Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Recovering Edges of an Image from Noisy Tomographic Data

Alexander Goldenshluger [1] Vladimir Spokoiny [2]

[1] Department of Statistics,
Haifa University,
Haifa 31905, Israel
E-Mail: goldensh@stat.haifa.ac.il

[2] Weierstrass Institute
and Humboldt University Berlin,
Mohrenstr. 39, 10117 Berlin, Germany
E-Mail: spokoiny@wias-berlin.de

No. 909
Berlin 2004

1991 *Mathematics Subject Classification.* 62G20; 62C20; 94A08.

Key words and phrases. Radon transform, optimal rates of convergence, support function, edge detection, minimax estimation.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

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.

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.

The problem of edge detection arises in numerious imaging applications. For example, images with discontinuitites 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 detection in tomographic images is extensively studied in the literature, both from the theoretical and applied perspective. Different techniques of edge reconstruction based on local inversion formulas are proposed in Faridani et al. (1992), Katsevich and Ramm (1995), Katsevich and Ramm (1996), and Faridani et al. (1997). The main idea underlying these proposals is to recover a transformation of f which admits local inversion and has the same set of singularities as f itself. This literature concentrates on mathematical properties of the local inversion formulate but ignores the effect of the measurement noise which may be significant in practice. We refer to Ramm and Katsevich (1996) for comprehensive review of this line of research and further references. There is a large amount of papers in the image processing literature where many practically useful algorithms for estimating edges from tomographic data are proposed. A representative publication from this area is, e.g., Srinivasa, Ramakrishnan and Rajgopal (1992). This literature, however, focuses exclusively on algorithmic and implementation aspects and lacks theoretical justification.

Although various methods and proposals are widely used in practice, theoretical limitations in the problem of edge detection 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? The goal of the present paper is to provide a theoretical perspective on these questions and to develop easily implemented nearly–optimal algorithm for edge recovery in tomographic images.

Recently Candés and Donoho (2002) considered the problem of recovering images with edges from the Radon data contaminated by Gaussian white noise with zero mean and variance σ^2. 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 $O(\sigma^{2/5})$. A curvelets–based estimator is developed in the cited paper, and it is shown to be nearly optimal in the sense of the convergence rates. In this paper we consider the model of Candés and Donoho (2002), but our focus is on direct recovery of the edge rather than on estimating the whole image. 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. We show that convex edge can be estimated with the pointwise risk of the order $O(\sigma^{4/5})$, and this rate cannot be essentially improved. It turns out that the main factor determining the rate of convergence is the curvature of the edge curve, and not its smoothness (provided that the edge curve is at least twice differentiable). Numerical examples illustrate reasonable practical behavior of the proposed estimator.

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, while Section 4 deals with theoretical analysis of its properties. In Section 5 we present some numerical examples; Section 6 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

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

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 exists $L > 0$ such that for sufficiently small $h > 0$

$$|(\mathcal{R}f)(g(\varphi), \varphi) - (\mathcal{R}f)(g(\varphi) - h, \varphi)| \ge Lh^{1/2}, \quad \forall\varphi. \tag{2}$$

This can be explained using simple geometrical argument which is illustrated in Figure 1 for $f = \tilde{f}\mathbf{1}_G$, $\tilde{f} \ge c > 0$. In this case the Radon transform $\mathcal{R}f$ is supported on the set $\{(s, \varphi) : 0 \le s \le g(\varphi), \varphi \in [0, 2\pi)\}$, and $(\mathcal{R}f)(g(\varphi) - h, \varphi)$ equals to the "weighted" length of the chord AB. Since ∂G has non–zero curvature and $\tilde{f} \ge c > 0$, this "weighted" length

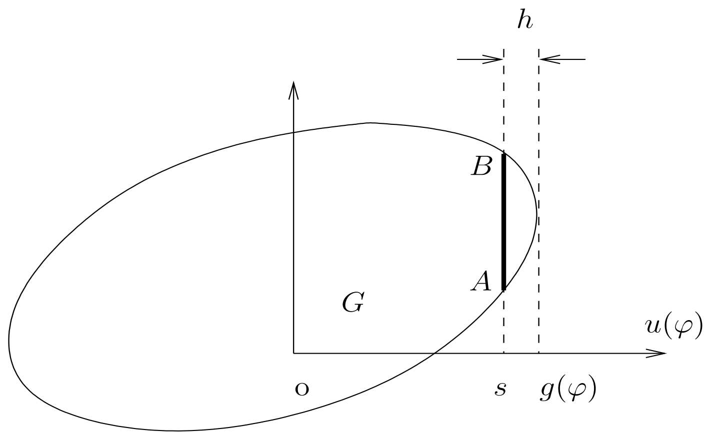

Figure 1: An illustration of the Radon transform behavior near the edge for $f(x) = \tilde{f}(x)\mathbf{1}_G(x)$, $\tilde{f}(x) \geq c > 0$.

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 $\varphi \in [0, 2\pi)$:

$$|(\mathcal{R}f)(\tau,\varphi) - (\mathcal{R}f)(t,\varphi)| \leq L_{\mathcal{R}}|\tau - t|, \quad \forall \tau < t \ \text{ s.t. } \ g(\varphi) \notin [\tau, t]. \tag{3}$$

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). We note, however, that in our setup g is the support function of a convex set; this fact leads to results which are not directly comparable to those in the cited papers.

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

Functional class. We say that function f on $B^2(o,1)$ belongs to the class $\mathcal{F} := \mathcal{F}(r, R)$ if it satisfies the following assumptions

(A) $|f(x)| \leq M$, $\forall x \in B^2(o,1)$, and f is smooth apart from a discontinuity jump along a curve which is the boundary 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

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

The collection of convex sets satisfying (B) will be designated $\mathcal{G} := \mathcal{G}(r, R)$.

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). For example, if $f = \tilde{f}\mathbf{1}_G$, $\tilde{f} \geq c > 0$ then (2) is valid with $L = c\sqrt{r}$. Inequality

(4) states the lower and upper bounds on the radius of curvature $\rho(\varphi) = g(\varphi) + g''(\varphi)$ of the boundary [see (iii)]. In what follows we always assume that $R \gg r$ so that the class $\mathcal{F}$ is rich enough. Note that when $r = R$ the class $\mathcal{F}$ contains only discs of the radius r. The lower bound in (4) implies that G is the r*–convex set.* We recall that a set G is called r*-convex* 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–convex if $g(\cdot) - r$ is the support function of a convex set.

Properties of convex sets from the class $\mathcal{G}$. First we investigate some properties of sets G from the class $\mathcal{G}$.

Lemma 1 *Let G and K be convex sets from $\mathcal{G}(r, R)$ such that*

$$g_G(\theta) - g_K(\theta) \geq h, \quad \text{for some } \theta \in [0, 2\pi) \text{ and } h > 0; \tag{5}$$

here $g_A(\cdot)$ stands for the support function of a convex set A. Then for sufficiently small h and any $\delta \leq \sqrt{2h/R}$

$$g_G(\varphi) - g_K(\varphi) \geq r\psi(\varphi), \quad \varphi \in [\theta - \delta, \theta + \delta], \tag{6}$$

where

$$\psi(\varphi) := \begin{cases} 1 - \cos(\varphi - (\theta - \delta)), & \varphi \in [\theta - \delta, \theta], \\ 1 - \cos(\varphi - (\theta + \delta)), & \varphi \in [\theta, \theta + \delta]. \end{cases} \tag{7}$$

We note that under the premise of the lemma $r\psi(\varphi) \leq r\psi(\theta) = r(1 - \cos\delta) \leq hr/R \leq h$. The lemma essentially states that if the support functions of of two sets G and K from $\mathcal{G}(r, R)$ are separated at least by h at a single point $\theta \in [0, 2\pi)$, then there exists a $O(\sqrt{h})$–vicinity of θ where the support functions are also "well–separated"; the separating function ψ_θ is given by (7). As the proof shows, r–convexity of sets from $\mathcal{G}(r, R)$ is essential here; if only convexity is assumed the vicinity size of the order $O(\sqrt{h})$ cannot be ensured.

The probe functional. Using the property established in Lemma 1 we introduce the following definition. Assume that $f \in \mathcal{F}(r, R)$, let $\delta > 0$ and $I_\delta := [\theta - \delta, \theta + \delta]$. For $t \in (0, 1)$ we define *the probe functional*

$$\ell_\delta(t) := \int_{I_\delta} \int_{t - r\psi(\varphi)}^{t} (\mathcal{R}f)(\tau, \varphi) d\tau d\varphi - \int_{I_\delta} \int_{t}^{t + r\psi(\varphi)} (\mathcal{R}f)(\tau, \varphi) d\tau d\varphi, \tag{8}$$

where $\psi(\cdot)$ is given in (7). The functional $\ell_\delta(t)$ will be used for detecting the location of the cusp curve g at a single given point θ. The region of integration in (8) defines a *diamond–like template* in the Radon domain whose profile is adjusted to the properties of the sought cusp curve, see Lemma 1. The next statement shows that the absolute value of $\ell_\delta(t)$ is large when $t = g(\theta)$ and small when t is separated from $g(\theta)$; the localization accuracy of $\ell_\delta(t)$ is not less than $R\delta^2/2$.

Lemma 2 *Let $G \in \mathcal{G}(r, R)$, and δ be sufficiently small.*

(i) Then

$$|\ell_\delta(g(\theta))| \geq C_1 L r^{3/2} \delta^4 - C_2 L_{\mathcal{R}}\, r^2 \delta^5, \tag{9}$$

where L and $L_{\mathcal{R}}$ are given in (2) and (3), and C_1, C_2 are absolute constants.

(ii) Let $h = R\delta^2/2$ and $|t - g(\theta)| > h$; then

$$|\ell_\delta(t)| \leq C_3 L_{\mathcal{R}}\, r^2 \delta^5, \tag{10}$$

where C_3 is an absolute constant.

We note that other probe functionals can be used for the cusp curve detection. For example, one can define $\ell_\delta(t)$ equal to the second integral on the right hand side of (8). In this case the same separation rates as in (9)–(10) are valid although the constants will depend on the magnitude of f and not on L and $L_{\mathcal{R}}$. The important features of the probe functional construction are profile and scaling of the vertical and horizontal size of the template. As it will be shown, the template profile allows the "maximal smoothing" along the angles while preserving "good" localization properties in the vertical direction.

3 Estimation procedure

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. The location of this singularity can be described as the point of maximum of the probe functional $\ell_\delta(t)$. This leads to the following procedure. For fixed angle $\theta \in [0, 2\pi)$ we estimate the probe functional $\ell_\delta(t)$ for different values of $t \in [h, 1-h]$, $h = R\delta^2/2$, and define $\hat{g}(\theta)$ as the value of t where the maximum of the estimated probe functional is achieved. More formally, we denote

$$\hat{\ell}_\delta(t) := \int_{I_\delta} \int_{t - r\psi(\varphi)}^{t} Y(d\tau, d\varphi) - \int_{I_\delta} \int_{t}^{t + r\psi(\varphi)} Y(d\tau, d\varphi)\ , \tag{11}$$

where the function $\psi(\varphi)$ is defined in (7), and let

$$\hat{g}(\theta) = \hat{t} = \arg \max_{t \in [h, 1-h]} |\hat{\ell}_\delta(t)|. \tag{12}$$

The definition of $\hat{g}(\theta)$ depends on parameters δ, and r. With these parameters Lemma 2 states that for any $G \in \mathcal{G}(r, R)$ the localization accuracy of ℓ_δ is at least $h = R\delta^2/2$. In what follows δ will be chosen as a function of σ ensuring the optimal rate of convergence of the mean squared error as $\sigma \to 0$. As for the choice of r, we assume that $\mathcal{F}(r, R)$ is given so that r is known. It can be shown that the choice of r does not affect the rates of convergence in the following sense. If we set in the estimation procedure $r = r_0$, then the statement of Theorem 1 below remains true for any class $\mathcal{F}(r, R)$ with $r \geq r_0$. It is important to realize however that the choice of these parameters is crucial for practical implementation of the proposed estimation scheme. We discuss these issues in Section 6.

4 Theoretical properties

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

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

$$\delta = C_1\Big\{\sigma\sqrt{\ln\frac{1}{\sigma}}\Big\}^{2/5} \quad \textit{for some } C_1 > 0, \tag{13}$$

and $h = R\delta^2/2$. Then for sufficiently small σ

$$\sup_{f\in\mathcal{F}(r,R)} \mathbb{E}|\hat{g}(\theta) - g(\theta)|^\nu \le C_2\left\{\sigma^{4/5}\Big(\ln\frac{1}{\sigma}\Big)^{2/5}\right\}^\nu, \quad \forall \nu > 0,$$

where C_2 depends on r, R, and ν.

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

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

where C 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. It is interesting to note that the rate of the order $O(\sigma^{4/5})$ is determined by the curvature properties of edge, and not by the smoothness of the edge curve (as long as the corresponding support function is twice differentiable). In particular, the lower bound shows that this rate cannot be essentially improved even if the the edge curve is infinitely differentiable.

Global accuracy measures. 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 (11) and (12). The estimate of the boundary ∂G is given by (14) with the inequality sign replaced by equality. Although construction of $\hat{g}$ does not ensure that $\hat{g}$ is the support function of a convex set, $\hat{G}$ is always convex because it is defined as the envelope of the estimated supporting lines. 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 following statement is an immediate consequence of Theorem 1.

Theorem 3 *Let $\hat{G}$ be given by (14); then under conditions of Theorem 1 for sufficiently small σ*

$$\sup_{f\in\mathcal{F}(r,R)} \Big\{\mathbb{E}[d_p^p(\hat{G}, G)]\Big\}^{1/p} \le C_4\sigma^{4/5}\Big(\ln\frac{1}{\sigma}\Big)^{2/5}, \quad p \in [1,\infty], \tag{15}$$

where C_4 is the constant depending on r, R and p. In the case $p = \infty$ the left hand side of (15) is interpreted as $\sup_{f\in\mathcal{F}(r,R)} \mathbb{E}d_\infty(\hat{G}, G)$.

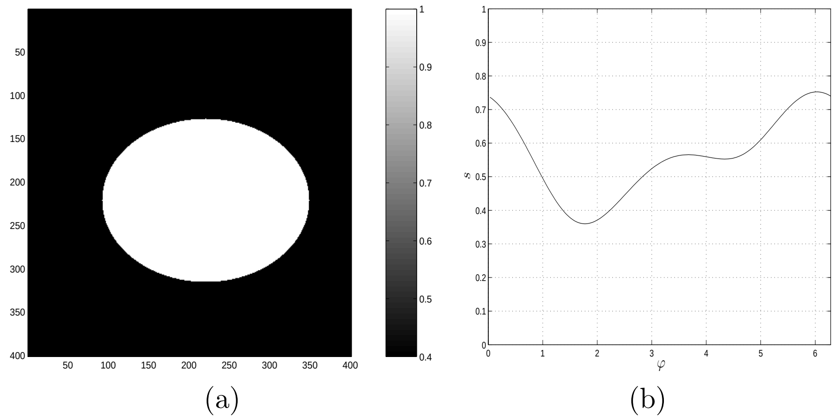

Figure 2: (a) The original image; (b) the support function of the ellipse.

5 Numerical examples

We conducted a small numerical experiment in order to illustrate practical behavior 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 200×200 regular grid 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.

The following version of the proposed estimator was implemented. Recall that the diamond–like template in the construction of the probe functional (8) is defined via function ψ [see (7)] and depends on the two design parameters, δ and r. Our numerical experience indicates that a larger class of the probe functional templates allows to obtain better practical results. In particular, the template used in simulations is obtained by the intersection of the rectangle having vertical and horizontal sizes $2h$ and 2δ respectively with the diamond–like template of (8). The new template is different from that defined in (8) only if $h \le r(1 - \cos\delta)$. Thus three design parameters h, δ and r should be selected. In the numerical examples below we always set $r = 0.3$ while the values of h and δ were selected to achieve the best visual appearance of the estimated edge. Because the data are available on the 200×200 regular grid, we specify the bandwidths h and δ in pixels, i.e., for example, $h = 5$ means that $h = 5/200$ while $\delta = 5$ means $\delta = 5(2\pi/200)$. We would like to emphasize that our goal here is merely to demonstrate practical potential of the proposed estimation scheme. The question of data–driven selection of parameters is

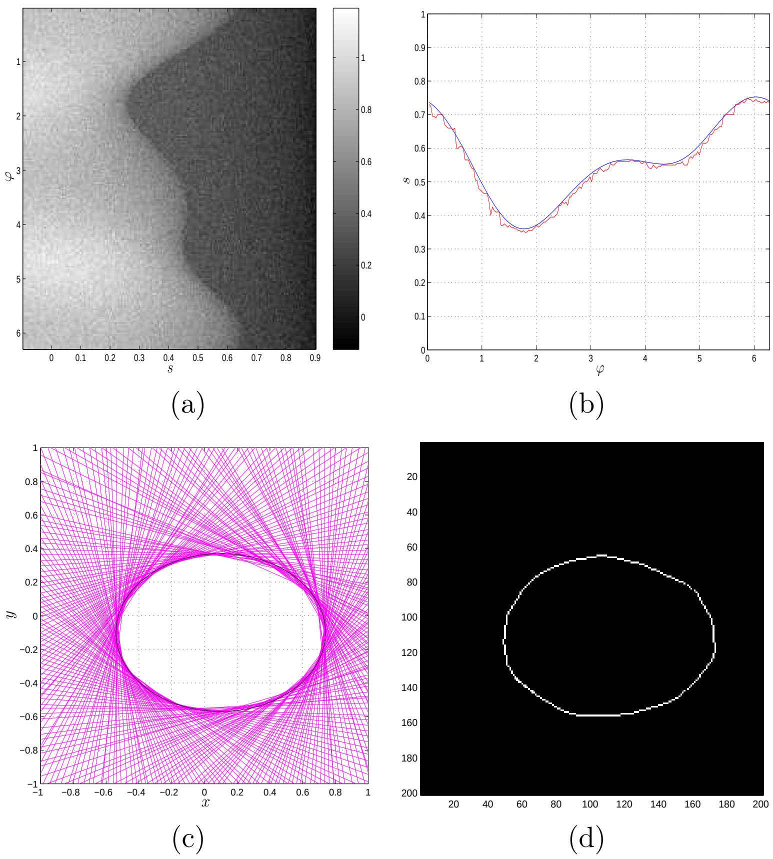

Figure 3: Edge recovery for the low noise level ($\sigma = 0.05$): (a) The noisy observations in the Radon domain. (b) The support function estimate. (c) The "true" edge (solid line) along with the estimated supporting lines. (d) The extracted estimate of the edge.

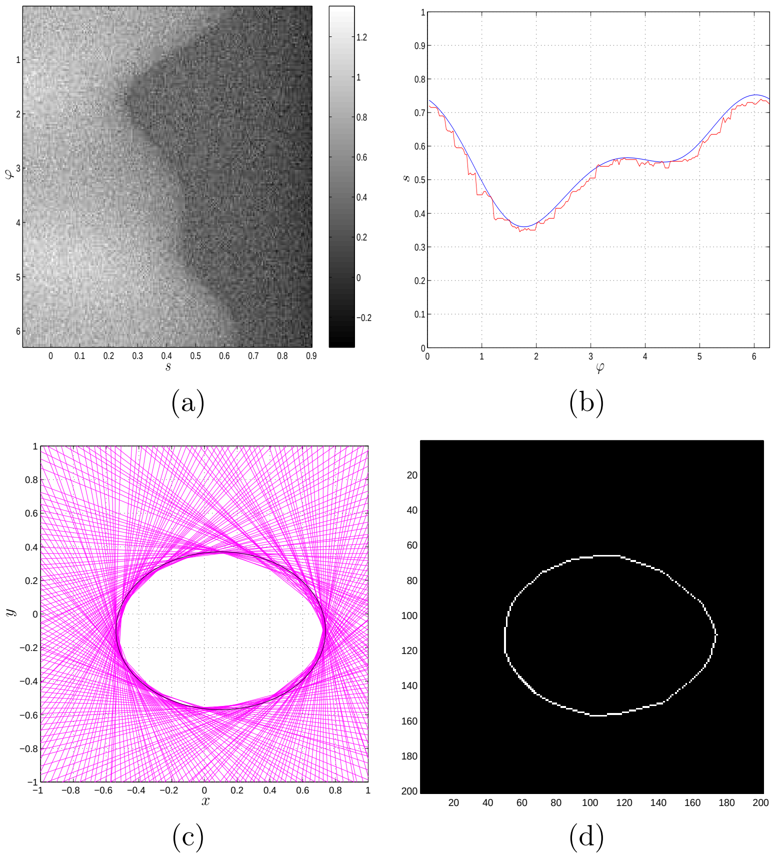

Figure 4: Edge recovery for the medium noise level ($\sigma = 0.1$): (a) The noisy observations in the Radon domain. (b) The support function estimate. (c) The "true" edge (solid line) along with the estimated supporting lines. (d) The extracted estimate of the edge.

crucial for applications. This subject however is beyond the scope of the present paper.

Figure 3 displays the results obtained for the case of low noise level conditions, $\sigma = 0.05$. Here the values $h = 4$ and $\delta = 6$ 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 $\sigma = 0.15$ respectively. We note only that in the case of the medium noise level we selected $h = 7$ and $\delta = 9$, while in the case of high noise level $h = 12$ and $\delta = 12$.

The numerical results demonstrate reasonable practical behavior of the proposed estimation scheme. By construction our support function estimator has a negative bias so that the recovered set tends to be smaller than the original one. The numerical results also clearly demonstrate that the estimation accuracy is better at the points of low curvature.

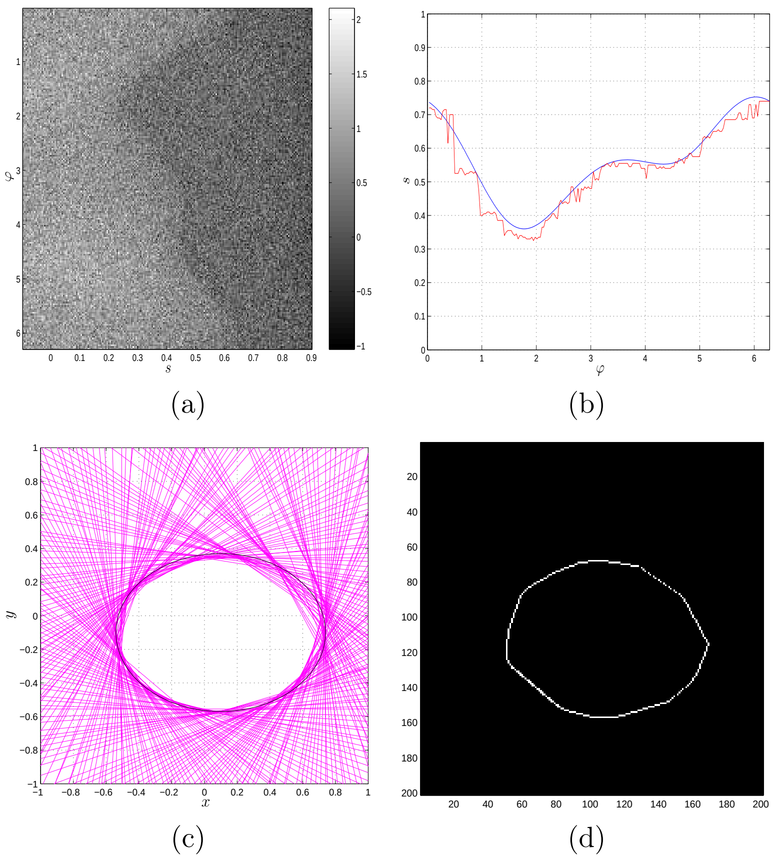

Figure 5: Edge recovery for the high noise level ($\sigma = 0.3$): (a) The noisy observations in the Radon domain. (b) The support function estimate. (c) The "true" edge (solid line) along with the estimated supporting lines. (d) The extracted estimate of the edge.

6 Concluding remarks

1. The theoretical results show that our estimator is nearly optimal in order within a logarithmic in σ^{-1} factor. Numerical examples of Section 5 also indicate that the proposed procedure may be rather useful in practical applications. We note however that the implementation depends on design parameters that should be chosen in some way. Even though this choice does not affect the asymptotic properties of the proposed procedure, it has a crucial effect on the estimator performance for finite sample sizes. The question of data–driven selection of parameters can be approached using recently developed adaptation techniques. In particular, more practical edge recovery procedure can be based on the fully adaptive AWS algorithm of Polzehl and Spokoiny (2000). It can be used for adaptive estimation of the cusp curve in the Radon domain.

2. Our approach establishes an interesting connection between the problem of edge detection in tomography and the boundary fragment model of Korostelev and Tsybakov (1993). Indeed, the problem of estimating the cusp curve in the Radon domain resembles to the problem of estimating non–sharp boundaries as in Härdle, Park and Tsybakov (1995) and Wang (1998). Still there is a fundamental difference. Because the cusp curve is the support function of a convex set with smooth boundary, the main factor affecting the estimating accuracy is the boundary curvature, and not the boundary smoothness. Our lower bound shows that the rate $O(\sigma^{4/5})$ cannot be essentially improved even for infinitely differentiable boundaries. This is in the striking contrast with estimating non–sharp boundaries, where the rates are essentially determined by the boundary smoothness [see Härdle, Park and Tsybakov (1995)].

3. As it might be expected, the edge can be recovered more accurately than the whole image f in $\mathbb{L}_2$. In particular, it follows from the results of Candés and Donoho (2002) that the best rate of convergence for recovering an image which is twice differentiable apart from a twice differentiable smooth curve in $\mathbb{L}_2$ is $O(\sigma^{2/5})$, while the convex edge can be estimated with the faster rate $O(\sigma^{4/5})$.

4. 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 this case it is natural to consider the Radon transform supported on $[0, \pi) \times [-1, 1]$; then the support function of a convex set is represented by two curves in the Radon domain. If convex sets of the image have empty intersection these curves are well separated in the Radon domain. In addition, 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.

5. The points of zero curvature on the boundary correspond to sharper cusps in the Radon domain. We note however that one cannot improve accuracy of estimating the support function in these particular directions because the set of points where the curvature vanishes has zero Lebesgue measure. Thus our results are valid for the class of functions with edge along a smooth convex closed curve which can have points of zero curvature.

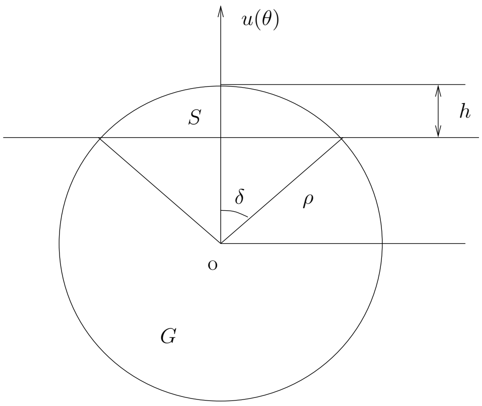

Figure 6: Illustration of the proof of Lemma 1.

Appendix

Proof of Lemma 1

First we establish the lemma for the case where G is a disc. Then we will show how this result can be extended for general sets from the class $\mathcal{G}(r, R)$.

The proof is based on the following simple geometrical argument. Let G be the disc of radius $\rho \in [r, R]$ centered at the origin. Clearly in this case $g_G(\varphi) = \rho$, $\forall\varphi$. Fix $\theta \in [0, 2\pi)$ and for sufficiently small h let

$$\tilde{G} := G \backslash S, \quad S := \{(x, y) : g(\theta) - h \leq x\cos\varphi + y\sin\varphi \leq g(\theta),\ \forall\varphi \in [0, 2\pi)\},$$

see Figure 6. It is evident that for any convex set $K \in \mathcal{G}(r, R)$ satisfying (6),

$$g_K(\varphi) \leq g_{\tilde{G}}(\varphi), \quad \forall\varphi \in [\theta - \delta, \theta + \delta]$$

where $\delta := \arccos(1-h/\rho)$ [see Figure 6]. It is easily verified that $g_{\tilde{G}}(\varphi)$ equals $\rho\cos(\varphi-(\theta-\delta))$ whenever $\varphi \in [\theta-\delta, \theta]$, and $\rho\cos(\varphi-(\theta+\delta))$ whenever $\varphi \in [\theta, \theta+\delta]$. Then (6) holds with r replaced by ρ and $\delta = \arccos(1-h/\rho)$. Because $\cos\delta = 1-(\delta^2/2!)+(\delta^4/4!)-(\delta^6/6!)+\cdots$, (6) is also valid for every $\delta \leq \sqrt{2h/\rho}$ as claimed in the statement of the lemma. Thus (6) holds when G is the disc of radius ρ centered in the origin.

For general sets $G \in \mathcal{G}(r, R)$ we use the same reasoning as before. For sufficiently small h we replace the boundary of the set G in the vicinity of the support value in direction $u(\theta)$ by the arc of the disc B centered at $e(\theta) = g(\theta)u'(\theta) - g''(\theta)u(\theta)$ and having the radius $\rho(\theta) = g_G(\theta) + g''_G(\theta)$ [see (iii) of Section 2]. Then the support function of any convex set $K \in \mathcal{G}(r, R)$ satisfying (6) must be less that the support function of $B\backslash S$ in a small vicinity of the support value (for sufficiently small h). Hence the situation differs from that discussed above only in that B is not centered at the origin. However, the difference between support functions of two convex sets is preserved under translation of these sets. Therefore (6) follows from the above considerations taking into account that $r \leq \rho(\varphi) \leq R$, $\forall\varphi \in [0, 2\pi)$. $\blacksquare$

Proof of Lemma 2

Proof (i). We have

$$
\begin{aligned}
\ell_\delta(g(\theta)) &= \int_{I_\delta}\Big\{\int_{g(\theta)-r\psi(\varphi)}^{g(\theta)}(\mathcal{R}f)(\tau,\varphi)d\tau - \int_{g(\varphi)}^{g(\varphi)+r\psi(\varphi)}(\mathcal{R}f)(\tau,\varphi)d\tau\Big\}d\varphi \\
&= \int_{I_\delta}\int_{g(\theta)-r\psi(\varphi)}^{g(\theta)}[(\mathcal{R}f)(\tau,\varphi)-(\mathcal{R}f)(g(\theta),\varphi)]d\tau d\varphi \\
&\quad - \int_{I_\delta}\int_{g(\theta)}^{g(\theta)+r\psi(\varphi)}[(\mathcal{R}f)(\tau,\varphi)-(\mathcal{R}f)(g(\theta),\varphi)]d\tau d\varphi =: J_1 - J_2.
\end{aligned}
$$

By (2),

$$
\begin{aligned}
|J_1| &\geq \int_{I_\delta}\int_{g(\theta)-r\psi(\varphi)}^{g(\theta)} L|g(\theta)-\tau|^{1/2}d\tau d\varphi \\
&= \frac{2}{3}\int_{I_\delta} L[r\psi(\varphi)]^{3/2}d\varphi = C_1 L r^{3/2}\delta^4 ,
\end{aligned}
$$

where C_1 is an absolute constant. Further, because $(\mathcal{R}f)$ is smooth on the set $\{(s,\varphi) : g(\theta)\leq s\leq g(\theta)+r\psi(\varphi),\ \varphi\in I_\delta\}$, (3) implies

$$
\begin{aligned}
|J_2| &\leq \int_{I_\delta}\int_{g(\theta)}^{g(\theta)+r\psi(\varphi)} L_{\mathcal{R}}|\tau-g(\theta)|d\tau d\varphi \\
&= \frac{r^2}{2}L_{\mathcal{R}}\int_{I_\delta}\psi^2(\varphi)d\varphi = C_2 r^2 L_{\mathcal{R}}\delta^5,
\end{aligned}
$$

where C_2 is an absolute constant. Hence (9) follows.

(ii). Similarly to the above considerations,

$$
\begin{aligned}
\ell_\delta(t) &= \int_{I_\delta}\int_{t-r\psi(\varphi)}^{t}[(\mathcal{R}f)(\tau,\varphi)-(\mathcal{R}f)(g(\theta),\varphi)]d\tau d\varphi \\
&\quad - \int_{I_\delta}\int_{t}^{t+r\psi(\varphi)}[(\mathcal{R}f)(\tau,\varphi)-(\mathcal{R}f)(g(\theta),\varphi)]d\tau d\varphi. \qquad (16)
\end{aligned}
$$

Assume, e.g., that $t > g(\theta)+h$. Then by Lemma 1, $t-r\psi(\varphi)\geq g(\varphi)$, $\forall\varphi\in I_\delta$, and therefore the cusp curve $\{(s,\varphi) : s = g(\varphi),\ \varphi\in[0,2\pi)\}$ is separated from the set $\{(s,\varphi) : t-r\psi(\varphi) < s\leq t+r\psi(\varphi),\ \varphi\in I_\delta\}$. Thus the integrals in (16) are bounded similarly to J_2 above. In the case $t < g(\theta)-h$ separation of $g(\varphi)$ from the set $\{(s,\varphi) : t\leq s < t+r\psi(\varphi),\ \varphi\in I_\delta\}$ is shown similarly. ■

Proof of Theorem 1

In the proof $c_1, c_2, \ldots$ denote constants that may depend on parameters r, R, L, $L_{\mathcal{R}}$ and ν only.

Because $\mathrm{supp}(f)\subseteq B^2(o,1)$, we can write for any $\nu > 0$

$$
\begin{aligned}
\mathbb{E}|\hat{g}(\theta)-g(\theta)|^\nu &\leq h^\nu + \mathbb{E}\Big[|\hat{g}(\theta)-g(\theta)|^\nu \mathbf{1}\{|\hat{g}(\theta)-g(\theta)| > h\} \\
&\leq h^\nu + \mathbb{P}\Big\{|\hat{g}(\theta)-g(\theta)| > h\Big\}. \qquad (17)
\end{aligned}
$$

Our goal is to bound the probability on the RHS of (17).

We have for small enough σ

$$
\begin{aligned}
\mathbb{P}\{|\hat{g}(\theta)-g(\theta)|>h\} &\le \mathbb{P}\Big\{\max_{t:|t-g(\theta)|>h}|\hat{\ell}_\delta(t)|\ge|\hat{\ell}_\delta(g(\theta))|\Big\}\\
&\le \mathbb{P}\Big\{\max_{t\in[h,1-h]}|\hat{\ell}_\delta(t)-\ell_\delta(t)|+\max_{t:|t-g(\theta)|>h}|\ell_\delta(t)|\ge|\hat{\ell}_\delta(g(\theta))|\Big\}\\
&\overset{(a)}{\le} \mathbb{P}\Big\{\max_{t\in[h,1-h]}|\hat{\ell}_\delta(t)-\ell_\delta(t)|+c_1L_{\mathcal{R}}\,r^2\delta^5\ge|\hat{\ell}_\delta(g(\theta))|\Big\}\\
&\le \mathbb{P}\Big\{2\max_{t\in[h,1-h]}|\hat{\ell}_\delta(t)-\ell_\delta(t)|+c_1L_{\mathcal{R}}\,r^2\delta^5\ge|\ell_\delta(g(\theta))|\Big\}\\
&\overset{(b)}{\le} \mathbb{P}\Big\{\max_{t\in[h,1-h]}|\hat{\ell}_\delta(t)-\ell_\delta(t)|\ge c_2Lr^{3/2}\delta^4\Big\}\ , \qquad (18)
\end{aligned}
$$

where (a) follows from Lemma 2 (ii), and (b) is a consequence of Lemma 2 (i) and the fact that σ is small. Thus it remains to bound from above the probability $\mathbb{P}\{\max_{t\in[h,1-h]}|X_t|\ge c_2Lr^{3/2}\delta^4\}$, where

$$
X_t:=\sigma\Big\{\int_{I_\delta}\int_{t-r\psi(\varphi)}^{t}W(d\tau,d\varphi)-\int_{I_\delta}\int_{t}^{t+r\psi(\varphi)}W(d\tau,d\varphi)\Big\} \qquad (19)
$$

is the zero mean Gaussian process indexed by $t\in[h,1-h]$. 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

$$
\max_{t\in[h,1-h]}\mathbb{E}|X_t|^2\le c_3r\sigma^2\delta^3.
$$

Further, it is straightforward to see that $\mathbb{E}|X_t-X_s|^2\le c_4\sigma^2\delta|t-s|$. Therefore one needs no more than $N(\epsilon)=c_5(\sigma^2\delta\epsilon^{-1})^{1/2}$ balls of the radius ϵ in the natural semimetric in order to cover the index set $[h,1-h]$. In addition, it follows from (13) that $\sigma\delta^{3/2}=o(\delta^4)$ as $\sigma\to 0$. Therefore applying the exponential inequality of Proposition A.2.7 from van der Vaart and Wellner (1996) [with $V=1/2$, $\epsilon_0\sim\sigma\delta^{3/2}$, $K\sim\sigma^2\delta$ and $\lambda\sim\delta^4$] we obtain

$$
\begin{aligned}
\mathbb{P}\{\max_{t\in[h,1-h]}|X_t|\ge c_2r\delta^4\} &\le c_6\delta\exp\Big\{-\frac{c_5\delta^8}{\sigma^2\delta^3}\Big\}\\
&\le c_6\delta\exp\{-c_7h^{5/2}\sigma^{-2}\}\le c_8h^\nu
\end{aligned}
$$

where the last inequality follows by choice of C_1 in (13). ■

Proof of Theorem 2

In the proof below $c_1,c_2,\dots$ stand for absolute constants or constants depending on r and R only. We assume that $R\gg r$ so that class $\mathcal{G}(r,R)$ is sufficiently rich (e.g., $r=R$ implies that $\mathcal{G}(r,R)$ contains only discs of radius $r=R$). Without loss of generality we assume $\theta=\pi/2$, and let G_0 be the disc $B^2(o,c_1r)$ of radius c_1r, $c_1>1$ centered at the origin $o=(0,0)$. The support function of G_0 is $g_{G_0}(\varphi)=g_0(\varphi)=c_1r$, $\forall\varphi$. For some $h>0$ define $\tilde{G}_1=G_0\cap B^2(\eta,R)$, where $B^2(\eta,R)$ is the disc of radius R centered at

$\eta = (0, -R - h + c_1 r)$. By construction, $g_{\tilde{G}_1}(\pi/2) + h = g_{G_0}(\pi/2)$; note however, that $\tilde{G}_1 \notin \mathcal{G}(r, R)$, because $\partial\tilde{G}_1$ is not differentiable at the points of intersection of $\partial B^2(\eta, R)$ and ∂G_0. We define $G_1 \in \mathcal{G}(r, R)$ by replacing the boundary of $\tilde{G}_1$ by circular arcs of the radius r in vicinity of the singularity points; this is always possible if $\mathcal{G}(r, R)$ is rich enough. For such a set $G_1 \subset G_0$ we have $g_{G_1}(\pi/2) = g_{G_0}(\pi/2) - h$.

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 distance between the probability measures $\mathbb{P}_0$ and $\mathbb{P}_1$ corresponding to the processes

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

is given by

$$\begin{aligned} KL(\mathbb{P}_0, \mathbb{P}_1) &= \mathbb{E}_0 \ln \frac{d\mathbb{P}_1}{d\mathbb{P}_0}(Y_0) \\ &= \frac{1}{2\sigma^2} \int_0^{2\pi} \int_0^1 |(\mathcal{R}(f_0 - f_1))(\tau, \varphi)|^2 d\tau d\varphi \; . \end{aligned} \tag{20}$$

To bound $KL(\mathbb{P}_0, \mathbb{P}_1)$ we use the idea similar to that in Candés and Donoho (2002); namely, we bound the Radon transform of the set $G_0 \backslash G_1$ by the Radon transform of some ellipse. Indeed, because G_1 is convex, it necessarily belongs to the set $G_0 \backslash \{(x, y) : c_1 r - h \leq y \leq c_1 r\}$. Therefore $G_0 \backslash G_1$ contains an ellipse with semi-axes of the size $c_2 h$ and $c_3\sqrt{h}$. On the other hand, it is easily checked that $G_0 \backslash G_1$ is also contained in some ellipse with semi-axes $c_4 h$ and $c_5\sqrt{h}$. Recall that the Radon transform of the ellipse $\mathcal{E}(a, b)$ with semi-axes a and b is given by

$$(\mathcal{R}\mathbf{1}_{\mathcal{E}(a,b)})(\tau, \varphi) = \frac{ab}{s}\Big(1 - \frac{\tau^2}{s^2}\Big)_+^{1/2}, \quad s^2 := a^2\cos^2\varphi + b^2\sin^2\varphi. \tag{21}$$

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

$$\{\mathcal{R}\mathbf{1}_{\mathcal{E}(a,b)}(Qx - e)\}[\tau, u(\varphi)] = \{\mathcal{R}\mathbf{1}_{\mathcal{E}(a,b)}\}[\tau - e^T Qu(\varphi), Qu(\varphi)].$$

Using this property and (21) we bound the integral on the RHS of (20) as follows

$$\begin{aligned} \|\mathcal{R}(f_0 - f_1)\|_2^2 &\leq c_6 \int_0^{2\pi} \int_0^1 \frac{h^3}{s^2}\Big(1 - \frac{\tau^2}{s^2}\Big)_+ d\tau d\varphi \\ &= c_7 h^3 \int_0^{2\pi} \frac{d\varphi}{\sqrt{h^2\cos^2\varphi + h\sin^2\varphi}} \\ &\leq c_8 h^{5/2} \ln\frac{1}{h} \; . \end{aligned}$$

Thus if we choose $h = c_9[\sigma^2(\ln\frac{1}{\sigma})^{-1}]^{2/5}$ the Kullback–Leibler distance will be of the order of $O(1)$; this implies the lower bound. ■

Proof of Theorem 3

First we note that

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

Therefore in the case of $p \in [1, \infty)$ the statement follows immediately from Theorem 1 by integrating the risk upper bound over $\theta \in [0, 2\pi)$. We indicate only modifications in the proof of Theorem 1 needed to obtain the announced upper bound when $p = \infty$. In this case the argument similar to (18) leads to

$$\mathbb{P}\Big\{ \max_{\theta \in [0,2\pi)} |\hat{g}(\theta) - g(\theta)| > h \Big\} \le \mathbb{P}\Big\{ \max_{\theta \in [0,2\pi)} \max_{t \in [h, 1-h]} |X_t(\theta)| \ge cLr^{3/2}\delta^4 \Big\},$$

where the Gaussian process $\{X_t(\theta)\}$ is again given by (19), but now its index set is $(\theta, t) \in [0, 2\pi) \times [h, 1-h]$. The general exponential inequality is applied to bound this probability. The announced result follows by appropriate choice of the constant C_1 in (13). ■

References

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

Faridani, A., Ritman, E., and Smith, K. (1992). Local tomography. *SIAM J. Appl. Math.* **52**, 459–484.

Faridani, A., Finch, D., Ritman, E., and Smith, K. (1997). Local tomography II. *SIAM J. Appl. Math.* **57**, 1095–1127.

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

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.

Katsevich, A. and Ramm, A. (1995). New methods for finding values of the jumps of a function from its local tomographic data. *Inverse Problems* **11**, 1005–1023.

Katsevich, A. and Ramm, A. (1996). Pseudolocal tomography. *SIAM J. Appl. Math.* **56**, 167–191.

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

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 Katsevich, A. (1996). *The Radon Transform and Local Tomography.* CRC Press, Boca Raton.

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

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.