POINT SPREAD FUNCTION ESTIMATION AND UNCERTAINTY QUANTIFICATION

By

Kevin Thomas Joyce

B.S., Montana State University, Bozeman, MT, 2006
M.S., Montana State University, Bozeman, MT, 2009

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctorate of Philosophy
in Mathematics

The University of Montana
Missoula, MT

May 2016

Approved by:

Sandy Ross, Associate Dean of the Graduate School
Graduate School

Dr. Johnathan Bardsley, Co-Chair
Mathematical Sciences

Dr. Aaron Luttman, Co-Chair
Signals Processing and Applied Mathematics
National Security Technologies, LLC, Las Vegas, NV

Dr. Jon Graham
Mathematical Sciences

Dr. Leonid Kalechev
Mathematical Sciences

Joyce, Kevin, Doctorate of Philosophy, May 2016 Mathematics

Point Spread Function Estimation and Uncertainty Quantification

Committee Chair: Johnathan Bardsley, Ph.D.

Acknowledgments

Notations

Contents

List of Tables

List of Figures

Chapter 1

Images and Blur

1.1 Introduction

In addition to being a rich source of artistic and creative value, images (or more precicely, visual data from projections of light) are an important source of scientific information. Even the word "observation" generally connotes the visual perception, and its use as a catch-all for the measured verification of a hypothesis exemplifies the central role of vision in science. Many important scientific results have used visual data to discover and explain natural phenomena; for example, visual observations such as the color and shape of various plant organs in Gregor Mendel's famous experiments on hybridized peas formed the primary source of data for developing his famous model for genetic inheritance [Magner, 2002]. Arthur Eddington's famous 1919 image of the gravitationally lensed path of a comet during a solar eclipse [Dyson et al., 1920] provided the first experimental evidence supporting Albert Einstein's general theory of relativity. Yet, in these cases, only the *qualitative* components of the visual response and their relationship to the experiment were relevant. With the advent of the camera, photosensitive chemistry, and later digital imaging technology, high-fidelity recording of visual observations

became possible, allowing the potential to *quantitatively* analyze visual information. The ever-progressing technology in optical science and engineering are rapidly increasing the amount of data that can be measured in an image; yet, the extreme quantity and spatial organization that comes with high resolution data makes a rigorous quantitative analysis quite challenging.

Images, when viewed quantitatively, can be described as the response of the incidence of light, and the subsequent exchange of energy, on a grid of regularly spaced grid elements, which we refer to as pixels. The domain of interest for imaging systems can vary greatly in both spatial dimension and the spectrum of light captured. The energy response of light is spectrum dependent, and typically is analyzed at fixed bands of frequency with very high spatial resolution. For example, astronomical images captured by the interstellar robotic probes Voyagers I and II measured 5 bands in the visible spetrum on a pixel grid with dimension 800×800 [Showalter et al., 2006]. In medical imaging, computed tomography (CT) is an imaging process where a series of axial measurements of attenuation of electromagnetic radiation are used to reconstruct a cross-section of a scanned object [Epstein, 2008]. The primary focus in this work are pulsed X-ray measurements, referred to as radiographs, that are used as an experimental diagnostic of high-energy physics experiments. In this case, X-rays are pulsed through an experiment, then the attenuated X-rays excite a crystal that responds by luminescing visible light at an intensity related to the energy of the attenuated wave-front. The light is then measured on a high resolution array (on the order of 1000×1000 or more) of carge coupled devices (CCD) calibrated to count photons at a specified spectral band.

Having established the form of image data, i.e. *what* the data is, *how* does one analyze it? Although images often provide a large volume of data, their spatial nature make measurements at each pixel highly dependent on measurments of adjacent pixels. A quantitative analysis of the image cannot assume that measured values are independent, because it is precisely this lack of independence that makes an image interesting – independent image data is *white noise* (perhaps more appropriately, "gray noise") from which one can only infer the average of the measured pixels.

From a statistical point of view, the field of spatial statistics considers broadly the analysis of data that is structured and spatially correlated in this way. This field has had much development over the past half-century and has inspired a wealth of theory and computational tools, but is far from complete. Moreover, a broad field of scientific disciplines have considered image data, or more generally spatially correlated data, in one way or another; fields such as astronomy, astrophysics, biology, medicine, geology, computer science, and nuclear physics to name a few. The book [Cressie, 1993] provides an excellent overview of the history and current methods for statistical methods for spatial data. Although much work has been done, it is still a very active research area and is far from the level of consensus and understanding that analysis of independently sampled data has achieved. The aim of this work is to develop and adapt current models and methods for estimation and quantifying uncertainty to a small component of image analysis related specifically to the *system for capturing images*. Understanding this component is an important preliminary step to developing methods for quantitatively analyzing the images themselves.

1.1.1 Organization

The following sections outline the modeling of image blur via the *point spread function*, the primary quantity of interest in this work. In particular, we will introduce a model for the blurring of an opaque edge. The model can be cast in various mathematical forms without changing the physical assumptions. We will exploit these different mathematical formulations (e.g. by changing variables) to show that a blurred edge is sufficient for estimation and to show that the problem is "ill-posed".

In the next chapter, we will derive the necessary theory and technical definitions so that the problem is well-defined on a Hilbert space. Chapter 2 will be mainly theoretical, but the explicit forms for the discrete model of the forward operator and prior information are motivated and derived there. Chapter 3 will give details on how to carry out the estimation

on a computer. There, we will deal with how to discretely represent each of the necessary components in the estimation problem. We will also motivate and present the design of a detailed algorithm for carrying out statistical estimation. Finally, in Chapter 4, we present the results of an implementation on synthetically derived data and on measured data from a high-energy X-ray imaging system at the U.S. Department of Energy's Nevada National Security Site. We will end with a discussion of conclusions and possible future work.

1.2 Modeling blur with a PSF

One major component of the spatial relationship of neighboring pixels of an image is due to blur from the imaging instrumentation. That is, under the assumption that arbitrary images are consistently measured by the modeled system, what contribution does this system have on how pixels are related, and how can we quantify this relationship? A widely used model for blurring [Hansen, 2010, Jain, 1989, Vogel, 2002, Epstein, 2008] expresses this relationship as a linear filter that maps the ideal image f to b by integrating

$$b(x,y) = \iint_{\mathbb{R}^2} k(x,y;s,t) f(s,t)\, dsdt, \tag{1.1}$$

where $b(x,y)$ represents the intensity of the blurred image at (x,y); $f(s,t)$ represents the intensity of the ideal un-blurred image at (s,t); and k is the kernel of the filter, which characterizes the blurring process. Informally, we can view the effect of blur point-wise by observing the system response of a "point-source" at $(\overline{x},\overline{y})$ (formally, take $f = \delta_{\overline{x},\overline{y}}$, Dirac's delta translated to $(\overline{x},\overline{y})$), then $b(x,y) = k(x,y;\overline{x},\overline{y})$ represents the "spread" at the point source. The function k is referred to as the *point spread function* (PSF) of the system at $(\overline{x},\overline{y})$. When the effect of blurring does not depend on the location of this point, that is, translating f by $(\overline{x},\overline{y})$

results in b translated by $(\overline{x}, \overline{y})$, we say that it is *spatially invariant*. In (1.1), this means

$$\begin{aligned} b(x-\overline{x}, y-\overline{y}) &= \iint_{\mathbb{R}^2} k(x,y;s,t) f(s-\overline{x}, t-\overline{y}) ds dt \\ &= \iint_{\mathbb{R}^2} k(x,y;s'+\overline{x}, t'+\overline{y}) f(s',t') ds' dt'. \end{aligned} \tag{1.2}$$

On the other hand

$$b(x-\overline{x}, y-\overline{y}) = \iint_{\mathbb{R}^2} k(x-\overline{x}, y-\overline{y}, s, t) f(s,t) ds dt. \tag{1.3}$$

Since (1.2) and (1.3) hold for all f, we have for each $x, y, \overline{x}, \overline{y}, s$, and t,

$$k(x,y;s+\overline{x}, t+\overline{y}) = k(x-\overline{x}; y-\overline{y}, s, t) \tag{1.4}$$

and in particular when we fix $(s,t) = (0,0)$,

$$k(x,y;\overline{x},\overline{y}) = k(x-\overline{x}, y-\overline{y}; 0, 0). \tag{1.5}$$

Let us denote $k(x,y;\overline{x},\overline{y}) = k(x-\overline{x}, y-\overline{y})$, then the linear filter in (1.1) reduces to

$$b(x,y) = \iint_{\mathbb{R}^2} k(x-s, y-t) f(s,t)\, ds dt. \tag{1.6}$$

Equation (1.6) is called the *convolution* of f by k. In fact, a general result about arbitrary linear filters defined on L^p spaces states that the action of any linear filter can be expressed through convolution with some generalized function k [Grafakos, 2014]. In any case, when blur is assumed to be spatial invariant, it results in solving the convolution equation (1.6). Mathematical methods that estimate f given b and k are referred to as *deconvolution* techniques.

Note that a change of variables by $s' = x - s$ and $t' = y - t$ results in a convolution of k by

f, which is to say that convolution, as an operation, is symmetric. That is

$$b(x,y) = \iint_{\mathbb{R}^2} k(s,t) f(x-s, y-t)\, dsdt. \tag{1.7}$$

This dual relationship between the PSF and the image will allow us to use the framework and many of the tools of deconvolution for the problem of PSF estimation.

Typically, deconvolution methods assume that the form of the PSF can be accurately described by modeling the imaging system [Jain, 1989, Hansen, 2010], but for complex imaging systems such as the one described for X-ray radiography, this is not realistic. Instead, if the imaging system is designed so that repeated images can be taken under consistent conditions, then by convolution symmetry in (1.7), the blurring of a *known calibration image* can be cast as deconvolving the PSF from the ideal f corresponding to the known image.

Recall that the PSF models the blurring response of a single point. So, a direct estimate of k can be obtained by imaging a bright point-source, which approximates the impulse response to (1.7). In astronomical imaging, the point-source can be a bright distant star, or in a controlled setting where visible light is measured, a focused laser beam provides a good point-source estimate. However, in the spectral regime of X-rays, focusing the high-frequency light is notoriously difficult and usually is impractical in situations of interest, so a point-source estimate of the PSF is usually unavailable at these frequencies. Instead, the system response of a uniformly opaque calibration object with a simple geometry can be measured. Under the assumption that the object is sufficiently thick so that X-rays are completely attenuated on the profile of the object, then f is given by an indicator function on a set $E \subseteq \mathbb{R}^2$ determined by the object's profile. Calibration objects typically have simple geometry and reduce the complexity of solving the deconvolution problem in (1.7). For example, the object could be a circular aperture or two perpendicular edges aligned with the imaging plane [Doering et al., 1992, Watson, 1993]. An additional assumption of *radial symmetry* on k, will allow for a very simple calibration object – an edge. That is, if the calibration object completely

attenuates X-rays along a vertical edge at the fixed location at $s = 0$ in the imaging plane, then $E = \{(s,t) : s \geq 0\}$ and $f(s,t) \stackrel{\text{def}}{=} f_E(s) = 1$ if $s \geq 0$ and 0 if $s < 0$; see Figures 1.1 and 1.2 for a schematic of the calibration object and measurement system and an example of recorded intensity data. The model for blur in (1.7) reduces to

$$b(x,y) = \iint_{\mathbb{R}^2} k(s,t) f_E(x-s) ds dt. \tag{1.8}$$

Note that b does not depend on y in (1.8), so denoting $b(x,0) = b(x)$, (1.8) reduces to

$$b(x) = \iint_{\mathbb{R}^2} k(s,t) f_E(x-s) ds dt. \tag{1.9}$$

In general, estimating k from b in (1.9) is underdetermined, since there are many distinct k that can result in the same output b. To see this, note that

$$\int_{-\infty}^{\infty} t e^{-t^2-s^2} dt = 0 \tag{1.10}$$

for all s since the integrand is odd in t. So given any solution k, $k(s,t) + te^{-t^2-s^2}$ is also a solution but is *not* radially symmetric. We will see in the next section that the assumption of radial symmetry on k is sufficient for a well-defined analytic solution to (1.9).

In summary, we've assumed that the effect of blur is modelled by a spatially invariant linear filter with a radially symmetric kernel, and (1.9) describes the blurring of a calibration object whose profile is an edge. Solving the integral equation in (1.9) will be the primary focus of this work.

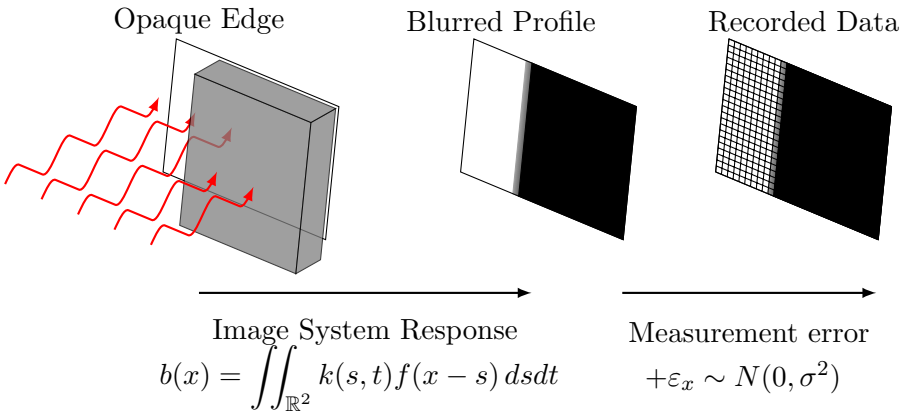

Figure 1.1: A schematic of the measurement model for an X-Ray image of an edge. An opaque block aligned with the imaging plane blocks light on the half plane to produce a blurred edge.

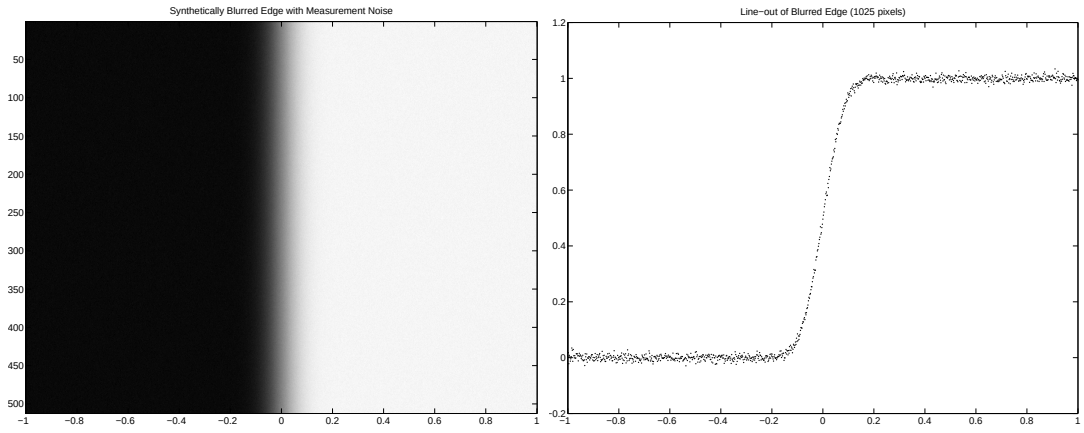

Figure 1.2: A synthetically blurred edge with simulated measurement error and a line-out (horizontal cross-section) from the data.

1.3 The Abel transform and a deterministic solution

In this section, we will show that radial symmetry is sufficient to guarantee a unique analytic answer. However, the suggested analytic method will be problematic as will be seen in the following section.

A large part of designing a system for imaging is to minimize the effect of blur. Often, limitations due to physical laws put a lower bound on the measurement precision so that even an optimal design cannot ignore the effect of blur. Although arbitrary resolution may be impossible, engineering effort can still optimize accuracy, or equivalently, minimize bias. Hence, an optimally designed imaging system should exhibit isotropic blur (an absence of direction bias), so that, in the convolution model, the PSF is radial symmetric. In fact, many parametrically modeled PSFs assume radially symmetry [Doering et al., 1992, Jain, 1989, Kundur and Hatzinakos, 1996, Watson, 1993].

When one assumes that the PSF of their system is radially symmetric, then it has a unique one-dimensional representation; that is, there exists a function p defined on $[0, \infty)$ so that $k(s,t) = p\left(\sqrt{s^2+t^2}\right)$. The function p is referred to as the radial profile of k. Viewing (1.9) as iterated integration first in t allows $f_E(s-x)$ to be factored out of the inner integral. Then substituting p for k, the inner integral has the form

$$\ell(s) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} p\left(\sqrt{s^2+t^2}\right)\, dt. \tag{1.11}$$

The function $\ell(s)$ is the integration along a line perpendicular to the edge E, and its form is commonly encountered in other imaging applications with radial geometry, such as tomographic imaging science. The transformation that takes p to ℓ is known as the *Abel transform*, and for its study in imaging science, see [Bracewell, 1965, Epstein, 2008, Knill et al., 1993].

Substituting (1.11) for k into (1.9) and changing the bounds of integration in s according

to $f_E(x-s)$, results in

$$\begin{aligned} b(x) &= \int_{-\infty}^{\infty} f_E(x-s)\left(\int_{-\infty}^{\infty} p\left(\sqrt{s^2+t^2}\right)\,dt\right)ds \\ &= \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} p\left(\sqrt{s^2+t^2}\right)\,dt\right)ds \\ &= \int_{-\infty}^{x} \ell(s)ds. \end{aligned} \tag{1.12}$$

Observe that b exhibits point symmetry about $b(0)$. To see this analytically, note that when $x > 0$

$$\begin{aligned} b(x) - \int_{-\infty}^{0} \ell(s)ds &= \int_{0}^{x} \ell(s)ds \\ &= \int_{0}^{-x} \ell(s')ds' \\ &= b(-x) - b(0). \end{aligned} \tag{1.13}$$

Since the function $\widetilde{b}(x) = b(x) - b(0)$ is odd, $b(x)$ has reflection symmetry about $(0, b(0))$. Hence, data defined on either $x \in (-\infty, 0]$ or $x \in [0, \infty)$ should be sufficient for estimating p. This observation will be important in the next section.

The Abel transform has an explicit expression for its inverse [Epstein, 2008] given by

$$p(r) = -\frac{1}{\pi r}\frac{d}{dr}\left(\int_r^{\infty} \frac{\ell(s)sds}{(s^2-r^2)^{1/2}}\right). \tag{1.14}$$

The following calculations verify (1.14).

Proposition 1.3.1. *Suppose that $p(r)$ is such $\lim_{r\to\infty} rp(r) = 0$ and $\ell(s)$ in (1.11) is pointwise defined and the integral in (1.14) is finite for each r, then equation (1.14) holds.*

Proof. We can express the inner integral in (1.12) as

$$\ell(s) = 2\int_{|s|}^{\infty} \frac{p(t)t}{(t^2 - s^2)^{1/2}}\, dt \tag{1.15}$$

by symmetry of the integrand (it is even) and a change of variable by $r = s^2 + t^2$. Now, interchanging the order of integration in (1.14) results in

$$\begin{aligned}\left(\int_r^{\infty} \frac{\ell(s)sds}{(s^2 - r^2)^{1/2}}\right) &= \int_r^{\infty}\int_s^{\infty} \frac{2p(t)ts}{(s^2 - r^2)^{1/2}(t^2 - s^2)^{1/2}}\, dtds \\ &= \int_r^{\infty} p(t)t \int_r^{t} \frac{2s}{(s^2 - r^2)^{1/2}(t^2 - s^2)^{1/2}}\, dsdt. \end{aligned} \tag{1.16}$$

In the second step, we have interchanged variables and the integral's support can be expressed

$$\{(t, s) : r \le s, s \le t\} = \{(s, t) : r \le s \le t, r \le t\}. \tag{1.17}$$

Another change of variables by $s^2 = \tau t^2 + (1 - \tau)r^2$ (note $s \ge 0$) results in $2sds = (t^2 - r^2)d\tau$, so that the inner integral in (1.16) is

$$\int_r^{t} \frac{2s}{(s^2 - r^2)^{1/2}(t^2 - s^2)^{1/2}}\, ds = \int_0^1 \frac{1}{\tau^{1/2}(1 - \tau)^{1/2}}\, d\tau. \tag{1.18}$$

Note that the resulting integral is independent of both r and t, and hence, is constant. To evaluate it, recall the Gamma function identity

$$\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} = \int_0^1 \tau^{-\alpha}(1 - \tau)^{\alpha - 1}\, d\tau, \tag{1.19}$$

from which the expression in (1.18) reduces to $\Gamma(1/2)^2 = \pi$. Collecting these results and applying the fundamental theorem of calculus with the assumption that $\lim_{r\to\infty} rp(r) = 0$ to

(1.16) implies

$$
\begin{aligned}
-\frac{\pi}{r}\frac{d}{dr}\left(\int_r^\infty \frac{\ell(s)sds}{(s^2-r^2)^{1/2}}\right) &= -\frac{\pi}{r}\frac{d}{dr}\int_r^\infty p(t)t\pi dt \\
&= \frac{1}{r}\left(p(r)r - \lim_{T\to\infty} p(T)T\right) \\
&= p(r), \qquad (1.20)
\end{aligned}
$$

which proves the identity in (1.14). □

With this result, p can be analytically recovered from b in (1.12) as follows; given $b(x)$, the fundamental theorem of calculus gives $\ell(x) + \lim_{X\to\infty} \ell(X)$ by differentiating $b(x)$. Since $\lim_{r\to\infty} tp(t) = 0$ symmetry and the change of variables in (1.15) implies $\lim_{X\to-\infty} \ell(X) = 0$. Then, applying the inversion formula in (1.14) to $b'(x)$ gives the radial profile $p(r)$. Hence, the assumption of radial symmetry sufficiently constrains the problem to uniquely determine the PSF from an edge calibration object illustrated in Figure 1.2.

In theory, we have outlined a solution to the problem, but there is one more component to the noise-free model that has not been addressed – random effects due to measurement error – for which a direct application of outlined method on measured data will fail spectacularly, due to the estimation problem being "ill-posed" which we address in the next section.

1.4 PSF reconstruction as an ill-posed inverse problem

The analytic solution outlined in the last section will not be sufficient when measurement errors are introduced. One way to see that the solution method outlined will be insufficient is that it requires taking derivatives of measured data, which is known to be problematic [Hanke and Scherzer, 2001]. In this section, we will return to (1.9), and perform a different variable transformation to explicitly illustrate the instability, and in doing so, will derive a form that

is more suitable for analysis and numerical discretization.

The measurements of the imaging system are generally not deterministic and are subject to measurement noise. Precisely modeling the stochastic effect of measurement error is system dependent and can be quite complicated. In the X-ray radiography example, uncertainty can enter into the system at the luminescing crystal response, at the counts of CCD array, or through the electrical transmission of the signal. In order to be broadly applicable, and appealing generally to various central-limit-theorem-like results in probability [Durrett, 2010], we model the stochastic measurement effect in aggregate as an additive, independent Gaussian noise process with zero mean and unknown variance. For now, this assumption can be viewed as a small perturbation from the model, but its form will be important for the inference techniques developed in subsequent chapters.

Estimating a quantity of interest, in our case k, from indirect and noisy measurements, b, with a model where an operator takes k *to* b (referred to as the *forward operator*) is called an inverse problem. The problem is called well-posed when the forward operator is invertible, and the inverse is continuous. These famous conditions were laid in the early 20th century in [Hadamard, 1902], but a number of important applications have arisen (among those computational imaging) where these conditions are violated, enough to the extent that the term "inverse problems", as it refers to the mathematical research area, is exclusively devoted to solving these *ill-posed* problems. In particular, most cases of interest exhibit a model where the inverse of the forward operator is discontinuous.

The discussion thus far for PSF reconstruction has been somewhat informal, as we have not defined a space for the PSF or its radial representation, so we have not formally defined the forward operator of the model. Defining these spaces in detail is technical and will be addressed in in Chapter 2; however, assuming these spaces have been defined, we can illustrate that the problem of reconstructing the radial profile is ill-posed.

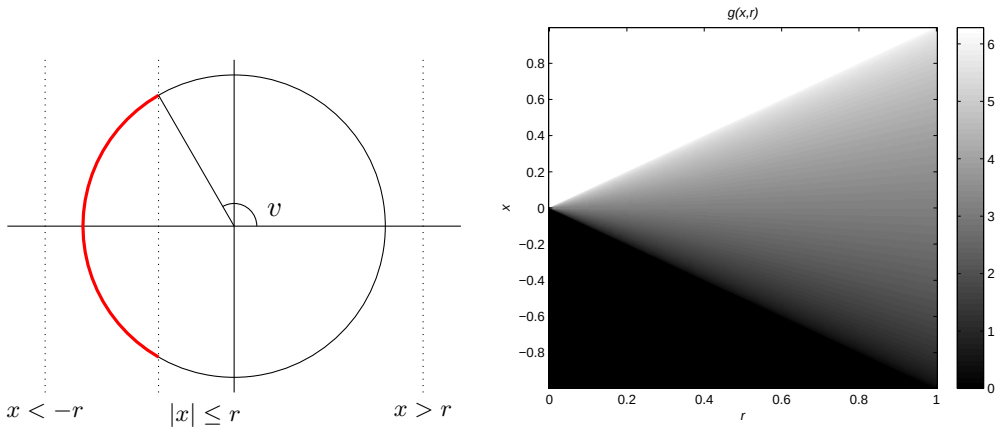

Figure 1.3: The PSF forward integral operator kernel $g(x, r)$ represented as the arc measure of v in $(-\pi, \pi)$ where $x \geq r \cos v$.

Returning to (1.9), a variable transformation by $(s, t) = T(r, v) = (r \cos v, r \sin v)$, has $|dT(r, v)| = r$ and

$$\begin{aligned} b(x) &= \int_0^\infty p(r) \left(\int_{-\pi}^{\pi} f_E(x - r\cos v) dv \right) r dr \\ &= \int_0^\infty p(r) g(x, r) r dr, \end{aligned} \tag{1.21}$$

where

$$g(x, r) \stackrel{\text{def}}{=} \begin{cases} 0 & x < -r \\ 2(\pi - \operatorname{acos}(x/r)) & |x| < r \\ 2\pi & x > r \end{cases} . \tag{1.22}$$

To see that g has this form, note that integrating $f_E(x - r\cos v)$ is the radian measure of the set $\{v \in (-\pi, \pi) : r \cos v \leq x\}$; see Figure 1.3.

There are three key observations to make. From this viewpoint, the forward model is now a one-dimensional integral equation on the radial profile as opposed to the two-dimensional problem in (1.9). Second, note that $g(x, r)$ is continuous (although it has a discontinuity in its partial derivatives across $r = s$). Finally, recall that the graph of $b(x)$ exhibits reflection

symmetry about $b(0)$. So, $b(x)$ defined on either $(-\infty, 0]$ or $[0, \infty)$ completely determines p.

With these observations, we can define $G : \mathcal{H}_1 \to \mathcal{H}_2$ is an operator between closed isometric subspaces of $L^2([0, \infty))$ that acts by the integral equation in (1.21), i.e.,

$$[Gp](x) = \int_0^\infty p(r)g(x, r)rdr. \tag{1.23}$$

Hence, G is a compact Hilbert-Schmidt operator since g is continuous and $\mathcal{H}_1$ and $\mathcal{H}_2$ are seperable Hilbert spaces. Moreover, G is an injective operator since we showed that $b = Gp$ has an analytic solution. The spectral theorem for such operators implies that G has a countable spectrum which has zero as a limit point, and hence, its inverse G^{-1} is unbounded. See one of many texts on functional analysis [Bachman and Narici, 1966, Rudin, 1991] and [Tikhonov, 1963, Vogel, 2002, Morozov and Stessin, 1993] for the analytic treatment of ill-posed problems.

Solving ill-posed inverse problems requires *regularization* of the unbounded inverse. Recall that the original formulation of the problem is cast in terms of deconvolution, and much of the literature of inverse problems is devoted to this subject. This work draws heavily from techniques for that purpose. In particular, we will take a Bayesian approach to the inverse problem so that, in addition to estimating k, uncertainties in the estimate can be quantified by analysing the so-called *posterior distribution*. These methods have been the subject of much recent research (see the books [Calvetti and Somersalo, 2007, Kaipio and Somersalo, 2005, Stuart, 2010]), and the problem of PSF reconstruction fits neatly into that framework once the spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ have been defined, which we address in the next chapter.

Chapter 2

Reconstruction on the Continuum

Chapter 3

Computational Methods

In the last chapter, we established the theory for defining the inverse problem in an infinte dimensional Hilbert space. Of course, to carry out numerical estimation, the data and the estimate for the PSF must be represented by a fintite set of numbers on computer. The discrete representations will be point-based on equally spaced grids, and each operator defined in the Chapter 2 will be estimated using either numerical quadrature or finite differencing. Hence, to each operator we will define a matrix that defines their action on the point-wise estimates, and they will be derived in Section 3.1 We then develop the *discrete* probability spaces associated with the matrix-operators defined in Section 3.1, which will serve as our discrete approximation of the infinite dimensional space defined in Chapter 2. In this development, we will add "uninformative" prior assumptions for the parameters λ and δ, forming a hierarchical Bayesian model. From there, the discrete posterior distribution can be expressed in terms of conditional distributions in such a way so that Markov Chain Monte Carlo sampling techniques (e.g. Gibbs' sampling) can be applied to provide estimates and quantification of uncertainty. In Section 3.2, we will derive a standard Gibbs' sampling algorithm Geman and Geman [1984] and improve upon it using a technique called *partial collapse*, which can motivated by several recent theoretical and practial analyses [Van Dyk and Park, 2008, Agapiou et al., 2014, Fox

and Norton, 2015]. We will also briefly review standard convergence diagnostics for comparing Markov Chain based sampling algorithms, which will show that our adapted alorithm is indeed an enhancement of standard Gibbs' sampling.

3.1 Discrete representation of the inference problem

Recall, the radial Laplacian $R : \mathcal{H}_1 \to \mathcal{H}_1$ by

$$[Rp](r) = r^{-1} \frac{d}{dr} \left(r \frac{d}{dr} p(r) \right). \tag{3.1}$$

Recall also that the inner product for the regularization is also affected by the change of variables, i.e.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k(s', t') \, (\Delta^n k)(s', t') \, ds' dt' \quad = 2\pi \int_0^{\infty} x(r)(R^n x)(r) \, r^{1-n} dr, \tag{3.2}$$

so the Laplacian regularization of order n smoothness on k induces a regularization operator on x of the form $r^{1-n} R^n$. We take $n = 2$, which in the probabilistic framework, guarantees that the corresponding prior covariance operator on k is trace class **?**.

For numerical estimation, we discretize the forward integral operator in (**??**) using midpoint quadrature, and, imposing the same resolution on the reconstructed radial profile as the data results in $\mathbf{A}$ being an $N \times (N-1)/2 = N \times M$ matrix. The differential operator R is discretized using centered differencing, i.e. for $h = 1/M$ denote $r_{j \pm 1/2} = h \cdot j \pm h/2$, $x_j = x(r_j)$ and

$$[\boldsymbol{R}\boldsymbol{x}]_j \stackrel{\text{def}}{=} \frac{1}{h^2} \Big(r_{j+1/2}(x_{j+1} - x_j) - r_{j-1/2}(x_j - x_{j-1}) \Big). \tag{3.3}$$

or as a matrix stencil

$$\frac{1}{h^2}\begin{bmatrix} -(r_{j-3/2}+r_{j-1/2}) & r_{j-1/2} & 0 \\ r_{j-1/2} & -(r_{j-1/2}+r_{j+1/2}) & r_{j+1/2} \\ 0 & r_{j+1/2} & -(r_{j+1/2}+r_{j+3/2}) \end{bmatrix}\begin{bmatrix} x_{j-1} \\ x_j \\ x_{j+1} \end{bmatrix}. \tag{3.4}$$

The discretization of $\boldsymbol{R}$ has a zero right boundary condition since $\lim_{r\to\infty} x(r) = 0$ (we assume the domain for the radial profile is sufficiently large so that r_M is less than machine epsilon) and a reflective left boundary condition because of the assumption of radial symmetry. We then take $\boldsymbol{L} = \boldsymbol{r}^{-1} \odot \boldsymbol{R}^2$, i.e., the coordinate wise multiplication of reciprocals of the grid points r_j composed with $\boldsymbol{R}^2$. Finally, the discretization of the transformed inverse problem is

$$\boldsymbol{b} = \mathbf{A}x + \boldsymbol{\varepsilon}, \tag{3.5}$$

with $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \lambda^{-1}\boldsymbol{I})$ and $\boldsymbol{x} \sim \mathcal{N}(\mathbf{0}, \delta^{-1}(\boldsymbol{L}^T\boldsymbol{L})^{-1})$.

3.1.1 The discrete hierarchical Bayesian model

3.1.2 The discrete posterior distribution

3.2 Markov Chain Monte Carlo estimation

3.2.1 Preliminaries

3.2.2 Gibbs sampling

3.2.3 The partially collapsed Gibbs sampler

3.3 Evaluating Convergence

3.3.1 Theoretical justification for partial collapse

3.3.2 Estimating Burn-in

3.3.3 Integrated Autocorrelation

Chapter 4

Results

Bibliography

Sergios Agapiou, Johnathan M Bardsley, Omiros Papaspiliopoulos, and Andrew M Stuart. Analysis of the gibbs sampler for hierarchical inverse problems. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):511–544, 2014.

George Bachman and Lawrence Narici. *Functional Analysis*. Academic Press, 1966.

R. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, 1965.

Daniela Calvetti and Erkki Somersalo. *An Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing*, volume 2. Springer Science & Business Media, 2007.

Noel AC Cressie. *Statistics for spatial data*. Wiley New York, 1993.

Edward R. Doering, Joseph Gray, and John P. Basart. Point spread function estimation of image intensifier tubes. *Review of Progress in Quantitative Nondestructive Evaluation*, 11: 323–329, 1992.

Richard Durrett. *Probability: theory and examples*. Cambridge university press, 2010.

F.W. Dyson, A.S. Eddington, and Davidson C. A determination of the deflection of light by the sun's gravitational field, from observations made at the total eclipse of 29 may 1919. *Philosophical Transactions of the Royal Society*, 220A:291–333, 1920.

Charles L. Epstein. *Introduction to the Mathematics of Medical Imaging*. Society for Industrial and Applied Mathematics, 2008. ISBN 9780898716429.

Colin Fox and Richard A Norton. Fast sampling in a linear-gaussian inverse problem. *arXiv preprint arXiv:1507.01614*, 2015.

Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):721–741, 1984.

L. Grafakos. *Classical Fourier Analysis*. Graduate Texts in Mathematics. Springer New York, 2014. ISBN 9781493911943.

Jacques Hadamard. Sur les problmes aux drives partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52, 1902.

Martin Hanke and Otmar Scherzer. Inverse problems light: numerical differentiation. *The American Mathematical Monthly*, 108(6):512–521, 2001.

P. C. Hansen. *Discrete Inverse Problems: Insight and Algorithms*. SIAM, Philadelphia, 2010.

A.K. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall information and system sciences series. Prentice Hall, 1989.

J. Kaipio and E. Somersalo. *Statistical and Computational Methods for Inverse Problems*. Springer, 2005.

O. Knill, R. Dgani, and M. Vogel. A new approach to abel's integral operator. *Astronomy and Astrophysics*, 274:1002–1008, 1993.

D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Magazine*, 13(3):43–64, 1996.

L.N. Magner. *History of the Life Sciences*. New York: Marcel Dekker, Inc, 2002.

Vladimir Aleseevich Morozov and Michael Stessin. *Regularization methods for ill-posed problems*. CRC Press, 1993.

Walter Rudin. *Functional analysis. International series in pure and applied mathematics*. McGraw-Hill, Inc., New York, 1991.

MR Showalter, MK Gordon, and D Olson. Vg1/vg2 saturn iss processed images v1.0. NASA Planetary Data System, 2006.

A. M. Stuart. Inverse problems: A bayesian perspective. *Acta Numerica*, 19:451–559, 2010.

A.N. Tikhonov. Regularization of incorrectly posed problems. *Soviet Mathematics Doklady*, 4:1624–1627, 1963.

David A Van Dyk and Taeyoung Park. Partially collapsed gibbs samplers: Theory and methods. *Journal of the American Statistical Association*, 103(482):790–796, 2008.

Curtis R. Vogel. *Computational Methods of Inverse Problems*. Society for industrual and applied mathematics, 2002.

Scott A. Watson. Real-time spot size measurement for pulsed high-energy radiographic machines. *Proceedings of the 1993 Particle Accelerator Converence*, 4:2447–2449, 1993.