Minimax Optimal Estimator in a Stochastic Inverse Problem for Exponential Radon Transform

Anuj Abhishek

UNC Charlotte, Charlotte, USA

Abstract

In this article, we consider the problem of inverting the exponential Radon transform of a function in the presence of noise. We propose a kernel estimator to estimate the true function. Such an estimator is closely related to filtered backprojection type inversion formulas in the noise-less setting. For the estimator proposed in this article, we then show that the convergence to the true function is at a minimax optimal rate.

AMS (2000) subject classification. Primary 62G05; Secondary 44A12.

Keywords and phrases. Exponential Radon Transform, Non-parametric estimation

1 Introduction

Exponential Radon transform (ERT), which is the object of study in this article, can be thought of as a generalization of the classical Radon transform. In fact, the ERT of a compactly supported function $f(x)$ in $\mathbb{R}^2$ is given by:

$$T_\mu f(\theta, s) = \int_{x \cdot \theta = s} e^{\mu x \cdot \theta^\perp} f(x) dx. \quad (1.1)$$

Here $s \in \mathbb{R}$, $\theta \in S^1$ where $S^1$ is the unit circle in $\mathbb{R}^2$, $\mu$ is a constant and $\theta^\perp$ denotes a unit vector perpendicular to $\theta$. Recall that lines in $\mathbb{R}^2$ can be parameterized as $L(\theta, s) = \{x : x \cdot \theta = s\}$ where we represent the inner product of two vectors $a$ and $b$ by $a \cdot b$. The classical Radon transform is a special case of ERT and is obtained when $\mu = 0$ in Eq. 1.1. While the classical Radon transform arises in imaging modalities such as X-Ray CT (Computerized Tomography) and PET (Positron Emission Tomography), ERT arises quite naturally in imaging modalities such as SPECT (single
photon emission computed tomography) imaging (Wen and Liang, 2006) and nuclear magnetic resonance imaging (Louis, 1982).

Inversion methods for the exponential Radon transform were derived in Natterer (1979) and in Tretiak and Metz (1980). Authors in Hazou and Solmon (1989) gave filtered backprojection (FBP) type formulas for inversion of ERT using a class of filters. We recall that FBP is an analytical reconstruction method which uses a convolutional filter (e.g. a ‘Ram-Lak filter’ or a ‘Shepp-Logan filter’) to overcome the blurring that is a common feature of backprojection algorithm without filtering. Such FBP type inversion formulas are based on the method of approximate inverse which was developed systematically in Louis and Maass (1990) & Louis (1995). An exhaustive treatment of the method of approximate inverse can be found in the book Schuster (2007). The method of approximate inverse was used to derive Sobolev estimates for attenuated Radon transform in Rigaud and Lakhal (2015), these estimates were central to proving some of the theorems in this article. Furthermore, in Novikov (2002) and Natterer (2001a), the authors gave an inversion formula for the more general attenuated Radon transform. There is extensive literature available on this subject and we give now a partial list of references where an interested reader may find important insights and advances made in the study of exponential and attenuated Radon transforms, see e.g. Aguilar et al. (1996), Bal and Moireau (2004), Boman and Strömberg (2004), Guillement et al. (2002), Holman et al. (2018), Rullgå (2004), Salo and Uhlmann (2011), Shneïberg (1994), & Shneïberg et al. (1994).

Inversion of classical Radon transform has also been extensively studied in the stochastic setting, i.e. when the data obtained is corrupted by noise. Studying the inversion of tomographic problems in a noisy setting is of great practical significance as the data obtained is never free of noise. A detailed discussion of positron emission tomography (PET) in presence of noise can be found in the seminal article (Johnstone and Silverman, 1990). In Hahn and Quinto (1985), the authors established upper and lower bounds for the convergence of two probability measures in terms of the rates of convergence of their Radon transforms. In Korostelëv and Tsybakov (1991) & Korostelëv and Tsybakov (1992), the authors showed that optimal minimax convergence rates are attained by kernel type estimators, which are closely linked to FBP inversion methods. An exhaustive coverage of the non-parametric estimation methods that are used to establish the optimal convergence rates in this article and elsewhere can be founds in the books (Korostelëv and Tsybakov, 1993; Tsybakov, 2009). In Cavalier (1998, 2000), the author obtained results on efficient estimation of density in the non-parametric setting for stochastic
PET problem. However, to the best of our knowledge, a kernel estimator for the stochastic exponential Radon transform has not been proposed before. In this article, we propose a statistical kernel estimator for the ERT problem and show that it attains the minimax optimal rate of convergence. We would like to point out that in addition to the non-parametric kernel type estimators, one can also devise Bayesian estimators for the stochastic problem of X-ray tomography and properties of such Bayesian estimators have been studied in Siltanen et al. (2003), Vänskä et al. (2009), & Monard et al. (2019).

The organization of the article is as follows: in Section 2, we describe the mathematical set up of the stochastic problem for ERT and recall some standard definitions from the literature. In Section 3, we recall the FBP type inversion in the noise-less (deterministic) setting and propose a kernel type estimator. We establish that the estimator is asymptotically unbiased. Finally, in Section 4 we show that this estimator attains minimax optimal rates of convergence.

2 Mathematical Set-Up and Definitions

In this section, we will describe the mathematical framework for the problem and recall some standard definitions from the literature that will help us assess the optimality of the estimator proposed in this article.

Let \( f(x) : \mathbb{R}^2 \to \mathbb{R} \) be a function that satisfies the following assumptions:

**Assumption 1 (A1):** Let \( B^1(x) = \{ x : ||x|| \leq 1 \} \) be the unit ball in \( \mathbb{R}^2 \). We assume that \( f(x) \) is supported in the unit ball \( B^1(x) \). This assumption makes sense in a practical setting as the objects that we would like to image will necessarily have compact supports.

**Assumption 2 (A2):** Let \( \tilde{f}(\xi) \) represent the Fourier transform of \( f(x) \), i.e. \( \tilde{f}(\xi) = \int_{\mathbb{R}^2} f(x) e^{-i\xi \cdot x} dx \). We assume that the Fourier transform of \( f(x) \) satisfies the following inequality,

\[
\int_{\mathbb{R}^2} (1 + ||\xi||^2)^\beta |\tilde{f}(\xi)|^2 d\xi \leq L
\]

for some fixed positive numbers \( L \) and \( \beta > 1 \).

Such an assumption guarantees that the function \( f \) will be necessarily continuous and higher \( \beta \) would imply a higher degree of (Sobolev) smoothness of the function \( f \). We will denote by \( H(\beta, L) \), the class of functions satisfying assumptions A1 and A2. We note also that such assumptions are quite
natural to make in the setting of statistical inverse problems, (Korostelëv
and Tsybakov, 1991, chapter 9).

**Definition 1.** Let $S^1$ denote the unit circle in $\mathbb{R}^2$ and $Z = S^1 \times [-1, 1]$ be the cylinder whose points are given by $(\theta, s)$ where $s \in [-1, 1]$ and $\theta \in S^1$. By $\theta ^\perp$, we will denote a unit vector perpendicular to $\theta$. The exponential Radon transform of $f \in H(\beta, L)$ is defined as the following function on $Z$:

$$T_\mu f(\theta, s) = \int_{x \cdot \theta = s} e^{\mu x \cdot \theta ^\perp} f(x) dx$$

where $\mu$ is a fixed constant. It is clear that if $\mu = 0$, then the exponential Radon transform reduces to the case of the classical Radon transform of such compactly supported functions $f \in H(\beta, L)$.

We note here that for functions $f \in H(\beta, L)$ compactly supported on the unit ball, one can simply consider ERT data on a cylinder $Z$ as above because the ERT vanishes over lines that are more than a unit distance away from the origin, i.e. for $|s| > 1$.

**Definition 2.** Associated to the exponential Radon transform, is its dual transform

$$T_\mu ^* g(x) = \int_{S^1} e^{\mu x \cdot \theta ^\perp} g(\theta, x \cdot \theta) d\theta.$$

Clearly, for $\mu = 0$, this is the backprojection operator for the classical Radon transform.

Now we will describe the stochastic problem of exponential Radon transform. Let $\{(\theta_i, s_i)\}_{i=1}^{n}$ be $n$ random points on the observation space $Z$ and let the observations be of the form:

$$Y_i = T_\mu f(\theta_i, s_i) + \epsilon_i. \quad (2.1)$$

We assume that the points $(\theta_i, s_i)$ are independent and identically distributed (i.i.d.) on $Z$ and $\epsilon_i$ are i.i.d. random variables with zero mean and some finite positive variance $\sigma^2$. The collection of the random points $\{(\theta_i, s_i)\}_{i=1}^{n}$ where observations are made is called the design and will be denoted by $D_n$. Through out this article, we will assume uniform design, i.e. $(\theta_i, s_i)$ will be assumed to have a uniform distribution. In the observation model given by Eq. 2.1, the random variables $\epsilon_i$ account for noise. The stochastic inverse problem for exponential Radon transform is to then estimate the function $f(x)$ based on the observations $Y_i$ for $i = \{1, 2, \ldots, n\}$. This problem is non-parametric in the sense that the function $f$ itself is not assumed to be of any parametric form but is rather assumed to belong to a general class of functions, say $\mathcal{F}$. In this article we have assumed
f \in H(\beta, L). Suppose one devises an estimator \( \hat{f}_n(x) \) based on the observed data. One is then naturally led to ask the question, if this estimator is optimal in some sense? The most popular of such approaches to assess the optimality of estimators in a non-parametric setting is the minimax approach, which we will describe below. Let the nonparametric class of functions \( \mathcal{F} \) be equipped with a semi-norm \( d \). Thus, the semi-distance between two elements \( f, g \in \mathcal{F} \) will be represented as \( d(f, g) \) and we will use the quantity \( d^2(\hat{f}, f) = (d(\hat{f}, f))^2 \) as a measure of error between an estimator \( \hat{f} \) and the true function \( f \). First of all, note that, as any such estimator \( \hat{f}_n(x) \) will depend on the random observation points \( \{(s_i, \theta_i)\}_{i=1}^{n} \) and observations \( \{Y_i\}_{i=1}^{n} \), it is better to consider the expected value of the error between the estimator and the true function (under the chosen semi-norm) as a measure of accuracy. The following definitions are standard in the literature.

**Definition 3** (Johnstone and Silverman (1990) & Tsybakov (2009)). The risk function of an estimator \( \hat{f}_n \) is defined as:

\[
R(\hat{f}_n, f) = E_f (d^2(\hat{f}_n, f)).
\]

From here on, \( E_f \) will be used to denote the expectation with respect to the joint distribution of random variables \( (s_i, \theta_i, Y_i), i = \{1, \ldots, n\} \) satisfying the model given by Eq. 2.1. Ideally, one would like to devise an estimator that would minimize the risk function. However, as the definition of the risk function depends on \( f \) as well, one tries instead to find an overall measure of risk such as the minimax risk.

**Definition 4.** (Tsybakov, 2009, Page 78) Let \( f(x) \) belong to some nonparametric class of functions \( \mathcal{F} \). The maximum risk of an estimator \( \hat{f}_n \) is defined as:

\[
r(\hat{f}_n) = \sup_{f \in \mathcal{F}} R(\hat{f}_n, f).
\]

Finally, the minimax risk on \( \mathcal{F} \) is defined as:

\[
r_n(\mathcal{F}) = \inf_{\hat{f}_n} \sup_{f \in \mathcal{F}} R(\hat{f}_n, f)
\]

where the infimum is taken over the set of all possible estimators \( \hat{f}_n \) of \( f \). Clearly,

\[
r_n(\mathcal{F}) \leq r(\hat{f}_n).
\]

**Definition 5.** (Tsybakov, 2009, Page 78) Let \( \{\Psi^2_n\}_{n=1}^{\infty} \) be a positive sequence converging to zero. An estimator \( \hat{f}_n^* \) is said to be minimax rate optimal if there exist finite positive constants \( C_1 \) and \( C_2 \) such that,

\[
C_1 \Psi^2_n \leq r_n(\mathcal{F}) \leq r(\hat{f}_n^*) \leq C_2 \Psi^2_n.
\]
Furthermore, $\Psi^2_n$ is said to be the optimal rate of convergence.

In this article, whenever we refer to the optimality of an estimator, we will mean its minimax rate optimality. In Section 4, we will propose an estimator for $f(x) \in H(\beta, L)$ based on the model (2.1) and establish its optimality in the following (semi) norms:

1. $d_1(f, g) = |f(x_0) - g(x_0)|$ (where $x_0$ is an arbitrary fixed point in $B^1(x)$)
2. $d_2(f, g) = \left( \int |f(x) - g(x)|^2 \, dx \right)^{1/2}$

as per Definition 5 above. We also note that the risk function defined using semi-norm $d_1$ is called the mean squared error (MSE), while the risk function defined using $d_2$ is referred to in the literature as the mean integrated squared error (MISE) of the estimator. Thus:

$$\text{MSE}(\hat{f}_n, f) = E_f(d_1^2(\hat{f}_n, f)), \quad \text{MISE}(\hat{f}_n) = E_f(d_2^2(\hat{f}_n, f)).$$

Finally, we recall the Kullback distance between two probability measures on a measurable space:

**Definition 6.** (Tsybakov, 2009, Page 84) Let $P$ and $Q$ be two probability measures on some measurable space $(X, \mathcal{A})$. The Kullback distance between the two measures is given by,

$$I(P, Q) = \int \log \frac{dP}{dQ} \, dP \quad \text{if } P \text{ is absolutely continuous with respect to } Q$$

$$= \infty \quad \text{otherwise.}$$

### 3 An asymptotically Unbiased Estimator for Class $H(\beta,L)$

In this section we propose a statistical estimator for $f \in H(\beta, L)$ based on the model (2.1) in the stochastic problem of exponential Radon transform. To motivate the particular form of the proposed estimator, we first recall the FBP reconstruction in the noise-less setting i.e. when the observations as per the model given by Eq. 2.1 are not corrupted by noise. Let $\rho > 0$ such that $0 < |\mu| < 1/\rho$. Consider the function $K_\rho(\theta, s) = K_\rho(s)$ defined as:

$$K_\rho(s) = \frac{1}{\pi} \int_{|\mu|}^{\sqrt{(1/\rho^2) + \mu^2}} r \cos(sr) \, dr.$$  

(3.1)

These kind of functions have been used in the context of filtered backprojection formulas for Radon transforms, see e.g. (Korostelëv and Tsybakov,
Let $I_p(t)$ denote the indicator function:

$$I_p(t) = \begin{cases} 1, & |t| < 1/p \\ 0, & |t| \geq 1/p. \end{cases}$$

The one dimensional Fourier transform of $K_\rho(\theta, s)$ (in the $s$-variable) is:

$$\tilde{K}_\rho(\theta, t) = |t|, \quad |\mu| < |t| < \sqrt{(1/\rho^2) + \mu^2}$$
$$= 0, \quad \text{otherwise.} \quad (3.2)$$

The quantity $1/\rho$ here represents the ‘filter bandwidth’. In fact, to understand why it is called so, assume for simplicity’s sake that $\mu = 0$. The function $\tilde{K}_\rho(\theta, t)$ then acts as a ramp ‘band-pass’ filter in the frequency (Fourier) domain with the bandwidth equal to $1/\rho$. In the following analysis, $\ast$ will represent the operation of convolution of functions. Furthermore, whenever the convolution of two functions $f$ and $g$ defined on the cylinder $Z = S^1 \times \mathbb{R}$ is considered, the convolution will be understood to be taken with respect to their second variable, i.e.

$$f \ast g(\theta, s) = \int_{\mathbb{R}} f(\theta, s - t)g(\theta, t)dt.$$ 

**Theorem 1.** (Natterer, 2001b, Page 49) Let $f_\rho(x) = \frac{1}{4\pi}T^\sharp_{-\mu}(K_\rho \ast T_{\mu}f)$. Then,

$$f(x) = \lim_{\rho \to 0} f_\rho(x).$$

**Proof.** The proof of this theorem is well known, see e.g. (Natterer, 2001b, Section II.6). However, we will reproduce it here for the sake of completeness. First of all recall that from (Natterer, 2001b, (6.2), Page 47), we know that:

$$T^\sharp_{-\mu}(g \ast T_{\mu}f) = (T^\sharp_{-\mu}g) \ast f.$$ 

Thus, if we can show that $\frac{1}{4\pi}T^\sharp_{-\mu}K_\rho$ is an approximate Dirac-delta function, then we are done. Let us then compute:

$$T^\sharp_{-\mu}K_\rho(x) = \int_{S^1} e^{-\mu x \cdot \theta} K_\rho(\theta, x \cdot \theta)d\theta$$

$$= \frac{1}{2\pi} \int_{S^1} e^{-\mu x \cdot \theta} \int_{\mathbb{R}} e^{ix \cdot \theta} K_\rho(\theta, t)dtd\theta$$

$$= \frac{1}{2\pi} \int_{|\mu| < |t| < \sqrt{(1/\rho^2) + \mu^2}} |t| \int_{S^1} e^{-\mu x \cdot \theta + i(x \cdot \theta)t} d\theta dt.$$
In what follows, by $J_0$ we will denote the Bessel function of first kind of integer order 0. Now from (Natterer, 2001b, VII.3.17) $\int_{S^1} e^{-\mu x \cdot \theta + i(x \cdot \theta)t} d\theta = 2\pi J_0(|x|(t^2 - \mu^2)^{1/2})$. Thus,

$$T_{-\mu}^\# K_\rho(x) = \int_{|\mu|<|t|<\sqrt{(1/\rho^2)+\mu^2}} |t| J_0(|x|(t^2 - \mu^2)^{1/2})dt$$

$$= 2 \int_0^{1/\rho} \sigma J_0(|x|\sigma) d\sigma \quad (\sigma = (t^2 - \mu^2)^{1/2})$$

$$= 4\pi \left( \frac{1}{2\pi} \int_0^{1/\rho} \sigma J_0(|x|\sigma) d\sigma \right)$$

$$= 4\pi \delta^{1/\rho}(x) \quad [32, (1.3), Page 183]$$

where

$$\delta^{1/\rho}(x) = \frac{1}{2\pi} \int_{|t|<1/\rho} e^{ix \cdot t}dt = \frac{1}{2\pi} \int_{\mathbb{R}} I_\rho(t)e^{ix \cdot t}dt$$

is an approximate Dirac-delta function that converges to Dirac distribution $\delta(x)$ pointwise (in the space of tempered distributions) as $\rho \to 0$. This completes the proof.

Now we are ready to propose a statistical estimator for $f \in H(\beta, L)$ in the stochastic problem of exponential Radon transform. Inspired by Theorem 1 and similar to the estimator proposed in (Korostelëv and Tsybakov, 1991), let us consider the statistical estimator:

$$f_n^*(x) = \frac{1}{n} \sum_{i=1}^{n} e^{-\mu x \cdot \theta_i} K_{\rho_n}(\langle x \cdot \theta_i \rangle - s_i)Y_i$$

(3.3)

where $\theta_i, s_i$ and $Y_i$ are i.i.d. random variables as per the model (2.1) and $\rho_n \to 0$ as $n \to \infty$. We will call $\rho_n$ as the bandwidth of the estimator. Note that the MSE of the estimator in the non-parametric setting can be broken down in to two terms a “bias term” and a “variance term”:

$$\text{MSE}(f_n^*, f) = E_f[(f_n^*(x) - f(x))^2]$$

$$= (E_f(f_n^*(x)) - f(x))^2 + E_f[(f_n^*(x) - E_f(f_n^*(x)))^2]$$

$$= B_n^2(x) + V_n^2(x)$$

(3.4)

where $B_n(x)$ is the bias of the estimator and $V_n^2(x)$ is its variance. Note that

$$\text{MISE}(f_n^*, f) = ||B_n(x)||_2^2 + ||V_n(x)||_2^2$$
where $\|\cdot\|_2$ denotes $L^2$ norm. Recall that an estimator is said to be asymptotically unbiased if its bias goes to zero pointwise as the number of observations (samples) $n$ grows. We will now show that the estimator proposed above is asymptotically unbiased.

**Theorem 2.** Let $(\theta_i, s_i), i = \{1, \ldots, n\}$ be i.i.d. random variables uniformly distributed on $Z = S^1 \times [-1, 1]$ and these points be independent of the errors $(\epsilon_1, \ldots, \epsilon_n)$. If we consider the kernel estimator $f_n^*(x) = \frac{1}{n} \sum_{i=1}^{n} e^{-\mu x \cdot \theta_i} K_{\rho_n}(\langle x \cdot \theta_i \rangle - s_i)Y_i$, then for each $x$ in the unit ball $B(x)$, the bias term, $B_n(x) = (E(f_n^*(x)) - f(x))$ goes to zero as $n \to \infty$.

**Proof.** It suffices to show that $E(f_n^*(x)) = f_{\rho_n}(x)$ where $f_{\rho_n}(x)$ is given by Theorem (1). Then since $\rho_n \to 0$ as $n \to \infty$, hence $E(f_n^*(x)) = f_{\rho_n} \to f(x)$ pointwise. In what follows, we will say that the i.i.d random variables $\theta_i$ have the same uniform distribution as some random variable $\theta$, all $s_i$ are distributed with the same uniform distribution as some random variable $s$ and similarly $Y$ and $\epsilon$ are random variables with the same distribution as random variables $Y_i$ and $\epsilon_i$ respectively. We will also denote by $E_{\theta,s}(\cdot)$ the expected value of a random variable with respect to the joint distribution of $(\theta, s)$ and by $E_{f_{\theta,s}}(\cdot)$ the conditional expectation of a random variable given $(\theta, s)$. Consider,

\[
E_f(f_n^*(x)) = \frac{1}{n} E_f\left( \sum_{i=1}^{n} e^{-\mu x \cdot \theta_i} K_{\rho_n}(\langle x \cdot \theta_i \rangle - s_i)Y_i \right)
\]

\[
= E_f(e^{-\mu x \cdot \theta} K_{\rho_n}(\langle x \cdot \theta \rangle - s)Y)
\]

\[
= E_{\theta,s}(E_{f_{\theta,s}}(e^{-\mu x \cdot \theta} K_{\rho_n}(\langle x \cdot \theta \rangle - s)(T_{\mu}f(\theta, s) + \epsilon)))
\]

\[
= E_{\theta,s}(E_{f_{\theta,s}}(e^{-\mu x \cdot \theta} K_{\rho_n}(\langle x \cdot \theta \rangle - s)(T_{\mu}f(\theta, s)))(\epsilon \text{ has mean 0})
\]

\[
= E_{\theta,s}(e^{-\mu x \cdot \theta} K_{\rho_n}(\langle x \cdot \theta \rangle - s)(T_{\mu}f(\theta, s)))
\]

\[
= \frac{1}{4\pi} \int_{S^1} e^{-\mu x \cdot \theta} \int_{-1}^{1} K_{\rho_n}(\langle x \cdot \theta \rangle - s)(T_{\mu}f(\theta, s))d\theta ds
\]

\[
= f_{\rho_n}(x).
\]

### 4 Optimality of the Estimator

In this section we will show first of all that while the bias of the estimator decreases as bandwidth goes to zero, the variance increases as bandwidth decreases. Thus an optimal rate of convergence can be obtained by finding a suitable bandwidth $\rho_n$ which balances the bias and the variance term. Furthermore, we will establish the optimality of the proposed estimator under
both semi-norms $d_1$ and $d_2$ as defined in Section 2. Let us now analyze the bias and the variance terms one by one. It is easy to check that for $\beta > 1$ the following relations hold,

\begin{equation}
|I_{\rho_n}(t) - 1| \leq (|t|\rho_n)^\beta \quad (4.1)
\end{equation}

\begin{equation}
|I_{\rho_n}(t) - 1| \leq \left[\frac{2|t|\rho_n}{1+|t|\rho_n}\right]^\beta. \quad (4.2)
\end{equation}

Consider first the bias term, $B_n(x) = f_{\rho_n}(x) - f(x) = \delta^{1/\rho_n} \ast f(x) - f(x)$. Then for any fixed point $x \in B^1(x)$ and $\beta > 1$:

\begin{equation}
B_n(x) = |(\delta^{1/\rho_n} \ast f(x_0) - f(x_0))|
\leq \frac{1}{2\pi} \int_{\mathbb{R}} |(I_{\rho_n}(|\xi|) - 1)||\tilde{f}(\xi)|d\xi
\leq \frac{1}{2\pi} \int_{\mathbb{R}} |\tilde{f}(\xi)||2(|\xi|\rho_n)^\beta/(1 + (|\xi|\rho_n)^\beta)d\xi
\leq \frac{\rho_n^\beta}{\pi} \left[\int_{\mathbb{R}^2}|\tilde{f}(\xi)|^2|\xi|^{2\beta}d\xi\right]^{\frac{1}{2}} \left[\int_{\mathbb{R}^2} (1 + (|\xi|\rho_n)^\beta)^{-2}d\xi\right]^{\frac{1}{2}} \quad \text{(using (8))}
= c_1 \rho_n^{\beta-1}, \quad c_1 > 0. \quad (4.3)
\end{equation}

Anticipating the calculations required to show optimality using norm $d_2$, we also find an estimate for $||B_n(x)||_2^2$.

\begin{equation}
||B_n(x)||_2^2 = ||\delta^{1/\rho_n} \ast f(x) - f(x)||_2^2
\leq \frac{1}{2\pi} \int_{\mathbb{R}^2} |(I_{\rho_n}(|\xi|) - 1)|^2|\tilde{f}(\xi)|^2d\xi \quad \text{(using Parseval’s theorem)}
\leq \frac{1}{2\pi} \int_{\mathbb{R}^2} |\tilde{f}(\xi)|^2(|\xi|\rho_n)^{2\beta}d\xi \quad \text{(using 7)} \quad (4.4)
\leq \frac{L\rho_n^{2\beta}}{2\pi} = c_2 \rho_n^{2\beta} \quad (4.5)
\end{equation}

where $c_2 = L/2\pi$. Now we estimate the variance.

**Lemma 1.** $V_n^2(x) = E_f \left( (f_n^*(x) - E_f(f_n^*(x))^2) \right) \leq c_3/n\rho_n^3$ for $x \in B^1(x)$ and for some constant $c_3 > 0$. From this it also follows that for $x \in B^1(x)$, $||V_n(x)||_2^2 \leq c_4/n\rho_n^3$ for some constant $c_4$. 

Proof. In the following, $\text{Var}$ will denote the variance as per standard notation. First of all, note that $E(f_n^*(x)) = f_{\rho_n}(x)$ and $s_i, \theta_i$ and $Y_i$ are i.i.d. random variables. Thus,

$$V_n^2(x) = E_f\left((f_n^*(x) - E(f_n^*(x)))^2\right)$$

$$= \frac{1}{n}\left(\text{Var}_f(e^{-\mu(x, \theta^*)}K_{\rho_n}(x \cdot \theta - s)T_\mu f(\theta, s)) + \frac{1}{n}\left(E(f_n^*(x))^2\right)\right)$$

$$\leq \frac{\sigma^2 + 4e^{3|\mu|}L^2}{4\pi n} \int_{-1}^{1} e^{-2\mu(x, \theta^*)} \int_{-1}^{1} K_{\rho_n}^2(x \cdot \theta - s) dsd\theta$$

where we use the fact that since $f \in H(\beta, L)$ is compactly supported in $B^1(x)$, we get $|T_\mu f(\theta, s)| \leq 2e^{3|\mu|}L$. Let us now estimate:

$$\int_{-1}^{1} K_{\rho_n}^2(x \cdot \theta - s) ds \leq \int_{-\infty}^{\infty} |K_{\rho_n}(s)|^2 ds$$

$$\leq \int_{-\infty}^{\infty} |\tilde{K}_{\rho_n}(s)|^2 ds$$

$$\leq \frac{1}{3}\left[ ((1/\rho_n^2) + \mu^2)^{3/2} - |\mu|^3 \right]$$

$$= \frac{1}{3}\left[ ((1/\rho_n^2) + \mu^2)^{1/2} - |\mu| \right] \left( ((1/\rho_n^2) + 2\mu^2 + |\mu|) \right)$$

$$\leq \frac{(3 + \sqrt{2})(1/\rho_n^3)}{3}$$

where we have used the fact that we choose $|\mu| \leq (1/\rho_n)$. Thus

$$V_n^2(x) \leq \frac{(3 + \sqrt{2})(\sigma^2 + 4e^{3|\mu|}L^2)}{4\pi n \rho_n^3} \int_{-1}^{1} e^{-2\mu x, \theta^*} d\theta$$

$$\leq \frac{c_3}{n \rho_n^3}$$

for $x \in B^1(x)$ (4.6)

where $c_3 > 0$ is a constant. Now $||V_n(x)||_2^2 = \int_{x \in B^1(x)} V_n^2(x) dx \leq c_4/n \rho_n^3$ for some constant $c_4$.

Theorem 3. Let $f \in H(\beta, L)$ where $\beta > 1$ and $f_n^*(x)$ be the estimator defined in Section 3. Let $\theta_i, s_i$ for $i = 1, \ldots, n$ be i.i.d. uniform random variables and the observation model corresponding to the problem of ERT be given by (2.1). Let $x_0 \in B^1(x)$ be some fixed point. In the the definition of
risk in Section 2, let us use the seminorm $d_1(f, g) = |f(x_0) - g(x_0)|$ where $x_0 \in B^1(x)$ is some arbitrary point. Let $\rho_n = \alpha_1 n^{-1/(2\beta+1)}$ for some constant $\alpha_1$, then the following upper bound holds:

$$\sup_{f \in H(\beta, L)} \psi_n^{-2} \text{MSE}(f^*_n, f) \leq C_0$$

where $\psi_n = n^{-\frac{\beta-1}{2\beta+1}}$. 

**Proof.**

$$\text{MSE}(f^*_n, f) = B_n^2(x_0) + V_n^2(x_0) \leq c_1^2 \rho_n^{2\beta-2} + \frac{c_3}{n \rho_n^{2\beta}}.$$ 

The minimum of the RHS is obtained for $\rho_n^* = \left(\frac{3c_3}{2c_1^2(\beta-1)}\right)^{\frac{1}{2\beta+1}} [n^{-\frac{1}{2\beta+1}}]$. With this choice of $\rho_n = \rho_n^*$, $\text{MSE}(f^*_n, f) = O(n^{-(2\beta-2)/(2\beta+1)})$.

**Theorem 4.** Let $f \in H(\beta, L)$ where $\beta > 1$ and $f_n^*(x)$ be the estimator defined in Section 3. Let $\theta_i, s_i$ for $i = 1, \ldots, n$ be uniform i.i.d. random variables and the observation model corresponding to the problem of ERT be given by Eq. 2.1. Consider the seminorm given by $d_2(f, g) = \|f - g\|_2$ where $\|(\cdot)\|_2$ indicates the $L_2$ norm as usual. Let $\rho_n = \alpha_2 n^{-1/(2\beta+3)}$, where $\alpha_2$ is a constant. Then the following upper bound holds,

$$\sup_{f \in H(\beta, L)} \psi_n^{-2} \text{MISE}(f^*_n, f) \leq C_1$$

where $\psi_n = n^{-\beta/(2\beta+3)}$ and a positive constant $C_1$.

**Proof.**

$$\text{MISE}(f^*_n, f) = \|B_n(x)\|_2^2 + \|V_n(x)\|_2^2 \leq c_2 \rho_n^{2\beta} + c_4/n \rho_n^3.$$ 

Note that the minimum of the RHS above is attained for $\rho_n^* = \left(\frac{3c_3}{2c_1^2(\beta-1)}\right)^{\frac{1}{2\beta+3}} [n^{-\frac{1}{2\beta+3}}]$. With this choice of $\rho_n = \rho_n^*$, $\text{MISE}(f^*_n, f) = O(n^{-2\beta/(2\beta+3)})$. This completes our proof.

The upper bounds established in Theorems 3 and 4 above imply that the minimax risks for the estimator using the two seminorms $d_2$ and $d_1$ is bounded above by $C_1 \Psi_n^2$ and $C_2 \psi_n^2$ respectively where $\Psi_n$ and $\psi_n$ are sequences that go to zero as $n \to \infty$. As per Definition (5), to establish the optimality of the estimator we need to show that each of the two minimax risks also satisfy the corresponding lower bounds. To that end, at first we make the following additional assumptions for the observation model Eq. 2.1:
Assumption on the distribution of noise (B1): The random variables $\epsilon_i$ are i.i.d having a distribution $G(\cdot)$ that satisfies:

$$\int_{-\infty}^{\infty} \ln \frac{dG(u)}{dG(u+v)} dG(u) \leq I_0 v^2, \quad |v| \leq v_0$$  \hspace{1cm} (4.7)

where $I_0 > 0$ and $v_0 > 0$ are some constants. Note that, this assumption on the distribution of noise has the effect of saying that the noise is ‘nearly’ Gaussian. Indeed, it can be easily shown (see, (Tsybakov, 2009, corollary 2.1)) that, if $\phi(x)$ is the density of the standard Normal distribution $\mathcal{N}(0, 1)$, then

$$\int_{-\infty}^{\infty} \ln \frac{\phi(x)}{\phi(x+t)} \phi(x) dx = \frac{t^2}{2}, \quad t \in \mathbb{R}.$$  

Theorem 5. Let $\beta, f, f_1^*, \theta_i, s_i$ as in Theorem 3. If in addition, assumption B1 is satisfied, then the following inequality holds:

$$\liminf_{n \to \infty} \inf_{\hat{f}_n} \sup_{f \in H(\beta, L)} \psi_n^{-2} \text{MSE}(\hat{f}_n, f) \geq c_0$$

where $\psi_n$ is the same sequence as in Theorem 3, $\inf$ denotes the infimum over all estimators and $c_0 > 0$ is some constant.

Proof. The proof method follows that in (Korostelëv and Tsybakov, 1991, Theorem 4) and we will adapt their proof wherever needed. As noted there, using standard reduction techniques for establishing lower bounds on the minimax risk of regression estimators in a non-parametric setting, the problem can be reduced to showing that the Kullback distance between the two probability measures corresponding to two appropriately chosen functions (hypothesis) is bounded, see also (Tsybakov, 2009, section 2.5). Thus consider the functions (hypothesis) $f_0(x) = 0$ and $f_1(x) = Ah^{\beta - 1} \eta_0((x - x_0)/h)$ where $h = n^{-\frac{1}{2\beta + 1}}$, $\eta_0(x) \in H(\beta, L)$ is a compactly supported bounded function such that $\eta_0(0) > 0$ and $0 < A < 1$ is a constant. Following (Korostelëv and Tsybakov, 1991), we will first show that $f_1(x) \in H(\beta, L)$. Note that:

$$\tilde{f}_1(\xi) = Ah^{\beta - 1} \int_{\eta_0((x - x_0)/h)} e^{i\xi x} dx = Ah^{\beta - 1} e^{i\xi x_0} \int_{\mathbb{S}^1} \int_{0}^{\infty} (u) \eta_0(u\theta/h) e^{i\xi u\theta} dud\theta$$

$$= Ah^{\beta + 1} e^{i\xi x_0} \int_{\mathbb{S}^1} (\tilde{u}) \eta_0(\tilde{u}\theta) e^{i(h\xi - \theta)\tilde{u}} d\tilde{u} \theta = Ah^{\beta + 1} e^{i\xi x_0} \tilde{\eta}_0(h\xi).$$
Thus,
\[
\int (1 + |\xi^2|)^\beta |\tilde{f}_1(\xi)|^2 d\xi = A^2 h^{2(\beta+1)} \int (1 + |\xi^2|)^\beta |\tilde{\eta}_0(h\xi)|^2 d\xi
\]
\[
= A^2 \int (h^2 + |\xi|^2)^\beta |\tilde{\eta}_0(\xi)|^2 d\xi \leq L
\]

where we have used the fact that $0 < h, A < 1$ and $\eta_0(x) \in H(\beta, L)$. Also observe that $|f_1(x_0) - f_0(x_0)| = Ah^{\beta-1}\eta_0(0)$ and $\eta_0(0) > 0$ by assumption.

Now let $P_0$ and $P_1$ be probability measures corresponding to the experiments with observations given by the regression model (2.1) for $f = f_0$ and $f = f_1$ respectively and $p_0$ and $p_1$ be the densities corresponding to the measures $P_0$ and $P_1$ respectively. Then to complete the proof of the theorem it suffices to show the Kullback information distance between the two measures, $I(P_0, P_1) \leq 1/2$, see (Korostelëv and Tsybakov, 1991, equation (14)). Now consider,
\[
I(P_0, P_1) = \int \ln \left( \frac{dP_0}{dP_1} \right) dP_0 = E_{f_0} \int \ln \left( \frac{dP_0}{dP_1} \right) d\nu \quad (\nu \text{ is the Lebesgue measure})
\]
\[
= E_{(\theta, s)} \left( E_{f_0(\theta, s)} \int \ln \left( \frac{dP_0}{dP_1} \right) d\nu \right)
\]
\[
= E_{(\theta, s)} \left[ \sum_{i=1}^n \int \ln \left( \frac{dG(v - T_{\mu}f_0(\theta_i, s_i))}{dG(v - T_{\mu}f_1(\theta_i, s_i))} \right) dG(v - T_{\mu}f_0(\theta_i, s_i)) \right] (26, (2.36))
\]
\[
\leq C_3 n I_{f_0} \int |T_{\mu}f_1(\theta, s)|^2 ds d\theta \quad \text{(using B1).} \quad (4.8)
\]

To estimate $\int Z |T_{\mu}f_1(\theta, s)|^2 ds d\theta$, we will follow (Rigaud and Lakhal, 2015, section 4). Consider a function $\phi(x) \in S(\mathbb{R}^2)$ (i.e. Schwartz class) such that $\phi(x) = 1$ for $x \in B^1(x)$. Let us introduce
\[
\tilde{\phi}(x, \theta) = \phi(x)e^{\mu x \cdot \theta \perp} \quad (4.9)
\]
Clearly for any function $f_1(x)$ supported in $B^1(x)$,
\[
T_{\mu}f_1(\theta, s) = T\tilde{\phi}f_1(\theta, s) = \int_{\mathbb{R}^2} \tilde{\phi}(x, \theta) f(x) \delta(x \cdot \theta - s) dx.
\]
Taking the Fourier transform of $T\tilde{\phi}f_1(\theta, s)$ with respect to the $s$-variable we get the following inequality (Rigaud and Lakhal, 2015, equation 27),
\[
|\tilde{T}\tilde{\phi}f(\theta, t)|^2 \leq (2\pi)^{-1} |W_{\tilde{\phi}} \ast \tilde{f}(\xi)|^2 \quad (4.10)
\]
where $W_{\tilde{w}} = \sup_{\theta \in S^1} |\tilde{w}(\theta, t)|$ and $(\cdot)$ indicates the corresponding Fourier transform (either 1-d or 2-d) as usual. Now from (Rigaud and Lakhal, 2015, equation 29),

$$||T_{\mu} f_1(\theta, s)||^2_{L^2(Z)} \leq ||T_{\mu} f_1(\theta, s)||^2_{H^{1/2}(Z)} \leq K||W_{\tilde{w}}||^2_{L^1(\mathbb{R}^2)} ||f_1||^2_{L^2(\mathbb{R}^2)} = \tilde{K} ||f_1||^2_{L^2(\mathbb{R}^2)} \quad (4.11)$$

where $\tilde{K} = K||W_{\tilde{w}}||_{L^1(\mathbb{R}^2)}$. We note in passing that since $\tilde{w}(x, \theta)$ is given by (4.9), $||W_{\tilde{w}}||_{L^1(\mathbb{R}^2)}$ is finite.

Now $||f_1||^2_{L^2(\mathbb{R}^2)} = A^2 h^{2\beta - 2} \int_{\mathbb{R}^2} |\eta_0((x-x_0)/h)|^2 dx = A^2 h^{2\beta + 1} \int_{\mathbb{R}^2} |\eta_0(y)|^2 dy$.

Since $\eta_0 \in H(\beta, L)$ is compactly supported bounded function, thus $||\eta_0(y)||^2_2$ is finite. Thus,

$$I(P_0, P_1) \leq C_3 I_0 \tilde{K} A^2 ||\eta_0(y)||^2_2 nh^{2\beta + 1} = C_3 I_0 \tilde{K} A^2 ||\eta_0(y)||^2_2 \quad (h = n^{-\frac{1}{2\beta+1}}). \quad (4.12)$$

Thus if we choose $A$ to be small enough, $I(P_0, P_1) \leq 1/2$.

**Remark 1.** Note that Theorems 3 and 5 together establish the optimality of the convergence rate of minimax risk for the estimator proposed in Section 3 under the seminorm $d_1$.

**Theorem 6.** Let $\beta, f, f^*_n, \theta_i, s_i$ as in Theorem 4. If in addition, assumption B1 is satisfied by the observation model (2.1) then the following inequality holds:

$$\liminf_{n \to \infty} \inf_{f_n} \sup_{f \in H(\beta, L)} \Psi_n^{-2} MISE(\hat{f}_n, f) \geq c_1$$

where $\Psi_n$ is the same sequence as in Theorem 4, $\inf$ denotes the infimum over all estimators and $c_1 > 0$ is some constant.

**Proof.** First of all, we recall from (Tsybakov, 2009, section 2.6) that to establish lower bounds for the convergence rate of the estimators in $L_p$ seminorms requires us to work with many hypotheses (M-hypotheses) instead of just two as we did in the proof of Theorem 5 above. The proof of this theorem follows that of (Korostelev and Tsybakov, 1991, Theorem 5). All the geometric arguments in this proof are identical to the geometrical arguments in Korostelev and Tsybakov (1991) and we only need to change the argument wherever an estimate for the usual Radon transform is to be replaced with an analogous estimate for the exponential Radon transform. For the sake of completeness, we outline the proof given in Korostelev and Tsybakov (1991) here, adapting it to the case of ERT wherever needed.
Consider a collection of non-intersecting balls $\Delta_k, k \in \{1, \ldots, M\}$ inscribed in $B_1^1(x)$ with center $a_k$ and of radius $1/m$ such that $m$ and $M$ are sequences and $m \to \infty$ as $n \to \infty$. Furthermore, one can choose $m$ and $M$ (the precise choice for $m$ is described later) such that the following relation is satisfied:

$$C_4 m^2 \leq M \leq C_5 m^2.$$ (4.13)

Let $\eta(x)$ be a smooth function supported in $B_1^1(x)$. Then each function $\eta_k(x) = \eta(m(x - a_k))$ is supported respectively in $\Delta_k$. To each $m$-tuple $b = (b_1, \ldots, b_m)$ where $b_k$ is either 0 or 1, we associate a function $f(x, b)$ supported in $B_1^1(x)$ such that:

$$f(x, b) = Am^{-\beta} \sum_{k=1}^{M} b_k \eta_k(x)$$

where $A > 0$ will be chosen in a manner described below. We state without proof the following two lemmas from (Korostel’ev and Tsybakov, 1991):

**Lemma 2.** (Korostel’ev and Tsybakov, 1991, Lemma 3) There exists $A_\beta > 0$ such that for $A < A_\beta$, the function $f(x, b) \in H(\beta, L)$ for any $m$-tuple $b$.

Consider any design $D_n = \{(\theta_i, s_i)\}_{i=1}^{n}$, and consider the lines $L_i = \{x \in \mathbb{R}^2 : x \cdot \theta_i = s_i\}$. Take the set of balls $\Delta_k$ such that each ball intersects at most $C_6 n/m$ lines where $C_6 > 0$ is a constant, whose choice is described in Lemma 3 below. Let the set of indices $J$ be defined as:

$$J = J(D_n) = \{k \in \{1, \ldots, M\} : \text{number of lines corresponding to } D_n \text{ that intersect with } \Delta_k \text{ is less than or equal to } C_6 n/m\}.$$ 

**Lemma 3.** (Korostel’ev and Tsybakov, 1991, Lemma 4) There exists $C_6 > 0$ such that for any design $D_n$, we have the inequality:

$$\text{card}J > M/2.$$ 

In what follows, $C_6$ is chosen such that Lemma 3 is satisfied.

Following Korostel’ev and Tsybakov (1991), we use, $b^{(k,0)}_0 = \{b_1, \ldots, b_{k-1}, 0, b_{k+1}, \ldots, b_M\}$ and $b^{(k,1)}_0 = \{b_1, \ldots, b_{k-1}, 0, b_{k+1}, \ldots, b_M\}$ to indicate $M$-tuples with fixed $k$-th elements. Furthermore, we use the following notation for functions:

$$f_{k_0} = f(x, b^{(k,0)}_0) \quad \text{and} \quad f_{k_1} = f(x, b^{(k,1)}_0).$$
Let $g_k(x) = f_{k_0}(x) - f_{k_1}(x)$ which is supported only on $\Delta_k$ by construction. Let $P_{k_0}$ and $P_{k_1}$ be the probability measures corresponding to the model 2.1 for $f = f_{k_0}$ and $f = f_{k_1}$. Let $I(P_{k_0}, P_{k_1})$ be the Kullback information distance between these two probability measures. Thus from (Korostelëv and Tsybakov, 1991, section 3), the desired lower bound for the minimax rate will be obtained if we can show that for a sufficiently small $C_8 > 0$ such that $m = (C_8 n)^{\frac{1}{2\beta+3}}$, $I(P_{k_0}, P_{k_1}) < 1/2$. Just as in (Korostelëv and Tsybakov, 1991) and similar to the proof of Theorem 5 above, from assumption B1, we get:

$$I(P_{k_0}, P_{k_1}) \leq I_0 \sum_{i=1}^{n} (T_{\mu} g_k(\theta_i, s_i))^2. \tag{4.14}$$

Now from the definition of ERT and from the fact that $\eta_k(x)$ is supported in $\Delta_k \subset B^1(x)$,

$$|E_{\mu} g_k(\theta_i, s_i)| = \left| \int_{L_i \cap \Delta_k} e^{\mu x \cdot \theta_i^+} A^{-\beta} \eta(m(x - a_k)) dx \right| \leq C_9 \int_{L_i \cap \Delta_k} |A^{-\beta} \eta(m(x - a_k))| dx \quad (C_9 = \sup_{x \in B^1(x)} e^{\mu x \cdot \theta^+})$$

$$\leq C_{10} m^{-\beta - 1}. \tag{4.15}$$

Now note that since $k \in J$, thus at most $C_6 n/m$ of the terms in the sum on RHS of (4.14) are non zero. Putting it all together, we have:

$$I(P_{k_0}, P_{k_1}) \leq I_0 C_6 (C_{10})^2 \frac{n}{m} m^{-2\beta - 2} \leq I_0 C_6 C_{10}^2 C_8. \tag{4.16}$$

Thus if we choose $C_8 \leq \frac{I_0 C_6 C_{10}^2}{2}$, then we get $I(P_{k_0}, P_{k_1}) \leq 1/2$ as desired. This completes the proof of the theorem.

Remark 2. Note that Theorems 4 and 6 together establish the optimality of the estimator in the $d_2$ semi-norm setting.

Compliance with Ethical Standards. There is no conflict of interest or funding information to report for this article.

References

AGUILAR, V., EHRENPRIES, L. and KUCHMENT, P. (1996). Range conditions for the exponential Radon transform. J. Anal. Math. 68, 1–13. https://doi.org/10.1007/BF02790201

BAL, G. and MOIREAU, P. (2004). Fast numerical inversion of the attenuated Radon transform with full and partial measurements. Inverse Problems 20, 4, 1137–1164. https://doi.org/10.1088/0266-5611/20/4/009
boman, J. and Strömberg, J.-O. (2004). Novikov’s inversion formula for the attenuated Radon transform—a new approach. J. Geom. Anal. 14, 2, 185–198. https://doi.org/10.1007/BF02922067

Cavalier, L. (1998). Asymptotically efficient estimation in a problem related to tomography. Math. Methods Statist. 7, 4, 445–4561999.

Cavalier, L. (2000). Efficient estimation of a density in a problem of tomography. Ann. Statist. 28, 2, 630–647. https://doi.org/10.1214/aos/1016218233

Guillement, J. -P., Jauberteau, F., Kunyansky, L., Novikov, R. and Trebossen, R. (2002). On single-photon emission computed tomography imaging based on an exact formula for the nonuniform attenuation correction. Inverse Problems 18, 6, 11–19. https://doi.org/10.1088/0266-5611/18/6/101

Hahn, M. G. and Quinto, E. T. (1985). Distances between measures from 1-dimensional projections as implied by continuity of the inverse Radon transform. Z. Wahrsch. Verw. Gebiete 70, 3, 361–380. https://doi.org/10.1007/BF00534869

Hazou, I. A. and Solmon, D. C. (1989). Filtered-backprojection and the exponential Radon transform. J. Math. Anal. Appl. 141, 1, 109–119. https://doi.org/10.1016/0022-247X(89)90209-6

Holman, S, Monard, F and Stefanov, P (2018). The attenuated geodesic x-ray transform. Inverse Problems 34, 6, 064003–26. https://doi.org/10.1088/1361-6420/aab8bc

Johnstone, I. M. and Silverman, B. W. (1990). Speed of estimation in positron emission tomography and related inverse problems. Ann. Statist. 18, 1, 251–280. https://doi.org/10.1214/aos/1176347500

Korostelëv, A. P. and Tsybakov, A. B. (1991). Optimal rates of convergence of estimators in a probabilistic setup of tomography problem. Probl. Inf. Transm. 27, 73–81.

Korostelëv, A.P. and Tsybakov, A.B. (1992). Asymptotically minimax image reconstruction problems. In: Topics in nonparametric estimation. Adv. Soviet math., vol. 12, pp. 45–86. Amer. Math. Soc., providence, RI, ???.

Korostelëv, A.P. and Tsybakov, A. B. (1993). Minimax Theory of Image Reconstruction. Lecture Notes in Statistics, vol. 82, p. 258. Springer ???. https://doi.org/10.1007/978-1-4612-2712-0

Louis, A. K. (1982). Optimal sampling in nuclear magnetic resonance (NMR) tomography. J. Comput. Assist. Tomogr. 6, 2, 334–340. https://doi.org/10.1097/00004728-198204000-00019

Louis, A. K. (1995). Approximate inverse for linear and some nonlinear problems. Inverse Problems 11, 6, 1211–1223.

Louis, A. K. and Maass, P. (1990). A mollifier method for linear operator equations of the first kind. Inverse Problems 6, 3, 427–440.

Monard, F., Nickl, R. and Paternain, G. P. (2019). Efficient nonparametric Bayesian inference for X-ray transforms. Ann. Statist. 47, 2, 1113–1147. https://doi.org/10.1214/18-AOS1708

Natterer, F. (1979). On the inversion of the attenuated Radon transform. Numer. Math. 32, 4, 431–438. https://doi.org/10.1007/BF01401046

Natterer, F. (2001a). Inversion of the attenuated Radon transform. Inverse Problems 17, 1, 113–119. https://doi.org/10.1088/0266-5611/17/1/309

Natterer, F. (2001b). The mathematics of computerized tomography. Society for industrial and applied mathematics ???. https://doi.org/10.1137/1.9780898719284

Novikov, R.G. (2002). An inversion formula for the attenuated x-ray transformation. Arkiv för Matematik 40, 1, 145–167. https://doi.org/10.1007/bf02384507
RIGAUD, G. and LAKHAL, A. (2015). Approximate inverse and Sobolev estimates for the attenuated Radon transform. Inverse Problems 31, 10, 105010–21. https://doi.org/10.1088/0266-5611/31/10/105010

RULLGÅ, R.D.H. (2004). An explicit inversion formula for the exponential Radon transform using data from 180°. Ark. Mat. 42, 2, 353–362. https://doi.org/10.1007/BF02385485

SALO, M. and UHLMANN, G. (2011). The attenuated ray transform on simple surfaces. J. Differential Geom. 88, 1, 161–187.

SCHUSTER, T. (2007). The Method of Approximate inverse: Theory and Applications. Lecture Notes in Mathematics, vol. 1906, p. 198. Springer.

SHNEİBERG, I.Y. (1994). Exponential Radon transform. In: Applied problems of radon transform. Amer. Math. Soc. Transl. Ser. 2, vol. 162, pp. 235–245. Amer. Math. Soc., providence, RI.

SHNEİBERG, I.Y., PONOMAREV, I.V., DMITRICHENKO, V.A. and KALASHNIKOV, S.D. (1994). On a new reconstruction algorithm in emission tomography. In: Applied problems of radon transform. Amer. Math. Soc. Transl. Ser. 2, vol. 162, pp. 247–255. Amer. Math. Soc., providence, RI.

SILTANEN, S., KOLEHMAINEN, V., RVENP, S.J., KAIPIO, J.P., KOISTINEN, P., LASSAS, M., PIRTILT, J. and SOMERSALO, E. (2003). Statistical inversion for medical x-ray tomography with few radiographs: i. general theory. Phys. Med. Biol. 48, 10, 1437–1463. https://doi.org/10.1088/0031-9155/48/10/314

TRETIAK, O. and METZ, C. (1980). The exponential Radon transform. SIAM. J. Appl. Math. 39, 2, 341–354. https://doi.org/10.1137/0139029

TSYBAKOV, A. B. (2009). Introduction to Nonparametric Estimation. Springer Series in Statistics. Springer, Berlin, p. 214. https://doi.org/10.1007/b13794. Revised and extended from the 2004 French original Translated by Vladimir Zaiats. https://doi.org/10.1007/b13794

VÄNSKÄ, S., LASSAS, M. and SILTANEN, S. (2009). Statistical X-ray tomography using empirical Besov priors. Int. J. Tomogr. Stat. 11, S09, 3–32.

WEN, J. and LIANG, Z. (2006). An inversion formula for the exponential radon transform in spatial domain with variable focal-length fan-beam collimation geometry. Med. Phys. 33, 3, 792–798. https://doi.org/10.1118/1.2170596

Publisher’s Note. Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ANUJ ABHISHEK
DEPARTMENT OF MATHEMATICS AND
STATISTICS, UNC CHARLOTTE, 9201
UNIVERSITY CITY BLVD., CHARLOTTE,
28213, NORTH CAROLINA, USA
E-mail: anuj.abhishek@uncc.edu

Paper received: 27 September 2021; accepted 10 March 2022.