

CHANGE CURVES IN THE PRESENCE OF DEPENDENT NOISE

BY MADAN L. PURI* AND FRITS H. RUYMGAART
Indiana University and Texas Tech University

In image processing the assumption of independent errors is not always realistic. Therefore we consider the problem of change curve estimation when the error process is in a class of stationary random fields. This class contains ARMA fields as special cases. A speed of a.s. uniform convergence is obtained.

1. Introduction and Outline. In this paper a nonparametric method will be discussed to detect the curve along which a random field in the plane is supposed to change its average level. Let us assume that we observe random variables

$$X_{i,j} = \mu_{i,j} + E_{i,j}, \quad (i, j) \in \{1, \dots, n\}^2, \quad (1.1)$$

defined for all i, j , and n on one and the same probability space $(\Omega, \mathcal{W}, \mathbb{P})$. The indices (i, j) divide into two groups: within each group the $\mu_{i,j}$ are constant but between groups they are different. The error terms $E_{i,j}$ form a stationary asymptotically decomposable random field; such random fields include in particular linear random fields and more specifically ARMA fields as special cases. Further specifications will be given below. The assumption that the indices i and j have the same range is purely a matter of convenience.

For a further specification of the numbers $\mu_{i,j}$ as well as for asymptotic considerations it turns out to be useful to rescale the two-dimensional “time” and to define the observable process ($n \in \mathbb{N}$, $t = (t_1, t_2)$)

$$X_n(t) = X_{i,j}, \text{ for } \frac{i-1}{n} < t_1 \leq \frac{i}{n}, \frac{j-1}{n} < t_2 \leq \frac{j}{n}, (i, j) \in \{1, \dots, n\}^2, \quad (1.2)$$

on $(0, 1]^2$. It may occasionally be convenient to define $X_n = 0$ elsewhere. Similarly we may construct $\mu_n(t)$ and $E_n(t)$ for $t = (0, 1]^2$ from the $\mu_{i,j}$ and the $E_{i,j}$, setting them 0 elsewhere.

* Research supported by the Office of Naval Research Grant N0014-91-J-1020

AMS 1991 Subject Classification: Primary 62G05; Secondary 62M40

Key words and phrases: Change curves, random fields, nearly black objects.

The curve Γ is assumed to be the graph of a sufficiently smooth monotonically nonincreasing function from a to b ($0 \leq a < b \leq 1$). We also assume that $(0, 1]^2 \setminus \Gamma$ is a disconnected set with two components C_1, C_2 having interiors C_1^o, C_2^o and closures $\overline{C}_1, \overline{C}_2$. We index these sets in such a way that $(0, 0) \in C_1, (1, 1) \in C_2$; note that $\overline{C}_1 \cup \overline{C}_2 = [0, 1]^2$. These assumptions entail in particular that, for any $t \in \Gamma$ and $\varepsilon > 0$,

$$\begin{cases} (0, 1]^2 \cap [t_1 - \varepsilon, t_1] \times [t_2 - \varepsilon, t_2] \subset \overline{C}_1, \\ (0, 1]^2 \cap [t_1, t_1 + \varepsilon] \times [t_2, t_2 + \varepsilon] \subset \overline{C}_2. \end{cases} \tag{1.3}$$

For μ_n , we assume that

$$\mu_n(t) \longrightarrow \mu(t) = c_1 I_{C_1}(t) + c_2 I_{C_2}(t), \text{ as } n \rightarrow \infty, t \in (0, 1]^2 \setminus \Gamma, \tag{1.4}$$

where $I(\cdot)$ denotes the indicator function.

The assumption that the $E_{i,j}$ form a stationary asymptotically decomposable random field means that there exist parameters $\rho, \sigma, \tau \in (0, \infty)$ such that for every $n \in \mathbb{N}$ the following holds: there exist $r(n) \geq 1, \gamma(n) \in (0, \infty), \delta(n) \in (0, \infty)$ and decompositions satisfying

$$\begin{cases} E_{i,j} = E_{i,j}^{(n)} + \overline{E}_{i,j}^{(n)}, \\ \text{with } \mathcal{L}(E_{i,j}), \mathcal{L}(E_{i,j}^{(n)}) \text{ independent of } (i, j) \in \{1, \dots, n\}^2; \end{cases} \tag{1.5}$$

$$\begin{aligned} E_{i,j}^{(n)} \perp E_{k,\ell}^{(n)} \text{ for all } (i, j), (k, \ell) \in \{1, \dots, n\}^2 \\ \text{with } |i - k| \vee |j - \ell| \geq r(n); \end{aligned} \tag{1.6}$$

$$\mathbb{P} \left\{ |\overline{E}_{i,j}^{(n)}| \geq \gamma(n) \right\} \leq \delta(n); \tag{1.7}$$

$$r(n) = O(n^{2\rho}), \quad \gamma(n) = O(n^{-2\sigma}), \quad \delta(n) = O(n^{-2-\tau}), \text{ as } n \rightarrow \infty. \tag{1.8}$$

If the $E_{i,j}$ happen to be independent, the conditions are trivially satisfied with all the $\overline{E}_{i,j}^{(n)} = 0$. For applications it is convenient to further specify the orders in (1.8). We will assume that

$$0 \leq \rho < \frac{1}{2}, \quad \sigma > 1, \quad \tau > 1. \tag{1.9}$$

Although at first sight the definition of asymptotic decomposability might appear rather technical, the conditions are tailor-made and typically fulfilled for random fields with a Volterra expansion of finite order (Priestley (1981)) and thus it covers a wide range of interesting random fields. By way of an example, let us give simple sufficient conditions for a linear random field to

satisfy the conditions; a proof can be found in Puri and Ruymgaart (1991). Let $\xi_{i,j}$, $(i, j) \in \mathbb{Z}^2$, be i.i.d. random variables with

$$\mathbb{E}|\xi_{i,j}|^\nu < \infty, \text{ for some } 0 < \nu \leq 1, \quad (1.10)$$

and let us consider the linear field

$$E_{i,j} = \sum_{k \in \mathbb{Z}} \sum_{\ell \in \mathbb{Z}} a_{k,\ell} \xi_{i-k,j-\ell}, \quad (i, j) \in \{1, \dots, n\}^2, \quad (1.11)$$

where the real numbers $a_{k,\ell}$ satisfy

$$|a_{k,\ell}| = \mathcal{O}((k \vee \ell)^{-\lambda}), \quad \text{as } k \vee \ell \rightarrow \infty, \text{ for some } \lambda \geq \frac{\tau + \sigma\nu + 2\rho}{\nu\rho}. \quad (1.12)$$

Then the field satisfies the asymptotic decomposability conditions (1.5)–(1.9). Of course the conditions will also be satisfied for $\nu > 1$ in (1.10) but we focussed on the interesting case where not even a first moment might exist. See also Chanda et al. (1990), Chanda and Ruymgaart (1991), and Ruymgaart (1991) for further information about asymptotic decomposability.

Having specified the assumptions regarding the random field, let us now turn to the problem that we want to consider, viz. the construction of estimators for the change curve Γ and investigation of the convergence of the estimators. Rather than estimating the curve itself our procedure boils down to constructing a sequence of stochastic processes $\{\hat{\Gamma}_n, n \in \mathbb{N}\}$ on $(0, 1]^2$. For almost every $\omega \in \Omega$ there appears to exist $n(\omega) \in \mathbb{N}$ such that $\hat{\Gamma}_n$ is 0 everywhere except on a strip around the curve Γ , provided that $n \geq n(\omega)$. As the sample size tends to ∞ the width of these strips tends to zero. On Γ itself the processes $\hat{\Gamma}_n$ are almost everywhere uniformly close to the difference in level $c_1 - c_2$ in the case where the error distribution is symmetric about 0. Various methods are proposed in the literature to detect change but usually for one-dimensional time and i.i.d. errors; see, e.g., Pettitt (1979), Wolfe and Schechtman (1984), and Csörgö and Horvath (1989). Van de Geer (1988) considers, more generally, change in regression including some results for multidimensional regression but with independent errors. For independent variables and one dimensional time, Carlstein (1988) proposed a very general method not restricted to differences in location. Recently Tsybakov (1991) studied the problem of image estimation, which subsumes multidimensional time, but also this author assumes the errors to be i.i.d.. A closely related reference is also Carlstein and Krishnamoorthy (1992). See Parzen (1991) for general remarks regarding the analysis of change.

The method proposed here consists in reducing the model to a “nearly black object” as studied in Donoho et al. (1991). The procedure consists of

three steps where the first is a robust smoothing of the data like in Tsybakov (1991). See also the survey paper by Pitas and Venetsanopoulos (1992) on nonlinear smoothing and the use of linear combinations of order statistics. The smoothed field is supposed to be pretty much stabilized around a certain fixed value on C_1° and around a different fixed value on C_2° . In the second step the smoothed process is convoluted with a special type of kernel, having an effect similar to differentiation but applicable to any function. Consequently the result of this convolution is supposed to be a process which is close to 0 everywhere except for a strip around the curve Γ . In fact a more sophisticated form of this second step is known as a wavelet transform. In the recent literature such wavelet transforms are extensively studied. For the relation with change detection in a deterministic setting we refer in particular to Mallat and Zhong (1992). The purpose of the third step is to further enhance the features of the last process by replacing each value by 0 except when it exceeds a certain suitably chosen threshold. The use of such a threshold is intuitively clear and can be mathematically justified as arising from a nonlinear (L_1 -) penalty function in a least squares setting (Donoho et al. (1991)).

Section 2 is devoted to a precise description of the procedure sketched above, and in Section 3 we present results on the speed of uniform almost sure convergence and some other asymptotic results. In Section 4 we briefly comment on the results and the assumptions.

2. Description of the Estimation Procedure. At stage n , let us choose $\varepsilon = \varepsilon_n \in (0, 1)$ such that

$$\lim_{n \rightarrow \infty} n^{\delta/2} \varepsilon = 1, \text{ for some } 0 < \delta < 2 \tag{2.1}$$

and, for convenience, $n \varepsilon, \varepsilon^{-1} \in \mathbb{N}$.

See below for a choice of δ . We divide $(0, 1]^2$ into intervals

$$T_n(i, j) = (\varepsilon(i - 1), \varepsilon i] \times (\varepsilon(j - 1), \varepsilon j], \quad (i, j) \in \{1, \dots, 1/\varepsilon\}^2. \tag{2.2}$$

Let us introduce the index sets

$$\mathcal{I}_{n,\Gamma} = \{(i, j) : T_n(i, j) \cap \Gamma \neq \emptyset\}; \text{ note that } \#\mathcal{I}_{n,\Gamma} \leq 2 n^{\delta/2}; \tag{2.3}$$

$$\mathcal{I}_{n,\alpha} = \{(i, j) : T_n(i, j) \cap C_\alpha \neq \emptyset\} \setminus \mathcal{I}_{n,\Gamma}, \text{ for } \alpha = 1, 2. \tag{2.4}$$

Of course we have $\mathcal{I}_{n,\Gamma} \cup \mathcal{I}_{n,1} \cup \mathcal{I}_{n,2} = \{1, \dots, n\}^2$. It will also be convenient to define the subsets

$$\Gamma_\varepsilon = \bigcup_{(i,j) \in \mathcal{I}_{n,\Gamma}} T_n(i, j); \tag{2.5}$$

$$C_{\alpha,\varepsilon} = \bigcup_{(i,j) \in \mathcal{I}_{n,\alpha}} T_n(i,j), \quad \text{for } \alpha = 1, 2. \tag{2.6}$$

It should be noted that

$$\Gamma_\varepsilon \downarrow \Gamma, \quad C_{\alpha,\varepsilon} \uparrow C_\alpha^\circ, \quad \text{for } \alpha = 1, 2, \quad \text{as } n \uparrow \infty. \tag{2.7}$$

Step 1. Let $J : [0, 1] \rightarrow [0, \infty)$ be a score function which is 0 in neighborhoods of 0 and 1 (but not identically equal to 0) and which has a bounded continuous first derivative on $(0, 1)$. We define

$$\hat{F}_{\varepsilon,t}(x) = \left(\frac{1}{\varepsilon n}\right)^2 \sum_{\left(\frac{k}{n}, \frac{\ell}{n}\right) \in T_n(i,j)} \sum I_{(-\infty, x]}(X_{k,\ell}), \quad x \in \mathbb{R}, \tag{2.8}$$

for $t \in T_n(i, j)$ and $(i, j) \in \{1, \dots, 1/\varepsilon\}^2$. This is the empirical d.f. for the block of observables with $(k, \ell) \in \{(i - 1)\varepsilon n + 1, \dots, i\varepsilon n\} \times \{(j - 1)\varepsilon n + 1, \dots, j\varepsilon n\}$, containing $(\varepsilon n)^2$ elements. The original process X_n will now be replaced with

$$Y_n(t) = \int_0^1 \hat{F}_{\varepsilon,t}^{-1}(s) J(s) ds, \quad t \in T_n(i, j), \quad (i, j) \in \{1, \dots, 1/\varepsilon\}^2. \tag{2.9}$$

This process is a step function with constant values on the $T_n(i, j)$.

For $t \in C_{\alpha,\varepsilon}$, the $X_n(t)$ have a common d.f. F_α , say, and the $Y_n(t)$ are robust estimators of $\int_0^1 F_\alpha^{-1}(s) J(s) ds$, $\alpha = 1, 2$. The precise values of these location functionals is not important: the only thing that matters is that they are different. Under the present assumptions F_1 and F_2 are obviously translates of each other and hence we have indeed

$$\int_0^1 F_1^{-1}(s) J(s) ds \neq \int_0^1 F_2^{-1}(s) J(s) ds. \tag{2.10}$$

More specifically if F_1 is continuous and symmetric about some point (so that F_2 enjoys these properties as well) we even have

$$\int_0^1 F_\alpha^{-1}(s) J(s) ds = c_\alpha, \quad \alpha = 1, 2, \tag{2.11}$$

provided that J is symmetric about $1/2$, where the c_α are defined in (1.4).

Step 2. For the same ε as in (2.1), let us define the kernel $K_\varepsilon : \mathbb{R}^2 \rightarrow \mathbb{R}$ by

$$K_\varepsilon(t) = \frac{1}{\varepsilon^{2q}} \{I_{[0,\varepsilon^q]^2}(t) - I_{[-\varepsilon^q,0]^2}(t)\} = K_\varepsilon^+(t) - K_\varepsilon^-(t), \quad t \in \mathbb{R}^2, \tag{2.12}$$

for some $0 < q < 1$, a suitable value of which will be given below. This function is particularly well suited to detect discontinuities along curves running

“northwest”, like the present curve Γ , when used as a convolution kernel. It satisfies $K_\varepsilon * c = 0$ for any constant function c on \mathbb{R}^2 and it has the property that $K_\varepsilon * \mu$, with μ defined in (1.4), will be 0 outside a strip with a width of order $\varepsilon^q = \varepsilon_n^q$ (as $n \rightarrow \infty$). The smooth process Y_n that stabilizes around the two values in (2.10) is now replaced by the process

$$Z_n(t) = K_\varepsilon * Y_n(t), \quad t \in (0, 1]^2, \tag{2.13}$$

$Z_n = 0$ elsewhere. The present choice of kernel is easy to work with but a smooth version with essentially the same properties might occasionally be more desirable since convolution with a smooth kernel yields a process Z_n that is also smooth.

Step 3. The process Z_n is likely to reveal the position of the change curve Γ by a ridge in its surface. It is possible to further improve on the signal to noise ratio by applying a nonlinear L_1 -smoothing technique employed by Donoho et al. (1991), leading to

$$\hat{\Gamma}_n(t) = Z_n(t)I_{[\lambda_n, \infty)}(|Z_n(t)|), \quad t \in (0, 1]^2, \tag{2.14}$$

for a suitably chosen threshold $\lambda_n \in (0, \infty)$. It is intuitively clear that $\hat{\Gamma}_n$ should usually better display the change curve than Z_n .

3. Some Asymptotics. The following facts about linear combinations of order statistics that turn out to be expedient in the analysis of the process Y_n can be found in Boos (1979); see also Ruymgaart (1981). Let \mathcal{F} denote the class of all d.f.’s on \mathbb{R} and consider the functional $\theta : \mathcal{F} \rightarrow \mathbb{R}$, defined by

$$\theta(F) = \int_0^1 F^{-1}(s)J(s)ds, \quad F \in \mathcal{F}. \tag{3.1}$$

Note that $Y_n(t) = \theta(\hat{F}_{\varepsilon,t})$. For any $F \in \mathcal{F}$ we have

$$Y_n(t) - \theta(F) = \int_{-\infty}^{\infty} \left\{ \hat{F}_{\varepsilon,t}(x) - F(x) \right\} J(F(x))dx + R_{\varepsilon,t}, \tag{3.2}$$

where the remainder $R_{\varepsilon,t}$ equals

$$R_{\varepsilon,t} = \frac{1}{2} \int_{-\infty}^{\infty} \left\{ \hat{F}_{\varepsilon,t}(x) - F(x) \right\}^2 J'(F(x))dx, \tag{3.3}$$

with $\tilde{F}_{\varepsilon,t}(x)$ between $\hat{F}_{\varepsilon,t}(x)$ and $F(x)$ for every $x \in \mathbb{R}$.

ASSUMPTION 3.1. It will be assumed, often without explicit reference that (1.4)–(1.9), as well as (2.1) and the assumption that F_1 and F_2 are continuous, are fulfilled.

LEMMA 3.1. *Provided that $0 < \zeta < (1 - \rho)/2$ we have, for $\alpha = 1, 2$,*

$$n^{(2-\delta)\zeta} \sup_{t \in C_{\alpha,\varepsilon}} \sup_{x \in \mathbf{R}} |\hat{F}_{\varepsilon,t}(x) - F_{\alpha}(x)| \xrightarrow{a.s.} 0, \quad \text{as } n \rightarrow \infty. \quad (3.4)$$

PROOF. Recall that $t \mapsto \hat{F}_{\varepsilon,t}(x)$ is constant for $t \in T_n(i, j)$. Let us take $T_n(i, j) \subset C_{\alpha,\varepsilon}$ and observe that to prove the present asymptotic properties we may as well assume that $c_{\alpha} = 0$ so that we are dealing with the $E_{i,j}$. Each $\hat{F}_{\varepsilon,t}$ being based on $(\varepsilon n)^2$ observations, it follows from a modification for two-dimensional indices of Chanda and Ruymgaart (1991, Theorem 2.1) that, for any $c \in (0, \infty)$,

$$\begin{aligned} & \mathbb{P} \left(\left\{ \sup_{x \in \mathbf{R}} n^{(2-\delta)\zeta} |\hat{F}_{\varepsilon,t}(x) - F_{\alpha}(x)| \geq c \right\} \cap \Omega_n \right) \\ & \leq C n^{2-\delta} \exp \left(-A c^2 n^{(2-\delta)(1-\rho-2\zeta)} \right), \end{aligned} \quad (3.5)$$

where $\Omega_n = \{\max_{i,j} |\bar{E}_{i,j}^{(n)}| < \gamma(n)\}$ (see (1.7)) and $A, C \in (0, \infty)$ are fixed numbers (independent of n).

Writing $(i\varepsilon, j\varepsilon) = t_{i,j}$, it follows that

$$\begin{aligned} & \mathbb{P} \left(\left\{ \sup_{t \in C_{\alpha,\varepsilon}} \sup_{x \in \mathbf{R}} n^{(2-\delta)\zeta} |\hat{F}_{\varepsilon,t}(x) - F_{\alpha}(x)| \geq c \right\} \cap \Omega_n \right) \\ & \leq n^{\delta} \mathbb{P} \left(\left\{ \sup_{x \in \mathbf{R}} n^{(2-\delta)\zeta} |\hat{F}_{\varepsilon,t_{i,j}}(x) - F_{\alpha}(x)| \geq c \right\} \cap \Omega_n \right) \\ & \leq C n^{\delta} n^{2-\delta} \exp \left(-A c^2 n^{(2-\delta)(1-\rho-2\zeta)} \right). \end{aligned} \quad (3.6)$$

Because $\mathbb{P}(\Omega_n^c) = O(n^{-2\tau})$, as $n \rightarrow \infty$, and by assumption $(2-\delta)(1-\rho-2\zeta) > 0$ the claimed almost sure convergence follows. ■

LEMMA 3.2. *For $0 < \zeta < (1 - \rho)/2$, the processes Y_n satisfy*

$$n^{(2-\delta)\zeta} \sup_{t \in C_{\alpha,\varepsilon}} |Y_n(t) - \theta(F_{\alpha})| \xrightarrow{a.s.} 0, \quad \text{as } n \rightarrow \infty. \quad (3.7)$$

PROOF. For each $t \in C_{\alpha,\varepsilon}$, the absolute value of the first term on the right in (3.2), with $\theta(F)$ replaced by $\theta(F_{\alpha})$, is obviously bounded by

$$\begin{aligned} & \sup_{t \in C_{\alpha,\varepsilon}} \sup_{x \in \mathbf{R}} |\hat{F}_{\varepsilon,t}(x) - F_{\alpha}(x)| \int_{-\infty}^{\infty} J(F(x)) dx \\ & = o \left(n^{-(2-\delta)\zeta} \right), \quad \text{as } n \rightarrow \infty, \quad \text{a.s.}, \end{aligned} \quad (3.8)$$

according to (3.4). For $R_{n,t}$ in (3.3), let us note that the assumptions on J entail that $J' = 0$ on $[0, a] \cup [1 - a, 1]$ for some $0 < a < \frac{1}{2}$. Since $\tilde{F}_{\varepsilon,t}$ is between $\hat{F}_{\varepsilon,t}$ and F_α , it follows once more from (3.4) that the integration in (3.3) is effectively restricted to an interval contained in $[F_\alpha^{-1}(a/2), F_\alpha^{-1}(1 - a/2)]$ for n sufficiently large, a.s. It follows that

$$\begin{aligned} \sup_{t \in C_{\alpha,\varepsilon}} |R_{\varepsilon,t}| &\leq \left\{ \frac{1}{2} \sup_{t \in C_{\alpha,\varepsilon}} \sup_{x \in \mathbf{R}} |\hat{F}_{\varepsilon,t}(x) - F_\alpha(x)|^2 \right\} \\ &\quad \times \{F_\alpha^{-1}(1 - a/2) - F_\alpha^{-1}(a/2)\} \sup_{a/2 \leq s \leq 1 - a/2} |J'(s)| \\ &= o(n^{-2(2-\delta)\zeta}), \text{ as } n \rightarrow \infty, \text{ a.s.} \end{aligned} \tag{3.9}$$

Hence the order in (3.8) prevails. ■

Let us define strips

$$\Gamma_{\varepsilon,k} = \bigcup_{(t_1,t_2) \in \Gamma} [t_1 - k\varepsilon^q, t_1 + k\varepsilon^q] \times [t_2 - k\varepsilon^q, t_2 + k\varepsilon^q], \quad k \in \mathbb{N}, \tag{3.10}$$

around the curve and briefly write

$$\bar{\Gamma}_{\varepsilon,k} = (0, 1] \times (0, 1] \setminus \Gamma_{\varepsilon,k}. \tag{3.11}$$

Lemma 3.3. For $0 < \zeta < (1 - \rho)/2$, the processes Z_n satisfy

$$n^{(2-\delta)\zeta} \sup_{t \in \bar{\Gamma}_{\varepsilon,2}} |Z_n(t)| \xrightarrow{a.s.} 0, \text{ as } n \rightarrow \infty. \tag{3.12}$$

Proof. According to (3.7), there exists a measurable $N \subset \Omega$ with $\mathbb{P}(N) = 0$, and $c_n = c_n(\omega) \downarrow 0$ (as $n \rightarrow \infty$) for each $\omega \in \Omega \setminus N$ such that we have

$$\begin{aligned} \theta(F_\alpha) - c_n n^{-(2-\delta)\zeta} &\leq K_\varepsilon^\pm * (\theta(F_\alpha) - c_n n^{-(2-\delta)\zeta}) \\ &\leq K_\varepsilon^\pm * Y_n \leq K_\varepsilon^\pm * (\theta(F_\alpha) + c_n n^{-(2-\delta)\zeta}) \leq \theta(F_\alpha) + c_n n^{-(2-\delta)\zeta}, \end{aligned} \tag{3.13}$$

for $t \in C_{\alpha,\varepsilon} \cap \bar{\Gamma}_{\varepsilon,2}$. This entails

$$|Z_n(t)| \leq 2 c_n n^{-(2-\delta)\zeta}, \quad t \in \bar{\Gamma}_{\varepsilon,2}, \tag{3.14}$$

and the lemma follows. ■

In order to investigate the behavior of the estimators at points t on the curve Γ we need some further preparations. First of all it should be noted that all we have been doing so far for the observations $X_{i,j}$ could also be done – mutatis mutandis – for the absolute values $|X_{i,j}|$. Let us denote the

successive processes based on the absolute values by $|X|_n(t)$, $|Y|_n(t)$, and $|Z|_n(t)$, $t \in (0, 1]^2$, and let us write $|F|_\alpha$ for the d.f. of $|X|_n(t)$ for $t \in C_{\alpha,\varepsilon}$. In particular we have

$$n^{(2-\delta)\zeta} \sup_{t \in C_{\alpha,\varepsilon}} ||Y|_n(t) - \theta(|F|_\alpha)| \xrightarrow{a.s.} 0, \quad \text{as } n \rightarrow \infty, \tag{3.15}$$

and of course $|Y_n(t)| \leq |Y|_n(t)$ for all t .

For an arbitrary $t \in \Gamma$ let us write

$$\begin{cases} S_{1,\varepsilon}(t) = [t_1 - \varepsilon^q, t_1] \times [t_2 - \varepsilon^q, t_2], \\ S_{2,\varepsilon}(t) = [t_1, t_1 + \varepsilon^q] \times [t_2, t_2 + \varepsilon^q]; \end{cases} \tag{3.16}$$

$$\begin{cases} Q_{1,\varepsilon}(t) = [t_1 - \varepsilon^q, t_1 - \varepsilon] \times [t_2 - \varepsilon^q, t_2 - \varepsilon], \\ Q_{2,\varepsilon}(t) = [t_1 + \varepsilon, t_1 + \varepsilon^q] \times [t_2 + \varepsilon, t_2 + \varepsilon^q]. \end{cases} \tag{3.17}$$

Note the relationship between the $S_{1,\varepsilon}(t)$, $S_{2,\varepsilon}(t)$ and $\Gamma_{\varepsilon,1}$. Finally, let us write

$$\Gamma^{(\varepsilon)} = \Gamma \cap [\varepsilon^q, 1 - \varepsilon^q]^2, \tag{3.18}$$

for the part of the curve that doesn't come too close to the boundary of the unit square and let us note that

$$C_{\alpha,\varepsilon} \supset S_{\alpha,\varepsilon}(t) \cap \Gamma_\varepsilon^c \supset Q_{\alpha,\varepsilon}(t), \quad t \in \Gamma^{(\varepsilon)}, \quad \alpha = 1, 2. \tag{3.19}$$

LEMMA 3.4. For $0 < \xi < \delta(1 - q)/2 \wedge (2 - \delta)\zeta$, the processes Z_n satisfy

$$n^\xi \sup_{t \in \Gamma^{(\varepsilon)}} |Z_n(t) - \{\theta(F_1) - \theta(F_2)\}| \xrightarrow{a.s.} 0, \quad \text{as } n \rightarrow \infty. \tag{3.20}$$

PROOF. First observe that, for $t \in \Gamma^{(\varepsilon)}$,

$$\begin{aligned} & |Z_n(t) - \{\theta(F_1) - \theta(F_2)\}| \tag{3.21} \\ & \leq \left| \varepsilon^{-2q} \int \int_{S_{1,\varepsilon}(t)} Y_n(s) ds - \theta(F_1) \right| + \left| \varepsilon^{-2q} \int \int_{S_{2,\varepsilon}(t)} Y_n(s) ds - \theta(F_2) \right|. \end{aligned}$$

Since both terms on the right in (3.21) can be dealt with similarly, let us focus on the first one. Application of (3.7), (3.15), and (3.19) yields that, almost surely,

$$\left| \varepsilon^{-2q} \int \int_{S_{1,\varepsilon}(t)} Y_n(s) ds - \theta(F_1) \right| \tag{3.22}$$

$$\begin{aligned}
 &\leq \left| \varepsilon^{-2q} \int \int_{Q_{1,\varepsilon}(t)} Y_n(s) ds - \theta(F_1) \right| + \varepsilon^{-2q} \int \int_{S_{1,\varepsilon}(t) \setminus Q_{1,\varepsilon}(t)} |Y|_n(s) ds \\
 &\leq \{1 - \varepsilon^{-2q}(\varepsilon^q - \varepsilon)^2\} \theta(F_1) + \varepsilon^{-2q}(\varepsilon^q - \varepsilon)^2 n^{-(2-\delta)\zeta} + \varepsilon^{-2q} \int \int_{S_{1,\varepsilon}(t) \setminus Q_{1,\varepsilon}(t)} |Y|_n(s) ds \\
 &\leq 2\varepsilon^{1-q} \theta(F_1) + n^{-(2-\delta)\zeta} + \varepsilon^{-2q} \int \int_{S_{1,\varepsilon}(t) \setminus Q_{1,\varepsilon}(t)} |Y|_n(s) ds.
 \end{aligned}$$

It remains to obtain an a.s. upper bound for the process $|Y|_n(s)$ for $s \in \Gamma_\varepsilon$. For any such s , this process is based on $(\varepsilon n)^2$ sample elements $|X_{i,j}|$ a fraction of which has d.f. $|F|_1$ and the remainder fraction of which has d.f. $|F|_2$. Since all observations are nonnegative, so is the linear combination of order statistics. By completing the fractions to two samples each of size $(\varepsilon n)^2$, and one with sample elements having d.f. $|F|_1$ and the other with sample elements having d.f. $|F|_2$, we may add the corresponding linear combinations of order statistics and arrive at a process $|Y^*|_n(s)$ which clearly satisfies, for almost every $\omega \in \Omega$,

$$0 \leq |Y|_n(s) \leq |Y^*|_n(s) \leq \theta(|F|_1) + \theta(|F|_2) + c_n^*(\omega) n^{-(2-\delta)\zeta}, \tag{3.23}$$

where $c_n^*(\omega) \downarrow 0$, as $n \rightarrow \infty$, and where $s \in \Gamma_\varepsilon$. It follows that, almost surely,

$$\begin{aligned}
 &\varepsilon^{-2q} \int \int_{S_{1,\varepsilon}(t) \setminus Q_{1,\varepsilon}(t)} |Y|_n(s) ds \\
 &\leq 2\varepsilon^{1-q} \{ \theta(|F|_1) + \theta(|F|_2) + n^{-(2-\delta)\zeta} \}.
 \end{aligned} \tag{3.24}$$

The range for ξ is easily obtained by combining the orders of magnitude appearing in (3.22) and (3.24). ■

THEOREM 3.1. *Let Assumption 3.1 be satisfied. Then for almost every $\omega \in \Omega$,*

$$\begin{cases} \text{there exists } n(\omega) \text{ such that } \sup_{t \in \bar{\Gamma}_{\varepsilon,2}} |\hat{\Gamma}_n(t)| = 0, & \text{for } n \geq n(\omega), \\ \sup_{t \in \Gamma(\varepsilon)} n^\xi |\hat{\Gamma}_n(t) - \{\theta(F_1) - \theta(F_2)\}| \longrightarrow 0, & \text{as } n \rightarrow \infty, \end{cases} \tag{3.25}$$

for any $0 < \zeta < (1 - \rho)/2$, and $0 < \xi < \delta(1 - q)/2 \wedge (2 - \delta)\zeta$, provided that we choose the threshold $\lambda_n = \lambda n^{-\eta}$, for some $0 < \eta < (2 - \delta)\zeta$ and $0 < \lambda < \infty$.

PROOF. This is immediate from the lemmas. For $t \in \bar{\Gamma}_{\varepsilon,2}$ we have $|\hat{\Gamma}_n(t)| \leq |Z_n(t)|$ so that the first part of the equality in (3.25) is immediate from (3.12). We see from (3.20) that for n sufficiently large $|Z_n(t)| \geq |\theta(F_1) - \theta(F_2)| - n^{-\xi}$, uniformly for $t \in \Gamma(\varepsilon)$, a.s. Hence eventually we have

$|Z_n(t)| > \lambda_n$ for all $t \in \Gamma^{(\varepsilon)}$ simultaneously, meaning that $\hat{\Gamma}_n(t) = Z_n(t)$, a.s., for all $t \in \Gamma^{(\varepsilon)}$ simultaneously provided that n is sufficiently large. ■

The theorem enables us to eventually identify the strip $\Gamma_{\varepsilon,3}$ in which $\Gamma^{(\varepsilon)}$ must lie. As an estimator of $\Gamma^{(\varepsilon)}$ we might take any curve in this strip that satisfies the prior assumptions. Since the width of the strip is of order $n^{-\delta q/2}$, as $n \rightarrow \infty$, this settles a speed of almost sure convergence. For practical purposes data-driven parameter selection is one of the problems that remains to be considered. Although a slow convergence rate of the process Z_n due to strong dependence (i.e. large ρ) doesn't seem to affect the order of ε asymptotically, it is very likely to have an impact on the constants determining the actual strip width.

4. Some Comments.

1. The curves Γ that we consider here run “northwest”. When we know this we should choose kernels that run “northeast”. Conversely, curves that run “northeast” can be better detected by kernels running “northwest”. When we don't know what kind of curve we are dealing with, or when the regions of constant values are separated by a simple closed curve we might screen the entire domain twice: once with a northeast and once with a northwest kernel.

2. As we observed already in the introduction, the asymptotic decomposability condition is trivially fulfilled in the case of i.i.d. errors. Under mild assumptions, linear random fields satisfy the conditions, as we have seen. Since bilinear processes for one-dimensional time are shown to be asymptotically decomposable (Chanda and Ruyngaert (1991)) we may conjecture that bilinear fields also satisfy that condition, in which case an important class of nonlinear processes would be included. Of course, asymptotic decomposability in its present generality is hard to verify. Our main purpose was, however, to show that our search procedure has rather good asymptotic properties for a broad class of error fields that goes far beyond the i.i.d. case.

Acknowledgments. We are grateful to Manny Parzen for sending a reprint of his paper and to David Donoho for pointing out the usefulness of kernels like those in (2.12) to detect boundaries and for sending a copy of his paper. We would also like to thank Ed Carlstein for some useful remarks.

REFERENCES

- Boos, D. D. (1979). A differential for L -statistics. *Ann. Statist.* **7**, 955–959.
 CARLSTEIN, E. (1988). Nonparametric change point estimation. *Ann. Statist.* **16**, 188–197.

- CARLSTEIN, E. and KRISHNAMOORTHY, C. (1992). Boundary estimation. *J. Amer. Statist. Assoc.* **87**, 430–438.
- CHANDA, K. C., PURI, M. L. and RUYMGAART, F. H. (1990). Asymptotic normality of L -statistics based on $m(n)$ -decomposable time series. *J. Multivariate Anal.* **35**, 260–275.
- CHANDA, K. C. and RUYMGAART, F. H. (1991). Curve estimation for $m(n)$ -decomposable time series including bilinear processes. *J. Multivariate Anal.* **38**, 149–166.
- CSÖRGÖ, M. and HORVÁTH, L. (1988). Invariance principles for changepoint problems. *J. Multivariate Anal.* **27**, 151–168.
- DONOHO, D. L., JOHNSTONE, I. M., HOCH, J. C. and STERN, A. S. (1991). Maximum entropy and the nearly black object. *J. Roy. Statist. Soc. Ser. B* **53**, 41–81.
- MALLAT, S. and ZHONG, S. (1992). Wavelet transform maxima and multiscale edges. In: *Wavelets and Their Applications*, pp. 67–104 (M.B. Ruskai, Ed.). Jones and Bartlett, Boston.
- PARZEN, E. (1991). Comparison change analysis. To appear in: *Proc. Internat. Symp. Nonparametric Statist.* (A. K. Saleh, Ed.). Elsevier, Amsterdam.
- PETTITT, A. N. (1979). A non-parametric approach to the change point problem. *Appl. Statist.* **28**, 126–135.
- PITAS, I. and VENETSANOPOULOS, A. N. (1992). Order statistics in digital image processing. *Proc. IEEE* **80**, 1892–1921.
- PRIESTLEY, M. B. (1981). *Spectral Analysis and Time Series*, Vol. 1. Academic Press, New York.
- PURI, M. L. and RUYMGAART, F. H. (1991). Nonparametric prediction for random fields. *Stochastic Process. Appl.* **48**, 139–156.
- RUYMGAART, F. H. (1981). A robust principal component analysis. *J. Multivariate Anal.* **11**, 485–497.
- RUYMGAART, F. H. (1991). Conditional empirical, quantile and difference processes for a large class of time series with applications. *J. Statist. Pl. Inf.*, to appear.
- TSYBAKOV, A. B. (1991). Nonparametric techniques in image estimation. In: *Nonparametric Functional Estimation and Related Topics*, pp. 669–677 (G. Roussas, Ed.). Kluwer, Dordrecht.
- VAN DE GEER, S. (1988). *Regression Analysis and Empirical Processes*. CWI-tract 45, Centre for Mathematics and Computer Science, Amsterdam.

WOLFE, D. A. and SCHECHTMAN, E. (1984). Nonparametric statistical procedures for the changepoint problem. *J. Statist. Plann. Inference* **9**, 389–396.

DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
BLOOMINGTON, IN 47405

DEPARTMENT OF MATHEMATICS
TEXAS TECH UNIVERSITY
LUBBOCK, TX 79409