Central Limit Theorems and Uniform Laws of Large Numbers for Arrays of Random Fields

Nazgul Jenish[1] and Ingmar R. Prucha[2]

Department of Economics
University of Maryland
College Park, MD 20742

September 12, 2007

[1]Department of Economics, University of Maryland, College Park, MD 20742. Tel.: 301-405-6589, Email: jenish@econ.umd.edu

[2]Department of Economics, University of Maryland, College Park, MD 20742. Tel.: 301-405-3499, Email: prucha@econ.umd.edu

Abstract

Spatial-interaction models are being increasingly considered in economics, and have a long tradition in geography, regional science and urban economics. In this paper, we derive a new central limit theorem, a new law of large numbers and a new uniform law of large numbers for spatial processes, or random fields. Such limit theorems form the essential building blocks towards developing an asymptotic theory of M-estimators for spatial processes, including maximum likelihood and generalized method of moments estimators. The development of a general estimation theory has been hampered by lack of general limit theorems. In this paper, we establish limit theorems that are applicable to a broad range of data processes in economics and other fields. In particular, we extend the literature by considering weakly dependent random fields located on arbitrary unevenly spaced lattices in d-dimensional Euclidean space, and allow for spatial processes that are non-stationary, possibly with unbounded moments. We provide weak, yet primitive, sufficient conditions for each of the theorems.

JEL Classification: C10, C21, C31

Key words: Random field, spatial process, central limit theorem, uniform law of large numbers, law of large numbers

1 Introduction[1]

Spatial-interaction models have a long tradition in geography, regional science and urban economics. For the last two decades spatial-interaction models have also been increasingly considered in economics and the social sciences, in general. Applications range from their traditional use in agricultural, environmental, urban and regional economics to other branches of economics including international trade, industrial organization, labor, public economics, political economics, and macroeconomics.[2]

The proliferation of spatial-interaction models in economics was accompanied by an upsurge in contributions to a rigorous theory of estimation and testing of spatial-interaction models.[3] Much of those developments have focused on Cliff-Ord type models; cp. Cliff and Ord (1973, 1981). However, the development of a general theory of estimation for (possibly) nonlinear spatial-interaction models under sets of assumptions that are both general and accessible for interpretation by applied researchers has been hampered by a lack of pertinent central limit theorems (CLTs), uniform laws of large numbers (ULLNs), and laws of large numbers (LLNs). Evidently, such limit theorems form the basic modules one would typically employ in deriving the asymptotic properties of M-estimators for nonlinear spatial-interaction models, such as maximum likelihood (ML) and generalized method of moments (GMM) estimators. The purpose of this paper is to introduce a new CLT, ULLN and LLN for spatial processes (or random fields or multi-dimensional processes) under assumptions appropriate for many spatial processes in economics. As discussed in more detail below, our assumptions allow for nonstationary processes; in particular we allow processes to be heteroskedastic, and to have trending moments. Our assumptions also allow for sample regions of general configuration and, more importantly, for unevenly spaced locations. To accommodate Cliff-Ord type processes, we fur-

[1]We are grateful to Benedikt M. Pötscher for his valuable comments. We also thank the participants of the First World Conference of the Spatial Econometrics Association, Cambridge, July 2007, as well as the seminar participants at University of Maryland for helpful discussions. This research benefitted from a University of Maryland Ann G. Wylie Dissertation Fellowship for the first author and from financial support from the National Institute of Health through SBIR grant 1 R43 AG027622 for the second author.

[2]Some recent applications include Audretsch and Feldmann (1996), Baltagi, Egger and Pfaffermayr (2005), Bell and Bockstael (2000), Betrand, Luttmer and Mullainathan (2000), Besley and Case (1995), Brock and Durlauf (2001), Case (1991), Cohen and Morrison Paul (2004), Conley and Dupor (2003), De Long and Summers (1991), Hanushek et al (2003), Holtz-Eakin (1994), Ionnides and Zabel (2003), Keller and Shiue (2007) , Kling, Ludwig and Katz (2007), Pinkse, Slade and Brett (2002), Rees, Zaks and Herries (2003), Sacredote (2001), Shroder (1995) and Topa (2001).

[3]Some recent contributions to the theoretical econometrics literature include Baltagi and Li (2001a,b), Baltagi, Song, Jung and Koh (2005), Baltagi, Song and Koh (2003), Bao and Ullah (2007), Brock and Durlauf (2007), Conley (1999), Conley and Molinari (2007), Conley and Topa (2007), Das, Kelejian and Prucha (2003), Driscol and Kraay (1998), LeSage and Pace (2007), Kapoor, Kelejian and Prucha (2007), Kelejian and Prucha (2007a,b, 2004, 2002, 2001, 1999, 1998), Korniotis (2005), Lee (2007a,b,c, 2004, 2003, 2002), Pinkse and Slade (1998), Pinkse, Slade, and Brett (2002), Robinson (2007a,b), Sain and Cressie (2007), Su and Yang (2007), Yang (2005), and Yu, de Jong and Lee (2006).

thermore permit random variables to depend on the sample, i.e., to form triangular arrays. For short, we consider arrays of weakly dependent nonstationary random fields on irregular lattices in $\mathbb{R}^d$.

To put our contribution into context, we begin by highlighting key differences of the limit theory of random fields, where the index set is a subset of the $\mathbb{R}^d$, $d > 1$, from the limit theory for one-dimensional processes, i.e., processes where the index set is a subset of the real line such as time series processes. First, there is no natural order in the $\mathbb{R}^d$. Moreover, the higher dimensionality and more complex geometry of the index sets give rise to different modes of convergence, e.g., Van Hove or Fischer modes of convergence.[4] In contrast to the one-dimensional case, restrictions on the configuration and growth behavior of the index sets therefore play an important role in the limit theory of random fields. Second, there is also a wider choice over definition of weak dependence, and in particular, mixing. Unlike mixing coefficients in the standard time series literature, those of random fields depend not only on the distance between two datasets, but also their sizes. Given a distance, it is natural to expect more dependence between two larger sets than between two smaller sets. Failure to take into account the cardinalities of index sets may result in trivial notions of dependence and leave out many dependent processes encountered in applications. For instance, Dobrushin (1968a,b) demonstrates that the multidimensional analogue of the standard time series α-mixing condition is not satisfied by simple two-state Markov chains on $\mathbb{Z}^2$.

There is a vast literature on CLTs for weakly dependent random fields under various mixing conditions, including Neaderhouser (1978, α-mixing), Nahapetian (1980, 1987, α- and ϕ-mixing), Bolthausen (1982, α-mixing), Bradley (1992, ρ^*-mixing), Guyon (1995, α-mixing), and McElroy and Politis (2000, α-mixing). These results have been obtained for random fields on the integer lattice $\mathbb{Z}^d$ and are, therefore, not immediately applicable to many spatial processes of interest, e.g., real estate prices, given that housing units are frequently unevenly spaced. Moreover, some of these theorems, e.g., Neaderhouser (1978) and McElroy and Politis (2000) rest on more stringent moment and mixing assumptions.

Apart from allowing for unevenly spaced locations, our CLT differs from the previous results in other critical aspects. First, our CLT relies only on fairly minimal assumptions with respect to the geometry and growth behavior of sample regions. This is in contrast to the existing CLTs, e.g., Nahapetian (1980, 1987), McElroy and Politis (2000) who restrict the sample regions to rectangles and adopt, respectively, Van Hove and Fischer modes of convergence. Neaderhouser (1978) also exploits the Van Hove mode of convergence. Bolthausen (1982) and Guyon (1995) require the sample regions to form a strictly increasing sequence, in which each subsequent set contains the preceding one, and Bolthausen (1982) additionally requires the size of the border to be negligible relative to that of the whole region.

Second, spatial processes encountered in applications are often nonstationary

[4] For formal definitions, see, e.g., Nahapetian (1991).

and, in particular, heteroskedastic, since spatial units often differ in various important dimensions such as size. However, most of the available results, e.g., Bolthausen (1982), Nahapetian (1980, 1987) maintain strict stationarity.[5] Our CLT accommodates nonstationary processes. Furthermore, to the best of our knowledge, there seem to be no results that allow for processes with unbounded moments, to which we will also refer to as trending spatial processes in analogy with time series processes. Spatial processes with unbounded moments may arise in a wide range of economic applications. For instance, real estate prices usually shoot up as one moves from the periphery to the center of a big city. Individual incomes in the European Union countries rise in the northwestern direction.[6]

Third, our CLT handles arrays of random fields, i.e., allows random variables to depend on the sample. This is important since spatial processes defined by the widely used class of Cliff-Ord models depend on the sample.[7]

ULLNs are essential tools for establishing consistency of nonlinear estimators; cp., eg., Gallant and White (1988), p. 19, and Pötscher and Prucha (1997), p. 17. Generic ULLN for time series processes have been introduced by Andrews (1987, 1992), Newey (1991) and Pötscher and Prucha (1989, 1994a,b). These ULLNs are generic in the sense that they transform pointwise LLNs into uniform ones, given some form of stochastic equicontinuity of the summands.[8]. ULLNs for time series processes, by their nature, assume evenly spaced observations on a line. They are not immediately suitable for fields on unevenly spaced lattices. The generic ULLN for random fields introduced in this paper is an extension of the one-dimensional ULLNs given in Pötscher and Prucha (1994a) and Andrews (1992). In addition to the generic ULLN, we also provide low level sufficient conditions for stochastic equicontinuity that are easy to check.[9]

Our pointwise weak LLN for spatial processes on general lattices in $\mathbb{R}^d$ is based on a subset of the assumptions maintained for our CLT, which facilitates their joint use in the proof of consistency and asymptotic normality of spatial estimators. The overwhelming majority of the existing LLNs[10] are strong laws

[5] Conley (1999) makes an important contribution towards developing an estimation theory of GMM estimators for spatial processes. In deriving the limiting distribution of his estimators, he utilizes Bolthausen's (1982) CLT, and thus maintains stationarity of the spatial processes.

[6] Cressie (1993) provides numerous examples of trending spatial processes.

[7] The recent CLT proposed by Pinske, Shen and Slade (2006) also allows for nonstationarity and dependence on the sample. This CLT relies on a set of high level assumptions including conditions on the rates of decay of the correlation among Bernstein's blocks, and the ability to select appropriate blocks. Of course, a crucial step in verifying a CLT for a particular process using Bernstein's blocking method is to demonstrate that it is indeed possible to form appropriate blocks. We note that there are α-mixing processes that are covered by our CLT but not by Pinske, Shen and Slade (2006). Thus, on a technical level, neither of the CLTs contains nor dominates the other.

[8] For different definitions of stochastic equicontinuity see Section 3 of the present paper or Pötscher and Prucha (1994a).

[9] The existing literature on the estimation of nonlinear spatial models has maintained high-level assumptions such as first moment continuity to imply uniform convergence; cp., e.g., Conley (1999). The results in this paper are intended to be more accessible, and in allowing, e.g., for nonstationarity, to cover larger classes of processes.

[10] See, e.g., Smythe (1973), Moricz (1978), Klesov (1981), Peligrad and Gut (1999), Noczaly

for fields on partially ordered rectangles in $\mathbb{Z}^d$, which prevents their use in more general settings.

The remainder of the paper is organized as follows. Section 2 defines the underlying weak dependence concepts and provides essential mixing inequalities. Our CLT for arrays of nonstationary α- and ϕ-mixing random fields on irregular lattices is presented in Section 3. The generic ULLN, pointwise LLN and various sufficient conditions are discussed in Section 4. All proofs are relegated to the appendix.

2 Weak Dependence Concepts and Mixing Inequalities

Establishing limit theorems for fields on irregular lattices poses various technical problems stemming from the higher dimensionality and intricate geometry of sample regions. First, there is a myriad of ways in which index sets can grow. The two basic asymptotic structures commonly used in the spatial literature are the so-called increasing domain and infill asymptotics, see, e.g., Cressie (1993), p. 480. In this paper, we employ increasing domain asymptotics. Second, for a CLT to hold, the variance of partial sums must obey a certain growth behavior (Ibragimov, 1962). Deriving a CLT hence involves determining the growth rate of the variances of partial sums, which in turn requires establishing bounds on the cardinalities of some basic sets on the lattice. Finally, there is also a wider choice over definition of mixing. As discussed below, not all of them are sensible and useful in applications, and therefore, should be handled with caution.

Before presenting our main results, we therefore tackle these issues. We consider spatial processes located on a (possibly) unevenly spaced lattice $D \subseteq \mathbb{R}^d$, $d \geq 1$. It proves convenient to consider $\mathbb{R}^d$ as endowed with the metric $\rho(i,j) = \max_{1 \leq l \leq d} |j_l - i_l|$, and the corresponding norm $|i| = \max_{1 \leq l \leq d} |i_l|$, where i_l denotes the l-th component of i. Let $B(i,h) = \{j \in \mathbb{R}^d : \rho(i,j) \leq h\}$ denote the d-dimensional ball of radius $h > 0$ centered in $i \in \mathbb{R}^d$. Note that given our metric $B(i,h)$ represents a d-dimensional hyper-cube. For any subsets $U, V \subset D$ we define the distance between them as $\rho(U,V) = \inf\{\rho(i,j) : i \in U \text{ and } j \in V\}$. Furthermore, for any finite subset $U \subset D$ we denote its cardinality by $|U|$.

Throughout the sequel, we maintain the following assumption concerning D.

Assumption 1 *The lattice $D \subset \mathbb{R}^d$, $d \geq 1$, is infinite countable. All elements in D are located at distances of at least $d_0 > 0$ from each other, i.e., $\forall$ i, $j \in D : \rho(i,j) \geq d_0$; w.l.o.g. we assume that $d_0 > 1$.*

The assumption of a minimum distance has also been used by Conley (1999). It assures unbounded expansion of sample regions, and rules out infill asymptotics. It turns out that this *single* restriction on irregular lattices also provides sufficient structure for the index sets to permit the derivation of our limit results. Based on Assumption 1, Lemma A.1 in the Appendix establishes bounds

and Tomacs (2000).

on the cardinalities of some basic sets in D that will be used in the proof of the limit theorems.

We now turn to the weak dependence concepts employed in our theorems. Let $X = \{X_{i,n}; i \in D_n, n \in \mathbb{N}\}$ be a triangular array of real random fields defined on a common probability space $(\Omega, \mathfrak{F}, P)$, where D_n is a finite subset of D, and D satisfies Assumption 1. Further, let $\mathfrak{A}$ and $\mathfrak{B}$ be two sub- σ-algebras of $\mathfrak{F}$. Two common measures of dependence between $\mathfrak{A}$ and $\mathfrak{B}$, are α- and ϕ-mixing introduced, respectively, by Rosenblatt (1956) and Ibragimov (1962), defined as:

$$\begin{aligned} \alpha(\mathfrak{A}, \mathfrak{B}) &= \sup(|P(A \cap B) - P(A)P(B)|, A \in \mathfrak{A}, B \in \mathfrak{B}), \\ \phi(\mathfrak{A}, \mathfrak{B}) &= \sup(|P(A \mid B) - P(A)|, A \in \mathfrak{A}, B \in \mathfrak{B}, P(B) > 0). \end{aligned}$$

The concepts of α- and ϕ-mixing have been used extensively in the time series literature as measures of weak dependence. Recall that a time series process $\{X_t\}_{-\infty}^{\infty}$ is α-mixing [ϕ-mixing] if

$$\begin{aligned} \lim_{m\to\infty} \sup_t \alpha(\mathfrak{F}_{-\infty}^t, \mathfrak{F}_{t+m}^{+\infty}) &= 0 \\ [\lim_{m\to\infty} \sup_t \phi(\mathfrak{F}_{-\infty}^t, \mathfrak{F}_{t+m}^{+\infty}) &= 0], \end{aligned}$$

where $\mathfrak{F}_{-\infty}^t = \sigma(..., X_{t-1}, X_t)$ and $\mathfrak{F}_{t+m}^{\infty} = \sigma(X_{t+m}, X_{t+m+1}...)$. This definition captures the basic idea of diminishing dependence between different events as the distance between them increases.

To generalize these concepts to random fields, one could use formulations in close analogy with those employed for time-series processes. For instance, let H_k^a be a collection of all half-spaces of the type $\{i = (i_1, ..., i_d) \in \mathbb{R}^d, i_k \leq a\}$ and let H_k^b be a collection of all half-spaces of the type $\{i = (i_1, ..., i_d) \in \mathbb{R}^d, i_k \geq b\}$, with $a < b, a, b \in \mathbb{R}$, which are formed by the hyperplanes perpendicular to the k-th coordinate axis, $k = 1, .., d$. Define α-mixing coefficient in the k-th direction as

$$\alpha^k(r) = \sup\{\alpha(V_1, V_2) : V_1 \in H_k^a, V_2 \in H_k^b, \rho(V_1, V_2) \geq r\},$$

where $\alpha(V_1, V_2) = \alpha(\sigma(X_i; i \in V_1), \sigma(X_i; i \in V_2))$. The multidimensional counterpart to the conventional α-mixing coefficient is then obtained by taking supremum over all $d$directions, i.e.,

$$\widehat{\alpha}(r) = \sup_{1\leq k\leq d} \alpha^k(r).$$

These conditions were considered by Eberlein and Csenki (1979) and Hegerfeldt and Nappi (1977), who showed that some Ising ferromagnet lattice systems satisfy the condition $\widehat{\alpha}(r) \to 0$ as $r \to \infty$. However, as demonstrated by Dobrushin (1968a,b), the latter condition is generally restrictive for $d > 1$. It is violated even for simple two-state Markov chains on $D = \mathbb{Z}^2$. The problem with definitions of this ilk is that they neglect potential accumulation of dependence between σ-algebras $\sigma(X_i; i \in V_1)$ and $\sigma(X_i; i \in V_2)$ as sets V_1 and V_2 expand while the distance between them is kept fixed. Given a fixed distance, it is natural to expect more dependence between two larger sets than between two smaller sets.

Thus, extending mixing concepts to random fields in a practically useful way requires accounting for the sizes of subsets on which σ-algebras reside. Mixing conditions that depend on subsets of the lattice date back to Dobrushin (1968b). They were further expanded by Nahapetian (1980, 1987) and Bolthausen (1982). Following these authors, we adopt the following definitions of mixing:

Definition 1 *For $U \subseteq D_n$ and $V \subseteq D_n$, let $\sigma_n(U) = \sigma(X_{i,n}; i \in U)$, $\alpha_n(U,V) = \alpha(\sigma_n(U), \sigma_n(V))$ and $\phi_n(U,V) = \phi(\sigma_n(U), \sigma_n(V))$. Then the α- and ϕ-mixing coefficients for the array of random fields X are defined as follows:*

$$\begin{aligned} \alpha_n(k,l,r) &= \sup(\alpha_n(U,V), |U| \leq k, |V| \leq l, \rho(U,V) \geq r), \\ \phi_n(k,l,r) &= \sup(\phi_n(U,V), |U| \leq k, |V| \leq l, \rho(U,V) \geq r), \end{aligned}$$

with $k,l,r,n \in \mathbb{N}$. Furthermore, we will refer to

$$\begin{aligned} \overline{\alpha}(k,l,r) &= \sup_n \alpha_n(k,l,r), \\ \overline{\phi}(k,l,r) &= \sup_n \phi_n(k,l,r), \end{aligned}$$

as the corresponding uniform α- and ϕ-mixing coefficients.

As shown by Dobrushin (1968a,b), the weak dependence conditions based on the above mixing coefficients are satisfied by a large class of random fields including Gibbs fields. These mixing coefficients were also used by Doukhan (1994) and Guyon (1995), albeit without dependence on the sample. Given the array formulation, our definition allows for the latter dependence. The α-mixing coefficients for arrays of random fields used in McElroy and Politis (2000) are identical to ours. Doukhan (1994) provides an excellent overview of various mixing concepts.

We further note that if $Y_{i,n} = f(X_{i,n})$ is a Borel-measurable function of $X_{i,n}$, then $\sigma_n^Y(U) = \sigma(Y_{n,i}, i \in U) \subseteq \sigma_n^X(U)$, and hence $c_n^Y(U,V) \leq c_n^X(U,V)$, $c_n^Y(k,l,r) \leq c_n^X(k,l,r)$, $\overline{c}^Y(k,l,r) \leq \overline{c}^X(k,l,r)$ for $c \in \{\alpha, \phi\}$. Thus α- and ϕ-mixing conditions are preserved under transformation.

The value of the above mixing concepts in establishing limit theorems stems from the availability of corresponding moment inequalities. For convenience and ease of reference, we collect in the following lemma the covariance inequalities for α- and ϕ-mixing fields, which are central for proving our limit theorems.

Lemma 1 *Suppose U and V are finite sets in D_n with $|U| = \overline{k}, |V| = \overline{l}$ and $h = \rho(U,V)$, and let f and g be respectively $\sigma_n(U)$- and $\sigma_n(V)$-measurable.*

(i) *If $E|f|^p < \infty$ and $E|g|^q < \infty$ with $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$, $p,q > 1$ and $r > 0$, then*

$$|E(fg) - E(f)E(g)| < 8\alpha_n^{\frac{1}{r}}(\overline{k}, \overline{l}, h)(E|f|^p)^{\frac{1}{p}}(E|g|^q)^{\frac{1}{q}}$$

(ii) *If* $E|f|^p < \infty$ *and* $E|g|^q < \infty$ *with* $\frac{1}{p} + \frac{1}{q} = 1$, $p, q > 1$, *then*

$$|E(fg) - E(f)E(g)| < 2\phi_n^{\frac{1}{p}}(\overline{k}, \overline{l}, h)(E|f|^p)^{\frac{1}{p}}(E|g|^q)^{\frac{1}{q}}$$

(iii) *If* $|f| < C_1 < \infty$ *and* $|g| < C_2 < \infty$ *a.s., then*

$$|E(fg) - E(f)E(g)| < 4C_1C_2\alpha_n(\overline{k}, \overline{l}, h)$$

$$|E(fg) - E(f)E(g)| < 2C_1C_2\phi_n(\overline{k}, \overline{l}, h)$$

For a proof of the above inequalities, see, e.g., Hall and Heyde (1980), p. 277. The inequalities were originally derived by Ibragimov (1962).

3 Central Limit Theorem

Let $Z = \{Z_{i,n}; i \in D_n, n \in \mathbb{N}\}$ be an array of centered real random fields on a probability space $(\Omega, \mathfrak{F}, P)$, where the index sets D_n are finite subsets of $D \subset \mathbb{R}^d$, $d \geq 1$, which is assumed to satisfy Assumption 1. In the following, let $S_n = \sum_{i \in D_n} Z_{i,n}$ and $\sigma_n^2 = Var(S_n)$.

In this section, we provide a CLT for the normalized partial sums $\sigma_n^{-1} S_n$ of the array Z with possibly unbounded moments. Our CLT focuses on α- and ϕ-mixing fields and is based, respectively, on the following sets of assumptions.

Assumption 2 *(Uniform $L_{2+\delta}$ integrability) There exists an array of positive real constants $\{c_{i,n}\}$ such that*

$$\lim_{k \to \infty} \sup_n \sup_{i \in D_n} E[|Z_{i,n}/c_{i,n}|^{2+\delta} \mathbf{1}(|Z_{i,n}/c_{i,n}| > k)] = 0,$$

where $\mathbf{1}(\cdot)$ is the indicator function.

Assumption 3 *(α-mixing) The uniform α-mixing coefficients satisfy*

(a) $\sum_{m=1}^{\infty} m^{d-1} \overline{\alpha}(1, 1, m)^{\delta/(2+\delta)} < \infty$,

(b) $\sum_{m=1}^{\infty} m^{d-1} \overline{\alpha}(k, l, m) < \infty$ *for* $k + l \leq 4$,

(c) $\overline{\alpha}(1, \infty, m) = O(m^{-d-\varepsilon})$ *for some* $\varepsilon > 0$.

Assumption 4 *(ϕ-mixing) The uniform ϕ-mixing coefficients satisfy*

(a) $\sum_{m=1}^{\infty} m^{d-1} \overline{\phi}(1, 1, m)^{(1+\delta)/(2+\delta)} < \infty$,

(b) $\sum_{m=1}^{\infty} m^{d-1} \overline{\phi}(k, l, m) < \infty$ *for* $k + l \leq 4$,

(c) $\overline{\phi}(1, \infty, m) = O(m^{-d-\varepsilon})$ *for some* $\varepsilon > 0$.

Assumption 5 $\liminf_{n\to\infty} |D_n|^{-1} M_n^{-2} \sigma_n^2 > 0$*, where* $M_n = \max_{i\in D_n} c_{i,n}$.

Based on the above set of assumptions, we can formulate the following CLT for arrays of nonstationary random fields with possibly unbounded moments.

Theorem 1 *Suppose* $\{D_n\}$ *is a sequence of arbitrary finite subsets of* D*, satisfying Assumption 1, with* $|D_n| \to \infty$ *as* $n \to \infty$*. Suppose further that* $Z = \{Z_{i,n}; i \in D_n, n \in \mathbb{N}\}$ *is an array of real random fields with zero mean, where* Z *is either*

(a) α*-mixing satisfying Assumptions 2 and 3 for some* $\delta > 0$*, or*

(b) ϕ*-mixing satisfying Assumptions 2 and 4 for some* $\delta \geq 0$*.*

Suppose also that Assumption 5 holds, then

$$\sigma_n^{-1} S_n \Longrightarrow N(0,1).$$

The uniform L_p integrability condition postulated in Assumption 2 is a standard moment assumption seen in the CLTs for one-dimensional trending processes, e.g., Wooldridge (1986), Wooldridge and White (1988), Davidson (1992, 1993), and de Jong (1997). It ensures the existence of the $(2+\delta)$-th absolute moments of $Z_{i,n}$. A sufficient condition implying uniform $L_{2+\delta}$ integrability of $Z_{i,n}/c_{i,n}$ is their uniform L_r boundedness for some $r > 2+\delta$, i.e., $\sup_n \sup_{i\in D_n} |Z_{i,n}/c_{i,n}|^r < \infty$, see, e.g., Billingsley (1986), pp. 219.

The constants $c_{i,n}$ are scale factors that account for potentially unbounded moments of summands. For example, in the case of unbounded variances $v_{i,n}^2 = EZ_{i,n}^2$ the scale factors may be chosen as $c_{i,n} = \max(v_{i,n}, 1)$, and Assumption 2 would require uniform $L_{2+\delta}$ integrability of the array $Z_{i,n}/v_{i,n}$ for some $\delta > 0$. Within the context of time series processes, Davidson (1992) refers to the case with unbounded variances as global nonstationarity to distinguish it from the case of asymptotic covariance stationarity where the variance of normalized partial sums converges. In case the $Z_{i,n}$ are uniformly L_r bounded for some $r > 2$ the scale factors $c_{i,n}$ can be set to 1. While this case allows for some heterogeneity of the marginal distributions of $Z_{i,n}$, it would, e.g., not accommodate unbounded variances.

Spatial processes with unbounded moments, which correspond to trending processes in the time series literature, arise frequently in economics, geostatistics, epidemiology, regional and urban studies. A simple example from economics is real estate prices in a big city which frequently spike up as one moves from the outskirts of the city to its center.[11] Cressie (1993) contains numerous examples of spatial data exhibiting considerable heterogeneity and trend.

[11] For example, Bera and Simlai (2005) report on such spikes in housing prices and the variance of housing prices for central Boston.

These applications thus call for limit theorems covering spatial processes with unbounded moments.

Presently, to the best of our knowledge, there are no limit results for such spatial processes. All CLTs in the random fields literature rely on some form of uniform boundedness of Z_i. Therefore, when comparing our CLT with the existing results for $d > 1$, we shall always refer to the case $c_{i,n} = 1$. For the reference case, our moment Assumption 2 is slightly stronger than that in Bolthausen (1982), who assumes $L_{2+\delta}$ boundedness instead of integrability. This is not surprising since Bolthausen (1982) deals with strictly stationary processes, whereas our result allows for nonstationarity.[12]

Assumptions 3 and 4 restrict the dependence structure of the process Z. Assumption 3 is identical to the α-mixing conditions in Bolthausen (1982), seemingly, with the exception of Assumption 3c, in place of which Bolthausen postulates $\overline{\alpha}(1, \infty, m) = o(m^{-d})$. However, as pointed out by Goldie and Morrow (1986), p. 278, Bolthausen (1982) assumes polynomial decay of mixing coefficients. Therefore, our assumption and those in Bolthausen (1982) are equivalent. Assumption 4a parallels the ϕ-mixing condition used by Nahapetian (1991) to derive a CLT for strictly stationary ϕ-mixing random fields, see Theorem 7.2.2. Since ϕ-mixing is generally stronger than α-mixing, the rate of decay of mixing coefficients in Assumption 4a is slower than in Assumption 3a, and the corresponding moment condition (Assumption 2 with $\delta = 0$) in the ϕ-mixing case is weaker than that in the α-mixing case (Assumption 2 with $\delta > 0$). As shown in Bolthausen (1982), Assumptions 3 can be replaced with the following single condition: $\sum_{m=1}^{\infty} m^{d-1}\overline{\alpha}(2, \infty, m)^{\delta/(2+\delta)} < \infty$. Similarly, it is easy to see that the condition $\sum_{m=1}^{\infty} m^{d-1}\overline{\phi}(2, \infty, m)^{(1+\delta)/(2+\delta)} < \infty$ subsumes Assumptions 4.

Finally, Assumption 5 limits the growth behavior of $v_{i,n}^2 = EZ_{i,n}^2$.[13] For example, consider the case where $D_n = [-n; n]^d \subset \mathbb{Z}^d$, $Z_{i,n}$ satisfies Assumption 2 with $c_i = \max(v_i, 1)$, the $Z_{i,n}$ are uncorrelated, and $v_{i,n}^2$ grows with $|i|$. Then, Assumption 5 rules out exponential growth of the variances. However, Assumption 5 allows $v_{i,n}^2$ to grow at the rate of any finite nonnegative power of $|i|$. To see this, let $v_i^2 \sim |i|^\gamma$ for some $\gamma > 0$, then $M_n \sim n^{\gamma/2}$ and $\sigma_n^2 = \sum_{i \in D_n} v_i^2 \sim n^{(\gamma+d)}$. Observing that $|D_n| = (2n+1)^d$, it is then readily seen that Assumption 5 holds for arbitrary $\gamma > 0$. In the reference case, where $v_i^2 = O(1)$ and hence $M_n = O(1)$, Assumption 5 reduces to $\liminf_{n\to\infty} |D_n|^{-1}\sigma_n^2 > 0$, which is the condition employed by Bolthausen (1982). It rules out asymptotically degenerate distributions. In the literature on CLTs for time series processes with unbounded moments, similar assumptions were used by Wooldridge (1986) and

[12] Guyon (1995), p. 111, gives a CLT for nonstationary α-mixing random fields based on the conditions of Bolthausen (1982). He, too, assumes uniform $L_{2+\delta}$ boundedness. However, an important step of the proof that depends critically on the moment conditions is missing (for more details, see discussion below), thus raising concerns about his assumptions. His result is for random fields on the regular grid $\mathbb{Z}^d$ with $D_n \uparrow \mathbb{Z}^d$. Furthermore, it does not allow for trending moments and arrays.

[13] Extensions to variables with asymptoticaly vanishing variances is one direction for future research. In the one-dimensional literature, this task was accomplished by Davidson (1993a) and de Jong (1997).

Davidson (1992). These authors assume $\sup_n nM_n^2 < \infty$, while adopting the normalization $\sigma_n^2 = 1$. We note that in the case of $D = \mathbb{Z}$ and normalized variances $\sigma_n^2 = 1$, Assumption 5 becomes $\liminf_{n\to\infty} n^{-1}M_n^{-2} > 0$, or equivalently $\limsup_{n\to\infty} nM_n^2 < \infty$.

Of course, the above CLT can be readily generalized to vector-valued random fields using the standard Cramér-Wold device. We note that as a special case our CLT also contains a CLT for time series processes. As for the above-cited recent results in the one-dimensional literature, no strict comparison can be drawn as they are formulated for variables that are near-epoch dependent on mixing processes. Nevertheless, if one considers their special case where variables themselves are mixing, our CLT would include the CLTs of Davidson (1992, Theorem 3.6) and Wooldridge (1986, Theorem 3.13).

In the spatial context, the α-mixing part of our CLT extends Bolthausen's (1982) CLT in a number of important directions. First, it allows for the empirically important case where $D \subset \mathbb{R}^d$ is an unevenly spaced lattice. Second, it relaxes the assumption of stationarity. This is important, since many spatial processes considered in applied work may exhibit heteroskedasticity and other forms of nonstationarity, as sample units may often vary in size and other dimensions. Moreover, our CLT permits unbounded moments, and can thus also be applied to spatial processed that exhibit spikes in some of the moments. Third, it allows for the random variables to depend on the sample, as is, e.g., the case for the widely used class of Cliff-Ord type spatial processes. Finally, the CLT lifts Bolthausen's restrictions on the growth behavior of sets, namely that $D_n \uparrow D$ and $|\partial D_n| / |D_n| \to 0$, where ∂D_n is the border of D_n. The latter condition requires sets to grow in at least two non-opposing directions, and as a result, rules out sets that stretch in one direction. These patterns may arise under various spatial sampling procedures described in Ripley (1981), p. 19.

To provide additional insights and explain the rationale for strengthening the moment condition, we outline the structure of our proof. Perhaps, the most popular approach to proving CLTs for weakly dependent variables is Bernstein's blocking method. It involves splitting the sum into alternating big-small blocks and showing that the big blocks behave asymptotically as independent or martingale difference variables. In the spatial literature, this approach was, e.g., taken by Neaderhouser (1978), Nahapetian (1980, 1987), McElroy and Politis (2000). This method has, however, some undesirable features in the spatial context. First, as pointed out by Bolthausen (1982), this method leads to mixing conditions of the type: $\lim_{r\to\infty} \alpha(\infty,\infty,r) \to 0$. As noted earlier, this type of conditions are violated in many applications. Second, sectioning the sum into blocks and accounting for the sizes of blocks and remaining edges, already tedious on $\mathbb{Z}^d$, becomes a daunting task on unevenly spaced lattices. It seems that as a result, the existing results based on Bernstein's method typically impose quite stringent restrictions on the configuration and growth behavior of D_n. For instance, Nahapetian (1980, 1987), McElroy and Politis (2000) restrict D_n to rectangles in $\mathbb{Z}^d$ and adopt, respectively, some variant of Van Hove and Fischer mode of tendency of sets to infinity. Loosely speaking, these conditions require D_n to expand in all d directions and also assume that $|\partial D_n|/|D_n \to 0$. For exact definitions, see, e.g., Nahapetian (1991). Neaderhouser (1978) also relies on the Van Hove mode of convergence. In passing, we remark that the above results are also more restrictive than our CLT in other aspects: Nahapetian (1980, 1987) considers stationary fields, Neaderhouser (1978) and McElroy and Politis (2000), while permitting nonstationarity, rest on stronger moment and mixing assumptions.

In contrast, following Bolthausen (1982), our proof is based on Stein's lemma (1972); see Lemma B.1 in the Appendix. It exploits the differential equation satisfied by the characteristic function of the standard normal law. Stein's method allows us to circumvent mixing conditions of the type $\alpha(\infty, \infty, r)$ and to accommodate sample regions of arbitrary configuration and growth behavior. Our proof consists of three major steps. First, we demonstrate that the variances of the appropriately normalized partial sums are bounded from above. Second, we consider approximations of the partial sums $S_n = \sum_{i \in D_n} X_{i,n}$ in terms of partial sums $S_n^k = \sum_{i \in D_n} X_{i,n} \mathbf{1}(|X_{i,n}| \leq k)$, which correspond to the truncated versions of the scaled random variables with truncation point k. We then show that the limiting distribution of normalized S_n can be obtained as the sequential limiting distribution of S_n^k by letting first n and then k to infinity. The last and crucial step associated with Stein's method is to verify that, when properly normalized, S_n^k have the standard normal limiting distribution.

Bolthausen's proof builds on the arguments by Ibragimov and Linnik (1971), p. 345, to demonstrate that given stationarity and a regular lattice, the partial sums based on the truncated and original random variables converge to the same limiting distribution. Their argument exploits the fact that for stationary variables $\lim_{n \to \infty} |D_n|^{-1} \sigma_n^2 = \sigma_0^2 < \infty$. In our setting that allows for nonstationarity and irregular lattices, $|D_n|^{-1} \sigma_n^2$ need **not** converge, and therefore, we provide a different argument justifying the reduction to truncated variables.[14]

4 Uniform and Pointwise Law of Large Numbers

Uniform laws of large numbers (ULLNs) are a key tool for establishing consistency of nonlinear estimators. Suppose the true parameter of interest is $\theta_0 \in \Theta$, where Θ is the parameter space, and $\widehat{\theta}_n$ is a corresponding estimator defined as the maximizer of some real valued objective function $Q_n(\theta)$ defined on Θ, where the dependence on the data is suppressed. Suppose further that $EQ_n(\theta)$ is maximized at θ_0 and that θ_0 is identifiably unique. Then for $\widehat{\theta}_n$ to be consistent for θ_0, it suffices to show that $Q_n(\theta) - EQ_n(\theta)$ converge to zero uniformly over the parameter space; see, e.g., Gallant and White (1988), pp. 18, and Pötscher and Prucha (1997), pp. 16, for precise statements, which also allow the maximizers of $EQ_n(\theta)$ to depend on n. For many estimators the uniform convergence of $Q_n(\theta) - EQ_n(\theta)$ is established from a ULLN.

[14] The proof given in Guyon (1995) p.112 is similar to Bolthausen's, but does not furnish an explicit justification for the reduction to truncated variables in the nonstationary case. In supplying a rigorous argument for such reduction, we had to place slightly stronger conditions on the moments than Guyon.

In the following, we give a generic ULLN for spatial processes. The ULLN is generic in the sense that it turns a pointwise LLN into the corresponding uniform LLN. This generic ULLN assumes (i) that the random functions are stochastically equicontinuous in the sense made precise below, and (ii) that the functions satisfy a LLN for a given parameter value. For stochastic processes this approach was taken by Newey (1991), Andrews (1992), and Pötscher and Prucha (1994a).[15] Of course, to make the approach operational for random fields we need an LLN, and therefore we also introduce a new LLN for random fields. This LLN matches well with our CLT in that it holds under a subset of the conditions maintained for the CLT. We also report on two sets of sufficient conditions for stochastic equicontinuity that are fairly easy to verify.

As for our CLT, we consider again arrays of random fields residing on a (possibly) unevenly spaced lattice D, where $D \subset \mathbb{R}^d$, $d \geq 1$, is assumed to satisfy Assumption 1. However, for the ULLN the array is not assumed to be real-valued. More specifically, in the following let $\{Z_{i,n}; i \in D_n, n \in \mathbb{N}\}$, with D_n a finite subset of D, denote a triangular array of random fields defined on a probability space $(\Omega, \mathfrak{F}, P)$ and taking their values in Z, where $(Z, \mathcal{Z})$ is a measurable space. In applications, Z will typically be a subset of $\mathbb{R}^s$, i.e., $Z \subset \mathbb{R}^s$, and $\mathcal{Z} \subset \mathfrak{B}^s$, where $\mathfrak{B}^s$ denotes the s-dimensional Borel σ-field. We remark, however, that it suffices for the ULLN below if $(Z, \mathcal{Z})$ is only a measurable space. Further, in the following let $\{f_{i,n}(z,\theta), i \in D_n, n \in \mathbb{N}\}$ and $\{q_{i,n}(z,\theta), i \in D_n, n \in \mathbb{N}\}$ be doubly-indexed families of real-valued functions defined on $Z \times \Theta$, i.e., $f_{i,n}$: $Z \times \Theta \to \mathbb{R}$ and $q_{i,n}$: $Z \times \Theta \to \mathbb{R}$, where (Θ, ν) is a metric space with metric ν. Throughout the paper, the $f_{i,n}(\cdot, \theta)$ and $q_{i,n}(\cdot, \theta)$ are assumed $\mathcal{Z}/\mathfrak{B}$-measurable for each $\theta \in \Theta$ and for all $i \in D_n$, $n \geq 1$. Finally, let $B(\theta', \delta)$ be the open ball $\{\theta \in \Theta : \nu(\theta', \theta) < \delta\}$.

4.1 Generic Uniform Law of Large Numbers

The literature contains various definitions of stochastic equicontinuity. For a discussion of different stochastic equicontinuity concepts see, e.g., Andrews (1992) and Pötscher and Prucha (1994a). We note that apart from differences in the mode of convergence, the essential differences in those definitions relate to the degree of uniformity. We shall employ the following definition.[16]

Definition 2 *Consider array of random functions* $\{f_{i,n}(Z_{i,n}, \theta), i \in D_n, n \geq 1\}$. *Then* $f_{i,n}$ *is said to be*

[15] We note that the uniform convergence results of Bierens (1981), Andrews (1987), and Pötscher and Prucha (1989, 1994b) were obtained from closely related approach by verifying the so-called first moment continuity condition and from local laws of large numbers for certain bracketing functions. For a detailed discussion of similarities and differences see Pötscher and Prucha (1994a).

[16] All suprema and infima over subsets of Θ of random functions used below are assumed to be P-a.s. measurable. For sufficient conditions see, e.g., Pollard (1984), Appendix C, or Pötscher and Prucha (1994b), Lemma 2.

(a) L_0 *stochastically equicontinuous on* Θ *iff for every* $\varepsilon > 0$

$$\limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} P(\sup_{\theta'\in\Theta} \sup_{\theta\in B(\theta',\delta)} \left|f_{i,n}(Z_{i,n},\theta) - f_{i,n}(Z_{i,n},\theta')\right| > \varepsilon) \to 0 \text{ as } \delta \to 0;$$

(b) L_p *stochastically equicontinuous,* $p > 0$*, on* Θ *iff*

$$\limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} E(\sup_{\theta'\in\Theta} \sup_{\theta\in B(\theta',\delta)} \left|f_{i,n}(Z_{i,n},\theta) - f_{i,n}(Z_{i,n},\theta')\right|^p) \to 0 \text{ as } \delta \to 0;$$

(c) *a.s. stochastically equicontinuous on* Θ *iff*

$$\limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} \sup_{\theta'\in\Theta} \sup_{\theta\in B(\theta',\delta)} \left|f_{i,n}(Z_{i,n},\theta) - f_{i,n}(Z_{i,n},\theta')\right| \to 0 \text{ a.s. as } \delta \to 0.$$

Andrews (1992), within the context of one-dimensional processes, refers to L_0 stochastic equicontinuity as termwise stochastic equicontinuity. Pötscher and Prucha (1994a) refer to the stochastic equicontinuity concepts in Definition 2(a) *[*(b)*]*, *[[*(c)*]]* as asymptotic Cesàro L_0 *[*L_p*]*, *[[a.s.]]* uniform equicontinuity, and adopt the abbreviations ACL_0UEC *[*ACL_pUEC*]*, *[[*$a.s.ACUEC$*]]*. The following relationships among the equicontinuity concepts are immediate: $ACL_pUEC \Longrightarrow ACL_0UEC \Longleftarrow a.s.ACUEC$.

In formulating our ULLN, we will allow again for trending moments. We will employ the following domination condition.

Assumption 6 *(Domination Condition): There exists an array of positive real constants* $\{c_{i,n}\}$ *such that for some* $p \geq 1$ *:*

$$\limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} E(\ d_{i,n}^p \mathbf{1}(d_{i,n} > k)) \to 0 \text{ as } k \to \infty$$

where $d_{i,n}(\omega) = \sup_{\theta\in\Theta} |q_{i,n}(Z_{i,n}(\omega),\theta)| / c_{i,n}$.

We now have the following generic ULLN.

Theorem 2 *Suppose* $\{D_n\}$ *is a sequence of arbitrary finite subsets of* D*, satisfying Assumption 1, with* $|D_n| \to \infty$ *as* $n \to \infty$*. Let* (Θ, ν) *be a totally bounded metric space, and suppose* $\{q_{i,n}(z,\theta), i \in D_n, n \in \mathbb{N}\}$ *is a doubly-indexed family of real-valued functions defined on* $Z \times \Theta$ *satisfying Assumptions 6. Suppose further that the* $q_{i,n}(Z_{i,n},\theta)/c_{i,n}$ *are* L_0 *stochastically equicontinuous on* Θ*, and* *that for all $\theta \in \Theta_0$, where Θ_0 is a dense subset of Θ, the stochastic functions $q_{i,n}(Z_{i,n}, \theta)$ satisfy a pointwise LLN in the sense that*

$$\frac{1}{M_n |D_n|} \sum_{i \in D_n} [q_{i,n}(Z_{i,n}, \theta) - Eq_{i,n}(Z_{i,n}, \theta)] \to 0 \text{ i.p. [a.s.] as } n \to \infty, \quad (1)$$

where $M_n = \max_{i \in D_n} c_{i,n}$. Let $Q_n(\theta) = [M_n |D_n|]^{-1} \sum_{i \in D_n} q_{i,n}(Z_{i,n}, \theta)$, then

(a)

$$\sup_{\theta \in \Theta} |Q_n(\theta) - EQ_n(\theta)| \to 0 \text{ i.p. [a.s.] as } n \to \infty \quad (2)$$

(b) *$\overline{Q}_n(\theta) = EQ_n(\theta)$ is uniformly equicontinuous in the sense that*

$$\lim_{n \to \infty} \sup_{\theta' \in \Theta} \sup_{\theta \in B(\theta', \delta)} \left| \overline{Q}_n(\theta) - \overline{Q}_n(\theta') \right| \to 0 \text{ as } \delta \to 0. \quad (3)$$

The above ULLN adapts Corollary 4.3 in Pötscher and Prucha (1994a) to arrays of random fields, and also allows for trending moments. The case of bounded moments is covered as a special case with $c_{i,n} = 1$ and $M_n = 1$.

The ULLN allows for infinite-dimensional parameter spaces. It only maintains that the parameter space is totally bounded rather than compact. (Recall that a set of a metric space is totally bounded if for each $\varepsilon > 0$ it can be covered by a finite number of ε-balls). If the parameter space Θ is a finite-dimensional Euclidian space, then total boundedness is equivalent to boundedness, and compactness is equivalent to boundedness and closedness. By assuming only that the parameter space is totally bounded, the ULLN covers situations where the parameter space is not closed, as is frequently the case in applications.

Assumption 6 is implied by uniform integrability of individual terms, $d_{i,n}^p$, i.e., $\lim_{k \to \infty} \sup_n \sup_{i \in D_n} E(d_{i,n}^p \mathbf{1}(d_{i,n} > k)) = 0$, which, in turn, follows from their uniform L_r-boundedness for some $r > p$, i.e., $\sup_n \sup_{i \in D_n} \|d_{i,n}\|_r < \infty$.

Sufficient conditions for the pointwise LLN and the maintained L_0 stochastic equicontinuity of the normalized function $q_{i,n}(Z_{i,n}, \theta)/c_{i,n}$ are given in the next two subsection. The theorem only requires the pointwise LLN (1) to hold on a dense subset Θ_0, but, of course, also covers the case where $\Theta_0 = \Theta$.

As it will be seen from the proof, L_0 stochastic equicontinuity of $q_{i,n}(Z_{i,n}, \theta)/c_{i,n}$ and the Domination Assumption 6 jointly imply that $q_{i,n}(Z_{i,n}, \theta)/c_{i,n}$ is L_p stochastic equicontinuous for $p \geq 1$, which in turn implies uniform convergence of $Q_n(\theta)$ provided that a pointwise LLN is satisfied. Therefore, the weak part of ULLN will continue to hold if L_0 stochastic equicontinuity and Assumption 6 are replaced by the single assumption of L_p stochastic equicontinuity for some $p \geq 1$.

4.2 Pointwise Law of Large Numbers

The generic ULLN assumes a pointwise LLN for the stochastic functions $q_{i,n}(Z_{i,n}; \theta)$ for fixed $\theta \in \Theta$. In the following, we introduce a LLN for arrays of real random fields $\{Z_{i,n}; i \in D_n, n \in \mathbb{N}\}$ taking values in $Z = \mathbb{R}$ with possibly trending moments, which can in turn be used to establish a LLN for $q_{i,n}(Z_{i,n};\theta)$. The LLN below holds under a subset of assumptions of the CLT, Theorem 1, which facilitates their joint application. The CLT was derived under the assumption that the random field was uniformly $L_{2+\delta}$ integrable. As expected, for the LLN it suffices to assume uniform L_1 integrability.

Assumption 2 ** (Uniform L_1 integrability) There exists an array of positive real constants $\{c_{i,n}\}$ such that*

$$\lim_{k\to\infty} \sup_n \sup_{i\in D_n} E[|Z_{i,n}/c_{i,n}|\, \mathbf{1}(|Z_{i,n}/c_{i,n}| > k)] = 0,$$

where $\mathbf{1}(\cdot)$ is the indicator function.

A sufficient condition for Assumption 2* is $\sup_n \sup_{i\in D_n} E\,|Z_{i,n}/c_{i,n}|^{1+\eta} < \infty$ for some $\eta > 0$. We now have the following LLN.

Theorem 3 *Suppose $\{D_n\}$ is a sequence of arbitrary finite subsets of D, satisfying Assumption 1, with $|D_n| \to \infty$ as $n \to \infty$. Suppose further that $\{Z_{i,n}; i \in D_n, n \in \mathbb{N}\}$ is an array of real random fields satisfying Assumptions 2* and where the random field is either*

(a) *α-mixing satisfying Assumption 3(b) with $k = l = 1$, or*

(b) *ϕ-mixing satisfying Assumption 4(b) with $k = l = 1$.*

Then

$$\frac{1}{M_n\,|D_n|}\sum_{i\in D_n}(Z_{i,n} - EZ_{i,n}) \xrightarrow{L_1} 0,$$

where $M_n = \max_{i\in D_n} c_{i,n}$.

The existence of first moments is assured by the uniform L_1 integrability assumption. Of course, L_1-convergence implies convergence in probability, and thus the $Z_{i,n}$ also satisfies a weak law of large numbers. The theorem also covers uniformly bounded variables as a special case with $c_{i,n} = 1$ and $M_n = 1$. Comparing the LLN with the CLT reveals that not only the moment conditions employed in the former are weaker than those in the latter, but also the dependence conditions in the LNN are only a subset of the mixing assumptions maintained for the CLT.

There is a massive literature on weak LLNs for time series processes. Most recent contributions include Andrews (1988) and Davidson (1993b), among others. Andrews (1988) established an L_1-law for triangular arrays of L_1-mixingales. Davidson (1993b) extended the latter result to L_1-mixingale arrays with trending moments. Both results are based on the uniform integrability condition. In fact, our moment assumption is identical to that of Davidson (1993b). The mixingale concept, which exploits the natural order and structure of the time line, is formally weaker than that of mixing. It allows these authors to circumvent restrictions on the sizes of mixingale coefficients, i.e., rates at which dependence decays. Mixingales are not well-defined for random fields, without imposing a special order structure on the index space. Therefore, we cast our LLN in terms of mixing variables. Furthermore, due to the higher dimensionality and unevenness of the lattice, we have to make assumptions on the rates of decay of mixing coefficients.

The above LLN can be readily used to establish a pointwise LLN for stochastic functions $q_{i,n}(Z_{i,n};\theta)$ under the α- and ϕ-mixing conditions on $Z_{i,n}$ postulated in the theorem. For instance, suppose that $q_{i,n}(\cdot,\theta)$ is $\mathcal{Z}/\mathfrak{B}$-measurable and $\sup_n \sup_{i\in D_n} E\left|q_{i,n}(Z_{i,n};\theta)/c_{i,n}\right|^{1+\eta} < \infty$ for each $\theta \in \Theta$ and some $\eta > 0$, then $q_{i,n}(Z_{i,n};\theta)/c_{i,n}$ is uniformly L_1 integrable for each $\theta \in \Theta$. Recalling that the α- and ϕ-mixing conditions are preserved under measurable transformation, we see that $q_{i,n}(Z_{i,n};\theta)$ also satisfies a LNN for a given parameter value θ.

4.3 Stochastic Equicontinuity: Sufficient Conditions

In the previous sections, we saw that stochastic equicontinuity is a key ingredient of a ULLN. In this section, we explore various sufficient conditions for L_0 and *a.s.* stochastic equicontinuity of functions $f_{i,n}(Z_{i,n},\theta)$ as in Definition 2. These conditions place smoothness requirement on $f_{i,n}(Z_{i,n},\theta)$ with respect to the parameter and/or data. In the following, we will present two sets of sufficient conditions. The first set of conditions represent Lipschitz-type conditions, and only requires smoothness of $f_{i,n}(Z_{i,n},\theta)$ in the parameter θ. The second set requires less smoothness in the parameter, but maintains joint continuity of $f_{i,n}$ both in the parameter and data. These conditions should cover a wide range of applications and are relatively simple to verify. Lipschitz-type conditions for one-dimensional processes were proposed by Andrews (1987, 1992) and Newey (1991). Joint continuity-type conditions for one-dimensional processes were introduced by Pötscher and Prucha (1989). In the following we adapt those conditions to random fields.

We continue to maintain the setup defined at the beginning of the section.

4.3.1 Lipschitz in Parameter

Condition 1 *The array $f_{i,n}(Z_{i,n},\theta)$ satisfies for all $\theta,\theta' \in \Theta$ and $i \in D_n$, $n \geq 1$ the following condition:*

$$\left|f_{i,n}(Z_{i,n},\theta) - f_{i,n}(Z_{i,n},\theta')\right| \leq B_{i,n} h(\nu(\theta,\theta')) \text{ a.s.},$$

where h is a nonrandom function such that $h(x) \downarrow 0$ as $x \downarrow 0$, and $B_{i,n}$ are random variables that do not depend on θ such that for some $p > 0$

$$\limsup_{n\to\infty} |D_n|^{-1} \sum_{i\in D_n} EB_{i,n}^p < \infty \qquad [\limsup_{n\to\infty} |D_n|^{-1} \sum_{i\in D_n} B_{i,n} < \infty \text{ a.s.}]$$

Clearly, each of the above conditions on the Cesàro sums of $B_{i,n}$ is implied by the respective condition on the individual terms, i.e., $\sup_n \sup_{i \in D_n} EB_{i,n}^p < \infty$ *[*$\sup_n \sup_{i \in D_n} B_{i,n} < \infty$ *a.s.]*

Proposition 1 *Under Condition 1, $f_{i,n}(Z_{i,n}, \theta)$ is L_0 [a.s.] stochastically equicontinuous on Θ.*

4.3.2 Continuous in Parameter and Data

In this subsection, we assume additionally that Z is a metric space with metric τ and with $\mathcal{Z}$ the corresponding Borel σ-field. Also, let $B_\Theta(\theta, \delta)$ and $B_Z(z, \delta)$ denote δ-balls respectively in Θ and Z.

We consider functions of the form:

$$f_{i,n}(Z_{i,n}, \theta) = \sum_{k=1}^{K} r_{ki,n}(Z_{i,n}) s_{ki,n}(Z_{in}, \theta), \tag{4}$$

where $r_{ki,n} : Z \to \mathbb{R}$ and $s_{ki,n}(\cdot, \theta) : Z \to \mathbb{R}$ are real-valued functions, which are $\mathcal{Z}/\mathfrak{B}$-measurable for all $\theta \in \Theta$, $1 \leq k \leq K$, $i \in D_n$, $n \geq 1$. We maintain the following assumptions.

Condition 2 *The random functions $f_{i,n}(Z_{i,n}, \theta)$ defined in (4) satisfy the following conditions:*

(a) *For all* $1 \leq k \leq K$

$$\limsup_{n \to \infty} \frac{1}{|D_n|} \sum_{i \in D_n} E\left|r_{ki,n}(Z_{i,n})\right| < \infty.$$

(b) *For a sequence of sets $\{K_m\}$ with $K_m \in \mathcal{Z}$ the family of nonrandom functions $s_{ki,n}(z, \cdot)$, $1 \leq k \leq K$, satisfy the following uniform equicontinuity-type condition:*

$$\sup_n \sup_{i \in D_n} \sup_{z \in K_m} \sup_{\theta' \in \Theta} \sup_{\theta \in B(\theta', \delta)} \left|s_{ki,n}(z, \theta) - s_{ki,n}(z, \theta')\right| \to 0 \text{ as } \delta \to 0.$$

(c) *Also, for the sequence of sets $\{K_m\}$*

$$\lim_{m \to \infty} \limsup_{n \to \infty} \frac{1}{|D_n|} \sum_{i \in D_n} P\left(Z_{i,n} \notin K_m\right) = 0.$$

We now have the following proposition, which extends parts of Theorem 4.5 in Pötscher and Prucha (1994a) to arrays of random fields.

Proposition 2 *Under Condition 2, $f_{i,n}(Z_{i,n},\theta)$ is L_0 stochastically equicontinuous on Θ.*

We next discuss the assumptions of the above proposition and provide further sufficient conditions. We note that the $f_{i,n}$ are composed of two parts, $r_{ki,n}$ and $s_{ki,n}$, with the continuity conditions imposed only on the second part. Condition 2 allows for discontinuities in $r_{ki,n}$ with respect to the data. For example, the $r_{ki,n}$ could be indicator functions. A sufficient condition for Condition 2(a) is the uniform L_1 boundedness of $r_{ki,n}$, i.e., $\sup_n \sup_{i\in D_n} E\left|r_{ki,n}(Z_{i,n})\right| < \infty$.

Condition 2(b) requires *non*random functions $s_{ki,n}$ to be equicontinuous with respect to θ uniformly for all $z \in K_m$. This assumption will be satisfied if the functions $s_{ki,n}(z,\theta)$, restricted to $K_m \times \Theta$, are equicontinuous jointly in z and θ. More specifically, define the distance between the points (z,θ) and (z',θ') in the product space $Z \times \Theta$ by $r((z,\theta);(z',\theta')) = \max\left\{\nu(\theta,\theta'),\tau(z,z')\right\}$. This metric induces the product topology on $Z\times\Theta$. Under this product topology let $B((z',\theta'),\delta)$ be the open ball with center (z',θ') and radius δ in $K_m \times \Theta$. It is now easy to see that Condition 2(b) is implied by the following condition for each $1 \leq k \leq K$

$$\sup_n \sup_{i\in D_n} \sup_{(z',\theta')\in K_m\times\Theta} \sup_{(z,\theta)\in B((z',\theta'),\delta)} \left|s_{ki,n}(z,\theta) - s_{ki,n}(z',\theta')\right| \to 0 \text{ as } \delta \to 0,$$

i.e., the family of nonrandom functions $\{s_{ki,n}(z,\theta)\}$, restricted to $K_m \times \Theta$, is uniformly equicontinuous on $K_m\times\Theta$. Obviously, if both Θ and K_m are compact, the uniform equicontinuity is equivalent to equicontinuity, i.e.,

$$\sup_n \sup_{i\in D_n} \sup_{(z,\theta)\in B((z',\theta'),\delta)} \left|s_{ki,n}(z,\theta) - s_{ki,n}(z',\theta')\right| \to 0 \text{ as } \delta \to 0.$$

Of course, if the functions furthermore do not depend on i and n, then the condition reduces to continuity on $K_m\times\Theta$. Clearly if any of the above conditions holds on $Z \times \Theta$, then it also holds on $K_m \times \Theta$.

Finally, if the sets K_m can be chosen to be compact, then Condition 2(c) is an asymptotic tightness condition for the average of the marginal distributions of Z_{in}. Condition 2(c) can frequently be implied by a mild moment condition. In particular, the following is sufficient for Condition 2(c) in case $Z = \mathbb{R}^s$: $K_m \uparrow \mathbb{R}^s$ is a sequence of Borel measurable convex sets (for example, a sequence of open or closed balls), and $\limsup_{n\to\infty} |D_n|^{-1} \sum_{i\in D_n} Eh(Z_{in}) < \infty$ where $h : [0,\infty) \to [0,\infty)$ is a monotone function such that $\lim_{x\to\infty} h(x) = \infty$.[17]

We note that, in contrast to Condition 1, Condition 2 will generally not cover random fields with trending moments since in this case part (c) would typically not hold.

[17] For example $h(x) = x^p$ for some $p > 0$. The claim follows from lemma A4 in Pötscher and Prucha (1994b) with obvious modification to the proof.

5 Concluding Remarks

This paper derives a CLT, ULLN and LLN for arrays of random fields exhibiting considerable dependence and heterogeneity. The novel feature of these limit theorems is that they (i) allow for arrays of fields located on unevenly spaced lattices in $\mathbb{R}^d$, (ii) accommodate nonstationary fields with unbounded moments and (iii) place minimal restrictions on the configuration and growth behavior of index sets. The results are based on weak, yet primitive conditions which makes them applicable in a wide range of econometric contexts. They can readily be used to establish consistency and asymptotic normality of spatial estimators, and in particular, those arising from the Cliff-Ord-type models.

One direction for future research is to generalize the above CLT to random fields which are not mixing but can be approximated in some sense with mixing fields. This could be achieved, for example, by introducing the concept of near-epoch dependent random fields similar to the one used in the time-series literature. The authors are currently working in this direction. Another extension would be to obtain a CLT for fields with variances trending to zero.

A Appendix: Cardinalities of basic sets on irregular lattices in $\mathbb{R}^d$

This Appendix contains a series of calculations for the cardinalities of basic sets in D that will be used in the proof of the limit theorems. For any $i = (i_1, \ldots, i_d) \in \mathbb{R}^d$ let

$$\begin{aligned}(i, i+1] &= (i_1, i_1+1] \times ... \times (i_d, i_d+1], \\ [i, i+1] &= [i_1, i_1+1] \times ... \times [i_d, i_d+1],\end{aligned}$$

denote, respectively, the half-open and closed unitary cubes with "south-west" corner i. Note that given the metric, $[i, i+1] = B(j, 1/2)$, where $j = (i_1 + 1/2, ..., i_d + 1/2)$.

Lemma A.1 *Suppose that Assumption 1 holds. Then,*

(i) *Any unitary cube $B(i, 1/2)$ with $i \in \mathbb{R}^d$ contains at most one element of D, i.e., $|B(i, 1/2) \cap D| \leq 1$.*

(ii) *There exists a constant $C < \infty$ such that for $h \geq 1$*

$$\sup_{i \in R^d} |B(i, h) \cap D| \leq Ch^d,$$

i.e., the number of elements of D contained in a ball of radius h centered at $i \in \mathbb{R}^d$ is $O(h^d)$ uniformly in i.

(iii) *For $m \geq 1$ and $i \in \mathbb{R}^d$ let*

$$N_i(1, 1, m) = |\{j \in D : m \leq \rho(i, j) < m + 1\}|$$

be the number of all elements of D located at any distance $h \in [m, m+1)$ from i. Then, there exists a constant $C < \infty$ such that

$$\sup_{i \in R^d} N_i(1, 1, m) \leq Cm^{d-1}.$$

(iv) *Let U and V be some finite disjoint subsets of D. For $m \geq 1$ and $i \in U$ let*

$$\begin{aligned}N_i(2, 2, m) = &|\{(A, B) : |A| = 2, |B| = 2, A \subseteq U \text{ with } i \in A, \\ &B \subseteq V \text{ and } \exists\, j \in B \text{ with } m \leq \rho(i, j) < m + 1\}|\end{aligned}$$

be the number of all different combinations of subsets of U composed of two elements, one of which is i, and subsets of V composed of two elements, where for at least for one of the elements, say j, we have $m \leq \rho(i, j) < m + 1$. Then there exists a constant $C < \infty$ such that

$$\sup_{i \in U} N_i(2, 2, m) \leq Cm^{d-1} |U| |V|.$$

(v) *Let V be some finite subset of D. For $m \geq 1$ and $i \in \mathbb{R}^d$ let*

$$N_i(1,3,m) = |\{B : |B| = 3,\ B \subseteq V \text{ and } \exists\, j \in B \text{ with } m \leq \rho(i,j) < m+1\}|$$

be the number of the subsets of V composed of three elements, at least one of which is located at a distance $h \in [m, m+1)$ from i. Then there exists a constant $C < \infty$ such that

$$\sup_{i \in R^d} N_i(1,3,m) \leq C m^{d-1} |V|^2 .$$

Proof of Lemma A.1(i). We prove it by contradiction. Suppose that there is a unitary cube $B(i, 1/2)$ contains two elements of D, say, x and y. Then $\rho(x,i) \leq 1/2$ and $\rho(y,i) \leq 1/2$. Using the triangle inequality yields:

$$\rho(x,y) \leq \rho(x,i) + \rho(i,y) \leq 1/2 + 1/2 = 1 < d_0,$$

which contradicts Assumption 1. ■

Proof of Lemma A.1(ii). First, observe that for any $i \in \mathbb{R}^d$ and $h \geq 1$, we have $B(i,h) \subseteq B(i,[h]+1)$, where $[h]$ denotes the largest integer less than or equal to h. Note that $B(i,[h]+1)$ is a d-dimensional cube with sides of length $2[h]+2$. Clearly, $B(i,[h]+1)$ can be partitioned into $(2[h]+2)^d$ closed an half-open unitary cubes. Hence, in light of Lemma A.1(i)

$$\begin{aligned} |B(i,h) \cap D| &\leq |B(i,[h]+1) \cap D| \leq (2[h]+2)^d \\ &\leq 2^d (h+1)^d \leq C h^d \end{aligned}$$

with $C = 2^{2d+1} > 0$ observing that $h \geq 1$. Since C depends only on d and not on i, it follows that $\sup_{i \in R^d} |B(i,h) \cap D| \leq C h^d$. ■

Proof of Lemma A.1(iii). Consider the annulus $A(i,m) = \{j \in \mathbb{R}^d : m \leq \rho(i,j) < m+1\}$ of width 1, then

$$A(i,m) \subset B(i,m+1) \backslash B(i,m-1)$$

(If $m = 1$, the ball $B(i,m-1)$ collapses into a point.) Now observe that $B(i,m+1)$ is composed of exactly $[2\,(m+1)]^d$ closed an half-open unitary cubes, and $B(i,m-1)$ is composed of exactly $[2\,(m-1)]^d$ unitary cubes. Hence, the number of unitary cubes making up $B(i,m+1) \backslash B(i,m-1)$ is given by

$$\begin{aligned} & 2^d \left[(m+1)^d - (m-1)^d\right] = 2^d \left[\sum_{s=0}^{d} \binom{d}{s} m^{d-s} - \sum_{s=0}^{d} \binom{d}{s} m^{d-s} (-1)^s\right] \\ \leq\ & 2^{d+1} \left[m^{d-1} \sum_{s=1}^{d} \binom{d}{s} m^{-s+1}\right] \leq 2^{d+1} \left[\sum_{s=1}^{d} \binom{d}{s}\right] m^{d-1} \leq C m^{d-1} \end{aligned}$$

for some $C > 0$ that does not depend on i observing that $m^{-s+1} \leq 1$ for $s \geq 1$. By Lemma A.1(ii), we have

$$\begin{aligned} N_i(1,1,m) &= |\{j \in D : m \leq \rho(i,j) < m+1\}| \\ &= |A(i,m) \cap D| \leq |B(i,m+1) \backslash B(i,m-1)| \leq Cm^{d-1}, \end{aligned}$$

and hence $\sup_{i \in R^d} N_i(1,1,m) \leq Cm^{d-1}$. ■

Proof of Lemma A.1(iv). By Lemma A.1(iii), the number of the one-element subsets of V located at some distance $h \in [m, m+1)$ from $i \in U$ is less than or equal to $N_i(1,1,m) \leq Cm^{d-1}$, $C < \infty$. For each point $j \in V$ one can form at most $|V|$ different two-elements subsets of V that contain j. Thus, the number of the two-element subsets of V that have at least one element located at some distance $h \in [m, m+1)$ from i is less than or equal to $N_i(1,1,m)\,|V| \leq Cm^{d-1}\,|V|$. Furthermore, one can form at most $|U|$ different two-element subsets of U that include i. Hence, $N_i(2,2,m) \leq N_i(1,1,m)\,|V|\,|U| \leq Cm^{d-1}\,|V|\,|U|$. Thus, $\sup_{i \in U} N_i(2,2,m) \leq Cm^{d-1}\,|U|\,|V|$, where C does not depend on i. ■

Proof of Lemma A.1(v). By Lemma A.1(iii), the number of the one-element subsets of V located at some distance $h \in [m, m+1)$ from $i \in \mathbb{R}^d$ is less than or equal to $N_i(1,1,m) \leq Cm^{d-1}$, $C < \infty$. For each point $j \in V$, one can form at most $|V|^2$ different three-elements subsets of V that contain j. Then, the number of the three-element subsets of V that include at least one point located at some distance $h \in [m, m+1)$ from i, obeys: $N_i(1,3,m) \leq N_i(1,1,m)\,|V|^2 \leq Cm^{d-1}\,|V|^2$. Since C does not depend on i furthermore $\sup_{i \in R^d} N_i(1,3,m) \leq Cm^{d-1}\,|V|^2$. ■

B Appendix: Proof of CLT

The proof of Theorem 1 builds on the approach taken by Bolthausen (1982) towards establishing his CLT (for stationary random fields on regular lattices). In particular, rather than using the Bernstein blocking method, we will employ the following lemma to establish asymptotic normality.

Lemma B.1 *(Stein (1972), Bolthausen (1982), Lemma 2). Let $\{\mu_n\}$ be a sequence of probability measures on $(\mathbb{R}, \mathfrak{B})$, where $\mathfrak{B}$ is the Borel σ-field. Suppose the sequence $\{\mu_n\}$ satisfies (with $\mathbf{i}$ denoting the imaginary unit):*

(i) $\sup_n \int y^2 \mu_n(dy) < \infty$; *and*

(ii) $\lim_{n\to\infty} \int (\mathbf{i}\lambda - y) \exp(\mathbf{i}\lambda y) \mu_n(dy) = 0$ *for all* $\lambda \in \mathbb{R}$.

Then $\mu_n \implies N(0,1)$.

As part of the proof, we will also show that it suffices to establish the convergence of the normalized sums for bounded random variables. To that effect, we will utilize the following lemma.

Lemma B.2 *(Brockwell and Davis (1991), Proposition 6.3.9). Let Y_n, $n = 1, 2, \ldots$ and V_{nk}, $k = 1, 2, \ldots$; $n = 1, 2, \ldots$, be random vectors such that*

(i) $V_{nk} \Longrightarrow V_k$ *as* $n \to \infty$ *for each* $k = 1, 2, \ldots,$;

(ii) $V_k \Longrightarrow V$ *as* $k \to \infty$, *and*

(iii) $\lim_{k\to\infty} \limsup_{n\to\infty} P(|Y_n - V_{nk}| > \varepsilon) = 0$ *for every* $\varepsilon > 0$.

Then $Y_n \Longrightarrow V$ *as* $n \to \infty$.

Proof of Theorem 1. We give the proof for α-mixing fields. The argument for ϕ-mixing fields is analogous. The proof is lengthy, and for readability we break it up into several steps.

1. Notation and Reformulation. Consider

$$X_{i,n} = Z_{i,n}/M_n$$

where $M_n = \max_{i\in D_n} c_{i,n}$ is as in Assumption 5. Let $\sigma^2_{n,Z} = Var\left[\sum_{i\in D_n} Z_{i,n}\right]$ and $\sigma^2_{n,X} = Var\left[\sum_{i\in D_n} X_{i,n}\right] = M_n^{-2}\sigma^2_{n,Z}$. Since

$$\sigma^{-1}_{n,X} \sum_{i\in D_n} X_{i,n} = \sigma^{-1}_{n,Z} \sum_{i\in D_n} Z_{i,n},$$

to prove the theorem, it suffices to show that $\sigma_{n,X}^{-1} \sum_{i \in D_n} X_{i,n} \Longrightarrow N(0,1)$. In light of this, it proves convenient to switch notation from the text and to define

$$S_n = \sum_{i \in D_n} X_{i,n}, \qquad \sigma_n^2 = Var(S_n).$$

That is, in the following, S_n denotes $\sum_{i \in D_n} X_{i,n}$ rather than $\sum_{i \in D_n} Z_{i,n}$, and σ_n^2 denotes the variance of $\sum_{i \in D_n} X_{i,n}$ rather than of $\sum_{i \in D_n} Z_{i,n}$.

We next establish the moment and mixing conditions for $X_{i,n}$ implied by the assumptions of the CLT. Observe that by definition of M_n

$$\mathbf{1}(|X_{i,n}| > k) = \mathbf{1}(|Z_{i,n}/M_n| > k) \leq \mathbf{1}(|Z_{i,n}/c_{i,n}| > k),$$

and hence

$$E[|X_{i,n}|^{2+\delta} \mathbf{1}(|X_{i,n}| > k)] \leq E[|Z_{i,n}/c_{i,n}|^{2+\delta} \mathbf{1}(|Z_{i,n}/c_{i,n}| > k)].$$

Thus in light of Assumption 2,

$$\lim_{k \to \infty} \sup_n \sup_{i \in D_n} E[|X_{i,n}|^{2+\delta} \mathbf{1}(|X_{i,n}| > k)] = 0. \tag{B.1}$$

Clearly, the mixing coefficients for $X_{i,n}$ and $Z_{i,n}$ are identical, and hence Assumptions 3 also covers the $X_{i,n}$ process.

In light of our change in notation, Assumption 5 implies:

$$\lim \inf_{n \to \infty} |D_n|^{-1} \sigma_n^2 > 0. \tag{B.2}$$

2. Truncated Random Variables. In proving the CLT, we will consider truncated versions of the $X_{i,n}$. For $k > 0$ we define

$$X_{i,n}^k = X_{i,n} \mathbf{1}(|X_{i,n}| \leq k), \qquad \widetilde{X}_{i,n}^k = X_{i,n} \mathbf{1}(|X_{i,n}| > k),$$

and the corresponding variances as

$$\sigma_{n,k}^2 = Var\left[\sum_{i \in D_n} X_{i,n}^k\right], \qquad \widetilde{\sigma}_{n,k}^2 = Var\left[\sum_{i \in D_n} \widetilde{X}_{i,n}^k\right].$$

Since by (B.1) the $X_{i,n}$ are uniformly $L_{2+\delta}$ integrable, they are also uniformly $L_{2+\delta}$ bounded. Let

$$\|X\|_{2+\delta} = \sup_n \sup_{i \in D} \|X_{i,n}\|_{2+\delta},$$

then we have the following

$$\left\|X_{i,n}^k\right\|_{2+\delta} \leq \|X\|_{2+\delta} \text{ and } \left\|\widetilde{X}_{i,n}^k\right\|_{2+\delta} \leq \|X\|_{2+\delta}.$$

Furthermore, by (B.1)

$$\lim_{k\to\infty}\sup_n\sup_{i\in D}\|\widetilde{X}_{i,n}^k\|_{2+\delta}=\left\{\lim_{k\to\infty}\sup_n\sup_{i\in D}E\left|X_{i,n}\right|^{2+\delta}\mathbf{1}(|X_{i,n}|>k)\right\}^{1/(2+\delta)}=0. \tag{B.3}$$

3. Bounds and Limits for Variances and Variance Ratios. Using the mixing inequality of Lemma 1(i) with $\overline{k}=\overline{l}=1$, $p=q=2+\delta$, and $r=(2+\delta)/\delta$ gives:

$$|cov(X_{i,n},X_{j,n})|\leq 8\bar{\alpha}^{\delta/(2+\delta)}(1,1,\rho(i,j))\|X\|_{2+\delta}^2 \tag{B.4}$$

Since $X_{i,n}^k$ and $\widetilde{X}_{i,n}^k$ are measurable functions of $X_{i,n}$, their covariances and cross-covariances satisfy the same inequality.

We next derive bounds for σ_n^2. Let $K_1=\|X\|_{2+\delta}<\infty$ and observe that $K_2=\sum_{m\geq 1}m^{d-1}\bar{\alpha}^{\delta/(2+\delta)}(1,1,m)<\infty$ in light of Assumption 3(a). Utilizing Lemma A.1(iii), (B.4) and Lyapunov's inequality yields:

$$\begin{aligned}
\sigma_n^2 &\leq \sum_{i\in D_n}EX_{i,n}^2+\sum_{i,j\in D_n,j\neq i}|cov(X_{i,n},X_{j,n})| \qquad \text{(B.5)}\\
&\leq \sum_{i\in D_n}EX_{i,n}^2+8\sum_{i,j\in D_n,j\neq i}\bar{\alpha}^{\frac{\delta}{2+\delta}}(1,1,\rho(i,j))\|X\|_{2+\delta}^2\\
&\leq |D_n|\|X\|_{2+\delta}^2+8\|X\|_{2+\delta}^2\sum_{i\in D_n}\sum_{m=1}^{\infty}\sum_{j\in D_n:\rho(i,j)\in[m,m+1)}\bar{\alpha}^{\frac{\delta}{2+\delta}}(1,1,\rho(i,j))\\
&\leq |D_n|\|X\|_{2+\delta}^2+8\|X\|_{2+\delta}^2\sum_{i\in D_n}\sum_{m=1}^{\infty}N_i(1,1,m)\bar{\alpha}^{\frac{\delta}{2+\delta}}(1,1,m)\\
&\leq |D_n|\|X\|_{2+\delta}^2+8C\|X\|_{2+\delta}^2\sum_{i\in D_n}\sum_{m=1}^{\infty}m^{d-1}\bar{\alpha}^{\frac{\delta}{2+\delta}}(1,1,m)\\
&\leq |D_n|\left[1+8C\sum_{m=1}^{\infty}m^{d-1}\bar{\alpha}^{\frac{\delta}{2+\delta}}(1,1,m)\right]K_1^2\leq |D_n|B_2
\end{aligned}$$

with $B_2=[1+8CK_2]K_1^2<\infty$. In establishing the above inequality we also used the fact that for $\rho(i,j)\in[m,m+1)$: $\bar{\alpha}(1,1,\rho(i,j))\leq\bar{\alpha}(1,1,m)$.

Thus, $\limsup_n|D_n|^{-1}\sigma_n^2<\infty$. By condition (B.2)

$$\lim_{n\to\infty}\inf_{l\geq n}|D_l|^{-1}\sigma_l^2>0$$

and hence there exists an N_* and $B_1>0$ such that for all $n\geq N_*$, we have $B_1|D_n|\leq\sigma_n^2$. Combining the last two inequalities yields for $n\geq N_*$:

$$B_1|D_n|\leq\sigma_n^2\leq B_2|D_n|, \tag{B.6}$$

where $0<B_1\leq B_2<\infty$.

Using analogous arguments, one can bound the variances and covariances of $\sum_{D_n} X_{i,n}^k$, $\sum_{D_n} \widetilde{X}_{i,n}^k$ for each $k > 0$, as follows:

$$\sigma_{n,k}^2 = Var\left[\sum_{D_n} X_{i,n}^k\right] \leq B_2|D_n|,$$

$$\widetilde{\sigma}_{n,k}^2 = Var\left[\sum_{D_n} \widetilde{X}_{i,n}^k\right] \leq |D_n|B_2'\left[\sup_n \sup_{i\in D_n} \|\widetilde{X}_{i,n}^k\|_{2+\delta}\right]^2,$$

$$\left|cov\left\{\sum_{i\in D_n} X_{i,n}^k, \sum_{i\in D_n} \widetilde{X}_{i,n}^k\right\}\right| \leq |D_n|B_2''\left[\sup_n \sup_{i\in D_n} \|\widetilde{X}_{i,n}^k\|_{2+\delta}\right],$$

where $B_2' = [1+8CK_2] < \infty$ and $B_2'' = [2+8CK_2]\,K_1 < \infty$. Furthermore,

$$\begin{aligned}
\sigma_n^2 - \sigma_{n,k}^2 &= 2cov\left\{\sum_{i\in D_n} X_{i,n}^k, \sum_{i\in D_n} \widetilde{X}_{i,n}^k\right\} + \widetilde{\sigma}_{n,k}^2 \\
&\leq 2|D_n|B_2''\left[\sup_n \sup_{i\in D_n} \|\widetilde{X}_{i,n}^k\|_{2+\delta}\right] + |D_n|B_2'\left[\sup_n \sup_{i\in D_n} \|\widetilde{X}_{i,n}^k\|_{2+\delta}\right]^2
\end{aligned}$$

In light of (B.1), (B.6) and the above inequalities we have:

$$0 \leq \frac{\sigma_{n,k}^2}{\sigma_n^2} \leq \frac{B_2}{B_1} < \infty \text{ for all } n \geq N_* \text{ and all } k, \tag{B.7}$$

$$\begin{aligned}
\lim_{k\to\infty} \lim_{n\to\infty}\sup \frac{\widetilde{\sigma}_{n,k}^2}{\sigma_n^2} &\leq \lim_{k\to\infty} \lim_{n\to\infty}\sup \left\{\frac{B_2'}{B_1}\left[\sup_n \sup_{i\in D} \|\widetilde{X}_{i,n}^k\|_{2+\delta}\right]^2\right\} \qquad \text{(B.8)} \\
&= \frac{B_2'}{B_1}\left[\lim_{k\to\infty} \sup_n \sup_{i\in D} \|\widetilde{X}_{i,n}^k\|_{2+\delta}\right]^2 = 0.
\end{aligned}$$

and

$$\begin{aligned}
&\lim_{k\to\infty} \sup_{n\geq N_*} \left|\frac{\sigma_n^2 - \sigma_{n,k}^2}{\sigma_n^2}\right| \qquad \text{(B.9)} \\
\leq\; &\frac{2B_2''}{B_1}\left[\lim_{k\to\infty} \sup_n \sup_{i\in D} \|\widetilde{X}_{i,n}^k\|_{2+\delta}\right] + \frac{B_2'}{B_1}\left[\lim_{k\to\infty} \sup_n \sup_{i\in D} \|\widetilde{X}_{i,n}^k\|_{2+\delta}\right]^2 = 0.
\end{aligned}$$

4. Truncation Technique.[18] Our proof employs a truncation argument in conjunction with Lemma B.2. For $k > 0$ consider the decomposition

$$Y_n = \sigma_n^{-1} \sum_{i\in D_n} X_{i,n} = V_{nk} + (Y_n - V_{nk})$$

[18] We would like to thank Benedikt Pötscher for helpful discussions on this step of the proof.

with

$$V_{nk} = \sigma_n^{-1} \sum_{i \in D_n} (X_{i,n}^k - EX_{i,n}^k), \qquad Y_n - V_{nk} = \sigma_n^{-1} \sum_{D_n} (\widetilde{X}_{i,n}^k - E\widetilde{X}_{i,n}^k),$$

and let $V \sim N(0,1)$. We next show that $Y_n \Longrightarrow N(0,1)$ if

$$\sigma_{n,k}^{-1} \sum_{D_n} (X_{i,n}^k - EX_{i,n}^k) \Longrightarrow N(0,1) \tag{B.10}$$

for each $k = 1, 2, \ldots$ We note that the claim in (B.10) will be verified in subsequent steps.

We first verify condition (iii) of Lemma B.2. By Markov's inequality

$$P(|Y_n - V_{nk}| > \varepsilon) = P(\left|\sigma_n^{-1} \sum_{i \in D_n} (\widetilde{X}_{i,n}^k - E\widetilde{X}_{i,n}^k)\right| > \varepsilon) \leq \frac{\widetilde{\sigma}_{n,k}^2}{\varepsilon^2 \sigma_n^2}.$$

In light of (B.8)

$$\lim_{k\to\infty} \limsup_{n\to\infty} P(|Y_n - V_{nk}| > \varepsilon) \leq \lim_{k\to\infty} \limsup_{n\to\infty} \frac{\widetilde{\sigma}_{n,k}^2}{\varepsilon^2 \sigma_n^2} = 0,$$

which verifies the condition.

Next, observe that

$$V_{nk} = \frac{\sigma_{n,k}}{\sigma_n} \left[\sigma_{n,k}^{-1} \sum_{i \in D_n} (X_{i,n}^k - EX_{i,n}^k) \right].$$

Suppose $r(k) = \lim_{n\to\infty} \sigma_{n,k}/\sigma_n$ exists, then $V_{nk} \Longrightarrow V_k \sim N(0, r(k)^2)$ in light of (B.10). If furthermore, $\lim_{k\to\infty} r(k) \to 1$, then $V_k \Longrightarrow V \sim N(0,1)$, and the claim would follow by Lemma B.2. However, in the case of nonstationary variables $\lim_{n\to\infty} \sigma_{n,k}/\sigma_n$ need not exist, and therefore, we have to use a different argument to show that $Y_n \Longrightarrow V \sim N(0,1)$. We shall prove it by contradiction.

Let $\mathcal{M}$ be the set of all probability measures on $(\mathbb{R},\mathfrak{B})$. Observe that we can metrize $\mathcal{M}$ by, e.g., the Prokhorov distance, say $d(.,.)$. Let μ_n and μ be the probability measures corresponding to Y_n and V, respectively, then $\mu_n \Longrightarrow \mu$ iff $d(\mu_n, \mu) \to 0$ as $n \to \infty$. Now suppose that Y_n does not converge to V. Then for some $\varepsilon > 0$ there exist a subsequence $\{n(m)\}$ such that $d(\mu_{n(m)}, \mu) > \varepsilon$ for all $n(m)$. Observe that by (B.7) we have $0 \leq \sigma_{n,k}/\sigma_n \leq C < \infty$ for all $k > 0$ and all $n \geq N_*$, where N_* does not depend on k. W.l.o.g. assume that with $n(m) \geq N_*$, and hence $0 \leq \sigma_{n(m),k}/\sigma_{n(m)} \leq C < \infty$ for all $k > 0$ and all $n(m)$. Consequently, for $k = 1$ there exists a subsubsequence $\{n(m(l_1))\}$ such that such $\sigma_{n(m(l_1)),1}/\sigma_{n(m(l_1))} \to r(1)$ as $l_1 \to \infty$. For $k = 2$ there exists a subsubsubsequence $\{n(m(l_1(l_2)))\}$ such that $\sigma_{n(m(l_1(l_2))),2}/\sigma_{n(m(l_1(l_2)))} \to r(2)$ as $l_2 \to \infty$. The argument can be repeated for $k = 3, 4.....$ Now construct a subsequence $\{n_l\}$ such that n_1 corresponds to the first element of $\{n(m(l_1))\}$,

n_2 corresponds to the second element of $\{n(m(l_1(l_2)))\}$, and so on, then for $k = 1, 2, \ldots$,we have:

$$\lim_{l\to\infty} \frac{\sigma_{n_l,k}}{\sigma_{n_l}} = r(k). \tag{B.11}$$

Moreover, since by (B.9)

$$\begin{aligned}
\lim_{k\to\infty} \sup_{n\geq N_*} \left|1 - \frac{\sigma_{n,k}}{\sigma_n}\right| &\leq \lim_{k\to\infty} \sup_{n\geq N_*} \left|1 - \frac{\sigma_{n,k}}{\sigma_n}\right| \left|1 + \frac{\sigma_{n,k}}{\sigma_n}\right| \\
&= \lim_{k\to\infty} \sup_{n\geq N_*} \left|\frac{\sigma_n^2 - \sigma_{n,k}^2}{\sigma_n^2}\right| = 0
\end{aligned}$$

and

$$\begin{aligned}
|r(k) - 1| &= \left|r(k) - \frac{\sigma_{n_l,k}}{\sigma_{n_l}} + \frac{\sigma_{n_l,k}}{\sigma_{n_l}} - 1\right| \\
&\leq \left|r(k) - \frac{\sigma_{n_l,k}}{\sigma_{n_l}}\right| + \sup_{n_l\geq N_*} \left|\frac{\sigma_{n_l,k}}{\sigma_{n_l}} - 1\right|,
\end{aligned}$$

it follows from (B.11) that

$$\begin{aligned}
\lim_{k\to\infty} |r(k) - 1| &= \lim_{k\to\infty} \lim_{l\to\infty} |r(k) - 1| \qquad \text{(B.12)} \\
&\leq \lim_{k\to\infty} \lim_{l\to\infty} \left|r(k) - \frac{\sigma_{n_l,k}}{\sigma_{n_l}}\right| + \lim_{k\to\infty} \sup_{n\geq N_*} \left|\frac{\sigma_{n,k}}{\sigma_n} - 1\right| = 0.
\end{aligned}$$

Given (B.12), it follows that $V_{n_l k} \Longrightarrow V_k \sim N(0, r(k)^2)$.Then, by Lemma B.2, $Y_{n_l} \Longrightarrow V \sim N(0,1)$ as $l \to \infty$. Since $\{n_l\} \subseteq \{n(m)\}$, this contradicts the hypothesis that $d(\mu_{n(m)}, \mu) > \varepsilon$ for all $n(m)$.

Thus, we have shown that $Y_n \Longrightarrow N(0,1)$ if (B.10) holds. In light of this it suffices to prove the CLT for bounded variables.[19] In the following, we assume that $|X_{i,n}|| \leq C_X < \infty$.

5. Renormalization. Since $|D_n| \to \infty$ and $\bar{\alpha}(1, \infty, m) = O(m^{-d-\varepsilon})$ it is readily seen that we can choose a sequence m_n such that

$$\bar{\alpha}(1, \infty, m_n)|D_n|^{1/2} \to 0 \tag{B.13}$$

and

$$m_n^d |D_n|^{-1/2} \to 0 \tag{B.14}$$

as $n \to \infty$. Now, for such m_n define:

[19] Guyon (1995), p. 112, gives a CLT for non-stationary non-trending random fields on $\mathbb{Z}^d$. The proof in essence asserts, without giving detailed arguments, that (B.8) holds and that consequently it is sufficient to consider only the case of bounded random variables, while maintaining only $L_{2+\delta}$-boundedness. Our arguments verify the assertion, provided that $L_{2+\delta}$-boundedness is strengthened to $L_{2+\delta}$-uniform integrability.

$$\begin{aligned} a_n &= \sum_{i,j \in D_n, \rho(i,j) \leq m_n} E(X_{i,n}X_{j,n}), \\ b_n &= \sum_{i,j \in D_n, \rho(i,j) > m_n} E(X_{i,n}X_{j,n}), \end{aligned}$$

so that
$$\sigma_n^2 = Var(S_n) = \sum_{i,j \in D_n} E(X_{i,n}X_{j,n}) = a_n + b_n$$

Using the mixing inequality of Lemma 1(iii) with $\overline{k} = \overline{l} = 1$, Lemma A.1(ii), and argumentation analogous to that used in (B.5) yields

$$|b_n| \leq \sum_{i,j \in D_n, \rho(i,j) > m_n} |cov(X_{i,n}X_{j,n})| \leq 4CC_X^2 |D_n| \sum_{l=m_n}^{\infty} l^{d-1}\bar{\alpha}(1,1,l).$$

Since Assumption 3b implies $\sum_{l=m_n}^{\infty} l^{d-1}\bar{\alpha}(1,1,l) \to 0$ as $n \to \infty$, it follows that $b_n = o(|D_n|)$. Moreover, by (B.2) we have

$$\begin{aligned} &\liminf_{n\to\infty} |D_n|^{-1} a_n \\ \geq\ &\liminf_{n\to\infty} |D_n|^{-1}\sigma_n^2 + \liminf_{n\to\infty} \left\{ -|D_n|^{-1} b_n \right\} = \liminf_{n\to\infty} |D_n|^{-1}\sigma_n^2 > 0. \end{aligned}$$

Hence, for some $0 < B_1 < \infty$ and sufficiently large n we have $0 < B_1|D_n| < a_n$. From the inequalities established in (B.5) it follows furthermore that $|a_n| \leq \sum_{i,j \in D_n, \rho(i,j) \leq m_n} |cov(X_{i,n}, X_{j,n})| \leq B_2|D_n|$. Hence, for sufficiently large n, say $n \geq N_{**} \geq N_*$:

$$0 < B_1|D_n| \leq a_n \leq B_2|D_n|, \qquad 0 < B_1 \leq B_2 < \infty, \tag{B.15}$$

i.e., $a_n \sim |D_n|$ and, consequently,

$$\sigma_n^2 = a_n + o(|D_n|) = a_n + o(a_n) = a_n(1 + o(1)).$$

For $n \geq N_{**}$ define

$$\bar{S}_n = a_n^{-1/2} S_n = a_n^{-1/2} \sum_{i \in D_n} X_{i,n}.$$

To demonstrate that $\sigma_n^{-1} S_n \Longrightarrow N(0,1)$, it now suffices to show that $\bar{S}_n \Longrightarrow N(0,1)$.

6. Limiting Distribution of $\bar{S}_n$: From the above discussion $\sup_{n \geq N_{**}} E\bar{S}_n^2 < \infty$. In light of Lemma B.1, to establish that $\bar{S}_n \Longrightarrow N(0,1)$, it suffices to show that

$$\lim_{n\to\infty} E[(\mathbf{i}\lambda - \bar{S}_n)\exp(\mathbf{i}\lambda\bar{S}_n)] = 0$$

In the following, we take $n \geq N_{**}$, but will not indicate that explicitly for notational simplicity. Define

$$S_{j,n} = \sum_{i \in D_n, \rho(i,j) \leq m_n} X_{i,n} \qquad \text{and} \qquad \bar{S}_{j,n} = a_n^{-1/2} S_{j,n},$$

then

$$(\mathbf{i}\lambda - \bar{S}_n)\exp(\mathbf{i}\lambda \bar{S}_n) = A_{1,n} - A_{2,n} - A_{3,n},$$

with

$$\begin{aligned}
A_{1,n} &= \mathbf{i}\lambda e^{\mathbf{i}\lambda \bar{S}_n}(1 - a_n^{-1} \sum_{j \in D_n} X_{j,n} S_{j,n}), \\
A_{2,n} &= a_n^{-1/2} e^{\mathbf{i}\lambda \bar{S}_n} \sum_{j \in D_n} X_{j,n}[1 - \mathbf{i}\lambda \bar{S}_{j,n} - e^{-\mathbf{i}\lambda \bar{S}_{j,n}}], \\
A_{3,n} &= a_n^{-1/2} \sum_{j \in D_n} X_{j,n} e^{\mathbf{i}\lambda(\bar{S}_n - \bar{S}_{j,n})}.
\end{aligned}$$

To complete the proof we show that $E|A_{i,n}| \to 0$ as $n \to \infty$ for $i = 1, 2, 3$.

7. Proof that $E|A_{1,n}| \to 0$: Note that

$$\begin{aligned}
|A_1|^2 &= \left| i\lambda e^{i\lambda \bar{S}_n} \right| \left(1 - a_n^{-1} \sum_{j \in D_n} X_{j,n} S_{j,n} \right)^2 \\
&= \lambda^2 \left\{ 1 - 2a_n^{-1} \sum_{j \in D_n} X_{j,n} S_{j,n} + a_n^{-2} \left[\sum_{j \in D_n} X_{j,n} S_{j,n} \right]^2 \right\}
\end{aligned}$$

and hence, observing that $a_n = E \sum_{j \in D_n} X_{j,n} S_{j,n}$,

$$\begin{aligned}
E|A_1|^2 &= \lambda^2 \left\{ 1 - 2a_n^{-1} \sum_{j \in D_n} E X_{j,n} S_{j,n} \right. \\
&\quad \left. + a_n^{-2} \left[var\left(\sum_{j \in D_n} X_{j,n} S_{j,n} \right) + \left(\sum_{j \in D_n} E X_{j,n} S_{j,n} \right)^2 \right] \right\} \\
&= \lambda^2 \left\{ 1 - 2a_n^{-1} a_n + a_n^{-2} \left[var\left(\sum_{j \in D_n} X_{j,n} S_{j,n} \right) + a_n^2 \right] \right\} \\
&= \lambda^2 a_n^{-2} var\left(\sum_{j \in D_n} X_{j,n} S_{j,n} \right) = \lambda^2 a_n^{-2} var\left(\sum_{\substack{i \in D_n, j \in D_n \\ \rho(i,j) \leq m_n}} X_{i,n} X_{j,n} \right) \\
&= \lambda^2 a_n^{-2} \sum_{\substack{i \in D_n, j \in D_n, i' \in D_n, j' \in D_n \\ \rho(i,j) \leq m_n, \rho(i',j') \leq m_n}} cov\left(X_{i,n} X_{j,n}; X_{i',n} X_{j',n} \right).
\end{aligned}$$

By (B.15), we have

$$
\begin{aligned}
E|A_1|^2 &\leq C_*|D_n|^{-2} \sum_{\substack{i,j,i',j' \in D_n \\ \rho(i,j) \leq m_n, \rho(i',j') \leq m_n}} |cov\,(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})| \qquad \text{(B.16)} \\
&= C_*|D_n|^{-2} \sum_{\substack{i,j,i',j' \in D_n \\ \rho(i,j) \leq m_n, \rho(i',j') \leq m_n, \rho(i,i') \geq 3m_n}} |cov\,(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})| \\
&\quad + C_*|D_n|^{-2} \sum_{\substack{i,j,i',j' \in D_n \\ \rho(i,j) \leq m_n, \rho(i',j') \leq m_n, \rho(i,i') < 3m_n}} |cov\,(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})|\,,
\end{aligned}
$$

for some $C_* < \infty$. We next obtain bounds for the above inner sums for fixed $i \in D_n$ corresponding to $\rho(i,i') \geq 3m_n$ and $\rho(i,i') < 3m_n$, respectively.

7(a) First consider the case where $r = \rho(i,i') \geq 3m_n$. Since $\rho(i,j) \leq m_n$ and $\rho(i',j') \leq m_n$, clearly $\rho(i,j') \geq r - 2m_n$, $\rho(j,i') \geq r - 2m_n$ and $\rho(j,j') \geq r - 2m_n$. Take $U = \{i,j\}$ and $V = \{i',j'\}$, then $\rho(U,V) \geq r - 2m_n \geq 1$. Since $|X_{j,n}| \leq C_X$, using the first inequality of Lemma 1(iii) with $\overline{k} = \overline{l} = 2$, and observing that $\bar{\alpha}(\overline{k},\overline{l},\overline{h})$ is nonincreasing in $\overline{h}$ yields

$$
|cov\,(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})\,| \leq 4C_X^4 \bar{\alpha}(2,2,r-2m_n). \qquad \text{(B.17)}
$$

Now define $N_i(2,2,l)$ as the number of all different combinations consisting of subsets of $\{j : \rho(i,j) \leq m_n\}$ composed of two elements, one of which is i, and subsets of $\{j' : \rho(i',j') \leq m_n\}$ composed of two elements, one of which is i', where $\rho(i,i') \geq 3m_n$ and $l \leq \rho(i,i') < l+1$, $l \in \mathbb{N}$, i.e.,

$$
\begin{aligned}
N_i(2,2,l) \;\; = \;\; & |\{(A,B) : |A| = 2, |B| = 2, A \subseteq \{j : \rho(i,j) \leq m_n\} \text{ with } i \in A, \\
& B \subseteq \{j' : \rho(i',j') \leq m_n\} \text{ with } i' \in B \text{ and } 3m_n \leq l \leq \rho(i,i') < l+1\}|
\end{aligned}
$$

By Lemmata A.1(iv) and A.1(ii)

$$
\begin{aligned}
\sup_{i \in R^d} N_i(2,2,l) &\leq M l^{d-1}\, |\{j : \rho(i,j) \leq m_n\}|\, |\{j' : \rho(i',j') \leq m_n\}| \\
&\leq M_* m_n^{2d} l^{d-1} \qquad \text{(B.18)}
\end{aligned}
$$

for some $M < \infty$ and $M_* < \infty$. Note that if $l \leq r < l+1$, then $\bar{\alpha}(2,2,r-2m_n) \leq \bar{\alpha}(2,2,l-2m_n)$.

In light of (B.17) and (B.18), we now have for fixed $i \in D_n$:

$$
\begin{aligned}
& \sum_{\substack{j,i',j' \in D_n \\ \rho(i,j)\le m_n, \rho(i',j')\le m_n, \rho(i,i')\ge 3m_n}} |cov\,(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})| \quad \text{(B.19)} \\
\le\ & 4C_X^4 \left[\sum_{l=3m_n}^{\infty} N_i(2,2,l)\bar{\alpha}(2,2,l-2m_n) \right] \\
\le\ & 4C_X^4 M_* m_n^{2d} \sum_{l=3m_n}^{\infty} l^{d-1}\bar{\alpha}(2,2,l-2m_n) \\
\le\ & 3^{d-1} 4C_X^4 M_* m_n^{2d} \sum_{l=m_n}^{\infty} l^{d-1}\bar{\alpha}(2,2,l) \le C_1 m_n^{2d}
\end{aligned}
$$

for some $C_1 < \infty$.

7(b) Next consider the case where $r = \rho(i,i') < 3m_n$. Let $V_i = \{x \in D_n : \rho(x,i) \le 4m_n\}$ be the collection of the elements of D_n contained in the ball of the radius $4m_n$ centered in i. This set will necessarily include all points i', j, j' such that $\rho(i,i') < 3m_n$, $\rho(i,j) \le m_n$, and $\rho(i',j') \le m_n$. Further, let

$$h(j,i',j') = \min\{\rho(i,i'), \rho(i,j), \rho(i,j')\}\,.$$

Then using the first inequality of Lemma 1(iii) twice, first with $\overline{k} = 1, \overline{l} = 3$, and then with $\overline{k} = \overline{l} = 1$ gives

$$
\begin{aligned}
|cov\,(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})| \;\le\;& |E(X_{i,n}X_{j,n}X_{i',n}X_{j',n})| \quad \text{(B.20)} \\
& + |E(X_{i,n}X_{j,n})||E(X_{i',n}X_{j',n})| \\
\le\;& 4C_X^4\bar{\alpha}(1,3,h_i(j,i',j')) \\
& + 4C_X^4\bar{\alpha}(1,1,h_i(j,i',j'))\bar{\alpha}(1,1,\rho(i',j')) \\
\le\;& 4C_X^4\bar{\alpha}(1,3,h(j,i',j')) + 4C_X^4\bar{\alpha}(1,1,h(j,i',j')) \\
\le\;& 8C_X^4\bar{\alpha}(1,3,h(j,i',j')).
\end{aligned}
$$

observing that $\alpha(\overline{k},\overline{l},\overline{h})$ is less than or equal to one and nondecreasing in $\overline{k},\overline{l}$.

Now, let $W_i(l) = \{A \subseteq V_i : |A| = 3,\, l \le \rho(i,A) < l+1\}$ denote the set of three element subsets of V_i located at distances $\overline{h} \in [l, l+1)$ from i. Clearly, the number of such sets, $|W_i(l)|$ is no greater then $N_i(1,3,l)$, defined in Lemma A.1(v), and by Lemmata A.1(v) and A.1(ii), we have

$$\sup_{i\in R^d} |W_i(l)| \le \sup_{i \in R^d} N_i(1,3,l) \le \overline{M} l^{d-1} (4m_n)^{2d} = \overline{M}_* l^{d-1} m_n^{2d} \quad \text{(B.21)}$$

for some $\overline{M} < \infty$ and $\overline{M}_* = 2^{4d}\overline{M} < \infty$. Using (B.20) and (B.21) we have for fixed $i \in D_n$:

$$\sum_{\substack{j,i',j' \in D_n \\ \rho(i,j)\le m_n, \rho(i',j')\le m_n, \rho(i,i')<3m_n}} |cov\left(X_{i,n}X_{j,n}; X_{i',n}X_{j',n}\right)| \quad \text{(B.22)}$$

$$\begin{aligned}
&\le \sum_{j,i',j' \in V_i} |cov\left(X_{i,n}X_{j,n}; X_{i',n}X_{j',n}\right)| \\
&\le 8C_X^4 \sum_{j,i',j' \in V_i} \bar{\alpha}(1,3,h(j,i',j')) = 8C_X^4 \sum_{l=1}^{4m_n} \sum_{A \in W_i(l)} \bar{\alpha}(1,3,l) \\
&\le 8C_X^4 \overline{M}_* m_n^{2d} \sum_{l=1}^{4m_n} l^{d-1} \bar{\alpha}(1,3,l) \le C_2 m_n^{2d}
\end{aligned}$$

for some $C_2 < \infty$, using Assumption 3(b).

From (B.14), (B.16), (B.19) and (B.22) we have:

$$E|A_1|^2 \le C_* |D_n|^{-2} \sum_{i \in D_n} (C_1 + C_2) m_n^{2d} \le const * |D_n|^{-1} m_n^{2d} \to 0$$

as $n \to \infty$.

8. Proof that $E|A_{2,n}| \to 0$: Observe that by Lemma A.1(ii) and (B.15)

$$\begin{aligned}
|\bar{S}_{j,n}| &= a_n^{-1/2} |S_{j,n}| \le a_n^{-1/2} \sum_{i \in D_n, \rho(i,j) \le m_n} |X_{i,n}| \\
&\le C C_X a_n^{-1/2} m_n^d \le C_4 |D_n|^{-1/2} m_n^d.
\end{aligned}$$

for some $C_4 < \infty$. By (B.14) it follows that $|\bar{S}_{j,n}| \to 0$. Observe further that if z is a complex number with $|z| < 1/2$, then $|1 - z - e^{-z}| \le |z|^2$. Since $|\bar{S}_{j,n}| \to 0$, there exists $N_{***} \ge N_{**}$ such that for $n \ge N_{***}$ we have $|\bar{S}_{j,n}| < 1/2$ a.s. and hence

$$\left|1 - i\lambda \bar{S}_{j,n} - e^{-i\lambda \bar{S}_{j,n}}\right| \le \left|\bar{S}_{j,n}\right|^2 \text{ a.s.}$$

Using this inequality and the same arguments as before gives:

$$
\begin{aligned}
E|A_2| &\leq const * |D_n|^{-1/2} \sum_{j \in D_n} E\bar{S}^2_{j,n} \leq const * |D_n|^{-1/2} |D_n| \sup_{j \in D_n} E(\bar{S}^2_{j,n}) \\
&\leq const * |D_n|^{1/2} a_n^{-1} \sup_{j \in D_n} \sum_{\substack{i,i' \in D_n, \\ \rho(i,j) \leq m_n, \rho(i',j) \leq m_n}} |E(X_{i,n} X_{i',n})| \\
&\leq const * |D_n|^{-1/2} \sup_{j \in D_n} \sum_{\substack{i,i' \in D_n, \\ \rho(i,j) \leq m_n, \rho(i',j) \leq m_n}} \bar{\alpha}(1,1,\rho(i,i')) \\
&\leq const * |D_n|^{-1/2} \sup_{j \in D_n} \sum_{i \in D_n, \rho(i,j) \leq m_n} \sum_{1 \leq l \leq 2m_n} N_i(1,1,l) \bar{\alpha}(1,1,l) \\
&\leq const * |D_n|^{-1/2} m_n^d \sum_{1 \leq l \leq 2m_n} l^{d-1} \bar{\alpha}(1,1,l) \\
&\leq C_5 |D_n|^{-1/2} m_n^d
\end{aligned}
$$

for some $C_5 < \infty$. The last inequality used Assumption 3. Hence, by (B.14), $E|A_2| \to 0$ as $n \to \infty$.

9. Proof that $|EA_{3,n}| \to 0$: Note that

$$
|EA_3| = \left| E a_n^{-1/2} \sum_{j \in D_n} X_{j,n} e^{i\lambda(\bar{S}_n - \bar{S}_{j,n})} \right| \leq const * |D_n|^{-1/2} \sum_{j \in D_n} \left| EX_{j,n} e^{i\lambda(\bar{S}_n - \bar{S}_{j,n})} \right|
$$

and that $e^{i\lambda(\bar{S}_n - \bar{S}_{j,n})}$ is $\sigma(X_{i,n}, \rho(j,i) > m_n)$-measurable. Using the first inequality of Lemma 1(iii) with $\overline{k} = 1, \overline{l} = |D_n|$ gives $\left| EX_{j,n} e^{i\lambda(\bar{S}_n - \bar{S}_{j,n})} \right| \leq 4C_X \bar{\alpha}(1, |D_n|, m_n)$ and hence

$$
\begin{aligned}
|EA_3| &\leq const * |D_n|^{-1/2} |D_n| \bar{\alpha}(1, |D_n|, m_n) \\
&\leq const * |D_n|^{1/2} \bar{\alpha}(1, \infty, m_n) \to 0
\end{aligned}
$$

as $n \to \infty$ by (B.13).

This completes the proof of the CLT. ■

C Appendix: Proofs ULLN and LLN

Proof of Theorem 2: In the following we use the abbreviations ACL_0UEC *[*ACL_pUEC*] [[a.s.ACUEC]]* for L_0 [L_p], *[[a.s.]]* stochastic equicontinuity as defined in Definition 2. We first show that ACL_0UEC and the Domination Assumptions 6 for $g_{i,n}(Z_{i,n},\theta) = q_{i,n}(Z_{i,n},\theta)/c_{i,n}$ jointly imply that the $g_{i,n}(Z_{i,n},\theta)$ is ACL_pUEC, $p \geq 1$.

Given $\varepsilon > 0$, it follows from Assumption 6 that we can choose some $k = k(\varepsilon) < \infty$ such that

$$\limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i \in D_n} E(d_{i,n}^p \mathbf{1}(d_{i,n} > k) < \frac{\varepsilon}{3 \cdot 2^p}. \tag{C.1}$$

Let

$$Y_{i,n}(\delta) = \sup_{\theta' \in \Theta} \sup_{\theta \in B(\theta',\delta)} \left| g_{i,n}(Z_{i,n},\theta) - g_{i,n}(Z_{i,n},\theta') \right|^p,$$

and observe that $Y_{i,n}(\delta) \leq 2^p d_{i,n}^p$, then

$$\begin{aligned} E\left[Y_{i,n}(\delta)\right] &= E\left[Y_{i,n}(\delta))\mathbf{1}(Y_{i,n}(\delta) \leq \varepsilon/3)\right] + E\left[Y_{i,n}(\delta)\mathbf{1}(Y_{i,n}(\delta) > \varepsilon/3)\right] \\ &\leq \varepsilon/3 + EY_{i,n}(\delta)\mathbf{1}(Y_{i,n}(\delta) > \varepsilon/3, d_{i,n} > k) \qquad \text{(C.2)} \\ &+ \quad EY_{i,n}(\delta)\mathbf{1}(Y_{i,n}(\delta) > \varepsilon/3, d_{i,n} \leq k) \\ &\leq \varepsilon/3 + 2^p E d_{i,n}^p \mathbf{1}(d_{i,n} > k) + 2^p k^p P(Y_{i,n}(\delta) > \varepsilon/3) \end{aligned}$$

From the assumption that the $g_{i,n}(Z_{i,n},\theta)$ is ACL_0UEC, it follows that we can find some $\delta = \delta(\varepsilon)$ such that

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i \in D_n} P(Y_{i,n}(\delta) > \varepsilon) \qquad \text{(C.3)} \\ = \; &\limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i \in D_n} P\left(\sup_{\theta' \in \Theta} \sup_{\theta \in B(\theta',\delta)} \left| g_{i,n}(Z_{i,n},\theta) - g_{i,n}(Z_{i,n},\theta') \right| > \varepsilon^{\frac{1}{p}} \right) \\ \leq \; &\frac{\varepsilon}{3\,(2k)^p} \end{aligned}$$

It now follows from (C.1), (C.2) and (C.3) that for $\delta = \delta(\varepsilon)$,

$$\begin{aligned} &\limsup \frac{1}{|D_n|} \sum_{i \in D_n} EY_{i,n}(\delta) \\ \leq \; &\varepsilon/3 + 2^p \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i \in D_n} E d_{i,n}^p \mathbf{1}(d_{i,n} > k) \\ + \; &2^p k^p \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i \in D_n} P(Y_{i,n}(\delta) > \varepsilon/3) \leq \varepsilon, \end{aligned}$$

which implies that $g_{i,n}(Z_{i,n},\theta)$ is ACL_pUEC, $p \geq 1$.

We next show that this in turn implies that $Q_n(\theta)$ is AL_pUEC, $p \geq 1$, as defined in Pötscher and Prucha (1994a), i.e., we show that

$$\limsup_{n\to\infty} E\left\{\sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\left|Q_n(\theta)-Q_n(\theta')\right|^p\right\} \to 0 \text{ as } \delta \to 0.$$

To see this, observe that

$$\begin{aligned}
& E\sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\left|Q_n(\theta)-Q_n(\theta')\right|^p \\
=\ & E\sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\left|\frac{1}{M_n\left|D_n\right|}\sum_{i\in D_n}\left[q_{i,n}(Z_{i,n},\theta)-q_{i,n}(Z_{i,n},\theta')\right]\right|^p \\
\leq\ & E\sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\frac{1}{M_n^p\left|D_n\right|}\sum_{i\in D_n}\left|q_{i,n}(Z_{i,n},\theta)-q_{i,n}(Z_{i,n},\theta')\right|^p \\
\leq\ & \frac{1}{\left|D_n\right|}\sum_{i\in D_n}E\sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\left|q_{i,n}(Z_{i,n},\theta)-q_{i,n}(Z_{i,n},\theta')\right|^p/c_{i,n}^p \\
=\ & \frac{1}{\left|D_n\right|}\sum_{i\in D_n}EY_{i,n}(\delta)
\end{aligned}$$

where we have used inequality (1.4.3) in Bierens (1994). The claim now follows since the lim sup of the last term goes to zero as $\delta \to 0$, as demonstrated above. Moreover, by Theorem 2.1 in Pötscher and Prucha (1994a), $Q_n(\theta)$ is also AL_0UEC, i.e., for every $\varepsilon > 0$

$$\limsup_{n\to\infty} P\left\{\sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\left|Q_n(\theta)-Q_n(\theta')\right| > \varepsilon\right\} \to 0 \text{ as } \delta \to 0.$$

Given the assumed weak pointwise LLN for $Q_n(\theta)$ the i.p. portion of part (a) of the theorem now follows directly from Theorem 3.1(a) of Pötscher and Prucha (1994a).

For the a.s. portion of the theorem, note that by the triangle inequality

$$\begin{aligned}
& \limsup_{n\to\infty}\sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\left|Q_n(\theta)-Q_n(\theta')\right| \\
=\ & \limsup_{n\to\infty}\sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\frac{1}{M_n\left|D_n\right|}\left|\sum_{i\in D_n}q_{i,n}(Z_{i,n},\theta)-q_{i,n}(Z_{i,n},\theta')\right| \\
\leq\ & \limsup_{n\to\infty}\frac{1}{\left|D_n\right|}\sum_{i\in D_n}\sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\left|g_{i,n}(Z_{i,n},\theta)-g_{i,n}(Z_{i,n},\theta')\right|.
\end{aligned}$$

The r.h.s. of the last inequality goes to zero as $\delta \to 0$, since $g_{i,n}$ is $a.s.ACUEC$ by assumption. Therefore,

$$\limsup_{n\to\infty}\sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\left|Q_n(\theta)-Q_n(\theta')\right| \to 0 \text{ as } \delta \to 0 \text{ a.s.}$$

i.e., Q_n is $a.s.AUEC$, as defined in Pötscher and Prucha (1994a). Given the assumed strong pointwise LLN for $Q_n(\theta)$ the a.s. portion of part (a) of the theorem now follows from Theorem 3.1(a) of Pötscher and Prucha (1994a).

Next observe that since $a.s.ACUEC \Longrightarrow ACL_0UEC$ we have that $Q_n(\theta)$ is AL_pUEC, $p \geq 1$, both under the i.p. and a.s. assumptions of the theorem. This in turn implies that $\overline{Q}_n(\theta) = EQ_n(\theta)$ is $AUEC$, by Theorem 3.3 in Pötscher and Prucha (1994a), which proves part (b) of the theorem. ■

Proof of Theorem 3: Define $X_{i,n} = Z_{i,n}/M_n$, and observe that

$$[|D_n|\, M_n]^{-1} \sum_{i \in D_n} (Z_{i,n} - EZ_{i,n}) = |D_n|^{-1} \sum_{i \in D_n} (X_{i,n} - EX_{i,n}) .$$

Hence it suffices to prove the LLN for $X_{i,n}$.

We first establish mixing and moment conditions for $X_{i,n}$ from those for $Z_{i,n}$. Clearly, if $Z_{i,n}$ is α-mixing [ϕ-mixing], then $X_{i,n}$ is also α-mixing [ϕ-mixing] with the same coefficients. Thus, $X_{i,n}$ satisfies Assumption 3b with $k = l = 1$ [Assumption 4b with $k = l = 1$]. Furthermore, observe that by the definition of M_n we have $\mathbf{1}(|X_{i,n}| > k) = \mathbf{1}(|Z_{i,n}/M_n| > k) \leq \mathbf{1}(|Z_{i,n}/c_{i,n}| > k)$, and hence

$$\lim_{k \to \infty} \sup_n \sup_{i \in D_n} E[|X_{i,n}|\, \mathbf{1}(|X_{i,n}| > k)] \leq \lim_{k \to \infty} \sup_n \sup_{i \in D_n} E[|Z_{i,n}/c_{i,n}|\, \mathbf{1}(|Z_{i,n}/c_{i,n}| > k)] = 0, \tag{C.4}$$

i.e., $X_{i,n}$ is also uniformly L_1 integrable.

In proving the LLN we consider truncated versions of $X_{i,n}$. For $0 < k < \infty$ let

$$X_{i,n}^k = X_{i,n}\mathbf{1}(|X_{i,n}| \leq k), \qquad \widetilde{X}_{i,n}^k = X_{i,n} - X_{i,n}^k = X_{i,n}\mathbf{1}(|X_{i,n}| > k).$$

In light of (C.4)

$$\lim_{k \to \infty} \sup_n \sup_{i \in D_n} E\left|\widetilde{X}_{i,n}^k\right| = 0. \tag{C.5}$$

Clearly, $X_{i,n}^k$ is a measurable function of $X_{i,n}$, and thus $X_{i,n}^k$ is also α-mixing [ϕ-mixing] with mixing coefficients not exceeding those of $X_{i,n}$.

By Minkowski's inequality

$$\begin{aligned} & E\left|\sum_{i \in D_n} (X_{i,n} - EX_{i,n})\right| && \text{(C.6)} \\ \leq\ & E\left|\sum_{i \in D_n} \left(X_{i,n} - X_{i,n}^k\right)\right| + E\left|\sum_{i \in D_n} \left(X_{i,n}^k - EX_{i,n}^k\right)\right| + E\left|\sum_{i \in D_n} \left(EX_{i,n}^k - EX_{i,n}\right)\right| \\ \leq\ & 2E\left|\sum_{i \in D_n} \widetilde{X}_{i,n}^k\right| + E\left|\sum_{i \in D_n} \left(X_{i,n}^k - EX_{i,n}^k\right)\right| \end{aligned}$$

and thus

$$\lim_{n\to\infty}\left\| |D_n|^{-1}\sum_{i\in D_n}(X_{i,n}-EX_{i,n})\right\|_1 \tag{C.7}$$
$$\leq 2\lim_{k\to\infty}\sup_n\sup_{i\in D_n}E\left|\widetilde{X}_{i,n}^k\right| + \lim_{k\to\infty}\lim_{n\to\infty}\left\| |D_n|^{-1}\sum_{i\in D_n}\left(X_{i,n}^k-EX_{i,n}^k\right)\right\|_1$$

where $\|.\|_1$ denotes the L_1-norm. The first term on the r.h.s. of (C.7) goes to zero in light of (C.5). To complete the prove we now demonstrate that also the second term converges to zero. To that effect it suffices to show that $X_{i,n}^k$ satisfies a an L_1-norm LLN for fixed k.

Let $\sigma_{n,k}^2 = Var\left[\sum_{i\in D_n}X_{i,n}^k\right]$, then by Lyapunov's inequality

$$\left\| |D_n|^{-1}\sum_{i\in D_n}\left(X_{i,n}^k-EX_{i,n}^k\right)\right\|_1 \leq |D_n|^{-1}\sigma_{n,k}. \tag{C.8}$$

Using Lemma A.1(iii) and Lemma 1(iii), we have in the α-mixing case:

$$\begin{aligned}
\sigma_{n,k}^2 &\leq \sum_{i\in D_n}Var(X_{i,n}^k) + \sum_{\substack{i\in D_n, j\in D_n\\ j\neq i}}\left|Cov(X_{i,n}^k;X_{j,n}^k)\right| \\
&\leq 2k^2|D_n| + 4k^2\sum_{\substack{i\in D_n, j\in D_n\\ j\neq i}}\overline{\alpha}_X\left(1,1,\rho(i,j)\right) \\
&\leq 2k^2|D_n| + 4k^2\sum_{i\in D_n}\sum_{m=1}^{\infty}\sum_{j\in D_n:\rho(i,j)\in[m,m+1)}\overline{\alpha}_X\left(1,1,\rho(i,j)\right) \\
&\leq 2k^2|D_n| + 4k^2\sum_{i\in D_n}\sum_{m=1}^{\infty}N_i(1,1,m)\overline{\alpha}_X\left(1,1,m\right) \\
&\leq 2k^2|D_n| + 4k^2 C\sum_{i\in D_n}\sum_{m=1}^{\infty}m^{d-1}\overline{\alpha}_X\left(1,1,m\right) \\
&\leq |D_n|\left(k^2+4CKk^2\right).
\end{aligned}$$

with $C<\infty$, and $K=\sum_{m=1}^{\infty}m^{d-1}\overline{\alpha}_X(1,1,m)<\infty$ by Assumption 3b. Consequently, the r.h.s. of (C.8) is seen to go to zero as $n\to\infty$, which establishes that the $X_{i,n}^k$ satisfies an L_1-norm LLN for fixed k. The proof for the ϕ-mixing case is analogous. This completes the proof. ■

Proof of Proposition 1. Define the modulus of continuity of $f_{i,n}(Z_{i,n},\theta)$ as

$$w(f_{i,n},Z_{i,n},\delta) = \sup_{\theta'\in\Theta}\sup_{\theta\in B(\theta',\delta)}\left|f_{i,n}(Z_{i,n},\theta)-f_{i,n}(Z_{i,n},\theta')\right|.$$

Further observe that

$$\left\{\omega : w(f_{i,n},Z_{i,n},\delta)>\varepsilon\right\}\subseteq\left\{\omega : B_{i,n}h(\delta)>\varepsilon\right\}.$$

By Markov's inequality and the i.p. part of Condition 1, we have

$$\begin{aligned}
& \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} P\left[w(f_{i,n}, Z_{i,n}, \delta) > \varepsilon\right] \\
\leq & \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} P\left[B_{i,n} > \frac{\varepsilon}{h(\delta)}\right] \\
\leq & \left[\frac{h(\delta)}{\varepsilon}\right]^p \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} EB_{i,n}^p \leq C_1 \left[\frac{h(\delta)}{\varepsilon}\right]^p \to 0 \text{ as } \delta \to 0
\end{aligned}$$

for some $C_1 < \infty$, which establishes the i.p. part of the theorem. For the a.s. part, observe that by the a.s. part of Condition 1 we have a.s.

$$\begin{aligned}
& \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} w(f_{i,n}, Z_{i,n}, \delta) \\
\leq & h(\delta) \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} B_{i,n} \leq C_2 h(\delta) \to 0 \text{ as } \delta \to 0
\end{aligned}$$

for some $C_2 < \infty$, which establishes the a.s. part of the theorem. ■

Proof of Proposition 2. The proof is analogous to the first part of the proof of Theorem 4.5 in Pötscher and Prucha (1994a). We give an explicit proof for the convenience of the reader. Let

$$w(f_{i,n}, z, \delta) = \sup_{\theta'\in\Theta} \sup_{\theta\in B(\theta',\delta)} \left|f_{i,n}(z,\theta) - f_{i,n}(z,\theta')\right|$$

denote the modulus of continuity of $f_{i,n}(z,\theta)$, and let $w(s_{ki,n}, z, \delta)$ be defined analogously. First note that for any $\varepsilon > 0$, we have

$$\begin{aligned}
P(w(f_{in}, Z_{i,n}, \delta) & > \varepsilon) \leq P\left(\sum_{k-1}^{K} |r_{ki,n}(Z_{i,n})|\, w(s_{ki,n}, Z_{i,n}, \delta) > \varepsilon\right) \\
& \leq \sum_{k=1}^{K} P\left(|r_{ki,n}(Z_{i,n})|\, w(s_{ki,n}, Z_{i,n}, \delta) > \frac{\varepsilon}{K}\right) \\
& \leq \sum_{k=1}^{K} P\left(|r_{ki,n}(Z_{i,n})|\, w(s_{ki,n}, Z_{i,n}, \delta)\mathbf{1}_{K_m}(Z_{i,n}) > \frac{\varepsilon}{2K}\right) \\
& \quad + \sum_{k=1}^{K} P\left(|r_{ki,n}(Z_{i,n})|\, w(s_{ki,n}, Z_{i,n}, \delta)\mathbf{1}_{Z-K_m}(Z_{i,n}) > \frac{\varepsilon}{2K}\right).
\end{aligned}$$

For any m, $1 \leq k \leq K$, and $\eta > 0$ it follows form equicontinuity Condition 2(b), that there exists $\delta(m,\eta) > 0$ such that

$$\sup_n \sup_{i\in D_n} \sup_{z\in K_m} w(s_{ki,n}, z, \delta) < \eta.$$

By. Markov's inequality we now have for each $1 \leq k \leq K$:

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} P\left(|r_{ki,n}(Z_{i,n})|\, w(s_{ki,n}, Z_{i,n}, \delta)\mathbf{1}_{K_m}(Z_{i,n}) > \frac{\varepsilon}{2K}\right) \\
\leq\ & \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} P\left(|r_{ki,n}(Z_{i,n})|\, \eta > \frac{\varepsilon}{2K}\right) \\
\leq\ & \frac{2K\eta}{\varepsilon} \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} E\,|r_{ki,n}(Z_{i,n})| \leq \frac{2KB\eta}{\varepsilon}
\end{aligned}$$

where $B = \limsup_{n\to\infty} |D_n|^{-1} \sum_{i\in D_n} E\,|r_{ki,n}(Z_{i,n})|$, which is finite by Condition 2(a). Since η was arbitrary it follows that

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} P\left(|r_{ki,n}(Z_{i,n})|\, w(s_{ki,n}, Z_{i,n}, \delta)\mathbf{1}_{K_m}(Z_{i,n}) > \frac{\varepsilon}{2K}\right) = 0.$$

Also, for each $1 \leq k \leq K$ it follows from by Condition 2(b) that

$$\begin{aligned}
&\lim_{m\to\infty} \lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} P\left(|r_{ki,n}(Z_{i,n})|\, w(s_{ki,n}, Z_{i,n}, \delta)\mathbf{1}_{Z-K_m}(Z_{i,n}) > \frac{\varepsilon}{2K}\right) \\
\leq\ & \lim_{m\to\infty} \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} P\left(\mathbf{1}_{Z-K_m}(Z_{i,n})\right) = 0.
\end{aligned}$$

Hence

$$\begin{aligned}
&\lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} P(w(f_{in}, Z_{i,n}, \delta) > \varepsilon) \\
=\ & \lim_{m\to\infty} \lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{|D_n|} \sum_{i\in D_n} P(w(f_{in}, Z_{i,n}, \delta) > \varepsilon) = 0,
\end{aligned}$$

which completes the proof. ■

References

[1] Andrews, D.W.K. (1987): "Consistency in Nonlinear Econometric Models: a Generic Uniform Law of Large Numbers," *Econometrica*, 55, 1465-1471.

[2] Andrews, D.W.K. (1992): "Generic Uniform Convergence," *Econometric Theory*, 8, 241-257.

[3] Audretsch, D.B., and M.P. Feldmann (1996): "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 630-640.

[4] Baltagi, B.H., and D. Li, (2001a): "Double Length Artificial Regressions for Testing Spatial Dependence," *Econometric Reviews,* 20, 31-40.

[5] Baltagi, B.H., and D. Li, (2001b): "LM Test for Functional Form and Spatial Error Correlation," *International Regional Science Review,* 24, 194-225.

[6] Baltagi, B.H., S.H. Song, and W. Koh (2003): "Testing Panel Data Regression Models with Spatial Error Correlation," *Journal of Econometrics*, 117, 123-150.

[7] Baltagi, B.H., S.H. Song, B.C. Jung, and W. Koh (2006): "Testing for Serial Correlation, Spatial Autocorrelation and Random Effects Using Panel Data," forthcoming in *Journal of Econometrics*.

[8] Baltagi, B.H., P. Egger, and M. Pfaffermayr (2005): "Estimating Models of Complex FDI: Are There Third Country Effects?" Working Paper, Department of Economics, Texas A&M University, forthcoming in *Journal of Econometrics*.

[9] Bao, Y., and A. Ullah (2007): "Finite Sample Properties of Maximum Likelihood Estimators in Spatial Models," *Journal of Econometrics*, 137, 396-413.

Bell, K.P., and N.E. Blockstael (2000): "Applying the Generalized Method Estimation Approach to Spatial Problems Involving Microlevel Data," *Review of Economics and Statistics*, 82, 72-82.

[10] Bera, A.K, and P. Simlai: "Testing Spatial Autoregressive Model and a Formulation of Spatial ARCH (SARCH) Model with Applications. Paper presented at the Econometric Society World Congress, London, August 2005.

[11] Besley, T., and A. Case (1995): "Incumbent Behavior: Vote-seeking, Tax-setting, and Yardstick Competition," *American Economic Review*, 85, 25-45.

[12] Betrand, M., E.F.P. Luttmer, and S. Mullainathan (2000): "Network Effects and Welfare Cultures," *Quarterly Journal of Economics*, 115, 1019-1055.

[13] Bierens, H.J. (1994): *Topics in Advanced Econometrics.* Cambridge: Cambridge University Press.

[14] Billingsley, P. (1986): *Probability and Measure.* John Wiley, New York.

[15] Bolthausen, E. (1982): "On the Central Limit Theorem for Stationary Mixing Random Fields," *The Annals of Probability,* 10, 1047-1050.

[16] Bradley, R. (1992): "On the Spectral Density and Asymptotic Normality of Weakly Dependent Random Fields," *Journal of Theoretical Probability*, 5, 355-373.

[17] Brock, W., and S. Durlauf (2001): "Interactions-based Models," *Handbook of Econometrics*, 5, J. Heckman and E. Leamer eds., Amsterdam: North Holland.

[18] Brock, W., and S. Durlauf (2007): "Identification of Binary Choice Models with Social Interactions," forthcoming in *Journal of Econometrics.*

[19] Brockwell, P., and R. Davis (1991): *Times Series: Theory and Methods*, Springer Verlag.

[20] Case, A. (1991): "Spatial Patterns in Household Demand," *Econometrica*, 59, 953-966.

[21] Cliff, A., and J. Ord (1973): *Spatial Autocorrelation*, London: Pion.

[22] Cliff, A., and J. Ord (1981): *Spatial Processes, Models and Applications.* London: Pion.

[23] Cohen, J.P., and C.J. Morrison Paul (2004): "Public Infrastructure Investment, Interstate Spatial Spillovers and Manufacturing Costs," *Review of Economics and Statistics*, 86, 551-560.

[24] Conley, T. (1999): "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 92, 1-45.

[25] Conley, T., and G. Topa (2003): "Identification of Local Interaction Models with Imperfect Location Data," *Journal of Applied Econometrics*, 18, 605-618.

[26] Conley, T., and B. Dupor (2003): "Spatial Analysis of Sectoral Complementarity," *Journal of Political Economy,* 111, 311-352.

[27] Conley, T., and F. Molinari (2007): "Spatial Correlation Robust Inference with Errors in Location or Distance," forthcoming in *Journal of Econometrics.*

[28] Cressie, N.A.C. (1993): *Statistics for Spatial Data*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. New York: Wiley.

[29] Das, D., H.H. Kelejian, and I.R. Prucha (2003): "Small Sample Properties of Estimators of Spatial Autoregressive Models with Autoregressive Disturbances," *Papers in Regional Science,* 82, 1-26.

[30] Davidson, J. (1992): "A Central Limit Theorem for Globally Nonstationary Near-Epoch Dependent Functions of Mixing Processes," *Econometric Theory*, 8, 313-329.

[31] Davidson, J. (1993a): "A Central Limit Theorem for Globally Nonstationary Near-Epoch Dependent Functions of Mixing Processes: Asymptotically Degenerate Case," *Econometric Theory*, 9, 402-412.

[32] Davidson, J. (1993b): "An L_1 Convergence Theorem for Heterogenous Mixingale Arrays with Trending Moments," *Statistics and Probability Letters*, 8, 313-329.

[33] De Jong, R.M. (1997): "Central Limit Theorems for Dependent Heterogeneous Random Variables," *Econometric Theory*, 13, 353–367.

[34] De Long, J., and L. Summers (1991): "Equipment Investment and Economic Growth," *Quarterly Journal of Economics*, 106, 445-502.

[35] Driscoll, J. and A. Kraay (1998): "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *The Review of Economics and Statistics,* LXXX, 549-560.

[36] Dobrushin, R. (1968a): "The Description of a Random Field by its Conditional Distribution and its Regularity Condition," *Theory of Probability and its Applications,* 13, 197-227.

[37] Dobrushin, R. (1968b): "Gibbs Random Fields for Lattice Systems with Pairwise Interactions," *Functional Analysis and Applications,* 2, 292-301.

[38] Doukhan, P. (1994): *Mixing. Properties and Examples*. Springer-Verlag.

[39] Eberlein, E., and A. Csenki (1979): "A Note on Strongly Mixing Lattices of Random Variables," *Z. Wahrsch. verw. Gebiete*, 50, 135-136.

[40] Gallant, A.R., and H. White (1988): A *Unified Theory of Estimation and Inference for Nonlinear Dynamic Models.* New York: Basil Blackwell.

[41] Guyon, X. (1995): *Random Fields on a Network. Modeling, Statistics, and Applications.* Springer-Verlag.

[42] Hall, P., and C. Heyde (1980): *Martingale Limit Theory and its Applications,* Academic Press.

[43] Hanushek, E.A., J.F. Kain, J.M. Markman, and S.G. Rivkin (2003): "Does Peer Effect Ability Affect Student Achievement?" *Journal of Applied Econometrics,* 18, 527-544.

[44] Hegerfeldt, G.C., and C. R. Nappi (1977): "Mixing Properties in Lattice Systems," *Communications of Math. Physics,* 53, 1-7.

[45] Hotz-Eakin, D., (1994): "Public Sector Capital and the Productivity Puzzle," *Review of Economics and Statistic*s, 76, 12-21.

[46] Ibragimov, I.A. (1962): "Some Limit Theorems for Stationary Processes," *Theory of Probability and Applications,* 7, 349-382.

[47] Ibragimov, I.A., and Y. V. Linnik (1971): *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen.

[48] Ionnides, Y.M., and J.E. Zabel (2003): "Neighborhood Effects and Housing Demand," *Journal of Applied Econometrics*, 18, 563-584.

[49] Kapoor, M., H.H. Kelejian, and I.R. Prucha (2007): "Panel Data Models with Spatially Correlated Error Components," *Journal of Econometrics,* 140, 97-130

[50] Kelejian, H.H., and I.R. Prucha (1997): "Estimation of the Spatial Autoregressive Parameter by Two-Stage Least Squares Procedures: A Serous Problem," *International Regional Science Review,* 20, 103-111.

[51] Kelejian, H.H., and I.R. Prucha (1998): "A Generalized Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances," *Journal of Real Estate Finance and Economics,* 17, 99-121.

[52] Kelejian, H.H., and I.R. Prucha (1999): "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model," *International Economic Review,* 40, 509-533.

[53] Kelejian, H.H., and I.R. Prucha (2001): "On the Asymptotic Distribution of the Moran I Test Statistic with Applications," *Journal of Econometrics,* 104, 219-257.

[54] Kelejian, H.H., and I.R. Prucha (2002): "2SLS and OLS in a Spatial Autoregressive Model with Equal Spatial Weights," *Regional Science and Urban Economics,* 32, 691-707.

[55] Kelejian, H.H., and I.R. Prucha (2004): "Estimation of Simultaneous Systems of Spatially Interrelated Cross Sectional Equations," *Journal of Econometrics,* 118, 27-50.

[56] Kelejian, H.H., and I.R. Prucha (2007a): "HAC Estimation in a Spatial Framework," *Journal of Econometrics*,140, 131-154.

[57] Kelejian, H.H., and I.R. Prucha (2007b): "Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances," forthcoming in *Journal of Econometrics.*

[58] Keller, W., and C. Shiue (2007): "The Origin of Spatial Interaction," forthcoming in *Journal of Econometrics.*

[59] Klesov, O. (1981): "The Hajek-Renyi Inequality for Random Fields and Strong Laws of Large Numbers," *Theory of Probability and Mathematical Statistics*, 22, 58-66.

[60] Kling, J.R., J. Ludwig, and L. Katz (2007): "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75, 83-119.

[61] Korniotis, G.M. (2005): "A Dynamic Panel Estimator with Both Fixed and Spatial Effects," Department of Finance, University of Notre Dame.

[62] Lahiri, S. (1996): "On Inconsistency of Estimators Based on Spatial Data under Infill Asymptotics," *Sankhya*, 58 Ser. A, 403-417.

[63] Lee, L.-F. (2002): "Consistency and Efficiency of Least Squares Estimation for Mixed Regressive Spatial Autoregressive Models," *Econometric Theory,* 18, 252-277.

[64] Lee, L.-F. (2003): "Best Spatial Two-Stage Least Squares Estimator for a Spatial Autoregressive Model with Autoregressive Disturbances," *Econometric Reviews,* 22, 307-322.

[65] Lee, L.-F. (2004): "Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models," *Econometrica,* 72, 1899-1925.

[66] Lee, L.-F. (2007a): "GMM and 2SLS for Mixed Regressive, Spatial Autoregressive Models," *Journal of Econometrics,* 137, 489–514.

[67] Lee, L.-F. (2007b): "Method of Elimination and Substitution in the GMM Estimation of Mixed Regressive Spatial Autoregressive Models," forthcoming in *Journal of Econometrics.*

[68] Lee, L.-F. (2007c): "Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects," forthcoming in *Journal of Econometrics.*

[69] LeSage, J.P., and R.K. Pace (2007): "A Matrix Exponential Spatial Specification," forthcoming in *Journal of Econometrics.*

[70] McElroy, T., and D. Politis (2000): "A Central Limit Theorem for an Array of Strong Mixing Random Fields," *Asymptotics in Statistics and Probability. Papers in Honor of George Gregory Roussas*, 289-303.

[71] Moricz, F. (1978): "Multiparameter Strong Laws of Large Numbers. (Second Order Moment Restrictions)," *Acta Sci. Math. (Szeged)*, 40, 143-156.

[72] Nahapetian, B. (1980): "A Central Limit Theorem for Random Fields," *Multicomponent Random Systems* edited by Dobrushin and Sinai, 531-542.

[73] Nahapetian, B. (1987): "An Approach to Limit Theorems for Dependent Random Variables," *Theory of Probability and its Applications,* 32, 589-594.

[74] Nahapetian, B. (1991): *Limit Theorems and Some Applications in Statistical Physics.* Teubner-Texte zur Mathematik, Leipzig.

[75] Neaderhouser, C. (1978): "Limit Theorems for Multiply Indexed Mixing Random Variables," *Annals of Probability*, 6.

[76] Newey, W.K. (1991): "Uniform Convergence in Probability and Stochastic Equicontinuity," *Econometrica*, 59, 1161-1167.

[77] Noczaly, C., and T. Tomacs (2000): "A General Approach to Strong Laws of Large Numbers for Fields of Random Variables," *Annales Univ. Sci. Budapest*, 43, 61-78.

[78] Peligrad, M., and A. Gut (1999): "Almost-sure Results for a Class of Dependent Random Variables," *Journal of Theoretical Probability*, 12, 87-104.

[79] Pinske, J., and M.E. Slade (1998): "Contracting in Space: Application of Spatial Statistics to Discrete-Choice Models," *Journal of Econometrics*, 85, 1111-1153.

[80] Pinske, J., M.E. Slade, and C. Brett (2002): "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, 70, 1111-1153.

[81] Pinske, J., L. Shen, and M.E. Slade (2006): "A Central Limit Theorem for Endogenous Locations and Complex Spatial Interactions," forthcoming in *Journal of Econometrics.*

[82] Pollard, D. (1984): C*onvergence of Stochastic Processes.* Springer-Verlag, New York.

[83] Pötscher, B. M., and I. R. Prucha (1989): "A Uniform Law of Large Numbers for Dependent and Heterogeneous Data Processes," *Econometrica*, 57, 675-683.

[84] Pötscher, B.M., and I. R. Prucha (1994a): "Generic Uniform Convergence and Equicontinuity Concepts for Random Functions," *Journal of Econometrics*, 60, 23-63.

[85] Pötscher, B.M., and I. R. Prucha (1994b): "On the Formulation of Uniform Laws of Large Numbers: A Truncation Approach," *Statistics,* 25, 343-360.

[86] Pötscher, B.M. and I. R. Prucha (1997): *Dynamic Nonlinear Econometric Models*. Springer-Verlag, New York.

[87] Rees, D.I., J.S. Zaks, and J. Herries (2003): "Interdependence in Worker Productivity," *Journal of Applied Econometrics*, 18, 585-604.

[88] Ripley, B.D. (1981): *Spatial Statistics.* Wiley, New York.

[89] Robinson, P.M. (2007a): "Nonparametric Spectrum Estimation for Spatial Data," *Journal of Statistical Planning and Inference*, 137, 1024-1034.

[90] Robinson, P.M. (2007b): "Efficient Estimation of the Semiparametric Spatial Autoregressive Model," forthcoming in *Journal of Econometrics.*

[91] Rosenblatt, M. (1956): "A Central Limit Theorem and Strong Mixing Condition," *Proceedings of National Academie of Sciences of U.S.A.,* 42, 43-47.

[92] Sacredote, B. (2001): "Peer Effects with Random Assignment: Results of Darthmouth Roommates," *Quarterly Journal of Economics,* 116, 681-704.

[93] Sain, S.R., and N. Cressie (2007): "A Spatial Model for Multivariate Lattice Data," forthcoming in *Journal of Econometrics*.

[94] Shroder, M. (1995): "Games the States Don't Play: Welfare Benefits and the Theory of Fiscal Federalism," *Review of Economics and Statistics*, 77, 183-191.

[95] Smythe, R. (1973): "Strong Laws of Large Numbers for r-Dimensional Arrays of Random Variables," *Annals of Probability*, 1, 164-170.

[96] Smythe, R. (1974): "Sums of Independent Random Variables on Partially Ordered Sets," *Annals of Probability*, 2, 906-917.

[97] Stein, C. (1972): "A Bound for the Error in the Normal Approximation of a Sum of Dependent Random Variables," *Proc. Berkeley Symp.*, M.S.P.2, 583-603.

[98] Su, L., and Z. Yang (2007): "QML Estimation of Dynamic Panel Data Models with Spatial Errors," Working Paper, Peking University and Singapore Management University.

[99] Topa, G. (2001): "Social Interactions, Local Spillovers and Unemployment," *Review of Economic Studies*, 68, 261-295.

[100] Yang, Z. (2005): "Quasi-Maximum Likelihood Estimation for Spatial Panel Data Regressions," Paper presented at the Spatial Econometrics Workshop, Kiel Institute for World Economics, April 2005.

[101] Yu, J., R. de Jong and L-F. Lee (2006): "Quasi Maximum Likelihood Estimators for Spatial Dynamic Panel Data with Fixed Effects When Both n and T are Large," Working Paper, Ohio State University.

[102] Wooldridge, J. (1986): "Asymptotic Properties of Econometric Estimators," University of California San Diego, Department of Economics. Ph.D. Dissertation.

[103] Wooldridge, J., and H. White (1988): "Some Invariance Principles and Central Limit Theorems for Dependent Heterogeneous Processes," *Econometric Theory*, 4, 210–230.