FF

LBL-37032 Rev. 2
UC-405
Preprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

Submitted to SIAM Journal on Applied Mathematics

Efficient Estimation of Linear Functionals in Emission Tomography

A. Kuruc

February 1996

Donner Laboratory

Biology & Medicine Division

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

LBL-37032 Rev. 2
UC-405

Efficient Estimation of Linear Functionals in Emission Tomography

Alvin Kuruc

Life Sciences Division
Ernest Orlando Lawrence Berkeley National Laboratory
University of California
Berkeley, California 94720

February 1996

This work was supported in part by the Director, Office of Energy Research, Office of Health and Environmental Research, Medical Applications and Biophysical Research Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098, and by NIH Training Grant HL07367 and National Heart, Lung, and Blood Institute Grant P01-HL25840.

EFFICIENT ESTIMATION OF LINEAR FUNCTIONALS IN EMISSION TOMOGRAPHY *

ALVIN KURUC †

Abstract. In emission tomography, the spatial distribution of a radioactive tracer is estimated from a finite sample of externally-detected photons. We present an algorithm-independent theory of statistical accuracy attainable in emission tomography that makes minimal assumptions about the underlying image. Let f denote the tracer density as a function of position (i.e., f is the image being estimated). We consider the problem of estimating the linear functional $\Phi(f) \equiv \int \phi(x)f(x)\,dx$, where ϕ is a smooth function, from n independent observations identically distributed according to the Radon transform of f. Assuming only that f is bounded above and below away from 0, we construct statistically efficient estimators for $\Phi(f)$. By definition, the variance of the efficient estimator is a best-possible lower bound (depending on ϕ and f) on the variance of unbiased estimators of $\Phi(f)$. Our results show that, in general, the efficient estimator will have a smaller variance than the standard estimator based on the filtered-backprojection reconstruction algorithm. The improvement in performance is obtained by exploiting the range properties of the Radon transform.

Key words. Nonparametric estimation, Inverse problems, Ill-posed problems

AMS subject classifications. 92C55, 62G05, 44A12

1. Introduction. In emission tomography (ET), the goal is to characterize the density, f, of a radioactive tracer in a subject as a function of position by external detection of emitted photons. In this paper, we construct statistically efficient, i.e., minimum-variance unbiased, estimators of the linear functional $\Phi(f) \equiv \int \phi(x)f(x)\,dx$, where ϕ is a smooth function. This problem is motivated, for example, by the problem of quantifying the amount of tracer in a region of interest.

Our results can be summarized as follows. The standard estimator for $\Phi(f)$, based on the filtered backprojection (FB) reconstruction algorithm, is unbiased and may be expressed as a linear estimator of the form $n^{-1}\sum_{i=1}^{n}\psi(l_i)$, where ψ is a function on the observation space and the l_i are the observations. The efficient estimator is the linear estimator generated by the projection of ψ onto the range of the Radon transform, viewed in a suitable function space. In general, its variance is smaller than that of the standard estimator. Numerical results are given for the case where ϕ is a Gaussian density function.

1.1. Mathematical Model of ET. We start by proposing a simple mathematical model of ET. The model is highly idealized in that it ignores numerous secondary physical effects that occur in practice. However, it abstracts the basic problem of ET.

We consider the problem of characterizing the density, f, of a radioactive tracer on the unit (radius) disk $D \subset \mathbb{R}^2$, where $\mathbb{R}^2$ denotes 2-dimensional Euclidean space. A radioactive disintegration occurring at $x \in D$ results in the emission of one or (in the case of positron emitters) two photons which travel along a random line through x with uniformly distributed random orientation. (Positron emitters give off two photons that travel in antipodal directions, hence along the same line.) In most imaging systems, only photons traveling along lines lying in the plane of D are detected. We

*This work was supported in part by the Director, Office of Energy Research, Office of Health and Environmental Research, Medical Applications and Biophysical Research Division of the U.S. Department of Energy under contract No. DE-AC03-76SF00098 and in part by NIH training grant HL07367 and National Heart, Lung, and Blood Institute grant P01-HL25840.

†MS 55-121, Lawrence Berkeley National Laboratory, University of California, 1 Cyclotron Road, Berkeley, CA 94720 (A_Kuruc@lbl.gov).

will therefore consider the observations in ET to consist of these lines. In other words, we will ignore the 3-dimensional aspect of the problem and treat it as a problem in 2 dimensions.

We assume that f is normalized to unit area. The locations of the radioactive disintegrations are modeled as independent, identically distributed (i.i.d.) random variables with probability density function (p.d.f.) f. The observations are modeled as random lines in $\mathbb{R}^2$ through the locations of the radioactive disintegrations with uniformly distributed orientation.

Let $\mathbb{L}$ denote the set of lines in $\mathbb{R}^2$. We define the Radon transform of f to be the function $Rf\colon \mathbb{L} \to \mathbb{R}$ whose value at $l \in \mathbb{L}$ is the integral of f over l. We put coordinates on $\mathbb{L}$ by defining $\vec{\theta} \equiv (\cos\theta, \sin\theta) \in \mathbb{R}^2$ and assigning the coordinates (θ, s) to the line through $s\vec{\theta}$ that is perpendicular to $\vec{\theta}$. In this coordinate system, the observations in our model of ET are i.i.d. $\mathbb{L}$-valued random variables with p.d.f. Rf with respect to the dominating measure $\pi^{-1}\, ds\, d\theta$ on $\mathbb{L}$ [17, sec. 2.1] [20, sec. 2.3].

REMARK 1.1. Note that f is defined to be a p.d.f. on locations of radioactive disintegrations; it contains no information about the rate of disintegrations. (One can think of the p.d.f. as being obtained by dividing the disintegration rate per unit area by the total count rate.) Similarly, the observations are taken to be a sequence of elements of $\mathbb{L}$, there is no time information. Thus the way we have set up the problem defines away the problem of estimating the total count rate. This explains why the familiar Poisson distribution does not appear in our model. In practice, one would like to know the total count rate, but good estimates for this quantity are easy to construct.

1.2. Linear Functionals. For $\phi : \mathbb{R}^2 \to \mathbb{R}$, we term the function

$$f \mapsto \Phi(f) \equiv \int_{\mathbb{R}^2} \phi(x) f(x)\, dx$$

the linear functional generated by ϕ. We shall consider the problem of estimating $\Phi(f)$ from n i.i.d. observations distributed according to Rf.

A natural quantity of interest in ET is the fraction of the total tracer contained in some subset $S \subset D$, i.e., the linear functional generated by the indicator function of S. However, it turns out that estimating this quantity is a statistically ill-posed problem without strong regularity conditions on f. (Precise results along these lines may be found in [20, sec. 6.2] and [21, prop. 7.1].) To obtain a well-posed problem without strong assumptions about f, we therefore need to place some regularity conditions on ϕ. In what follows, we will assume that ϕ is a smooth, i.e., infinitely differentiable, function.

REMARK 1.2. Choosing ϕ to be a narrow Gaussian density gives a well-posed approximation to the problem of estimating the tracer density at a point. In this context, ϕ is sometimes referred to as an aperture function. The convolution of an indicator function with, for example, a narrow Gaussian density gives an approximation of the indicator function by a smooth function. One can therefore apply the results in this paper, in an approximate way, to the problem of estimating the fraction of tracer in a region of interest. Another obvious application is to reconstruction algorithms where the image is given by a series expansion with coefficients that are linear functionals of the image.

1.3. Image Model. We now need to specify our assumptions regarding the unknown p.d.f. f. We will assume only that f is a p.d.f. on $D \subset \mathbb{R}^2$ and that f is bounded above and below away from 0. We will denote the set of p.d.f.s satisfying this condition by $\mathcal{P}$.

REMARK 1.3. In our opinion, the restriction $f \in \mathcal{P}$ is sufficiently weak to cover almost all practical applications. The assumption that the density function is bounded above corresponds to the physical condition that the concentration of tracer in the subject is bounded above. Since this upper bound can be arbitrarily large, this condition will be satisfied in any practical application. The assumption that the density function is bounded below away from 0 is perhaps more problematic, but since this lower bound can be chosen to be arbitrarily small, we do not believe it alters the essence of the problem. Physically, one can think of it as postulating some positive level of background radiation. Alternatively, one could alter the experiment by adding some artificial observations that mimic those that would be obtained from a low-intensity uniform distribution.

1.4. Statistical Framework. We shall consider the statistical problem of estimating $\Phi(f)$ given n i.i.d. observations distributed according to Rf. For comparison, we shall also consider the simpler statistical problem of estimating $\Phi(f)$ given n i.i.d. observations distributed according to f itself. In other words, the observations are the locations of the radioactive disintegrations. We shall term the latter problem the planar-imaging problem.

In what follows, we shall construct efficient estimators for both the ET and planar imaging problems and compute their performance. By an efficient estimator, we mean an unbiased estimator whose variance at some $f \in \mathcal{P}$ is minimal in the class of all unbiased estimators. By definition, the variance of these estimators gives a best-possible lower bound on the variance attainable by any unbiased estimator. It therefore provides an algorithm-independent measure of how well $\Phi(f)$ may be estimated. This bound can be used as a benchmark in assessing the performance of image reconstruction and quantification algorithms. Appropriately generalized, it could also be used as a design tool for assessing the performance that is achievable by new imaging devices.

REMARK 1.4. The efficient estimator at f will be seen to depend on f. Since, in practice, f is unknown, one cannot attain the bound simply by just applying the efficient estimator. However, as will be discussed in §7, the analysis gives insight into how one might construct practical estimators whose performance approaches that of the efficient estimator.

1.5. Relation to Previous Work. During the early years of computed tomography, a number of authors characterized the propagation of measurement errors of specific reconstruction algorithms for both x-ray [4] [14] and emission [5] tomography. As a result, the statistical performance of the standard linear reconstruction algorithms, such as the FB algorithm, is fairly well understood. There have also been several papers on algorithm-independent lower bounds for the estimation of linear functionals in tomography. Tretiak [26] derived such a bound for the linear functional generated by a Gaussian aperture function centered at the origin in a model of x-ray computed tomography. He considered the case where f is the uniform distribution on D and showed that the FB algorithm comes close to attaining the bound in this case. More recently, Bickel and Ritov [3] considered the estimation of linear functionals using the same model of emission tomography considered here. However, they incorrectly assumed that the estimator based on the FB reconstruction algorithm is efficient. In this paper, we show that this estimator is not efficient except in certain special cases. One such special case is when both f and ϕ are radially symmetric, as was the case in the problem treated in [26]. In general, the construction of an efficient estimator needs to take into account the range properties of the Radon transform. The analysis of this point plays a large role in this paper.

1.6. Outline of Paper. In §2, we show that the non-tomographic planar imaging problem is rather trivial in the sense that there is only one unbiased linear estimator for $\Phi(f)$. In contrast, in §3 we show that there are many essentially different unbiased linear estimators for the ET problem. The construction of the efficient estimator, which depends on an analysis of the range properties of the Radon transform, is also carried out in this section. In §4, we show that the efficient linear estimators constructed in §2 and §3 are actually efficient without restriction of the class of estimators. In §5, we discuss concrete representations of the estimators constructed in §2 and §3. In §6, we consider the special case where ϕ is a Gaussian aperture function. Explicit numerical results for this case are given. Some concluding remarks are given in §7.

2. Efficient Linear Estimators for Planar Imaging. In order to develop some intuition for the problem, we start by considering a restricted class of estimators, the linear estimators. We shall construct estimators that are efficient within this restricted class, which we term efficient linear estimators. In §4, we shall see that these efficient linear estimators turn out to be efficient even when the class of estimators is not restricted.

In this section, we consider the estimation of $\Phi(f)$ given n i.i.d. observations distributed according to f, i.e., the planar-imaging problem. For this problem, there is an obvious estimator for $\Phi(f)$. Indeed, note that $\Phi(f)$ is the expected value of the random variable constructed by evaluating the function ϕ at a random sample distributed according to f. Thus an obvious estimate for $\Phi(f)$ is the sample mean of the derived observations obtained by evaluating the function ϕ at the original observations.

DEFINITION 2.1. Let $v : D \to \mathbb{R}$. We define the linear estimator generated by v to be the function given by $(x_1, \ldots, x_n) \mapsto n^{-1} \sum_{i=1}^{n} v(x_i)$, where x_i is the ith observation.

We shall now show that the estimator generated by ϕ is essentially the only unbiased linear estimator for $\Phi(f)$. It is thus an efficient linear estimator by default.

DEFINITION 2.2. Let $L^1(D)$ and $L^\infty(D)$ denote the usual spaces of absolutely integrable and bounded almost everywhere functions on D, respectively. We will assume some familiarity with the elementary facts about these spaces. The reader that is unfamiliar with these facts may consult, e.g., [11, sec. 6.1].

PROPOSITION 2.3. *The linear estimator generated by $v : D \to \mathbb{R}$ is an unbiased estimator of $\Phi(f)$ if and only if $v = \phi$ almost everywhere on D.*

Proof. At $f \in \mathcal{P}$, the linear estimator generated by v has expected value

$$\int_D \cdots \int_D n^{-1} \sum_{i=1}^{n} v(x_i) f(x_1) \cdots f(x_n)\, dx_1 \cdots dx_n = \int_D v(x) f(x)\, dx.$$

It is then clear that if $v = \phi$ almost everywhere on D, then v generates an unbiased estimator for $\Phi(f)$. To prove the converse, suppose v generates an unbiased estimator for $\Phi(f)$. We first claim that $v \in L^\infty(D)$. To prove the claim, assume, to obtain a contradiction, that $v \notin L^\infty(D)$. Define $S \equiv \{x \in D : v(x) \geq 3 \sup_D \phi\}$ and consider

$$f = \frac{2}{3 \int_S dx} 1_S + \frac{1}{3 \int_{D/S} dx} 1_{D/S},$$

where 1_S and $1_{D/S}$ denote the indicator functions of S and the difference of D and S, respectively. Then, $f \in \mathcal{P}$,

$$\begin{aligned}\int_D v(x) f(x)\,dx &\geq 3 \sup_D \phi \frac{2}{3\int_S dx} \int_S dx \\ &= 2 \sup_D \phi,\end{aligned}$$

but $\int_D \phi(x) f(x)\,dx \leq \sup_D \phi$, contradicting the unbiasedness of the estimator generated by v. This proves the claim that $v \in L^\infty(D)$. Now the unbiasedness assumption implies that the linear functional $\Upsilon : f \mapsto \Upsilon(f) \equiv \int_D v(x) f(x)\,dx$ agrees with the functional Φ on the linear span of $\mathcal{P}$. The linear span of $\mathcal{P}$ is dense in $L^1(D)$ since it contains all simple functions, i.e., functions taking only a finite number of values, and the set of simple functions are dense in $L^1(D)$ [11, prop. 6.7]. Thus the functionals Υ and Φ are equal on $L^1(D)$. The result now follows from the fact that $L^\infty(D)$ is the dual space to $L^1(D)$. □

PROPOSITION 2.4. *The variance of the estimator generated by ϕ at $f \in \mathcal{P}$ is*

$$n^{-1}\{\int_D \phi^2(x) f(x)\,dx - [\Phi(f)]^2\}.$$

Proof. The variance of the unbiased linear estimator generated by ϕ is

$$\begin{aligned}&\int_D \cdots \int_D [n^{-1} \sum_{i=1}^{n} \phi(x_i) - \Phi(f_0)]^2 f_0(x_1) \cdots f_0(x_n)\,dx_1 \cdots dx_n \\ &= n^{-1} \int_D [\phi(x) - \Phi(f_0)]^2 f_0(x)\,dx \\ &= n^{-1}\{\int_D \phi^2(x) f(x)\,dx - [\Phi(f)]^2\}. \qquad \square\end{aligned}$$

In the ET problem, it will be useful to compare linear estimators by viewing their generators as elements of suitable function spaces. To facilitate comparison, we will now introduce the corresponding function spaces for the planar imaging problem.

DEFINITION 2.5. Let $f_0 \in L^\infty(D)$. The space of real-valued functions f on D such that

$$||f||^2_{L^2(D,f_0)} \equiv \int_D f^2(x) f_0(x)\,dx < \infty$$

will be denoted by $L^2(D, f_0)$. When $f_0 = 1$, we simply write $L^2(D)$. $L^2(D, f_0)$ is a Hilbert space when equipped with the inner product

$$\langle f, g\rangle_{L^2(D,f_0)} \equiv \int_D f(x) g(x) f_0(x)\,dx$$

(see, e.g., [8, thm. 5.2.1]).

REMARK 2.6. When f_0 in definition 2.5 is a p.d.f., the space $L^2(D, f_0)$ corresponds to the space of random functions on D with finite second moment when the underlying random variable is distributed according to f_0.

The results of this section may now be summarized:

THEOREM 2.7. *The estimator generated by ϕ is an efficient linear estimator for the planar-imaging problem. Its variance at $f \in \mathcal{P}$ is*

$$n^{-1}\{||\phi||^2_{L^2(D,f)} - [\Phi(f)]^2\}. \qquad \square$$

3. Efficient Linear Estimators for ET. The efficient linear estimator for the planar imaging problem derived in §2 is not directly applicable to the ET problem since the observations do not lie in the domain of the function ϕ. To construct analogous estimators for the ET problem, it is natural to look for a function, ψ, on $\mathbb{L}$ such that $\Psi(Rf) = \Phi(f)$ for all $f \in \mathcal{P}$ and use the linear estimator generated by ψ. We call such a ψ an observation-domain representation of Φ. In this section, we shall construct efficient linear estimators for the ET problem using this approach. The major complication is that, due to the nature of the Radon transform, there are many such observation-domain representations of Φ.

3.1. Preliminaries.

DEFINITION 3.1. We write the integral of a function g on $\mathbb{L}$ as

$$\int_{\mathbb{L}} g(l)\,dl \equiv \pi^{-1} \int_0^{\pi} \int_{-\infty}^{\infty} g(\theta, s)\,ds\,d\theta.$$

REMARK 3.2. The coordinates (θ, s) may be identified with points in $S^1 \times \mathbb{R}$ in a obvious way, where S^1 denotes the unit circle. Under this identification, each $l \in \mathbb{L}$ maps to two points in $S^1 \times \mathbb{R}$. In the literature, the Radon transform is often defined to be the real-valued function on $S^1 \times \mathbb{R}$ obtained under this identification. It then becomes natural to write the integral of a function g on $\mathbb{L}$ as $\int_{S^1} \int_{-\infty}^{\infty} g(\theta, s)\,ds\,d\theta \equiv \int_0^{2\pi} \int_{-\infty}^{\infty} g(\theta, s)\,ds\,d\theta$. Since we use a different convention here, many of the formulas here will appear to differ from those in the literature by a factor of 2π. We shall henceforth make these conversions without further comment.

DEFINITION 3.3. We define C to be the subset of $\mathbb{L}$ consisting of lines whose distance from the origin is ≤ 1, i.e., lines with 's' coordinate of magnitude ≤ 1. For $g_0 \in L^\infty(C)$, we define $L^2(C, g_0)$ in an analogous way to the definition of $L^2(D, f_0)$ given in definition 2.5. Let $\psi \in L^2(C, g_0)$. We define the linear estimator generated by ψ to be the function $C^n \to \mathbb{R}$ given by $(l_1, \ldots, l_n) \mapsto n^{-1} \sum_{i=1}^n \psi(l_i)$. We define

$$\begin{aligned} \Psi(g) &\equiv \int_C \psi(l) g(l)\,dl \\ &\equiv \pi^{-1} \int_0^{\pi} \int_{-1}^{1} \psi(\theta, s) g(\theta, s)\,ds\,d\theta. \end{aligned}$$

PROPOSITION 3.4. *Let $f_0 \in \mathcal{P}$ and $\psi \in L^2(C, Rf_0)$. Then the linear estimator generated by ψ has mean $\Psi(Rf_0)$ and variance*

$$n^{-1} \left(\int_C \psi^2(l) Rf_0(l)\,dl - [\Psi(Rf_0)]^2 \right)$$

at f_0.

Proof. The calculations are essentially the same as those in the proofs of propositions 2.3 and 2.4. □

REMARK 3.5. The condition $\psi \in L^2(C, Rf_0)$ is clearly necessary and sufficient for the variance of an estimate of the form $(l_1, \ldots, l_n) \mapsto n^{-1} \sum_{i=1}^n \psi(l_i)$ to exist at $f_0 \in \mathcal{P}$. It is also necessary and sufficient for the variance to exist at any other $f \in \mathcal{P}$ since all the spaces $L^2(C, Rf)$ with $f \in \mathcal{P}$ are equal as sets.

3.2. Characterization of Unbiased Linear Estimators. Our next goal is to find the conditions under which an element $\psi \in L^2(C, Rf_0)$ with $f_0 \in \mathcal{P}$ generates an unbiased estimator for $\Phi(f)$. The intuitive content of the result is that ψ generates an unbiased estimator for $\Phi(f)$ if and only if ψ backprojects to ϕ. A mathematical formulation of this statement is that the adjoint operator of the Radon transform must map ψ to ϕ. We devote the remainder of this subsection to making these ideas precise.

Since it is natural to assess the variance of the linear estimator generated by ψ at $f_0 \in \mathcal{P}$ by viewing ψ as an element of $L^2(C, Rf_0)$, we reformulate R as an operator R_{f_0} whose range is $L^2(C, Rf_0)$.

DEFINITION 3.6. For $f_0 \in \mathcal{P}$, define the linear operator $R_{f_0}\colon L^2(D) \to L^2(C, Rf_0)$ by $f \mapsto Rf/Rf_0$. It is shown in [22, thm. II.1.6] that R_{f_0} is continuous when f_0 is the uniform density, from which it easily follows that R_{f_0} is continuous for any $f_0 \in \mathcal{P}$.

REMARK 3.7. $R_{f_0}f(l)$ is just the likelihood ratio of the observation l under the statistical hypotheses f and f_0.

The desired unbiasedness condition is naturally expressed in terms of the adjoint operator of R_{f_0}.

DEFINITION 3.8. If $A : H \to K$ is a continuous linear operator from a Hilbert space H to a Hilbert space K, there is a unique continuous linear operator, which is called the adjoint of A and denoted by A^*, such that $\langle Ax, y\rangle_K = \langle x, A^*y\rangle_H$ for all $x \in H$ and $y \in K$ [19, sec. 3.9].

PROPOSITION 3.9. *Let $f_0 \in \mathcal{P}$. The linear estimator generated by $\psi \in L^2(C, Rf_0)$ is an unbiased estimator of $\Phi(f)$ if and only if $R^*_{f_0}\psi = \phi$.*

Proof. Suppose $R^*_{f_0}\psi = \phi$ and $f \in \mathcal{P}$. Then, by the definition of the adjoint,

$$\begin{aligned} \Psi(Rf) &= \langle \psi, R_{f_0}f\rangle_{L^2(C,Rf_0)} \\ &= \langle R^*_{f_0}\psi, f\rangle_{L^2(D)} \\ &= \Phi(f). \end{aligned}$$

for all $f \in L^2(D)$, in particular for all $f \in \mathcal{P}$. Conversely, suppose $\Psi(Rf) = \Phi(f)$ for all $f \in \mathcal{P}$. Since Ψ and Φ are linear operators, $\Psi(Rf) = \Phi(f)$ for all f in the linear span of $\mathcal{P}$. Now the linear span of $\mathcal{P}$ is dense in $L^2(D)$ since it contains all simple functions and the set of simple functions are dense in $L^2(D)$ [11, prop. 6.7]. It follows that

$$\begin{aligned} \langle \psi, R_{f_0}f\rangle_{L^2(C,Rf_0)} &= \Psi(Rf) \\ &= \langle \phi, f\rangle_{L^2(D)} \end{aligned}$$

for all $f \in L^2(D)$. This says that $R^*_{f_0}\phi = \psi$. □

We conclude this subsection by showing that the adjoint of $R_{f_0} : L^2(D) \to L^2(C, Rf_0)$ is essentially the familiar backprojection operation used in the FB image reconstruction algorithm. The next proposition shows that the backprojection operation is the adjoint operator for R.

PROPOSITION 3.10. *The map $R^*\colon L^\infty(C) \to L^\infty(D)$ given by*

$$R^*g(x) = \pi^{-1}\int_0^\pi g(\theta, x\cdot\vec{\theta})\,d\theta \tag{3.1}$$

for $g \in L^\infty(C)$ is the adjoint of $R : L^1(D) \to L^1(C)$ in the sense that

$$\int_C Rf(l)\,g(l)\,dl = \int_D f(x)\,R^*g(x)\,dx$$

for all $f \in L^1(D)$ and $g \in L^\infty(C)$.

Proof. The corresponding result for $R : L^1(\mathbb{R}^2) \to L^1(\mathbb{L})$ is given in [13, p. 169]. The result for $R : L^1(D) \to L^1(C)$ follows easily by extending the functions on D and C to $\mathbb{R}^2$ and $\mathbb{L}$ with zeros. □

PROPOSITION 3.11. *The adjoint of $R_{f_0} : L^2(D) \to L^2(C, Rf_0)$ is given by the unique continuous extension of $R^* : L^\infty(C) \to L^\infty(D)$ to $L^2(C, Rf_0)$.*

Proof. Let $f \in L^2(D)$ and $g \in L^\infty(C)$. Then $f \in L^1(D)$ and proposition 3.10 shows that

$$\begin{aligned} \langle R_{f_0} f, g \rangle_{L^2(C,Rf_0)} &= \int_C Rf(l) g(l)\, dl \\ &= \langle f, R^* g \rangle_{L^2(D)}. \end{aligned}$$

Since $L^\infty(C)$ is dense in $L^2(C, Rf_0)$ [11, prop. 6.7], the result follows. □

REMARK 3.12. If $f, f_0 \in \mathcal{P}$, then it is easy to see from the proof of proposition 3.11 that $R_f^* = R_{f_0}^*$.

3.3. Existence of Unbiased Linear Estimators. Proposition 3.9 says that a function $\psi \in L^2(C, Rf_0)$ generates an unbiased linear estimator if and only if $R_{f_0}^* \psi = \phi$. We are therefore lead to investigate the existence and uniqueness of solutions to this equation. In this subsection, we answer the existence question in the affirmative by constructing a solution that is closely related to the standard FB reconstruction algorithm. We start by reviewing this algorithm.

DEFINITION 3.13. We denote the Schwartz space of smooth functions on $\mathbb{R}^d$ that, along with their derivatives, rapidly approach 0 as $|x| \to \infty$ by $\mathcal{S}(\mathbb{R})$ [27, ch. 10, ex. IV]. The Fourier transform $\mathcal{F}: \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is denoted by $f \mapsto \tilde{f}$, where $\tilde{f}$ is defined by

$$\tilde{f}(\xi) \equiv \int_{\mathbb{R}^d} e^{-i2\pi x \cdot \xi} f(x)\, dx$$

[27, p. 268, def. 25.1]. We define the Riesz potential operator I^{-1} for functions $f \in \mathcal{S}(\mathbb{R}^d)$ by

$$(I^{-1} f)^{\sim}(\xi) \stackrel{\text{def}}{=} 2\pi |\xi| \tilde{f}(\xi) \tag{3.2}$$

(cf., [25, sec. V.1]). Let $C^\infty(\mathbb{R}^d)$ denote the set of smooth functions on $\mathbb{R}^d$. For $f \in \mathcal{S}(\mathbb{R}^d)$, it is shown in lemma A.1 below that $I^{-1} f \in C^\infty(\mathbb{R}^d)$. For functions on $\mathbb{L}$, we define I^{-1} to act on the second, or 's' variable only.

REMARK 3.14. In the context of medical imaging, I^{-1} acting on functions on $\mathbb{L}$ is the "ramp filter" in the FB algorithm.

REMARK 3.15. For any $\phi \in C^\infty(\mathbb{R}^2)$, one can find a $\phi' \in \mathcal{S}(\mathbb{R}^2)$ that is equal to ϕ on D (cf. [11, lem 8.10]) and thus generates the same functional on $\mathcal{P}$ as ϕ. In what follows, we will therefore assume, without loss of generality, that $\phi \in \mathcal{S}(\mathbb{R}^2)$.

REMARK 3.16. Theorem 3.17 below is the basis for the FB algorithm. The Radon transform of ϕ, $R\phi$, is ramp-filtered (I^{-1}) and backprojected (R^*) to recover ϕ.

THEOREM 3.17. *Let $\phi \in \mathcal{S}(\mathbb{R}^2)$. Then $\phi = 2^{-1} R^* I^{-1} R\phi$ (recall that I^{-1} acts on the second or 's' variable of functions on $\mathbb{L}$).*

Proof. See [22, p. 18, thm. 2.1]. □

Theorem 3.17 immediately gives a solution to the unbiasedness condition $R_{f_0}^* \psi = \phi$.

DEFINITION 3.18. For $\phi \in \mathcal{S}(\mathbb{R}^2)$, define $F\phi \in C^\infty(\mathbb{L})$ by $F\phi = 2^{-1}I^{-1}R\phi$ (cf. lemma A.1).

PROPOSITION 3.19. *Let $\phi \in \mathcal{S}$. Then $R^*_{f_0}F\phi = \phi$.*

Proof. By propositions 3.11 and 3.17,

$$\begin{aligned} R^*_{f_0}F\phi &= R^*F\phi \\ &= \phi. \quad \square \end{aligned}$$

REMARK 3.20. Intuitively, proposition 3.19 says that an observation-domain representation of the functional generated by ϕ is obtained by ramp filtering the Radon transform of ϕ. This representation was considered in [3]. A discrete version of this representation for indicator functions may be found in [15].

REMARK 3.21. The standard approach to estimating linear functionals in medical imaging is to first obtain the FB estimate, $\hat{f}$, of f and then apply the functional to $\hat{f}$. We shall now give a heuristic argument to the effect that this standard procedure is essentially the linear estimator generated by $F\phi$.

Let δ_{l_i} denote a point mass at the ith observation l_i and define the empirical estimate $\widehat{Rf} \equiv n^{-1}\sum_{i=1}^{n}\delta_{l_i}$ of Rf. The FB estimate of f may be obtained by substituting $\widehat{Rf}$ in for $R\phi$ in the inversion formula of theorem 3.17:

$$\hat{f} = 2^{-1}R^*I^{-1}\widehat{Rf}.$$

Let $a \in \mathcal{S}(\mathbb{R}^2)$. Then, using the identity $a * R^*b = R^*(Ra * b)$, where convolution of functions on $\mathbb{L}$ is understood to taken with respect to the second or 's' variable only [22, eq. V.1.2],

$$\begin{aligned} a * \hat{f} &= 2^{-1}a * R^*I^{-1}\widehat{Rf} \\ &= 2^{-1}R^*(Ra * I^{-1}\widehat{Rf}) \\ &= 2^{-1}R^*(I^{-1}Ra * \widehat{Rf}) \\ &= R^*(Fa * \widehat{Rf}). \end{aligned} \tag{3.3}$$

We can apply equation 3.3 to the estimation of linear functionals by defining $\check{\phi}(x) \equiv \phi(-x)$ and noting that $\Phi(f) = \check{\phi} * f(0)$. Defining $\check{\psi}(\theta, s) \equiv \psi(\theta, -s)$, the corresponding FB estimate of $\Phi(f)$ is given by

$$\begin{aligned} \hat{\Phi}(f) &= R^*(\check{\psi} * n^{-1}\sum_{i=1}^{n}\delta_{l_i})(0) \\ &= n^{-1}\sum_{i=1}^{n} R^*[\delta(\theta - \theta_j)\check{\psi}(\theta_i, s - s_i)](0) \\ &= n^{-1}\sum_{i=1}^{n}\check{\psi}(\theta_i, -s_i) \\ &= n^{-1}\sum_{i=1}^{n}\psi(\theta_i, s_i), \end{aligned}$$

which is just the linear estimate generated by ψ. A discrete version of this argument may be found in [15].

3.4. Uniqueness of Unbiased Linear Estimators. Proposition 3.19 establishes the existence of an unbiased linear estimator. The next proposition addresses its uniqueness.

PROPOSITION 3.22. *Let $\phi \in \mathcal{S}$. Then $\psi \in L^2(C, Rf_0)$ satisfies $R^*_{f_0}\psi = \phi$ if and only if $\psi \in F\phi + \mathcal{N}(R^*_{f_0})$, where $\mathcal{N}(R^*_{f_0})$ denotes the nullspace of $R^*_{f_0}$.*

Proof. If $R^*_{f_0}\psi = \phi$, then $R^*_{f_0}(\psi - F\phi) = 0$, so $\psi - F\phi \in \mathcal{N}(R_{f_0})$. Conversely, if $\psi - F\phi \in \mathcal{N}(R^*_{f_0})$, then $R^*_{f_0}\psi = \phi$. $\square$

Intuitively, proposition 3.22 says that the functions that generate unbiased linear estimators form an affine linear space obtained by adding to $F\phi$ functions that backproject to 0. Thus the uniqueness question boils down to whether or not there exist nonzero functions on C that backproject to 0 on D. In fact, there are many such functions. For example,

$$\psi(\theta, s) = \begin{cases} 1 & \text{if } 0 \le \theta < \pi/2 \\ -1 & \text{if } \pi/2 \le \theta < \pi \end{cases}.$$

We conclude this subsection by pointing out that $\mathcal{N}(R^*_{f_0})$ is orthogonal to the range of R_{f_0}.

PROPOSITION 3.23. *Let $f_0 \in \mathcal{P}$. Then*

$$\mathcal{N}(R^*_{f_0}) = \overline{R_{f_0}L^2(D)}^{\perp},$$

where $\overline{R_{f_0}L^2(D)}^{\perp}$ denotes the orthogonal complement of the closure of $R_{f_0}L^2(D)$.

Proof. This is a special case of the following general fact. If $A : H \to K$ is a continuous linear operator from a Hilbert space H to a Hilbert space K, then $\mathcal{N}(A^*) = \overline{AH}^{\perp}$ (see, e.g., [6, secs. I.2, II.2]). $\square$

3.5. Construction of Efficient Linear Estimators. We conclude this section by showing that the efficient linear estimator is generated by the projection of any $\psi \in L^2(C, Rf_0)$ satisfying $R^*_{f_0}\psi = \phi$, e.g., $F\phi$, onto $\overline{R_{f_0}L^2(D)}$.

THEOREM 3.24. *Let $f_0 \in \mathcal{P}$ and $\phi \in \mathcal{S}(\mathbb{R}^2)$. Suppose $\psi \in L^2(C, Rf_0)$ satisfies $R^*_{f_0}\psi = \phi$. Let $p_{\overline{R_{f_0}L^2(D)}}\psi$ denote the projection of ψ on the subspace $\overline{R_{f_0}L^2(D)} \subset L^2(C, Rf_0)$. Then the estimator generated by $p_{\overline{R_{f_0}L^2(D)}}\psi$ is an efficient linear estimator for $\Phi(f)$ at f_0 in the ET problem. Its variance at f_0 is*

$$n^{-1}\left(||p_{\overline{R_{f_0}L^2(D)}}\psi||^2_{L^2(C,Rf_0)} - [\Psi(Rf_0)]^2\right). \tag{3.4}$$

Proof. We first note that $p_{\overline{R_{f_0}L^2(D)}}\psi$ generates a unbiased estimator since $\psi - p_{\overline{R_{f_0}L^2(D)}}\psi \in \overline{R_{f_0}L^2(D)}^{\perp}$ and hence $\psi - p_{\overline{R_{f_0}L^2(D)}}\psi \in \mathcal{N}(R^*_{f_0})$. Now suppose $\psi' \in L^2(C, Rf_0)$ generates an unbiased estimate of $\Phi(f)$. Then $R^*_{f_0}(\psi' - \psi) = 0$, so $\psi' - \psi \in \mathcal{N}(R^*_{f_0})$ and hence orthogonal to $\overline{R_{f_0}L^2(D)}$. It follows that $p_{\overline{R_{f_0}L^2(D)}}\psi' = p_{\overline{R_{f_0}L^2(D)}}\psi$. Thus

$$\begin{aligned} ||\psi'||^2_{L^2(C,Rf_0)} &= ||p_{\overline{R_{f_0}L^2(D)}}\psi||_{L^2(C,Rf_0)} + ||p_{\overline{R_{f_0}L^2(D)}^{\perp}}\psi||_{L^2(C,Rf_0)} \\ &\ge ||p_{\overline{R_{f_0}L^2(D)}}\psi||_{L^2(C,Rf_0)}. \qquad \square \end{aligned}$$

In effect, the unbiasedness condition determines the component of the generator of any unbiased linear estimator in the range of the Radon transform, but allows the component in the orthogonal complement to be arbitrary. The variance is minimized by setting the component in the orthogonal complement to be 0.

REMARK 3.25. In §4, we shall see that the estimator constructed in theorem 3.24 is an efficient estimator even when the class of estimators is not restricted. However, unlike the estimator for the planar imaging problem, the efficient estimator at f_0 depends on f_0 through Rf_0 since the projection operation depends on the weighting function Rf_0. Thus there is no estimator that is *uniformly* efficient over $\mathcal{P}$.

REMARK 3.26. Roughly speaking, the dependence of the efficient estimator on f_0 reflects the fact that Rf_0 is measured with a statistical uncertainty that varies over C with variance proportional to Rf_0. The efficient estimator constructed above is not a practical estimator since the weighting function Rf_0 is not known *a priori*. To construct a practical estimator, one would have to replace Rf_0 with a suitable estimate.

4. Efficient Estimators.

4.1. The Uniformly Efficient Estimator for Planar Imaging. In this subsection, we will prove the claim made in §2 that the linear estimator generated by ϕ is an uniformly efficient estimator for $\Phi(f)$ in the planar-imaging problem.

THEOREM 4.1. *Let $\phi \in \mathcal{S}(\mathbb{R}^2)$. The estimator generated by ϕ is an efficient estimator for $\Phi(f)$ uniformly over $f \in \mathcal{P}$ in the planar-imaging problem.*

Proof. Let $f_0 \in \mathcal{P}$ be given. For $|t| < \sup_{x\in D} \phi(x) - \Phi(f_0)$, define $f_t : D \to \mathbb{R}$ by

$$f_t \equiv \{1 + t[\phi - \Phi(f_0)]\} f_0.$$

Using the fact that

$$\int_D [\phi(x) - \Phi(f_0)] f_0 \, dx = 0, \tag{4.1}$$

it is clear that the $f_t \in \mathcal{P}$. Consider the one-dimensional subproblem of estimating $\Phi(f_t)$ from n i.i.d. observations distributed according to f_t. The Cramér-Rao inequality [16, art. 399D] states that the variance of any unbiased estimator of $\Phi(f_t)$ at f_0 is bounded below by

$$\frac{[\partial_t \Phi(f_t)|_{t=0}]^2}{n\|\partial_t \log f_t(x)|_{t=0}\|^2_{L^2(D,f_0)}}.$$

The denominator of this expression is commonly referred to as the Fisher information for t at $t = 0$. We have

$$\begin{aligned} \partial_t \log f_t(x)|_{t=0} &= \left.\frac{[\phi(x) - \Phi(f_0)] f_0(x)}{f_t(x)}\right|_{t=0} \\ &= \phi(x) - \Phi(f_0), \end{aligned}$$

and hence the Fisher information at $t = 0$ is equal to

$$n\|\phi(x) - \Phi(f_0)\|^2_{L^2(D,f_0)} = n\{\|\phi\|^2_{L^2(D,f_0)} - [\Phi(f_0)]^2\}.$$

Again using equation 4.1, we also have

$$\begin{aligned} \partial_t \Phi(f_t)|_{t=0} &= \partial_t \int_D \phi(x)\{1 + t[\phi(x) - \Phi(f_0)]\} f_0 \, dx|_{t=0} \\ &= \int_D \phi(x)[\phi(x) - \Phi(f_0)] f_0(x) \, dx \\ &= \|\phi\|^2_{L^2(D,f_0)} - [\Phi(f_0)]^2. \end{aligned}$$

It follows that the variance of an unbiased estimator of $\Phi(f_t)$ at f_0 must be $\geq n^{-1}\{\|\phi\|^2_{L^2(D,f_0)} - [\Phi(f_0)]^2\}$. By theorem 2.7, the linear estimator generated by ϕ is unbiased and achieves this lower bound. Since $f_0 \in \mathcal{P}$ was arbitrary, we conclude that the estimator generated by ϕ is an efficient estimator for $\Phi(f)$ uniformly over $f \in \mathcal{P}$. $\square$

4.2. Efficient Estimators for ET. In this subsection, we will show that the efficient linear estimator constructed in §3 is an efficient estimator for the ET problem.

THEOREM 4.2. *Let $f_0 \in \mathcal{P}$ and $\phi \in \mathcal{S}(\mathbb{R}^2)$. The estimator generated by $\psi \equiv p_{\overline{R_{f_0}L^2(D)}}F\phi$ is an efficient estimator for $\Phi(f)$ at f_0 in the ET problem.*

Proof. The proof is analogous to that of theorem 4.1 in that the Cramér-Rao inequality is applied to a one-dimensional subproblem to show that the variance of any unbiased estimator cannot be less than that of the efficient linear estimator. In the proof of theorem 4.1, we considered perturbations of f_0 in the direction $\phi(x) - \Phi(f_0)$. In the ET problem, we consider perturbations of g_0 (approximately) in the direction $p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)$.

We start by noting that constant functions are in $R_{f_0}L^2(D)$, since $R_{f_0}f_0 = 1$. It follows that $p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0) \in \overline{R_{f_0}L^2(D)}$. Define $g_0 \equiv Rf_0$ and let $L^2_0(D)$ and $L^2_0(C, g_0)$ denote the subspaces of $L^2(D)$ and $L^2(C, g_0)$ whose elements h satisfy $\int_D h(x)\,dx = 0$ and $\int_C h(l)g_0(l)\,dl = 0$, respectively. Then $p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0) \in \overline{R_{f_0}L^2(D)} \cap L^2_0(C, g_0)$, since $p_{\overline{R_{f_0}L^2(D)}}F\phi$ generates an unbiased linear estimator. We claim that $\overline{R_{f_0}L^2(D)} \cap L^2_0(C, g_0) \subseteq \overline{R_{f_0}L^2_0(D)}$. To prove the claim, let $f \in L^2(D)$. Then

$$(4.2) \qquad \begin{aligned} \int_D f(x)\,dx &= \int_D f(x)R^*1(x)\,dx \\ &= \int_C Rf(l)\,dl \\ &= \int_C R_{f_0}f(l)g_0\,dl. \end{aligned}$$

Thus if $f \in L^2(D)\backslash L^2_0(D)$, then $R_{f_0}f \notin L^2_0(C, g_0)$. This says that $R_{f_0}L^2(D) \cap L^2_0(C, g_0) \subseteq R_{f_0}L^2_0(D)$. Now suppose $g \in \overline{R_{f_0}L^2(D)} \cap L^2_0(C, g_0)$. Then there exists a sequence $\{g_n\}$ in $R_{f_0}L^2(D)$ such that $g_n \to g$ in $L^2(C, g_0)$. Since $g \in L^2_0(C, g_0)$, $\lim_{n\to\infty}\int_C g_n(l)g_0(l)\,dl = 0$. We can write each g_n as $R_{f_0}f_n$ for some $f_n \in L^2(D)$. By equation 4.2, $\lim_{n\to\infty}\int_D f_n(x)\,dx = 0$. Consider the sequence $\{f_n - \int_D f_n(x)\,dx\}$ in $L^2_0(D)$. We have

$$\begin{aligned} \lim_{n\to\infty} R_{f_0}[f_n - \int_D f_n(x)\,dx] &= \lim_{n\to\infty} g_n - R_{f_0}\lim_{n\to\infty}\int_D f_n(x)\,dx \\ &= g, \end{aligned}$$

so $g \in \overline{R_{f_0}L^2_0(D)}$. This proves the claim that $\overline{R_{f_0}L^2(D)} \cap L^2_0(C, g_0) \subseteq \overline{R_{f_0}L^2_0(D)}$. From the claim, we can now conclude that $p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0) \in \overline{R_{f_0}L^2_0(D)}$.

Let $\epsilon > 0$ be given. Since $p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0) \in \overline{R_{f_0}L^2_0(D)}$, we can choose $\eta \in L^2_0(D)$ such that $\|R_{f_0}\eta - p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)\|_{L^2(C,g_0)} < \epsilon$. Now the simple functions are dense in $L^2_0(D)$ [2, p. 52]. Using this fact and the fact that R_{f_0} is continuous, one may assume, without loss of generality, that η is a simple function. For

$|t| < \sup_{x\in D} |\eta(x)|$, define $g_t \in R\mathcal{P}$ by $g_t \equiv R(f_0 + t\eta)$. Consider the one-dimensional subproblem of estimating $\int_C F\phi(l)g_t(l)\,dl$ from n i.i.d. observations distributed according to g_t. (Note that, defining $f_t = f_0 + t\eta$, this is equivalent to the problem of estimating $\Phi(f_t)$.) We will now compute the Fisher information for the estimation of t at $t = 0$. We have

$$\begin{aligned}
\partial_t \log g_t|_{t=0} &= \left.\frac{\partial_t g_t}{g_t}\right|_{t=0} \\
&= \frac{R\eta}{Rf_0} \\
&= R_{f_0}\eta.
\end{aligned}$$

Thus the Fisher information for the estimation of t is equal to $n||R_{f_0}\eta||^2_{L^2(C,g_0)}$ at $t = 0$. Now

$$\begin{aligned}
&\partial_t \int_C F\phi(l)g_t(l)\,dl|_{t=0} \\
&= \int_C F\phi(l)R\eta(l)\,dl \\
&= \int_C F\phi(l)R_{f_0}\eta(l)g_0(l)\,dl \\
&= \int_C [F\phi(l) - \Psi(g_0)]R_{f_0}\eta(l)g_0(l)\,dl + \Psi(g_0) \\
&= \int_C p_{\overline{R_{f_0}L^2(D)}}[F\phi(l) - \Psi(g_0)]R_{f_0}\eta(l)g_0(l)\,dl + \Psi(g_0) \\
&= \int_C [p_{\overline{R_{f_0}L^2(D)}}F\phi(l) - \Psi(g_0)][p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)]g_0(l)\,dl \\
&\quad + \int_C [p_{\overline{R_{f_0}L^2(D)}}F\phi(l) - \Psi(g_0)]\{R_{f_0}\zeta(l) - [p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)]\}g_0(l)\,dl \\
&\geq ||p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)||^2_{L^2(C,g_0)} \\
&\quad - ||p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)||_{L^2(C,g_0)}||R_{f_0}\zeta(l) - p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)||_{L^2(C,g_0)} \\
&\geq ||p_{\overline{R_{f_0}L^2_0(D)}}F\phi - \Psi(g_0)||_{L^2(C,g_0)}(||p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)||_{L^2(C,g_0)} - \epsilon)
\end{aligned}$$

(the first inequality follows from the Cauchy-Schwarz inequality). It follows from the Cramér-Rao inequality [16, art. 389D] that the variance of an unbiased estimator of $\int_C F\phi(l)g_t(l)\,dl$ at g_0 must be at least

$$\begin{aligned}
&\frac{[\partial_t \int_C F\phi(l)g_t(l)\,dl|_{t=0}]^2}{n||R_{f_0}\zeta||^2_{L^2(C,g_0)}} \\
&\geq \frac{||p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)||^2_{L^2(C,g_0)}(||p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)||_{L^2(C,g_0)} - \epsilon)^2}{n(||p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)||_{L^2(C,g_0)} + \epsilon)^2}.
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, we conclude that the variance of an unbiased estimator of $\int_C F\phi(l)g_t(l)\,dl$ at g_0 must be $\geq n^{-1}||p_{\overline{R_{f_0}L^2(D)}}F\phi - \Psi(g_0)||^2_{L^2(C,g_0)}$. The unbiased linear estimator constructed in theorem 3.24 achieves this lower bound. □

REMARK 4.3. Unlike the estimator in theorem 4.1, the efficient estimator in theorem 4.2 depends on f_0. It is thus not a *uniformly* efficient estimator.

REMARK 4.4. The general approach used to construct lower bounds on the variance of estimators in theorems 4.1 and 4.2 is well known in the literature. In the terminology of Pfanzagl [23], the spaces $L_0^2(D)$ and $\overline{R_{f_0} L_0^2(D)}$ used in the proof are the tangent spaces to the statistical model in the planar imaging and ET problem, respectively. The functions $\phi - \Phi(f_0)$ and $p_{\overline{R_{f_0} L^2(D)}} F\phi - \Psi(g_0)$ are termed the canonical gradients of the functional Φ for these problems, respectively.

5. Construction of Projection Operators. In §4.2, we saw that the linear estimator generated by $p_{\overline{R_{f_0} L^2(D)}} F\phi$ is an efficient estimator for $\Phi(f)$ at f_0 in the ET problem. In this section, we will express the projection operator $p_{\overline{R_{f_0} L^2(D)}}$ in a concrete way that is suitable for numerical calculations.

We start in §5.1 by considering the special case where f_0 is the uniform distribution on D. It turns out that the analysis of this special case provides useful building blocks for the analysis of the general case with $f_0 \in \mathcal{P}$, which is carried out in §5.2.

5.1. The Uniform Distribution.

DEFINITION 5.1. Let f_u denote the uniform density on D, i.e., f_u is the constant function π^{-1} on D. Define $g_u = Rf_u$. Explicitly, $g_u(\theta, s) = 2\pi^{-1}\sqrt{1-s^2}$ (cf. [7, sec. 2.5, ex. 4], [17, sec. 2.1]). From this expression, we see that the marginal distribution of the s variable under g_u is given by the probability density $2\pi^{-1}\sqrt{1-s^2}$ on $[-1, 1]$. We shall denote this probability density by ρ_u.

DEFINITION 5.2. Let $\mathbb{Z}$, N, and N^+ denote the sets of integers, nonnegative integers, and positive integers, respectively. For $m \in \mathrm{N}$, define the functions U_m: $[-1, 1] \to \mathbb{R}$ by

$$U_m(\cos\theta) \equiv \frac{\sin[(m+1)\theta]}{\sin\theta}.$$

The U_m are called the Chebyshev polynomials of the second kind [7, sec. 7.6] [17, sec. 2]. They form an orthonormal basis for $L^2([-1, 1], \rho_u)$ [7, app. C.2]. As the name implies, the U_m are indeed polynomials, the first three are 1, $2s$, and $4s^2 - 1$. We extend the functions U_m to $S^1 \times \mathbb{R}$ by the formula $U_m(\theta, s) \equiv U_m(s)$. For $j \in \mathrm{N}^+$ and $m \in \mathrm{N}$, define the functions $a_{j,m} : S^1 \times \mathbb{R} \to \mathbb{R}$ and $b_{j,m} : S^1 \times \mathbb{R} \to \mathbb{R}$ by

$$a_{j,m}(\theta, s) \equiv \sqrt{2}U_m(s)\cos(j\theta),$$

and

$$b_{j,m}(\theta, s) \equiv \sqrt{2}U_m(s)\sin(j\theta).$$

DEFINITION 5.3. If $g(\theta, s)$ is an even function on $S^1 \times \mathbb{R}$, i.e., $g(-\theta, -s) = g(\theta, s)$, it can be viewed in a natural way as a function on $\mathbb{L}$ given by the formula $g(\theta, s)$. In particular, since U_m is even when m is even and $a_{j,m}$ and $b_{j,m}$ are even when $j + m$ is even, we shall view U_m, $a_{j,m}$, and $b_{j,m}$ as being defined on C by the formulas given in definition 5.2 when m and $j + m$ are even, respectively.

DEFINITION 5.4. Define

$$B_u \equiv \{U_m : m \in 2\mathrm{N}\} \cup \{a_{j,m}, b_{j,m} : j \in \mathrm{N}^+, m \in j + 2\mathrm{N}\}$$

and

$$B'_u \equiv \{a_{j,m}, b_{j,m} : j \in \mathrm{N}^+,\ m \in \{j \bmod 2, j \bmod 2 + 2, \ldots, j-2\}\}.$$

PROPOSITION 5.5. *B_u and B'_u are orthonormal bases for $\overline{R_{f_u}L^2(D)}$ and its orthogonal complement in $L^2(C, g_u)$, respectively.*

Proof. For $j \in \mathbb{Z}$ and $m \in \mathrm{N}$, define the functions $c_{j,m}$: $S^1 \times [-1,1] \to \mathbb{C}$ ($\mathbb{C}$ denotes the set of complex numbers) by

$$c_{j,m}(\theta, s) \equiv U_m(s)e^{ij\theta}.$$

Note that $c_{j,m}$ is even (odd) when $|m-j|$ is even (odd). An orthonormal basis for $\overline{R_{f_u}L^2(D)} \subset L^2(C, g_u)$ is given by $c_{j,m}$ with $j \in \mathbb{Z}$ and $m \in |j| + 2\mathrm{N}$ [7, sec. 7.6] [17, sec. 2.3], where the $c_{j,m}$ are interpreted as functions on C per definition 5.3.

Since we are dealing with (real) probability densities and linear functionals generated by real-valued functions, it is convenient to replace the above basis with one consisting of real-valued functions. Applying the standard procedure for converting a complex orthonormal basis to a real orthonormal basis to $\{c_{j,m} : j \in \mathrm{N},\ m \in j + 2\mathrm{N}\}$ yields B_u, which proves the first assertion.

To prove the second assertion, we show that an orthonormal basis for $L^2(C, g_u)$ is given by

$$B_u \cup B'_u = \{U_m : m \in 2\mathrm{N}\} \cup \{a_{j,m}, b_{j,m} : j \in \mathrm{N}^+, m \in j \bmod 2 + 2\mathrm{N}\}.$$

We start by noting that, since $\{(2\pi)^{-1/2}, \pi^{-1/2}\cos(j\theta), \pi^{-1/2}\sin(j\theta) : j \in \mathrm{N}^+\}$ is an orthonormal basis for $L^2(S^1)$ and $\{U_m : m \in \mathrm{N}\}$ is an orthonormal basis for $L^2([-1,1], 2\pi^{-1}\sqrt{1-s^2})$, an orthonormal basis for $L^2(S^1 \times [-1,1], g_u)$ is given by

$$\{(2\pi)^{-1/2}U_m : m \in \mathrm{N}\} \cup \{(2\pi)^{-1/2}a_{j,m}, (2\pi)^{-1/2}b_{j,m} : j \in \mathrm{N}^+, m \in \mathrm{N}\}.$$

If $\psi \in L^2(C, g_u)$, then

$$\begin{aligned}
\|\psi\|^2_{L^2(C,g_u)} &= \frac{1}{2\pi}\int_0^{2\pi}\int_{-1}^{1} \psi^2(\theta,s) g_u(\theta,s)\, ds\, d\theta \\
&= \frac{1}{4\pi^2}\left(\sum_{m\in\mathrm{N}} \langle U_m, \psi\rangle^2_{L^2(S^1\times[-1,1],g_u)}\right. \\
&\qquad \left. + \sum_{j\in\mathrm{N}^+}\sum_{m\in\mathrm{N}} \langle a_{j,m}, \psi\rangle^2_{L^2(S^1\times[-1,1],g_u)} + \langle b_{j,m}, \psi\rangle^2_{L^2(S^1\times[-1,1],g_u)}\right) \\
&= \left(\sum_{m\in 2\mathrm{N}} \langle U_m, \psi\rangle^2_{L^2(C,g_u)}\right. \\
&\qquad \left. + \sum_{j\in\mathrm{N}^+}\sum_{m\in j \bmod 2 + 2\mathrm{N}} \langle a_{j,m}, \psi\rangle^2_{L^2(C,g_u)} + \langle b_{j,m}, \psi\rangle^2_{L^2(C,g_u)}\right),
\end{aligned} \tag{5.1}$$

where the last equality follows since U_m is even (odd) when $j + m$ is even (odd) and $a_{j,m}$ and $b_{j,m}$ are even (odd) when $j+m$ is even (odd). Since $B_u \cup B'_u$ is easily verified to be an orthonormal subset of $L^2(C, g_u)$, it is an orthonormal basis for $L^2(C, g_u)$. Combining this fact with the first assertion gives the second assertion. □

Having established that B_u is an orthonormal basis for $\overline{R_{f_u}L^2(D)}$, it is now straightforward to express $p_{\overline{R_{f_u}L^2(D)}}$ in terms of this basis.

COROLLARY 5.6. *The efficient estimator at f_u for the ET problem is generated by*

$$
\begin{aligned}
p_{\overline{R_{f_u} L^2(D)}} F\phi &= \sum_{m\in 2\mathbb{N}} \langle U_m, F\phi\rangle_{L^2(C,g_u)} U_m + \sum_{j\in\mathbb{N}^+} \sum_{m\in j+2\mathbb{N}} \langle a_{j,m}, F\phi\rangle_{L^2(C,g_u)} a_{j,m} \\
&\quad + \sum_{j\in\mathbb{N}^+} \sum_{m\in j+2\mathbb{N}} \langle b_{j,m}, F\phi\rangle_{L^2(C,g_u)} b_{j,m}. \qquad (5.2)
\end{aligned}
$$

Its variance at f_u is given by

$$n^{-1} \sum_{m\in 2\mathbb{N}^+} \langle U_m, F\phi\rangle^2_{L^2(C,g_u)} \tag{5.3}$$

$$+n^{-1} \sum_{j\in\mathbb{N}^+} \sum_{m\in j+2\mathbb{N}} \left(\langle a_{j,m}, F\phi\rangle^2_{L^2(C,g_u)} + \langle b_{j,m}, F\phi\rangle^2_{L^2(C,g_u)}\right), \tag{5.4}$$

where

$$\langle g, F\phi\rangle_{L^2(C,g_u)} = 2\pi^{-2} \int_0^\pi \int_{-1}^1 g(\theta,s) F\phi(\theta,s)\sqrt{1-s^2}\, ds\, d\theta. \qquad \square \tag{5.5}$$

We noted in remark 3.21 that the estimator generated by $F\phi$ roughly corresponds to applying the functional Φ to the FB estimate of f. We shall therefore refer to it as the FB estimator. It is interesting to compare the performance of the FB estimator with that of the efficient estimator generated by $p_{\overline{R_{f_u} L^2(D)}} F\phi$.

COROLLARY 5.7. *The variance of the FB estimator at f_u is given by*

$$
\begin{aligned}
&n^{-1} \sum_{m\in 2\mathbb{N}^+} \langle U_m, F\phi\rangle^2_{L^2(C,g_u)} \\
&n^{-1} + \left(\sum_{j\in\mathbb{N}^+} \sum_{m\in j \bmod 2+2\mathbb{N}} \langle a_{j,m}, F\phi\rangle^2_{L^2(C,g_u)} + \langle b_{j,m}, F\phi\rangle^2_{L^2(C,g_u)} \right).
\end{aligned}
$$

The difference in variance at f_u between the FB estimator and the efficient estimator is

$$n^{-1} \sum_{j\in\mathbb{N}^+} \sum_{m=j\bmod, j\bmod 2+2,\ldots,j-2} \langle a_{j,m}, F\phi\rangle^2_{L^2(C,g_u)} + \langle b_{j,m}, F\phi\rangle^2_{L^2(C,g_u)}. \qquad \square \tag{5.6}$$

We conclude this section by showing there is an important special case where the FB estimator coincides with the efficient estimator.

PROPOSITION 5.8. *If $\phi \in \mathcal{S}(\mathbb{R}^2)$ is radial function, i.e., a function of $|x|$ alone, then $p_{\overline{R_{f_u} L^2(D)}} F\phi = F\phi$. Moreover, then $F\phi$ depends only on s.*

Proof. It is not difficult to verify that $F\phi$ is independent of θ and can be written as an even function of s. It follows that, for $j+m \in 2\mathbb{N}^+$, the inner product $\langle a_{j,m}, F\phi\rangle_{L^2(C,g_u)}$ reduces to

$$2\pi^{-2} \int_0^\pi \cos(j\theta)\, d\theta \int_{-1}^1 U_m(s) F\phi(s)\sqrt{1-s^2}\, ds = 0.$$

Similarly, $\langle b_{j,m}, F\phi\rangle_{L^2(C,g_u)} = 0$. Equation 5.2 thus reduces to

$$p_{\overline{R_{f_u} L^2(D)}} F\phi = \sum_{m\in 2\mathbb{N}} \langle U_m, F\phi\rangle_{L^2([-1,1],\rho_u)} U_m = F\phi, \tag{5.7}$$

where the last line follows since the set of functions $\{U_m; m \in 2\mathbb{N}\}$ is an orthonormal basis for the subspace of even functions in $L^2([-1,1],\rho_u)$.□

COROLLARY 5.9. *Suppose $\phi \in \mathcal{S}(\mathbb{R}^2)$ is radial. The estimator generated by $F\phi$ is an efficient estimator for $\Phi(f)$ at f_u in the ET problem. Its variance at f_u is given by*

$$n^{-1}\left(\|F\phi(0,\cdot)\|^2_{L^2([-1,1],\rho_u)} - [\Phi(f)]^2\right). \qquad \square \tag{5.8}$$

5.2. The General Case.

REMARK 5.10. Given a countable linearly-independent subset of a Hilbert space, there is a standard procedure, the Gram-Schmidt procedure, for constructing an orthonormal subset with the same linear span [6, I.4.6]. We shall refer to this procedure as orthonormalizing.

PROPOSITION 5.11. *An orthonormal basis for $\overline{R_{f_0} L^2(D)}$ is obtained by orthonormalizing the set of functions $\{\frac{Rf_u}{Rf_0} g_j\}_{g_j\in B_u}$ in $L^2(C, Rf_0)$.*

Proof. Defining the multiplication operator $M_{\frac{Rf_u}{Rf_0}} : L^2(C, Rf_u) \to L^2(C, Rf_0)$ by $g \mapsto \frac{Rf_u}{Rf_0} g$, it is clear that we can decompose R_{f_0} into the composition $M_{\frac{Rf_u}{Rf_0}} R_{f_u}$. Since $M_{\frac{Rf_u}{Rf_0}}$ is continuous, the linear span of $\{\frac{Rf_u}{Rf_0} g_j\}_{g_j\in B_u}$ is dense in $\overline{R_{f_0} L^2(D)}$. Orthonormalizing this set thus gives an orthonormal basis for $\overline{R_{f_0} L^2(D)}$. □

REMARK 5.12. Let $\{\eta_j\}$ be a orthonormal basis for $\overline{R_{f_0} L^2(D)}$, which may be constructed according to proposition 5.11. Then $p_{\overline{R_{f_0} L^2(D)}}\psi$ and $\|p_{\overline{R_{f_0} L^2(D)}}\psi\|^2_{L^2(C,Rf_0)}$ are given by the formulas

$$p_{\overline{R_{f_0} L^2(D)}}\psi = \sum_{j=1}^{\infty} \langle \psi, \eta_j\rangle_{L^2(C,Rf_0)} \eta_j$$

and

$$\|p_{\overline{R_{f_0} L^2(D)}}\psi\|^2_{L^2(C,Rf_0)} = \sum_{j=1}^{\infty} \langle \psi, \eta_j\rangle^2_{L^2(C,Rf_0)}.$$

One issue encountered with the use of these formulas in numerical calculation is that one can only compute a finite number of terms and it is difficult to assess how many terms are necessary. This issue may be addressed by constructing a basis for the orthogonal complement of $\overline{R_{f_0} L^2(D)}$.

PROPOSITION 5.13. *An orthonormal basis for $\overline{R_{f_0} L^2(D)}^{\perp}$ is obtained by orthonormalizing the set of functions B'_u in $L^2(C, Rf_0)$.*

Proof. By proposition 3.23, $\overline{R_{f_0} L^2(D)}^{\perp} = \mathcal{N}(R^*_{f_0})$. Thus, by proposition 3.23 and remark 3.12, $\overline{R_{f_0} L^2(D)}^{\perp} = \mathcal{N}(R^*_{f_u})$. An orthonormal basis for $\overline{R_{f_0} L^2(D)}^{\perp}$ is therefore obtained by orthonormalizing B'_u in $L^2(C, Rf_0)$. □

REMARK 5.14. We are now in a position to outline a numerical approach to the calculation of $p_{\overline{R_{f_0}L^2(D)}}\psi$ and $\|p_{\overline{R_{f_0}L^2(D)}}\psi\|^2_{L^2(C,Rf_0)}$. Orthonormal bases for $\overline{R_{f_0}L^2(D)}$ and $\overline{R_{f_0}L^2(D)}^{\perp}$ are given by propositions 5.11 and 5.13. Together they form a basis for $L^2(C, Rf_0)$. One can then expand ψ in terms of this basis. To check whether one has computed a sufficient number of terms, one can comparing the squared $L^2(C, Rf_0)$ norm of the expansion with $\|\psi\|^2_{L^2(C,Rf_0)}$, which can be computed by numerical integration.

6. Gaussian Functionals. In this section, we shall consider the special case where the linear functional is generated by a Gaussian density function. For brevity, we shall refer to such functionals as Gaussian functionals. In §6.1, we shall see that an observation-domain representation of Gaussian functionals can be given explicitly in terms of special functions. In §6.2, we specialize to the case of radial Gaussian functionals.

6.1. General Gaussian Functionals.

DEFINITION 6.1. *Let*

$$M(a; b; z) \equiv \sum_{k=0}^{\infty} \frac{\Gamma(a+k)/\Gamma(a)}{\Gamma(b+k)/\Gamma(b)} \frac{z^k}{k!}$$

denote Kummer's confluent hypergeometric function [24], where Γ *denotes the gamma function.*

PROPOSITION 6.2. *Define* $\phi_{a,\sigma} \in \mathcal{S}(\mathbb{R}^2)$ *to be the Gaussian density function*

$$\phi_{a,\sigma}(x) \equiv (2\pi\sigma^2)^{-1} e^{-|x-a|^2/2\sigma^2}$$

centered at $a \in \mathbb{R}^2$. *Then the observation-domain representation* $F\phi_{a,\sigma}$ *is given by*

$$F\phi_{a,\sigma}(\theta, s) = (2\pi\sigma^2)^{-1} e^{-(s-a\cdot\vec{\theta})^2/2\sigma^2} M(-1/2; 1/2; (s - a\cdot\vec{\theta})^2/2\sigma^2). \tag{6.1}$$

Proof. $R\phi_{a,\sigma}$ is given by

$$R\phi_{a,\sigma}(\theta, s) = (2\pi\sigma^2)^{-1/2} e^{-(s-a\cdot\vec{\theta})^2/2\sigma^2}$$

[7, sec. 3.5, eq. 5.3] and the Fourier transform of $R\phi_{a,\sigma}$ with respect to the second variable is given by

$$(R\phi)^{\sim}(\theta, \zeta) = e^{-i2\pi a\cdot\vec{\theta}} e^{-2\pi^2\sigma^2\zeta^2}$$

[11, prop. 8.24]. Thus, by definition 3.18,

$$\begin{aligned} (F\phi_{a,\sigma})^{\sim}(\theta, \zeta) &= 2^{-1}(I^{-1}R\phi_{a,\sigma})^{\sim}(\theta, \zeta) \\ &= \pi|\zeta|(R\phi_{a,\sigma})^{\sim}(\theta, \zeta) \\ &= \pi|\zeta| e^{-i2\pi a\cdot\vec{\theta}} e^{-2\pi^2\sigma^2\zeta^2}, \end{aligned} \tag{6.2}$$

where $(F\phi_{a,\sigma})^{\sim}$ denotes the Fourier transform of $F\phi_{a,\sigma}$ with respect to the second variable. The next task is to take the inverse Fourier transform of the above equation with respect to the second variable. First suppose that $a \cdot \vec{\theta} = 0$. By symmetry, we then have

$$\begin{aligned} F\phi_{a,\sigma}(\theta, s) &= \pi \int_{-\infty}^{\infty} |\zeta| e^{-2\pi^2\sigma^2\zeta^2} e^{i2\pi s\zeta}\, d\zeta \\ &= \pi \int_{-\infty}^{\infty} |\zeta| e^{-2\pi^2\sigma^2\zeta^2} \cos(2\pi s\zeta)\, d\zeta \\ &= 2\pi \int_{0}^{\infty} \zeta e^{-2\pi^2\sigma^2\zeta^2} \cos(2\pi s\zeta)\, d\zeta, \end{aligned}$$

which expresses $F\phi_{a,\sigma}$ in terms of a Fourier cosine transform. Using the Fourier cosine transform identity

$$\int_0^{\infty} x e^{-ax^2} \cos(xy)\, dx = (2a)^{-1} M(1; 1/2; -y^2/4a)$$

[10, eq. 1.4(14)], we get

$$F\phi_{a,\sigma}(\theta, s) = (2\pi\sigma^2)^{-1} M(1; 1/2; -s^2/2\sigma^2).$$

If $a \cdot \vec{\theta} \neq 0$, then the result for $a \cdot \vec{\theta} = 0$ and standard results on the effect of translation on the Fourier transform (see, e.g., [11, thm. 8.22]) give

$$F\phi_{a,\sigma}(\theta, s) = (2\pi\sigma^2)^{-1} M(1; 1/2; -(s - a \cdot \vec{\theta})^2/2\sigma^2).$$

We now apply the Kummer transformation identity $M(a, b, z) = e^z M(b - a; b; -z)$ [24, eq. 13.1.27] [9, eq. 6.4.7] to the last equation to obtain equation 6.1. □

REMARK 6.3. Defining the function $\chi_\sigma : \mathbb{R} \to \mathbb{R}$ by

$$\chi_\sigma(s) = (2\pi\sigma^2)^{-1} e^{-s^2/2\sigma^2} M(-1/2; 1/2; s^2/2\sigma^2),$$

we can write

$$F\phi_{a,\sigma}(\theta, s) = \chi_\sigma(s - a \cdot \vec{\theta}).$$

Thus for each fixed θ, $F\phi_{a,\sigma}$ as a function of s is a translate of the function χ_σ.

EXAMPLE 6.4. In figure 1, we illustrate the function χ_σ for $\sigma = 0.1$.

EXAMPLE 6.5. In figure 2, the graph on the upper left shows the zero-mean Gaussian density function with $\sigma = 0.5$, i.e., $\phi_{0,0.5}$. The graph on the upper right shows the observation–domain representation $F\phi_{0,0.5}$ of $\phi_{0,0.5}$, obtained using equation 6.1. The lower half of figure 2 is similar to the upper half, except that the Gaussian density function is centered at $a = (1, 0)$ instead of at the origin.

REMARK 6.6. By inserting the result of proposition 6.2 into corollary 5.6, we can explicitly compute the efficient estimator for Gaussian functionals at the uniform distribution along with its variance. Since the inner products in equations 5.2 and 5.3 are not available in closed form, it is necessary to evaluate them numerically. This comes down to evaluating a two-dimensional integral numerically, cf. equation 5.5. It is useful to start by computing the expansion given in equation 5.1. One can evaluate the left-hand side of this equation numerically and then verify that the right-hand side converges to the left-hand side. This provides a check on the accuracy of the numerical integrations and allows one to determine how many terms of the expansions in equations 5.2 and 5.3 are needed to achieve a given level of accuracy.

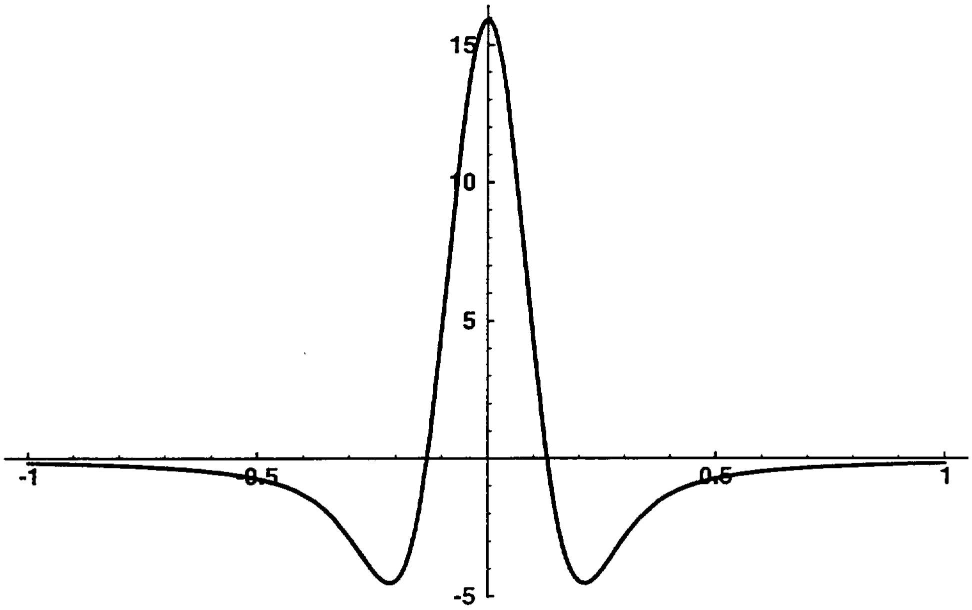

FIG. 6.1. *The function χ_σ evaluated for $\sigma = 0.1$. For each θ, the observation domain representation of a Gaussian functional with $\sigma = 0.1$ is a translate of this function.*

EXAMPLE 6.7. We consider the estimation of the functional generated by a Gaussian density function ϕ centered at $a = (1, 0)$ with $\sigma = 0.5$. (This was illustrated at the bottom of figure 2.) The FB estimator is generated by $F\phi_{(1,0),\sigma}$, which is shown at the left of figure 3. The efficient estimator at f_u is generated by $p_{\overline{R_{f_u} L^2(D)}} F\phi_{(1,0),\sigma}$, which is illustrated in the center of figure 3. The difference $F\phi_{(1,0),\sigma} - p_{\overline{R_{f_u} L^2(D)}} F\phi_{(1,0),\sigma}$, which backprojects to 0, is illustrated at the right of figure 3. In this particular case, the variance of the FB estimator is $0.087n^{-1}$ while the variance of the efficient estimator is $0.066n^{-1}$. Thus, in this case, the variance of the FB estimator is more than 30% higher than that of the efficient estimator.

6.2. Radial Gaussian Functionals.

REMARK 6.8. Efficient estimators for radial Gaussian functionals at f_u, along with their variance, can be computed by inserting the result of proposition 6.2 into corollary 5.9. Numerical evaluation of these quantities is much easier than in the nonradial case since no numerical integration is necessary to compute the estimator and its variance can be computed by evaluating a single one-dimensional integral numerically.

EXAMPLE 6.9. Figure 4 shows the variance of efficient estimators for a radial Gaussian functional given 10^6 observations evaluated at f_u as a function of σ. The lower curve is for the planar-imaging problem. It was obtained by numerically evaluating the formula given in theorem 4.1. The upper curve is for the ET problem. It was obtained by numerically evaluating the formula given in corollary 5.9.

For radial Gaussian functionals and the uniform distribution, the asymptotic behavior as $\sigma \to 0$ can be described very simply.

DEFINITION 6.10. For convenience, we define $\phi_\sigma \equiv \phi_{0,\sigma}$ and let Φ_σ denote the linear functional generated by ϕ_σ.

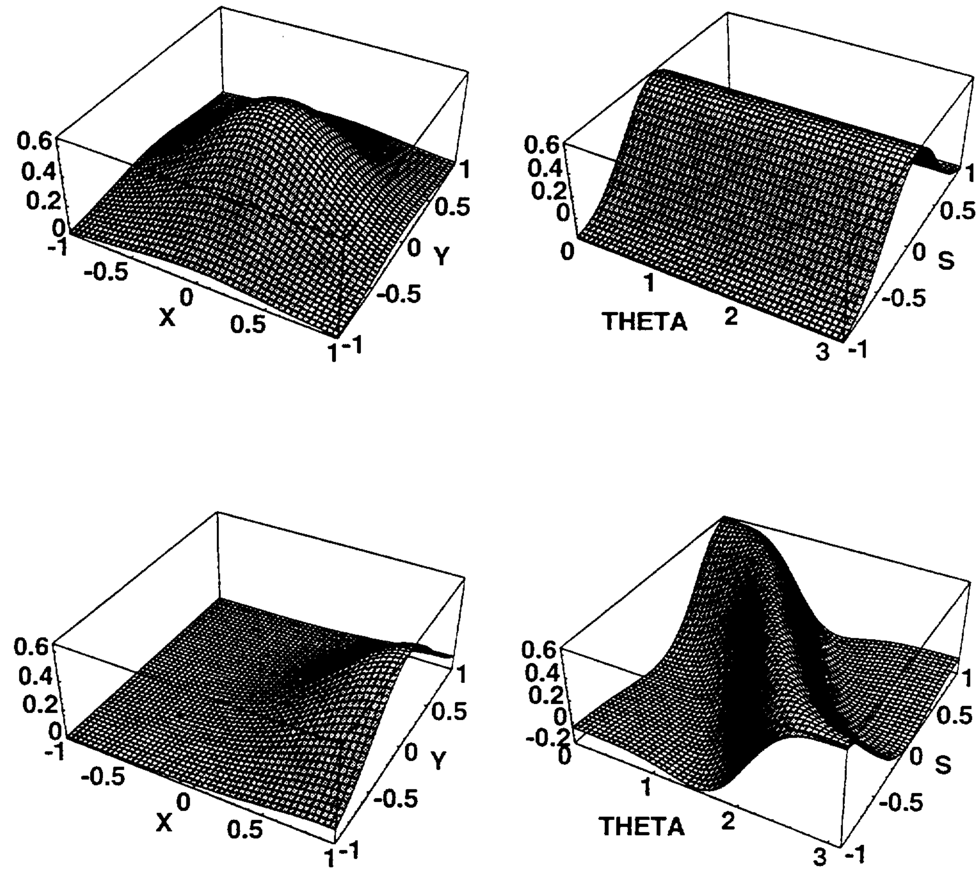

FIG. 6.2. *Gaussian functionals and their observation-domain representations. Gaussian densities are shown on the left and their respective observation-domain representations are shown on the right. The upper pair is for a radial Gaussian density, while the lower pair is for a Gaussian density centered at* $(1, 0)$*. For both pairs,* $\sigma = 0.5$.

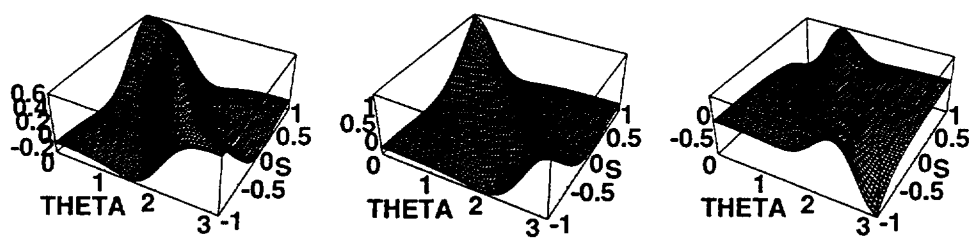

FIG. 6.3. *The functions which generate the FB and efficient estimators are shown on the left and middle, respectively. Their difference is shown on the right.*

PROPOSITION 6.11. *Asymptotically as* $\sigma \to 0$*, the variance of the efficient estimator for* $\Phi_\sigma(f)$ *at* f_u *is given by* $1/4\pi^2\sigma^2 n$ *in the planar imaging problem and* $1/8\pi^{3/2}\sigma^3 n$ *in the ET problem.*

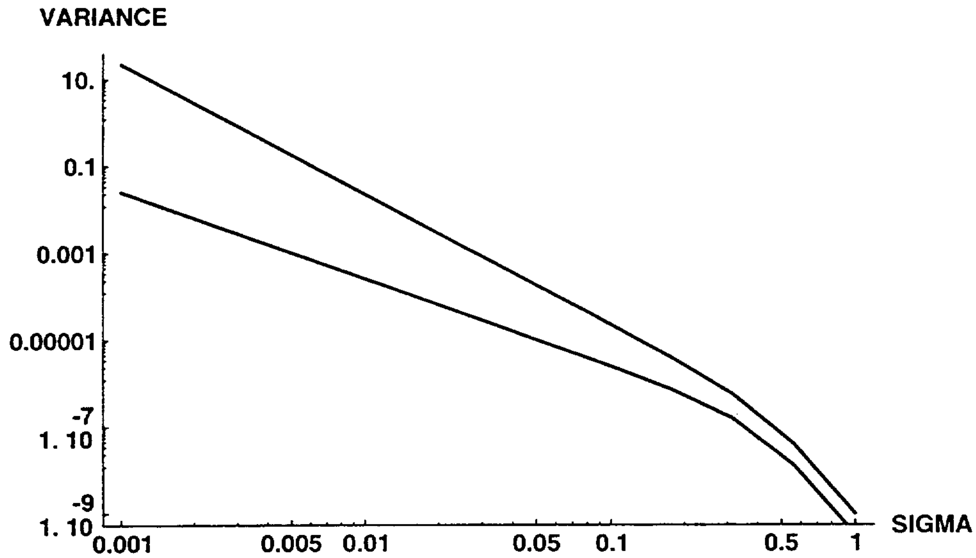

FIG. 6.4. *Variance of efficient estimator for radial Gaussian functionals at f_u with 10^6 observations (lower curve is for planar imaging, upper curve is for ET).*

Proof. Routine calculations show that

$$\begin{aligned}\Phi_\sigma(f_u) &= (2\pi^2\sigma^2)^{-1}\int_{|x|\le 1} e^{-|x|^2/2\sigma^2}\,dx \\ &= \pi^{-1}(1-e^{-1/2\sigma^2})\end{aligned}$$

and

$$\begin{aligned}\int_D \phi^2(x) f_u(x)\,dx &= \pi^{-1}\int_{|x|\le 1}\phi^2(x)\,dx \\ &= (4\pi^2\sigma^2)^{-1}(1-e^{-1/\sigma^2}).\end{aligned}$$

It follows from theorem 4.1 that the variance of the efficient estimator for the planar-imaging problem is given by

$$n^{-1}\left[(4\pi^2\sigma^2)^{-1}(1-e^{-1/\sigma^2}) - \pi^{-2}(1-e^{-1/2\sigma^2})^2\right],$$

which is clearly asymptotic to $1/4\pi^2\sigma^2 n$ in the limit as $\sigma \to 0$. Using theorem 4.2 and corollary 5.9, the variance of the efficient estimator for the ET problem is given by

$$n^{-1}\left[||F\phi_\sigma||^2_{L^2(C,Rf_u)} - \pi^{-2}(1-e^{-1/2\sigma^2})^2\right]. \tag{6.3}$$

It is shown below in lemma A.2 that $||F\phi_\sigma||^2_{L^2(C,Rf_u)}$ is asymptotic to $1/8\pi^{3/2}\sigma^3$ in the limit as $\sigma \to 0$. It follows that variance of the efficient estimator for the ET problem is asymptotic to $1/8\pi^{3/2}\sigma^3$. □

REMARK 6.12. The problem of estimating radial Gaussian functionals when f is uniform was considered in [26] in the context of x-ray computed tomography. While

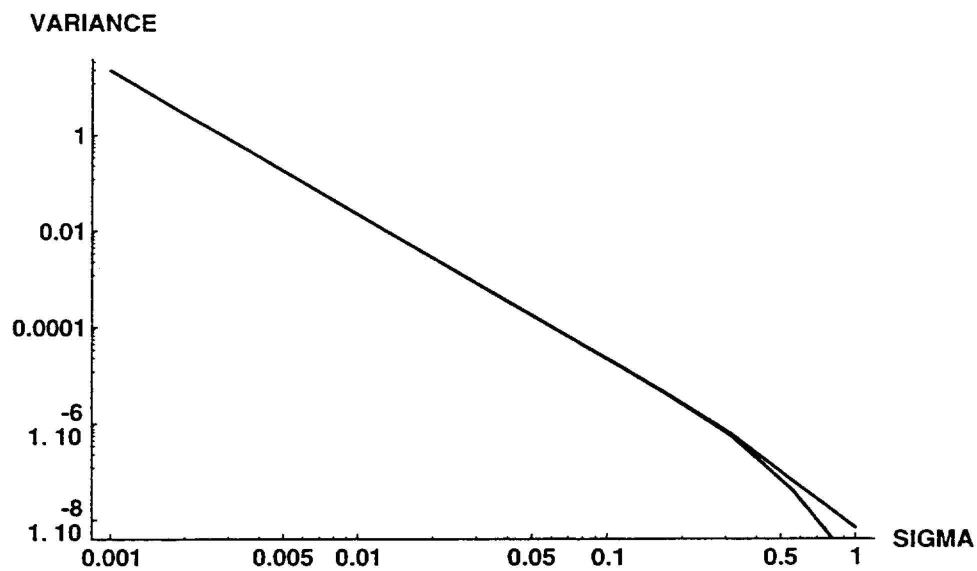

FIG. 6.5. *Small σ approximation to the variance of an efficient estimator for radial Gaussian functionals at f_u with 10^6 observations (lower curve is the true value, upper curve is the approximation described in example 6.13).*

the structure of the observations and the noise for this problem differ from that of the ET problem, it is interesting to note that the result in [26] reduces to $1/8\pi^{3/2}\sigma^3 n$ by taking appropriate limits, where n now denotes the number of transmitted photons. These results are roughly consistent with the performance of the FB algorithm seen in practice [4] [5].

EXAMPLE 6.13. Figure 5 shows the result of inserting the approximation

$$||F\phi_\sigma||^2_{L^2(C,Rf_u)} \approx 1/8\pi^{3/2}\sigma^3$$

into the expression for the variance of the efficient estimator given by equation 6.3. The approximation is shown in the upper curve, while the true value is shown in the lower curve. We see that the approximation is very accurate in the region $\sigma \leq 0.1$.

REMARK 6.14. For radial Gaussian functionals, the efficient estimator at f_u coincides with the FB estimator. We now present an example that shows this estimator can be very suboptimal at $f_0 \neq f_u$.

EXAMPLE 6.15. Let ϕ be radial Gaussian with $\sigma = 0.5$. Then the efficient estimator at f_u is generated by $F\phi$ and, by example 6.13, its variance at f_u is $\approx 0.087n^{-1}$. Suppose f_0 is highly concentrated about the point $(0, 0.63) \in D$. Then $g_0 = Rf_0$ is highly concentrated about the curve $\theta \mapsto (\theta, 0.63\sin\theta)$ and we can approximate the integral $\int_C g(l)g_0(l)\,dl$ by $\pi^{-1}\int_0^\pi g(\theta, 0.63\sin\theta)\,d\theta$. Using this approximation, the variance of the estimator generated by $F\phi$ at f_0 is $\approx 0.04683n^{-1}$. From proposition 5.13, the function $a_{2,0}/||a_{2,0}||_{L^2(C,g_0)}$ is a unit vector of $\overline{R_{f_0}L^2(D)}^{\perp}$. The squared inner product of this function with $F\phi$ in $L^2(C, g_0)$ is ≈ 0.04594, which implies that the variance of the efficient estimator at f_0 is at most $0.00089n^{-1}$. Thus in this, admittedly extreme, example, the variance of the efficient estimator is less than the variance of the FB estimator by a factor of more than 50.

7. Discussion. To summarize, the fact that not all functions on the observation space are Radon transforms of functions on D means that ET is, in a sense, an overdetermined problem. To construct an efficient estimator, it is necessary to weight the information obtained from the observations according to its statistical uncertainty. The resulting estimator is analogous to a weighted least squares procedure.

The results in this paper quantify the potential improvement over FB attainable by incorporation of information on the domain of the image and the statistical uncertainty of the observations into the estimation process. Our numerical results show that, at least in some cases, the efficient estimator has significantly less variance than the FB estimator. More extensive evaluation of the bound should help delineate the conditions under which significant improvement over the FB estimator is possible.

One potential application of the results on the estimation of functionals is as a benchmark for the performance of reconstruction algorithms. If the image estimator is denoted by $\hat{f}$, one can compare the variance of the implied estimate $\Phi(\hat{f})$ with the bound. The variance of this implied estimate gives a common performance index for reconstruction algorithms that is applicable even to algorithms that use different parameterizations of the image. Of course, parametric reconstruction algorithms may beat the bound if the image conforms to a parametric model. But the bound gives an index of the performance that might be expected from a "generic" reconstruction algorithm.

In terms of practical applications, it must be recognized that real tomographic data differs in two important ways from the mathematical model used here: they are generally discrete and have been modified by physical effects such as tissue attenuation. Thus to adequately model such data, it is necessary to replace the continuous Radon transform R used in our analysis with a discrete range operator that describes the imaging process. Discussion of such discrete-range operators may found in [1], [18], and [3]. There is some computational advantage to moving from the continuous formulation used here to a discrete one in that numerical evaluation of integrals is replaced by discrete sums. The discrete approach would therefore probably be our choice for a through assessment of the degree of improvement possible over the FB estimator in a practical problem. In our opinion, the main advantage of the continuous approach taken here is that, at least in the case $f = f_u$, the range structure of the Radon transform can be described explicitly, giving considerable insight into the problem. In addition, it is possible to derive an explicit formula for the observation-domain representation of a Gaussian functional and a compact asymptotic expression for its variance.

Appendix. Supporting Mathematical Results.

In this appendix, we collect the statements and proofs of some supporting results that were used in the body of this work.

LEMMA A.1. *If $f \in \mathcal{S}(\mathbb{R}^d)$, then $I^{-1}f \in C^\infty(\mathbb{R}^d)$.*

Proof. For $s \geq 0$, let $H_s(\mathbb{R}^d)$ denote the subspace of $L^2(\mathbb{R}^d)$ whose elements f have a Fourier transform satisfying

$$\begin{aligned} ||f||^2_{H_s(\mathbb{R}^d)} &\equiv \int_{\mathbb{R}^d} (1 + |\xi|^2)^{s/2} |\tilde{f}(\xi)|^2 \, d\xi \\ &< \infty. \end{aligned}$$

$H_s(\mathbb{R}^d)$ is termed the Sobolev space of order s. For each $k \in \mathrm{N}$, the Sobolev embedding theorem states that any function in $H_s(\mathbb{R}^d)$ for $s > k + d/2$ is k-times differentiable [11, thm. 8.54]. Now the Fourier transform of a function in $\mathcal{S}(\mathbb{R}^d)$ is in $\mathcal{S}(\mathbb{R}^d)$ [11, cor. 8.23]. Let $s \geq 0$ and $f \in \mathcal{S}(\mathbb{R}^d)$ be given. Since $\tilde{f} \in \mathcal{S}(\mathbb{R}^d)$, there exists $c > 0$

such that

$$(1+|\xi|^2)^{(s/2+d)/2}|\tilde{f}(\xi)| \leq c$$

on $\mathbb{R}^d$. It follows that

$$\begin{aligned}
||I^{-1}f||^2_{H_s(\mathbb{R}^d)} &= 4\pi^2 \int_{\mathbb{R}^d} (1+|\xi|^2)^{s/2}|\xi|^2|\tilde{f}(\xi)|^2 \, d\xi \\
&\leq 4\pi^2 c^2 \int_{\mathbb{R}^d} (1+|\xi|^2)^{-d} \, d\xi \\
&= 4\pi^2 c^2 \cdot \frac{2\pi^{d/2}}{\Gamma(d/2)} \int_0^\infty (1+r^2)^{-d} r^{d-1} \, d\xi \\
&\leq \frac{8c^2\pi^{2+d/2}}{\Gamma(d/2)} \int_0^\infty (1+r^2)^{-1} \, d\xi \\
&\leq \frac{8c^2\pi^{2+d/2}}{\Gamma(d/2)} (1 + \int_1^\infty r^{-2} \, d\xi) \\
&< \infty,
\end{aligned}$$

hence $I^{-1}f \in H_s(\mathbb{R}^d)$. Since $s \geq 0$ was arbitrary, it follows from the Sobolev embedding theorem that $I^{-1}f$ is k-times differentiable for all $k \in \mathbb{N}$, i.e., $I^{-1}f \in C^\infty(\mathbb{R}^d)$. □

LEMMA A.2. *The limiting behavior of* $||F\phi_\sigma||^2_{L^2(C,Rf_0)}$ *as* $\sigma \to 0$ *is described by*

$$\lim_{\sigma \to 0} \frac{||F\phi_\sigma||^2_{L^2(C,Rf_u)}}{2\pi^{-2}||F\phi_\sigma||^2_{L^2(\mathbb{L})}} = 1, \tag{A.1}$$

where

$$2\pi^{-1}||F\phi_\sigma||^2_{L^2(\mathbb{L})} = \frac{1}{8\pi^{3/2}\sigma^3}. \tag{A.2}$$

Proof. Starting with definition 3.18 and using equation 6.2, we obtain

$$\begin{aligned}
||F\phi_\sigma||^2_{L^2(\mathbb{L})} &= \pi^{-1} \int_0^\pi \int_{-\infty}^\infty [F\phi_\sigma(\theta, s)]^2 \, ds \, d\theta \\
&= \pi^{-1} \int_0^\pi \int_{-\infty}^\infty [(F\phi_\sigma)\hat{}(\theta, \zeta)]^2 \, d\zeta \, d\theta \\
&= \pi \int_0^\pi \int_{-\infty}^\infty \zeta^2 e^{-4\pi^2\sigma^2\zeta^2} \, d\zeta \, d\theta \\
&= 2\pi^2 \int_0^\infty \zeta^2 e^{-4\pi^2\sigma^2\zeta^2} \, d\zeta \, d\theta \\
&= \frac{1}{16\pi^{1/2}\sigma^3}
\end{aligned}$$

[12, 3.461.2], which gives equation A.2. To prove equation A.1, we will prove that

$$\lim_{\sigma \to 0} \frac{||F\phi_\sigma||^2_{L^2(C,Rf_u)} - 2\pi^{-1}||F\phi_\sigma||^2_{L^2(\mathbb{L})}}{2\pi^{-1}||F\phi_\sigma||^2_{L^2(\mathbb{L})}} = 0.$$

Let $\epsilon > 0$ be given. Since $\sqrt{1-s^2}$ is a continuous function at $s = 0$, we can choose $0 < \delta_s < 1$ such that $|\sqrt{1-s^2} - 1| < \epsilon/2$ if $|s| < \delta_s$. Then, writing

$$\begin{aligned}
&\frac{\left| ||F\phi_\sigma||^2_{L^2(C,Rf_u)} - 2\pi^{-1}||F\phi_\sigma||^2_{L^2(\mathbb{L})} \right|}{2\pi^{-1}||F\phi_\sigma||^2_{L^2(\mathbb{L})}} \\
&= \frac{\int_0^\pi \int_{-1}^1 (F\phi_\sigma)^2(\theta,s)|\sqrt{1-s^2}-1|\, ds\, d\theta + \int_0^\pi \int_{|s|>1} (F\phi_\sigma)^2(\theta,s)\, ds\, d\theta}{\pi||F\phi_\sigma||^2_{L^2(\mathbb{L})}} \\
&\le \frac{\frac{\epsilon}{2}\int_0^\pi \int_{|s|<\delta} (F\phi_\sigma)^2(\theta,s)\, ds\, d\theta}{\pi||F\phi_\sigma||^2_{L^2(\mathbb{L})}} + \frac{\int_0^\pi \int_{|s|\ge\delta} (F\phi_\sigma)^2(\theta,s)\, ds\, d\theta}{\pi||F\phi_\sigma||^2_{L^2(\mathbb{L})}} \\
&\le \frac{\epsilon}{2} + \frac{\int_0^\pi \int_{|s|\ge\delta} (F\phi_\sigma)^2(\theta,s)\, ds\, d\theta}{\pi||F\phi_\sigma||^2_{L^2(\mathbb{L})}},
\end{aligned}$$

we see that it suffices to prove that

$$\text{(A.3)} \qquad \lim_{\sigma\to 0} \frac{\int_0^\pi \int_{|s|\ge\delta} (F\phi_\sigma)^2(\theta,s)\, ds\, d\theta}{\pi||F\phi_\sigma||^2_{L^2(\mathbb{L})}} = 0.$$

Using the expansion

$$M(-1/2, 1/2, x) = -2^{-1}e^x x^{-1}[1 + O(|x|^{-1})],$$

which is valid for $x > 0$ [24, eq. 13.1.4], and equation 6.1, there exists a constant c such that

$$\begin{aligned}
|F\phi_\sigma(\theta,s)| &\le \frac{1}{4\pi\sigma^2}\cdot\frac{2\sigma^2}{s^2}[1 + c\sigma^2/s^2] \\
&\le \frac{1 + c\sigma^2/s^2}{2\pi s^2}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\frac{\int_0^\pi \int_{|s|\ge\delta} (F\phi_\sigma)^2(\theta,s)\, ds\, d\theta}{\pi||F\phi_\sigma||^2_{L^2(\mathbb{L})}} &\le \frac{\frac{[1+c\sigma^2/\delta^2]^2}{4\pi^2}\int_{|s|\ge\delta} s^{-4}\, ds}{||F\phi_\sigma||^2_{L^2(\mathbb{L})}} \\
&= \frac{16\sigma^3}{\pi^{1/2}}\frac{[1 + c\sigma^2/\delta^2]^2\delta^{-5}}{10\pi},
\end{aligned}$$

from which equation A.3 easily follows. □

REFERENCES

[1] J. R. Baker. *Spatially Variant Tomographic Imaging: Estimation, Identification, and Optimization.* PhD thesis, University of California at Berkeley, Berkeley, CA, 1991.

[2] P. J. Bickel, C. A. J. Klaassen, Y. Ritov, and J. A. Wellner. *Efficient and Adaptive Estimation for Semiparametric Models.* The Johns Hopkins University Press, Baltimore, 1993.

[3] P. J. Bickel and Y. Ritov. Estimating linear functionals of a PET image. *IEEE Trans. Med. Imaging*, 14:81–87, 1995.

[4] R. A. Brooks and G. Di Chiro. Statistical limitations in X-ray reconstruction tomography. *Med. Phys.*, 3:237–240, 1976.

[5] T. F. Budinger, S. E. Derenzo, W. L. Greenberg, G. T. Gullberg, and R. H. Huesman. Quantitative potentials of dynamic emission computed tomography. *J. Nucl. Med.*, 19:309–315, 1978.

[6] J. B. Conway. *A Course in Functional Analysis*. Springer-Verlag, New York, 1990.
[7] S. R. Deans. *The Radon Transform and Some of Its Applications*. John Wiley & Sons, New York, 1983.
[8] R. M. Dudley. *Real analysis and probability*. Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989.
[9] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi. *Higher Transcendental Functions*, volume 1. McGraw-Hill, New York, 1953.
[10] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi. *Tables of Integral Transforms*, volume 1. McGraw-Hill, New York, 1954.
[11] G. B. Folland. *Real Analysis*. John Wiley & Sons, New York, 1984.
[12] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, 1980.
[13] A. Hertle. Continuity of the Radon transform and its inverse on Euclidean space. *Math. Z.*, 184:165–192, 1983.
[14] R. H. Huesman. The effects of a finite number of projection angles and finite lateral sampling of projections on the propagation of statistical errors in transverse section reconstruction. *Phys. Med. Biol.*, 22:511–521, 1977.
[15] R. H. Huesman. A new fast algorithm for the evaluation of regions of interest and statistical uncertainly in computed tomography. *Phys. Med. Biol.*, 29:543–552, 1984.
[16] K. Itô. *Encyclopedic Dictionary of Mathematics*. MIT Press, Cambridge, MA, second edition, 1987.
[17] I. M. Johnstone and B. W. Silverman. Speed of estimation in positron emission tomography and related inverse problems. *Ann. Statist.*, 18:251–280, 1990.
[18] I. M. Johnstone and B. W. Silverman. Discretization effects in statistical inverse problems. *J. Complexity*, 7:1–34, 1991.
[19] E. Kreyszig. *Introductory Functional Analysis with Applications*. John Wiley & Sons, New York, 1989.
[20] A. Kuruc. *Probability Measure Estimation in Positron Emission Tomography Using Loss Functions Based on Sobolev Norms*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1994.
[21] A. Kuruc. Probability measure estimation using "weak" loss functions in positron emission tomography. In M. Cheney, P. Kuchment, and E. T. Quinto, editors, *Tomography, Impedance Imaging and Integral Geometry*, number 30 in Lectures in Applied Mathematics, pages 125–142. American Mathematical Society, Providence RI, 1994.
[22] F. Natterer. *The Mathematics of Computed Tomography*. John Wiley & Sons, New York, 1986.
[23] J. Pfanzagl. *Contributions to a General Asymptotic Statistical Theory*. Number 13 in Lecture Notes in Statistics. Springer-Verlag, New York, 1982.
[24] L. J. Slater. Confluent hypergeometric functions. In M. Abramowitz and I. A. Stegun, editors, *Handbook of Mathematical Functions*, pages 503–535. Dover Publications, New York, 1972.
[25] E. M. Stein. *Singular Integrals and Differentiability Properties of Integrals*. Princeton University Press, Princeton, NJ, 1970.
[26] O. J. Tretiak. Noise limitations in X-ray computed tomography. *J. Comp. Asst. Tomo.*, 2:477–480, 1978.
[27] F. Treves. *Topological Vector Spaces, Distributions, and Kernels*. Academic Press, New York, 1967.