MULTI-BANG REGULARIZATION AND APPLICATIONS

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Science and Engineering

2021

Philip J. Richardson
Department of Mathematics

Contents

Word count 23662

List of Figures

The University of Manchester

Philip J. Richardson
Doctor of Philosophy
Multi-bang Regularization and Applications
July 9, 2021

Inference of the internal structure of an object from passive radiation imaging has many applications in modern day life, ranging from Medical Imaging to Nuclear Security. In this thesis we focus on the joint reconstruction of attenuation a and radiation source density f from the Attenuated Radon Transform (AtRT) $R_a f$ which models Single-Photon Emission Computed Tomography (SPECT) data. Joint inversion in the general case is known to be impossible and we instead consider the setting were a takes only finitely many values, which we refer to as "multi-bang", and f is once differentiable with compact support. In this setting we are able to characterise singularities appearing in the AtRT. With constraints on the support of f in relation to the support of a and mild conditions on the boundaries of a, we are able to show unique recovery of the sets on which a is constant. When the sets making up a are nested convex sets, referred to here as *nicely multi-bang*, we show unique recovery of a and f. It is also possible to obtain partial results for the more general case ranging from the complete determination in special cases to situations where we can at least determine a in certain sets. We also propose a numerical algorithm to jointly compute a and f from $R_a f$ based on a weakly-convex regularizer, referred to as a multi-bang regularizer. Various numerical examples are given to show that the algorithm performs well on synthetic examples.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

I would like to thank my supervisor Sean Holman for suggesting this problem and for the unending support given to me throughout the entirety of my PhD studies. I would also like to thank my officemates Alex, Gabriele, Luke, Xiaoxi, Borys and Vicky for providing much needed conversation and fun in the office to keep me sane. Thank you to my friends Norris, Brad, Mhelody and KFC for being someone to talk to about things other than maths. Finally I would like to thank my parents and sister as well as my loving girlfriend Sophie for putting up with me over the last 3 and a half years.

Chapter 1

Introduction

The ability to infer internal geometry from passive radiation data has many applications in modern day society. These include, but are not limited to, Medical imaging; including PET [61] and SPECT imaging [61, 30, 31], Seismic Imaging [7] and Nuclear Security [59].

One important medical imaging technique is Single-Photon Emission Computed Tomography (SPECT) [30, 32, 55, 14]. In SPECT a radioactive tracer which produces gamma rays is inserted into the body and the intensity of the radiation is recorded by a gamma camera. This gamma camera is fitted with a collimator and so as well as recording intensity, it also records orientation of the rays measured. These measurements allow us to infer information about the internal geometry of objects. This thesis focuses on the simultaneous recovery of two quantities a, which will be used to denote the attenuation of the materials which make up that object, and f which will be the radiation source density. In two dimensions these two quantities are linked with passive radiation/SPECT measurements by the Attenuated Radon Transform (AtRT)

$$R_a f(s,\theta) = \int_{-\infty}^{\infty} f(s\theta^{\perp} + t\theta) e^{-Da(s\theta^{\perp}+t\theta,\theta)} \,\mathrm{d}t, \tag{1.1}$$

where D is the beam transform

$$Da(\boldsymbol{x},\theta) = \int_{0}^{\infty} a(\boldsymbol{x} + t\theta) \,\mathrm{d}t,$$

specific definitions of s, θ and $\theta^{\perp}$ are given in Definition 1.

The AtRT, given the attenuation and source density within an object, calculates the radiation emitted along a given line.

When a is fixed, the mapping $f \mapsto R_a f$ is known to be analytically invertible under certain mild conditions, and when these hold a closed form solution for the inverse is known, see [48, 14, 38]. Furthermore similar formulae have been found for attenuated tensor transforms [43] and the problem has been studied on surfaces both for functions and tensor fields [36, 40]. The problem of recovering both a and f from the AtRT is sometimes known as the SPECT identification problem [44, 56, 55], and a known result, given in [55], shows non-uniqueness for radial a and f. In particular [55] shows there are different pairs of a and f depending only on distance to the origin which give the same AtRT, in some cases a solution pair may exist with $a = 0$. This non-uniqueness property also holds for maps which are "close" to being radial, as shown in [30], and numerical investigations in [54] also show evidence of non-uniqueness in other situations. Another important assumption in the inversion of $R_a f$ is that f is non-zero, this is as if $f = 0$ then any choice of a leads to an AtRT which gives 0.

Despite some negative results, under additional hypotheses determination of a and f from $R_a f$ is still possible. In medical imaging literature there has been a lot of work on determining numerical methods for the SPECT identification problem (e.g. [58, 60, 31] and their references). Although we focus on simultaneous recovery of a and f it is common for practical applications of SPECT to obtain the attenuation through a separate transmission CT scan first; this is sometimes referred to as attenuation correction [58, 31]. The AtRT is nonlinear in a which can cause a large amount of computation strain in numerics and linearisation has been applied in order to obtain numerical reconstructions of attenuation from SPECT data alone in [12], as well as range conditions [13, 11]. Although we do not consider scattering effects here, some studies have also attempted to make use of scattered photons for attenuation identification in SPECT ([21, 22]). However, this case requires the forward model to be enriched with a scattering term, which yields a different mathematical problem.

There are not many positive theoretical results concerning recovery of both a and f from emission data alone. In the case when Da in (1.1) is replaced by a constant μ times t the transform is called the exponential Radon transform, and in [55] it is shown that μ can be determined from the exponential Radon transform when f is unknown, provided it is not radial. A linearisation of the problem is studied from a microlocal point of view in [56], and used to establish some results for the nonlinear

problem as well. A range characterisation of $f \mapsto R_a f$ for given a originally found in [48] and further explored in [49] was used in [3] to analyse recovery of both a and f. Another related result is given in [16] which shows that unique recovery of a and f is possible when a is a multiple of the characteristic function of a star shaped polygon. In this thesis we assume that a takes on only finitely many values, and refer to such a as *multi-bang*. As detailed below, we are able to show unique recovery of such a from $R_a f$ in some cases.

Recently the authors of [18, 19] introduced a convex multi-bang regularization technique intended to allow reconstructions of images in which there are only certain known values, and our line of research leading to this thesis was originally inspired by this technique. There are many applications where the multi-bang regularization technique might be useful, particularly in many forms of medical imaging, e.g. SPECT imaging [30] and X-ray Imaging [63], this is because there are only a handful of tissue types e.g muscle, bone, skin and tumorous expected in the body. Another potential application of multi-bang is in Seismic Imaging [7]. In Seismic Imaging geographical artefacts are scanned and often the materials expected are typically known a priori (for example scanning cliff faces of a known rock to find potential weaknesses [7]) and the internal structure of the artefact is the objective of the image reconstruction. The convex multi-bang technique of [18, 19] was applied numerically to the problem of recovering multi-bang a and f from $R_a f$ in [54] with mixed results, and we have since modified the method to use a weakly-convex (rather than convex) multi-bang regularization combined with Total Variation(TV) to promote the joint recovery of multi-bang a and f from the AtRT $R_a f$. We implement this by alternating updates between a and f using [1] to prove convergence. The a update is the most computationally intensive step due to the nonlinearity and using recent work by [32, 5] we apply a variant of the Alternating Direction Method of Multipliers (ADMM) with a non-convex multi-bang regularizer, which we show lends itself to promoting multi-bang solutions.

Furthermore, the combination of the weakly-convex multi-bang and TV regularizer allows for good recovery with few projections. Here a projection refers to data obtained for a single angle θ and many s. The Computerized Tomography (CT) problem [17, 58, 4], in which the aim is to instead invert the Radon Transform $R_0 f$ to recover

multi-bang f, with limited projections is an example of a wider range of tomography known as Discrete Tomography (DT) [34]. Multi-bang regularization seems to lend itself to the CT case, as the reconstructions of f are typically binary or take a small set of grayscale values which we could use as an admissible set. Since DT takes few projections between $4 - 32$ [17, 63] (with higher resolution and more complex images requiring more projections) the problem of numerically inverting is commonly underdetermined and the relationship between the known projection angles is important in ensuring a unique reconstruction. One particularly important result is given in [34] which shows that the angles should be chosen so that they are irrationally related in order to ensure the best recovery.

The novel contributions of this thesis and [37] are summarised as follows. The precise and rigorous versions of the theoretical results, which include a few other technical assumptions, are presented in Chapter 3 for Lemma 15 and Chapter 4 for Theorem 6 and Lemmas 18-20.

1. Lemma 15. Assuming that a is multi-bang, $f \in C_c^1(\mathbb{R}^2)$ is non-negative, and with some additional assumptions about the regularity of the boundaries of the regions of constant a, then we can uniquely determine the regions on which a is constant.

2. Theorem 6. If a and f are as in Lemma 15 with the regions of constant a being a sequence of nested convex sets and assuming $\text{supp}(f)$ satisfies certain hypotheses, then a and f can be uniquely determined from $R_a f$.

3. Lemmas 18- 20. If a and f are as in Lemma 15 then we can uniquely determine, in some cases, a and f from $R_a f$. Futhermore, Lemmas 18-20 give conditions required on a region Ω and f in order to uniquely determine the value of a in some of its regions. Again we have a requirement on the $\text{supp}(f)$.

4. We propose a numerical algorithm for joint recovery of multi-bang a and f for limited projection data, and demonstrate its utility with some numerical examples.

Point 2 in the above list is the main theoretical result from [37], and Algorithm 2 is also given in [37]. All results given in Chapter 3, excluding Chapter 3.5, (which relate

to points 1 and 3) are extensions of work from [37] are are novel contributions of this Thesis. The Discrete Tomography Algorithm given in Algorithm 3 is also novel to this thesis.

As it is presented in this thesis, the hypotheses of all results in Chapter 3 relating to Theorem 6 imply that $\text{supp}(f) \not\subset \text{supp}(a)$ which is not realistic in practical applications of SPECT. We would like to thank an anonymous referee for [37] who pointed out that in the case where $a = c$ on some circle of radius r, methods from [55] give a counterexample to Theorem 2 with radial f supported within that circle. In general, if $\text{supp}(f) \subset \text{supp}(a)$ then the outermost boundary on which a is discontinuous cannot be determined using techniques presented in this thesis. Numerically we have found that the only phantoms we examined which had non-unique solutions were radial in a and f and even small perturbations to make a non-radial produced unique solutions, this evidence is presented at the end of Chapter 6.

The proofs for Lemma 15, Lemmas 18-20 and Theorem 6 are based on careful analysis of the singularities which occur in $R_a f$ arising from the jumps of a, as well as a result that if a is known outside of a convex region, then $R_a f$ determines f uniquely also outside this convex region (see Lemma 17). We prove this latter result by reducing it to the problem considered in [15, Theorem 3.1], although other proofs using for example analytic microlocal analysis as in [29] should also be possible.

The rest of the thesis is structured as follows. Chapter 2 gives a literature review on relevant background topics. Chapter 3 contains results relating to the unique determination of the multi-bang regions of a and this Chapter is split into two sections. The first contains results in the general case and the second involves a special case where all the regions of a are nested convex sets, known as nicely multi-bang a. Chapter 4 is also split into two sections with the first outlining a complete proof of unique recovery of a and f from $R_a f$ in the nicely multi-bang case. The second section gives partial results on the unique recovery of a and f from $R_a f$ in the general multi-bang case. Chapter 5 focuses on the numerical methods involved in recovering a and f from the optimization problem. Chapter 6 gives some numerical reconstructions for various cases of a and f as well as some discrete tomography reconstructions. Finally, Chapter 7 summarizes the results obtained in the thesis as well as some potential directions for further research.

Chapter 2

Literature Review

2.1 Attenuated Radon Transform

Let a and f be compactly supported functions on $\mathbb{R}^2$ and let $\Omega \subset \mathbb{R}^2$ be a simply connected Lipschitz domain containing the supports. Then the 2-D photon transport equation, without scattering, is

$$\begin{aligned} \theta \cdot \nabla u(\boldsymbol{x}, \theta) + a(\boldsymbol{x}) u(\boldsymbol{x}, \theta) = f(\boldsymbol{x}), \quad (\boldsymbol{x}, \theta) \in \Omega \times \mathbb{S}^1, \\ u|_{\Gamma^-} = 0, \end{aligned} \tag{2.1}$$

where $u(x, \theta)$ is the photon flux through the point $\boldsymbol{x}$ in a unit direction $\theta \in \mathbb{S}^1$ and

$$\Gamma^- = \{(\boldsymbol{x}, \theta) \in \partial\Omega \times \mathbb{S}^1 \mid \theta \cdot \underline{n}(\boldsymbol{x}) \leq 0\},$$

where $\underline{n}(\boldsymbol{x})$ is the unit outward pointing normal to the boundary at $\boldsymbol{x}$. Intuitively the differential equation in (2.1) states that photons are created by a source with density f and then move along straight lines while being attenuated at a rate given by a. The boundary condition in (2.1) requires us to have no radiation entering the domain Ω. We can also define an analogous set

$$\Gamma^+ = \{(\boldsymbol{x}, \theta) \in \partial\Omega \times \mathbb{S}^1 \mid \theta \cdot \underline{n}(\boldsymbol{x}) > 0\}$$

which is the set of all points on the boundary and directions which point out of Ω.

Before we examine (2.1), we first explicitly define the Attenuated Radon transform and the Beam transform.

Definition 1 (**The Beam Transform).** *We define the beam transform of a via the map* $D : L_c^\infty(\mathbb{R}^2) \longrightarrow L^\infty(\mathbb{R}^2 \times \mathbb{S}^1)$

$$Da(\boldsymbol{x}, \theta) = \int_0^\infty a(\boldsymbol{x} + t\theta)\, \mathrm{d}t. \tag{2.2}$$

Here $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$ is an arbitrary starting point and θ is a direction of travel such that $\theta \in \mathbb{S}^1$. This transform represents the total attenuation over a ray. It is often more convenient and natural to work with rays which we parametrise by (s, θ), where s is a scalar quantity representing the closest approach a ray makes to the origin. We define $\theta^\perp := (-\theta_2, \theta_1)$ i.e, rotating θ by $\pi/2$ anticlockwise and then we can write any point $\boldsymbol{x} \in \mathbb{R}^2$ as a linear combination of θ and $\theta^\perp$. With this notation we can define the Attenuated Radon Transform.

Definition 2 (**The Attenuated Radon Transform).** *We define the Attenuated Radon Transform, for fixed* $a \in L_c^\infty(\mathbb{R}^2)$ *via the map* $R_a : L_c^\infty(\mathbb{R}^2) \longrightarrow L^\infty(\mathbb{R} \times \mathbb{S}^1)$

$$R_a f(s, \theta) = \int_{-\infty}^\infty f(s\theta^\perp + t\theta) e^{-Da(s\theta^\perp + t\theta, \theta)}\, \mathrm{d}t. \tag{2.3}$$

When $a = 0$ we recover the also well studied Radon transform [17, 56, 45, 34]

$$Rf(s, \theta) = \int_{-\infty}^\infty f(s\theta^\perp + t\theta)\, \mathrm{d}t.$$

The photon transport equation (2.1) can be solved by method of characteristics to produce the following result.

Theorem 1. *If* u *is the solution of* (2.1) *then* $\lim_{t\to\infty} u(\boldsymbol{x} + t\theta, \theta) = R_a f(\boldsymbol{x} \cdot \theta^\perp, \theta)$.

Proof. We use the method of characteristics to solve this first order partial differential equation. This leads us to the 3 first order ODEs,

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = \theta_1 \tag{2.4a}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = \theta_2 \tag{2.4b}$$

$$\frac{\mathrm{d}u}{\mathrm{d}t} = f(x) - a(x)u(x, \theta) \tag{2.4c}$$

Here we have set $\theta = (\theta_1, \theta_2)$ and $\boldsymbol{x} = (x_1, x_2)$. Solving (2.4a) and (2.4b) leads us to $\boldsymbol{x} = \boldsymbol{x}_0 + t\theta$, where $\boldsymbol{x}_0$ is some arbitrary fixed point. We can then replace $\boldsymbol{x}$ by $\boldsymbol{x}_0 + t\theta$ in order to solve (2.4c) . Using this we can rewrite (2.4c) as,

$$\frac{\mathrm{d}u}{\mathrm{d}t} = f(\boldsymbol{x}_0 + t\theta) - a(\boldsymbol{x}_0 + t\theta)u(\boldsymbol{x}_0 + t\theta, \theta). \tag{2.5}$$

This ODE can be solved using integrating factors. The integrating factor here is

$$\mathcal{I}(\boldsymbol{x}_0, \rho, \theta) = e^{\int_0^\rho a(\boldsymbol{x}_0 + p\theta)\mathrm{d}p}.$$

Suppose the oriented line $\boldsymbol{x} = \boldsymbol{x}_0 + t\theta$ enters Ω when $t = s$ then the boundary condition given in (2.1) tells us that $u(\boldsymbol{x}_0 + s\theta, \theta) = 0$. Multiplying (2.5) by $\mathcal{I}(\boldsymbol{x}_0, \rho, \theta)$ and using the fundamental theorem of calculus we get,

$$u(\boldsymbol{x}_0 + \tau\theta, \theta)e^{\int_0^\tau a(\boldsymbol{x}+p\theta)\mathrm{d}p} = \int_s^\tau f(\boldsymbol{x}_0 + t\theta)e^{\int_0^t a(\boldsymbol{x}_0+p\theta)\mathrm{d}p}\mathrm{d}t.$$

Rearranging this we get

$$u(\boldsymbol{x}_0 + \tau\theta, \theta) = \int_s^\tau f(\boldsymbol{x}_0 + t\theta)e^{-\int_t^\tau a(\boldsymbol{x}_0+p\theta)\mathrm{d}p}\mathrm{d}t.$$

Finally, using the substitution $\hat{p} = p - t$ in the exponential involving a and letting $\tau \longrightarrow \infty$ and the fact that $f(\boldsymbol{x} + t\theta) = 0$ for all $t \leq s$, we get

$$\lim_{\tau\to\infty} u(\boldsymbol{x}_0 + \tau\theta, \theta) = \int_{-\infty}^\infty f(\boldsymbol{x} + t\theta)e^{-\int_0^\infty a(\boldsymbol{x}_0+t\theta+\hat{p}\theta)\mathrm{d}\hat{p}}\,\mathrm{d}t.$$

□

With the Attenuated Radon Transform described we now move onto a section giving a precise description of what a multi-bang a is.

2.2 Multi-bang attenuation

We begin by defining the types of regions, specially the boundaries of the regions, over which a is constant that we consider in this thesis.

Definition 3. (Analytic boundary) *Suppose Ω is a bounded region in $\mathbb{R}^2$. For a point $x^* \in \partial\Omega$ we say that $\partial\Omega$ is analytic near x^* if there exists a neighbourhood V of x^* and a set of Cartesian coordinates centred at x^* such that on V the boundary is given by $(x, f(x))$ where f is a real analytic function. If $\partial\Omega$ is analytic near all $x^* \in \partial\Omega$ we say that $\partial\Omega$ is an analytic boundary.*

Definition 4. (Corner) *Let x^* be a point on the boundary of a region Ω. Then x^* is a corner point of Ω if there exists a neighbourhood V of x^* and Cartesian coordinates centred at x^* such that we can describe the boundary for $x < 0$ via $(x, f(x))$ and for $x > 0$ via $(x, g(x))$, where both f and g are analytic, and*

$$\lim_{x\to(x^*)^+} f'(x) \neq \lim_{x\to(x^*)^-} g'(x) \tag{2.6}$$

with the extra requirement that we do allow either of these limits to be $\pm\infty$.

Definition 5. (Piecewise analytic boundary with corners) *A set* Ω *has piecewise analytic boundaries with corners if for every point* x^* *in* $\partial\Omega$ x^* *is either a corner point for* Ω*, or* $\partial\Omega$ *is analytic near* x^*.

With these Definitions we are now ready to define precisely what we mean by a multi-bang a.

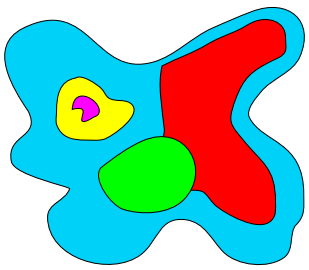
Definition 6. (Multi-bang) *We say that* $a \in L^\infty(\mathbb{R}^2)$ *is multi-bang if there exists a finite set* $\mathcal{A} = \{a_1, \ \dots \ , \ a_n\} \subset \mathbb{R}$*, called the admissible set, and a collection of disjoint bounded open sets* $\{\Omega_j\}_{j=1}^n$ *with piecewise analytic boundaries possibly having corners such that*

$$a = \sum_{j=1}^{n} a_j \chi_{\Omega_j}. \tag{2.7}$$

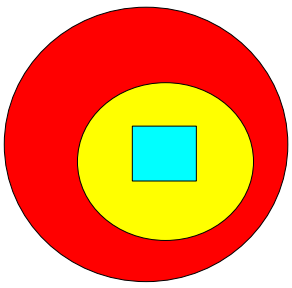
Here χ_{Ω_j} *is the characteristic function of the set* Ω_j*, and we assume that for all* Ω_j *the interior of the closure of* Ω_j *is equal to* Ω_j*. We also assume that any line only intersects the boundaries* $\cup_{j=1}^n \partial\Omega_j$ *finitely many times.*

Definitions 5 and 6 allow for a lot of different choices for a but eliminate some pathological cases such as any boundary which crosses any given straight line infinitely many times. We include Figure 2.1 for examples of multi-bang a.

Figure 2.1: Two examples of multi-bang a. The left is a general multi-bang example with 6 values (as 0 is also considered a multi-bang value). The right is a nicely multi-bang example, see Definition 14, with 4 multi-bang values



(a) General multi-bang example



(b) Nicely multi-bang example

With multi-bang a defined, we now give an outline of problem we wish to solve. In particular we show how we can form a Variational problem which can be optimized to simultaneously recover a and f.

2.3 Outline of problem

The aim of this project is as follows, given boundary data d on Γ^+ from multi-bang a (see Definition 6), with known admissible set $\mathcal{A}$, and $f \in C_c^1(\Omega)$ simultaneously and stably recover a and f. The boundary data here is the Attenuated Radon transform or equivalently the solution of the photon transport equation on Γ^+. Here stably means that the reconstruction depends continuously with respect to some norm on the data. We aim to do this by solving the following optimization problem

$$\underset{a,f\in BV(\Omega)}{\operatorname{argmin}} \|R_a f - d\|^2_{L^2(\Gamma^+)} + \alpha\mathcal{M}(a) + \gamma_a \mathrm{TV}(a) + \gamma_f \mathrm{TV}(f), \tag{2.8}$$

where $\mathcal{M}$ and TV denote the multi-bang penalty and total variation regularizer respectively (see (2.18) and (5.12)). Here α and γ_a, γ_f are regularization parameters associated to multi-bang and TV respectively and BV is the space of functions of bounded variation on, see [2]. An advantage of using (2.8) over an iterative regularisation method, as in [25], is that we obtain simpler expressions for updates, see Chapter 5.1.1-5.1.3, and have more control over the recovery of multi-bang images. Multi-bang regularization is the more recent development and Section 2.4 examines the literature relating to it in [19, 18, 54]. Total Variation is very well studied see [46, 18, 37, 54], and their references, and the exact form of Total Variation we use numerically will be examined in Chapter 5. In Chapters 5 and 6 we work in the discrete setting where the domain Ω is split into square pixels and we assume a and f to be piecewise constant over these pixels. This leads to the following discretised variational problem, which we consider in Chapter 5,

$$\underset{a,f}{\operatorname{argmin}} \|R[a]_a f - d\|_2^2 + \alpha\mathcal{M}(a) + \gamma_a \mathrm{TV}(a) + \gamma_f \mathrm{TV}(f), \tag{2.9}$$

where $R[a]_a f$ is a discretized version of the AtRT which depends on a.

Typical SPECT image sizes range from 256×256 to 2048×2048 [30, 17] or higher and even in the smallest case we have 256^{2k} possible multi-bang images, where k is

the number of multi-bang values. This makes the problem infeasible to be solved by just testing every combination of multi-bang values. Therefore it is necessary to solve the optimization problem (2.9) via numerical optimization techniques.

With the problem outlined we now move onto the literature review. We begin by looking at work carried out relating to multi-bang regularization.

2.4 Multi-bang regularization

This section of the Literature review is dedicated to multi-bang regularization first introduced in [19] and further developed in [18]. The general idea behind multi-bang regularization is as follows: Suppose that we have some quantity which takes only finitely many known values from an admissible set $\mathcal{A}$. Then produce a functional which penalizes functions containing values outside $\mathcal{A}$. In the literature [18, 19] a function is called multi-bang if it takes values from $\mathcal{A}$ almost everywhere. The multi-bang regularizer is then coupled with some data fidelity term to produce multi-bang reconstructions.

Before examining the multi-bang regularizer itself we need some preliminary definitions and results which we present in the next subsection.

2.4.1 Multi-bang regularization preliminaries

This subsection contains important definitions which we will use throughout the thesis when dealing with the multi-bang regularizer. We begin with the notion of Lower-semicontinuity.

Definition 7. (Lower-semicontinuous function) *A function $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}$, $n \in \mathbb{N}$, is called lower-semicontinuous at a point x_0 if $\forall \epsilon > 0, \exists U$ an open neighbourhood of x_0 such that $f(x) \geq f(x_0) - \epsilon \ \forall x \in U$ when $f(x_0) < \infty$, and $\lim_{x \to x_0} f(x) = \infty$ when $f(x_0) = \infty$. If f is lower-semicontinuous at all $x_0 \in \mathbb{R}^n$ we say that f is lower-semicontinuous. This can also be written as*

$$\liminf_{x \longrightarrow x_0} f(x) \geq f(x_0).$$

Lower semi-continuity is an important property relating to convergence of many numerical techniques to a critical point [6, 10] and this will be examined in more detail

in Chapter 5. We now give some important definitions related to convex functions.

Definition 8. (Proper convex function) *Let $X \subset \mathbb{R}^n$ be a convex set. We say that a function $f : X \to \mathbb{R} \cup \{\infty\}$ is convex if for all $x, y \in \mathbb{R}^n$ and for all $t \in [0, 1]$ we have*

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

Additionally, if there exists some x such that $f(x) < \infty$ and $f(y) > -\infty$ for all y then we say that f is proper convex function.

Convex functions may not necessarily be differentiable but they do admit a well-defined alternative called the convex subdifferential

Definition 9. (Convex subdifferential) *Let $X \subset \mathbb{R}^n$ be a convex set and $f : X \to \mathbb{R} \cup \{\infty\}$ be a convex function. We say that a vector $v \in X$ is a subgradient at a point $x_0 \in X$ if for all $x \in X$ we have*

$$f(x) - f(x_0) \geq v \cdot (x - x_0).$$

The set of all subgradients at x_0 is called the subdifferential of f at x_0 and is denoted $\partial f(x_0)$.

Note that if f is differentiable at a point x_0 then $\partial f(x_0) = \{f'(x_0)\}$ [54]. Although the convex-subdifferential is only defined for convex functions, we can extend this definition using techniques from [5] to the following class of functions known as weakly-convex functions.

Definition 10. (Weakly-convex function) *Let $X \subset \mathbb{R}^n$ be a convex set and $f : X \to \mathbb{R} \cup \{\infty\}$. We say that f is ρ-weakly convex, or just weakly-convex, if there exists some $\rho > 0$ such that the function*

$$h_\rho(x) := f(x) + \frac{\rho}{2}\|x\|^2 \tag{2.10}$$

is convex.

The formula (2.10) allows us to extend the notion of convex subdifferentials to weakly-convex functions.

Definition 11. (Weakly-convex subdifferential) *Let f be a ρ-weakly convex function. The weakly-convex subdifferential of f is given by*

$$\partial f(x) := \partial h_\rho(x) - \rho x \tag{2.11}$$

where $h_\rho(x)$ is as in (2.10) and $\partial h_\rho(x)$ is understood in the sense of Definition 9.

It is important to note that there may be many choices of ρ for a weakly-convex function which make h_ρ convex. The following lemma shows that the weakly-convex subdifferential is independent of the choice of ρ and therefore well-defined.

Lemma 1. *Let f be weakly-convex as in Definition 10. Let $\rho_1, \rho_2 > 0$ be such that $h_{\rho_1}(x)$ and $h_{\rho_2}(x)$ are convex. Then*

$$\partial h_{\rho_1}(x) - \rho_1 x = \partial h_{\rho_2}(x) - \rho_2 x$$

and therefore ∂f, as in Definition 11, is well-defined.

Proof. Let $x_0 \in X$ and suppose $p \in \partial h_{\rho_1}(x_0) - \rho_1 x_0$ and therefore $p + \rho_1 x_0 \in \partial h_{\rho_1}(x_0)$. By Definition 9 we have for any $x \in X$

$$f(x) - f(x_0) + \frac{\rho_1}{2}\|x\|^2 - \frac{\rho_1}{2}\|x_0\|^2 \geq (p + \rho_1 x_0) \cdot (x - x_0). \tag{2.12}$$

We can rewrite (2.12) to obtain

$$\begin{aligned} f(x) - f(x_0) + \frac{\rho_2}{2}\|x\|^2 - \frac{\rho_2}{2}\|x_0\|^2 &\geq (p + \rho_2 x_0) \cdot (x - x_0) + (\rho_1 - \rho_2) \cdot (x - x_0) \\ &\quad + \left(\frac{\rho_1 - \rho_2}{2}\right)\left(\|x_0\|^2 + \|x\|^2\right) \\ &= (p + \rho_2 x_0) \cdot (x - x_0) + \left(\frac{\rho_1 - \rho_2}{2}\right)\|x - x_0\|^2 \\ &\geq (p + \rho_2 x_0) \cdot (x - x_0), \end{aligned}$$

this gives $p + \rho_2 x_0 \in \partial h_{\rho_2}(x_0)$ and therefore $p \in \partial h_{\rho_2}(x_0) - \rho_2 x_0$. An identical argument can be used to show any $q \in \partial h_{\rho_2}(x_0) - \rho_2 x_0$ is also in $\partial h_{\rho_1}(x_0) - \rho_2 x_0$. Since $x_0 \in X$ was chosen arbitrarily this completes the proof. □

With these definitions and results established we are now ready to examine the multi-bang regularizer.

2.4.2 Multi-bang regularizer

Let Ω be a bounded domain in $\mathbb{R}^2$ and suppose that we expect some unknown quantity a to take values from the set $\mathcal{A} = \{a_1, a_2, ..., a_k\}$, with $0 = a_1 < a_2 < ... < a_k$. The goal is to then obtain reconstructions which take values only from this set. We refer to solutions only taking values from $\mathcal{A}$ as multi-bang solutions as in Definition 6[18]. This setup has real life applications, particularly in medical imaging since there are limited materials expected in the human body [63, 42]. In the case of SPECT imaging the quantity which we would expect to be multi-bang is attenuation. Other potential applications where certain materials can be expected would be in seismic imaging [7] and nuclear security [59]. In [19] the authors originally considered the following pointwise multi-bang penalty term

$$m_0(t) = \begin{cases} \frac{\alpha}{2}\left((a_i + a_{i+1})\,t - a_i a_{i+1}\right) & t \in [a_i, a_{i+1}], 1 \leq i < k, \\ \infty & \text{otherwise.} \end{cases} \tag{2.13}$$

This is the convex envelope of

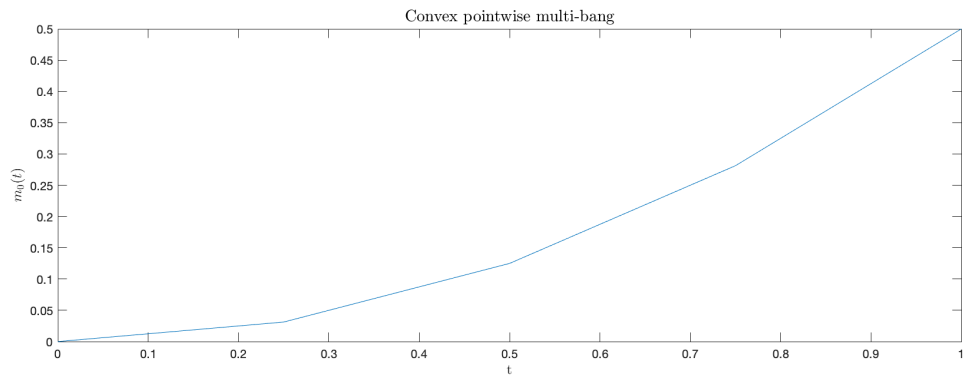
$$m_0^*(t) = \frac{\alpha}{2}t^2 + \beta\Pi_{i=1}^{k}|t - a_i|_0$$

where

$$|t|_0 = \begin{cases} 0 & t = 0 \\ 1 & \text{otherwise} \end{cases}$$

provided α and β satisfy certain criterion given in [19, 54]. See Figure 2.2 for an example plot of $m_0(t)$.

Figure 2.2: A plot of the convex pointwise multi-bang regularizer $m_0(t)$ defined in equation (2.13) with $\alpha = 1$ and $\mathcal{A} = \{0, 0.25, 0.5, 0.75, 1\}$.



The penalty term $m_0^*(t)$ favours values of t in the admissible set U, as they are local minimisers, however it is non-convex and so replacing $m_0^*(t)$ by its convex envelope $m_0(t)$ is computationally advantageous. Although, in the convex penalty each a_i is not a local minima and instead corresponds to a change in gradient. It is important to note that $m_0(t)$ is only the convex envelope of $m_0^*(t)$ when α and β are chosen in accordance with specific rules [19, 18], for more details on what happens in other cases see [19]. The global multi-bang regularizer for a function $v \in L^p(\Omega)$ is then given by

$$\mathcal{M}_0(v) = \int_\Omega m_0(v(t))\mathrm{d}x. \tag{2.14}$$

If we assume that v is piecewise constant over N pixels of length $\mathrm{d}x$ then (2.14) simply becomes

$$\mathcal{M}(v) = \mathrm{d}x \sum_{i=1}^{N} m_0(v_i) \tag{2.15}$$

where v_i is the value of v in the i-th pixel and $\mathrm{d}x$ is the area of each pixel. In the case of uniform pixels we can combine $\mathrm{d}x$ and α into one regularization parameter, which we again denote α.

The multi-bang penalty given in (2.14) is convex, because the pointwise formula (2.13) is convex. However, $\mathcal{M}$ does not promote spatial regularity in the variational problem. In order to obtain a balance between multi-bang solutions and obtain piecewise constant objects, multi-bang regularization can be combined with a Total Variation regularizer [46, 32] as in [18, 54].

The optimization problem (2.9) can be solved via a variety of methods but the approach we take is to solve via Alternating Direction Method of Multipliers(ADMM) [10]. In particular we need to make use of the proximal map of the pointwise multi-bang regularizer. The proximal map of $m_0(t)$, see Definition 25, is given by

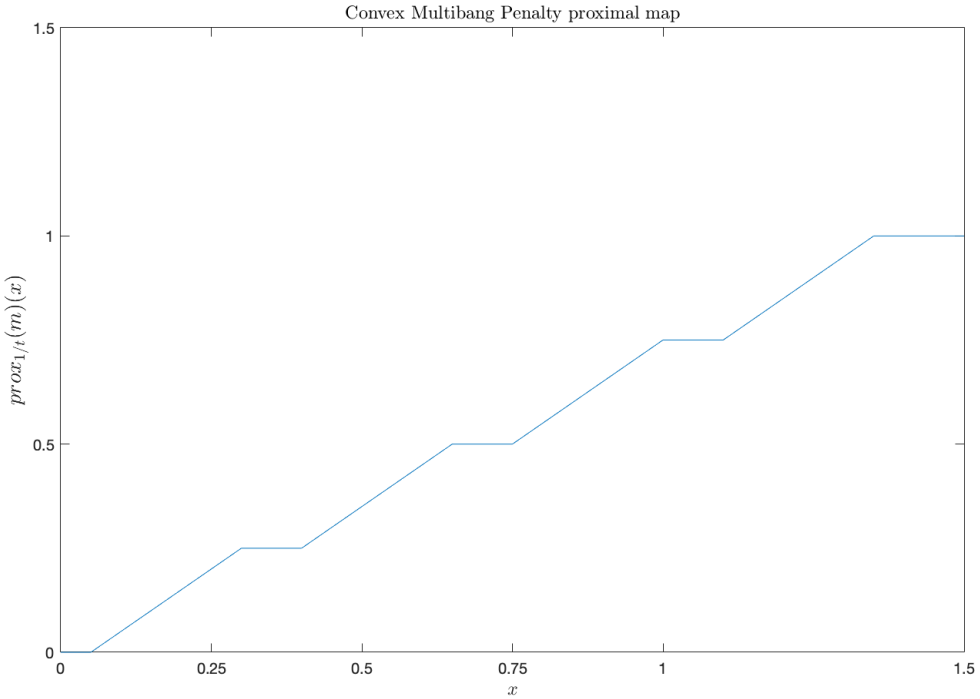
$$\mathrm{prox}_{\frac{1}{t}}(\alpha m_0)(x) = \begin{cases} x + t\alpha a_k & x < x_{0,-} \\ a_i & x_{i,-} \leq x \leq x_{i,+} \\ x - \frac{\alpha t}{2}\left(a_i + a_{i+1}\right) & x_{i,+} < x < x_{i+1,-} \quad \text{for } 1 \leq i \leq k-1 \\ x - t\alpha a_k & x > x_{k,+}, \end{cases} \tag{2.16}$$

where

$$x_{i,+} = a_i + \frac{t\alpha}{2}(a_i + a_{i+1}) \text{ for } i \in 1, ..., k-1 \text{ and } x_{k,+} = a_k + t\alpha a_k$$
$$x_{i,-} = a_i + \frac{t\alpha}{2}(a_{i-1} + a_i) \text{ for } i \in 1, ..., k \text{ and } x_{0,-} = a_0 - t\alpha a_k.$$

This has one important flaw when trying to obtain multi-bang solutions: the proximal map is not stationary at any multi-bang value except the first. This means that unless you choose α extremely small, in which case you are essentially not applying any multi-bang regularization, reconstructions will not be multi-bang. However, the authors of [18] obtain good results in a variety of linear inverse problems using the convex multi-bang regularizer given by equation (2.13). An example plot of the proximal map for m_0 is shown in Figure 2.3.

Figure 2.3: A plot of the proximal map of m_0 given in equation (2.16) with $t = \alpha = 1$ and $\mathcal{A} = \{0, 0.25, 0.5, 0.75, 1\}$.

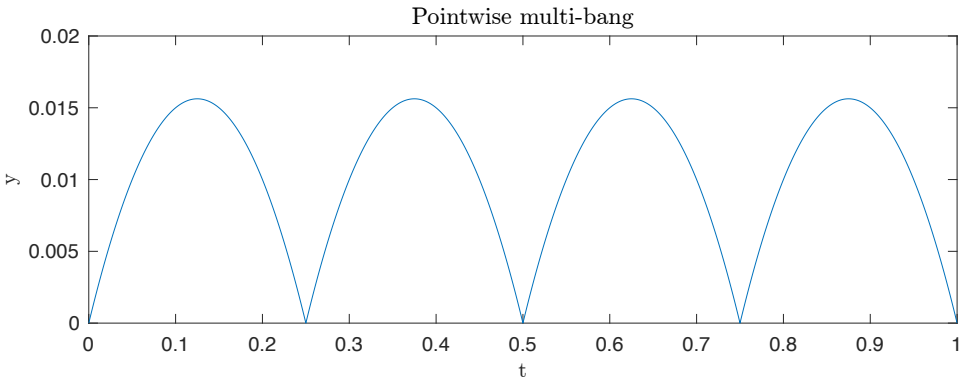


In order to remedy this as in [37], we instead consider the function $m : \mathbb{R} \to \mathbb{R}$ given by

$$m(t) = \begin{cases} (a_{i+1} - t)(t - a_i), & \text{if } t \in [a_i, a_{i+1}] \text{ for some } i \\ \infty, & \text{otherwise.} \end{cases} \tag{2.17}$$

This choice of multi-bang regularizer m is non-convex, and is made up of a series of upside down parabolas with each multi-bang value being a minimiser. A similar penalty is used in [5] when trying to assign integer values to a problem, although we derived this choice of m independently. A plot of m when $\mathcal{A} = \{0, 0.25, 0.5, 0.75, 1\}$ is given in Figure 2.4.

Figure 2.4: Plot of pointwise multi-bang penalty $m(t)$ given in equation (2.17) for $\mathcal{A} = \{0, 0.25, 0.5, 0.75, 1\}$



The multi-bang regularization term $\mathcal{M} : L^r(\Omega) \to \mathbb{R}$ is again given by

$$\mathcal{M}(u) := \int_\Omega m(u(x))\mathrm{d}x, \tag{2.18}$$

for $1 \leq r < \infty$.

Although m is non-convex it is still ρ-weakly convex with $\rho = 2$ as in Definition 10. This weak-convexity property of m allows us to calculate proximal maps, which is why we use this over something like multiple well potentials which are differentiable but not weakly-convex. With this choice of multi-bang regularizer we show that we obtain a proximal map which has multi-bang values as stationary points, this allows much larger values of α to be used in reconstructions and therefore promotes multi-bang solutions.

The next section examines some of the literature relating to when you can uniquely determine a and f from $R_a f$, this is often called the SPECT identification problem [13, 50, 14, 38].

2.5 SPECT identification problem

As mentioned in the introduction, without additional constraints on the allowed set of a and f, unique inversion of the AtRT is known to be impossible [24]. Furthermore, even when smoothness constraints are placed on a, and $a \neq 0$, with a and f radial it can be shown that a solution pair exists for any $R_a f$ of the form $(0, \tilde{f})$ [30, 53, 55, 24].

In the case of medical SPECT imaging we have the additional property that only a small number of tissues are likely to be found in a scan of organic matter. This raises the question: If we can restrict our inversion to the set of a taking values only from the set of known values of tissue attenuation then can we uniquely determine a and f simultaneously from the data? This problem is more commonly referred to as the Identification problem for SPECT [44, 56]. In [48, 49] it has been shown that for $a, f \in \mathrm{L}^{\infty,1+\epsilon}(\mathbb{R}^2, \mathbb{R})$, for some $\epsilon > 0$ where

$$\mathrm{L}^{\infty,\sigma}(\mathbb{R}^2, \mathbb{R}) := \{u \in \mathrm{L}^{\infty}(\mathbb{R}^2, \mathbb{R}) \mid \|u\|_{0,\sigma} < +\infty\},$$
$$\|u\|_{0,\sigma} = \text{ess sup}_{x \in \mathbb{R}^2}(1 + |x|^2)^{\sigma}|u(x)|, \quad \sigma \geq 0,$$

there is an explicit formula for f, when a is known a priori (so it is the identification problem of f only), in terms of the beam transform (2.2) and the following projection operators.

Definition 12. $(\mathcal{H})^{\pm}$ *Let $u \in L_c^{\infty}(\mathbb{R})$ then we define the projection operators $\mathcal{H}^{\pm}$ via*

$$\mathcal{H}^{\pm}u(s) := \frac{1}{\pi}\int_{\mathbb{R}} \frac{v(t)}{s \pm 0i - t}\mathrm{d}t = \lim_{\epsilon \to 0^{\pm}} \frac{1}{\pi}\int_{\mathbb{R}} \frac{v(t)}{s \pm \epsilon i - t}\mathrm{d}t.$$

Assuming that a is known the AtRT uniquely determines f on $\mathbb{R}^2$ by

$$\begin{aligned}
&f(\boldsymbol{x}) = -\frac{1}{4\pi}\left(\frac{\partial}{\partial x_1} - i\frac{\partial}{\partial x_2}\right)\int_{\mathbb{S}^1} \psi(\boldsymbol{x}, \theta)(\theta_1 + \theta_2 i)\ \mathrm{d}\theta,\\
&\psi(\boldsymbol{x}, \theta) = \exp\left\{-Da(\boldsymbol{x}, -\theta)\right\} m(\boldsymbol{x} \cdot \theta^{\perp}, \theta),\\
&m(s, \theta)\\
&= -i\mathrm{Re}\left(\exp\left\{\left(\frac{i}{2}\right)\mathcal{H}^{+}R_0 a(s, \theta)\right\}\mathcal{H}^{+}\left[\exp\left\{-\left(\frac{i}{2}\right)\mathcal{H}^{-}R_0 a(s, \theta)\right\} R_a f\right](s, \theta)\right),
\end{aligned} \tag{2.19}$$

where $\boldsymbol{x} = (x_1, x_2)$ and $R_0 f(s, \theta)$ is the Radon transform of f given by Definition 2 when a is 0. In the case where a is multi-bang, a is clearly bounded and has compact support and therefore $a \in \mathrm{L}^{\infty,1+\epsilon}(\mathbb{R}^2, \mathbb{R})$ for any $\epsilon > 0$. It is worth noting that this

result is independent of the multi-bang values $\mathcal{A}$. Furthermore the assumption that $f \in C_c^1(\Omega)$ (which is an assumption made at the start of Chapter 2.3), as well as the fact that Ω is bounded, gives $f \in \mathrm{L}^{\infty,1+\epsilon}(\mathbb{R}^2, \mathbb{R})$. This means that for the multi-bang case we consider it is enough to be able to determine a and then use (2.19) to uniquely determine f.

In [49, 48] they also obtain a set of conditions known as data consistency conditions or range conditions which are requirements on a which must hold when the data is given. In particular, for $a, f \in \mathrm{L}^{\infty,1+\epsilon}(\mathbb{R}^2, \mathbb{R})$ we have the following range condition

$$\int_{\mathbb{S}^1} \exp\left[-Da(\boldsymbol{x}, -\theta)\right] \mathrm{Re}\left\{g(\boldsymbol{x}, \theta)\right\} \mathrm{d}\theta = 0, \tag{2.20}$$

where

$$g(\boldsymbol{x}, \theta) := \\ \exp\left\{\left(\frac{i}{2}\right) \mathcal{H}^+ R_0 a(\boldsymbol{x} \cdot \theta^\perp, \theta)\right\} \mathcal{H}^+ \exp\left\{\left(\frac{-i}{2} \mathcal{H}^- R_0 a(\boldsymbol{x} \cdot \theta^\perp, \theta)\right)\right\} R_a f(\boldsymbol{x} \cdot \theta^\perp, \theta),$$

for all $x \in \mathbb{R}^2$. Apart from the data $R_a f(x \cdot \theta^\perp, \theta)$, (2.20) is independent of f and so is really a condition on a. There have been multiple papers [3, 38] exploring joint reconstruction of a and f by determining a second independent equation involving f, which they achieve by coupling the formula for f given in (2.19) with a linearization of the consistency condition (2.20). Note that these simultaneous recovery results require a and f to be smooth and compactly supported.

In terms of multi-bang a and differentiable f, some previous work has also been carried out. We discuss two results for when a takes only 2 multi-bang values, i.e is binary. The first result given in [57, 24] states that; if Ω^* is an open convex bounded set, f is supported in Ω^* and a is constant in Ω^* and 0 outside, then the AtRT can be reduced to the exponential X-ray transform [45, 57, 24].

Definition 13 (Exponential X-ray transform). *We define the exponential X-ray transform, for some fixed $c \in \mathbb{R}$ via the map $\mathcal{E}_c : L_c^\infty(\mathbb{R}^2) \longrightarrow L^\infty(\mathbb{R} \times \mathbb{S}^1)$*

$$\mathcal{E}_c f(s, \theta) = \int_{-\infty}^{\infty} e^{ct} f(s\theta^\perp + t\theta) \mathrm{d}t.$$

Let $x + \tau(s, \theta^\perp)\theta$ denote the point at which the ray starting at $x \in \Omega^*$ with orientation θ hits the boundary of Ω^*, note that as f has support in Ω^* which is convex there

is only ever one point where the ray hits the boundary. If $a = c\chi_{\Omega^*}$, c times the characteristic function on Ω^*, then

$$Da(x,\theta) = c(\tau(x\cdot\theta^\perp,\theta^\perp) - x\cdot\theta)$$

for $x \in \Omega^*$, $x\cdot\theta^\perp = s$. Setting $x\cdot\theta = t$ gives

$$\begin{aligned} R_a f(s,\theta) =& \int_{-\infty}^{\infty} e^{-c(\tau(s,\theta^\perp)-t)} f(s\theta^\perp + t\theta)\mathrm{d}t \\ =& e^{-c(\tau(s,\theta^\perp)} \int_{-\infty}^{\infty} e^{ct} f(s\theta^\perp + t\theta)\mathrm{d}t \\ =& e^{-c(\tau(s,\theta^\perp)} \mathcal{E}_c f(s,\theta) \end{aligned} \tag{2.21}$$

where $\mathcal{E}_c f(s,\theta)$ is the exponential X-ray transform as in Definition 13. The exponential X-ray transform has been examined extensively [45, 55, 57, 24] and we refer the reader to these references for more in depth results. We note that although we assume multi-bang a outside of Ω^*, i.e that $a = 0$, this assumption is not necessary outside the convex domain Ω^* for the results in [45, 57, 24] to hold. The results given in [45, 55, 57, 24] do not apply to radial f or circular Ω. In particular [55] proves and gives a method to construct a counterexample to the identification problem for the exponential radon transform for radial f. Indeed, if a is constant on the circle of radius 1 and 0 outside and we take $f(x) = 1 + \cos(\pi|x|)$ inside the unit circle and 0 outside. This choice of a and f are radial and a is nicely multi-bang. Then, by [55] the function

$$u(r) := -\frac{1}{\pi}\int_r^1 \frac{R_a f(s)}{\sqrt{s^2 - r^2}}\mathrm{d}s$$

defines a radial function f, with $|x| = r$, which satisfies $R_0 u(s) = R_a f(s)$ for all $s \in \mathbb{R}$ and is positive and smooth enough [57]. The second result related to the multi-bang a identification problem in SPECT is presented in [16] and shows unique recovery of a and f when a is a multiple of the characteristic function of a star-shaped polygon, f is smooth and the support of f is contained in the support of a.

For the more general case multi-bang a there has been some progress in the identification problem. Sections of Chapters 3 and 4 relate to our most recent work in [37]. In [37] we consider a subset of multi-bang a known as nicely multi-bang.

Definition 14 (Nicely multi-bang). *Let a be multi-bang, we say that a is nicely multi-bang if it can written in the form*

$$a = \sum_{i=1}^{N} c_j \chi_{C_j} \tag{2.22}$$

where the sets C_j are all convex, bounded, open with smooth boundary possibly having corners, as in Definition 5, and nested in the sense that

$$C_n \Subset C_{n-1} \Subset \ \ldots \ \Subset C_1,$$

where $C_j \Subset C_{j-1}$ means that C_j is contained in a compact set that is contained in C_{j-1}.

When a is nicely multi-bang and f is C^1 then we can uniquely determine a and f by the behaviour and direction of the jumps in the AtRT and its derivatives with respect to s and θ across a boundary. This is provided there is at least some section of the outer most boundary ∂C_1 contained in the support of f. Examining these jumps allows recovery of non-zero curvature parts of the boundary as well as corners. An argument involving the Gauss-Bonnet theorem and taking convex hulls as given in [37] allows unique recovery of the boundary. Then applying integral geometry results from [15] we can recover f in regions where a is known. It is important to stress that the integral geometry results rely on the convex nature of the boundaries $\partial\Omega_j$. However, in Chapter 3 we show that examining jumps in the derivatives of the data allows us to uniquely determine the boundary of a multi-bang a in a large variety of non-convex cases. Moreover, we can use the sign of jump in the derivative across the boundary to determine which side of the boundary has a larger attenuation.

The next section examines the literature relating to the inversion formula for the AtRT when a is known.

2.6 Inversion formula for the AtRT

This section relates to the material required to derive the inversion formula for f when a is known given in (2.19).

This technique was originally shown in [28] and involved analyzing the equation

$$\left(\frac{1}{2}\left(\lambda+\frac{1}{\lambda}\right)\partial_{x_1}+\frac{1}{2i}\left(\lambda-\frac{1}{\lambda}\right)\partial_{x_2}\right)\mu(x_1,x_2,\lambda)=f(x_1,x_2)\quad \lambda\in\mathbb{C}\setminus\{0\},\quad (2.23)$$

in order to determine an inversion formula for the Radon transform. Note that when λ is on the unit circle, i.e, $\lambda = 1$ (2.23) becomes

$$\left(\mathrm{Re}(\lambda)\partial_{x_1}+\mathrm{Im}(\lambda)\partial_{x_2}\right)\mu(x_1,x_2,\lambda)=f(x_1,x_2).$$

Using the ideas in [28] the authors of [26] derived the Attenuated Radon Transform inversion formula for f when a is known given equation (2.19). The formula agrees with the original inversion formula obtained by Novikov in [48]. The modification made in [26] is to instead examine the PDE

$$\left(\frac{1}{2}\left(\lambda+\frac{1}{\lambda}\right)\partial_{x_1}+\frac{1}{2i}\left(\lambda-\frac{1}{\lambda}\right)\partial_{x_2}+a(x_1,x_2)\right)\mu(x_1,x_2,\lambda)=f(x_1,x_2), \quad \lambda\in\mathbb{C}\setminus\{0\}. \tag{2.24}$$

The analysis of (2.24) in [26] requires some important results from complex analysis, which we present in the next subsection.

2.6.1 Preliminaries

This subsection contains some necessary lemmas and definitions in order to analyse (2.24). Further information on any of the results in here and detailed proofs can be found in [27].

We first define two important classes of functions called Schwartz functions and Hölder continuous functions.

Definition 15 (Schwartz Function). *The space of Schwartz functions $\mathcal{S}(\mathbb{R}^2)$ is defined as*

$$\mathcal{S}(\mathbb{R}^2)=\left\{f\in C^\infty(\mathbb{R}^2):\|f\|_{\alpha,\beta}<\infty\right\}$$

where $\alpha=(\alpha_1,\alpha_2)$, $\beta=(\beta_1,\beta_2)$ and

$$\|f\|_{\alpha,\beta}:=\sup_{x\in\mathbb{R}^2}\left|x_1^{a_1}x_2^{a_2}\left(\partial_{x_1}^{\beta_1}\partial_{x_2}^{\beta_2}f(x_1,x_2)\right)\right|. \tag{2.25}$$

The Schwartz space $\mathcal{S}(\mathbb{R}^2)$ consists of all smooth functions whose partial derivatives decrease rapidly as $|x|\to\infty$.

Definition 16 (Hölder condition). *A real or complex valued function $f:\mathbb{C}^m\to\mathbb{C}$, where m is the dimension of the Euclidean space, satisfies a Hölder condition if there exist real $C,\alpha>0$, such that*

$$|f(x)-f(y)|\leq C\|x-y\|^\alpha, \tag{2.26}$$

where $z,z'\in\mathbb{C}$ and $\|\cdot\|$ is the standard Euclidean 2-norm.

Functions which satisfy a Hölder condition are called Hölder continuous functions [8, 9, 48].

Throughout the analysis in this section we make use of the analytic property of a complex function. The following result is helpful in proving analyticity of a function and is called Morera's Theorem [20].

Theorem 2 (**Morera's Theorem**). *Let f be a continuous function $f : D \to \mathbb{C}$ for some open set D in the complex plane. Suppose that for every closed triangle $T \subset D$*

$$\int_{\partial T} f(z)\mathrm{d}z = 0 \tag{2.27}$$

where ∂T is the boundary of the triangle T in the anti-clockwise direction, then f is holomorphic on D.

Proof. A proof of this result can be found in [20] □

Further generalizations of Theorem 2 for complex functions of several variables can be found in [20]. We now present some important results from [26]. Let $z = x + iy$, as we can write $x = \frac{z+\bar{z}}{2}$ and $y = \frac{z-\bar{z}}{2}$, then we can consider any function $f(x, y)$ as $f(z, \bar{z})$. Note that as $\bar{z} = x - iy$ does not satisfy the Cauchy-Riemann equations anywhere, $\bar{z}$ is nowhere analytic. Therefore taking the derivative of a function with respect to $\bar{z}$ is really a measure of nonanalyticity. Conversely if $f(z, \bar{z})$ is analytic then $\partial_{\bar{z}} f(z) = 0$. There is rich history of analysis with respect to $\partial_{\bar{z}}$, often known as d-bar analysis [48, 26, 27]. In particular it is an important part of the following result, a generalization of Cauchy's integral theorem, also called Pompieu's formula [52].

Lemma 2 (**Pompieu's formula**). *Let $f(z, \bar{z})$ be a Schwartz function. Let D be a finite closed region with simple boundary in which f and its partial derivatives are continuous. Denote by ∂D the closed boundary of D which is oriented anticlockwise. Then $f(z, \bar{z})$ can be determined for any point $z \in \mathrm{int}D$ by*

$$f(z, \bar{z}) = \frac{1}{2i\pi}\left(\oint_{\partial D} f(\xi, \bar{\xi})\frac{\mathrm{d}\xi}{\xi - z} + \iint_D \frac{\partial f}{\partial \bar{\xi}}(\xi, \bar{\xi})\frac{\mathrm{d}\xi \wedge \mathrm{d}\bar{\xi}}{\xi - z}\right) \tag{2.28}$$

where the wedge product $\mathrm{d}\xi \wedge \mathrm{d}\bar{\xi}$, $\xi = \zeta + i\eta$ is given by

$$\mathrm{d}\xi \wedge \mathrm{d}\bar{\xi} = -2i\mathrm{d}\zeta\mathrm{d}\eta.$$

Proof. A proof of this can be found in [27]. □

In the derivation of the inversion formula for the Attenuated Radon transform we analyse the PDE (2.24) as $\lambda \in \mathbb{C}$ approaches the unit circle from the inside and outside. The following result, known as the Plemelj Formula [51], is extremely useful for analysis when dealing with taking limits as we aproach the unit circle from both directions.

Lemma 3 (Plemelj Formula). *Let Γ be a smooth, oriented, contour, which can open or closed. Suppose $\phi(t)$ satisfies a Hölder continuity condition on Γ. Then, the Cauchy-type integral*

$$\psi(z) = \frac{1}{2\pi i}\int_{\Gamma}\frac{\phi(\tau)}{\tau - z}\mathrm{d}\tau, \tag{2.29}$$

has well defined limiting values $\psi_-(t)$ and $\psi_+(t)$ as z approaches Γ from the right and left respectively, provided that t is not an endpoint of Γ. These limits are given by

$$\psi_{\pm}(t) = \pm\frac{1}{2}\phi(t) + \frac{1}{2\pi i}\mathrm{p.v}\int_{\Gamma}\frac{\phi(\tau)}{\tau - t}\mathrm{d}\tau. \tag{2.30}$$

Where

$$\mathrm{p.v}\int_{\Gamma} f(\tau)\mathrm{d}\tau = \lim_{\epsilon\to 0^+}\int_{\Gamma-\Gamma_\epsilon} f(\tau)\mathrm{d}\tau,$$

with Γ_ϵ the part of the contour Γ centred around t with length 2ϵ.

Proof. A proof of this can be found in [27]. □

A particularly useful application of Lemma 3 is when $\Gamma = \mathbb{R}$ as given below.

Corollary 1. *Let g satisfy a Hölder continuity condition on $\mathbb{R}$ and be integrable. Then the projection operators $\mathcal{H}^{\pm}g(s)$ given in Definition 12 can be written in the form*

$$(\mathcal{H}^{\pm}g)(s) = \mp ig(s) + \mathcal{H}g(s), \tag{2.31}$$

where

$$\mathcal{H}g(s) = \frac{1}{\pi}\mathrm{p.v.}\int_{-\infty}^{\infty}\frac{g(s')}{s - s'}\mathrm{d}s'$$

Proof. By Lemma 3 we have

$$\psi(z) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{g(s')}{z - s'}\mathrm{d}s'$$

has well-defined limits $\mathcal{H}^{\pm}$ as z approaches the real axis from above and below. Furthermore, formula (2.30) (after factoring out a multiple of $(-\frac{1}{2i})$) gives the required result. □

We will see later that the operators given in Corollary 1 occur naturally in the inversion formula for the Attenuated Radon Transform.

We can use Lemma 3 to solve the following Riemann-Hilbert scalar problem.

Lemma 4. *The unique solution of the Riemann-Hilbert problem*

$$\psi_+(t) - \psi_-(t) = g(t), \quad t \in \Gamma, \tag{2.32}$$

where ψ is holomorphic on $\mathbb{C}^1 \setminus \Gamma$,satisfies

$$\psi(z) = O\left(\frac{1}{z}\right), \quad |z| \to \infty, \ z \notin \Gamma, \tag{2.33}$$

and g satisfies as Hölder continuity condition on Γ and is integrable, is given by

$$\psi(z) = \frac{1}{2\pi i}\int_\Gamma \frac{g(\tau)}{\tau - z}\mathrm{d}\tau. \tag{2.34}$$

Proof. We begin by showing that (2.34) is a solution to (2.32) and (2.33). By lemma 3, for any $t \in \Gamma$

$$\psi_\pm(t) = \pm\frac{1}{2}g(t) + \frac{1}{2\pi i}\mathrm{p.v}\int_\Gamma \frac{g(\tau)}{\tau - t}\mathrm{d}\tau.$$

Therefore $\psi_+(t) - \psi_-(t) = g(t)$ for all $t \in \Gamma$ and so we satisfy condition (2.32). Furthermore, as g is bounded on Γ and satisfies a Hölder continuity condition on Γ we have that (2.33) is also satisfied. Hence, we have shown that $\psi(z)$ is a solution of the Riemann-Hilbert problem (2.32) and (2.33).

Now suppose that there is another solution $\nu(z)$ satisfying the Riemann-Hilbert problem. Consider $\psi(z) - \nu(z)$. This function is holomorphic on $\mathbb{C} \setminus \Gamma$, as both ψ and ν are, and is continuous on Γ. The continuity follows from the fact that both ψ and ν satisfy the boundary condition (2.34). We aim to apply Theorem 2 to $\psi(z) - \nu(z)$ in order to show that $\psi(z) - \nu(z)$ is entire. Let T be any triangle which does not contain any part of the contour Γ,i.e., T lies entirely to the right or the left of the contour Γ. As $\psi(z) - \nu(z)$ is holomorphic in T, we have

$$\int_{\partial T} \psi(z) - \nu(z)\mathrm{d}z = 0.$$

The remaining case to consider is when T contains some portion of the contour Γ. In this case we can approximate the integral by splitting the triangle into two parts T_1 and T_2, which lie to the left and right of the contour respectively. Again, as $\psi(z) - \nu(z)$ is holomorphic off the contour Γ the only possible contributions around the contours

∂T_1 and ∂T_2 will come from $\int_{\partial T_i \cap \Gamma} \psi(z) - \nu(z) \mathrm{d}z$ for $i = 1, 2$. By construction we have $\partial T_1 \cap \Gamma = -\partial T_2 \cap \Gamma$, where $-$ refers to the reverse path. Finally by continuity of $\psi(z) - \nu(z)$ we have that

$$\int_{\partial_{T_1} \cap \Gamma} \psi(z) - \nu(z) \mathrm{d}z = -\int_{\partial_{T_2} \cap \Gamma} \psi(z) - \nu(z) \mathrm{d}z$$

and so $\int_{\partial T} \psi(z) - \nu(z) \mathrm{d}z = 0$ for any closed triangle T in $\mathbb{C}$. Hence by Theorem 2 we have $\psi(z) - \nu(z)$ is an entire function. Furthermore we have $\psi(z) - \nu(z)$ is bounded and so by the Liouville theorem we have that $\psi(z) - \nu(z)$ is a constant. Finally the constraint that $\psi(z) - \nu(z) = O(\frac{1}{z})$ as $z \to \infty$ gives that this constant must be 0 and so $\psi(z) = \nu(z)$, as required. □

With these preliminary results established we now derive the inversion formula for the Attenuated Radon Transform.

2.6.2 Inversion formula for Attenuated Radon Transform

Although the inversion formula is obtained by analysis of the PDE (2.24), we need the following result from [27] obtained by analyzing (2.23) first.

Lemma 5. *Define a complex variable z by*

$$z := \frac{1}{2i}\left(\lambda - \frac{1}{\lambda}\right) x_1 - \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right) x_2 \tag{2.35}$$

where x_1, x_2 are the real Cartesian coordinates, and λ is a complex variable with $\lambda \neq 0$. Suppose that $a(x_1, x_2) \in \mathcal{S}(\mathbb{R}^2)$. Let $\mu(x_1, x_2, \lambda)$ satisfy the PDE (2.24) with the boundary condition $\mu = O\left(\frac{1}{z}\right)$ as $|x_1| + |x_2| \to \infty$. Let λ^+ and λ^- denote the limits of λ as it approaches the unit circle from inside and outside the unit circle respectively, i.e.

$$\lambda^{\pm} = \lim_{\epsilon \to 0^+} (1 \mp \epsilon) e^{i\omega}, \quad \epsilon > 0, \quad 0 \leq \omega \leq 2\pi. \tag{2.36}$$

Then

$$\mu(x_1, x_2, \lambda^{\pm}) = \pm \frac{1}{2i} \mathcal{H}^{\mp} R_0 a(s, \theta) - \int_{\tau}^{\infty} a(s\theta^{\perp} + t'\theta) \mathrm{d}t', \tag{2.37}$$

where $(x_1, x_2) = s\theta^{\perp} + t\theta$ with $\theta = (\cos(\omega), \sin(\omega)) := (\theta_1, \theta_2)$, $\theta^{\perp} = (-\theta_2, \theta_1)$ and $\mathcal{H}^{\pm}$ are projection operators in the variable s

$$(\mathcal{H}^{\pm} g)(s, \theta) = \lim_{\epsilon \to 0^+} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{g(s', \theta)}{(s \pm i\epsilon) - s'} \mathrm{d}s' = \mp i g(s, \theta) + \mathcal{H} g(s, \theta), \tag{2.38}$$

where

$$\mathcal{H}g(s,\theta) = \frac{1}{\pi}\text{p.v.}\int_{-\infty}^{\infty}\frac{g(s',\theta)}{s-s'}\mathrm{d}s'.$$

Proof. We follow the proof given in [27]. Note that the PDEs (2.23) and (2.24) motivate the choice of z. Taking the complex conjugate of (2.35) gives

$$\bar{z} := -\frac{1}{2i}\left(\bar{\lambda}-\frac{1}{\bar{\lambda}}\right)x_1 - \frac{1}{2}\left(\bar{\lambda}+\frac{1}{\bar{\lambda}}\right)x_2. \tag{2.39}$$

Then $z, \bar{z}$ describes a change of coordinates from (x_1, x_2). We can use this to rewrite the PDE (2.23). By the chain rule we have

$$\begin{aligned}\partial_{x_1} &= \frac{1}{2i}\left(\lambda-\frac{1}{\lambda}\right)\partial_z - \frac{1}{2i}\left(\bar{\lambda}-\frac{1}{\bar{\lambda}}\right)\partial_{\bar{z}}\\ \partial_{x_2} &= -\frac{1}{2}\left(\lambda+\frac{1}{\lambda}\right)\partial_z - \frac{1}{2}\left(\bar{\lambda}+\frac{1}{\bar{\lambda}}\right)\partial_{\bar{z}}.\end{aligned}$$

This then gives

$$\begin{aligned}&\frac{1}{2}\left(\lambda+\frac{1}{\lambda}\right)\left\{\frac{1}{2i}\left(\lambda-\frac{1}{\lambda}\right)\partial_z - \frac{1}{2i}\left(\bar{\lambda}-\frac{1}{\bar{\lambda}}\right)\partial_{\bar{z}}\right\} =\\ &\frac{1}{4i}\left\{\left(\lambda+\frac{1}{\lambda}\right)\left(\lambda-\frac{1}{\lambda}\right)\partial_z - \left(\lambda+\frac{1}{\lambda}\right)\left(\bar{\lambda}-\frac{1}{\bar{\lambda}}\right)\partial_{\bar{z}}\right\}\end{aligned} \tag{2.40}$$

and

$$\begin{aligned}&\frac{1}{2}\left(\lambda+\frac{1}{\lambda}\right)\left\{\frac{1}{2i}\left(\lambda-\frac{1}{\lambda}\right)\partial_z - \frac{1}{2i}\left(\bar{\lambda}-\frac{1}{\bar{\lambda}}\right)\partial_{\bar{z}}\right\} =\\ &\frac{-1}{4i}\left\{\left(\lambda+\frac{1}{\lambda}\right)\left(\lambda-\frac{1}{\lambda}\right)\partial_z - \left(\lambda-\frac{1}{\lambda}\right)\left(\bar{\lambda}+\frac{1}{\bar{\lambda}}\right)\partial_{\bar{z}}\right\}.\end{aligned} \tag{2.41}$$

Using (2.40) and (2.41) we have

$$\begin{aligned}\frac{1}{2}\left(\lambda+\frac{1}{\lambda}\right)\partial_{x_1} + \frac{1}{2i}\left(\lambda-\frac{1}{\lambda}\right)\partial_{x_2} &= \\ &-\frac{1}{4i}\left\{\left(\lambda+\frac{1}{\lambda}\right)\left(\bar{\lambda}-\frac{1}{\bar{\lambda}}\right) + \left(\lambda-\frac{1}{\lambda}\right)\left(\bar{\lambda}+\frac{1}{\bar{\lambda}}\right)\right\}\partial_{\bar{z}}\\ &= \frac{1}{2i}\left(\frac{1}{|\lambda|^2}-|\lambda|^2\right)\partial_{\bar{z}}.\end{aligned} \tag{2.42}$$

Then rewriting (2.23) using (2.42) means $\mu(x_1, x_2, \lambda)$ must satisfy

$$\frac{1}{2i}\left(\frac{1}{|\lambda|^2}-|\lambda|^2\right)\frac{\partial\mu(x_1,x_2,\lambda)}{\partial\bar{z}} = a(x_1,x_2), \quad |\lambda|\neq 1, \quad (x_1,x_2)\in\mathbb{R}^2. \tag{2.43}$$

For now keep λ fixed and assume that

$$\frac{\partial\mu(z_r,z_i)}{\partial\bar{z}} = g(z_r,z_i)$$

with $z = z_r + iz_i$ with both z_r and z_i real and finite. By assumption we also have the boundary condition $\mu = O(1/z)$ as $|z| \to \infty$. Provided that g is a Schwartz function, we can apply Lemma 2 to obtain

$$\mu(z_R, z_I) = -\frac{1}{\pi} \iint_{R^2} \frac{g(z'_R, z'_I)}{z' - z} \mathrm{d}z'_R \mathrm{d}z'_I.$$

Since

$$\begin{aligned} z_R &= \mathrm{Re}\left(\frac{1}{2i}\left(\lambda - \frac{1}{\lambda}\right) x_1 - \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right) x_2\right) \\ &= \frac{x_1}{2}\mathrm{Im}\left(\lambda - \frac{1}{\lambda}\right) - \frac{x_2}{2}\mathrm{Re}\left(\lambda + \frac{1}{\lambda}\right) \end{aligned}$$

and

$$\begin{aligned} z_I &= \mathrm{Im}\left(\frac{1}{2i}\left(\lambda - \frac{1}{\lambda}\right) x_1 - \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right)\right) \\ &= \frac{-x_1}{2}\mathrm{Re}\left(\lambda - \frac{1}{\lambda}\right) - \frac{x_2}{2}\mathrm{Im}\left(\lambda + \frac{1}{\lambda}\right) \end{aligned}$$

we have that

$$\mathrm{d}z'_R \mathrm{d}z'_I = \frac{1}{4}\left|\frac{1}{|\lambda|^2} - |\lambda|^2\right| \mathrm{d}x'_1 \mathrm{d}x'_2. \tag{2.44}$$

Combining (2.44), setting $g(z_R, z_I) = \frac{2ia}{\frac{1}{|\lambda|^2} - |\lambda|^2}$ and the fact that $|z|/z = \mathrm{sgn}(z)$ gives

$$\mu(x_1, x_2, \lambda) = \frac{1}{2\pi i}\mathrm{sgn}\left(\frac{1}{|\lambda|^2} - |\lambda|^2\right) \iint_{R^2} \frac{a(x'_1,' x'_2)}{z' - z} \mathrm{d}x'_1 \mathrm{d}x'_2, \quad |\lambda| \neq 1. \tag{2.45}$$

This gives us an integral representation of μ in terms of the function a for all complex values of λ, provided $|\lambda| \neq 1$. Note that the only dependence of μ on λ is in the sign of the integral and in $z - z'$, hence we see that μ is piecewise analytic inside and outside the unit circle and there is a jump in μ across the unit circle $|\lambda| = 1$.

We next analyse the analyticity properties of μ with respect to λ in order to determine an alternative representation for μ.

Firstly note that by (2.35) we have

$$z' - z = \frac{1}{2i}\left(\lambda - \frac{1}{\lambda}\right)(x'_1 - x_1) - \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right)(x'_2 - x_2). \tag{2.46}$$

Combining (2.46) with (2.45) implies that $\mu = O(1/\lambda)$ as $|\lambda| \to \infty$ and hence μ is bounded for all complex λ. In order to examine the behaviour of μ as λ approaches the unit circle we set

$$\lambda^{\pm}(\epsilon) = (1 \mp \epsilon)e^{i\omega}, \quad 0 \leq \omega < 2\pi, \quad \epsilon > 0. \tag{2.47}$$

and let $\lambda^{\pm} = \lim_{\epsilon \to 0^+} \lambda^{\pm}(\epsilon)$. Note that all $\lambda^{+}(\epsilon)$ are inside the unit circle and all $\lambda^{-}(\epsilon)$ are outside the unit circle. Therefore, taking the limit as $\epsilon \to 0^+$ for λ^+ and λ^- would give us the limits of the function μ as λ approaches the unit circle from the inside and outside respectively. Using (2.47) we have

$$\lambda^{+} \mp \frac{1}{\lambda^{+}} = (1-\epsilon)e^{i\omega} \mp (1+\epsilon)e^{-i\omega} + O(\epsilon^2) \tag{2.48}$$

and

$$\lambda^{-} \mp \frac{1}{\lambda^{-}} = (1+\epsilon)e^{i\omega} \mp (1-\epsilon)e^{-i\omega} + O(\epsilon^2) \tag{2.49}$$

as $\epsilon \to 0^+$, where we have used the Taylor series expansion for $\frac{1}{1 \mp \epsilon}$. Putting $\lambda^{\pm}$ into (2.46) gives

$$\begin{aligned} z - z' =& (x_1' - x_1)\sin(\omega) - (x_2' - x_2)\cos(\omega) \\ & \pm i\epsilon\left((x_1' - x_1)\cos(\omega) + (x_2' - x_2)\sin(\omega)\right) + O(\epsilon^2). \end{aligned} \tag{2.50}$$

Recall that any $\boldsymbol{x} = (x_1, x_2)$ can be written in terms of the orthogonal coordinate vectors $\theta = (\cos(\omega), \sin(\omega))$ and $\theta^{\perp} = (-\sin(\omega), \cos(\omega))$ via

$$\boldsymbol{x} = s\theta^{\perp} + t\theta. \tag{2.51}$$

Equation (2.51) gives

$$\begin{aligned} t &= \cos(\omega)x_1 + \sin(\omega)x_2 \\ s &= -\sin(\omega)x_1 + \cos(\omega)x_2. \end{aligned} \tag{2.52}$$

Combining (2.52) and (2.50) yields

$$z' - z = (s - s') \pm i\epsilon(t' - t) + O(\epsilon^2). \tag{2.53}$$

Suppose that, for now, we only consider λ^+. Substituting (2.53) into (2.45) and taking the limit as $\epsilon \to 0$ gives

$$\mu(x_1, x_2, \lambda^{+}) = \lim_{\epsilon \to 0^+} -\frac{1}{2\pi i} \iint_{\mathbb{R}^2} \frac{a(x_1', x_2')}{s' - (s - i\epsilon(t' - t))} \mathrm{d}x_1' \mathrm{d}x_2'. \tag{2.54}$$

As the Jacobian of the transformation $(x_1, x_2) \mapsto (s, t)$ is 1 we have that $\mathrm{d}x_1'\mathrm{d}x_2' = \mathrm{d}t'\mathrm{d}s'$ and hence

$$\mu(x_1, x_2, \lambda^{+}) = \lim_{\epsilon \to 0^+} -\frac{1}{2\pi i} \iint_{\mathbb{R}^2} \frac{a(s'\theta^{\perp} + t'\theta)}{s' - (s - i\epsilon(t' - t))} \mathrm{d}t' \mathrm{d}s'. \tag{2.55}$$

We next simplify the expression by splitting the integral over $\mathrm{d}t'$ into two parts

$$\int_{-\infty}^{\infty} \mathrm{d}t' = \int_{-\infty}^{t} \mathrm{d}t' + \int_{t}^{\infty} \mathrm{d}t'.$$

The advantage gained in splitting the integral in this way is that in the first $t' - t < 0$ and in the second $t' - t > 0$. Applying Corollary 1 to

$$\mu(s,t,\lambda^+) = -\frac{1}{2\pi i} \lim_{\epsilon \to 0^+} \int_{-\infty}^{\infty} \Bigg\{ \int_{-\infty}^{t} \frac{a(s'\theta^\perp + t'\theta)\mathrm{d}t'}{s' - (s - i\epsilon(t'-t))} + \int_{t}^{\infty} \frac{a(s'\theta^\perp + t'\theta)\mathrm{d}t'}{s' - (s - i\epsilon(t'-t))} \Bigg\} \mathrm{d}s'$$

gives, after swapping orders of integration,

$$\begin{aligned} \mu(s,t,\lambda^+) &= -\frac{1}{2\pi i} \int_{-\infty}^{t} \left\{ -\pi i a(s\theta^\perp + t'\theta) + \pi \mathcal{H} a(s,t',\theta) \right\} \mathrm{d}t' \\ &\quad - \frac{1}{2\pi i} \int_{t}^{\infty} \left\{ \pi i a(s\theta^\perp + t'\theta) + \pi \mathcal{H} a(s,t',\theta) \right\} \mathrm{d}t' \\ &= -\frac{1}{2i} \int_{-\infty}^{\infty} \mathcal{H} a(s,t',\theta)\mathrm{d}t' + \frac{1}{2} \int_{-\infty}^{t} a(s\theta^\perp + t'\theta)\mathrm{d}t' \\ &\quad - \frac{1}{2} \int_{t}^{\infty} a(s\theta^\perp + t'\theta)\mathrm{d}t'. \end{aligned} \tag{2.56}$$

Then adding and subtracting $\frac{1}{2}\int_t^\infty a(s\theta^\perp + t'\theta)\mathrm{d}t'$ to (2.56) gives

$$\mu(s,t,\lambda^+) = -\frac{1}{2i} \int_{-\infty}^{\infty} \mathcal{H} a(s,t',\theta)\mathrm{d}t' + \frac{1}{2} R_0 a(s,\theta) - \int_{t}^{\infty} a(s\theta^\perp + t'\theta)\mathrm{d}t'.$$

The first two terms on the right hand side of this equal $\frac{1}{2i}\mathcal{H}^- R_0 a$. A similar argument can be made for the case involving λ^-. □

Although Lemma 5 relates more directly to (2.23), we have the following result as an immediate consequence of Lemma 5 which is useful in analysis of (2.24).

Corollary 2. *Let z and $\bar{z}$ be as in Lemma 5. Let $\lambda^\pm$ as defined in* (2.47) *be the limits of λ as λ approaches the unit circle from the inside and outside respectively and suppose that $\mu(x_1, x_2, \lambda)$ satisfies* (2.24)*. Set $\gamma(\lambda) = \frac{1}{2i}\left(\frac{1}{|\lambda|^2} - |\lambda|^2\right)$. Then,*

$$\lim_{\lambda \to \lambda^\pm} \left(\partial_{\bar{z}}^{-1} \left\{ \frac{a(x_1,x_2)}{\gamma(\lambda)} \right\} \right) = \pm \frac{1}{2i} (\mathcal{H}^\mp R_0 a)(s,\theta) - \int_t^\infty a(s\theta^\perp + t'\theta)\mathrm{d}t' \tag{2.57}$$

for $(x_1, x_2) = s\theta^\perp + t\theta$.

Proof. In the proof of Lemma 5 we obtained (2.43) which implies that

$$\mu(x_1, x_2, \lambda) = \partial_{\bar{z}}^{-1} \left\{ \frac{a(x_1, x_2)}{\gamma(\lambda)} \right\}. \tag{2.58}$$

Taking the limit of (2.58) as λ approaches the unit circle from the inside and the outside yields

$$\mu(x_1, x_2, \lambda^{\pm}) = \lim_{\lambda \to \lambda^{\pm}} \partial_{\bar{z}}^{-1} \left\{ \frac{a(x_1, x_2)}{\gamma(\lambda)} \right\}. \tag{2.59}$$

Finally combining (2.37) and (2.59) completes the proof. □

Note that Corollary 2 gives a formula for the operator $\partial_{\bar{z}}^{-1}$ in the limit where $\lambda \to \lambda^{\pm}$. We can use the formula give in Corollary 2 directly in analyzing (2.24). Before presenting the final result we need the following lemma relating to the recovery of μ from its jump across the unit circle.

Lemma 6. *Let z and $\bar{z}$ be as in Lemma 5. Again suppose that μ satisfies*

$$\frac{1}{2i} \left(\frac{1}{|\lambda|^2} - |\lambda|^2 \right) \frac{\partial \mu(x_1, x_2, \lambda)}{\partial \bar{z}} = a(x_1, x_2), |\lambda| \neq 1, \qquad (x_1, x_2) \in \mathbb{R}^2$$

with the boundary condition $\mu = O(\frac{1}{z})$ as $|z| \to \infty$. Then

$$\mu(x_1, x_2, \lambda) = \frac{1}{2\pi} \int_0^{2\pi} \frac{J(x_1, x_2, \omega) e^{i\omega}}{e^{i\omega} - \lambda} \mathrm{d}\omega \tag{2.60}$$

where $J(x_1, x_2, \omega) = \mu(x_1, x_2, \lambda^+) - \mu(x_1, x_2, \lambda^-)$.

Proof. From the proof of Lemma 5 we have the following formula for μ

$$\mu(x_1, x_2, \lambda) = \frac{1}{2\pi i} \mathrm{sgn} \left(\frac{1}{|\lambda|^2} - |\lambda|^2 \right) \iint_{R^2} \frac{a(x_1', ' x_2')}{z' - z} \mathrm{d}x_1' \mathrm{d}x_2', \quad |\lambda| \neq 1.$$

Note that the dependence of μ on λ in the above is only in the sign function and in $z' - z$. As shown in the proof of Lemma 5 we have that

$$z' - z = O(\lambda)$$

as $\lambda \to \infty$. Therefore we have that $\mu = O\left(\frac{1}{\lambda}\right)$ as $\lambda \to \infty$. Now we define

$$J(x_1, x_2, \lambda) = \mu(x_1, x_2, \lambda^+) - \mu(x_1, x_2, \lambda^-) \tag{2.61}$$

with $\lambda^{\pm}$ as in Lemma 5. Combining (2.61) and the boundary condition $\mu = O\left(\frac{1}{\lambda}\right)$ forms a scalar Riemann-Hilbert problem. By Lemma 4 we have

$$\mu(x_1, x_2, \lambda) = \frac{1}{2\pi i} \int_{|\lambda'|=1} \frac{J(x_1, x_2, \lambda')}{\lambda' - \lambda} \mathrm{d}\lambda'. \tag{2.62}$$

As $\lambda' = e^{i\omega}$, $\mathrm{d}\lambda' = ie^{i\omega}\mathrm{d}\omega$. We can therefore rewrite (2.62) as

$$\mu(x_1, x_2, \lambda) = \frac{1}{2\pi}\int_0^{2\pi} \frac{J(x_1, x_2, \omega)e^{i\omega}}{e^{i\omega} - \lambda}\mathrm{d}\omega,$$

as required. □

With these results established we now ready to present the inversion formula for the attenuated Radon transform with known a.

Theorem 3. *Let $R_a f$ be the attenuated Radon transform of a Schwartz function $f(x_1, x_2)$ attenuated by the Schwartz function $a(x_1, x_2)$. Then*

$$\begin{aligned} &f(x_1, x_2) = \\ &\frac{1}{4\pi}\left(\partial_{x_1} - i\partial_{x_2}\right)\int_{\mathbb{S}^1} \exp\left(\int_t^\infty a(s\theta^\perp + t'\theta)\mathrm{d}t'\right) M(x_1, x_2, \theta)(\theta_1 + i\theta_2)\mathrm{d}\theta, \end{aligned} \tag{2.63}$$

where

$$\begin{aligned} &M(x_1, x_2, \omega) \\ &= \frac{-1}{2i}\Bigg\{ \exp\left(\frac{-1}{2i}\mathcal{H}^- R_0 a(s,\theta)\right)\ \mathcal{H}^-\left\{\exp\left(\frac{1}{2i}\mathcal{H}^- R_0 a(s,\theta)\right) R_a f\right\}(s,\theta) \\ &+ \exp\left(\frac{1}{2i}\mathcal{H}^+ R_0 a(s,\theta)\right)\mathcal{H}^+\left\{\exp\left(\frac{-1}{2i}\mathcal{H}^+ R_0 a(s,\theta)\right) R_a f\right\}(s,\theta)\Bigg\} \end{aligned} \tag{2.64}$$

$\theta = (\theta_1, \theta_2) = (\cos(\omega), \sin(\omega))$ *and* $(x_1, x_2) = s\theta^\perp + t\theta$.

Proof. Let z and $\bar{z}$ be defined by equations (2.35) and (2.39) respectively. For simplicity, as in Corollary 2, we again set

$$\gamma(\lambda) = \frac{1}{2i}\left(\frac{1}{|\lambda|^2} - |\lambda|^2\right).$$

Then we can rewrite (2.24), in a similar manner to (2.23) in the proof of Lemma 5, to find

$$\frac{\partial}{\partial\bar{z}}\mu(x_1, x_2, \lambda) + \frac{a(x_1, x_2)\mu(x_1, x_2, \lambda)}{\gamma(\lambda)} = \frac{f(x_1, x_2)}{\gamma(\lambda)}. \tag{2.65}$$

We can multiply (2.65) by the integrating factor $\exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a(x_1,x_2)}{\gamma(\lambda)}\right)\right)$ to get

$$\begin{aligned} &\frac{\partial}{\partial\bar{z}}\left\{\mu(x_1, x_2, \lambda)\exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a(x_1, x_2)}{\gamma(\lambda)}\right)\right)\right\} \\ &\quad = \exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a(x_1, x_2)}{\gamma(\lambda)}\right)\right)\frac{f(x_1, x_2)}{\gamma(\lambda)}. \end{aligned} \tag{2.66}$$

In a similar manner to Lemma 5, by setting $z = z_r + iz_i$ (provided that the right hand side of (2.66) is a Schwartz function) we can apply Lemma 2 to find

$$\begin{aligned}&\mu(x_1,x_2,\lambda)\exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a(x_1,x_2)}{\gamma(\lambda)}\right)\right)=\\&\frac{1}{2\pi i}\mathrm{sgn}\left(\frac{1}{|\lambda|^2}-|\lambda|^2\right)\iint_{\mathbb{R}^2}\frac{\exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a(x_1',x_2')}{\gamma(\lambda)}\right)\right)f(x_1',x_2')}{z'-z}\mathrm{d}x_1'\mathrm{d}x_2'.\end{aligned}\tag{2.67}$$

We can divide both sides of (2.67) by the integrating factor $\exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a}{\gamma(\lambda)}\right)\right)$ to obtain a representation for μ. There are two important conclusions which can be made from (2.67). The first is that, as in the case of Lemma 5, μ is a sectionally analytic function of λ (with there being two regions: one inside and the other outside the unit circle) with a jump over the unit circle $|\lambda| = 1$. The second is that the dependence of μ on λ is only in the sign function, the integrating factor and $z' - z$. Since $z' - z = O\left(\frac{1}{\lambda}\right)$ as $\lambda \to \infty$ and the order of λ of the integrating factor cancels on both sides we find that $\mu(x_1, x_2, \lambda) \to O\left(\frac{1}{\lambda}\right)$ as $\lambda \to \infty$. Therefore we can apply the same logic as in Lemma 6 to deduce that

$$\mu(x_1,x_2,\lambda)=\frac{1}{2\pi}\int_0^{2\pi}\frac{J(x_1,x_2,\omega)e^{i\omega}}{e^{i\omega}-\lambda}\mathrm{d}\omega,\tag{2.68}$$

where $J(x_1, x_2, \omega) = \mu(x_1, x_2, \lambda^+) - \mu(x_1, x_2, \lambda^-)$ with $\lambda^\pm$ as in Lemma 5. We now determine this jump function J. From (2.66) we have

$$\begin{aligned}&\mu(x_1,x_2,\lambda)\exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a(x_1,x_2)}{\gamma(\lambda)}\right)\right)=\\&\partial_{\bar{z}}^{-1}\left\{\exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a(x_1,x_2)}{\gamma(\lambda)}\right)\right)\frac{f(x_1,x_2)}{\gamma(\lambda)}\right\}.\end{aligned}\tag{2.69}$$

In the limit $\lambda \to \lambda^\pm$, applying Corollary 2 and using the parametrization of (x_1, x_2) in terms of s and t gives

$$\begin{aligned}&\mu(x_1,x_2,\lambda^\pm)\exp\left(\pm\frac{1}{2i}\mathcal{H}^\mp R_0a(s,\theta)-\int_t^\infty(a\theta^\perp+t'\theta)\mathrm{d}t'\right)=\\&\lim_{\lambda\to\lambda^\pm}\partial_{\bar{z}}^{-1}\left\{\exp\left(\pm\frac{1}{2i}\mathcal{H}^\mp R_0a(s,\theta)-\int_t^\infty(a\theta^\perp+t'\theta)\mathrm{d}t'\right)\frac{f(x_1,x_2)}{\gamma(\lambda)}\right\}.\end{aligned}\tag{2.70}$$

We can then apply Corollary 2 to the right hand side of (2.70) to find

$$\begin{aligned}&\mu(x_1,x_2,\lambda^\pm)\exp\left(\pm\frac{1}{2i}\mathcal{H}^\mp R_0a(s,\theta)-\int_t^\infty a(s\theta^\perp+t'\theta)\mathrm{d}t'\right)=\\&\pm\frac{1}{2i}\mathcal{H}^\mp\left\{\exp\left(\pm\frac{1}{2i}\mathcal{H}^\mp R_0a(s,\theta)\right)R_af\right\}(s,\theta)\\&-\int_t^\infty f(s\theta^\perp+t'\theta)\exp(\pm\frac{1}{2i}\mathcal{H}^\mp R_0a(s,\theta))\exp\left(-\int_{t'}^\infty a(s\theta^\perp+\tau\theta)\mathrm{d}\tau\right)\mathrm{d}t'.\end{aligned}\tag{2.71}$$

Note that the second line in (2.71) is valid because $\mathcal{H}^{\mp}R_0a(s,\theta)$ is independent of t', so can be brought outside the Radon transform and the remaining terms are precisely the definition of the Attenuated Radon Transform. Since $\exp(\pm\frac{1}{2i}\mathcal{H}^{\mp}R_0a(s,\theta))$ is independent of t' we can factor it out of the integral on the right hand side of (2.71). After doing this and dividing to isolate $\mu(x_1,x_2,\lambda^{\pm})$ we obtain

$$\begin{aligned}
&\mu(x_1,x_2,\lambda^{\pm}) = \\
&\pm\frac{1}{2i}\exp\left(\int_t^{\infty} a(s\theta^{\perp}+t'\theta)\mathrm{d}t'\right)\times \\
&\left[\left\{\exp\left(\mp\frac{1}{2i}\mathcal{H}^{\mp}R_0a(s,\theta)\right)\ \mathcal{H}^{\mp}\left\{\exp\left(\pm\frac{1}{2i}\mathcal{H}^{\mp}R_0a(s,\theta)\right)R_af\right\}(s,\theta)\right] \\
&-\int_t^{\infty} f(s\theta^{\perp}+t'\theta)\exp\left(-\int_{t'}^{\infty} a(s\theta^{\perp}+\tau\theta)\mathrm{d}\tau\right)\mathrm{d}t'\Bigg\}.
\end{aligned} \tag{2.72}$$

Therefore the jump function $J(x_1,x_2,\omega)$ is given by

$$\begin{aligned}
&J(x_1,x_2,\omega) = \\
&\frac{1}{2i}\exp\left(\int_t^{\infty} a(s\theta^{\perp}+t'\theta)\mathrm{d}t'\right)\times \\
&\left\{\exp\left(\frac{-1}{2i}\mathcal{H}^{-}R_0a(s,\theta)\right)\ \mathcal{H}^{-}\left\{\exp\left(\frac{1}{2i}\mathcal{H}^{-}R_0a(s,\theta)\right)R_af\right\}(s,\theta)\right. \\
&\left.+\exp\left(\frac{1}{2i}\mathcal{H}^{+}R_0a(s,\theta)\right)\mathcal{H}^{+}\left\{\exp\left(\frac{-1}{2i}\mathcal{H}^{+}R_0a(s,\theta)\right)R_af\right\}(s,\theta).\right\}
\end{aligned} \tag{2.73}$$

By (2.68) and the Laurent series of $\frac{1}{1-\frac{e^{i\omega}}{\lambda}}$ for $\lambda\in\mathbb{C}^1$ with $|\lambda|>1$ we have

$$\mu(x_1,x_2,\lambda) = -\left(\frac{1}{2\pi}\int_0^{2\pi} J(x_1,x_2,\omega)e^{i\omega}\mathrm{d}\omega\right)\frac{1}{\lambda}+O\left(\frac{1}{\lambda^2}\right) \tag{2.74}$$

as $\lambda\to\infty$. Subbing (2.74) into the PDE (2.24) and letting $\lambda\to\infty$ gives

$$f(x_1,x_2) = -\frac{1}{4\pi}(\partial_{x_1}-i\partial_{x_2})\int_0^{2\pi} J(x_1,x_2,\omega)e^{i\omega}\mathrm{d}\omega. \tag{2.75}$$

Finally setting

$$\begin{aligned}
&M(x_1,x_2,\theta) = \\
&\frac{-1}{2i}\left\{\exp\left(\frac{-1}{2i}\mathcal{H}^{-}R_0a(s,\theta)\right)\ \mathcal{H}^{-}\left\{\exp\left(\frac{1}{2i}\mathcal{H}^{-}R_0a(s,\theta)\right)R_af\right\}(s,\theta)\right. \\
&\left.+\exp\left(\frac{1}{2i}\mathcal{H}^{+}R_0a(s,\theta)\right)\mathcal{H}^{+}\left\{\exp\left(\frac{-1}{2i}\mathcal{H}^{+}R_0a(s,\theta)\right)R_af\right\}(s,\theta)\right\}
\end{aligned}$$

and noting $\theta=\cos(\omega)+i\sin(\omega)$ completes the proof. □

We note that Theorem 3 only holds for Schwartz a and f, which is not the case under a multi-bang assumption on a, although it is shown that this formula holds for the more general case of $a, f \in L^{\infty,1+\epsilon}(\mathbb{R}^2, \mathbb{R})$ with $\epsilon > 0$ in [48].

In all previous analysis we have been examining μ as $\lambda \to \infty$, we have seen that μ is analytic, with respect to λ, inside the unit circle and so we can also examine the behaviour of μ as $\lambda \to 0$. By examining μ as $\lambda \to 0$ we can derive the data consistency condition (2.20) [48, 49] .

Theorem 4. *Let f, a and M be as in Theorem 3, then*

$$0 = \int_{\mathbb{S}^1} \exp\left(\int_t^\infty a(s\theta^\perp + t'\theta)\mathrm{d}t'\right) M(x_1, x_2, \theta)\mathrm{d}\theta. \tag{2.76}$$

Proof. From the proof of Theorem 3 we obtained the following formula

$$\begin{aligned} &\mu(x_1, x_2, \lambda) \exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a(x_1, x_2)}{\gamma(\lambda)}\right)\right) = \\ &\frac{1}{2\pi i}\mathrm{sgn}\left(\frac{1}{|\lambda|^2} - |\lambda|^2\right) \iint_{\mathbb{R}^2} \frac{\exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a(x_1', x_2')}{\gamma(\lambda)}\right)\right) f(x_1', x_2')}{z' - z} \mathrm{d}x_1'\mathrm{d}x_2'. \end{aligned} \tag{2.77}$$

Note that as $\exp\left(\partial_{\bar{z}}^{-1}\left(\frac{a(x_1,x_2)}{\gamma(\lambda)}\right)\right)$ appears on both sides (and we are only integrating with respect to x_1' and x_2') that the behaviour of μ as $\lambda \to 0$ is determined by $z' - z$. In the proof of Lemma 5 we found that

$$z' - z = \frac{1}{2i}\left(\lambda - \frac{1}{\lambda}\right)(x_1' - x_1) - \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right)(x_2' - x_2), \tag{2.78}$$

which shows that $z' - z = O\left(\frac{1}{z}\right)$ as $\lambda \to 0$. Therefore (2.77) combined with the two comments above gives $\mu = O(\lambda)$ as $\lambda \to 0$, i.e

$$\lim_{\lambda \to 0} \mu(x_1, x_2, \lambda) = 0. \tag{2.79}$$

Now recall that in the proof of Theorem 3 we obtained the following alternative representation of μ

$$\mu(x_1, x_2, \lambda) = \frac{1}{2\pi}\int_0^{2\pi} \frac{J(x_1, x_2, \omega)e^{i\omega}}{e^{i\omega} - \lambda}\mathrm{d}\omega, \tag{2.80}$$

where $J(x_1, x_2, \omega) = -\exp\left(\int_t^\infty a(s\theta^\perp + t'\theta)\mathrm{d}t'\right) M(x_1, x_2, \theta)$. As we are interested in the behaviour of (2.80) as $\lambda \to 0$ we can use the Laurent series expansion of $\frac{1}{1-\frac{\lambda}{e^{i\omega}}}$ to find

$$\mu(x_1, x_2, \lambda) = \frac{1}{2\pi}\int_0^{2\pi} J(x_1, x_2, \omega)\mathrm{d}\omega + O(\lambda), \tag{2.81}$$

as $\lambda \to 0$. Combining (2.79),(2.81) and setting $\theta = e^{i\omega}$ yields

$$0 = \int_{\mathbb{S}^1} \exp\left(\int_t^\infty a(s\theta^\perp + t'\theta)\mathrm{d}t'\right) M(x_1, x_2, \theta)\mathrm{d}\theta.$$

as required. □

The final section of the literature review examines some of the material relating to Discrete Tomography.

2.7 Discrete Tomography

As mentioned in the introduction, a consequence of the inclusion of the joint multi-bang and TV regularizer into the variation problem (2.9) is that we can obtain good recovery of a and f whilst using a limited number of projections. At least numerically, this property continues to hold when we instead consider the Computerized Tomography (CT) case in which instead of the AtRT we aim to invert the Radon Transform

$$Rf(s, \theta) := \int_{-\infty}^{\infty} f(s\theta^\perp + t\theta)\mathrm{d}t \tag{2.82}$$

for f.

The problem of recovering binary or f taking values from a discrete set, i.e multi-bang f, from a limited number of angle projections is often called Discrete Tomography (DT), see [34] and its references. DT has many practical uses, reduction in the number of projections needed both reduces time taken to scan and reduces radiation doses significantly [34, 47]. A reduction in the number of projections also results in being able to perform numerical algorithms quicker. The downside to this is that there are typically many solutions [47, 34] and a priori information about f, such as convexity [17], is required to restrict the number of solutions [47].

There are some widely known results relating to sampling rates of projections in order to be able to uniquely recover the original f from its Radon transform. The Nyquist-Shannon sampling theorem implies that provided the change in angle between each projection is sufficiently small compared to the largest non-zero power in the Fourier series of $Rf(s, \theta)$, with respect to θ, then it is possible to recover f with no knowledge lost [39]. From numerical evidence given in [35, 39] it is recommended that the number of projections used is of the same order as the number of pixels a ray with

$s = 0$ passes through, i.e a ray passing through the origin. When f is multi-bang this number can likely be significantly reduced by the use of the multi-bang regularizer.

A common approach to projection sampling in the DT problem is to choose the projections so that they are equally spaced around 360 degrees with some perturbation. This perturbation is done to avoid data redundancy (Since $Rf(s,\theta) = Rf(s,-\theta)$)[47, 17] and to make the angles irrationally related, which is known to improve the reconstructions [47, 63, 34, 17]. Using fewer projections typically means the optimization problem (2.9) is under determined. One major issue with the lack of data is in recovery of discontinuities of f. It is well known [17, 63, 62] that in order to recover some part of the boundary of a CT image you need data from the ray which intersects that boundary tangentially. This means that more complicated f require more projections to recover. Work in [62] and their references uses TV and Generalized Total Variation regularizers combined with a data fidelity term in order to make up for reduced amount of available data and still recover the important boundary features.

As the Discrete tomography problem for CT concerns recovery of f which takes discrete values between 0 and 1 [35], the multi-bang regularizer seems like a natural choice for this application. Since the Radon Transform (2.82) is linear in f, determining the gradient with respect to f of the data fidelity term in (2.9) is numerically less intensive than in the joint a and f recovery methods. A description of the inversion algorithm for the DT case is given in Chapter 5. Chapter 6 gives a numerical example of recovery with 16 projections for a binary image. Other examples of binary recovery from sparse view CT data is given in [63, 17].

This concludes the Literature review, the next chapter presents novel results on the recovery of boundaries of a multi-bang a with $f \in C_c^1(\Omega)$ from knowledge of $R_a f$.

Chapter 3

Recovery of boundaries in the SPECT Identification problem

This chapter focuses on extending the work carried out in [37] related to the identification problem for SPECT when a is multi-bang, as in definition 6. The proof presented in [37] regarding unique recovery of a and f is for nested convex multi-bang a and $f \in C^1(\mathbb{R}^2)$. The specific extension given in this section is assuming $f \in C^1(\mathbb{R}^2)$ and that the boundaries $\partial\Omega_j$ are piecewise analytic with corners then we can uniquely determine the boundaries $\partial\Omega_j$ from the data $R_a f(s,\theta)$. The method is inspired by results of microlocal analysis, although does not exactly use microlocal analysis (e.g. see [56]), to show that jumps in a across boundaries of Ω_j lead to singularities in $R_a f$ at rays (s,θ) corresponding to lines which are either tangent to the boundaries $\partial\Omega_j$, or pass through corners of $\partial\Omega_j$.

In particular we show that if a ray (s^*,θ^*) does not contain any line segment which is also a boundary of a then $R_a f$ is continuous at (s^*,θ^*). This allows us to distinguish between rays which contain edges and all other rays. For a ray (s^*,θ^*) containing an edge, we can then pick a point p on the ray and then rotate about this point by an angle ω and examine $\partial_{p\cdot\theta^*} \lim_{\omega\to\omega^*} R_a f(p\cdot(\theta^*)^\perp,\theta)$ with $\theta^* = (\cos(\omega^*),\sin(\omega^*)$ to find that $\partial_{p\cdot\theta^*} \lim_{\omega\to\omega^*} R_a f$ is only non-zero when $p\cdot(\theta^*)^\perp$ lies on an edge. Restricting ourselves to the case where $R_a f$ is continuous at (s^*,θ^*) we are able to show that provided this ray does not intersect a boundary tangentially or pass through a corner then $R_a f$ is C^1 at (s^*,θ^*). We then show that when (s,θ) passes through a corner $\partial_s R_a f$ is discontinuous but bounded and when (s^*,θ^*) passes the boundary tangentially $\partial_s R_a f$

is discontinuous and unbounded. Points on the analytic parts of the boundary are able to be determined by first finding rays for which $\partial_s R_a f$ is unbounded and then fixing a point p on the ray and rotating about that point by some angle ω and examining the behaviour of $\partial_\omega R_a f$. When (s^*, θ^*) passes only a single part of the boundary, if p is not on the boundary we find that $\partial_\omega R_a f$ is unbounded and if p is on the boundary then $\partial_\omega R_a f$ is bounded. If there are multiple points of tangency for a ray (s^*, θ^*) then we can examine the singularity of $\partial_\omega R_a f$ as we vary the point p and find that jumps in the derivative with respect to p coincide with points on the boundary. We can then remove all the highest singularity terms from $\partial_s R_a f$ and repeat until we end up with something which is at worst discontinuous but bounded at (s^*, θ^*), i,e the ray also passed through a corner. Therefore we can find all tangent points passed through by a ray (s, θ). For the corner case, after determining all points on the analytical parts of the boundary and removing their contributions to $\partial_s R_a f$ is necessary, all but two rays (the rays which are tangential to either side of the corner), passing through it will have discontinuous but bounded derivative with respect to s and this allows us to determine corner points. Exactly how all of this is done is given in detail Chapter 3.1-3.3.

If we have access to the complete data then we are able to determine where the data or its derivatives jump and this information allows us to determine whether a ray passes the boundary tangentially, at a corner or not at all. As mentioned previously it is important to stress the requirement that every ray which either passes through a corner or intersects the boundary tangentially also intersects the support of f. If this condition is not met some parts of the boundary will be unrecoverable with the method given in this Chapter.

We begin by setting up some useful definitions, the first clarifies what we mean by a boundary point.

Definition 17. ($\mathcal{P}_a$) *For a multi-bang a with boundaries $\{\partial\Omega_j\}_{j=1}^n$, which may overlap, there are three types of boundary points $\boldsymbol{x}$. Either $\boldsymbol{x}$*

1. *lies only on analytic and non-flat parts of the boundaries $\partial\Omega_j$, i.e at least locally to $\boldsymbol{x}$ we can define a Taylor series expansion, which is not simply a constant, about $\boldsymbol{x}$ which describes the boundary, or*

2. *lies on a straight edge of the boundary, we note that we do distinguish this case separately to case 1, or*

3. *a corner for at least one of the boundaries $\partial\Omega_j$.*

Furthermore we separate these boundary points into two sets, straight edges $\mathcal{P}_a^e$ which are case 2 and let $\mathcal{P}_a$ be the set of $\boldsymbol{x} \in \partial\Omega_j$ which fall into case 1 or 3.

We aim show that we can recover all points on the boundary. Note that it is not entirely necessary to be able to determine the points which are corners; this is because there can be only finitely many corners. Then taking the closure of the edge points and piecewise analytic points is enough to uniquely determine the boundary. We do however include the formulas and techniques related to corner points as this may be of independent interest and adds completeness to the result.

Following [37] we slightly modify the following definition in order to accommodate our more general case.

Definition 18. **($\mathcal{K}_a$)** *Suppose that a is multi-bang with sets $\{\Omega_j\}_{j=1}^n$ as in Definition 6 and $\mathcal{P}_a$ as in Definition 17. We define $\mathcal{K}_a$ to be the subset of $\{(s,\theta) : s \in \mathbb{R}, \theta \in \mathbb{S}^1\}$ such that the line corresponding of (s,θ) is either tangent to $\partial\Omega_j$, contains a flat edge of the boundary or passes through a corner of $\partial\Omega_j$ for some j. Furthermore we define $\mathcal{K}_a^e \subset \mathcal{K}_a$ to be the set of all rays in $\mathcal{K}_a$ which also do not contain any flat edges of the boundary. Finally, we define subsets of $\mathcal{K}_a$ which are the set $\mathcal{K}_{a,1}$ of lines tangent at a point which lies on an analytic part of the boundary, and $\mathcal{K}_{a,2}$ the set of lines passing through corners. (Note that $\mathcal{K}_{a,1}$ and $\mathcal{K}_{a,2}$ may not be disjoint.)*

Before examining the behaviour of the jumps in $R_a f(s,\theta)$ we first state the following result about the continuity of the $R_a f(s,\theta)$ at a ray (s,θ) which is not tangential to any part of boundary nor passes through a corner.

Lemma 7. *Suppose that $f \in C_c^1(\mathbb{R}^2)$, a is multi-bang, and $(s^*,\theta^*) \in (\mathcal{K}_a)^c$. Then the mapping $(s,\theta) \mapsto R_a f(s,\theta)$ is C^1 at (s^*,θ^*).*

Proof. We can prove this result in a similar manner as in [37, Lemma 1]. Since $(s^*,\theta^*) \in (\mathcal{K}_a^0)^c$ the line tangent to θ^* with distance s^* from the origin is neither tangent to one of the boundaries $\partial\Omega_j$, nor passes through a corner of one of these boundaries.

W.L.O.G we can rotate and translate the axis so that the line corresponding to (s^*, θ^*) lies on the x-axis and $\theta^* = (1, 0)$ or $\omega^* = 0$. Then we have

$$R_a f(s, \theta^*) = \int_{-\infty}^{\infty} f(x, s) e^{-Da((x,s),\theta^*)} \mathrm{d}x. \tag{3.1}$$

For $x \in \mathbb{R}$, recall that

$$Da((x, s), \theta^*) = \int_x^{\infty} a(t, s) \mathrm{d}t.$$

By assumption of a being multi-bang, the line given by (s, θ^*) only crosses the jumps of a finitely many times and we label the ordered values of t for which these crossings occur (when the line is parametrised as $t \mapsto (t, s)$) as $\{t_i(s, \theta)\}_{i=1}^N$. Since $(s^*, \theta^*) \in (\mathcal{K}_a)^c$ these functions t_i are differentiable in a neighbourhood of $(0, 0)$. Next we introduce the functions

$$\phi_i(x, s, \omega) = \begin{cases} t_i(s, \omega) & x < t_i(s, \omega) \\ x & x \geq t_i(s, \omega). \end{cases} \tag{3.2}$$

Note that for all i, ϕ_i is continuous with bounded first derivative in a neighbourhood of $(s, \omega) = (0, 0)$ which is also continuous when $x \neq t_i(s, \omega)$. Using these functions we have

$$Da((x, s), \theta) = \sum_{i=1}^{N-1} c_i(\phi_{i+1}(x, s, \omega) - \phi_i(x, s, \omega)) = \sum_{i=1}^{N} (c_{i-1} - c_i)\phi_i(x, s, \omega) \tag{3.3}$$

where for each i, c_i is one of the admissible values or possibly zero (in particular c_0 and c_N are always zero). We thus see that $Da((x, s), \theta^*)$ also has bounded derivatives in a neighbourhood of $(s, \theta^*) = (0, (1, 0))$ that are continuous except when $x = t_i(s, \theta)$ for some i, and differentiating this with respect to s and ω gives

$$\partial_s Da((x, s), \theta) = \sum_{i=1}^{N} (c_{i-1} - c_i)\partial_s \phi_i(x, s, \omega)$$

and

$$\partial_\omega Da((x, s), \theta) = \sum_{i=1}^{N} (c_{i-1} - c_i)\partial_\omega \phi_i(x, s, \omega).$$

Since $f \in C^1(\mathbb{R}^2)$ as well, we can therefore differentiate under the integral sign in (3.1) to get

$$\begin{aligned} \partial_s R_a f(s, \theta^*) &= \int_{-\infty}^{\infty} \left(\partial_s f(x, s) - \partial_s Da((x, s), \theta^*) f(x, s)\right) e^{-Da((x,s),\theta^*)} \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \partial_s f(x, s) e^{-Da((x,s),\theta^*)} \mathrm{d}x \\ &\quad + \sum_{i=1}^{N} (c_i - c_{i-1}) \partial_s t_i(s, \omega) \int_{-\infty}^{t_i(s)} f(x, s) e^{-Da((x,s),\theta^*)} \mathrm{d}x. \end{aligned} \tag{3.4}$$

Since $f \in C_c^1(\mathbb{R}^2)$ and the derivatives $\partial_s t_i$ are all continuous, we see that in fact $\partial_s R_a f(s,\theta^*)$ is also continuous with respect to s at $s=0$. It remains to be shown what happens in the case of ω at $\omega = 0$.

For ω a similar argument as in the s case shows we can differentiate under the integral sign to obtain

$$\begin{aligned}\partial_\omega R_a f(s,\theta) &= \int_{-\infty}^{\infty} \left(\partial_\omega f(x,s) - \partial_\omega Da((x,s),\theta) f(x,s)\right) e^{-Da((x,s),\theta^*)} \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \partial_\omega f(x,s) e^{-Da((x,s),\theta^*)} \mathrm{d}x \\ &\quad + \sum_{i=1}^{N} (c_i - c_{i-1}) \partial_\omega t_i(s,\omega) \int_{-\infty}^{t_i(s)} f(x,s) e^{-Da((x,s),\theta^*)} \mathrm{d}x.\end{aligned} \tag{3.5}$$

Again as $f \in C_c^1(\mathbb{R}^2)$ and the derivatives $\partial_\omega t_i$ are all continuous we have $R_a f(s,\theta)$ is continuous with respect to ω at $\omega = 0$. Combining (3.4) and (3.5) shows that $\nabla R_a f(s,\omega)$ is continuous in a neighbourhood of $(s,\omega) = (0,0)$ as required. □

The next section contains our analysis of the jumps of $R_a f$ resulting from a ray passing through some part of a piecewise analytic boundary of $\partial\Omega$.

3.1 Boundary points on a piecewise analytic part of the boundary

We now begin the analysis of the jumps in $R_a f$ by looking at a ray which is tangent to only a single point on a piecewise analytic part of the boundary.

Lemma 8. *Suppose that $f \in C_c^1(\mathbb{R}^2)$ and a is multi-bang. Suppose that the ray $(s^*,\theta^*) \in \mathcal{K}_{a,1}$ is tangential to a piecewise analytic boundary at a single point $s^*(\theta^*)^\perp + t^*\theta^*$. Suppose that locally to the point $s^*(\theta^*)^\perp + t^*\theta^*$, with $\theta^* = (\cos(\omega^*), \sin(\omega^*))$, the Cartesian co-ordinates of the boundary can be described via $y'(x') = (x' - x_1')^n g(x')$, where $x' = x\cos(\omega^*) + y\sin(\omega^*)$, $y' = -x\sin(\omega^*) + y\cos(\omega^*)$ and x_1' is the x' coordinate of the tangent point, for some $g(x')$ with $g(0) = \kappa \neq 0$. Furthermore, suppose that $a = c$ on the side $(\theta^*)^\perp = (-\sin(\omega^*), \cos(\omega^*))$ points to at the point of tangency, and $a = c_0$ on the other side. Additionally assume that $\boldsymbol{x}^*$ is a point on the line*

corresponding to (s^*, θ^*)*, then*

$$\begin{aligned}&\lim_{\omega \to (\omega^*)^{s_1}} |\sin(\omega - \omega^*)|^{\frac{n-1}{n}} \partial_\omega R_a(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \\ &2s_2(c_0 - c) \left(\frac{|\boldsymbol{x}^* \cdot \theta^* - t^*|}{n|\kappa|} \right)^{\frac{1}{n}} \int_{-\infty}^{t^*} f(s^*(\theta^*)^\perp + t\theta^*) e^{-Da(s^*(\theta^*)^\perp + t\theta^*, \theta^*)} \, \mathrm{d}t\end{aligned} \tag{3.6}$$

if n *is even, and*

$$\begin{aligned}&\lim_{\omega \to (\omega^*)^{s_1}} |\sin(\omega - \omega^*)|^{\frac{n-1}{n}} \partial_\omega R(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \\ &s_1(c_0 - c) \left(\frac{|\boldsymbol{x}^* \cdot \theta^* - t^*|}{n|\kappa|} \right)^{\frac{1}{n}} \int_{-\infty}^{t^*} f(s^*(\theta^*)^\perp + t\theta^*) e^{-Da(s^*(\theta^*)^\perp + t\theta^*, \theta^*)} \, \mathrm{d}t\end{aligned} \tag{3.7}$$

if n *is odd, where* s_1 *is the sign of* $(t^* - \boldsymbol{x}^* \cdot \theta^*)\kappa$ *and* s_2 *is the sign of* κ*.*

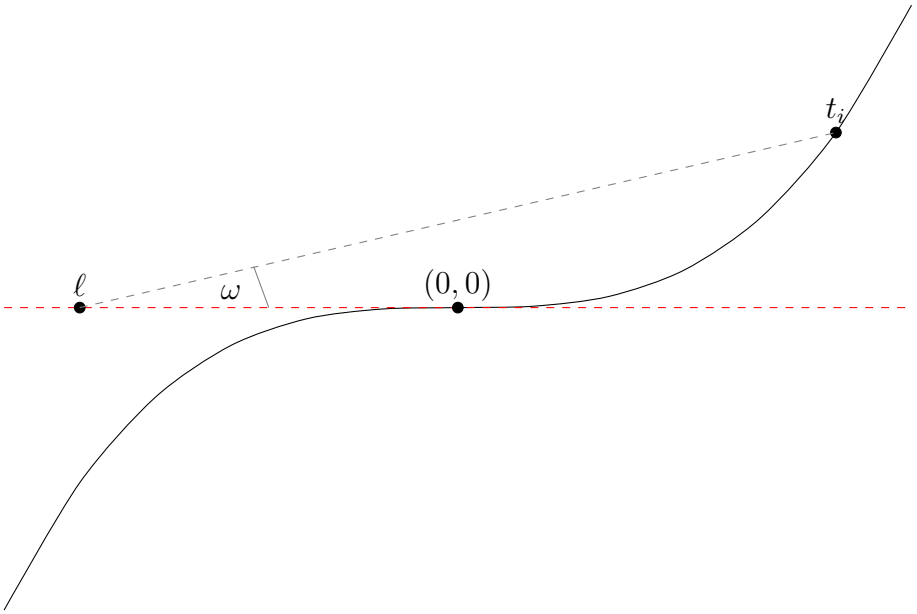
Proof. As in [37] by rotating, translating and possibly reflecting about the x-axis we assume W.L.O.G. that $\theta^* = (1, 0)$, $(\theta^*)^\perp = (0, 1)$ and the point of tangency is at the origin (i.e. $t^* = 0$). We also assume that $\boldsymbol{x}^* = (\ell, 0)$ and for now only consider the case $\ell \neq 0$. Note that the line corresponding to $(x^* \cdot \theta^\perp, \theta)$ is precisely the line through $\boldsymbol{x}^*$ tangent to θ, and we can change the parametrization of this line in the integral definition of the AtRT so that $t = 0$ always corresponds with $\boldsymbol{x}^*$. After doing this we have

$$R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \int_{-\infty}^{\infty} f(x^* + t\theta) e^{-Da(\boldsymbol{x}^* + t\theta, \theta)} \, \mathrm{d}t. \tag{3.8}$$

Now we use the same notation as in [37] and label the ordered values of t along the line $t \mapsto \boldsymbol{x}^* + t\theta$, for $\mathrm{sgn}(\ell)\omega > 0$ and $|\omega|$ sufficiently small, at which the line intersects one of the boundaries $\partial\Omega_j$ as $\{t_i(\omega)\}_{i=1}^N$. The importance of $|\omega|$ being sufficient small is to ensure that as $\omega \to 0^+$ the intersection points, which depend on ω, behave continuously. By hypothesis, at least locally to $(0, 0)$ we can describe the boundary by $y = x^n g(x)$. If n is even then this corresponds to a locally convex or concave part of the boundary. By locally convex or concave we mean that if we restrict the domain enough the boundary is convex or concave. If the boundary is locally convex or concave there are either two points or zero points of intersection, if n is odd then there will be just the single intersection point for a suitable oriented line with angle ω. Therefore at most two of the t_i will correspond to the point of tangency and these will satisfy $t_i(\omega) \to -\ell$ as $\omega \to 0^{-\mathrm{sgn}(\ell)}$.

W.L.O.G we orient the boundary such that it can be written $y = x^n g(x)$ such that $\kappa := g(0) > 0$. This orientation gives us two cases to consider: either $(0, 0)$ is on a

Figure 3.1: Visual aid for Lemma 8. Here the tangent point is an inflexion point on the boundary with $y(x)$ increasing. We can see that $(x, y) = (-\ell + t_i \cos(\omega), t_i \sin(\omega))$.



convex part of the boundary, or it is an inflexion point for an increasing part of the boundary(in the sense that $y(x)$ is increasing). In either case we have

$$\frac{\mathrm{d}y}{\mathrm{d}\omega} = \left(nx^{n-1}g(x) + x^n g'(x)\right)\frac{\mathrm{d}x}{\mathrm{d}\omega}. \tag{3.9}$$

Furthermore, for t_i corresponding to an intersection we have

$$y = t_i \sin(\omega)$$
$$x = t_i \cos(\omega) - \ell$$

see Figure 3.1.

Therefore we have

$$\frac{\mathrm{d}y}{\mathrm{d}\omega} = t_i \cos(\omega) + \frac{\mathrm{d}t_i}{\mathrm{d}\omega}\sin(\omega)$$
$$\frac{\mathrm{d}x}{\mathrm{d}\omega} = -t_i \sin(\omega) + \frac{\mathrm{d}t_i}{\mathrm{d}\omega}\cos(\omega).$$

Dividing (3.9) by $\cos(\omega)$ we find

$$t^* + \frac{\mathrm{d}t_i}{\mathrm{d}\omega}\tan(\omega) = \left(nx^{n-1}g(x) + x^n g'(x)\right)\left(-t^*\tan(\omega) + \frac{\mathrm{d}t_i}{\mathrm{d}\omega}\right),$$
$$\implies \frac{\mathrm{d}t_i}{\mathrm{d}\omega}\left(\tan(\omega) - \left(nx^{n-1}g(x) + x^n g'(x)\right)\right) = -(nx^{n-1}g(x) + x^n g'(x))t_i\tan(\omega) - t_i,$$
$$\implies \frac{\mathrm{d}t_i}{\mathrm{d}\omega} = \frac{-(nx^{n-1}g(x) + x^n g'(x))t_i\tan(\omega) - t_i}{\tan(\omega) - (nx^{n-1}g(x) + x^n g'(x))}.$$

We now note we have the following geometric relations

$$\begin{aligned}\tan(\omega) &= \frac{y}{x-\ell} = \frac{x^n g(x)}{|\ell|+x} \\ \sin(\omega) &= \frac{y}{t_i} = \frac{x^n g(x)}{t_i}.\end{aligned} \tag{3.10}$$

Since $\lim_{\omega\to 0^+} t_i = -\ell = |\ell|$ and $\lim_{\omega\to 0^+} x = 0$, taking the limit as $\omega \to 0^+$ we obtain

$$\begin{aligned}\lim_{\omega\to 0^+} \frac{\mathrm{d}t_i}{\mathrm{d}\omega} &= \frac{\ell}{\lim_{\omega\to 0^+} x^{n-1}\left(\frac{\tan(\omega)}{x^{n-1}} - (ng(x)+xg'(x))\right)}, \\ &= \frac{\ell}{\lim_{\omega\to 0^+} x^{n-1}\left(\frac{xg(x)}{x+|\ell|} - (ng(x)+xg'(x))\right)} \\ &= \frac{|\ell|}{\lim_{\omega\to 0^+} x^{n-1}\left(ng(x)\right)}.\end{aligned}$$

Note that the limit of the denominator in the above is 0 but it is left in this form for the following reason. Using (3.10), the fact that $g(0) = \kappa$ and labelling $t_{i\pm}$ as the values of t corresponding to an intersection

$$\lim_{\omega\to 0^+} |\sin(\omega)|^{\frac{n-1}{n}} \frac{\mathrm{d}t_{i\pm}}{\mathrm{d}\omega} = \pm\left(\frac{|\ell|}{n\kappa}\right)^{\frac{1}{n}} \tag{3.11}$$

if n is even and

$$\lim_{\omega\to 0^+} |\sin(\omega)|^{\frac{n-1}{n}} \frac{\mathrm{d}t_{i}}{\mathrm{d}\omega} = \left(\frac{|\ell|}{n\kappa}\right)^{\frac{1}{n}} \tag{3.12}$$

when n is odd. When $\ell = 0$ we can show in a similar manner to the previous derivation that

$$\begin{aligned}\lim_{\omega\to 0^+} |\sin(\omega)|^{\frac{n-1}{n}} \frac{\mathrm{d}t_{i}}{\mathrm{d}\omega} &= 0 \\ \lim_{\omega\to 0^+} |\sin(\omega)|^{\frac{n-1}{n}} \frac{\mathrm{d}t_{i\pm}}{\mathrm{d}\omega} &= 0.\end{aligned} \tag{3.13}$$

Note that if we considered the point of tangency to be at a point $\beta > 0$ rather than 0 the formula obtained above would be the same but with the term $|\ell|$ replaced by $|\ell - \beta|$.

Now as in previous lemmas, define functions ϕ_i

$$\phi_i(t,\omega) = \begin{cases} t_i(\omega) & t < t_i(\omega) \\ t & t \geq t_i(\omega), \end{cases}$$

for $-\mathrm{sgn}(\ell)\omega > 0$. This gives

$$Da(x^* + t\theta, \theta) = \sum_{i=1}^{N} (c_{i-1} - c_i)\phi_i(t,\omega),$$

where c_i are the multi-bang values inbetween the boundaries passed through by the line passing through $(l,0)$ and oriented at $\omega > 0$, for $-\mathrm{sgn}(\ell)\omega > 0$. In the case $\ell = 0$ we still have versions of the previous formula for $\omega \neq 0$, but it will change depending on the sign of ω. At this point we have 2 cases to consider, when n is even and when n is odd, we begin by considering the case when n is even.

When n is even, we write $\phi_\pm$ for those $\phi_\pm$ corresponding to $t_\pm$. Now let us take the derivative of (3.8) in the case when $-\mathrm{sgn}(\ell)\omega > 0$ if $\ell \neq 0$ or $\omega \neq 0$ if $\ell = 0$. We then have

$$\partial_\omega R_a f(x^* \cdot \theta^\perp, \theta) = \int_{-\infty}^{\infty} \Big(\partial_\omega f(\boldsymbol{x}^* + t\theta) - \partial_\omega Da(\boldsymbol{x}^* + t\theta, \theta) f(\boldsymbol{x}^* + t\theta)\Big) e^{-Da(\boldsymbol{x}^* + t\theta, \theta)} \, \mathrm{d}t. \tag{3.14}$$

First consider the case $\ell = 0$. In this case when we multiply by $|\sin(\omega)|^{\frac{n-1}{n}}$ and take the limit as $\omega \to 0$, using (3.13) we see that the limit is zero. Since $\ell = \boldsymbol{x}^* \cdot \theta^* - t^*$, this proves the result when $\ell = 0$. Now consider when $\ell \neq 0$. In this case we multiply by $|\sin(\omega)|^{\frac{n-1}{n}}$ and take the limit as $\omega \to 0^{-\mathrm{sgn}(\ell)}$. The only terms that are not bounded in (3.14) for ω close to zero are those that involve derivatives of $\phi_\pm$. We therefore have

$$\begin{aligned}
&\lim_{\omega \to 0^{-\mathrm{sgn}(\ell)}} |\sin(\omega)|^{\frac{n-1}{n}} \partial_\omega R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \\
&(c - c_0) \lim_{\omega \to 0^{-\mathrm{sgn}(\ell)}} \int_{-\infty}^{t_-(\omega)} |\sin(\omega)|^{\frac{n-1}{n}} \partial_\omega t_-(\omega) f(\boldsymbol{x}^* + t\theta) e^{-Da(\boldsymbol{x}^* + t\theta, \theta)} \, \mathrm{d}t \\
&+ (c_0 - c) \lim_{\omega \to 0^{-\mathrm{sgn}(\ell)}} \int_{-\infty}^{t_+(\omega)} |\sin(\omega)|^{\frac{n-1}{n}} \partial_\omega t_+(\omega) f(\boldsymbol{x}^* + t\theta) e^{-Da(\boldsymbol{x}^* + t\theta, \theta)} \, \mathrm{d}t.
\end{aligned}$$

Applying (3.11) and (3.12) to this we finally obtain

$$\begin{aligned}
&\lim_{\omega \to 0^{-\mathrm{sgn}(\ell)}} |\sin(\omega)|^{\frac{n-1}{n}} \partial_\omega R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \\
&\qquad 2(c_0 - c) \left(\frac{|\ell|}{n\kappa}\right)^{\frac{1}{n}} \int_{-\infty}^{-\ell} f(\boldsymbol{x}^* + t\theta^*) e^{-Da(\boldsymbol{x}^* + t\theta^*, \theta^*)} \, \mathrm{d}t.
\end{aligned} \tag{3.15}$$

Taking into account the translations, rotation and reflection from the beginning of the proof, this formula agrees with (3.6).

The only remaining case to consider is when n is odd. As there is only one intersection point of interest here, we write ϕ for the ϕ_i corresponding to t^*. As before, let us take the derivative of (3.8) in the case when $-\mathrm{sgn}(\ell)\omega > 0$ if $\ell \neq 0$ or $\omega \neq 0$ if $\ell = 0$. We then have

$$\begin{aligned}
&\partial_\omega R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \\
&\int_{-\infty}^{\infty} \Big(\partial_\omega f(\boldsymbol{x}^* + t\theta) - \partial_\omega Da(\boldsymbol{x}^* + t\theta, \theta) f(\boldsymbol{x}^* + t\theta)\Big) e^{-Da(\boldsymbol{x}^* + t\theta, \theta)} \, \mathrm{d}t.
\end{aligned} \tag{3.16}$$

First consider the case $\ell = 0$. Again, in this case when we multiply by $|\sin(\omega)|^{\frac{n-1}{n}}$ and take the limit as $\omega \to 0$, using (3.13) we see that the limit is zero. Since $\ell = \boldsymbol{x}^* \cdot \theta^* - t^*$, this proves the result when $\ell = 0$. Now consider when $\ell \neq 0$. In this case we multiply by $|\sin(\omega)|^{\frac{n-1}{n}}$ and take the limit as $\omega \to 0^{-\text{sgn}(\ell)}$. The only term that is not bounded in (3.16) for ω close to zero is the term involving the derivative of ϕ. By a similar argument to the even n case we obtain

$$\begin{aligned} \lim_{\omega \to 0^{-\text{sgn}(\ell)}} & |\sin(\omega)|^{\frac{n-1}{n}} \partial_\omega R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) \\ & = (c_0 - c) \left(\frac{|\ell|}{n\kappa} \right)^{\frac{1}{n}} \int_{-\infty}^{-\ell} f(\boldsymbol{x}^* + t\theta^*) e^{-Da(\boldsymbol{x}^* + t\theta^*, \theta^*)} \, \mathrm{d}t. \end{aligned} \tag{3.17}$$

Taking into account the translations, rotation and reflection from the beginning of the proof, the formula given in (3.17) agrees with (3.7), and so completes the proof. □

Note that when $n = 2$ Lemma 8 agrees with Lemma 3 in [37]. Lemma 8 gives the following useful corollary.

Corollary 3. *Let f and a satisfy all the assumptions of Lemma 8, including the assumption that the Cartesian co-ordinates of the boundary can be described via $y'(x') = (x' - x_1')^n g(x')$, where $x' = x\cos(\omega^*) + y\sin(\omega^*), y' = -x\sin(\omega^*) + y\cos(\omega^*)$ and x_1^1 is the x' co-ordinate of the tangent point , for some $g(x')$ with $g(0) = \kappa \neq 0$. Then, for any $m < n$, with $m \geq 1$ we have*

$$\lim_{\omega \to (\omega^*)^{s_1}} |\sin(\omega - \omega^*)|^{\frac{m-1}{m}} \partial_\omega R(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = s_2 \infty \tag{3.18}$$

if n is even and

$$\lim_{\omega \to (\omega^*)^{s_1}} |\sin(\omega - \omega^*)|^{\frac{m-1}{m}} \partial_\omega R(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = s_1 \infty \tag{3.19}$$

if n is odd where s_1 is the sign of $(t^* - \boldsymbol{x}^* \cdot \theta^*)\kappa$, and s_2 is the sign of κ.

Proof. W.L.O.G we can again consider the case where we can describe the boundary locally to the tangent point via $y = x^n g(x)$ with $g(0) = \kappa \neq 0$. Note that for any value of $m < n$ we can alternatively describe the y co-ordinate of the boundary via $y = x^m \hat{g}(x)$ except now $\hat{g}(0) = x^{n-m} g(x)|_{x=0} = 0$. Writing y in this way we can perform identical analysis as in Lemma 8 and note that $\kappa = 0$ and so the limits given in (3.6) and (3.7) tend to $\pm\infty$, as required. □

With the generalization for a single tangent point established, the next case to consider is the possibility of a single ray being tangent to a finite number of distinct points on analytic parts of the boundary.

Lemma 9. *Suppose that $f \in C_c^1(\mathbb{R}^2)$ and a is multi-bang. Suppose that $(s^*, \theta^*) \in \mathcal{K}_{a,1}$ passes the boundary tangentially at series of points $s^*(\theta^*)^\perp + t_i^*\theta^*$, which are all distinct and none of which are corners or lie on a straight edge. Suppose that local to each point $s^*(\theta^*)^\perp + t_i^*\theta^* = (x_i, y_i)$ the boundary can be described by $y'(x') = (x' - x_i')^{n_i} g_i(x')$ for some $g_i(x')$ with $g_i(x_i) = \kappa_i \neq 0$, $n_i \in \mathbb{N}$, x', y' as in Lemma 8 and x_i' the x' coordinate of each of the tangent points. Furthermore, suppose that $a = c_i$ on the side $(\theta^*)^\perp = (-\theta_2, \theta_1)$ points to at the point $s^*(\theta^*)^\perp + t_i^*\theta^*$ and d_i on the opposite side. Additionally assume that $\omega^* \in [0, 2\pi]$ satisfies $\theta^* = (\cos(\omega^*), \sin(\omega^*))$ and let $\boldsymbol{x}^*$ be a point on the line corresponding to (s^*, θ^*) and set $m = \max n_i$. If m is odd let $\mathcal{I}_m$ be the set of t values which correspond with tangent points of order m then*

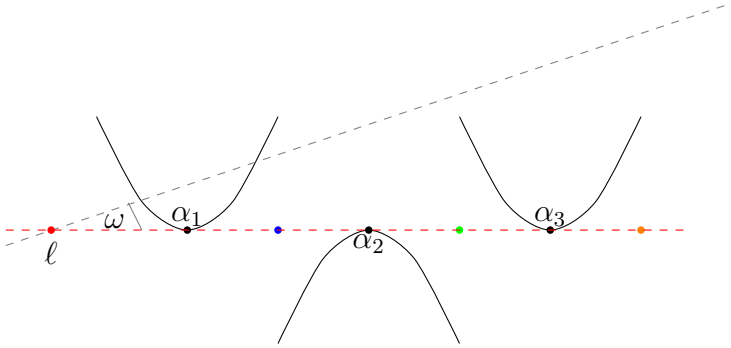
$$\lim_{\omega \to (\omega^*)^+} |\sin(\omega - \omega^*)|^{\frac{m-1}{m}} \partial_\omega R(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \sum_{t_i^* \in \mathcal{I}_m(\boldsymbol{x}^*)} s_i(d_i - c_i) \left(\frac{|\boldsymbol{x}^* \cdot \theta^* - t_i^*|}{m|\kappa_i|} \right)^{\frac{1}{m}} \int_{-\infty}^{t_i^*} f(s^*(\theta^*)^\perp + t\theta^*) e^{-Da(s^*(\theta^*)^\perp + t\theta^*, \theta^*)} \, \mathrm{d}t, \tag{3.20}$$

where s_i is the sign of $(t_i^* - \boldsymbol{x}^* \cdot \theta^*)\kappa_i$. If m is even then for $\boldsymbol{x}^*$ on the line (s^*, θ^*) define $\mathcal{I}_m^1(x)$ to be the set of t values corresponding to tangent points with $\kappa_i > 0$ and $t_i \geq \boldsymbol{x}^* \cdot \theta^*$, that is to say the convex points after $\boldsymbol{x}^*$ on the oriented ray, and $\mathcal{I}_m^2(\boldsymbol{x}^*)$ be the set of t values corresponding to tangent points with $\kappa < 0$ and $t_i \leq \boldsymbol{x}^* \cdot \theta^*$ then

$$\begin{aligned} \lim_{\omega \to (\omega^*)^+} |\sin(\omega - \omega^*)|^{\frac{m-1}{m}} \partial_\omega R(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \\ \sum_{t_i^* \in \mathcal{I}_m^1(\boldsymbol{x}^*)} 2(d_i - c_i) \left(\frac{|\boldsymbol{x}^* \cdot \theta^* - t_i^*|}{m|\kappa_i|} \right)^{\frac{1}{m}} \int_{-\infty}^{t_i^*} f(s^*(\theta^*)^\perp + t\theta^*) e^{-Da(s^*(\theta^*)^\perp + t\theta^*, \theta^*)} \, \mathrm{d}t \\ + \sum_{t_i^* \in \mathcal{I}_m^2(\boldsymbol{x}^*)} 2(c_i - d_i) \left(\frac{|\boldsymbol{x}^* \cdot \theta^* - t_i^*|}{m|\kappa_i|} \right)^{\frac{1}{m}} \int_{-\infty}^{t_i^*} f(s^*(\theta^*)^\perp + t\theta^*) e^{-Da(s^*(\theta^*)^\perp + t\theta^*, \theta^*)} \, \mathrm{d}t. \end{aligned} \tag{3.21}$$

Proof. As in lemma 8 by rotating, translating and possibly reflecting about the x-axis we assume W.L.O.G. that $\theta^* = (1, 0)$, $\theta_\perp^* = (0, 1)$ and the points of tangency are at

Figure 3.2: Demonstrating the effect of changing ℓ on the visible tangent points as $\omega \to 0^+$. With ℓ in the red position, as $\omega \to 0^+$, α_1 and α_2 are visible and α_2 is invisible. In the blue position only α_3 is visible. In the green position α_2 and α_3 are visible. In the orange position only α_2 is visible.



$(t^*) = 0 = \alpha_0, \alpha_1, \alpha_2, ..., \alpha_k$, for some $k \geq 1$. As before, by suitable re parametrization we have

$$R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \int_{-\infty}^{\infty} f(\boldsymbol{x}^* + t\theta) e^{-Da(\boldsymbol{x}^* + t\theta, \theta)} \, \mathrm{d}t, \tag{3.22}$$

where $t = 0$ corresponds with $\boldsymbol{x}^*$. Again we assume that $\boldsymbol{x}^* = (\ell, 0)$ and for the moment consider only the case $\ell \neq 0$. As in lemma 8 label the ordered values of t along the line $t \mapsto \boldsymbol{x}^* + t\theta$, for $\omega > 0$ and $|\omega|$ sufficiently small, at which the line intersects one of the boundaries $\partial\Omega_j$ as $\{t_i(\omega)\}_{i=1}^N$, note that $N \geq k$ with equality only if every point of tangency is an inflexion point.

By assumption we know that local to each α_i the boundary can be described by $y(x) = (x - \alpha_i)^{n_i} g_i(x)$. By Lemma 8 we know that the orders of the singularities produced by this as $\omega \to 0^+$ are of the order $\sin(\omega)^{\frac{n-1}{n}}$. Let m be the highest order of singularity, that is $m = \max_{n_i}$. If there is only one such singularity then using the fact that all other tangency points will contribute nothing when multiplied by $\sin(\omega)^{\frac{m-1}{m}}$ we are in the setting of lemma 8 and so the result holds. Therefore we can assume there are at least two tangents with order m. Assume that we have n points of tangency, each with order m and label the set of tangents of order m $T = \{\alpha_{m_1}, \alpha_{m_2}, ..., \alpha_{m_n}\} \subseteq \{\alpha_1, ..., \alpha_k\}$. Local to each of the α_{m_i} the boundaries can be described by $(x - \alpha_{m_i})^m g_i(x)$ with $g_i(\alpha_{m_i}) := \kappa_{m_i} \neq 0$. As in lemma 8 the ray oriented with small but positive ω about $(\ell, 0)$, with $\ell < 0$ can intersect the boundaries related to α_{m_i} in 0,1 or 2 places. 0 places corresponds to an α_i on a locally concave part of the boundary, 1 place corresponds to an inflexion point and 2 a convex part of the

boundary. Since an α_i corresponding to a locally concave/convex part of the boundary must have even order and an inflexion point odd order then the set T contains either only inflexion points or only a mix of locally concave/convex points. This gives us two cases to consider.

Case 1- The set T consists of inflexion points.

As in lemma 8 define functions, for fixed ℓ, ϕ_i

$$\phi_i(t,\omega) = \begin{cases} t_i(\omega) & t < t_i(\omega) \\ t & t \geq t_i(\omega), \end{cases}$$

then if we define b_i (note some of these will be c_i and d_i) to be the multi-bang values in the regions between adjacent $t_i(\omega)$, the beam transform can be written as

$$Da(\boldsymbol{x}^* + t\theta, \theta) = \sum_{i=1}^{N} (b_{i-1} - b_i)\phi_i(t,\omega). \tag{3.23}$$

Now taking the derivative of (3.22) with respect to ω yields, as in lemma 8

$$\partial_\omega R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \int_{-\infty}^{\infty} \Big(\partial_\omega f(\boldsymbol{x}^* + t\theta) - \partial_\omega Da(\boldsymbol{x}^* + t\theta, \theta) f(\boldsymbol{x}^* + t\theta)\Big) e^{-Da(\boldsymbol{x}^* + t\theta, \theta)} \, \mathrm{d}t. \tag{3.24}$$

Since $f \in C_c^1(\mathbb{R}^2)$ the term involving $\partial_\omega f$ will be bounded as $\omega \to 0$ therefore any unbounded terms will come from the derivative of the beam transform. Taking the derivative of (3.23) with respect to ω we obtain

$$\partial_\omega Da(\boldsymbol{x}^* + t\theta, \theta) = \sum_{i=1}^{N} (b_{i-1} - b_i)\partial_\omega \phi_i(t,\omega).$$

If we multiply by $\sin(\omega)^{\frac{m-1}{m}}$ and then take the limit as $\omega \to 0^+$ the only terms which do not vanish will be those which involve derivatives related to one of the α_{m_i}. Note that in this case the relevant b_i will be given by d_{m_i} and c_{m_i}. This gives us, for $\ell < 0$

$$\begin{aligned} &\lim_{\omega \to 0^+} |\sin(\omega)|^{\frac{m-1}{m}} \partial_\omega R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \\ &\sum_{i=1}^{n} (d_{m_i} - c_{m_i}) \left(\frac{|\ell - \alpha_{m_i}|}{m\kappa_{m_i}} \right)^{\frac{1}{m}} \int_{-\infty}^{-\ell + \alpha_{m_i}} f(\boldsymbol{x}^* + t\theta^*) e^{-Da(\boldsymbol{x}^* + t\theta^*, \theta^*)} \, \mathrm{d}t. \end{aligned} \tag{3.25}$$

Now consider the case where $\alpha_{m_i} < \ell < \alpha_{m_{i+1}}$ for some i. As $\omega \to 0^+$ the contribution to the derivative from ϕ_{m_j}, for $N \geq j \geq i+1$ remains the same as in (3.25). There will also be singular contributions to the derivative from the $\partial_\omega \phi_{m_j}$ for $1 \leq j \leq i$ but

as these are approached from below the sign on these contributions will be swapped, therefore we have

$$\lim_{\omega\to 0^+} |\sin(\omega)|^{\frac{m-1}{m}}\partial_\omega R_a f(\boldsymbol{x}^*\cdot\theta^\perp,\theta) = \sum_{i=1}^{n}\operatorname{sign}(\alpha_{m_i}-\ell)(d_{m_i}-c_{m_i})\left(\frac{|\ell-\alpha_{m_i}|}{m\kappa_{m_i}}\right)^{\frac{1}{m}}\int_{-\infty}^{-\ell+\alpha_{m_i}} f(\boldsymbol{x}^*+t\theta^*)e^{-Da(\boldsymbol{x}^*+t\theta^*,\theta^*)}\,\mathrm{d}t. \tag{3.26}$$

Taking into account the translations, rotation and reflection from the beginning of the proof this formula agrees with (3.20). We now consider the case where α_{m_i} are on either locally convex or locally concave parts of the boundary.

Case 2- The set $\{\alpha_{m_1},\alpha_{m_2},...,\alpha_{m_n}\}$ consists of points on locally convex or locally concave parts of the boundary. Define $\mathcal{I}_m^1(\ell)$ to be the set of t values which correspond to the tangent points of order m α_{m_i} which are on locally convex parts of the boundary, see Figure 3.2.

We first consider the case where $\ell<0$ and suppose that ω is sufficiently small and positive. As in odd m case we have

$$Da(\boldsymbol{x}^*+t\theta,\theta)=\sum_{i=1}^{N}(b_{i-1}-b_i)\phi_i(t,\omega). \tag{3.27}$$

This again gives

$$\partial_\omega R_a f(\boldsymbol{x}^*\cdot\theta^\perp,\theta)=\int_{-\infty}^{\infty}\Big(\partial_\omega f(\boldsymbol{x}^*+t\theta)-\partial_\omega Da(\boldsymbol{x}^*+t\theta,\theta)f(\boldsymbol{x}^*+t\theta)\Big)e^{-Da(\boldsymbol{x}^*+t\theta,\theta)}\,\mathrm{d}t. \tag{3.28}$$

and

$$\partial_\omega Da(\boldsymbol{x}^*+t\theta,\theta)=\sum_{i=1}^{N}(b_{i-1}-b_i)\partial_\omega\phi_i(t,\omega).$$

Again, if we multiply by $\sin(\omega)^{\frac{m-1}{m}}$ and then take the limit as $\omega\to 0^+$ the only terms which do not vanish will correspond to the $\alpha_{m_i}^1\in\mathcal{I}_m^1(\ell)$. Note that as we approach from above the singular behaviour of locally concave parts of the boundary will not contribute to the derivative and so only the convex points contribute. The associated b_i will be given by d_{m_i} and c_{m_i}. This gives us for $\ell<0$

$$\lim_{\omega\to 0^+} |\sin(\omega)|^{\frac{m-1}{m}}\partial_\omega R_a f(\boldsymbol{x}^*\cdot\theta^\perp,\theta) = \sum_{i=1}^{|\mathcal{I}_m^1(\ell)|}(d_{m_i}-c_{m_i})\left(\frac{|\ell-\alpha_{m_i}|}{m\kappa_{m_i}}\right)^{\frac{1}{m}}\int_{-\infty}^{-\ell+\alpha_{m_i}} f(\boldsymbol{x}^*+t\theta^*)e^{-Da(\boldsymbol{x}^*+t\theta^*,\theta^*)}\,\mathrm{d}t. \tag{3.29}$$

Now consider the case where $\alpha_{m_i} < \ell < \alpha_{m_{i+1}}$ for some i. Let $\mathcal{I}_m^1(\ell)$ be the set of t values which correspond to a tangent point α_{m_j} which lies on a locally convex part of the boundary with $\alpha_{m_j} \geq \ell$, similarly define $\mathcal{I}_m^2(\ell)$ as the set of t values corresponding to a tangent point α_{m_j} which lies on a locally concave part of the boundary with $\alpha_{m_j} \leq \ell$. As $\omega \to 0^+$ the only contributions to the derivative the convex parts of the boundary make are those in $\mathcal{I}_m^1(\ell)$, note that the other convex parts of the boundary are not singular as $\omega \to 0^+$ (since for convex points $\alpha_{m_i} < \ell$ as $\omega \to 0^+$ we approach the associated boundary from below and the derivative with respect to s bounded in that limit). Furthermore the contributions the concave parts of the boundary make to the derivative will be those in $\mathcal{I}_m^2(\ell)$, as again the other concave parts do not contribute any singularities as $\omega \to 0^+$. Therefore we have

$$\begin{aligned}
&\lim_{\omega\to 0^+} |\sin(\omega)|^{\frac{m-1}{m}} \partial_\omega R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \\
&\sum_{i=1}^{|\mathcal{I}_m^1(\ell)|} (d_{m_i} - c_{m_i}) \left(\frac{|\ell - \alpha_{m_i}|}{m\kappa_{m_i}} \right)^{\frac{1}{m}} \int_{-\infty}^{-\ell+\alpha_{m_i}} f(\boldsymbol{x}^* + t\theta^*) e^{-Da(\boldsymbol{x}^*+t\theta^*,\theta^*)} \, \mathrm{d}t \\
&+ \sum_{i=1}^{|\mathcal{I}_m^2(\ell)|} (c_{m_i} - d_{m_i}) \left(\frac{|\ell - \alpha_{m_i}|}{m\kappa_{m_i}} \right)^{\frac{1}{m}} \int_{-\infty}^{-\ell+\alpha_{m_i}} f(\boldsymbol{x}^* + t\theta^*) e^{-Da(\boldsymbol{x}^*+t\theta^*,\theta^*)} \, \mathrm{d}t.
\end{aligned} \tag{3.30}$$

Note that if $\mathcal{I}_m^2(\ell)$ is empty then the formula in (3.30) holds with the concave part removed. A similar argument can be made for when $\mathcal{I}_m^1(\ell)$ is empty. Taking into account the translations, rotation and reflection from the beginning of the proof this formula agrees with (3.21), as required. □

In the case of lemma 8 we can easily determine the tangent point as this will be only point where the RHS of (3.7) or (3.6) will be 0. Unfortunately in the case involving multiple tangent points the formulas in lemma 9 cannot be used in the same way. This is because there are contributions from each tangent point so it will be non-zero at most values of ℓ. We can however determine the points of tangency by taking the derivatives with respect to $\boldsymbol{x}^* \cdot \theta$ of formulas obtained in lemma 9. This idea is the focus of the next corollary

Corollary 4. *Assume the same hypothesis as in lemma 9. If m is odd*

$$\begin{aligned}
&\partial_{\boldsymbol{x}^*\cdot\theta^*} \lim_{\omega\to(\omega^*)^+} |\sin(\omega - \omega^*)|^{\frac{m-1}{m}} \partial_\omega R(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \\
&\sum_{\mathcal{I}_m(\boldsymbol{x}^*)} \frac{s_i(d_i - c_i)}{(m|\kappa_i|)^{\frac{1}{m}}} (|\boldsymbol{x}^* \cdot \theta^* - t_i^*|)^{m^*} \int_{-\infty}^{t_i^*} f(s^*(\theta^*)^\perp + t\theta^*) e^{-Da(s^*(\theta^*)^\perp+t\theta^*,\theta^*)} \, \mathrm{d}t,
\end{aligned} \tag{3.31}$$

and

$$\partial_{\boldsymbol{x}^*,\theta^*} \lim_{\omega\to(\omega^*)^+} |\sin(\omega-\omega^*)|^{\frac{m-1}{m}}\partial_\omega R(\boldsymbol{x}^*\cdot\theta^\perp,\theta) =$$
$$\sum_{t_i^*\in\mathcal{I}_m^1(\boldsymbol{x}^*)} \frac{2(d_i-c_i)}{(m|\kappa_i|)^{\frac{1}{m}}}(|\boldsymbol{x}^*\cdot\theta^*-t_i^*|)^{m^*}\int_{-\infty}^{t_i^*} f(s^*(\theta^*)^\perp+t\theta^*))e^{-Da(s^*(\theta^*)^\perp+t\theta^*,\theta^*)}\,\mathrm{d}t+$$
$$\sum_{t_i^*\in\mathcal{I}_m^2(\boldsymbol{x}^*)} \frac{2(c_i-d_i)}{(m|\kappa_i|)^{\frac{1}{m}}}(|\boldsymbol{x}^*\cdot\theta^*-t_i^*|)^{m^*}\int_{-\infty}^{t_i^*} f(s^*(\theta^*)^\perp+t\theta^*))e^{-Da(s^*(\theta^*)^\perp+t\theta^*,\theta^*)}\,\mathrm{d}t \tag{3.32}$$

if m is even, where $m^* = -\left(1-\frac{1}{m}\right)$.

Proof. W.L.O.G we can work with the formulas presented in the proof of lemma 9 involving ℓ. The derivative we consider in this case will be with respect to ℓ. We first consider the inflexion points, i.e when m is odd. Recall that for $\ell<0$ we have

$$\lim_{\omega\to0^+} |\sin(\omega)|^{\frac{m-1}{m}}\partial_\omega R_a f(\boldsymbol{x}^*\cdot\theta^\perp,\theta) =$$
$$\sum_{i=1}^{n}(d_{m_i}-c_{m_i})\left(\frac{|\ell-\alpha_{m_i}|}{m\kappa_{m_i}}\right)^{\frac{1}{m}}\int_{-\infty}^{-\ell+\alpha_{m_i}} f(\boldsymbol{x}^*+t\theta^*)e^{-Da(\boldsymbol{x}^*+t\theta^*,\theta^*)}\,\mathrm{d}t.$$

Note that there are only two where ℓ appears in the RHS of (3.25) is in the upper limit and coefficient of each integral and in x^*. By re parametrizing we can remove the dependence of ℓ in the upper limit. Therefore the only contributions to the derivative with respect of ℓ of (3.25) will be from the coefficient so for $\ell<0$,

$$\partial_\ell \lim_{\omega\to0^+} |\sin(\omega)|^{\frac{m-1}{m}}\partial_\omega R_a f(\boldsymbol{x}^*\cdot\theta^\perp,\theta) =$$
$$\sum_{i=1}^{n}\frac{(d_{m_i}-c_{m_i})}{(m\kappa_{m_i})^{\frac{1}{m}}}(|\ell-\alpha_{m_i}|)^{-\left(1-\frac{1}{m}\right)}\int_{-\infty}^{-\ell+\alpha_{m_i}} f(\boldsymbol{x}^*+t\theta^*)e^{-Da(\boldsymbol{x}^*+t\theta^*,\theta^*)}\,\mathrm{d}t.$$

By an identical argument we obtain for $\ell>0$

$$\partial_\ell \lim_{\omega\to0^+} |\sin(\omega)|^{\frac{m-1}{m}}\partial_\omega R_a f(\boldsymbol{x}^*\cdot\theta^\perp,\theta) =$$
$$\sum_{i=1}^{n}\frac{s_i(d_{m_i}-c_{m_i})}{(m\kappa_{m_i})^{\frac{1}{m}}}(|\ell-\alpha_{m_i}|)^{-\left(1-\frac{1}{m}\right)}\int_{-\infty}^{-\ell+\alpha_{m_i}} f(\boldsymbol{x}^*+t\theta^*)e^{-Da(\boldsymbol{x}^*+t\theta^*,\theta^*)}\,\mathrm{d}t.$$

Taking into account the rotation and translation this agrees with the formula presented in (3.31). By an identical argument using the formulas relating to (3.21) taking into account the rotation and translation we obtain the result for even m, as required. □

Note that this derivative gives us a way of determining tangent points as these will be precisely the points where $\partial_{\boldsymbol{x}^*,\theta^*}\lim_{\omega\to0^+}|\sin(\omega)|^{\frac{m-1}{m}}\partial_\omega R_a f(\boldsymbol{x}^*\cdot\theta^\perp,\theta)$ is discontinuous.

This method of looking at derivatives with respect to ℓ, as in the proof of Corollary 4 gives us a way to tackle straight edges. Applying results similar to those given in this section but to the case of straight edges is the focus of the next section

3.2 Straight edge segments of the boundary

We first examine what happens when we have a single straight edge.

Note that without much analysis we can determine a ray which contains some number of straight edges. This is because the only possibility for a jump in the AtRT across a ray is that there is at least one straight edge on it. This alone is not enough to determine start and end points or indeed how many edges actually lie on the ray. We can however use the following lemmas in order to determine start and end points. We first examine what happens when we have a single straight edge on a ray.

Lemma 10. *Let $f \in C_c^1(\mathbb{R}^2)$ and a be multi-bang. Let (s^*, θ^*) be a ray which contains only a single straight edge $E := \{s^*(\theta^*)^\perp + t\theta^* : \alpha < t < \beta\}$ of a boundary and assume E does not contain any corners for any other part of the boundary. Suppose that $a = c$ on the side $\theta^\perp$ points to at any point on E and c_0 on the opposite side. Additionally assume that $\theta^* = (\cos(\omega^*), \sin(\omega^*))$ and $\boldsymbol{x}^*$ is a point on the ray (s^*, θ^*) then*

$$\partial_{\boldsymbol{x}^* \cdot \theta^*} \lim_{\omega \to (\omega^*)^+} R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) =$$
$$\begin{cases} (c - c_0) \int_{-\infty}^{0} f(\boldsymbol{x}^* + t\theta^*) e^{-\lim \omega \to (\omega^*)^+ Da(\boldsymbol{x}^*)^\perp + t\theta^*, \theta^*)} \mathrm{d}t. & \text{if } \alpha < \boldsymbol{x}^* \cdot \theta^* < \beta \\ 0 & \text{otherwise.} \end{cases} \tag{3.33}$$

Proof. Assume that the intersection of the ray $\{s^*(\theta^*)^\perp + t\theta^* \mid t < \beta\}$ and $\{f > 0\}$ is non-empty, otherwise the result is trivially true. After a rotational transformation, translation and possible reflection we can assume WLOG that $\theta^* = (1, 0)$, $\theta^{*\perp} = (0.1)$ and the point $s^*\theta_\alpha^\perp \theta\alpha = 0$ is the origin i.e, $\omega^* = 0$ and the edge $E := \{(t, 0) : 0 < t < \beta\}$. Following the proof of lemma 8, we assume that $\boldsymbol{x}^* = (\ell, 0)$ and for the moment consider only the case $\ell \neq 0$. Then the line corresponding to $(\boldsymbol{x}^* \cdot \theta^\perp, \theta)$ is precisely the line through $\boldsymbol{x}^*$ tangent to θ, and we will change the parametrization of this line in the integral definition of the AtRT so that $t = 0$ always corresponds with $\boldsymbol{x}^*$. Then we have

$$R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \int_{-\infty}^{\infty} f(\boldsymbol{x}^* + t\theta) e^{-Da(\boldsymbol{x}^* + t\theta, \theta)} \, \mathrm{d}t.$$

We then label the ordered values of t along the line $t \mapsto \boldsymbol{x}^* + t\theta$, for $\omega > 0$ at which the line intersects one of the boundaries $\partial\Omega_j$, other than in the edge, as $\{t_i(\omega)\}_{i=1}^N$. We also set, for $y \cdot \theta^\perp > x \cdot \theta^\perp$,

$$Da^y(\boldsymbol{x}^* + t\theta, \theta) = Da(\boldsymbol{x}^* + t\theta, \theta) - Da(y + t\theta, \theta).$$

We now have 3 cases to consider, either $\ell < 0$, $0 \leq \ell \leq \beta$ or $\beta < \ell$. If $\ell < 0$ then as $\omega \to 0^+$ one of the $t_i = -\ell$ and since E contains no corners for any of the other boundaries then $t_{i+1} = \beta - \ell$. Therefore we have, noting that $(\ell, 0) = \boldsymbol{x}^*$ and $\theta = (1, 0)$

$$\begin{aligned}\lim_{\omega \to 0^+} R_a f(\boldsymbol{x}^*, \theta) = &\int_{-\infty}^{-\ell} f(\boldsymbol{x}^* + t\theta) e^{-Da^{(-\ell,0)}(\boldsymbol{x}^* + t\theta, \theta) - c\beta - Da((\beta,0) - (\ell,0) + t\theta, \theta)} \mathrm{d}t \\ &+ \int_{-\ell}^{\beta - \ell} f(\boldsymbol{x}^* + t\theta) e^{-c(\beta - \ell - t) - Da((\beta,0) - (\ell,0) + t\theta, \theta)} \mathrm{d}t \\ &+ \int_{\beta - \ell}^{\infty} f(\boldsymbol{x}^* + t\theta) e^{-Da(x^* + t\theta, \theta)} \mathrm{d}t.\end{aligned}$$

If we make the substitution $t \mapsto t - \ell$ then the above formula has no dependence on ℓ and so

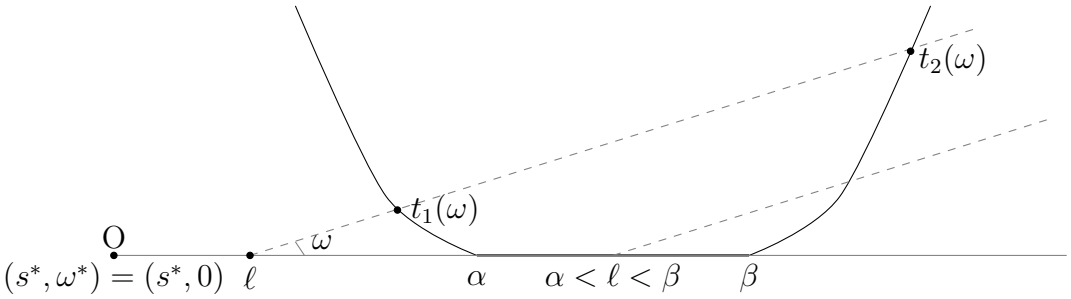
$$\partial_\ell \lim_{\omega \to 0^+} R_a f(\boldsymbol{x}^*, \theta) = 0$$

when $\ell < 0$. Now suppose $\ell > \beta$, then since we are interested in taking the limit as $\omega \to 0^+$ the edge E is no longer part of the beam transform(see Figure 3.3) and so none of the intersection points t_i will correspond with the start and end of E. Furthermore the intersection points t_i as $\omega \to 0^+$ are simply translations by ℓ therefore by making the substitution $t \mapsto t - \ell$, we can see that $\lim_{\omega \to 0^+} R_a f(\boldsymbol{x}^*, \theta)$ does not depend on ℓ and we find

$$\partial_\ell \lim_{\omega \to 0^+} R_a f(\boldsymbol{x}^*, \theta) = 0$$

when $\ell > \beta$.

Figure 3.3: Demonstrating the effect of changing ℓ on the visible edge.



Edge fully visible as $\omega \to 0^+$

Only portion between ℓ and β is visible as $\omega \to 0^+$

The final case to consider is when $0 < \ell < \beta$. Here the edge E is split into two segments one between ℓ and β will be approached from above as $\omega \to 0^+$ and one from 0 to ℓ which will be approached from below, again see Figure 3.3. In particular one of the $t_i = 0$ and $t_{i+1} = \beta - \ell$. In this case we have

$$\begin{aligned}\lim_{\omega\to 0^+} R_a f(\boldsymbol{x}^*, \theta) = &\int_{-\infty}^{0} f(\boldsymbol{x}^* + t\theta) e^{-Da^{(-\ell,0)}(\boldsymbol{x}^*+t\theta,\theta) - c_0\ell - c(\beta-\ell) - Da((\beta,0)-(\ell,0)+t\theta,\theta)} \mathrm{d}t \\ &+ \int_{0}^{\beta-\ell} f(\boldsymbol{x}^* + t\theta) e^{-c(\beta-\ell-t) - Da((\beta,0)-(\ell,0)+t\theta,\theta)} \mathrm{d}t \\ &+ \int_{\beta-\ell}^{\infty} f(\boldsymbol{x}^* + t\theta) e^{-Da(\boldsymbol{x}^*+t\theta,\theta)} \mathrm{d}t.\end{aligned}$$

The term $-c_0\ell - c(\beta - \ell)$ in the exponent of e arises from the fact that we see the edge E from above for all points after $(\ell, 0)$ and from below for all points before $(\ell, 0)$, since the multi-bang value is c above and c_0 below this gives the term $-c_0\ell - c(\beta - \ell)$. The third integral does not depend on ℓ, after a substitution, and so makes no contribution to the derivative with respect to ℓ. The first and second do depend on ℓ, although the only contribution the second integral makes is to cancel the term obtained by evaluating the limits when differentiating under the integral of the first. Therefore we have

$$\begin{aligned}&\partial_\ell \lim_{\omega\to 0^+} R_a f(\boldsymbol{x}^*, \theta) = \\ &(c - c_0) \int_{-\infty}^{0} f(\boldsymbol{x}^* + t\theta) e^{-Da^{(0,0)}(\boldsymbol{x}^*+t\theta,\theta) - c(\beta-\ell) - Da((\beta,0)-(\ell,0)+t\theta,\theta)} \mathrm{d}t.\end{aligned}$$

After taking into account the rotation translation and possible reflection this proves the result. □

Remark 1. *In the proof of lemma 10 we use the fact that changing ℓ translates all of the intersection points in the limit as $\omega \to 0^+$ in order to show independence of the AtRT from ℓ. The importance of ℓ when it lies between 0 and β is that as well as translating the intersection points it also changes the length of the edge which is visible as $\omega \to 0+$, as shown in Figure 3.3. In particular the contribution is only $c(\beta - \ell)$ and it is this change in visible edge which causes the jump in the derivative with respect to ℓ. This property of ℓ is very helpful when we move onto proving results with multiple edges.*

If f and a satisfy the conditions given in lemma 10 then we can determine the start and end points of a single edge by examining the derivative with respect to $\boldsymbol{x}^* \cdot \theta^*$ of the AtRT. Whenever this derivative jumps corresponds with either entering or exiting the edge. We can extend the proof and techniques used in lemma 10 to prove the following result about rays which pass through multiple edges.

Lemma 11. *Let $f \in C_c^1(\mathbb{R}^2)$ and a be multi-bang. Let (s^*, θ^*) be a ray which contains a finite number M of straight edges $E_i := \{s^*(\theta^*)^\perp + t\theta^* : \alpha_i < t < \beta_i\}$ of the boundary, also E_i does not contain any corners for any other part of the boundary. Suppose that $a = c_i$ on the side $\theta^\perp$ points to for any non end point of E_i and $a = d_i$ on the opposite side. Additionally assume that $\theta^* = (\cos(\omega^*), \sin(\omega^*))$ and $\boldsymbol{x}^*$ is a point on the ray (s^*, θ^*) then*

$$\partial_{\boldsymbol{x}^* \cdot \theta^*} \lim_{\omega \to (\omega^*)^+} R_a f(\boldsymbol{x}^* \cdot \theta^{*\perp}, \theta) =$$

$$\begin{cases} (c_i - d_i) \int_{-\infty}^{0} f(\boldsymbol{x}^* + t\theta^*) e^{-\lim \omega \to (\omega^*)^+ Da(\boldsymbol{x} + t\theta^*, \theta^*)} \mathrm{d}t. & \text{if } \alpha < \boldsymbol{x}^* \cdot \theta^* < \beta \\ 0 & \text{otherwise.} \end{cases} \tag{3.34}$$

Proof. As in Lemma 10 we assume that all E_i intersect the set $\{f > 0\}$ otherwise the result is trivially true. As in lemma 10 after a rotational transformation, translation and possible reflection we can assume WLOG that $\theta^* = (1, 0)$, $(\theta^*)^\perp = (0.1)$ and the $\alpha_1 = 0$ is the origin i.e, $\omega^* = 0$ and the edge $E_1 := \{(t, 0) : 0 < t < \beta_1\}$. Following the proof of lemma 10, we assume that $\boldsymbol{x}^* = (\ell, 0)$. Then the line corresponding to $(\boldsymbol{x}^* \cdot \theta^\perp, \theta)$ is precisely the line through $\boldsymbol{x}^*$ tangent to θ, and we will change the parametrization of this line in the integral definition of the AtRT so that $t = 0$ always corresponds with $\boldsymbol{x}^*$. Then we have

$$R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = \int_{-\infty}^{\infty} f(\boldsymbol{x}^* + t\theta) e^{-Da(\boldsymbol{x}^* + t\theta, \theta)} \, \mathrm{d}t.$$

We then label the ordered values of t along the line $t \mapsto \boldsymbol{x}^* + t\theta$, for $\omega > 0$ at which the line intersects one of the boundaries $\partial\Omega_j$ as $\{t_i(\omega)\}_{i=1}^N$. Again for $y \cdot \theta^\perp > x \cdot \theta^\perp$, set

$$Da^y(\boldsymbol{x}^* + t\theta, \theta) = Da(\boldsymbol{x}^* + t\theta, \theta) - Da(y + t\theta, \theta).$$

There are four cases of importance: $\ell < \alpha_1$, $\alpha_i < \ell < \beta_i$ for some $M \geq i \geq 1$, $\beta_i < \ell < \alpha_{i+1}$ for some $M - 1 \geq \ell \geq 1$ and $\ell > \beta_M$.

We first examine the case where $\ell < \alpha_1$. Then we have M pairs of intersection points which correspond with the start and end of the edges E_i. In particular, since each edge contains no corners, the pairs of intersections points are adjacent to each other in the set of intersection points. Furthermore, as each intersection point is of the form $\beta_i - \ell$ or $\alpha_i - \ell$ the only effect that ℓ, provided $\ell < 0$, has on the data is to shift all of the intersection points. By the translation $t \mapsto t - \ell$ we can again see that the AtRT is independent of ℓ and so, recalling that $\boldsymbol{x}^* = (\ell, 0)$,

$$\partial_\ell R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = 0$$

for $\ell < \alpha_1$.

An identical argument as given in the case of $\ell > \beta$ in Lemma 10 can be used in the case of multiple edges when $\ell > \beta_M$ and so we again find that

$$\partial_\ell R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = 0$$

for $\ell > \beta_M$.

This leaves us with two cases to consider $\alpha_i < \ell < \beta_i$ for some $1 \leq i \leq M$, i.e the point $\boldsymbol{x}^*$ lies on one of the edges E_i or $\beta i < \ell < \alpha_{i+1}$ for some $1 \leq i \leq M - 1$, i.e the point $\boldsymbol{x}^*$ lies between two edges. We first consider the case when $\boldsymbol{x}^*$ lies between two edges. In this case the only edges which will be visible are the edges which lie to the right of $\boldsymbol{x}^*$ i.e E_k where $M \geq k \geq i + 1$. By remark 1 and the proof of lemma 10 we find that the only effect that ℓ has, provided $\beta_i < \ell < \alpha_{i+1}$, is to translate the intersection points. Again by substituting we find that the AtRT is independent of ℓ. Therefore we have

$$\partial_\ell R_a f(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = 0$$

when $\beta_i < \ell < \alpha_{i+1}$ for $1 \leq i \leq M-1$.

We now move onto the final case which is when $\boldsymbol{x}^*$ lies on one of the edges E_i. Let $\boldsymbol{x}^*$ lie on the ith edge so that $\alpha_i < \ell < \beta_i$. As in the previous case only edges which lie to the right of E_i and the portion of E_i between ℓ and β_i will be visible as $\omega \to 0^+$. In this case we have

$$\begin{aligned}\lim_{\omega\to 0^+} R_a f(\boldsymbol{x}^*,\theta) = &\int_{-\infty}^{0} f(\boldsymbol{x}^*+t\theta) e^{-Da^{(-\ell,0)}(\boldsymbol{x}^*+t\theta,\theta)-d_i(\ell)-c_i(\beta_i-\ell)-Da((\beta_i,0)-(\ell,0)+t\theta,\theta)}\mathrm{d}t\\ &+\int_{0}^{\beta_i-\ell} f(\boldsymbol{x}^*+t\theta) e^{-c_i(\beta_i-\ell-t)-Da((\beta_i,0)-(\ell,0)+t\theta,\theta)}\mathrm{d}t\\ &+\int_{\beta_i-\ell}^{\infty} f(\boldsymbol{x}^*+t\theta) e^{-Da(\boldsymbol{x}^*+t\theta,\theta)}\mathrm{d}t.\end{aligned}$$

Note that the edges on the right of E_i are completely visible and the intersection points are just translated by ℓ so after translation they do not depend on ℓ. As in remark 1 in addition to the translation ℓ effects the visible portion of E_i and so this has a non-zero contribution to the derivative. The second and third integrals, exactly the same as in Lemma 10 make, no contribution and so we have

$$\begin{aligned}&\partial_\ell \lim_{\omega\to 0^+} R_a f(\boldsymbol{x}^*,\theta) =\\ &(c_i-d_i)\int_{-\infty}^{0} f(\boldsymbol{x}^*+t\theta) e^{-Da^{(0,0)}(\boldsymbol{x}^*+t\theta,\theta)-c_i(\beta_i-\ell)-Da((\beta_i,0)-(\ell,0)+t\theta,\theta)}\mathrm{d}t\end{aligned}$$

for $\alpha_i < \ell < \beta_i$. After taking into account the rotation translation and possible reflection this agrees with formula (3.34), as required. □

Remark 2. *Note that the proof of lemma 11 only requires the interior of the edges to be disjoint and not their endpoints. Therefore we can have $\alpha_{i+1} = \beta_i$ for some i and the result still holds.*

Lemma 11 gives us the desired technique to both determine how many edges on the boundary lie on a single ray as well as their endpoints. The only case remaining to deal with is when a ray passes through some number of corners. As stated at the start of Chapter 3 it is not necessary to determine corners using jumps as we can simply take the closure of all the edges and points on an analytic part of the boundary. Nonetheless, we do include analysis even in the case as this may be on independent interest.

3.3 Corner points

The next few lemmas which we present relate to corners of the boundary. The first result is also presented in [37] for nested convex domains but the result and the proof extend naturally to the more general multi-bang setting.

Lemma 12. *Suppose that $f \in C_c^1(\mathbb{R}^2)$, a is multi-bang, and $(s^*, \theta^*) \in \mathcal{K}_{a,2} \setminus \mathcal{K}_{a,1}$ and passes though exactly one corner . Then $s \mapsto \partial_s R_a f(s, \theta^*)$ is bounded near s^*.*

Suppose additionally that the corner point occurs at $s^*(\theta^*)^\perp + t^*\theta^*$, is a corner for N different components of the regions Ω_j, and the boundaries of these regions make angles $\{\alpha_k\}_{k=1}^N$ with $(\theta^*)^\perp$ where the orientation is chosen so that θ^* is at a positive angle. Also suppose that the jump in a across the boundary with angle α_k in the direction of increasing angle is b_k (see figure 3.4 and caption). Then there is a jump in $\partial_s R_a f(s, \theta^*)$ across $s = s^*$ given by

$$\left[\partial_s R_a f(s,\theta^*)\right]_{s_-^*}^{s_+^*} = \left(\sum_{k=1}^{N} b_k \tan(\alpha_k)\right) \int_{-\infty}^{t^*} f(s^*(\theta^*)^\perp + t\theta^*) e^{-Da(s^*(\theta^*)^\perp + t\theta^*, \theta^*)} \, \mathrm{d}t. \tag{3.35}$$

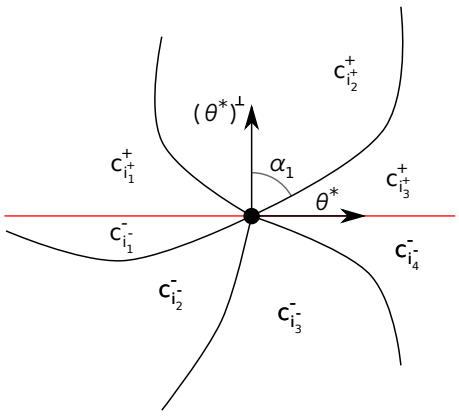


Figure 3.4: This figure illustrates some of the notation used in the statement in lemma 12, this figure is taken from [37]. The line corresponding to (s^*, θ^*) is shown in red. The jump b_1 across the boundary corresponding to the angle α_1 in the figure is $b_1 = c^+_{i_3^+} - c^+_{i_2^+}$.

Proof. The proof we follow for this result is given in [37, Lemma 2]. We use the same set-up and notation as in the proof of Lemma 7. By translating if necessary, we also assume W.L.O.G. that the corner occurs at the origin. There may be a different number of boundary crossings N when $s > 0$ and $s < 0$, and so we introduce corresponding functions $\{t_i^\pm\}_{i=1}^{N^\pm}$ giving the crossing points where the t_i^+ are defined for

$s > 0$ and the t_i^- are defined for $s < 0$. As in Lemma 7 these $t_i^\pm$ will all have bounded derivatives up to $s = 0$ (this is because the line given by (s^*, θ^*) is not tangent to any boundary). We also introduce the corresponding $\phi_i^\pm$ defined as in (3.2) but only for $s > 0$ and $s < 0$ respectively. The formula (3.38) still holds with $\pm$ added in appropriate places, and we can still see that $Da((x,s),\theta^*)$ is continuous for s close to zero. For the derivative $\partial_s Da((x,s),\theta^*)$ we have for $s \neq 0$, where $\pm$ is the sign of s,

$$\partial_s Da((x,s),\theta^*) = \sum_{i=1}^{N^\pm} (c_{i-1}^\pm - c_i^\pm)\partial_s \phi_i^\pm(x,s). \tag{3.36}$$

Since these derivatives are all bounded (but not necessarily continuous at $s = 0$) we still have (3.40) for $s \neq 0$, and we see that $\partial_s R_a f(s,\theta^*)$ is bounded thus proving the first statement of the theorem. It remains to analyse the jump at $s = 0$.

Let us first consider the jump in $\partial_s Da((x,s),\theta^*)$ across $s = 0$. The only terms contributing to this jump in (3.36) will be those with $\phi_i^\pm$ where $t_i^\pm(s) \to 0$ as $s \to 0^\pm$ since the others correspond to boundaries which do not have corners along (s^*,θ^*) and the line is not tangent to any of the boundaries. Let us reindex the indices i corresponding to such t_i using a new index k as $\{i_k^\pm\}_{k=1}^{\tilde{N}^\pm}$. Then the jump in $\partial_s Da((x,s),\theta^*)$ is given by

$$\begin{aligned}\Big[\partial_s Da((x,s),\theta^*)\Big]_{0^-}^{0^+} &= \lim_{s\to 0^+} \partial_s Da((x,s),\theta^*) - \lim_{s\to 0^-} \partial_s Da((x,s),\theta^*) \\ &= \left[\sum_{k=1}^{\tilde{N}^+} (c_{i_k^+-1}^+ - c_{i_k^+}^+)\partial_s \phi_{i_k^+}^+(x,0^+)\right] \\ &\quad - \left[\sum_{k=1}^{\tilde{N}^-} (c_{i_k^--1}^- - c_{i_k^-}^-)\partial_s \phi_{i_k^-}^-(x,0^-)\right].\end{aligned}$$

Using (3.40) we find that the jump of $\partial_s R_a f(s,\theta^*)$ across $s = 0$ will be

$$\begin{aligned}&\Big[\partial_s R_a f(s,\theta^*)\Big]_{0^-}^{0^+} \\ &= \lim_{s\to 0^+} \partial_s R_a f(s,\theta^*) - \lim_{s\to 0^-} \partial_s R_a f(s,\theta^*) \\ &= -\left(\left[\sum_{k=1}^{\tilde{N}^+} (c_{i_k^+-1}^+ - c_{i_k^+}^+)\partial_s t_{i_k^+}^+(0^+)\right] - \left[\sum_{k=1}^{\tilde{N}^-} (c_{i_k^--1}^- - c_{i_k^-}^-)\partial_s t_{i_k^-}^-(0^-)\right]\right) \\ &\quad \times \int_{-\infty}^0 f(x,0) e^{-Da((x,0),\theta^*)}\,\mathrm{d}x.\end{aligned}$$

Taking into account the rotation and translation used at the beginning we see that this corresponds with (3.35) and so completes the proof. □

This result can be extended to the case where a ray passes through multiple corners, provided that the ray is not tangent to any boundary $\partial\Omega_j$. Note than the tangency condition comes from the fact that, as shown in lemmas 8 and 9, the mapping $s \mapsto \partial_s R_a f(s, \theta^*)$ blows up near s^*. This blow up on first sight appears to obscure any information about possible corners but we shall see later that even the case where a ray passes through any number of corners and is tangent to some parts of the boundary can be resolved. The next result tackles multiple corner cases.

Lemma 13. *Suppose that $f \in C_c^1(\mathbb{R}^2)$, a is multi-bang, and $(s^*, \theta^*) \in \mathcal{K}_{a,2} \setminus \mathcal{K}_{a,1}$ and passes though M corners. Then $s \mapsto \partial_s R_a f(s, \theta^*)$ is bounded near s^*.*

Suppose additionally that the corner points occur at $s^*(\theta^*)^\perp + {t_i}^*\theta^*$ and each is a corner for N_i different components of the regions Ω_j, and the boundaries of these regions make angles $\{\alpha_k^i\}_{k=1}^{N_i}$ with $(\theta^*)^\perp$ where the orientation is chosen so that θ^* is at a positive angle. Also suppose that the jump in a across the boundary with angle α_k^i in the direction of increasing angle is b_k^i. Then there is a jump in $\partial_s R_a f(s, \theta^*)$ across $s = s^*$ given by

$$\begin{aligned}&\left[\partial_s R_a f(s, \theta^*)\right]_{s_-^*}^{s_+^*} = \\ &\sum_{i=1}^{M} \left\{ \left(\sum_{k=1}^{N_i} b_k^i \tan(\alpha_k^i) \right) \int_{-\infty}^{t_i^*} f(s^*(\theta^*)^\perp + t\theta^*) e^{-Da(s^*(\theta^*)^\perp + t\theta^*, \theta^*)} \, \mathrm{d}t \right\}.\end{aligned} \tag{3.37}$$

Proof. We use the same set-up and notation as in the previous lemmas, and we also assume W.L.O.G. that the corner corresponding to t_1^* is now at the origin and the other corners are located at $\{\boldsymbol{\alpha_i}\}_{i=2}^M$, with $\boldsymbol{\alpha_i} = (\alpha_i, 0)$ with $\alpha_i > 0$ and $\boldsymbol{\alpha_1} = (0, 0)$. As in Lemma 12, there may be a different number of boundary crossings at $s^*(\theta^*)^\perp + t_i\theta^*$ when $s > 0$ and $s < 0$. Again, we introduce functions $\{t_i^\pm\}_{i=1}^{N^\pm}$ giving the crossing points where the t_i^+ are defined for $s > 0$ and the t_i^- are defined for $s < 0$. These $t_i^\pm$ will all have bounded derivatives up to $s = 0$ (this is because the line given by (s^*, θ^*) is not tangent to any boundary). We also introduce the corresponding $\phi_i^\pm$ defined as in lemma 9 but only for $s > 0$ and $s < 0$ respectively. In this case we have

$$Da((x, s), \theta^*) = \sum_{i=1}^{N^\pm - 1} c_i(\phi_{i+1}^\pm(x, s) - \phi_i^\pm(x, s)) = \sum_{i=1}^{N^\pm} (c_{i-1} - c_i)\phi_i^\pm(x, s) \tag{3.38}$$

and we have that $Da((x, s), \theta^*)$ is continuous for s close to zero. For the derivative

$\partial_s Da((x,s),\theta^*)$ we have for $s \neq 0$, where $\pm$ is the sign of s,

$$\partial_s Da((x,s),\theta^*) = \sum_{i=1}^{N^\pm} (c_{i-1}^\pm - c_i^\pm)\partial_s \phi_i^\pm(x,s). \tag{3.39}$$

Since these derivatives are all bounded (but not necessarily continuous at $s=0$) we have

$$\begin{aligned}\partial_s R_a f(s,\theta^*) &= \int_{-\infty}^{\infty} \left(\partial_s f(x,s) - \partial_s Da((x,s),\theta^*) f(x,s)\right) e^{-Da((x,s),\theta^*)} \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \partial_s f(x,s) e^{-Da((x,s),\theta^*)} \mathrm{d}x \\ &\quad + \sum_{i=1}^{N^\pm} (c_i - c_{i-1}) t^{\pm\prime}{}_i(s) \int_{-\infty}^{t_i^\pm(s)} f(x,s) e^{-Da((x,s),\theta^*)} \mathrm{d}x\end{aligned} \tag{3.40}$$

for $s \neq 0$, and we see that $\partial_s R_a f(s,\theta^*)$ is bounded thus proving the first part of the theorem. We now analyse the jump at $s=0$.

We first consider the jump in $\partial_s Da((x,s),\theta^*)$ across $s=0$. The terms contributing to this jump in (3.39) will be those arising from the corners i.e, with $\phi_i^\pm$ where $t_i^\pm(s) \to 0$ as $s \to 0^\pm$ since the others correspond to boundaries which do not have corners along (s^*,θ^*) and the line is not tangent to any of the boundaries. We relabel the indices which are related to the jth corner by $\{i^j{}_k^\pm\}_{k=1}^{\tilde{N}_j^\pm}$. Note here that $\tilde{N}_i^+ + \tilde{N}_i^- = N_i$. With this notation the jump in $\partial_s Da((x,s),\theta^*)$ is given by

$$\begin{aligned}\Big[\partial_s Da((x,s),\theta^*)\Big]_{0^-}^{0^+} &= \lim_{s\to 0^+} \partial_s Da((x,s),\theta^*) - \lim_{s\to 0^-} \partial_s Da((x,s),\theta^*) \\ &= \sum_{j=1}^{M} \Bigg\{ \left[\sum_{k=1}^{\tilde{N_j}^+} (c^+_{i^j{}_k^+ - 1} - c^+_{i^j{}_k^+}) \partial_s \phi^+_{i^j{}_k^+}(x,0^+)\right] - \\ &\qquad \left[\sum_{k=1}^{\tilde{N_j}^-} (c^-_{i^j{}_k^- - 1} - c^-_{i^j{}_k^-}) \partial_s \phi^-_{i^j{}_k^-}(x,0^-)\right] \Bigg\}.\end{aligned}$$

By (3.40) we find that the jump of $\partial_s R_a f(s,\theta^*)$ across $s=0$ is given by

$$\begin{aligned}\Big[\partial_s R_a f(s,\theta^*)\Big]_{0^-}^{0^+} &= \lim_{s\to 0^+} \partial_s R_a f(s,\theta^*) - \lim_{s\to 0^-} \partial_s R_a f(s,\theta^*) \\ &= -\sum_{j=1}^{M} \Bigg\{ \Bigg(\left[\sum_{k=1}^{\tilde{N}_j^+} (c^+_{i^j{}_k^+ - 1} - c^+_{i^j{}_k^+}) \partial_s t^+_{i^j{}_k^+}(0^+)\right] - \\ &\qquad \left[\sum_{k=1}^{\tilde{N}_j^-} (c^-_{i^j{}_k^- - 1} - c^-_{i^j{}_k^-}) \partial_s t^-_{i^j{}_k^-}(0^-)\right] \Bigg) R_j \Bigg\}\end{aligned}$$

where

$$R_j = \int_{-\infty}^{\alpha_j} f(x,0)e^{-Da((x,0),\theta^*)}.$$

Taking into account the rotation and translation used at the beginning we see that this corresponds with (3.35) as required. □

Something we have not yet talked about is the possibility for the formulas given in lemmas 9 and 13 to be 0 for some x^* on a line. This can only happen when certain singularities or corners along a ray sum together to cancel out for a specific x^*, this is not an issue however as they cannot cancel for all x^* on a ray.

This concludes the corner section and the following small section brings all of the results from Chapter 3.1-3.3 together to prove one of the main results.

3.4 Unique recovery of boundaries $\{\partial\Omega\}_{j=1}^n$

With Corollary 4, Lemmas 11 and 13 we are ready to present the main result in this section.

Theorem 5. *Suppose that $f \in C_c^1(\mathbb{R}^2)$ is non-negative, a is multi-bang as in definition 6. If for every point on the boundary $\boldsymbol{x}$ there exists some ray $(s,\theta) \in \mathcal{K}_a$ which passes the boundary tangentially at $\boldsymbol{x}$ and intersects the support of f, then we can uniquely recover the boundaries $\{\partial\Omega_j\}_{j=1}^n$.*

Proof. We begin by focussing our attention on the tangent points on piecewise analytic parts of the boundary. Let $(s,\theta) \in \mathcal{K}_{a,1}$, then using Corollary 4 we can determine all points of tangency on the boundary with the largest singularity order. Once these points have been identified we can use Lemma 9 to remove these singularities influence on the data. We can then recursively determine tangency points of lower singularity order until we have determined them all, with the possible exception of a countable set. Note that since tangency points cause a blow up in the derivative along rays in $\mathcal{K}_{a,1}$ any corners which such a ray might pass through has no impact, since the jumps for corners are bounded. Thus we are able to determine all points on the boundary which lie on a piecewise analytic part of the boundary.

With the tangent points dealt with we can then use Lemma 11 to determine the start and end points of all edges on the boundary.

The only set we have left to deal with is the set of rays $\mathcal{K}_{a,2}$, i.e those passing through the corners. By lemmas 12 and 13, after recursively removing singularities from the data any remaining corner on the boundary has the property that all but two rays passing through that point have a bounded jump in the derivative with respect to s. This is since the two tangents going into and out of the corner produce unbounded jumps and have already been removed. Therefore we are able to determine all corners of the boundaries $\{\Omega_j\}_{j=1}^n$. Since piecewise analytic boundaries with corners are made up of precisely corners, flat edges and points on an analytic boundary we can uniquely recover $\{\partial\Omega_j\}_{j=1}^n$, as required. □

Before moving onto the next chapter containing results about recovering a and f from SPECT data, we include a small section about the special case of nicely multi-bang a as in Definition 14.

3.5 Recovery of boundaries in the nicely multi-bang case

The main result in this section Lemma 15 was presented as an intermediate step in the proof of Theorem 6 in [37]. Before we go into detail we first define the notion of curvature for the graph of a function.

Definition 19. *Suppose we have a function $y = f(x)$ with $f(x)$ at least twice differentiable. Then the curvature $\kappa(x)$ is given by*

$$\kappa(x) = \frac{f''(x)}{(1 + f'(x))^{\frac{3}{2}}}. \tag{3.41}$$

We can evaluate the signed curvature $\kappa(x)$ at a point x^* to obtain the curvature of the graph (x, y) at the point x^*.

From Definition 5 for a point x^* on an analytic part of a boundary the boundary is given, at least locally, by $y = (x - x^*)^n g(x)$ with $g(x^*) \neq 0$ and $n \geq 2$. In this geometry we have $\frac{\mathrm{d}y}{\mathrm{d}x}|_{x=x^*} = 0$. This gives

$$\kappa(x^*) = \begin{cases} 0 & \text{if } n \geq 3 \\ 2g(x^*) & \text{if } n = 2. \end{cases} \tag{3.42}$$

Equation (3.42) shows we only have non-zero curvature when $n = 2$, we refer to any x^* with $\kappa(x^*) \neq 0$ as a point with non-zero curvature.

Remark 3. *The choice to use the parameter κ in chapter 3 is to signify that curvature is just a special case where $n = 2$.*

Theorem 5 shows that we can recover the boundaries of a in the general multi-bang case, and therefore the nicely multi-bang setting in which the formulas are significantly simpler, but we make use of a potentially very large amount of the boundary. With the notion of points of non-zero curvature we are able to recreate nicely multi-bang boundaries without resorting to points on the boundary with zero curvature.

As boundaries of nicely multi-bang a are boundaries of nested convex sets, and we intend to recover these sets with non-zero curvature points, a helpful concept is that of a convex hull.

Definition 20 (**Convex hull).** *Let $\mathcal{P}$ be a bounded set of points in 2D space. The convex hull of $\mathcal{P}$ denoted* conhull$(\mathcal{P})$ *is defined as the unique minimal convex set containing all points in $\mathcal{P}$.*

In order to prove the main result in this section we first need the following geometric lemma.

Lemma 14. *Suppose that $C \subset \mathbb{R}^2$ is closed, convex, bounded and has smooth boundary possibly with corners. Also let $\mathcal{P}$ be the subset of ∂C consisting of points which are either corners of ∂C, or where ∂C has nonzero curvature. Then*

$$C = \overline{\text{conhull}\,(\mathcal{P})}$$

where conhull$(\mathcal{P})$ *is the convex hull of $\mathcal{P}$ as in Definition 20.*

Proof. Since C is closed and convex, and $\mathcal{P} \subset C$, we have $\overline{\text{conhull}\,(\mathcal{P})} \subset C$. Thus it only remains to show the opposite inclusion. Suppose that $\boldsymbol{x} \in \partial C \setminus \overline{\text{conhull}\,(\mathcal{P})}$. Then there must be a neighbourhood U of $\boldsymbol{x}$ such that $U \cap \partial C$ does not intersect $\mathcal{P}$. Therefore the curvature of ∂C is zero at all points in $U \cap \partial C$, and since $\boldsymbol{x}$ is also not a corner for ∂C, this implies that $U \cap \partial C$ must contain a line segment containing $\boldsymbol{x}$ in its relative interior. There must then be some maximally extended line segment containing $\boldsymbol{x}$ which is contained in ∂C. At least one of the end points of this maximal line segment

must also be in $\partial C \setminus \overline{\text{conhull}\,(\mathcal{P})}$ since otherwise we would have $\boldsymbol{x} \in \overline{\text{conhull}\,(\mathcal{P})}$ by convexity. However this is a contradiction since by the argument we have already given this endpoint would be in the relative interior of a line segment contained in $\partial C \setminus \overline{\text{conhull}\,(\mathcal{P})}$. Thus $\partial C \setminus \overline{\text{conhull}\,(\mathcal{P})} = \emptyset$, which then implies the result. □

Lemma 14 allows the recovery of a single closed convex bounded set from a subset of its boundary. Lemma 14 is also useful in determining multiple nested convex sets, which is the case for nicely multi-bang a. Suppose that we have two nested closed convex and bounded sets $C_2 \subset C_1$ and know the set of all points $\mathcal{P}$ which are either points of non-zero curvature or corners of ∂C_1 or ∂C_2. We can take the convex hull of $\mathcal{P}$ and note that as $C_2 \subset C_1$ the points in $\mathcal{P}$ relating to ∂C_2 will lie in the interior of $\mathring{conhull}\,(\mathcal{P})$ and so will have no effect on the boundary of the convex hull. Therefore, by Lemma 14 the convex hull of $\mathcal{P}$ is equal to C_1 and we therefore know ∂C_1. We can then replace $\mathcal{P}$ by $\mathcal{P} \setminus \partial C_1$ to remove any points relating to the boundary of C_1. We can then use the convex hull of $\mathcal{P} \setminus \partial C_1$ and Lemma 14 to determine C_2. This technique can be generalized to any finite number of nested sets and this idea is the basis of the proof of the following result.

Lemma 15. *Suppose that a is nicely multi-bang (see Definition 14) and $f \in C_c^1(\mathbb{R}^2)$ is non-negative. Also assume that*

1. *for all $\boldsymbol{x} \in \mathcal{K}_{a,1}$ the line tangent to a boundary at x passes through the set $\{f > 0\}$, and*

2. *for all $\boldsymbol{x} \in \mathcal{K}_{a,2}$ there is a line passing through x that also passes through the set $\{f > 0\}$,*

where $\mathcal{K}_{a,1}$ and $\mathcal{K}_{a,2}$ are as in Definition 18. Then we can determine from $R_a f$ the sets C_j appearing in Definition 14 for a.

Proof. We first note that by Lemmas 7 and 12 and Corollary 3 the set of (s^*, θ^*) such that $\partial_s R_a f(s, \theta^*)$ for s near s^* is not bounded gives the set of lines which are tangent to some boundary ∂C_j, possibly missing some of the lines which intersect a boundary in a line segment. We can get rid of all of the (s^*, θ^*) corresponding to lines which intersect a boundary ∂C_j in a line segment by looking at the continuity of $R_a f(s, \theta^*)$ near s^* and using Lemma 10. Thus we can determine the set of (s^*, θ^*) such that the

corresponding lines are tangent to a boundary ∂C_j at some point, and since the C_j are nested convex sets the point of tangency along each such line must be unique. We can determine the point of tangency along each line that is tangent at a point where the curvature of ∂C_j is not zero using Lemma 8 or we can determine if at the point of tangency the curvature is zero using Corollary 3. Thus we can identify all points in the boundaries of the C_j at which the curvature of ∂C_j is not zero. Next we will show that we can also find the corners of the boundaries ∂C_j.

By Lemma 12 and the hypotheses, for every corner point $\boldsymbol{x}$ for some ∂C_j there will infinitely many lines passing through $\boldsymbol{x}$ such that for at least (s^*, θ^*) corresponding to these lines $\partial_s R_a f(s, \theta^*)$ is bounded, but has a jump at $s = s^*$. This allows us to determine the corner points, and combining this with the previous paragraph we see that we can determine the $\mathcal{P}_a$ from $R_a f$ under the given hypotheses. We next show that this is sufficient to determine all of the C_j.

By Lemma 14 the closure of the convex hull of $\mathcal{P}_a$ is equal to the closure of C_1. Therefore we can determine C_1. The rest of the sets C_j can now be determined inductively. Indeed, suppose that we know C_l for all $l < j$. Then by Lemma 14 again

$$\overline{C_j} = \overline{\text{conhull}\left(\mathcal{P}_a \setminus \bigcup_{l=1}^{j-1} \partial C_l\right)},$$

and so we can determine C_j. This completes the proof. □

This concludes the chapter on boundary recovery of a from SPECT data. The next chapter contains results relating to the recovery of a and f SPECT data.

Chapter 4

Joint recovery of a and f from SPECT data

This chapter contains novel results on the recovery of a and f from SPECT data when a is multi-bang and $f \in C_c^1(\Omega)$. This chapter is split into two sections. The first section contains results relating to nicely multi-bang a, where each of the regions of a are nested convex sets. The results from the first section are also given in [37]. The second section relates to the more general not nicely multi-bang case. In the case of nicely multi-bang functions uniqueness of a and f is explicitly shown. Unfortunately the non-convex case is more involved and a full uniqueness proof is currently unknown. However, we still have some novel intermediate results relating to the general multi-bang case and also some conditions on a which guarantee a unique recovery.

4.1 Nicely multi-bang recovery

This section contains results from the paper [37] relating to the unique recovery of a and f from SPECT data. In the case of nicely multi-bang a, as in Definition 14, Lemma 15 shows that we can uniquely determine the sets C_j which make up a. In theory we would only need to uniquely determine a and then apply Theorem 3 to determine f. Since the outermost layer must contain a point of non-zero curvature we can use either Lemma 12 or 8 to determine a in this region. Unfortunately, determining the values that a should take in any other region requires analysing the jumps in the derivative of the data, and these depend on integrals of f. In the case of nicely multi-bang a it is

useful to use the convex nature of the boundaries of a to determine f in the outermost region of a and then use the knowledge of f to determine a in the next region.

In order to do this we make heavy use of the following integral geometry result which is given in [15, Theorem 3.1].

Lemma 16. *Suppose that in the strip $D = \{\xi, \eta | \xi \in \mathbb{R}, \eta \in (0, h)\}$ we have the set of parabolas*

$$\gamma(x, y) : \eta = y - (\xi - x)^2, \quad \xi \in (\xi_1, \xi_2)$$

with vertices $x, y \in D$ and end points on the $\eta = 0$ axis. For a continuous weight function $w(\xi, x, y)$ and a finite continuous function $u \in C_0(\Omega)$ set

$$Pu(x, y) = \int w(\xi, x, y) u(\xi, y - (\xi - x)^2) \mathrm{d}\xi. \tag{4.1}$$

Suppose that w is continuous with respect to $y \in [0, h]$ analytic with respect to $\xi, \boldsymbol{x}$ in the complex plane

$$|\mathrm{Im} x| < \delta, \qquad |\xi - x| < \delta, \quad x, \xi \in \mathbb{C},$$

and there exists some $\rho > 0$ such that $|w(x, x, y)| \geq \rho$ for all $x \in \mathbb{C}, |\mathrm{Im} x| < \delta, y \in [0, h]$. Then the solution of the equation $Pu = g$ is unique in $L^1(D)$.

Proof. A proof of this result is given in [15, Theorem 3.1]. □

The general idea when trying to apply Lemma 16 to the problem of SPECT is to consider first the outermost layer. First imagine that the support of a is above the x axis. Suppose that we can find a shallow parabola which lies under the outermost boundary of a and that we know a under this shallow parabola. Since f is compactly supported we know that we can find some equally shallow parabola under the support of f and the aim is to recover f in-between these two parabolas. From the data we know all straight lines passing through the region between the two parabolas. We can perform a change of variables to make the two parabolas flat and all the straight line integrals become integrals along parabolas instead. With this setup we can then apply the result in Lemma 16. Exactly how this is done is given in detail in proof of Lemma 17. The property of convexity in nicely multi-bang a ensures that finding parabolas in this way is possible for every point on the outermost boundary, at which the boundary is not flat. The nested property is needed to ensure that the weight function w, which in our case is related to a, is analytic in the required region.

Before presenting the major result of unique recovery of a and f we first present the following lemma which allows us to determine f in regions where we know a.

Lemma 17. *Suppose that $f \in C_c^1(\mathbb{R}^2)$, and a is nicely multi-bang with sets $\{C_j\}_{j=1}^n$, and let C_0 be an open ball centred at the origin that is sufficiently large so that $C_1 \Subset C_0$ and* $\operatorname{supp}(f) \Subset C_0$. *Then for $j \geq 1$, $f|_{C_{j-1} \setminus C_j}$ is uniquely determined if we know all of*

1. *$R_a f$,*
2. *the sets C_j,*
3. *$a|_{\mathbb{R}^2 \setminus C_j}$, and*
4. *$f|_{\mathbb{R}^2 \setminus C_{j-1}}$.*

Proof. For this proof we will write (x, y) as Cartesian coordinates for points in $\mathbb{R}^2$. By translating and rotating as necessary we assume W.LO.G. that C_j is contained in the upper half plane $\{y > 0\}$, and show that we can then uniquely determine f restricted to the lower half-plane $\{y < 0\}$. By translating to bring C_j arbitrarily close to $\{y = 0\}$ and rotating this then shows we can determine f everywhere outside of C_j and so will complete the proof. See Figure 4.1.

Having done the transformations described in the previous paragraph, we also assume that $C_{j-1} \subset \{y > -h\}$. Now choose $\omega > \epsilon > 0$ such that the parabola $\{y = \epsilon x^2\}$ lies entirely outside of C_j and the parabola $\{y = \omega x^2 - h\}$ lies entirely outside of C_{j-1}. It is possible to find such ω and ϵ since the C_j are all bounded. Now choose $\phi \in C^\infty(\mathbb{R}^2)$ such that $\phi(x, y) = 1$ on C_{j-1} and $\phi(x, y) = 0$ on the set $\{y < \omega x^2 - h\}$. Then define $\tilde{f} = \phi f$ so that $\tilde{f}$ has support contained in the set $\{y \geq \omega x^2 - h\}$ and is such that $\tilde{f}|_{C_{j-1}} = f|_{C_{j-1}}$. Also, supposing that $a = c$ on $C_{j-1} \setminus C_j$, we set

$$\tilde{a} = \begin{cases} \tilde{a}(x, y) = a(x, y) & (x, y) \in C_{j-1} \\ \tilde{a}(x, y) = c & (x, y) \in (\mathbb{R}^2 \setminus C_{j-1}) \cap \{y > \epsilon x^2 - h - 1\} \\ \tilde{a}(x, y) = 0 & \text{otherwise.} \end{cases}$$

The setup described in the last few lines is illustrated in figure 4.1.

Our next step is to show that we can determine $R_{\tilde{a}} \tilde{f}(s, \theta)$ if (s, θ) corresponds to a line contained in the set $\{y < \epsilon x^2\}$ given the hypotheses of the lemma. If this line

does not pass through C_{j-1}, then there is no problem since we know $\tilde{f}$ and $\tilde{a}$ outside of C_{j-1}. Suppose on the other hand that the line does pass through C_{j-1}, and let the two points of intersection between the line and ∂C_{j-1} be denoted $t_1 < t_2$ (note there will always be two such points by convexity and these can be determined from C_j). We then have

$$R_a f(s,\theta) = \int_{-\infty}^{t_1} f(s\theta_\perp + t\theta) e^{-Da(s\theta_\perp + t\theta,\theta)}\, \mathrm{d}t + \int_{t_1}^{t_2} f(s\theta_\perp + t\theta) e^{-Da(s\theta_\perp + t\theta,\theta)}\, \mathrm{d}t + \int_{t_2}^{\infty} f(s\theta_\perp + t\theta) e^{-Da(s\theta_\perp + t\theta,\theta)}\, \mathrm{d}t.$$

The first and third terms on the right side of the last equation only involve $a|_{\mathbb{R}^2\setminus C_j}$ and $f|_{\mathbb{R}^2\setminus C_{j-1}}$ as well as t_1 and t_2, and thus are known functions of (s,θ) under the given hypotheses. We combine these together, and also $R_a f$, into one function $G(s,\theta)$, and so, since also $f|_{C_{j-1}} = \tilde{f}|_{C_{j-1}}$, we have

$$\int_{t_1}^{t_2} \tilde{f}(s\theta_\perp + t\theta) e^{-Da(s\theta_\perp + t\theta,\theta)}\, \mathrm{d}t = G(s,\theta)$$

where G is a function which can be determined from the known information. Next note that for $t \in (t_1, t_2)$, $F = -Da(s\theta_\perp + t\theta,\theta) + D\tilde{a}(s\theta_\perp + t\theta,\theta)$ only depends on s and θ, and can be determined under the hypotheses. Therefore we have

$$\int_{t_1}^{t_2} \tilde{f}(s\theta_\perp + t\theta) e^{-D\tilde{a}(s\theta_\perp + t\theta,\theta)}\, \mathrm{d}t = e^{-F(s,\theta)} G(s,\theta).$$

Finally, we can add back in the integrals with $\tilde{f}$ and $\tilde{a}$ from $-\infty$ to t_1 and t_2 to ∞ since these only involve $\tilde{a}|_{\mathbb{R}^2\setminus C_j}$, and $\tilde{f}|_{\mathbb{R}^2\setminus C_{j-1}}$ which are assumed to be known. Doing this we see that $R_{\tilde{a}}\tilde{f}$ can be determined given the hypotheses. The problem has now been reduced to determining $\tilde{f}|_{C_{j-1}\cap\{y<0\}}$.

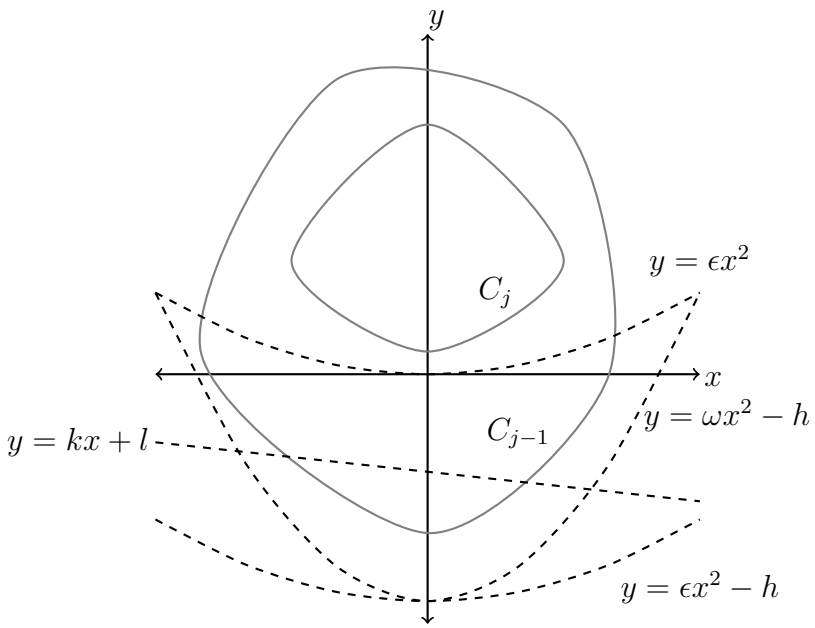


Figure 4.1: This illustrates the setup and some of the notation used in the proof of Lemma 17. Note that we assume $a = c$ in the region $C_{j-1} \setminus C_j$, and $\tilde{a} = c$ in the region above the lowest parabola translated downwards by 1 and outside of C_j. f is assumed to be known outside of C_{j-1}, and then $\tilde{f}$ is supported in the region above the middle parabola.

Our final step is to change variables in order to reduce the problem to one for which we can apply Lemma 16. For this we consider only $R_{\tilde{a}}\tilde{f}(s,\theta)$ for (s,θ) corresponding to lines contained in $\{y < \epsilon x^2\}$. We reparametrise such lines using $y = kx + l$ where the slope k and intercept l replace θ and s respectively. We will also write $x_+(k,l) = \frac{k+\sqrt{k^2+4\epsilon(l+h+1)}}{2\epsilon}$ for the larger value of x at which $\{y = kx + l\}$ intersects $\{y = \epsilon x^2 - h - 1\}$. When we parametrise the lines in this way, the beam transform becomes

$$D\tilde{a}((x, kx+l), k) = \int_0^\infty \tilde{a}(x+s, k(x+s)+l)\sqrt{1+k^2}\,\mathrm{d}s = c(x_+(k,l) - x)\sqrt{1+k^2}$$

and so the AtRT becomes

$$R_{\tilde{a}}\tilde{f}(k,l) = \int_{-\infty}^\infty \tilde{f}(x, kx+l)e^{c(x-x_+(k,l))\sqrt{1+k^2}}\sqrt{1+k^2}\,\mathrm{d}x.$$

We now introduce the new coordinates (z,w) defined by $z = \sqrt{\epsilon}x$ and $w = y - \epsilon x^2 + h$. With this change, the region $\{\epsilon x^2 > y > \epsilon x^2 - h\}$ becomes the strip $\{h > w > 0\}$, and abusing notation slightly by writing $\tilde{f}$ also for the same function in these coordinates

the AtRT becomes

$$R_{\tilde{a}}\tilde{f}(k,l) = \int_{-\infty}^{\infty} \tilde{f}\left(z, -\left(z - \frac{k}{2\sqrt{\epsilon}}\right)^2 + \frac{k^2}{4\epsilon} + l - h\right) e^{c(\sqrt{\epsilon}z - x_+(k,l))\sqrt{1+k^2}} \sqrt{\epsilon(1+k^2)}\, \mathrm{d}z$$

for any (k,l) corresponding to a line contained in $\{y < \epsilon x^2\}$ and passing through the region $\{y > \epsilon x^2 - h\}$. The uniqueness of $\tilde{f}$ in the region $\{y < \epsilon x^2\}$, and therefore also f in the same region, now follows from Lemma 16 since we have that

$$w = e^{c(\sqrt{\epsilon}z - x_+(k,l))\sqrt{1+k^2}} \sqrt{\epsilon(1+k^2)}$$

is an analytic function of $k/(2\sqrt{\epsilon})$ and z provided the imaginary part of k is sufficiently small. □

We are now ready to present the uniqueness result for nicely multi-bang a.

Theorem 6. *Suppose that a is nicely multi-bang (see Definition 14) and $f \in C_c^1(\mathbb{R}^2)$ is non-negative. Also assume that*

1. *for all $\boldsymbol{x} \in \mathcal{K}_{a,1}$ the line tangent to a boundary at x passes through the set $\{f > 0\}$, and*
2. *for all $\boldsymbol{x} \in \mathcal{K}_{a,2}$ there is a line passing through x that also passes through the set $\{f > 0\}$.*

Then a and f are uniquely determined by $R_a f$.

Proof. By Lemma 15 we can recover the sets C_j which make up the nicely multi-bang a. We can determine the $\mathrm{supp}(f)$ from the set of rays (s,θ) which give $R_a f(s,\theta) = 0$. Therefore we can pick some C_0 such that $\mathrm{supp}(f) \Subset C_0$ and note that we know $f|_{\mathbb{R}^2 \setminus C_0} = 0$ and $a|_{\mathbb{R}^2 \setminus C_1} = 0$. Therefore applying Lemma 17 we can determine $f|_{\mathbb{R}^2 \setminus C_1}$. With f known in this region we can use either of the formulas in Lemma 8 or 12 to determine $a_{\mathbb{R}^2 \setminus C_2}$. We can then use Lemma 17 as the inductive step to complete the proof. □

Remark 4. *As stated at the end of the introduction, the constraint that tangent lines and lines passing through corners have non-empty intersection with the set $\{f > 0\}$ requires* $\mathrm{supp}(f) \not\subset \mathrm{supp}(a)$. *Furthermore, by the inversion formula for the AtRT given in Theorem 3 it is not necessary to calculate f in the final region C_n as we know a completely at this stage and can just apply the inversion formula.*

The proof of Theorem 6 makes use of both the nested and convex property of the sets making up a. Unfortunately when we include potentially nonconvex piecewise analytic boundaries with corners we can no longer end up with a problem which can be transformed to one which satisfies the conditions given in Lemma 16. In the case where some $x \in \partial\Omega$ is a boundary point for multiple regions Lemma 16 can not be applied due to the fact that a will have discontinuities and hence not be analytic in the required regions. That being said there are still some partial results which can be useful in the more general multi-bang case and the next section presents these in detail.

4.2 General multi-bang results

The first partial result we include relates to the ability to determine which side of a boundary has the larger multi-bang value.

Lemma 18. *Suppose that $f \in C_c^1(\mathbb{R}^2)$, $f \geq 0$ and a is multi-bang. Suppose that the ray $(s^*, \theta^*) \in \mathcal{K}_{a,1}$ is tangential to a piecewise analytic boundary at a single point $s^*(\theta^*)^\perp + t^*\theta^*$ and that the ray $(s^*, \theta^*) \cap \operatorname{supp}(f) \neq \emptyset$. Let $a = c$ on side $\theta^\perp = (-\theta_2, \theta_1)$ points to at the point of tangency, and $a = c_0$ on the other side.*

Then we can determine whether $c_0 > c$ or $c < c_0$ and hence which side of the boundary has the larger value of a.

Proof. This result follows almost immediately from Lemma 8. Indeed, by Lemma 8 we have

$$\lim_{\omega \to (\omega^*)^{s_1}} |\sin(\omega - \omega^*)|^{\frac{n-1}{n}} \partial_\omega R(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = 2s_2(c_0 - c) \left(\frac{|\boldsymbol{x}^* \cdot \theta^* - t^*|}{n|\kappa|} \right)^{\frac{1}{n}} \int_{-\infty}^{t^*} f(s^*(\theta^*)^\perp + t\theta^*) e^{-Da(s^*(\theta^*)^\perp + t\theta^*, \theta^*)} \, \mathrm{d}t$$

if n is even, and

$$\lim_{\omega \to (\omega^*)^{s_1}} |\sin(\omega - \omega^*)|^{\frac{n-1}{n}} \partial_\omega R(\boldsymbol{x}^* \cdot \theta^\perp, \theta) = s_1(c_0 - c) \left(\frac{|\boldsymbol{x}^* \cdot \theta^* - t^*|}{n|\kappa|} \right)^{\frac{1}{n}} \int_{-\infty}^{t^*} f(s^*(\theta^*)^\perp + t\theta^*) e^{-Da(s^*(\theta^*)^\perp + t\theta^*, \theta^*)} \, \mathrm{d}t$$

if n is odd, where s_1 is the sign of $(t^* - \boldsymbol{x}^* \cdot \theta^*)\kappa$, and s_2 is the sign of κ. By Lemma 5 we know the boundary of a, and therefore we know κ and n. Since f is non-negative

the integrals in these two limits are both non-negative. Hence the only unknown in these limits is $(c_0 - c)$. As everything else is known we can determine the sign of $c_0 - c$ and therefore whether $c_0 > c$ or $c < c_0$, as required. □

Unfortunately Lemma 18 is difficult to generalize to the case where we have multiple tangent points along a single ray. This is because in that case the jump in the derivative is a sum of contributions and, without further knowledge on f, means that the sign of the jump does not uniquely determine the sign of each $(c_i - d_i)$.

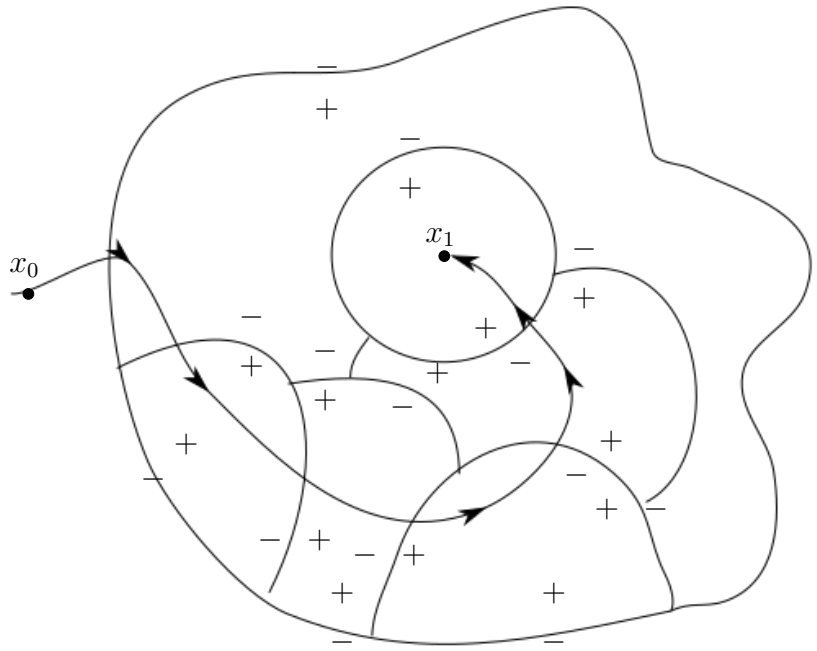
The upside to a result such as Lemma 18 is that we only need to have a single point on the boundary of a region to satisfy the constraints in order to tell the direction in which a increases over a part of the boundary. Even knowing the regions that make up a and in which direction across each boundary is not always enough to uniquely determine a, this is because we can only determine the sign and not the size of $c_0 - c$ in Lemma 18. We can however present a uniqueness proof in a special case. Before presenting this we need the following technical definitions.

Definition 21 (**Multi-bang pair** (a, f) **with directional jumps known).** *Let a be multi-bang as in Definition 6 and that $f \in C_c^1(\mathbb{R}^2)$ is non-negative. Furthermore, suppose that for every boundary between two regions there exists some ray (s^*, θ^*) satisfying the assumptions for Lemma 18, then we call (a, f) a multi-bang with directional jumps known.*

If a is multi-bang with directional jumps known then we can also define the following.

Definition 22 (**Ascending/Descending path).** *Let (a, f) be a multi-bang pair with directional jumps known. Let Γ be a path from a point x_0 to x_1 with $x_0, x_1 \in \mathbb{R}^2$. We say that Γ is ascending if every boundary of a crossed from x_0 to x_1 is in the direction of increasing a. We call a path descending if every boundary of a crossed is in the direction of decreasing a. See Figure 4.2 for an example.*

Figure 4.2: Demonstration of an ascending path from x_0 to x_1. At each boundary the pair + and − shows which side has the larger value of a.



We can now present a uniqueness proof for some multi-bang (a, f) with directional jumps known.

Lemma 19. *Let (a, f) be multi-bang with directional jumps known. Let $\mathcal{A}$, the admissible set of values a takes, contain k elements. Suppose there exists a finite collection of ascending or descending paths $\{\Gamma_n\}$ with each Γ_n passing through k distinct regions and the set of regions passed through is equal to the sets making up a. Then a and therefore f are uniquely determined from $R_a f$*

Proof. By Theorem 5 we can determine the boundaries of a. Then by by the assumption that a is multi-bang with directional jumps known we have all the jump directions of a. If Γ_n is an ascending or descending path and passes through k regions we know that at each boundary crossing a either increases or decreases. Since there are k regions and we have k admissible values, and there has to be a jump so no two adjacent regions can take the same value, every region passed through by Γ_n can be uniquely

assigned an admissible value. By assumption every region of a is passed through by some ascending or descending path and therefore its multi-bang value can by uniquely determined. Hence, we can uniquely determine a and by Theorem 3 f, as required. □

Lemma 19 has strengths and weaknesses when compared to Theorem 6. Not every nicely multi-bang a is also multi-bang with directional jumps known and so this is not a direct strengthening of Theorem 6. However, Lemma 19 does allow for points which are on the boundary for multiple regions. It also gives a uniqueness proof for non-nested sets, as the region where $a = 0$ can be used in a path. However we do also require the admissible set to be correct and to not contain any extra values for the hypothesis to hold.

Lemma 19 is a little restrictive and so we present a weaker but more general result relating to when we can determine the multi-bang value in a single region of a multi-bang a from SPECT. It uses a similar idea as the one used in the proof of Lemma 17.

Lemma 20. *Let a be multi-bang and $f \in C_c^1(\mathbb{R}^2)$. Let $x \in \mathbb{R}^2$ and suppose that there exists a unique point $\bar{z}$ which is the point which achieves the supremum of the set $\{\|y - x\| : y \in \cup_{j=1}^n \partial\Omega_j\}$. If $\bar{z}$ is not a corner then we can determine the value of a in the region $\bar{z}$ lies on the boundary of.*

Proof. Let z be a point satisfying the hypothesis and $\bar{z}$ be the corresponding point on the boundary. The requirement of the existence of a closed ball containing no other point of any of the sets Ω_j requires the boundary $\bar{z}$ lies on to be the outermost layer, .i.e that one side of the boundary $\bar{z}$ lies on is $\mathbb{R}^2 \setminus \cup_{j=1}^n \Omega_j$.

Similar to the proof of 17, we write (x, y) as Cartesian coordinates for points in $\mathbb{R}^2$. By translating and rotating as necessary we assume W.LO.G. that $\Omega = \sum_{j=1}^n \Omega_j$ is contained in the upper half plane $\{y > 0\}$, and show that we can then uniquely determine f restricted to the lower half-plane $\{y < 0\}$. By translating to bring Ω arbitrarily close to $\{y = 0\}$ we can determine f in $\{y < 0\}$. The proof of this follows directly from the proof of Lemma 17, the only difference being in this case that we cannot necessarily rotate to determine f uniquely outside Ω. The assumption that $\bar{z}$ is not a corner is required in order to apply Lemma 16, as this would make a non-analytic in the necessary region. □

Lemma 20 is useful when determining values of a uniquely in the outer regions but cannot be used to determine any interior values in the general case. Since we only need the hypothesis in Lemma 20 to hold for a single z, Lemma 20 does provide at least partial recovery for a from $R_a f$.

Unfortunately the results presented in this section do not provide uniqueness for all multi-bang a. One potential method of solving this problem could come from examining the range condition (2.20) in order to prove that only one possible combination of admissible values produces a with the desired properties in f.

One final remark to make in regards to uniqueness in the general multi-bang case is that is it possible to use Lemma 20 to weaken the hypothesis required in Lemma 19. For example if we can determine all the outer most values of a then the number of regions and ascending/descending paths required for a unique solution can be reduced significantly.

This concludes the joint recovery of a and f from SPECT data part of the thesis. The next chapters focus on the numerical problem of inverting the AtRT from data and the theory behind the numerical convergence of the algorithm.

Chapter 5

Numerical methods

This chapter focuses on the numerical methods which we use to solve the optimization problem

$$\operatorname*{argmin}_{a,f\in BV(\Omega)} \|R_a f - d\|^2_{L^2(\Gamma^+)} + \alpha\mathcal{M}(a) + \gamma_a \mathrm{TV}(a) + \gamma_f \mathrm{TV}(f). \tag{5.1}$$

We begin by outlining how we descritize the problem and give explicit formulas for the multi-bang regularizer $\mathcal{M}$ and total variation TV. In the previous chapters we have been working under the assumption that $R_a f(s,\theta)$ is known for all (s,θ). In practice there can be only finitely many measurements.

Split Ω, the domain of interest, into M^2 square pixels of resolution $\mathrm{d}x$ and order the pixels lexicographically from the top left to the bottom right. We then assume that a and f are piecewise constant over each pixel. Recall that for an oriented line given by (s,θ) we defined the AtRT by

$$R_a f(s,\theta) = \int_{-\infty}^{\infty} f(s\theta^\perp + t\theta) e^{-Da(s\theta^\perp + t\theta,\theta)}\,\mathrm{d}t \tag{5.2}$$

where s is the signed closest approach to the origin and θ is a unit direction tangent to the ray providing orientation. Since a and f are piecewise constant on each pixel, we can discretize (5.2) exactly as follows. Let P be an ordered list of the pixels passed, the order being in which order they are passed, along the oriented ray and let N be the total size of this set. Note this is only well-defined when the ray is oriented. Let K be the ordered set of t values which correspond to an intersection with an edge of a pixel in the grid and let Z be the set of Euclidean distances between adjacent entries

in K. With this notation we have

$$R_a f(s,\theta) = \sum_{i=1}^{N} \int_{K(i)}^{K(i+1)} f(s\theta^{\perp} + t\theta) e^{-Da(s\theta^{\perp}+t\theta,\theta)} \mathrm{d}t. \tag{5.3}$$

Considering both a and f to be piecewise constant over the pixels we can compute the beam transform exactly. For the last pixel the ray passes through, i.e for t $\in$ $[K(N), K(N+1)]$ we have

$$Da(s\theta^{\perp} + t\theta, \theta) = \int_0^{K(N+1)-t} a_{P(N)} \mathrm{d}\eta = a_{P(N)}(K(N+1) - t).$$

Plugging this information into (5.3) for the last pixel the ray passes through, i.e the N-th pixel, gives,

$$\begin{aligned} \int_{K(N)}^{K(N+1)} f(s\theta^{\perp} + t\theta) e^{-Da(s\theta^{\perp}+t\theta,\theta)} \mathrm{d}t &= \int_{K(N)}^{K(N+1)} f_{P(N)} e^{-a_{P(N)}(K(N+1)-t)} \mathrm{d}t \\ &= f_{P(N)} \frac{1 - e^{-Z(N)a_{P(N)}}}{a_{P(N)}}. \end{aligned} \tag{5.4}$$

This form is not appropriate for calculations if a is very small, since then the numerical calculation of the formula in (5.4) would include catastrophic cancellation. It turns out that with a little algebraic manipulation we can rewrite (5.4) as,

$$f_{P(N)} Z(N) e^{-\frac{Z(N)a_{P(N)}}{2}} \mathrm{sinhc}\left(\frac{Z(N)a_{P(N)}}{2\pi}\right), \tag{5.5}$$

where we define, in an analogous way to the sinc function,

$$\mathrm{sinhc}(z) = \begin{cases} \frac{\sinh(z)}{z} & z \neq 0 \\ 1 & z = 0. \end{cases}$$

This form avoids catastrophic cancellation for small a values and can make use of the sinc function which is already built in for programs such as MATLAB. We can use the above information to build up the transform given in (5.3). For the j-th summand in (5.3), $1 \leq j \leq N$,

$$\int_{K(j)}^{K(j+1)} f(s\theta^{\perp} + t\theta) e^{-Da(s\theta^{\perp}+t\theta,\theta)} \mathrm{d}t =$$

$$f_{P(j)} Z(j) e^{-\frac{Z(j)a_{P(i)}}{2}} \mathrm{sinhc}\left(\frac{Z(j)a_{P(j)}}{2}\right) \Pi_{k=j+1}^{N} e^{-Z(k)a_{P(k)}}$$

where we define a new vector S by $S(N) = 1$ and $S(i-1) = S(i)e^{-Z(i)a_{P(i)}}$. We can then use the above formula to rewrite (5.3) as

$$R_a f(s,\theta) = \sum_{i=1}^{N} f_{P(i)} Z(i) e^{-\frac{Z(i)a_{P(i)}}{2}} \mathrm{sinhc}\left(\frac{Z(i)a_{P(i)}}{2}\right) S(i). \tag{5.6}$$

The formula given in (5.6) allows us to rewrite the AtRT as a vector equation involving a and f. If we are given a data vector d for a set $\mathcal{I}$ of oriented rays $(s_i, \theta_i)_{i\in\mathcal{I}}$, then we can combine the vector equations in (5.6) into the matrix equation

$$R[a]f = d. \tag{5.7}$$

The discretised problem of interest is then to determine both a and f from d given by (5.7) where a is multi-bang with the admissible set $\mathcal{A} = \{a_0, a_1, ..., a_n\}$ known (note that for notational convenience we have reindexed the admissible values relative to Definition 6). We next look at the regularizers for the discrete case.

Recall that the non-convex multi-bang regularizer is given by

$$\mathcal{M}(a) := \int_\Omega m(a(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} \tag{5.8}$$

where

$$m(t) = \begin{cases} (a_{i+1} - t)(t - a_i), & \text{if } t \in [a_i, a_{i+1}] \text{ for some } i \\ \infty, & \text{otherwise.} \end{cases} \tag{5.9}$$

When a is piecewise constant over pixels we can evaluate (5.8) exactly. In particular we find

$$\mathcal{M}(a) = \mathrm{d}x^2 \sum_{i=1}^{M^2} m(a(p_i)) \tag{5.10}$$

where $\mathrm{d}x$ is the resolution of the pixel and $a(p_i)$ is the attenuation of a in the p_ith pixel. In practice we omit the term $\mathrm{d}x^2$ and absorb it into the regularization parameter α. We note that each of the summands in (5.10) only depend on the value of a at a single pixel. This has the advantage that formulas involving the multi-bang regularizer can typically be separated and we need only deal with m instead of $\mathcal{M}$. Unfortunately this separability property means that multi-bang regularization gives no spatial regularity and this is the main reason why we include total variation which is well known to produce piecewise constant images with minimal perimeter. There are two common types of total variation; anisotropic TV with multi-bang regularization is studied in [18, 54], anisotropic TV is separable (after a change of variables linked by a matrix equation) and so again is numerically easier to manipulate. However numerical experiments in [54] show that anisotropic TV, which for an image $x \in \mathbb{R}^n \times \mathbb{R}^n$ is given by

$$\begin{aligned}\mathrm{TV}_{\mathrm{an}}(x) = &\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} |x_{i+1,j} - x_{i,j}| + |x_{i,j+1} - x_{i,j}| + \sum_{i=1}^{n-1} |x_{i+1,n} - x_{i,n}| \\ &+ \sum_{j=1}^{n-1} |x_{n,j+1} - x_{n,j}|,\end{aligned} \tag{5.11}$$

struggles to distinguish circular and square objects, see Remark 5. Instead we use isotropic TV, as in [37], which is given, for an image x, by

$$\begin{aligned}\mathrm{TV}(x) = &\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2} + \sum_{i=1}^{n-1} |x_{i+1,n} - x_{i,n}| \\ &+ \sum_{j=1}^{n-1} |x_{n,j+1} - x_{n,j}|.\end{aligned} \tag{5.12}$$

Using (5.12) as the total variation regularizer as in [54] allows recovery of circular objects but at the cost of being more computationally challenging, particularly when it comes to writing (5.1) in an appropriate form for performing an Alternating Direction Method of multipliers update on a.

Remark 5. *For a binary image, i.e each pixel is either 0 or 1, and with square pixels the anisotropic TV is simply the perimeter of the shape as proven in general in [46]. This makes computation simple but if we take the image x to be a square with perimeter $4n$ and remove the bottom right pixel then the perimeter of this shape remains the same. This argument holds for any pixel which lies on a corner. In particular an approximation of a circle on any number of pixels has the same anisotropic TV given by* (5.11) *as a square of the same diameter. Therefore we cannot distinguish shapes based purely on their anisotropic TV.*

After discretising, the optimization problem we aim to solve is

$$\underset{a,f\in\mathbb{R}^{M^2}}{\operatorname{argmin}} \mathcal{R}(a,f) := \|R[a]f - d\|_2^2 + \alpha\mathcal{M}(a) + \gamma_a \mathrm{TV}(a) + \gamma_f \mathrm{TV}(f). \tag{5.13}$$

The optimization problem (5.13) is non-convex for two reasons. Firstly it is clear that $\mathcal{M}$ is non-convex as m is non-convex. The second and perhaps more interesting reason is related to the data fidelity term

$$\|R[a]f - d\|^2.$$

In particular this term is only guaranteed to be convex, with f fixed, when each of entries of the vector $R[a]f - d \geq 0$. This is because we can consider the data fidelity term as a sum of squares of the individual vector entries, $(R[a]f - d)_j$ is a monotone decreasing function (that is increasing any entry in a reduces or maintains the value of $(R[a]f - d)_j$). Then by using results from [54, 37] we have that the composition of two non-negative convex monotone functions is a convex monotone function. This is only satisfied when $(R[a]f - d)_j \geq 0$. In order to solve this non-convex optimization problem we use the following alternating algorithm as in [1]

$$\begin{aligned} a^{k+1} &\in \operatorname*{argmin}_a \mathcal{R}(a, f^k) + \frac{1}{2\xi^k}\|a - a^k\|^2, \\ f^{k+1} &\in \operatorname*{argmin}_f \mathcal{R}(a^{k+1}, f) + \frac{1}{2\xi^k}\|f - f^k\|^2, \end{aligned} \tag{5.14}$$

for sufficiently small $\{\xi_k\}_{k=1}^{\infty}$. In particular [1] shows that for any problem which can be written in the form

$$\operatorname*{argmin}_{x,y} f(x) + Q(x, y) + g(y)$$

with f and g proper lower-semicontinuous, Q a C^1 function with Lipschitz continuous partial derivatives on bounded subsets of $\mathbb{R}^{n\times m}$, the iterates converge to a critical point of the system, provided that the sequence ξ^k is sufficiently small and f and g are so called Kurdyka-Łojasiewicz(KL)[32] functions.

Definition 23 (Kurdyka-Łojasiewicz(KL) functions). *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper lower-semicontinuous function. For $-\infty < c_1 < c_2 < \infty$, define*

$$[c_1 < f < c_2] = \{x \in \mathbb{R}^n : c_1 < f(x) < c_2\}.$$

The function f has the KL property at $x^* \in \mathrm{dom}\partial f$, where $\mathrm{dom}\partial f$ is the domain of ∂_f, if there exist $c \in (0, \infty]$, a neighbourhood U of x^*, and a continuous concave function $\phi : [0, c) \to \mathbb{R}^+$ such that

1. *$\phi(0) = 0$;*

2. *ϕ is C^1 on $(0, c)$ and continuous at 0;*

3. $\phi'(s) > 0$, $\forall s \in (0, c)$;

4. for all x *in* $U \cap [f(x^*) < f < f(x^*)+c]$, *the Kurdyka-Łojasiewicz inequality holds*

$$\phi'(f(x) - f(x^*))\mathrm{d}(0, \partial f(x)) \geq 1.$$

where $d(0, \partial f(x)) = \inf_{y \in \partial_{f(x)}} \|0 - y\|$. *Let* Φ_c *denote the set of functions which satisfy conditions 1-3. If every* $\phi \in \Phi_c$ *also satisfies the 4th condition then we call* f *a Kurdyka-Łojasiewicz(KL) function.*

The general idea behind KL functions is that there is a change of variables such that resulting range of the function has a kink at the minimum and is steep close to this minimum point.

Before showing that these conditions for convergence of the alternating method are satisfied in (5.13), we first define what we mean by a Lipschitz continuous function which is a special case of Definition 16 with $\alpha = 1$.

Definition 24. (Lipschitz continuous) *Let* g *be a function* $g : \mathbb{R}^n \to \mathbb{R}^m$, *where* $n, m \in \mathbb{N}$, *we say that* g *is Lipschitz continuous if there exists some* $L > 0$ *such that for all* $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$

$$\|g(\boldsymbol{x}_1) - g(\boldsymbol{x}_2)\| \leq L\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|.$$

In the case of $\mathcal{R}(a, f)$, the only mixed term is the data fidelity term

$$\|R[a]f - d\|_2^2.$$

This term is differentiable in both a and f, this can be shown explicitly using (5.6) as follows. Since $R[a]f$ is linear in f the derivative is simply

$$\nabla_f \|R[a]f - d\|_2^2 = 2R[a]^T(R[a]f - d),$$

which is Lipschitz continuous for bounded subsets of $\mathbb{R}^{N \times N}$. Unfortunately, as $\mathcal{F}$ is non-linear in a, finding $\nabla_a \mathcal{F}(a, f)$ is a significantly more computationally expensive process. We tackle this gradient by looking at one ray at a time. In order to take this approach, we first note that we have

$$da := \nabla_a \|R[a]f - d\|_2^2 = 2\left\langle \frac{\partial R[a]}{\partial a_L} f, R[a]f - d \right\rangle.$$

We set $h = R[a]f - d$, so that the value of h corresponding to the j-th ray is h_j. Then as described earlier, let P be the ordered set of pixels passed through by the j-th ray with the orientation θ, and let Z be the set of Euclidean lengths of the intersections with each of these pixels. Let N be the total number of pixels that the j-th ray passes through. Then we have by (5.6),

$$(R[a]f)_j = \sum_{i=1}^{N} f_{P(i)} Z(i) e^{-\frac{Z(i)a_{P(i)}}{2}} \operatorname{sinhc}\left(\frac{Z(i)a_{P(i)}}{2}\right) S(i)$$

where we define $S(i)$ exactly as before. In order to simplify this expression we set

$$g_i = e^{-\frac{Z(i)a_{P(i)}}{2}} \operatorname{sinhc}\left(\frac{Z(i)a_{P(i)}}{2}\right).$$

Now from the definition of S we have that

$$S(i) = e^{-\sum_{k=i+1}^{N} Z(k)a_{P(k)}}.$$

Then the derivative with respect to a is given by,

$$\begin{aligned}\frac{\partial (R[a]f)_j}{\partial a_{P(l)}} = \frac{f_{P(l)} Z(l)^2}{2}\left(-g_l + S(l) e^{-\frac{Z(l)a_{P(l)}}{2}} \operatorname{sinhc}'\left(\frac{Z(l)a_{p(l)}}{2}\right)\right) \\ - \sum_{i=1}^{l-1} f_{P(j)} Z(j) Z(l) g_j.\end{aligned} \tag{5.15}$$

Evaluating sinhc' causes an issue when a is very small, once again due to catastrophic cancellation, and so we make use of the Taylor series for sinhc in order to rewrite this derivative as a function which remains small as a is small. It can be shown that provided $a(l) < 1/Z(l)^5$ then we can approximate sinhc' by

$$\operatorname{sinhc}'\left(\frac{Z(l)a_{P(l)}}{2}\right) \simeq \left(\frac{Z(l)a_{P(l)}}{6} + \frac{(Z(l)a_{P(l)})^3}{240}\right).$$

Then we find that the individual contribution to $\mathrm{d}a_{P(l)}$ from the j-th ray is given by

$$da^j_{P(l)} := 2\frac{\partial (R[a]f)_j}{\partial a_{P(l)}} h_j.$$

This is simply the formula in (5.15) multiplied by h_j. If we then set $\mathrm{d}a^j_p = 0$ for all $p \notin P$, then finally we have $\nabla_a \mathcal{F}(a, f)$ by summing over all lines (s, θ) in our data set, i.e

$$\nabla_a \|R[a]f - d\|_2^2 = \sum_r da^r. \tag{5.16}$$

The formula given in (5.15) is again differentiable for each a_j and the derivative is continuous and therefore (5.16) is also continuously differentiable and so Lipschitz continuous on any bounded subset of $\mathbb{R}^{N\times N}$. Hence, the conditions on Q are satisfied. Then the terms involving just a are

$$F_1(a) := \alpha\mathcal{M}(a) + \gamma_a \mathrm{TV}(a)$$

and the terms involving just f are

$$F_2(f) := \gamma_f \mathrm{TV}(f).$$

In can be shown that F_1 and F_2 are subanalytic [41], the details are very technical and beyond the scope of this research. We can apply results directly from [32] to show that both F_1 and F_2 are KL. Therefore in our setting the alternating method given in (5.14) converges, provided ξ^k is sufficiently small.

Remark 6. *Although theoretical convergence of* (5.14) *has been proved in [1] for sufficiently small ξ_k, we have had success in reconstruction when the terms $\|a-a^k\|^2$ and $\|f-f^k\|^2$ are omitted which corresponds to $\xi_k \to \infty$. We have not yet been able to prove this result theoretically but for the examples given in this report, and indeed all the examples we have currently tested, the algorithm converges typically in fewer iterations for the same range of initial guesses of parameters when $\xi_k \to \infty$ as when it is taken to be small.*

We first turn our attention to the more computationally challenging part of the algorithm which involves updating a.

5.1 Updating a

After removing terms which only depend on f, we can find a^{k+1} via the first order optimality condition

$$a^{k+1} \in \operatorname*{argmin}_a \|R[a]f^k - d\|^2 + \alpha\mathcal{M}(a) + \gamma_a \mathrm{TV}(a) + \frac{1}{2\xi^k}\|a-a^k\|^2. \tag{5.17}$$

This sort of problem can be solved by splitting the objective function into two parts each depending on separate parameters which are linked together by a matrix equation.

We split the objective function, and a, into two parts

$$f(x) := \|R[x]f^k - d\|^2 + \alpha\mathcal{M}(x) + \frac{1}{2\xi^k}\|x - a^k\|^2 \tag{5.18}$$

where x is related to the data fidelity and multi-bang part of a and

$$g(y) := \gamma_a \sum_{i=1}^{M^2-1} \|\boldsymbol{y_i}\|_2 \tag{5.19}$$

where $\boldsymbol{y_i}$ is a discrete derivative of a in the i-th pixel. We then concatenate the $\boldsymbol{y_i}$ to obtain y which is an approximation of the gradient of a. Precisely how we do this split and obtaining linear constraints on the separation will be examined at the end of section 5.1. Once the objective function is split, optimizing can then be done via the Alternating Direction Method of multipliers algorithm [10] which we outline below.

Suppose that an optimization problem can be written in the form

$$\begin{aligned} &\underset{x,y}{\operatorname{argmin}}\, f(x) + g(y) \\ &\text{s.t } Ax + y = 0 \end{aligned} \tag{5.20}$$

where $Ax + y = 0$ is a linear constraint and f and g are proper lower-semicontinuous and convex functions. For $\beta > 0$ we construct the augmented Lagrangian

$$\mathrm{L}_\beta(x, y, \lambda) := f(x) + g(y) - \lambda^T(Ax + y) + \frac{\beta}{2}\|Ax + y\|_2^2$$

where λ is a vector of Lagrange multipliers; here augmented is used to refer to the additional term $\frac{\beta}{2}\|Ax + y\|^2$. The ADMM algorithm is then given by the following update scheme

$$\begin{aligned} x^{k+1} &= \underset{x}{\operatorname{argmin}}\, \mathrm{L}_\beta(x, y^k, \lambda^k) \\ y^{k+1} &= \underset{y}{\operatorname{argmin}}\, \mathrm{L}_\beta(x^{k+1}, y, \lambda^k) \\ \lambda^{k+1} &= \lambda^k - \beta(Ax^{k+1} + y^{k+1}), \end{aligned}$$

which terminates when the primal residual $r^k := \|Ax + y\|^2$ and an approximate dual residual $s^k := \beta D^T(y^{k+1} - y^k)$ originally obtained and used in [10] are sufficiently small. It is worth noting that in the general case the x^{k+1} and y^{k+1} can belong to a potential set of minimisers, as so we would need to replace $=$ with $\in$ for the first two lines of the ADMM algorithm. With an additional requirement on A given the

list below, since g is assumed convex and the weak convexity of f can be countered by the convexity of $\frac{\beta}{2}\|Ax+y\|^2$, provided β is suitably large, we do have equality for the x^{k+1} and y^{k+1} with our hypothesis. This algorithm is a specific example of the more general class of saddle-point algorithms studied in [10], as the x and y updates are descent directions and the λ update is an ascent update. If both f and g are convex [10] gives a series of mild conditions which will guarantee convergence of the sequence $\{x^{k+1}, y^{k+1}, \lambda^{k+1}\}_{k\geq 1}$. Recently [32] has extended convergence of the iterates $\{x^{k+1}, y^{k+1}, \lambda^{k+1}\}_{k\geq 1}$ to the case where f and g are non-convex, with the constraints that the derivative of g is Lipschitz continuous and both f and g are KL functions. Before continuing we give an important result relating to convergence of the ADMM algorithm in the non-convex case based on [32]. A proof is also included because it introduces an important quantity we discuss later in the numerical section.

Suppose that we have the following conditions on (5.20)

- f is weakly-convex and lower-semicontinuous as in Definitions 10 and 7.
- g is convex, with Lipschitz continuous gradient as in Definition 24, i,e there exists some $L > 0$ such that

$$\|\nabla g(y) - \nabla g(y')\| \leq L\|y - y'\|.$$

- $A^TA \geq \mu\mathcal{I}$, for some $\mu > 0$. Here this means the matrix $A^TA - \mu\mathcal{I}$ is positive semi-definite and guarantees that the augmented term is strongly convex.
- The dual update β satisfies both

$$\left(\frac{\beta}{2} - \frac{L^2}{2}\right) > 0 \tag{5.21}$$

and

$$\left(\frac{\beta\mu}{2} - \frac{\rho}{2}\right) > 0 \tag{5.22}$$

where ρ is the associated weak convexity of f.

With these conditions we have the following result given in [32].

Lemma 21. *Let f, g and A satisfy the conditions given above. Then there exists $\delta > 0$ such that for β satisfying* (5.21) *and* (5.22)*, the ADMM iterates satisfy*

$$\mathrm{L}_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) \leq \mathrm{L}_\beta(x^k, y^k, \lambda^k) - \delta\left(\|x^{k+1} - x^k\|^2 + \|y^{k+1} - y^k\|^2\right).$$

Proof. To prove this result we follow the proof given in [32]. We aim to show the result in the lemma by rewriting

$$\begin{aligned}\mathrm{L}_\beta(x^{k+1},y^{k+1},\lambda^{k+1})-\mathrm{L}_\beta(x^k,y^k,\lambda^k)=&\big(\mathrm{L}_\beta(x^{k+1},y^{k+1},\lambda^{k+1})-\mathrm{L}_\beta(x^{k+1},y^{k+1},\lambda^k)\big)\\&+\big(\mathrm{L}_\beta(x^{k+1},y^{k+1},\lambda^k)-\mathrm{L}_\beta(x^k,y^{k+1},\lambda^k)\big)\\&+\big(\mathrm{L}_\beta(x^k,y^{k+1},\lambda^k)-\mathrm{L}_\beta(x^k,y^k,\lambda^k)\big).\end{aligned}\tag{5.23}$$

From the definition of the augmented Lagrangian function $\mathrm{L}_\beta(\cdot)$, it follows that

$$\begin{aligned}\mathrm{L}_\beta(x^{k+1},y^{k+1},\lambda^{k+1})&=\mathrm{L}_\beta(x^{k+1},y^{k+1},\lambda^k)+\langle\lambda^k-\lambda^{k+1},Ax^{k+1}+y^{k+1}\rangle\\&=\mathrm{L}_\beta(x^{k+1},y^{k+1},\lambda^k)+\frac{1}{\beta}\|\lambda^k-\lambda^{k+1}\|^2,\end{aligned}$$

where the last line follows from $\lambda^{k+1}=\lambda^k-\beta(Ax^{k+1}+y^{k+1})$. We have already found an equality for the first of the three terms in (5.23) and we proceed by bounding the other two parts.

Consider

$$\begin{aligned}&\mathrm{L}_\beta(x^k,y^{k+1},\lambda^k)-\mathrm{L}_\beta(x^{k+1},y^{k+1},\lambda^k)\\&\quad=f(x^k)+g(y^{k+1})-\langle\lambda^k,Ax^k+y^{k+1}\rangle+\frac{\beta}{2}\|Ax^k+y^{k+1}\|^2\\&\qquad-\left\{f(x^{k+1})+g(y^{k+1})-\langle\lambda^k,Ax^{k+1}+y^{k+1}\rangle+\frac{\beta}{2}\|Ax^{k+1}+y^{k+1}\|^2\right\}\\&\quad=f(x^k)-f(x^{k+1})+\langle\lambda^k,A(x^{k+1}-x^k)\rangle+\frac{\beta}{2}\|Ax^k+y^{k+1}\|^2\\&\qquad-\frac{\beta}{2}\|Ax^{k+1}+y^{k+1}\|^2.\end{aligned}\tag{5.24}$$

We now make use of the weakly-convex property of f. Let $\rho>0$ be chosen so that

$$f(x)+\frac{\rho}{2}\|x-x^{k+1}\|^2$$

is convex. Then by the first order optimality condition

$$A^T\lambda^{k+1}+\beta A^T(y^{k+1}-y^k)+\rho(x-x^{k+1})\Big|_{x=x^{k+1}}\in\partial\left(f(x)+\frac{\rho}{2}\|x-x^{k+1}\|^2\right)\Big|_{x=x^{k+1}}$$

we have

$$\begin{aligned}&f(x^k)-f(x^{k+1})+\frac{\rho}{2}\|x^k-x^{k+1}\|^2\geq\langle A^T(\lambda^{k+1}+\beta(y^{k+1}-y^k)),x^k-x^{k+1}\rangle\\\implies&f(x^k)-f(x^{k+1})\geq\langle A^T(\lambda^{k+1}+\beta(y^{k+1}-y^k)),x^k-x^{k+1}\rangle-\frac{\rho}{2}\|x^k-x^{k+1}\|^2.\end{aligned}\tag{5.25}$$

Combining (5.24) and (5.25) gives

$$\begin{aligned}\mathrm{L}_\beta(x^k, y^{k+1}, \lambda^k) - \mathrm{L}_\beta(x^{k+1}, y^{k+1}, \lambda^k)\\ \geq \langle(\lambda^{k+1} - \lambda^k) + \beta(y^{k+1} - y^k), A(x^k - x^{k+1})\rangle - \frac{\rho}{2}\|x^k - x^{k+1}\|^2\\ + \frac{\beta}{2}\|Ax^k + y^{k+1}\|^2 - \frac{\beta}{2}\|Ax^{k+1} + y^{k+1}\|^2.\end{aligned} \tag{5.26}$$

Similarly, we have

$$\begin{aligned}\mathrm{L}_\beta(x^k, y^k, \lambda^k) - \mathrm{L}_\beta(x^k, y^{k+1}, \lambda^k)\\ = g(y^k) - g(y^{k+1}) + \langle\lambda^k, y^{k+1} - y^k\rangle\\ + \frac{\beta}{2}\|Ax^k + y^k\|^2 - \frac{\beta}{2}\|Ax^k + y^{k+1}\|^2.\end{aligned} \tag{5.27}$$

As g is convex and by the optimality condition $\lambda^{k+1} = \nabla g(y^{k+1})$ we have

$$g(y^k) - g(y^{k+1}) \geq \langle\lambda^{k+1}, y^k - y^{k+1}\rangle. \tag{5.28}$$

Combining (5.27) and (5.28) we obtain

$$\begin{aligned}\mathrm{L}_\beta(x^k, y^k, \lambda^k) - \mathrm{L}_\beta(x^k, y^{k+1}, \lambda^k)\\ \geq \langle\lambda^{k+1} - \lambda^k, y^k - y^{k+1}\rangle + \frac{\beta}{2}\|Ax^k + y^k\|^2 - \frac{\beta}{2}\|Ax^k + y^{k+1}\|^2.\end{aligned} \tag{5.29}$$

If we combine (5.29) and (5.26) we have

$$\begin{aligned}\mathrm{L}_\beta(x^k, y^k, \lambda^k) - \mathrm{L}_\beta(x^{k+1}, y^{k+1}, \lambda^k)\\ \geq \langle\lambda^{k+1} - \lambda^k, y^k - y^{k+1}\rangle + \frac{\beta}{2}\|Ax^k + y^k\|^2 - \frac{\beta}{2}\|Ax^{k+1} + y^{k+1}\|^2\\ \langle(\lambda^{k+1} - \lambda^k) + \beta(y^{k+1} - y^k), A(x^k - x^{k+1})\rangle - \frac{\rho}{2}\|x^k - x^{k+1}\|^2.\end{aligned} \tag{5.30}$$

Recall

$$\begin{aligned}\lambda^{k+1} &= \lambda^k - \beta(Ax^{k+1} + y^{k+1})\\ \implies Ax^{k+1} + y^{k+1} &= \frac{1}{\beta}(\lambda^k - \lambda^{k+1})\\ \implies Ax^k + y^k &= \frac{1}{\beta}(\lambda^k - \lambda^{k+1}) + (y^k - y^{k+1}) + A(x^k - x^{k+1}),\end{aligned}$$

combining this with the fact that

$$2\langle x, y\rangle + \|x - y\|^2 = \|x\|^2 + \|y\|^2$$

gives

$$\begin{aligned}
&\beta\left\langle\frac{1}{\beta}(\lambda^{k+1}-\lambda^k)+(y^{k+1}-y^k),A(x^k-x^{k+1})\right\rangle+\frac{\beta}{2}\|Ax^k+y^k\|^2=\\
&\beta\left\langle\frac{1}{\beta}(\lambda^{k+1}-\lambda^k)+(y^{k+1}-y^k),A(x^k-x^{k+1})\right\rangle\\
&+\frac{\beta}{2}\left\|\frac{1}{\beta}(\lambda^k-\lambda^{k+1})+(y^k-y^{k+1})+A(x^k-x^{k+1})\right\|^2\\
=&\frac{\beta}{2}\left\|A(x^k-x^{k+1})\right\|^2+\frac{\beta}{2}\left\|\frac{1}{\beta}(\lambda^{k+1}-\lambda^k)+(y^{k+1}-y^k)\right\|.
\end{aligned}\tag{5.31}$$

Repeating this process again we find

$$\begin{aligned}
\langle\lambda^{k+1}-\lambda^k,y^k-y^{k+1}\rangle+\frac{\beta}{2}\left\|\frac{1}{\beta}(\lambda^{k+1}-\lambda^k)+(y^{k+1}-y^k)\right\|^2\\
=\frac{\beta}{2}\left\|y^k-y^{k+1}\right\|^2+\frac{1}{2\beta}\left\|\lambda^{k+1}-\lambda^k\right\|^2.
\end{aligned}\tag{5.32}$$

Combining (5.31),(5.32) and (5.30) we have

$$\begin{aligned}
&\mathrm{L}_\beta(x^k,y^k,\lambda^k)-\mathrm{L}_\beta(x^{k+1},y^{k+1},\lambda^k)\\
&\geq\frac{\beta}{2}\left\|y^k-y^{k+1}\right\|^2+\frac{1}{2\beta}\left\|\lambda^{k+1}-\lambda^k\right\|^2+\left\|A(x^k-x^{k+1})\right\|^2\\
&-\frac{\beta}{2}\left\|Ax^{k+1}+y^{k+1}\right\|^2-\frac{\rho}{2}\|x^k-x^{k+1}\|^2.
\end{aligned}\tag{5.33}$$

Again, from

$$\lambda^{k+1}=\lambda^k-\beta(Ax^{k+1}+y^{k+1}),$$

we find

$$\begin{aligned}
\lambda^{k+1}-\lambda^k&=-\beta(Ax^{k+1}+y^{k+1})\\
\implies\left\|\lambda^{k+1}-\lambda^k\right\|^2&=\beta^2\left\|Ax^{k+1}+y^{k+1}\right\|^2\\
\implies\frac{1}{2\beta}\left\|\lambda^{k+1}-\lambda^k\right\|^2&=\frac{\beta}{2}\left\|Ax^{k+1}+y^{k+1}\right\|^2.
\end{aligned}$$

Therefore we can write (5.33) as

$$\begin{aligned}
&\mathrm{L}_\beta(x^k,y^k,\lambda^k)-\mathrm{L}_\beta(x^{k+1},y^{k+1},\lambda^k)\\
&\geq\frac{\beta}{2}\left\|y^k-y^{k+1}\right\|^2+\left\|A(x^k-x^{k+1})\right\|^2-\frac{\rho}{2}\|x^k-x^{k+1}\|^2.
\end{aligned}\tag{5.34}$$

Now since $A^TA\geq\mu\mathcal{I}$ we have

$$\|A(x^k-x^{k+1})\|^2\geq\mu\left\|x^k-x^{k+1}\right\|^2,$$

and so (5.34) can be written as

$$\begin{aligned} &\mathrm{L}_\beta(x^k, y^k, \lambda^k) - \mathrm{L}_\beta(x^{k+1}, y^{k+1}, \lambda^k) \\ &\qquad \geq \frac{\beta}{2}\left\|y^k - y^{k+1}\right\|^2 + \left(\frac{\beta\mu}{2} - \frac{\rho}{2}\right)\|x^k - x^{k+1}\|^2. \end{aligned} \tag{5.35}$$

Therefore we have

$$\begin{aligned} &\mathrm{L}_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) - \mathrm{L}_\beta(x^k, y^k, \lambda^k) \\ &\qquad \leq -\frac{\beta}{2}\left\|y^k - y^{k+1}\right\|^2 - \left(\frac{\beta\mu}{2} - \frac{\rho}{2}\right)\|x^k - x^{k+1}\|^2 \\ &\qquad\quad + \frac{1}{\beta}\|\lambda^k - \lambda^{k+1}\|^2. \end{aligned} \tag{5.36}$$

Finally using the fact the gradient of g is Lipschitz continuous, as well as the first order optimality condition for y, we have

$$\frac{1}{\beta}\|\lambda^{k+1} - \lambda^k\|^2 \leq \frac{L^2}{\beta}\|y^{k+1} - y^k\|^2.$$

Combining all of this then gives

$$\begin{aligned} &\mathrm{L}_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) - \mathrm{L}_\beta(x^k, y^k, \lambda^k) \\ &\qquad \leq -\left(\frac{\beta}{2} - \frac{L^2}{\beta}\right)\left\|y^k - y^{k+1}\right\|^2 - \left(\frac{\beta\mu}{2} - \frac{\rho}{2}\right)\|x^k - x^{k+1}\|^2. \end{aligned} \tag{5.37}$$

Since we assume that both $\left(\frac{\beta}{2} - \frac{L^2}{\beta}\right)$ and $\left(\frac{\beta\mu}{2} - \frac{\rho}{2}\right)$ are positive, taking $\delta = \min\left\{\left(\frac{\beta}{2} - \frac{L^2}{\beta}\right), \left(\frac{\beta\mu}{2} - \frac{\rho}{2}\right)\right\}$ gives

$$\mathrm{L}_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) \leq \mathrm{L}_\beta(x^k, y^k, \lambda^k) - \delta\left(\|x^{k+1} - x^k\|^2 + \|y^{k+1} - y^k\|^2\right),$$

as required. □

The proof that the iterates converge then follows using an identical method of analyzing KL inequalities as in [32, Lemmas 3.1-3.4 and Theorem 3.1].

Remark 7. *We remark that the isotropic TV regularizer given in* (5.12) *is non-smooth and therefore does not have Lipschitz continuous gradient. This means that the proof given in lemma 21 no longer holds. One potential method of dealing with this could be to consider the Bregman distance on g rather than the two norm as in [25] and this is potential further avenue of research. At this time we are unable to prove theoretical convergence of the ADMM algorithm when the Lipschitz conitinuity condition on g*

is removed, the condition on g being convex is unchanged even with non-smooth TV. However most of the proof of Lemma 21 is unchanged when we consider the non-smooth version, in fact we can follow identical logic all the way to (5.36)*. We do examine a parameter arising from* (5.36) *which would guarantee convergence at the end of this section. Numerically using the non-smooth TV gives convergence for a vast array of examples with large variations in initial guesses. The smoothed TV given by* (5.43) *is continuously differentiable and therefore Lipschitz continuous for any bounded subset of* $\mathbb{R}^N$ *and so satisfies the assumptions of Lemma 21.*

With this result established we return to solving the optimization problem (5.13), first we obtain the matrix equation relating x the data fidelity and multi-bang part of a and the total variation part y.

Recall that the entry in the i-th position of x corresponds to the value of x in the i-th pixel and we order lexicographically from top left to bottom right. We can define a series of matrices $D_i \in \mathbb{R}^{2,N^2}$ which have two rows. The first row contains a 1 in the position corresponding to x_i and a -1 in the position corresponding to the pixel directly to the right of x_i. If the pixel x_i is on the right hand edge of the grid then the first row is simply all zeros. Similarly the second row contains a 1 in the position corresponding to x_i and a -1 corresponding to the pixel directly below of x_i. Again if the pixel is on the bottom of the grid we set the second row to be all zeros. Combining this together, for a given vector $x \in \mathbb{R}^{N^2}$ we see that

$$D_i x = \begin{cases} \begin{pmatrix} x_i - x_{i+1} \\ x_i - x_{i+M} \end{pmatrix} & \text{if } 1 \leq i \leq M^2 - M \ \& \quad \mod(i, N) \neq 0 \\ \begin{pmatrix} 0 \\ x_i - x_{i+M} \end{pmatrix} & \text{if } 1 \leq i \leq M^2 - M \ \& \quad \mod(i, M) = 0 \\ \begin{pmatrix} x_i - x_{i+1} \\ 0 \end{pmatrix} & \text{if } M^2 - M + 1 \leq i \leq M^2 - 1. \end{cases} \tag{5.38}$$

In particular this allows us to rewrite (5.12) as

$$\mathrm{TV}(x) = \sum_{i}^{M^2-1} \|D_i x\|_2. \tag{5.39}$$

If we stack all of these D_i matrices on top of each other and compare with the formula given in (5.12) we obtain the linear constraint

$$y = Dx.$$

We further split y into a series of 2 by 1 column vectors which are linked to x by the matrix equations

$$D_i x = \boldsymbol{y_i}.$$

Remark 8. *As it is written D is not positive semi-definite, its null space is the set of all constant vectors, we can eliminate this problem by fixing the top left most pixel, i,e the first entry in the vector x, to be 0. With this modification we then satisfy the requirements of Lemma 21.*

Then we can rewrite the variational problem for a as

$$\begin{aligned} &\underset{x}{\operatorname{argmin}} \|R[x]f - d\|_2^2 + \alpha\mathcal{M}(x) + \gamma_a \sum_{i=1}^{M^2-1} \|\boldsymbol{y_i}\|_2 + \frac{1}{2\xi^k}\|x - a^k\|^2 \\ &\text{s.t } \boldsymbol{y_i} = D_i x, \text{ for } i = 1, 2, ..., M^2 - 1. \end{aligned} \tag{5.40}$$

The augmented Lagrangian corresponding to (5.40) is given by

$$\begin{aligned} \mathrm{L}_\beta(x, y, \lambda) = &\sum_i \left(\gamma_a \|\boldsymbol{y}_i\|_2 - \boldsymbol{\lambda}_i^T(\boldsymbol{y}_i - D_i x) + \frac{\beta}{2}\|\boldsymbol{y}_i - D_i x\|^2 \right) \\ &+ \|R[x]f - d\|_2^2 + \alpha\mathcal{M}(x) + \frac{1}{2\xi^k}\|x - a^k\|^2 \end{aligned}$$

where each $\boldsymbol{\lambda_i} \in \mathbb{R}^2$, and $\lambda \in \mathbb{R}^{2N(N-1)}$ is a reordering of $\boldsymbol{\lambda}_i$, similar to y. With this we are now ready to start outlining the numerical method of updating a. We begin by first considering the update of x, the data fidelity and multi-bang part of a.

5.1.1 Updating x

If we only consider the terms involving x in (5.13), we have

$$\begin{aligned} \underset{x}{\operatorname{argmin}} &||R[x]f - d||_2^2 - \lambda^T(y - Dx) + \frac{\beta}{2}||Dx - y||_2^2 + \frac{1}{2\xi^k}\|x - a^k\|^2 \\ &+ \alpha\mathcal{M}(x). \end{aligned} \tag{5.41}$$

The top line of (5.41) is the sum of functions with Lipschitz continuous gradients, Lipschitz continuity follows from the fact that the data fidelity term is Frechet differentiable [23], and the second line is a weakly-convex. This type of update can be performed by Iterative Shrinking Thresholding algorithm (ISTA) [6] or its variants such as Fast Iterative Shrinking Thresholding Algorithm (FISTA) [6]. In order to apply the FISTA algorithm we make use of the proximal map of $\mathcal{M}$. The proximal map is typically only well-defined when both parts of (5.41) are convex, recently [5] extended the definition of a proximal map to weakly-convex functions for an interval of proximal parameter.

Definition 25. (Proximal map) *Let f be a lower-semicontinuous, convex function. Then the proximal map of f for a proximal parameter t is given by*

$$\text{prox}_t(f)(y) := \underset{x}{\text{argmin}} \left(f(x) + \frac{t}{2}||x-y||^2 \right).$$

In the case of where f is only weakly-convex the proximal map defined by the formula in Definition 25 is still well-defined provided that $t < \frac{1}{\rho}$ [5]. We now determine the proximal map of $\mathcal{M}$.

Since $\mathcal{M}(x)$ is a sum over pixels and each term is independent from any others, we can determine the proximal map of $\mathcal{M}(x)$ by first finding the proximal map of $m(t)$. The proximal map of $m(t)$ is given in the following theorem.

Theorem 7. *Let m be as in (2.17). The proximal map for m, for $t < \frac{1}{2}$ is given by*

$$\text{prox}_{\frac{1}{t}} m(x) = \begin{cases} a_0 & \text{if } x \leq x_{0,+} \\ a_i & \text{if } x_{i,-} \leq x \leq x_{i,+} \text{ for } i \in \{1, ..., k-1\} \\ a_k & \text{if } x_{k,-} \leq x \\ \frac{1}{1-2t}\Big(x - t(a_{i+1} + a_i)\Big) & \text{if } x_{i,+} < x < x_{i+1,-} \text{ for } i \in \{0, ..., k-1\} \end{cases} \tag{5.42}$$

where

$$x_{i,-} = a_i - t(a_i - a_{i-1}) \text{ for } i = 1, ..k$$

$$x_{i,+} = a_i + t(a_{i+1} - a_i) \text{ for } i = 0, ..., k-1.$$

Proof. Recall from Definition 25 that for a given x we define the proximal map as the unique y with the property

$$\text{prox}_{\frac{1}{t}}(m)(x) = \underset{y}{\text{argmin}} \left\{ m(y) + \frac{1}{2t}\|y - x\|^2 \right\}.$$

We can find this by considering the first order optimality condition

$$\begin{aligned}
0 &\in \partial\left\{m(y) + \frac{1}{2t}\|y-x\|^2\right\} \implies \\
0 &\in \partial\left\{m(y) + \frac{\rho}{2}\|y\|^2 + \frac{1}{2t}\|y-x\|^2 - \frac{\rho}{2}\|y\|^2\right\} \implies \\
0 &\in \partial\left\{h_\rho(y)\right\} + \partial\left\{\frac{1}{2t}\|y-x\|^2 - \frac{\rho}{2}\|y\|^2\right\} \implies \\
0 &\in \partial h_\rho(y) + \frac{1}{t}\Big(y-x\Big) - \rho y \implies \\
0 &\in \partial m(y) + \frac{1}{t}\Big(y-x\Big) \implies \\
\frac{1}{t}\Big(x-y\Big) &\in \partial m(y),
\end{aligned}$$

where $\partial m(y)$ is understood in the sense of Definition 11. The left hand side of this optimality condition is a line of gradient $-\frac{1}{t}$ in y with intercept $\frac{x}{t}$. The right hand side of this condition is given by

$$\partial m(y) = \begin{cases} (a_i + a_{i+1}) - 2y & y \in [a_i, a_{i+1}] \\ [-\infty, a_1 - a_0] & y = a_0 \\ [a_{i-1} - a_i, a_{i+1} - a_i] & y = a_i, \quad i = 1, ..., k-1 \\ [a_{k-1} - a_k, \infty] & y = a_k. \end{cases}$$

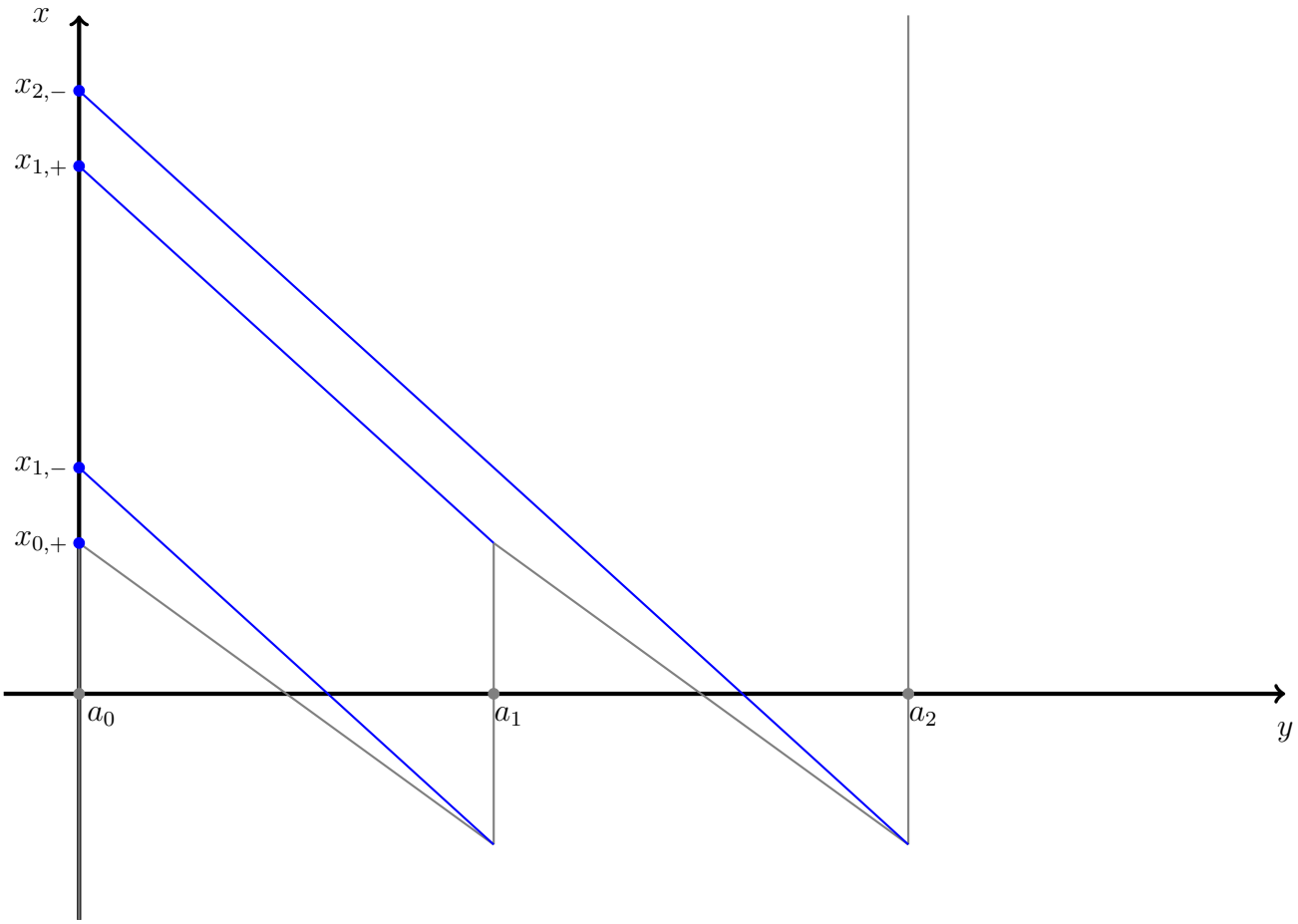
We can determine $\text{prox}_{\frac{1}{t}}(m(x))$ by finding intersection points between $\frac{1}{t}(x-y)$ and the subderivative of $m(y)$. These intersection points can be found by using a similar approach to the case of dealing with the convex $m_0(y)$ examined in [54]. We focus initially on finding what values of x lead to intersections at "corners" of ∂m.

The first corner of ∂m occurs at $(a_0, a_1 - a_0)$. If we wish to find $x_{0,+}$, which we define to be the intercept leading to intersection at this first corner, we can use the formula for a straight line as follows

$$\begin{aligned}
\Big(\frac{x_{0,+}}{t} - (a_1 - a_0)\Big) = -\frac{1}{t}(0 - a_0) &\implies \\
x_{0,+} = a_0 + t(a_1 - a_0)&
\end{aligned}.$$

Then for any value of $x \leq x_{0,+}$ the intersection will occur at $y = a_0$. Figure 5.1 gives a graphical representation of how we find these intersection points.

Figure 5.1: Demonstration of finding $x_{\pm,i}$ graphically. Here the set of admissable values is $\{a_0 = 0, a_1, a_2\}$. Each blue line has slope $-\frac{1}{t}$.



We now define some important x values: $x_{i,-}$ is the intercept leading to an intersection at the lowest point on the graph of $\partial m(y)$ at $y = a_i$ and $x_{i,+}$ the intercept leading to an intersection at the the highest point of the graph of $\partial m(y)$ at $y = a_i$. We then have any x satisfying $x_{i,-} \leq x \leq x_{i,+}$ leads to an intersection at $y = a_i$. Again we can use the formula for a straight line to find, for $1 \leq i < k - 1$,

$$\left(\frac{x_{i,-}}{t} - (a_{i-1} - a_i)\right) = -\frac{1}{t}\Big(0 - a_i\Big) \implies$$
$$x_{i,-} = a_i - t(a_i - a_{i-1})$$

and

$$\left(\frac{x_{i,+}}{t} - (a_{i+1} - a_i)\right) = -\frac{1}{t}\Big(0 - a_i\Big) \implies$$
$$x_{i,+} = a_i + t(a_{i+1} - a_i).$$

For $i = k$ we have

$$\Big(\frac{x_{i,-}}{t} - (a_{k-1} - a_k)\Big) = -\frac{1}{t}\Big(0 - a_k\Big) \implies$$
$$x_{k,-} = a_k - t(a_k - a_{k-1}).$$

Note that in order to avoid overlapping regions, and hence lead to a non well-defined proximal map, we must have $t < 0.5$

Finally we need to determine where the intersection point occurs for $x_{i,+} \leq x \leq x_{i+1,-}$ for some i. In this case the intersection occurs somewhere on the non-vertical parts of the graph of $\partial m(y)$. The intersection point occurs at $(y, (a_{i+1} + a_i) - 2y)$

$$\Big(\frac{x}{t} - \big((a_{i+1} + a_i) - 2y\big)\Big) = \frac{1}{t}\Big(y\Big) \implies$$
$$x - t(a_{i+1} + a_i) = (1 - 2t)y \implies$$
$$y = \frac{1}{1-2t}\Big(x - t(a_{i+1} + a_i)\Big).$$

Combining all of this gives

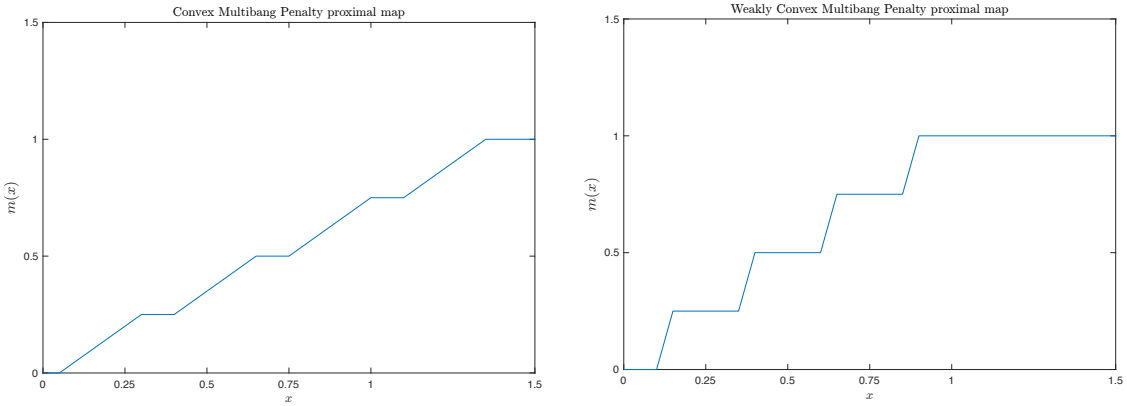
$$\text{prox}_{\frac{1}{t}} m(x) = \begin{cases} a_0 & \text{if } x \leq x_{0,+} \\ a_i & \text{if } x_{i,-} \leq x \leq x_{i,+} \text{ for } i \in \{1, ..., k-1\} \\ a_k & \text{if } x_{k,-} \leq x \\ \frac{1}{1-2t}\Big(x - t(a_{i+1} + a_i)\Big) & \text{if } x_{i,+} < x < x_{i+1,-} \text{ for } i \in \{0, ..., k-1\} \end{cases}$$

as required. □

This shows that the choice of a weakly-convex multi-bang regularizer gives the desirable property that $\text{prox}_{\frac{1}{t}} m(a_i) = a_i$, at the cost of limiting the values of t which give a well-defined proximal map. Figure 5.2 shows a comparison of the proximal maps for both the convex multi-bang regularizer m_0 and the weakly-convex multi-bang regularizer for a specific example. By the definition of the proximal map we have, for any $\alpha > 0$

$$\begin{aligned} \text{prox}_{\frac{1}{t}}(\alpha f)(y) :=& \operatorname*{argmin}_x \{\alpha f(x) + \frac{1}{2t}\|x - y\|^2\} \\ =& \operatorname*{argmin}_x \{f(x) + \frac{1}{2\alpha t}\|x - y\|^2\} \\ =& \text{prox}_{\frac{1}{t\alpha}}(f)(y). \end{aligned}$$

Figure 5.2: Comparing proximal maps. The left shows the proximal map obtained using the convex multi-bang penalty m_0 used in [19] and the right is the weakly convex multi-bang penalty m proximal map. Here $\mathcal{A} = \{0, 0.25, 0.5, 0.75, 1\}$.



(a) Convex proximal map

(b) Weakly convex proximal map

Therefore, the proximal map for αm is just the formula for the proximal map of m with t replaced by αt, with the additional change that the proximal map is only well-defined for $\alpha t < \frac{1}{2}$. As stated at the start of the section, we can sum over all pixels to obtain the proximal map for $\alpha\mathcal{M}$. The gradient, with respect to x of the function

$$\mathcal{F}(x) := \|R[x]f - d\|_2^2 - \lambda^T(y - Dx) + \frac{\beta}{2}\|Dx - y\|_2^2 + \frac{1}{2\xi^k}\|x - a^k\|^2$$

is

$$\nabla_x(\|R[x]f - d\|_2^2) + D^T\lambda + \rho D^T(Dx - y) + \frac{1}{\xi^k}(x - a^k),$$

and with this we are ready to explicitly write out the Fast Iterative Shrinkage Thresholding Algorithm(FISTA)[6] for our problem.

Algorithm 1 FISTA algorithm

1: Input x_0 an initial guess, a step size t and tolerance δ_1. Set $p_1 = x_0$.

2: Let $s_1 = 1$.

3: **for** $k \geq 1$ **do**

$$\begin{aligned} u_k &= p_k - t\nabla\mathcal{F}(p_k) \\ x_k &= \text{prox}_{\frac{1}{t}}(\alpha\mathcal{M})(u_k), \\ s_{k+1} &= \frac{1 + \sqrt{1 + 4s_k^2}}{2} \\ p_{k+1} &= x_k + \frac{s_k - 1}{s_{k+1}}\left(x_k - x_{k-1}\right). \end{aligned}$$

4: Terminate when $\|x_k - x_{k-1}\|_2 < \delta_1$.

Note that our choice of $\frac{1}{t}$ rather than t makes the first step in FISTA more like gradient descent. Since

$$||R[x]f - d||_2^2 + \lambda^T(Dx - y) + \frac{\beta}{2}||Dx - y||_2^2 + \frac{1}{2\xi^k}\|x - a^k\|^2$$

has Lipschitz continuous gradient, as each term is smooth, and $\alpha\mathcal{M}(x)$ is a ρ-weakly convex lower-semicontinuous function, provided that we choose t sufficiently large [6, Section A: Proposition 5], we have convergence of FISTA to a critical point of

$$\begin{aligned} \underset{x}{\text{argmin}} ||R[x]f - d||_2^2 + \lambda^T(Dx - y) + \frac{\beta}{2}||Dx - y||_2^2 + \frac{1}{2\xi^k}\|x - a^k\|^2 \\ + \alpha\mathcal{M}(x). \end{aligned}$$

Therefore to perform the x minimisation step of the ADMM algorithm we use FISTA given above. Since FISTA only gives an approximate minimiser for x, and theoretical convergence of ADMM requires exact minimisers, we use a small tolerance $\delta_1 > 0$ to terminate when successive iterates become sufficiently close with respect to the 2-norm. We now look at updating the y or total variation part of a.

5.1.2 Updating y

As mentioned at the beginning of chapter 5 we use the non-smooth variant of TV (5.12) in the numerics given in section 6, we do however give details on how we update y when we use a smoothed TV given by

$$\mathrm{TV}_c(a) = \sum_{i=1}^{M^2-1} \sqrt{\|D_i a\|_2^2 + c}. \tag{5.43}$$

We first examine the non-smooth case. Omitting terms not involving y and after some algebraic manipulation as in [46], we can write the y update as

$$\boldsymbol{y}_i^{k+1} = \operatorname{argmin} \left\{ \gamma_a \|\boldsymbol{y}_i\|_2 + \frac{\beta}{2} \left\| \boldsymbol{y}_i - \left(D_i x + \frac{1}{\beta} \boldsymbol{\lambda}_i^k \right) \right\|_2^2 \right\}. \tag{5.44}$$

Problems of the form (5.44) have been studied before [10, 32], one particularly useful result comes from [32].

Proposition 1. *For any $\beta > 0$ and $v, y \in \mathbb{R}^n$, the minimiser of*

$$\mathcal{J}(y) = \|y\|_2 + \frac{\beta}{2}\|y - v\|_2^2$$

is given by

$$y(v) = \max\left\{ \|v\|_2 - \frac{1}{\beta}, 0 \right\} \frac{v}{\|v\|_2}.$$

Proof. A proof of this can be found in [32]. □

By applying Proposition 1 with $v = D_i x$ we can explicitly write out the y update step. For each $\boldsymbol{y_i}$ we find that

$$\boldsymbol{y}_i^{k+1} = \max\left\{ \left\| D_i x^{k+1} + \frac{1}{\beta} \boldsymbol{\lambda}_i^k \right\| - \frac{\gamma_a}{\beta}, 0 \right\} \frac{D_i x^{k+1} + \frac{1}{\beta}\boldsymbol{\lambda}_i^k}{\|D_i x^{k+1} + \frac{1}{\beta}\boldsymbol{\lambda}_i^k\|} \tag{5.45}$$

where $\frac{0}{0} = 0$. Since y is simply a reordering of entries in all of the $\boldsymbol{y}_i$ updating the $M(M-1)$ $\boldsymbol{y}_i$ allows us to fully update y. This completes the non-smooth update. We now move onto the case where we use

$$TV_c = \sum_{i=1}^{M^2-1} \sqrt{\|D_i a\|_2^2 + c}.$$

We again split the y update into sub problems involving $\boldsymbol{y_i}$. The first order optimality condition for each of the $\boldsymbol{y_i}$ updates in this case gives

$$0 = \gamma_a \frac{y_i}{\sqrt{\|y_i\|_2^2 + c}} - \boldsymbol{\lambda_i} + \beta(y_i - D_i x^{l+1}) \tag{5.46}$$

for all i. Whilst this cannot be explicitly solved for y_i easily, we can make use of the gradient, which is the right hand side of (5.46) to solve the y update via gradient

descent. Again we use a small tolerance $\delta_2 > 0$ for the maximum difference between successive iterates i.e, $\|\boldsymbol{y_i}^{k+1} - \boldsymbol{y_i}^{k}\|_2$ before we terminate, in order to speed up convergence of ADMM. Once we have updated all of the y_i in this way we can combine them to update y. With the x and y updates examined the λ update is simply given by

$$\lambda^{k+1} = \lambda^k + \beta(y^{k+1} - Dx^{k+1}).$$

To complete the algorithm we set $a^{k+1} = x^k$, where x^k is the final x output from performing the ADMM algorithm. In practice having an adaptive technique for varying β leads to improved speed of convergence [10], although the theoretical proof does rely on constant β. For all of our numerical reconstructions we use the following simple update for β as in [10]

$$\beta^{k+1} := \begin{cases} \tau^+\beta^k & \text{if} \|r_k\|_2 > \mu\|s_k\|_2 \\ \frac{\beta^k}{\tau^-} & \text{if} \|s_k\|_2 > \mu\|r_k\|_2 \\ \beta^k & \text{otherwise} \end{cases} \tag{5.47}$$

where $r_k := Dx^k - y$ and $s_k := \beta D^T(y^{k+1} - y^k)$ as before. We now move onto the f update.

5.1.3 Updating source radiation density f

Since the AtRT is linear in f updating f is much less computationally expensive. Removing the parts of the objective function (5.13) we obtain

$$f^{k+1} \in \underset{f}{\operatorname{argmin}} \|R[a^{k+1}]f - d\|^2 + \eta \sum_{i=1}^{M^2-1} \sqrt{\|D_i f\|_2^2 + c} + \frac{1}{2\xi^k}\|f - f^k\|^2 \tag{5.48}$$

when we use smoothed TV and

$$f^{k+1} \in \underset{f}{\operatorname{argmin}} \|R[a^{k+1}]f - d\|^2 + \eta \sum_{i=1}^{M^2-1} \|D_i f\|_2 + \frac{1}{2\xi^k}\|f - f^k\|^2 \tag{5.49}$$

when using the non-smooth version. As in the a update we find numerically that using the non-smooth TV yields convergence over a wider range of regularization parameters and so use the non-smooth version. We can then split either (5.48) or (5.49) as in the case of a and solve the resulting problem via ADMM, although since we do not have a

multi-bang regularizer in this case the computation is faster and does not require the use of a proximal map. We can now fully describe the joint recovery algorithm

Algorithm 2 Joint reconstruction algorithm

Input a^0 as initial guess, step sizes t, β^0, tolerances $\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6$ and regularization parameters α, λ and μ.
Set f^0 to be the least squares solution of $\|R[a^0]f - d\|^2$.
for $k \geq 0$ **do**
    Set $x^0 = a^k$ and $y^0 = Dx^0$.
    for $l \geq 0$ **do**
        Update x^{l+1} via ISTA or FISTA with δ_1 as a tolerance on $\|x^{l+1} - x^l\|$.
        Update y^{l+1} via gradient descent on (5.46) terminate when $\|y^{l+1} - y^l\| < \delta_2$.
        Set $\mu^{l+1} = \mu^l + \beta^l(y^{l+1} - Dx^{l+1})$.
        Update β^{l+1} via (5.47)
        Terminate when $r^l < \delta_3$ and $s^l < \delta_4$ and output $a^{k+1} = x^{l+1}$.
    Update f^{k+1} via (5.48) or (5.49) using ADMM with tolerance δ_5.
    Terminate when $\|a^{k+1} - a^k\|_2 < \delta_6$ and $\|f^{k+1} - f^k\|_2 < \delta_6$.

We point out that in this algorithm β^0 is reset to the same initialised value whenever the inner iterations aimed at the a update in (5.14) (those indexed by l) restart. Before presenting numerical results we also cover the Discrete Tomography case discussed in section 2.5.

5.2 Discrete Tomography algorithm

In the case when we work with the Radon transform we are only inverting to recover f and so do not need to alternate updates. Note that when working with the Radon transform the matrix $R[0]$ is not changed at each iteration and can be calculated before beginning. As we are working in the DT setting we expect f to take values from a discrete set, as stated before this lends itself nicely into using multi-bang regularization with the admissible set containing the expected values. In this case our optimization problem is given by

$$\underset{f\in\mathbb{R}^{M^2}}{\operatorname{argmin}} \|R[0]f - d\|_2^2 + \alpha\mathcal{M}(f) + \gamma_f \mathrm{TV}(f) \tag{5.50}$$

This optimization problem can be split into two parts x and y, in a similar way to both a and f in the joint reconstruction setting, and solved using ADMM. The algorithm for updating in this case is given below

Algorithm 3 Discrete Tomography algorithm

1: Input f^0 as initial guess, step sizes t, β^0, tolerances $\delta_1, \delta_2, \delta_3, \delta_4$ and regularization parameters α, λ and β.
2: **for** $k \geq 0$ **do**
3: Set $x^0 = f^k$ and $y^0 = Dx^0$.
4: **for** $l \geq 0$ **do**
5: Update x^{l+1} via ISTA or FISTA with δ_1 as a tolerance on $\|x^{l+1} - x^l\|$.
6: Update y^{l+1} via gradient descent on (5.46).
7: Set $\lambda^{l+1} = \lambda^l + \beta^l(y^{l+1} - Dx^{l+1})$.
8: Update β^{l+1} via (5.47)
9: Terminate when $r^l < \delta_2$ and $s^l < \delta_3$ and output $f^{k+1} = x^{l+1}$.
10: Terminate when $\|f^{k+1} - f^k\|_2 < \delta_4$.

Before presenting some numerical reconstructions obtained by using Algorithms 2 and 3, we first briefly discuss a possible parameter which controls behaviour of both algorithms 2 and 3 when using non-smooth TV.

5.2.1 μ^k and a possible condition on convergence

Recall that in the proof of lemma 21 we obtained the following inequality

$$\begin{aligned} \mathrm{L}_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) - \mathrm{L}_\beta(x^k, y^k, \lambda^k) \\ \leq -\frac{\beta}{2}\left\|y^k - y^{k+1}\right\|^2 - \left(\frac{\beta\mu}{2} - \frac{\rho}{2}\right)\|x^k - x^{k+1}\|^2 \\ + \frac{1}{\beta}\|\lambda^k - \lambda^{k+1}\|^2. \end{aligned} \tag{5.51}$$

In the case where $g(y)$ has Lipschitz continuous gradient, we were able to replace $\|\lambda^k - \lambda^{k+1}\|$ by $L\|y^k - y^{k+1}\|$ and obtain the result we needed. However, when the

gradient of $g(y)$ is not Lipschitz continuous things are not so straightforward. Even though there is currently no proof of theoretical convergence of ADMM when g does not have Lipschitz continuous gradient we have, at least in all the numerical experiments we have examined, observed numerical convergence for a wide variety of examples. In fact if we define a parameter

$$\mu^k := \frac{1}{4}\left\{\frac{\beta\|y^k - y^{k+1}\|^2 + (\beta\mu - \rho)\|x^k - x^{k+1}\|^2}{\|r^{k+1}\|^2}\right\}, \tag{5.52}$$

where $r^{k+1} = Ax^{k+1} + y^k$ then, using (5.52) and the λ^{k+1} update we can rewrite (5.51) as

$$\begin{aligned} \mathrm{L}_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) - \mathrm{L}_\beta(x^k, y^k, \lambda^k) & \\ \leq -\frac{\beta}{4}\left\|y^k - y^{k+1}\right\|^2 - \left(\frac{\beta\mu}{4} - \frac{\rho}{4}\right)\|x^k - x^{k+1}\|^2 & \\ + (\beta - \mu^k)\|r^{k+1}\|^2. & \end{aligned} \tag{5.53}$$

The quantity μ^k therefore gives access to a method of proving convergence, provided $\beta - \mu^k < 0$, the exact behaviour of μ^k for convergent numerical examples is something which could be studied further. The quantity μ^k has the following very important property. If $(\beta - \mu_k) < 0$ for all $k \geq \tilde{k}$, for some $\tilde{k}$ then in a similar manner as in the proof of lemma 21 choosing

$$\delta = \min\left\{\frac{\beta}{4}, \left(\frac{\beta\mu}{2} - \frac{\rho}{2}\right), \left(\mu^k - \beta\right)\right\}$$

gives the following corollary.

Corollary 5. *Let f and A satisfy the conditions given in Lemma 21 and suppose g is a proper convex function. Additionally suppose that $(\mu^k - \beta) > 0$ for all k sufficiently large, then there exists $\delta > 0$ such that the ADMM iterates satisfy*

$$\mathrm{L}_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) \leq \mathrm{L}_\beta(x^k, y^k, \lambda^k) - \delta\left(\|x^{k+1} - x^k\|^2 + \|y^{k+1} - y^k\|^2 + \|r^{k+1}\|^2\right).$$

The proof of convergence would then follow by a similar argument involving KL functions as for the Lipschitz continuous case given in [32]. Accounting for the fact that we actually update β at each iteration if we have $\mu^k - \beta^k > 0$ at every iteration after some finite step we would be able to ensure convergence. Further work needs to be done to precisely identify the relationship between $(\mu^k - \beta^k)$ and the convergence of the algorithm but we present an example that shows if this quantity is negative for

a large number of iterations the algorithms diverge and when it is positive after some number of iterations the method converges.

We now present the numerical reconstructions.

Chapter 6

Reconstructions

Throughout this section we have produced data on a 340 by 340 pixel grid and reconstructed on a 200 by 200 grid to avoid inverse crimes. The size of each pixel in all examples presented is 0.2 by 0.2 (i.e. $dx = 0.2$), the size of the domain is of particular importance in the joint reconstruction case as large attenuation values typically cause reconstructions of a from noisy data to be poor [17]. All of the following examples have 5% added Gaussian white noise and were produced on a standard 4 core laptop using MATLAB. Note that much of the computational time is spent computing and recomputing the matrix representation of $R[a]$ when a is updated, and many of the steps in this reconstruction can be done using parallel computing toolboxes. When we consider the Radon case where we only reconstruct f this computation time is no longer a concern. Unless otherwise stated the following reconstructions use 16 parallel ray projections which are equally spaced with some small perturbation to make the angles irrationally related (i.e. unless otherwise stated we only use data with 16 different values of θ). Irrationally related angles have been shown to reduce the number of projections required to obtain good reconstructions [17, 63]. For 16 projections the simultaneous reconstruction algorithm takes approximately 15 minutes and the Discrete Tomography algorithm takes approximately 3 minutes.

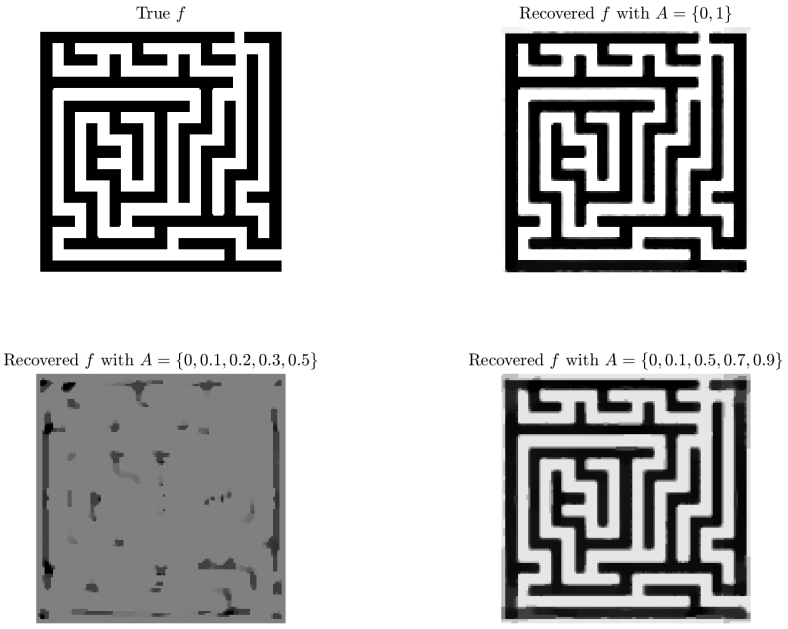
In both algorithms there are a large number of parameters to control. This gives good flexibility but does require extensive parameter tuning in order to obtain optimal results. Therefore, parameter selection is an important part of the algorithm. Although optimal results require fine tuning, we typically find that a large range of the parameters $\beta^0, t, \alpha, \gamma_a$ and γ_f lead to good results. Firstly, the step sizes t and β^0 can

be chosen quite widely, and only the number of iterations required for convergence are affected. Since it is actually the quantity $\frac{1}{t}$ that we use as the step size, small values of t lead to large step sizes. For a given α, the strength of the multi-bang penalty, there is an upper bound on the value of t that can be used which is determined by (5.42), since the proximal map is not well-defined if αt is too large. As we update β^0 at each iteration, the initial choice of β^0 can be widely varied and still produce good results. Perhaps the most interesting parameter choice concerns α. Since we are interested in the quantity αt, taking α too large causes the proximal map to be ill-defined. If α is close to the limit allowed by αt then in numerical experiments we found the algorithm produced completely multi-bang reconstructions but the boundaries were poorly recovered, whereas taking α too small tends to give non multi-bang reconstructions. Again there is a large range, several orders of magnitude, for which good reconstructions are obtained. For example, in the reconstructions shown in this section a choice of $\alpha \in [10^{-3}, 0.5]$ and $t \in [0.001, 0.1]$ would yield similar results. There is some interaction between the choice of α and γ_a; in particular if one is much larger than the other the reconstructions essentially become purely multi-bang or total variation reconstructions. Keeping the orders of α and γ_a the same, at least in the experiments we have tried, produces a with both good multi-bang and shape recovery. Finally a large range of γ_f yield good recovery and so this parameter is not extensively tuned.

In the following examples we use initial guesses where a is constant. In practice convergence is obtained for all tested phantoms for any constant initial guess of a, provided the constant value lies between a_0 and a_n in the admissable set. We therefore use initial guess $a^0 = 0$ for all numerical results presented here. The last general comment we make is that if we set $\xi = \infty$, effectively removing the added terms $\|a - a^k\|^2$ and $\|f - f^k\|^2$ from (5.14) we still obtain convergence. In many cases removing this part improves the speed of convergence, although the theoretical proof of convergence does not hold in this case. Throughout this chapter we fix $\xi = 50$ and set all tolerances δ_i to 1×10^{-3}.

Figure 6.1 shows an example of numerical reconstruction in the DT case. The true phantom for f is a binary image of a maze with admissible set $A = \{0, 1\}$. The top right reconstruction is obtained with $\alpha = \gamma_f = 0.05$ with the correct set

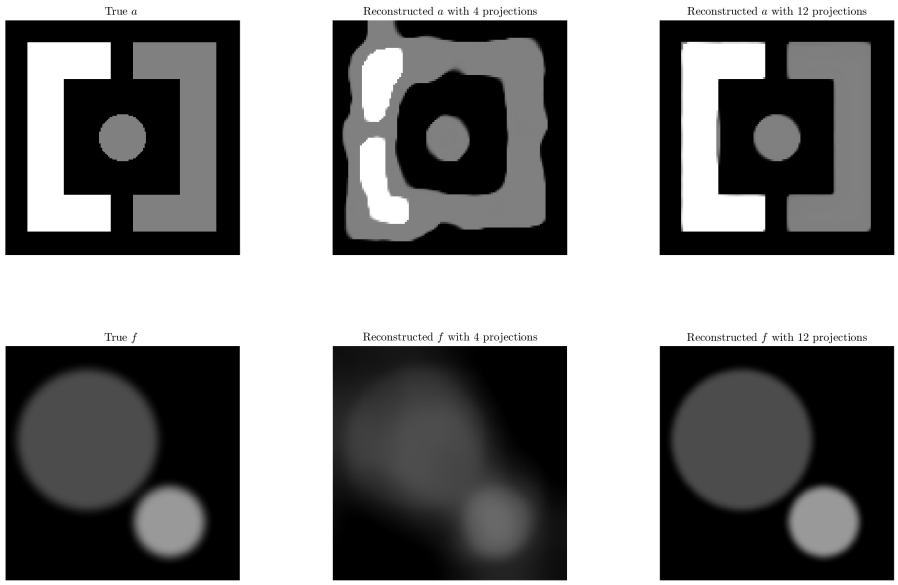
Figure 6.1: Discrete Tomography phantom recovery, here the true image (top left) is binary with $A = \{0, 1\}$.



A and initial step sizes $t = 0.025$ and $\beta^0 = 0.1$. The bottom left reconstruction is obtained with $\alpha = 0.1$, $\gamma_f = 0.05$ and initial step sizes $t = \beta^0 = 0.05$. The bottom right reconstruction is obtained with $\alpha = 0.075$, $\gamma_f = 0.05$ and initial step sizes $t = \beta^0 = 0.075$. When the correct set A is used the reconstruction is multi-bang and recovery is good with only a few pixels on the upper most edge not being correctly assigned. When we input incorrect but "close" multi-bang values into the admissible set, as in the bottom right reconstruction in Figure 6.1 (here the largest value is 0.9), we again obtain multi-bang reconstructions however there is more variation and the image is no longer binary. When the incorrect values are further away, such as the bottom left reconstruction where the largest value is 0.5, we see that the reconstruction is much poorer, the recovered f is still multi-bang but the much lower values in A have caused the recovered f to grow in order to compensate. This behaviour is good in some aspects as, at least in this example, you can obtain good results provided that you have a good approximation of the correct set A.

Figure 6.2 shows the effect of the number of projections on reconstruction quality in a joint reconstruction example. Here the phantom for a is made up of 3 regions with $A = \{0, 50, 100\}$. The left column shows the true phantoms for a and f. The middle

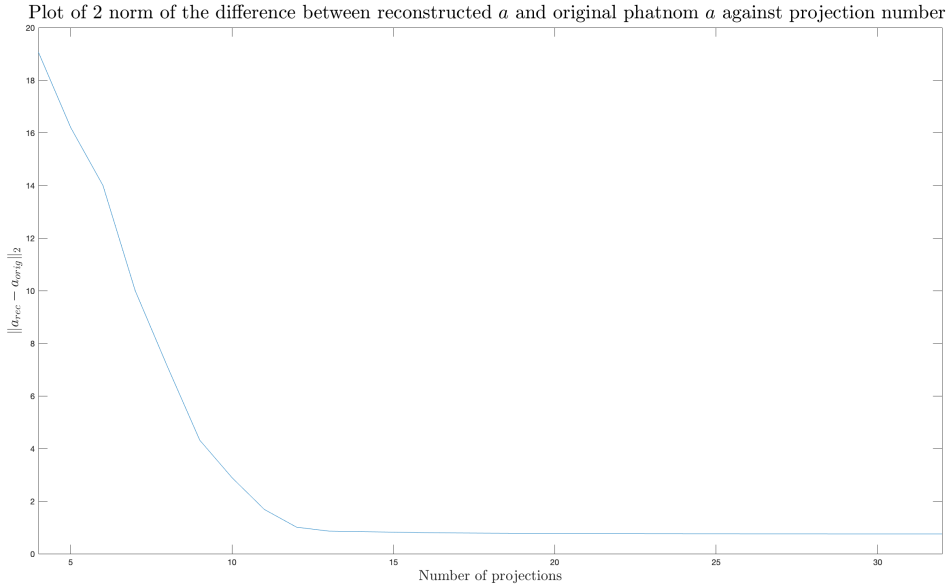
Figure 6.2: Demonstration of the effect the number of projections has on the joint reconstruction of a and f. In this case f is continuous and a is multi-bang with admissible set $A = \{0, 50, 100\}$



column is an optimized reconstruction using 4 projections with $\alpha = 0.1$ and $\lambda = \eta = 0.05$. The right column shows an optimized reconstruction using 12 projections with $\alpha = 0.1, \lambda_a = 0.05$ and $\lambda_f = 0.15$. In both reconstructions for a we obtain multi-bang solutions. The middle column shows a poor recovery of the structure of a and f. The recovered a has a lot of misclassification and has been unable to separate the regions. The inaccuracies in a have an impact on the recovery of f, with the outer most regions of f being poorly recovered. The rightmost column is a very good recovery of both a and f, with just a small section on the left bracket being misclassified. The matching f is also very well recovered.

Figure 6.3 shows a graph of the 2-norm error between an optimized reconstruction and the true a against the number of projections used. As stated earlier in this Chapter, the angles used are evenly spread and then perturbed slightly to make them irrationally related. The reconstructions obtained for fewer projections than 4 are very poor and not included in the plot. We can see that the increasing the number of projections rapidly increases reconstruction quality up until 12 projections. After 12 projections the improvement is minimal and there is no improvement after 20. It is important to mention that we found more complex phantoms generally required more

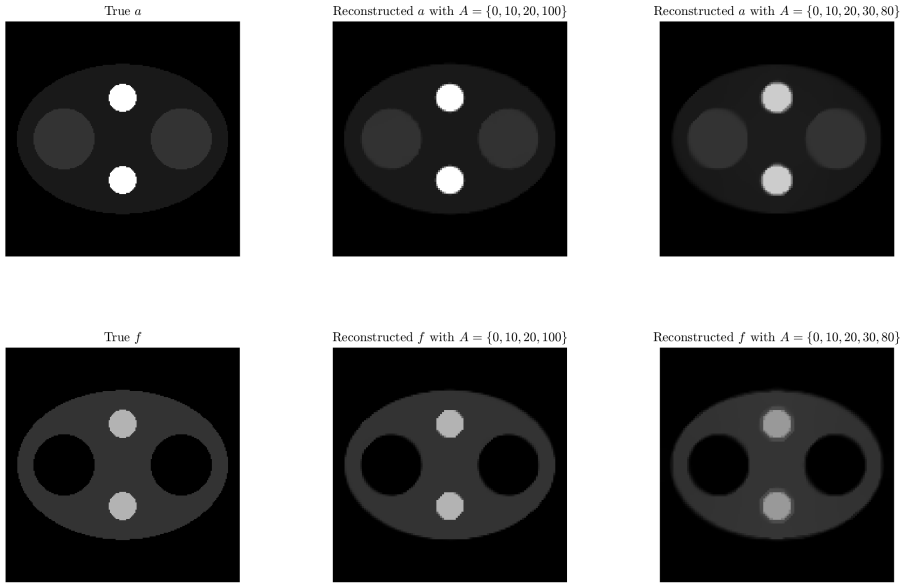
Figure 6.3: Plot of the 2-norm error of the reconstructed a against number of projections used. Here the true a is the same as in Figure 6.2.



projections to obtain a good reconstruction than the more simple examples.

Figure 6.4 gives a joint reconstruction example when a and f have the same support and are both multi-bang. Note that even though f is multi-bang we do not include multi-bang regularization in the update steps of f in Algorithm 2; it is only TV regularized. The left hand column gives the true a and f, the correct admissible set for a is $\{0, 10, 20, 100\}$ and for f is $\{0, 0.2, 0.7\}$. The choice of a being much larger than f is to make the chest model more realistic [42]. The middle column gives an optimized reconstruction when the admissible set for A is known, here the parameters are $\alpha = 0.05$, $\gamma_a = 0.05$, $\gamma_f = 0.5$ and initial step sizes $t = \beta^0 = 0.1$. The right hand column gives an optimized reconstruction when the admissible set is incorrectly chosen as $\{0, 10, 20, 30, 80\}$; here the parameters are $\alpha = 0.075$, $\gamma_a = 0.05$, $\gamma_f = 0.5$ and initial step sizes $t = \beta^0 = 0.1$. In both reconstructions the multi-bang property of a is recovered. More interestingly this also seems to have given f the multi-bang property without actually having to apply multi-bang regularization directly to f. Figure 6.5 shows the absolute error in the reconstructions given in 6.4 divided by 100. In the case where A is correct the reconstruction of a and f is good with only a few pixels on the boundary of the highest attenuation circles being mislabelled. When we reduce the maximum known value to 80 the reconstruction still captures the features

Figure 6.4: Joint reconstruction where a and f are both multi-bang. Here a is a chest phantom from in [42] with admissible set $A = \{0, 10, 20, 30, 100\}$.



of a but the circles of highest attenuation have a constant error across them, which is to be expected. This reduction in maximum a value causes f to generally be smaller than it should. This may be because with lower a a lower value of f along a ray (s, θ) would be able to produce the same $R_a f(s, \theta)$.

Figure 6.6 gives a plot of the objective function against iteration number for the chest reconstructions given in Figure 6.4. Here the objective function at the $i - th$ iteration is found by evaluating (5.13) at a_i and f_i. The top graph corresponds to the middle column and the bottom graph corresponds to the right column of Figure 6.4. In both cases the general behaviour is the same, we have a decreasing objective function with a sharp decrease in the first 10 or so iterations and a slower decrease after this. There are two differences however, the first is in the value of the objective function itself. In the top graph we have a much lower value achieved at around 0.5 whereas in the bottom the objective function reaches 1.5. The second difference is in the number of iterations. The top graph corresponding to the correct admissible set terminates after 38 iterations whereas the reconstruction with a partially incorrect set terminates after 53. When the joint algorithm converges, this behaviour of the objective function is typical of the numerical reconstructions we have computed throughout this project.

Figure 6.5: The absolute error of reconstructions given in Figure 6.4 divided by 100. The left hand column shows reconstructions obtained using the correct admissible set and the right hand column shows reconstructions obtained using a spurious admissible set.

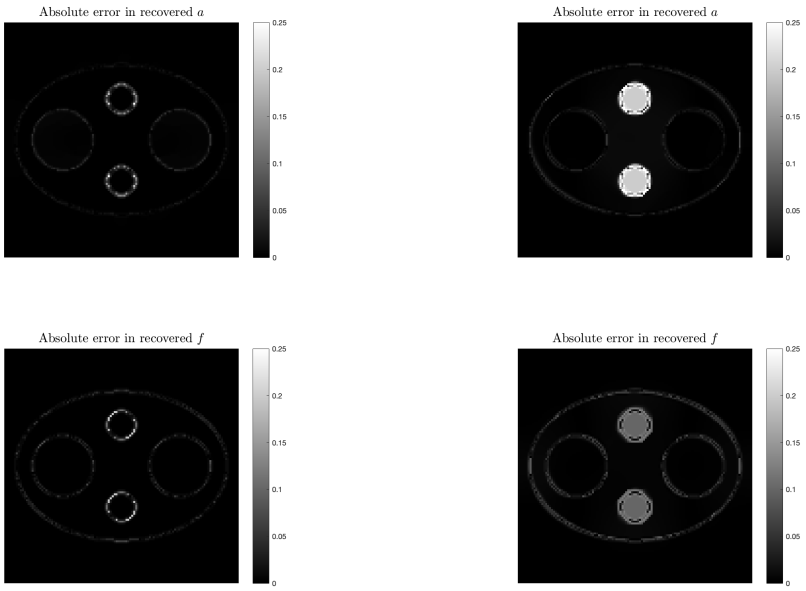


Figure 6.7 gives a joint reconstruction example where again both a and f have the same support but here f is continuous and not multi-bang. The left hand column gives the true a and f, and the admissible set for a is $\{0, 50, 100\}$. The middle column gives an optimized reconstruction when the admissible set for A is known. Here the parameters are $\alpha = 0.05$, $\gamma_a = 0.15$, $\gamma_f = 0.25$ and initial step sizes $t = \beta^0 = 0.1$. The right hand column gives an optimized reconstruction when the admissible set is incorrectly set as $\{0, 10, 20, 30, 50, 80\}$, here the parameters are $\alpha = 0.075$, $\gamma_a = 0.05$, $\gamma_f = 0.5$ and initial step sizes $t = 0.05$ $\beta^0 = 0.15$. Again in both reconstructions the multi-bang property of a is well recovered. When the admissible set is incorrect we again lose more detail in both a and f. We also see that the non-zero regions of a grow when the maximum multi-bang value is lowered. The growth in a is also shown in f where it has also grown to compensate. Figure 6.8 gives the absolute error between the reconstructions in a and f and the true phantom divided by 100. The left column corresponds with the middle column of Figure 6.7 and the right column corresponds with the right column of Figure 6.7. In both cases we see that the error is

Figure 6.6: Plot of objective function against iteration number for the example given in Figure 6.4. Here the top graph corresponds to the middle column and the bottom the right column in Figure 6.4.

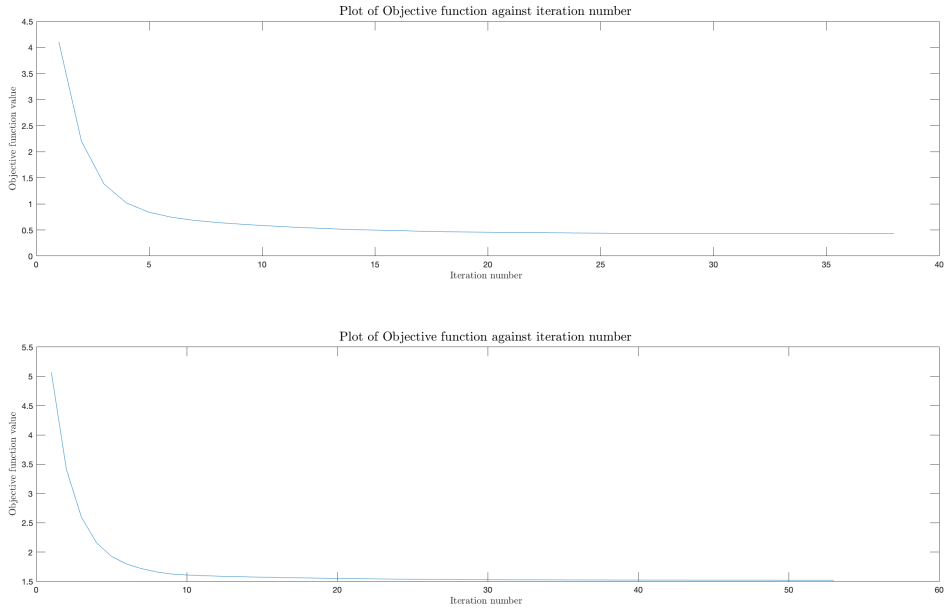


Figure 6.7: Joint reconstruction where only a is multi-bang with a and f sharing the same support. Here a and f are based on a walnut phantom from [33]. The true admissible set is $A = \{0, 50, 100\}$.

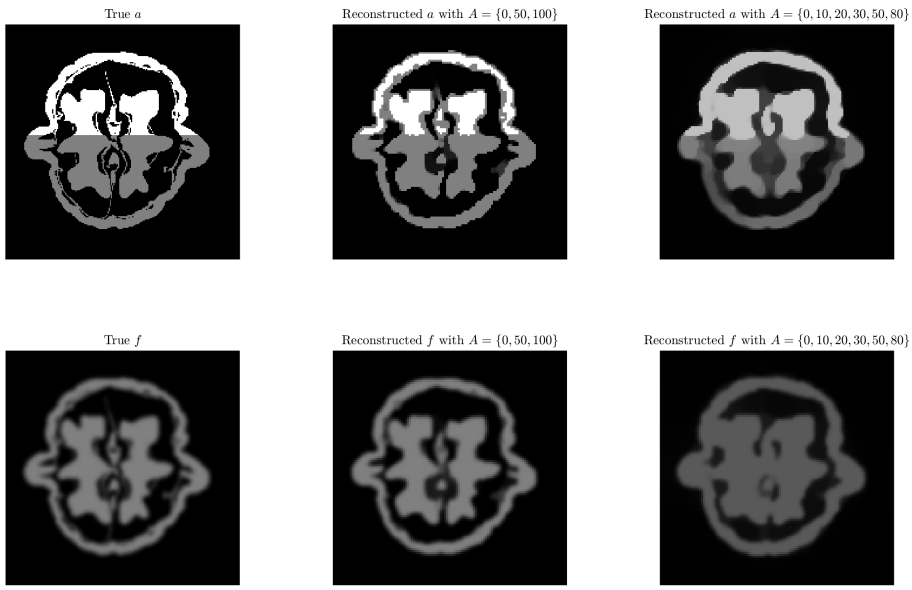
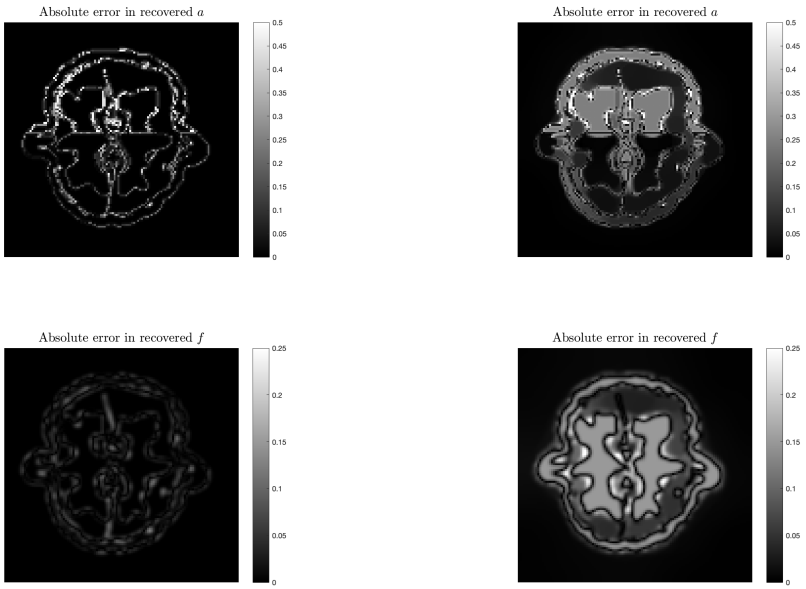


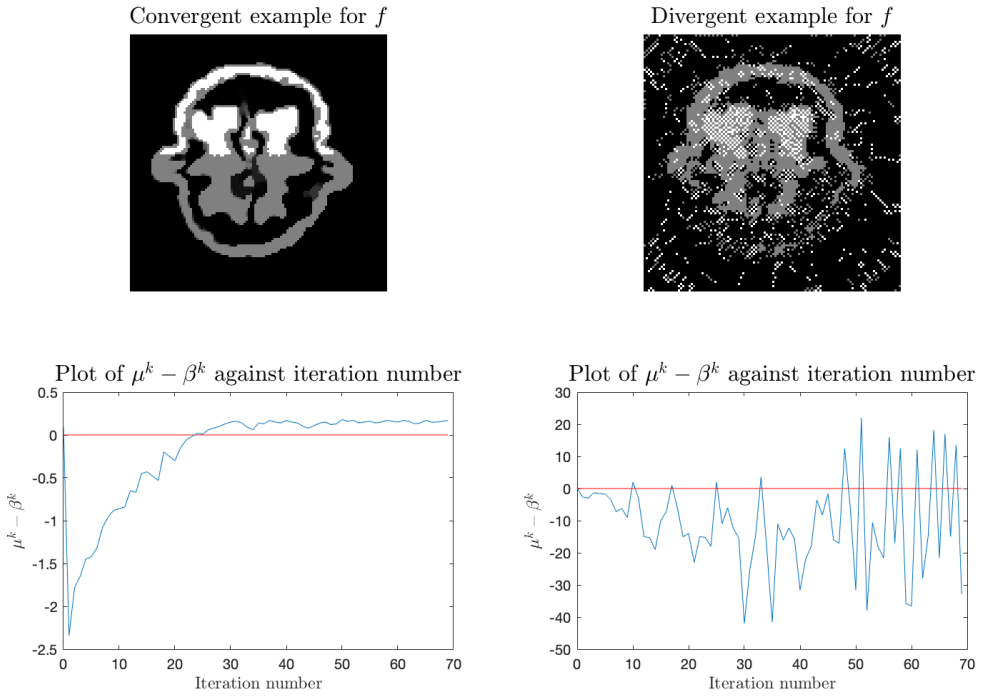
Figure 6.8: The absolute error of reconstructions given in Figure 6.7 divided by 100. The left column shows reconstructions obtained using the correct admissible set and the right hand column shows reconstructions obtained using a spurious admissible set.



concentrated on the boundaries of a and f, with there being significantly more error in the case where the admissible set is incorrect. The areas in which a and f are 0 are well recovered in both examples.

The next figure we examine relates to the parameter μ^k which discussed in Chapter 5.2.1. The first row Figure 6.9 shows two examples of reconstructions for the DT case where the true phantom for f is the same as the true phantom for a in Figure 6.9. The second row shows a plot of the quantity $\mu^k - \beta^k$ as defined in (5.52). In the left hand column the algorithm converged in 69 iterations, the right hand column does not converge and was terminated at 69 iterations. In both cases we choose initial parameters so that $\mu^0 = \beta^0$ is positive and at this time we are unsure as to whether this aids convergence or not. All parameters bar two remain the same in each of these reconstructions, on the left we use $t = 0.1$ and $\alpha = 0.15$ and on the right $t = 0.5$ and $\alpha = 0.001$. In the convergent case we see that after about 25 iterations $\mu^k - \beta^k$ remains positive and seems to converge after about 40 iterations. When the algorithm diverges the value of $\mu^k - \beta^k$ rapidly oscillates and the oscillations get worse as the iterations increase. It is worth noting that even though the algorithm is diverging some features of f are recovered particularly in the upper half of the image which corresponds to

Figure 6.9: A plot of $\mu^k - \beta^k$ at each iteration for a Discrete Tomography reconstruction. Here the phantom for f is the same as the phantom for a in Figure 6.7. The left hand column shows a convergence example and the right hand side a divergent.



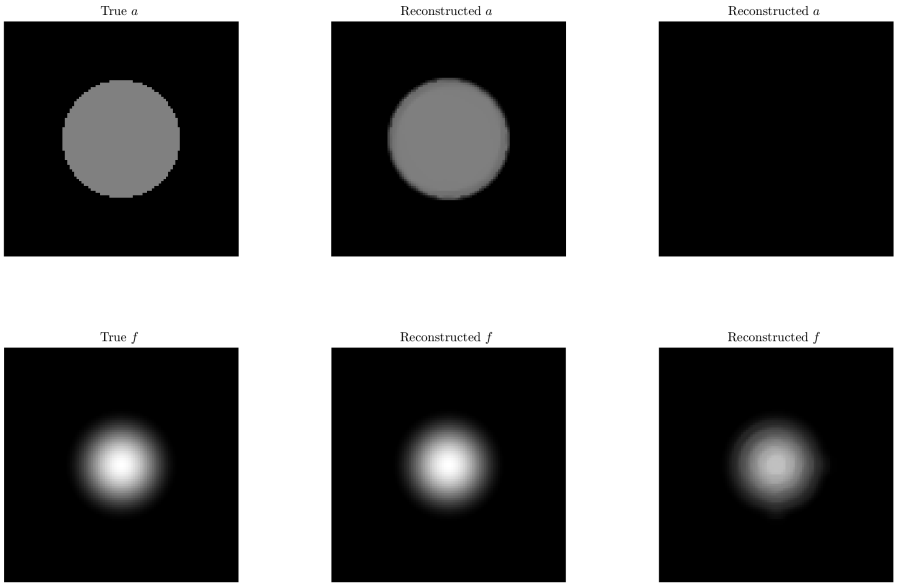
the larger values of a. Further study into the properties of μ^k during reconstruction is something which could be an area for future research.

The last set of results relate to the non-unique solution for radial a and f which was discussed in Introduction and Literature review.

6.1 Radial and close to radial reconstructions

This section contains results relating to the case where the data used for reconstruction is produced from a which is multi-bang and consists of a single circle. The f used is supported on the circle and goes to 0 on the edge. The exact f we use in this section is given by $f = 1 + \cos(|\boldsymbol{x}|)$. As discussed in section 2.5, [55] gives a method which shows that with this choice of a and f a solution exists with $a = 0$. We present three figures. The first shows that numerically in the case of the radial a and f this can lead to a second solution being recovered. The other two figures show that, using the same parameters which led to different solutions in the first case, when a is perturbed slightly we can only obtain one reconstruction. This does not contradict the result in

Figure 6.10: Joint reconstruction of a and f where both are radial and share the same support. The admissible set here is $A = \{0, 50\}$.

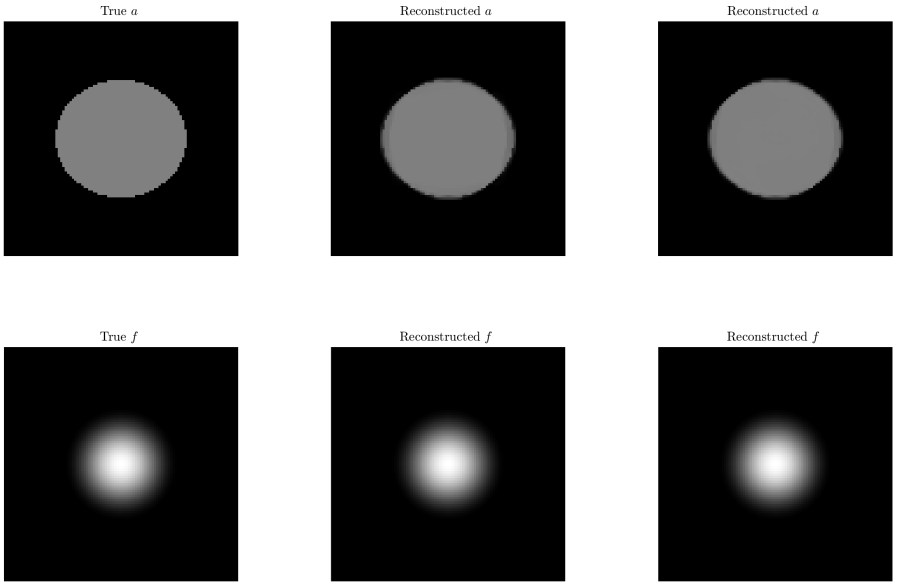


[55] as the theoretical results used to obtain a second solution requires a and f to be radial.

Figure 6.10 shows reconstructions in the case where a and f are both radial. The admissible set is $A = \{0, 50\}$. The middle column gives a reconstruction where we obtain a solution close to the true phantom and the right column gives a reconstruction which recovers $a = 0$ and a different, but still radial f. For the middle column the parameters are $\alpha = 0.1$, $\gamma_a = 0.15$, $\gamma_f = 0.025$ and initial step sizes $t = \beta^0 = 0.05$. For the right column we have $\alpha = 0.005$, $\gamma_a = 0.5$, $\gamma_f = 0.025$ and initial step sizes $t = \beta^0 = 0.075$. In all reconstructions shown in this section we use the initial guess $a = 0$ in every pixel. This demonstration of non-unique recovery matches the theory about radial a and f given in [55] and section 2.5. In the right hand column we have a similar effect to that shown in Figure 6.4 where the value of the recovered f is smaller than that of the true f to compensate for the lower attenuation. It is worth noting that the vast majority of parameter choices which lead to a convergent solution produce reconstructions resembling the middle column. Numerically the ADMM algorithm we use based on [32] only guarantees convergence of the iterations to a critical point of the system and this is one possible explanation as to why we can obtain two convergent

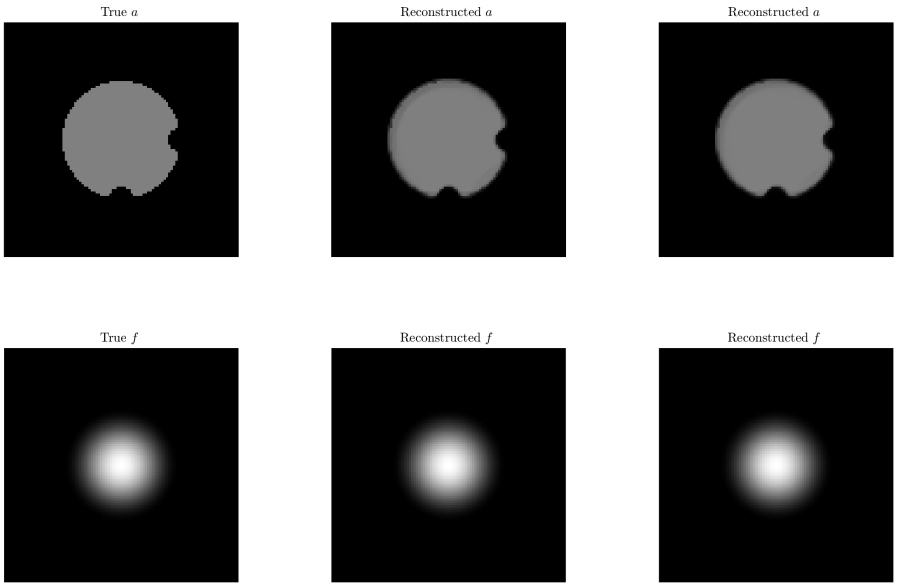
but distinct solutions.

Figure 6.11: Joint reconstruction of a and f where a is an ellipse and f is the same as in Figure 6.10. The admissible set here is $A = \{0, 50\}$.



Figures 6.11 and 6.12 show reconstructions where f is the same radial example as in Figure 6.10 but with a slightly perturbed. In Figure 6.11 a is an ellipse with major axis 1.1 and minor axis 1 and in Figure 6.12 two small circles have been removed from the edge of a. In both cases the middle and right column use the same parameters as the middle and right columns of Figure 6.10. For both modified phantoms we recover the same a and f and do not obtain a second solution. Furthermore, at least from the numerical testing we carried out, we were unable to choose convergent parameters which led to another solution being obtained.

Figure 6.12: Joint reconstruction of a and f where a is close to radial and f is the same as in Figure 6.10.The admissible set here is $A = \{0, 50\}$.



6.2 Comparison with another joint approach

This short section makes a visual comparison between another joint approach given in [25]. In [25] they aim to simultaneous recover an image and a segmentation from Magnetic Resonance Imaging (MRI) data, see [25] and their references. Here a segmentation refers to a colour map which identifies all points on the same object as one colour, i.e they obtain a purely multi-bang image for the segmentation. The authors of [25] make use an iterative regularization technique in contrast to the variational problem we solve given in (2.8) and physical factors mean that instead of under sampling the projections they under sample the k-space (Fourier Space). Both our approach and the one given in [25] make use of Total Variation but we also include multi-bang regularization. At the additional a priori cost of needing to know the admissible set $\mathcal{A}$ we can obtain an image which is very close to being purely multi-bang, even when restricted to limited projections (although the numerics given in this Chapter have roughly equal spacing rather than being focussed in some region). This leads to the production of an a image which is easy to segment by inspection. Performing the

reconstruction of an image from some physical data and then performing the segmentation on the reconstruction is one standard approach to segmentation, which makes this approach very susceptible to errors in the construction. One advantage that [25] has over the joint multi-bang TV regularizer we have used is that they do not need prior knowledge of the values to obtain a good reconstruction. As shown in Figure 6.1, which is a Discrete tomography example which uses the joint regularizer, an incorrect admissible set leads to bloating and then the multi-bang images we recover are poor and the resulting image cannot be segmented accurately. In practice the physical application to which we are applying the joint regularizer is important, if the admissible set is known very well then we recover good reconstructions and if the set is not so well known an approach such as the one given in [25] may be more suitable.

Chapter 7

Conclusions

Throughout this thesis we have considered recent work by [54, 19, 18, 37] relating to multi-bang regularization. There have been new results obtained both in theory with the nicely multi-bang SPECT uniqueness result as well as non nicely multi-bang partial recovery and a proposed joint reconstruction method making use of a variety of known algorithms.

A number of new theoretical results have been produced during this project. Originally presented in [37] we have shown unique recovery of a and f from $R_a f$ when a is nicely multi-bang. Partial recovery results for more general multi-bang a have also been obtained and some complete recovery results have been shown in special cases. More generally, we have shown that the unique recovery of the boundaries of multi-bang a is possible, provided that the support of f contains a sufficient amount of each boundary of a. The method of boundary recovery has potential uses in other fields, although it would take further research to examine the case where some of the boundaries are invisible, i.e when the lines tangential to the boundary do not intersect the support of f.

Another addition is the successful implementation of the joint reconstruction technique with relatively few projections. The results for DT are comparable to those in [63] in terms of the number of projections and number of rays used to obtain a "usable image". Extensive improvements to the coding of algorithms has seen a significant improvement in both reconstruction quality and speed. Much of the code uses parallel computing and so would gain a significant speed increase when using a GPU.

A complete list of inversion parameters is:

1. a^0, the initial guess for a.
2. α, the multi-bang penalty weight.
3. γ_a, the TV penalty weight on a.
4. γ_f, the TV penalty weight on f.
5. t, primal step size for the ADMM algorithm for a and f.
6. β^0, initial guess for dual step size in ADMM algorithm.
7. λ, as the initial Lagrange variable.
8. δ_1, the tolerance for difference in iterates to terminate the FISTA algorithm.
9. δ_2, the tolerance for difference in iterates to terminate the gradient descent algorithm for y.
10. δ_3, the tolerance for the primal residual r in the ADMM algorithm for a.
11. δ_4, the tolerance for the approximate dual residual s in the ADMM algorithm for a.
12. δ_5, the tolerance for difference in iterates in the ADMM algorithm for f.
13. δ_6, the tolerance for the difference in iterates of a and f to terminate joint inversion algorithm.

Due to the large number of inversion parameters, some work needs to be done in order to determine suitable bounds for good choices of inversion parameters. This could be done by the use of L-curves and finding methods to calculate the Lipschitz constant of the gradient of the objective function (since this provides optimal FISTA convergence [6]). Another option would be to use machine learning in order to try to learn the best parameters for convergence and this could be a potential avenue for further research.

Although the joint reconstruction algorithm obtains good results, there is room for improvement in the speed at which we can achieve a good reconstruction. The largest time sink in the joint reconstruction is the need to recreate the discrete attenuated

Radon transform matrix for each FISTA step in the inner iteration when recovering a. One calculation of this for 16 projections, which corresponds to 1600 rays, takes just under 2 seconds and we typically have to calculate this several hundred times per reconstruction. Any reduction in the time to calculate this sparse matrix would cause a huge time reduction for the overall algorithm.

Finally, we have yet to prove convergence in the non-smooth TV case even though it produces convergent results in our numerical investigations. From a proof of convergence given in [32], we have found the quantity $\mu^k - \beta^k$ which, if it remains positive, coincides with successful reconstructions and theoretical convergence. It is currently not yet known whether $\mu^k - \beta^k$ being positive for sufficiently large k is necessary for convergence. Proving convergence in this case where $\mu^k - \beta^k$ is potentially negative would have wide applications and is another possible avenue for further research.

Bibliography

[1] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality. *Mathematics of Operations Research*, 35(2):438–457, May 2010.

[2] H. Attouch, G. Buttazzo, and G. Michaille. *Variational Analysis in Sobolev and BV Spaces: Application to PDEs and Optimization*. SIAM, 2014.

[3] G. Bal and A. Jollivet. Combined source and attenuation reconstructions in SPECT. In *Tomography and Inverse Transport Theory*, volume 559 of *Contemporary Mathematics*, pages 13–28. American Mathematical Society, 2011.

[4] K. J. Batenburg and J. Sijbers. Dart: A practical reconstruction algorithm for discrete tomography. *IEEE Transactions on Image Processing*, 20(9):2542–2553, 2011.

[5] I. Bayram. On the convergence of the iterative shrinkage/thresholding algorithm with a weakly convex penalty. *IEEE Transactions on Signal Processing*, 64(6):1597–1608, 2016.

[6] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.

[7] B. L. Biondi. *3D Seismic Imaging*. Society of Exploration Geophysicists, 2006.

[8] J. Bolte, S. Sabach, and M. Teboulle. Nonconvex Lagrangian-based optimization: Monitoring schemes and global convergence. *Mathematics of Operations Research*, 43(4):1210–1232, Nov 2018.

[9] J. Bolte, S. Sabach, M. Teboulle, and Y. Vaisbourd. First order methods beyond convexity and Lipschitz gradient continuity with applications to quadratic inverse problems. *SIAM Journal on Optimization*, 28(3):2131–2151, 2018.

[10] S. Boyd. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2010.

[11] A. V . Bronnikov. Reconstruction of attenuation map using discrete consistency conditions. *IEEE Transactions on Nuclear Science*, 19(5):451–462, 2000.

[12] A. V. Bronnikov. Approximate reconstruction of attenuation map in spect imaging. *IEEE Transactions on Nuclear Science*, 42(5):1483–1488, 1995.

[13] A. V. Bronnikov. Numerical solution of the identification problem for the attenuated radon transform. *Inverse Problems*, 15(5):1315–1324, 1999.

[14] A. A. Bukhgeim and S. G. Kazantsev. Inversion formula for the fan-beam, attenuated Radon transform in a unit disk. *Russian Academy of Science Siberian Branch: The Sobolev Institute of Mathematics, Preprint No.99*, 2002.

[15] A. L. Bukhgeim. *Introduction to the Theory of Inverse Problems*. Number 19 in Inverse and Ill-Posed Problems. V.S.P. International Science Publishers, 2000.

[16] A. L. Bukhgeim. Inverse gravimetry approach to attenuated tomography. In *Tomography and Inverse Transport Theory*, volume 559 of *Contemporary Mathematics*. American Mathematical Society, 2011.

[17] G. Chen, J. Tang, and S. Leng. Prior image constrained compressed sensing (piccs): A method to accurately reconstruct dynamic ct images from highly undersampled projection data sets. *Medical Physics*, 35(2):660–663, Jan 2008.

[18] C. Clason, F. Kruse, and K. Kunisch. Total variation regularization of multimaterial topology optimization. *ESAIM: Mathematical Modelling and Numerical Analysis*, 52(1):275–303, Jan 2018.

[19] C. Clason and K. Kunisch. A convex analysis approach to multi-material topology optimization. *ESAIM: Mathematical Modelling and Numerical Analysis*, 50(6):1917–1936, Nov 2016.

[20] J. B. Conway. *Functions of One Complex Variable I.* Number 11 in Graduate Texts in Mathematics. Springer Verlag, 1973.

[21] M. Courdurier, F. Monard, A. Osses, and F. Romero. Simultaneous source and attenuation reconstruction in SPECT using ballistic and single scattering data. *Inverse Problems*, 31, 2015.

[22] E. Cueva, A. Osses, J. C. Quintana, C. Tejos, M. Courdurier, and P. Irarrazaval. Algebraic reconstruction of source and attenuation in SPECT using first scattering measurements. In B. Hofmann, A. Leitão, and J. Zubelli, editors, *New Trends in Parameter Identification for Mathematical Models*. Birkhäuser, Cham, 2018.

[23] D. Drusvyatskiy and C. Paquette. Efficiency of minimizing compositions of convex functions and smooth maps. *Mathematical programming*, 178(1):503–558, 2019.

[24] L. Ehrenpreis, P. Kuchment, and A. Pacnchenko. The exponential x-ray transform and fritz john's equation: I. range description. *Contemporary Mathematics*, 251(1):173–188, 2000.

[25] Corona V. et al. Enhancing joint reconstruction and segmentation with non-convex bregman iteration. *Inverse problems*, 35(5):778–783, 2019.

[26] A. S. Fokas, A. Iserles, and V. Marinakis. Reconstruction algorithms for positron emission tomography and single photon emission computed tomography and their numerical implementation. arXiv:physics/0412030[physics.med-ph], December 2004.

[27] A. S. Fokas, G. A. Kastis, and N. E. Protonotarios. *A new approach for the Inversion of the Attenuated Radon Transform.* Number 154 in Mathematical Analysis and Applications. Cham:Springer International Publishing, 2019.

[28] A. S. Fokas and R. G. Novikov. Discrete analogues of $\bar{\delta}$-equations and of radon transform. *C.RAcad. Sci. Paris Ser. I. Math*, 313(75), 1991.

[29] B. Frigyik, P. Stefanov, and G. Uhlmann. The X-ray transform for a generic family of curves and weights. *Journal of Geometric Analysis*, 18(89), 2008.

[30] D. Gourion and D. Noll. The inverse problem of emission tomography. *Inverse Problems*, 18(5):1435–1460, Sep 2002.

[31] D. Gourion, D. Noll, P. Gantet, A. Celler, and J.P. Esquerre. Attenuation correction using SPECT emission data only. *IEEE Transactions on Nuclear Science*, 49(5):2172–2179, 2002.

[32] K. Guo, D. R. Han, and T. T. Wu. Convergence of alternating direction method for minimizing sum of two nonconvex functions with linear constraints. *International Journal of Computer Mathematics*, 94(8):1653–1669, Sep 2016.

[33] K. Hämäläinen, L. Harhanen, A. Kallonen, A. Kujanpää, E. Niemi, and S. Siltanen. Tomographic x-ray data of a walnut. arXiv:1502.04064, February 2015.

[34] G. T Herman and Kuba. A. *Advances in Discrete Tomography and its Applications*. Birkhäuser Boston, 2007.

[35] G.T. Herman. *Fundamentals of Computerized Tomography*. Springer London, 2009.

[36] S. Holman, F. Monard, and P. Stefanov. The attenuated geodesic x-ray transform. *Inverse Problems*, 34(6), 2018.

[37] S. Holman and P. Richardson. Emission tomography with a multi-bang assumption on attenuation. *Inverse Problems*, 36(12), 2020.

[38] Q Huang, G. L. Zeng, and J. Wu. An alternative proof of Bukhgeim and Kazantsev's inversion formula for attenuated fan-beam projections. *Medical Physics*, 33(11):3983–3987, Oct 2006.

[39] F. Kharfi. *Mathematics and Physics of Computed Tomography (CT): Demonstrations and Practical Examples*. IntechOpen, 2013.

[40] V. P. Krishnan, R. K. Mishra, and F. Monard. On solenoidal-injective and injective ray transforms of tensor fields on surfaces. *Journal of Inverse and Ill-Posed Problems*, 4(27):527–538, 2019.

[41] K. Kurdyka and K. Parusiński. w_f-stratification of subanalytic functions and the łojasiewicz inequality. *Comptes rendus de I'Académie des Sciences Paris*, 318:129–133, 1994.

[42] S. H. Manglos, R. J. Jaszczak, and C. E. Floyd. Weighted backprojection implemented with a non-uniform attenuation map for improved spect quantization. *IEEE Transactions on Nuclear Science*, 35(1):625–628, 1988.

[43] F. Monard. Efficient tensor tomography in fan-beam coordinates. ii: Attenuated transforms. *Inverse Problems and Imaging*, 12(2):433–460, 2018.

[44] F. Natterer. The identification problem in emission computed tomography. *In Mathematical Aspects of Computerized Tomography*, 1(1):45–56, 1981.

[45] F. Natterer. *The Mathematics of Computerized Tomography*. Society for Industrial and Applied Mathematics, 2001.

[46] D. Needell and R. Ward. Stable image reconstruction using total variation minimization. *SIAM Journal on Imaging Sciences*, 6(2):1035–1058, 2013.

[47] S. Niu, Y. Gao, Z. Bian, J. Huang, W. Chen, G. Yu, Z. Liang, and J. Ma. Sparse-view x-ray CT reconstruction via total generalized variation regularization. *Physics in Medicine and Biology*, 59(12):2997–3017, may 2014.

[48] R. G. Novikov. An inversion formula for the attenuated X-ray transformation. *Arkiv för Matematik*, 40(1):145–167, 2002.

[49] R. G. Novikov. On the range characterization for the two-dimensional attenuated X-ray transformation. *Inverse Problems*, 18(3):677–700, Apr 2002.

[50] R. G. Novikov. On the range characterization for the two-dimensional attenuated x-ray transformation. *Inverse Problems*, 18, 2002.

[51] J. Plemelj. *Problems in the sense of Riemann and Klein*. New York: Interscience Publishers, 1964.

[52] D. Pompeiu. Sur la continuité des fonctions de variables complexes. *Annales de la Faculté des Sciences de Toulouse*, 7(2):265–315, 1905.

[53] E.T Quinto. The invertibility of rotation invariant radon transforms. *Journal of Mathematical Analysis and Applications*, 91(2):510 – 522, 1983.

[54] P. Richardson. Applications of multi-bang regularisation in passive radiation emission tomography. Master's thesis, University of Manchester, 2017.

[55] D. C. Solmon. The identification problem for the exponential Radon transform. *Mathematical Methods in the Applied Sciences*, 18(9):687–695, Jul 1995.

[56] P. Stefanov. The identification problem for the attenuated X-ray transform. *American Journal of Mathematics*, 136(5):1215–1247, 2014.

[57] O. J Tretiak and C. E Metz. The exponential radon transform. *SIAM J. Appl. Math*, 39(1):341–354, 1980.

[58] B. M. W. Tsui, H. Hu, and D. R. Gilland. Implementation of simultaneous attenuation and detector response correction in SPECT. *IEEE Transactions on Nuclear Science*, 35(1):778–783, 1988.

[59] K. Vetter, R. Barnowksi, A. Haefner, T.H.Y Joshi, R. Pavlovsky, and B. J Quiter. Gamma-ray imaging for nuclear security and safety: Towards 3-d gamma-ray vision. *Nuclear Inst. and Methods in Physics Research*, 10(878), 2018.

[60] A. Welch, R. Clack, F. Natterer, and G. Gullberg. Toward accurate attenuation correction in SPECT without transmission measurements. *IEEE Transactions on Medical Imaging*, 16(5):532–541, 1997.

[61] M.N. Wernick and J.N. Aarsvold. *Emission Tomography: The Fundamentals of PET and SPECT*. Elsevier Science, 2004.

[62] Z. Zhu, K. A. Wahid, and P. Babyn. Ct image reconstruction from partial angular measurements via compressed sensing. In *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–4, 2012.

[63] X. Zhuge, W. J. Palenstijn, and K. J. Batenburg. TVR-DART: A more robust algorithm for discrete tomography from limited projection data with automated gray value estimation. *IEEE Transactions on Image Processing*, 25(1):455–468, 2016.