

Tikhonov Regularization for Ill-Posed Poisson Likelihood Estimation: Analysis and Computation

Johnathan M. Bardsley*[†] and N'djekornom Laobeul*
(received December 2006)

The noise contained in images collected by a charge coupled device (CCD) camera is predominantly of Poisson type. This motivates the use of the negative logarithm of the Poisson likelihood in place of the ubiquitous least squares fit-to-data. However if the underlying mathematical model is assumed to have the form $z = Au$, where A is a linear, compact operator, Poisson likelihood estimation is ill-posed, and hence some form of regularization is required. In [1], a numerical method is presented and analyzed for Tikhonov regularized Poisson likelihood estimation, but no theoretical justification of the approach is given. Our primary objective in this paper is to provide such a theoretical justification. We then briefly present the computational method of [1], which is very effective and computationally efficient for this problem. The practical validity of the approach is then demonstrated on a synthetic example from astronomical imaging.

Keywords: regularization, ill-posed problems, maximum likelihood estimation, image reconstruction, nonnegatively constrained minimization.

AMS Subject Classification: 65J22, 65K10, 65F22.

1 Introduction

The following problem is very common in imaging science: given a blurred, noisy $N \times N$ image array \mathbf{z} , obtain an estimate of the underlying $N \times N$ true object array $\mathbf{u}_{\text{exact}}$, by approximately solving a linear system of the form

$$\mathbf{z} = \mathbf{A}\mathbf{u}. \quad (1)$$

Here \mathbf{z} has been column stacked so that it is $N^2 \times 1$, and \mathbf{A} is a known $N^2 \times N^2$ ill-conditioned matrix. We assume that $\mathbf{A}\mathbf{u} \geq \mathbf{0}$ whenever $\mathbf{u} \geq \mathbf{0}$.

*Department of Mathematical Sciences, University of Montana, Missoula, MT.

[†]To whom correspondence should be addressed. Email: bardsleyj@mso.umt.edu. This work was done during this author's visit to the University of Helsinki, Finland in 2006-07 under the University of Montana Faculty Exchange Program.

In practice, the image \mathbf{z} is collected by a charge couple device (CCD) camera and contains random noise. Thus \mathbf{z} is a realization of a random vector $\hat{\mathbf{z}}$. A statistical model for $\hat{\mathbf{z}}$ is given by (c.f. [8,9])

$$\hat{\mathbf{z}} \sim \text{Poiss}(\mathbf{A}\mathbf{u}_{\text{exact}}) + \text{Poiss}(\gamma \cdot \mathbf{1}) + N(\mathbf{0}, \sigma^2 \mathbf{I}). \quad (2)$$

Here $\mathbf{1}$ is an $N^2 \times 1$ vector of all ones, and \mathbf{I} is the $N^2 \times N^2$ identity matrix. By (2) we mean that each element \hat{z}_i of the vector $\hat{\mathbf{z}}$ is a random variable with distribution

$$\hat{z}_i \sim n_{\text{obj}}(i) + n_0(i) + g(i), \quad i = 1, \dots, N^2, \quad (3)$$

where

- $n_{\text{obj}}(i)$ is the number of object dependent photoelectrons measured by the i th detector in the CCD array. It is a Poisson random variable with Poisson parameter $[\mathbf{A}\mathbf{u}_{\text{true}}]_i$.
- $n_0(i)$ is the number of background photoelectrons, which arise from both natural and artificial sources, measured by the i th detector in the CCD array. It is a Poisson random variable with a fixed positive Poisson parameter γ .
- $g(i)$ is the so-called readout noise, which is due to random errors caused by the CCD electronics and errors in the analog-to-digital conversion of measured voltages. It is a Gaussian random variable with mean 0 and fixed variance σ^2 .

The random variables $n_{\text{obj}}(i)$, $n_0(i)$, and $g(i)$ are assumed to be independent of one another and of $n_{\text{obj}}(j)$, $n_0(j)$, and $g(j)$ for $i \neq j$.

As in [8], following [3, pp. 190 and 245], we use the approximation

$$N(\sigma^2, \sigma^2) \approx \text{Poiss}(\sigma^2). \quad (4)$$

From this together with the independence properties of the random variables in (3) we obtain the following approximation of (2):

$$\hat{\mathbf{z}} + \sigma^2 \cdot \mathbf{1} \sim \text{Poiss}(\mathbf{A}\mathbf{u}_{\text{exact}} + \gamma \cdot \mathbf{1} + \sigma^2 \cdot \mathbf{1}). \quad (5)$$

The maximum likelihood estimate of $\mathbf{u}_{\text{exact}}$ given a realization \mathbf{z} from $\hat{\mathbf{z}}$ defined in (5) is the minimizer with respect to \mathbf{u} of the Poisson likelihood functional

$$T_0(\mathbf{u}) \stackrel{\text{def}}{=} \sum_{i=1}^{N^2} ([\mathbf{A}\mathbf{u}]_i + \gamma + \sigma^2) - \sum_{i=1}^{N^2} (z_i + \sigma^2) \log([\mathbf{A}\mathbf{u}]_i + \gamma + \sigma^2). \quad (6)$$

Here $[\mathbf{A}\mathbf{u}]_i$ and z_i are the i th components of $\mathbf{A}\mathbf{u}$ and \mathbf{z} respectively.

In order to perform our analysis, we need the analogue of T_0 as a functional on $L^p(\Omega)$. To obtain this, we make the standard assumption that $\mathbf{A}\mathbf{u}$ is a discretization of the compact operator

$$Au \stackrel{\text{def}}{=} \int_{\Omega} a(x, y; \xi, \eta) u(\xi, \eta) d\xi d\eta, \quad (7)$$

where $a \in L^2(\Omega)$. As in the discrete case, we assume that $Au \geq 0$ for all $u \geq 0$. Then as $N \rightarrow \infty$, we have $\mathbf{A}\mathbf{u} \rightarrow Au$. Given the continuous true image u_{exact} , we define $z := Au_{\text{exact}}$. Since u_{exact} denotes image intensity, it is nonnegative, and hence, $z \geq 0$.

For N fixed, the stochastic errors in the data \mathbf{z} can be removed by computing the mean of samples from (2) as the sample size tends to infinity. The result, via the Central Limit Theorem, is $\mathbf{A}\mathbf{u}_{\text{exact}} + \gamma$.

Simultaneously removing discretization and stochastic errors gives us the analogue of $T_0(\mathbf{u})$ on $L^p(\Omega)$ that we seek:

$$T_0(Au; z + \gamma) = \int_{\Omega} ((Au + \gamma + \sigma^2) - (z + \gamma + \sigma^2) \log(Au + \gamma + \sigma^2)) dx. \quad (8)$$

To see that u_{exact} is the unique minimizer of $T_0(u)$, we compute the gradient and Hessian of T_0 with respect to u , which are given, respectively, by

$$\nabla T_0(Au; z + \gamma) = A^* \left(\frac{Au - z}{Au + \gamma + \sigma^2} \right), \quad (9)$$

$$\nabla^2 T_0(Au; z + \gamma) = A^* \left(\text{diag} \left(\frac{z + \gamma + \sigma^2}{(Au + \gamma + \sigma^2)^2} \right) \right) A, \quad (10)$$

where “ $*$ ” denotes operator adjoint, and $\text{diag}(v)$ is defined by $\text{diag}(v)w = vw$. Since $z \geq 0$, (10) is a positive definite operator, implying T_0 is strictly convex [11, Theorem 2.42] with unique minimizer. By (9), this minimizer is u_{exact} as desired.

Since A is compact, however, the problem of minimizing $T_0(u)$ is ill-posed [10, 11], and hence, regularization is required. If standard Tikhonov regularization is used and the associated minimization problem is component-wise nonnegatively constrained - a natural constraint for our problem - we obtain

$$\min_{u \in \mathcal{C}} T_{\alpha}(Au; z + \gamma) \stackrel{\text{def}}{=} T_0(Au; z + \gamma) + \frac{\alpha}{2} \|u\|_2^2, \quad (11)$$

where

$$\mathcal{C} = \{u \in L^p(\Omega) \mid u \geq 0 \text{ almost everywhere}\}, \quad (12)$$

and Ω denotes the computational domain.

Our main objective in this paper is to provide a theoretical justification for using (11), (12). This involves first proving that (11), (12) is a well-posed problem, i.e. that (11) has a unique solution $u_\alpha \in \mathcal{C}$ for each $\alpha > 0$, and that given a sequence of perturbed compact operator equations

$$z_n(x, y) = A_n u(x, y) \stackrel{\text{def}}{=} \int_{\Omega} a_n(x, y; \xi, \eta) u(\xi, \eta) d\xi d\eta, \quad (13)$$

where $a_n \in L^2(\Omega)$ and solutions $u_{\alpha, n}$ of the corresponding minimization problems

$$\arg \min_{u \in \mathcal{C}} T_\alpha(A_n u; z_n), \quad (14)$$

$u_{\alpha, n} \rightarrow u_\alpha$ provided $A_n \rightarrow A$ and $z_n \rightarrow z + \gamma$.

We also must show that a sequence of positive regularization parameters $\{\alpha_n\}$ can be chosen so that $u_{\alpha_n, n} \rightarrow u_{\text{exact}}$ as $\alpha_n \rightarrow 0$, where $u_{\alpha_n, n}$ is the minimizer of $T_{\alpha_n}(A_n u; z_n)$ over \mathcal{C} . The theoretical analysis is the focus of Section 2.

We note that these two convergence results are not only of academic interest, since in practice one always deals with some perturbed approximation of an (assumed) exact underlying model.

In [1], an effective computational method for solving

$$\arg \min_{\mathbf{u} \geq \mathbf{0}} T_\alpha(\mathbf{u}) \stackrel{\text{def}}{=} T_0(\mathbf{u}) + \frac{\alpha}{2} \|\mathbf{u}\|^2, \quad (15)$$

where T_0 is defined in (6), is presented and is compared, with favorable results, against several other methods and approaches. We briefly describe this method and present a numerical experiment in Section 3.

We end the paper with conclusions in Section 4.

2 Theoretical Analysis

We begin with definitions and results that will be needed in our later analysis.

Let Ω and \mathcal{C} be as defined in the introduction. Then $|\Omega| = \int_{\Omega} dx < \infty$. We will assume that $u_{\text{exact}} \in \mathcal{C}$ and that $Au \geq 0$ for every $u \in \mathcal{C}$. Then $z \geq 0$.

We make the further assumption that $z \in L^\infty(\Omega)$. We note that these are reasonable assumptions on A and z for imaging applications.

Let $|\cdot|$ denote the Euclidean norm in \mathbb{R}^d and $\|\cdot\|_p$ the Banach space norm on $L^p(\Omega)$ for $1 \leq p \leq \infty$. Since Ω is bounded, $L^p(\Omega) \subset L^1(\Omega)$ for $p > 1$. Also, the following result will hold.

THEOREM 2.1 *If \mathcal{S} is a bounded set in $L^2(\Omega)$ it is a bounded set in $L^p(\Omega)$ for all $p \geq 1$.*

The following definition will be needed in our analysis below.

Definition 2.2 A functional $T : L^p(\Omega) \rightarrow \mathbb{R}$ is coercive if

$$T(u) \rightarrow \infty \quad \text{whenever} \quad \|u\|_2 \rightarrow \infty. \quad (16)$$

2.1 Well-Posedness

We now prove that problem (11) is well-posed for $\alpha > 0$. In order to simplify the notation in our arguments, we will use $T_\alpha(u)$ to denote $T_\alpha(Au; z + \gamma)$ and $T_{\alpha,n}(u)$ to denote $T_\alpha(A_n u; z_n)$ throughout the remainder of the paper.

In order to prove the existence and uniqueness of solutions of (11), we will use the following theorem, which is similar to [11, Theorem 2.30].

THEOREM 2.3 *If $T : L^p(\Omega) \rightarrow \mathbb{R}$ is strictly convex and coercive, then it has a unique minimizer on \mathcal{C} .*

Proof Let $\{u_n\} \subset \mathcal{C}$ be such that $T(u_n) \rightarrow T_* \stackrel{\text{def}}{=} \inf_{u \in \mathcal{C}} T(u)$. Then, by (16), the sequence $\{u_n\}$ is bounded in $L^2(\Omega)$. By Theorem 2.1, this implies that $\{u_n\}$ has a subsequence $\{u_{n_j}\}$ that converges to some $u_* \in \mathcal{C}$. Now, since T is strictly convex, it is weakly lower semi-continuous [11], and hence,

$$T(u_*) \leq \liminf T(u_{n_j}) = \lim T(u_n) = T_*.$$

Thus u_* minimizes T on \mathcal{C} and is unique since T is a strictly convex functional and \mathcal{C} is a convex set. \square

COROLLARY 2.4 *(Existence and Uniqueness of Minimizers) T_α has a unique minimizer over \mathcal{C} .*

Proof The strict convexity of T_α follows from the strict convexity of both T_0 – shown in the introduction – and $\|u\|^2$.

For coercivity, we note that by Jensen's inequality and the properties of the

function $x - c \log x$ for $c > 0$,

$$\begin{aligned} T_0(u) &\geq \|Au + \gamma + \sigma^2\|_1 - \|z + \gamma + \sigma^2\|_\infty \log \|Au + \gamma + \sigma^2\|_1, \\ &\geq \|z + \gamma + \sigma^2\|_\infty - \|z + \gamma + \sigma^2\|_\infty \log \|z + \gamma + \sigma^2\|_\infty. \end{aligned} \quad (17)$$

Since $z > 0$, T_0 is bounded below. The coercivity of $T_\alpha(u) = T_0(u) + \alpha \|u\|_2^2$ then follows immediately.

By Theorem 2.3, T_α has a unique minimizer in \mathcal{C} . \square

Let u_α be the unique solution of T_α over \mathcal{C} given by Corollary 2.4. A similar analysis yields the existence and uniqueness of solutions $u_{\alpha,n}$ of (14) for $\alpha \geq 0$ provided that for each $n \in \mathbb{N}$, $A_n u \geq 0$ for all $u \in \mathcal{C}$ and $z_n \geq 0$. Problem (11) is then said to be stable provided $A_n \rightarrow A$ and $z_n \rightarrow z + \gamma$ implies $u_{\alpha,n} \rightarrow u_\alpha$.

The following theorem gives conditions that guarantee this result.

THEOREM 2.5 *Let $u_{\alpha,n}$ be the unique minimizer of $T_{\alpha,n}$ over \mathcal{C} , and suppose that*

(i) *for any sequence $\{u_n\} \subset L^p(\Omega)$,*

$$\lim_{n \rightarrow \infty} T_{\alpha,n}(u_n) = +\infty \quad \text{whenever} \quad \lim_{n \rightarrow \infty} \|u_n\|_2 = +\infty; \quad (18)$$

(ii) *given $B > 0$ and $\epsilon > 0$, there exists N such that*

$$|T_{\alpha,n}(u) - T_\alpha(u)| < \epsilon \quad \text{whenever} \quad n \geq N, \|u\|_2 \leq B. \quad (19)$$

Then

$$\lim_{n \rightarrow \infty} \|u_{\alpha,n} - u_\alpha\|_p = 0. \quad (20)$$

Proof Note that $T_{\alpha,n}(u_{\alpha,n}) \leq T_{\alpha,n}(u_\alpha)$. From this and (19), we have

$$\liminf T_{\alpha,n}(u_{\alpha,n}) \leq \limsup T_{\alpha,n}(u_{\alpha,n}) \leq T_\alpha(u_\alpha) < \infty. \quad (21)$$

Thus by (18), the $u_{\alpha,n}$'s are bounded in $L^2(\Omega)$. By Theorem 2.1, we know that there exists a subsequence u_{n_j} that converges to some $\hat{u} \in L^p(\Omega)$. Furthermore, by the weak lower semicontinuity of T_α , (19), and (21) we have

$$\begin{aligned} T_\alpha(\hat{u}) &\leq \liminf T_\alpha(u_{n_j}), \\ &= \liminf (T_\alpha(u_{n_j}) - T_{\alpha,n_j}(u_{n_j})) + \liminf T_{\alpha,n_j}(u_{n_j}), \\ &\leq T_\alpha(u_\alpha). \end{aligned}$$

By uniqueness of minimizers, $\hat{u} = u_\alpha$. Thus every convergent subsequence of $\{u_{\alpha,n}\}$ converges to u_α , and hence, we have (20). \square

The following corollary of Theorem 2.5 is the stability result for (11) that we seek.

COROLLARY 2.6 (*Stability of Minimizers*) *Assume $\|z_n - (z + \gamma)\|_\infty \rightarrow 0$, and that $A_n \rightarrow A$ in the $L^1(\Omega)$ operator norm. Then*

$$\lim_{n \rightarrow \infty} \|u_{\alpha,n} - u_\alpha\|_p = 0.$$

Proof It suffices to show that conditions (i) and (ii) from Theorem 2.5 hold. For condition (i), note that the analogue of inequality (17) for $T_{0,n}$ is given by

$$T_{0,n}(u_{\alpha,n}) \geq \|z_n + \sigma^2\|_\infty - \|z_n + \sigma^2\|_\infty \log \|z_n + \sigma^2\|_\infty,$$

which has a lower bound for all n since $\|z_n - (z + \gamma)\|_\infty \rightarrow 0$ and $z \in L^\infty(\Omega)$ is nonnegative. Thus $T_{\alpha,n}(u_n) = T_{0,n}(u_n) + \alpha \|u_n\|_2^2 \rightarrow \infty$ whenever $\|u_n\|_2 \rightarrow \infty$, and hence, (18) is satisfied.

For condition (ii), note that, using Jensen's inequality and the properties of the logarithm,

$$\begin{aligned} |T_{\alpha,n}(u) - T_\alpha(u)| &= \left| \int_\Omega ((A_n - A)u - (z_n + \sigma^2) \log(A_n u + \gamma + \sigma^2)) \, dx \right. \\ &\quad \left. + \int_\Omega ((z + \gamma + \sigma^2) \log(Au + \gamma + \sigma^2)) \, dx \right|, \\ &\leq \|A_n - A\|_1 \|u\|_1 \\ &\quad + \|z_n - (z + \gamma)\|_\infty \log(\|A_n\|_1 \|u\|_1 + (\gamma + \sigma^2)|\Omega|) \quad (22) \\ &\quad + \|z + \gamma + \sigma^2\|_\infty \log \left\| \frac{(Au + \gamma + \sigma^2)}{(A_n u + \gamma + \sigma^2)} \right\|_1. \end{aligned}$$

By assumption, $\|A_n - A\|_1, \|z_n - (z + \gamma)\|_\infty \rightarrow 0$. Furthermore, by the Banach-Steinhaus Theorem, $\|A_n\|_1$ is uniformly bounded, and since we are assuming that $\|u\|_2$ is bounded, by Theorem 2.1 we have that $\|u\|_1$ is bounded as well. Thus the first two terms on the right-hand side in (22) tend to zero as $n \rightarrow \infty$. For the third term note that

$$\left\| \frac{Au + \gamma + \sigma^2}{A_n u + \gamma + \sigma^2} - 1 \right\|_1 \leq \left\| \frac{1}{A_n u + \gamma + \sigma^2} \right\|_1 \|A_n - A\|_1 \|u\|_1,$$

which converges to zero since $\|1/(A_n u + \gamma + \sigma^2)\|_1$ is bounded and $\|A_n - A\|_1 \rightarrow$

0. Thus $\log(\|(Au + \gamma + \sigma^2)/(Au_n + \gamma + \sigma^2)\|_1) \rightarrow \log(1) = 0$, and hence

$$|T_{\alpha,n}(u) - T_\alpha(u)| \rightarrow 0. \quad (23)$$

The desired result now follows from Theorem 2.5. \square

Finally, the main result of this subsection now follows directly from Corollaries 2.4 and 2.6.

THEOREM 2.7 (*Well-Posedness*) *Problem (11) is well-posed.*

2.2 Convergence of Minimizers

It remains to show that a sequence of positive regularization parameters $\{\alpha_n\}$ can be chosen so that $u_{\alpha_n,n} \rightarrow u_{\text{exact}}$ as $\alpha_n \rightarrow 0$.

THEOREM 2.8 (*Convergence of Minimizers*) *Suppose $\|z_n - (z + \gamma)\|_\infty \rightarrow 0$, $A_n \rightarrow A$ in the $L^1(\Omega)$ operator norm, and $\alpha_n \rightarrow 0$ at a rate such that*

$$(T_{0,n}(u_{\text{exact}}) - T_{0,n}(u_{0,n}))/\alpha_n \quad (24)$$

is bounded. Then $u_{\alpha_n,n} \rightarrow u_{\text{exact}}$ strongly in $L^p(\Omega)$ for $p \geq 1$.

Proof Since $u_{\alpha_n,n}$ minimizes $T_{\alpha_n,n}$, we have

$$T_{\alpha_n,n}(u_{\alpha_n,n}) \leq T_{\alpha_n,n}(u_{\text{exact}}). \quad (25)$$

Since $\{z_n\}$ and $\{A_n\}$ are uniformly bounded and $A_n \rightarrow A$ in the $L^1(\Omega)$ operator norm, $\{T_{\alpha_n,n}(u_{\text{exact}})\}$ is a bounded sequence. Hence $\{T_{\alpha_n,n}(u_{\alpha_n,n})\}$ is bounded by (25).

Subtracting $T_{0,n}(u_{0,n})$ from both sides of (25) and dividing by α_n yields

$$(T_{0,n}(u_{\alpha_n,n}) - T_{0,n}(u_{0,n})) / \alpha_n + \|u_{\alpha_n,n}\|_2^2 \leq (T_{0,n}(u_{\text{exact}}) - T_{0,n}(u_{0,n})) / \alpha_n + \|u_{\text{exact}}\|_2^2. \quad (26)$$

By (24), the right-hand side is bounded, implying the left hand side is bounded. Since $T_{0,n}(u_{\alpha_n,n}) - T_{0,n}(u_{0,n})$ is nonnegative, this implies that $\{u_{\alpha_n,n}\}$ is bounded in $L^2(\Omega)$.

We now show that $u_{\alpha_n,n} \rightarrow u_{\text{exact}}$ in $L^p(\Omega)$ by showing that every subsequence of $\{u_{\alpha_n,n}\}$ contains a subsequence that converges to u_{exact} . Since $\{u_{\alpha_n,n}\}$ is bounded in $L^2(\Omega)$, by Theorem 2.1, each of its subsequences in turn has a subsequence that converges strongly in $L^p(\Omega)$. Let $\{u_{\alpha_{n_j},n_j}\}$ be such a

sequence and \hat{u} its limit. Then

$$\begin{aligned} T_0(\hat{u}) &= \int_{\Omega} (A(\hat{u} - u_{\alpha_{n_j}, n_j}) + (A - A_{n_j})u_{\alpha_{n_j}, n_j}) dx dy \\ &\quad + \int_{\Omega} (z_{n_j} - (z + \gamma)) \log(A\hat{u} + \gamma + \sigma^2) dx dy \\ &\quad - \int_{\Omega} (z_{n_j} + \sigma^2) \log((A_{n_j}u_{\alpha_{n_j}, n_j} + \gamma + \sigma^2)/(A\hat{u} + \gamma + \sigma^2)) dx dy \\ &\quad + T_{0, n_j}(u_{\alpha_{n_j}, n_j}), \end{aligned}$$

which, as in previous arguments, yields

$$\begin{aligned} |T_{0, n_j}(u_{\alpha_{n_j}, n_j}) - T_0(\hat{u})| &\leq \int_{\Omega} A(\hat{u} - u_{\alpha_{n_j}, n_j}) dx dy \\ &\quad + \|z_{n_j} - (z + \gamma)\|_{\infty} \log(\|A\|_1 \|\hat{u}\|_1 + \gamma |\Omega|) \\ &\quad + \|z_{n_j} + \sigma^2\|_{\infty} \log \|(A_{n_j}u_{\alpha_{n_j}, n_j} + \gamma + \sigma^2)/(A\hat{u} + \gamma + \sigma^2)\|_1 \\ &\quad + \|A - A_{n_j}\|_1 \|u_{\alpha_{n_j}, n_j}\|_1. \end{aligned}$$

Then

$$\|z_{n_j} - (z + \gamma)\|_{\infty} \log(\|A\|_1 \|\hat{u}\|_1 + (\gamma + \sigma^2)|\Omega|) \rightarrow 0,$$

since $\|z_{n_j} - (z + \gamma)\|_{\infty} \rightarrow 0$ and $\log(\|A\|_1 \|\hat{u}\|_1 + (\gamma + \sigma^2)|\Omega|)$ is constant, and

$$\|A - A_{n_j}\|_1 \|u_{\alpha_{n_j}, n_j}\|_1 \rightarrow 0$$

since $\|A - A_{n_j}\|_1 \rightarrow 0$ and $\|u_{\alpha_{n_j}, n_j}\|_1$ is bounded by Theorem 2.1.

Since A is a bounded linear operator and Ω is a set of finite measure, $F(u) = \int_{\Omega} Au dx$ is a bounded linear functional on $L^p(\Omega)$. The convergence of $\{u_{\alpha_{n_j}, n_j}\}$ then implies $\int_{\Omega} Au_{\alpha_{n_j}, n_j} dx \rightarrow \int_{\Omega} A\hat{u} dx$, which yields

$$\int_{\Omega} A(\hat{u} - u_{\alpha_{n_j}, n_j}) dx \rightarrow 0.$$

Since A is compact, it is completely continuous, i.e. $u_{\alpha_{n_j}, n_j} \rightarrow \hat{u}$ implies that $\|Au_{\alpha_{n_j}, n_j} - A\hat{u}\|_1 \rightarrow 0$ (cf. [2, Prop. 3.3]). Thus, since $\left\| \frac{1}{A\hat{u} + \gamma + \sigma^2} \right\|_1$ is bounded,

and

$$\begin{aligned} \left\| \frac{A_{n_j} u_{\alpha_{n_j}, n_j} + \gamma + \sigma^2}{A\hat{u} + \gamma + \sigma^2} - 1 \right\|_1 &\leq \left\| \frac{1}{A\hat{u} + \gamma + \sigma^2} \right\|_1 \|A_{n_j} u_{\alpha_{n_j}, n_j} - A\hat{u}\|_1, \\ &\leq \left\| \frac{1}{A\hat{u} + \gamma + \sigma^2} \right\|_1 \times \\ &\quad \left(\|A_{n_j} - A\|_1 \|u_{\alpha_{n_j}, n_j}\|_1 + \|A u_{\alpha_{n_j}, n_j} - A\hat{u}\|_1 \right), \end{aligned}$$

we have that $\|z_{n_j} + \sigma^2\|_\infty \log \|(A_{n_j} u_{\alpha_{n_j}, n_j} + \gamma + \sigma^2)/(A\hat{u} + \gamma + \sigma^2)\|_1 \rightarrow 0$. Therefore

$$T_0(\hat{u}) = \lim_{n_j \rightarrow \infty} T_{0, n_j}(u_{\alpha_{n_j}, n_j}).$$

Invoking (26), (24), and (23), respectively, yields

$$\lim_{n_j \rightarrow \infty} T_{0, n_j}(u_{\alpha_{n_j}, n_j}) = \lim_{n_j \rightarrow \infty} T_{0, n_j}(u_{\alpha_{n_j}, n_j}) = \lim_{n_j \rightarrow \infty} T_{0, n_j}(u_{\text{exact}}) = T_0(u_{\text{exact}}).$$

Thus $T_0(\hat{u}) = T_0(u_{\text{exact}})$. Since u_{exact} is the unique minimizer of T_0 , we have $\hat{u} = u_{\text{exact}}$. Therefore $\{u_{\alpha_{n_j}, n_j}\}$ converges strongly to u_{exact} in $L^p(\Omega)$. \square

3 An Efficient Numerical Method

In this section, we give a brief outline of the method introduced in [1] for solving the nonnegatively constrained minimization problem (15).

3.1 Preliminaries

The projection of a vector $\mathbf{u} \in \mathbb{R}^{N^2}$ onto the feasible set $\{\mathbf{u} \in \mathbb{R}^{N^2} \mid \mathbf{u} \geq \mathbf{0}\}$ can be conveniently expressed as

$$\mathcal{P}(\mathbf{u}) \stackrel{\text{def}}{=} \arg \min_{\mathbf{v} \in \Omega} \|\mathbf{v} - \mathbf{u}\| = \max\{\mathbf{u}, \mathbf{0}\},$$

where $\max\{\mathbf{u}, \mathbf{0}\}$ is the vector whose i th component is zero if $u_i < 0$ and is u_i otherwise. The active set for a vector $\mathbf{u} \geq \mathbf{0}$ is defined

$$\mathcal{A}(\mathbf{u}) = \{i \mid u_i = 0\},$$

and the complementary set of indices, $\mathcal{I}(\mathbf{u})$, is known as the inactive set.

The gradient of T_α defined in (15) has the form

$$\nabla T_\alpha(\mathbf{u}) = \mathbf{A}^T ((\mathbf{A}\mathbf{u} + \gamma - \mathbf{z}) / (\mathbf{A}\mathbf{u} + \gamma + \sigma^2)) + \alpha\mathbf{u}, \quad (27)$$

where “ / ” denotes component-wise division. Its Hessian is

$$\nabla^2 T_\alpha(\mathbf{u}) = \mathbf{A}^T \text{diag}(\mathbf{w}(\mathbf{u})) \mathbf{A} + \alpha\mathbf{I}, \quad (28)$$

where the components of $\mathbf{w}(\mathbf{u})$ are given by

$$[\mathbf{w}(\mathbf{u})]_i = \frac{z_i + \sigma^2}{([\mathbf{A}\mathbf{u}]_i + \gamma + \sigma^2)^2}. \quad (29)$$

Note that for moderate to large values of σ^2 , say $\sigma^2 \geq 3^2$, it is extremely unlikely for the Gaussian $g(i) + \sigma^2$ to take on negative values. Then since Poisson random variables take on only nonnegative integer values, the random variable $z_i + \sigma^2$ is also highly unlikely to be nonpositive. Furthermore, γ and σ^2 are both positive parameters and by assumption $\mathbf{A}\mathbf{u} \geq \mathbf{0}$ whenever $\mathbf{u} \geq \mathbf{0}$. Thus it can be reasonably assumed that $\nabla^2 T_\alpha(\mathbf{u})$ is strictly positive definite for all $\mathbf{u} \geq \mathbf{0}$, in which case T_α has a unique minimizer over the set $\{\mathbf{u} \in \mathbb{R}^{N^2} \mid \mathbf{u} \geq \mathbf{0}\}$.

The reduced gradient of T_α at $\mathbf{u} \geq \mathbf{0}$ is given by

$$[\nabla_{\text{red}} T_\alpha(\mathbf{u})]_i = \begin{cases} \frac{\partial T_\alpha(\mathbf{u})}{\partial x_i}, & i \in \mathcal{I}(\mathbf{u}) \\ 0, & i \in \mathcal{A}(\mathbf{u}), \end{cases}$$

the projected gradient of T_α by

$$[\nabla_{\text{proj}} T_\alpha(\mathbf{u})]_i = \begin{cases} \frac{\partial T_\alpha(\mathbf{u})}{\partial x_i}, & i \in \mathcal{I}(\mathbf{u}), \text{ or } i \in \mathcal{A}(\mathbf{u}) \text{ and } \frac{\partial T_\alpha(\mathbf{u})}{\partial x_i} < 0, \\ 0, & \text{otherwise,} \end{cases}$$

and the reduced Hessian by

$$[\nabla_{\text{red}}^2 T_\alpha(\mathbf{u})]_{ij} = \begin{cases} \frac{\partial^2 T_\alpha(\mathbf{u})}{\partial x_i \partial x_j}, & \text{if } i \in \mathcal{I}(\mathbf{u}) \text{ and } j \in \mathcal{I}(\mathbf{u}) \\ \delta_{ij}, & \text{otherwise.} \end{cases}$$

Finally, we define $\mathbf{D}_{\mathcal{I}}(\mathbf{u})$ to be the diagonal matrix with components

$$[\mathbf{D}_{\mathcal{I}}(\mathbf{u})]_{ii} = \begin{cases} 1, & i \in \mathcal{I}(\mathbf{u}) \\ 0, & i \in \mathcal{A}(\mathbf{u}), \end{cases} \quad (30)$$

and $\mathbf{D}_{\mathcal{A}}(\mathbf{u}) = \mathbf{I} - \mathbf{D}_{\mathcal{I}}(\mathbf{u})$. Note then that

$$\nabla_{\text{red}} T_{\alpha}(\mathbf{u}) = \mathbf{D}_{\mathcal{I}}(\mathbf{u}) \nabla T_{\alpha}(\mathbf{u}), \quad (31)$$

$$\nabla_{\text{red}}^2 T_{\alpha}(\mathbf{u}) = \mathbf{D}_{\mathcal{I}}(\mathbf{u}) \nabla^2 T_{\alpha}(\mathbf{u}) \mathbf{D}_{\mathcal{I}}(\mathbf{u}) + \mathbf{D}_{\mathcal{A}}(\mathbf{u}), \quad (32)$$

3.2 Gradient Projection Iteration

A key component of the iterative method introduced in [1] is the gradient projection iteration [5], which we present now. Given $\mathbf{u}_k \geq \mathbf{0}$, we compute \mathbf{u}_{k+1} via

$$\mathbf{p}_k = -\nabla T_{\alpha}(\mathbf{u}_k) \quad (33)$$

$$\lambda_k = \arg \min_{\lambda > 0} T_{\alpha}(\mathcal{P}(\mathbf{u}_k + \lambda \mathbf{p}_k)) \quad (34)$$

$$\mathbf{u}_{k+1} = \mathcal{P}(\mathbf{u}_k + \lambda_k \mathbf{p}_k) \quad (35)$$

In practice, subproblem (34) is solved inexactly using a projected backtracking line search. In the implementation used here, we take the initial step length parameter to be

$$\lambda_k^0 = \frac{\|\mathbf{p}_k\|^2}{\langle \nabla^2 T_{\alpha}(\mathbf{u}_k) \mathbf{p}_k, \mathbf{p}_k \rangle}. \quad (36)$$

The quadratic backtracking line search algorithm found in [6] is then used to create a sequence of line search parameters $\{\lambda_k^j\}_{j=0}^m$, where m is the smallest positive integer such that the sufficient decrease condition

$$T_{\alpha}(\mathbf{u}_k(\lambda_k^j)) \leq T_{\alpha}(\mathbf{u}_k) - \frac{\mu}{\lambda_k^j} \|\mathbf{u}_k - \mathbf{u}_k(\lambda_k^j)\|^2 \quad (37)$$

holds. Here $\mu \in (0, 1)$ and

$$\mathbf{u}_k(\lambda) = \mathcal{P}_{\Omega}(\mathbf{u}_k + \lambda \mathbf{p}_k). \quad (38)$$

The approximate solution of (34) is then taken to be λ_k^m .

In [1], the following theorem is proved regarding the convergence of the projected gradient iteration when applied to (15).

THEOREM 3.1 *The gradient projection iteration creates a sequence \mathbf{u}_k that converges to the unique solution of (15).*

3.3 The Reduced Newton Step and Conjugate Gradient

In practice, the gradient projection iteration is very slow to converge. However, a robust method with good convergence properties results if gradient projection iterations are interspersed with steps computed from the reduced Newton system

$$\nabla_{\text{red}}^2 T_\alpha(\mathbf{u}_k) \mathbf{p} = -\nabla_{\text{red}} T_\alpha(\mathbf{u}_k). \quad (39)$$

Approximate solutions of (39) can be efficiently obtained using conjugate gradient iteration (CG) [7] applied to the problem of minimizing

$$q_k(\mathbf{p}) = T_\alpha(\mathbf{u}_k) + \langle \nabla_{\text{red}} T_\alpha(\mathbf{u}_k), \mathbf{p} \rangle + \frac{1}{2} \langle \nabla_{\text{red}}^2 T_\alpha(\mathbf{u}_k) \mathbf{p}, \mathbf{p} \rangle. \quad (40)$$

The result is a sequence $\{\mathbf{p}_k^j\}$ that converges to the minimizer of (40). Even with rapid CG convergence, for large-scale problems it is important to choose effective stopping criteria to reduce overall computational cost. We have found that the following stopping criterion from Moré and Toraldo [6] is very effective:

$$q_k(\mathbf{p}_k^{j-1}) - q_k(\mathbf{p}_k^j) \leq \gamma_{CG} \max\{q_k(\mathbf{p}_k^{i-1}) - q_k(\mathbf{p}_k^i) \mid i = 1, \dots, j-1\}, \quad (41)$$

where $0 < \gamma_{CG} < 1$. Then the approximate solution of (40) is taken to be the $\mathbf{p}_k^{m_{CG}}$ where m_{CG} is the smallest integer such that (41) is satisfied.

With $\mathbf{p}_k := \mathbf{p}_k^{m_{CG}}$, we again apply a projected backtracking line search, only this time we use the much less stringent acceptance criteria

$$T_\alpha(\mathbf{u}_k(\lambda_k^m)) < T_\alpha(\mathbf{u}_k). \quad (42)$$

3.4 The Numerical Algorithm

In the first stage of our algorithm we need stopping criteria for the gradient projection iterations. Borrowing from Moré and Toraldo [6], we stop when

$$T_\alpha(\mathbf{u}_{k-1}) - T_\alpha(\mathbf{u}_k) \leq \gamma_{GP} \max\{T_\alpha(\mathbf{u}_{i-1}) - T_\alpha(\mathbf{u}_i) \mid i = 1, \dots, k-1\}, \quad (43)$$

where $0 < \gamma_{GP} < 1$.

Gradient Projection-Reduced Newton-CG (GPRNCG) Algorithm

Step 0: Select initial guess \mathbf{u}_0 , and set $k = 0$.

Step 1: Given \mathbf{u}_k .

- (1) Take gradient projection steps until either (43) is satisfied or GP_{\max} iterations have been computed. Return updated \mathbf{u}_k .

Step 2: Given \mathbf{u}_k .

- (1) Do CG iterations to approximately minimize the quadratic (40) until either (41) is satisfied or CG_{\max} iterations have been computed. Return $\mathbf{p}_k = \mathbf{p}_k^{m_{\text{CG}}}$.
- (2) Find λ_k^m that satisfies (42), and return $\mathbf{u}_{k+1} = \mathbf{u}_k(\lambda_k^m)$.
- (3) Update $k := k + 1$ and return to Step 1.

Since at each outer GPRNCG iteration at least one gradient projection step, with sufficient decrease condition (42), is taken, by Theorem 3.1 we have the following result.

THEOREM 3.2 *The iterates $\{\mathbf{u}_k\}$ generated by GPRNCG are guaranteed to converge to the unique solution of problem (15).*

3.5 A Numerical Experiment

In this section we present results obtained when GPRNCG is used for solving (15) with simulated data generated according to statistical model (2). We do not present comparisons with other approaches since this is done extensively in [1]. In particular, there it is shown that GPRNCG is more computationally efficient than a number of other algorithms – both standard and state of the art. It is also shown that reconstruction quality can be expected to be better in certain cases when the Poisson likelihood is used instead of the least squares likelihood (both weighted and unweighted, and with and without nonnegativity constraints).

In [1], the simulated object $\mathbf{u}_{\text{exact}}$ is a star field. Thus here we perform tests on a different type of object; namely, the 64×64 simulated satellite seen on the left-hand side in Figure 1. Generating corresponding blurred noisy data requires a discrete PSF \mathbf{a} , which we compute using the Fourier optics [4] PSF model

$$\mathbf{a} = \left| \mathbf{F}^{-1} \left\{ \mathbf{p} \odot e^{i\phi} \right\} \right|^2,$$

where \mathbf{p} is the $N \times N$ indicator array for the telescopes pupil; “ \odot ” denotes Hadamard (component-wise) product; ϕ is the $N \times N$ array that represents the aberrations in the incoming wavefronts of light; $i = \sqrt{-1}$; and \mathbf{F} is the two-dimensional discrete Fourier transform matrix. The $64^2 \times 64^2$ blurring matrix \mathbf{A} obtained from \mathbf{a} is block Toeplitz with Toeplitz blocks (BTTB) [11]. For efficient computations, \mathbf{A} is embedded in a $128^2 \times 128^2$ block circulant

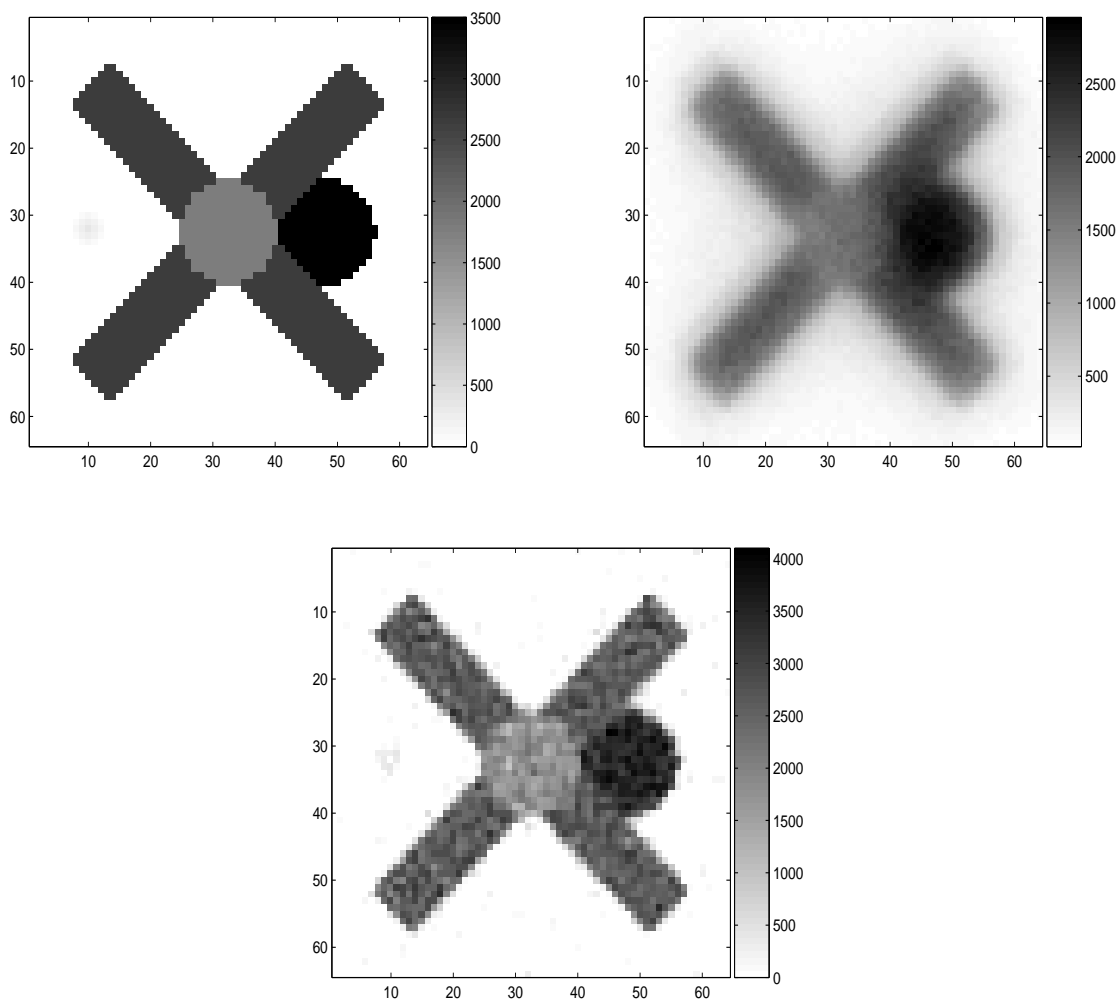


Figure 1. On the upper-left is the true object \mathbf{u}_{true} . On the upper-right, is the blurred, noisy image \mathbf{z} . The lower image is the reconstruction obtained by GPRNCG applied to (15) with $\alpha = 2 \times 10^{-6}$.

with circulant block (BCCB) matrix, which can be diagonalized by the two-dimensional discrete Fourier and inverse discrete Fourier transform matrices [11]. Data \mathbf{z} with a signal-to-noise ratio of approximately 35 is then generated using (2) with $\sigma^2 = 25$ and $\gamma = 10$ – physically realistic values for these parameters. To generate Poisson noise, the `poissrnd` function in MATLAB’s Statistics Toolbox is used. The corresponding blurred, noisy data \mathbf{z} is given on the right hand side in Figure 1.

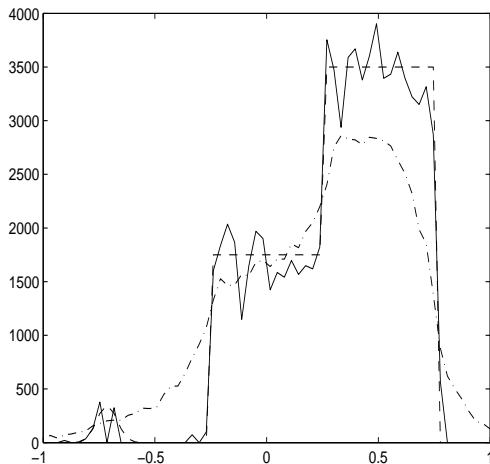


Figure 2. Plots of the 32nd row of the 64×64 arrays $\mathbf{u}_{\text{exact}}$, \mathbf{z} and \mathbf{u}_α . The dashed line corresponds to $\mathbf{u}_{\text{exact}}$; the dash-dotted line corresponds to \mathbf{z} ; and the solid line corresponds to \mathbf{u}_α with $\alpha = 2 \times 10^{-6}$.

We use GPRNCG to obtain the solution of (15). The regularization parameter $\alpha = 2 \times 10^{-6}$ was chosen so that $\|\mathbf{u}_\alpha - \mathbf{u}_{\text{exact}}\|$ is minimized. Our choice of GPRNCG parameters included $\text{GP}_{\text{max}} = 1$, since more gradient projection iterations did not appreciably improve the convergence properties of GPRNCG, $\text{CG}_{\text{max}} = 50$, and $\gamma_{\text{CG}} = 0.25$. We stopped GPRNCG iterations after a 10 orders of magnitude decrease in the norm of $\nabla_{\text{proj}} T_\alpha(\mathbf{u}_k)$. In this example, this stopping criteria was satisfied after only 12 GPRNCG iterations. The reconstruction is given on the bottom in Figure 1.

Finally, we note that one of the benefits of using the Poisson likelihood in place of least squares is that it is sensitive to changes in the low intensity regions of images. This is illustrated by Figure 2 where a cross section of $\mathbf{u}_{\text{exact}}$, \mathbf{z} and \mathbf{u}_α are plotted corresponding to the 32nd row of the respective arrays. The low intensity feature, which can also be seen in the two dimensional images in Figure 1, is reconstructed with reasonable accuracy using the Poisson likelihood. A more extensive analysis of this type, with comparisons between Poisson likelihood and least squares reconstructions, is given in [1].

4 Conclusions

The main results of the paper are the proofs of the well-posedness of (11), (12) (see Theorem 2.7) and of the convergence of a sequence of the type $\{u_{\alpha_n, n}\}$ to u_{exact} , where $u_{\alpha_n, n}$ is defined by (14) and $\{\alpha_n\}$ is chosen so that (24) is

bounded (see Theorem 2.8).

Following the theoretical analysis, we presented the computational method of [1]. This method is efficient and very effective for nonnegatively constrained Poisson likelihood estimation problems. It is also guaranteed to converge to the unique minimizer of (15) by Theorem 3.2. We demonstrate the effectiveness of Tikhonov-regularized Poisson likelihood estimation in general, and of the computational method in particular, on simulated astronomical data generated using statistical model (2).

5 Acknowledgements

The authors acknowledge the influence of Professor Anatoly Yagola of Moscow State University. Professor Yagola gave a course on ill-posed problems at the University of Montana during the Fall 2005 semester. Both authors participated in the course, and the idea for this work came to the first author as a result of his participation. The first author must also acknowledge the support of the University of Montana International Exchange Program.

References

- [1] J. M. Bardsley and C. R. Vogel, *A Nonnegatively Constrained Convex Programming Method for Image Reconstruction*, SIAM Journal on Scientific Computing, 25(4), 2004, pp. 1326-1343.
- [2] John B. Conway, *A Course in Functional Analysis*, Second Edition, Springer, 1990.
- [3] W. Feller, *An Introduction to Probability Theory and Its Applications*, Wiley, New York, 1971.
- [4] J. W. Goodman, *Introduction to Fourier Optics, 2nd Edition*, McGraw-Hill, 1996.
- [5] C. T. Kelley, *Iterative Methods for Optimization*, SIAM, Philadelphia, 1999.
- [6] J. J. Moré and G. Toraldo, *On the Solution of Large Quadratic Programming Problems with Bound Constraints*, SIAM Journal on Optimization, 1 (1991), pp. 93-113.
- [7] J. Nocedal and S. Wright, *Numerical Optimization*, Springer 1999.
- [8] D. L. Snyder, A. M. Hammoud, and R. L. White, *Image recovery from data acquired with a charge-coupled-device camera*, Journal of the Optical Society of America A, 10 (1993), pp. 1014-1023.
- [9] D. L. Snyder, C. W. Helstrom, A. D. Lanterman, M. Faisal, and R. L. White, *Compensation for readout noise in CCD images*, Journal of the Optical Society of America A, 12 (1995), pp. 272-283.
- [10] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Kluwer Academic Publishers, 1990.
- [11] C. R. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002.