Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

On the shape–from–moments problem and recovering edges from noisy Radon data

Goldenshluger, Alexander and Spokoiny, Vladimir

submitted: 20. December 2002

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

Weierstrass Institute
and Humboldt University Berlin,
Mohrenstr. 39, 10117 Berlin, Germany
E-Mail: spokoiny@wias-berlin.de
URL: http://www.wias-berlin.de/~spokoiny

No. 802
Berlin 2003

1991 *Mathematics Subject Classification.* 62C20, 62G20, 94A12.

Key words and phrases. minimax estimation, optimal rates of convergence, shape, moments, support function, Radon transform, tomography.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
D — 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 reconstructing a planar convex set from noisy observations of its moments. An estimation method based on pointwise recovering of the support function of the set is developed. We study intrinsic accuracy limitations in the shape–from–moments estimation problem by establishing a lower bound on the rate of convergence of the mean squared error. It is shown that the proposed estimator is near–optimal in the sense of the order. An application to tomographic reconstruction is discussed, and it is indicated how the proposed estimation method can be used for recovering edges from noisy Radon data.

1 Introduction

In this paper we consider the problem of reconstructing a planar region from noisy measurements of its moments. Let G denote a simply connected compact set on the plane belonging to the interior of the unit disc D. Assume that complex

$$\iint_D z^m \mathbf{1}_G(x,y)\,dxdy, \quad z = x + \imath y, \qquad m = 0, 1, \ldots \tag{1}$$

or geometric

$$\iint_D x^k y^l \mathbf{1}_G(x,y)\,dxdy, \qquad k, l = 0, 1, \ldots\ , \tag{2}$$

moments can be observed with gaussian noise having zero mean and variance σ^2. The shape–from–moments problem is to reconstruct the set G from noisy measurements of its moments.

The problem of reconstructing the shape of a planar object or region from indirect measurements arises in numerous applications. Milanfar et al. (1995) study recovery of polygons from the moment data and establish close connections of the shape–from–moments problem to array processing. Milanfar, Karl and Willsky (1996) develop a moment–based approach to tomographic reconstruction. An application in geophysics is discussed in Golub, Milanfar and Varah (1999). For detailed literature survey on the shape–from–moments problem we refer to Elad, Milanfar and Golub (2002).

Most reconstruction methods described in this literature deal with reconstructing polygons and are based on the so-called Motzkin–Schoenberg formula. If $z_1, \ldots, z_n$ designate the vertices of a polygon G in the complex plane, and if f is an analytic function in an open set containing G, then the Motzkin–Schoenberg formula states that

$$\iint_G f''(z)\,dxdy = \sum_{j=1}^{n} a_j f(z_j), \tag{3}$$

where coefficients $\{a_j\}$ do not depend on f, and are determined completely by the vertices $z_1, \ldots, z_n$ (and the way they are connected) [cf. Milanfar et al. (1995)]. Choosing $f(z) = z^k$ in (3) we obtain

$$k(k-1) \iint_G z^{k-2} dxdy = \sum_{j=1}^{n} a_j z_j^k \, , \tag{4}$$

so that *the weighted complex moments* [cf. (1)] are expressed directly through the vertices $z_1, \ldots, z_n$ of the polygon G. The next step is to observe that the sequence of the weighted complex moments in (4) satisfies a linear homogeneous difference equation whose characteristic polynomial has the roots $z_1, \ldots, z_n$. In this way the problem is reduced to estimating the roots of a characteristic polynomial from noisy observations of a sequence satisfying the corresponding linear homogeneous difference equation. This idea underlies the Prony method widely used in signal processing. Elad, Milanfar and Golub (2002) describe different estimation techniques (including Prony–based) and provide extensive simulation results.

Although various algorithms has been developed in the aforementioned literature, their statistical properties have not been studied thoroughly. Most studies focus exclusively on algorithmic and implementation aspects for reconstructing polygons. Recovery of quadrature domains from exact moment measurements is considered in Gustafsson et al. (2000). We note, however, that in practically all applications involving reconstruction of shapes from moments the effect of noise is significant (Golub, Milanfar and Varah, 1999). Therefore understanding intrinsic accuracy limitations in the shape–from–moments estimation problem is important.

The goal of this paper is to develop an optimal and computationally efficient algorithm for estimating convex compact planar regions from noisy observations of their moments. Our approach to the shape–from–moments problem is based on pointwise estimation of *the support function*. It is well–known that the boundary of a planar convex set is completely characterized as the envelope of *the support lines* that graze the set in different directions. The distance between a support line and the origin as function of the angle (direction) is *the support function*. Thus pointwise estimation of the support function leads to a pointwise estimate of the set boundary. Closely related problem of reconstructing a convex set from noisy data on its support function has been considered in Prince and Willsky (1990) and Fisher et.al. (1997). We refer also to Korostelev and Tsybakov (1993) for various models related to estimating sets from noisy data.

The main contributions of this paper are the following. First we develop an estimator of the support function based on noisy measurements of the geometric moments (2). It is shown that the mean squared error of this estimator converges to zero at a very slow logarithmic rate as $\sigma \to 0$. We argue that this rate cannot be essentially improved in the sense of the order. Therefore the shape–from–moments problem is effectively insoluble in practical terms whenever noisy measurements of geometric moments are given. The reason is that the design functions $x^k y^l$, $k, l = 0, 1, \ldots$ are non–orthogonal. Considering the choice of the design functions as a part of our estimation procedure, we develop a method with fast polynomial rate of convergence. In particular, we show the mean squared error of our pointwise estimator converges to zero at the rate $O([\sigma^2 \ln \sigma^{-2}]^{1/\alpha})$ as $\sigma \to 0$, where $\alpha \in [1, 2]$ is a constant depending on the local behavior of the set G in the vicinity of the estimated support value. We establish a lower bound showing that the proposed estimator is near–optimal in order within a logarithmic $\ln(\sigma^{-2})$ factor. We discuss application of the proposed procedure for reconstructing a convex set from noisy Radon data and demonstrate that the same rates of convergence are achieved in this particular inverse problem.

It is interesting to compare our results with results obtained for the problem of recovering edges from indirect observations. The shape–from–moments problem can be considered as a problem of estimating the indicator function of a planar set from noisy moments measurements. In this setup reconstruction methods based on orthogonal series expansions have been developed [see, e.g. Liao and Pawlak (1996, 2002)]. It is important to emphasize, however, that traditional linear methods behave poorly when the function to be estimated has edges. This fact is reflected by slow rates at which the estimation error tends to zero as $\sigma \to 0$. Recently Candés and Donoho (2002) developed a method for recovering bivariate functions with edges from noisy Radon data. The method is based on recently introduced curvlet decomposition of the Radon operator. This technique applied to the problem of estimating the indicator function $\mathbf{1}_G(\cdot,\cdot)$ from noisy Radon data yields an estimator with the mean integrated squared error of the order $O(\sigma^{4/5})$ as $\sigma \to 0$, provided that the boundary of the set G is twice differentiable. In Section 4 we show that in this particular setup for convex G the edge can be estimated with pointwise mean squared error of the order $O([\sigma^2 \ln \sigma^{-2}]^{1/\alpha})$ for some $\alpha \in [1,2]$. This is much faster than the rate indicated above.

The rest of the paper is organized in the following way. In Section 2 we consider the problem of reconstructing a convex set from noisy measurements of its geometric moments. The case of orthogonal design is treated in Section 3. In Section 4 we discuss application of the proposed algorithm to tomographic reconstruction. Section 5 contains the proofs.

2 Reconstruction from geometric moments

Let $\{\mu_{k,l}\}$ be the geometric moments of G given by

$$\mu_{k,l} = \iint_D x^k y^l \mathbf{1}_G(x,y)dx\,dy, \quad k,l = 0,1,\dots\ .$$

The objective is to reconstruct the set G using noisy observations

$$y_{k,l} = \mu_{k,l} + \sigma\varepsilon_{k,l}\,, \quad k,l = 0,1,\dots\ , \tag{5}$$

where $\{\varepsilon_{k,l}\}$ are independent standard gaussian random variables. In what follows we always assume that the origin belongs to the interior of the set G.

It is well–known that the boundary of a convex planar set G can be characterized as an envelope of the support lines $\ell_G(\theta)$ of the set G in directions $\omega = (\cos\theta, \sin\theta)'$, $\theta \in [0, 2\pi)$. The line $\ell_G(\theta)$ is orthogonal to ω and tangent to the set G in ω-direction. *The support function* $\tau = \tau(\theta)$, $\theta \in [0, 2\pi)$ is defined as the distance from the origin to the corresponding support line $\ell_G(\theta)$ [cf. Figure 1]. More formally, the support line $\ell_G(\theta)$ at angle θ for the closed and bounded planar set G is given by

$$\ell_G(\theta) = \{(x,y) : x\cos\theta + y\sin\theta = \tau(\theta)\},$$

where

$$\tau(\theta) = \sup_{(x,y)\in G} \{x\cos\theta + y\sin\theta\}$$

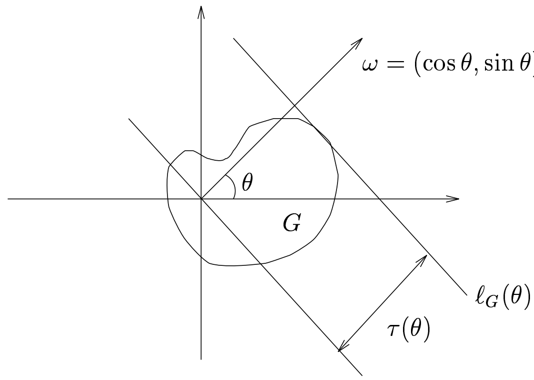

Figure 1: The geometry of support lines

is the support function. We note that the support function $\tau(\cdot)$ takes values in $[0,1]$ for $\theta \in [0, 2\pi)$. In what follows we concentrate on pointwise estimation of the support function $\tau(\cdot)$ of the set G using noisy observations of its moments. We call the value of support function $\tau(\cdot)$ at a single direction given by θ, the support value.

From now on, for the sake of definiteness, we assume that $\theta \in [0,\pi)$ and define the function

$$g_\theta(t) = \iint_D \mathbf{1}_{[t,1]}(x\cos\theta + y\sin\theta)\mathbf{1}_G(x,y)dx\,dy, \quad \text{for } 0 \le t \le 1. \tag{6}$$

If $\theta \in [\pi, 2\pi)$ then we define $g_\theta(\cdot)$ by (6) with $\mathbf{1}_{[t,1]}(\cdot)$ replaced by $\mathbf{1}_{[-1,t]}(\cdot)$ under the integral sign. Clearly, $g_\theta(t)$ is the Lebesgue measure (denoted by $\mathcal{L}\{\cdot\}$) of the intersection of G with the half–plane $\{(x,y) \in D : x\cos\theta + y\sin\theta \ge t\}$:

$$g_\theta(t) = \mathcal{L}\{G_\theta(t)\}, \quad G_\theta(t) := \{(x,y) \in D : x\cos\theta + y\sin\theta \ge t\} \cap G. \tag{7}$$

It follows from (7) that $g_\theta(\cdot) = 0$ for all $t \in (\tau(\theta), 1]$ and grows monotonically as t decreases from $\tau(\theta)$ to zero. This property of $g_\theta(\cdot)$ underlies construction of our estimator.

Let $\{p_n(x)\}_{n=0,1,\dots}$ be the orthonormal Legendre polynomials on $[-1,1]$, and let

$$p_n(x) = \sum_{j=0}^{n} \beta_{n,j} x^j, \quad \text{and} \quad p_n(x) = \sqrt{\frac{2n+1}{2}} P_n(x).$$

Denoting $u = x\cos\theta + y\sin\theta$ and expanding the function $\mathbf{1}_{[t,1]}(\cdot)$ into Fourier series with respect to this orthonormal system we can write $\mathbf{1}_{[t,1]}(u) = \sum_{n=0}^{\infty} a_n p_n(u)$, where for $n \ge 1$

$$\begin{aligned} a_n = a_n(t) = -\int_{-1}^{t} p_n(u)du &= -\sqrt{\frac{2n+1}{2}} \int_{-1}^{t} P_n(u)du \\ &= \frac{1}{\sqrt{4n+2}} \Big[P_{n-1}(t) - P_{n+1}(t) \Big], \end{aligned} \tag{8}$$

and the series converge in $L_2(-1,1)$. Here we used the following well–known properties of the Legendre polynomials [see, e.g., Erdéyi et al. (1953, v. II, Chapter X)]

$$(2n+1)P_n(x) = P'_{n+1}(x) - P'_{n-1}(x), \qquad P_{n+1}(-1) = P_{n-1}(-1), \ \forall n.$$

Then (6) is rewritten as

$$\begin{aligned} g_\theta(t) &= \sum_{n=0}^{\infty} a_n \iint_D p_n(x\cos\theta + y\sin\theta)\mathbf{1}_G(x,y)dxdy \\ &= \sum_{n=0}^{\infty} a_n \sum_{j=0}^{n} \beta_{n,j} \iint_D (x\cos\theta + y\sin\theta)^j \mathbf{1}_G(x,y)dxdy \\ &= \sum_{n=0}^{\infty} a_n \sum_{j=0}^{n} \beta_{n,j} \sum_{m=0}^{j} \binom{j}{m} \cos^m(\theta)\sin^{j-m}(\theta)\mu_{m,j-m} \ . \end{aligned} \tag{9}$$

These considerations lead to the following natural estimator of the function $g_\theta(t)$. We define

$$\hat{g}_\theta^N(t) = \sum_{n=0}^{N} a_n \sum_{j=0}^{n} \beta_{n,j} \sum_{m=0}^{j} \binom{j}{m} \cos^m(\theta)\sin^{j-m}(\theta) y_{m,j-m} \ , \tag{10}$$

where $\{y_{k,l}\}$ are given by (5), and N is a natural number to be chosen.

Theorem 1 *Let G be a convex set in the interior of the closed disc D_{1-h} of the radius $1-h$ centered at the origin. Let $\hat{g}_\theta^*(t)$ be the estimator defined in (10) and associated with*

$$N = N_* := \Big\lfloor \frac{1}{\ln 32}\Big\{\ln\Big(\frac{1}{\sigma^2 h^2}\Big) - \ln\ln\Big(\frac{1}{\sigma^2 h^2}\Big)\Big\}\Big\rfloor. \tag{11}$$

Then for any $\theta \in [0,\pi)$ and σ small enough

$$\sup_{t\in(0,1-h]} \mathbb{E}|\hat{g}_\theta^*(t) - g_\theta(t)|^2 \le C_1\Big[h^2 \ln\Big(\frac{1}{\sigma^2 h^2}\Big)\Big]^{-1},$$

where C_1 is an absolute constant.

Now we define the estimator of the support value $\tau = \tau(\theta)$ at angle $\theta \in [0,\pi)$. For fixed $r = r_\sigma > 0$ let

$$\hat{\tau}(\theta) = \max\{t \in (0, 1-h] : \hat{g}_\theta^*(t) \ge r\}, \tag{12}$$

where $\hat{g}_\theta^*(t)$ is given by (10) and (11). Observe that for small enough σ and $r < \mathcal{L}\{G\}$ the estimate $\hat{\tau}(\theta)$ is well–defined. It follows from (8) and (10) that $\hat{g}_\theta^*(\cdot)$ is a continuous function of t; hence $\hat{g}_\theta^*(\hat{\tau}(\theta)) = r$.

To analyze accuracy of the above estimator we introduce assumptions on the local behavior of the boundary of the set G in the vicinity of the support point $\tau = \tau(\theta)$.

> We say that G belongs to the class $\mathcal{G}_\theta(\alpha, L)$ if there exist positive numbers α, L, and Δ such that
>
> $$g_\theta(t) \ge L|\tau - t|^\alpha, \quad \text{for } \ t \in (\tau - \Delta, \tau). \tag{13}$$

It is important to emphasize that the class $\mathcal{G}_\theta(\alpha, L)$ is defined for a fixed direction $\omega = (\cos\theta, \sin\theta)'$, so that in general constants L, α, and Δ depend on θ. For simplicity we omit this dependence from the notation.

Because $g_\theta(t)$ is the Lebesgue measure of the set $G_\theta(t)$ given by (7), the above condition specifies the rate at which this measure increases as t decreases from $\tau(\theta)$ to zero. It is easily verified that if G is convex then necessarily $1 \le \alpha \le 2$ for any angle θ. Next examples illustrate how parameters α and L of the class $\mathcal{G}_\theta(\alpha, L)$ are related to geometrical properties of the set G.

Examples:

1. Let G be a convex polygon. Then for any direction ω which is not perpendicular to the sides of the polygon, G belongs to class $\mathcal{G}_\theta$ with $\alpha = 2$. The constant L depends in an evident way on the angle between two adjacent sides of the polygon corresponding to the support vertex. This situation corresponds to the minimal increase of the Lebesgue measure of $G_\theta(t)$ as t varies in an open left vicinity of the support value $\tau(\theta)$. If the direction ω is perpendicular to a side of the polygon, then the corresponding support line contains that side. In this case G belongs to $\mathcal{G}_\theta$ with $\alpha = 1$, and we have the maximal increase of $\mathcal{L}\{G_\theta(t)\}$ as t varies in an open left vicinity of $\tau(\theta)$.

2. If G is a circle or an ellipse, then $G \in \mathcal{G}_\theta$ with $\alpha = 3/2$ for any direction $\omega = (\cos\theta, \sin\theta)'$. More generally, let (η, ω) denote the coordinate system on the plane associated with direction $\omega = (\cos\theta, \sin\theta)'$. If in some vicinity of the point $\eta = \eta_0$ where the support value $\tau = \tau(\theta)$ is attained, the boundary ∂G of the set G can be represented as $\{(\eta, \omega) : \omega = \tau - c(\eta - \eta_0)^q\}$ for some $c > 0$ and $q > 1$, then $G \in \mathcal{G}_\theta$ with $\alpha = 1 + 1/q$.

Now we are in position to state the main result of this section.

Theorem 2 *Let conditions of Theorem 1 be fulfilled. Let $\hat{\tau}$ be the estimator associated with $N = N_*$ given by (11) and*

$$r = r_\sigma := \left[\frac{4\ln\ln\left(\frac{1}{\sigma^2 h^2}\right)}{h^2 \ln\left(\frac{1}{\sigma^2 h^2}\right)}\right]^{1/2} .$$

Then for σ small enough

$$\sup_{G \in \mathcal{G}_\theta(\alpha, L)} \mathbb{E}|\hat{\tau}(\theta) - \tau(\theta)|^2 \le C_2 (hL)^{-2/\alpha} \left[\frac{\ln\ln\left(\frac{1}{\sigma^2 h^2}\right)}{\ln\left(\frac{1}{\sigma^2 h^2}\right)}\right]^{1/\alpha} , \tag{14}$$

where C_2 is an absolute constant.

Theorem 2 indicates that the estimator $\hat{\tau}$ converges to the support value $\tau(\theta)$ at a very slow logarithmic rate. In fact, it can be argued that this rate cannot be substantially improved, see remark immediately after Theorem 5 in Section 3. As proofs of the Section 5 indicate, this slow convergence rate is a consequence of the fact that the monomials $x^k y^l$, $k, l = 0, 1, \ldots$ are highly non–orthogonal, and each geometric moment brings a small amount of information about the set to be estimated. It was recognized widely in the literature that even if exact measurements of the moments are available, this non–orthogonality leads to unstable reconstruction algorithms.

3 Reconstruction from Legendre moments

In this section we show that the estimation accuracy can be substantially improved by more careful choice of design functions. Typically in applications involving reconstructing shapes from moments design functions can be selected; geometric and/or complex moments are usually used only for the sake of simplicity and convenience. For discussion of these issues we refer to Milanfar et. al. (1995), Milanfar, Karl and Willsky (1995), and Golub, Milanfar and Varah (1999). We explore the situation where the moments with respect to the Legendre polynomials can be observed with gaussian noise.

As before, we consider the problem of pointwise estimation of the support value $\tau(\theta)$ at a single fixed direction $\omega = (\cos\theta, \sin\theta)'$. Suppose that for given ω the Legendre moments

$$\nu_n = \nu_n(\theta) = \iint_D p_n(x\cos\theta + y\sin\theta)\mathbf{1}_G(x,y)dxdy, \quad n = 0, 1, \dots \tag{15}$$

can be observed with noise, i.e.,

$$y_n(\theta) = \nu_n(\theta) + \sigma\varepsilon_n(\theta), \quad n = 0, 1, \dots, \tag{16}$$

where $\{\varepsilon_n(\theta)\}$ are independent standard gaussian random variables. We construct an estimate of the support function $\tau = \tau(\theta)$ based on observations (16).

With the above notation, considerations similar to those preceding (9) lead to $g_\theta(t) = \sum_{n=0}^{\infty} a_n(t)\nu_n(\theta)$, where $a_n(t)$ are given by (8). For fixed integer N we define

$$\hat{g}_\theta^N(t) = \sum_{n=0}^{N} a_n(t) y_n(\theta). \tag{17}$$

The next statement is obtained as an immediate consequence of Theorem 1.

Theorem 3 *Let G be a convex set in the interior of the closed disc D_{1-h} of the radius $1-h$ centered at the origin. Let $g_\theta(t)$ be given by (17); then for any N and $\theta \in [0, \pi)$*

$$\sup_{t\in(0,1-h]} \mathbb{E}|\hat{g}_\theta^N(t) - g_\theta(t)|^2 \le 2\sigma^2\Big(1 + \frac{\pi}{h^2N}\Big) + \frac{8\pi}{h^2N} \ .$$

The estimator $\hat{\tau}$ of the support value $\tau = \tau(\theta)$ is defined as follows. Fix $N = N_* = [\sigma^{-2}]$, and let $\hat{g}_*(\cdot) = \hat{g}_\theta^{N_*}(\cdot)$. For $r = r_\sigma := 2\sigma\sqrt{\ln(1/\sigma^2)}$ we define

$$\hat{\tau} = \max\{t \in (0, 1-h] : \hat{g}_*(t) \ge r\}. \tag{18}$$

Theorem 4 *Let conditions of Theorem 3 hold, $\alpha \ge 1$ and $\hat{\tau}$ be given by (18). Then for σ small enough*

$$\sup_{G\in\mathcal{G}_\theta(\alpha,L)} \mathbb{E}|\hat{\tau}(\theta) - \tau(\theta)|^2 \le C_3\Big[\frac{\sigma^2}{L^2}\ln\Big(\frac{1}{\sigma^2}\Big)\Big]^{1/\alpha} \ ,$$

where C_3 is an absolute constant.

Proof of Theorem 4 goes along the same lines as the proof of Theorem 2, and therefore it is omitted.

Theorem 4 shows that the rates given in (14) can be substantially improved provided that for a fixed direction moments with respect to the correspondingly rotated Legendre polynomials can be observed. The next statement establishes a lower bound showing that the proposed estimator $\hat{\tau}$ is near–optimal in order up to a logarithmic factor.

Theorem 5 *Let G be a convex set in the interior of the closed disc D_{1-h} of the radius $1-h$ centered at the origin. For any estimator $\hat{\tau}$ of $\tau = \tau(\theta)$ based on observations (15)-(16) and for σ small enough*

$$\sup_{G \in G_\theta(\alpha, L)} \mathbb{E}|\hat{\tau}(\theta) - \tau(\theta)|^2 \geq C_4 \Big(\frac{\sigma^2}{L^2}\Big)^{1/\alpha} \quad ,$$

where C_4 depends on α and h only.

We remark that the lower bound of Theorem 5 remains valid when the moments with respect to any orthonormal system on $[-1, 1]$ rotated correspondingly are observed in (16). The proof of the lower bound exploits equivalence between the gaussain white noise model and the gaussian sequence space model for Fourier coefficients with respect to an orthonormal system of functions. By the same token, gaussian sequence space model with respect to a non–orthonormal system of functions is equivalent to a continuous model with correlated gaussian noise. Using this idea and the same reasoning as in the proof of Theorem 5 one can show that if the geometric moments are observed with gaussian noise then the risk of the pointwise estimation is bounded from below by $O([\ln(\frac{1}{\sigma^2 h^2})]^{-1})$. Hence the upper bound of Theorem 1 cannot be substantially improved.

4 Application to tomography

We consider the problem of reconstructing a convex set G from noisy Radon data given by the white noise model:

$$Y(dt, d\theta) = (R\,\mathbf{1}_G)(t, \theta) + \sigma W(dt, d\theta). \tag{19}$$

Here $W(t, \theta)$ denotes the Wiener sheet on $[-1, 1] \times [0, \pi]$ and $R : L_2(D) \to L_2([-1, 1] \times [0, \pi])$ is the Radon transform,

$$(Rf)(t, \theta) = \iint_D f(x, y) \delta(t - x\cos\theta - y \sin\theta) dx dy,$$

where $\delta(\cdot)$ is the delta–function. The continuous observation model (19) means that for any function $s(\cdot, \cdot) \in L_2([-1, 1] \times [0, \pi])$ we can observe integrals $\iint s(t, \theta)(Rf)(t, \theta) dt\, d\theta$ with zero mean gaussian noise having the variance $\sigma^2 \iint s^2(t, \theta) dt\, d\theta$.

Although many different methods for restoring functions from noisy Radon data have been analyzed in the literature, the focus is usually on estimation of smooth functions [see, e.g., Johnstone and Silverman (1990), Korostelev and Tsybakov (1993) and references therein]. Recently Candés and Donoho (2002) considered the problem of recovering a function which is smooth apart from a discontinuity along a twice differentiable curve on the plane. For the observation model similar to (19) they develop an estimator based on the curvlet decomposition of the Radon operator. Applying that estimator we obtain that the indicator function $\mathbf{1}_G(\cdot,\cdot)$ can be estimated from observations (19) with the mean integrated squared error of the order $O(\sigma^{4/5})$ as $\sigma \to 0$, provided that the boundary of the set G is twice differentiable. It turns out that the edge of a convex set can be estimated with much better accuracy from noisy Radon data. In particular, we demonstrate below that the estimator developed in Section 3 achieves the rate $O([\sigma^2 \ln \sigma^{-2}]^{1/\alpha})$ with $\alpha \in [1,2]$ in the problem of pointwise edge estimation from noisy Radon data.

In practice the data are usually discretely sampled, and the continuous white noise model (19) is only a useful idealization. We assume that discretization with respect to the angle variable θ is performed, i.e. we can observe

$$Y_{\theta_j}(dt) = (R\,\mathbf{1}_G)(t,\theta_j)dt + \sigma W_{\theta_j}(dt) \tag{20}$$

for angles $\theta_j \in [0,\pi]$, $j = 1,\dots,n_\theta$. In what follows we consider the problem of estimating the support function $\tau = \tau(\theta)$ of G at an angle $\theta \in \{\theta_1,\dots,\theta_{n_\theta}\}$ using the data (20).

It follows immediately from the definition of the Radon transform that for any square integrable on $[-1,1]$ function $F(\cdot)$

$$\int_{-1}^{1} (Rf)(t,\theta)F(t)dt = \iint_D f(x,y)F(x\cos\theta + y\sin\theta)dxdy\ . \tag{21}$$

In particular, the choice $F(t) = e^{-i\omega t}$ leads to the well–known Projection Slice Theorem.

Let $F(t) = p_n(t)$, where $p_n(\cdot)$ is the Legendre orthogonal polynomial of the degree n on $[-1,1]$. Then applying (21) for $f(x,y) = \mathbf{1}_G(x,y)$ we obtain from (20) for given $\theta \in \{\theta_1,\dots,\theta_{n_\theta}\}$

$$\begin{aligned} y_n(\theta) &:= \int_{-1}^{1} p_n(t)Y_\theta(dt) \\ &= \int_{-1}^{1} p_n(t)(R\,\mathbf{1}_G)(t,\theta)dt + \sigma\int_{-1}^{1} p_n(t)W_\theta(dt) \\ &= \iint_D p_n(x\cos\theta + y\sin\theta)\mathbf{1}_G(x,y)dxdy + \sigma\varepsilon_n(\theta), \end{aligned}$$

where $\varepsilon_n(\theta)$ is a sequence of independent standard gaussian random variables. This shows that the observation model (15)-(16) is equivalent to (20). An immediate consequence of this equivalence is that the upper bound of Theorem 4 is valid for estimating the support function of the set G from noisy Radon data (20).

5 Proofs

In the proofs below we use well–known properties of the Legendre polynomials; all these facts can be found, e.g., in Natanson (1949, Part 2, Chapter V) and Erdéyi et al. (1953, v. II, Chapter X).

5.1 Proof of Theorem 1

For fixed N we have

$$\begin{aligned}\mathbb{E}|\hat{g}_\theta^N(t)-g_\theta(t)|^2 &= v_N+b_N^2\\ &= \sigma^2\mathbb{E}\Big(\sum_{n=0}^N a_n\sum_{j=0}^n \beta_{n,j}\xi_j\Big)^2\\ &\quad +\Big(\sum_{n=N+1}^\infty a_n\iint_D p_n(x\cos\theta+y\sin\theta)\mathbf{1}_G(x,y)dxdy\Big)^2, \quad (22)\end{aligned}$$

where

$$\xi_j=\xi_j(\theta):=\sum_{m=0}^j\binom{j}{m}\cos^m(\theta)\sin^{j-m}(\theta)\varepsilon_{m,j-m}\ ,\quad j=0,\dots,n\ . \tag{23}$$

First we bound the variance term v_N. To this end we observe that ξ_j, $j=0,\dots,n$ are independent zero mean gaussian random variables with variances

$$\gamma_j^2:=\mathrm{var}\{\xi_j(\theta)\}=\sum_{m=0}^j\binom{j}{m}^2\cos^{2m}(\theta)\sin^{2(j-m)}(\theta)\ .$$

Therefore the variance term v_N can be written in the form $v_N=\sigma^2\mathbf{a}_N'B\Gamma^2B'\mathbf{a}_N$, where $\mathbf{a}_N=(a_0,a_1,\dots,a_N)'$, $\Gamma=\mathrm{diag}(\gamma_0,\dots,\gamma_N)$, and B is the $(N+1)\times(N+1)$ lower triangular matrix with non-zero elements given by

$$B=\begin{bmatrix}\beta_{0,0} & & & & \\ \beta_{1,0} & \beta_{1,1} & & & \\ \beta_{2,0} & \beta_{2,1} & \beta_{2,2} & & \\ \vdots & \vdots & \vdots & & \\ \beta_{N,0} & \beta_{N,1} & \beta_{N,2} & \cdots & \beta_{N,N}\end{bmatrix}.$$

Noting that $\gamma_j^2\le 2^j$ for all $j=0,\dots,n$ we obtain $v_N\le\sigma^2 2^N\|\mathbf{a}_N\|^2\lambda_{\max}[BB']$, where $\lambda_{\max}[\cdot]$ stands for the maximal eigenvalue of a matrix. Because of (8) and the well-known fact that

$$|P_n(t)|\le\frac{1}{h}\sqrt{\frac{\pi}{2n}},\quad \forall t\in[-1+h,1-h],\quad n=1,2,\dots$$

we have

$$\begin{aligned}|a_n| &\le \sqrt{\frac{\pi}{h(4n+2)}}\Big[\frac{1}{\sqrt{2(n+1)}}+\frac{1}{\sqrt{2(n-1)}}\Big]\\ &\le \frac{\sqrt{\pi}}{h\sqrt{(2n+1)(n-1)}}\quad \text{for } n=2,3,\dots\ . \end{aligned}\tag{24}$$

In addition, $a_0=(1-t)/\sqrt{2}$ and $a_1=\sqrt{3/8}(1-t^2)$. Thus,

$$\begin{aligned}\|\mathbf{a}_N\|^2 &\le 2+\frac{\pi}{h^2}\sum_{n=2}^N\frac{1}{(n-1)^2}\le 2+\frac{\pi}{h^2}\Big(1+\int_1^N x^{-2}dx\Big)\\ &\le 2\Big(1+\frac{\pi}{h^2N}\Big)\ .\end{aligned}\tag{25}$$

To bound $\lambda_{\max}[BB']$ we note that $\mathrm{trace}[BB'] = \sum_{n=0}^{N} S_n^2$ where S_n^2 is the sum of squared coefficients of the polynomial $p_n(x)$: $S_n^2 = \sum_{j=0}^{n} \beta_{n,j}^2$. It is well–known that

$$P_n(x) = \frac{1}{2^n} \sum_{m=0}^{[n/2]} (-1)^m \binom{n}{m} \binom{2n-2m}{n} x^{n-2m}$$

where $[\cdot]$ denotes the integer part. Therefore

$$\begin{aligned} S_n^2 &\leq \frac{2n+1}{2} \frac{1}{4^n} \sum_{m=0}^{[n/2]} \left\{ \binom{n}{m} \binom{2n-2m}{n} \right\}^2 \\ &\leq \frac{2n+1}{2} \frac{1}{4^n} \left\{ \binom{2n}{n} \right\}^2 (2^n)^2 \leq \frac{4^{2n}}{\pi} \Big(1 + \frac{1}{2n}\Big), \end{aligned}$$

where we have used the fact that $(2n)!(n!)^{-2} \leq 4^n (n\pi)^{-1/2}$ [see Natanson (1949, p. 666)]. Therefore $\mathrm{trace}[BB'] = \sum_{n=0}^{N} S_n^2 \leq 4^{2N}(10\pi)^{-1}$ and combining this inequality with (25) we finally obtain

$$v_N \leq \bar{v}_N := \frac{2^{5N}\sigma^2}{5\pi} \Big(1 + \frac{\pi}{h^2 N}\Big) . \tag{26}$$

Now we bound the bias term in (22). The orthogonal transformation of the coordinate system results in

$$\begin{aligned} c_n &:= \iint_D p_n(x\cos\theta + y\sin\theta) \mathbf{1}_G(x,y)\,dxdy \\ &= \int_{-1}^{1} p_n(u) \int_{\varphi_1(u)}^{\varphi_2(u)} \mathbf{1}_G(u,w)\,dwdu \\ &= \int_{-1}^{1} p_n(u)[\varphi_2(u) - \varphi_1(u)]du, \end{aligned} \tag{27}$$

where $u = x\cos\theta + y\sin\theta$, $w = -x\sin\theta + y\cos\theta$, and $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$ are the w–coordinates of the intersection points of the lines $u = \mathrm{const}$ with the boundary of G. We note that the function $\varphi_2(\cdot) - \varphi_1(\cdot)$ is defined on $[-1,1]$, takes values in $[0,2]$ and is continuous because G is a convex simply connected set. Therefore $\varphi_2(\cdot) - \varphi_1(\cdot)$ belongs to $L_2(-1,1)$, and c_n in (27) is nothing but the n-th Fourier coefficient of this function with respect to the Legendre orthonormal system on $L_2(-1,1)$; hence by the Parseval formula

$$\sum_{n=0}^{\infty} c_n^2 = \int_{-1}^{1} [\varphi_2(u) - \varphi_1(u)]^2 du \leq 8.$$

This along with (22) and (24) yields

$$b_N^2 \leq \Big(\sum_{n=N+1}^{\infty} a_n c_n\Big)^2 \leq 8 \sum_{n=N+1}^{\infty} a_n^2 \leq \frac{8\pi}{h^2 N} := \bar{b}_N^2 . \tag{28}$$

Combining (26), (28) and (11) we complete the proof. ■

5.2 Proof of Theorem 2

First we prove an auxiliary lemma. Denote

$$X_N(t) = \sigma \sum_{n=0}^{N} a_n(t) \sum_{j=0}^{n} \beta_{n,j}\xi_j, \quad 0 \le t \le 1-h, \tag{29}$$

where $a_n = a_n(t)$, $n = 1, 2, \ldots$ and $\xi_j = \xi_j(\theta)$ are given by (8) and (23) respectively. We note that $\{X_N(\cdot)\}$ is a zero mean gaussian process with continuous sample paths, and

$$\sup_{t\in[0,1-h]} \mathbb{E}|X_N(t)|^2 = v_N \le \bar{v}_N < \infty,$$

where $\bar{v}_N = (5\pi)^{-1}2^{5N}\sigma^2(1+\pi h^{-2}N^{-1})$ [cf. the proof of Theorem 1]. In the sequel we write v_* and $\bar{v}_*$ for v_{N_*} and $\bar{v}_{N_*}$ respectively, where N_* is given by (11).

Lemma 1 *There exists an absolute constant c_1 such that for fixed N and all $\delta \ge 2\sqrt{\bar{v}_N}$*

$$\mathbb{P}\Big\{ \sup_{t\in[0,1-h]} |X_N(t)| \ge \delta \Big\} \le c_1 N \sqrt{\frac{\bar{v}_N}{v_N}} \exp\Big\{ -\frac{\delta^2}{2v_N} \Big\} . \tag{30}$$

In particular, if $N = N_*$ and σ is small enough then for $\delta = \sqrt{2\varkappa \bar{v}_* \ln(1/\bar{v}_*)}$ with $\varkappa \ge 1$ we have

$$\mathbb{P}\Big\{ \sup_{t\in[0,1-h]} |X_{N_*}(t)| \ge \sqrt{2\varkappa\bar{v}_* \ln\Big(\frac{1}{\bar{v}_*}\Big)} \Big\} \le c_1 N_* \bar{v}_*^{\varkappa} . \tag{31}$$

Proof The proof is based on Theorem 2.4 from Talagrand (1994). Below we use the notation introduced in the proof of Theorem 1. We have for $0 \le s < t \le 1-h$

$$\begin{aligned} \rho^2(X_N(s), X_N(t)) &:= \mathbb{E}[X_N(s) - X_N(t)]^2 \\ &= \sigma^2[\mathbf{a}_N(s) - \mathbf{a}_N(t)]' B\Gamma^2 B' [\mathbf{a}_N(s) - \mathbf{a}_N(t)] \\ &\le \sigma^2 2^N \|\mathbf{a}_N(s) - \mathbf{a}_N(t)\|^2 \lambda_{\max}[BB']. \end{aligned}$$

As it was shown in the proof of Theorem 1, $\sigma^2 2^N \lambda_{\max}[BB'] \le \bar{v}_N$. Moreover, by (8)

$$\begin{aligned} \|\mathbf{a}_N(s) - \mathbf{a}_N(t)\|^2 &= \sum_{n=0}^{N} |a_n(s) - a_n(t)|^2 \\ &\le \sum_{n=0}^{N} \frac{2n+1}{2} \Big| \int_s^t P_n(t)dt \Big|^2 \le |t-s|^2 N^2 . \end{aligned}$$

Therefore $\rho^2(X_N(s), X_N(t)) \le N^2|t-s|^2\bar{v}_N$, and the minimal number of balls of the radius ε (with respect to the semi–norm ρ) covering the index set $[0, 1-h]$ does not exceed $N\varepsilon^{-1}\sqrt{\bar{v}_N}$, for any $\varepsilon \in (0, \sqrt{\bar{v}_N})$. Applying Theorem 2.4 from Talagrand (1994) we obtain that for all $\delta \ge 2\sqrt{\bar{v}_N}$

$$\mathbb{P}\Big\{ \sup_{t\in[0,1-h]} |X_N(t)| \ge \delta \Big\} \le c_1 N \sqrt{\frac{\bar{v}_N}{v_N}} \exp\Big\{ -\frac{\delta^2}{2v_N} \Big\} ,$$

which completes the proof of (30).

To derive (31) we set $N = N_*$ in (30) and choose $\delta = \sqrt{2\varkappa \bar{v}_* \ln(1/\bar{v}_*)}$ with $\varkappa \geq 1$. For σ small enough we have

$$\begin{aligned}
\mathbb{P}\Big\{\sup_{t\in[0,1-h]} |X_{N_*}(t)| \geq \sqrt{2\varkappa \bar{v}_* \ln(1/\bar{v}_*)}\Big\} &\leq \mathbb{P}\Big\{\sup_{t\in[0,1-h]} |X_{N_*}(t)| \geq \sqrt{2\varkappa v_* \ln(1/v_*)}\Big\} \\
&\leq c_1 N_* \bar{v}_*^{\varkappa} .
\end{aligned}$$

The lemma is proved. ■

Proof of Theorem 2 We write

$$\begin{aligned}
\mathbb{E}|\hat{\tau} - \tau|^2 &= I_1 + I_2 \\
&:= \mathbb{E}\Big[|\hat{\tau} - \tau|^2 \mathbf{1}\{\hat{\tau} \leq \tau\}\Big] + \mathbb{E}\Big[|\hat{\tau} - \tau|^2 \mathbf{1}\{\hat{\tau} > \tau\}\Big]. \qquad (32)
\end{aligned}$$

We bound I_1 and I_2 separately.

First we bound I_1. In view of Theorem 1, $\hat{g}_\theta^*(t)$ converges in probability to $g_\theta(t)$ for any fixed θ and t as $\sigma \to 0$. This implies that $\hat{\tau}$ converges in probability to τ as $\sigma \to 0$ provided $r = r_\sigma \to 0$. Therefore without loss of generality we can assume that $\hat{\tau} \in (\tau - \Delta, \tau)$ for σ small enough. Under this assumption we have from (13)

$$\begin{aligned}
|\hat{\tau} - \tau|^2 &\leq L^{-2/\alpha} |g_\theta(\hat{\tau})|^{2/\alpha} \\
&\leq L^{-2/\alpha} \Big(|g_\theta(\hat{\tau}) - \hat{g}_\theta^*(\hat{\tau})| + |\hat{g}_\theta^*(\hat{\tau})|\Big)^{2/\alpha} \\
&= L^{-2/\alpha} \Big(|g_\theta(\hat{\tau}) - \hat{g}_\theta^*(\hat{\tau})| + r\Big)^{2/\alpha} .
\end{aligned}$$

Applying Theorem 1 we obtain

$$\begin{aligned}
I_1 &\leq 2^{1/\alpha} L^{-2/\alpha} \Big(\mathbb{E}|g_\theta(\hat{\tau}) - \hat{g}_\theta^*(\hat{\tau})|^2 + r^2\Big)^{1/\alpha} \\
&\leq 2^{1/\alpha} L^{-2/\alpha} \Big\{ r^{2/\alpha} + C_1^{1/\alpha} \Big[h^2 \ln\Big(\frac{1}{\sigma^2 h^2}\Big)\Big]^{-1/\alpha} \Big\} . \qquad (33)
\end{aligned}$$

To bound I_2 we note that

$$\begin{aligned}
I_2 &\leq 4\,\mathbb{P}\{\hat{\tau} > \tau\} \leq 4\,\mathbb{P}\{\hat{g}_\theta^*(t) \geq r \text{ for some } t \in (\tau, 1-h]\} \\
&= 4\,\mathbb{P}\Big\{\sup_{t\in(\tau,1-h]} \hat{g}_\theta^*(t) \geq r\Big\} \\
&\leq 4\,\mathbb{P}\Big\{\sup_{t\in[0,1-h]} |X_{N_*}(t)| \geq r - |\bar{b}_{N_*}|\Big\},
\end{aligned}$$

where $X_N(t)$ is defined (29), and $\bar{b}_N$ is given by (28). Clearly, for σ small enough $r - |\bar{b}_{N_*}| \geq r/2$. By Lemma 1 for our choice of r we have

$$\begin{aligned}
I_2 &\leq 4\,\mathbb{P}\Big\{\sup_{t\in[0,1-h]} |X_{N_*}(t)| \geq r/2\Big\} \leq c_2 N_* \bar{v}_*^{\varkappa} \\
&\leq c_3 \ln\Big(\frac{1}{\sigma^2 h^2}\Big) \Big[h^2 \ln\Big(\frac{1}{\sigma^2 h^2}\Big)\Big]^{-\varkappa} . \qquad (34)
\end{aligned}$$

Combining (34), (33), and (32), and taking into account that I_1 dominates I_2 for $\varkappa = 2$ and σ small enough, we complete the proof. ■

5.3 Proof of Theorem 5

Without loss of generality we assume that $\theta=0$. Let G_0 be a convex set in the interior of the unit disc with support value $\tau_0=\tau_0(0)$ in the direction associated with angle $\theta=0$. Denote

$$g_0(t):=g_{G_0}(t)=\iint_D \mathbf{1}_{[t,1]}(x)\mathbf{1}_{G_0}(x,y)dxdy$$

and assume that for some $\Delta>0$

$$g_0(t)=L|t-\tau_0|^\alpha,\quad \text{for } t\in(\tau_0-\Delta,\tau_0).$$

In addition, let

$$\nu_{n,0}=\iint_D p_n(x)\mathbf{1}_{G_0}(x,y)dxdy,\quad n=0,1,\ldots$$

denote the Legendre moments of G_0 associated with the angle $\theta=0$. It follows from (9) that $g_0(t)=\sum_{n=0}^\infty a_n(t)\nu_{n,0}$, where functions $a_n(t)$ are given by (8). It is important to emphasize here that $g_0(\cdot)$ depends on the underlying set G_0 only through the moments $\nu_{n,0}$.

Fix $\delta\in(0,\Delta)$, and let G_δ denote the translate of G_0 by vector $(-\delta,0)'$: $G_\delta=G_0-(\delta,0)'$. Clearly, support value τ_δ of the set G_δ in the direction $\theta=0$ is $\tau_\delta=\tau_\delta(0)=\tau_0-\delta$, and $g_\delta(t):=g_{G_\delta}(t)=g_0(t+\delta)$. In addition, we can write $g_\delta(t)=\sum_{n=0}^\infty a_n(t)\nu_{n,\delta}$, where

$$\nu_{n,\delta}=\iint_D p_n(x)\mathbf{1}_{G_\delta}(x,y)dxdy,\quad n=0,1,\ldots\ .$$

Using the aforementioned definitions we obtain $g_0(\tau_0-\delta)-g_\delta(\tau_0-\delta)=g_0(\tau_0-\delta)=L\delta^\alpha$, and therefore

$$g_0(\tau_0-\delta)-g_\delta(\tau_0-\delta)=\sum_{n=0}^\infty a_n(\tau_0-\delta)[\nu_{n,0}-\nu_{n,\delta}]=L\delta^\alpha. \tag{35}$$

Now we evaluate the Kullback–Leibler distance $\mathcal{K}(\cdot,\cdot)$ between the probability measures Q_0 and Q_δ corresponding to the observations (16) associated with sets G_0 and G_δ. For this purpose we note that by definition

$$\begin{aligned}\nu_{n,0} &= \int_{-1}^1 p_n(x)\left[\overline{\varphi}_0(x)-\underline{\varphi}_0(x)\right]dx\\ \nu_{n,\delta} &= \int_{-1}^1 p_n(x)\left[\overline{\varphi}_\delta(x)-\underline{\varphi}_\delta(x)\right]dx\end{aligned}$$

where $\overline{\varphi}_0,\underline{\varphi}_0$, and $\overline{\varphi}_\delta,\underline{\varphi}_\delta$ are the y–coordinates of the intersection points of the lines $x=$ const with the boundary of G_0 and G_δ respectively. Hence $\{\nu_{n,0}\}$ and $\{\nu_{n,\delta}\}$ are noting but the Fourier coefficients of the functions $\psi_0=\overline{\varphi}_0-\underline{\varphi}_0$ and $\psi_\delta=\overline{\varphi}_\delta-\underline{\varphi}_\delta$ with respect to the Legendre orthonormal system on $[-1,1]$. Therefore, by equivalence of the model (16) and the standard white noise model, we obtain

$$\mathcal{K}(Q_0,Q_\delta)=\frac{1}{2\sigma^2}\sum_{n=0}^\infty|\nu_{n,0}-\nu_{n,\delta}|^2\ . \tag{36}$$

Now we note that $\mathcal{K}(Q_0,Q_\delta)\le c_4\sigma^{-2}L^2\delta^{2\alpha}$, where c_4 depends on h only. This follows from the fact that the norm of the sequence $\{a_n(\tau_0-\delta)\}$ is bounded away from zero for any fixed

δ, and the maximal value of the Kullback–Leibler distance given by (36) under restriction (35) equals $L^2\delta^{2\alpha}[2\sigma^2\sum_{n=0}^{\infty}a_n^2(\tau_0-\delta)]^{-1}$. Therefore choosing δ so that $\sigma^{-2}L^2\delta^{2\alpha}\asymp O(1)$ [or, equivalently, $\delta\asymp O(1)(\sigma/L)^{1/\alpha}$] as $\sigma\to 0$, we obtain that the probability of the error in distinguishing between the sets G_0 and G_δ on the basis of observations (16) is of the order $O(1)$. This completes the proof. ■

References

[1] Candès, E. J, and Donoho, D. L. (2002). Recovering edges in ill–posed inverse problems: optimality of curvlet frames. *Ann. Statist.* **30**, 784–842.

[2] Elad, M., Milanfar, M. and Golub, G. H. (2002). Shape from moments – an estimation theory perspective. Submitted to *IEEE Trans. Signal Processing.*

[3] Erdélyi, A., Magnus, W., Oberhettinger, F. and Tricomi F. G. (1953). *Higher transcendental functions.* Vol. II. Based, in part, on notes left by Harry Bateman. McGraw-Hill, New York.

[4] Fisher, N. I., Hall, P., Turlach, B.A. and Watson, G.S. (1997). On the estimation of a convex set from noisy data on its support function. *J. Amer. Statist. Assoc.* **92**, 84–91.

[5] Golub, G. H., Milanfar, P. and Varah, J. (1999). A stable numerical method for inverting shape from moments. *SIAM J. Sci. Comput.* **21**, 1222–1243.

[6] Gustafsson, B., He, Chiyu, Milanfar, P. and Putinar, M. (2000). Reconstructing planar domains from their moments. *Inverse Problems* **16**, 1053–1070.

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

[8] Korostelev, A. and Tsybakov, A. (1993). *Minimax Theory of Image Reconstruction.* Lecture Notes in Statist. **82**. New York, Springer.

[9] Liao, S. X. and Pawlak, M. (1996). On image analysis by moments. *IEEE Trans. Pattern Anal. Machine Intell.* **18**, 254–266.

[10] Liao, S. X. and Pawlak, M. (2002). On the recovery of a function on a circular domain. *IEEE Trans. Inform. Theory* **48**, 2736–2753.

[11] Milanfar, P., Karl, W. C. and Willsky, A. S. (1996). A moment–based variational approach to tomographic reconstruction. *IEEE Trans. Image Processing* **5**, 459–470.

[12] Milanfar, P., Verghese, G. C., Karl, W. C. and Willsky, A. S. (1995). Reconstructing polygons from moments with connections to array processing, *IEEE Trans. Signal Processing* **43**, 432–443.

[13] Natanson, I. P. (1949). *Constructive Function Theory.* Gostechizdat, Moscow (in Russian).

[14] Prince, J.L. and Willsky, A.S. (1990). Reconstructing convex sets from support line measurements. *IEEE Trans. Pattern Anal. Machine Intell.* **12**, 377–389.

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