LBL-37032
UC-405
Preprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

Submitted to SIAM Journal on Applied Mathematics

Estimation of Smooth Integral Functionals in Emission Tomography

A. Kuruc

March 1995

U. C. Lawrence Berkeley Laboratory
Library, Berkeley

FOR REFERENCE

Not to be taken from this room

Donner Laboratory

Biology & Medicine Division

REFERENCE COPY
Does Not
Circulate
Bldg. 50 Library.
LBL-37032
Copy 1

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

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-37032
UC-405

Estimation of Smooth Integral Functionals in Emission Tomography

Alvin Kuruc

Life Sciences Division
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

March 1995

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 National Institutes of Health Training Grant HL07367.

Estimation of Smooth Integral Functionals in Emission Tomography

Alvin Kuruc
Center for Functional Imaging, Lawrence Berkeley Laboratory,
University of California, Berkeley, CA

March 30, 1995

Running head: Estimation of Smooth Functionals in Tomography

Address for correspondence:

Alvin Kuruc
Lawrence Berkeley Laboratory
University of California
1 Cyclotron Road
MS 55-121
Berkeley, CA 94720
A_Kuruc@lbl.gov
510-486-6868
510-486-4768 (FAX)

Abstract

Emission tomography amounts to the statistical estimation of the spatial distribution of a radioactive tracer based on a finite sample of externally-detected photons. We present is 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 $\dot{\phi}(f) \equiv \int \phi(x) f(x)\, dx$, where ϕ is a smooth function. Assuming only that f is bounded above and below away from 0, we construct minimum-variance unbiased (MVU) estimators for $\dot{\phi}(f)$. By definition, the variance of the MVU estimator is a best-possible lower bound (depending on ϕ and f) on the variance of unbiased estimators of $\dot{\phi}(f)$. It gives a benchmark for the performance of image reconstruction and quantification algorithms. The analysis gives a geometrical explanation of when and by how much estimators based on the standard filtered-backprojection reconstruction algorithm may be improved.

Keywords: Radon transform, Unbiased estimation, Density estimation

AMS Subject Classifications: 92C55, 62G07

Contents

1 Introduction **5**

1.1 Mathematical Model of ECT 6

1.2 Integral Functionals 8

1.3 Basic Statistical Notions 9

1.4 Image Model 11

1.5 Outline of Paper 12

2 The Radon Transform **13**

2.1 Definition 13

2.2 The Adjoint 15

2.3 L^2 Theory 16

2.4 Inversion 19

3 MVU Linear Estimators **22**

3.1 The Planar-Imaging Problem 22

3.2 The ECT Problem 24

4 MVU Estimators **33**

4.1 Uniformly MVU Estimators for Planar Imaging 33

4.2 MVU Estimators for ECT 34

5 Construction of Projection Operators **38**

5.1 The Uniform Distribution 38

5.2 The General Case 43

6 Gaussian Functionals **45**

6.1 General Gaussian Functionals 45

6.2 Radial Gaussian Functionals 48

7 Extension of the Model **51**

8 Discussion **53**

A Supporting Mathematical Results **54**

Acknowledgements **57**

References **58**

Index **61**

1 Introduction

Computed tomography is based on the fundamental mathematical result that, under suitable regularity conditions, a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is uniquely determined by its line integrals.

Definition 1.1 Let $\mathbb{R}^d$ denote d-dimensional Euclidean space. Let G denote the set of lines in $\mathbb{R}^2$ (G is mnemonic for Grassmann manifold). The Radon transform of $f\colon \mathbb{R}^2 \to \mathbb{R}$ is defined to be the function $Rf : G \to \mathbb{R}$ which maps $l \in G$ to the line integral of f over l.

In emission computed tomography (ECT), the goal is to determine the density, f, of a radioactive tracer in a subject as a function of position by external detection of emitted photons. Since the probability of a photon traveling along a specific line through the subject is proportional to the integral of f along that line, the probability density function (p.d.f.) of the observations is given by the (appropriately scaled) Radon transform of f. Because only a finite number of observations distributed according to the Radon transform of f are available, not the Radon transform of f itself, a statistical estimate of f must be made from the available data. In this paper, we will quantify the statistical difficulty of estimating quantities of the form $\int \phi(x) f(x)\, dx$, where ϕ is a smooth function. We term these quantities integral functionals of f generated by smooth functions or, more briefly, smooth functionals.

Example 1.2 Let $D \subseteq \mathbb{R}^2$ denote the closed unit (radius disk). Suppose that f is square integrable and 0 outside D and that ϕ is an element of an orthonormal basis for the square-integrable functions on D. Then $\int \phi(x) f(x)\, dx$ is the coefficient of ϕ in the expansion of f with respect to that basis.

Example 1.3 If ϕ approximates the indicator function of a pixel or region of interest (ROI), then $\int \phi(x) f(x)\, dx$ approximates the fraction of activity in that pixel or ROI.

Under minimal assumptions about f, we shall construct minimum-variance unbiased (MVU) estimators for smooth functionals and compute their performance. 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 such functionals may be estimated. Moreover, it turns out that there is a simple geometrical relationship between the MVU estimator and an estimator based upon the filtered-backprojection reconstruction algorithm commonly used in medical imaging. The analysis provides an explanation of when and how much estimators based on filtered backprojection may be improved.

Remark 1.4 The MVU estimator will be seen to depend on f. Since, in practice, f is unknown, one cannot attain the bound simply by just applying the MVU estimator. However, as will be discussed in sec. 8, the analysis gives insight into how one might construct practical estimators whose performance approaches that of the MVU estimator.

1.1 Mathematical Model of ECT

There are two principal types of ECT used in medical imaging: single-photon emission computed tomography (SPECT) and positron emission tomography (PET). The very idealized mathematical model we will consider here is applicable to both. To motivate the model, we will consider the specific example of PET.

As the name suggests, PET images are produced using a positron-emitting radioactive tracer. Shortly after a positron is emitted, it combines with an electron in an annihilation reaction. On the average, a positron travels only a very short distance between emission and annihilation, so the position of the annihilation reaction is approximately that of the positron's emission. The annihilation reaction results in the emission of a pair of annihilation photons that travel in (approximately) opposite directions along a line that has uniformly-distributed random orientation. These pho-

tons subsequently strike detectors at approximately the same time, forming what is known as a coincident pair. From a coincident pair, one can thus infer that a positron was emitted (approximately) on the line segment between the two detectors which detected the pair.

The positions of the positron emissions and the paths of the annihilation photons are naturally modeled as points and lines, respectively, in $\mathbb{R}^3$. However, the most common detector configurations actually only detect coincident pairs whose paths are close to a fixed plane (and which therefore must originate from an annihilation reaction near that plane). In this case, it makes sense to think of the problem as being 2-dimensional. While in what follows we will, for simplicity, restrict ourselves to the 2-dimensional problem, our approach could equally be applied to the full 3-dimensional problem or any other detector configuration.

To construct a mathematical idealization of the PET problem, we model the locations of positron emissions as independent random variables distributed according to a p.d.f., f, D. Let $\Gamma \subset G$ denote the set of lines in $\mathbb{R}^2$ that pass through D. Given that the ith positron emission occurs at $x_i \in D$, the line along which the resulting annihilation photons travel will be modeled as a random line $l_i \in \Gamma$ through x_i with uniformly-distributed random orientation. We model these orientations as being independent of the x_i and each other so that the l_i are independent random variables. It can then be shown that the l_i are distributed according to the probability density $\frac{1}{\pi}Rf$ on Γ [14, sec. 2.1][15, sec. 2.3]. For convenience, we define $Tf \stackrel{\text{def}}{=} \frac{1}{\pi}Rf$.

Remark 1.5 In sec. 2, we will redefine R and T as operators on spaces of functions.

Remark 1.6 It should be emphasized that the model described here is a very idealized model of the PET problem. It ignores a number of physical effects which complicate PET in practice. (A brief discussion of how one might extend the model to include some of these effects may be found in sec. 7.) Nevertheless, it abstracts the fundamental problem of ECT.

1.2 Integral Functionals

The problem of estimating smooth functionals is motivated, in part, by their being a statistically well-posed substitute for the more natural problem of estimating the fraction of the total tracer contained in some subset $S \subset D$. Define the indicator function of S by

$$1_S(x) \stackrel{\text{def}}{=} \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases} .$$

Then the fraction of the total tracer in S is given by $\int_D 1_S(x) f(x)\, dx$. This suggests considering the more general problem of estimating quantities of the form $\dot{\phi}(f) \stackrel{\text{def}}{=} \int_D \phi(x) f(x)\, dx$, which we term integral functions, where the indicator function is replaced by a more general function. We shall consider the problem of estimating $\dot{\phi}(f)$ from n independent observations distributed according to the probability density Tf in the case where ϕ is a smooth function, i.e., ϕ is infinitely differentiable.

Remark 1.7 It is natural to ask why we treat only integral functionals generated by smooth functions and thereby exclude, in particular, indicator functions. The reason is that, under the very minimal assumptions that we want to make about f, it is impossible to construct estimators for integral functionals generated by indicator functions that perform well for all possible f. In other words, the problem of estimating integral functionals generated by indicator functions is statistically ill-posed without strong regularity conditions on f. Precise results along these lines may be found in [15, sec. 6.2] and [16, prop. 7.1]. Previous workers have avoided this difficulty by making strong assumptions about the functional form of f (see section 1.4). In a sense, the approach taken here obviates the need for strong assumptions about f by placing strong conditions on ϕ.

Remark 1.8 While indicator functions are not smooth, any indicator function can be closely approximated by a smooth function (cf. [8, prop. 8.17]). Under suitable regularity conditions on f, the integral functional generated by an indicator function

will then be close to the integral functional generated by the approximating smooth function. Thus, under suitable regularity conditions, the results presented here may be carried over to the problem of estimating the fraction of the total tracer in a pixel or ROI.

Remark 1.9 Suppose f is square integrable and $\{\phi_j\}$ is an orthonormal basis for the square-integrable functions on D. Then $f = \sum_j \dot{\phi}_j(f)\phi_j$ and the problem of estimating f is closely related to the problem of estimating the $\dot{\phi}_j(f)$. Thus, understanding how well one can estimate the integral functionals of f provides an approach to understanding how well one can estimate f. The point of this and the preceding remark is that understanding the estimation of smooth functionals of f provides a technical tool for understanding the estimation of the basic quantities of interest in ECT.

Remark 1.10 Note well that f is defined to be a p.d.f. on emission locations; it contains no information about the *rate* of emissions. (One can think of the p.d.f. as being obtained by dividing the emission rate per unit area by the total count rate.) Similarly, the data are taken to be a sequence of elements of Γ, 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 known the total count rate, but good estimates for this quantity are easy to construct.

1.3 Basic Statistical Notions

In this subsection, we introduce some basic statistical notions. We follow [22].

Definition 1.11 We define a statistical experiment to be an ordered pair $(\Omega, \{p_f : f \in \mathcal{P}\})$, where Ω is a probability (sample) space and $\{p_f : f \in \mathcal{P}\}$ is a set of p.d.f.s on Ω indexed by the set $\mathcal{P}$. The set Ω is the set of possible observations. In our model of the ECT problem, Ω will be the n-fold Cartesian product of Γ, Γ^n. The set

$\mathcal{P}$ is called the parameter space (cf. [22, def. 7.1]) and represents the set of possible states of nature. Under the statistical hypothesis $f \in \mathcal{P}$, the observation is a random variable distributed according to the p.d.f. p_f. In our model of the ECT problem, $\mathcal{P}$ will be taken to be a set of p.d.f.s on D. Under the statistical hypothesis $f \in \mathcal{P}$, the observations will be distributed according to the n-fold product of Tf, $(Tf)^n$, i.e., the probability density at $l = (l_1, l_2, \ldots, l_n) \in \Gamma^n$ is given by $\prod_{i=1}^n Tf(l_i)$. In other words, we shall consider the statistical experiment $(\Gamma^n, \{(Tf)^n : f \in \mathcal{P}\})$. For comparison, we shall also consider the simpler statistical experiment, the so-called planar-imaging problem, in which the observations under f are elements of the n-fold Cartesian product D^n distributed according to the n-fold product of f, i.e., the statistical experiment $(D^n, \{f^n : f \in \mathcal{P}\})$. In other words, the observations are the locations of the radioactive disintegrations.

Definition 1.12 We are interested in estimating smooth functionals of f given observations in Γ^n (or D^n). We will therefore consider estimators that are functions $\hat{\phi}_n$: $\Gamma^n \to \mathbb{R}$ (or $\hat{\phi}_n : D^n \to \mathbb{R}$). If $\hat{\phi}_n(l)$ is an estimate of $\dot{\phi}(f)$, the "badness" of the estimate will be quantified by the squared error $[\hat{\phi}_n(l) - \dot{\phi}(f)]^2$. The expected value of the squared error under f is termed the mean squared error (MSE) of the estimator $\hat{\phi}_n$ under f.

Roughly speaking, we would like construct estimators whose MSE is small for many of the $f \in \mathcal{P}$. It is easy to construct estimators whose MSE is small at any particular $f_0 \in \mathcal{P}$, indeed the estimator $\hat{\phi}_n = \dot{\phi}(f_0)$ has MSE 0 at f_0. But such estimators perform poorly at other $f \in \mathcal{P}$ and are of little interest. In order to obtain useful results, we must restrict our attention to estimators that satisfy some impartiality criterion. One such criterion is unbiasedness.

Definition 1.13 An estimator $\hat{\phi}_n$ is said to be an unbiased estimator of $\dot{\phi}(f)$ if $E_f \hat{\phi}_n = \dot{\phi}(f)$ for all $f \in \mathcal{P}$, where E_f denotes mathematical expectation under the statistical hypothesis f. The MSE of an unbiased estimator is equal to its variance.

An estimator is said to be minimum-variance unbiased (MVU) at f_0 if its variance at f_0 is minimal with respect to that of all unbiased estimators. An estimator is said to be uniformly MVU if it is MVU at all $f \in \mathcal{P}$.

In what follows, we will construct MVU estimators for both the planar-imaging and ECT problems.

1.4 Image Model

As discussed in sec. 1.3, useful estimators need to satisfy some impartiality criterion with respect to the various $f \in \mathcal{P}$. This makes any useful notion of statistical optimality dependent upon $\mathcal{P}$. While a number of authors have obtained results on the statistical difficulty of the ECT problem, the work here differs from previous work in that it makes very weak assumptions about the unknown density f, i.e., $\mathcal{P}$ is allowed to be a very "large" subset of the set of p.d.f.s on D.

In most statistical models for ECT, f is assumed to belong to some predetermined finite-dimensional family of probability densities. Such models are termed "parametric". Most commonly, the domain of interest is "pixelized", i.e., divided into a finite number of squares. It is then assumed that f is constant on each square (e.g., [24, sec. 1.2]). Another parametric approach, based on a singular value decomposition of the tomography process, is described in [1]. The advantage of parametric models is that one can use standard methods of parametric statistics to develop and evaluate estimators (e.g., maximum likelihood estimators and Cramér-Rao bounds). The main disadvantage of parametric models in ECT is that, in practice, the actual f is unlikely to conform to the assumed model, creating an error which falls outside the assumed statistical framework.

Definition 1.14 In the remainder of this paper, we will take $\mathcal{P}$ to be the set of p.d.f.s on D that are bounded above and below away from 0.

Remark 1.15 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.

1.5 Outline of Paper

In sec. 2, we review concepts related to the Radon transform. In sec. 3, we discuss the representation of smooth functionals on $\mathcal{P}$ by functions on Γ and construct MVU estimators in the class of linear estimators for both the planar-imaging and ECT problems. In sec. 4, we show that the linear estimators constructed in sec. 3 are actually MVU estimators in the class of all estimators. In sec. 5, we discuss concrete representations of the estimators constructed in sec. 3. In sec. 6, we consider the special case where ϕ is a Gaussian p.d.f.. Explicit numerical results for this case are given. In sec. 7, we sketch a simple generalization of our model to include the physical effects of photon attenuation and detector nonhomogeneity. Some concluding remarks are given in sec. 8.

2 The Radon Transform

In this section, we review some notions related to the Radon transform that are used in what follows. In sec. 2.1, we introduce some notation and define the Radon transform. Sec. 2.2 discusses the adjoint of the Radon transform, which is commonly referred to as the backprojection operator in medical imaging. In sec. 2.3, we consider the Radon transform on spaces of square-integrable functions. Sec. 2.4 gives the inversion formula for the Radon transform.

Remark 2.1 On a first reading, the reader may wish to just skim this section, focusing mainly on the notation and definitions. The motivation for many of the definitions will become apparent in sec. 3.

2.1 Definition

Definition 2.2 Let $S^1 \subset \mathbb{R}^2$ denote the unit circle. Define the map $\omega: \mathbb{R} \to S^1$ by $\theta \mapsto (\cos\theta, \sin\theta)$. We shall use the map ω to define local coordinate systems on S^1, the local coordinate being denoted by θ. We define the integral of f over S^1 as $\int_{S^1} f(\omega(\theta))\, d\theta \stackrel{\text{def}}{=} \int_0^{2\pi} f(\omega(\theta))\, d\theta$.

Definition 2.3 Let G denote the set of lines in $\mathbb{R}^2$. We define the double covering π: $S^1 \times \mathbb{R} \to G$ by taking $\pi(\omega, s) \in G$ to be the line in $\mathbb{R}^2$ through $s\omega$ which is perpendicular to ω. Note that $\pi(\omega, s) = \pi(-\omega, -s)$. It will be convenient to use the coordinates (θ, s) to refer to the line $\pi(\omega(\theta), s) \in G$. We will then write the integral of a function g on G as $\int_G g(\theta, s)\, ds\, d\theta \stackrel{\text{def}}{=} \int_0^{\pi} \int_{-\infty}^{\infty} g(\theta, s)\, ds\, d\theta$.

Definition 2.4 For any function $g : S^1 \times \mathbb{R} \to \mathbb{R}$, define its drop $\check{g}$: $G \to \mathbb{R}$ by $\check{g}(\pi(\omega, s)) \stackrel{\text{def}}{=} g(\omega, s) + g(-\omega, -s)$. For $g : G \to \mathbb{R}$, define the lift of g to $S^1 \times \mathbb{R}$ by $\tilde{g} \stackrel{\text{def}}{=} g \circ \pi : S^1 \times \mathbb{R} \to \mathbb{R}$.

Remark 2.5 We are now ready to define the Radon transform. In essence, the Radon transform of a function $f : \mathbb{R}^2 \to \mathbb{R}$ is the real-valued function on G whose

value on a given line $l \in G$ is equal to the integral of f over l. It is, however, more standard to define the Radon transform to be the real-valued function on $S^1 \times \mathbb{R}$ which is the lift of the Radon transform just defined. The resulting Radon transform has a redundancy due to the fact that each point in G is represented by two points in $S^1 \times \mathbb{R}$. While working on $S^1 \times \mathbb{R}$ is convenient in most (deterministic) contexts, in our statistical problem the observations are naturally elements of G and it is awkward to represent them as elements of $S^1 \times \mathbb{R}$. The major downside of this is that to use any of the previously obtained results on the Radon transform, we have to go back and forth between the two definitions. We shall refer to the Radon transform which gives functions on $S^1 \times \mathbb{R}$ as the standard Radon transform and the transform which gives functions on G simply as the Radon transform.

Definition 2.6 We shall be dealing with a number of sets on which there is some fixed notion of integration, e.g., the integral on S^1 defined in def. 2.2, the integral on G defined in def. 2.3, the usual integrals on $\mathbb{R}^d$ and D, etc. In order to avoid repetition, we shall generically denote a set equipped with some fixed notion of integration by X. This idea can be made precise by taking X to be a measure space. The reader who is unfamiliar with measure spaces will lose little by using the preceding intuitive definition of X.

Definition 2.7 Let X be as in def. 2.6. Define $L^1(X)$ to be the set of absolutely-integrable real-valued functions on X.

Definition 2.8 The Radon transform on $L^1(\mathbb{R}^2)$, which we shall denote by R, maps $f \in L^1(\mathbb{R}^2)$ to the function $Rf : G \to \mathbb{R}$ whose value at $l \in G$ is equal to the integral of f on l. The *standard* Radon transform on $L^1(\mathbb{R}^2)$, which we shall denote by $\widetilde{R}$, maps $f \in L^1(\mathbb{R}^2)$ to the lift of Rf to $S^1 \times \mathbb{R}$. It can be shown that $\widetilde{R}f \in L^1(S^1 \times \mathbb{R})$ [11, p. 168]. It follows at once that $Rf \in L^1(G)$.

2.2 The Adjoint

In this subsection, we will develop the adjoint of R. In the context of medical imaging, this operator is commonly referred to as the backprojection operator.

Definition 2.9 Let X be as in def. 2.6. Define $L^\infty(X)$ to be the set of bounded real-valued functions on X.

Proposition 2.10 *The map $\widetilde{R}^* : L^\infty(S^1 \times \mathbb{R}) \to L^\infty(\mathbb{R}^2)$ given by*

$$\widetilde{R}^* g(x) = \int_{S^1} g(\omega(\theta), x \cdot \omega(\theta))\, d\theta$$

for $g \in L^\infty(S^1 \times \mathbb{R})$ is the adjoint of $\widetilde{R} : L^1(\mathbb{R}^2) \to L^1(S^1 \times \mathbb{R})$ in the sense that

$$\int_{S^1 \times \mathbb{R}} \widetilde{R} f(\theta, s)\, g(\theta, s)\, ds\, d\theta = \int_{\mathbb{R}^2} f(x) \widetilde{R}^* g(x)\, dx$$

for all $f \in L^1(\mathbb{R}^2)$ and $g \in L^\infty(S^1 \times \mathbb{R})$. The map R^*: $L^\infty(G) \to L^\infty(\mathbb{R}^2)$ given by

$$\begin{aligned} R^* g(x) &= \frac{1}{2} \widetilde{R}^* \tilde{g}(x) \\ &= \int_0^\pi g(\theta, x \cdot \omega(\theta))\, d\theta \end{aligned} \tag{2.1}$$

for $g \in L^\infty(G)$ is the adjoint of $R : L^1(\mathbb{R}^2) \to L^1(G)$ in the sense that

$$\int_G R f(\theta, s)\, g(\theta, s)\, ds\, d\theta = \int_{\mathbb{R}^2} f(x)\, R^* g(x)\, dx$$

for all $f \in L^1(\mathbb{R}^2)$ and $g \in L^\infty(G)$.

Proof. The result for $\widetilde{R}$ is given in [11, p. 169]. The result for R follows from this by the definitions. ■

We conclude this subsection with a simple identity which relates the adjoints of R and $\tilde{R}$.

Lemma 2.11 *If $g \in L^\infty(S^1 \times \mathbb{R})$, we have the identity*

$$R^* \check{g} = \tilde{R}^* g,$$

where $\check{g} \in L^\infty(G)$ is defined as in def. 2.4.

Proof. The lemma follows by chasing through the definitions. The details may be found in [15, lem. 2.1.16]. ■

2.3 L^2 Theory

It is very natural to define the Radon transform as an operator from $L^1(\mathbb{R}^2)$ to $L^1(G)$. However, in our statistical investigations we will often need to consider the Radon transform on spaces of square-integrable functions. We start by recalling the definition of a Hilbert space, which is a (possibly infinite dimensional) generalization of Euclidean space.

Definition 2.12 Let V be a (real) vector space. An inner product on V is a function $V \times V \to \mathbb{R}$, which we denote by $(x, y) \mapsto (x|y)_V$, that is symmetric in its arguments, separately linear in its arguments, and satisfies $||x||_V^2 \equiv (x, x) \geq 0$ for all $x \in V$ with equality if and only if $x = 0$. The distance between x and y is then defined to be $||x - y||$.

Example 2.13 $\mathbb{R}^d$ becomes an inner product space by taking the inner product to be the usual dot product.

Definition 2.14 Let V be an inner product space. A sequence $\{x_n\}$ of elements of V is said be a Cauchy sequence if $\lim_{n,m\to\infty} ||x_n - x_m||_V = 0$. (Intuitively, the condition means that the x_n stay close to each other as $n \to \infty$.) The sequence $\{x_n\}$ is said to converges to $x \in V$ if $\lim_{n\to\infty} ||x_n - x||_V = 0$. V is said to be a Hilbert space if every Cauchy sequence in V converges to an element of V.

Example 2.15 $\mathbb{R}^d$ is a Hilbert space (see, e.g., [18, thm. 3.11(c)]). Other important examples are the spaces of square-integrable functions introduced in def. 2.16.

Definition 2.16 Let X be as in def. 2.6 and $f_0 \in L^\infty(X)$. The space of real-valued functions f on X such that

$$||f||^2_{L^2(X,f_0)} \stackrel{\text{def}}{=} \int_X f^2(x) f_0(x)\, dx < \infty$$

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

$$(f|g)_{L^2(X,f_0)} \stackrel{\text{def}}{=} \int_X f(x) g(x) f_0(x)\, dx$$

(see, e.g., [5, thm. 5.2.1]).

Remark 2.17 If $f \in L^2(X, f_0)$, then the Cauchy-Schwarz inequality [3, thm. I.1.4] implies that the function $L^2(X, f_0) \to \mathbb{R}$ defined by $g \mapsto (f|g)_{L^2(X,f_0)}$ is continuous. We shall make frequent use of this fact.

Definition 2.18 Define Γ to be the subset of G consisting of lines whose distance from the origin is ≤ 1, i.e., lines with s coordinate of magnitude ≤ 1.

Definition 2.19 Let $p \leq q$ be 1, 2, or ∞. It is frequently convenient to consider a function in $L^q(D)$ as a function in $L^p(\mathbb{R}^2)$ by defining it to be zero outside D. Similarly, it is frequently convenient to consider a function in $L^q(\mathbb{R}^2)$ as a function in $L^p(D)$ by restriction. To avoid confusion, we introduce some notation for this. If $f \in L^q(D)$, we define $\iota_{L^q(D),L^p(\mathbb{R}^2)} f \in L^p(\mathbb{R}^2)$ to agree with f on D and equal 0 outside D. If $f \in L^q(\mathbb{R}^2)$, we define $\pi_{L^q(\mathbb{R}^2),L^p(D)} f \in L^p(D)$ to be the restriction of f to D. We make similar definitions for $L^q(\Gamma)$ and $L^p(G)$.

Definition 2.20 We define $T\colon L^2(D) \to L^2(\Gamma)$ by

$$T : f \mapsto \frac{1}{\pi} \pi_{L^1(G),L^2(\Gamma)} \circ R \circ \iota_{L^2(D),L^1(\mathbb{R}^2)} f.$$

In words, Tf is obtained by considering f as an element of $L^1(\mathbb{R}^2)$, taking the Radon transform and scaling (giving a function in $L^1(G)$), and restricting the result to

Γ. That the definition makes sense, and, in fact, T is continuous, can be see from the stronger result that $R : L^2(D) \to L^2(\Gamma, (1-s^2)^{-1/2})$ is a continuous map [17, thm. II.1.6].

We now turn our attention to the adjoint of T.

Definition 2.21 The adjoint operator of $T : L^2(D) \to L^2(\Gamma)$ is defined to be the unique continuous linear operator T^*: $L^2(\Gamma) \to L^2(D)$ such that $(f|T^*g)_{L^2(D)} = (Tf|g)_{L^2(\Gamma)}$ for all $f \in L^2(D)$ and $g \in L^2(\Gamma)$ [3, sec. II.2].

Remark 2.22 In def. 2.20, we were able to obtain the Radon transform on $L^2(D)$ from the Radon transform on $L^1(\mathbb{R}^2)$ since any square-integrable function on D is absolutely integrable. In the case of the adjoint operator, we cannot simply obtain T^* from R^* since the set of square integrable functions on Γ strictly contains the bounded functions on Γ and R^* is only defined on the latter. However, it turns out that a square-integrable function on Γ can be approximated arbitrarily closely by a bounded function and thus T^* can be obtained essentially by a continuous extension of $\frac{1}{\pi}R^*$. The following definition formalizes this idea of arbitrarily closed approximation.

Definition 2.23 Let L be a linear subspace of a Hilbert space H. L is said to be dense in H if for each $x \in L$ there exists a sequence $\{x_n\}$ in L which converges to x. A continuous function on a Hilbert space is clearly determined by its value on a dense subspace.

Remark 2.24 All the notation should not distract the reader from the point of the next proposition, which is that T^* is essentially the scaled backprojection operator $\frac{1}{\pi}R^*$.

Proposition 2.25 *The adjoint $T^* : L^2(\Gamma) \to L^2(D)$ is given by the unique continuous extension of $\frac{1}{\pi}\pi_{L^\infty(\mathbb{R}^2),L^2(D)} \circ R^* \circ \iota_{L^\infty(\Gamma),L^\infty(G)} : L^\infty(\Gamma) \to L^\infty(D)$ to $L^2(\Gamma)$.*

Proof. Let $f \in L^2(D)$ and $g \in L^\infty(\Gamma)$. Then

$$\begin{aligned}
&(f|\frac{1}{\pi}\pi_{L^\infty(\mathbb{R}^2),L^2(D)} \circ R^* \circ \iota_{L^\infty(\Gamma),L^\infty(G)} g)_{L^2(D)} \\
&= \frac{1}{\pi}\int_D f(x)\pi_{L^\infty(\mathbb{R}^2),L^2(D)} \circ R^* \circ \iota_{L^\infty(\Gamma),L^\infty(G)} g(x)\,dx \\
&= \frac{1}{\pi}\int_{\mathbb{R}^2} \iota_{L^2(D),L^1(\mathbb{R}^2)} \circ f(x) R^* \circ \iota_{L^\infty(\Gamma),L^\infty(G)} g(x)\,dx \\
&= \frac{1}{\pi}\int_G R\iota_{L^2(D),L^1(\mathbb{R}^2)} \circ f(\theta,s)\iota_{L^\infty(\Gamma),L^\infty(G)} g(\theta,s)\,ds\,d\theta \\
&= \frac{1}{\pi}\int_\Gamma \pi_{L^1(\mathbb{R}^2),L^2(D)} \circ R \circ \iota_{L^2(D),L^1(\mathbb{R}^2)} f(\theta,s) g(\theta,s)\,ds\,d\theta \\
&= \int_\Gamma Tf(\theta,s)g(\theta,s)\,ds\,d\theta \\
&= (Tf|g)_{L^2(\Gamma)},
\end{aligned}$$

so

$$T^*g = \frac{1}{\pi}\pi_{L^\infty(\mathbb{R}^2),L^2(D)} \circ R^* \circ \iota_{L^\infty(\Gamma),L^\infty(G)}$$

on $L^\infty(\Gamma)$. Since $L^\infty(\Gamma)$ is dense in $L^2(\Gamma)$, the result follows.

2.4 Inversion

We will now develop an inversion formula for the Radon transform.

Definition 2.26 Let X be as in def. 2.6. Define $C^\infty(X)$ to be the set of smooth, i.e., infinitely differentiable, functions on X.

Definition 2.27 The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is defined to be the subspace of $C^\infty(\mathbb{R}^d)$ consisting of functions f such that

$$|f|_{m,k} \stackrel{\text{def}}{=} \sup_{\beta_1+\cdots+\beta_n \le m} \sup_{x\in\mathbb{R}^d} (1+|x|^2)^k |\partial_{x_1}^{\beta_1}\cdots\partial_{x_n}^{\beta_n} f(x)|$$

is finite for all $m, k \in \mathbb{N}$ [23, ch. 10, ex. IV]. Intuitively, $\mathcal{S}(\mathbb{R}^d)$ is the set of smooth functions that, along with their derivatives, rapidly approach 0 as $|x| \to \infty$. The Fourier transform $\mathcal{F}\colon \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is denoted by $f \mapsto \hat{f}$, where $\hat{f}$ is defined by

$$\hat{f}(\xi) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} e^{-i2\pi x\cdot\xi} f(x)\,dx$$

[23, p. 268, def. 25.1]. The Schwartz space $\mathcal{S}(G)$ is defined to be the subspace of $C^\infty(G)$ consisting of functions g such that

$$|g|_{m,k} \stackrel{\text{def}}{=} \sup_{\beta+\gamma\leq m} \sup_{(\theta,s)\in G} (1+s^2)^k |\partial_\theta^\beta \partial_s^\gamma g(\theta,s)|$$

is finite for all $m, k \in \mathbb{N}$, cf. [10, eq. 2.10].

Remark 2.28 The use of the circumflex to denote the Fourier transform conflicts with its use in denoting an estimate of an unknown quantity. Since both notations are quite standard, we shall just endure this ambiguity. It should be clear from context which meaning is intended.

Definition 2.29 We define the Riesz potential operator I^{-1} for functions $f \in \mathcal{S}(\mathbb{R}^d)$ by

$$(I^{-1}f)\hat{}\,(\xi) \stackrel{\text{def}}{=} 2\pi|\xi|\hat{f}(\xi) \tag{2.2}$$

(cf., [21, sec. V.1]). For $f \in \mathcal{S}(\mathbb{R}^d)$, it is shown in lem. A.1 below that $I^{-1}f \in C^\infty(\mathbb{R}^d)$. For functions on $S^1 \times \mathbb{R}$ and G, we define I^{-1} to act on the second, or "s" variable.

Remark 2.30 In the context of medical imaging, I^{-1} acting on functions on $S^1 \times \mathbb{R}$ or G is the "ramp filter" in the "filtered-backprojection" algorithm.

We are now ready to give the inversion formula for the Radon transform.

Theorem 2.31 *Let $\phi \in \mathcal{S}(\mathbb{R}^2)$. Then*

$$\phi = \frac{1}{2\pi} R^* I^{-1} R\phi$$

(recall that I^{-1} acts on the second or "s" variable of functions on G).

Proof. The inversion formula for the standard Radon transform,

$$\phi = \frac{1}{4\pi} \tilde{R}^* I^{-1} \tilde{R}\phi,$$

is proved in [17, p. 18, thm. 2.1]. From this, using lem. 2.11, it follows that

$$\begin{aligned}\phi &= \frac{1}{4\pi}R^*(I^{-1}\tilde{R}\phi)\check{} \\ &= \frac{1}{2\pi}R^*I^{-1}R\phi. \quad \blacksquare\end{aligned}$$

Remark 2.32 Thm. 2.31 is just the "filtered backprojection" algorithm in medical imaging. The Radon transform of ϕ, $R\phi$, is ramp-filtered (I^{-1}) and backprojected (R^*) to recover ϕ.

3 MVU Linear Estimators

In order to develop some intuition for the problem, we start by considering a restricted class of estimators, the so-called linear estimators, for smooth functionals. We shall construct estimators which are MVU within this restricted class, which we term MVU linear estimators. In sec. 4, we shall see that these MVU linear estimators turn out to be MVU estimators in the class of all estimators.

3.1 The Planar-Imaging Problem

We start by considering the estimation of $\dot{\phi}(f)$ given n independent, identically distributed (i.i.d.) observations distributed according to f, i.e., the planar-imaging problem. For this problem, there is an obvious estimator for $\dot{\phi}(f)$. Indeed, note that $\dot{\phi}(f)$ 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 $\dot{\phi}(f)$ is the sample mean of the derived observations obtained by evaluating the function ϕ at the original observations. We can write this estimator as $(x_1, \ldots, x_n) \mapsto \frac{1}{n}\sum_{i=1}^{n} \phi(x_i)$, where x_i is the ith observation.

Definition 3.1 Let $v \in L^2(D)$. We define the linear estimator generated by v to be the function given by $(x_1, \ldots, x_n) \mapsto n^{-1}\sum_{i=1}^{n} v(x_i)$.

We shall now show that the estimator generated by ϕ is in fact the MVU linear estimator.

Remark 3.2 It is useful to describe the variance of the estimator generated by v by embedding v in $L^2(D, f_0)$. This is the motivation for introducing the weighted L^2 space $L^2(X, f_0)$ in def. 2.16. When f_0 in def. 2.16 is a p.d.f. on the probability space X, the weighting of the integrals in def. 2.16 corresponds to a probabilistic weighting. The space $L^2(X, f_0)$ corresponds to the space of random functions on X with finite second moment when X is distributed according to f_0.

Proposition 3.3 *Let $f_0 \in \mathcal{P}$ and $v \in L^2(D)$. Then the linear estimator generated by v has mean $\dot{v}(f_0)$ and variance*

$$n^{-1}\{||v||^2_{L^2(D,f_0)} - [\dot{v}(f_0)]^2\}.$$

at f_0. The variance is finite since $v \in L^2(D, f_0)$ if and only if $v \in L^2(D)$.

Proof. At f_0, the linear estimator generated by v has mean

$$\begin{aligned}\int_D \cdots \int_D \frac{1}{n}\sum_{i=1}^n v(x_i) f_0(x_1)\cdots f_0(x_n)\,dx_1\cdots dx_n &= \int_D v(x) f_0(x)\,dx \\ &= \dot{v}(f_0)\end{aligned}$$

and variance

$$\begin{aligned}&\int_D \cdots \int_D [\frac{1}{n}\sum_{i=1}^n v(x_i) - \dot{v}(f_0)]^2 f_0(x_1)\cdots f_0(x_n)\,dx_1\cdots dx_n \\ &= \int_D \cdots \int_D \{\frac{1}{n}\sum_{i=1}^n [v(x_i) - \dot{v}(f_0)]\}^2 f_0(x_1)\cdots f_0(x_n)\,dx_1\cdots dx_n \\ &= n^{-2}\int_D \cdots \int_D [\sum_{i=1}^n v(x_i) - \dot{v}(f_0)]^2 f_0(x_1)\cdots f_0(x_n)\,dx_1\cdots dx_n \\ &= n^{-1}\int_D [v(x) - \dot{v}(f_0)]^2 f_0(x)\,dx \\ &= n^{-1}\{\int_D v^2(x) f_0(x)\,dx - [\dot{v}(f_0)]^2\} \\ &= n^{-1}\{||v||^2_{L^2(D,f_0)} - [\dot{v}(f_0)]^2\}.\end{aligned}$$

The condition $f_0 \in \mathcal{P}$ implies that $v \in L^2(D)$ if and only if $v \in L^2(D, f_0)$. It follows that both $||v||^2_{L^2(D,f_0)}$ and $\dot{v}(f_0)$ are finite. ■

Remark 3.4 The condition $v \in L^2(D)$ is clearly a necessary and sufficient for the variance of an estimate of the form $(x_1, \ldots, x_n) \mapsto \frac{1}{n}\sum_{i=1}^n v(x_i)$ to exist at $f_0 \in \mathcal{P}$.

The next proposition shows that ϕ is the only function in $L^2(D)$ that generates an unbiased estimator.

Proposition 3.5 *Let $v \in L^2(D)$ and suppose the linear estimator generated by v is an unbiased estimator of $\dot{\phi}(f)$, i.e., $\dot{v}(f) = \dot{\phi}(f)$ for all $f \in \mathcal{P}$. Then $v = \phi$.*

Proof. Since $\dot{v}$ and $\dot{\phi}$ are linear functions on $L^2(D)$, they must agree on the linear span of $\mathcal{P}$. A routine argument shows that the linear span of $\mathcal{P}$ is dense in $L^2(D)$. (It is well-known that the step functions are dense in $L^2(D)$ (see, e.g., [8, prop. 6.7]). Now all the step functions are in the linear span of $\mathcal{P}$ except those taking a value of 0. But it is clear that step functions taking a value of 0 can be approximated arbitrarily closely in $L^2(D)$ by a function in the linear span of $\mathcal{P}$). Since $\dot{v}$ and $\dot{\phi}$ are continuous, the result follows. ■

To summarize:

Theorem 3.6 *The estimator generated by ϕ is the unique MVU linear estimator for the planar-imaging problem. Its variance at f_0 is*

$$n^{-1}\{||\phi||^2_{L^2(D,f_0)} - [\dot{v}(f_0)]^2\}.$$

Proof. Combine props. 3.3 and 3.5. ■

We shall see in sec. 4.1 that the estimator generated by ϕ is actually a uniformly MVU estimator for $\dot{\phi}(f)$ in the class of all estimators.

3.2 The ECT Problem

We now turn our attention to the construction of MVU linear estimators for the ECT problem.

Definition 3.7 Let f_u denote the uniform distribution on D, i.e., f_u is the constant function $1/\pi$ on D. Define $g_u = Tf_u$.

Definition 3.8 Let $\psi \in L^2(\Gamma, g_u)$. We define the linear estimator generated by ψ to be the function $\Gamma^n \to \mathbb{R}$ given by $(l_1, \ldots, l_n) \mapsto n^{-1}\sum_{i=1}^n \psi(l_i)$.

Proposition 3.9 *Let $f_0 \in \mathcal{P}$ and $\psi \in L^2(\Gamma, g_u)$. Then the linear estimator generated by ψ has mean $\dot{\psi}(Tf_0)$ and variance*

$$n^{-1}\left(\int_\Gamma \psi^2(\theta,s)Tf_0(\theta,s)\,ds\,d\theta - [\dot{\psi}(Tf_0)]^2\right).$$

at f_0. The variance is finite since $\psi \in L^2(\Gamma, Tf_0)$ if and only if $\psi \in L^2(\Gamma, g_u)$.

Proof. Essentially the same as the proof of prop. 3.3. ■

Remark 3.10 The condition $\psi \in L^2(\Gamma, g_u)$ is clearly necessary and sufficient for the variance of an estimate of the form $(l_1, \ldots, l_n) \mapsto \frac{1}{n}\sum_{i=1}^n \psi(l_i)$ to exist at $f_0 \in \mathcal{P}$.

Our next goal is to find a criterion for an element of $L^2(\Gamma, g_u)$ to generate an unbiased estimator. The desired result is given by prop. 3.16. In the interim, we shall define the operator T_{f_u} and see that the condition $T_{f_u}^*\psi = \phi$ given in prop. 3.16 has the intuitive content that ψ backprojects to ϕ.

Since we are working in $L^2(\Gamma, g_u)$, we need to reformulate T as an operator T_{f_u} whose range is $L^2(\Gamma, g_u)$. The next definition does this in slightly greater generality.

Definition 3.11 For $f_0 \in \mathcal{P}$, define the linear operator $T_{f_0}\colon L^2(D) \to L^2(\Gamma, Tf_0)$ by $f \mapsto \frac{Tf}{Tf_0}$. It is shown in [17, thm. II.1.6] that T_{f_u} is continuous, from which it easily follows that T_{f_0} is continuous.

Remark 3.12 $T_{f_0}f(\theta, s)$ is just the likelihood ratio of the observation (θ, s) under the statistical hypotheses f and f_0.

The next proposition shows that the adjoint operator $T_{f_u}^* : L^2(\Gamma, g_u) \to L^2(D)$ is just the continuous extension of T^*. Its intuitive content is that $T_{f_u}^*$ is just the backprojection operator. (The situation is analogous to the one discussed in rem. 2.22.)

Proposition 3.13 *The adjoint of $T_{f_0} : L^2(D) \to L^2(\Gamma, Tf_0)$ is given by the unique continuous extension of $T^* : L^2(\Gamma) \to L^2(D)$ to $L^2(\Gamma, Tf_0)$.*

Proof. Let $f \in L^2(D)$ and $g \in L^2(\Gamma)$. Then

$$\begin{aligned}
(T_{f_0} f | g)_{L^2(\Gamma, g_0)} &= \int_\Gamma \frac{Tf(\theta,s)}{Tf_0(\theta,s)} g(\theta,s) Tf_0(\theta,s)\, ds\, d\theta \\
&= \int_\Gamma Tf(\theta,s) g(\theta,s)\, ds\, d\theta \\
&= (Tf|g)_{L^2(\Gamma)} \\
&= (f|T^* g)_{L^2(D)}.
\end{aligned}$$

Since $L^2(\Gamma)$ is dense in $L^2(\Gamma, Tf_0)$, the result follows. ■

Corollary 3.14 *If $f_0 \in \mathcal{P}$, then $T^*_{f_0} = T^*_{f_u}$.*

Proof. By prop. 3.13, $T^*_{f_0}$ and $T^*_{f_u}$ agree on the dense subset $L^2(\Gamma)$. ■

Corollary 3.15 *Let $\phi \in \mathcal{S}(\mathbb{R}^2)$, $f_0 \in \mathcal{P}$, and $\psi \in L^2(\Gamma, g_u)$. Then $T^*_{f_0}\psi = \phi$ if and only if $T^*_{f_u}\psi = \phi$.* ■

Proposition 3.16 *Let $\phi \in \mathcal{S}(\mathbb{R}^2)$ and $\psi \in L^2(\Gamma, g_u)$. Then the linear estimator generated by ψ is unbiased if and only if $T^*_{f_u}\psi = \phi$.*

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

$$\begin{aligned}
\dot{\psi}(Tf) &= \int_\Gamma \psi(\theta,s) Tf(\theta,s) ds\, d\theta \\
&= \int_\Gamma \psi(\theta,s) \frac{Tf(\theta,s)}{Tf_u(\theta,s)} Tf_u(\theta,s)\, ds\, d\theta \\
&= (\psi, T_{f_u} f)_{L^2(\Gamma, g_u)} \\
&= (T^*_{f_u}\psi, f)_{L^2(D)} \\
&= (\phi, f)_{L^2(D)} \\
&= \dot{\phi}(f).
\end{aligned}$$

for all $f \in L^2(D)$, in particular for all $f \in \mathcal{P}$. Conversely, suppose $\dot{\psi}(Tf) = \dot{\phi}(f)$ for all $f \in \mathcal{P}$. Since $\dot{\psi}$ and $\dot{\phi}$ are linear operators, $\dot{\psi}(Tf) = \dot{\phi}(f)$ for all f in the linear

span of $\mathcal{P}$. Since the linear span of $\mathcal{P}$ is dense in $L^2(D)$, it follows that $\dot{\psi}(Tf) = \dot{\phi}(f)$ for all $f \in L^2(D)$. This says that $T^*\phi = \psi$. The result now follows from prop. 3.13. ■

Intuitively, prop. 3.16 says that a function ψ on Γ generates an unbiased linear estimator if and only if it backprojects to ϕ, i.e., if $T^*_{f_u}\psi = \phi$. We are therefore lead to investigate the existence and uniqueness of solutions to the equation $T^*_{f_u}\psi = \phi$. We start by answering the existence question in the affirmative.

Definition 3.17 For $\phi \in \mathcal{S}(\mathbb{R}^2)$, define $F\phi \in C^\infty(G)$ by $F\phi \stackrel{\text{def}}{=} \frac{1}{2}I^{-1}R\phi$ (cf. lem. A.1).

Proposition 3.18 *Let $\phi \in \mathcal{S}$. Then $T^*_{f_u}F\phi = \phi$.*

Proof. By props. 3.13, 2.25, and 2.31,

$$\begin{aligned} T^*_{f_u}F\phi &= T^*F\phi \\ &= \frac{1}{\pi}R^*F\phi \\ &= \phi. \quad \blacksquare \end{aligned}$$

Remark 3.19 The standard approach to quantifying the fraction of activity in a ROI may be termed the "filtered backprojection and integration" algorithm, i.e., one estimates the image by an estimate $\hat{f}$ using the filtered backprojection algorithm and computes the integral of $\hat{f}$ over the ROI. Since this algorithm is usually implemented using a discrete version of the Radon transform, it does not fit precisely into our framework. However, we shall give a heuristic argument to the effect that this standard procedure is essentially the linear estimator generated by $F\phi$.

The idea of filtered backprojection is to replace the quantity $R\phi$ in the inversion formula given in thm. 2.31 with a smooth approximation to the observations. Let $\omega \in \mathcal{S}(G)$ be an approximate delta function in the second variable, i.e., for fixed θ, $\omega(\theta, s)$ as a function of s approximates a delta function. Let g_n be the smooth function on G obtained by convolving the sum of n point masses of mass n^{-1} at

the observations $l_1, \ldots, l_n$, so that g_n is a smooth approximation to the empirical distribution generated by the observations. (The convolution is taken with respect to the second variable only.) Then the filtered-backprojection estimate, $\hat{f}$, of f is $\frac{1}{2}R^* I^{-1} g_n$ and

$$\begin{aligned}
\dot{\phi}(\hat{f}) &= \frac{1}{2}\int_{\mathbb{R}^2} \phi(x) R^* I^{-1} g_n(x)\, dx \\
&= \frac{1}{2}\int_0^\pi \int_{-\infty}^{\infty} R\phi(\theta, s) I^{-1} g_n(\theta, s)\, ds\, d\theta \\
&= \frac{1}{2}\int_0^\pi \int_{-\infty}^{\infty} \overline{\widehat{R\phi}(\theta,\eta)} (I^{-1} g_n)^\wedge(\theta,\eta)\, d\eta\, d\theta \\
&= \frac{1}{2}\int_0^\pi \int_{-\infty}^{\infty} \overline{\widehat{R\phi}(\theta,\eta)} \hat{g}_n(\theta,\eta) |\eta|\, d\eta\, d\theta \\
&= \frac{1}{2}\int_0^\pi \int_{-\infty}^{\infty} \overline{(I^{-1} R\phi)^\wedge(\theta,\eta)} g_n(\theta,\eta)\, d\eta\, d\theta \\
&= \frac{1}{2}\int_0^\pi \int_{-\infty}^{\infty} I^{-1} R\phi(\theta, s) g_n(\theta, s)\, ds\, d\theta \\
&= \int_0^\pi \int_{-\infty}^{\infty} F\phi(\theta, s) g_n(\theta, s)\, ds\, d\theta \\
&= (F\phi)^{\cdot}(g_n),
\end{aligned}$$

where the overbar indicates complex conjugate and third equality follows from the Plancherel theorem (see, e.g., [8, thm. 8.29]). In the limit as g_n approaches a delta function in the second variable, this quantity will approach the estimator generated by $F\phi$. A discrete version of this argument may be found in [12].

Prop. 3.18 establishes the existence of an unbiased linear estimator. The next proposition addresses its uniqueness.

Definition 3.20 If A is a linear operator, let $\mathcal{N}(A)$ denote its nullspace, i.e., the set of vectors x such that $Ax = 0$.

Proposition 3.21 *Let $\phi \in \mathcal{S}$. Then $\psi \in L^2(\Gamma, g_u)$ satisfies $T^*_{f_u}\psi = \phi$ if and only if $\psi \in F\phi + \mathcal{N}(T^*_{f_u})$.*

Proof. If $T^*_{f_u}\psi = \phi$, then $T^*_{f_u}(\psi - F\phi) = 0$, so $\psi - F\phi \in \mathcal{N}(T_{f_u})$. Conversely, if $\psi - F\phi \in \mathcal{N}(T^*_{f_u})$, then $T^*_{f_u}\psi = \phi$. ■

Intuitively, prop. 3.21 says that the functions which generate unbiased linear estimators form an affine linear space obtained by adding functions which backproject to 0 to $F\phi$. Thus the uniqueness question boils down to whether or not there exist nonzero functions on Γ that backproject to 0. The next proposition shows that this question is equivalent to whether or not $T_{f_u}L^2(D)$ is dense in $L^2(\Gamma, Tf_0)$, i.e., whether or not any function on Γ can be approximated arbitrarily closely by the Radon transform of a function on D.

Definition 3.22 Let L be a linear subspace of a Hilbert space H. The closure of L, which we denote by $\bar{L}$, is the subspace of H consisting of elements f such that there exists a sequence in L converging to f. It is obvious that $L \subseteq \overline{L}$ (consider constant sequences); the point is that a sequence $\{f_n\} \subseteq L$ may converge to an $f \notin L$. By definition, L is dense in $\bar{L}$.

Definition 3.23 Let H be a Hilbert space and $x, y \in H$. Then x is said to be orthogonal to y if $(x|y)_H = 0$. If L is a subspace of H, then the subspace of H whose elements are orthogonal to every element of L is termed the orthogonal complement to L and denoted by $L^\perp$. We note the identity $(\bar{L}^\perp)^\perp = \bar{L}$ [3, cor. I.2.9].

Proposition 3.24 *Let $f_0 \in \mathcal{P}$. Then*

$$\mathcal{N}(T_{f_0}^*) = \overline{T_{f_0}L^2(D)}^\perp.$$

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., [3, secs. I.2, II.2]). (This is just the Hilbert space analogue of the familiar fact from linear algebra that the range space of a real matrix is orthogonal to the nullspace of its transpose.) ■

Corollary 3.25 $\mathcal{N}(T_{f_0}^*) = \{0\}$ *if and only if* $T_{f_0}L^2(D)$ *is dense in* $L^2(\Gamma, Tf_0)$.

Proof. If $T_{f_0}L^2(D)$ is dense in $L^2(\Gamma, Tf_0)$, then $\overline{T_{f_0}L^2(D)} = L^2(\Gamma, Tf_0)$, so

$$\begin{aligned}\mathcal{N}(T^*_{f_0}) &= L^2(\Gamma, Tf_0)^\perp \\ &= \{0\}.\end{aligned}$$

Conversely, if $\mathcal{N}(T^*_{f_0}) = \{0\}$, then $\overline{T_{f_0}L^2(D)}^\perp = \{0\}$, so

$$\begin{aligned}\overline{T_{f_0}L^2(D)} &= \left[\overline{T_{f_0}L^2(D)}^\perp\right]^\perp \\ &= \{0\}^\perp \\ &= L^2(\Gamma, Tf_0). \ \blacksquare\end{aligned}$$

Intuitively, prop. 3.24 tells us that the generators of unbiased linear estimators differ by functions that are orthogonal to the range of the Radon transform. In sec. 5, we shall see that $\overline{T_{f_u}L^2(D)}^\perp$ is infinite dimensional, implying there is an infinite-dimensional space of unbiased linear estimators. In the remainder of this section, we shall consider, in a somewhat abstract manner, the construction of MVU linear estimators for the ECT problem. It turns out that the MVU linear estimator is generated by the projection of any $\psi \in L^2(\Gamma, g_u)$ that generates an unbiased estimator onto $\overline{T_{f_0}L^2(D)}$.

Definition 3.26 Let L be a subspace of a Hilbert space H. L is said to be closed if $\bar{L} = L$. If L is closed and $g \in H$, there is a unique element $p_L g \in L$, termed the projection of g onto L, such that

$$||g - p_L g||_H = \inf_{g' \in L} ||g - g'||_H$$

[8, thm. 5.24], i.e., $p_L g$ is the element of L that is closest to g. It has the property $g - p_l g \in L^\perp$. Thus if $g' \in L$,

$$(g|g')_H = (p_L g|g')_H. \tag{3.1}$$

We also have the formulas

$$g = p_L g + p_{L^\perp} g$$

and

$$||g||_H^2 = ||p_L g||_H^2 + ||p_{L^\perp} g||_H^2.$$

Theorem 3.27 *Let $f_0 \in \mathcal{P}$ and $\phi \in \mathcal{S}(\mathbb{R}^2)$. Suppose $\psi \in L^2(\Gamma, Tf_0)$ satisfies $T_{f_0}^* \psi = \phi$. Then the estimator generated by $p_{\overline{T_{f_0} L^2(D)}} \psi$ is an MVU linear estimator for $\dot{\phi}(f)$ at f_0 in the ECT problem. Its variance is*

$$n^{-1} \left(||p_{\overline{T_{f_0} L^2(D)}} \psi||_{L^2(\Gamma, Tf_0)}^2 - [\dot{\psi}(Tf_0)]^2 \right) \tag{3.2}$$

at f_0.

Proof. We first note that $p_{\overline{T_{f_0} L^2(D)}} \psi$ generates a unbiased estimator since $\psi - p_{\overline{T_{f_0} L^2(D)}} \psi \in \overline{T_{f_0} L^2(D)}^\perp$ and hence $\psi - p_{\overline{T_{f_0} L^2(D)}} \psi$ is in the nullspace of $T_{f_0}^*$. Now suppose $\psi' \in L^2(\Gamma, Tf_0)$ generates an unbiased estimate of $\dot{\phi}(f)$. Then $T_{f_0}^*(\psi' - \psi) = 0$, so $\psi' - \psi$ is in the nullspace of $T_{f_0}^*$ and hence orthogonal to $\overline{T_{f_0} L^2(D)}$. By eq. 3.1, it follows that $p_{\overline{T_{f_0} L^2(D)}} \psi' = p_{\overline{T_{f_0} L^2(D)}} \psi$. Thus

$$\begin{aligned} ||\psi'||_{L^2(\Gamma, Tf_0)}^2 &= ||p_{\overline{T_{f_0} L^2(D)}} \psi||_{L^2(\Gamma, Tf_0)} + ||p_{\overline{T_{f_0} L^2(D)}^\perp} \psi||_{L^2(\Gamma, Tf_0)} \\ &\geq ||p_{\overline{T_{f_0} L^2(D)}} \psi||_{L^2(\Gamma, Tf_0)}. \quad \blacksquare \end{aligned}$$

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 3.28 In sec. 4, we shall see that the estimator constructed in thm. 3.27 is an MVU estimator in the class of all estimators. However, since the minimizing ψ will change as a function of f_0, such an estimator can no longer be a *uniformly* MVU estimator.

Remark 3.29 From the above, $\psi \in L^2(\Gamma, f_u)$ generates an unbiased estimator if and only if $\dot{\psi}(Tf) = \dot{\phi}(f)$ for all $f \in \mathcal{P}$. This says that the functional generated by ϕ on

$\mathcal{P}$ is the same as the functional generated by ψ on $T\mathcal{P}$. This idea can be expressed by saying that ψ is an observation-domain representation of ϕ.

Remark 3.30 Under the assumption that $f \in \mathcal{P}$, the above theory extends to any $\phi \in C^\infty(\mathbb{R}^2)$. Indeed, 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. [8, lem 8.10]) and thus generates the same functional on $\mathcal{P}$ as ϕ.

4 MVU Estimators

In this section, we will show that the MVU linear estimators constructed in sec. 3 are actually MVU estimators in the class of all estimators.

4.1 Uniformly MVU Estimators for Planar Imaging

In this subsection, we will prove the claim made in sec. 3 that the linear estimator generated by ϕ is an uniformly MVU estimator for $\dot{\phi}(f)$ in the planar-imaging problem.

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

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

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

Using the fact that

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

it is clear that the $f_t \in \mathcal{P}$. Consider the one-dimensional subproblem of estimating $\dot{\phi}(f_t)$ from n i.i.d. observations distributed according to f_t for $|t| < \sup_{x\in D} |\phi(x) - \dot{\phi}(f_0)|$. The Cramér-Rao inequality [13, art. 399D] states that the variance of any unbiased estimator of $\dot{\phi}(f_t)$ at f_0 is bounded below by

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

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

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

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

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

Again using eq. 4.1, we also have

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

It follows that the variance of an unbiased estimator of $\dot{\phi}(f_t)$ at f_0 must be $\geq n^{-1}\{||\phi||^2_{L^2(D)} - [\dot{\phi}(f_0)]^2\}$. By thm. 3.6, 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 uniformly MVU estimator for $\dot{\phi}(f)$. ■

4.2 MVU Estimators for ECT

In this subsection, we will show that the MVU linear estimator constructed in sec. 3.2 is an MVU estimator for the ECT problem.

Theorem 4.2 *Let $f_0 \in \mathcal{P}$, $g_0 = Tf_0$, and $\phi \in \mathcal{S}(\mathbb{R}^2)$. The estimator generated by $\psi \stackrel{\text{def}}{=} p_{\overline{T_{f_0}L^2(D)}} F\phi$ is an MVU estimator for $\dot{\phi}(f)$ at f_0 in the ECT problem.*

Proof. The proof is analogous to that of thm. 4.1 in that the Cramér-Rao inequality is applied to one-dimensional subproblems to show that the variance of any unbiased estimator cannot be less than that of the MVU linear estimator. In thm. 4.1, we considered perturbations of f_0 in the direction $\phi(x) - \dot{\phi}(f_0)$. In the ECT problem, we consider perturbations of g_0 (approximately) in the direction $p_{\overline{T_{f_0}L^2(D)}} F\phi - \dot{\psi}(g_0)$. For this to make sense, the first step is to show that $p_{\overline{T_{f_0}L^2(D)}} F\phi - \dot{\psi}(g_0) \in \overline{T_{f_0}L^2_0(D)}$.

We start by noting that constant functions are in $T_{f_0}L^2(D)$, since $T_{f_0}f_0 = 1$. It follows that $p_{\overline{T_{f_0}L^2(D)}} F\phi - \dot{\psi}(g_0) \in \overline{T_{f_0}L^2(D)}$. Let $L^2_0(D)$ and $L^2_0(\Gamma, g_0)$ denote the

subspaces of $L^2(D)$ and $L^2(\Gamma, g_0)$, respectively, whose elements h satisfy $\int_D h(x)\,dx = 0$ and $\int_\Gamma h(s,\theta)g_0(s,\theta)\,ds\,d\theta = 0$. Then $p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot{\psi}(g_0) \in \overline{T_{f_0}L^2(D)} \cap L_0^2(\Gamma, g_0)$, since $p_{\overline{T_{f_0}L^2(D)}}F\phi$ generated an unbiased linear estimator. We claim that $\overline{T_{f_0}L^2(D)} \cap L_0^2(\Gamma, g_0) \subseteq \overline{T_{f_0}L_0^2(D)}$. To prove the claim, let $f \in L^2(D)$. Then

$$\begin{aligned} \int_D f(x)\,dx &= \int_D f(x)T^*1(x)\,dx \\ &= \int_\Gamma Tf(\theta,s)\,ds\,d\theta \\ &= \int_\Gamma T_{f_0}f(\theta,s)g_0\,ds\,d\theta. \end{aligned} \tag{4.2}$$

Thus if $f \in L^2(D)\backslash L_0^2(D)$, then $T_{f_0}f \notin L_0^2(\Gamma, g_0)$. This says that $T_{f_0}L^2(D) \cap L_0^2(\Gamma, g_0) \subseteq T_{f_0}L_0^2(D)$. Now suppose $g \in \overline{T_{f_0}L^2(D)} \cap L_0^2(\Gamma, g_0)$. Then there exists a sequence $\{g_n\}$ in $T_{f_0}L^2(D)$ such that $g_n \to g$ in $L^2(\Gamma, g_0)$. Since $g \in L_0^2(\Gamma, g_0)$, $\lim_{n\to\infty} \int_\Gamma g_n(\theta,s)g_0(\theta,s)\,ds\,d\theta = 0$. We can write each g_n as $T_{f_0}f_n$ for some $f_n \in L^2(D)$. By eq. 4.2, $\lim_{n\to\infty} \int f_n(x)\,dx = 0$. Consider the sequence $\{f_n - \int_D f_n(x)\,dx\}$ in $L_0^2(D)$. We have

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

so $g \in \overline{T_{f_0}L_0^2(D)}$. This prove the claim that $\overline{T_{f_0}L^2(D)} \cap L_0^2(\Gamma, g_0) \subseteq \overline{T_{f_0}L_0^2(D)}$. From the claim, we can now conclude that $p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot{\psi}(g_0) \in \overline{T_{f_0}L_0^2(D)}$.

Let $\epsilon > 0$ be given. Since $p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot{\psi}(g_0) \in \overline{T_{f_0}L_0^2(D)}$, we can choose $\eta \in L_0^2(D)$ such that $||T_{f_0}\eta - p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot{\psi}(g_0)||_{L^2(\Gamma,g_0)} < \epsilon/2$. Now the simple functions, i.e., functions taking only a finite number of values, are dense in $L_0^2(D)$ [2, p. 52]. Using this fact and the fact that T_{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 $g_t \in T\mathcal{P}$ by $g_t = T(f_0 + t\eta)$. Consider the one-dimensional subproblem of estimating $(F\phi)^{\cdot}(g_t)$ 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 $\dot{\phi}(f_t)$.) We will now compute the Fisher

information for the estimation of t at $t = 0$. We have

$$\begin{aligned}\partial_t \log g_t|_{t=0} &= \left.\frac{\partial_t g_t}{g_t}\right|_{t=0} \\ &= \frac{T\eta}{Tf_0} \\ &= T_{f_0}\eta\end{aligned}$$

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

$$\begin{aligned}
&\partial_t (F\phi)^{\cdot}(g_t)|_{t=0} \\
&= \partial_t \int_\Gamma F\phi(\theta,s)T[f_0(\theta,s) + t\eta(\theta,s)]\,ds\,d\theta|_{t=0} \\
&= \int_\Gamma F\phi(\theta,s)T\eta(\theta,s)\,ds\,d\theta \\
&= \int_\Gamma F\phi(\theta,s)T_{f_0}\eta(\theta,s)g_0(\theta,s)\,ds\,d\theta \\
&= \int_\Gamma [F\phi(\theta,s) - \dot\psi(g_0)]T_{f_0}\eta(\theta,s)g_0(\theta,s)\,ds\,d\theta + \dot\psi(g_0) \\
&= \int_\Gamma p_{\overline{T_{f_0}L^2(D)}}[F\phi(\theta,s) - \dot\psi(g_0)]T_{f_0}\eta(\theta,s)g_0(\theta,s)\,ds\,d\theta + \dot\psi(g_0) \\
&= \int_\Gamma [p_{\overline{T_{f_0}L^2(D)}}F\phi(\theta,s) - \dot\psi(g_0)][p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot\psi(g_0)]g_0(\theta,s)\,ds\,d\theta \\
&\quad + \int_\Gamma [p_{\overline{T_{f_0}L^2(D)}}F\phi(\theta,s) - \dot\psi(g_0)] \\
&\quad \times\{T_{f_0}\zeta(\theta,s) - [p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot\psi(g_0)]\}g_0(\theta,s)\,ds\,d\theta \\
&\geq ||p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot\psi(g_0)||^2_{L^2(\Gamma,g_0)} \\
&\quad -||p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot\psi(g_0)||_{L^2(\Gamma,g_0)}||T_{f_0}\zeta(\theta,s) - p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot\psi(g_0)||_{L^2(\Gamma,g_0)} \\
&\geq ||p_{\overline{T_{f_0}L^2_0(D)}}F\phi - \dot\psi(g_0)||_{L^2(\Gamma,g_0)}(||p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot\psi(g_0)||_{L^2(\Gamma,g_0)} - \epsilon/2)
\end{aligned}$$

(the first inequality follows from the Cauchy-Schwarz inequality). It follows from the Cramér-Rao inequality [13, art. 389D] that the variance of an unbiased estimator of $(F\phi)^{\cdot}(g_t)$ at g_0 must be at least

$$\begin{aligned}
&\frac{[\partial_t (F\phi)^{\cdot}(g_t)|_{t=0}]^2}{n||T_{f_0}\zeta||^2_{L^2(\Gamma,g_0)}} \\
&\geq \frac{||p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot\psi(g_0)||^2_{L^2(\Gamma,g_0)}(||p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot\psi(g_0)||_{L^2(\Gamma,g_0)} - \epsilon/2)^2}{n(||p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot\psi(g_0)||_{L^2(\Gamma,g_0)} + \epsilon/2)^2}.
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, we conclude that the variance of an unbiased estimator of $(F\phi)^{\cdot}(g_t)$ at g_0 must be $\geq n^{-1}||p_{\overline{T_{f_0}L^2(D)}}F\phi - \dot{\psi}(g_0)||^2_{L^2(\Gamma,g_0)}$. The unbiased linear estimator constructed in thm. 3.27 achieves this lower bound. ■

Remark 4.3 Unlike the estimator in thm. 4.1, the MVU estimator in thm. 4.2 depends on f_0. It is thus not a *uniformly* MVU estimator.

5 Construction of Projection Operators

In sec. 4.2, we saw that the linear estimator generated by $\psi_{\phi,f_0} \stackrel{\text{def}}{=} p_{\overline{T_{f_0}L^2(D)}} F\phi$ is an MVU estimator for $\dot{\phi}(f)$ at f_0 in the ECT problem. In this section, we will express the projection operator $p_{\overline{T_{f_0}L^2(D)}}$ in a concrete way that is suitable for numerical calculations.

We start in sec. 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, which is carried out in sec. 5.2.

5.1 The Uniform Distribution

Definition 5.1 Let f_u and g_u be as in def. 3.7. Explicitly, $g_u(\theta, s) = \frac{2}{\pi^2}\sqrt{1-s^2}$ (cf. [4, sec. 2.5, ex. 4], [14, sec. 2.1]). For future reference, we note that the marginal distribution of the "s" variable is given by the probability density $\frac{2}{\pi}\sqrt{1-s^2}$ on $[-1,1]$. We shall denote this probability density by ρ_u.

Definition 5.2 A subset of a Hilbert space H is said to be orthonormal if each of its elements have norm 1 and distinct elements are orthogonal to each other. An orthonormal subset is said to be an orthonormal basis if its linear span in dense in H. If $\{x_i\}$ is an orthonormal basis for a Hilbert space H, then each $x \in H$ can be expressed as $x = \sum_i (x_i|x)_H x_i$ and $||x||_H^2 = \sum_i (x_i|x)_H^2$.

Definition 5.3 Let $\mathbb{Z}$, $\mathbb{N}$, and $\mathbb{N}^+$ denote the sets of integers, nonnegative integers, and positive integers, respectively. For $m \in \mathbb{N}$, define the functions U_m: $[-1,1] \to \mathbb{R}$ by

$$U_m(\cos\theta) \stackrel{\text{def}}{=} \frac{\sin[(m+1)\theta]}{\sin\theta}.$$

The U_m are called the Chebyshev polynomials of the second kind [4, sec. 7.6] [14, sec. 2]. They form an orthonormal basis for $L^2([-1,1],\rho_u)$ [4, 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 $U_m(\theta, s) \stackrel{\text{def}}{=} U_m(s)$. For $l \in \mathbb{N}^+$ and $m \in \mathbb{N}$, define the functions $a_{l,m} : S^1 \times \mathbb{R} \to \mathbb{R}$ and $b_{l,m} : S^1 \times \mathbb{R} \to \mathbb{R}$ by

$$a_{l,m}(\theta, s) \stackrel{\text{def}}{=} \sqrt{2} U_m(s) \cos(l\theta),$$

and

$$b_{l,m}(\theta, s) \stackrel{\text{def}}{=} \sqrt{2} U_m(s) \sin(l\theta).$$

Definition 5.4 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 G given by the formula $g(\theta, s)$. In particular, since U_m is even when m is even and $a_{l,m}$ and $b_{l,m}$ are even when $l + m$ is even, we shall view U_m, $a_{l,m}$, and $b_{l,m}$ as being defined on Γ by the formulas given in def. 5.3 when m and $l + m$ are even, respectively.

Definition 5.5 Define

$$B_u = \{U_m : m \in 2\mathbb{N}\} \cup \{a_{l,m}, b_{l,m} : l \in \mathbb{N}^+, m \in l + 2\mathbb{N}\}$$

and

$$B'_u = \{a_{l,m}, b_{l,m} : l \in \mathbb{N}^+,\ m \in \{l \bmod 2, l \bmod 2 + 2, \ldots, l - 2\}\}.$$

Proposition 5.6 *B_u and B'_u are orthonormal bases for $\overline{T_{f_u} L^2(D)}$ and its orthogonal complement in $L^2(\Gamma, T f_u)$, respectively.*

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

$$c_{l,m}(\theta, s) \stackrel{\text{def}}{=} U_m(s) e^{il\theta}.$$

Note that $c_{l,m}$ is even (odd) when $|m - l|$ is even (odd). An orthonormal basis for $\overline{T_{f_u} L^2(D)} \subset L^2(\Gamma, g_u)$ is given by $c_{l,m}$ with $l \in \mathbb{Z}$ and $m \in |l| + 2\mathbb{N}$ [4, sec. 7.6] [14, sec. 2.3], where the $c_{l,m}$ are interpreted as functions on Γ per def. 5.4.

Since we are dealing with (real) probability densities and integral 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 $\{c_{l,m} : l \in \mathbb{N},\ m \in l + 2\mathbb{N}\}$ yields B_u, which proves the first assertion.

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

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

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

$$\{\frac{1}{\sqrt{2}}U_m : m \in \mathbb{N}\} \cup \{\frac{1}{\sqrt{2}}a_{l,m}, \frac{1}{\sqrt{2}}b_{l,m} : l \in \mathbb{N}^+, m \in \mathbb{N}\}.$$

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

$$\begin{aligned}
||\psi||^2_{L^2(\Gamma,g_u)} &= \int_\Gamma \psi^2(\theta,s) g_u(\theta,s)\, ds\, d\theta \\
&= \frac{1}{2}\int_{S^1\times[-1,1]} \tilde{\psi}^2(\theta,s)\tilde{g}_u(\theta,s)\, ds\, d\theta \\
&= \frac{1}{4}\Bigg(\sum_{m\in\mathbb{N}} (U_m|\tilde{\psi})^2_{L^2(S^1\times[-1,1],\tilde{g}_u)} \\
&\qquad + \sum_{l\in\mathbb{N}^+}\sum_{m\in\mathbb{N}} (a_{l,m}|\tilde{\psi})^2_{L^2(S^1\times[-1,1],\tilde{g}_u)} + (b_{l,m}|\tilde{\psi})^2_{L^2(S^1\times[-1,1],\tilde{g}_u)}\Bigg) \\
&= \frac{1}{4}\Bigg(\sum_{m\in 2\mathbb{N}} (U_m|\tilde{\psi})^2_{L^2(S^1\times[-1,1],\tilde{g}_u)} \\
&\qquad + \sum_{l\in\mathbb{N}^+}\sum_{m\in l \bmod 2+2\mathbb{N}} (a_{l,m}|\tilde{\psi})^2_{L^2(S^1\times[-1,1],\tilde{g}_u)} + (b_{l,m}|\tilde{\psi})^2_{L^2(S^1\times[-1,1],\tilde{g}_u)}\Bigg) \\
&= \Bigg(\sum_{m\in 2\mathbb{N}} (U_m|\psi)^2_{L^2(\Gamma,g_u)} \qquad (5.1) \\
&\qquad + \sum_{l\in\mathbb{N}^+}\sum_{m\in l \bmod 2+2\mathbb{N}} (a_{l,m}|\psi)^2_{L^2(\Gamma,g_u)} + (b_{l,m}|\psi)^2_{L^2(\Gamma,g_u)}\Bigg),
\end{aligned}$$

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

Having established that B_u is an orthonormal basis for $\overline{T_{f_u} L^2(D)}$, it is now straightforward to express $p_{\overline{T_{f_0} L^2(D)}}$ in terms of this basis.

Corollary 5.7 *The MVU estimator at f_u for the ECT problem is generated by*

$$p_{\overline{T_{f_u} L^2(D)}} F\phi = \sum_{m \in 2\mathbb{N}} (U_m | F\phi)_{L^2(\Gamma, g_u)} U_m \tag{5.2}$$
$$+ \sum_{l \in \mathbb{N}^+} \sum_{m \in l+2\mathbb{N}} (a_{l,m} | F\phi)_{L^2(\Gamma, g_u)} a_{l,m} + (b_{l,m} | F\phi)_{L^2(\Gamma, g_u)} b_{l,m}.$$

Its variance at f_u is given by

$$n^{-1} \left(\sum_{m \in 2\mathbb{N}^+} (U_m | F\phi)^2_{L^2(\Gamma, g_u)} \right. \tag{5.3}$$
$$\left. + \sum_{l \in \mathbb{N}^+} \sum_{m \in l+2\mathbb{N}} (a_{l,m} | F\phi)^2_{L^2(\Gamma, g_u)} + (b_{l,m} | F\phi)^2_{L^2(\Gamma, g_u)} \right),$$

where

$$(g | F\phi)_{L^2(\Gamma, g_u)} = \frac{2}{\pi^2} \int_0^\pi \int_{-1}^1 g(\theta, s) F\phi(\theta, s) \sqrt{1-s^2}\, ds\, d\theta. \quad ■ \tag{5.4}$$

We noted in rem. 3.19 that the estimator generated by $F\phi$ roughly corresponds to applying the filtered-backprojection algorithm and taking a weighed integral of the result. We shall therefore refer to it as the standard algorithm. It is interesting to compare the performance of the standard algorithm with that of the MVU estimator generated by $p_{\overline{T_{f_u} L^2(D)}} F\phi$.

Corollary 5.8 *The variance of the standard algorithm at f_u is given by*

$$n^{-1} \left(\sum_{m \in 2\mathbb{N}^+} (U_m | F\phi)^2_{L^2(\Gamma, g_u)} \right.$$
$$\left. + \sum_{l \in \mathbb{N}^+} \sum_{m \in l \bmod 2 + 2\mathbb{N}} (a_{l,m} | F\phi)^2_{L^2(\Gamma, g_u)} + (b_{l,m} | F\phi)^2_{L^2(\Gamma, g_u)} \right).$$

The difference in variance at f_u between the standard algorithm and the MVU estimator is

$$n^{-1} \sum_{l \in \mathbb{N}^+} \sum_{m = l \bmod, l \bmod 2+2, \ldots, l-2} (a_{l,m} | F\phi)^2_{L^2(\Gamma, g_u)} + (b_{l,m} | F\phi)^2_{L^2(\Gamma, g_u)}. \quad \blacksquare \tag{5.5}$$

We conclude this section by showing there is an important special case where the standard algorithm coincides with the MVU algorithm.

Definition 5.9 A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is said to be radial if it can be written as a function of $|x|$ alone, i.e., if $\phi(x_1) = \phi(x_2)$ for all $x_1, x_2 \in \mathbb{R}^2$ such that $|x_1| = |x_2|$.

Proposition 5.10 *Suppose $\phi \in \mathcal{S}(\mathbb{R}^2)$ is radial. Then $p_{\overline{T_{f_u} L^2(D)}} F\phi = F\phi$. Moreover, $p_{\overline{T_{f_u} L^2(D)}} F\phi$ depends only on s, i.e., $p_{\overline{T_{f_u} L^2(D)}} F\phi(\theta_1, s) = p_{\overline{T_{f_u} L^2(D)}} F\phi(\theta_2, s)$ for all θ_1, θ_2, $s \in \mathbb{R}$.*

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 $l + m \in 2\mathbb{N}^+$, the inner product $(a_{l,m} | F\phi)_{L^2(\Gamma, g_u)}$ reduces to

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

Similarly, $(b_{l,m} | F\phi)_{L^2(\Gamma, g_u)} = 0$. Eq. 5.2 thus reduces to

$$\begin{aligned} p_{\overline{T_{f_u} L^2(D)}} F\phi &= \sum_{m \in 2\mathbb{N}} (U_m | F\phi)_{L^2([-1,1], \rho_u)} U_m \\ &= F\phi, \end{aligned} \tag{5.6}$$

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 $L^2([-1, 1], \rho_u)$. $\blacksquare$

Corollary 5.11 *Suppose $\phi \in \mathcal{S}(\mathbb{R}^2)$ is radial. The estimator generated by $F\phi$ is an MVU estimator for $\dot{\phi}(f)$ at f_u in the ECT problem. Its variance at f_u is given by*

$$n^{-1} \left(||F\phi(0, \cdot)||^2_{L^2([-1,1], \rho_u)} - [\dot{\phi}(f)]^2 \right). \quad \blacksquare \tag{5.7}$$

5.2 The General Case

Remark 5.12 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 [3, I.4.6]. We shall refer to this procedure as orthonormalizing.

Proposition 5.13 *An orthonormal basis for $\overline{T_{f_0}L^2(D)}$ is obtained by orthonormalizing the set of functions $\{\frac{Tf_u}{Tf_0}g_j\}_{g_j\in B_u}$ in $L^2(\Gamma, Tf_0)$.*

Proof. Defining the multiplication operator $M_{\frac{Tf_u}{Tf_0}} : L^2(\Gamma, Tf_u) \to L^2(\Gamma, Tf_0)$ by $g \mapsto \frac{Tf_u}{Tf_0}g$, it is clear that we can decompose T_{f_0} into the composition $M_{\frac{Tf_u}{Tf_0}}T_{f_u}$. Since $M_{\frac{Tf_u}{Tf_0}}$ is continuous, the linear span of $\{\frac{Tf_u}{Tf_0}g_j\}_{g_j\in B_u}$ is dense in $\overline{T_{f_0}L^2(D)}$. Orthonormalizing this set thus gives an orthonormal basis for $\overline{T_{f_0}L^2(D)}$.

Remark 5.14 Let $\{\eta_j\}$ be a orthonormal basis for $\overline{T_{f_u}L^2(D)}$, which may be constructed according to prop. 5.13. Then $p_{\overline{T_{f_0}L^2(D)}}\psi$ and $||p_{\overline{T_{f_0}L^2(D)}}\psi||^2_{\overline{T_{f_0}L^2(D)}}$ are given by the formulas

$$p_{\overline{T_{f_0}L^2(D)}}\psi = \sum_{j=1}^{\infty}(\psi|\eta_j)_{\overline{T_{f_0}L^2(D)}}\eta_j$$

and

$$||p_{\overline{T_{f_0}L^2(D)}}\psi||^2_{L^2(\Gamma,Tf_0)} = \sum_{j=1}^{\infty}(\psi|\eta_j)^2_{\overline{T_{f_0}L^2(D)}}.$$

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 constructed a basis for the orthogonal complement of $\overline{T_{f_0}L^2(D)}^{\perp}$.

Proposition 5.15 *An orthonormal basis for $\overline{T_{f_0}L^2(D)}^{\perp}$ is obtained by orthonormalizing the set of functions B'_u in $L^2(\Gamma, Tf_0)$.*

Proof. By prop. 3.24, $\overline{T_{f_0}L^2(D)}^{\perp} = \mathcal{N}(T^*_{f_0})$. Thus, by cor. 3.14, $\overline{T_{f_0}L^2(D)}^{\perp} = \mathcal{N}(T^*_{f_u})$. An orthonormal basis for $\overline{T_{f_0}L^2(D)}^{\perp}$ is therefore obtained by orthonormalizing B'_u in $L^2(\Gamma, Tf_0)$. ■

Remark 5.16 We are now in a position to outline a numerical approach to the calculation of $p_{\overline{T_{f_0}L^2(D)}}\psi$ and $||p_{\overline{T_{f_0}L^2(D)}}\psi||^2_{L^2(\Gamma,Tf_0)}$. Orthonormal bases for $\overline{T_{f_0}L^2(D)}$ and $\overline{T_{f_0}L^2(D)}^{\perp}$ are given by props. 5.13 and 5.15. Together they form a basis for $L^2(\Gamma, Tf_0)$. 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 $L^2(\Gamma, Tf_0)$ norm of the expansion with $||\psi||^2_{L^2(\Gamma,Tf_0)}$, which can be computed by numerical integration.

6 Gaussian Functionals

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

6.1 General Gaussian Functionals

Definition 6.1 Let

$$\Phi(a;b;z) \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \frac{\Gamma(a+k)/\Gamma(a)}{\Gamma(b+k)/\Gamma(b)} \frac{z^k}{k!}$$

denote Kummer's confluent hypergeometric function [19], where, in this equation, Γ denotes the gamma function.

Proposition 6.2 *Let $\phi_{a,\sigma} \in \mathcal{S}(\mathbb{R}^2)$ be given by the Gaussian p.d.f.*

$$\phi_{a,\sigma}(x) \stackrel{\text{def}}{=} \frac{1}{2\pi\sigma^2} e^{-|x-a|^2/2\sigma^2}$$

with mean $a \in \mathbb{R}^2$ and covariance $\sigma^2 I$. Then the observation-domain representation $F\phi_{a,\sigma}$ is given by

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

Proof. $R\phi$ is given by

$$R\phi(\theta,s) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(s-a\cdot\omega(\theta))^2/2\sigma^2}$$

[4, sec. 3.5, eq. 5.3] and the Fourier transform of $R\phi$ with respect to the second variable is given by

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

[8, prop. 8.24]. Thus, by def. 3.17,

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

where $(F\phi_{a,\sigma})\hat{}$ 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 $a\cdot\omega(\theta)=0$. By symmetry, we then have

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

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

$$\int_0^\infty xe^{-ax^2}\cos(xy)\,dx = \frac{1}{2a}\Phi(1;1/2;-y^2/4a)$$

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

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

If $a\cdot\omega(\theta)\neq 0$, then the result for $a\cdot\omega(\theta)=0$ and standard results on the effect of translation on the Fourier transform (see, e.g., [8, thm. 8.22]) give

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

Eq. 6.1 now follows by applying the Kummer transformation identity $\Phi(a,b,z) = e^z\Phi(b-a;b;-z)$ [19, eq. 13.1.27] [6, eq. 6.4.7] to the last equation. ■

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

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

we can write

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

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 fig. 1, we illustrate the function χ_σ for $\sigma = 0.1$. fig. 1

Example 6.5 In fig. 2, the graph on the upper left shows the zero-mean Gaussian p.d.f. with $\sigma = 0.1$, i.e., $\phi_{0,0.1}$. The graph on the upper right shows the observation–domain representation $F\phi_{0,0.1}$ of $\phi_{0,0.1}$, obtained using eq. 6.1. The lower half of fig. 1 is similar to the upper half, except that the Gaussian p.d.f. is centered at $a = (1, 0)$ instead of at the origin. Fig. 3 is identical to fig. 2 except that $\sigma = 0.5$ instead of $\sigma = 0.1$. fig. 2 fig. 3

Remark 6.6 By inserting the result of prop. 6.2 into cor. 5.7, we can now explicitly compute the MVU estimator for Gaussian functionals at the uniform distribution along with its variance. Since the inner products in eqs. 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. eq. 5.4. It is useful to start by computing the expansion given in eq. 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 eqs. 5.2 and 5.3 are needed to achieve a given level of accuracy.

Example 6.7 We consider the estimation of the functional generated by a Gaussian p.d.f. ϕ centered at $a = (1, 0)$ with $\sigma = 0.5$. (This was illustrated at the bottom of fig. 3.) The standard estimator is generated by $F\phi_{(1,0),\sigma}$, which is shown at the top of

fig. 4. The MVU estimator at f_u is generated by $p_{\overline{T_{f_u}L^2(D)}}F\phi_{(1,0),\sigma}$, which is illustrated in the middle of fig. 4. The difference $F\phi_{(1,0),\sigma} - p_{\overline{T_{f_u}L^2(D)}}F\phi_{(1,0),\sigma}$, which backprojects to 0, is illustrated at the bottom of fig. 4. In this particular case, the variance of the standard estimator is $0.087n^{-1}$ while the variance of the MVU estimator is $0.066n^{-1}$. Thus, in this case, the variance of the standard estimator is more than 30% higher than that of the MVU estimator. fig. 4

6.2 Radial Gaussian Functionals

Remark 6.8 MVU estimators for radial Gaussian functionals at the uniform distribution along with their variance can be computed by inserting the result of prop. 6.2 into cor. 5.11. 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 Fig. 5 shows the variance of MVU estimators for a radial Gaussian functional given 10^6 observations evaluated at the uniform distribution as a function of σ. The lower curve is for the planar-imaging problem. It was obtained by numerically evaluating the formula given in thm. 4.1. The upper curve is for the ECT problem. It was obtained by numerically evaluating the formula given in cor. 5.11.

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

Proposition 6.10 *Let ϕ_σ be a zero-mean Gaussian p.d.f. with covariance matrix $\sigma^2 I$ and f_u the uniform density on D. Asymptotically as $\sigma \to 0$, the variance of the MVU estimator for $\dot{\phi}(f)$ 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 ECT problem.*

Proof. Routine calculations show that

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

and

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

It follows from thm. 4.1 that the variance of the MVU estimator for the planar-imaging problem is given by

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

which is clearly asymptotic to $1/4\pi^2\sigma^2 n$ in the limit as $\sigma \to 0$. Using thm. 4.2 and cor. 5.11, the variance of the MVU estimator for the ECT problem is given by

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

It is shown below in lem. A.2 that $||F\phi_\sigma||^2_{L^2(\Gamma,Tf_u)}$ is asymptotic to $1/8\pi^{3/2}\sigma^3$ in the limit as $\sigma \to 0$. It follows that variance of the MVU estimator for the ECT problem is asymptotic to $1/8\pi^{3/2}\sigma^3$. ■

Example 6.11 Fig. 6 shows the result of inserting the approximation

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

into the expression for the variance of the MVU estimator given by eq. 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 6.12 For radial Gaussian functionals, the MVU estimator at f_u coincides with the standard estimator. We now present an example that shows this estimator can be very suboptimal at $f_0 \neq f_u$.

Example 6.13 Let ϕ be radial Gaussian with $\sigma = 0.5$. Then the MVU estimator at f_u is generated by $F\phi$ and, by ex. 6.11, its variance at f_u is $\approx 0.087n^{-1}$. Now suppose that f_0 is highly concentrated about the point $(0, 0.63) \in D$. Then $g_0 = Tf_0$ is highly concentrated about the curve $\theta \mapsto (\theta, 0.63 \sin\theta)$ and we can approximate the integral $\int_\Gamma g(\theta, s) g_0(\theta, s)\, ds\, d\theta$ by $\frac{1}{\pi}\int_0^\pi g(\theta, 0.63 \sin\theta)\, d\theta$. Using this approximation, the variance of the estimator generated by $F\phi$ at f_0 is $\approx 0.04683n^{-1}$. From prop. 5.15, the function $a_{2,0}/||a_{2,0}||_{L^2(\Gamma, g_0)}$ is a unit vector of $\overline{T_{f_0} L^2(D)}^\perp$. The squared inner product of this function with $F\phi$ in $L^2(\Gamma, g_0)$ is ≈ 0.04594, which implies that the variance of the MVU estimator at f_0 is at most $0.00089n^{-1}$. Thus in this, admittedly extreme, example, the variance of the MVU estimator is less than the variance of the standard estimator by a factor of more than 50.

7 Extension of the Model

The model of ECT given in sec. 1.1 and analyzed in this paper is very idealized. However, the general approach can be extended to more realistic models. Of course, extending the model will introduce varying degrees of additional mathematical complications. In this section, we will briefly illustrate a relatively simple extension of the model to include the effects of photon attenuation and detector nonhomogeneity. Again, we motivate our model by the physical problem of PET.

In PET, some of the emitted photons are absorbed by the subject and thus not detected. It can be shown that the fraction of photons pairs traveling along a given line which are not absorbed is independent of the position of the positron emission which gave rise to them [20, p. 416]. That is, the number of observed photon pairs observed along a given line (θ, s) is proportional to $\gamma(\theta, s)Tf(\theta, s)$, where $\gamma(\theta, s)$ is the fraction of photon pairs traveling along (θ, s) which are not absorbed. Thus the probability density of the observations is given by

$$\Upsilon f(\theta, s) \stackrel{\text{def}}{=} \frac{\gamma(\theta, s)Tf(\theta, s)}{\int_\Gamma \gamma(\theta, s)Tf(\theta, s)\, ds\, d\theta}. \tag{7.1}$$

Note that the normalizing factor $\int_\Gamma \gamma(\theta, s)Tf(\theta, s)\, ds\, d\theta$ is the expected value of the fraction of positron emissions whose photon pairs are detected. For simplicity, we will henceforth assume that γ is positive on Γ.

The approach in this paper is based on observation-domain representations of smooth functionals as was done in prop. 3.16 for the transformation T. To get a similar result for the transformation Υ, we solve eq. 7.1 for Tf, obtaining

$$Tf(\theta, s) = \int_\Gamma \gamma(\theta, s)Tf(\theta, s)\, ds\, d\theta \frac{\Upsilon f(\theta, s)}{\gamma(\theta, s)}. \tag{7.2}$$

Since $\int_\Gamma Tf(\theta, s)\, ds\, d\theta = 1$, it follows that

$$\int_\Gamma \frac{\Upsilon f(\theta, s)}{\gamma(\theta, s)}\, ds\, d\theta = \frac{1}{\int_\Gamma \gamma(\theta, s)Tf(\theta, s)\, ds\, d\theta}.$$

Inserting this relation into eq. 7.2 gives

$$Tf(\theta, s) = \frac{\Upsilon f(\theta, s)}{\gamma(\theta, s) \int_\Gamma \frac{\Upsilon f(\theta, s)}{\gamma(\theta, s)}\, ds\, d\theta}.$$

We now write

$$\begin{aligned}\int_{\mathbb{R}^2} \phi(x) f(x)\, dx &= \frac{\pi}{2} \int_{\mathbb{R}^2} T^* I^{-1} T\phi(x) f(x)\, dx \\ &= \frac{\pi}{2} \int_{\Gamma} I^{-1} T\phi(\theta, s) T f(\theta, s)\, ds\, d\theta \\ &= \frac{\pi}{2} \left(\int_{\Gamma} \frac{\Upsilon f(\theta, s)}{\gamma(\theta, s)}\, ds\, d\theta \right)^{-1} \int_{\Gamma} \frac{I^{-1} T\phi(\theta, s)}{\gamma(\theta, s)} \Upsilon f(\theta, s)\, ds\, d\theta.\end{aligned}$$

We now see that $\frac{\pi}{2} \left(\int_{\Gamma} \frac{\Upsilon f(\theta,s)}{\gamma(\theta,s)}\, ds\, d\theta \right)^{-1} \frac{I^{-1}T\phi}{\gamma}$ is an observation-domain representation of the functional generated by ϕ for the transformation Υ except for the complication that the normalizing factor $\int_{\Gamma} \frac{\Upsilon f(\theta,s)}{\gamma(\theta,s)}\, ds\, d\theta$ depends on the unknown distribution f. There are two approaches to dealing with this. One approach is to note that $\int_{\Gamma} \frac{\Upsilon f(\theta,s)}{\gamma(\theta,s)}\, ds\, d\theta$ can generally be estimated with a high degree of certainty (an obvious estimate is $\frac{1}{n}\Sigma_{i=1}^{n} \gamma^{-1}(l_i)$) and therefore can be taken as a known constant. The other approach is to rephrase the problem to one of estimating the functional generated by $\int_{\Gamma} \frac{\Upsilon f(\theta,s)}{\gamma(\theta,s)}\, ds\, d\theta^{-1} \cdot \phi$, for which $\frac{\pi}{2} \frac{I^{-1}T\phi}{\gamma}$ is an observation-domain representation. Since the factor $\int_{\Gamma} \frac{\Upsilon f(\theta,s)}{\gamma(\theta,s)}\, ds\, d\theta$ is independent of ϕ, this will not change the ratios of the various functionals being estimated, which is generally the main point at issue. (E.g., if we think of the functionals as approximating pixels, the effect would be a scale factor applied to the image.) In either case, having obtained the observation-domain representation of the functional, the extension of our approach to this model is now straightforward. It is evident that detector nonhomogeneity effects may be incorporated into the function γ and thus be subsumed under the same mathematical model.

8 Discussion

The results in this paper give best-possible lower bounds on the variance of unbiased estimation of smooth functionals in the planar imaging and ECT problems. They can be used as a benchmark in assessing the performance of image reconstruction and quantification algorithms. Appropriately generalized, they can also be used as a design tool for assessing the performance that is achievable by new imaging devices.

One attractive feature of the approach taken here is that the resulting bounds may be applied to parametric reconstruction algorithms, regardless of how the image is parametrized. E.g., if the estimate of the image is represented by $\hat{f}(x) = \sum_{i=1}^{k} \hat{a}_i b_i(x)$, where the $\hat{a}_i$ are parameter estimates and the b_i are basis functions, one can compare the variance of the implied estimate of $\dot{\phi}(f)$, i.e., $\hat{\dot{\phi}}(f) \stackrel{\text{def}}{=} \sum_{i=1}^{k} \hat{a}_i \dot{\phi}(b_i)$, with the bound.

The numerical results given in ex. 6.7 show that, at least in some cases, the MVU estimator for $\dot{\phi}(f)$ may have significantly less variance than the estimator generated by $F\phi$. Since, as mentioned in rem. 3.19, the latter estimator corresponds roughly to the standard algorithm used in medical imaging practice, the results suggest that some improvement of the standard estimator may be possible. More extensive evaluation of the bound should help delineate the conditions under which significant improvement over the standard algorithm is possible.

It should be emphasized that the MVU estimators for the ECT problem constructed in this paper are only minimum variance at particular points $f_0 \in \mathcal{P}$. There is no guarantee that will perform well for all $f \in \mathcal{P}$. Nevertheless, they suggest an heuristic approach to constructing algorithms which will perform well at all $f \in \mathcal{P}$. The idea is that, given a sufficient number of observations, one can construct a good auxiliary estimate $\hat{f}$ of f and then apply the MVU estimator at $\hat{f}$. We are currently investigating this issue.

A 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

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

$H_s(\mathbb{R}^d)$ is termed the Sobolev space of order s. For each $k \in \mathbb{N}$, the Sobolev embedding theorem states that any function in $H_s(\mathbb{R}^d)$ for $s > k+d/2$ is k-times differentiable [8, thm. 8.54]. Now the Fourier transform of a function in $\mathcal{S}(\mathbb{R}^d)$ is in $\mathcal{S}(\mathbb{R}^d)$ [8, cor. 8.23]. Let $s \geq 0$ and $f \in \mathcal{S}(\mathbb{R}^d)$ be given. Since $\hat{f} \in \mathcal{S}(\mathbb{R}^d)$, there exists $c > 0$ such that

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

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

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

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 *Let ϕ_σ denote the zero-mean Gaussian p.d.f. with covariance matrix $\sigma^2 I$. The limiting behavior of $||F\phi_\sigma||^2_{L^2(\Gamma,Tf_0)}$ as $\sigma \to 0$ is described by*

$$\lim_{\sigma\to 0} \frac{||F\phi_\sigma||^2_{L^2(\Gamma,Tf_u)}}{\frac{2}{\pi^2}||F\phi_\sigma||^2_{L^2(G)}} = 1 \tag{A.1}$$

where

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

Proof. Starting with def. 3.17 and using eq. 6.2, we obtain

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

[9, 3.461.2], which gives eq. A.2. To prove eq. A.1, we will prove that

$$\lim_{\sigma\to 0} \frac{||F\phi_\sigma||^2_{L^2(\Gamma,Tf_u)} - \frac{2}{\pi^2}||F\phi_\sigma||^2_{L^2(G)}}{\frac{2}{\pi^2}||F\phi_\sigma||^2_{L^2(G)}} = 0.$$

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

$$\begin{aligned}
&\frac{\left|||F\phi_\sigma||^2_{L^2(\Gamma,Tf_u)} - \frac{2}{\pi^2}||F\phi_\sigma||^2_{L^2(G)}\right|}{\frac{2}{\pi^2}||F\phi_\sigma||^2_{L^2(G)}} \\
&= \frac{\int_0^\pi \int_{-1}^1 (F\phi_\sigma)^2(\theta,s)|\sqrt{1-s^2}-1|\, ds\, d\theta + \int_0^\pi \int_{|s|>1} (F\phi_\sigma)^2(\theta,s)\, ds\, d\theta}{||F\phi_\sigma||^2_{L^2(G)}}
\end{aligned}$$

$$\leq \frac{\frac{\epsilon}{2}\int_0^\pi \int_{|s|<\delta}(F\phi_\sigma)^2(\theta,s)\,ds\,d\theta}{||F\phi_\sigma||^2_{L^2(G)}} + \frac{\int_0^\pi \int_{|s|>\delta}(F\phi_\sigma)^2(\theta,s)\,ds\,d\theta}{||F\phi_\sigma||^2_{L^2(G)}}$$
$$\leq \frac{\epsilon}{2} + \frac{\int_0^\pi \int_{|s|\geq\delta}(F\phi_\sigma)^2(\theta,s)\,ds\,d\theta}{||F\phi_\sigma||^2_{L^2(G)}},$$

we see that it suffices to prove that

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

Using the expansion

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

which is valid for $x > 0$ [19, eq. 13.1.4], and eq. 6.1, there exists a constant c such that

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

It follows that

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

from which eq. A.3 easily follows. ■

Acknowledgements

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.

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] J. B. Conway. *A Course in Functional Analysis.* Springer-Verlay, New York, 1990.

[4] S. R. Deans. *The Radon Transform and Some of Its Applications.* John Wiley & Sons, New York, 1983.

[5] R. M. Dudley. *Real analysis and probability.* Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989.

[6] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi. *Higher Transcendental Functions*, volume 1. McGraw-Hill, New York, 1953.

[7] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi. *Tables of Integral Transforms*, volume 1. McGraw-Hill, New York, 1954.

[8] G. B. Folland. *Real Analysis.* John Wiley & Sons, New York, 1984.

[9] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products.* Academic Press, New York, 1980.

[10] M. G. Hahn and E. T. Quinto. Distances between measures from 1-dimensional projections as implied by continuity of the inverse Radon transform. *Z. Wahrsch. verw. Gebiete*, 70:361–380, 1985.

[11] A. Hertle. Continuity of the Radon transform and its inverse on Euclidean space. *Math. Z.*, 184:165–192, 1983.

[12] 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.

[13] K. Itô. *Encyclopedic Dictionary of Mathematics.* MIT Press, Cambridge, MA, second edition, 1987.

[14] I. M. Johnstone and B. W. Silverman. Speed of estimation in positron emission tomography and related inverse problems. *Ann. Statist.*, 18:251–280, 1990.

[15] A. Kuruc. *Probability Measure Estimation in Positron Emission Tomography Using Loss Functions Based on Sobolev Norms.* PhD thesis, Massachusetts Institute of Technology, Cambridge, CA, 1994.

[16] 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. American Mathematical Society, Providence RI, 1994.

[17] F. Natterer. *The Mathematics of Computed Tomography.* John Wiley & Sons, New York, 1986.

[18] W. Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, New York, third edition, 1976.

[19] 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.

[20] J. A. Sorenson and M. E. Phelps. *Physics in Nuclear Medicine.* Grune & Stratton, Inc., New York, second edition, 1987.

[21] E. M. Stein. *Singular Integrals and Differentiability Properties of Integrals.* Princeton University Press, Princeton, NJ, 1970.

[22] H. Strasser. *Mathematical Theory of Statistics.* Walter de Gruyter, New York, 1985.

[23] F. Treves. *Topological Vector Spaces, Distributions, and Kernels.* Academic Press, New York, 1967.

[24] Y. Vardi, L. A. Shepp, and L. Kaufman. A statistical model for positron emission tomography. *J. Amer. Stat. Assn.*, 80:8–20, 1985.

Index

A

$a_{l,m}$, 39

adjoint, 15, 18

B

B'_u, 39

B_u, 39

$b_{l,m}$, 39

backprojection, 15

C

$\check{g}$, 13

C^∞, 19

$c_{l,m}$, 39

closed, 30

closure, 29

D

D, 5

$\dot{\phi}$, 8

D^n, 10

dense, 18

drop, 13

E

E_f, 10

ECT, 5

F

f, 5

$\hat{\phi}_n$, 10

f_0, 10

f_u, 24

F, 27

Φ, 45

$\phi_{a,\sigma}$, 45

ϕ_σ, 48

filtered backprojection, 21

Fourier transform, 19

G

Γ, 7

Γ^n, 9

G, 13

Γ, 17

g_u, 24

γ, 51

Gaussian functionals, 45

H

$\hat{f}$, 19

Hilbert space, 16

I

ι_{L^q,L^p}, 17

I^{-1}, 20

i.i.d., 22

indicator function, 8
inner product, 16

L

l_i, 7
L^1, 14
L^∞, 15
L^2, 17
$L^2(X, f_0)$, 17
lift, 13
linear estimator
    for ECT, 24
    for planar imaging, 22

M

MSE, 10
MVU, 11
MVU linear estimator, 22

N

$||f||_{L^2(X,f_0)}$, 17
$\mathcal{N}$, 28
$\mathbb{N}$, 38
$\mathbb{N}^+$, 38

O

1_S, 8
$\overline{L}$, 29
orthogonal, 29
orthogonal complement, 29
orthonormal, 38
orthonormal basis, 38

P

$\mathcal{P}$, 9
π, 13
$(f|g)_{L^2(X,f_0)}$, 17
π_{L^q,L^p}, 17
$L^\perp$, 29
p_L, 30
p.d.f., 5
parameter space, 10
parametric, 11
PET, 6
pixel, 11
planar imaging, 10

R

$\mathbb{R}^d$, 5
$\tilde{R}$, 14
R, 14
R^*, 15
ρ_u, 38
radial, 42
Radon transform, 5
ramp filter, 20
Riesz potential operator, 20
ROI, 5

S

S^1, 13

$\mathcal{S}(\mathbb{R}^d)$, 19

$\mathcal{S}(G)$, 20

singular value decomposition, 11

smooth functionals, 5

standard algorithm, 41

statistical experiment, 9

T

$\tilde{g}$, 13

T, 17

T^*, 18

T_{f_0}, 25

U

U_m, 38

Υ, 51

unbiased estimator, 10

W

Ω, 9

ω, 13

X

x_i, 7

X, 14

Z

$\mathbb{Z}$, 38

Figures

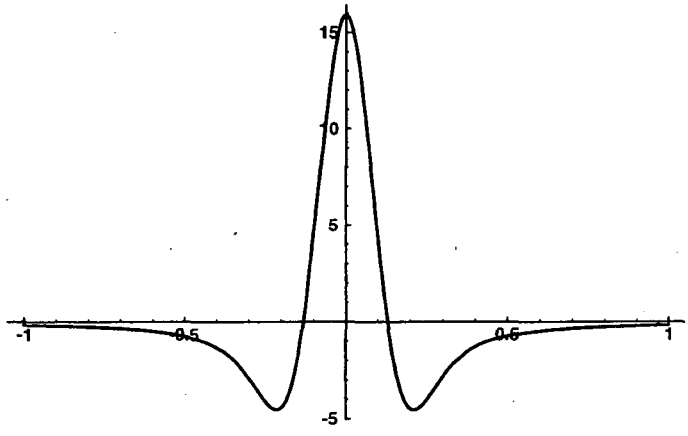

Figure 1: The function χ_σ evaluated for $\sigma = 0.1$

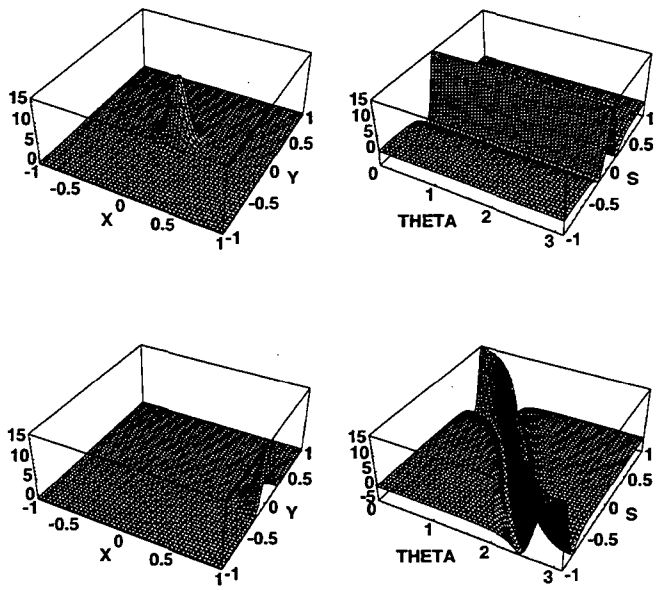

Figure 2: Gaussian functionals and their observation-domain representations. Gaussian p.d.f.s are shown on the left and their respective observation-domain representations are shown on the right. The upper pair is for a radial Gaussian p.d.f., while the lower pair is for a Gaussian p.d.f. centered at $(1, 0)$. For both pairs, $\sigma = 0.1$.

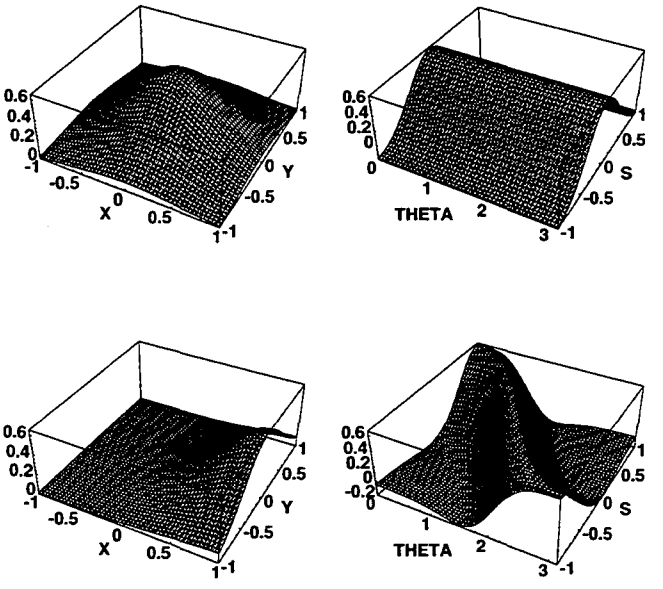

Figure 3: Gaussian functionals and their observation-domain representations. Everything is as in fig. 1, except that $\sigma = 0.5$.

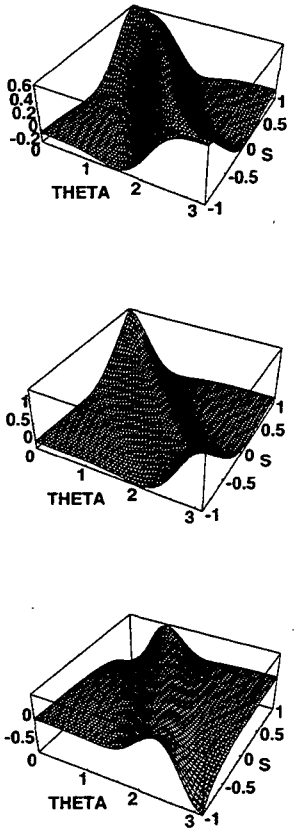

Figure 4: The functions which generate the standand and MVU estimators are shown in the top and middle, respectively. Their difference is shown at the bottom.

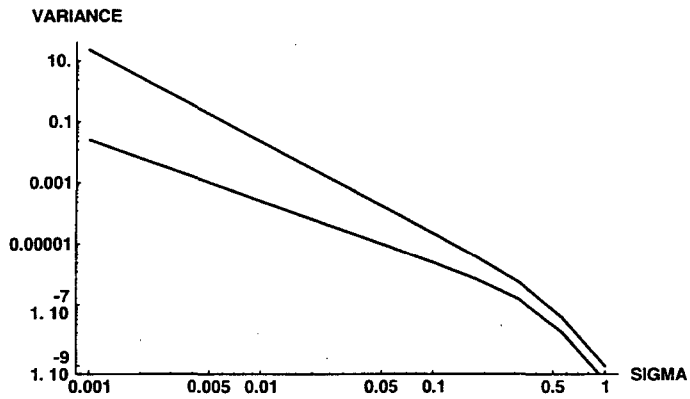

Figure 5: Variance of MVU estimator for radial Gaussian functionals at the uniform distribution with 10^6 photon pairs (lower curve is for planar imaging, upper curve is for ECT).

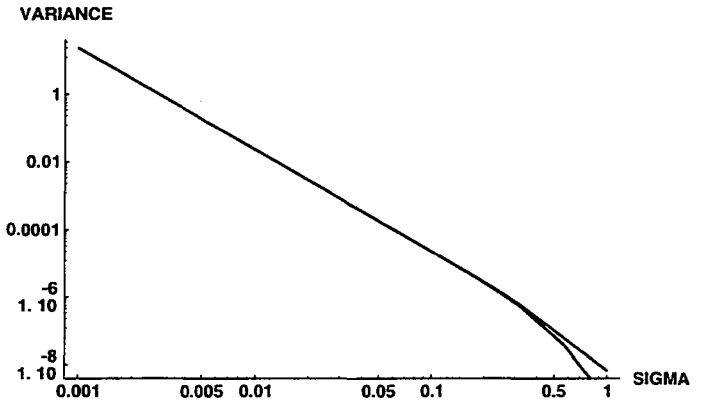

Figure 6: Approximation to the variance of MVU estimator for radial Gaussian functionals at the uniform distribution with 10^6 photon pairs (lower curve is the true value, upper curve is the approximation).

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720