

**NONPARAMETRIC ESTIMATION OF
THE SINGULARITIES OF A SIGNAL
FROM NOISY MEASUREMENTS**

A.I. KATSEVICH, A.G. RAMM

ABSTRACT. We study a problem of locating and estimating singularities of a signal measured with noise on a discrete set of points (fixed-design model). The signal consists of a smooth part with bounded first derivative and of finite number of singularities of the type $(x - t_i)_{\pm}^p d_i, 0 \leq p \leq \frac{1}{2}$. The case $p = 0$ corresponds to a piecewise continuous function. The algorithm is based on convolving the data with a kernel having compact support. Optimal bandwidth of the kernel is calculated, the consistency of the algorithm is proved. The results of testing of the proposed algorithm on model examples are presented.

1.INTRODUCTION

The problem of locating singular points of a function is of interest in many applications [1-9]. For example, the recovery of discontinuity curves of a function $f(x), x \in \mathbb{R}^2$, from its Radon transform $\hat{f}(\alpha, t), 0 \leq \alpha \leq \pi, p \in \mathbb{R}$, leads to the problem of finding the singular locus of $\hat{f}(\alpha, t)$ [6,8], such that in the neighborhood of point $(\alpha, t_0^{(i)}(\alpha))$

$$\hat{f}(\alpha, t) = s(\alpha, t) + d(\alpha)(t - t_0^{(i)}(\alpha))_{\pm}^p, \quad p = \frac{1}{2}, \quad d(\alpha) \neq 0,$$

where $s(\alpha, t)$ is a smooth function, and

$$(x - t_0)_{\pm}^p = \begin{cases} |x - t_0|^p, & \text{if } \pm(x - t_0) \geq 0 \\ 0, & \text{if } \pm(x - t_0) \leq 0. \end{cases} \quad (1.1)$$

In this paper we develop an algorithm for locating and estimating singularities of a signal $s(x)$ measured with noise at a discrete equidistant set of points. The signal is supposed to have a finite number of singularities of the type (1.1) with known exponent $p, 0 \leq p \leq \frac{1}{2}$. The idea of the algorithm resembles that of the nonparametric kernel estimation [4] and consists of convolving the data with an odd piecewise constant kernel having compact support.

Similar problems have been studied in papers [2,3] using global approach based on modified smoothing splines. In the particular case $p = 0$ our algorithm reduces to the template matching technique which is used within a different framework in computer vision for edge detection[1,5].

The paper is organized as follows. In Section 2 we describe the algorithm and calculate the optimal kernel bandwidth. Section 3 is devoted to the investigation of basic properties of the proposed algorithm. We show that this algorithm is consistent in the following sense: for any a priori fixed probability $\epsilon, 0 < \epsilon < 1$, of the first kind error – "there is a singularity at a given point when there is none" and

1991 *Mathematics Subject Classification.* Primary 62G05.

Key words and phrases. Noisy data, singularity localization, kernel estimation.

The authors thank ONR for support and Prof. G.Wahba for pointing out several references.

for any "size" $d \neq 0$ of a singularity $d(x - x_0)_\pm^p$ the probability of missing it and the accuracy of its localization go to zero as the stepsize goes to zero. In Section 4 we briefly discuss the case $p = 0$ (localization and estimation of discontinuities of a piecewise differentiable signal) as a particular case of the general theory developed in Sections 2 and 3. The numerical results of testing of the proposed algorithm are presented in Section 5. The proofs of all lemmas from Section 3 are given in the Appendix.

2. FINDING SINGULARITIES OF A PIECEWISE-SMOOTH FUNCTION

We use the following model. Let the known data be

$$u_i = s_i + n_i, \quad s_i := s(x_i), \quad x_i = ih, \quad i \in \mathbb{Z}, \quad (2.1)$$

where $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, the random errors n_i are independent, identically distributed with

$$E(n_i n_j) = \begin{cases} \sigma^2, & \text{if } i = j \\ 0, & \text{if } i \neq j, \end{cases} \quad E(n_i) = 0. \quad (2.2)$$

The signal is assumed to be of the form

$$s(x) = \hat{s}(x) + \sum_k (x - t_k)_\pm^p d_k, \quad d_k = \text{const}, \quad 0 \leq p \leq \frac{1}{2}, \quad x \in \mathbb{R}^1.$$

Here $\hat{s}(x)$ is a smooth part of the signal with the known estimate of its first derivative

$$|\hat{s}'(x)| \leq C. \quad (2.3)$$

The points t_k will be called singularities, or singular points (that is the points at which one-sided limit of $s'(x)$ equals infinity if $p > 0$ or $s(x)$ has a jump if $p = 0$), d_k will be called coefficients of the singularities and p will be called the exponent of singularity. The singularities $(x - t_k)_\pm^p$ will be called right and left respectively, or "+" and "-" type singularities. We assume that t_k are located sufficiently far from each other. This restriction will be formulated precisely later. The problem is to find points t_k , values d_k and determine the type of singularity: if it is $(x - t_k)_+^p$ or $(x - t_k)_-^p$. First we derive an algorithm for the solution of this problem and then investigate its properties.

Let us assume that the singular point t_0 is at the middle of the interval $[x_j, x_{j+1}]$: $t_0 = \frac{x_j + x_{j+1}}{2}$ and consider the "+" type singularity

$$\begin{cases} s(x) = \hat{s}(x) + (x - t_0)^p d_0, & x \geq t_0 \\ s(x) = \hat{s}(x), & x < t_0, \end{cases} \quad 0 \leq p \leq \frac{1}{2}. \quad (2.4)$$

We do not include other singularities ($k \neq 0$) in (2.4), since we assume that they are located outside the neighborhood Δ_0 of t_0 in which we work. In particular, it is assumed that

$$|(\hat{s}(x) + \sum_{k \neq 0} (x - t_k)_\pm^p d_k)'| = |\hat{s}'(x) + \sum_{k \neq 0} \left(\pm \frac{p d_k}{(x - t_k)_\pm^{1-p}} \right)| \leq C, \quad x \in \Delta_0.$$

The last inequality holds if the distance between two neighboring singularities satisfies

$$|t_{k+1} - t_k| \geq A \left(\frac{pd^{max}}{C} \right)^{\frac{1}{1-p}}, \quad (2.5)$$

where A is a sufficiently large number and d^{max} is the upper bound of the singularity coefficients. Similarly it is assumed that in a neighborhood of each singularity the derivative of a smooth part of a signal still has the bound C . Let us define the following estimate \tilde{d}_j of d_0 :

$$\tilde{d}_j = \frac{\sum_{i=1}^M u_{i+j} - \sum_{i=0}^{M-1} u_{-i+j}}{h^p \sum_{i=1}^M (i - \frac{1}{2})^p}. \quad (2.6)$$

Representing the signal on both sides of t_0 by the Lagrange formula, one has

$$\begin{aligned} s_{i+j} &= \hat{s}(t_0) + (i - \frac{1}{2})h\hat{s}'(\xi_i) + ((i - \frac{1}{2})h)^p d_0, \quad \xi_i \in (t_0, t_0 + (i - \frac{1}{2})h), \\ s_{-i+j} &= \hat{s}(t_0) - (i + \frac{1}{2})h\hat{s}'(\xi_{-i}), \quad \xi_{-i} \in (t_0 - (i + \frac{1}{2})h, t_0). \end{aligned} \quad (2.7)$$

Substituting (2.1) and (2.7) into (2.6) and using (2.3), one obtains a formula for the error of the estimate (2.6):

$$\begin{aligned} |\tilde{d}_j - d_0| &= \frac{\left| \sum_{i=1}^M h(i - \frac{1}{2})[\hat{s}'(\xi_i) + \hat{s}'(\xi_{-i+1})] + \sum_{i=1}^M (n_{j+i} - n_{j-i+1}) \right|}{h^p \sum_{i=1}^M (i - \frac{1}{2})^p} \leq \\ &= \frac{2Ch \sum_{i=1}^M (i - \frac{1}{2}) + \left| \sum_{i=1}^M (n_{j+i} - n_{j-i+1}) \right|}{h^p \sum_{i=1}^M (i - \frac{1}{2})^p} = \frac{CM^2h + |\nu|}{h^p \sum_{i=1}^M (i - \frac{1}{2})^p}, \end{aligned} \quad (2.8)$$

where

$$\nu = \sum_{i=1}^M (n_{j+i} - n_{j-i+1}). \quad (2.9)$$

Let us fix ϵ , $0 < \epsilon < 1$, and find x_ϵ such that $P\{|\nu| > x_\epsilon\} = \epsilon$, where P stands for probability. Assuming that M is sufficiently large and using the Central Limit Theorem (or assuming that n_i are Gaussian), one gets from (2.2)

$$\nu \in \mathcal{N}(0, 2M\sigma^2), \quad (2.10)$$

where $\mathcal{N}(0, \sigma^2)$ is the class of the Gaussian random variables with zero mean value and variance σ^2 . Thus x_ϵ satisfies the equation

$$\operatorname{erfc} \left(\frac{x_\epsilon}{2\sqrt{M}\sigma} \right) = \epsilon,$$

where $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2) dt$. Hence $x_\epsilon = 2\sqrt{M}\sigma \operatorname{erfc}^{-1}(\epsilon)$. From (2.8) it follows that the inequality

$$|\tilde{d}_j - d_0| \leq \frac{CM^2h + 2\sqrt{M}\sigma \operatorname{erfc}^{-1}(\epsilon)}{h^p \sum_{i=1}^M (i - \frac{1}{2})^p} := \Delta d(M) \quad (2.11)$$

holds with probability $1 - \epsilon$. Minimizing the right-hand side of (2.11) with respect to M yields the optimal value M^{opt} and the optimal value of the estimate Δd^{opt} . It is easy to see that (2.11) holds also for the "–" type singularity $(x - t_0)_-^p d_0$. If there is no singularity in the interval $\Delta_j := [x_{j-M^{opt}+1}, x_{j+M^{opt}}]$, the largest possible value of the estimate (2.11) is also bounded by Δd^{opt} with probability $1 - \epsilon$ (take $d_0 = 0$ in (2.11)). Therefore one can formulate the following

A) Algorithm for locating and estimating singularities of a signal:

Consider the points x_{k_n} at which $|\tilde{d}_{k_n}| > \Delta d^{opt}$, and assume that these points are arranged so that $|\tilde{d}_{k_1}| > |\tilde{d}_{k_2}| > \dots$. The point $\tilde{x}_1 := \frac{x_{k_1} + x_{k_1+1}}{2}$ is taken as the estimate of the singular point. The corresponding value \tilde{d}_{k_1} is taken as the estimate of the coefficient of the singularity. Delete all the points from the set $\{x_{k_n}\}$ such that $|x_{k_n} - \tilde{x}_1| < L$, where $L = \max\left(M^{opt}h, A\left(\frac{pd^{max}}{C}\right)^{\frac{1}{1-p}}\right)$ and A is the constant in (2.5). Let x_{k_j} be the point with the largest value of $|\tilde{d}_{k_j}|$ among the remaining ones. Then $\tilde{x}_2 := \frac{x_{k_j} + x_{k_j+1}}{2}$ and \tilde{d}_{k_j} are taken as the estimates of the next singular point and the coefficient of the singularity, respectively. Repeating this process, one estimates all singular points and corresponding coefficients.

If the conclusion have been made that there is a singularity at the middle of the current interval, we need to determine if it is of "+" or "-" type. To this end let us calculate the variation of the observed data for the left and right parts of the interval Δ_j separately. Since the derivative of the signal in the neighborhood of singularity is large compared to C , the variance of the signal will be large too, and one can conclude that the side with larger variance corresponds to the side with singularity. Note that this does not hold if $p = 0$, i.e. if the signal has discontinuity. But in this case one needs only to determine the sign of a jump which is given by the sign of the estimate (2.6). Let us also note that the exponents p of the singularities can vary from a singular point to another singular point. However, one has to assume that these exponents are known.

3. INVESTIGATION OF THE BASIC PROPERTIES OF THE ALGORITHM

First let us calculate the probabilities of errors of each kind. From (2.8) – (2.11) it follows that the probability of the conclusion "there is a singularity at a given point" when there is none is less than ϵ . To find the probability δ of missing a singularity let us assume that the singularity is of the type $(x - t_0)_+^p d_0$ with $d_0 > 0$.

Then one has from (2.3), (2.9) and (2.10)

$$\begin{aligned}
\delta &= P\{\tilde{d}_j < \Delta d^{opt}\} = P\left\{d_0 + \frac{\sum_{i=1}^{M^{opt}} (i - \frac{1}{2})h [\hat{s}'(\xi_i) + \hat{s}'(\xi_{-i+1})] + \nu}{h^p \sum_{i=1}^{M^{opt}} (i - \frac{1}{2})^p} < \Delta d^{opt}\right\} \leq \\
&\leq P\left\{-\nu > d_0 h^p \sum_{i=1}^{M^{opt}} (i - \frac{1}{2})^p - 2C(M^{opt})^2 h - 2\sqrt{M^{opt}}\sigma \operatorname{erfc}^{-1}(\epsilon)\right\} = \\
&= \frac{1}{2} \operatorname{erfc}\left(\frac{d_0 h^p \sum_{i=1}^{M^{opt}} (i - \frac{1}{2})^p - 2C(M^{opt})^2 h}{2\sqrt{M^{opt}}\sigma} - \operatorname{erfc}^{-1}(\epsilon)\right). \tag{3.1}
\end{aligned}$$

To study the asymptotics of δ as $h \rightarrow 0$, one needs the following lemma (the proofs of all lemmas from this Section are given in the Appendix):

Lemma 1. *One has*

$$M^{opt} = c_1 h^{-2/3}(1+o(1)), \quad h^p \sum_{i=1}^{M^{opt}} (i - \frac{1}{2})^p = c_2 h^{\frac{p-2}{3}}(1+o(1)) \quad \text{as } h \rightarrow 0, \tag{3.2}$$

where the constants $c_1, c_2 > 0$ depend only on C, σ, p , and do not depend on h .

Substituting (3.2) into (3.1), one gets

$$\begin{aligned}
\delta &\leq \frac{1}{2} \operatorname{erfc}\left(\frac{d_0 c_2 h^{\frac{p-2}{3}}(1+o(1)) - 2C c_1^2 (1+o(1))^2 h^{-4/3} h}{2\sqrt{c_1 h^{-2/3}(1+o(1))}\sigma} - \operatorname{erfc}^{-1}(\epsilon)\right) = \\
&= \frac{1}{2} \operatorname{erfc}(c_3 h^{\frac{p-1}{3}} - c_4) \leq \frac{1}{\sqrt{\pi}} \frac{\exp\{-(c_3 h^{\frac{p-1}{3}} - c_4)^2\}}{(c_3 h^{\frac{p-1}{3}} - c_4)} \rightarrow 0 \quad \text{as } h \rightarrow 0,
\end{aligned}$$

since $p < 1$. Here c_3 and c_4 are some constants and ϵ , $0 < \epsilon < 1$, is fixed. Clearly, the same result holds for $d_0 < 0$. Hence we proved the following

Theorem 1. *For any fixed ϵ , $0 < \epsilon < 1$, and $d_0 \neq 0$, the probability of missing singularity goes to zero as $h \rightarrow 0$.*

To find the localization accuracy, let us assume that the singularity is of the form $(x - t_0)_+^p d_0$ with $d_0 > 0$ and let us compare the values \tilde{d}_j and \tilde{d}_{j+N} , that is consider the cases when the singularity is at the middle of the interval and when the interval has been shifted N points to the right, $N \leq M^{opt}$. When the singularity is at the middle of the interval, \tilde{d}_j is given by (2.6) with $M = M^{opt}$. When the interval has been shifted N points to the right, the estimate is

$$\tilde{d}_{j+N} = \frac{\sum_{i=1}^{M^{opt}} u_{i+j+N} - \sum_{i=0}^{M^{opt}-1} u_{-i+j+N}}{h^p \sum_{i=1}^{M^{opt}} (i - \frac{1}{2})^p}. \tag{3.3}$$

The localization accuracy is N_0h if $\tilde{d}_j > \tilde{d}_{j+N}$ for $N_0 < N \leq M^{opt}$. Hence, from (2.6) and (3.3) one gets

$$\sum_{i=1}^{M^{opt}} u_{i+j} - \sum_{i=0}^{M^{opt}-1} u_{-i+j} > \sum_{i=1}^{M^{opt}} u_{i+j+N} - \sum_{i=0}^{M^{opt}-1} u_{-i+j+N} \quad (3.4)$$

for all $N_0 < N \leq M^{opt}$, where N_0 is to be determined. Canceling the similar terms on the left and right sides of (3.4) and using (2.1), (2.7), one obtains

$$\begin{aligned} & d_0 h^p \left(2 \sum_{i=1}^N \left(i - \frac{1}{2}\right)^p - \sum_{i=1}^N \left(i + M^{opt} - 1/2\right)^p \right) > \\ & \left\{ - \left(\sum_{i=1}^N \left(i - \frac{1}{2}\right) h \hat{s}'(\xi_i) + \sum_{i=M^{opt}-N+1}^{M^{opt}} \left(i - \frac{1}{2}\right) h \hat{s}'(\xi_{-i+1}) \right) + M^{opt} h \sum_{i=1}^N \hat{s}'(\eta_i) \right\} + \tilde{\nu}, \\ & \eta_i \in (x_{i+j}, x_{i+j+M^{opt}}), \end{aligned} \quad (3.5)$$

where

$$\tilde{\nu} = \sum_{i=M^{opt}-N}^{M^{opt}-1} n_{-i+j} + \sum_{i=1+M^{opt}}^{N+M^{opt}} n_{i+j} - 2 \sum_{i=1}^N n_{i+j}. \quad (3.6)$$

The term in the braces on the right-hand side of (3.5) can be estimated as follows:

$$\begin{aligned} & \left| - \left(\sum_{i=1}^N \left(i - \frac{1}{2}\right) h \hat{s}'(\xi_i) + \sum_{i=M^{opt}-N+1}^{M^{opt}} \left(i - \frac{1}{2}\right) h \hat{s}'(\xi_{-i+1}) \right) + M^{opt} h \sum_{i=1}^N \hat{s}'(\eta_i) \right| \leq \\ & \leq Ch \left(\frac{N(N+1)}{2} - \frac{N}{2} + \frac{M^{opt}(M^{opt}+1)}{2} - \frac{(M^{opt}-N)(M^{opt}-N+1)}{2} - \frac{N}{2} \right. \\ & \quad \left. + M^{opt} N \right) = 2CM^{opt}Nh. \end{aligned} \quad (3.7)$$

Let us fix ϵ' , $0 < \epsilon' < 1$, and find $x_{\epsilon'}$ such that

$$P\{\tilde{\nu} > x_{\epsilon'}\} = \epsilon'. \quad (3.8)$$

Let us assume that n_i are Gaussian. Since linear combination of Gaussian random variables is Gaussian, one gets from (2.2) and (3.6)

$$E(\tilde{\nu}) = 0, \quad \sigma^2(\tilde{\nu}) = 6N\sigma^2, \quad \tilde{\nu} \in \mathcal{N}(0, 6N\sigma^2). \quad (3.9)$$

Hence

$$x_{\epsilon'} = \sqrt{12N}\sigma \operatorname{erfc}^{-1}(2\epsilon'). \quad (3.10)$$

Substitution of (3.7) and (3.10) into (3.5) yields

$$d_0 h^p f_p(N, M^{opt}) > 2CM^{opt}hN + \sqrt{12N}\sigma \operatorname{erfc}^{-1}(2\epsilon'), \quad N_0 < N \leq M^{opt}, \quad (3.11)$$

where

$$f_p(N, M) = 2 \sum_{i=1}^N \left(i - \frac{1}{2}\right)^p - \sum_{i=1}^N \left(i + M - 1/2\right)^p. \quad (3.12)$$

The smallest value N_0 such that (3.11) holds for all N , $N_0 < N \leq M^{opt}$, guarantees the localization accuracy N_0h with probability $1 - \epsilon'$. The existence of such N_0 for sufficiently large d_0 follows from lemma 2:

Lemma 2. *If $0 \leq p \leq 1/2$ and $M \geq 1$, then $f_p(M, M) > 0$.*

Therefore if we need the accuracy $N_0 h$, it will be guaranteed with probability $1 - \epsilon'$ if the following two conditions hold

$$d_0 > \max_{N_0+1 \leq N \leq M^{opt}} \frac{2CM^{opt}hN + \sqrt{12N}\sigma \operatorname{erfc}^{-1}(2\epsilon')}{h^p f_p(N, M^{opt})}, \quad (3.13a)$$

$$f_p(N, M^{opt}) > 0 \quad \text{for } N_0 + 1 \leq N \leq M^{opt}. \quad (3.13b)$$

From (3.12) it follows that $f_p(N, M)$ as a function of N is increasing for

$$N > y_p(M) := \frac{1}{2} + \frac{M}{2^{1/p} - 1} \geq \frac{M}{3} + \frac{1}{2}.$$

Therefore if $f_p(N_0, M^{opt}) > 0$ and $N_0 > y_p(M^{opt})$, then $f_p(N, M^{opt}) > 0$ for all $N \geq N_0$. Note that inequality (3.13) by itself is not sufficient to guarantee the localization accuracy $N_0 h$, because it may happen that the local maximum of the estimate (2.6) is situated to the left of singular point t_0 , or farther than M^{opt} points to the right of t_0 . Hence one needs lemma 3 which asserts that this is not the case. Repeating calculations similar to (3.3)-(3.13) one obtains

Lemma 3. *a) The case when the interval has been shifted N points to the right, $M^{opt} + 1 \leq N \leq 2M^{opt}$, leads to an inequality weaker than (3.13) with $N_0 = M^{opt} - 1$, $N = M^{opt}$. b) The case when the interval has been shifted N points to the left, $1 \leq N \leq M^{opt}$, leads to an inequality weaker than (3.13) with the same value of N .*

In fact, lemma 3 means that the singularity $(x - t_0)_\pm^p d_0$ with the exponent $0 \leq p \leq 1/2$ does not "propagate" too far: the maximum value of the estimate (2.6) is in the interval $[t_0 - M^{opt}h, t_0 + M^{opt}h]$ with the probability $1 - \epsilon'$, provided that d_0 is sufficiently large (see (3.13)). The reason for considering only the cases formulated in lemma 3 is clear from (2.6): the estimate \tilde{d}_j is proportional to the average value of the derivative of the signal in the interval $\Delta_j = [x_{j-M^{opt}+1}, x_{j+M^{opt}}]$. If one shifts Δ_j more than M^{opt} points to the left, then $t_0 \notin \Delta_j$ and the derivative of the signal is bounded by C . Hence the estimate (2.6) will exceed the threshold Δd^{opt} with probability less than ϵ . The situation is different if one shifts the interval to the right of t_0 , since the derivative of the signal is large in the neighborhood of singularity. But, since the derivative of the singular term is decreasing as the distance from t_0 is increasing, one sees that there is no need to consider $N > 2M^{opt}$, because $\Delta_{j+2M^{opt}}$ is the closest interval to t_0 such that $\Delta_{j+N} \cap \Delta_j = \emptyset$.

Let us now study the behavior of the right-hand side of (3.13a) as $h \rightarrow 0$. Consider the worst case when $N_0 + 1 = M^{opt}$. In this case (3.13a) reduces to

$$d_0 > d_0^{min}(h) := \frac{2C(M^{opt})^2 h + \sqrt{12M^{opt}}\sigma \operatorname{erfc}^{-1}(2\epsilon')}{h^p f_p(M^{opt}, M^{opt})}. \quad (3.14)$$

Lemma 4.

$$d_0^{min}(h) = c_3 h^{\frac{1-p}{3}} (1 + o(1)) \quad \text{as } h \rightarrow 0, \quad (3.15)$$

where the constant c_3 does not depend on h .

The above lemmas yield the following result.

Theorem 2. *For any $d_0 \neq 0$ and accuracy $a > 0$ of the localization of the singularity $(x - t_0)_\pm^p d_0$, $0 \leq p \leq 1/2$, the probability of localization error larger than a goes to zero as $h \rightarrow 0$.*

Proof. Suppose for simplicity that $d_0 > 0$. The case when $d_0 < 0$ can be considered similarly. Let us fix any $\epsilon' > 0$ and find h_0 such that the probability of an error larger than a is less than ϵ' for any $h \leq h_0$. Lemma 3 asserts that if one wants the accuracy $(M^{opt} - 1)h$ with probability $1 - \epsilon'$, it is sufficient to satisfy (3.14). Using lemmas 1 and 4, find h_1 and h_2 such that

$$M^{opt}h_1 = c_1 h_1^{1/3}(1 + o(1)) < a, \quad d_0^{min}(h_2) = c_3 h_2^{\frac{1-p}{3}}(1 + o(1)) < d_0.$$

Let $\Delta(h)$ be the localization error and let $h_0 = \min(h_1, h_2)$. Then, since $d_0^{min}(h) \leq d_0^{min}(h_0) < d_0$ and $M^{opt}h \leq M^{opt}h_0 < a$, one has

$$P\{\Delta(h) > a\} < P\{\Delta(h) > M^{opt}h\} < \epsilon' \quad \text{for } h \leq h_0.$$

Theorem 2 is proved.

4. LOCALIZATION OF JUMPS

Since the problem of jump localization from noisy data is practically important, we consider the case $p = 0$ separately. Many formulas from previous sections can be considerably simplified if $p = 0$. Letting $p = 0$ in (2.11), one gets

$$|\tilde{d}_j - d_0| \leq \Delta d(M) = CMh + \frac{2}{\sqrt{M}} \sigma \operatorname{erfc}^{-1}(\epsilon). \quad (4.1)$$

Minimization of the right-hand side of (4.1) with respect to M yields

$$M^{opt} = \left[\left(\frac{\sigma \operatorname{erfc}^{-1}(\epsilon)}{Ch} \right)^{2/3} \right], \quad \Delta d^{opt} = \Delta d(M^{opt}) \approx 3(\operatorname{erfc}^{-1}(\epsilon))^{2/3} (Ch)^{1/3} \sigma^{2/3}. \quad (4.2)$$

Note that (4.2) agrees with lemma 1, which asserts $M^{opt} = O(h^{-2/3})$ as $h \rightarrow 0$. Substitution of $p = 0$ and (4.2) into (3.1) (with d_0 replaced by $|d_0|$ to cover cases $d_0 > 0$ and $d_0 < 0$) gives the probability of missing a jump given its value d_0 :

$$\delta \leq \frac{1}{2} \operatorname{erfc} \left\{ \frac{|d_0|}{2} \left(\frac{\operatorname{erfc}^{-1}(\epsilon)}{Ch\sigma^2} \right)^{1/3} - 2\operatorname{erfc}^{-1}(\epsilon) \right\}. \quad (4.3)$$

To find the accuracy of jump localization, put $p = 0$ in (3.11) (with d_0 replaced by $|d_0|$) and (3.12). One has $f_0(N, M) = N$ and

$$|d_0| > 2CM^{opt}h + \sqrt{\frac{12}{N}} \sigma \operatorname{erfc}^{-1}(2\epsilon'), \quad N_0 + 1 \leq N \leq M^{opt}. \quad (4.4)$$

Since the right-hand side of (4.4) is strictly decreasing function of N , one concludes that the localization accuracy $N_0 h$ is guaranteed with probability $1 - \epsilon'$ if the jump magnitude d_0 satisfies (4.4) for $N = N_0 + 1$ (this inequality is much easier verifiable than the general inequality (3.13)). Conversely, the accuracy of the jump localization, given the jump magnitude d_0 , is

$$N_0 h = \left[12 \left(\frac{\sigma \operatorname{erfc}^{-1}(2\epsilon')}{|d_0| - 2CM^{opt}h} \right)^2 \right] h. \quad (4.5)$$

Define $N_0 := N_0(\epsilon')$ to be the expression in the brackets in equation (4.5). Let us summarize the results.

Theorem 3. *Let $p = 0$. Then the algorithm A detects jumps with the probability $1 - \delta$, where the estimate (4.3) of δ depends on the jump magnitude and on the fixed probability ϵ of the "false alarm" error. For any ϵ' with $0 < \epsilon' < 1$ and $N_0(\epsilon') \leq M^{opt}$, formula (4.5) gives the accuracy of the jump localization, such that the localization error larger than $N_0 h$ occurs with the probability less than ϵ' .*

5. NUMERICAL EXPERIMENTS

The algorithms described in Sections 2 and 4 have been tested on model examples. Figure 1 shows the results of applying algorithm for singularity localization. The test function (solid line)

$$s(x) = 0.02(x - 5)^2 + (x - 5)_+^{0.3},$$

is corrupted at each point by noise (dotted line) with uniform distribution, zero mean and standard deviation $\sigma = 0.15$. The observation points are $x_i = 0.05i$, $0 \leq i \leq 200$. The probability of "false singularities" is chosen $\epsilon = 0.005$, the corresponding optimal bandwidth, obtained by minimizing (2.11), is $M^{opt} = 17$. The x -coordinate of the big dot represents the position of the found singularity $t_0 = 5.1$ (the true position is $t_0 = 5$). The y -coordinate of the dot is equal to the coefficient of the singularity estimated by (2.6) with $p = 0.3$: $d_0 = 1.04$ (the true value is $d_0 = 1$). No "false singularities" have been detected in this model. The probability of missing the singularity calculated by formula (3.1) is $\delta \leq 1.4 \cdot 10^{-9}$.

The results of testing the algorithm for jump localization are presented in Figure 2. The test function (solid line)

$$s(x) = \begin{cases} 0.02(x - 5)^2 - 0.5, & x < 5, \\ 0.02(x - 5)^2 + 0.5, & x > 5 \end{cases}$$

is corrupted by the uniformly distributed noise (dotted line) with zero mean value and $\sigma = 0.4$. The observation points are $x_i = \frac{1}{15}i$, $0 \leq i \leq 150$. The probability ϵ is equal to 0.01 and the corresponding M^{opt} , calculated by (4.2), is 14. The x -coordinate of the big dot represents the position of the estimated jump $t_0 = 5$ (which coincides with its true position), the y -coordinate of the dot equals to the estimate of the jump magnitude $d_0 = 1.14$ (the true value is 1). No "false jumps" have been detected in this example. The probability of missing the jump calculated by formula (4.3) is $\delta \leq 6.1 \cdot 10^{-2}$.

APPENDIX

Proof of lemma 1. Let us first prove the following simple formula

$$\sum_{i=1}^M \left(i - \frac{1}{2}\right)^p = \frac{M^{p+1}}{p+1} (1 + o(1)) \quad \text{as } M \rightarrow \infty. \quad (\text{A.1})$$

One has

$$\int_1^M (x - 1/2)^p dx < \sum_{i=1}^M \left(i - \frac{1}{2}\right)^p < \int_1^{M+1} (x - 1/2)^p dx.$$

Thus

$$\frac{\left| \sum_{i=1}^M (i - \frac{1}{2})^p - \int_0^M x^p dx \right|}{\int_0^M x^p dx} \leq \frac{\int_0^{M+1} (x - 1/2)^p dx + \left(\int_0^M x^p dx - \int_1^M (x - 1/2)^p dx \right)}{\int_0^M x^p dx} =$$

$$\frac{(M + 1/2)^{p+1} - (M - 1/2)^{p+1} + M^{p+1} - (M - 1/2)^{p+1} + (1/2)^{p+1}}{M^{p+1}} \rightarrow 0$$

as $M \rightarrow \infty$.

Hence

$$\sum_{i=1}^M (i - \frac{1}{2})^p = \int_0^M x^p dx (1 + o(1)) = \frac{M^{p+1}}{p+1} (1 + o(1)) \quad \text{as } M \rightarrow \infty.$$

Now let us prove that $M^{opt} \rightarrow \infty$ as $h \rightarrow 0$. Let us argue by contradiction. Assume that there exists $M_0 > 0$ such that $M^{opt} < M_0$ as $h \rightarrow 0$. Pick any $\delta \ll 2\sigma \operatorname{erfc}^{-1}(\epsilon)$, put $h_0 = \frac{\delta}{CM_0^2}$ and fix any $h < h_0$ in (2.11). The numerator of (2.11) grows like $M^{1/2}$ (for $M \leq M_0$), the denominator grows like M^{p+1} (see (A.1)). Hence

$$\Delta d(M_0 + 1) < \Delta d(M) \quad \forall M, 1 \leq M \leq M_0,$$

in particular, $\Delta d(M_0 + 1) < \Delta d(M^{opt})$, which contradicts the definition of M^{opt} . To prove the first assertion of lemma 1, denote $t = Mh$ and substitute (A.1) into (2.11) to get:

$$\Delta d(t) = \frac{C \frac{t^2}{h} + \sqrt{\frac{4t}{h}} \sigma \operatorname{erfc}^{-1}(\epsilon)}{\frac{1}{h} \frac{t^{p+1}}{p+1} (1 + o(1))} = (p+1)(1 + o(1)) [Ct^{1-p} + \sqrt{4ht}^{-1/2-p} \sigma \operatorname{erfc}^{-1}(\epsilon)].$$
(A.2)

The expression in brackets attains its minimum at $t = t^{opt}$, where

$$t^{opt} = \left(\frac{2\sigma \operatorname{erfc}^{-1}(\epsilon)(p+1/2)}{C(1-p)} \right)^{2/3} h^{1/3}.$$

Thus M^{opt} is given by the formula

$$M^{opt} = c_1 h^{-2/3} (1 + o(1)),$$
(A.3)

where

$$c_1 = \left(\frac{2\sigma \operatorname{erfc}^{-1}(\epsilon)(p+1/2)}{C(1-p)} \right)^{2/3}.$$

Substituting (A.3) into (A.2), one obtains

$$\Delta d(M^{opt}) = O(h^{\frac{1-p}{3}}) \quad \text{as } h \rightarrow 0.$$

Combining (A.1) and (A.3) yields the last assertion of lemma 1.

Proof of lemma 2. Let us consider first the case $p = 1/2$.

$$\begin{aligned}
& f_{1/2}(M+1, M+1) - f_{1/2}(M, M) = \\
& 2\sqrt{M+1/2} - \sqrt{2M+3/2} - \sum_{i=1}^M \left(\sqrt{i+M+1/2} - \sqrt{i+M-1/2} \right) > \\
& 2\sqrt{M+1/2} - \sqrt{2M+3/2} - \frac{1}{2} \sum_{i=1}^M \frac{1}{\sqrt{i+M-1/2}} > \\
& 2\sqrt{M+1/2} - \sqrt{2M+3/2} - \frac{1}{2} \int_0^M \frac{dx}{\sqrt{x+M-1/2}} = \\
& \left(2\sqrt{M+1/2} + \sqrt{M-1/2} \right) - \left(\sqrt{2M+3/2} + \sqrt{2M-1/2} \right) > 0.
\end{aligned}$$

One has $f_{1/2}(1, 1) = \frac{2-\sqrt{3}}{\sqrt{2}} > 0$, hence $f_{1/2}(M, M) > 0$ for all $M \geq 1$. Now let $0 \leq p \leq 1/2$. One has

$$f_p(M, M) = 2 \sum_{i=1}^M \frac{\sqrt{i-1/2}}{(i-\frac{1}{2})^{1/2-p}} - \sum_{i=1}^M \frac{\sqrt{i+M-1/2}}{(i+M-1/2)^{1/2-p}}.$$

Since $(i-\frac{1}{2})^{1/2-p} < M^{1/2-p} < (i+M-1/2)^{1/2-p}$ for $1 \leq i \leq M$, it follows that

$$f_p(M, M) > \frac{1}{M^{1/2-p}} f_{1/2}(M, M) > 0.$$

Proof of lemma 3. Let us consider the first claim, $M^{opt} + 1 \leq N \leq 2M^{opt}$. The basic inequality is

$$\sum_{i=1}^{M^{opt}} u_{i+j} - \sum_{i=0}^{M^{opt}-1} u_{-i+j} > \sum_{i=1}^{M^{opt}} u_{i+j+N} - \sum_{i=0}^{M^{opt}-1} u_{-i+j+N}. \quad (\text{A.4})$$

Canceling similar terms on the left and right sides of (A.4), using (2.1) and (2.4), one obtains

$$\begin{aligned}
& d_0 h^p \left(\sum_{i=1}^{M^{opt}} (i-\frac{1}{2})^p + \sum_{i=N-M^{opt}+1}^N (i-\frac{1}{2})^p - \sum_{i=N+1}^{M^{opt}+N} (i-\frac{1}{2})^p \right) \geq \\
& - \sum_{i=1}^{M^{opt}} (i-\frac{1}{2}) h [\hat{s}'(\xi_i) + \hat{s}'(\xi_{-i+1})] + M^{opt} h \sum_{i=1}^{M^{opt}} \hat{s}'(\eta_i) + \tilde{\nu}, \quad (\text{A.5})
\end{aligned}$$

where

$$\tilde{\nu} = \sum_{i=N+1}^{M^{opt}+N} n_{j+i} + \sum_{i=0}^{M^{opt}-1} n_{j-i} - \sum_{i=1}^{N-M^{opt}} n_{j+i} - \sum_{i=1+M^{opt}}^N n_{j+i} - 2 \sum_{i=N-M^{opt}+1}^{M^{opt}} n_{j+i}. \quad (\text{A.6})$$

The last sum in (A.6) is defined to be zero if $N = 2M^{opt}$. Estimating the right-hand side of (A.5) with probability $1 - \epsilon$, one gets

$$d_0 h^p F_p(N, M^{opt}) > 2C(M^{opt})^2 h + x_\epsilon, \quad M^{opt} + 1 \leq N \leq 2M^{opt}, \quad (\text{A.7})$$

where

$$F_p(N, M) = \sum_{i=1}^M (i - \frac{1}{2})^p + \sum_{i=N-M+1}^N (i - \frac{1}{2})^p - \sum_{i=N+1}^{M+N} (i - \frac{1}{2})^p, \quad (\text{A.8})$$

and x_ϵ is determined from the condition $P\{\tilde{\nu} > x_\epsilon\} = \epsilon$. From (A.6) and (2.2) it follows that $\tilde{\nu} \in \mathcal{N}(0, (8M^{opt} - 2N)\sigma^2)$. Hence

$$x_\epsilon = 2\sqrt{4M^{opt} - N}\sigma \operatorname{erfc}^{-1}(2\epsilon)$$

and (A.7) takes the form

$$\begin{aligned} d_0 h^p F_p(N, M^{opt}) &> 2C(M^{opt})^2 h + 2\sqrt{4M^{opt} - N}\sigma \operatorname{erfc}^{-1}(2\epsilon), \\ M^{opt} + 1 &\leq N \leq 2M^{opt}. \end{aligned} \quad (\text{A.9})$$

Comparing (A.9) and (3.11) with $N = M^{opt}$, one sees that to prove the first assertion of lemma 3 it is sufficient to show $F_p(N, M) \geq f_p(M, M)$ for $M + 1 \leq N \leq 2M$. From (3.12) and (A.8) it follows that one needs to prove the inequality

$$2 \sum_{i=1}^M (i - \frac{1}{2})^p - \sum_{i=M+1}^{2M} (i - \frac{1}{2})^p \leq \sum_{i=1}^M (i - \frac{1}{2})^p + \sum_{i=N-M+1}^N (i - \frac{1}{2})^p - \sum_{i=N+1}^{M+N} (i - \frac{1}{2})^p.$$

Cancellation of the similar terms yields the equivalent inequality

$$\sum_{i=1}^{N-M} (i - \frac{1}{2})^p + \sum_{i=2M+1}^{M+N} (i - \frac{1}{2})^p \leq 2 \sum_{i=M+1}^N (i - \frac{1}{2})^p,$$

which can be rewritten as

$$\sum_{i=1}^{N-M} \left\{ [(x_i + 2M)^p - (x_i + M)^p] - [(x_i + M)^p - x_i^p] \right\} \leq 0, \quad x_i = i - 1/2. \quad (\text{A.10})$$

Since $0 \leq p \leq 1/2$ and $M > 0$, one has

$$(x_i + 2M)^p - (x_i + M)^p < (x_i + M)^p - x_i^p,$$

which proves (A.10) and the first claim in lemma 3.

In a similar fashion it is easy to show that the case when the interval has been shifted N points to the left, $1 \leq N \leq M^{opt}$, leads to an inequality weaker than (3.11) with the same N . Indeed, the resulting inequality is

$$d_0 h^p \sum_{i=M^{opt}-N+1}^{M^{opt}} (i - \frac{1}{2})^p > 2CM^{opt} h N + \sqrt{12N}\sigma \operatorname{erfc}^{-1}(2\epsilon), \quad 1 \leq N \leq M^{opt},$$

and from the equation

$$\sum_{i=M-N+1}^M \left(i - \frac{1}{2}\right)^p - \left(2 \sum_{i=1}^N \left(i - \frac{1}{2}\right)^p - \sum_{i=1}^N (i + M - 1/2)^p\right) = \sum_{i=1}^N \left(\left(i - 1/2 + M\right)^p + \left(i - 1/2 + M - N\right)^p - 2\left(i - 1/2\right)^p\right) > 0$$

one obtains the desired assertion

Proof of lemma 4. Using the approach from the proof of lemma 1, one obtains the following estimate

$$h^p \sum_{i=1}^{M^{opt}} (i + M^{opt} - 1/2)^p = \frac{(2M^{opt}h)^{p+1} - (M^{opt}h)^{p+1}}{(p+1)h} (1 + o(1)) \quad \text{as } h \rightarrow 0. \quad (\text{A.11})$$

Combining (3.2), (3.12), (3.14) and (A.11) yields

$$\begin{aligned} d_0^{min}(h) &= \frac{2C(M^{opt})^2 h + \sqrt{12M^{opt}} \sigma \operatorname{erfc}^{-1}(2\epsilon')}{2 \frac{(M^{opt}h)^{p+1}}{(p+1)h} (1 + o(1)) - \frac{(2M^{opt}h)^{p+1} - (M^{opt}h)^{p+1}}{(p+1)h} (1 + o(1))} = \\ &= (p+1) \frac{2C(M^{opt}h)^2 + \sqrt{12M^{opt}} h \sigma \operatorname{erfc}^{-1}(2\epsilon')}{(3 - 2^{p+1})(M^{opt}h)^{p+1} (1 + o(1))} = \\ &= \frac{\text{const}}{3 - 2^{p+1}} \frac{h^{2/3}}{h^{\frac{p+1}{3}}} (1 + o(1)) = \text{const}_1 h^{\frac{1-p}{3}} (1 + o(1)) \rightarrow 0 \quad \text{as } h \rightarrow 0, \end{aligned}$$

where we used the inequality $2^{p+1} < 3$ for $0 \leq p \leq 1/2$. Lemma 4 is proved.

REFERENCES

1. L.S. Davis, *A survey of edge detection techniques*, Comput. Graphics Image Processing **4** (1975), 248–270.
2. D.Girard, *From template matching to optimal approximation by piecewise smooth curves*, SPIE Vol. 1251, Curves and Surfaces in Computer Vision and Graphics (1990), 174–182.
3. D.Girard and P.J.Laurent, *Splines and estimation of nonlinear parameters*, Mathematical methods in computer aided geometric design (T.Lyche and L.L.Schumaker, eds.), Academic Press, 1989, pp. 273–298.
4. W. Härdle, *Applied nonparametric regression*, Cambridge University Press, Cambridge, 1990.
5. W.K. Pratt, *Digital Image Processing*, 2nd ed., Wiley - Interscience, New York, 1991.
6. A.G. Ramm, *Multidimensional Inverse Scattering Problems*, Longman, Wiley, New York, 1992, (Expanded Russian edition, Mir, Moscow, 1993).
7. A.G. Ramm, *Random fields estimation theory*, Longman, Wiley, New York, 1990.
8. A.G. Ramm and A.I. Zaslavsky, *Reconstructing singularities of a function from its Radon transform*, Bull. Amer. Math. Soc. **25** (1993), 109–115.
9. G. Wahba, *Spline models for observational data*, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59, SIAM, Philadelphia, PA, 1990.

MATHEMATICS DEPARTMENT, KANSAS STATE UNIVERSITY, MANHATTAN, KS 66506-2602

MATHEMATICS DEPARTMENT, KANSAS STATE UNIVERSITY, MANHATTAN, KS 66506-2602
E-mail address: ramm@math.ksu.edu