Recovering Convex Edges of an Image from Noisy Tomographic Data

A. Goldenshluger *
Haifa University

V. Spokoiny†
Weierstass Institute

This version: July 10, 2005

Abstract

We consider the problem of recovering edges of an image from noisy tomographic data. The original image is assumed to have a discontinuity jump (edge) along the boundary of a compact convex set. The Radon transform of the image is observed with noise, and the problem is to estimate the edge. We develop an estimation procedure which is based on recovering support function of the edge. It is shown that the proposed estimator is nearly optimal in order in a minimax sense. Numerical examples illustrate reasonable practical behavior of the estimation procedure.

Short Title: Recovering edges of tomographic images

Keywords: Radon transform, optimal rates of convergence, support function, edge detection, minimax estimation

2000 AMS Subject Classification (Primary): 62G20; 62C20; 94A08

1 Introduction

In this paper we address the problem of recovering edges of an image from noisy tomographic data. The original image is modeled by function f defined on the unit disc $B^2(o,1) \subset \mathbb{R}^2$. Assume that f is smooth apart from a discontinuity jump along a smooth curve. The problem of edge recovery from tomographic data is to estimate the discontinuity curve from noisy measurements of line integrals of f.

*Department of Statistics, Haifa University, Haifa 31905, Israel, e-mail: goldensh@stat.haifa.ac.il
Supported by the Israel Science Foundation grant No. 300/04.
†Weierstrass Institute of Applied Analysis and Stochastics, Mohrenstrasse 39, D-10117 Berlin, Germany, e-mail: spokoiny@wias-berlin.de

From statistical perspective the problem of image reconstruction from tomographic data was studied in Vardi, Shepp, and Kaufman (1985), Johnstone and Silverman (1990), Johnstone and Silverman (1991), Korostelev and Tsybakov (1993), Bickel and Ritov (1995), Ritov (1998) and Cavalier and Koo (2002), among many others. In these papers typically the reconstruction of the whole function f or a linear functional thereof is considered. In many applications, however, one is interested in reconstruction of certain geometric features of the image such as edges, boundaries, shapes etc. In particular, the problem of edge detection arises in numerous imaging applications. For example, images with discontinuities along edges are ubiquitous in medical applications; here edges bring important information about body regions with different levels of metabolic activity. Thus edge recovery is an important step in processing tomographic images.

The problem of edge recovery in tomographic images is extensively studied in applied mathematics and image processing literature [see, e.g., Faridani, Ritman and Smith (1992) and Srinivasa et al. (1992), for representative publications]. This literature however concentrates either on mathematical properties of reconstruction formulas, or on algorithmic and implementation aspects. Typically the presence of observation noise is ignored, and statistical properties of reconstruction procedures are not analyzed. Recently Hero et al. (1999) and Ye, Bresler and Moulin (2000) discussed information–theoretic aspects in parametric modeling of edge shapes using tomographic data. In these papers boundary estimation procedures are proposed, and fundamental bounds on the performance of parametric estimation are established via the Cramer–Rao lower bounds.

In the present paper we address information–theoretic issues arising in nonparametric boundary estimation from tomographic data. Although various methods and proposals are widely used in practice, theoretical limitations in the problem of nonparametric edge estimation from the Radon data are yet to be understood. What is the best attainable accuracy in recovering edges from noisy observations of projections? Which methods can achieve this optimal performance? Our goal is to provide a theoretical perspective on these questions and to develop easily implemented nearly–optimal algorithm for edge recovery in tomographic images. We assume that the edge can be represented as the boundary of a convex set, and propose a method for estimating support function of this set. Then the boundary is recovered as the envelope of the estimated supporting lines. We analyze theoretical properties of the proposed estimation scheme and show that it is nearly optimal in order in the sense of the rates of convergence. Assuming that the Radon transform of an image is observed with additive Gaussian white noise of variance σ^2, we prove that convex edges can be estimated with the pointwise risk of the order $\sigma^{4/5}$ up to a logarithmic factor. Our lower bound on the estimation accuracy demonstrates that this rate cannot be essentially improved. Numerical examples illustrate reasonable practical behavior of the proposed estimator.

We would like to emphasize that the proposed recovery procedure does not involve inversion of the Radon transform. We establish a close connection between the problem of edge recovery and the boundary fragment model of Korostelev and Tsybakov (1993) [see also Härdle, Park and Tsybakov (1995), Wang (1998)]. Exploiting this connection, we show that the problem of recovering a convex edge from tomographic data can be approached via detection of a cusp curve in the Radon domain from noisy observations. This cusp curve is determined by the support function of the edge in the original image. Thus our reconstruction procedure operates in the Radon domain where noisy observations are directly available.

Recently Candés and Donoho (2002) considered the problem of recovering images with edges from the Radon data contaminated by the Gaussian white noise. It was shown there that if the image f is twice continuously differentiable except for a discontinuity along a twice differentiable smooth curve, then the best achievable rate of convergence in estimating f in $\mathbb{L}_2$–norm is $\sigma^{2/5}$ up to a logarithmic in σ^{-1} factor. As we show in this paper, in the same model the convex edge can be estimated with the rate $\sigma^{4/5}$ (ignoring multiplicative $\ln(1/\sigma)$–factors). Our technique can be extended to other models with indirect observations; see, e.g., Goldenshluger and Spokoiny (2004) and Goldenshluger and Zeevi (2004).

The rest of the paper is organized as follows. In Section 2 we formulate the problem of edge recovery from noisy tomographic data, introduce definitions and discuss some preliminary results. Section 3 describes construction of our estimation procedure, and presents main theoretical results. In Section 4 we present numerical examples; Section 5 contains concluding remarks. Proofs are given in Appendix.

2 Problem formulation and preliminaries

The observation model. Let f be a square–integrable function on the unit disc $B^2(o,1) \subset \mathbb{R}^2$. The Radon transform $\mathcal{R} : \mathbb{L}_2(B^2(o,1)) \to \mathbb{L}_2([0,1] \times [0,2\pi))$ of f is defined by integration of f along the lines $l_{s\varphi}$ parametrized by angle $\varphi \in [0,2\pi)$ and distance to the origin $s \in [0,1]$:

$$(\mathcal{R}f)(s,\varphi) = \int_{l_{s\varphi}} f(x,y)\, dt,$$

here dt is the Lebesgue measure on $l_{s\varphi}$. Consider the the following white noise model

$$Y(ds,d\varphi) = (\mathcal{R}f)(s,\varphi) ds\, d\varphi + \sigma W(ds,d\varphi), \tag{1}$$

where $W(s,\varphi)$ denotes the Wiener sheet, and σ is the noise level. The model (1) specifically means that for any function $v \in \mathbb{L}_2([0,1]\times[0,2\pi))$ the integral $\iint v(s,\varphi)(\mathcal{R}f)(s,\varphi)ds\,d\varphi$ can be observed with Gaussian error having zero mean and variance $\sigma^2 \iint v^2(s,\varphi)ds\,d\varphi$. Assume that f is smooth apart from a discontinuity jump along a smooth curve which is the boundary ∂G of a convex set $G \subset B^2(o,1)$; for simplicity, we suppose that $o \in int(G)$. The goal is to estimate the boundary of G.

Support function of convex sets. It is well known that there is a one-to-one correspondence between convex sets and their *support functions*. Therefore our approach to estimating the edge ∂G from observations (1) will be based on pointwise recovering the support function of G. Below we collect some preliminary results and definitions that will be repeatedly used in what follows. These results can be found, e.g., in Schneider (1993), Gardner (1995), and Groemer (1996).

If G is a nonempty compact convex set in $\mathbb{R}^2$, the *support function* g_G of G is defined by $g_G(u) = g(u) := \max\{x^T u : \ x \in G\}$ for $u \in S^1 := \{(\cos\varphi, \sin\varphi) : \varphi \in [0,2\pi)\}$. Every compact convex set is uniquely determined by its support function:

$$G = \{x \in \mathbb{R}^2 : x^T u \le g(u), \ \ u \in S^1\}.$$

If $u \in S^1$ then $H_u := \{x : x^T u = g(u)\}$ is the *supporting line* to G with outward normal u. Support function $g(u)$ gives the *signed* distance from the origin $o = (0,0)$ to H_u. For simplicity we assume that $o \in G$ so that $g(u)$ gives the actual distance from the origin o to H_u. In the planar case it is natural to view the support function as function of $\varphi \in [0,2\pi)$ and write $g(\varphi)$ rather than $g(u)$ or $g(u(\varphi))$. Basic properties of support functions are summarized as follows.

(I) The support function $g(\varphi)$ is 2π–periodic. If $G \subset B^2(o,1)$ then

$$|g(\varphi_1) - g(\varphi_2)| \le |\varphi_1 - \varphi_2|.$$

Thus g is absolutely continuous and $|g'(\varphi)| \le 1$ almost everywhere on $[0,2\pi)$.

(II) A twice differentiable 2π–periodic function $g(\varphi)$ is the support function of some convex domain if $g(\varphi) + g''(\varphi) > 0$ for all $\varphi \in [0,2\pi)$.

(III) The position vector $q(\varphi)$ of the closed convex curve ∂G is given by

$$q(\varphi) = g'(\varphi)u'(\varphi) + g(\varphi)u(\varphi),$$

where as before $u(\varphi) = (\cos\varphi, \sin\varphi)$. The radius of curvature $\rho(\varphi)$ of ∂G at the point $q(\varphi)$ is given by $\rho(\varphi) = g(\varphi) + g''(\varphi)$ and the center of curvature $e(\varphi)$ is

$$e(\varphi) = g'(\varphi)u'(\varphi) - g''(\varphi)u(\varphi).$$

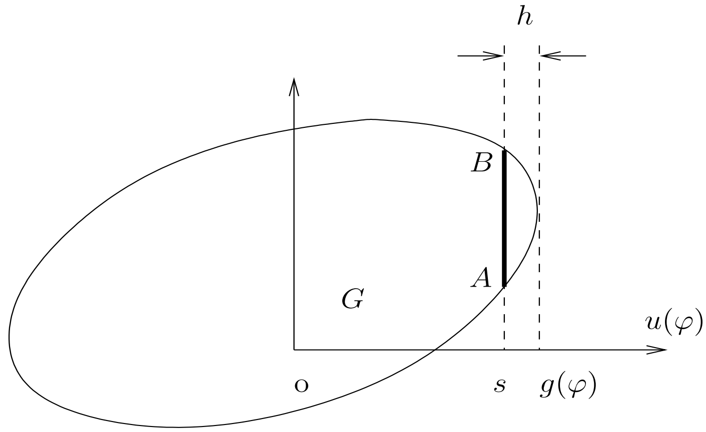


Figure 1: An illustration of the Radon transform behavior near the edge for $f(x) = \tilde{f}(x)\mathbf{1}_G(x)$, $\tilde{f}(x) \geq c > 0$.

Properties of the Radon transform. It turns out that estimating support function of the edge is rather natural when noisy Radon observations are available. According to general results on singularities of the Radon transform of discontinuous functions [Quinto (1993), Ramm and Zaslavsky (1993)], the Radon transform $(\mathcal{R}f)(s, \varphi)$ is smooth at every point (s, φ) if and only if the line $l_{s\varphi}$ with coordinates (s, φ) *is not tangent* to the discontinuity curve of f. If f is discontinuous along the boundary ∂G of a convex set G with support function g then supporting lines have coordinates $(g(\varphi), \varphi)$, and they are tangent to the discontinuity curve of f. Therefore $\mathcal{R}(s, \varphi)$ has a singularity along the curve $\{(s, \varphi) : s = g(\varphi), \varphi \in [0, 2\pi)\}$. The type of this singularity is essentially determined by geometrical properties of the boundary ∂G. In particular, if ∂G has everywhere positive curvature then the Radon transform $\mathcal{R}f$ has the one–sided singularity cusp of the order $1/2$ along the curve $\{(s, \varphi) : s = g(\varphi), \varphi \in [0, 2\pi)\}$, i.e. there exist $0 < L_0 \leq L_1$ such that for some $h_0 > 0$

$$L_0 h^{1/2} \leq |(\mathcal{R}f)(g(\varphi), \varphi) - (\mathcal{R}f)(g(\varphi) - h, \varphi)| \leq L_1 h^{1/2}, \quad 0 < h \leq h_0, \quad \forall \varphi. \tag{2}$$

This can be explained using simple geometrical argument which is illustrated in Figure 1 for $f = \tilde{f}\mathbf{1}_G$, $\tilde{f} \geq c > 0$. In this case the Radon transform $\mathcal{R}f$ is supported on the set $\{(s, \varphi) : 0 \leq s \leq g(\varphi), \varphi \in [0, 2\pi)\}$, and $(\mathcal{R}f)(g(\varphi) - h, \varphi)$ equals to the "weighted" length of the chord AB. Since ∂G has non–zero curvature and $\tilde{f} \geq c > 0$, this "weighted" length is at least of the order $O(h^{1/2})$ for sufficiently small h; hence (2) follows. The Radon transform $\mathcal{R}f$ is smooth apart from the set $\{(s, \varphi) : s = g(\varphi), \ \varphi \in [0, 2\pi)\}$; for general results on local smoothness of $\mathcal{R}f$ we refer to Quinto (1993). However, for our purposes it will be sufficient to assume the Lipschitz condition for every fixed $\varphi \in [0, 2\pi)$:

$$|(\mathcal{R}f)(\tau, \varphi) - (\mathcal{R}f)(t, \varphi)| \leq L_2 |\tau - t|, \quad \forall \tau, t \in [0, g(\varphi) - h_0] \cup [g(\varphi), 1]. \tag{3}$$

The above considerations show that the problem of recovering a convex edge from observations (1) can be viewed as the problem of estimating the cusp curve in the Radon domain. This is similar to the boundary fragment model of Korostelev and Tsybakov (1993); see also Härdle, Park and Tsybakov (1995) and Wang (1998).

In the rest of the paper we assume that the underlying function f belongs to some class of functions f with edges.

Functional class. We say that function f on $B^2(o,1)$ belongs to the class $\mathcal{F}$ if it can be represented as $f = f_0 + a\mathbf{1}_G$, where

(A) f_0 is supported on $B^2(o,1)$ and satisfy the Lipschitz condition, $a \in \mathbb{R}$, $|a| > 0$ is a fixed constant, and $\mathbf{1}_G$ is the indicator function of a convex set $G \subset B^2(o,1)$, $o \in int(G)$;

(B) the convex set G has smooth boundary with everywhere non-zero curvature and support function g which is twice continuously differentiable and satisfies

$$0 < r \le g(\varphi) + g''(\varphi) \le R < \infty, \quad \forall \varphi \in [0, 2\pi). \tag{4}$$

The collection of convex sets satisfying (B) will be designated $\mathcal{G}$.

Several remarks on the above definition are in order. First, (A) along with the assumption of non–zero boundary curvature in (B) implies that the Radon transform $\mathcal{R}f$ obeys (2) and (3). In particular, for some $h_0 > 0$,

$$(\mathcal{R}f)(s,\varphi) = \sqrt{2\rho(\varphi)}a(g(\varphi) - s)_+^{1/2} + Q(s,\varphi), \;\; g(\varphi) - h_0 \le s \le g(\varphi), \tag{5}$$

where $\rho(\varphi)$ is the radius of curvature of ∂G at $q(\varphi)$ [see (III)], and $Q(s,\varphi)$ is a smooth function; see, e.g., Gelfand, Graev and Vilenkin (1966, §1.7). Inequality (4) states the lower and upper bounds on the radius of curvature $\rho(\varphi) = g(\varphi) + g''(\varphi)$ of the boundary [see (III)]. Thus (5) along with (4) implies that the left inequality in (2) is valid with $L_0 = \sqrt{2r}|a|$. In what follows we always assume that $R \gg r$ so that the class $\mathcal{G}$ is rich enough. Note that when $r = R$ the class $\mathcal{G}$ contains only discs of the radius r. The lower bound in (4) implies that G is the *r–smooth set* [see, e.g., Groemer (1996, p. 19)]. We recall that a set G is called *r–smooth* if it can be written as $G = \tilde{G} + rB^2(o,1)$ for some convex set $\tilde{G}$ and $r > 0$. In other words, a convex set G with support function g is r–smooth if $g(\cdot) - r$ is the support function of a convex set. Observe that L_i, $i = 0,1,2$ in (2) and (3) are determined entirely by parameters of the functional class $\mathcal{F}$.

3 Estimation procedure and main results

Our approach to estimating the convex edge is based on pointwise recovery of its support function. As mentioned in the previous section, the Radon transform $\mathcal{R}f$ has a cusp-type singularity along the curve given by the support function of the edge. We will use *a probe functional* in order to detect the location of this singularity.

The probe functional. We focus on estimating the support function g of the edge at a single given point $\theta \in [0, 2\pi)$. Let $\delta > 0$, $h > 0$, and $I_\delta := [\theta - \delta, \theta + \delta]$. For $t \in [\delta, 1 - \delta]$, and $b \in [-1, 1]$ we let $w_{t,b}(\varphi) := t + b(\varphi - \theta)$ and define

$$\ell_{\delta,h}[t, b] := \int_{I_\delta} \int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} (\mathcal{R}f)(s, \varphi) ds d\varphi - \int_{I_\delta} \int_{w_{t,b}(\varphi)}^{w_{t,b}(\varphi)+h} (\mathcal{R}f)(s, \varphi) ds d\varphi. \tag{6}$$

The next statement establishes detection properties of *the probe functional* $\ell_{\delta,h}[t, b]$.

Lemma 1 *Suppose that h and δ are sufficiently small and*

$$h \geq (16L_1/L_0)^{2/3} R\delta^2, \tag{7}$$

where constants L_0, L_1 and R are given in (2) and (4) respectively.

(i) Then

$$|\ell_{\delta,h}[g(\theta), g'(\theta)]| \geq \frac{2}{3} L_0 h^{3/2} \delta. \tag{8}$$

(ii) Let $\varkappa > 6$. Then

$$\sup_{t:|t-g(\theta)|>\varkappa h} \; \sup_{|b|\leq 1} |\ell_{\delta,h}[t, b]| \leq C_* \varkappa^{-1/2} h^{3/2} \delta, \tag{9}$$

where C_* is a positive constant that may depend on L_i, $i = 0, 1, 2,$ and R only.

The lemma shows that the localization accuracy of the probe functional $\ell_{\delta,h}[t, b]$ is $\varkappa h$: if distance between t and the target value $g(\theta)$ (measured as a multiple of h) grows by $\varkappa$, the absolute values of $\ell_{\delta,h}[t, b]$ decreases by $\varkappa^{-1/2}$. The important feature of the probe functional is the scaling restriction (7) on the horizontal and vertical size of the probe functional template. As it will be shown, the scaling $h \asymp \delta^2$ allows "maximal smoothing" along the angles while preserving "good" localization properties in the vertical direction. It is interesting to note that the similar scaling law underlies construction of the curvelet frames used for recovering functions with singularities along smooth curves [Candés and Donoho (2002)].

Estimation procedure. We define the estimator $\hat{g}(\theta)$ of the cusp curve at point $\theta \in [0, 2\pi)$ as follows. For fixed $t \in [\delta, 1-\delta]$ and $|b| \le 1$ let

$$\hat{\ell}_{\delta,h}[t,b] := \int_{I_\delta} \int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} Y(ds, d\varphi) - \int_{I_\delta} \int_{w_{t,b}(\varphi)}^{w_{t,b}(\varphi)+h} Y(ds, d\varphi) \; ; \tag{10}$$

$\hat{\ell}_{\delta,h}[t,b]$ is the estimate of $\ell_{\delta,h}[t,b]$ based on observations (1). Define

$$(\hat{t}, \hat{b}) = \arg \max_{(t,b) \in [\delta, 1-\delta] \times [-1,1]} |\hat{\ell}_{\delta,h}[t,b]|, \tag{11}$$

and let

$$\hat{g}(\theta) := \hat{t}. \tag{12}$$

Bounds on the risk. The main results of this paper are given in the following theorems.

Theorem 1 *Let $\hat{g}(\theta)$ be given by (10), (11) and (12) with*

$$\delta = C_1^* \Big\{ \sigma \sqrt{\ln \frac{1}{\sigma}} \Big\}^{2/5}, \quad h = C_2^* \delta^2 \tag{13}$$

for some positive constants C_1^* and C_2^*. Then there exists a constant $C_3^* < \infty$ depending on L_i, $i = 0, 1, 2$ and R only such that

$$\sup_{f \in \mathcal{F}} \Big\{ \mathbb{E} |\hat{g}(\theta) - g(\theta)|^2 \Big\}^{1/2} \le C_3^* \sigma^{4/5} \Big(\ln \frac{1}{\sigma} \Big)^{2/5} .$$

for all sufficiently small σ.

Theorem 2 *Let $\tilde{g}(\theta)$ be an arbitrary estimator of $g(\theta)$ based on observations (1). Then for sufficiently small σ*

$$\sup_{f \in \mathcal{F}} \Big\{ \mathbb{E} |\tilde{g}(\theta) - g(\theta)|^2 \Big\}^{1/2} \ge C_4^* \sigma^{4/5} \Big(\ln \frac{1}{\sigma} \Big)^{-2/5} ,$$

where C_4^* depends on r and R.

These results show that our estimator $\hat{g}(\theta)$ is nearly optimal in order within a logarithmic in σ^{-1} factor.

Based on the pointwise estimates of the edge support function we define the estimator of the set G as follows

$$\hat{G} = \{(x,y) \in B^2(o,1) : x \cos\varphi + y \sin\varphi \le \hat{g}(\varphi), \;\; \forall \varphi \in [0, 2\pi)\}, \tag{14}$$

where $\hat{g}(\varphi)$ is given by (10) and (11). The estimate of the boundary ∂G is given by (14) with the inequality sign replaced by equality. Note that $\hat{G}$ is a convex set by construction.

Therefore global accuracy of $\hat{G}$ may be measured using metrics for classes of convex sets. In particular, global distances between two convex sets G_1 and G_2 in $\mathbb{R}^2$ with support functions g_1 and g_2 can be defined by

$$d_p(G_1, G_2) := \Big\{ \frac{1}{2\pi} \int_0^{2\pi} |g_1(\varphi) - g_2(\varphi)|^p d\varphi \Big\}^{1/p}, \quad p \in [1, \infty]$$

with d_∞ being the well–known Hausdorff distance [see, e.g., Groemer (1996)]. The next statement establishes an upper bound on the accuracy of $\hat{G}$.

Theorem 3 *Let $\hat{G}$ be the estimate of G defined in (14) and assume that (13) is valid for some constants C_1^* and C_2^*. Then*

$$\sup_{f \in \mathcal{F}} \Big\{ \mathbb{E}[d_p^p(\hat{G}, G)] \Big\}^{1/p} \le C_5^* \sigma^{4/5} \Big(\ln \frac{1}{\sigma} \Big)^{2/5}, \quad p \in [1, \infty]. \tag{15}$$

In the case $p = \infty$ the left hand side of (15) is interpreted as $\sup_{f \in \mathcal{F}} \mathbb{E} d_\infty(\hat{G}, G)$.

4 Numerical examples

We conducted a small numerical experiment in order to illustrate practical potential of the proposed estimation scheme. Although the theoretical properties have been investigated for the idealized continuous white noise model, the estimator can be easily implemented for more realistic discrete observations model.

The original image used in our experiments is displayed in Figure 2(a). It is given by the function that equals 1 inside the ellipse G with center $(0.1, -0.1)$ and semi-axes $a = 0.64$ and $b = 0.47$, and 0.4 outside G. Thus f has a discontinuity jump of size 0.6 along the boundary of the ellipse; support function of G is depicted in Figure 2(b). In our experiments the Radon transform of the original image is observed with noise at the points of the regular grid with step size 0.01 on $[0, 2\pi] \times [0, 1]$. We assume that the noise is zero mean Gaussian and consider the low, medium and high noise level conditions when the noise standard deviation σ equal to 0.05, 0.1 and 0.3 respectively. For instance, the Radon transform observations with added Gaussian noise of standard deviation $\sigma = 0.05$ is shown in Figure 3(a). As it was indicated in Section 2, the cusp curve visible in Figure 3(a) corresponds to the support function of the ellipse in Figure 2.

In our implementation for any fixed angle φ, we compute the value of the probe functional $\ell_{h,\delta}[t, b]$ for all (t, b) from the discrete set of 200×20 regular grid points on $[0, 1] \times [-1, 1]$. The pair (t, b) corresponding to the maximum of the probe functional is selected, and its t–component is taken as the estimate of the support function at point φ.

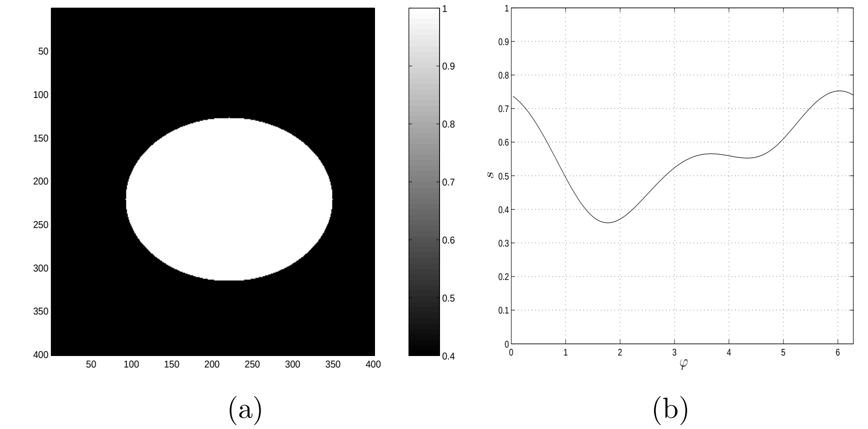


Figure 2: (a) The original image; (b) the support function of the ellipse.

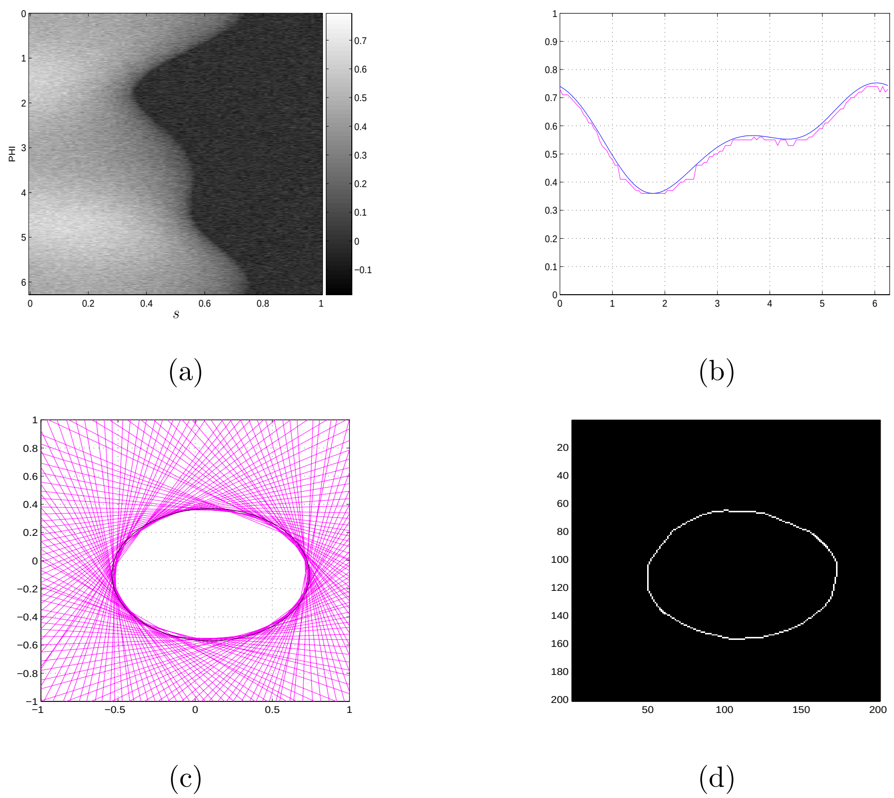


Figure 3: Edge recovery for the low noise level ($\sigma = 0.05$): (a) The noisy observations in the Radon domain. (b) The support function estimate. (c) The "true" edge (solid line) along with the estimated supporting lines. (d) The extracted estimate of the edge.

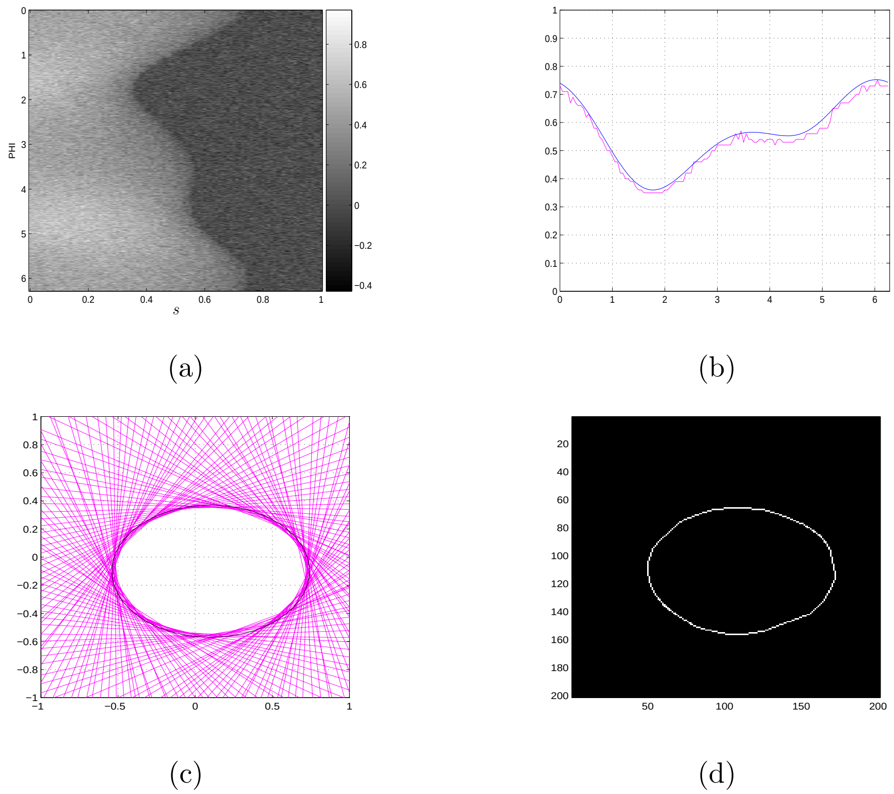
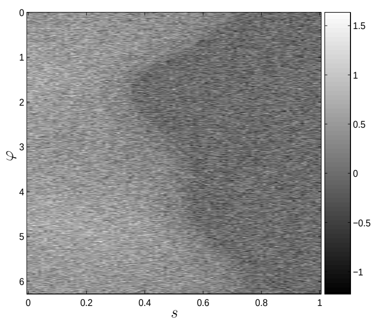


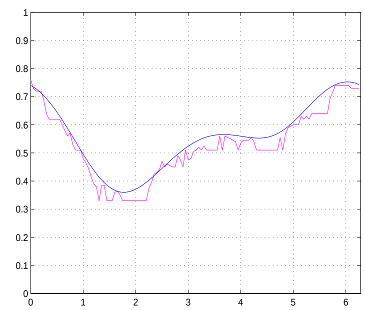
Figure 4: Edge recovery for the medium noise level ($\sigma = 0.1$): (a) The noisy observations in the Radon domain. (b) The support function estimate. (c) The "true" edge (solid line) along with the estimated supporting lines. (d) The extracted estimate of the edge.

In the numerical examples below the bandwidths h and δ were selected to achieve good visual appearance of the estimated edge. Because the data are available on the regular grid, we specify the bandwidths h and δ in the grid step size units; for example, $h = 5$ means that that the actual bandwidth is $5 \times 0.01 = 0.05$.

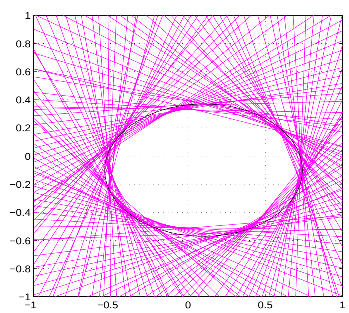
Figure 3 displays the results obtained for the case of low noise level conditions, $\sigma = 0.05$. Here the values $h = 5$ and $\delta = 7$ were selected. The panel (a) shows noisy observations in the Radon domain; (b) presents the estimate of the support function along with the "true" curve. The reconstructed set can be seen in Figure 3(c) as the inner envelope of the estimated supporting lines; the original set is also presented (solid line). Finally, panel (d) displays the extracted boundary. The similar graphs are presented in Figure 4 and 5 for $\sigma = 0.1$ and $\sigma = 0.3$ respectively. In the case of the medium noise level we selected $h = 6$ and $\delta = 8$, while in the case of high noise level $h = 9$ and $\delta = 12$. The numerical results demonstrate reasonable practical behavior of the proposed estimation scheme.



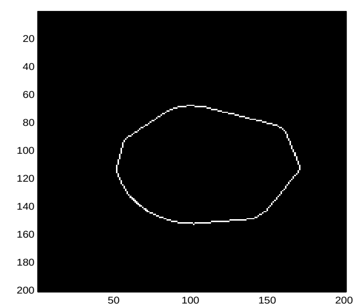
(a)



(b)



(c)



(d)

Figure 5: Edge recovery for the high noise level ($\sigma = 0.3$): (a) The noisy observations in the Radon domain. (b) The support function estimate. (c) The "true" edge (solid line) along with the estimated supporting lines. (d) The extracted estimate of the edge.

5 Concluding remarks

1. Our approach to edge recovery from tomographic data exploits duality between convex sets and their support functions. Using this duality relationship we reduce the inverse ill–posed problem of edge recovery from tomographic data to estimation of the cusp curve from direct noisy observations in the Radon domain. We note that many recently developed practical procedures for image denoising can be used in order to estimate the cusp curve in the Radon domain. In particular, the fully adaptive AWS algorithm of Polzehl and Spokoiny (2000) can be applied for this purpose.

2. It is interesting to compare our results with the results on estimation of convex boundaries for models with direct observations. Korostelev and Tsybakov (1994), Korostelev, Simar and Tsybakov (1995) and Mammen and Tsybakov (1995) study this problem under various assumptions. In particular, Korostelev and Tsybakov (1994) show that if we are given a sample of n independent data points, uniformly distributed over a convex planar region G with smooth boundary, then the best achievable accuracy in estimating ∂G in Hausdorff metric is $n^{-2/3}$. With usual calibration $n^{-1/2} = \sigma$, this corresponds to the rate $\sigma^{4/3}$ for the white noise model. Our results indicate that the convergence rate slows down to $\sigma^{4/5}$ when noisy observations in the Radon transform are available.

3. The minimax rates of convergence derived here depend crucially on the assumption that the boundary has everywhere positive curvature. This assumption guarantees that the Radon transform has a singularity of the order 1/2 along the curve given by the support function of the edge. The points of zero curvature on the boundary correspond to sharper cusps in the Radon domain. However, one cannot improve accuracy of estimation in these particular directions because the set of points where the curvature vanishes has zero Lebesgue measure.

4. If instead of the class of convex boundaries with positive curvature we consider the class of all convex boundaries, the minimax rates of convergence will change. For instance, if G is a convex polygon then the Radon transform will have singularity of the order 1. In other words, the first partial derivative of the Radon transform with respect to the distance variable will have a jump along the curve determined by the support function of the edge. In this case a sensible estimation procedure could be based on the search for the change curve in the first partial derivative of the Radon transform. Under these circumstances the minimax rates of convergence are slower than $\sigma^{4/5}$ and depend on smoothness of the Radon transform away from the change curve.

5. Although we considered functions with a single edge along the boundary of a convex

set, our technique can be extended to more general images comprised of several convex domains with different intensities. Such images are usually serve as phantoms in numerical studies, see, e.g., Vardi, Shepp, and Kaufman (1985). In particular, the proposed procedure can be applied to images with several convex edges having well–separated support functions. If the boundaries have everywhere non–zero curvature then the problem is reduced to estimating cusp curves of the order 1/2 in the Radon domain. This can be pursued by the method developed in this paper.

Appendix

In the proofs $c_1, c_2, \ldots$ stand for positive constants whose values can be different on different occasions. These constants may depend only on parameters L_i, $i = 0, 1, 2$, R and r characterizing the functional class $\mathcal{F}$. In what follows $\mathrm{meas}(\cdot)$ stands for the Lebesgue measure of a set on the real line.

Proof of Lemma 1

Proof (i). Let $p_\theta(\varphi) := g(\theta) + g'(\theta)(\varphi - \theta)$ for $\varphi \in I_\delta$, and $\Delta(s,\varphi) := (\mathcal{R}f)(s,\varphi) - (\mathcal{R}f)(g(\varphi),\varphi)$. Then

$$\begin{aligned}
\ell_{\delta,h}[g(\theta), g'(\theta)] &= \int_{I_\delta}\int_{p_\theta(\varphi)-h}^{p_\theta(\varphi)} (\mathcal{R}f)(s,\varphi)ds d\varphi - \int_{I_\delta}\int_{p_\theta(\varphi)}^{p_\theta(\varphi)+h} (\mathcal{R}f)(s,\varphi)dsd\varphi \\
&= \int_{I_\delta}\int_{p_\theta(\varphi)-h}^{p_\theta(\varphi)} \Delta(s,\varphi)dsd\varphi \;-\; \int_{I_\delta}\int_{p_\theta(\varphi)}^{p_\theta(\varphi)+h} \Delta(s,\varphi)dsd\varphi \\
&=: J_1 - J_2.
\end{aligned}$$

We have

$$\begin{aligned}
J_1 &= \int_{I_\delta}\int_{g(\varphi)-h}^{g(\varphi)} \Delta(s,\varphi)dsd\varphi - \int_{I_\delta}\int_{A_1} \Delta(s,\varphi)dsd\varphi \\
&=: J_{11} - J_{12},
\end{aligned}$$

where $A_1 := [g(\varphi) - h, g(\varphi)] \triangle [p_\theta(\varphi) - h, p_\theta(\varphi)]$, and $\triangle$ denotes the symmetric difference. By (2)

$$|J_{11}| \;\geq\; \int_{I_\delta}\int_{g(\varphi)-h}^{g(\varphi)} L_0|g(\varphi) - s|^{1/2} dsd\varphi = \frac{4}{3}L_0 h^{3/2}\delta.$$

Because $\sup_{\varphi\in I_\delta}|g(\varphi) - p_\theta(\varphi)| \leq R\delta^2$, $\mathrm{meas}(A_1) \leq 2R\delta^2$ and hence

$$|J_{12}| \leq 2\int_{I_\delta}\int_0^{R\delta^2} L_1 s^{1/2} dsd\varphi \leq \frac{4}{3}L_1\int_{I_\delta}(R\delta^2)^{3/2} \leq \frac{8}{3}L_1 R^{3/2}\delta^4.$$

Combining two last inequalities we obtain

$$|J_1| \geq \frac{4}{3} L_0 h^{3/2} \delta - \frac{8}{3} L_1 R^{3/2} \delta^4.$$

Note that the difference on the RHS is positive in view of (7).

Now we bound J_2 from above. Similarly, we write

$$J_2 = \int_{I_\delta} \int_{g(\varphi)}^{g(\varphi)+h} \Delta(s,\varphi) ds d\varphi - \int_{I_\delta} \int_{A_2} \Delta(s,\varphi) ds d\varphi,$$

where $A_2 = [g(\varphi), g(\varphi)+h] \triangle [p_\theta(\varphi), p_\theta(\varphi)+h]$. It follows from (3) that the absolute value of the first term on the RHS is less than $\frac{1}{2} L_2 h^2 \delta$, while the absolute value of the second term does not exceed $\frac{8}{3} L_1 R^{3/2} \delta^4$ by the same argument as in bounding J_{12}. Combining these inequalities and taking into account (7) we come to (8).

(ii). Introduce the following notation: $S^* := \{(s,\varphi) : s = g(\varphi),\ \varphi \in I_\delta\}$,

$$\begin{aligned} S^+ &:= \{(s,\varphi) : s > g(\varphi),\ \varphi \in I_\delta\}, \quad S^- := \{(s,\varphi) : s < g(\varphi),\ \varphi \in I_\delta\} \\ T_{t,b} &:= \{(s,\varphi) : w_{t,b}(\varphi) - h \leq s \leq w_{t,b}(\varphi) + h,\ \ \varphi \in I_\delta\}. \end{aligned}$$

We will prove the statement of the lemma considering different subsets of the set $\{(t,b) : |t - g(\theta)| > \varkappa h,\ |b| \leq 1\}$.

1. First assume that $t - g(\theta) > \varkappa h$ and $T_{t,b} \subset S^+$, i.e. *the template* $T_{t,b}$ of the probe functional $\ell_{\delta,h}[t,b]$ lies entirely above the cusp curve. Condition $T_{t,b} \subset S^+$ imposes restrictions on the slope b of the template. In particular, simple argument shows that if

$$|b - g'(\theta)| \leq \frac{1}{\delta}[(\varkappa - 1)h - R\delta^2] =: c_0 \varkappa h \delta^{-1} \tag{16}$$

then $T_{t,b} \subset S^+$. Under these conditions we have

$$\begin{aligned} \ell_{\delta,h}[t,b] &= \int_{I_\delta} \int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} (\mathcal{R}f)(s,\varphi) d\tau d\varphi - \int_{I_\delta} \int_{w_{t,b}(\varphi)}^{w_{t,b}(\varphi)+h} (\mathcal{R}f)(s,\varphi) ds d\varphi \\ &= \int_{I_\delta} \int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} [(\mathcal{R}f)(s,\varphi) - (\mathcal{R}f)(w_{t,b}(\varphi),\varphi)] ds d\varphi \\ &\quad - \int_{I_\delta} \int_{w_{t,b}(\varphi)}^{w_{t,b}(\varphi)+h} [(\mathcal{R}f)(s,\varphi) - (\mathcal{R}f)(w_{t,b}(\varphi),\varphi)] ds d\varphi. \end{aligned}$$

Applying (3) to the both integrals on the RHS we obtain

$$|\ell_{\delta,h}[t,b]| \leq 2 \int_{I_\delta} \int_{w_{t,b}(\varphi)}^{w_{t,b}(\varphi)+h} L_2 |s - w_{t,b}(\varphi)| ds d\varphi = 2 L_2 h^2 \delta. \tag{17}$$

2. Now let $g(\theta) - t > \varkappa h$ and $T_{t,b} \subset S^-$; here the template $T_{t,b}$ lies entirely under the cusp curve $\{(s,\varphi) : s = g(\varphi),\ \varphi \in I_\delta\}$. Condition (16) along with $g(\theta) - t > \varkappa h$ guarantees

that $T_{t,b} \subset S^-$. In view of (5) we have

$$\begin{aligned}
\ell_{\delta,h}[t,b] &= \int_{I_\delta} L(\varphi) \int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} (g(\varphi)-s)^{1/2} ds d\varphi \\
&\quad - \int_{I_\delta} L(\varphi) \int_{w_{t,b}(\varphi)}^{w_{t,b}(\varphi)+h} (g(\varphi)-s)^{1/2} ds d\varphi - \ell^Q_{\delta,h}[t,b] \\
&=: K - \ell^Q_{h,\delta}[t,b],
\end{aligned}$$

where for brevity we denoted $L(\varphi) = \sqrt{2\rho(\varphi)}a$, and $\ell^Q_{\delta,h}[t,b]$ is the probe functional $\ell_{\delta,h}[t,b]$ applied to the function $Q(s,\varphi)$ [see (5)]. Using the same reasoning as in the proof of (17) we obtain

$$|\ell^Q_{\delta,h}[t,b]| \le 2L_2 h^2 \delta. \tag{18}$$

Further, integrating with respect to s we have

$$\begin{aligned}
K &= \frac{2}{3}\int_{I_\delta} L(\varphi)\Big\{(g(\varphi)-w_{t,b}(\varphi)+h)^{3/2} - 2(g(\varphi)-w_{t,b}(\varphi))^{3/2} \\
&\qquad + (g(\varphi)-w_{t,b}(\varphi)-h)^{3/2}\Big\} d\varphi.
\end{aligned} \tag{19}$$

Note that

$$\begin{aligned}
g(\varphi) - w_{t,b}(\varphi) &\ge g(\theta) - t + (g'(\theta)-b)(\varphi-\theta) - R\delta^2 \\
&\ge g(\theta) - t - |g'(\theta)-b|\delta - R\delta^2 > h
\end{aligned}$$

provided that $g(\theta) - t > \varkappa h$, and (16) is valid. Therefore expanding the integrand in (19) in the Taylor series we obtain for some $\eta \in (0,1)$

$$\begin{aligned}
|K| &\le L_1 h^2 \int_{I_\delta} \frac{d\varphi}{(g(\varphi)-w_{t,b}(\varphi))^{1/2}} + \frac{3}{8} L_1 h^4 \int_{I_\delta} \frac{d\varphi}{(g(\varphi)-w_{t,b}(\varphi)-\eta h)^{5/2}} \\
&=: K_1 + K_2.
\end{aligned}$$

In order to bound from above K_1 and K_2 under $g(\theta) - t > \varkappa h$ and (16), we consider separately the cases where $|g'(\theta) - b| \le 2h$, and $2h \le |g'(\theta)-b| \le c_0 \varkappa h \delta^{-1}$.

If $|g'(\theta) - b| \le 2h$, then

$$\begin{aligned}
K_1 &\le L_1 h^2 \int_{-\delta}^{\delta} [g(\theta) - t + (g'(\theta)-b)\varphi - R\delta^2]^{-1/2} d\varphi \\
&\le 2\delta L_1 h^2 [\varkappa h - 2h\delta - R\delta^2]^{-1/2} \le c_1 L_1 h^{3/2} \delta \varkappa^{-1/2}
\end{aligned}$$

and

$$\begin{aligned}
K_2 &\le \frac{3}{8} L_1 h^4 \int_{-\delta}^{\delta} [g(\theta) - t + (g'(\theta)-t)\varphi - \eta h - R\delta^2]^{-5/2} d\varphi \\
&\le \frac{3}{4} L_1 h^4 \delta [\varkappa h - 2h\delta - \eta h - R\delta^2]^{-5/2} \le c_2 L_1 h^{3/2} \delta \varkappa^{-5/2}
\end{aligned}$$

so that

$$|K| \le c_3 L_1 h^{3/2}\delta\varkappa^{-1/2}. \tag{20}$$

If $2h \le |b-g'(\theta)| \le c_0\varkappa h\delta^{-1}$ then integrating w.r.t. φ we obtain

$$\begin{aligned}
K_1 &\le L_1 h^2 \int_{-\delta}^{\delta} [g(\theta)-t+(g'(\theta)-b)\varphi - h - R\delta^2]^{-1/2} d\varphi \\
&\le \frac{L_1 h^2}{|g'(\theta)-b|}\Big\{[g(\theta)-t+|g'(\theta)-b|\delta - h - R\delta^2]^{1/2} \\
&\qquad\qquad - [g(\theta)-t-|g'(\theta)-b|\delta - h - R\delta^2]^{1/2}\Big\}.
\end{aligned}$$

Noting that function $x \mapsto [(a+\delta x)^{1/2}-(a-\delta x)^{1/2}]/x$, $a>0$, $\delta>0$ is monotone increasing in $(0, a/\delta)$ we conclude that

$$K_1 \le \frac{L_1 h^2 \delta}{(\varkappa-1)h - R\delta^2}(2\varkappa h)^{1/2} \le c_4 L_1 h^{3/2}\delta\varkappa^{-1/2}.$$

Similar argument shows that $K_2 \le c_5 L_1 h^{3/2}\delta\varkappa^{-5/2}$; hence (20) holds also for $2h \le |b - g'(\theta)| \le c_0\varkappa h\delta^{-1}$. Combining these inequalities with (18) and (17) we finally get

$$\max_{|t-g(\theta)|>\varkappa h} \; \max_{|g'(\theta)-b|\le c_0\varkappa h\delta^{-1}} |\ell_{\delta,h}[t,b]| \le c_6 L_1 h^{3/2}\delta\varkappa^{-1/2}.$$

3. To complete the proof of the lemma it remains to establish (9) for (t,b) such that

$$|t-g(\theta)| > \varkappa h, \quad \text{and} \quad |b-g'(\theta)| > c_0\varkappa h\delta^{-1}. \tag{21}$$

This corresponds to the case where $T_{t,b}\cap S_* \ne \emptyset$, i.e. the template of the probe functional intersects the cusp curve. For definiteness we assume here that $g(\theta)-t > \varkappa h$, and define $\Phi^+ := \{\varphi \in I_\delta : (s,\varphi) \in T_{t,b}\cap S^+\}$, $\Phi^- := \{\varphi \in I_\delta : (s,\varphi)\in T_{t,b}\cap S^-\}$, $\Phi^* = \Phi^+\cap\Phi^-$. In words, Φ^+ and Φ^- denote the set of those φ for which the template $T_{t,b}$ lies above and below the cusp curve respectively.

Note that $I_\delta = \Phi^+\cup\Phi^-$ so that $\ell_{\delta,h}[t,b]$ can be written as a sum of three integrals over the sets $\Phi^-\backslash\Phi^*$, Φ^*, and $\Phi^+\backslash\Phi^*$:

$$\begin{aligned}
\ell_{\delta,h}[t,b] &= \int_{\Phi^-\backslash\Phi^*}\Big\{\int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} (\mathcal{R}f)(s,\varphi)ds - \int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} (\mathcal{R}f)(s,\varphi)ds\Big\}d\varphi \\
&\quad + \int_{\Phi^*}\Big\{\int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} (\mathcal{R}f)(s,\varphi)ds - \int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} (\mathcal{R}f)(s,\varphi)ds\Big\}d\varphi \\
&\quad + \int_{\Phi^+\backslash\Phi^*}\Big\{\int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} (\mathcal{R}f)(s,\varphi)ds - \int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} (\mathcal{R}f)(s,\varphi)ds\Big\}d\varphi \\
&=: J_1 + J_2 + J_3.
\end{aligned}$$

The absolute values of J_1 and J_3 are bounded from above exactly as $|\ell_{\delta,h}[t,b]|$ in the cases $T_{t,b} \subset S^-$ and $T_{t,b} \subset S^+$ respectively. In particular, $|J_3| \le c_7 L_1 h^2 \delta$ and $|J_1| \le c_8 L_1 h^{3/2} \delta \varkappa^{-1/2}$.

To bound $|J_2|$ we argue first that

$$\mathrm{meas}(\Phi^*) \le c_9 \delta \varkappa^{-1}. \tag{22}$$

Indeed, let φ_1 and φ_2 be defined by equations

$$g(\theta) + g'(\theta)(\varphi_1 - \theta) - h = t + b(\varphi_1 - \theta), \quad g(\theta) + g'(\theta)(\varphi_2 - \theta) + h = t + b(\varphi_2 - \theta).$$

Then, in view of (21),

$$|\varphi_1 - \varphi_2| \le \frac{2h}{|b - g'(\theta)|} \le \frac{2\delta}{c_0 \varkappa}. \tag{23}$$

Let $\varphi_- \in I_\delta$ be the minimal solution to the equation $|g(\varphi) - t - b(\varphi - \theta)| = h$. In words, φ_- is the left endpoint of the set $T_{t,b} \cap S^*$. Using the definition of φ_1 we have

$$\begin{aligned} h &= |g(\varphi_-) - t - b(\varphi_- - \theta)| \\ &= |g(\varphi_-) - t - b(\varphi_1 - \theta) + b(\varphi_1 - \varphi_-)| \\ &= |g(\varphi_-) - g(\theta) - g'(\theta)(\varphi_1 - \theta) + h + b(\varphi_1 - \varphi_-)| \\ &= |g(\varphi_-) - g(\theta) - g'(\theta)(\varphi_- - \theta) + h + (g'(\theta) - b)(\varphi_- - \varphi_1)|; \end{aligned}$$

hence

$$h \ge |g'(\theta) - b|\,|\varphi_- - \varphi_1| - R\delta^2 - h \quad \Rightarrow \quad |\varphi_- - \varphi_1| \le \frac{2h + R\delta^2}{|g'(\theta) - b|} \le c_{10}\delta\varkappa^{-1}.$$

Similarly, the same inequality is established for $|\varphi_+ - \varphi_2|$, where φ_+ stands for the right endpoint of the set $T_{t,b} \cap S^*$. Using these inequalities and (23) we get (22). Therefore for sufficiently small h and δ

$$|J_2| \le 4L_1 \int_{\Phi_*} \int_{w_{t,b}(\varphi) - h}^{g(\varphi)} (g(\varphi) - s)^{1/2} ds d\varphi \le c_{11} L_1 h^{3/2} \delta \varkappa^{-1}.$$

This completes the proof of the lemma. ■

Proof of Theorem 1

Let $\varkappa > 6$ be a fixed constant, large enough so that $C_* \varkappa^{-1/2} \le L_0/3$, where C_* appears on the RHS of (9). Because $\mathrm{supp}(f) \subseteq B^2(o,1)$, we can write

$$\begin{aligned} \mathbb{E}|\hat{g}(\theta) - g(\theta)|^2 &\le (\varkappa h)^2 + \mathbb{E}\Big[|\hat{g}(\theta) - g(\theta)|^2 \mathbf{1}\{|\hat{g}(\theta) - g(\theta)| > \varkappa h\} \\ &\le (\varkappa h)^2 + \mathbb{P}\Big\{|\hat{g}(\theta) - g(\theta)| > \varkappa h\Big\}. \end{aligned} \tag{24}$$

Our goal is to bound the probability on the RHS of (24).

Define $u := (t, b)$, $u_* = (g(\theta), g'(\theta))$, $U_0 := [\delta, 1-\delta] \times [-1, 1]$, and $U := \{(t, b) \in U_0 : |t - g(\theta)| > \varkappa h, |b| \leq 1\}$. For brevity throughout the proof we write $\ell_{\delta,h}[u]$ for $\ell_{\delta,h}[t, b]$. We have for small enough σ

$$
\begin{aligned}
\mathbb{P}\{|\hat{g}(\theta) - g(\theta)| > \varkappa h\} &\leq \mathbb{P}\Big\{\sup_{u\in U} |\hat{\ell}_{\delta,h}[u]| \geq |\hat{\ell}_{\delta,h}[u_*]|\Big\} \\
&\leq \mathbb{P}\Big\{\sup_{u\in U} |\hat{\ell}_{\delta,h}[u] - \ell_{\delta,h}[u]| + \sup_{u\in U} |\ell_{\delta,h}[u]| \geq |\hat{\ell}_{\delta,h}[u_*]|\Big\} \\
&\overset{(a)}{\leq} \mathbb{P}\Big\{\sup_{u\in U_0} |\hat{\ell}_{\delta,h}[u] - \ell_{\delta,h}[u]| + c_1\varkappa^{-1/2}h^{3/2}\delta \geq |\hat{\ell}_{\delta,h}[u_*]|\Big\} \\
&\leq \mathbb{P}\Big\{2\sup_{u\in U_0} |\hat{\ell}_{\delta,h}[u] - \ell_{\delta,h}[u]| + c_1\varkappa^{-1/2}h^{3/2}\delta \geq |\ell_{\delta,h}[u_*]|\Big\} \\
&\overset{(b)}{\leq} \mathbb{P}\Big\{\sup_{u\in U_0} |\hat{\ell}_{\delta,h}[u] - \ell_{\delta,h}[u]| \geq c_2 h^{3/2}\delta\Big\}\ , \qquad (25)
\end{aligned}
$$

where (a) follows from Lemma 1 (ii), and (b) is a consequence of Lemma 1 (i) and the fact that $c_1\varkappa^{-1/2}h^{3/2}\delta < \frac{2}{3}L_0h^{3/2}\delta$ if $\varkappa$ is sufficiently large. Thus it remains to bound from above the probability $\mathbb{P}\{\sup_{u\in U_0} |X_u| \geq c_2h^{3/2}\delta\}$, where

$$
X_u = X_{(t,b)} := \sigma\Big\{\int_{I_\delta}\int_{w_{t,b}(\varphi)-h}^{w_{t,b}(\varphi)} W(ds, d\varphi) - \int_{I_\delta}\int_{w_{t,b}(\varphi)}^{w_{t,b}(\varphi)+h} W(ds, d\varphi)\Big\} \qquad (26)
$$

is the zero mean Gaussian process indexed by $u = (t, b) \in U_0 = [\delta, 1-\delta] \times [-1, 1]$. For this purpose we apply the exponential inequality of Talagrand (1994) for general Gaussian processes [see also van der Vaart and Wellner (1996)]. First we note that

$$
\sup_{u\in U_0} \mathbb{E}|X_u|^2 \leq c_3\sigma^2 h\delta. \qquad (27)
$$

Let $u_i = (t_i, b_i)$, and for a fixed $\varphi \in I_\delta$ we define $M_i^+ = [w_{t_i,b_i}(\varphi) - h, w_{t_i,b_i}(\varphi)]$, $M_i^- = [w_{t_i,b_i}(\varphi), w_{t_i,b_i}(\varphi) + h]$, $i = 1, 2$. Then

$$
\begin{aligned}
X_{u_1} - X_{u_2} &= \sigma\Big\{\int_{I_\delta}\int_{M_1^+\setminus M_2^+} W(ds, d\varphi) - \int_{I_\delta}\int_{M_2^+\setminus M_1^+} W(ds, d\varphi)\Big\} \\
&\quad - \sigma\Big\{\int_{I_\delta}\int_{M_1^-\setminus M_2^-} W(ds, d\varphi) - \int_{I_\delta}\int_{M_2^-\setminus M_1^-} W(ds, d\varphi)\Big\}
\end{aligned}
$$

and

$$
\mathbb{E}|X_{u_1} - X_{u_2}|^2 \leq c_4\sigma^2[\delta|t_1 - t_2| + \delta^2|b_1 - b_2|].
$$

This implies that one needs no more that $N(\epsilon) = c_5\sigma^2\delta^3\epsilon^{-4}$ balls of radius ϵ in the natural semimetric in order to cover the index set $U_0 = [\delta, 1-\delta] \times [-1, 1]$. Now we apply the exponential inequality of Proposition A.2.7 from van der Vaart and Wellner (1996) [with

$K \sim \sigma^{1/2}\delta^{3/4}$, $V = 4$, $\epsilon_0 \sim \sigma\sqrt{h\delta}$ and $\lambda \sim h^{3/2}\delta$]. With this choice of the parameters all conditions of the proposition are fulfilled so that we obtain

$$\begin{aligned} \mathbb{P}\{\sup_{u\in U_0} |X_u| \geq c_2 h^{3/2}\delta\} &\leq \Big(\frac{c_6\sigma^{1/2}\delta^{3/4}h^{3/2}\delta}{\sigma^2 h\delta}\Big)^4 \exp\Big\{-\frac{c_2h^3\delta^2}{\sigma^2 h\delta}\Big\} \\ &= c_7\frac{\delta^3h^2}{\sigma^6}\exp\{-c_2h^2\delta\sigma^{-2}\} \leq c_8(\varkappa h)^2 \end{aligned}$$

where the last inequality follows by appropriate choice of C_1^* and C_2^* in (13). ■

Proof of Theorem 2

In the proof below $c_1, c_2, \ldots$ stand for constants that may depend on r and R only. We assume that $R \gg r$ so that class of sets $\mathcal{G}$ is sufficiently rich (e.g., $r = R$ implies that $\mathcal{G}$ contains only discs of radius $r = R$).

Without loss of generality we assume $\theta = \pi/2$, and let G_0 be the disc $B^2(o,\rho)$ of radius $r < \rho < R$, centered at the origin $o = (0,0)$. The support function of G_0 is $g_{G_0}(\varphi) = g_0(\varphi) = \rho$, $\forall\varphi$. For some $h > 0$ define $\tilde{G}_1 = G_0 \cap B^2(A,R)$, where $B^2(A,R)$ is the disc of radius R centered at $A = (0, -R-h+\rho)$; see Figure 6. By construction, $g_{\tilde{G}_1}(\pi/2) + h = g_{G_0}(\pi/2)$; note however, that $\tilde{G}_1 \notin \mathcal{G}$, because $\partial\tilde{G}_1$ is not differentiable at the points E and F of intersection of $\partial B^2(A,R)$ and ∂G_0. We define $G_1 \in \mathcal{G}$ by replacing the boundary of $\tilde{G}_1$ by circular arcs of the radius r in vicinity of the singularity points E and F as shown in Figure 6; this is always possible if $\mathcal{G}$ is rich enough. For such a set $G_1 \subset G_0$ we have $g_{G_1}(\pi/2) = g_{G_0}(\pi/2) - h$.

Now assume that $f_0(x) = \mathbf{1}_{G_0}(x)$ so that it has a discontinuity jump along the boundary of G_0, and let $f_1(x) = \mathbf{1}_{G_1}(x)$. Assume that we have observations (1). The Kullback–Leibler divergence between the probability measures $\mathbb{P}_0$ and $\mathbb{P}_1$ corresponding to the processes

$$Y_i(ds, d\varphi) = (\mathcal{R}f_i)(s,\varphi)dsd\varphi + \sigma W(ds,d\varphi), \quad i = 0,1$$

is given by

$$\begin{aligned} \mathcal{K}(\mathbb{P}_0,\mathbb{P}_1) &= \mathbb{E}_0 \ln\frac{d\mathbb{P}_0}{d\mathbb{P}_1}(Y_0) \\ &= \frac{1}{2\sigma^2}\int_0^{2\pi}\int_0^1 |(\mathcal{R}(f_0-f_1))(s,\varphi)|^2 dsd\varphi \,. \end{aligned} \tag{28}$$

To bound $\mathcal{K}(\mathbb{P}_0,\mathbb{P}_1)$ we use the idea similar to that in Candés and Donoho (2002); namely, we show that $G_0 \backslash G_1$ contains an ellipse with certain ratio of semi–axes and is contained in another ellipse with the same ratio of the semi-axes. Then the Radon transform of $\mathbf{1}_{G_0\backslash G_1}$ can be bounded in terms of the Radon transform of an appropriate ellipse. Indeed,

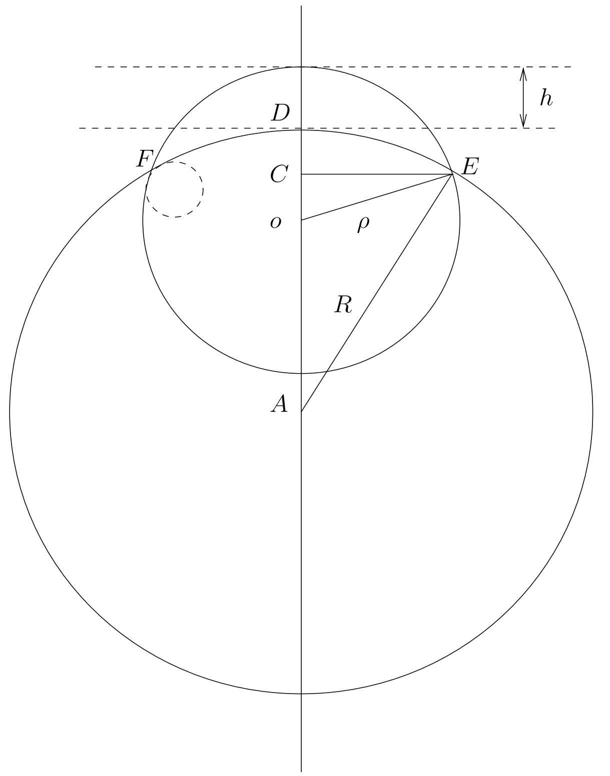


Figure 6: Illustration of the proof of Theorem 2

G_1 belongs to the set $G_0\backslash\{(x,y) : \rho-h \le y \le \rho\}$. Therefore $G_0\backslash G_1$ contains an ellipse with semi-axes of the size c_2h and $c_3\sqrt{h}$. On the other hand, $G_0\backslash G_1$ is also contained in some ellipse with semi-axes c_4h and $c_5\sqrt{h}$. In order to show this it suffices to verify that $|CD| \le c_6h$ for some constant c_6 [see Figure 6]. Setting $\alpha = \angle oAE$ we note that $|CD| = R-R\cos\alpha$; considering the triangle AoE we find that $\rho^2 = R^2+(R-\rho+h)^2-2R(R-\rho+h)\cos\alpha$ and after straightforward calculations we obtain $|CD| = R-R\cos\alpha = \frac{1}{2}h(2\rho-h)/(R-\rho+h)$. Thus for small enough h, $|CD| \le c_6h$ and this, in turn, implies that $|CE| \le c_7\sqrt{h}$. Therefore the set $G_0\backslash G_1$ can be covered by an ellipse with semi-axes of the size c_2h and $c_3\sqrt{h}$ as claimed.

Recall that the Radon transform of the indicator of ellipse $\mathcal{E}(a,b)$ with semi-axes a and b is given by

$$(\mathcal{R}\mathbf{1}_{\mathcal{E}(a,b)})(s,\varphi) = \frac{ab}{p}\Big(1-\frac{s^2}{p^2}\Big)_+^{1/2}, \quad p^2 := a^2\cos^2\varphi + b^2\sin^2\varphi. \tag{29}$$

Further, if V is the orthogonal matrix representing the planar rotation by θ, $e \in \mathbb{R}^2$, and $u(\varphi) = (\cos\varphi, \sin\varphi)$ then

$$\{\mathcal{R}\mathbf{1}_{\mathcal{E}(a,b)}(Vx-e)\}[s,u(\varphi)] = \{\mathcal{R}\mathbf{1}_{\mathcal{E}(a,b)}\}[s-e^TVu(\varphi), Vu(\varphi)].$$

Using this property and (29) we bound the integral on the RHS of (28) as follows

$$\begin{aligned}
\|\mathcal{R}(f_0 - f_1)\|_2^2 &\leq c_8 \int_0^{2\pi} \int_0^p \frac{h^3}{p^2}\Big(1 - \frac{s^2}{p^2}\Big)_+ ds d\varphi \\
&= c_9 h^3 \int_0^{2\pi} \frac{d\varphi}{\sqrt{h^2\cos^2\varphi + h\sin^2\varphi}} \\
&\leq c_{10} h^{5/2} \ln \frac{1}{h} \ .
\end{aligned}$$

Thus if we choose $h = c_{11}[\sigma^2(\ln \frac{1}{\sigma})^{-1}]^{2/5}$ the Kullback–Leibler divergence will be of the order of $O(1)$; this completes the proof of the lower bound. ■

Proof of Theorem 3

Noticing that

$$d_p(\hat{G}, G) = \Big\{\frac{1}{2\pi}\int_0^{2\pi} |\hat{g}(\varphi) - g(\varphi)|^p d\varphi\Big\}^{1/p} \ .$$

we conclude that in the case of $p \in [1, \infty)$ the statement of the theorem follows immediately from Theorem 1 by integrating the risk upper bound over $\theta \in [0, 2\pi)$. Therefore only the case $p = \infty$ should be considered.

First we note that the statement of Lemma 1 holds for all $\theta \in [0, 2\pi)$, in particular,

$$\begin{aligned}
|\ell_{\delta,h}[g(\theta), g'(\theta)]| &\geq \frac{2}{3} L_0 h^{3/2}\delta, \quad \forall\theta \in [0, 2\pi) \\
\sup_{t:|t-g(\theta)|>\varkappa h} \ \sup_{|b|\leq 1} |\ell_{\delta,h}[t, b]| &\leq C\varkappa^{-1/2} h^{3/2}\delta, \quad \forall\theta \in [0, 2\pi).
\end{aligned}$$

Further, the argument similar to (25) leads to

$$\mathbb{P}\Big\{\sup_{\theta\in[0,2\pi)} |\hat{g}(\theta) - g(\theta)| > \varkappa h\Big\} \leq \mathbb{P}\Big\{\sup_{u\in U_0} |X_u| \geq c_1 h^{3/2}\delta\Big\},$$

where the Gaussian process $\{X_u\}$ is again given by (26), but now the index set is different: $u = (\theta, t, b)$, $U_0 = [0, 2\pi) \times [\delta, 1-\delta] \times [-1, 1]$. We again apply the general exponential inequality in order to bound this probability. To this end, we first observe that (27) is valid. Setting $u_i = (\theta_i, t_i, b_i)$, $i = 1, 2$ we find from straightforward geometrical considerations that

$$\mathbb{E}|X_{u_1} - X_{u_2}|^2 \leq c_2\sigma^2\Big[\delta^2|b_1 - b_2| + \delta|t_1 - t_2| + |\theta_1 - \theta_2|\,|t_1 - t_2| + h|\theta_1 - \theta_2|\Big]$$

so that one needs no more than $N(\epsilon) = c_3\sigma^2 h\delta^3\epsilon^{-6}$ balls of radius ϵ in the natural semimetric in order to cover the index set $[0, 2\pi) \times [\delta, 1-\delta] \times [-1, 1]$. Then by Proposition A.2.7 from van der Vaart and Wellner (1996) [with $K \sim \sigma^{1/3}h^{1/6}\delta^{1/2}$, $V = 6$, $\epsilon_0 \sim \sigma\sqrt{h\delta}$ and $\lambda \sim h^{3/2}\delta$]

$$\mathbb{P}\Big\{\sup_{u\in U_0} |X_u| \geq c_1 h^{3/2}\delta\Big\} \leq c_3\frac{h^4\delta^3}{\sigma^{10}} \exp\{-c_4 h^2\delta\sigma^{-2}\} \leq c_5(\varkappa h)^2,$$

where the last inequality follows by appropriate choice of C_1^* and C_2^* in (13). ■

References

BICKEL, P. and RITOV, Y. (1995). Estimating a linear functional of a PET image. *IEEE Trans. Med. Imaging* **14**, 81-87.

CANDÉS, E. J. and DONOHO, D. (2002). Recovering edges in ill-posed inverse problems: optimality of curvelet frames. *Ann. Statist.* **30**, 784–842.

CAVALIER, L. and KOO, J.-Y. (2002). Poisson intensity estimation for tomographic data using a wavelet shrinkage approach. *IEEE Trans. Inform. Theory* **48** , 2794–2802.

FARIDANI, A., RITMAN, E., and SMITH, K. (1992). Local tomography. *SIAM J. Appl. Math.* **52**, 459–484.

GARDNER, R. (1995). *Geometric Tomography.* Cambridge University Press, Cambridge.

GELFAND, I. M., GRAEV, M. I., and VILENKIN, N. YA. (1966). *Generalized Functions. Vol. 5. Integral geometry and representation theory.* Academic Press, New York-London, 1966.

GOLDENSHLUGER, A., and SPOKOINY, V. (2004). On the shape–from–moments problem and recovering edges from noisy Radon data. *Probab. Theory Rel. Fields* **128**, 123-140.

GOLDENSHLUGER, A. and ZEEVI, A. (2004). Recovering convex boundaries from blurred and noisy observations. *Manuscript.*

GROEMER, H. (1996). *Geometric Applications of Fourier Series and Spherical Harmonics.* Cambridge University Press, Cambridge.

HÄRDLE, W., PARK, B. U. and TSYBAKOV, A. (1995). Estimation of non–sharp support boundaries. *J. Multivariate Anal.* **55**, 205–218.

HERO, A. O., PIRAMUTHU, R., FESSLER, J. A. and TITUS, S. R. (1999). Minimax emission computed tomography using high–resolution anatomical side information and B–spline models. *IEEE Trans. Inform. Theory* **45**, 920-938.

JOHNSTONE, I. and SILVERMAN, B. (1990). Speed of estimation in the positron emission tomography and related inverse problems. *Ann. Statist.* **18**, 251-280.

JOHNSTONE, I. and SILVERMAN, B. (1991). Discretization effects in statistical inverse problems. *J. Complexity* **7**, 1–34.

KOROSTELEV, A. and TSYBAKOV, A. (1993). *Minimax Theory of Image Reconstruction.* Lectures notes in statistics, Springer, New York.

Korostelev, A. and Tsybakov, A. (1994). Asymptotic efficiency in estimation of a convex set. *Problems of Information Transmission* **30**, 317–327.

Korostelev, A., Simar, L. and Tsybakov, A. (1995). On estimation of monotone and convex boundaries. *Publ. Inst. Stat. Univ. Paris.* **39**, 3–18.

Mammen, E. and Tsybakov, A. (1995). Asymptotical minimax recovery of sets with smooth boundaries. *Ann. Statist.* **23**, 502-524.

Polzhel, J. and Spokoiny, V. (2000). Adaptive weights smoothing with applications to image restoration. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **62**, 335–354.

Quinto, E. T. (1993). Singularities of the X–ray transform and limited data tomography in $\mathbb{R}^2$ and $\mathbb{R}^3$. *SIAM J. Math. Analysis* **24**, 1215–1225.

Ramm, A. and Zaslavsky, A. (1993). Reconstructing singularities of a function from its Radon transform. *Math. Comput. Modelling* **18**, 109–138.

Ritov, Y. (1998). Estimating mass and shape of domains in PET imaging. *J. Nonparametric Statist.* **10**, 47-66,

Schneider, R. (1993). *Convex Bodies: The Brunn–Minkowski Theory.* Cambridge University Press, Cambridge.

Srinivasa, N., Ramakrishnan, K. R. and Rajgopal, K. (1992). Edge detection from projections. *IEEE Trans. Med. Imaging* **11**, 76-80.

Talagrand, M. (1994). Sharper bounds for Gaussian and empirical processes. *Ann. Probab.* **22**, 28–76.

van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes.* Springer, New York.

Vardi, Y., Shepp, L. and Kaufman, L. (1985). A statistical model for positron emission tomography. With discussion. *J. Amer. Statist. Assoc.* **80**, 8–37.

Wang, Y. (1998). Change curve estimation via wavelets. *J. Amer. Statst. Assoc.* **93**, 163–172.

Ye, J. C., Bresler, Y. and Moulin, P. (2000). Asymptotic global confidence regions in parametric shape estimation problems. *IEEE Trans. Inform. Theory* **46**, 1881-1895.