the Bregmanised TGV^2 both visually and in terms of SSIM.

We perform the same experiments for the circle image. If we examine the middle row profiles that appear in Figure 4.22, we realise that there is no need to treat β as a space dependent parameter since the slope is almost the same in the inner circle and in the background region. We collect all reconstructions regarding the TVL^∞ and the Bregmanised TVL^∞ in Figure 4.30. Firstly, we demonstrate that we can eliminate the staircasing with high values of α and β . Secondly, we can apply the Bregmanised approach (4.92) and obtain similar reconstruction to TGV^2 . In fact, spike-like structures that are present in the origin of the circle are better preserved using the TVL^∞ regulariser, see for instance the surface plots in Figures 4.30c–4.30f as well as the absolute difference in Figure 4.30f. Compared to the finite p case, see Figure 4.23f, we observe that in the $p = \infty$ case, the spike structure in the origin is even better reconstructed.

4.6.5 Open problems

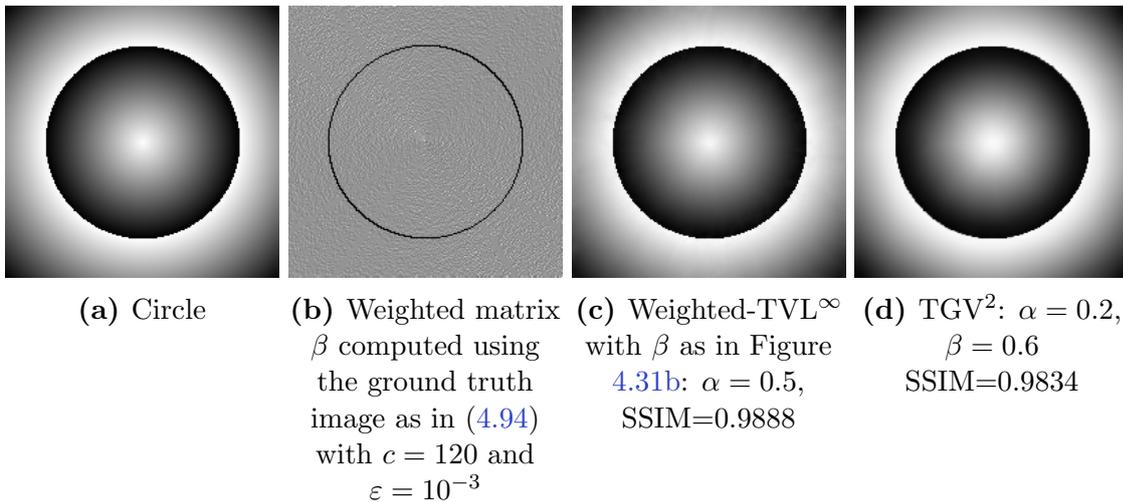


Figure 4.31: Comparison between weighted- TVL^∞ with β as in (4.94) and TGV^2 regularisers. A TVL^∞ reconstruction is visually closer to the ground truth, since TGV^2 tends to smooth spike-like structures, hence SSIM is slightly less.

In general images with complex structures, finding the correct matrix β is not an easy task and remains an open question. However, we would like to emphasise the capabilities of our regulariser compared to high-order methods as TGV^2 . As we verified numerically in the previous section, β has to be selected appropriately so that the slope of the regularised solution is close to the slope of the ground truth image g . In fact, β has to be inversely proportional to the gradient of the ground truth. Hence, one reasonable choice is to define β as

$$\beta = \frac{c}{|Dg| + \epsilon} \quad (4.94)$$

where c is a positive constant and ε a small positive constant. This choice of β can be interpreted as follows: in *flat-constant* regions, i.e., where $|Dg| \rightarrow 0$, $\beta \rightarrow \infty$ meaning that ROF-type solution is enforced which is preferable to reconstruct piecewise constant regions and preserve the edges on the image. Also, a suitable choice of the parameter c based on the values $|Dg|$ can approximate the correct slope for the solution u . In Figures 4.31b-4.31c, we compute the space dependent matrix β using the gradient of the ground truth image g and compare it with the corresponding TGV² reconstruction. Our reconstruction is better compared to TGV² both visually and in terms of the SSIM. In addition, we can avoid considering the Bregmanised approach of our model since a sophisticated choice of β is enough to reduce on its own the loss of contrast and hence the computational cost, in order to arrive to a satisfying reconstruction, is also reduced. In our final numerical experiments, we follow the same procedure to real world images as in the *parrot* and the *ladybug*. The noisy version of parrot is corrupted with Gaussian noise of $\sigma = 0.001$, see Figure 4.32. The image resolution is 200x200 pixels. Since, there is no a priori information about the ground truth image, it is reasonable to apply a *smoothing* to our noisy version and use this version to compute the matrix β . We perform a *Gaussian* filtering using the Matlab's built-in functions *fspecial* and *imfilter* with standard deviation $\sigma = 2$ and window size of 13x13 pixels. The filtered image is denoted by f_σ . The corresponding β is now

$$\beta = \frac{c}{|\nabla f_\sigma| + \varepsilon}. \quad (4.95)$$

First, we observe that due to the smoothing we perform in the noisy image, we have lost all the edge information as it is appeared in Figure 4.32c. This is far more convincing if one compares between the gradient images of the ground truth and the smoothed versions in Figures 4.32d-4.32e. However, we are not interested so much on losing the edges since this can be guaranteed by our proposed regulariser and the selection of the constant c . In fact, a large constant c will result to a similar reconstruction where one employs TV as the regulariser and hence edges will be preserved, see Figure 4.33a. The proof is identical to Proposition 4.4.6 that if

$$c \geq \alpha|\Omega|(\|Df_\sigma\|_\infty + \varepsilon) \quad (4.96)$$

then a ROF-type solution is produced. Simply, notice that

$$\frac{|Df_\sigma| + \varepsilon}{|Df_\sigma| + \varepsilon} \circ |w| \leq \frac{(\|Df_\sigma\|_\infty + \varepsilon)}{c} \|\beta \circ w\|_\infty$$

where all the actions are considered componentwise and ε is a relatively small constant. Certainly is not an accurate threshold but can help us on finding the best constants α and c that produce the highest SSIM. On the other hand, if we use the weighted-TVL[∞] where β is computed via the gradient of the filtered image, we have reduced the staircasing especially in *smooth* regions on the background and in its beak. Now, if we compute β

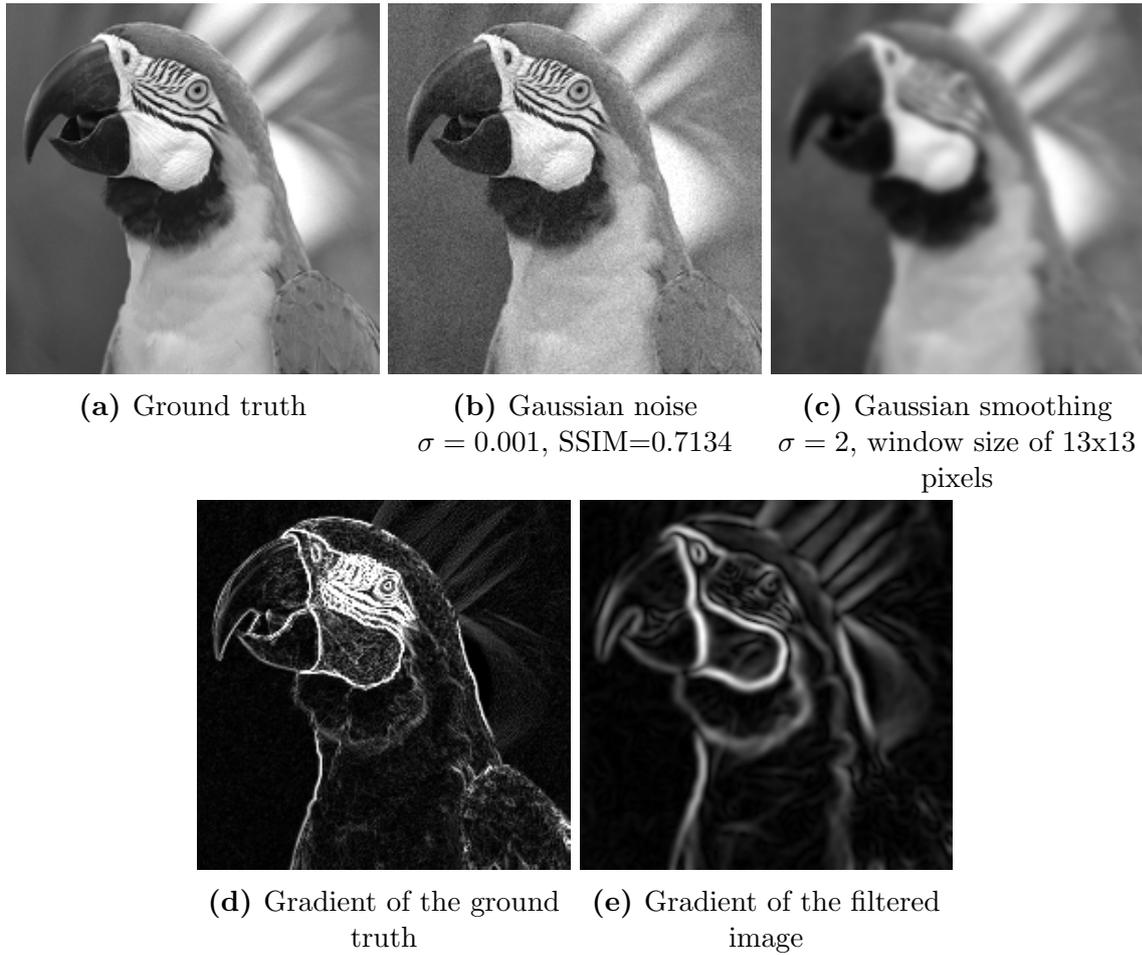


Figure 4.32: Parrot image corrupted with Gaussian noise and the gradients of the filtered and the ground truth images.

based on the gradient of the noiseless image, the solution we obtain is almost perfect and visually close to the original image with a very high SSIM value, compare also the SSIMs of TV, TGV² and TVL[∞] computed via the filtered version reconstructions. We perform similar numerical experiments for the image in Figure 4.34. Its resolution is 200x183 pixels and is contaminated by Gaussian noise of $\sigma = 0.005$. The best reconstructions in terms of SSIM are presented in Figure 4.35.

Let us mention that all the experiments regarding the weighted-TV L[∞] functional are again solved by the split Bregman algorithm performed in (4.93). Similar to (4.88), here we need to solve

$$\min_{w \in (\mathbb{R}^{n \times m})^2} \|\beta \circ w\|_{\infty} + \frac{1}{2} \|w - \eta\|_2^2. \quad (4.97)$$

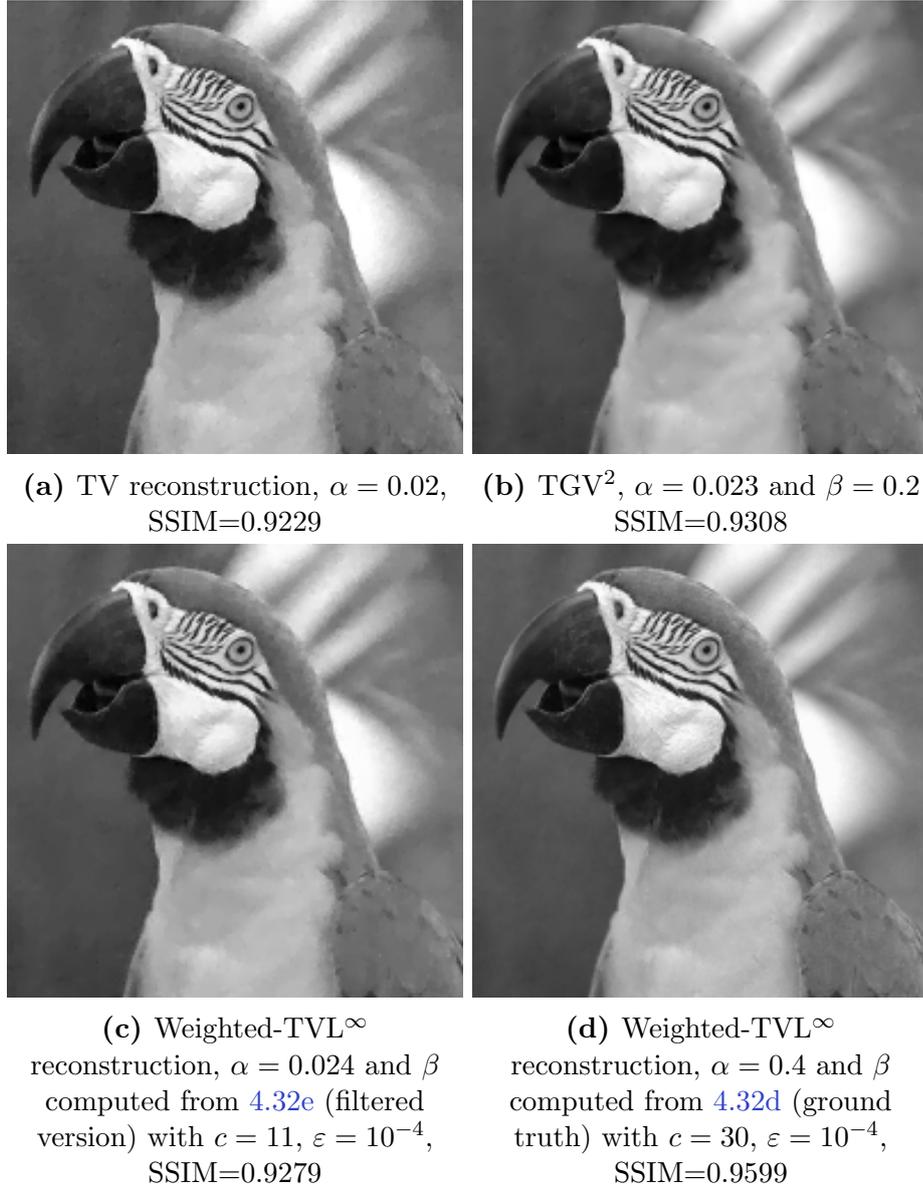


Figure 4.33: Best reconstruction in terms of SSIM using TV, TGV^2 and weighted- TVL^∞ regularisers computed via the filtered and the ground truth image.

This can be rephrased as a projection onto weighted ℓ^1 balls, since

$$\left(\|\beta \circ (\cdot)\|_\infty \right)^*(w) = \mathcal{I}_{\left\{ \left\| \frac{1}{\beta} \circ (\cdot) \right\|_1 \leq 1 \right\}}(w),$$

where

$$\left\| \frac{1}{\beta} \circ w \right\|_1 = \sum_{i,j}^{n,m} \left| \frac{1}{\beta_{i,j}} \circ w_{i,j} \right|.$$

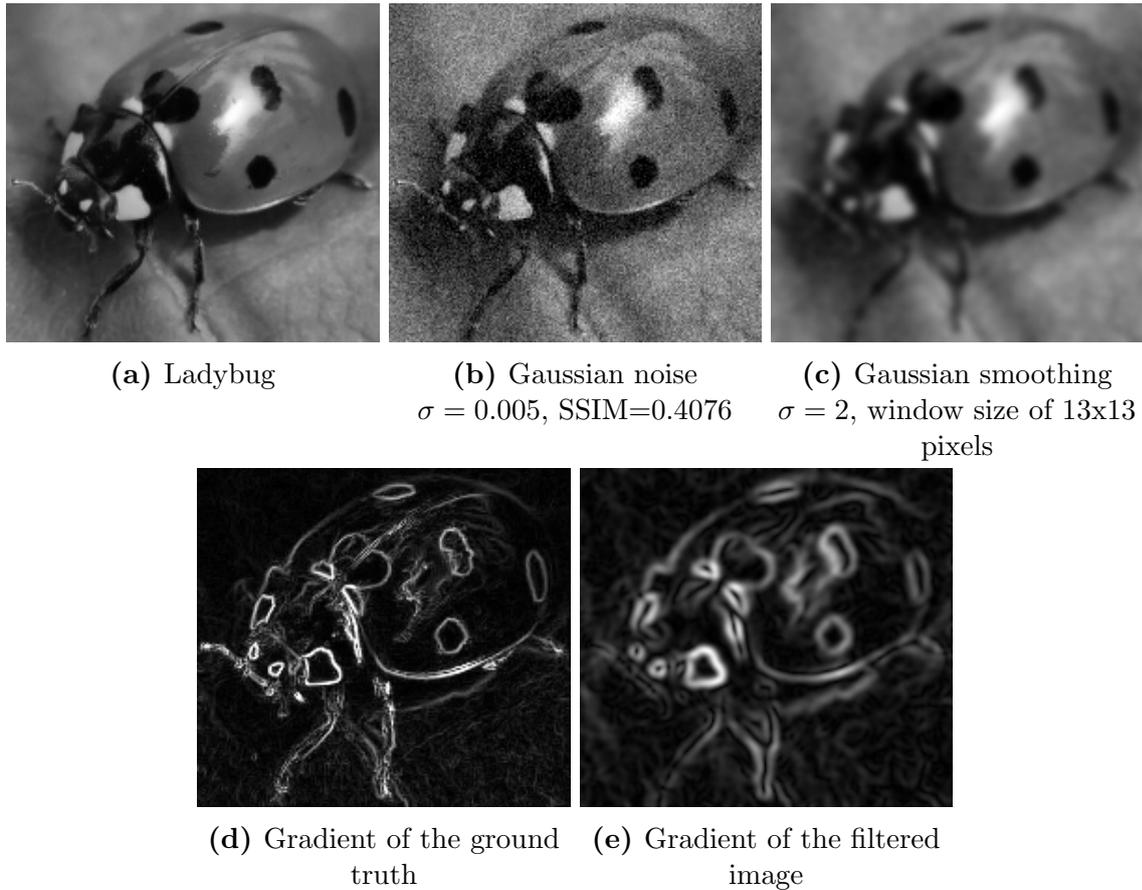


Figure 4.34: Ladybug image corrupted with Gaussian noise and the gradients of the filtered and the ground truth images.

We use the algorithm proposed in [KST11] with a slight modification to solve the projection onto weighted ℓ_1 balls and using (4.89), we obtain the solution of (4.97). Moreover, we compare numerically its computational cost with other regularisers such as TV and TGV^2 . In order to have a fair comparison, we choose to solve all the minimisation models for TV and TGV^2 regularisers using the same numerical approach. For this purpose, we use CVX. In Table 4.2, we collect all the computational (CPU) times for the reconstructions presented in Figure 4.33 and Figure 4.35 with medium *cvx precision*. As we mentioned before, CVX is significantly computational slow and is not suitable for medium-large two dimensional images. Recall that the resolutions for the parrot and ladybug images are 200x200 and 200x183 respectively and the CPU time is affected by the size of the images, see Table 4.2. For large scale problems, MOSEK uses a presolver in order to reduce some constraints and variables for the interior-point method. This is very computationally expensive and for instance images of size close to 512x512 are the limit that MOSEK can handle with a 2GB of memory. However, it provides with very accurate and trustworthy

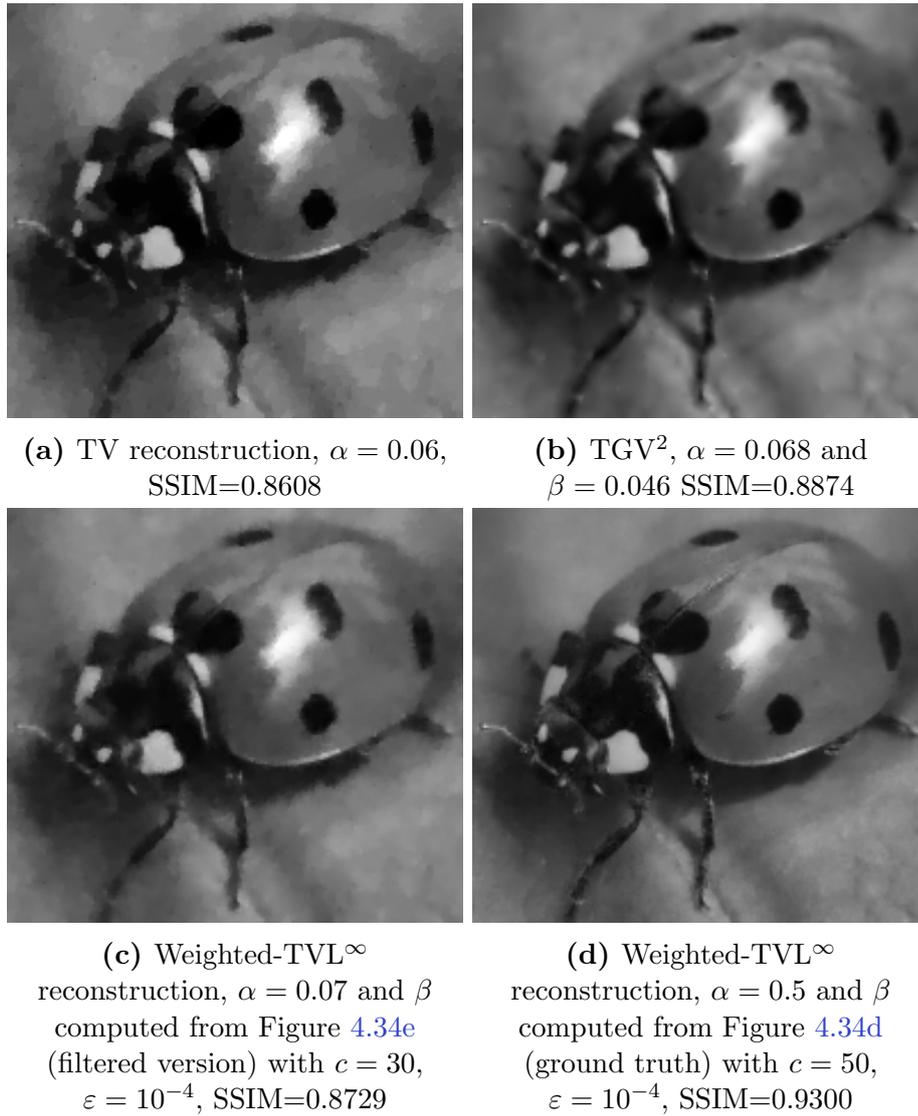


Figure 4.35: Best reconstruction in terms of SSIM using TV, TGV^2 and weighted- TVL^∞ regularisers computed via the filtered and the ground truth image.

reconstructions and furthermore accurate conclusions regarding the computational speed. Clearly, we observe that TGV^2 reconstructions are indeed the slowest compared to TV and weighted- TVL^∞ reconstructions for both the parrot and the ladybug images. The computational time difference between the weighted- TVL^∞ and TGV^2 is over 100 seconds.

Therefore, we believe that by finding a suitable choice of β , first-order TV-related regularisers are in fact sufficient and efficient so as to obtain comparable reconstructions with high-order regularisers. There is still room for improvement for better quality reconstructions as one can observe by comparing between a TGV^2 reconstruction with SSIM=0.9284 and a weighted TVL^∞ reconstruction with SSIM=0.9599 in Figures 4.33b-4.33d respec-

	CPU time with CVX (sec)				
	TV	w-TVL $^\infty$ (filtered)	w-TVL $^\infty$ (ground truth)	TGV 2	Size
Parrot	92.1	235.48	215.74	357.73	200x200
Ladybug	79.66	189.26	186.78	284.16	200x183

Table 4.2: Computational times using CVX under MOSEK for the reconstructions presented in Figures 4.33–4.35. The *cvx precision* is medium. The implementations were done using MATLAB (2013) with 2.4 GHz Intel Core 2 Duo and 2 GB of memory.

tively. This difference in the SSIM value is mainly due to the fact that TGV 2 regulariser besides eliminating the noise and the staircasing it also oversmooths texture structures specially in the wing of the parrot. Hence, a correct choice of β provides not only a suitable gradient information to the final solution but also enforce a better treatment to texture regions, a significant image property that is responsible to obtain high-quality image reconstructions.

Chapter 5

Conclusions

In this thesis we have presented novel image reconstruction methods based on the total variation regularisation that was introduced in [ROF92]. These kind of methods have been proposed in the context of general imaging denoising as well as for reconstructing tomographic images obtained via Positron emission tomography (PET).

We begin, in Chapter 2, with an introduction and a review of basic mathematical tools needed in this dissertation. We emphasise mainly the space of functions of bounded variation and how it can be beneficial in imaging applications.

In Chapter 3, we present our variational method related to medical imaging and PET reconstruction. It is a combined regularisation between the image and the sinogram space that one encounters during a PET scan. Our explicit reconstruction of total variation regularisation, directly on the sinogram space, provides us with a new insight on how PET reconstruction could be improved and in which cases. This is focused on enhancing object boundaries and especially for images where thin structures are present. We illustrate this behaviour with an extensive analysis at both theoretical and numerical levels where we realise that one can achieve a significant improvement on reconstructing tomographic images governed by thin and elongated structures. This property is illustrated for simple phantoms as well as realistic PET phantoms that depict the activity of the human heart as the XCAT cardiac-torso phantom.

In Chapter 4, we begin with a brief review of first and second order TV-based regularisation methods such as the classical ROF model, the infimal convolution regularisation proposed in [CL97], and the second order total generalised variation proposed in [BKP10]. In order to remove the staircasing that TV usually introduces, one common and successful path is to enforce high-order derivatives to the variational model with an additional computational cost. However, this is not always the case as we demonstrate in Chapter 4. We propose a family of first-order infimal convolution regularisation TVL^p functionals of BV and L^p spaces where $1 < p \leq \infty$. We examine the type of solutions that TVL^p promotes depending on its regularising parameters and the value of p . In addition, we

show that our regulariser is either equivalent to the ROF regularisation when certain conditions are fulfilled or to the Huber-TV regularisation, see [Hub64], for $p = 2$. However, the most important outcome of this analysis is in the case of $p = \infty$. We acquire piecewise affine structures, where to the best of our knowledge this is a novel result regarding first-order methods. The reconstructions are not only comparable with those obtained from high-order methods such as TGV² but also in certain cases are visually better e.g. for spike-like structures. Finally, we state some open questions related to the weighted-TVL[∞] regulariser and how β can be tuned appropriately so as to recover an even better quality image.

Appendix A

Exact solutions of TVL^∞ : Piecewise affine function

In this section, we present our analytical computations of $\text{TVL}_{\alpha,\beta}^\infty$ regulariser that are omitted from Chapter 4 in Section 4.4.3 for the piecewise affine function g with one discontinuity. Recall that

$$g(x) = \begin{cases} \lambda x, & , x \in (-L, 0], \\ \lambda x + h, & , x \in (0, L), \end{cases}$$

with $\lambda \neq 0$ and the corresponding optimality conditions for $p = \infty$ that a solution of (\mathcal{P}) should satisfy for $\phi \in \text{H}_0^1(\Omega)$ are the following:

$$\begin{aligned} \phi' &= u - g, \\ \phi &\in \alpha \text{Sgn}(Du - w), \\ \phi &\in \text{L}^1(\Omega), \|\phi\|_{\text{L}^1(\Omega)} \leq \beta, \quad \text{if } w = 0, \\ \phi &\in \text{L}^1(\Omega), \langle \phi, w \rangle = \beta \|w\|_\infty, \|\phi\|_{\text{L}^1(\Omega)} \leq \beta, \quad \text{if } w \neq 0. \end{aligned} \tag{A.1}$$

We begin with the computations of the three ROF solutions where $w = 0$ in $\Omega = (-L, L)$. For a solution that is constant and equal to the mean value i.e., $u = \frac{1}{\Omega} \int_\Omega g(x) dx = \frac{h}{2}$, we require that ϕ is a quadratic function which vanishes at $x = \pm L$. Indeed, for every $x \in (-L, 0)$, we have that $\phi'(x) = \frac{h}{2} - \lambda x$ which by the boundary conditions and its symmetry can be expressed as $\phi(x) = \frac{\lambda}{2}(L^2 - x^2) + \frac{h}{2}(L - |x|)$ for every $x \in \Omega$. Moreover,

$$\begin{aligned} \|\phi\|_{\text{L}^1(\Omega)} \leq \beta &\Leftrightarrow \beta \geq \frac{2\lambda L^3}{3} + \frac{hL^2}{2}, \\ \phi(0) < \alpha &\Leftrightarrow \frac{\lambda L^2 + hL}{2} < \alpha, \end{aligned} \tag{A.2}$$

hence we obtain $(\text{ROF}_{\text{III}} : \text{affine})$.

On the other hand for a piecewise constant solution u with $u(x) \neq g(x)$ in Ω , we obtain $\phi(x) = -c_1|x| - \frac{\lambda}{2}x^2 + c_2$. Since $\phi(-L) = 0$ and $\phi(0) = \alpha$ due to the discontinuity at $x = 0$, we conclude that $c_1 = \frac{2\alpha - \lambda L^2}{2L}$ and $c_2 = \alpha$. In addition, due to the structure of the solution, we require that

$$\begin{aligned} u(0) < \frac{h}{2} &\Leftrightarrow \alpha < \frac{hL + \lambda L^2}{2}, \\ u(-L) = c_1 \geq 0 &\Leftrightarrow \alpha \geq \frac{hL}{2}, \\ \|\phi\|_{L^1(\Omega)} \leq \beta &\Leftrightarrow \beta \geq \frac{\lambda L^3}{6} + \alpha L, \end{aligned} \tag{A.3}$$

which coincides with (ROF_{II} : affine).

For a piecewise constant function in $I_1 = (-L, -l_1)$, $I_2 = (l_1, L)$, $0 \leq l_1 < L$ and $u(x) = g(x)$ otherwise, we have that $\phi(x) = -c_1|x| - \frac{\lambda}{2}x^2 + c_2$ with $\phi(-l_1) = \alpha$ and $\phi(-L) = 0$. Then,

$$\begin{aligned} c_1 &= -\frac{\alpha}{l_1 - L} - \frac{\lambda}{2}(l_1 + L), \\ c_2 &= -\frac{\alpha L}{l_1 - L} - \frac{\lambda l_1 L}{2}. \end{aligned} \tag{A.4}$$

In order to compute l_1 , we require also that $\phi'(-l_1) = 0 \Leftrightarrow c_1 = -\lambda l_1$ and by (A.4) one needs to solve

$$\lambda l_1^3 - 2\lambda L l_1^2 - (2L^2 - 2\alpha)l_1 = 0. \tag{A.5}$$

Either $l_1 = 0$ or $l_1 = L - \sqrt{\frac{2\alpha}{\lambda}}$. Notice that for the case $l_1 = 0$, $c_1 = 0$ we obtain a (ROF_{II} : affine) solution with $u(x) = 0$ for $x \in (-L, 0)$ and $u(x) = h$ for $x \in (0, L)$. Hence,

$$\begin{aligned} l_1 \geq 0 &\Leftrightarrow \alpha \leq \frac{\lambda L^2}{2}, \\ \|\phi\|_{L^1(\Omega)} \leq \beta &\Leftrightarrow \beta \geq 2\alpha L - \frac{2\alpha}{3} \sqrt{\frac{2\alpha}{\lambda}}, \end{aligned} \tag{A.6}$$

which is exactly (ROF_I : affine).

Now, for the non-ROF solutions where $w \neq 0$ in Ω , we follow the same strategy using also Propositions 4.4.7 and 4.4.10. For a piecewise affine solution, such that $u(x) = g(x)$ in $I = (-L, -l_2) \cup (l_2, L)$, we have that

$$\phi(x) = \begin{cases} 0, & x \in (-L, l_2), \\ (c_1 - \lambda)\frac{x^2}{2} - c_2|x| + c_3, & x \in (-l_2, l_2), \\ 0, & x \in (l_2, L). \end{cases} \tag{A.7}$$

Since the solution preserve the discontinuity at $x = 0$, we have that $c_3 = \alpha$ and c_1, c_2 can

be found from

$$\begin{aligned} \phi(-l_2) = 0, \\ \phi'(-l_2) = 0, \end{aligned} \Leftrightarrow \begin{cases} c_1 = \frac{2\alpha}{l_2} + \lambda, \\ c_2 = \frac{2\alpha}{l_2}. \end{cases} \quad (\text{A.8})$$

In order to compute l_2 , we define

$$w(x) = \begin{cases} \lambda, & x \in (-L, l_2), \\ c_1, & x \in (-l_2, l_2), \\ \lambda, & x \in (l_2, L), \end{cases} \quad (\text{A.9})$$

and $\langle \phi, w \rangle = \beta \|w\|_\infty$ implies that $l_2 = \frac{3\beta}{2\alpha}$. Notice that w is not unique in general and an equivalent expression is still valid, e.g. $w = 0$ in $I = (-L, -l_2) \cup (l_2, L)$ namely when $\phi = 0$. Finally, we require that

$$\begin{aligned} l_2 < L &\Leftrightarrow \beta \leq \frac{2\alpha L}{3}, \\ u(0) < \frac{h}{2} &\Leftrightarrow c_2 < \frac{h}{2} \Leftrightarrow \beta > \frac{8\alpha^2}{3h}, \end{aligned} \quad (\text{A.10})$$

which results to [\(TVL_I[∞] : affine\)](#). A solution with the same structure but without any discontinuity in Ω has $u(0) = \frac{h}{2}$, and therefore

$$\phi(x) = \begin{cases} 0, & x \in (-L, l_3), \\ (c_1 - \lambda)\frac{x^2}{2} - \frac{h}{2}|x| + c_3, & x \in (-l_3, l_3), \\ 0, & x \in (l_3, L), \end{cases} \quad , \quad w(x) = \begin{cases} \lambda, & x \in (-L, l_3), \\ \frac{h}{2l_3} + \lambda, & x \in (-l_3, l_3), \\ \lambda, & x \in (l_3, L). \end{cases} \quad (\text{A.11})$$

Again, we can compute c_1, c_3 and l_3 by

$$\begin{aligned} \phi(-l_3) = 0, \\ \phi'(-l_3) = 0, \\ \langle \phi, w \rangle = \beta \|w\|_\infty \end{aligned} \Leftrightarrow \begin{cases} c_1 = \frac{h}{2l_3} + \lambda, \\ c_3 = \frac{hl_3}{4}, \\ l_3 = \sqrt{\frac{6\beta}{h}}. \end{cases} \quad (\text{A.12})$$

Then, $l_3 \leq L \Leftrightarrow \beta \leq \frac{hL^2}{6}$ and $\phi(0) < \alpha \Leftrightarrow \beta < \frac{8\alpha^2}{3h}$ which are the necessary and sufficient condition for [\(TVL_{II}[∞] : affine\)](#).

For a discontinuous solution such that $u(x) \neq g(x)$ in Ω we have that $\phi(x) = (c_1 - \lambda)\frac{x^2}{2} - c_2|x| + c_3$ and $w(x) = c_1$. Then, we find that

$$\begin{aligned} \phi(0) = \alpha, \\ \phi(-L) = 0, \\ \langle \phi, w \rangle = \beta \|w\|_\infty \end{aligned} \Leftrightarrow \begin{cases} c_1 = \frac{6(\alpha L - \beta)}{L^3} + \lambda, \\ c_2 = \frac{4\alpha L - 3\beta}{L^2}, \\ c_3 = \alpha. \end{cases} \quad (\text{A.13})$$

Finally, we require that

$$\begin{aligned}
 c_1 > 0 &\Leftrightarrow \beta < \alpha L + \frac{\lambda L^3}{6}, \\
 u(0) < \frac{h}{2} &\Leftrightarrow c_2 < \frac{h}{2} \Leftrightarrow \beta > \frac{4\alpha L}{3} - \frac{hL^2}{6}, \\
 \phi'(-L) > 0 &\Leftrightarrow \beta > \frac{2\alpha L}{3}, \\
 c_2 > 0 &\Leftrightarrow \beta < \frac{4\alpha L}{3},
 \end{aligned} \tag{A.14}$$

which is equivalent to $(\text{TVL}_{\text{III}}^\infty : \text{affine})$. Certainly, we have similar results for a continuous solution that $u(x) \neq g(x)$ in Ω . Then, $\phi(x) = (c_1 - \lambda)\frac{x^2}{2} - \frac{h}{2}|x| + c_3$ with $w(x) = c_1$ and one has

$$\begin{aligned}
 \phi(-L) = 0, \\
 \langle \phi, w \rangle = \beta \|w\|_\infty
 \end{aligned}
 \Leftrightarrow
 \begin{cases}
 c_1 = \frac{3(\frac{hL^2}{2} - \beta)}{2L^3} + \lambda, \\
 c_3 = \frac{hL}{8} + \frac{3\beta}{4L}
 \end{cases} \tag{A.15}$$

where the conditions

$$\begin{aligned}
 \phi(0) < \alpha &\Leftrightarrow c_3 < \alpha \Leftrightarrow \beta < \frac{4\alpha L}{3} - \frac{hL^2}{6}, \\
 c_1 > 0 &\Leftrightarrow \beta < \frac{hL^2}{2} + \frac{2\lambda L^3}{3}, \\
 \phi'(-L) > 0 &\Leftrightarrow \beta > \frac{hL^2}{6}
 \end{aligned} \tag{A.16}$$

provide us with the same solution as $(\text{TVL}_{\text{V}}^\infty : \text{affine})$.

Our final solution is piecewise affine where $u(x) = g(x)$ in $I = (-l_4, l_4)$ and can be seen as the complementary case of $(\text{TVL}_{\text{I}}^\infty : \text{affine})$. Therefore, one has that $w(x) = c_1$ and

$$\phi(x) = \begin{cases} (c_1 - \lambda)\frac{x^2}{2} + c_2x + c_3, \\ c_4, \\ (c_1 - \lambda)\frac{x^2}{2} - c_2x + c_3. \end{cases} \tag{A.17}$$

We can compute c_i , $i = 1, 2, 3, 4$ and l_4 by the conditions above

$$\begin{aligned}
 \phi(0) = \alpha, \\
 \phi(-L) = 0, \\
 \phi(-l_4) = \alpha, \\
 \phi'(-l_4) = 0, \\
 \langle \phi, w \rangle = \beta \|w\|_\infty
 \end{aligned}
 \Leftrightarrow
 \begin{cases}
 c_4 = \alpha, \\
 c_1 = \lambda - \frac{2\alpha}{(L-l_4)^2}, \\
 c_3 = \alpha - \frac{\alpha l_4^2}{(L-l_4)^2}, \\
 c_2 = -\frac{2\alpha l_4}{(L-l_4)^2}, \\
 l_4 = \frac{3\beta}{2\alpha} - 2L
 \end{cases} \tag{A.18}$$

and conclude to (TVL_{IV}[∞] : affine) by

$$\begin{aligned}c_1 > 0 &\Leftrightarrow \beta < 2\alpha L - \frac{2\alpha}{3} \sqrt{\frac{2\alpha}{\lambda}}, \\c_2 \leq 0 &\Leftrightarrow l_4 \geq 0 \Leftrightarrow \beta \geq \frac{4\alpha L}{3}.\end{aligned}\tag{A.19}$$

Bibliography

- [AA05] G. Aubert and J.-F. Aujol, *Modeling Very Oscillating Signals. Application to Image Processing*, Applied Mathematics and Optimization **51** (2005), no. 2, 163–182, <http://dx.doi.org/10.1007/s00245-004-0812-z>. (Cited on pages 48 and 143).
- [AABFC05] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle, *Image Decomposition into a Bounded Variation Component and an Oscillating Component*, Journal of Mathematical Imaging and Vision **22** (2005), no. 1, 71–88, <http://dx.doi.org/10.1007/s10851-005-4783-8>. (Cited on page 107).
- [AB86] H. Attouch and H. Brezis, *Duality for the sum of convex functions in general Banach spaces*, North-Holland Mathematical Library **34** (1986), 125–133. (Cited on pages 49, 111, and 122).
- [ABM14] H. Attouch, G. Buttazzo, and G. Michaille, *Variational Analysis in Sobolev and BV Spaces*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2014, <http://dx.doi.org/10.1137/1.9781611973488>. (Cited on pages 114 and 121).
- [ABT08] R. Abraham, M. Bergounioux, and E. Trélat, *A Penalization Approach for Tomographic Reconstruction of Binary Axially Symmetric Objects*, Applied Mathematics and Optimization **58** (2008), no. 3, 345–371, <http://dx.doi.org/10.1007/s00245-008-9039-8>. (Cited on page 66).
- [AF03] R. A. Adams and J. F. Fournier, *Sobolev spaces, 2nd edition*, Academic Press, 2003. (Cited on page 38).
- [AFP00] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, USA, 2000. (Cited on pages 34 and 127).
- [AGCO] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, *Structure-Texture Image Decomposition Modeling, Algorithms, and Parameter Selection*, International

- Journal of Computer Vision **67**, no. 1, 111–136, <http://dx.doi.org/10.1007/s11263-006-4331-z>. (Cited on page 107).
- [AHH06] U.M. Ascher, E. Haber, and H. Huang, *On Effective Methods for Implicit Piecewise Smooth Surface Recovery*, SIAM Journal on Scientific Computing **28** (2006), no. 1, 339–358, <http://dx.doi.org/10.1137/040617261>. (Cited on page 116).
- [AV94] R. Acar and C. R. Vogel, *Analysis of Bounded Variation Penalty Methods for Ill-Posed Problems*, Inverse Problems **10** (1994), no. 6, 1217–1229, <http://dx.doi.org/10.1088/0266-5611/10/6/003>. (Cited on pages 19, 68, 70, and 113).
- [BB03] J. Bourgain and H. Brezis, *On the Equation $\operatorname{div} Y = f$ and Application to Control of Phases*, Journal of the American Mathematical Society **16** (2003), no. 2, pp. 393–426, <http://www.jstor.org/stable/30041437>. (Cited on page 48).
- [BBBM13] M. Benning, C. Brune, M. Burger, and J. Müller, *Higher-order TV methods – Enhancement via Bregman iteration*, Journal of Scientific Computing **54** (2013), no. 2-3, 269–310, <http://dx.doi.org/10.1007/s10915-012-9650-3>. (Cited on pages 110 and 159).
- [BBS⁺09] C. Brune, M. Burger, A. Sawatzky, T. Kösters, and F. Wübbeling, *Forward-Backward EM-TV methods for inverse problems with Poisson noise*, Preprint (2009), 5133 – 5137. (Cited on pages 24, 52, 63, and 64).
- [BC11] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer New York, 2011, <http://dx.doi.org/10.1007/978-1-4419-9467-7>. (Cited on pages 44 and 49).
- [BFK09] P. E. Barbano, A. S. Fokas, and G. A. Kastis, *Analytical reconstructions for PET and SPECT employing L^1 -denoising*, Digital Signal Processing, 2009 16th International Conference (2009), 1–5, <http://dx.doi.org/10.1109/ICDSP.2009.5201187>. (Cited on page 64).
- [BFS11] P. E. Barbano, A. Fokas, and C.-B. Schönlieb, *Alternating regularisation in measurement- and image space for PET reconstruction*, Proceeding International conference on Sampling Theory and Applications (2011), 1–4. (Cited on pages 9 and 63).

- [BKP10] K. Bredies, K. Kunisch, and T. Pock, *Total Generalized Variation*, SIAM Journal on Imaging Sciences **3** (2010), no. 3, 492–526, <http://dx.doi.org/10.1137/090769521>. (Cited on pages 27, 107, 110, and 175).
- [BKV12] K. Bredies, K. Kunisch, and T. Valkonen, *Properties of L^1 -TGV²: The one-dimensional case*, Journal of Mathematical Analysis and Applications (2012), <http://dx.doi.org/10.1016/j.jmaa.2012.08.053>. (Cited on pages 120, 122, 123, and 127).
- [BLZ08] M. Bertero, H. Lanteri, and L. Zanni, *Iterative image reconstruction: a point of view*, Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT), (CRM Series vol 7) ed Y Censor, M Jiang and A K Louis (2008), 37–63. (Cited on pages 24 and 52).
- [BMPS14] M. Burger, J. Müller, E. Papoutsellis, and C.-B. Schönlieb, *Total variation regularization in measurement and image space for PET reconstruction*, Inverse Problems **30** (2014), no. 10, 105003, <http://stacks.iop.org/0266-5611/30/i=10/a=105003>. (Cited on pages 9, 25, and 51).
- [BO04] M. Burger and S. Osher, *Convergence rates of convex variational regularization*, Inverse Problems **20** (2004), 1411–1420, <http://dx.doi.org/10.1088/0266-5611/20/5/005>. (Cited on pages 73 and 76).
- [BO13] M. Burger and S. Osher, *A Guide to the TV Zoo: Level Set and PDE Based Reconstruction Methods in Imaging*, Springer, 2013. (Cited on pages 48 and 141).
- [Bor98] R. Boris, *The Calderon reproducing formula, windowed X-ray transforms, and radon transforms in L^p -spaces*, Journal of Fourier Analysis and Applications **4** (1998), no. 2, 175–197, <http://dx.doi.org/10.1007/BF02475988>. (Cited on page 58).
- [BP10] M. Bergounioux and L. Piffet, *A Second-Order Model for Image Denoising*, Set-Valued and Variational Analysis **18** (2010), no. 3-4, 277–306, <http://dx.doi.org/10.1007/s11228-010-0156-6>. (Cited on page 107).
- [BPPS15a] M. Burger, K. Papafitsoros, E. Papoutsellis, and C.-B. Schönlieb, *Infimal convolution regularisation functionals of BV and L^p spaces, Part I: The finite p case*, Journal of Mathematical Imaging and Vision, (Accepted) (2015), <http://arxiv.org/abs/1504.01956>. (Cited on pages 9, 25, and 105).
- [BPPS15b] ———, *Infimal convolution regularisation functionals of BV and L^p Spaces. The case $p = \infty$* , submitted (2015), <http://arxiv.org/abs/1510.09032>. (Cited on pages 9 and 105).

- [Bre67] L. M. Bregman, *The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming*, USSR Computational Mathematics and Mathematical Physics **7** (1967), no. 3, 200–217, <http://www.sciencedirect.com/science/article/pii/S0041555367900407>. (Cited on page 73).
- [Bre11] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer-Verlag New York, 2011, <http://dx.doi.org/10.1007/978-0-387-70914-7>. (Cited on pages 38, 45, 120, and 141).
- [BSB10] C. Brune, A. Sawatzky, and M. Burger, *Primal and Dual Bregman Methods with Application to Optical Nanoscopy*, International Journal of Computer Vision (2010), 211–229, <http://dx.doi.org/10.1007/s11263-010-0339-5>. (Cited on page 63).
- [BT10] M. Bergounioux and E. Trélat, *A variational method using fractional order Hilbert spaces for tomographic reconstruction of blurred and noised binary images*, Journal of Functional Analysis **259** (2010), no. 9, 2296 – 2332, <http://dx.doi.org/10.1016/j.jfa.2010.05.016>. (Cited on page 66).
- [CCC⁺10] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock, *An introduction to total variation for image analysis*, Theoretical Foundations and Numerical Methods for Sparse Recovery, De Gruyter, 2010. (Cited on pages 48 and 141).
- [CCDIR⁺15] L. Calatroni, C. Cao, J. C. De los Reyes, C.B. Schönlieb, and Valkonen T., *Bilevel approaches for learning of variational imaging models*, arXiv preprint (2015), <http://arxiv.org/abs/1505.02120>. (Cited on page 24).
- [CCN07] V. Caselles, A. Chambolle, and M. Novaga, *The Discontinuity Set of Solutions of the TV Denoising Problem and Some Extensions*, Multiscale Modeling & Simulation **6** (2007), no. 3, 879–894, <http://dx.doi.org/10.1137/070683003>. (Cited on page 78).
- [CDIRS14] L. Calatroni, J.C. De los Reyes, and C.B. Schönlieb, *Dynamic sampling schemes for optimal noise learning under multiple nonsmooth constraints*, arXiv preprint (2014), <http://arxiv.org/abs/1403.1278>. (Cited on page 24).
- [CE05] T.F. Chan and S. Esedoglu, *Aspects of Total Variation Regularized L^1 Function Approximation*, SIAM Journal on Applied Mathematics (2005), 1817–1837, <http://dx.doi.org/10.1137/040604297>. (Cited on page 24).

- [CEP07] T.F. Chan, S. Esedoglu, and F.E. Park, *Image decomposition combining staircase reduction and texture extraction*, Journal of Visual Communication and Image Representation **18** (2007), no. 6, 464–486, <http://dx.doi.org/10.1016/j.jvcir.2006.12.004>. (Cited on page 107).
- [CEP10] ———, *A fourth order dual method for staircase reduction in texture extraction and image restoration problems*, ICIP'10 (2010), 4137–4140. (Cited on page 107).
- [Cha04] A. Chambolle, *An Algorithm for Total Variation Minimization and Applications*, Journal of Mathematical Imaging and Vision **20** (2004), no. 1-2, 89–97, <http://dx.doi.org/10.1023/B:JMIV.0000011325.36760.1e>. (Cited on page 81).
- [CL97] A. Chambolle and P. L. Lions, *Image recovery via total variation minimization and related problems*, Numerische Mathematik **76** (1997), 167–188, <http://dx.doi.org/10.1007/s002110050258>. (Cited on pages 27, 106, and 175).
- [CP08] P. L. Combettes and J-C. Pesquet, *Proximal Thresholding Algorithm for Minimization over Orthonormal Bases*, SIAM Journal on Optimization **18** (2008), no. 4, 1351–1376, <http://dx.doi.org/10.1137/060669498>. (Cited on page 151).
- [CP11] A. Chambolle and T. Pock, *A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging*, Journal of Mathematical Imaging and Vision **40** (2011), no. 1, 120–145, <http://dx.doi.org/10.1007/s10851-010-0251-1>. (Cited on pages 108 and 160).
- [CS05] T. Chan and J. Shen, *Image Processing and Image Analysis, Variational, PDE, Wavelet and Stochastic Methods*, SIAM, Philadelphia, 2005, <http://epubs.siam.org/doi/abs/10.1137/1.9780898717877>. (Cited on page 64).
- [CW05] P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting*, Multiscale Modeling & Simulation **4** (2005), 1168–1200, <http://dx.doi.org/10.1137/050626090>. (Cited on page 151).
- [Dea83] S. R. Deans, *The Radon Transform and Some of Its Applications*, New York: John Wiley & Sons, 1983. (Cited on page 57).
- [Dem85] F. Demengel, *Fonctions à Hessien borné*, Annales de l'Institut Fourier **34** (1985), 155–190. (Cited on page 106).

- [DLR] A. P. Dempster, N. M. Laird, and D. B. Rubin, *Maximum Likelihood from Incomplete Data via the EM Algorithm*, Journal of the Royal Statistical Society. Series B (Methodological) **39**, no. 1, 1–38, <http://www.jstor.org/stable/2984875>,. (Cited on page 64).
- [DSYT08] J. Duchi, S. S. Shai, S. Yoram, and C. Tushar, *Efficient Projections Onto the L1-ball for Learning in High Dimensions*, Proceedings of the 25th International Conference on Machine Learning, ICML '08, 2008, <http://doi.acm.org/10.1145/1390156.1390191>, pp. 272–279. (Cited on page 151).
- [EG92] L.C. Evans and R.F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992. (Cited on page 34).
- [Eps07] C. L. Epstein, *Introduction to the Mathematics of Medical Imaging, Second Edition*, 2007, <http://dx.doi.org/10.1137/9780898717792.fm>. (Cited on pages 55 and 57).
- [ET99] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, 1999, <http://epubs.siam.org/doi/abs/10.1137/1.9781611971088>. (Cited on pages 44, 47, 49, and 123).
- [EV07] C. Elion and L. Vese, *An image decomposition model using the total variation and the infinity laplacian*, Proc. SPIE **6498** (2007), 64980W–64980W–10, <http://dx.doi.org/10.1117/12.716079>. (Cited on page 108).
- [Eva10] L.C. Evans, *Partial Differential Equations, volume 19 of Graduate Studies in Mathematics, Second Edition*, American Mathematical Society, 2010. (Cited on pages 22 and 38).
- [EZC10] E. Esser, X. Zhang, and T. Chan, *A General Framework for a Class of First Order Primal-Dual Algorithms for Convex Optimization in Imaging Science*, SIAM Journal on Imaging Sciences **3** (2010), no. 4, 1015–1046, <http://dx.doi.org/10.1137/09076934X>. (Cited on page 83).
- [FL07] I. Fonseca and G. Leoni, *Modern Methods in the Calculus of Variations: Lp Spaces*, Springer, 2007. (Cited on pages 34 and 39).
- [FP11] J. M. Fadili and G. Peyré, *Total Variation Projection With First Order Schemes*, Image Processing, IEEE Transactions on **20** (2011), no. 3, 657–669, <http://dx.doi.org/10.1109/TIP.2010.2072512>. (Cited on page 151).
- [GB14] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming, version 2.1*. (Cited on page 148).

- [GLML07] J. B. Garnett, T. M. Le, Y. Meyer, and Vese L., *Image decompositions using bounded variation and generalized homogeneous Besov spaces*, Applied and Computational Harmonic Analysis **23** (2007), no. 1, 25 – 56, <http://dx.doi.org/10.1016/j.acha.2007.01.005>. (Cited on page 107).
- [GO09] T. Goldstein and S. Osher, *The split Bregman method for L1 regularized problems*, SIAM Journal on Imaging Sciences **2** (2009), no. 2, 323–343, <http://dx.doi.org/10.1137/080725891>. (Cited on pages 81, 83, 84, 87, 149, and 151).
- [Had05] A. Haddad, *Ph.D thesis: Méthodes variationnelles en traitement d'image*. (Cited on page 143).
- [Han10] P. Hansen, *Discrete Inverse Problems*, Society for Industrial and Applied Mathematics, 2010, <http://dx.doi.org/10.1137/1.9780898718836>. (Cited on page 150).
- [HS06a] W. Hinterberger and O. Scherzer, *Variational Methods on the Space of Functions of Bounded Hessian for Convexification and Denoising*, Computing **76** (2006), no. 1-2, 109–133, <http://dx.doi.org/10.1007/s00607-005-0119-1>. (Cited on page 107).
- [HS06b] M. Hintermüller and G. Stadler, *An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration*, SIAM Journal on Scientific Computing **28** (2006), no. 1, 1–23, <http://dx.doi.org/10.1137/040613263>. (Cited on page 118).
- [Hub64] P. J. Huber, *Robust estimation of a location parameter*, Ann. Math. Statist. **35** (1964), no. 1, 73–101, <http://dx.doi.org/10.1214/aoms/1177703732>. (Cited on pages 116 and 176).
- [KGSF11] G.A. Kastis, A. Gaitanis, T. Skouras, and A.S. Fokas, *Evaluation of a spline reconstruction technique for spect: Comparison with fbp and osem*, Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE, 2011, <http://dx.doi.org/10.1109/NSSMIC.2011.6153751>. (Cited on page 64).
- [KGSF15] G. A. Kastis, A. Gaitanis, A. P. Samartzis, and A. S. Fokas, *The SRT reconstruction algorithm for semiquantification in PET imaging*, Medical Physics **42** (2015), no. 10, 5970–5982, <http://dx.doi.org/10.1118/1.4931409>. (Cited on page 64).
- [KST11] Y. Kopsinis, K. Slavakis, and S. Theodoridis, *Online Sparse System Identification and Signal Reconstruction Using Projections Onto Weighted ℓ_1*

- Balls*, Signal Processing, IEEE Transactions on **59** (2011), no. 3, 936–952, <http://dx.doi.org/10.1109/TSP.2010.2090874>. (Cited on page 172).
- [Kui07] A. Kuijper, *p-Laplacian Driven Image Processing*, IEEE International Conference on Image Processing **5** (2007), 257–260, <http://dx.doi.org/10.1109/ICIP.2007.4379814>. (Cited on pages 108 and 128).
- [KV09] Y. Kim and L. Vese, *Image recovery using functions of bounded variation and Sobolev spaces of negative differentiability*, Inverse Problems and Imaging **3** (2009), 43–68, <http://dx.doi.org/10.3934/ipi.2009.3.43>. (Cited on page 142).
- [LCA07] T. Le, R. Chartrand, and T. J. Asaki, *A variational approach to reconstructing images corrupted by Poisson noise*, Journal of Mathematical Imaging and Vision **27** (2007), no. 3, 257–263, <http://dx.doi.org/10.1007/s10851-007-0652-y>. (Cited on pages 24, 25, and 64).
- [Leo09] G. Leoni, *A First Course in Sobolev Spaces, Vol. 105, Graduate Studies in Mathematics*, American Mathematical Society, 2009. (Cited on pages 38 and 39).
- [LL01] E. H. Lieb and M. Loss, *Analysis, volume 14 of graduate studies in mathematics*, vol. 4, American Mathematical Society, Providence, RI., 2001. (Cited on page 38).
- [LV05] T. Le and L. Vese, *Image Decomposition Using Total Variation and $div(BMO)$* , Multiscale Modeling & Simulation **4** (2005), no. 2, 390–423, <http://dx.doi.org/10.1137/040610052>. (Cited on pages 107 and 142).
- [LV08] L. Lieu and L. Vese, *Image Restoration and Decomposition via Bounded Total Variation and Negative Hilbert-Sobolev Spaces*, Applied Mathematics and Optimization **58** (2008), no. 2, 167–193, <http://dx.doi.org/10.1007/s00245-008-9047-8>. (Cited on pages 107, 142, and 145).
- [Mar14] A. Markoe, *Analytic Tomography*, Cambridge University Press, 2014, <http://dx.doi.org/10.1017/CB09780511530012>. (Cited on pages 57 and 60).
- [MBS⁺11] J. Müller, C. Brune, A. Sawatzky, T. Kusters, K.P. Schafers, and M. Burger, *Reconstruction of short time PET scans using Bregman iterations*, Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE, 2011, <http://dx.doi.org/10.1109/NSSMIC.2011.6153884>, pp. 2383–2385. (Cited on page 65).

- [Mey01] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*, vol. 22, American Mathematical Society, 2001. (Cited on pages 48, 140, 142, and 143).
- [MFI06] V. Marinakis, A. S. Fokas, and A. Iserles, *Reconstruction Algorithm for single photon emission computed tomography and its numerical implementation*, *Journal of the Royal Society Interface* **6** (2006), 45–54, <http://dx.doi.org/10.1098/rsif.2005.0061>. (Cited on page 64).
- [Mos08] A. Mosek, *Mosek: a full featured software package intended for solution of large scale optimization problems*. (Cited on page 148).
- [MS89] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, *Communications on Pure and Applied Mathematics* **42** (1989), no. 5, 577–685, <http://dx.doi.org/10.1002/cpa.3160420503>. (Cited on page 19).
- [Mül13] J. Müller, *Advanced Image Reconstruction and Denoising – Bregmanized (Higher Order) Total Variation and Application in PET*, Ph.D. thesis, University of Münster, 2013. (Cited on pages 65, 107, and 162).
- [Nat01] F. Natterer, *The Mathematics of Computerized Tomography*, Society for Industrial and Applied Mathematics, 2001, <http://dx.doi.org/10.1137/1.9780898719284>. (Cited on pages 57 and 61).
- [Nik04] M. Nikolova, *A variational approach to remove outliers and impulse noise*, *Journal of Mathematical Imaging and Vision* **20** (2004), no. 1, 99–120, <http://dx.doi.org/10.1023/B:JMIV.0000011326.88682.e5>. (Cited on page 24).
- [NW01] F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction*, Society for Industrial and Applied Mathematics, 2001, <http://dx.doi.org/10.1137/1.9780898718324>. (Cited on pages 57 and 62).
- [OBG⁺05] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, *An iterative regularization method for total variation-based image restoration*, *Multiscale Modeling & Simulation* **4** (2005), 460–489, <http://dx.doi.org/10.1137/040605412>. (Cited on pages 83, 149, and 159).
- [OSV03] S. Osher, A. Solé, and L. Vese, *Image Decomposition and Restoration Using Total Variation Minimization and the H^{-1}* , *Multiscale Modeling & Simulation* **1** (2003), no. 3, 349–370, <http://dx.doi.org/10.1137/S1540345902416247>. (Cited on pages 107 and 142).

- [P. 12] P. Getreuer, *Rudin-Osher-Fatemi Total Variation Denoising using Split Bregman*, Image Processing On Line (2012), <http://dx.doi.org/10.5201/ipol.2012.g-tvd>. (Cited on page 87).
- [PB15] K. Papafitsoros and K. Bredies, *A study of the one dimensional total generalised variation regularisation problem*, Inverse Problems and Imaging **9** (2015), no. 2, 511–550, <http://dx.doi.org/10.3934/ipi.2015.9.511>. (Cited on pages 120, 125, and 154).
- [PCBC10] T. Pock, D. Cremers, H. Bischof, and A. Chambolle, *Global Solutions of Variational Models with Convex Regularization*, SIAM Journal on Imaging Sciences **3** (2010), no. 4, 1122–1145, <http://dx.doi.org/10.1137/090757617>. (Cited on page 116).
- [Phe06] E. M. Phelps, *PET: Physics, Instrumentation and Scanners*, Springer New York, 2006, <http://dx.doi.org/10.1007/0-387-34946-4>. (Cited on page 53).
- [Pou10] A.D. Poularikas, *Transforms and application handbook, 3rd edition*, CRC Press, 2010. (Cited on pages 69 and 80).
- [PS08] C. Pöschl and O. Scherzer, *Characterization of minimisers of Convex regularization functionals*, Contemporary mathematics **451** (2008), 219–248. (Cited on page 107).
- [PS13] K. Papafitsoros and C.-B. Schönlieb, *A Combined First and Second Order Variational Approach for Image Reconstruction*, Journal of Mathematical Imaging and Vision (2013), 1–31, <http://dx.doi.org/10.1007/s10851-013-0445-4>. (Cited on page 108).
- [PW90] J. L. Prince and A. S. Willsky, *A geometrical Projection-Space reconstruction algorithm*, Linear Algebra and its Applications **130** (1990), 151–191, [http://dx.doi.org/10.1016/0024-3795\(90\)90211-T](http://dx.doi.org/10.1016/0024-3795(90)90211-T). (Cited on page 63).
- [Rad17] J. Radon, *Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten*, Akad. Wiss. **69** (1917), 262–277. (Cited on page 57).
- [Rin00] W. Ring, *Structural Properties of Solutions to Total Variation Regularisation Problems*, ESAIM: Mathematical Modelling and Numerical Analysis **34** (2000), 799–810, <http://dx.doi.org/10.1051/m2an:2000104>. (Cited on page 125).

-
- [Roc70] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970. (Cited on page 44).
- [ROF92] L.I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, *Physica D: Nonlinear Phenomena* **60** (1992), no. 1-4, 259–268, [http://dx.doi.org/10.1016/0167-2789\(92\)90242-F](http://dx.doi.org/10.1016/0167-2789(92)90242-F). (Cited on pages 19, 24, and 175).
- [Rud87] W. Rudin, *Real and complex analysis*, WCB, McGraw-Hill, Boston, 1987. (Cited on page 34).
- [Saw11] A. Sawatzky, *(Nonlocal) total variation in Medical imaging*, Ph.D. thesis, University of Münster, 2011. (Cited on pages 52 and 65).
- [SBMB09] A. Sawatzky, C. Brune, J. Müller, and M. Burger, *Total variation processing of images with Poisson statistics*, *Computer Analysis of Images and Patterns*, Springer, 2009, http://dx.doi.org/10.1007/978-3-642-03767-2_65, pp. 533–540. (Cited on pages 25 and 63).
- [SBW⁺08] A. Sawatzky, C. Brune, F. Wübbeling, T. Kusters, K. Schafers, and M. Burger, *Accurate EM-TV algorithm in PET with low SNR*, *Nuclear Science Symposium Conference Record*, 2008. NSS '08. IEEE, 2008, <http://dx.doi.org/10.1109/NSSMIC.2008.4774392>, pp. 5133–5137. (Cited on page 63).
- [Sch98] O. Scherzer, *Denoising with higher order derivatives of bounded variation and an application to parameter estimation*, *Computing* **60** (1998), no. 1, 1–27, <http://dx.doi.org/10.1007/BF02684327>. (Cited on page 107).
- [Sch09] C.-B. Schönlieb, *Total variation minimization with an H^{-1} constraint*, *Singularities in Nonlinear Evolution Phenomena and Applications*, 2009, pp. 201–232. (Cited on page 107).
- [Set09] S. Setzer, *Split Bregman Algorithm, Douglas-Rachford Splitting and Frame Shrinkage*, *Scale Space and Variational Methods in Computer Vision* **5567** (2009), 464–476, http://dx.doi.org/10.1007/978-3-642-02256-2_39. (Cited on page 83).
- [Set11] S. Setzer, *Operator Splittings, Bregman Methods and Frame Shrinkage in Image Processing*, *International Journal of Computer Vision* **92** (2011), no. 3, 265–280, <http://dx.doi.org/10.1007/s11263-010-0357-3>. (Cited on page 83).

- [SST11] S. Setzer, G. Steidl, and T. Teuber, *Infimal convolution regularizations with discrete ℓ_1 -type functionals*, Communications in Mathematical Sciences **9** (2011), 797–872, <http://dx.doi.org/10.4310/CMS.2011.v9.n3.a7>. (Cited on page 107).
- [SSW77] K. T. Smith, D. C. Solmon, and S. L. Wagner, *Practical and mathematical aspects of the problem of reconstructing objects from radiographs*, Bulletin of the American Mathematical Society **83** (1977), no. 6, 1227–1270, <http://projecteuclid.org/euclid.bams/1183539851>. (Cited on page 59).
- [SV82] L.A. Shepp and Y. Vardi, *Maximum Likelihood Reconstruction for Emission Tomography*, Medical Imaging, IEEE Transactions on **1** (1982), no. 2, 113–122, <http://dx.doi.org/10.1109/TMI.1982.4307558>. (Cited on page 64).
- [TA77] A. N. Tikhonov and V. Y. Arsenin, *Solutions of ill-posed problems*, Winston and Sons, 1977. (Cited on page 24).
- [Thi91] J. Ph. Thirion, *A geometric alternative to computed tomography*, Technical report 1463, INRIA (1991). (Cited on page 63).
- [TPT⁺13] C. Tsoumpas, I. Polycarpou, K. Thielemans, C. Buerger, A. P. King, T. Schaeffter, and P. K. Marsden, *The effect of regularization in motion compensated PET image reconstruction: a realistic numerical 4D simulation study*, Physics in Medicine and Biology **58** (2013), no. 6, 1759–1773, <http://stacks.iop.org/0031-9155/58/i=6/a=1759>. (Cited on page 54).
- [TS80] R. Temam and G. Strang, *Functions of bounded deformation*, Archive for Rational Mechanics and Analysis **75** (1980), no. 1, 7–21, <http://dx.doi.org/10.1007/BF00284617>. (Cited on page 107).
- [Ves01] L. Vese, *A Study in the BV Space of a Denoising-Deblurring Variational Problem*, Applied Mathematics and Optimization **44** (2001), no. 2, 131–161. (Cited on pages 68 and 113).
- [VO96] C. Vogel and M. Oman, *Iterative Methods for Total Variation Denoising*, SIAM Journal on Scientific Computing **17** (1996), no. 1, 227–238, <http://dx.doi.org/10.1137/0917016>. (Cited on page 151).
- [WA04] M. N. Wernick and J. N. Aarsvold, *Emission tomography : the fundamentals of PET and SPECT*, Elsevier Academic Press, 2004. (Cited on pages 53 and 55).
- [WBSS04] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, *Image quality assessment: From error visibility to structural similarity*, IEEE Transactions

- on Image Processing **13** (2004), no. 4, 600–612, <http://dx.doi.org/10.1109/TIP.2003.819861>. (Cited on page 158).
- [WYYZ08] Y. Wang, J. Yang, W. Yin, and Y. Zhang, *A new alternating minimization algorithm for total variation image reconstruction*, SIAM Journal of Imaging Sciences (2008), 248–272, <http://dx.doi.org/10.1137/080724265>. (Cited on pages 86 and 151).
- [YOGD08] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, *Bregman Iterative Algorithms for ℓ_1 -Minimization with Applications to Compressed Sensing*, SIAM Journal on Imaging Sciences **1** (2008), no. 1, 143–168, <http://dx.doi.org/10.1137/070703983>. (Cited on page 83).
- [Yos80] K. Yosida, *Functional Analysis*, 6th edition, Springer-Verlag, New York, 1980. (Cited on pages 39 and 124).