METHODS FOR TRACKING SUPPORT BOUNDARIES WITH CORNERS

Ming-Yen Cheng[1,2] Peter Hall[2]

ABSTRACT. In a range of practical problems the boundary of the support of a bivariate distribution is of interest, for example where it describes a limit to efficiency or performance, or where it determines the physical extremities of a spatially distributed population in forestry, marine science, medicine, meteorology or geology. We suggest a tracking-based method for estimating a support boundary when it is composed of a finite number of smooth curves, meeting together at corners. The smooth parts of the boundary are assumed to have continuously turning tangents and bounded curvature, and the corners are not allowed to be infinitely sharp; that is, the angle between the two tangents should not equal π. In other respects, however, the boundary may be quite general. In particular it need not be uniquely defined in Cartesian coordinates, its corners my be either concave or convex, and its smooth parts may be neither concave nor convex. Tracking methods are well suited to such generalities, and they also have the advantage of requiring relatively small amounts of computation. It is shown that they achieve optimal convergence rates, in the sense of uniform approximation.

KEYWORDS. Bandwidth, boundary, corner, curvature, frontier, kernel method, local linear, nonparametric curve estimation, support.

SHORT TITLE. Boundary tracking.

[1] Department of Mathematics, National Taiwan University, Taipei 106, Taiwan

[2] Centre for Mathematics and its Applications, Australian National University, Canberra, ACT 0200, Australia

1. Introduction. In standard problems of univariate nonparametric curve estimation, for example density estimation or regression, one usually constructs the estimator by starting at one end of the real line and moving steadily towards the other, until the curve estimate has been traced out. Of course, in the discrete environment of computation one moves in steps rather than in the continuum, but nevertheless the estimate is calculated in a one-dimensional setting, not a two-dimensional one. This results in computational savings.

In the present paper we propose a similar one-dimensional approach to estimating the boundary of the support of a bivariate distribution. The aim is to estimate the boundary by steadily following a univariate track generated by the estimate itself. The calculations involve reaching a bandwidth radius into the plane from the current point estimate, gathering the data within that radius, and using this information to compute the next point estimate. Thus, all the calculations are confined to data that lie in a tube of width equal to twice the bandwidth, whose axis is the curve estimate. There are consequent computational savings, relative to methods that require the data to be analysed in a larger region of the plane. Moreover, the method is coordinate-independent, and in particular may be applied to curves that cannot be represented, in a Cartesian coordinate system (x, y), in the form $y = g(x)$ for a single-valued function g.

Even when the boundary is smoothly curved, and that assumption is exploited when constructing the estimator, the tracking problem is complex because of the difficulty of estimating support-boundary tangents. The case of a boundary with corners is substantially more difficult, since a decision about where the corners are located has to be made from information gained through tracking the boundary into the corner, and from data within a bandwidth of the corner. The algorithm necessarily involves decisions about using left or right smooths in different places along the boundary estimate, and switching between them. Furthermore, a purely tracking method should not involve backtracking and recomputing the curve estimate after it was found that a corner had been mistakenly omitted.

If we are tracking the boundary by circumnavigating it in a clockwise direction then it is necessary to use a right-hand smooth on leaving a corner, and switch to a left smooth by the time we approach the next corner. Between corners we should use information in both left and right smooths; and the transition from one type of smooth to another should be achieved gently, without introducing jumps that could by misconstrued as additional corners.

The method that we suggest achieves these goals, is practicable for implementation, and enjoys theoretically optimal convergence rates. It is based on kernel methods (e.g. Wand and Jones, 1992), implemented in a very nonstandard way. In particular, it uses extrapolation methods to reach into the corners, where data are often sparse. No conditions are imposed on the convexity, or otherwise, of either the smooth parts of the curve or the corners.

Our methods can be generalised to $d \geq 3$ dimensions, where they produce estimators with the optimal convergence rate, $(\nu^{-1} \log \nu)^{2/(d+1)}$. However, the case $d \geq 3$ is difficult to motivate, since the advantages of tracking methods are significantly reduced if one is estimating a surface (for example), rather than a curve. One cannot simply track a surface from its beginning to its end. Instead, a zig zag path must be constructed, moving backwards and forwards across the boundary, constructing an approximation based on polygons. The procedure is more cumbersome, and less attractive, than its analogue for $d = 2$.

A variety of methods is used for solving jump regression problems, in the absence of corners. See, for example, Müller and Song (1994), O'Sullivan and Qian (1994), Qiu (1997, 1998, 2002, 2004) and Qiu and Yandell (1997). Tracking methods are suggested by Hall and Rau (2000) and Hall, Peng and Rau (2001). All these approaches are in spatial or multivariate settings. Qiu (2005, Chaps. 4, 5) gives a thorough review of this work, and also (Chap. 3) of contributions to the related problem of one-dimensional change-point analysis.

There is an extensive literature on non-tracking methods for boundary-support estimation when the curve contains no corners. A significant part of it is in the area of econometrics, where the boundaries are often interpreted as "production frontiers". See for example work of Aigner, Lovell and Schmidt (1977), who discuss parametric methods, and Kneip and Simar (1996), Seiford (1996), Kneip, Park and Simar (1998) and Park, Sickles and Simar (1998), who address nonparametric approaches. Many nonparametric techniques are based on enveloping the data in some sense, and include "data envelopment analysis" (Farrell, 1957) and the "free disposal hull" (Deprins, Simar and Tulkens, 1984). Theoretical performance (including convergence rates) and numerical properties of these and other methods, in the context of statistics rather than econometrics, have been investigated by Ripley and Rasson (1977), Hartigan (1987), Carlstein and Krishnamoorthy (1992), Korostelev and Tsybakov (1993a), Rudemo and Stryhn (1994), Härdle, Park and Tsybakov (1995), Korostelev, Simar and Tsybakov (1995a,b), Mammen and Tsybakov (1995), Hall, Nussbaum and Stern (1997), Hall, Park and Stern (1998), Donoho (1999), Gijbels, Mammen, Park and Simar (1999) and Baíllo and Cuevas (2001). Optimality theory developed in the context of image analysis, and initiated by Korostelev and Tsybakov (1993b,c), is also relevant. Some change-point methods, for example those proposed by Deshayes and Picard (1981) and Picard (1985), are related.

Qiu (2005, Chap. 6) gives an excellent survey of edge detection and estimation from an image-processing viewpoint. There, methodologies vary from early techniques based on differencing, to recent approaches founded on wavelets and wedgelets (e.g. Donoho, 1999). Some of the mathematical work of this type, and some of the work discussed in the previous paragraph, permits the boundary to have tangent discontinuities. For example, if the boundary satisfies a Hölder condition, but does not have a derivative, it may have corners. Part of the novelty of the work in the present paper is that the methodology makes explicit use of smoothness of the boundary between corners, which are taken to be separated and only finite in number. As a result, convergence rates are faster than they would be if only a Hölder condition were assumed; there, infinitely many arbitrarily-close discontinuities can be present.

2. Methodology

2.1. Overview. Assume we observe data $\mathcal{X} = \{X_1, X_2, \ldots\}$ from a realisation of a point process in the plane. Let $\partial\mathcal{S}$ denote the boundary of the support of the intensity function for $\mathcal{X}$. We shall refer to the points of $\mathcal{X}$ as lying "below" $\partial\mathcal{S}$. The boundary will be traced in a clockwise direction, and so "below" may equivalently be thought of as lying to the right of the direction of travel, although we shall use "left" and "right" for another purpose. The notion of a short line segment that has no points above it is intuitively clear in many cases. More generally, section 3.2 will give a formal definition and discuss the effects of stochastic errors in determining whether the segment has no points above it.

Our boundary estimator is piecewise linear, and in particular consists of line segments joining adjacent estimators $\widehat{Q}_j$ of points on $\partial\mathcal{S}$, indexed in such a manner that we move around $\partial\mathcal{S}$ in a clockwise sense. We pass from $\widehat{Q}_j$ to $\widehat{Q}_{j+1}$ by first moving to a preliminary point $\widehat{Q}'_{j+1}$, calculated by fitting either a left smooth or a right smooth to the boundary at $\widehat{Q}_j$; and then we refine $\widehat{Q}'_{j+1}$ to $\widehat{Q}_{j+1}$ by fitting two short line segments to data in the vicinity of $\widehat{Q}'_{j+1}$.

This procedure by itself produces a boundary estimator that tends to cut across the corners, however, rather than reach into them. That is, in the vicinity of a corner the sequence $\widehat{Q}_j, \widehat{Q}_{j+1}, \ldots$ generally slips from one side of the corner to the other, by passing inside the boundary. As a result, this simple form of the boundary estimator does not enjoy the desired level of accuracy. To overcome this problem we suggest a threshold technique for deciding when the sequence $\{\widehat{Q}_j\}$ has cut a corner. We discard a subsequence that cuts a corner, close up the remaining members of the sequence, and estimate the corner by extrapolating to it from points $\widehat{Q}_j$ that lie on either side of the discarded sequence. These operations are conducted completely sequentially, and in particular do not involve drawing the boundary and then erasing part of it. Our algorithm tells the curve estimator unambiguously when to mark time, i.e. to stop confirming boundary points $\widehat{Q}_j$, and when to start confirming them again.

Next we give an overview of the methods for calculating $\widehat{Q}_j$ from $\widehat{Q}'_j$. Starting from a preliminary approximation Q' (in particular, $\widehat{Q}'_j$) to a point on $\partial\mathcal{S}$, we compute a refinement Q (in particular, $\widehat{Q}_j$) in two stages. First we construct "rough" approximations $\widehat{V}^{\mathrm{L}}$ and $\widehat{V}^{\mathrm{R}}$ to points on $\partial\mathcal{S}$; the superscripts L and R indicate that they quantities are constructed to the left- or right-hand sides, respectively, of the current position. Estimates of the tangent angles at these points are denoted by $\widehat{\omega}^{\mathrm{L}}$ and $\widehat{\omega}^{\mathrm{R}}$, respectively. Next we smooth $\widehat{V}^{\mathrm{L}}$, $\widehat{V}^{\mathrm{R}}$, $\widehat{\omega}^{\mathrm{L}}$ and $\widehat{\omega}^{\mathrm{R}}$ to $\widehat{W}^{\mathrm{L}}$, $\widehat{W}^{\mathrm{R}}$, $\hat{\theta}^{\mathrm{L}}$ and $\hat{\theta}^{\mathrm{R}}$, using kernel techniques. Depending on whether Q' is calculated from on the left- or right-hand side, we take either $\widehat{W}^{\mathrm{L}}$ or $\widehat{W}^{\mathrm{R}}$ to be the refined version Q of Q'. To construct the next version of Q', employing a method for switching from the left to the right, or vice versa, we move a short way from the current version of Q, travelling a fixed distance in the direction of either $\hat{\theta}^{\mathrm{L}}$ or $\hat{\theta}^{\mathrm{R}}$.

As can be seen from this discussion, "handedness" is important. If, when we move from $Q = \widehat{Q}_j$, we travel in the direction of $\hat{\theta}^{\mathrm{L}}_j = \hat{\theta}^{\mathrm{L}}$, we say we are using a left smooth; and if the direction is that of $\hat{\theta}^{\mathrm{R}}_j = \hat{\theta}^{\mathrm{R}}$, we are using a right smooth. We should use a left smooth as we approach a corner (travelling around the boundary in a clockwise direction), but we must change to a right smooth after leaving the corner. And, before reaching the next corner, we should switch again, back to a left smooth. We use a threshold argument, similar to that for deciding when corners are present, to make these parity changes.

Technical details are needed in order to fully specify the procedure. They include the methods for calculating the rough approximations $\widehat{V}^{\mathrm{L}}$, $\widehat{V}^{\mathrm{R}}$, $\widehat{\omega}^{\mathrm{L}}$ and $\widehat{\omega}^{\mathrm{R}}$, and the methods for obtaining the smoothed forms $\widehat{W}^{\mathrm{L}}$, $\widehat{W}^{\mathrm{R}}$, $\hat{\theta}^{\mathrm{L}}$, $\hat{\theta}^{\mathrm{R}}$. These will be give in sections 2.2 and 2.3, respectively. In section 2.4 we shall give the full algorithm, referring back to sections 2.2 and 2.3 for concise definitions.

2.2. Rough estimators $\widehat{V}^{\mathrm{L}}$, $\widehat{V}^{\mathrm{R}}$, $\widehat{\omega}^{\mathrm{L}}$, $\widehat{\omega}^{\mathrm{R}}$. To locate a starting point Q', we lay a line, $\mathcal{L}$ say, across the spatial region, and, as we move along $\mathcal{L}$, conduct a sequence of tests for a discontinuity. In this manner we determine a point Q' that approximates a place where $\mathcal{L}$ cuts $\partial\mathcal{S}$. A simple difference-based method suffices; we do not require the starting point to be particularly accurate. If the tests indicate that $\partial\mathcal{S}$ and $\mathcal{L}$ do not intersect, we draw another line and try again. In practice, $\mathcal{L}$ is often determined from prior belief as to the location of $\partial\mathcal{S}$. See Hall and Rau (2000, section 2) for further discussion.

Let $\mathcal{M}^{\mathrm{L}}$ and $\mathcal{M}^{\mathrm{R}}$ denote line segments of length $2h$ lying to the left and right of $\mathcal{L}$, with their right- and left-hand ends, respectively, located at Q'. (Here h denotes a bandwidth, and will be chosen to decrease to zero as the intensity of the point process increases.) Each line segment is placed so that the acute angle it makes to $\mathcal{L}$ exceeds a small, given value $\Delta \in (0, \pi/2)$, no point of $\mathcal{X}$ lies above it, and at least one point of $\mathcal{X}$ lies on it. Let u^{L} and u^{R} denote the midpoints of $\mathcal{M}^{\mathrm{L}}$ and $\mathcal{M}^{\mathrm{R}}$, respectively. Subject to these constraints, choose $\widehat{\mathcal{M}}^{\mathrm{L}} = \mathcal{M}^{\mathrm{L}}$ and $\widehat{\mathcal{M}}^{\mathrm{R}} = \mathcal{M}^{\mathrm{R}}$ so as to minimise kernel-weighted distances to data values:

$$\begin{aligned} &\sum\nolimits_i \|X_i - Y_i^{\mathrm{L}}\|\, K(\|u^{\mathrm{L}} - Y_i^{\mathrm{L}}\|/h)\, K(\|u^{\mathrm{L}} - X_i\|/2h)\,, \\ &\sum\nolimits_i \|X_i - Y_i^{\mathrm{R}}\|\, K(\|u^{\mathrm{R}} - Y_i^{\mathrm{R}}\|/h)\, K(\|u^{\mathrm{R}} - X_i\|/2h)\,, \end{aligned} \tag{2.1}$$

respectively, where Y_i^{L} and Y_i^{R} denote the projections of X_i onto the infinite lines of which $\mathcal{M}^{\mathrm{L}}$ and $\mathcal{M}^{\mathrm{R}}$ form respective parts, and K is a smooth, nonincreasing, nonnegative function on $[0, \infty)$ satisfying $K(0) = 1$, $K(1) = 0$ and $K > 0$ on $[0, 1)$. Details of finding the numerical solution to the minimisation problem (2.1) are given in section 4.

Write $\widehat{V}^{\mathrm{L}}$ and $\widehat{V}^{\mathrm{R}}$ for the points at which $\widehat{\mathcal{M}}^{\mathrm{L}}$ and $\widehat{\mathcal{M}}^{\mathrm{R}}$, respectively, intersect $\mathcal{L}$. The angles $\widehat{\omega}^{\mathrm{L}}$ and $\widehat{\omega}^{\mathrm{R}}$ made by $\widehat{\mathcal{M}}^{\mathrm{L}}$ and $\widehat{\mathcal{M}}^{\mathrm{R}}$ to a given direction (for example, to the positive x-axis of a Cartesian system) are our "rough" estimators of the orientations of boundary tangent lines to the left and right, respectively, of Q'. For some realisations, and some configurations of $\mathcal{L}$, $\widehat{\omega}^{\mathrm{L}}$ and $\widehat{\omega}^{\mathrm{R}}$ will not be well defined, but this seldom causes difficulties; see section 3.3 for discussion.

The form of (2.1) requires explanation. Omitting the second kernel weight factor, expressed in terms of $\|u^{\mathrm{L}} - X_i\|$ or $\|u^{\mathrm{R}} - X_i\|$, produces a criterion that can give high weight to distances $\|X_i - Y_i^{\mathrm{L}}\|$ or $\|X_i - Y_i^{\mathrm{R}}\|$ which involve points X_i that are a long way from $\mathcal{M}^{\mathrm{L}}$ or $\mathcal{M}^{\mathrm{R}}$, respectively. Moreover, replacing $2h$ by h in the second kernel weight can result in tangent estimates that tend towards being parallel to $\mathcal{L}$. The bandwidth multiplier 2 is somewhat arbitrary, and in asymptotic terms any multiplier greater than 1 is adequate. Based on our numerical experience, taking it to be 2 gives good results, better than using values in (1,1.5).

2.3. Smoothed estimators $\widehat{W}^{\mathrm{L}}$, $\widehat{W}^{\mathrm{R}}$, $\hat{\theta}^{\mathrm{L}}$ *and* $\hat{\theta}^{\mathrm{R}}$. Both $\widehat{V}^{\mathrm{L}}$ and $\widehat{V}^{\mathrm{R}}$ can convey useful information about the location of $\partial\mathcal{S}$ in the current vicinity, and both $\widehat{\omega}^{\mathrm{L}}$ and $\widehat{\omega}^{\mathrm{R}}$ can give useful information about the slope of the tangent to $\partial\mathcal{S}$ there. Only when $\widehat{\omega}^{\mathrm{L}}$ and $\widehat{\omega}^{\mathrm{R}}$ are some distance apart, indicating the presence of a corner, would we want to place particular emphasis on one of the smooths rather than the other. The present section suggests a way of allocating emphasis, depending on the size of $|\widehat{\omega}^{\mathrm{L}} - \widehat{\omega}^{\mathrm{R}}|$.

Let $B_1 > 0$ and put

$$\hat{\rho} = \tfrac{1}{2}\, K\big(|\widehat{\omega}^{\mathrm{L}} - \widehat{\omega}^{\mathrm{R}}|/B_1 h\big)\,, \tag{2.2}$$

which is a nonincreasing function of the distance between the estimated angles. (The constants $B_1, B_2, \ldots$, as well as the bandwidth h, are tuning parameters of our algorithm.) Our smoothed versions of $\widehat{V}^{\mathrm{L}}$ and $\widehat{V}^{\mathrm{R}}$ are

$$\widehat{W}^{\mathrm{L}} = (1-\hat{\rho})\,\widehat{V}^{\mathrm{L}} + \hat{\rho}\,\widehat{V}^{\mathrm{R}} \quad \text{and} \quad \widehat{W}^{\mathrm{R}} = \hat{\rho}\,\widehat{V}^{\mathrm{L}} + (1-\hat{\rho})\,\widehat{V}^{\mathrm{R}}\,, \tag{2.3}$$

respectively. If $\widehat{\omega}^{\mathrm{L}}$ and $\widehat{\omega}^{\mathrm{R}}$ are close, which will generally be the case except in neighbourhoods of corners, $\hat{\rho}$ will be close to $\frac{1}{2}$ and so $\widehat{W}^{\mathrm{L}}$ and $\widehat{W}^{\mathrm{R}}$ will both be close to the simple average, $\frac{1}{2}(\widehat{V}^{\mathrm{L}} + \widehat{V}^{\mathrm{R}})$, of the left- and right-hand estimates. On the other hand, if $\widehat{\omega}^{\mathrm{L}}$ and $\widehat{\omega}^{\mathrm{R}}$ are some distance apart then $\hat{\rho}$ will be close to zero, and $\widehat{W}^{\mathrm{L}}$ and $\widehat{W}^{\mathrm{R}}$ will be close to $\widehat{V}^{\mathrm{L}}$ and $\widehat{V}^{\mathrm{R}}$, respectively. Likewise, the smoothed versions of $\widehat{\omega}^{\rm L}$ and $\widehat{\omega}^{\rm R}$ are

$$\hat\theta^{\rm L} = (1-\hat\rho)\,\widehat{\omega}^{\rm L} + \hat\rho\,\widehat{\omega}^{\rm R} \quad \text{and} \quad \hat\theta^{\rm R} = \hat\rho\,\widehat{\omega}^{\rm L} + (1-\hat\rho)\,\widehat{\omega}^{\rm R}\,.$$

Since K vanishes outside $(0,1)$ (see section 2.2) then $\hat\rho = 0$ if $|\widehat{\omega}^{\rm R} - \widehat{\omega}^{\rm L}| > B_1 h$, and in this case $(\widehat{V}^{\rm L}, \widehat{V}^{\rm R}, \widehat{\omega}^{\rm L}, \widehat{\omega}^{\rm R})$ is identical to $(\widehat{W}^{\rm L}, \widehat{W}^{\rm R}, \hat\theta^{\rm L}, \hat\theta^{\rm R})$. This degeneracy warns of an approaching corner, and will in fact form the basis of our procedure for identifying corners.

2.4. Details of algorithm. The algorithm that we shall give below depends on several tuning parameters: the bandwidth h introduced in section 2.1, constants $B_1 \geq B_2$ appearing in the definitions of $\hat\rho$ (see section 2.3) and in thresholds (see step (iv) below), a constant $\epsilon > 0$ determining the approximate distance between successive estimates of points on the boundary (step (ii) below), and $\Delta > 0$ governing the choice of $\mathcal{L}$ in section 2.2. In asymptotic terms, if the point process in $\mathcal{S}$ is homogeneous Poisson with intensity ν per unity area, it suffices to put $h = B_3\,(\nu^{-1}\log\nu)^{1/3}$ where $B_3 \geq B$, say, and to take $B_1 \geq B_2 \geq B$, $\epsilon \in (0,1]$, and Δ strictly less than the acute angle that the initial line $\mathcal{L}_1$ makes to the tangent to $\partial\mathcal{S}$ (see step (i) below). For this choice of tuning parameters, the probability that the boundary estimator defined by linearly interpolating between adjacent confirmed points $\widehat{Q}_j$, prescribed by the algorithm given below, lies uniformly within a fixed constant multiple of h^2 of the true boundary over any finite portion, is no less than $1 - O(\nu^{-\lambda})$, where $\lambda = \lambda(B)$ increases without bound as $B \to \infty$. Theorem 3.1 will give details.

Next we give the algorithm, which is in a sequence of four steps. Key parts of the respective steps are depicted in panels (a)–(d) Figure 1 below.

Step (i): Initiation. We start the algorithm on a line $\mathcal{L}_1$ which intersects the boundary. Taking the line $\mathcal{L}$ in section 2.2 to be $\mathcal{L}_1$, use the procedure there to calculate $\widehat{V}^{\rm L}$, $\widehat{\omega}^{\rm L}$ and $\widehat{\omega}^{\rm R}$, giving each of them the subscript 1; and take $\widehat{Q}_1 = \widehat{V}_1^{\rm L}$. (This is the only instance where $\widehat{Q}_j$ is a "rough" estimator of location, as defined in section 2.2; for $j \geq 2$ it is a "smooth" estimator, as defined in section 2.3.) Let $(\hat\theta_1^{\rm L}, \hat\theta_1^{\rm R})$ be the version of $(\hat\theta^{\rm L}, \hat\theta^{\rm R})$ computed from $(\widehat{\omega}^{\rm L}, \widehat{\omega}^{\rm R})$. For definiteness, immediately after $\widehat{Q}_1$ we use the left smooth.

Step (ii): Calculating $\widehat{Q}'_{j+1}$ from $\widehat{Q}_j$. The jth step is defined to be the step that takes $\widehat{Q}_{j-1}$ to $\widehat{Q}_j$. Suppose we have calculated $(\hat\theta_i^{\rm L}, \hat\theta_i^{\rm R})$ for $i \leq j$. If, immediately

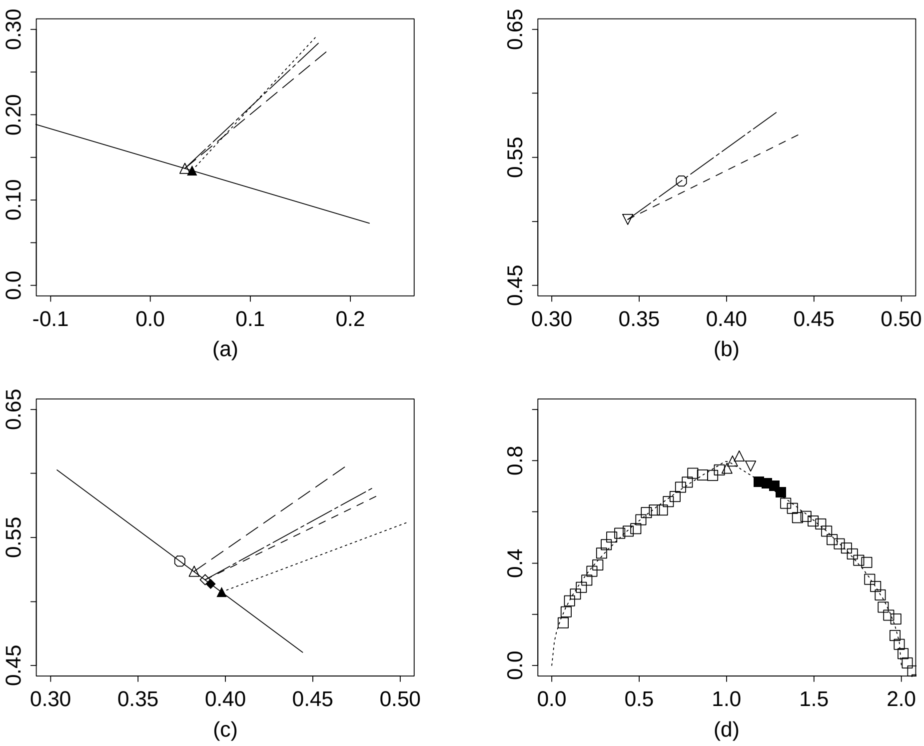

Figure 1. *Illustration of the algorithm.* Panels (a)–(d) respectively illustrate steps (i)–(iv). In panel (a), the unbroken line depicts $\mathcal{L}_1$, the triangle and filled triangle locate $\widehat{V}_1^{\mathrm{L}}$ and $\widehat{V}_1^{\mathrm{R}}$ ($\widehat{Q}_1 = \widehat{V}_1^{\mathrm{L}}$), and $\widehat{\omega}_1^{\mathrm{L}}$, $\widehat{\omega}_1^{\mathrm{R}}$ and $\hat{\theta}_1^{\mathrm{L}}$ are the angles made by the dashed, dotted and dotted-dashed lines. Panel (b) depicts the points $\widehat{Q}_j$ (inverted triangle) and $\widehat{Q}'_{j+1}$ (circle), and the angles $\hat{\theta}_j^{\mathrm{L}}$ and $\hat{\theta}_j^{\mathrm{R}}$ made by the dot-dashed and short-dashed lines. (At this particular step $\hat{\theta}_j = \hat{\theta}_j^{\mathrm{L}}$.) Panel (c) shows $\mathcal{L}_{j+1}$ (unbroken line), $\widehat{Q}'_{j+1}$ (circle), the rough estimates $\widehat{V}_{j+1}^{\mathrm{L}}$ (triangle) and $\widehat{V}_{j+1}^{\mathrm{R}}$ (filled triangle), the smooth estimates $\widehat{W}_{j+1}^{\mathrm{L}}$ (diamond) and $\widehat{W}_{j+1}^{\mathrm{R}}$ (filled diamond), and the angles $\widehat{\omega}_{j+1}^{\mathrm{L}}$, $\widehat{\omega}_{j+1}^{\mathrm{R}}$, $\hat{\theta}_{j+1}^{\mathrm{L}}$ and $\hat{\theta}_{j+1}^{\mathrm{R}}$ made by the dashed, dotted, dot-dashed and short-dashed lines. Panel (d) shows the true boundary (dotted line) and $\widehat{Q}_1, \widehat{Q}_2, \cdots$ computed from a sample generated by model (4.1). To illustrate step (iv), the squares and triangles respectively indicate cases (a) and (b), the inverted triangle (or the last in a sequence of filled squares) indicates where a switch from left (or right) smooth to right (or left) smooth occurs.

after the jth step, we are using the left (respectively, right) smooth, let $\widehat{Q}'_{j+1}$ be the point in the plane reached by moving distance ϵh along the line passing through $\widehat{Q}_j$ at angle $\hat{\theta}_j^{\mathrm{L}}$ (respectively $\hat{\theta}_j^{\mathrm{R}}$), travelling in a clockwise direction relative to the part of the boundary that has already been tracked.

Step (iii): Calculating $\widehat{Q}_{j+1}$ from $\widehat{Q}'_{j+1}$. Let $\hat{\theta}_j = \hat{\theta}_j^{\mathrm{L}}$ or $\hat{\theta}_j^{\mathrm{R}}$, according as the left or right smooth, respectively, was used to transit from $\widehat{Q}_j$ to $\widehat{Q}'_{j+1}$. Compute

$(\widehat{V}^{\mathrm{L}}_{j+1}, \widehat{V}^{\mathrm{R}}_{j+1}, \widehat{\omega}^{\mathrm{L}}_{j+1}, \widehat{\omega}^{\mathrm{R}}_{j+1})$ by applying the method of section 2.2 with $Q' = \widehat{Q}'_{j+1}$ and with $\mathcal{L}$ equal to the line passing through $\widehat{Q}'_{j+1}$ and making angle $\hat{\theta}_j + \frac{\pi}{2}$. Then, calculate $(\widehat{W}^{\mathrm{L}}_{j+1}, \widehat{W}^{\mathrm{R}}_{j+1}, \hat{\theta}^{\mathrm{L}}_{j+1}, \hat{\theta}^{\mathrm{R}}_{j+1})$ using the method of section 2.3. Put $\widehat{Q}_{j+1} = \widehat{W}^{\mathrm{L}}_{j+1}$ or $\widehat{W}^{\mathrm{R}}_{j+1}$ according as $\hat{\theta}_j = \hat{\theta}^{\mathrm{L}}_j$ or $\hat{\theta}^{\mathrm{R}}_j$.

Step (iv): Changing smooths and incorporating corners. Let $0 < B_2 \le B_1$. If we used a right smooth to calculate $\widehat{Q}_{j+1}$, and if $|\widehat{\omega}^{\mathrm{R}}_{j+1} - \widehat{\omega}^{\mathrm{L}}_{j+1}| < B_2 h$, then we switch to a left smooth immediately after $\widehat{Q}_{j+1}$, and otherwise we continue using a right smooth. If we used a left smooth to calculate $\widehat{Q}_{j+1}$, and if either (a) $|\widehat{\omega}^{\mathrm{R}}_i - \widehat{\omega}^{\mathrm{L}}_i|$ has not exceeded $B_1 h$ since the last corner, or (b) $|\widehat{\omega}^{\mathrm{R}}_i - \widehat{\omega}^{\mathrm{L}}_i|$ has exceeded $B_1 h$ but has not dropped back to a value not exceeding $B_1 h$ since that point, then we use a left smooth immediately after $\widehat{Q}_{j+1}$ and we do not declare a corner to have been rounded. On the other hand, if we used a left smooth to calculate $\widehat{Q}_{j+1}$, and if (a) and (b) both fail (implying that $|\widehat{\omega}^{\mathrm{L}}_i - \widehat{\omega}^{\mathrm{R}}_i|$ has just, for $i = j+1$, dropped below $B_1 h$ for the first time in a sequence of consecutive values of i), then we declare a corner to have been rounded and switch to a right smooth. We estimate the position of the corner by extrapolation backward from $\widehat{Q}_{j+1}$ at angle $\hat{\theta}^{\mathrm{R}}_{j+1}$, and forward from the next-most-recent value of $\widehat{Q}_i$ ($\widehat{Q}_k$, say) for which $|\widehat{\omega}^{\mathrm{R}}_i - \widehat{\omega}^{\mathrm{L}}_i| \le B_1 h$. The angle of extrapolation there is $\hat{\theta}^{\mathrm{L}}_k$.

Note that changing smooths and identifying corners involves two thresholds. The first, $B_1 h$, is for switching from a left to a right smooth, which will happen not long after a corner is rounded. The second, $B_2 h$, is for switching from a right to a left smooth, and this will occur midway between two corners. In each case the switch occurs when a value of $|\widehat{\omega}^{\mathrm{R}}_i - \widehat{\omega}^{\mathrm{L}}_i|$ falls below the relevant threshold. It may not be clear that the tuning parameters and kernel can be chosen such that, despite all the parity changes, the boundary estimator is smooth between corners. However, section 2.5 will show that this can, in fact, be achieved.

2.5. Smoothness of the boundary estimate. The switch from left to right smooths, which is determined by the threshold $B_2 h$, can be achieved very gently by taking the kernel K at (2.2) to be flat and identically equal to 1 in a nonvanishing interval immediately to the right of the origin. Indeed, in that case the switch can be effected at a place where the left and right smooths are identical, simply by choosing B_1 and B_2 sufficiently large.

To appreciate why this is possible, note that the probability that $|\widehat{\omega}^{\rm L} - \widehat{\omega}^{\rm R}| \leq Ch$, uniformly in smooth parts of the boundary, converges to 1 as $C \to \infty$. Indeed, if $\lambda > 0$ is given, and h is chosen to produce the optimal convergence rate, then the probability equals $1 - O(\nu^{-\lambda})$ for fixed $C = C(\lambda)$ sufficiently large, where ν denotes the intensity of the point process. Therefore, if K is identically 1 on the interval $[0, \xi]$ (where $0 < \xi < 1$), and if we choose $\xi B_1 > B_2 \geq C$, then (with probability $1 - O(\nu^{-\lambda})$) the weight $\hat{\rho}$ defined at (2.2) equals $\frac{1}{2}$ at any point at which $|\widehat{\omega}^{\rm L} - \widehat{\omega}^{\rm R}|$ is close to $B_2 h$, where a switch is made from the right to the left smooth. In consequence, when the switch occurs both the left and right smooths are equal, to $\frac{1}{2}(\widehat{V}^{\rm L} + \widehat{V}^{\rm R})$.

If this regime applies, and if the kernel K is continuous, then the curve estimate is smooth in the following sense. If each X_i is perturbed by the addition of a small 2-vector δ_i then, as the δ_i's converge uniformly to 0, the curve estimator converges to its counterpart with each $\delta_i = 0$.

3. Theoretical properties

3.1. Regularity conditions. First we define what we mean by corners in, and ends of, a segment of a support boundary. Suppose a function f of two variables has support $\mathcal{S}$, with boundary $\partial\mathcal{S}$, and that there exists a finite number of distinct points, $P_0, \ldots, P_k$ say (with $k \geq 1$), in this order in a clockwise sense along the boundary, such that: (a) $\partial\mathcal{S}$ has a continuously turning tangent and uniformly bounded curvature between P_j and P_{j+1} (for $0 \leq j \leq k-1$); (b) the tangent angles have well-defined limits as P_j is approached from the direction of P_{j+1} and as P_{j+1} is approached from the direction of P_j (for $0 \leq j \leq k-1$); (c) if $0 \leq j_1 < j_2 \leq k-1$ then the boundary segment between P_{j_1} and P_{j_1+1}, and the boundary segment between P_{j_2} and P_{j_2+1}, do not intersect except possibly at just one of their ends, and in this case $j_2 = j_1 + 1$; and (d) if $k \geq 2$ then the difference between the limits of tangent angles on either side of P_j, for $1 \leq j \leq k-1$, is assumed not to equal π. Assuming these conditions to hold for some $k \geq 2$, we define P_0 and P_k to be the *ends* of the boundary segment, and $P_1, \ldots, P_{k-1}$ to be the *corners*.

We assume of f that

> f is a nonnegative, compactly supported function of two variables, supported and with a bounded derivative on $\mathcal{S}$, bounded away from zero on $\partial\mathcal{S}$, and such that a segment of $\partial\mathcal{S}$ contains just $k-1$ corners $P_1, \ldots, P_{k-1}$ between its ends P_0 and P_k. $\quad (\mathrm{C}_f)$

Since $P_1, \ldots, P_{k-1}$ are, in (C_f), assumed to be "corners" between P_0 and P_k, then (C_f) also implies the properties asserted in the definition of corners, i.e. in (a)–(d) in the previous paragraph. In this regard the following consequences of (C_f) should be stressed. First, the tangent angle varies continuously except at a finite number of points $P_1, \ldots, P_k$. Secondly, these points are distinct, and so corners cannot coincide. Thirdly, although, in an infinite class of boundaries satisfying (C_f), corners can be arbitrarily close together and arbitrarily large in number, for any single boundary in the class the corners are distinct and finite in number. Theorem 3.1 applies to this setting, where there is a fixed boundary with a fixed number of distinct corners. Theorem 3.2, which asserts a bound that applies uniformly over different boundaries, is formulated in the context of the smooth boundary-fragment model, where corners do not arise.

Assume the point process $\mathcal{X} = \{X_1, X_2, \ldots\}$ in the plane is Poisson with intensity νf. We allow ν to diverge to infinity and take $h = h(\nu)$ to be a positive quantity satisfying

$$h \to 0 \text{ as } \nu \to \infty, \text{ and } h \geq B_3\,(\nu^{-1} \log \nu)^{1/3}, \qquad (\mathrm{C}_h)$$

where $B_3 > 0$. We suppose too that

> K has a bounded derivative on the positive real line, is nonincreasing there, and satisfies $K(0) = 1$, $K(1) = 0$ and $K > 0$ on $[0, 1)$. $\quad (\mathrm{C}_K)$

Theorems 3.1 and 3.2 continue to hold if we assume in addition that f is a probability density, and ask that instead of $\mathcal{X}$ being a Poisson process, it is a set of exactly ν independent random variables each distributed with density f. However, when considering the performance of tracking methods it is arguably more appropriate to consider Poisson distributed points, since the tracking algorithm is motivated by the fact that we do not need to treat all the data in $\mathcal{X}$. Indeed, tracking methods use only points very close to $\partial\mathcal{S}$, and even if the data were independent and identically distributed we would likely never know the number of points.

3.2. Defining "above". The estimator definitions in section 2 are unambiguous if we specify what we mean by saying that a line segment $\mathcal{M}$ of length $2h$ lies "above" the point cloud represented by $\mathcal{X}$. To this end, let $B > 0$ be a large constant and construct open rectangles, with dimensions $2h \times Bh^2$, on either side of $\mathcal{M}$, in each case with $\mathcal{M}$ as one of their longer sides. If no more than one of the rectangles contains points of $\mathcal{X}$ then we say that $\mathcal{M}$ lies above $\mathcal{X}$, and if exactly one of the rectangles contains points of $\mathcal{X}$ then we say that the direction of the opposite rectangle, relative to $\mathcal{M}$, is "above" $\mathcal{X}$. With this interpretation we say that "no point of $\mathcal{X}$ lies above $\mathcal{M}$". This definition of "above" is of course local to $\mathcal{M}$, and subject to statistical error.

If a $2h \times Bh^2$ rectangle lies entirely within $\mathcal{S}$ then the probability that it contains at least one point of $\mathcal{X}$ equals $1 - O(\nu^{-\lambda})$, where λ can be made arbitrarily large by choosing B_3 sufficiently large. The number of steps needed to track the boundary will be only polynomially large in ν, and so ambiguities that arise with probability $O(\nu^{-\lambda})$, for sufficiently large λ, in specifying what we mean by "above", are adequately small. Arguments such as this indicate that our definition is adequate, despite its inherent statistical error, and that fact will be confirmed by Theorem 3.1.

3.3. Penetrating into corners. In this section we discuss difficulties that are inherent to estimating the boundary at corners, and show that extrapolation methods such as that suggested in section 2 are essential for solving this problem.

For simplicity, let us take $h = B_3 (\nu^{-1} \log \nu)^{1/3}$ in condition (C_h); this is the size of bandwidth that gives optimal convergence rates. A simple calculation based on the Poisson distribution shows that the probability that no points of $\mathcal{X}$ lie within radius $h^{3/2}$ of the corner at P_j converges to 1 as $\nu \to \infty$. Therefore, there are effectively no data within radius $O(h^2)$ of a corner. This implies that any estimator of the corner which is accurate to within $O(h^2)$, as we shall claim ours to be, has to be based on extrapolation from an order of magnitude further away. The fact that linear extrapolation is adequate, even over distances as large as $O(h)$, follows from the property that the tangent angle at a distance $O(h)$ from the corner can be estimated with accuracy $O(h)$; see points (IV) and (V) in the following section.

In practice there are occasionally problems, in $O(h)$-neighbourhoods of corners, with formal definitions of $\mathcal{M}^{\mathrm{L}}$ and $\mathcal{M}^{\mathrm{R}}$. They arise when no points of $\mathcal{X}$ lie to the left or right, respectively, of the transect $\mathcal{L}$ on which $\mathcal{M}^{\mathrm{L}}$ and $\mathcal{M}^{\mathrm{R}}$ are based, and are readily overcome by making minor adjustments. They cause no problems in our theoretical work, however, since, given any $\lambda > 0$ the difficulties arise only with probability $O(\nu^{-\lambda})$, for sufficiently large values of the tuning constants B_j. The main reason is that $\mathcal{L}$ is taken to be approximately perpendicular to the boundary at the current point; it is perpendicular to the previous tangent estimate.

3.4. Main results. We shall trace the boundary segment in a clockwise direction. To initiate the algorithm, draw a line $\mathcal{L}_1$ that cuts the boundary strictly between P_0 and P_1, at a point Q_1 where $\mathcal{L}_1$ is not tangential to $\partial\mathcal{S}$. Construct the first point estimate, $\widehat{Q}_1$, of the boundary estimate by arguing as in section 2.4. Immediately after this point the left smooth is used. With high probability, the difference between tangent angles will not exceed a certain constant multiple of h until the boundary has been tracked to within $2h$ of the next corner, during which time the "handedness" of the smooth will have switched from right to left. The algorithm specified in section 2.4 is now followed until a vertical line $\mathcal{L}_2$ is first reached or crossed; $\mathcal{L}_2$ is assumed to cut the boundary strictly between P_{k-1} and P_k and not be tangential to $\partial\mathcal{S}$ there. At that stage we terminate the algorithm.

Recall that Δ is used in the definition of tangent angles in section 2.2; B_1 and $B_2 \leq B_1$ are used in the definitions of the tangent-angle weight $\hat{\rho}$ and the tangent-angle threshold $B_2 h$, respectively; and B_3 is employed in condition (C_h). Given $C > 0$, let $\mathcal{E}(C)$ denote the event that the following six properties all hold:

(I) the algorithm correctly determines that there are just $k-1$ corners and k smooth sections of the boundary,

(II) each corner estimate is no further than Ch^2 from its true location,

(III) if a tube is constructed as the union of the continuum of discs of radius Ch^2 with their centres along the piecewise-linear boundary estimate, then the true boundary lies within the tube and leaves it only at the tube's beginning and end,

(IV) the left- and right-hand tangent-angle estimates $\hat{\theta}^{\mathrm{L}}$ and $\hat{\theta}^{\mathrm{R}}$, at confirmed points in the sequence $\widehat{Q}_j$ introduced in section 2.6, are both in error by no more than Ch,

(V) all points $\widehat{Q}_j$ that lie further than Ch from each corner are confirmed, and

(VI) the number of steps taken before the algorithm terminates is not greater than Ch^{-1}.

Theorem 3.1. *Assume conditions* (C_f), (C_h) *and* (C_K), *and that constants* $\epsilon > 0$ *(for the step length,* ϵh*) and* $\lambda > 0$ *are given. Then, provided* B_1, B_2, B_3 *and* C *are chosen sufficiently large, and* Δ *is sufficiently small,* $P\{\mathcal{E}(C)\} = 1 - O(\nu^{-\lambda})$.

Instead of being fixed, the value of ϵ (in the definition of step length) may decrease at a rate no faster than a polynomial in ν^{-1}, although then property (VI) in the definition of the event $\mathcal{E}(C)$ should be altered by stating that the number of steps is no more than $C(\epsilon h)^{-1}$. In this case the proof in section 5 needs to be substantially revised, and can be based on Bernstein-type inequalities for high-order differences of centred forms of the weighted counts at (2.1).

The technique described by Theorem 3.1 is adaptive, in that it detects corners.

However, choice of the constants B_1, B_2 and B_3 requires knowledge of the maximum curvature between corners. An empirically adaptive algorithm can be developed to remove this difficulty. It requires the boundary to have two continuous derivatives between consecutive corners, and involves estimating curvature there. In this way one can construct an empirical upper bound, $\hat{\gamma}$ say, to maximum curvature, γ, having the property that $P(\gamma < \hat{\gamma} < \gamma + 1) = 1 - O(\nu^{-\lambda})$ for all $\lambda > 0$. Replacing γ by $\hat{\gamma}$ in formulae for tuning constants we obtain an empirically adaptive version of Theorem 3.1.

Taking $h = B_3\,(\nu^{-1}\log\nu)^{1/3}$, and choosing B_1, B_2 and B_3 large, we deduce from Theorem 3.1 that the uniform convergence rate is $(\nu^{-1}\log\nu)^{2/3}$. This is the minimax optimal rate for estimating boundaries with continuously turning tangents and bounded curvature, even in the absence of corners. To define the optimal rate, suppose $B > 0$ is given and let $\mathcal{P}_{\nu,g}$ denote a Poisson process with constant intensity ν per unit area, supported in the region $\{(x^{(1)}, x^{(2)}) : x^{(2)} \le g(x^{(1)}),\ 0 \le x^{(1)} \le 1\}$, where g is a member of the class $\mathcal{G}(B)$ of functions on the interval $\mathcal{I} = [0,1]$ that have two derivatives there and satisfy $\sup|g^{(j)}| \le B$ for $j = 0, 1, 2$. This is the "boundary fragments" model used extensively by Korostelev and Tsybakov (1993c). The following result expresses $(\nu^{-1}\log\nu)^{2/3}$ as a lower bound to the convergence rate of estimators of boundaries in $\mathcal{G}(B)$.

Theorem 3.2. *If $\widehat{\mathcal{G}}$ denotes the class of measurable estimators $\hat{g}$ of g based on the data $\mathcal{P}_{\nu,g}$, then*

$$\lim_{\xi\to 0}\ \liminf_{\nu\to\infty}\ \inf_{\hat{g}\in\widehat{\mathcal{G}}}\ \sup_{g\in\mathcal{G}(B)}\ P_{\nu,g}\Big\{\sup_{x\in\mathcal{I}}|\hat{g}(x) - g(x)| > \xi\,(\nu^{-1}\log\nu)^{2/3}\Big\} = 1\,, \tag{3.1}$$

where the probability measure $P_{\nu,g}$ is that corresponding to the process $\mathcal{P}_{\nu,g}$.

The particular estimator $\hat{g}$, defined by our tracking algorithm with $h = B_3 \times (\nu^{-1}\log\nu)^{1/3}$, initiated by the right smooth at the line $\mathcal{L}_1$ defined by the equation $x^{(1)} = 0$, and terminated at the line $\mathcal{L}_2$ given by $x^{(1)} = 1$, satisfies

$$\lim_{\xi\to\infty}\ \limsup_{\nu\to\infty}\ \sup_{g\in\mathcal{G}(B)}\ P_{\nu,g}\Big\{\sup_{x\in\mathcal{I}}|\hat{g}(x) - g(x)| > \xi\,(\nu^{-1}\log\nu)^{2/3}\Big\} = 0 \tag{3.2}$$

for $B_1 \geq B_2 \geq B$, $B_3 \geq B$, B sufficiently large and Δ sufficiently small. (To avoid edge effects, the obvious modifications should be made to the "handedness" of smooths used near either boundary.) Therefore our estimator achieves the uniform convergence rate implicit in (3.1) in the absence of corners. The rate in the presence of corners must therefore also be optimal.

Theorem 3.1 will be derived in section 5, and result (3.2) may be proved similarly. With the minor alteration that our point process is Poisson, rather than the result of distributing a given number of independent random variables, Theorem 3.2 follows from results in section 5.3 of Korostelev and Tsybakov (1993c); see in particular their Theorem 5.3.3. Clearly, our Theorem 3.2 implies an analogous result in the presence of corners.

4. Numerical illustration.

First, numerical solution to the minimisation problem (2.1) needs explanation. Finding $\widehat{\mathcal{M}}^{\mathrm{L}}$ and finding $\widehat{\mathcal{M}}^{\mathrm{R}}$ are analogous so we only discuss the latter. Note that the angle made by $\mathcal{M}^{\mathrm{R}}$ has to be in the range $\Theta = (\hat{\theta} - \frac{\pi}{2} + \Delta, \hat{\theta} + \frac{\pi}{2} - \Delta)$, where $\hat{\theta} + \frac{\pi}{2}$ is the angle made by $\mathcal{L}$. Viewed as a function of the angle of $\mathcal{M}^{\mathrm{R}}$, the kernel-weighted distance defined in (2.1) can fluctuate significantly. Therefore it is suggested to perform a grid search in Θ, subject to the constraints that no point of $\mathcal{X}$ lies above $\mathcal{M}^{\mathrm{R}}$ and at least one point of $\mathcal{X}$ lies on it, to find the solution $\widehat{\mathcal{M}}^{\mathrm{R}}$.

Throughout this section we consider the following model. Suppose V is exponentially distributed with mean 1/3, and X is uniformly distributed on the interval $[0, 2]$. Put

$$Y = 0.8 \left\{ \sqrt{X}\, I_{(0<X<1)} \;+\; \sqrt{2-X}\, I_{(1 \leq X < 2)} \right\} \exp(-V)\,. \tag{4.1}$$

We sample 100 independent points (X, Y) from this model. The support boundary is described by the equation $y = 0.8 \left\{ \sqrt{x}\, I_{(0<x<1)} + \sqrt{2-x}\, I_{(1 \leq x < 2)} \right\}$. It has a sharp corner at the point $(1, 0.8)$; see Figure 3.

Figure 2 illustrates jth step of the algorithm. The crosses represent data points,

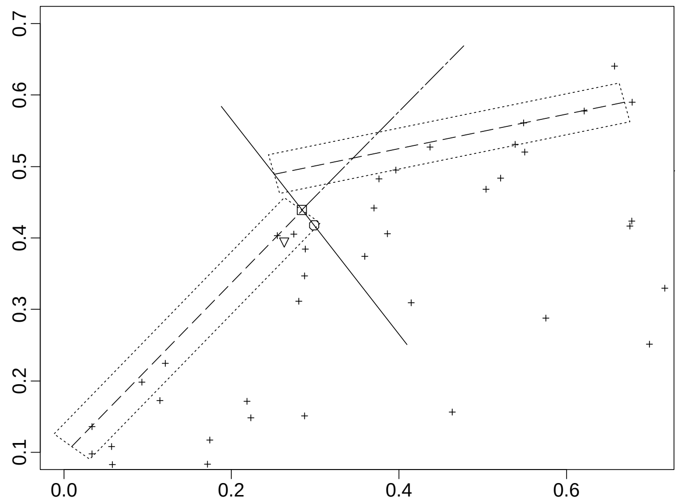

Figure 2. *Illustration at a particular step.* The crosses represent part of a sample of size 100 from model (4.1). The inverted triangle, circle and square respectively label $\widehat{Q}_j$, $\widehat{Q}'_{j+1}$ and $\widehat{Q}_{j+1}$. The unbroken line is $\mathcal{L}$ and the dashed lines are $\widehat{\mathcal{M}}^{\mathrm{L}}$ and $\widehat{\mathcal{M}}^{\mathrm{R}}$. The dotted rectangles, each with dimensions $2h \times Bh^2$, are used to confirm that $\widehat{\mathcal{M}}^{\mathrm{L}}$ and $\widehat{\mathcal{M}}^{\mathrm{R}}$ lie above $\mathcal{X}$. The dot-dashed line leaves $\widehat{Q}_{j+1}$ at the angle $\hat{\theta}_{j+1}$.

and the centres of the inverted triangle and circle indicate $\widehat{Q}_j$ and $\widehat{Q}'_{j+1}$, respectively. The unbroken line represents $\mathcal{L}$; it passes through $\widehat{Q}'_{j+1}$ and makes angle $\hat{\theta}_j + \frac{1}{2}\pi$. The dashed lines (i.e. the longer axes of the two rectangles) are $\widehat{\mathcal{M}}^{\mathrm{L}}$ and $\widehat{\mathcal{M}}^{\mathrm{R}}$. The intersection of $\mathcal{L}$ and $\widehat{\mathcal{M}}^{\mathrm{L}}$ (or of $\mathcal{L}$ and $\widehat{\mathcal{M}}^{\mathrm{R}}$) forms the rough estimate, $\widehat{V}^{\mathrm{L}}$ (or $\widehat{V}^{R}$). In the case of this step, the left smooth was used immediately after the jth step, implying that the left smooth $\widehat{W}^{\mathrm{L}}$ (indicated by the centre of the square situated very close to $\widehat{V}^{\mathrm{L}}$) is the confirmed smooth boundary-point estimate $\widehat{Q}_{j+1}$. Since the left smooth was used to calculate $\widehat{Q}_{j+1}$ and $|\widehat{\omega}_i^R - \widehat{\omega}_i^{\mathrm{L}}|$ has not exceeded $B_1 h$ since last corner, we continue using a left smooth in the next step, i.e. $\hat{\theta}_{j+1} = \hat{\theta}^{\mathrm{L}}_{j+1}$. The ray that leaves $\widehat{Q}_{j+1}$ at angle $\hat{\theta}_{j+1}$, and points in the direction of travel around the boundary, is indicated by the dot-dashed line. The next tentative point estimate, $\widehat{Q}'_{j+2}$, will lie on it and be distance ϵh from $\widehat{Q}_{j+1}$.

Figure 3 shows the support-tracking algorithm at work for a given sample. The true boundary and data points are shown in panel (a). Panels (b) and (c) plot successive tentative boundary estimates $\widehat{Q}'_j$, and smooth estimates $\widehat{Q}_j$, respectively.

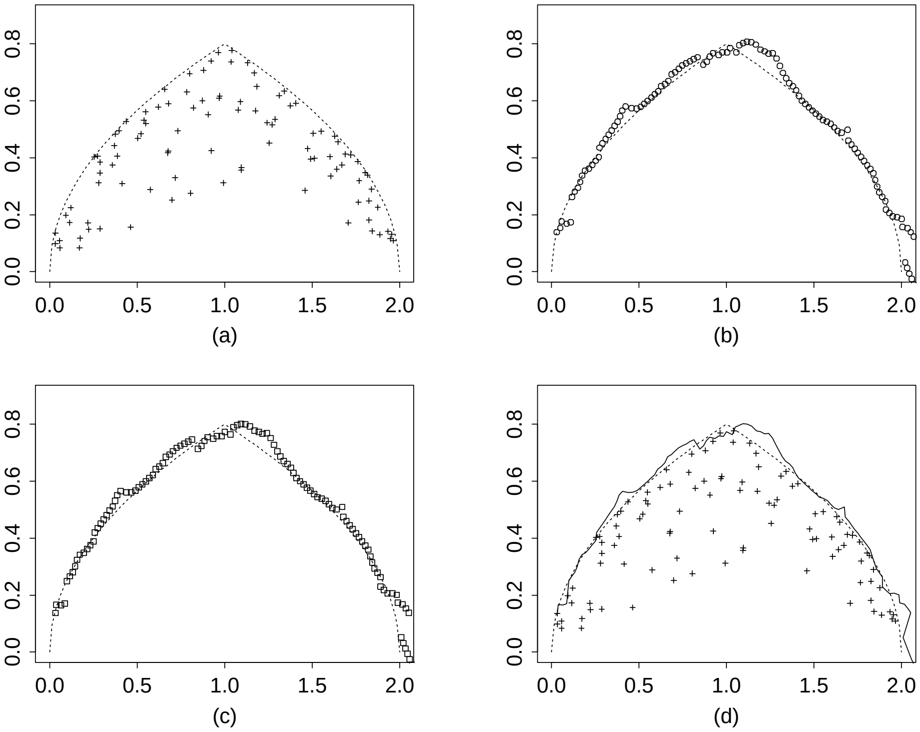

Figure 3. *Numerical results for a given sample.* The data, generated from model (4.1), are depicted in panels (a) and (d). The dotted line represents the true support edge of the model. The sequences of estimates $\widehat{Q}'_j$ and $\widehat{Q}_j$ are given in panels (b) and (c), respectively. The final support edge estimate is given by the unbroken line in panel (d).

Panel (d) plots the final boundary estimate, with the confirmed corner obtained by linear extrapolation. Figure 3 reveals an appealing feature of our tracking method: it moves smoothly, even at "gaps" in the point cloud, effectively guarding against fluctuations. This behavior is a result of the restriction on the difference between adjacent tangent angle approximations, which is controlled by the parameter Δ, see section 2.2.

A slight modification of the algorithm further strengthens the above mentioned smoothness property: the angle between the line segment connecting $\widehat{Q}_j$ and $\widehat{V}^{\mathrm{L}}_{j+1}$ (or $\widehat{V}^{\mathrm{R}}_{j+1}$) and the line segment containing $\widehat{Q}_j$ and $\widehat{Q}'_{j+1}$ cannot exceed a given amount, say $\pi - 2\Delta$; ensuring that the "approximate tangent line," formed by the successive boundary estimates $\widehat{Q}_j$ and $\widehat{Q}_{j+1}$, cannot be too far apart from the tangent smooth $\hat{\theta}_j$. The effect of this modification is shown in Figure 4, which presents the results for 10 samples. The left (or right) panel shows the results with (or without) implementing the modification. Clearly, the modification produces smoother estimates. Nevertheless, the modification has little impact on L^p performance of the estimator; even without the modification the estimates are still quite close to the true support boundary. Here, $\Delta = 0.35\pi$, $B_2 = 1.5$ and $B_1 = 2.5$. The bandwidth constant B_3 is allowed to vary in $\{0.8, 1, 1.1\}$. The reason is that larger values of B_3 prevents the tracking algorithm from being occasionally trapped in the point cloud, and we picked the smallest value that it did not occur. In addition, $B = 0.6$ when $B_3 = 1$ or 1.1, and equals 0.8 when $B_3 = 0.8$.

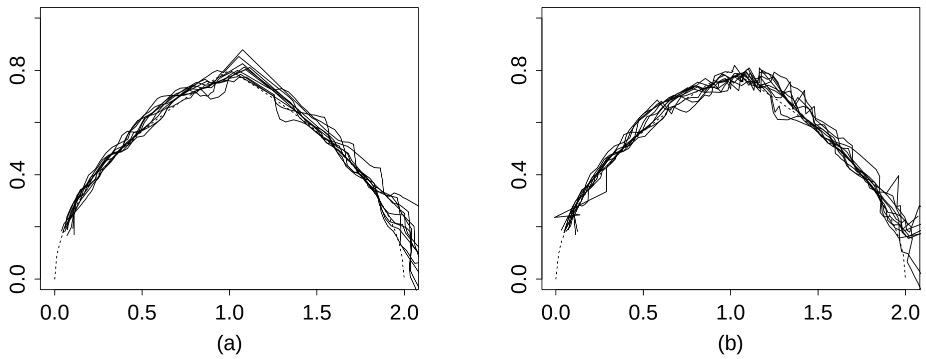

Figure 4. *Effect of the modification.* The data were generated from model (4.1) and the sample size was 100. The estimates in panel (a) were constructed using the modification and the estimates in panel (b) did not involve the modification.

If it happens that we are using left smooths but the right-hand estimates are not available, due to our restriction on the searched angle range, we nevertheless allow the algorithm to continue. We continue using the left-hand estimates, and operate as if $|\widehat{\omega}^{\mathrm{R}} - \widehat{\omega}^{\mathrm{L}}| > B_1 h$. Occasionally it happens that the left- and right-hand estimates are both missing, particularly near a corner. In that case one can either try another set of parameters $(\Delta, B,\ B_1, B_2, B_3)$, or change to a smaller value of Δ at the point (so that the eligible tangent angle range is made larger).

The method involves five tuning parameters $(\Delta, B,\ B_1, B_2, B_3)$. Unless there is strong indication that the smooth parts of the boundary have high curvatures, we suggest to take relatively large values of Δ, e.g. $\Delta \in [0.3\pi, 0.4\pi]$. Since B is employed to decide "above" and may not exceed B_2 or B_3, it is preferable to take small values of B and it can be simply set as $\min\{B_2, B_3\}$ or a slightly smaller value. Choice of B_2, used to determine changing from right to left smooths, is not critical and can be fixed at around $0.5B_1$. The tuning parameter B_1 is used to construct the smoothed estimators, see section 2.3, and controls construction of corner estimates. In general, letting B_1 to be anything in $(2, 4)$ or taking B_1 such that $B_1 h \approx 0.8(\pi - 2\Delta)$ would do well. Among the five tuning parameters, B_3 decides most of the smoothness of the final estimate and often determines whether the algorithm can successfully track the entire support boundary. Once the other parameters fixed as above, one can take B_3 to be the smallest one, among possible values, such that the second qualification meets.

5. Proof of Theorem 3.1

5.1. Preliminaries, and summary of proof. The direction referred to as "below" the support boundary is defined to be the direction to the right of $\partial\mathcal{S}$ as the boundary is traced clockwise. Condition (C_f) implies that there exists a constant $c > 0$ such that for sufficiently small $\delta > 0$, $f(x) > c$ whenever x lies below $\partial\mathcal{S}$ and is within δ of some point on $\partial\mathcal{S}$ between P_0 and P_k. For simplicity we shall suppose that $\partial\mathcal{S}$ has bounded curvature for a nonzero distance on the opposite side of P_0 from P_1, and on the opposite side of P_k from P_{k-1}. (This condition is used only early in section 5.2, where it simplifies the definition of the density p_Q.)

Our proof will treat only the case $h = B_3\,(\nu^{-1}\log\nu)^{1/3}$; the case where $h \to 0$ more slowly than $(\nu^{-1}\log\nu)^{1/3}$ is simpler. We shall separately address (a) estimation of tangent angles uniformly in smooth parts of the curve (section 5.2), (b) estimation of points on the boundary uniformly in places where the boundary is smooth (section 5.4), (c) switching from right to left smooths between corners (section 5.4), and (d) identifying and traversing corners (section 5.5). In each of these four problems we shall show that if $\lambda > 0$ is given then specific levels of accuracy can be achieved, uniformly in those parts of the curve to which the problem applies, with probability $1 - O(\nu^{-\lambda})$. It will also follow that solving each of the problems, along the length of boundary from Q_1 to any point between P_{k-1} and P_k, involves no more than $O(h^{-1})$ steps; see sections 5.4 and 5.5. Therefore, putting the results together and simply adding the probabilities of events where desired levels of accuracy are not present, we obtain the desired accuracy with probability $1 - O(\nu^{(1/3)-\lambda})$. Since λ is arbitrarily large then Theorem 3.1 is proved.

5.2. Tangent angle estimators. A corner in $\partial\mathcal{S}$ will be said to be convex if, on tracking through it (this time following the true boundary), the tangent to $\partial\mathcal{S}$ turns through more than π radians. It is concave if the turning angle is less than π.

For specificity we shall treat the right-hand tangent estimate; the left-hand case is similar. Write $\partial\mathcal{S}_j$ for that part of the boundary between P_j and P_{j+1}, where $0 \leq j \leq k-1$, and let $\partial\mathcal{S}_j(h)$ denote the set of points in $\partial\mathcal{S}_j$ that are at least $2h$ from P_{j+1}. (By considering right-hand tangents at points in $\partial\mathcal{S}_j(h)$, rather than simply points in $\partial\mathcal{S}_j$, we avoid problems caused by edge effects.)

Let $p = c_0 f$ denote the unique probability density on $\mathcal{S}$ that is proportional to f, and let c_Q be the value taken by p at $Q \in \mathcal{S}$. Without loss of generality, $c_0 = 1$; ensuring this property involves only a scale change. Condition (C_f) allows us to choose a constant $\delta > 0$ such that the following is true: for each point Q on $\cup_j \partial\mathcal{S}_j$ there exists a probability density p_Q whose support equals $\mathcal{S}$, is such that $p = p_Q = c_Q$ at Q, and is constant in the region $\{x : x \in \mathcal{S}$ and $\|x - y\| \leq \delta$ for some $y \in \partial\mathcal{S}\}$.

We consider first a deterministic setting. Let $\mathcal{L} = \mathcal{L}(Q, \psi)$ denote the line that intersects $\partial\mathcal{S}_j(h)$ at a point Q and whose normal makes angle $\psi \in (-\pi/2, \pi/2)$ to the tangent to $\partial\mathcal{S}_j$ at Q. Given $C_1 > 0$, let $\mathcal{M}^{\mathrm{R}}$ be a line segment of length $2h$ lying to the right of $\mathcal{L}$, with its left end on $\mathcal{L}$ and placed so that its right-hand end lies within $C_1 h^2$ of $\partial\mathcal{S}$, no part of it lies further than $C_1 h^2$ below $\partial\mathcal{S}$, and the acute angle ϕ that it makes to $\mathcal{L}$ satisfies $|\phi| > \Delta$ for some $\Delta \in (0, \pi/2)$. Call these conditions $(\mathrm{C}^{\mathrm{R}}_{\mathcal{M}})$, and let u^{R} denote the centre of $\mathcal{M}^{\mathrm{R}}$. Define

$$\begin{aligned} \mu(\mathcal{M}^{\mathrm{R}}) &= E\big\{\|X - Y^{\mathrm{R}}\|\, K(\|u^{\mathrm{R}} - Y^{\mathrm{R}}\|/h)\, K(\|u^{\mathrm{R}} - X\|/2h)\big\}\,, \\ \mu_Q(\mathcal{M}^{\mathrm{R}}) &= E\big\{\|X_Q - Y^{\mathrm{R}}_Q\|\, K(\|u^{\mathrm{R}} - Y^{\mathrm{R}}_Q\|/h)\, K(\|u^{\mathrm{R}} - X_Q\|/2h)\big\} \\ &= c_Q \int_{\mathcal{S}} \|x - y^{\mathrm{R}}\|\, K(\|u^{\mathrm{R}} - y^{\mathrm{R}}\|/h)\, K(\|u^{\mathrm{R}} - x\|/2h)\, dx\,, \end{aligned} \tag{5.1}$$

where the random variables X and X_Q are distributed on $\mathcal{S}$ with densities p and p_Q, respectively, and Y^{R}, Y^{R}_Q and y^{R} denote the projections of X, X_Q and x, respectively, onto the infinite line of which $\mathcal{M}^{\mathrm{R}}$ forms a part.

Let $\mathcal{A}^0$ be the set of pairs (Q, ψ) such that $Q \in \cup_j \partial\mathcal{S}_j(h)$ and $\psi \in [-\frac{1}{2}\pi + \Delta, \frac{1}{2}\pi - \Delta]$, and let $\mathcal{A}^1$ denote the set of triples $(Q, \psi, \mathcal{M}^{\mathrm{R}})$ such that $(Q, \psi) \in \mathcal{A}^0$ and $\mathcal{M}^{\mathrm{R}}$ satisfies $(\mathrm{C}^{\mathrm{R}}_{\mathcal{M}})$. Given a set $\mathcal{A}^j$ for $j = 1, 2, \ldots$, write $\sup^j$ for the supremum over quantities in $\mathcal{A}^j$, and note that

$$\sup{}^1 \big\{\mu(\mathcal{M}^{\mathrm{R}}) + \mu_Q(\mathcal{M}^{\mathrm{R}})\big\} = O\big(h^3\big) \tag{5.2}$$

as $h \to 0$. For example, to prove that $\sup^1 \mu(\mathcal{M}^{\mathrm{R}}) = O(h^3)$ we note that

$$\begin{aligned} \mu(\mathcal{M}^{\mathrm{R}}) &= \int_{\mathcal{S}} \|x - y^{\mathrm{R}}\|\, K(\|u^{\mathrm{R}} - y^{\mathrm{R}}\|/h)\, K(\|u^{\mathrm{R}} - x\|/h)\, f(x)\, dx \\ &\le 3\,h\,(\sup K) \int_{\mathcal{S}} K(\|u^{\mathrm{R}} - x\|/2h)\, f(x)\, dx \\ &\le 12\,h^3\,(\sup K)\,(\sup f) \int K(\|z\|)\, dz = O(h^3)\,. \end{aligned} \tag{5.3}$$

(The first inequality follows from the fact that $K(\|u^{\mathrm{R}} - y^{\mathrm{R}}\|/h)\, K(\|u^{\mathrm{R}} - x\|/h) \ne 0$ implies $\|u^{\mathrm{R}} - y^{\mathrm{R}}\| \le h$ and $\|u^{\mathrm{R}} - x\| \le 2h$, whence $\|x - y^{\mathrm{R}}\| \le 3h$.)

By construction of p_Q, and property (C_f) (particularly the fact that f has a bounded derivative on $\mathcal{S}$),

$$\sup{}^1 \big|\mu(\mathcal{M}^{\mathrm{R}}) - \mu_Q(\mathcal{M}^{\mathrm{R}})\big| = O\big(h^4\big)\,. \tag{5.4}$$

(This result follows from a short Taylor expansion of (5.3) around (5.1).) Let $\mathcal{M}^{\mathrm{R}}_0$ denote the unique line segment that satisfies $(\mathrm{C}^{\mathrm{R}}_{\mathcal{M}})$, lies above $\partial\mathcal{S}$, is distant just

C_1h^2 from $\partial\mathcal{S}$ at its closest point, and is inclined at the same angle as the tangent to $\mathcal{S}$ at Q. (Provided $\mathcal{L}$ intersects $\partial\mathcal{S}$ at an angle sufficiently close to $\frac{\pi}{2}$, $\mathcal{M}_0^{\rm R}$ is well defined.) The condition on angle intersection is equivalent to $|\phi| > \Delta$, where Δ is sufficiently small. However, the condition is important only at the very beginning of the procedure, where $\mathcal{L} = \mathcal{L}_1$, since at all later steps ϕ is a realisation of a random variable whose absolute value is less than any given positive Δ with probability $O(\nu^{-\lambda})$, for all $\lambda > 0$. There is clearly no more than a polynomial number of steps; our proof will show that the number is in fact no more than $O(h^{-1})$.

In view of (5.4),

$$\sup{}^1 \left|\{\mu(\mathcal{M}^{\rm R}) - \mu(\mathcal{M}_0^{\rm R})\} - \{\mu_Q(\mathcal{M}^{\rm R}) - \mu_Q(\mathcal{M}_0^{\rm R})\}\right| = O\big(h^4\big)\,. \tag{5.5}$$

Let $\alpha = \frac{1}{2}$ or 1, and $C_2 > 0$. Write $\inf^2_{\alpha,C_2}$ to denote the infimum over the set $\mathcal{A}^2(\alpha, C_2)$ of triples $(Q, \psi, \mathcal{M}^{\rm R})$ such that $(Q, \psi) \in \mathcal{A}^0$, $\mathcal{M}^{\rm R}$ satisfies $({\rm C}^{\rm R}_{\mathcal{M}})$, and the absolute value of the difference between the angles of inclination of $\mathcal{M}^{\rm R}$ and $\mathcal{M}_0^{\rm R}$ exceeds C_2h^α. It may be deduced from the integral at (5.1) that

$$\inf{}^2_{\alpha,C_2} \{\mu_Q(\mathcal{M}^{\rm R}) - \mu_Q(\mathcal{M}_0^{\rm R})\} \geq Ch^{\alpha+3} + o\big(h^{\alpha+3}\big)\,,$$

where $C = C(C_2) > 0$ increases without bound as C_2 increases. From this property and (5.5) it follows that

$$\inf{}^2_{\alpha,C_2} \{\mu(\mathcal{M}^{\rm R}) - \mu(\mathcal{M}_0^{\rm R})\} \geq C_3h^{\alpha+3} + o\big(h^{\alpha+3}\big)\,, \tag{5.6}$$

where $C_3 = C_3(C_2) > 0$ for sufficiently large $C_2 > 0$. This result is clear when $\alpha = \frac{1}{2}$, and in fact $C_3 = C$ there. To treat $\alpha = 1$, let Ah^4 denote an upper bound to the right-hand side of (5.5). Then (5.6) holds if its right-hand sided is altered to $\{C(C_2) - A\}h^4 + o(h^4)$. Taking C_2 so large that $C(C_2) \geq 2A$ we obtain (5.6) in its stated form. For both $\alpha = \frac{1}{2}, 1$ the value of C_3 increases without bound as C_2 increases.

Let $\mathcal{A}^3 = \mathcal{A}^3(C_2)$ denote the class of quadruples $(Q, \psi, \mathcal{M}^{\rm R}_{(1)}, \mathcal{M}^{\rm R}_{(2)})$ such that $(Q, \psi, \mathcal{M}^{\rm R}_{(i)}) \in \mathcal{A}^1$ for $i = 1, 2$, and the absolute value of the difference between the inclinations of $\mathcal{M}_{(1)}$ and $\mathcal{M}_{(2)}$ is no greater than $C_2 h^{1/2}$. Let N represent the number of points in $\mathcal{X}$, and put

$$\widehat{\mu}(\mathcal{M}^{\mathrm{R}}) = N^{-1} \sum_{i=1}^{N} \|X_i - Y_i^{\mathrm{R}}\| \, K(\|u^{\mathrm{R}} - Y_i^{\mathrm{R}}\|/h) \, K(\|u^{\mathrm{R}} - X_i\|/2h) \, .$$

(This is a normalised form of the second series at (2.1).) Then $E\{\widehat{\mu}(\mathcal{M}^{\mathrm{R}})|N\} = \mu(\mathcal{M}^{\mathrm{R}})$. (Here we have used the assumption that, in the relation $p = c_0 f$, $c_0 = 1$.) Bernstein's inequality implies that for $t > 0$ and all $(Q, \psi, \mathcal{M}^{\mathrm{R}}) \in \mathcal{A}^1$,

$$P\big\{|\widehat{\mu}(\mathcal{M}^{\mathrm{R}}) - \mu(\mathcal{M}^{\mathrm{R}})| > N^{-1/2}\, t h^2 \,\big|\, N\big\} \leq 2 \exp\big\{-D_1 t^2 / \big(1 + N^{-1/2} t h^3\big)\big\} \, , \quad (5.7)$$

where $D_1 > 0$ depends only on f, K and Δ, the latter through the definition of $\mathcal{A}^1$. (To derive (5.7), note that conditional on N, $\widehat{\mu}(\mathcal{M}^{\mathrm{R}})$ equals a sum of independent random variables with mean $\mu(\mathcal{M}^{\mathrm{R}})$.)

Similarly, for $t > 0$ and all $(Q, \psi, \mathcal{M}^{\mathrm{R}}_{(1)}, \mathcal{M}^{\mathrm{R}}_{(2)}) \in \mathcal{A}^3$,

$$\begin{aligned} P\Big[\big|\{\widehat{\mu}(\mathcal{M}^{\mathrm{R}}_{(1)}) - \widehat{\mu}(\mathcal{M}^{\mathrm{R}}_{(2)})\} - \{\mu(\mathcal{M}^{\mathrm{R}}_{(1)}) - \mu(\mathcal{M}^{\mathrm{R}}_{(2)})\}\big| > N^{-1/2}\, t\, h^{5/2} \,\Big|\, N\Big] \\ \leq 2 \exp\big\{ - D_2 t^2 / \big(1 + N^{-1/2} t h^{7/2}\big)\big\} \, , \qquad (5.8) \end{aligned}$$

where $D_2 > 0$ depends only on f, K, Δ and C_2.

Recall from assumption (C_h) that $h \geq B_3 (\nu^{-1} \log \nu)^{1/3}$ for a large constant B_3. If $C_4 = \int f$ then the probability that $|N - C_4 \nu| > \frac{1}{2} C_4 \nu$ equals $O(\nu^{-\lambda})$ for all $\lambda > 0$. Therefore, taking $t = \frac{1}{2} C_3 \, (\frac{1}{2} C_4)^{1/2} \nu^{1/2} \, h^{3/2}$ we deduce from (5.6) (with $\alpha = \frac{1}{2}$) and (5.7) that given $\lambda > 0$ we have for all sufficiently large B_3, and a constant $C_5 > 0$,

$$P\big\{\widehat{\mu}(\mathcal{M}^{\mathrm{R}}) \leq \widehat{\mu}(\mathcal{M}^{\mathrm{R}}_0) + C_5 h^{7/2}\big\} = O\big(\nu^{-\lambda}\big) \qquad (5.9)$$

uniformly in $(Q, \psi, \mathcal{M}^{\mathrm{R}}) \in \mathcal{A}^2(\frac{1}{2}, C_2)$. Exploiting the smoothness of $\widehat{\mu}$ as a function of $\mathcal{M}^{\mathrm{R}}$, and choosing B_3 larger, we may use a standard continuity argument (involving approximation by a polynomially large number of elements of $\mathcal{A}^2$) to show from this result that

$$P\Big\{ \widehat{\mu}(\mathcal{M}^{\mathrm{R}}) \leq \widehat{\mu}(\mathcal{M}^{\mathrm{R}}_0) \quad \text{for some} \quad (Q, \psi, \mathcal{M}^{\mathrm{R}}) \in \mathcal{A}^2(\tfrac{1}{2}, C_2) \Big\} = O\big(\nu^{-\lambda}\big) \, . \quad (5.10)$$

In this and the next two paragraphs we treat a class of triples $(Q, \psi, \mathcal{M}^{\mathrm{R}})$ which does not intersect $\mathcal{A}^2(\frac{1}{2}, C_2)$. Let $\mathcal{A}^4$ denote the analogue of $\mathcal{A}^3(C_2)$ for triples; it is the set of $(Q, \psi, \mathcal{M}^{\mathrm{R}})$ such that $(Q, \psi) \in \mathcal{A}^0$, $\mathcal{M}^{\mathrm{R}}$ satisfies $(\mathrm{C}^{\mathrm{R}}_{\mathcal{M}})$, and the absolute value of the difference between the angles of inclination of $\mathcal{M}^{\mathrm{R}}$ and $\mathcal{M}^{\mathrm{R}}_0$ does not exceed $C_2 h^{1/2}$. Letting $t = \xi \nu^{1/2} h^{3/2}$ for $\xi > 0$ fixed but arbitrarily small, and taking $(Q, \psi, \mathcal{M}^{\mathrm{R}}_{(1)}) \in \mathcal{A}^4$ and $\mathcal{M}^{\mathrm{R}}_{(2)} = \mathcal{M}^{\mathrm{R}}_0$ in (5.8), we deduce from (5.8) that for any given $\eta, \lambda > 0$ we may choose B_3 (in condition (C_h)) so large that

$$\sup{}^4\, P\Big[\big|\big\{\widehat{\mu}(\mathcal{M}^{\mathrm{R}}) - \widehat{\mu}(\mathcal{M}^{\mathrm{R}}_0)\big\} - \big\{\mu(\mathcal{M}^{\mathrm{R}}) - \mu(\mathcal{M}^{\mathrm{R}}_0)\big\}\big| > \eta\, h^4\Big] = O\big(\nu^{-\lambda}\big)\,. \quad (5.11)$$

Replace (C_2, C_3) (in the context of (5.6)) by (C_2', C_3'), to distinguish the case $\alpha = 1$ from $\alpha = \frac{1}{2}$ treated in the previous paragraph; apply (5.6) for $\alpha = 1$; and note the remark immediately following (5.6). Arguing in this way we may deduce that if $C_6 > 0$ is given then $C_3' \geq C_6$ for sufficiently large C_2', and thence for such values of C_2' in the definition of $\mathcal{A}^2(1, C_2')$,

$$\inf{}^2_{1,C_2'}\, \big\{\mu(\mathcal{M}^{\mathrm{R}}) - \mu(\mathcal{M}^{\mathrm{R}}_0)\big\} \geq C_6 h^4 + o\big(h^4\big)\,. \quad (5.12)$$

Let $\mathcal{A}^5(C_2, C_2')$ denote the set of triples $(Q, \psi, \mathcal{M}^{\mathrm{R}})$ such that $(Q, \psi) \in \mathcal{A}^0$, $\mathcal{M}^{\mathrm{R}}$ satisfies $(\mathrm{C}^{\mathrm{R}}_{\mathcal{M}})$, and the absolute value of the difference between the angles of inclination of $\mathcal{M}^{\mathrm{R}}$ and $\mathcal{M}^{\mathrm{R}}_0$ lies between $C_2 h^{1/2}$ and $C_2' h$. Combining (5.11) and (5.12) we deduce that for a constant $C_7 > 0$,

$$\sup{}^5\, P\big\{\widehat{\mu}(\mathcal{M}^{\mathrm{R}}) \leq \widehat{\mu}(\mathcal{M}^{\mathrm{R}}_0) + C_7\, h^4\big\} = O\big(\nu^{-\lambda}\big)\,. \quad (5.13)$$

Applying to (5.13) the continuity argument that produced (5.10) from (5.9) we obtain on the present occasion,

$$P\Big\{\,\widehat{\mu}(\mathcal{M}^{\mathrm{R}}) \leq \widehat{\mu}(\mathcal{M}^{\mathrm{R}}_0) \quad \text{for some} \quad (Q, \psi, \mathcal{M}^{\mathrm{R}}) \in \mathcal{A}^5(C_2, C_2')\,\Big\} = O\big(\nu^{-\lambda}\big)\,. \quad (5.14)$$

Note that $\mathcal{A}^5(C_2, C_2')$ does not intersect $\mathcal{A}^2(\frac{1}{2}, C_2)$. The latter set appeared in the analogue (5.10) of (5.14).

Let $\mathcal{A}^6(C_2')$ denote the set of triples $(Q, \psi, \mathcal{M}^{\rm R})$ such that $(Q, \psi) \in \mathcal{A}^0$, $\mathcal{M}^{\rm R}$ satisfies $({\rm C}^{\rm R}_{\mathcal{M}})$, and the absolute value of the difference between the angles of inclination of $\mathcal{M}^{\rm R}$ and $\mathcal{M}^{\rm R}_0$ exceeds $C_2'h$. Combining the complementary results (5.10) and (5.14) we see that given any $C_1, \lambda > 0$ we may choose B_3 (in condition $({\rm C}_h)$) and C_2' so large that

$$P\Big\{\, \widehat{\mu}(\mathcal{M}^{\rm R}) \leq \widehat{\mu}(\mathcal{M}^{\rm R}_0) \quad \text{for some} \quad (Q, \psi, \mathcal{M}^{\rm R}) \in \mathcal{A}^6(C_2') \,\Big\} = O\big(\nu^{-\lambda}\big)\,. \qquad (5.15)$$

Finally we convert this result to one for a line segment $\widehat{\mathcal{N}}^{\rm R}$ that is similar to $\widehat{\mathcal{M}}^{\rm R}$, defined in section 2.2. Recall that $\mathcal{L} = \mathcal{L}(Q, \psi)$ denote a line that intersects $\partial\mathcal{S}_j(h)$ at a point Q and whose normal makes angle $\psi \in (-\pi/2, \pi/2)$ to the tangent to $\partial\mathcal{S}_j$ at Q. Let $\mathcal{N}^{\rm R}$ be of length $2h$, lie to the right of $\mathcal{L}$ with its left end on $\mathcal{L}$, and be placed so that no point of $\mathcal{X}$ lies above it, at least one point of $\mathcal{X}$ lies on it, and the acute angle ϕ that it makes to $\mathcal{L}$ satisfies $|\phi| > \Delta$. Call these conditions $({\rm C}^{\rm R}_{\mathcal{N}})$. Let $\mathcal{A}^7$ represent the class of segments $\mathcal{N}^{\rm R}$ that satisfy $\widehat{\mu}(\mathcal{N}^{\rm R}) \leq \widehat{\mu}(\mathcal{M}^{\rm R}_0)$. If $\lambda > 0$ is given then we may choose C_1 (in the definition of $({\rm C}^{\rm R}_{\mathcal{M}})$) so large that

$$P\Big\{\text{all segments } \mathcal{N}^{\rm R} \text{ in } \mathcal{A}^7 \text{ satisfying } ({\rm C}^{\rm R}_{\mathcal{N}}) \text{ also satisfy } ({\rm C}^{\rm R}_{\mathcal{M}})\Big\} = 1 - O\big(\nu^{-\lambda}\big). \qquad (5.16)$$

See section 5.3 for a proof. Given $C_8 > 0$, denote by $\mathcal{A}^8(C_8)$ the set of triples $(Q, \psi, \mathcal{N}^{\rm R})$ such that $(Q, \psi) \in \mathcal{A}^0$, $\mathcal{N}^{\rm R}$ satisfies $({\rm C}^{\rm R}_{\mathcal{N}})$, and the absolute value of the difference between the angles of inclination of $\mathcal{N}^{\rm R}$ and $\mathcal{M}^{\rm R}_0$ exceeds C_8h. It follows from (5.15) and (5.16) that for sufficiently large B_3, C_1 and C_8,

$$P\Big\{\, \widehat{\mu}(\mathcal{N}^{\rm R}) \leq \widehat{\mu}(\mathcal{M}^{\rm R}_0) \quad \text{for some} \quad (Q, \psi, \mathcal{N}^{\rm R}) \in \mathcal{A}^8(C_8) \,\Big\} = O\big(\nu^{-\lambda}\big)\,.$$

Therefore, if $\widehat{\mathcal{N}}^{\rm R}$ denotes the minimiser of $\widehat{\mu}(\mathcal{N}^{\rm R})$ over line segments $\mathcal{N}^{\rm R}$ that satisfy $({\rm C}^{\rm R}_{\mathcal{N}})$, and if anglediff$(\widehat{\mathcal{N}}^{\rm R}, \mathcal{M}^{\rm R}_0)$ represents the difference between the angles of inclination of $\widehat{\mathcal{N}}^{\rm R}$ and $\mathcal{M}^{\rm R}_0$, then

$$P\Big\{ \sup_{(Q,\psi)\,:\,(Q,\psi)\in\mathcal{A}^0} \big|\text{anglediff}\big(\widehat{\mathcal{N}}^{\rm R}, \mathcal{M}^{\rm R}_0\big)\big| > C_8h \Big\} = O\big(\nu^{-\lambda}\big)\,. \qquad (5.17)$$

5.3. Derivation of (5.16). Let $\text{segs}(\mathcal{N})$ denote the set of line segments $\mathcal{M}$ that satisfy $(\mathrm{C}^{\mathrm{R}}_{\mathcal{N}})$ and intersect $\partial\mathcal{S}$. For each $\mathcal{M} \in \text{segs}(\mathcal{N})$, denote by $\text{angle}(\mathcal{M})$ the angle that $\mathcal{M}$ makes to $\partial\mathcal{S}$, and let $D(\mathcal{M})$ denote the maximum distance that $\mathcal{M}$ protrudes below $\partial\mathcal{S}$. Let $\text{segs}(\mathcal{N} \mid C_9)$ denote the set of all $\mathcal{M} \in \text{segs}(\mathcal{N})$ for which $\text{angle}(\mathcal{M}) \le C_9 h$. In the next two paragraphs we shall outline a proof that if $C_9, \lambda > 0$ are given, and $C_1 = C_1(C_9, \lambda) > 0$ is sufficiently large, then

$$P\left\{ \sup_{\mathcal{M}\,:\,\mathcal{M}\in\text{segs}(\mathcal{N}|C_9)} D(\mathcal{M}) > C_1\, h^2 \right\} = 1 - O\big(\nu^{-\lambda}\big)\,. \tag{5.18}$$

Let $\xi, \eta \in (0,1)$ be arbitrarily small, and place rectangles with dimensions $\xi h \times (1-\eta) C_1 h^2$ below $\partial\mathcal{S}$ in a regular fashion, with their long sides parallel to $\partial\mathcal{S}$ and their short sides in the perpendicular direction, and such that adjacent rectangles touch one another and $\partial\mathcal{S}$ at edges or corners. The total number of rectangles involved equals $O(h^{-1})$, and the probability that any particular rectangle contains at least one point of $\mathcal{X}$ equals $1 - \exp\{-\xi(1-\eta)c\nu C_1 h^3\}$, where c is as in section 5.1. Therefore if C_1 is sufficiently large, depending on ξ, η and λ, then the probability of the event $\mathcal{F}$ that each rectangle contains at least one point of $\mathcal{X}$ equals $1 - O(\nu^{-\lambda})$.

Suppose ξ is sufficiently small, depending on C_9 and on the value of Δ in the assumption that the absolute value of the acute angle ϕ that a line in $(\mathrm{C}^{\mathrm{R}}_{\mathcal{N}})$ makes to $\mathcal{L}$ must exceed Δ. Then for all sufficiently small h, whenever $\mathcal{F}$ holds, any segment satisfying $(\mathrm{C}^{\mathrm{R}}_{\mathcal{N}})$ and intersecting $\partial\mathcal{S}$ at a steeper angle than $C_9 h$ cannot protrude more than $C_1 h^2$ below $\partial\mathcal{S}$. (The condition that h be small, and also the factor $1-\eta$ in the definition of the longer side length, are needed to overcome slight anomalies caused by the sides of the rectangles being straight and the boundary $\partial\mathcal{S}$ being possibly curved.) Result (5.18) follows from this result and that in the previous paragraph.

However, (5.18) does not hold for steep lines, which may protrude relatively deeply. We deal with that case in two parts. The first is treated by the following result, which has a similar proof to (5.18): for all $\lambda > 0$,

$$P\Big\{ \sup_{\mathcal{M}\,:\,\mathcal{M}\in \mathrm{segs}(\mathcal{N})} D(\mathcal{M}) > 2h \Big\} = 1 - O\big(\nu^{-\lambda}\big)\,. \tag{5.19}$$

Next we consider line segments that protrude between $C_1 h^2$ and $2h$. Let $(\mathrm{C}^{\mathrm{R}}_{\mathcal{M}})'$ be the version of $(\mathrm{C}^{\mathrm{R}}_{\mathcal{M}})$ in which "$C_1 h^2$" is replaced by "$2h$" in the restriction that "the right-hand end [of $(\mathrm{C}^{\mathrm{R}}_{\mathcal{M}})$] lies within $C_1 h^2$ of $\partial\mathcal{S}$ [and] no part of it lies further than $C_1 h^2$ below $\partial\mathcal{S}$". Denote by $\mathcal{A}^9(C_9)$ the class of triples $(Q, \psi, \mathcal{M}^{\mathrm{R}})$ such that $(Q,\psi) \in \mathcal{A}^0$, $\mathcal{M}^{\mathrm{R}}$ satisfies $(\mathrm{C}^{\mathrm{R}}_{\mathcal{M}})'$, and the absolute value of the difference between the angles of inclination of $\mathcal{M}^{\mathrm{R}}$ and $\mathcal{M}^{\mathrm{R}}_0$ exceeds $C_9 h$. The argument leading to (5.15) continues to hold and now gives the following analogue of that result: For any $C_1, \lambda > 0$ we may choose B_3 and C_9 so large that

$$P\Big\{ \widehat{\mu}(\mathcal{M}^{\mathrm{R}}) \le \widehat{\mu}(\mathcal{M}^{\mathrm{R}}_0) \quad \text{for some} \quad (Q,\psi,\mathcal{M}^{\mathrm{R}}) \in \mathcal{A}^9(C_9) \Big\} = O\big(\nu^{-\lambda}\big)\,. \tag{5.20}$$

(As in the derivation of (5.15), the technique involves splitting $\mathcal{A}^9(C_9)$ into two non-overlapping sets of triples $(Q, \psi, \mathcal{M}^{\mathrm{R}})$, establishing the version of (5.20) in the case where $\mathcal{A}^9(C_9)$ is replaced by either of these sets, and adding the result.)

The desired result (5.16) follows from (5.18)–(5.20).

5.4. Left and right smooths. Define $\widehat{\mathcal{N}}^{\mathrm{R}}$ as in section 5.2, let $\widehat{Z}^{\mathrm{R}}$ denote the point at which $\widehat{\mathcal{N}}^{\mathrm{R}}$ intersects $\mathcal{L}(Q,\psi)$, and let z be the vector represented by Q. Result (5.16) implies that if $\lambda > 0$ is given, and C_1 is chosen sufficiently large, then the probability that $\widehat{\mathcal{N}}^{\mathrm{R}}$ protrudes no further than $C_1 h^2$ below $\cup_j\, \partial\mathcal{S}_j$ equals $1 - O(\nu^{-\lambda})$. It follows from this property and (5.18) that for sufficiently large B_3, C_1 and C_9,

$$P\big\{ \|\widehat{Z}^{\mathrm{R}} - z\| > C_9\, h^2 \quad \text{for some} \quad (Q,\psi) \in \mathcal{A}^0 \big\} = O\big(\nu^{-\lambda}\big)\,. \tag{5.21}$$

Next we consider the uniform accuracy of tangent-angle estimators. Result (5.17) has of course an analogue for left-hand tangent-angle estimators. Both left and right estimators are identical to those described in section 2.1, except for the constraint that "the acute angle ϕ that $\mathcal{N}^{\mathrm{R}}$ makes to $\mathcal{L}$ satisfies $|\phi| > \Delta$"; see the definition of condition ($\mathrm{C}^{\mathrm{R}}_{\mathcal{N}}$) below (5.15). If Δ is sufficiently small (as assumed in Theorem 3.1) then this constraint will be fulfilled at the initial step, where $\widehat{Q}_1$ is calculated using the transect $\mathcal{L}_1$. Given that this is the case then in all other steps, uniformly along smooth parts of the curve (that is, for $(Q,\psi) \in \mathcal{A}^0$ in the case of right-hand tangent-angle estimators, and analogously for left-hand estimators), (5.17) and its left-hand analogue imply that the condition stated just above in quotation marks is satisfied with probability $1 - O(\nu^{-\lambda})$, for any given $\lambda > 0$ provided the constants are chosen sufficiently large. Since Theorem 3.1 refers only to events whose probability of not occurring is of order $O(\nu^{-\lambda})$, then we have proved that, uniformly along smooth parts of the curve, there exists a constant $C_{10} > 0$ such that, with probability $1 - O(\nu^{-\lambda})$, both

$$\sup_{(Q,\psi)\in\mathcal{A}^0} |\widehat{\omega}^{\mathrm{R}} - \omega^{\mathrm{R}}| \le C_{10} h\,, \tag{5.22}$$

and its analogue for left-hand tangent-angle estimators, hold.

Results (5.21) and (5.22) establish properties (III) and (IV) in event $\mathcal{E}(C)$, except in $O(h^2)$ neighbourhoods of corners. (That case will be treated in section 5.5.) It also implies that if $\lambda > 0$ is given then for some $C_{11} = C_{11}(\lambda) > 0$, with probability $1 - O(\nu^{-\lambda})$ the number of steps taken to traverse each smooth segment of the boundary (i.e. from P_j to P_{j+1}, or the fragment of that curve which we estimate when $j = 0$ or $k-1$), is bounded by $C_{11} h^{-1}$. It will follow from results in section 5.5 that with the same probability, no more than a bounded number of steps is spent negotiating each corner. Together these results imply property (VI) in event $\mathcal{E}(C)$.

Result (5.22), and its left-hand counterpart, imply that for some $C_{12} > 0$ and with probability $1 - O(\nu^{-\lambda})$, $|\widehat{\omega}^{\mathrm{L}} - \widehat{\omega}^{\mathrm{R}}| \le C_{12} h$ for each step along smooth parts of the curve. We take B_1 and B_2, in the definitions of the weight $\hat{\rho}$ and the tangent-angle threshold $B_2 h$, to be constants exceeding C_{10}. Choosing B_1 larger ensures smoother right-to-left-hand switches of smoothing algorithms, as discussed in section 2.5.

5.5. Corners. Recall that $\mathcal{L} = \mathcal{L}(Q,\psi)$ denotes a line that intersects $\partial\mathcal{S}$ at a point

Q and whose normal makes angle $\psi \in (-\pi/2, \pi/2)$ to the tangent to $\partial\mathcal{S}$ at Q. In this section we allow Q to be a corner; at such locations there are of course two values of ψ.

Versions of the results in section 5.2 may be developed in arbitrarily close neighbourhoods of corners, with the proviso that the line segment $\mathcal{M}_0^{\mathrm{R}}$ introduced there is redefined as the minimiser of $\mu(\mathcal{M}^{\mathrm{R}})$ over line segments $\mathcal{M}^{\mathrm{R}}$ that intersect $\cup_j \partial\mathcal{S}_j$ and have area $C_{13}h^3$ of $\mathcal{S}$ between the segment and the boundary, where "between" means in the direction perpendicular to the boundary, and C_{13} is a large positive constant. With the new definition of $\mathcal{M}_0^{\mathrm{R}}$ we may prove instead of (5.15) that if $C_{14} > 0$ is sufficiently large then

$$P\Big\{ \widehat{\mu}(\mathcal{M}^{\mathrm{R}}) \le \widehat{\mu}(\mathcal{M}_0^{\mathrm{R}}) \quad \text{for some} \quad (Q, \psi, \mathcal{M}^{\mathrm{R}}) \in \mathcal{A}^{10}(C_{14}) \Big\} = O\big(\nu^{-\lambda}\big)\,, \quad (5.23)$$

where $\mathcal{A}^{10}(C_{14})$ denotes the set of triples $(Q, \psi, \mathcal{M}^{\mathrm{R}})$ such that $(Q, \psi) \in \mathcal{A}^{11}$ and the absolute value of the difference between the angles of inclination of $\mathcal{M}^{\mathrm{R}}$ and $\mathcal{M}_0^{\mathrm{R}}$ exceeds $C_{14}h$, and $\mathcal{A}^{11}$ is the set of pairs (Q, ψ) such that $Q \in \cup_j \partial\mathcal{S}_j$ and is distant at least $2h$ from both P_0 and P_k, and $\psi \in [-\frac{1}{2}\pi + \Delta, \frac{1}{2}\pi - \Delta]$. (The latter restriction can be ill defined within distance δh, say, of a corner $(0 < \delta < 2)$, but there it can be greatly relaxed since there is a much wider range of orientations of $\mathcal{M}^{\mathrm{R}}$ for which, with probability $1 - O(\nu^{-\lambda})$, points of $\mathcal{S}$ lie below the segment.) Derivation of (5.23) uses the same technique employed to establish (5.15) and (5.22). It involves splitting the set of triples $(Q, \psi, \mathcal{M}^{\mathrm{R}})$ — here $\mathcal{A}^{10}(C_4)$ — into two nonoverlapping subsets, bounding the probability for each, and adding the results.

To convert (5.23) to a result for a line segment $\widehat{\mathcal{N}}^{\mathrm{R}}$ that is similar to $\widehat{\mathcal{M}}^{\mathrm{R}}$, let condition $(\mathrm{C}_{\mathcal{N}}^{\mathrm{R}})$ on line segments $\mathcal{N}^{\mathrm{R}}$ be as in section 5.2, and let $\widehat{\mathcal{N}}^{\mathrm{R}}$ be as defined there. The following analogue of (5.17) may be proved via (5.23): if $\lambda > 0$ is given then for sufficiently large B_3, C_1 and C_{15},

$$P\bigg\{ \sup_{(Q,\psi)\,:\,(Q,\psi)\in\mathcal{A}^{11}} \big|\mathrm{anglediff}\big(\widehat{\mathcal{N}}^{\mathrm{R}}, \mathcal{M}_0^{\mathrm{R}}\big)\big| > C_{15}h \bigg\} = O\big(\nu^{-\lambda}\big)\,. \qquad (5.24)$$

Define $\widehat{\mathcal{N}}^{\mathrm{L}}$ analogously to $\widehat{\mathcal{N}}^{\mathrm{R}}$, let $\hat{\xi}^{\mathrm{L}}$ and $\hat{\xi}^{\mathrm{R}}$ be the respective angles of inclination of $\widehat{\mathcal{N}}^{\mathrm{L}}$ and $\widehat{\mathcal{N}}^{\mathrm{R}}$, and let ξ^{L} and ξ^{R} be the corresponding angles for $\mathcal{M}_0^{\mathrm{L}}$ and $\mathcal{M}_0^{\mathrm{R}}$.

Put $\chi = |\xi^{\mathrm{L}} - \xi^{\mathrm{R}}|$ and $\widehat{\chi} = |\hat{\xi}^{\mathrm{L}} - \hat{\xi}^{\mathrm{R}}|$. We may deduce from (5.24), and its analogue in the left-hand case, that for some $C_{16} > 0$,

$$P\Big\{|\widehat{\chi} - \chi| > C_{16}\, h \text{ for some } (Q,\psi) \in \mathcal{A}^{11}\Big\} = O\big(\nu^{-\lambda}\big)\,. \tag{5.25}$$

If A and B are events, let $P(A \,\|\, B) = P(A \cap B)$. Given a point $Q \in \partial\mathcal{S}$, let $d(Q)$, $d^{\mathrm{L}}(Q)$ and $d^{\mathrm{R}}(Q)$ denote the distance from Q to the nearest corner of $\partial\mathcal{S}$, to the nearest corner on the left, and to the nearest corner on the right, respectively. Let ω denote the true tangent angle at Q; it is well defined except when Q is a corner. Put $\chi^{\mathrm{L}} = |\xi^{\mathrm{L}} - \omega|$ and $\chi^{\mathrm{R}} = |\xi^{\mathrm{R}} - \omega|$. Algebraic and geometric arguments may be used to show that for all sufficiently l arge ν: (a) there exists $C_{17} > 0$ such that $\chi^{\mathrm{L}} \le C_{17}h$ [respectively, $\chi^{\mathrm{R}} \le C_{17}h$] if Q is more than $2h$ from the nearest corner on the left [right]; and (b) if $C_{18} > 0$ is sufficiently large then there exist $C_{19}, C_{20} > 0$ such that (α) $\chi^{\mathrm{L}} \le C_{19}h$ and $\chi^{\mathrm{R}} \le C_{19}h$ if $\chi \le C_{18}h$, and (β) the distance from Q to the nearest corner is less than $C_{20}h$ if $\chi > C_{18}h$. Combining these properties with (5.24), (5.25) and the left-hand analogue of (5.24), we deduce that for some $C_{21} > 0$, and given any sufficiently large $C_{22} > 0$, there exist $C_{23}, C_{24} > 0$ such that

$$\begin{aligned}
P\Big\{\sup_{(Q,\psi)} |\hat{\xi}^{\mathrm{L}} - \omega| > C_{21}\, h \quad\text{and}\quad d^{\mathrm{L}}(Q) > 2h\Big\} &= O\big(\nu^{-\lambda}\big)\,,\\
P\Big\{\sup_{(Q,\psi)} |\hat{\xi}^{\mathrm{R}} - \omega| > C_{21}\, h \quad\text{and}\quad d^{\mathrm{R}}(Q) > 2h\Big\} &= O\big(\nu^{-\lambda}\big)\,,\\
P\Big\{\sup_{(Q,\psi)\in\mathcal{A}^{11}} \big(|\hat{\xi}^{\mathrm{L}} - \omega| + |\hat{\xi}^{\mathrm{R}} - \omega|\big) > C_{23}h \quad\text{and}\quad \widehat{\chi} \le C_{22}\, h\Big\} &= O\big(\nu^{-\lambda}\big)\,,\\
P\Big\{\sup_{(Q,\psi)\in\mathcal{A}^{11}} d(Q) > C_{23}h \quad\text{and}\quad \widehat{\chi} > C_{22}\, h\Big\} &= O\big(\nu^{-\lambda}\big)\,,
\end{aligned} \tag{5.26}$$

Given a point Q and a set $\mathcal{R}$, let $D(Q,^{\mathrm{R}}\,)$ equal the infimum of distances from Q to points of $\mathcal{R}$. In this notation, the following result is a consequence of the manner of construction of point estimates $\widehat{Q}_j$: for some $C_{25} > 0$,

$$P\Big\{\sup_j D(\widehat{Q}_j, \partial\mathcal{S}) \le C_{25}\, h\Big\} = 1 - O\big(\nu^{-\lambda}\big)\,. \tag{5.27}$$

We can effectively equate $(\widehat{\mathcal{N}}^{\mathrm{L}}, \widehat{\mathcal{N}}^{\mathrm{R}})$ with $(\widehat{\mathcal{M}}^{\mathrm{L}}, \widehat{\mathcal{M}}^{\mathrm{R}})$, since (for any $\lambda > 0$) the probability that they are not equal on all occasions where they are computed can be made equal to $O(\nu^{-\lambda})$ by choosing the constants large. Likewise, we can effectively equate $\widehat{\chi}$ and $|\widehat{\omega}^{\rm L} - \widehat{\omega}^{\rm R}|$. Therefore, results (5.26) and (5.27) imply that the following properties hold with probability $1 - O(\nu^{-\lambda})$, uniformly in any polynomially large number of steps. (A) If a point estimate $\widehat{Q}_j$ is confirmed then both $\widehat{\omega}^{\rm L}$ and $\widehat{\omega}^{\rm R}$ differ from the true tangent angle by no more than $C_{26}h$. (B) If a point estimate $\widehat{Q}_j$ is not confirmed then that point is within $C_{27}h$ of a corner. (C) If $C_{28} > 0$ then the number of steps taken from when the point estimate is distant $C_{28}h$ to the left of a given corner $\mathcal{C}$, to when it is distant $C_{28}h$ to the right of $\mathcal{C}$, is no more than C_{29}.

Next we show that for each corner $\mathcal{C}$, the associated sequence of unconfirmed point estimates is a consecutive sequence, with probability $1 - O(\nu^{-\lambda})$. Call this property (D). (If (D) failed then we could mistakenly determine, with non-negligible probability, that more than one corner existed in that vicinity of $\mathcal{C}$, and in particular that property (I) in the definition of event $\mathcal{E}(C)$ could be violated.) It suffices to show that for each $\mathcal{C}$ and all sufficiently small $C_{30} > 0$, the probability that a sequence of point estimates $\ldots, \widehat{Q}_i, \widehat{Q}_{i+1}, \ldots$ all of which lie within distance $C_{30}h$ of $\mathcal{C}$, includes a subsequence for which the confirmation status has the order "unconfirmed, confirmed, ..., confirmed, unconfirmed", equals $O(\nu^{-\lambda})$. Bearing in mind the following consequence of the results in the previous paragraph:

$$P\left[\#\left\{\widehat{Q}_j \,:\, D(\widehat{Q}_j, \mathcal{C}) \le C_{28}\, h\right\} \le C_{29}\right] = 1 - O\big(\nu^{-\lambda}\big)\,. \tag{5.28}$$

we see that all cases may be treated in the same way as that where the sequence is of length three with the following confirmation statuses: "unconfirmed, confirmed, unconfirmed".

If the point estimates corresponding to the latter sequence are $\widehat{Q}_j, \widehat{Q}_{j+1}, \widehat{Q}_{j+2}$ then the estimated turning-angle differences for $\widehat{Q}_j$ and $\widehat{Q}_{j+2}$ have absolute values exceeding B_1h, whereas their counterpart for $\widehat{Q}_{j+1}$ does not exceed B_1h. For any $0 < \epsilon_1 < \epsilon < \epsilon_2$ it can be shown that with probability $1 - O(\nu^{-\lambda})$, the distance of $\widehat{Q}_{j+1}$ from each of $\widehat{Q}_j$ and $\widehat{Q}_{j+2}$ lies between $\epsilon_1 h$ and $\epsilon_2 h$. Also it may be proved that if ϵ_1 and ϵ_2 are sufficiently close to ϵ, and if C_{30} is sufficiently small, then for some $C_{31} > 0$, $C_{31}h$ is exceeded by the absolute value of the difference between two versions of $|\xi^{\mathrm{L}} - \xi^{\mathrm{R}}|$, computed at points Q and Q' (say) that are distant between $\epsilon_1 h$ and $\epsilon_2 h$ apart, and which are each distant no more than $C_{30}h$ from $\mathcal{C}$. (The value of C_{31} can be made arbitrarily large by choosing C_{30} small.) By selecting B_3 sufficiently large the Bernstein-inequality arguments in section 5.2 may be used to prove that with probability $1-O(\nu^{-\lambda})$, both $|\hat{\xi}^{\mathrm{L}} - \xi^{\mathrm{L}}| \le \frac{1}{3}C_{31}h$ and $|\hat{\xi}^{\mathrm{R}} - \xi^{\mathrm{R}}| \le \frac{1}{3}C_{31}h$, uniformly in pairs (Q, ψ) with Q no more than $C_{30}h$ from $\mathcal{C}$. Consequently,

$$\Big| |\hat{\xi}^{\mathrm{L}} - \hat{\xi}^{\mathrm{R}}| - |\xi^{\mathrm{L}} - \xi^{\mathrm{R}}| \Big| \le \tfrac{2}{3}\, C_{31}\, h$$

uniformly in the same range, with probability $1 - O(\nu^{-\lambda})$. Hence, with probability $1 - O(\nu^{-\lambda})$ the differences between the versions of $|\hat{\xi}^{\mathrm{L}} - \hat{\xi}^{\mathrm{R}}|$ for $\widehat{Q}_j$ and $\widehat{Q}_{j+1}$, and for $\widehat{Q}_{j+1}$ and $\widehat{Q}_{j+2}$, exceeds $\frac{1}{3}C_{31}h$. (Note result (5.28).)

This establishes the desired result about confirmation statuses of triples, and so proves (D). In conjunction with (A)–(C) noted three paragraphs above, it also completes the proof of property (I) in the definition of event $\mathcal{E}(C)$.

The $O(h^2)$ accuracy of point estimates computed using right smooths when Q is not closer than $2h$ to a corner on the right, was established at (5.21). It of course has an analogue for left smooths. Result (A) implies that with probability $1 - O(\nu^{-\lambda})$, tangent-angle estimates (we are focusing on those near a corner) are accurate to $O(h)$ for confirmed points, and so must be computed using a left smooth (to the left of a corner) or a right smooth (to the right). Result (D) implies that with probability $1-O(\nu^{-\lambda})$, the switch from a left smooth to a right smooth occurs only once for each corner. Therefore, with probability $1-O(\nu^{-\lambda})$, confirmed points either lie to the left of a corner, are computed using a left smooth, and are within $O(h^2)$ of the true boundary, or lie to the right, are computed via a right smooth, and are accurate to $O(h^2)$.

Therefore, with probability $1 - O(\nu^{-\lambda})$, the tangent-angle estimates used for extrapolation to corners are accurate to $O(h)$, and the point estimates from which the extrapolations are made are accurate to $O(h^2)$. Result (5.26) implies that those points are also within $O(h)$ of the true corners. It follows that the corner estimates obtained by extrapolation are accurate to within $O(h^2)$, and moreover that both the line segments used in the extrapolation are uniformly within $O(h^2)$ of the boundary. Call this result (E).

Property (II), in the definition of event $\mathcal{E}(C)$, follows from (E). Property (V) is implied by (B), and the proofs of (III), (IV) and (VI), which were commenced in section 5.4, are completed using (E), (A) and (C) respectively.

REFERENCES

AIGNER, D.J., LOVELL, C.A.K. AND SCHMIDT, P. (1977). Formulation and estimation of stochastic frontier models. *J. Econometrics* **6**, 21–37.

BAÍLLO, A. AND CUEVAS, A. (2001). On the estimation of a star-shaped set. *Adv. Appl. Probab.* **33**, 717–726.

CARLSTEIN, E. AND KRISHNAMOORTHY, C. (1992). Boundary estimation. *J. Amer. Statist. Assoc.* **87**, 430–438.

DEPRINS, D., SIMAR, L. AND TULKENS, H. (1984). Measuring labor inefficiency in post offices. In *The Performance of Public Enterprises: Concepts and measurements*. M. Marchand, P. Pestieau and H. Tulkens (eds.), Amsterdam, North-Holland, 243–267.

DESHAYES, J. AND PICARD, D. (1981). Convergence de processus à double indice: application aux tests de rupture dans un modèle. *Comptes Rendus* **292**, 642–669.

DONOHO, D.L. (1999) Wedgelets: nearly minimax estimation of edges. *Ann. Statist.* **27**, 859–897.

FARRELL, M.J. (1957). The measurement of productive efficiency. *J. Roy. Statist. Soc.* Ser. A **120**, 253–281.

GIJBELS, I., MAMMEN, E., PARK, B.U. AND SIMAR, L. (1999). On estimation of monotone and concave frontier functions. *J. Amer. Statist. Assoc.* **94**, 220–228.

HALL, P., NUSSBAUM, M. AND STERN, S.E. (1997). On the estimation of a support curve of indeterminate sharpness. *J. Multivariate Anal.* **62**, 204–232.

HALL, P., PARK, B.U. AND STERN, S.E. (1998). On polynomial estimators of frontiers and boundaries. *J. Multivariate Anal.* **66**, 71–98.

HALL, P., PENG, L. AND RAU, C. (2001). Local likelihood tracking of fault lines and boundaries. *J. Royal Statist. Soc. Ser. B* **63**, 569–582.

HALL, P. AND RAU, C. (2000). Tracking a smooth fault line in a response surface. *Ann. Statist.* **28**, 713–733.

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

HARTIGAN, J.A. (1987). Estimation of a convex density contour in two dimensions. *J. Amer. Statist. Assoc.* **82**, 267–270.

KNEIP, A., PARK, B.U. AND SIMAR, L. (1998). A note on the convergence of nonparametric DEA estimators for production efficiency scores. *Econometric Theory* **14**, 783–793.

KNEIP, A. AND SIMAR, L. (1996). A general framework for frontier estimation with panel data. *J. Productivity Anal.* **7**, 187–212.

KOROSTELEV, A., SIMAR, L. AND TSYBAKOV, A.B. (1995a). Efficient estimation of monotone boundaries. *Ann. Statist.* **23**, 476–489.

KOROSTELEV, A., SIMAR, L. AND TSYBAKOV, A.B. (1995b). On estimation of monotone and convex boundaries. *Pub. Inst. Stat. Univ. Paris* **39**, 3–18.

KOROSTELEV, A.P. & TSYBAKOV, A.B. (1993a). Estimating the support of a density and functionals of it. *Problems Inform. Transmission* **29**, 1–15.

KOROSTELEV, A.P. & TSYBAKOV, A.B. (1993b). Minimax linewise algorithm for image reconstruction. In *Computer Intensive Methods in Statistics*, Eds W. Härdle and L. Simar, pp. 113–33. Vienna: Physica.

KOROSTELEV, A.P. AND TSYBAKOV, A.B. (1993c). *Minimax Theory of Image Reconstruction.* Springer Lecture Notes in Statistics **82**. Springer, New York.

MAMMEN, E. AND TSYBAKOV, A.B. (1995). Asymptotical minimax recovery of sets with smooth boundaries. *Ann. Statist.* **23**, 502–524.

MÜLLER, H.-G. AND SONG, K.S. (1994). Maximin estimation of multidimensional boundaries. *J. Multivariate Anal.* **50**, 265–281.

O'SULLIVAN, F. AND QIAN, M. (1994). A regularized contrast statistic for object boundary estimation. *IEEE Trans. Patt. Anal. Machine Intell.* **16**, 561–570.

PARK, B.U., SICKLES, R.C. AND SIMAR, L (1998). Stochastic panel frontiers: a semiparametric approach. *J. Econometrics* **84**, 273–301.

PICARD, D. (1985). Testing and estimating change-points in time series. *Adv. Appl. Probab.* **17**, 841–867.

QIU, P. (1997). Nonparametric estimation of jump surface. *Sankhyā* Ser. A **59**, 268–294.

QIU, P. (1998). Discontinuous regression surfaces fitting. *Ann. Statist.* **26**, 2218–2245.

QIU, P. (2002). A nonparametric procedure to detect jumps in regression surfaces. *J. Comput. Graph. Statist.* **11**, 799–822.

QIU, P. (2004). The local piecewisely linear kernel smoothing procedure for fitting jump regression surfaces. *Technometrics* **46**, 87–98.

QIU, P. (2005). *Image Processing and Jump Regression Analysis.* Wiley, New York.

QIU, P. AND YANDELL, B. (1997). Jump detection in regression surfaces. *J. Comput. Graph. Statist.* **6**, 332–354.

RIPLEY, B.D. AND RASSON, J.-P. (1977). Finding the edge of a Poisson forest. *J. Appl. Probab.* **14**, 483–491.

RUDEMO, M. AND STRYHN, H. (1994). Approximating the distribution of maximum likelihood contour estimators in two-region images. *Scand. J. Statist.* **21**, 41–55.

SEIFORD, L.M. (1996). Data envelopment analysis: the evolution of the state-of-the-art (1978–1995). *J. Productivity Anal.* **7**, 99–137.

WAND, M.P. AND JONES, M.C. (1995). *Kernel Smoothing*. Chapman and Hall, London.