Probab. Theory Relat. Fields 114, 55–84 (1999)

Probability Theory and Related Fields

Asymptotic distribution of the empirical spatial cumulative distribution function predictor and prediction bands based on a subsampling method⋆

S.N. Lahiri

Department of Statistics, Iowa State University, Ames, IA 50011, USA

Received: 14 July 1997 / Revised version: 2 June 1998

Abstract. A spatial cumulative distribution function $\hat{F}_\infty$ (say) is a random distribution function that provides a statistical summary of random field over a given region. This paper considers the empirical predictor of $\hat{F}_\infty$ based on a finite set of observations from a region in $\mathbb{R}^d$ under a uniform sampling design. A functional central limit theorem is proved for the predictor as a random element of the space $D[-\infty, \infty]$. A striking feature of the result is that the rate of convergence of the predictor to the predictand $\hat{F}_\infty$ depends on the location of the data-sites specified by the sampling design. A precise description of the dependence is given. Furthermore, a subsampling method is proposed for integral-based functionals of random fields, which is then used to construct large sample prediction bands for $\hat{F}_\infty$.

Mathematics Subject Classification (1991): Primary 62E20; Secondary 62G30, 60G60

1. Introduction

Let $Z(\cdot)$ be a measurable random field (r.f.) on $\mathbb{R}^d$, $d \geq 1$, that serves as a model for a spatially distributed univariate quantity of interest. For a given region R in $\mathbb{R}^d$,

Key words and phrases: Functional Limit Theorem, Spatial Cumulative Distribution Function, Prediction Band, Subsampling

⋆ Research partially supported by EPA contract no. CR82291901 and NSF grant no. DMS 9505124.

$$\hat{F}_\infty(z; R) = |R|^{-1} \int_R I(Z(\mathbf{s}) \le z)\, \mathbf{ds}, \quad z \in \mathbb{R} \tag{1.1}$$

is called the spatial cumulative distribution function (CDF) of the process $Z(\cdot)$ over the region R. Here, $|R|$ denotes the volume of R and $I(\cdot)$ denotes the indicator function. Note that $\hat{F}_\infty(\cdot; R)$ given by (1.1) is a random CDF in the sense that it is a nondecreasing, right continuous function with $\lim_{z\to\infty} \hat{F}_\infty(z; R) = 1$ and $\lim_{z\to\infty} \hat{F}_\infty(-z; R) = 0$ for every realization of $\{Z(\mathbf{s}), \mathbf{s} \in R\}$. Thus, just as a probability distribution summarizes important characteristics of a population, the spatial CDF $\hat{F}_\infty(z; R)$ quite effectively summarizes the statistical information on the process $Z(\cdot)$ over the region R. For example, if $Z(\mathbf{s})$ represents the concentration of a pollutant at site $\mathbf{s}$ in the geographical region of interest R, one of the statistical summaries that one might be interested in is the average concentration of the pollutant over the region R, given by $Z(R) \equiv \int_R Z(\mathbf{s})\, \mathbf{ds}/|R|$. This can be very easily recovered from the knowledge of the spatial CDF $\hat{F}_\infty(z; R)$, since $Z(R) = \int_{\mathbb{R}} z\, d\hat{F}_\infty(z; R)$. If, on the other hand, one is interested in the proportion of area within the region R where the concentration level of the pollutant exceeds a prescribed safety level z_0, say, the relevant quantity is given by $\int_R I(Z(\mathbf{s}) > z_0)\, \mathbf{ds}/|R| = 1 - \hat{F}_\infty(z_0; R)$, which is again summarized by the spatial CDF $\hat{F}_\infty(\cdot; R)$. Because of its summarization capability and visual appeal (cf. Majure, Cook, Cressie, Kaiser, Lahiri, and Syamanzik, 1995), the spatial CDF $\hat{F}_\infty(\cdot; R)$ is a basic functional of the underlying process $Z(\cdot)$ that is of interest in many applications involving spatially distributed random processes. For an example of an application of the spatial CDF in the context of ecological resource monitoring, see Lahiri, Kaiser, Hsu, and Cressie (1999).

Note that the spatial CDF $\hat{F}_\infty(\cdot; R)$, as defined above, depends on the entire collection of random variables (r.v.s) $Z(\mathbf{s}), \mathbf{s} \in R$ and hence remains unknown to the statistician. Typically, observations on the process $Z(\mathbf{s})$ are taken at finitely many data-sites in R and $\hat{F}_\infty$ is to be predicted from the observed values. Suppose $\{Z(\mathbf{s}_1), \ldots, Z(\mathbf{s}_N)\}$ denote the available data, observed at sampling sites $\mathbf{s}_1, \ldots, \mathbf{s}_N \in R$. The most commonly used predictor of $\hat{F}_\infty$ is the empirical spatial CDF predictor given by

$$\hat{F}_N(z; R) = N^{-1} \sum_{i=1}^{N} I(Z(\mathbf{s}_i) \le z), \quad z \in \mathbb{R}\ . \tag{1.2}$$

In this paper, we address the problems of determining the asymptotic distribution of the centered spatial predictor $\hat{F}_N - \hat{F}_\infty$ (with a suitable scaling sequence) and of constructing large sample prediction bands for the underlying random CDF $\hat{F}_\infty$ based on $\hat{F}_N$. For both the problems, the structure of the sampling design plays a crucial role. The sampling structure we are going to assume is a combination of what are known as the "increasing domain" and the "infill" asymptotic structures (see Section 2 for more details on these notions). The increasing domain component of it involves letting the sampling region $R \equiv R_n$ grow as the sample size $N \equiv N_n$ increases, while the other part requires "filling in" any given bounded subregion of R_n with an increasingly densely placed points from a *uniform* sampling design. A similar formulation has been used by Hall and Patil (1994) in the context of estimating the autocovariance function of a r.f. However, here we work under a fixed design framework as compared to their stochastic design formulation.

We allow the sampling region R_n to have a fairly irregular shape. Essentially, any subset of $\mathbb{R}^d$ that can be obtained by magnifying a member of a large class of Borel subsets of $(-1/2, 1/2]^d$ (that contain the origin) can serve as a sampling region (see Section 2 for details). This, in particular, covers polyhedrons, spheres, and many non-convex regions in $\mathbb{R}^d$. The magnification is achieved through scaling up the subset of $(-1/2, 1/2]^d$ by a factor λ_n that tends to infinity with n and thus, makes the sampling region grow with the sample size.

One of the main results of the paper is a Functional Central Limit Theorem for the normalized predictor process

$$\xi_n \equiv b_n(\hat{F}_n(\cdot;\, R_n) - \hat{F}_\infty(\cdot;\, R_n))$$

where b_n is a scaling constant and where, for notational simplicity, we write $\hat{F}_n(\cdot;\, R_n) \equiv F_{N_n}(\cdot;\, R_n)$. In deriving the asymptotic distributional result for $\hat{F}_n(\cdot;\, R_n)$, one of the issues that require some nontrivial consideration is the determination of the right sequence $\{b_n\}$ of scaling constants that ensure weak convergence of the process ξ_n to a nondegenerate limit. To appreciate why, consider the following analogous situation involving a deterministic function f on an interval $[a, b]$. Let $t_i = a + i(b-a)/m$, $i = 0, 1, \ldots, m$ denote a partition of $[a, b]$ by $(m + 1)$ equispaced points, and let $\nu_i = a + [i(b-a)+c]/m$ be a point in the interval $[t_i, t_{i+1}]$, $i = 0, 1, \ldots, m-1$, where $c \in (0, b-a)$. Then, the rate at which the partial sum $\sum_{i=0}^{m-1} f(\nu_i)(b-a)/m$ approximates $\int_a^b f(t)\,dt$ depends on c. Indeed, if the function f is sufficiently "well-behaved", then using Taylor's expansion, one can show that the difference

$$\sum_{i=0}^{m-1} f(\nu_i)(b-a)/m - \int_a^b f(t)\,dt$$

goes to zero at the rate $0(m^{-1})$ if $c \neq (b-a)/2$, and it decays much faster, viz., at the rate $0(m^{-2})$ when $c = (b-a)/2$.

A similar phenomenon occurs when we consider the rate of convergence of the difference $\hat{F}_n(\cdot; R_n) - \hat{F}_\infty(\cdot; R_n)$. However, unlike the simplistic situation of a *smooth* and *deterministic* function f over a *fixed* finite interval $[a, b]$, for establishing weak convergence of the process ξ_n, we have to deal with some additional complications arising from

(i) the *nonsmooth* character of the indicator function defining $\hat{F}_n(\cdot; R_n) - \hat{F}_\infty(\cdot; R_n)$;
(ii) the *randomness* of the function $I(Z(\mathbf{s}) \le z)$ involved; and
(iii) the *relative growth* rate of the sampling region, compared to the rate of 'infilling' by the sampling design.

In Section 2, we show that depending on the starting point **c**, say, of the cubic sampling grid (that determines our sampling design), the right choice of the normalizing constant b_n can be

$$b_n = \lambda_n^{d/2} h_n^{-1}$$

or

$$b_n = \lambda_n^{d/2} h_n^{-2} \quad ,$$

where $h_n \to 0$ is a sequence of constants specified by sampling design. The larger normalizing constant, given by the second equation, is appropriate only for a specific value of **c** that enjoys certain symmetry properties (like, $c = (b-a)/2$ in the deterministic case). An important implication of this is that the accuracy of the empirical spatial CDF predictor can be significantly enhanced simply by choosing a single design parameter **c** suitably.

The second problem considered in this paper is the construction of prediction regions for the underlying spatial CDF $\hat{F}_\infty(\cdot; R_n)$. Though, in principle, the large sample distribution of the normalized predictor can be used to get a prediction region, it may not be very convenient for practical applications. As follows from Theorem 2.1 below, the asymptotic covariance function of ξ_n is an integral of certain higher order partial derivatives of the bivariate joint probability distribution functions of the $Z(\mathbf{s})$s, and is rather awkward for direct estimation. Furthermore, that approach would call for additional smoothness assumptions if the integrals of the derivatives are estimated nonparametrically. Instead, we propose a generalization of the standard subsampling method, used in the context of time-series and lattice processes (cf. Possolo, 1991; Politis and Romano, 1994; Hall and Jing, 1996; and Sherman and Carlstein, 1994), to r.f.s with a continuous spatial index. The basic idea is to construct smaller subregions within the given sampling region R_n that have similar shape as R_n (cf. Sherman and Carlstein, 1994) and, then, use the sampling design to recreate the effects of a "sample" and the "population" at the level of the subregions. In Section 3, we construct a large sample prediction band for $\hat{F}_\infty(\cdot;\, R_n)$ using the proposed subsampling method and show that it attains the desired confidence level asymptotically. This form of subsampling can also be used effectively in inference problems concerning other integral-based functionals of r.f.s. Some related recent works in this direction include Lahiri, Kaiser, Cressie, and Hsu (1999), Bertail, Politis, and Rhomari (1996), and Politis, Paparoditis, and Romano (1996).

The rest of the paper is organized as follows. Section 2 states the assumptions and the main results on the asymptotic distribution of the spatial CDF predictor. In Section 3, the subsampling method is introduced and is applied to construct prediction bands for $\hat{F}_\infty(\cdot;\, R_n)$. Proofs of the results in Sections 2 and 3 are given in Section 4.

2. Asymptotic distribution

This section is divided into three parts. In Section 2.1, we describe the sampling design and the structure of the sampling region and, in Section 2.2, we state the assumptions used in the paper. Asymptotic distributional results on the spatial CDF predictor $\hat{F}_n$ are given in Section 2.3.

2.1. The sampling structure

There are essentially two basic sampling structures for studying asymptotic properties of estimators and predictors based on spatial data. When all sampling sites are separated by a fixed positive distance, and the sampling region R_n becomes unbounded as the sample size increases, the resulting structure leads to what is known as the 'increasing domain asymptotics' (cf. Cressie, 1993). This is the most common framework used for asymptotics for spatial data, and often leads to results similar to those obtained in time series. The other form, known as the 'infill asymptotic structure' (cf. Cressie, 1993), is inherently different and is more suitable for inference for continuous parameter r.f.s. observed on bounded regions. When an increasing number of samples are collected from within a sampling region R_n that does not become unbounded with the sample size, we obtain the 'infill' structure. However, for the purpose of our study, neither the 'infill' nor the 'increasing domain' structure seems suitable. We assume a sampling scheme that is a mixture of both in the sense that we let the sampling region R_n to grow, and at the same time, allow 'infilling' of any fixed bounded subregion of R_n. This structure appears to be the natural one for this problem, because of the following reasons. Since $\hat{F}_\infty$ is defined in terms of an *integral* on R_n, $\hat{F}_\infty$ cannot be predicted consistently without infilling. On the other hand, if the region R_n remains bounded, then there is not enough information to allow consistent estimation of population quantiles for constructing prediction bands for $\hat{F}_\infty$ (cf. Lahiri, 1996). A similar "mixed" structure has been used by Hall and Patil (1994) in the context of nonparametric estimation of the auto-covariance function of an r.f. However, unlike their stochastic design, here we will consider a fixed sampling design.

To describe the sampling structure used in this paper, suppose that $\mathscr{P}_n$ denotes a partition of $\mathbb{R}^d$ by equal-volume cubes with sides h_n, where $h_n \downarrow 0$ as $n \to \infty$. We can identify $\mathscr{P}_n$ with the set:

$$\mathscr{P}_n = \{(\mathbf{i} + \Delta_0)h_n : \mathbf{i} \in \mathscr{Z}^d\}$$

where $\mathscr{Z}$ denotes the set of all integers and $\Delta_0 \equiv (0, 1]^d$ denotes the unit cube in $\mathbb{R}^d$. Note that just as it partitions $\mathbb{R}^d$, $\mathscr{P}_n$ also induces a simultaneous partition of any given region $R_n \subset \mathbb{R}^d$ by cubes of volume h_n^d. We assume that the sampling sites are selected on a *regular* grid such that there is *exactly* one potential sampling site in each cube $\Gamma_\mathbf{i} \equiv (\mathbf{i} + \Delta_0)h_n$. As mentioned in the Introduction, the choice of starting point of the grid, or equivalently, the choice of the sampling site within a $\Gamma_\mathbf{i}$ has a nontrivial effect on the rate of convergence of the spatial CDF predictor $\hat{F}_n$. As a consequence, we need to specify exactly how these sites are selected. Let $\mathbf{c}$ be an arbitrary point in the interior of the unit cube Δ_0. Then, the sampling sites within the region R_n are given by the points on the grid $\{(\mathbf{i}+\mathbf{c})h_n : \mathbf{i} \in \mathscr{Z}^d\}$ that lie inside R_n. To be more specific, let $J(R_n)$ denote the set $J(R_n) = \{\mathbf{i} \in \mathscr{Z}^d : \mathbf{s_i} \in R_n\}$, where $\mathbf{s_i} = (\mathbf{i} + \mathbf{c})h_n$. Then,

$$\{\mathbf{s_i} : \mathbf{i} \in J(R_n)\}$$

gives the collection of sampling sites in R_n. It is evident from the above description that this yields a nonstochastic uniform sampling design where the sampling sites are located on a grid whose "starting" point is an arbitrary point in the cube $(0, h_n]^d$.

Next we specify the structure of the regions R_n, which quantifies the "increasing domain" component of our sampling structure. Here we adopt a formulation similar to that of Sherman and Carlstein (1994) (see also Hall and Patil, 1994). Let R_0 be a Borel subset of $(-1/2, 1/2]^d$ containing an open neighbourhood of the origin such that for any sequence of positive real numbers $a_n \to 0$, the number of cubes of the integer lattice $a_n\mathscr{Z}^d$ that intersect both R_0 and R_0^c is $O((a_n^{-1})^{d-1})$ as $n \to \infty$. Also, let $\{\lambda_n\}$ be a sequence of real numbers that goes to infinity with n. Then, the sampling region R_n is obtained by "inflating" the set R_0 by the scaling sequence λ_n, i.e.,

$$R_n = \lambda_n R_0 \ .$$

Since the origin is assumed to lie inside R_0, the shape of the sampling region is preserved for different values of n. Furthermore, the requirements on R_0 guarantee that the effect of the data points lying on the boundary of R_n is negligible compared to the totality of data values.

The formulation given above allows the sampling region R_n to have a fairly irregular shape. Some common examples of such regions are spheres, ellipsoids, polyhedrons, and star-shaped regions (which can be non-convex sets with irregular boundaries). Sherman and Carlstein (1994) considers a rich subclass of such regions in the plane (i.e., for $d = 2$) where the boundaries of the sets R_0 are delineated by simple rectifiable curves with finite lengths.

Note that in our formulation λ_n acts as a *common* scaling factor in all directions. As a consequence, under the uniform sampling design described above, the number of sampling sites $N(R_n)$, say, in R_n satisfies the growth condition:

$$N(R_n) \sim |R_0| \cdot \lambda_n^d / h_n^d$$

as $n \to \infty$, where $|R_0|$ denotes the volume of the set R_0, and for any two sequences $\{r_n\}$ and $\{t_n\}$ of positive real numbers, we write $r_n \sim t_n$ if $r_n/t_n \to 1$ as $n \to \infty$. Thus, with different choices of the factors λ_n, h_n and the set R_0, the sampling structure adopted here provides a flexible framework for handling varying degrees of "infilling" of "increasing domains" that may have a wide variety of shapes.

Next we state the assumptions used in the paper.

2.2. Assumptions

For stating the assumptions, we need to introduce some notation. For a vector $\mathbf{x} = (x_1, \ldots, x_k)' \in \mathbb{R}^k$ $(k \geq 1)$, let $\|\mathbf{x}\| = \sum_{i=1}^k x_i^2$ and $|\mathbf{x}| = \sum_{i=1}^k |x_i|$ denote the (usual) Euclidean and the ℓ^1 norms of $\mathbf{x}$, respectively. We shall use the notation $|\cdot|$ also in two other cases: for a countable set J, $|J|$ would denote the cardinality of the set J, and for an uncountable set $A \subset \mathbb{R}^k$, $|A|$ would refer to the volume (i.e., the Lebesgue measure) of A. Let $\mathscr{Z}_+$ be the set of all nonnegative integers. If f is a function from $\mathbb{R}^k \to \mathbb{R}$, and $\mathbf{x} \in \mathbb{R}^q$ $(q < k)$, then $f(\mathbf{x}; \cdot)$ denotes the function from $\mathbb{R}^{k-q} \to \mathbb{R}$ that takes the value $f((\mathbf{x}', \mathbf{y}')')$ at $\mathbf{y} \in \mathbb{R}^{k-q}$. Let $D_j f$ denote the partial derivative of f with respect to its jth argument. For $\mathbf{x} = (x_1, \ldots, x_k)' \in \mathbb{R}^k$ and $\alpha = (\alpha_1, \ldots, \alpha_k)' \in \mathbb{R}^k \quad (k \geq 1)$, write $\mathbf{x}^\alpha = \prod_{j=1}^k x_j^{\alpha_j}$, $\alpha! = \prod_{j=1}^k \alpha_j!$, and D^α for the differential operator $D_1^{\alpha_1} \cdots D_k^{\alpha_k}$.

Under the stationarity assumption on the random field $\{Z(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d\}$, let F_0 denote the (common) marginal distribution of $Z(\mathbf{s})$. Next define the functions G and G_1 based on the bivariate joint distribution of $Z(\mathbf{0})$ and $Z(\mathbf{s})$ by

$$G(z_1, z_2; \mathbf{s}) = P(Z(\mathbf{0}) \leq z_1, Z(\mathbf{s}) \leq z_2),$$

$$G_1(z_1, z_2; \mathbf{s}) = P(z_1 < Z(\mathbf{0}) \leq z_2, z_1 < Z(\mathbf{s}) \leq z_2) \ , \tag{2.1}$$

$z_1, z_2 \in \mathbb{R}$ and $\mathbf{s} \in \mathbb{R}^d$.

Let $\mathscr{L}_2'(A)$ be the collection of all random variables with zero mean and finite second moment that are measurable with respect to the σ-field generated by $\{Z(\mathbf{s}) : \mathbf{s} \in A\}$, $A \subset \mathbb{R}^d$. For $A, B \subset \mathbb{R}^d$, write

$$\rho_1(A, B) = \sup\{|E\xi\eta|/(E\xi^2)^{1/2}(E\eta^2)^{1/2} : \xi \in \mathscr{L}_2'(A), \eta \in \mathscr{L}_2'(B)\} \ .$$

Then, define the ρ-mixing co-efficient (cf. Doukhan, 1994) of the r.f. $\{Z(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d\}$ by

$$\rho(k; m) = \sup\{\rho_1(A, B) : |A| \leq m, |B| \leq m, d(A, B) \geq k\} \tag{2.2}$$

where $d(A, B)$ denotes the distance between the sets $A, B \subset \mathbb{R}^d$ in the $|\cdot|$-norm, given by $d(A, B) = \inf\{|x - y| : x \in A, y \in B\}$. Let $\mathbf{c}_o = (\frac{1}{2}, \ldots, \frac{1}{2})'$ denote the midpoint of the unit cube Δ_o. Define

$$\kappa = 2 \quad \text{if} \quad \mathbf{c} \neq \mathbf{c}_o \quad \text{and} \quad \kappa = 4 \quad \text{if} \quad \mathbf{c} = \mathbf{c}_o \ . \tag{2.3}$$

The following conditions will be assumed to be in effect in the rest of the paper.

Assumptions:

(A.1) There exist positive real numbers C, τ, θ satisfying $\tau > 3d$ and $\theta d < \tau$ such that

$$\rho(k; m) \leq C k^{-\tau} m^\theta \ .$$

(A.2) $\{Z(\mathbf{s}) : \mathbf{s} \in \mathbb{R}^d\}$ is stationary and $F_0(\cdot)$ is continuous on $\mathbb{R}$.

(A.3) For each $z_1, z_2 \in \mathbb{R}$, $G(z_1, z_2; \cdot)$ has

(i) bounded and Lebesgue integrable partial derivatives of order κ on $\mathbb{R}^d$, and

(ii) for $|\alpha| = \kappa$, there exist nonnegative integrable functions $H_\alpha(z_1, z_2; \cdot)$ such that for all $\mathbf{s}, \mathbf{t} \in \mathbb{R}^d$, $\|\mathbf{t}\| \leq 1$,

$$|D^\alpha G(z_1, z_2; \mathbf{s}+\mathbf{t}) - D^\alpha G(z_1, z_2; \mathbf{s})| \leq \|\mathbf{t}\|^\eta H_\alpha(z_1, z_2, \mathbf{s})$$

for some $\eta > 0$ (which does not depend on z_1, z_2).

(A.4) There exist constants $C > 0$, $1/2 < \gamma \leq 1$ such that

$$\sum_{|\alpha|=2}^{\kappa} |D^\alpha G_1(z_1, z_2; \mathbf{s})| \leq C|F_0(z_2) - F_0(z_1)|^\gamma$$
$$\text{for all} \quad z_1, z_2 \in \mathbb{R} \quad \text{and} \quad \mathbf{s} \in \mathbb{R}^d \ .$$

(A.5) $(h_n^{(2\gamma+1)\kappa}\lambda_n^d)^{-1} + (h_n\lambda_n/\log\lambda_n)^{-1} \to 0$ as $n \to \infty$, where γ is as in (A.4) and κ is given by (2.3).

Some comments about the assumptions are in order.

It is well known (cf. Bradley, 1989; Doukhan, 1994) that for a r.f. on $\mathbb{R}^d$ with $d > 1$, if the sizes of the sets A and B in the definition of the ρ-mixing co-efficient in (2.2) are unrestricted, a ρ-mixing condition requiring $'\lim_{k\to\infty} \rho(k; \infty) = 0'$ forces the r.f. to be m-dependent. Thus, to ensure validity of our results for a large class of r.f.s, we adopt the standard convention (see, for example, Doukhan, 1994) that for any fixed distance k between two sets A and B of indices, the ρ-mixing co-efficient becomes unbounded as the sizes of these sets tend to infinity.

Assumption (A.5) is a condition on the sampling design parameters λ_n and h_n, which specifies a viable range of "infilling" and "increasing domain" components of the spatial sampling design. Note that the growth condition on both terms within parentheses in (A.5) restricts h_n to take too small values and thus rules out an arbitrary amount of infilling of the sampling region.

Assumptions (A.3) (i) and (ii) are smoothness conditions on the bivariate probability distribution functions $P(Z(\mathbf{0}) \leq z_1, Z(\mathbf{s}) \leq z_2)$, considered as a function of $\mathbf{s} \in \mathbb{R}^d$ with (z_1, z_2) *fixed*. Note that these conditions are almost minimal, since (cf. Theorem 2.1 below) the covariance function of the asymptotic Gaussian process depends on the κ-th order partial derivatives of $G(z_1, z_2; \cdot)$. Assumption (A.3) (ii) can also be viewed as a Lipschitz condition of order $\eta > 0$ for the functions $D^\alpha G(z_1, z_2; \cdot)$, $|\alpha| = \kappa$, in the $L^1(\mathbb{R}^d)$-norm.

Assumption (A.4) is used exclusively in the context of proving tightness of the process ξ_n considered as a random element of the space of functions $D[-\infty, \infty]$. Note that we can express the function $G_1(\cdot, \cdot; \mathbf{s})$ in terms of the function $G(\cdot, \cdot; \mathbf{s})$ as

$$G_1(z_1, z_2; \mathbf{s}) = G(z_2, z_2; \mathbf{s}) - G(z_2, z_1; \mathbf{s}) - G(z_1, z_2; \mathbf{s}) + G(z_1, z_1; \mathbf{s}) \ , \tag{2.4}$$

for $z_1, z_2 \in \mathbb{R}$, and $\mathbf{s} \in \mathbb{R}^d$. Write $\tilde{G}(t_1, t_2; \mathbf{s}) = G(F_0^{-1}(t_1), F_0^{-1}(t_2); \mathbf{s})$ for $t_1, t_2 \in (0, 1)$, where $F_0^{-1}(\cdot)$ denotes the inverse of the CDF $F_0(\cdot)$, defined by $F_0^{-1}(t) = \inf\{z : F_0(z) \geq t\}, \quad t \in (0, 1)$. Then, it is easy to check that (A.4) holds if the functions $D^\alpha \tilde{G}(t_1, t_2; \mathbf{s}), \ |\alpha| = 2, \ldots, \kappa$ satisfy a Lipschitz condition of order γ in each of t_1 and t_2, for all $\mathbf{s} \in \mathbb{R}^d$.

For an example of a r.f., where the results of the paper are applicable, consider the random process

$$Z(\mathbf{s}) = g(\epsilon(\mathbf{s})), \quad \mathbf{s} \in \mathbb{R}^d \ ,$$

where $g : \mathbb{R} \to \mathbb{R}$ is a Borel measurable function and $\{\epsilon(\mathbf{s}) : \ \mathbf{s} \in \mathbb{R}^d\}$ is a stationary Gaussian r.f. with $E\epsilon(\mathbf{s}) = 0$, $E(\epsilon(\mathbf{s}))^2 = 1$, and autocorrelation function $\rho(\cdot)$. Then, the function $G(\cdot, \cdot; \cdot)$ can be written as

$$G(z_1, z_2; \mathbf{s}) = \int_{g^{-1}(-\infty, z_1]} \left[\int_{g^{-1}(-\infty, z_2]} (1 - \rho(\mathbf{s})^2)^{-1/2} \phi \left(\frac{v - \rho(\mathbf{s})u}{\sqrt{1 - \rho(\mathbf{s})^2}} \right) dv \right] d\Phi(u) \, , \tag{2.5}$$

$z_1, z_2 \in \mathbb{R}$, $\mathbf{s} \in \mathbb{R}^d$, where ϕ and Φ respectively denote the density and the distribution function of a $N(0, 1)$ random variable. Next, assume that $\lim_{\|\mathbf{s}\| \to 0} |\rho(\mathbf{s})| < 1$ and that $\rho(\mathbf{s})$ has bounded, Lebesgue integrable partial derivatives of order $\kappa + 1$ on $\mathbb{R}^d \backslash \{\mathbf{0}\}$. Then, by straight-forward algebra, it follows that the inequality in Assumption (A.3) (ii) holds with $\eta = 1$ for any $z_1, z_2 \in \mathbb{R}$ and for $\mathbf{s}, \mathbf{s} + \mathbf{t} \in \mathbb{R}^d \backslash \{\mathbf{0}\}$, $\|\mathbf{t}\| \leq 1$. At the points $\mathbf{s}, \mathbf{s} + \mathbf{t} \in \{\mathbf{0}\}$, the function $G(z_1, z_2; \cdot)$ is *not* necessarily differentiable and hence, the inequality in (A.3) (ii) can not be verified. However, the effect of the violation of (A.3) (ii) at these points can be shown to be negligible under (A.5), and hence, all the steps in the proofs of the results of the paper go through under this weaker form of (A.3) (ii).

As for Assumption (A.4), in addition to the Conditions on $\rho(\cdot)$, suppose further that the transformation g is piece-wise strictly monotone with a piece-wise continuous derivative on $\mathbb{R}$. Then, the marginal distribution F_0 of the process $Z(\cdot)$ has a density. Using (2.4) and (2.5), and computing the κ-th order partial derivatives of the function $G(z_1, z_2; \mathbf{s})$ at $\mathbf{s} \neq \mathbf{0}$, one can readily establish the inequality in (A.4) with $\gamma = 1$, for all $z_1, z_2 \in \mathbb{R}$ and for all $\mathbf{s} \in \mathbb{R}^d \backslash \{\mathbf{0}\}$. For $\mathbf{s} = \mathbf{0}$, a comment similar to (A.3) (ii) applies. In summary, the results of the paper hold for the class of r.f.s $Z(\cdot)$ that are instantaneous functions of a stationary Gaussian r.f., provided Assumption (A.1), (A.5), and the conditions on $\rho(\cdot)$ and $g(\cdot)$ specified above are satisfied.

2.3. *Asymptotic Distribution*

The main result of this section is a Functional Central Limit Theorem for the normalized CDF predictor:

$$\xi_n(z) \equiv b_n(\hat{F}_n(z; R_n) - \hat{F}_\infty(z; R_n)), \quad z \in \mathbb{R} \ . \tag{2.6}$$

Note that for each n,

$$P\left(\lim_{|z|\to\infty} \xi_n(z) = 0\right) = 1 \ .$$

Hence, $\xi_n(\cdot)$ may be considered as a random element of the space $D[-\infty, \infty]$ of all real valued functions on $[-\infty, \infty]$ that are right continuous with left hand limits. We equip $D[-\infty, \infty]$ with the Skorohod metric and show that under the assumptions of Section 2.2, $\xi_n(\cdot)$ converges weakly to a continuous Gaussian process $W(\cdot)$ on $[-\infty, \infty]$ with $W(\infty) = 0 = W(-\infty)$ a.s.

The normalizing constant b_n and the covariance function of the limiting process $W(\cdot)$ depend on the choice of the vector $\mathbf{c} \in \Delta_0$. Recall that $\mathbf{c}$ determines the starting point of the sampling grid of our uniform sampling design. It turns out that when $\mathbf{c} \neq \mathbf{c}_o$, the right choice of b_n is given by $\lambda_n^{d/2} h_n^{-1}$ while for $\mathbf{c} = \mathbf{c}_o$, b_n should be taken as $\lambda_n^{d/2} h_n^{-2}$. The covariance functions of $W(\cdot)$ in these cases are, respectively, given by certain integrals of the second and the fourth order partial derivatives of the function $G(\cdot, \cdot; \mathbf{s})$, defined in (2.1). To state the result formally, define

$$a(\alpha) = \int_{\Delta_0}\int_{\Delta_0} \{(\mathbf{x}-\mathbf{s})^\alpha - (\mathbf{x}-\mathbf{c})^\alpha - (\mathbf{c}-\mathbf{s})^\alpha\}\, \mathbf{dx}\, \mathbf{ds}, \quad \alpha \in (\mathscr{Z}_+)^d \ .$$

Then we have the following theorem.

Theorem 2.1. *Suppose that Assumptions* (A.1)–(A.5) *hold. Let* $b_n = \lambda_n^{d/2} h_n^{-\kappa/2}$, *where* κ *is given by* (2.3). *Then,*

$$\xi_n(\cdot) \to^d W(\cdot)$$

where $\to^d$ *denotes weak convergence and where* $W(\cdot)$ *is a zero mean Gaussian process with the covariance function*

$$\sigma(z_1, z_2) = |R_0|^{-1} \sum_{|\alpha|=\kappa} a(\alpha)(\alpha!)^{-1} \int_{\mathbb{R}^d} D^\alpha G(z_1, z_2; \mathbf{s})\, \mathbf{ds} \ , \qquad (2.7)$$

$z_1, z_2 \in \mathbb{R}$. *Moreover,* $W(+\infty) = W(-\infty) = 0$ *a.s. and* $W(\cdot)$ *has continuous sample paths with probability one.*

Thus, Theorem 2.1 shows that with the proper choice of b_n, the centered spatial CDF predictor $\hat{F}_n$ does indeed have a nondegenerate limit distribution on $D[-\infty, \infty]$. For sums of random variables from r.f.s on the integer lattice $\mathscr{Z}^d$, Central Limit Theorems (CLTs) have been proved by Bolthausen (1982), Bulinskii and Zhurbenko (1976), and Guyon and Richardson (1984) under different sets of moment and mixing conditions. For r.f.s with a continuous spatial index, Ivanov and Leonenko (1989) obtains a CLT for certain weighted integrals of the field, assuming a strong-mixing condition. Theorem 2.1 takes a step towards proving functional CLTs for triangular arrays of $D[-\infty, \infty]$-valued random elements generated by r.f.s. One of the main technical problems that arise in proving Theorem 2.1 is to establish tightness of the process ξ_n in $D[-\infty, \infty]$. In Section 4, we obtain some auxiliary results in this context, which can be useful also for establishing similar functional and finite dimensional limit theorems for estimators and predictors based on finite samples from continuous parameter r.f.s.

Although Theorem 2.1 provides a very precise description of the right normalizing constant b_n, in practice, however, it may not always be clear as to which of the two different rates should be used. This is because in an application, given the locations of the sampling sites $\{\mathbf{s}_1, \ldots, \mathbf{s}_{N_n}\}$ on a regular grid and the sampling region R_n, there may not be a unique reference point that can be identified as the origin, and hence, one may not be able to discriminate between the cases $\mathbf{c} = \mathbf{c}_o$ and $\mathbf{c} \neq \mathbf{c}_o$ from this information. An exceptional situation, where one can discriminate between these two possibilities solely from the knowledge of the sampling sites and the sampling region R_n, occurs when the sampling region can be *tessellated* by cubes of side h_n with centers either at the given sampling sites (leading to the case $\mathbf{c} = \mathbf{c}_o$) or at points other than the sampling sites (corresponding to $\mathbf{c} \neq \mathbf{c}_o$). In the following, we exploit this observation to formulate a variant of Theorem 2.1 that can be used in practice for sampling regions that are not tessellated exactly.

Suppose that $\{\mathbf{s}_1, \ldots, \mathbf{s}_{N_n}\}$ are the sampling sites within the sampling region R_n, coming from a regular grid $\mathscr{P}_n$ for *some* $\mathbf{c} \in \Delta_0$. For $\mathbf{s} \in \mathbb{R}^d$ and $\mathbf{c}_1 \in \Delta_0$, let $\Gamma_n(\mathbf{s}; \mathbf{c}_1)$ denote the cube with sides of length h_n and with center at $\mathbf{s} + (\mathbf{c}_1 - \mathbf{c}_o)h_n$. Then, it follows that $\Gamma_n(\mathbf{s}_i; \mathbf{c}_1)$ is a cube of volume h_n^d containing the sampling site $\mathbf{s}_i$, and has its center at $\mathbf{s}_i$ if and only if $\mathbf{c}_1 = \mathbf{c}_o$. Next define the region $R_n(\mathbf{c}_1) = \cup_{i=1}^{N_n} \Gamma_n(\mathbf{s}_i; \mathbf{c}_1)$. (cf. Figure 1 below).

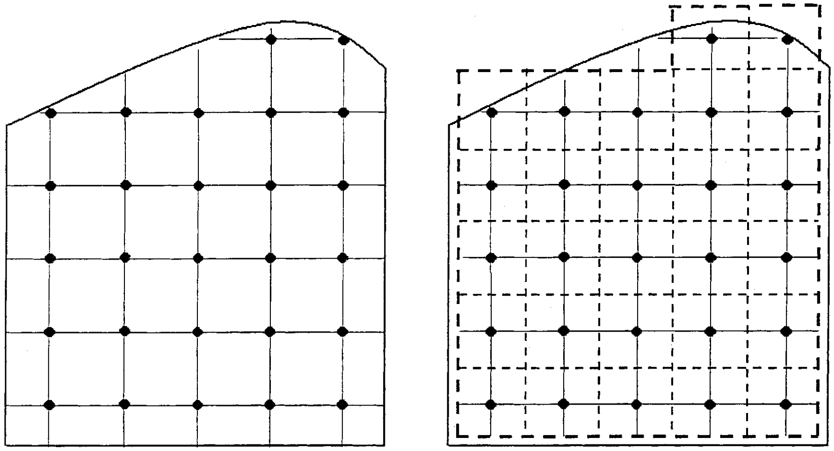

Fig. 1. (Left) Sampling region R_n with the sampling sites $\mathbf{s}_i$s denoted by solid circles; (Right) The augmented region $R_n(\mathbf{c}_o)$ with its boundary shown in boldface dashed line, and the squares $\Gamma_n(\mathbf{s}_i; \mathbf{c}_o)$ centered at the sampling sites $\mathbf{s}_i$ with their boundaries in dashed lines.

Since the sampling sites $\{\mathbf{s}_1, \ldots, \mathbf{s}_{N_n}\}$ lie on the regular grid $\mathscr{P}_n$ which has an increment h_n in *each* direction, the cubes $\{\Gamma_n(\mathbf{s}_i; \mathbf{c}_1),\ i = 1, \ldots, N_n\}$ *tessellate* $R_n(\mathbf{c}_1)$. Furthermore, since it is obtained from R_n by an augmentation of the cubes only at the *boundary* sampling sites, the region $R_n(\mathbf{c}_1)$ differs from the original sampling region R_n by a volume that is negligible compared to the total volumes of both R_n and $R_n(\mathbf{c}_1)$. As a result, one can use the empirical predictor $\hat{F}_n(\cdot; R_n)$ based on observations on the process $Z(\cdot)$ at the sampling sites $\{\mathbf{s}_i : i = 1, \ldots, N_n\} \subset R_n(\mathbf{c}_1)$ also as a predictor of the spatial CDF $\hat{F}_\infty(\cdot; R_n(\mathbf{c}_1))$ over the augmented region $R_n(\mathbf{c}_1)$. Then, the following version of Theorem 2.1 holds.

Theorem 2.1′. *Assume that assumptions* (A.1)–(A.5) *hold and that* $b_{1n} = \lambda_n^{d/2} h_n^{-2}$ *if* $\mathbf{c}_1 = \mathbf{c}_o$, *and* $b_{1n} = \lambda_n^{d/2} h_n^{-1}$ *if* $\mathbf{c}_1 \neq \mathbf{c}_o$. *Then,*

$$b_{1n}(\hat{F}_n(\cdot; R_n) - \hat{F}_\infty(\cdot; R_n(\mathbf{c}_1))) \to^d W(\cdot; \mathbf{c}_1)$$

where $W(\cdot; \mathbf{c}_1)$ *is a zero mean Gaussian process having the covariance function* (2.7) *with* $\kappa = 4$ *if* $\mathbf{c}_1 = \mathbf{c}_o$, *and* $\kappa = 2$ *if* $\mathbf{c}_1 \neq \mathbf{c}_o$. *Moreover,* $W(+\infty; \mathbf{c}_1) = W(-\infty; \mathbf{c}_1) = 0$ *a.s. and* $W(\cdot; \mathbf{c}_1)$ *has continuous sample paths with probability one.*

Thus, given the sampling sites $\{\mathbf{s}_1, \ldots, \mathbf{s}_{N_n}\}$ from a regular grid, one can *choose* a desired value for $\mathbf{c}_1$ and apply Theorem 2.1′ with the *corresponding* scaling constant b_{1n}. Hence, unlike the choice of b_n, one can *always* determine the right scaling constant b_{1n} for predicting the spatial CDF $\hat{F}_\infty(\cdot; R_n(\mathbf{c}_1))$ over the region $R_n(\mathbf{c}_1)$.

Secondly, Theorem 2.1′ also shows that under Assumptions (A.1)–(A.5), the empirical predictor $\hat{F}_n$ achieves the higher level of accuracy for predicting the spatial CDF $\hat{F}_\infty(\cdot;\, R_n(\mathbf{c}_o))$ over the region $R_n(\mathbf{c}_o)$, *irrespective* of its rate of convergence to the spatial CDF $\hat{F}_\infty(\cdot;\, R_n)$ over the original region R_n. Consequently, inference procedures based on the empirical predictor $\hat{F}_n$ would be most accurate for the region $R_n(\mathbf{c}_1)$ if one chooses $\mathbf{c}_1 = \mathbf{c}_o$. In particular, prediction bands based on $\hat{F}_n$ would be narrower for $\hat{F}_\infty(\cdot;\, R_n(\mathbf{c}_o))$ than for $\hat{F}_\infty(\cdot;\, R_n(\mathbf{c}_1))$ with $\mathbf{c}_1 \neq \mathbf{c}_o$.

Finally, it should be noted that for a sampling region R_n that is *not* tessellated by the cubes from the grid $\mathscr{P}_n$, Theorem 2.1′ yields a large sample result for the spatial CDFs only over the regions $R_n(\mathbf{c}_1), \mathbf{c}_1 \in \Delta_0$, which are *different* from the original sampling region R_n. Though the volume of the symmetric difference $R_n \Delta R_n(\mathbf{c}_1)$ is negligible compared to the total volumes of R_n and $R_n(\mathbf{c}_1)$ in $\mathbb{R}^d$, in general, we can not replace $\hat{F}_\infty(\cdot;\, R_n(\mathbf{c}_1))$ by $\hat{F}_\infty(\cdot;\, R_n)$ in Theorem 2.1′. This is because the *scaled* difference $b_{1n}(\hat{F}_\infty(\cdot;\, R_n(\mathbf{c}_1)) - \hat{F}_\infty(\cdot;\, R_n))$ is not necessarily negligible when the dimension d of the sampling region R_n is 2 or more.

We conclude this section with a result that will be used in the context of obtaining large sample prediction bands for $\hat{F}_\infty$ in Section 3. Let $w(\cdot)$ be a nonnegative integrable function on $\mathbb{R}$. Define the weighted L^p-norm of an element $x \in D[-\infty, \infty]$ by

$$\|x\|_p = \left(\int_{-\infty}^{\infty} |x(z)|^p w(z)\, dz \right)^{1/p}$$

for $p \in [1, \infty)$. Also, corresponding $p = \infty$, let

$$\|x\|_\infty = \sup\{|x(z)| : z \in [-\infty, \infty]\} \ .$$

Then, the following result holds.

Theorem 2.2. *Assume that the conditions of Theorem* 2.1 *hold. Then*, *as* $n \to \infty$,

$$\|\xi_n\|_p \to^d \|W\|_p$$

and

$$\|b_{1n}(\hat{F}_n(\cdot;\, R_n) - \hat{F}_\infty(\cdot;\, R_n(\mathbf{c}_1)))\|_p \to^d \|W(\cdot;\, \mathbf{c}_1)\|_p \ ,$$

for all $p \in [1, \infty]$.

3. Subsampling prediction band

In this section we consider the problem of constructing valid large sample prediction bands for the underlying spatial CDF $\hat{F}_\infty$. For simplicity of exposition, we restrict attention to the spatial CDF $\hat{F}_\infty(\cdot;\, R_n)$ corresponding to the sampling region R_n. Prediction bands for $\hat{F}_\infty(\cdot; R_n(\mathbf{c}_1))$ can be constructed by straightforward modifications (cf. Remark 3.1 below). Hence, unless explicitly mentioned, we write $\hat{F}_\infty$ to denote the spatial CDF $\hat{F}_\infty(\cdot; R_n)$ over R_n.

For $0 < \alpha < 1$, let q_α denote the α quantile of $\|W\|_p$. Then, by Theorem 2.2 it follows that

$$I_\alpha = \{F : b_n \|\hat{F}_n - F\|_p < q_\alpha\} \tag{3.1}$$

is a prediction band for $\hat{F}_\infty$ that attains the nominal coverage level α asymptotically. However, the difficulty with using (3.1) is that the quantity q_α depends on the bivariate population CDF G, and hence is unknown in practice. In principle, it is possible to estimate the covariance function of the process W and use orthogonal decomposition (cf. Anderson, 1993) to obtain an estimator of q_α. However, given the complex structure of the covariance function, this approach does not seem very convenient for practical applications. Instead, here we propose an extension of the standard subsampling method, used in the context of purely "increasing domain asymptotics" for time-series and lattice processes (cf. Possolo, 1991; Politis and Romano, 1994; Hall and Jing, 1996; and Sherman and Carlstein, 1994), to allow "infill sampling" of continuous parameter r.f.s and apply it to construct valid large sample prediction sets for $\hat{F}_\infty$.

The main idea behind the proposed subsampling method is to use several smaller regions within R_n of *similar* shape (cf. Sherman and Carlstein, 1994) and exploit the "infill" component of the sampling design to recreate the effect of "sample" and "population" at the level of the subsamples. The modification is required for defining the subsample versions of the sample based predictor $\hat{F}_n$ and the unobservable random predictand $\hat{F}_\infty$ that depends on the "population" of the entire r.f. $Z(z)$ over R_n.

First we define the subregions. Let $l (\equiv l_n)$ be an integer such that $l/n \to 0$ as $n \to \infty$. Here l determines the "window width" for the subsamples. Let $K_1 \equiv K_{1n}$ denote the largest integer not exceeding λ_n/λ_l. Then, $(-\lambda_n, \lambda_n]^d$ contains $(2K_1)^d$ smaller cubes of the form $(\mathbf{i} + \Delta_0)\lambda_l$ where $\mathbf{i} \in \mathscr{Z}^d$. Let $S_1, \ldots, S_K$ denote the set of all such subcubes that lie *inside* the given region R_n. Then, define the subsampling regions

$$R_{*1}, \ldots, R_{*K}$$

by inscribing the translate of the region $\lambda_l R_0$ inside each one of the subregions $S_1, \ldots, S_K$ such that the origin is mapped onto the midpoint of the given subcube. (cf. Figure 2 below). Then, this gives us a collection

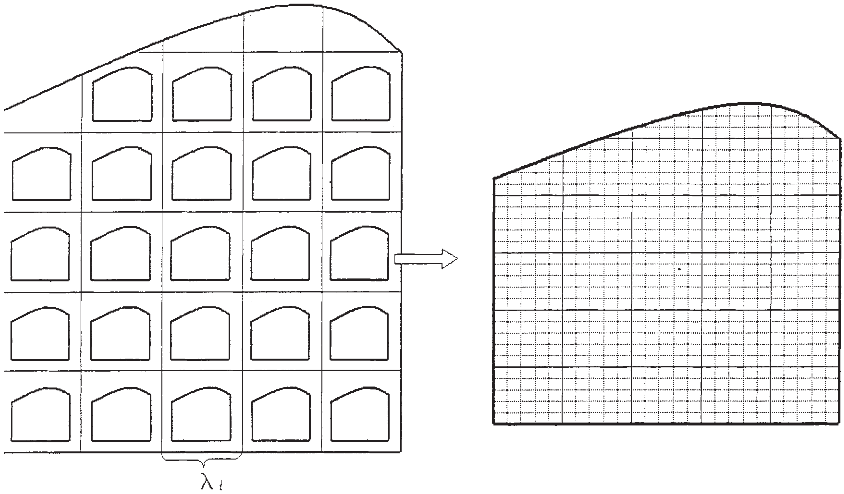

Fig. 2. (Left) The sampling region R_n with the subregions R_{*i}s; (Right) A given subregion R_{*i} (magnified) with the finer and the coarser partitions, given respectively by the dotted lines and the solid lines.

of nonoverlapping subregions that are of the same shape as the original sampling region R_n, and are contained in R_n.

Next we define a copy of $\hat{F}_n$ and $\hat{F}_\infty$ on each R_{*i}. Note that corresponding to l and n, $\mathscr{P}_l$ and $\mathscr{P}_n$ induce partitions of R_{*i}s at two levels of resolution. We use the partition $\mathscr{P}_l$ to define the subsample version of $\hat{F}_n$ and $\mathscr{P}_n$ to define that of $\hat{F}_\infty$. Let $\{\mathbf{s_j} : \mathbf{j} \in N_i\}$ denote the ($\mathscr{P}_n$-level) sampling sites in R_{*i}, $i = 1, \ldots, K$. Next for simplicity, assume that l is such that h_l/h_n is an integer. Then, the sampling sites corresponding to $\mathscr{P}_l$ and $\mathscr{P}_n$ are nested. Let $\{\mathbf{s_j} : \mathbf{j} \in L_i\}$ denote the collection of $\mathscr{P}_l$-level sampling sites in R_{*i}, $i = 1, \ldots, K$. Then, define the versions of $\hat{F}_n$ and $\hat{F}_\infty$ on R_{*i} respectively by

$$F_n^{*i}(z) = |L_i|^{-1} \sum_{\mathbf{j} \in L_i} I(Z(\mathbf{s_j}) \leq z)$$

and

$$F_\infty^{*i}(z) = |N_i|^{-1} \sum_{\mathbf{j} \in N_i} I(Z(\mathbf{s_j}) \leq z) \ ,$$

$z \in \mathbb{R}$, $i = 1, \ldots, K$. Thus, F_n^{*i} plays the role of $\hat{F}_n(\cdot; R_n)$ and F_∞^{*i} plays the role of $\hat{F}_\infty(\cdot; R_n)$ at the level of the subregion R_{*i}. The version of the process ξ_n on R_{*i} is given by

$$\xi_n^{*i}(z) = b_l(F_n^{*i}(z) - F_\infty^{*i}(z)), \quad z \in \mathbb{R} \ .$$

Since the subregions R_{*i}s are nonoverlapping, $\xi_n^{*1}, \ldots, \xi_n^{*K}$ behave like "approximately independent" copies of the process ξ_n. Hence, we define the subsampling estimator of the CDF $H_n(\cdot;\, p)$, say, of $\|\xi_n\|_p$ by

$$\hat{H}_n(z;\, p) = K^{-1} \sum_{i=1}^{K} 1(\|\xi_n^{*i}\|_p \le z), \quad z \in \mathbb{R} \ .$$

The following theorem justifies the use of the proposed subsampling method for constructing prediction bands for $\hat{F}_\infty$.

Theorem 3.1. *Suppose that conditions of Theorem* 2.1 *hold and that the weight function* w *is such that* $\|W\|_p$ *has a continuous distribution on* $\mathbb{R}$. *Also, let* l *be such that* $h_n/h_l \to 0$ *as* $n \to \infty$, *and Assumption* (A.5) *holds with* λ_n *replaced by* λ_l. *Then, for any* $p \in [1, \infty]$,

$$\sup_{z \in \mathbb{R}} |\hat{H}_n(z;\, p) - H_n(z;\, p)| \to 0 \quad \textit{in probability,} \quad \textit{as} \quad n \to \infty \ .$$

Thus, under the conditions of Theorem 3.1, the subsampling estimator of the sampling distribution of $\|\xi_n\|_p$ provides a valid approximation. To construct a $100\alpha\%$ $(0 < \alpha < 1)$ prediction region for $\hat{F}_\infty$ based on the subsampling method, let $\hat{q}_\alpha$ be the α quantile of $\hat{H}_n(\cdot;\, p)$. Define

$$I_\alpha^S = \{F : b_n \|\hat{F}_n(\cdot;\, R_n) - F\|_p < \hat{q}_\alpha\} \ .$$

Then, I_α^S gives the desired subsampling prediction band for the spatial CDF $\hat{F}_\infty$. Note that by Theorem 3.1,

$$P(\hat{F}_\infty \in I_\alpha^S) \to \alpha \quad \text{as} \quad n \to \infty \ .$$

Hence, the prediction region I_α^S attains the target coverage probability α asymptotically. In practice, implementation of the subsampling procedure is quite simple. To find the quantile $\hat{q}_\alpha$, note it is given by the $(K\alpha)$-th order statistic of the values $\|\xi_n^{*i}\|_p$, $i = 1, \ldots, K$ and hence, can be easily found by arranging the $\|\xi_n^{*i}\|_p$'s in an increasing order. Finite sample performance of the method and the choice of the subsampling parameters can be found in Kaiser, Hsu, Cressie, and Lahiri (1997). Also, for an application of the subsampling method to a real data set and for overlapping versions of the subsampling method, see Lahiri, Kaiser, Cressie, and Hsu (1999).

Remark 3.1. A variant of Theorem 3.1 also holds for the process $\xi_{1n} \equiv b_{1n}(\hat{F}_n(\cdot;\, R_n) - \hat{F}_\infty(\cdot;\, R_n(\mathbf{c}_1)))$ considered in Theorem 2.1′, provided the subsampling estimator is redefined appropriately. The main modification is that on each subregion R_{*i}, a copy the process ξ_{1n} is now defined as

$$\xi_{1n}^{*i} \equiv b_{1l}(F_n^{*i}(\cdot) - F_\infty^{*i}(\cdot;\, R_{*i}(\mathbf{c}_1)))$$

where $F_\infty^{*i}(\cdot;\, R_{*i}(\mathbf{c}_1)))$ is the empirical distribution function of the observations located on the *finer* grid $\mathscr{P}_n$ in the *augmented* subregion $R_{*i}(\mathbf{c}_1) \equiv \cup_{\mathbf{s}\in R_{*i}\cap \mathscr{P}_l} \Gamma_l(\mathbf{s};\, \mathbf{c}_1)$. Thus, to define the subsample "copy" of ξ_{1n} on R_{*i}, we replace $\hat{F}_\infty(\cdot;\, R_n(\mathbf{c}_1))$ by $F_\infty^{*i}(\cdot;\, R_{*i}(\mathbf{c}_1))$, which is the right subsample version of the spatial CDF $\hat{F}_\infty(\cdot;\, R_n(\mathbf{c}_1))$. The subsampling estimator of the sampling distribution $H_{1n}(\cdot;\, p)$, say, of $\|\xi_{1n}\|_p$ is then given by

$$K^{-1} \sum_{i=1}^{K} 1(\|\xi_{1n}^{*i}\|_p \leq \cdot) \enspace ,$$

which provides a valid approximation to $H_{1n}(\cdot;\, p)$ under the conditions of Theorem 3.1. Hence, one can construct valid prediction bands for the spatial CDF $\hat{F}_\infty(\cdot;\, R_n(\mathbf{c}_1))$ as above.

Recently, Bertail, Politis, and Romano (1995) and Sherman and Carlstein (1997) have developed methods for constructing confidence intervals for population parameters of time series data, when the rates of convergence of the corresponding estimators are unknown. Hence, as an alternative, one may *adapt* these methods to the present problem to construct prediction bands for $\hat{F}_\infty(\cdot;\, R_n)$ itself, if the choice of the scaling constant b_n is not obvious.

4. Proofs

For proving the theorems, we need a few lemmas. We will use C, $C(\cdot)$ to denote generic positive constants that depend on their arguments (if any). Also, unless otherwise specified, limits in order symbols are taken letting n tend to infinity.

Lemma 4.1. *Let $B_n = t_n B_0$ and $\tilde{Z}(\mathbf{i}) = \int f_{\mathbf{i}}(Z(\mathbf{s}))1(\mathbf{s} \in (\mathbf{i}+\Delta_0) \cap B_n)\, \mathbf{ds}$, $\mathbf{i} \in J_n \equiv \{\mathbf{i} \in \mathscr{Z}^d : \mathbf{i} + \Delta_0 \cap B_n \neq \phi\}$ be random variables satisfying*

$$E\tilde{Z}(\mathbf{i}) = 0, \quad |\tilde{Z}(\mathbf{i})| \leq 1 \quad \textit{and}$$

$$E|\tilde{Z}(\mathbf{i})|^2 \leq \delta_n \quad \textit{for all} \quad \mathbf{i} \in J_n \enspace ,$$

where $t_n \to \infty$, B_0 is a Borel subset of $(-1/2,\, 1/2)^d$ that satisfies the same boundary condition as the set R_0, and $f_{\mathbf{i}} : \mathbb{R} \to \mathbb{R}$ is a Borel measurable function such that $Ef_{\mathbf{i}}(Z(\mathbf{0}))^4 < \infty, \mathbf{i} \in J_n$. Then, under assumption (A.1),

$$E\left(\sum_{\mathbf{i}\in J_n} \tilde{Z}(\mathbf{i})\right)^4 \leq C(d, \rho(\cdot))\left[t_n^{2d}\delta_n^2 + t_n^d\delta_n\right] \enspace .$$

Proof . Clearly,

$$
\begin{aligned}
&E\left(\sum_{\mathbf{i}\in J_n}\tilde{Z}(\mathbf{i})\right)^4 \\
&\quad \le C(d)\Bigg\{\sum_{\mathbf{i}\in J_n} E\tilde{Z}(\mathbf{i})^4+\sum_{\mathbf{i}_1\neq\mathbf{i}_2}|E\tilde{Z}(\mathbf{i}_1)^3\tilde{Z}(\mathbf{i}_2)|+\sum_{\mathbf{i}_1\neq\mathbf{i}_2}E\tilde{Z}(\mathbf{i}_1)^2\tilde{Z}(\mathbf{i}_2)^2 \\
&\quad +\sum_{\mathbf{i}_1\neq\mathbf{i}_2\neq\mathbf{i}_3}|E\tilde{Z}(\mathbf{i}_1)^2\tilde{Z}(\mathbf{i}_2)\tilde{Z}(\mathbf{i}_3)|+\sum_{\mathbf{i}_1\neq\mathbf{i}_2\neq\mathbf{i}_3\neq\mathbf{i}_4}|E\tilde{Z}(\mathbf{i}_1)\tilde{Z}(\mathbf{i}_2)\tilde{Z}(\mathbf{i}_3)\tilde{Z}(\mathbf{i}_4)|\Bigg\} \\
&\equiv I_1+I_2+I_3+I_4+I_5,\ \text{say}\ .
\end{aligned}
$$

By the ρ-mixing condition,

$$
\begin{aligned}
&I_1+I_2+I_3 \\
&\quad \le C(d)\left[\left(1+\sum_{k=1}^{\infty}k^{d-1}\rho(k-1;1)\right)|J_n|\delta_n+\sum_{\mathbf{i}_1\neq\mathbf{i}_2}E\tilde{Z}(\mathbf{i}_1)^2E\tilde{Z}(\mathbf{i}_2)^2\right] \\
&\quad \le C(d,\rho)\left[t_n^d\delta_n+t_n^{2d}\delta_n^2\right]\ .
\end{aligned}
$$

Next we obtain bounds on I_4 and I_5. In both cases, the key step in applying the mixing condition involves counting the number of different indices (viz., $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ for I_4 and $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{i}_4$ for I_5) that correspond to a given maximal gap. We consider the case I_5 first. For $\mathbf{i}_1 \neq \mathbf{i}_2 \neq \mathbf{i}_3 \neq \mathbf{i}_4$, write $J = \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{i}_4\}$. Next define

$$
d_j(J)=\max\{d(I, J\backslash I): I\subset J, |I|=j\},\quad j=1,2\ .
$$

Thus, $d_2(J)$ denotes the maximal distance between any two (out of six possible) pairs of indices in J, while $d_1(J)$ gives the maximum distance of a single index in J from the rest of the indices. Depending on the elements in J, $d_1(J)$ can be larger or smaller than $d_2(J)$. We claim that for given integers $1 \le d_{01}, d_{02} \le |J_n|$,

$$
|\{J: d_1(J)=d_{01}\quad\text{and}\quad d_2(J)=d_{02}\}|\le C(d)[(d_{01}+d_{02})]^{3d-1}|J_n|\ . \tag{4.1}
$$

To see this, fix any $\mathbf{i}_1 \in J_n$, say, $\mathbf{i}_1 = \mathbf{0}$ (for notational convenience), and suppose that for $J = \{\mathbf{0}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$, $d_r(J) = d_{0r}$, $r = 1, 2$. Then, there exists a set $I \subset J$, $|I| = 2$ such that $d_{02} = d(I, J \backslash I)$. Without loss of generality (w.l.g.), assume that $I = \{\mathbf{0}, \mathbf{j}\}$, $I^c = \{\mathbf{i}, \mathbf{k}\}$, and $d(I, I^c) = |\mathbf{i}|$. This implies $|\mathbf{i}| = d_{02}$, $|\mathbf{k}| \geq d_{02}$, $|\mathbf{j} - \mathbf{i}| \geq d_{02}$. By virtue of its definition, a bound on $d_2(J)$ restricts the possible choices of at least one of $\mathbf{j}$ and $\mathbf{k}$. Indeed,

$$\text{if} \quad d_{02} = d(I, I^c), \ \text{ then either } |\mathbf{j}| \leq 2d_{02} \ \text{ or } |\mathbf{k}| \leq 2d_{02} \ . \tag{4.2}$$

To prove this, note that if $|\mathbf{j}| \leq 2d_{02}$, then (4.2) is trivially true. Hence, suppose that $|\mathbf{j}| > 2d_{02}$. Then, it is easy to check that if $|\mathbf{k}|$ is also bigger than $2d_{02}$, then for the set $I_1 = \{\mathbf{0}, \mathbf{i}\}$ $\quad d(I_1, J \backslash I_1) > d_{02}$, contradicting the maximality of d_{02}. Thus, $d_{02} = d(I, I^c) = |\mathbf{i}|$ implies at least two of the indices $\mathbf{i}, \mathbf{j}, \mathbf{k}$ do not exceed $2d_{02}$.

In general, for a given value d_{02} alone, there can be as many as $O(t_n^d)$ possible choices for the third index. (For example, take $\mathbf{k} = 2\mathbf{i} = (2d_{02}, 0, \ldots, 0)'$, and let all components of $\mathbf{j}$ be less than $-2d_{02}$.) But if the value of $d_1(J)$ is also specified, then the norm of the otherwise uncontrollable third index admits a bound in terms of $d_1(J)$ and $d_2(J)$. Indeed, in the case when $|\mathbf{j}| > 2d_{02}$ and $|\mathbf{k}| \leq 2d_{02}$ in (4.2), we must have $|\mathbf{j}| \leq 2d_{02} + d_{01}$. This follows from the inequalities: $d_{01} \geq d(\{\mathbf{j}\}, \{\mathbf{0}, \mathbf{i}, \mathbf{k}\}) = \min\{|\mathbf{j}|, |\mathbf{i} - \mathbf{j}|, |\mathbf{j} - \mathbf{k}|\}$, $|\mathbf{j}| \leq |\mathbf{i} - \mathbf{j}| + d_{02}$ and $|\mathbf{j}| \leq |\mathbf{k} - \mathbf{j}| + 2d_{02}$. In the other case, i.e. when $|\mathbf{j}| \leq 2d_{02}$, we get a similar bound for $\mathbf{k}$. Since $|\{\mathbf{i} \in \mathscr{Z}^d : |\mathbf{i}| \leq a\}| = O(a^d)$ and $|\{\mathbf{i} \in \mathscr{Z}^d : |\mathbf{i}| = a\}| = O(a^{d-1})$ as $a \to \infty$, this proves (4.1).

Now, writing $\sum^{(1)}$ ($\sum^{(2)}$) for the summation over all $J = \{\mathbf{i}_1 \neq \mathbf{i}_2 \neq \mathbf{i}_3 \neq \mathbf{i}_4\}$ with $d_2(J) \geq d_1(J)$ ($d_2(J) < d_1(J)$, respectively), we get

$$\begin{aligned}
I_5 &\leq C(d) \sum^{(1)} \{|E\tilde{Z}(I)E\tilde{Z}(I^c)| + \rho(d_2(J) - 1; 2)\delta_n\} \\
&\quad + C(d) \sum^{(2)} \rho(d_1(J) - 1; 3)\delta_n \\
&\leq C(d) \left\{ |J_n|^2 \delta_n^2 \left(1 + \sum_{k=1}^{\infty} \rho(k-1; 1)\right)^2 + |J_n|\delta_n \sum_{k=1}^{\infty} k^{3d-1} \rho(k-1; 2) \right\} \\
&\quad + C(d)|J_n|\delta_n \sum_{k=1}^{\infty} k^{3d-1} \rho(k-1; 3) \\
&\leq C(d, \rho(\cdot)) \left\{ t_n^{2d} \delta_n^2 + t_n^d \delta_n \right\} \ ,
\end{aligned}$$

where for any index-set A, $\tilde{Z}(A) = \prod_{\mathbf{i} \in A} \tilde{Z}(\mathbf{i})$.

Next, we consider I_4. Define $d_3(J) = \max\{d(\{\mathbf{i}_2\}, \{\mathbf{i}_2\}^c), d(\{\mathbf{i}_3\}, \{\mathbf{i}_3\}^c)\}$ for any collection of indices $J = \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3\}$ in I_4. By an argument similar to the above, it can be shown that $|\{J : d_3(J) = d_{03}\}| \le C(d)\, d_{03}^{2d-1}|J_n|$. Hence, it follows that

$$I_4 \le C(d)\left(\sum_{k=1}^{\infty} k^{2d-1}\rho(k-1;2)\right)|J_n|\delta_n$$
$$\le C(d,\rho(\cdot))t_n^d\delta_n \quad .$$

Lemma 4.1 now follows combining the bounds for $I_1, \ldots, I_5$.

Lemma 4.2. *Suppose that B_n is a region in R_n satisfying the conditions of Lemma* 4.1 *and that assumption* (A.5) *holds with λ_n replaced by t_n. If, in addition, assumptions* (A.1) *and* (A.3) *hold, then for any $z_1, z_2 \in \mathbb{R}$,*

$$E\left(\sum_{\mathbf{i}\in J_n} Y_1(\mathbf{i})\right)\left(\sum_{\mathbf{i}\in J_n} Y_2(\mathbf{i})\right)$$
$$= |B_0|^{-1}t_n^d h_n^\kappa\left[\sum_{|\alpha|=\kappa} a(\alpha)(\alpha!)^{-1}\int_{\mathbb{R}^d} D^\alpha G(z_1,z_2;\mathbf{x})\,\mathbf{dx}\right](1+o(1))$$

where $J_n \equiv \{\mathbf{i} \in \mathscr{Z}^d : \Gamma_{\mathbf{i}} \cap B_n \ne \phi\}$ (different from the J_n of Lemma 4.1) *and $Y_j(\mathbf{i}) = \int_{\Gamma(1,\mathbf{i})}(1(Z(\mathbf{s_i}) \le z_j) - 1(Z(\mathbf{s}) \le z_j))\,\mathbf{ds}$, with $\Gamma(1,\mathbf{i}) = \Gamma_{\mathbf{i}} \cap B_n,\ \ \mathbf{i} \in J_n$, and $j = 1, 2$.*

Proof. We consider the case $\mathbf{c} = \mathbf{c}_o$ only. For $\mathbf{c} \ne \mathbf{c}_o$, similar arguments can be used to prove the conclusions of the lemma. By Taylor's expansion,

$$I_n \equiv E\left(\sum_{\mathbf{i}\in J_n} Y_1(\mathbf{i})\right)\left(\sum_{\mathbf{i}\in J_n} Y_2(\mathbf{i})\right)$$
$$= \sum_{\mathbf{i}}\sum_{\mathbf{j}}\int_{\Gamma(1,\mathbf{i})}\int_{\Gamma(1,\mathbf{j})}[G(z_1,z_2;\mathbf{s_j}-\mathbf{s_i}) - G(z_1,z_2;\mathbf{x}-\mathbf{s_i})$$
$$- G(z_1,z_2;\mathbf{s_j}-\mathbf{s}) + G(z_1,z_2;\mathbf{x}-\mathbf{s})]\,\mathbf{ds}\,\mathbf{dx}$$

$$= \sum_{\mathbf{i}}\sum_{\mathbf{j}}\int_{\Gamma(1,\mathbf{i})}\int_{\Gamma(1,\mathbf{j})}\left[\sum_{|\alpha|=2}^{4}(D^\alpha G(z_1,z_2;\mathbf{s_j}-\mathbf{s_i})/\alpha!)\right.$$
$$\times\left\{-(\mathbf{x}-\mathbf{s_j})^\alpha - (\mathbf{s_i}-\mathbf{s})^\alpha + ((\mathbf{x}-\mathbf{s})-(\mathbf{s_j}-\mathbf{s_i}))^\alpha\right\}$$
$$\left. + r_n(\mathbf{i},\mathbf{j};\mathbf{x},\mathbf{s})\right]\mathbf{ds}\,\mathbf{dx} \quad ,$$

where $|r_n(\mathbf{i},\mathbf{j};\mathbf{x},\mathbf{s})| \le C(d)(\|\mathbf{x}-\mathbf{s_j}\|^{4+\eta} + \|\mathbf{s}-\mathbf{s_i}\|^{4+\eta}) \sum_{|\alpha|=4} H_\alpha(z_1, z_2; \mathbf{x}-\mathbf{s})$.

Next define $a_n(\mathbf{i},\mathbf{j};\alpha) \equiv \int_{\Gamma(1,\mathbf{i})}\int_{\Gamma(1,\mathbf{j})}\{-(\mathbf{x}-\mathbf{s_j})^\alpha - (\mathbf{s_i}-\mathbf{s})^\alpha + ((\mathbf{x}-\mathbf{s}) - (\mathbf{s_j}-\mathbf{s_i}))^\alpha\}\,\mathbf{ds}\,\mathbf{dx}$, $J_{1n} = \{\mathbf{i} \in J_n : \Gamma_\mathbf{i} \cap B_n \neq \phi, \Gamma_\mathbf{i} \cap B_n^c \neq \phi\}$ and $J_{2n} = J_n \backslash J_{1n}$. Also, for $|\alpha| = 2$, let β and $\gamma \in \mathscr{Z}_+^d$ be such that $|\beta| = |\gamma| = 1$ and $\beta + \gamma = \alpha$. Then it follows that for $|\alpha| = 2$ and $\mathbf{i},\mathbf{j} \in J_{2n}$,

$$\begin{aligned} a_n(\mathbf{i},\mathbf{j};\alpha) &= \int_{\Gamma_\mathbf{i}}\int_{\Gamma_\mathbf{j}}\{(\mathbf{x}-\mathbf{s_j})^\beta(\mathbf{s_i}-\mathbf{s})^\gamma + (\mathbf{x}-\mathbf{s_j})^\gamma(\mathbf{s_i}-\mathbf{s})^\beta\}\,\mathbf{ds}\,\mathbf{dx} \\ &= h_n^{2d+2}[(1/2-\mathbf{c}^\beta)(\mathbf{c}^\gamma - 1/2) + (1/2-\mathbf{c}^\gamma)(\mathbf{c}^\beta - 1/2)] \\ &= -2h_n^{2d+2}(\mathbf{c}_o - \mathbf{c})^\alpha \ . \end{aligned}$$

Hence, for $\mathbf{c} = \mathbf{c}_o$, terms corresponding to $|\alpha| = 2$ vanish for all $\mathbf{i},\mathbf{j} \in J_{2n}$.

Next we turn to the cases $|\alpha| = 3, 4$. Note that for $\mathbf{i},\mathbf{j} \in J_{2n}$,

$$\begin{aligned} a_n(\mathbf{i},\mathbf{j};\alpha) &= h_n^{2d+|\alpha|}\int_{\Delta_0}\int_{\Delta_0}\{-(\mathbf{x}-\mathbf{c}_o)^\alpha - (\mathbf{c}_o-\mathbf{s})^\alpha + (\mathbf{x}-\mathbf{s})^\alpha\}\,\mathbf{dx}\,\mathbf{ds} \\ &\equiv h_n^{2d+|\alpha|}a(\alpha) \ . \end{aligned}$$

It is easy to check that $a(\alpha) = 0$ for all $|\alpha| = 3$ and $a(\alpha) \neq 0$ for $|\alpha| = 4$. Next write $\sum^{(1)}$ for summation over all $\mathbf{i} \in J_{1n}$ and $\sum^{(2)}$ for all $\mathbf{i} \in J_{2n}$. Then, noting that $|J_{1n}| = O(t_n^{d-1}h_n^{1-d})$, we have

$$\begin{aligned} I_{1n} &\equiv \sum_{|\alpha|=4}(\alpha!)^{-1}\sum_{\mathbf{i}}^{(2)}\sum_{\mathbf{j}}^{(2)} D^\alpha G(z_1, z_2; \mathbf{s_i}-\mathbf{s_j})a_n(\mathbf{i},\mathbf{j};\alpha) \\ &= |B_0|t_n^d h_n^4 \sum_{|\alpha|=4} a(\alpha)(\alpha!)^{-1}\left[\int_{\mathbb{R}^d} D^\alpha G(z_1, z_2; \mathbf{s})\,\mathbf{ds}\right](1+o(1)) \ , \end{aligned} \tag{4.3}$$

and

$$\begin{aligned} I_{2n} &\equiv \sum_{\mathbf{i}}\sum_{\mathbf{j}}\int_{\Gamma(1,\mathbf{i})}\int_{\Gamma(1,\mathbf{j})} |r_n(\mathbf{i},\mathbf{j};\mathbf{x},\mathbf{s})|\,\mathbf{dx}\,\mathbf{ds} \\ &\le C(d)t_n^d h_n^{4+\eta}\sum_{|\alpha|=4}\int H_\alpha(z_1, z_2; \mathbf{s})\,\mathbf{ds} \ . \end{aligned} \tag{4.4}$$

Next note that for any $\mathbf{i},\mathbf{j} \in J_n$,

$$\begin{aligned}\mathbf{s}_i - \mathbf{s}_j &\in \{\mathbf{x} - \mathbf{y} : \mathbf{x} \in \Gamma_{\mathbf{i}},\ \ \mathbf{y} \in \Gamma_{\mathbf{j}}\} \\ &= \{(\mathbf{i} + \mathbf{x})h_n - (\mathbf{j} + \mathbf{y})h_n : \mathbf{dx}, \mathbf{y} \in \Delta_0\} \\ &\subset ((\mathbf{i} - \mathbf{j}) + \Delta_1)h_n\end{aligned}$$

where $\Delta_1 = (-1, 1)^d$. Then, for any $p, q \in \{1, 2\}$,

$$\begin{aligned}J_{pqn} &\equiv \{\mathbf{k} \in \mathscr{Z}^d : \mathbf{k} = \mathbf{i} - \mathbf{j} \quad \text{for some} \quad \mathbf{i} \in J_{pn}, \mathbf{j} \in J_{qn}\} \\ &\subset \{\mathbf{k} \in \mathscr{Z}^d : \|\mathbf{k}\| \le C(d) t_n^d h_n^{-d}\}\end{aligned}$$

and for any $\mathbf{k} \in J_{pqn}$,

$$\begin{aligned}&|\{(\mathbf{i}, \mathbf{j}) \in J_{pn} \times J_{qn} : \mathbf{i} - \mathbf{j} = \mathbf{k}\}| \\ &\qquad \le \min\{|J_{pn}|, |J_{qn}|\}\ .\end{aligned}$$

Hence, by (4.3) and (4.4), it follows that

$$\begin{aligned}&|I_n - I_{1n} - I_{2n}| \\ &= \left| \sum_{|\alpha|=2}^{4} \left(\sum_{\mathbf{i}}^{(1)} \sum_{\mathbf{j}}^{(1)} + \sum_{\mathbf{i}}^{(1)} \sum_{\mathbf{j}}^{(2)} + \sum_{\mathbf{i}}^{(2)} \sum_{\mathbf{j}}^{(1)} \right) D^\alpha G(\mathbf{s}_j - \mathbf{s}_i) a_n(\mathbf{i}, \mathbf{j}; \alpha) \right| \\ &\le C(d) \sum_{|\alpha|=2}^{4} h_n^{2d+2} \left(\sum_{\mathbf{k} \in J_{11n}} + \sum_{\mathbf{k} \in J_{12n}} + \sum_{\mathbf{k} \in J_{21n}} \right) \\ &\quad \cdot |J_{1n}| \sup\{|D^\alpha G(z_1, z_2; \mathbf{x})| : \mathbf{x} \in (\mathbf{k} + \Delta_1)h_n\} \\ &\le C(d, D(G)) t_n^{d-1} h_n^3\ .\end{aligned}$$

Therefore, in view of condition (A.5) (with $\lambda_n = t_n$), Lemma 4.2 is proved.

Lemma 4.3. *Let* $R_{1n} = (a_{1n}, b_{1n}) \times \cdots \times (a_{dn}, b_{dn})$ *be a rectangle with* $|R_{1n}| \ge c > 0$ *for all* n, *and* $R_{1n} \cap R_n \ne \phi$. *Also, let* $J_{3n} = \{\mathbf{i} \in \mathscr{Z}^d : \Gamma_{\mathbf{i}} \cap R_{1n} \cap R_n \ne \phi\}$. *Then, under Assumptions* (A.1), (A.2), *and* (A.3) (i), *for any* $-\infty \le z_1 < z_2 < \infty$,

$$E\left(\sum_{\mathbf{i} \in J_{3n}} Y_3(\mathbf{i})\right)^2 \le C(d, \rho(\cdot), G_1(z_1, z_2; \cdot)) h_n^\kappa |R_{1n}|$$

where $Y_3(\mathbf{i}) = \int_{\Gamma(2,\mathbf{i})} (1(z_1 < Z(\mathbf{s}_{\mathbf{i}}) \le z_2) - 1(z_1 < Z(\mathbf{s}) \le z_2))\, \mathbf{ds}$, *and* $\Gamma(2, \mathbf{i}) = \Gamma_{\mathbf{i}} \cap R_{1n} \cap R_n$, *for* $\mathbf{i} \in J_{3n}$.

Proof. Lemma 4.3 can be proved using arguments similar to the proof of Lemma 4.2. We omit the details.

Lemma 4.4. *Assume that the conditions of Theorem* 2.1 *hold. Then for any real numbers* $a_1, \ldots, a_r$ *and* $z_1, \ldots, z_r$ $(r \geq 1)$,

$$\sum_{i=1}^{r} a_i \xi_n(z_i) \to^d N\left(0, \sum_{i=1}^{k}\sum_{j=1}^{k} a_i a_j \sigma(z_i, z_j)\right)$$

where $\sigma(z_i, z_j)$ *is as in the statement of Theorem* 2.1.

Proof. Let $\{\lambda_{1n}\}$ and $\{\lambda_{2n}\}$ be two sequences of positive numbers (to be specified later) such that $\lambda_{1n}/h_n \in \mathscr{Z}$, $\lambda_{2n}/h_n \in \mathscr{Z}$, and that

$$\lambda_{1n}^{-1} + \lambda_{2n}^{-1} + \lambda_{2n}/\lambda_{1n} + \lambda_{1n}/\lambda_n \to 0 \quad \text{as} \quad n \to \infty \ . \tag{4.5}$$

Here we employ the "blocking" method of Bernstein (1944), with λ_{1n} defining the "big" blocks (cubes) and λ_{2n} defining the "little" blocks (parallelepipeds). To define the blocks, we use a grid along each axis defined by the points $s_{1i} = i\lambda_{3n}, \quad s_{2i} = s_{1i} + \lambda_{1n}, i = 0, \pm 1, \ldots$ where $\lambda_{3n} = \lambda_{1n} + \lambda_{2n}$. Let $I(1, i) = [s_{1i}, s_{2i})$, $I(2, i) = [s_{2i}, s_{1i+1})$ and $I(i) = [s_{1i}, s_{1i+1})$. Next define the parallelepipeds in $\mathbb{R}^d$ using the points s_{1i}, s_{2i} as

$$\Delta_n(\mathbf{i}; \mathbf{0}) = (\mathbf{i} + \Delta_o)\lambda_{3n}$$

$$\Delta_n(\mathbf{i}; \boldsymbol{\epsilon}) = I(\epsilon_1, i_1) \times \cdots \times I(\epsilon_d, i_d)$$

for $\mathbf{i} \in \mathscr{Z}^d$ and $\boldsymbol{\epsilon} \in \{1, 2\}^d \equiv \Theta$. Note that for any $\mathbf{i}$, $\Delta_n(\mathbf{i}; \mathbf{0}) = \cup_{\boldsymbol{\epsilon} \in \Theta} \Delta_n(\mathbf{i}; \boldsymbol{\epsilon})$, and that corresponding to the 2^d $\boldsymbol{\epsilon}$'s in Θ, we get 2^d "types" of parallelepipeds $\Delta_n(\cdot; \boldsymbol{\epsilon})$s. Furthermore, the volume of a parallelepiped of type $\boldsymbol{\epsilon} \in \Theta$ is given by $|\Delta_n(\mathbf{i}; \boldsymbol{\epsilon})| = \lambda_{1n}^q \lambda_{2n}^{d-q}$ where $q = q(\boldsymbol{\epsilon}) = |\{1 \leq j \leq d : \epsilon_j = 1\}|$. Hence, for all $\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_o \equiv (1, \ldots, 1)'$, and for all $\mathbf{i} \in \mathscr{Z}^d$,

$$|\Delta_n(\mathbf{i}; \boldsymbol{\epsilon})| = o(|\Delta_n(\mathbf{i}; \boldsymbol{\epsilon}_o)|) \quad \text{as} \quad n \to \infty \ . \tag{4.6}$$

We shall consider the cubes $\Delta_n(\mathbf{i}; \boldsymbol{\epsilon}_o)$ as defining our big "blocks" and the rest as little "blocks". Let $\sum^{(1)}$ denote summation over all $\mathbf{i} \in J_{6n} \equiv \{\mathbf{i} \in \mathscr{Z}^d : \Delta_n(\mathbf{i}; \mathbf{0}) \subset R_n\}$, and $\sum^{(2)}$ extend over all $\mathbf{i} \in J_{7n} \equiv \{\mathbf{i} \in \mathscr{Z}^d : \Delta_n(\mathbf{i}; \mathbf{0}) \cap R_n \neq \phi$ and $\Delta_n(\mathbf{i}; \mathbf{0}) \cap R_n^c \neq \phi\}$. Next define the random variables

$$Y_{\mathbf{i}} = (N_n h_n^d)^{-1} b_n \sum_{j=1}^{r} a_j \int_{\Gamma_{\mathbf{i}} \cap R_n} (1(Z(\mathbf{s}_{\mathbf{i}}) \leq z_j) - 1(Z(\mathbf{s}) \leq z_j))\, \mathbf{ds}, \ \ \mathbf{i} \in \mathscr{Z}^d,$$

$$Y(\mathbf{i}; \boldsymbol{\epsilon}) = \sum_{\mathbf{j}: \Gamma_{\mathbf{j}} \subset \Delta_n(\mathbf{i}; \boldsymbol{\epsilon})} Y_{\mathbf{j}}, \quad \mathbf{i} \in J_{6n}, \ \ \boldsymbol{\epsilon} \in \Theta$$

and

$$Y(\mathbf{i};\mathbf{0}) = \sum_{\mathbf{j}:\Gamma_{\mathbf{j}}\cap R_n\cap\Delta_n(\mathbf{i};\mathbf{0})\neq\phi} Y_{\mathbf{j}}, \quad \mathbf{i}\in J_{7n} \ .$$

Then it follows that

$$\sum_{i=1}^{r} a_i\xi_n(z_i) = \sum_{\boldsymbol{\epsilon}\in\Theta}\sum_{\mathbf{i}}^{(1)} Y(\mathbf{i};\boldsymbol{\epsilon}) + \sum_{\mathbf{i}}^{(2)} Y(\mathbf{i};\mathbf{0}) + \sum_{i=1}^{r} a_i((N_n h_n^d)^{-1} - |R_n|^{-1}) \times b_n \int_{R_n} (I(Z(\mathbf{s})\leq z_i) - F_0(z_i))\, \mathbf{ds} \ . \tag{4.7}$$

By Lemma 1.8.1 of Ivanov and Leonenko (1989) and the boundary condition on R_0, the third term goes to zero in L^2. We now show that the second term and the terms corresponding to $\boldsymbol{\epsilon}\neq\boldsymbol{\epsilon}_o$ in the first summation also tend to zero in L^2. Using the ρ-mixing condition and Lemma 4.3, and noting that $K_2\equiv|J_{7n}|\leq C(d)(\lambda_n/\lambda_{3n})^{d-1}$, we get

$$\begin{aligned}
&E\left(\sum_{\mathbf{i}}^{(2)} Y(\mathbf{i};\mathbf{0})\right)^2 \\
&\leq \sum_{k=0}^{K_2-1} |\{(\mathbf{i},\mathbf{j}) : \mathbf{i},\mathbf{j}\in J_{7n}, |\mathbf{i}-\mathbf{j}| = k\}|\rho((k-1)_+\lambda_{3n};\lambda_{3n}^d) \\
&\qquad \times \max\{EY(\mathbf{i};\mathbf{0})^2 : \mathbf{i}\in J_{7n}\} \\
&\leq C(d,\rho(\cdot),r)(\lambda_n/\lambda_{3n})^{d-1}\left(1+\sum_{k=1}^{K_2} k^{(d-1)}\rho(k\lambda_{3n};\lambda_{3n}^d)\right)\lambda_{3n}^d h_n^\kappa b_n^2\lambda_n^{-2d} \\
&\leq C(d,\rho(\cdot),r)(\lambda_{3n}^{d\theta-\tau})(\lambda_{3n}/\lambda_n) \ .
\end{aligned} \tag{4.8}$$

By similar arguments, noting that the distance between $\Delta(\mathbf{i};\boldsymbol{\epsilon})$ and $\Delta(\mathbf{j};\boldsymbol{\epsilon})$ is $(|\mathbf{i}-\mathbf{j}|-1)\lambda_{3n}+\lambda_{2n}$, we get

$$\begin{aligned}
&E\left(\sum_{\boldsymbol{\epsilon}\neq\boldsymbol{\epsilon}_o}\sum_{\mathbf{i}}^{(1)} Y(\mathbf{i};\boldsymbol{\epsilon})\right)^2 \leq 2^{2^d-2}\sum_{\boldsymbol{\epsilon}\neq\boldsymbol{\epsilon}_o} E\left(\sum_{\mathbf{i}}^{(1)} Y(\mathbf{i};\boldsymbol{\epsilon})\right)^2 \\
&\leq C(d,\rho(\cdot),r)\left[1+\sum_{1\leq k<L} k^{d-1}\rho(k\lambda_{3n}+\lambda_{2n};\lambda_{1n}^{d-1}\lambda_{2n})\right](\lambda_n/\lambda_{3n})^d \\
&\quad \times (\lambda_{1n}^{d-1}\lambda_{2n})h_n^\kappa b_n^2\lambda_n^{-2d} \\
&\leq C(d,\rho(\cdot),r)(\lambda_n/\lambda_{3n})^{d\theta}\lambda_{3n}^{d\theta-\tau}(\lambda_{2n}/\lambda_n)^\theta(\lambda_{2n}/\lambda_{3n})
\end{aligned} \tag{4.9}$$

where $L = |J_{6n}|$. Next note that by the ρ-mixing condition

$$
\begin{aligned}
&|E\exp\left(it\sum_{\mathbf{j}}^{(1)} Y(\mathbf{j};\epsilon_o)\right) - \prod_{\mathbf{j}\in J_{6n}} E\exp(itY(\mathbf{i};\epsilon_o))| \\
&\quad \le C(d)(\lambda_n/\lambda_{3n})^d \rho(\lambda_{2n}; |R_n|) \le C(d)(\lambda_n/\lambda_{3n})^{d+d\theta}\lambda_{3n}^{d\theta}\lambda_{2n}^{-\tau} \ . \qquad (4.10)
\end{aligned}
$$

Now choose $\{\lambda_{1n}\}$ and $\{\lambda_{2n}\}$ such that

$$\lambda_{1n} \sim \lambda_n(\log\lambda_n)^{-1}, \quad \text{and}$$

$$\lambda_{2n} \sim \lambda_n^{d\theta/\tau}\log\lambda_n \ .$$

Check that with this choice of λ_{1n} and λ_{2n}, (4.5) holds and that, the expressions in the final steps of the inequalities (4.8), (4.9) and (4.10) tend to zero as $n \to \infty$. Hence, from (4.8)–(4.10), it follows that

$$E\left(\sum_{j=1}^{r} a_j\xi_n(z_j) - \sum_{\mathbf{i}}^{(1)} Y(\mathbf{i};\epsilon_o)\right)^2 \to 0 \quad \text{as} \quad n\to\infty \ . \qquad (4.11)$$

Thus, $\sum_{i=1}^{r} a_i\xi_n(z_i)$ has the same limiting distribution as $\sum_{\mathbf{j}}^{(1)} X(\mathbf{j})$, where $X(\mathbf{j})$'s are iid random variables with $X(\mathbf{j}) =^d Y(\mathbf{j};\epsilon_o)$. Since by Lemmas 4.1–4.3, $EX(\mathbf{j})^4/(EX(\mathbf{j})^2)^2 = O(1)$, it now follows that $X(\mathbf{j})$'s satisfy the Lyapounov's Condition. Hence, in view of (4.11) and Lemma 4.2, Lemma 4.4 is proved.

Proof of Theorem 2.1. Since F_0 is continuous, by standard arguments, it is enough to show that the time-scaled process $\tilde{\xi}_n(t) \equiv \xi_n(F_0^{-1}(t))$, $t\in[0,1]$ converges in distribution to $W(F_0^{-1}(t))$ as random elements of $D[0,1]$, where $D[0,1]$ is the space of all right continuous functions on $[0,1]$ with left hand limits, equipped with the Skorohod metric. Note that by Lemma 4.4, the finite dimensional distributions of $\tilde{\xi}_n(\cdot)$ converge weakly to those of $W(F_0^{-1}(\cdot))$. Hence, it remains to establish the tightness of the sequence $\{\tilde{\xi}_n, n\ge 1\}$ and the almost sure continuity of the sample paths of $W(F_0^{-1}(\cdot))$. By Theorem 15.5 and the proof of Theorem 22.1 (cf. pages 198–199) of Billingsley (1968), both assertions would follow if we showed that for every $\epsilon > 0$, $\eta > 0$, there exists a $0 < \delta < 1$ such that for all sufficiently large n,

$$P(\sup\{|\tilde{\xi}_n(t) - \tilde{\xi}_n(s)| : s \le t \le (s+\delta)\wedge 1\} \ge \epsilon) \le \eta\delta \qquad (4.12)$$

where $0 \le s \le 1$, and for two real numbers a, b, $a\wedge b$ denotes the minimum of a and b. Fix $s\in[0,1]$, and $0 < \epsilon, \eta < 1$. Using the monotonicity of $\hat{F}_n$ and $\hat{F}_\infty$, it is easy to show that for any $0 \le t_1 \le t \le t_2 \le 1$, $\tilde{\xi}_n(t) \le \tilde{\xi}_n(t_2) + b_n(\tilde{F}_\infty(t_2) - \tilde{F}_\infty(t_1))$ and $\tilde{\xi}_n(t) \ge \tilde{\xi}_n(t_1) + b_n(\tilde{F}_\infty(t_1) - \tilde{F}_\infty(t_2))$ where $\tilde{F}_\infty(u) = \hat{F}_\infty(F_0^{-1}(u)), u \in [0, 1]$. Let $p \equiv p(n, \epsilon, \eta) = C(\epsilon, \eta)(\lambda_n^d h_n^\kappa)^{-1/\gamma}$, and for a positive integer m (to be determined later), let $\delta = mp$. Then, it follows that

$$\begin{aligned} &P(\sup\{|\tilde{\xi}_n(t) - \tilde{\xi}_n(s)| : s \le t \le (s + \delta) \wedge 1\} > \epsilon) \\ &\quad \le P(\max\{|\tilde{\xi}_n(s + ip) - \tilde{\xi}_n(s)| : 1 \le i \le m\} > \epsilon/6) \\ &\quad + P(\max\{|b_n(\tilde{F}_\infty(s + ip) - \tilde{F}_\infty(s + (i-1)p))| : 1 \le i \le m\} > \epsilon/2) , \end{aligned} \tag{4.13}$$

where we set $s + ip = 1$, if it exceeds 1. Next using Lemma 4.1 (with $B_n = R_n$), Lemma 4.3 (with $R_{1n} = \mathbf{j} + \Delta_0$), assumption (A.4), and the fact that $E\tilde{F}_\infty(t) = t, 0 \le t \le 1$, we get

$$\begin{aligned} &E(\tilde{\xi}_n(t_2) - \tilde{\xi}_n(t_1))^4 \\ &\quad \le C(d, \rho(\cdot), \kappa)\lambda_n^{-2d} h_n^{-2\kappa}[\lambda_n^{2d} h_n^{2\kappa}|t_2 - t_1|^{2\gamma} + \lambda_n^d h_n^\kappa |t_2 - t_1|^\gamma] \\ &\quad \le C(d, \rho(\cdot), \kappa)|t_2 - t_1|^{2\gamma}, \quad \text{if} \quad |t_2 - t_1| \ge p \end{aligned} \tag{4.14}$$

and

$$\begin{aligned} &E\{b_n[(\tilde{F}_\infty(t_2) - \tilde{F}_\infty(t_1)) - E(\tilde{F}_\infty(t_2) - \tilde{F}_\infty(t_1))]\}^4 \\ &\quad \le C(d, \rho(\cdot), \kappa)\lambda_n^{-2d} h_n^{-2\kappa}[\lambda_n^{2d}|t_2 - t_1|^2 + \lambda_n^d |t_2 - t_1|] \end{aligned} \tag{4.15}$$

for $0 \le t_1, t_2 \le 1$. Next by Assumption (A.5) and (4.15), it follows that there exists $n_1 = n_1(\epsilon, \eta) \ge 1$ such that for all $n \ge n_1$,

$$\begin{aligned} b_n p &= (\lambda_n^{d(\gamma-2)} h_n^{-\kappa(2+\gamma)})^{1/2\gamma} \\ &= ((\lambda_n^d h_n^{(2\gamma+1)\kappa})(\lambda_n^d h_n^\kappa)^{1-\gamma})^{-1/2\gamma} \\ &\le (\lambda_n^d h_n^{(2\gamma+1)\kappa})^{-(2-\gamma)/2\gamma} \\ &< \epsilon/4 \ , \end{aligned} \tag{4.16}$$

and

$$
\begin{aligned}
&P(\max\{|b_n(\tilde F_\infty(s+ip)-\tilde F_\infty(s+(i-1)p))|:1\le i\le m\}>\epsilon/2)\\
&\quad\le\sum_{i=1}^{m}P(b_n|\tilde F_\infty(s+ip)-\tilde F_\infty(s+ip-p)-p|>\epsilon/4)\\
&\quad\le C(d,\rho(\cdot),\kappa)\epsilon^{-4}m\lambda_n^{-2d}h_n^{-2\kappa}[\lambda_n^{2d}p^2+\lambda_n^{d}p]\\
&\quad\le C(d,\rho(\cdot),\kappa)\epsilon^{-4}(mp)\big([\lambda_n^{d}h_n^{(2\gamma+1)\kappa}]^{-1/\gamma}+[\lambda_n^{d}h_n^{2\kappa}]^{-1}\big)\\
&\quad\le\eta\delta/2\ . \qquad (4.17)
\end{aligned}
$$

Next, using Theorem 12.2 of Billingsley (1968) and inequality (4.14) above, we get

$$
\begin{aligned}
&P(\max\{|\tilde\xi_n(s+ip)-\tilde\xi_n(s)|:1\le i\le m\}>\epsilon/6)\\
&\quad\le C(d,\rho(\cdot),\kappa)\epsilon^{-4}(mp)^{2\gamma}\ . \qquad (4.18)
\end{aligned}
$$

Now, let $\delta=\delta(\eta,\epsilon)>0$ be such that for the constant $C(d,\rho(\cdot),\kappa)$ appearing in the last inequality, $C(d,\rho(\cdot),\kappa)\epsilon^{-4}\delta^{2\gamma-1}\le\eta/2$, and δ/p is an integer. (This is possible, since by assumption, $2\gamma-1>0$.) Therefore, taking $m=\delta/p$, (4.12) follows from (4.13)–(4.18). This completes the proof of Theorem 2.1.

Proof of Theorem 2.1$'$. Similar to the proof of Theorem 2.1.

Proof of Theorem 2.2. Use the Continuous Mapping Theorem (cf. Theorem 15.1, Billingsley (1968) and Theorem 4.2.12, Pollard (1984)).

Proof of Theorem 3.1. For $i=1,\dots,K$, define

$$\tilde\xi_n^{*i}(z)=b_l(F_n^{*i}(z)-\hat F_\infty(z;R_{*i})),\quad z\in\mathbb{R}$$

and

$$\tilde H_n(z,p)=K^{-1}\sum_{i=1}^{K}1(\|\tilde\xi_n^{*i}\|_p\le z),\quad z\in\mathbb{R}\ .$$

Then, by the stationarity of $Z(\cdot)$s and the ρ-mixing condition,

$$
\begin{aligned}
&E(\tilde H_n(z,p)-H_n(z,p))^2\\
&\quad\le C(d)K^{-2}K_1^d\Big(\sum_{k\le 2dK_1}k^{d-1}\rho(k\lambda_l;\lambda_l^d)\Big)\to 0\quad\text{as}\quad n\to\infty\ . \qquad (4.19)
\end{aligned}
$$

Let $\epsilon > 0$ be given. Since the CDF $H(\cdot; p)$, say, of $\|W\|_p$ is (uniformly) continuous on $\mathbb{R}$, there exists a $\eta > 0$ such that $\sup\{|H(z+\eta; p) - H(z-\eta; p)| : z \in \mathbb{R}\} < \epsilon$. Hence, using (4.19) and the monotonicity of $\hat{H}_n$ and $\tilde{H}_n$, it can be shown that

$$\begin{aligned}
&\lim_{n\to\infty} P(|\hat{H}_n(z,p) - \tilde{H}_n(z,p)| > 4\epsilon) \\
&\quad \le \lim_{n\to\infty} P(|\tilde{H}_n(z+\eta,p) - \tilde{H}_n(z-\eta,p)| > 3\epsilon) \\
&\qquad + \lim_{n\to\infty} P\left(K^{-1}\sum_{i=1}^{K} 1(|\|\xi_n^{*i}\|_p - \|\tilde{\xi}_n^{*i}\|_p| > \eta) > \epsilon\right) \\
&\quad \le C\epsilon^{-2} \lim_{n\to\infty} [E(\tilde{H}_n(z+\eta;p) - H_n(z+\eta;p))^2 \\
&\qquad + E(\tilde{H}_n(z-\eta;p) - H_n(z-\eta;p))^2] \\
&\qquad + \lim_{n\to\infty} P(b_l\|F_\infty^{*1} - \hat{F}_\infty(\cdot;R_\ell)\|_p > \eta)/\epsilon \\
&\quad = 0\ .
\end{aligned}$$

In the last step, we have used the fact that under the assumed conditions on l, $\lambda_l^{d/2} h_n^{-\kappa/2} \cdot (F_\infty^{*1}(\cdot) - \hat{F}_\infty(\cdot; R_\ell))$ converges weakly to $W(\cdot)$ on $D[-\infty, \infty]$. This follows from Theorem 2.1 by replacing the sequence $\{\lambda_n\}$ by $\{\lambda_{\ell_n}\}$.

Therefore, it follows that for all $z \in \mathbb{R}$,

$$\hat{H}_n(z,p) - H_n(z,p) \to_p 0$$

as $n \to \infty$. Since $\|W\|_p$ has a continuous distribution on $\mathbb{R}$ and $H_n(z,p)$ converges in distribution to it, an argument similar to (the proof of) Polyà's Theorem can now be used to complete the proof of Theorem 3.1.

Acknowledgements. The author would like to thank N. Cressie, N.J. Hsu, M. Kaiser, and T. Olshen for many interesting discussions and the referees for constructive suggestions that improved the presentation of the paper. Special thanks are due to Yoon-Dong Lee for his assistance with the figures.

References

Anderson, T.W.: Goodness of fit tests for spectral distributions. Ann. Statist., **21**, 830–847 (1993)

Bernstein, S.N.: An extension of a limit theorem in probability theory to sums of dependent random variables. Uspekhi Mat. Nauk., **10**, 65–114 (1944) (In Russian)

Bertail, P., Politis, D., Romano, J.: On subsampling estimators with with unknown rate of convergence. Technical Report no. 95-20, Department of Statistics, Stanford University (1995)

Bertail, P., Politis, D., Rhomari, N.: Subsampling continuous parameter random fields. Technical Report TR-96-18, Department of Mathematics and Statistics, University of Cyprus (1996)

Billingsley, P.: Convergence of Probability Measures, Wiley, New York (1968)

Bolthausen, E.: On the Central Limit Theorem for Stationary Mixing Random Fields. Ann. Probab., **10**, 1047–1050 (1982)

Bradley, R.C.: A Caution on Mixing Conditions for Random Fields. Statist. and Probab. Letters, **8**, 489–491 (1989)

Bulinskii, A.V., Zhurbenko, I.G.: A Central Limit Theorem for Random Fields. Dokl. Akad. Nauk. USSR, **226**, 23–25 (1976)

Cressie, N.A.C.: Statistics for Spatial Data. Wiley, New York (1993)

Doukhan, P.: Mixing. Lecture Notes in Statistics, No. 85. Springer-Verlag, New York (1994)

Guyon, X., Richardson, S.: Vitesse de convergence du theoreme de la limite centrale pour des champs faiblement dependants. Z. Wahrsch. Verw. Gebeite, **66**, 297–314 (1984)

Hall, P., Jing, B-Y.: On the sample Re-use Methods. J. Royal Statist. Soc., Ser B, **58**, 727–738 (1996)

Hall, P., Patil, P.: Properties of nonparametric estimators of autocovariance function for stationary random fields. Probability Theory and Related Fields **99**, 399–424 (1994)

Ivanov, A.V., Leonenko, N.N.: Statistical Analysis of Random Fields. Kluwer Academic Publishers, Boston (1989)

Kaiser, M, N.J. Hsu, Cressie, N., Lahiri, S.N.: Inference for spatial processes using subsampling: A simulation study. Environmetrics **8**, 485–502 (1997)

Lahiri, S.N.: On inconsistency of estimators under infill asymptotics for spatial data. Sankhya, Ser A **58**, 403–417 (1996)

Lahiri, S.N., Kaiser, M., Cressie, N., Hsu, N-J.: Prediction of spatial cumulative distribution functions using subsampling. J. Amer. Statist. Assoc. (To appear) (1999)

Majure, J.J., Cook, D., Cressie, N., Kaiser, M.S., Lahiri, S.N., Symanzik, J.: Spatial CDF estimation and visualization with applications to forest health monitoring. Computing Science and Statistics **27**, 93–101 (1995)

Politis, D., Romano, J.P.: Large sample confidence regions based on subsamples under minimal assumptions. Annals of Statistics **22**, 2031–2050 (1994)

Politis, D., Paparoditis, E., Romano, J.: Large sample inference for irregularly spaced dependent observations based on subsampling. Technical Report TR-96-06, Department of Mathematics and Statistics, University of Cyprus (1996)

Pollard, D.: Convergence of Stochastic Processes. Springer Verlag, New York (1984)

Possolo, A.: Subsampling a random field. In Spatial Statistics and Imaging, IMS Lecture Notes Monograph Series, **20**, Ed.- A. Possolo, Hayward, CA. Institute of Mathematical Statistics, pp. 286–294 (1991)

Sherman, M., Carlstein, E.: Nonparametric Estimation of the moments of a general statistic computed from spatial data. Journal of American Statistical Association **89**, 496–500 (1994)

Sherman, M., Carlstein, E.: Omnibus confidence intervals. Preprint, Department of Statistics, Texas A & M University (1997)