

ABSTRACT

Title of Dissertation: ASYMPTOTIC THEORY FOR SPATIAL PROCESSES

Nazgul Jenish, Doctor of Philosophy, 2008

Directed by: Professor Ingmar R. Prucha
Department of Economics

Recent years have seen a marked increase in the application of spatial models in economics and the social sciences, in general. However, the development of a general asymptotic estimation and inference theory for spatial estimators has been hampered by a lack of central limit theorems (CLTs), uniform laws of large numbers (ULLNs) and pointwise laws of large number (LLN) for random fields under the assumptions relevant to economic applications. These limit theorems are the basic building blocks for the asymptotic theory of M-estimators, including maximum likelihood and generalized method of moments estimators. The dissertation derives new CLTs, ULLNs and LLNs for weakly dependent random fields that are applicable to a broad range of data processes in economics and other fields. Relative to the existing literature, the contribution of the dissertation is threefold. First, the proposed limit theorems accommodate nonstationary random fields with asymptotically unbounded or trending moments. Second, they cover a larger class of weakly dependent spatial processes than mixing random fields. Third, they allow for arrays of fields located on unevenly spaced lattices, and place minimal restrictions on the configuration and growth behavior of index sets. Each of the theorems is provided with weak yet primitive sufficient conditions.

ASYMPTOTIC THEORY FOR SPATIAL PROCESSES

By

Nazgul Jenish

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2008

Advisory Committee:
Professor Ingmar R. Prucha, Chair
Professor Harry Kelejian
Professor John Rust
Professor John Chao
Professor Dilip Madan

© Copyright by
Nazgul Jenish
2008

Dedication

To my grandparents, Isaakun and Guljan.

Acknowledgements

I would like to express my deepest gratitude to my mentor Professor Ingmar Prucha for his teaching, advice and encouragement. Without his continuous support, this dissertation could not have been written.

I have also had the pleasure of learning from and working with Professors Harry Kelejian, John Rust and John Chao. I would like to thank them all for their valuable comments and encouragement. This research was supported by a University of Maryland Ann G. Wylie Dissertation Fellowship.

Table of Contents

1. Introduction.....	1
2. Review of Literature	6
3. Mixing Spatial Processes	15
3.1. Introduction.....	15
3.2. Mixing Definitions and Inequalities	19
3.3. Central Limit Theorem	24
3.4. Uniform and Pointwise Laws of Large Numbers	29
3.4.1. Generic Uniform Law of Large Numbers.....	30
3.4.2. Pointwise Law of Large Numbers	34
3.4.3. Stochastic Equicontinuity: Sufficient Conditions.....	36
4. Near-Epoch Dependent Spatial Processes	41
4.1. Introduction	41
4.2. Definition and Examples of NED Processes	46
4.3. Central Limit Theorem for NED Processes.....	57
4.4. Law of Large Numbers for NED Processes.....	65
5. Conclusion	68
A. Appendix. Cardinalities of Basic Sets on Irregular Lattices.....	70
B. Appendix. Proof of CLT for Mixing Processes	75
C. Appendix. Proofs of ULLN and LLN	91
D. Appendix. Proofs for Section 4.2	99
E. Appendix. Proof of CLT for NED Processes	102
F. Appendix. Proof of LLN for NED Processes	115

References..... 118

1 Introduction

Recent years have seen a marked increase in the application of spatial models in economics and the social sciences, in general. Apart from traditional applications in agricultural economics and economic geography, spatial methods have increasingly been used to model and estimate interaction among economic agents in various other fields of economics including IO, labor and public economics, international economics, political economy and macroeconomics. In these models, economic agents are located in some space with a specified metric.¹ Strategic behavior of agents as well as common factors such as shared resources, shocks and trade induce dependence in agents' characteristics. Furthermore, economic agents are often heterogenous in various dimensions, e.g., size. All these diverse economic applications thus share a common mathematical structure: an agent's observation can be viewed as a realization of a dependent heterogenous spatial process indexed by a point in a finite-dimensional metric space, or a *random field*.

Statistical inference in these models is typically based on the large sample properties of estimators. To the present date, the asymptotic properties of spatial estimators have been established, to the best of our knowledge, only for special classes of models: (i) linear first-order spatial autoregressive

¹The space and metric here are not restricted to physical space and distance, but refer to more general spaces and notions of proximity. For instance, in their study of the productivity co-movements across sectors, Conley and Dupor (2003) define the distance between industries in terms of input shares, i.e., two industries are deemed to be close if they use the same inputs in the same proportions.

models;² and (ii) stationary models.³ Clearly, many interesting empirical models are nonlinear, e.g., discrete choice and limited dependent variable models, and their underlying data-generating processes are often nonstationary. Thus, the existing asymptotic foundation of spatial models has become increasingly inadequate to sustain the growing complexity and diversity of empirical applications. These applications call for a less restrictive large sample theory that accommodates nonlinearity and more general forms of heterogeneity and dependence.

The development of such an asymptotic theory has been hampered by a lack of uniform laws of large numbers (ULLNs), pointwise laws of large number (LLN), and central limit theorems (CLTs) for random fields under the assumptions relevant to economic applications. These limit theorems are the key tools for analyzing the large sample properties of estimators, i.e., for establishing consistency and the asymptotic distribution of estimators.

There exists a vast literature on limit theorems for random fields. A detailed review of this literature is provided in the next section. Here, we highlight only the generic features of these results that prevent their application in socioeconomic models. First, all CLTs and LLNs for discrete-index random fields are, to the best of our knowledge, for processes on evenly spaced

²The asymptotic theory for spatial autoregressive models has been developed in a number of important contributions by Kapoor, Kelejian and Prucha (2007); Kelejian and Prucha (1998, 1999, 2001, 2004, 2007a,b); Lee (2002, 2004, 2007a); Pinkse and Slade (1998); Pinkse, Shen and Brett (2002); Pinkse, Shen and Slade (2006); Robinson (2007b); Yu, de Jong and Lee (2006).

³Consistency and asymptotic normality of nonlinear GMM estimators for stationary α -mixing random fields are established in Conley (1999).

lattices, while locations of economic data are rarely evenly spaced. Second, the existing theorems rely either on homogeneity/stationarity of the random field or the existence of uniform bounds on moments, which restricts heterogeneity. However, spatial processes encountered in economic applications often exhibit different forms of heterogeneity/nonstationarity. For example, many spatial processes in applications are heteroskedastic. Furthermore, similar to time series processes, their moments may increase or "trend" with indices so that there is no uniform bound on them. Spatial processes with trending moments arise frequently in applications. For instance, real estate prices often shoot up as one moves from the periphery towards the center of a big city. Bera and Simlai (2005) report sharp spikes in the variances of housing prices in Boston. Such irregular behavior of second moments may cause problems for the CLT and therefore should be dealt with explicitly. The third hurdle is a lack of ULLNs for random fields, which are essential for proving consistency of nonlinear optimization or M-estimators, including generalized method of moments (GMM) and maximum likelihood (MLE) estimators.

These features of the random fields limit theory are mirrored in the existing large sample theory of spatial M-estimators, and in particular, in Conley (1999) paper, which is one of the first important applications of the random field apparatus in econometrics. The paper makes use of Bolthausen's (1982) CLT for α -mixing stationary random fields on \mathbb{Z}^d to show asymptotic normality of nonlinear GMM estimators. It assumes that the data-generating process is (i) stationarity; and (ii) evenly spaced. Furthermore, to prove consistency, it effectively postulates uniform convergence instead of proving it

from low-level conditions.

To fill this gap in the literature, the first part of the dissertation derives a CLT, ULLN and LLN for nonstationary mixing random fields suitable for econometric applications. In contrast to the existing literature, the proposed limit theorems (i) accommodate nonstationary random fields with asymptotically unbounded or trending moments, (ii) allow for doubly-indexed arrays of random fields on unevenly spaced lattices in d -dimensional spaces, and (iii) relax assumptions on the configuration and growth behavior of sample regions imposed by existing theorems. As discussed earlier, all these features are critical for many econometric models.

Mixing is perhaps the most common notion of weak dependence employed in the literature. It dates back to Rosenblatt (1956), Ibragimov (1962) and Billingsley (1968). Loosely speaking, under the mixing property, autocorrelation of the process decays with the distance. It is quite a reasonable assumption satisfied in many econometric applications. However, it has one undesirable feature: it is not preserved under general data transformations, and in particular, those involving an infinite number of lags. Yet, there are spatial processes that are generated as infinite lag transformations of some input mixing process, e.g., infinite moving average random fields, which are also referred to as linear random fields. Therefore, limit theorems for mixing random fields are not directly applicable to such processes.

To address this problem, the second part of the dissertation extends the concept of near-epoch dependent (NED) processes used in the time series literature to spatial processes, and obtains a CLT and LLN for such processes. The basic idea is that a NED process can be approximated by a process

defined as a function of only a finite number of spatial lags of some input process. As a result, the approximating process inherits the mixing property of the input process and the limit theorems for mixing fields can be applied to infer the limiting behavior of the approximating process. Under some mild conditions, the approximation error can be shown to be asymptotically negligible, i.e., to have no effect on the limiting behavior of the NED process. Consequently, the NED process will satisfy a CLT or an LLN. The NED property is compatible with considerable amount of heterogeneity and dependence. The class of NED spatial processes, which subsumes mixing processes, is sufficiently broad to cover many spatial processes of interest, for example, ARMA random fields and Cliff-Ord type processes used widely in applications.

The CLT and LLN for NED random fields thus cover a larger class of dependent spatial processes than mixing fields. As the theorems of the first part of the dissertation, they also allow for nonstationary processes with asymptotically unbounded moments, located on unevenly spaced lattices. To the best of our knowledge, there have been no similar results in the random fields literature. In the time series literature, CLTs for NED processes have been obtained by Wooldridge (1986), Davidson (1992, 1993), and de Jong (1997). Interestingly, our CLT contains as a special case the CLTs of Wooldridge (1986) and Davidson (1992), whereby establishing direct connection and consistency in the asymptotic properties of spatial (multi-dimensional) and time-series (one-dimensional) processes.

The proposed limit results can be readily used to investigate the large sample properties of nonlinear econometric estimators and test statistics in

a wide range of spatial models. More generally, they will also be useful in cross-sectional and panel data models (especially those with small T and large N) with cross-sectionally dependent observations when distances between observations are known. In separate work, building on these limit theorems, the author establishes consistency and asymptotic normality of spatial M-estimators including MLE and GMM estimators. These results form a fundamental basis for statistical inference, e.g., testing hypothesis and constructing confidence intervals, in a broad range of spatial models. However, we do not pursue them here. Aside from the asymptotic theory of econometric estimators, the areas of potential applications also include biology, psychology, sociology, political and environmental sciences.

The dissertation is organized as follows. Section 2 provides a review of the literature on CLTs and LLNs for mixing random fields. Section 3 presents a CLT, ULLN and LLN for mixing random fields. Section 4 introduces the concept of near-epoch dependent random fields and establishes the corresponding CLT and LLN. Section 5 concludes. All proofs are contained in appendices.

2 Review of Literature

The literature on limit theory of weakly dependent random fields is truly massive. We will therefore restrict our attention to *discrete-index mixing* random fields.

Unlike to what one might expect, limit theorems for multi-dimensional processes or random fields are *not* straightforward generalizations of those

for one-dimensional or time series processes. There are a number of distinguishing characteristics of the limit theory of random fields. First, there are two principally different ways in which the sample can grow to the limit, or in other words, asymptotic structures: *increasing domain* and *infill* asymptotics, see, e.g., Cressie (1993), p. 480. Under increasing domain asymptotics, the growth of sample is ensured by unbounded expansion of the sample region. In contrast, under infill asymptotics, the sample region remains fixed, and the growth of the sample size is achieved by sampling points arbitrarily dense in the given region. Second, unlike \mathbb{R} , there is no natural order in \mathbb{R}^d . Consequently, some of the dependence structures commonly used in the time series literature such as martingales and mixingales are not well-defined (without imposing additional structure on \mathbb{R}^d). Third, there are also differences in the definition of 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 sizes/cardinalities of index sets may result in trivial notions of dependence and leave out many dependent processes encountered in applications. For instance, Dobrushin (1968a) demonstrated 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 . The mixing coefficients in this condition are defined over two half-spaces containing infinite number of elements, and as such, they do not account for cardinalities index sets. Later, Bradey (1993) proved that this condition in the case of stationary random fields reduces to ρ -mixing, which is a more

restrictive form of dependence. The formal definitions and a more detailed discussion of mixing conditions are given in Section 3. For a comprehensive review of various mixing conditions, see Doukhan (1994). Finally, configuration and growth behavior of (multi-dimensional) sample regions also play an important role in the limit theory of random fields. This has to do with the need to obtain bounds on the variances or other moments of partial sums over the sample region. For example, the rates of convergence in the strong laws of large numbers depend on the configuration of sample regions, see Smythe (1974). Thus, the limit theorems for time series processes are not directly applicable to spatial processes or random fields.

Central limit theorems establish convergence in distribution of normalized partial sums to a normal law. Their primary application in statistics is to ascertain asymptotic normality of various estimators and test statistics, which in turn provides the basis for inference. In general, CLTs for weakly dependent random fields rely on three sets of conditions: (i) conditions restricting the degree of heterogeneity of the processes; (ii) conditions restricting the range of dependence of the process, and (iii) conditions on the index sets. Various central limit theorems differ mainly in these three major dimensions. Therefore, we will focus on these conditions in our subsequent discussion of CLTs.

Early central limit theorems for random fields were motivated by the study of Gibbs states of lattice systems in statistical physics. The central limit results for α - and ϕ -mixing conditions satisfied by Gibbs fields first appeared in the works of Neaderhouser (1978a,b, 1981); Nahapetian (1980, 1987, 1991); McElroy and Politis (2000). The common feature of these CLTs

is that they consider random fields on \mathbb{Z}^d and impose quite stringent conditions on the configuration and growth behavior of sample regions. They require the index set to expand in all directions and the border of the index set to be asymptotically negligible in size relative to the size of the entire set. Furthermore, Nahapetian (1980, 1987), McElroy and Politis (2000) restrict sample regions to rectangles in \mathbb{Z}^d . These CLTs are also more restrictive than our CLTs in other dimensions. Nahapetian (1980, 1987, 1991) exploits stationarity. Neaderhouser (1978a,b), and McElroy and Politis (2000), while permit nonstationarity, rely on stronger moment and mixing assumptions.

In passing, we note that the restrictions on configuration of sample regions in these CLTs stem from their method of proof – Bernstein’s blocking method – a common approach to prove CLTs for weakly dependent variables. The method involves splitting the sum into alternating big-small blocks and showing that the big blocks behave asymptotically as independent or martingale difference random variables.

Bolthausen (1982) obtains a CLT for strictly stationary α -mixing random fields on \mathbb{Z}^d . The CLT relies on finite $2 + \delta$ moments and stationarity. In contrast to the above-cited results, the proof of Bolthausen’s (1982) CLT is based on Stein’s lemma (1972); see Lemma B.1 in Appendix B. It exploits the differential equation satisfied by the characteristic function of the standard normal law. Stein’s method allows to circumvent mixing conditions in which both index sets are of infinite cardinality as well as to relax conditions on the sample regions. We follow Bolthausen in using Stein’s lemma to prove our CLT for mixing random fields.

Guyon and Richardson (1984), and Guyon (1995), p. 11, derive CLTs for

nonstationary α -mixing random fields on \mathbb{Z}^d . Both results assume *uniformly* (over the index space) bounded $2 + \delta$ moments, which restricts heterogeneity of the random field. As such, they do not allow for asymptotically unbounded moments and unevenly spaced locations. Moreover, the CLT of Guyon and Richardson (1984) exploits mixing conditions in which both index sets are of infinite cardinality. As discussed earlier, these conditions are generally restrictive.

Bulinskii (1988), see also Bulinskii (1989), establishes a CLT for nonstationary α -mixing fields on \mathbb{Z}^d . This CLT improves on some of earlier results in the literature including Neaderhouser (1978a,b; 1981), Nahapetian (1980, 1987), Bolthausen (1982) and Bulinskii (1986). While the CLT accommodates nonstationarity, it does not, however, allow for unevenly spaced locations. Bulinskii and Doukhan (1990) further examine the rate of the convergence in Bulinskii's (1988) CLT.

Bradley (1992) proves a CLT under the condition $\rho(r) \rightarrow 0$ as $r \rightarrow \infty$ and some additional restrictions on the spectral density of the process. As is well-known, ρ -mixing is a stronger dependence concept than α -mixing. The CLT is for strictly stationary fields on \mathbb{Z}^d .

Nahapetian and Petrossian (1992), and Nahapetian (1995) generalize the notion of martingales and martingale differences to random fields by introducing partial order structure in \mathbb{Z}^d . They propose two CLTs for martingale difference random fields. The first CLT deals with strictly stationary ergodic martingale difference random fields. Nahapetian and Petrossian show that the existence of finite second moments is sufficient to guarantee a CLT for such fields. This is in fact a multi-dimensional analogue of

the Billingsley-Ibragimov CLT for strictly stationary ergodic martingale difference sequences. The second theorem considers nonstationary martingale difference fields with finite $2 + \delta$ moments under an additional α -mixing condition. This combination of two different dependence structures seems unnecessarily restrictive. Furthermore, the scope of potential applications of martingale random fields in econometrics may be limited due to the lack of obvious order structure in many economic models.

Comets and Janžura (1998) do not use any mixing conditions. Instead, they consider a special class of conditionally centered random fields on \mathbb{Z}^d derived from some underlying "well-behaved" random field. Their CLT allows for nonstationarity, but assumes uniformly bounded *fourth* moments, which is quite restrictive. Moreover, conditional centering may not be satisfied by mixing fields. As noted in Dedecker (1998), conditions of Bolthausen's CLT cannot be inferred from this CLT.

Perera (1997) relaxes the moment conditions in Bolthausen's (1982) CLT. Yet, the result is still for stationary fields on \mathbb{Z}^d . Dedecker (1998) obtains a CLT for stationary random fields under an alternative projective criterion which, roughly speaking, involves convergence of the sum of conditional covariances. This criterion enables him to further refine the moment and dependence conditions in Bolthausen's (1982) CLT. However, the proof depends critically on stationarity, and therefore, it is not clear if the CLT could extend to the nonstationary case. Dedecker (2001) establishes exponential inequalities and uses them to derive a functional form of his CLT for stationary fields.

A different kind of CLT that does not employ any mixing coefficients is

proposed by Pinkse, Shen and Slade (2006). The CLT allows for nonstationarity and dependence on the sample. However, it 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 proving a CLT by 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 Pinkse, Shen and Slade (2006). Thus, on a technical level, neither of the CLTs contains nor dominates the other.

More recent literature focuses on functional central limit theorems (FCLTs) for partial sums indexed with general Borel sets. El Machkouri (2002) proves a set-indexed FCLT for *stationary* fields under a finite *exponential moment* condition. It covers ϕ -mixing but not α -mixing fields. El Machkouri and Ouchti (2005) provide a FCLT for stationary martingale-difference fields. Given their complex nature, all these results are for *stationary* fields and employ stronger moment, dependence, and additional metric entropy conditions. Establishing general positive-entropy-set-indexed FCLTs for L_p -bounded ($0 < p < \infty$) α -mixing fields remains an open problem, see counterexamples in El Machkouri and Volný (2002).

There is an equally extensive literature on laws of large numbers for random fields. Vast majority of this literature focuses on almost sure convergence rather than convergence in probability. Just as with central limit theorems, extension of one-parameter laws of large numbers to random fields is fraught with technical difficulties stemming from complex geometry of index sets. Yet, these complications become even more pronounced in the case of strong

laws of large numbers (SLLNs). The proof of CLTs for mixing processes makes use of mixing covariance inequalities which do not depend on ordering of index sets. In contrast, maximal inequalities, which are the key tools for proving SLLNs, depend critically on ordering of index sets. Maximal inequalities place a bound on the extreme behavior of the maximum of partial sums over a succession of steps, and hinge upon ordering of index sets. As a consequence, SLLNs for random fields impose restrictions on configuration of index sets. Most SLLNs are formulated for partially ordered rectangular sets. We now briefly discuss these results.

Early results in the SLLN literature are concerned with independent random fields. Smythe (1973, 1974) establishes SLLNs for i.i.d. random fields on \mathbb{Z}^d . Fazekas (1983) generalizes Smythe's (1973) SLLN to Banach space valued i.i.d random fields indexed by partially ordered rectangular sets in \mathbb{Z}_+^d . Gut (1978) extends Marcinkiewicz-Zygmund SLLN to the case of i.i.d. random fields. Marcinkiewicz-Zygmund SLLNs are generalization of the classical Kolmogorov SLLN to the case of finite moments of order $0 < p < 2$.

Using the concept of ρ -mixing fields introduced by Bradley (1992), Peligrad and Gut (1999) obtain a maximal inequality for ρ -mixing random fields and, building on it, establish a SLLN on \mathbb{Z}^d . Moricz (1977) derives more general maximal inequalities for rectangular sets in \mathbb{Z}_+^d . The distinguishing feature of his inequalities is that no assumptions are made with respect to the dependence structure and the degree of heterogeneity of the random field. This allows Moricz (1978) to derive a SLLN under quite general dependence assumptions. However, the SLLN requires finite second moments and imposes restrictions on the norming factor.

A slightly different approach is taken by Klesov (1981), Noczaly and Tomacs (2000). It is based on the modified version of maximal inequalities known as Hajek-Renyi type inequalities. The main advantage of Hajek-Renyi type maximal inequalities over standard maximal inequalities is that they allow for arbitrary norming factors, for example, those increasing at a logarithmic rate. Like Moricz' inequalities, they are flexible to cover different dependence structures. Klesov (1981) establishes a SLLN for a larger class of random fields comprising martingales, orthogonal and stationary (in the wide sense) random fields on \mathbb{Z}^d . He also relaxes the moment condition to $1 < q \leq 2$.

The natural question that arises is whether the dependence conditions in the latter SLLN could be further relaxed, for example, whether a SLLN could be obtained for d -dimensional mixingales and how the conditions of such a SLLN would compare with those of SLLNs for one-dimensional processes, e.g., McLeish (1975a) and Hansen (1991). This problem is investigated by Noczaly and Tomacs (2000). They derive SLLNs for d -dimensional martingale difference sequences and d -dimensional mixingales on rectangular index sets in \mathbb{Z}^d . This result generalizes Hansen's (1991) SLLN for one-dimensional mixingales. While the mixingale conditions in Noczaly and Tomacs (2000) are similar to those in Hansen (1991), the former result rests on a stronger moment condition (moment of order $r \geq 2$) than its one-dimensional counterpart, which relies only on $r > 1$. This inefficiency stems from the lack of linear ordering and higher dimensionality of the index space. To date, there seems to be no known SLLNs for mixing random fields that are based on moments of strictly less than 2.

To summarize, the existing SLLNs for random fields rely on fairly strong moment conditions and ordering of index sets, which are typically assumed to be rectangular sets. Clearly, this limits applicability of the SLLN for random fields in econometrics, and in particular, their use in the proof of strong consistency of econometric estimators.

Fortunately, weak consistency of estimators is sufficient for showing validity of asymptotic inference procedures in many econometric applications. It is also easier to verify especially when the data-generating process is a complicated function of some underlying dependent process. Therefore, we derive weak LLNs for mixing random fields and their functions. Our LLNs do not depend on configuration of index sets and require finite moments of order slightly greater than 1. Furthermore, they hold under a subset of assumptions maintained for our CLTs, which facilitates their joint application in the proof of consistency and asymptotic normality of spatial estimators.

3 Mixing Spatial Processes

3.1 Introduction

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.

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.⁴ 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 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 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 configura-

⁴Some 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 (2001, 2007), Conley (1999), Conley and Molinari (2007), Conley and Topa (2007), Das, Kelejian and Prucha (2003), Driscoll 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).

tion and, more importantly, for unevenly spaced locations. To accommodate Cliff-Ord type processes, we furthermore 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 .

There is a vast literature on CLTs for weakly dependent random fields under various mixing conditions, including Neaderhouser (1978a,b, α -mixing), Nahapetian (1980, 1987, α - and ϕ -mixing), Bolthausen (1982, α -mixing), Guyon and Richardson (1984, α -mixing), Bulinskii (1988, α -mixing), Bradley (1992, ρ -mixing), Guyon (1995, α -mixing), Perera (1997, α -mixing), Dedecker (1998, 2001) 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 (1978a,b) 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 of index sets.⁵ Neaderhouser (1978a,b) 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

⁵For formal definitions, see, e.g., Nahapetian (1991).

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 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), Bradley (1992), Perera (1997), Dedecker (1998, 2001) maintain strict stationarity. Our CLT accommodates nonstationary processes. Furthermore, to the best of our knowledge, there seem to be no results that allow for processes with trending moments, to which we will also refer to as trending spatial processes in analogy with time series processes. Spatial processes with asymptotically 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. For more examples, see Cressie (1993).

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.

ULLNs are essential tools for establishing consistency of nonlinear estimators; cp., e.g., 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.⁶ 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.⁷

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 including, among others, Smythe (1973, 1974), Gut (1978), Moricz (1978), Klesov (1981), Peligrad and Gut (1999), Noczaly and Tomacs (2000) are strong laws for fields on partially ordered rectangles in \mathbb{Z}^d , which prevents their use in more general settings.

3.2 Mixing Definitions and Inequalities

We consider spatial processes located on a (possibly) unevenly spaced lattice $D \subseteq \mathbb{R}^d$, $d \geq 1$, where the index space \mathbb{R}^d is endowed with the maximum metric: $\rho(i, j) = \max_{1 \leq l \leq d} |j_l - i_l|$, and the corresponding norm $|i| =$

⁶For different definitions of stochastic equicontinuity see Section 3 of the present paper or Pötscher and Prucha (1994a).

⁷The 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.

$\max_{1 \leq l \leq d} |i_l|$, where i_l denotes the l -th component of i . The distance between any subsets $U, V \subset D$ is defined as $\rho(U, V) = \inf \{\rho(i, j) : i \in U \text{ and } j \in V\}$. Furthermore, let $|U|$ denote the cardinality of a finite subset $U \subset D$. 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 ensures unbounded expansion of sample regions, i.e., increasing domain asymptotics, 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 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:

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

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 if

$$\lim_{m \rightarrow \infty} \sup_t \alpha(\mathfrak{F}_{-\infty}^t, \mathfrak{F}_{t+m}^{+\infty}) = 0,$$

where $\mathfrak{F}_{-\infty}^t = \sigma(\dots, X_{t-1}, X_t)$ and $\mathfrak{F}_{t+m}^{+\infty} = \sigma(X_{t+m}, X_{t+m+1}, \dots)$. 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 $\{i = (i_1, \dots, i_d) \in \mathbb{R}^d, i_k \leq a\}$ and let H_k^b be a collection of all half-spaces of the type $\{i = (i_1, \dots, i_d) \in \mathbb{R}^d, i_k \geq b\}$, with $a < b, a, b \in \mathbb{R}$, which are formed by the hyperplanes perpendicular to the k -th coordinate axis, $k = 1, \dots, d$. 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.,

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

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 $\tilde{\alpha}(r) \rightarrow 0$ as $r \rightarrow \infty$. 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

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

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 $c_n^Y(U, V) \leq c_n^X(U, V)$, $c_n^Y(k, l, r) \leq c_n^X(k, l, r)$, $\bar{c}^Y(k, l, r) \leq \bar{c}^X(k, l, r)$ for $c \in \{\alpha, \phi\}$. Thus α - and ϕ -mixing conditions are preserved under transformation.

The key role in establishing CLTs for mixing processes is played by covariance inequalities. For convenience and ease of reference, we collect the covariance inequalities for α - and ϕ -mixing variables, which are central for the proof of our limit theorems, in the following lemma.

Lemma 1 *Suppose U and V are finite sets in D with $|U| = k, |V| = l$ and $h = \rho(U, V)$. Let f and g be respectively $\sigma_n(U)$ - and $\sigma_n(V)$ -measurable and let $\|f\|_p = (E|f|^p)^{1/p}$.*

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

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

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

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

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

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

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

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

3.3 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 $\sigma_n^2 = \text{Var}(S_n)$.

In this section, we present a CLT for the normalized partial sums $\sigma_n^{-1}S_n$ of the array Z with asymptotically 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*

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

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

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

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

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

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

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

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

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

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

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

Theorem 1 Suppose $\{D_n\}$ is a sequence of arbitrary finite subsets of D , satisfying Assumption 1, with $|D_n| \rightarrow \infty$ as $n \rightarrow \infty$. 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

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

Clearly, the CLT can be readily extended to vector-valued random fields using the standard Cramér-Wold device. 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., $\sup_n \sup_{i \in D_n} E |Z_{i,n}/c_{i,n}|^r < \infty$, see, e.g., Billingsley (1986), pp. 219.

The constants $c_{i,n}$ are scale factors that account for potentially trending 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 asymptotically unbounded variances.

Spatial processes with asymptotically 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. Cressie (1993) contains numerous examples of spatial data exhibiting considerable heterogeneity and trend.

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.

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 $\bar{\alpha}(1, \infty, m) = o(m^{-d})$. 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 $\delta > 0$).

Finally, Assumption 5 limits the growth behavior of $v_{i,n}^2 = EZ_{i,n}^2$. For example, consider the case where $D_n = [-n; n]^d \subset \mathbb{Z}^d$, $Z_{i,n}$ satisfies Assumption 2 with $c_{i,n} = \max(v_{i,n}, 1)$, 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 $v_{i,n}^2 \sim |i|^\gamma$ for some $\gamma > 0$, then $M_n \sim n^{\gamma/2}$

and $\sigma_n^2 = \sum_{i \in D_n} v_{i,n}^2 \sim n^{(\gamma+d)}$. 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 $v_{i,n}^2 = O(1)$ and hence $M_n = O(1)$, Assumption 5 reduces to $\liminf_{n \rightarrow \infty} |D_n|^{-1} \sigma_n^2 > 0$. In the stationary case, an analogous condition was 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 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 $\liminf_{n \rightarrow \infty} n^{-1} M_n^{-2} > 0$, or equivalently $\limsup_{n \rightarrow \infty} nM_n^2 < \infty$.

Thus, the α -mixing part of Theorem 1 extends Bolthausen's (1982) CLT in a number of important directions. In particular, it has the following attributes essential for economic applications discussed in Introduction: i) it allows moments to depend on indices, ii) it accommodates asymptotically unbounded or trending second moments, and iii) it allows for more general index sets than subsets of \mathbb{Z}^d , including unevenly spaced locations. In particular, it relaxes Bolthausen's restrictions on the growth behavior of sets, namely that $D_n \uparrow D$ and $|\partial D_n|/|D_n| \rightarrow 0$, 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 the best of our knowledge, there are no results in the literature that combine these features and/or contain Theorem 1 as a special case.

3.4 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.

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).⁸ Of course, to make the approach operational for random fields we need an LLN, and therefore we also introduce a new

⁸We 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).

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., $f_{i,n}: Z \times \Theta \rightarrow \mathbb{R}$ and $q_{i,n}: Z \times \Theta \rightarrow \mathbb{R}$, 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 $\{\theta \in \Theta : \nu(\theta', \theta) < \delta\}$.

3.4.1 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 will employ the following definition.⁹

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

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

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

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

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

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

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

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: $ACL_pUEC \implies ACL_0UEC \iff a.s.ACUEC$.

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

⁹All 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.

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

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

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 $|D_n| \rightarrow \infty$ as $n \rightarrow \infty$. 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 Assumption 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

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

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)

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

(b) $\bar{Q}_n(\theta) = E Q_n(\theta)$ is uniformly equicontinuous in the sense that

$$\limsup_{n \rightarrow \infty} \sup_{\theta' \in \Theta} \sup_{\theta \in B(\theta', \delta)} |\bar{Q}_n(\theta) - \bar{Q}_n(\theta')| \rightarrow 0 \text{ as } \delta \rightarrow 0. \quad (3)$$

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., $\lim_{k \rightarrow \infty} \sup_n \sup_{i \in D_n} E(d_{i,n}^p \mathbf{1}(d_{i,n} > k)) = 0$, 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 $q_{i,n}/c_{i,n}$ are given in the next two subsections. 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 $q_{i,n}/c_{i,n}$ and the Domination Assumption 6 jointly imply that $q_{i,n}/c_{i,n}$ 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 $p \geq 1$.

3.4.2 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*

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

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 $|D_n| \rightarrow \infty$ as $n \rightarrow \infty$. Suppose further that $\{Z_{i,n}; i \in D_n, n \in \mathbb{N}\}$ is an array of real random fields satisfying Assumption 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. Fur-

thermore, 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 $\sup_n \sup_{i \in D_n} E |q_{i,n}(Z_{i,n}; \theta)/c_{i,n}|^{1+\eta} < \infty$ 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 .

3.4.3 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.

Lipschitz in Parameter

Assumption 7 *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:*

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

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 $p > 0$

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

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 Assumption 7, $f_{i,n}(Z_{i,n}, \theta)$ is L_0 [a.s.] stochastically equicontinuous on Θ .*

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:

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

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

Assumption 8 *The random functions $f_{i,n}(Z_{i,n}, \theta)$ defined in (4) satisfy the following conditions:*

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

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

(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:*

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

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

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

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 Assumption 8, $f_{i,n}(Z_{i,n}, \theta)$ 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. Assumption 8 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 Assumption 8(a) is the uniform L_1 boundedness of $r_{ki,n}$, i.e., $\sup_n \sup_{i \in D_n} E |r_{ki,n}(Z_{i,n})| < \infty$.

Assumption 8(b) requires *nonrandom* 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 $r((z, \theta); (z', \theta')) = \max \{\nu(\theta, \theta'), \tau(z, z')\}$. 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 Assumption 8(b) is implied by the following condition for each $1 \leq k \leq K$

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

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.,

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

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 Assumption 8(c) is an asymptotic tightness condition for the average of the marginal distributions of Z_{in} . Assumption 8(c) can frequently be implied by a mild moment condition. In particular, the following is sufficient for Assumption 8(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 $\limsup_{n \rightarrow \infty} |D_n|^{-1} \sum_{i \in D_n} E h(Z_{in}) < \infty$ where $h : [0, \infty) \rightarrow [0, \infty)$ is a monotone function such that $\lim_{x \rightarrow \infty} h(x) = \infty$.¹⁰

We note that, in contrast to Assumption 7, Assumption 8 will generally not cover random fields with trending moments since in this case part (c) would typically not hold.

¹⁰For 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.

4 Near-Epoch Dependent Spatial Processes

4.1 Introduction

In Section 3, we established a CLT, uniform and pointwise LLNs for non-stationary α - and ϕ -mixing spatial processes. These limit theorems are the essential tools for analyzing the asymptotic properties of spatial estimators and test statistics. Our results are primarily motivated by the need to develop a more general asymptotic estimation and inference theory for the growing body of spatial statistical models in economics. Over the last decade, this has been an active area of research in spatial econometrics. While a significant progress has been made for some important classes of models including linear spatial autoregressive models and stationary models¹¹, the asymptotic properties of nonlinear estimators for nonstationary dependent spatial processes have not been formally examined. This is mainly due to the lack of CLTs, uniform and pointwise LLNs for spatial processes or random fields under the assumptions relevant to socioeconomic applications. As discussed in the previous section, existing limit theorems for random fields maintain stationarity or allow only for some restrictive forms of nonstationarity. Furthermore, they do not allow for processes with unevenly spaced locations and impose restrictions on the sample regions. These features of the existing limit theorems prevent their application to econometric models in which spatial processes

¹¹E.g., Kelejian and Prucha (1998, 1999, 2001, 2004, 2007a,b); Conley (1999); Pinkse and Slade (1998); Lee (2002, 2004, 2007a); Pinkse, Shen and Brett (2002); Pinkse, Shen and Slade (2006); Yu, de Jong and Lee (2006), Robinson (2007b); Kapoor, Kelejian and Prucha (2007).

are often nonstationary and located on unevenly spaced lattices. The CLT, ULLN and LLN for nonstationary α - and ϕ -mixing random fields derived in Section 3 relax these restrictions.

However, the class of mixing random fields may not be sufficiently large to cover some applications of interest such as spatial ARMA processes, also called linear processes, and Cliff-Ord (1971, 1981) type spatial processes. These important spatial processes are generated as functions of some input process, which is usually assumed to be spatially mixing. The function may involve infinitely many spatial lags of the input process. As is well-known, while measurable functions of finite number of lags of a mixing process are also mixing, the mixing property is not preserved under transformations involving infinite number of lags. For instance, in the time series context, Andrews (1984) showed that a simple first-order autoregressive process fails to be α -mixing although its innovation/input process is independent, and hence mixing.

In general, linear processes whose input process is α -mixing (ϕ -mixing) will not be α -mixing (ϕ -mixing) without further restrictions on the probability density of the input process and the rate of decay of its mixing coefficients. Sufficient conditions for preservation α -mixing property under moving average transformations for time series processes were established by Gorodetskii (1977) and Withers (1981). These conditions were generalized to moving average random fields by Doukhan and Guyon (1991), see also Guyon (1995). They include invertibility of the moving average process, restrictions on the mixing coefficients and smoothness of the probability density function of the input process. Clearly, these conditions are difficult to verify, and may not

be satisfied in many applications.

Aside from linear processes, there are also important classes of processes generated as nonlinear transformations of mixing processes. Bernoulli shifts are one of important examples of such processes. In the time series literature, Bernoulli shifts are defined as nonlinear functions of an infinite history of some independent input process. Doukhan and Louhichi (1999) demonstrated that under some mild conditions, these processes are weakly dependent, i.e., their autocovariances decay sufficiently fast with the distance. Yet, they are not generally mixing although their input process is independent, see Rosenblatt (1980).

To accommodate a larger class of weakly dependent spatial processes, we extend the concept of near-epoch dependent (NED) processes used in the time series literature to random fields. The NED concept dates back to Ibragimov (1962), Billingsley (1968), Ibragimov and Linnik (1971), although these authors did not employ the present term. It has been used extensively in the time series literature by McLeish (1975a, 1975b), Wooldridge (1986), Gallant and White (1988), Andrews (1988), Hansen (1991), Davidson (1992, 1993a,b), Pötscher and Prucha (1997), and de Jong (1997).

The basic idea is that a NED process can be approximated by a function of only a finite number of spatial lags of the input process, which is assumed mixing. As a result, the approximating process inherits the mixing property of the input process. One can then use the limit theorems for mixing fields of Section 3 to establish a CLT and an LLN for the approximating process. Under some weak conditions, the approximation error can be shown to be asymptotically negligible in the sense that it does not affect the limiting

behavior of the NED process. Consequently, the NED process will satisfy a CLT or an LLN.

The NED property is compatible with considerable heterogeneity and dependence, and is preserved under infinite lag transformations under fairly general conditions. Therefore, it is convenient for modeling many processes encountered in applications. The class of NED random fields is sufficiently broad to cover many important applications. It includes, but is not limited to, mixing processes. It is also shown to cover Cliff-Ord type processes, spatial infinite moving average processes and some infinite nonlinear transformations of random fields. All these processes need not be mixing. Thus, the class of NED spatial processes is strictly larger than that of mixing random fields.

In this part of the dissertation, we provide a central limit theorem and law of large numbers for NED spatial processes. Just as the limit theorems of Section 3, the CLT and LLN for NED spatial processes accommodate nonstationary random fields with trending moments. They also allow for more general unevenly spaced index sets. All these attributes are critical in many applied settings where unevenly spaced data observations exhibit considerable heterogeneity and dependence.

To the best of our knowledge, NED processes have not been considered in the spatial literature. In the time series literature, CLTs for NED processes have been obtained by Wooldridge (1986), Wooldridge and White (1988), Davidson (1992, 1993a,b), and de Jong (1997). Interestingly, our CLT contains as a special case the CLTs for time-series NED processes of Wooldridge (1986) [Theorem 3.13 and Corollary 4.4] and Davidson's (1992). As such, the proposed CLT reveals direct connection and consistency in the asymptotic

properties of spatial (multi-dimensional) and time-series (one-dimensional) processes.

The LLN is an L_1 -law based on the subset of assumptions maintained in the CLT, which facilitates their joint application in the proof of consistency and asymptotic normality of spatial estimators. It requires the existence of absolute moments of order slightly greater than one. This is a reasonably mild assumption commonly used in weak LLNs for NED processes, for example, in Andrews (1988) and Davidson (1993b). Andrews (1988) establishes an L_1 -law for triangular arrays of L_1 -NED processes. Davidson (1993b) extends the latter result to processes with trending moments.

Our limit theorems for NED spatial processes can be used to develop an asymptotic theory of spatial econometric estimators for dependent nonstationary data-generating processes. In particular, these results should allow extension of the asymptotic theory for spatial GMM estimators proposed by Conley (1999) in two critical directions: (i) from mixing processes to the larger class of weakly dependent NED processes; and (ii) from stationary to nonstationary processes. Some progress in this direction has been made by the author. Furthermore, these limit theorems can be also used to study the large sample properties of cross-sectional and panel data models with cross-sectionally dependent observations when data locations and distances are known.

4.2 Definition and Examples of NED Processes

We continue with the basic set-up of Section 3 and consider spatial processes located on a possibly unevenly spaced lattice $D \subseteq \mathbb{R}^d$, $d \geq 1$. The distance between elements $i = (i_1, \dots, i_d)$ and $j = (j_1, \dots, j_d)$ of D is defined as $\rho(i, j) = \max_{1 \leq l \leq d} |j_l - i_l|$. As in Section 3, we assume that there is a minimum positive distance between any two elements of the lattice D , i.e., we maintain Assumption 1.

In this section, we will introduce the notion of near-epoch dependent (NED) random fields, which encompasses mixing random fields and many non-mixing weakly dependent random fields. NED processes have nice properties similar to mixing processes, e.g., stability under smooth transformations. More importantly, they will be shown to satisfy a CLT and LLN under some fairly mild conditions.

Although some weakly dependent spatial processes are not mixing, they can often be represented as functions of mixing processes. To fix ideas, suppose that the random field $Z = \{Z_{i,n}\}$ can be written as

$$Z_{i,n} = f_{i,n}(X_{j,n}, j \in D)$$

for some α -mixing field $X = \{X_{i,n}\}$ and a measurable function $f_{i,n}$. Observe that Z need not be α -mixing since it depends on an infinite number of spatial lags of X . However, if the functions $f_{i,n}$ are "well-behaved", this structure of dependence is often sufficient to derive limit theorems for $Z_{i,n}$. Intuitively, we can expect a CLT (or an LLN) to hold if functions $f_{i,n}$ are such that they put "declining weights" on the spatial lags of $X_{j,n}$ that are remote from point i , thus effectively ensuring that the behavior $Z_{i,n}$ is mainly driven by

$X_{j,n}$ located in this some bounded neighborhood of i . The idea of "declining weights" can be formalized by approximating $f_{i,n}(X_{j,n}, j \in D)$ by a measurable function that depends on *finitely* many spatial lags of X , e.g., by $h_{i,n}^s(X_{j,n}, j \in D: \rho(i, j) \leq s)$. Since measurable functions of finitely many lags of a mixing process are also mixing, the approximating function $h_{i,n}^s$ will inherit the mixing property of the process X , and hence one can apply the limit theorems for mixing random fields to establish the asymptotic behavior of $h_{i,n}^s$. Finally, if the approximating error can be made arbitrarily small (in some norm) by increasing the size of the neighborhood, i.e.,

$$\|Z_{i,n} - h_{i,n}^s\|_p \rightarrow 0 \text{ as } s \rightarrow \infty, \quad (5)$$

then $Z_{i,n}$ will satisfy a CLT and LLN under some reasonable regularity conditions. This is, in a nutshell, the basic idea behind approximating concepts used in the time series literature. Various approximating concepts mainly differ in the choice of the approximating functions $h_{i,n}^s$ and the measure for the approximating error.

We now extend this approximation concept to random fields. We will use the conditional expectations of $Z_{i,n}$ as the approximating functions $h_{i,n}^s$. More specifically, let $Z = \{Z_{i,n}, i \in D_n, n \geq 1\}$ and $X = \{X_{i,n}, i \in T_n, n \geq 1\}$ be two vector-valued arrays of random fields defined on a common probability space $(\Omega, \mathfrak{F}, P)$ and taking their values in \mathbb{R}^{p_z} and \mathbb{R}^{p_x} , respectively.¹² In the following, we assume that \mathbb{R}^{p_y} is a normed metric space equipped with the Euclidian norm: $|y| = [\sum_{k=1}^{p_y} y_k^2]^{1/2}$. For any random vector Y , let $\|Y\|_p = [E|Y|^p]^{1/p}$, $p \geq 1$, denote its L_p -norm. Finally, let $\mathfrak{F}_{i,n}(s) = \sigma(X_{j,n}; j \in T_n :$

¹²Note that the vectors Z and X may have different dimensions. For the approximation to be well-defined, it is assumed that $D_n \subseteq T_n$.

$\rho(i, j) \leq s$) be the σ -algebra generated by X 's located in the s -neighborhood of $Z_{i,n}$.

Definition 3 *Random field $Z = \{Z_{i,n}, i \in D_n, n \geq 1\}$ is said to be L_p -near-epoch dependent ($p \geq 1$) on the field $X = \{X_{i,n}, i \in T_n, n \geq 1\}$ with $D_n \subseteq T_n$, if there exist positive constants $\{d_{i,n}\}$ such that*

$$\|Z_{i,n} - E(Z_{i,n} | \mathfrak{F}_{i,n}(s))\|_p \leq d_{i,n} \psi(s) \quad (6)$$

for some non-increasing sequence $\psi(s)$ such that $\lim_{s \rightarrow \infty} \psi(s) = 0$. The $\psi(s)$ are referred to as the NED coefficients, and the $d_{i,n}$ - as the NED magnitude indices. Z is said to be L_p -NED on X of size $-\lambda$ if $\psi(s) = O(s^{-\mu})$ for some $\mu > \lambda > 0$. Furthermore, if $\sup_n \sup_{i \in D_n} d_{i,n} \leq c$, then Z is said to be uniformly L_p -NED on X .

Typically, T_n will be an infinite subset of D , and often $T_n = D$. But, to cover the practically important case of triangular arrays, and in particular Cliff-Ord type processes, T_n is allowed to depend on n and to be finite provided that it increases in size with n .

The magnitude indices $\{d_{i,n}\}$ account for processes with potentially trending moments. Thus, the NED property is compatible with considerable amount of heterogeneity. In many cases, $d_{i,n}$ can be chosen as $d_{i,n} \leq 4 \|Z_{i,n}\|_p$. This follows from application of Minkowski's and conditional Jensen's inequalities to the left-hand side of (6):

$$\begin{aligned} \left\| Z_{i,n} - \tilde{Z}_{i,n}^s \right\|_p &\leq \|Z_{i,n} - EZ_{i,n}\|_p + \left\| EZ_{i,n} - \tilde{Z}_{i,n}^s \right\|_p \\ &\leq 2 \|Z_{i,n} - EZ_{i,n}\|_p \leq 4 \|Z_{i,n}\|_p \end{aligned}$$

where $\tilde{Z}_{i,n}^s = E(Z_{i,n} | \mathfrak{F}_{i,n}(s))$. Thus, we can choose $d_{i,n}$ such that $d_{i,n} \leq 4 \|Z_{i,n}\|_p$, and consequently, assume $\psi(s) \leq 1$, with no loss of generality. Clearly, if the $Z_{i,n}$ are uniformly L_p -bounded, then $\sup_n \sup_{i \in D_n} d_{i,n} \leq c$ and, hence, Z is uniformly L_p -NED on X .

Using the convention $d_{i,n} = 4 \|Z_{i,n}\|_p$, one can also avoid the possibility that condition (6) might be trivially satisfied by choosing $d_{i,n}$ arbitrarily big so that $d_{i,n}^{-1} \|Z_{i,n} - \tilde{Z}_{i,n}^s\|_p$ becomes arbitrarily small for *any* process, even if it does not have the required NED property. This situation is ruled out by imposing the restriction $d_{i,n} \leq 4 \|Z_{i,n}\|_p$. Furthermore, note that by Lyapunov's inequality, if $Z_{i,n}$ is L_p -NED, then it is also L_q -NED with the same coefficients $\{d_{i,n}\}$ and $\{\psi(s)\}$ for any $q \leq p$.

To the best of our knowledge, the NED concept has not yet been considered in the random fields literature. In the time series literature, the idea first appeared in the works of Ibragimov (1962), Ibragimov and Linnik (1971), and Billingsley (1968), although they did use the present term. The concept of time series NED processes was later formalized by McLeish (1975a, 1975b), Wooldridge (1986), Gallant and White (1988). These authors considered only L_2 -NED processes. Andrews (1988) generalized it to L_p -NED processes for $p \geq 1$. Davidson (1992, 1993a,b, 1994) further extended it to allow for trending time series processes.

A more general concept of approximation for time series processes was introduced by Pötscher and Prucha (1997). They call the process $Z_{t,n}$ L_p -approximable by $X_{t,n}$ if there exist a function $h_{t,n}^s = h_{t,n}^s(X_{t-s,n}, \dots, X_{t+s,n})$ such that

$$\limsup_{n \rightarrow \infty} n^{-1} \sum_{t=1}^n \|Z_{t,n} - h_{t,n}^s\|_p \rightarrow 0 \text{ as } s \rightarrow \infty, \quad (7)$$

A similar approximation concept is used by Davidson (1994). Loosely speaking, he defines L_p -approximable time series processes as those satisfying condition (5). Since the $h_{i,n}^s$ can be chosen as the conditional expectations, L_p -approximable processes include L_p -NED processes as a special case. Clearly, the conditional expectation is one of candidates for the approximating function, but other approximating functions are also possible. However, in the case $p = 2$ the conditional mean is the best approximator in the sense that it minimizes the mean squared error. Therefore, the CLT in the next section is derived for L_2 -NED processes, and our use of the NED concept is not restrictive.

We will now show that some practically important classes of spatial processes have the NED property. More specifically, we establish the NED property for the following classes of processes: (i) infinite moving average random fields, (ii) Cliff-Ord type processes and (iii) spatial Bernoulli shifts. These processes are used widely in applications.

Example 1 *Infinite Moving Average (or Linear) Random Fields*

Consider an infinite moving average random field $Y = \{Y_i, i \in D = \mathbb{Z}^d\}$ defined as:

$$Y_i = \sum_{j \in \mathbb{Z}^d} g_{i,j} X_j \tag{8}$$

where $X = \{X_i, i \in \mathbb{Z}^d\}$ is a vector-valued random field and $g_{i,j}$ are some real numbers. This class of linear random fields on \mathbb{Z}^d were studied by Doukhan and Guyon (1991), see also Doukhan (1994). The linear spatial processes on \mathbb{Z}^2 considered in the well-known paper by Whittle (1954) is a special case

of the linear field (8). Just as in the time series literature, Y_i is defined as the limit of the partial sum process $Y_i^s = \sum_{j \in \mathbb{Z}^d, \rho(i,j) \leq s} g_{i,j} X_j$, $s \in \mathbb{N}$, which exists under some weak conditions. More specifically, Doukhan (1994), pp. 75-81, proves the following lemma.

Lemma 2 [Doukhan, 1994] *The distribution of random field Y in (8) is well defined under the following assumptions:*

(a) $X = \{X_i, i \in \mathbb{Z}^d\}$ is uniformly L_p ($p \geq 1$) bounded, i.e.,

$$\sup_{i \in \mathbb{Z}^d} \|X_i\|_p < \infty \quad (9)$$

(b)

$$\limsup_{s \rightarrow \infty} \sup_{i \in \mathbb{Z}^d} \sum_{j \in \mathbb{Z}^d, \rho(i,j) > s} |g_{i,j}| = 0 \quad (10)$$

Moreover, the finite dimensional distributions of Y are limits of the those of the random fields $Y_i^s = \sum_{j \in \mathbb{Z}^d, \rho(i,j) \leq s} g_{i,j} X_j$.

We show that under the same conditions, the linear field Y is L_p -NED on the field X .

Lemma 3 *Under conditions (9)-(10) maintained in its definition, the linear field Y in (8) is uniformly L_p -NED on the field X .*

The nice feature of this result is that to verify the NED property, it does not impose any additional conditions over and above those incorporated in the definition of the linear field. Assumption (10) is satisfied if $|g_{i,j}| = O(|j|^{-\gamma})$ uniformly in i for some $\gamma > d$. In the case $d = 1$, the latter

condition reduces to $|g_{i,j}| = O(j^{-1-\epsilon})$, uniformly in i for some $\epsilon > 0$, which is a standard sufficient condition for the absolute convergence of series.

Example 2 *Cliff-Ord Type Spatial Processes*

Consider the following Cliff-Ord type model:

$$\begin{aligned} Y_n &= \lambda M_n Y_n + Z_n \beta + u_n \\ u_n &= \rho W_n u_n + \varepsilon_n \end{aligned}$$

where $Y_n = (Y_{1,n}, \dots, Y_{n,n})$ is n -vector of endogenous variables, Z_n is $n \times k$ matrix of regressors, M_n and W_n are known $n \times n$ weight matrices that generally depend on the sample size, λ, ρ are unknown scalar parameters and β is unknown k -vector of slope coefficients. The ε_n are n -vector of disturbances. The reduced form of the model is

$$Y_n = (I_n - \lambda M_n)^{-1} Z_n \beta + (I_n - \lambda M_n)^{-1} (I_n - \rho W_n)^{-1} \varepsilon_n,$$

Assume for simplicity that Z_n is a column vector. Let $X_{i,n} = (Z_{i,n}, \varepsilon_{i,n})'$ be uniformly (in i and n) L_p -bounded for some $p \geq 1$. Note that $X_{i,n}$ need not be independent. The reduced form of the model is

$$Y_{i,n} = \beta \sum_{j=1}^n a_{ij,n} Z_{j,n} + \sum_{j=1}^n b_{ij,n} \varepsilon_{j,n}, \quad i = 1, \dots, n \quad (11)$$

where $A_n = (a_{ij,n}) = (I_n - \lambda M_n)^{-1}$ and $B_n = (b_{ij,n}) = (I_n - \lambda M_n)^{-1} (I_n - \rho W_n)^{-1}$.

Although for *fixed* n , the output process $Y_{i,n}$ depends on only finite number of spatial lags of the input process $X_{i,n}$, the mixing property of $X_{i,n}$ may

not carry over to $Y_{i,n}$. The reason is that the number of spatial lags grows unboundedly with the sample size so that the mixing property can break down in the limit. This is especially important when analyzing the asymptotic properties of Cliff-Ord type processes.

Observations in Cliff-Ord models are typically indexed by natural numbers, e.g. $i, j \in \{1, \dots, n\}$. Although coordinates/locations in \mathbb{R}^d corresponding to various observations are not explicitly specified, the distances between observations are often known. They are used to construct the weighting matrices W_n and M_n .

We will now show that despite the lack of explicitly specified locations, the NED concept can be applied to Cliff-Ord type processes provided that the distances between observations are known. Suppose observations reside in some bounded region D_n of the lattice D satisfying Assumption 1, with the sample size $|D_n| = n$, and also suppose that the distances $\rho(i, j)$ between any two observations i and j are known. [Strictly speaking, one has to write $\rho(l_i, l_j)$ instead of $\rho(i, j)$ where $l_i, l_j \in \mathbb{R}^d$ are locations corresponding to observations i and j]. We emphasize that one does not have to know the sample region D_n and the locations since our results hold for *arbitrary* D_n and do *not* depend on ordering of observations, as shown below.

It turns out that the Cliff-Ord process (11) will satisfy the NED property under the following relatively weak conditions.

$$\begin{aligned} \lim_{s \rightarrow \infty} \sup_n \sup_{1 \leq i \leq n} \sum_{1 \leq j \leq n: \rho(i,j) > s} |a_{ij,n}| &= 0 \\ \lim_{s \rightarrow \infty} \sup_n \sup_{1 \leq i \leq n} \sum_{1 \leq j \leq n: \rho(i,j) > s} |b_{ij,n}| &= 0 \end{aligned} \tag{12}$$

and

$$\sup_n \sup_{1 \leq i \leq n} \|X_{i,n}\|_p < \infty. \quad (13)$$

Lemma 4 *Under conditions (12)-(13), the process $Y_{i,n}$ given by (11) is uniformly L_p -NED on the process $X_{i,n} = (Z_{i,n}, \varepsilon_{i,n})'$.*

Condition (12) implies the absolute summability of the weighting coefficients, i.e.,

$$\begin{aligned} \sup_n \sup_{1 \leq i \leq n} \sum_{1 \leq j \leq n} |a_{ij,n}| &< \infty \\ \sup_n \sup_{1 \leq i \leq n} \sum_{1 \leq j \leq n} |b_{ij,n}| &< \infty \end{aligned} \quad (14)$$

Since any re-arrangement of the terms of an absolutely convergent series converges to the same limit, the ordering of observations does not affect the NED property. Hence, one does not have to know the locations, and observations may be indexed in arbitrary way by naturals.

Condition (12) is satisfied if uniformly in i and n : $|a_{ij,n}| = O(j^{-\gamma})$ for some $\gamma > 1$, which can be verified by simple calculations as follows:

$$\sup_n \sup_{1 \leq i \leq n} \sum_{1 \leq j \leq n: \rho(i,j) > s} |a_{ij,n}| \leq C \int_s^\infty x^{-\gamma} dx \leq C' s^{1-\gamma}.$$

for some finite constants C and C' . Hence, condition (12) follows immediately if $\gamma > 1$.

We note that condition (12) on the weights of the Cliff-Ord type process is analogous to those used by Kelejian and Prucha (1998, 1999, 2001, 2004, 2007a,b). In particular, Kelejian and Prucha require the *row* sums of the weighting matrices $(a_{ij,n})$ and $(b_{ij,n})$ to be uniformly bounded, i.e., maintain

condition (14), which is implied by condition (12). So, condition (12) is slightly stronger than uniform boundedness of row sums of the weighting matrices. However, in addition, Kelejian and Prucha require the *column* sums of the weighting matrices to be also uniformly bounded. The latter condition is not imposed here. Thus, overall, condition (12) and the set of conditions used by Kelejian and Prucha are similar, but neither of them dominates the other. The uniform boundedness of row and column sums of the weighting matrices is the standard assumption maintained in the Cliff-Ord literature, e.g., Lee (2002, 2004, 2007a), Kapoor, Kelejian, and Prucha (2007). Thus, the NED concept fits quite naturally into the existing Cliff-Ord literature.

Example 3 *Spatial Bernoulli Shifts*

Consider a real-valued random field $Y = \{Y_i, i \in D\}$ defined as:

$$Y_i = H(X_{i+j}, j \in D) \tag{15}$$

where $D \subset \mathbb{R}^d$ is a lattice satisfying Assumption 1, $X = \{X_i, i \in D\}$ is a real-valued random field and $H : \mathbb{R}^D \rightarrow \mathbb{R}$ is a measurable function. (Note that H is a function of countably infinite number of scalar arguments u). The process (15) generalizes the one-dimensional Bernoulli shift process studied by Doukhan and Louhichi (1999) to random fields. A simple example of infinitely dependent Bernoulli shift is the infinite moving average (or linear) random field discussed in Example 1. In more general cases, H may be a complicated nonlinear function. Following Doukhan and Louhichi (1999), we

assume that H satisfies the following Lipschitz-type regularity condition:

$$|H(u_i, i \in D) - H(v_i, i \in D)| \leq \sum_{i \in D} w_i |u_i - v_i| \quad (16)$$

for some positive constants $\{w_j, j \in D\}$ such that

$$\lim_{s \rightarrow \infty} \sup_{i \in D} \sum_{j \in D: |j| > s} w_{i+j} = 0. \quad (17)$$

Finally, we assume that the innovation process $X = \{X_i, i \in D\}$ has uniformly bounded second moments, i.e.,

$$\sup_{i \in D} \|X_i\|_2 < \infty \quad (18)$$

Then, one can establish the following result.

Lemma 5 *Under conditions (16)-(18), the random field $Y = \{Y_i, i \in D\}$ defined by (15) is uniformly L_2 -NED on the random field $X = \{X_i, i \in D\}$.*

Conditions (16)-(18) are analogous to those used by Doukhan and Louhichi (1999), p. 324 for time-series Bernoulli shifts. Condition (16) is fulfilled if the function H is differentiable in each of its arguments and its partial derivatives with respect to each argument are bounded as follows:

$$\left| \frac{\partial H}{\partial u_i} \right| \leq w_i.$$

Condition (17) is in turn satisfied if $w_i = O(|i|^{-\gamma})$ for some $\gamma > d$.

To summarize, the class of NED random fields covers not only linear functions of mixing fields but also more general infinite-lag nonlinear transformations of mixing random fields under reasonably weak conditions.

4.3 Central Limit Theorem for NED Processes

In this section, we provide a CLT for arrays of L_2 -NED random fields. Let $Z = \{Z_{i,n}, i \in D_n, n \geq 1\}$ be a real-valued random field, which is L_2 -NED on a vector-valued field $X = \{X_{i,n}, i \in T_n, n \geq 1\}$ with the NED coefficients $\{\psi(s)\}$ and the magnitude indices $\{d_{i,n}\}$, where $D_n \subseteq T_n \subseteq D$ and the lattice D satisfies Assumption 1. In the following, we will use the following notation and conventions for the field Z :

$$EZ_{i,n} = 0; S_n = \sum_{i \in D_n} Z_{i,n}; \sigma_n^2 = \text{var}(S_n).$$

As for the input process X , we assume that X is either α - or ϕ -mixing. We employ the same definition of mixing as in Section 3, see Definition 1. According to this definition, the mixing coefficients of X - $\bar{\alpha}(k, l, r)$ or $\bar{\phi}(k, l, r)$ - depend not only on the distance, r , between two datasets but also sizes of the two index sets, k and l . As discussed in Section 3, mixing conditions with both $k = \infty$ and $l = \infty$ are restrictive. Bradley (1993) shows that for stationary random fields the condition $\lim_{r \rightarrow \infty} \bar{\alpha}(\infty, \infty, r) = 0$ is equivalent to ρ -mixing, which is a more restrictive form of dependence. Furthermore, stationary random fields satisfying the condition $\lim_{r \rightarrow \infty} \bar{\beta}(\infty, \infty, r) = 0$ (or $\lim_{r \rightarrow \infty} \bar{\phi}(\infty, \infty, r) = 0$) reduce to r -dependent processes, which is a trivial form of dependence. Nevertheless, this difficulty can be overcome if the size of at least one of the index sets in the mixing coefficient is finite. Bradley (2005), p.315, remarks: *These pitfalls can be avoided if in the definition of the dependence coefficients, at least one of the two index sets is finite and its cardinality plays a suitable role. Indeed, in the formulation of strong mixing conditions for random fields, that has been common practice at least since the*

paper of Dobrushin (1968a).

Therefore, researchers have employed mixing conditions of the following type:

$$\begin{aligned}\bar{\alpha}(k, l, m) &\leq f(k, l)\hat{\alpha}(m) \\ \bar{\phi}(k, l, m) &\leq f(k, l)\hat{\phi}(m)\end{aligned}\tag{19}$$

where $f(k, l)$ is some non-decreasing in both arguments function (which will be specified later) and $\hat{\alpha}(m)$ and $\hat{\phi}(m)$ are non-increasing functions. The idea is to account separately for two different aspects of dependence: (i) decay of dependence with the distance, and (ii) accumulation of dependence with the growth of the sample size. To derive limit theorems, researchers have further specialized the function $f(k, l)$. One of most common choices is $f(k, l) = (k + l)^\tau$ for some $\tau \geq 0$, e.g. Neaderhouser (1978a,b), Takahata (1983), Nahapetian (1987, 1991), Bulinskii (1989), Bulinskii and Doukhan (1990).

We follow this approach and assume that the mixing coefficients of the input random field X satisfy the following two sets of assumptions.

Assumption 9 (*α -mixing*) *The uniform α -mixing coefficients of X satisfy*

(a)

$$\bar{\alpha}(k, l, m) \leq (k + l)^\tau \hat{\alpha}(m)\tag{20}$$

for some $\tau \geq 0$ and $\hat{\alpha}(m)$ such that $\sum_{m=1}^{\infty} m^{d(\tau+1)-1} \hat{\alpha}^{\frac{\delta}{2(2+\delta)}}(m) < \infty$.

(b) for each given k and some $\varepsilon > 0$:

$$\bar{\alpha}(k, \infty, m) = O(m^{-d-\varepsilon})\tag{21}$$

Assumption 10 (*ϕ -mixing*) *The uniform ϕ -mixing coefficients of X satisfy*

(a)

$$\bar{\phi}(k, l, m) \leq (k + l)^\tau \hat{\phi}(m) \quad (22)$$

for some $\tau \geq 0$ and $\hat{\phi}(m)$ such that $\sum_{m=1}^{\infty} m^{d(\tau+1)-1} \hat{\phi}(m)^{(1+\delta)/(2+\delta)} < \infty$.

(b) $\bar{\phi}(k, \infty, m) = O(m^{-d-\varepsilon})$ for each given k and some $\varepsilon > 0$.

Mixing conditions of this type have been used extensively in the random fields literature. In particular, conditions (20) and (22) have been employed by Neaderhouser (1978a,b), Takahata (1983), Nahapetian (1987, 1991), Bulinskii (1989), Bulinskii and Doukhan (1990). Conditions similar to (21) have been exploited by Neaderhouser (1981), Tran (1990), Carbon, Tran and Wu (1997).

Unlike mixing conditions based on $\bar{\alpha}(\infty, \infty, m)$ or $\bar{\phi}(\infty, \infty, m)$, the above conditions are satisfied by large classes of random fields encountered in applications. For instance, Dobrushin (1968a) provides examples of Gibbs fields used widely in statistical physics that satisfy the α -mixing condition (21), but not the condition $\lim_{m \rightarrow \infty} \bar{\alpha}(\infty, \infty, m) = 0$. Bradley (1993) gives examples of random fields satisfying condition (20) with $k = l$ and $\tau = 1$. Bulinskii (1989) constructs, for any given $\hat{\alpha}(m)$, infinite moving average random fields satisfying the α -mixing condition (20) with $\tau = 1$. The standard mixing coefficients used in the time series literature are also covered by conditions (20) and (21) by setting $\tau = 0$ and $d = 1$.

The CLT relies on the same moment conditions as the CLT for mixing random fields in Section 3, see Theorem 1. For ease of reference, we restate

them below.

Assumption 11 (*Uniform $L_{2+\delta}$ integrability*) *There exists an array of positive constants $\{c_{i,n}\}$ and a $\delta \geq 0$ such that*

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

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

Assumption 12 $\liminf_{n \rightarrow \infty} |D_n|^{-1} M_n^{-2} \sigma_n^2 > 0$, where $M_n = \max_{i \in D_n} c_{i,n}$ and $\{D_n\}$ is a sequence of arbitrary finite subsets of D such that $|D_n| \rightarrow \infty$ as $n \rightarrow \infty$.

Finally, we need to control the sizes of the NED coefficients and NED magnitude indices:

Assumption 13 *NED coefficients satisfy $\sum_{m=1}^{\infty} m^{d-1} \psi(m) < \infty$.*

Assumption 14 *NED magnitude indices satisfy $\sup_n \sup_{i \in D_n} c_{i,n}^{-1} d_{i,n} \leq C < \infty$.*

We can now state our CLT for L_2 -NED random fields.

Theorem 4 *Suppose $\{D_n\}$ is a sequence of arbitrary finite subsets of the lattice D , satisfying Assumption 1, with $|D_n| \rightarrow \infty$ as $n \rightarrow \infty$. Let T_n be a sequence of subsets of D such that $D_n \subseteq T_n$. Let $Z = \{Z_{i,n}, i \in D_n, n \geq 1\}$ be an array of real-valued centered random fields, which is L_2 -NED on $X = \{X_{i,n}, i \in T_n, n \geq 1\}$ with the NED coefficients $\{\psi(s)\}$ satisfying Assumption 13 and magnitude indices $\{d_{i,n}\}$ satisfy Assumption 14. Suppose that either*

- (a) X is α -mixing satisfying Assumptions 9a-b, and Z satisfies Assumption 11 with $\delta > 0$, or
- (b) X is ϕ -mixing satisfying Assumptions 10a-b and Z satisfies Assumption 11 with $\delta \geq 0$.

If Assumption 12 holds, then

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

Clearly, the CLT can be extended to vector-valued fields using the standard Cramér-Wold device.

Assumptions 11 and 12 are identical to those of the CLT for mixing random fields in Section 3. Similar conditions have been used in the time-series literature by Wooldridge (1986), Wooldridge and White (1988), Davidson (1992, 1993a,b), and de Jong (1997).

Assumption 11 is satisfied if the $Z_{i,n}/c_{i,n}$ are uniformly L_r -bounded for some $r > 2 + \delta$, i.e., $\sup_n \sup_{i \in D_n} \|Z_{i,n}/c_{i,n}\|_r < \infty$. The nonrandom constants $c_{i,n}$ allow for processes with asymptotically unbounded (trending) moments. They can be thought of as upper bounds on the moments of the individual terms, i.e., $\|Z_{i,n}\|_r \leq c_{i,n} < \infty$. In the case of uniformly L_r -bounded variables, i.e., when $\sup_n \sup_{i \in D_n} \|Z_{i,n}\|_r \leq M < \infty$, constants $c_{i,n}$ can be set to 1, without loss of generality.

In the case of asymptotically unbounded moments, there is no finite uniform upper bound M on the moments. For example, such behavior is exhibited by some linear processes $Z_{i,n} = a_{i,n} X_{i,n}$, where $\sup_n \sup_{i \in D_n} \|X_i\|_{2+\delta} < \infty$ and $a_{i,n}$ are nonrandom constants increasing unboundedly with n . This

example suggests that it is often possible to choose the scaling constants $c_{i,n}$ so that the $Z_{i,n}/c_{i,n}$ are uniformly $L_{2+\delta}$ -bounded. For instance, this can be done by setting $c_{i,n} = \|Z_{i,n}\|_{2+\delta}$.

To obtain a CLT, the asymptotic behavior of individual terms' moments needs to be further restricted by some kind of asymptotic negligibility condition, which rules out situations in which individual summands influence disproportionately the entire sum. Assumption 12 serves precisely this purpose by limiting the growth behavior of moments. In the case of uniformly $L_{2+\delta}$ -bounded fields, Assumption 12 reduces to $\liminf_{n \rightarrow \infty} |D_n|^{-1} \sigma_n^2 > 0$, cp. Bolthausen (1982) and Guyon (1995).

We now illustrate Assumption 12 with two examples of random fields with asymptotically unbounded moments. Assumption 12 is satisfied in the first example, and is violated in the second example. First, let $\{Z_i, i \in D = \mathbb{N}^d\}$ be an independent random field with Z_i uniformly distributed on $[-|i|^\gamma; |i|^\gamma]$ for some $\gamma > 0$ and consider the sums $S_n = \sum_{D_n} Z_i$ on $D_n = [1; n]^d$. Then, $|D_n| = n^d$, $EZ_i^2 = 3^{-1}|i|^{2\gamma}$, $M_n^2 = \max_{i \in D_n} EZ_i^2 = 3^{-1}n^{2\gamma}$ and $\sigma_n^2 \sim n^{(2\gamma+d)}$. Clearly, Assumption 12 is satisfied for all $\gamma > 0$ in this example.

Second, consider an independent random field $\{Z_i, i \in \mathbb{N}^d\}$ with Z_i uniformly distributed on $[-2^{g_i}; 2^{g_i}]$ where $g_i = \sum_{p=1}^d i_p$. In this case, $EZ_i^2 = 3^{-1}4^{g_i}$, $M_n^2 = 3^{-1}4^{dn}$ and $\sigma_n^2 = 3^{-1-d}4^d(4^n - 1)^d$, and hence

$$\liminf_{n \rightarrow \infty} |D_n|^{-1} M_n^{-2} \sigma_n^2 = 0.$$

Thus, Assumption 12 is violated. It can easily be shown that the Lindeberg condition is also violated.

Assumptions 9 and 10 restrict the dependence structure of the input process X . They reflect the usual trade-off between the moment and mixing

conditions: lower moment conditions are associated with faster rates of decay of mixing coefficients. In addition, they capture the trade-off between decay of dependence with the distance, on the one-hand, and accumulation of dependence with the growth of the sample region, on the other hand.

Assumptions 9 and 10 are stronger than the mixing assumptions of the CLT for mixing fields, Theorem 1 of Section 3. First, the rate of decrease of the α -mixing coefficients with the distance in Theorem 4 is twice as faster as that in Theorem 1. These rates are the same for the ϕ -mixing case.

Second, in contrast to Theorem 1, Theorem 4 accounts explicitly for potential accumulation of dependence with the expansion of index sets. More specifically, mixing coefficients in Assumptions 9a and 10a are assumed to increase at the rate of τ in the cardinalities of index sets, while no such rates are assumed in Theorem 1. This strengthening of the mixing conditions is necessitated by the transition from mixing to NED random fields. Intuitively, this can be explained as follows. Recall that $Z_{i,n}$ need not be mixing, but can be approximated sufficiently well by X 's located in the s -neighborhood of $Z_{i,n}$. Under Assumption 1, the s -neighborhood of any point on the lattice D contains at most $(2[s] + 2)^d$ points of D (for proof see Lemma A.1(ii) in Appendix A). Therefore, to control for dependence between, say, $Z_{i,n}$ and $Z_{j,n}$, one has to check dependence between their approximating functions, each of which involves $(2[s] + 2)^d$ spatial lags of X . This leads to mixing coefficients of the type $\bar{\alpha}_X(k, k, h)$ with $k = (2s + 2)^d$, in which the cardinalities of index sets increase with s . Therefore, the mixing coefficients of the input field have to decline at a faster rate to compensate for the accumulation of dependence with the increase in s . Hence, we have the above mentioned

trade-off. In contrast, if Z is itself mixing, the dependence between any $Z_{i,n}$ and $Z_{j,n}$ is measured by $\bar{\alpha}_Z(1, 1, h)$ so that there is no need to account for the cardinalities of index sets.

Interestingly, for the case $d = 1$, Assumptions 9 and 10 imply the same rates of the decay of α - and ϕ -mixing coefficients as in the CLTs of Wooldridge (1986) for time series NED processes, see Theorem 3.13 and Corollary 4.4. In the time series case, mixing coefficients do not depend on sizes of index sets, and hence, $\tau = 0$. Setting $d = 1$ and $\tau = 0$ in Assumptions 9a and 10a gives $\sum_{m=1}^{\infty} \hat{\alpha}^{\frac{\delta}{2(2+\delta)}}(m) < \infty$ and $\sum_{m=1}^{\infty} \hat{\phi}(m)^{(1+\delta)/(2+\delta)} < \infty$, which are analogous to the conditions exploited by Wooldridge (1986).

Our ϕ -mixing condition in the case $d = 1$ is weaker than that in Davidson (1992), who requires $\sum_{m=1}^{\infty} \hat{\phi}(m)^{\delta/(2+\delta)} < \infty$. At the same time, our α -mixing conditions are slightly stronger compared to Davidson (1992) and de Jong (1997), who assume $\sum_{m=1}^{\infty} \hat{\alpha}^{\frac{\delta}{2(2+\delta)}}(m) < \infty$. This is due to the fact that these authors exploit the concept of mixingales for $d = 1$ and the related sharper inequalities. For $d > 1$, the concept of mixingales is not well-defined, and therefore, we cannot take advantage of the mixingale inequalities.

Assumption 13 controls for the size of the NED coefficients. The NED coefficients measure the error in the approximation of $Z_{i,n}$ by X . Intuitively, for a CLT to hold, the approximation errors have to decline sufficiently fast with each successive approximation. This idea is reflected in Assumption 13. It is satisfied if $\psi(m) = O(m^{-d-\gamma})$ for some $\gamma > 0$, i.e., if the size of the NED coefficients is $-d$, according to Definition 3. When $d = 1$, the required NED size is -1 , which is precisely the assumption maintained by Wooldridge (1986) and Davidson (1992).

Lastly, Assumption 14 is a technical condition, which ensures that the magnitudes of $2 + \delta$ moments and the NED magnitude indices grow at the same rate as the sample size increases. As discussed earlier, $c_{i,n}$ are, in most cases, chosen as $c_{i,n} = \|Z_{i,n}\|_{2+\delta}$, and the NED magnitude indices $d_{i,n}$ are usually chosen as $d_{i,n} = 4 \|Z_{i,n}\|_2$. By Lyapunov's inequality, $\|Z_{i,n}\|_2 \leq \|Z_{i,n}\|_{2+\delta}$. Hence, Assumption 14 is automatically satisfied. It has also been used by de Jong (1997) and Davidson (1992) in the time series context.

Theorem 4 is applicable to one-dimensional processes. It contains as a special case some of the CLTs for time series NED processes. In particular, it generalizes Theorem 3.13 and Corollary 4.4 of Wooldridge (1986). Our CLT also contains the ϕ -mixing part of Davidson's (1992) CLT. In the spatial context, Theorem 4 extends the CLT for α - or ϕ -mixing random fields of Section 3 to a larger class of weakly dependent random fields.

4.4 Law of Large Numbers for NED Processes

In the previous section, we established a CLT for NED random fields. We now give a LLN which holds under a subset of the assumptions used in that CLT. Thus, the two theorems can be used jointly in the proof of consistency and asymptotic normality of spatial estimators.

The LLN is an L_1 -norm LLN for L_1 -NED random fields. It relies on the following set of moment and mixing assumptions.

Assumption 15 *There exist nonrandom positive constants $\{c_{i,n}, i \in D_n, n \geq 1\}$ such that $Z_{i,n}/c_{i,n}$ is uniformly L_p -bounded for some $p > 1$, i.e.,*

$$\sup_n \sup_{i \in D_n} E |Z_{i,n}/c_{i,n}|^p < \infty.$$

Assumption 16 $\bar{\alpha}(k, l, m) \leq f(k, l)\hat{\alpha}(m)$ for some non-decreasing function $f(\cdot, \cdot)$ and $\hat{\alpha}(m)$ such that $\sum_{m=1}^{\infty} m^{d-1}\hat{\alpha}(m) < \infty$

Assumption 17 $\bar{\phi}(k, l, m) \leq f(k, l)\hat{\phi}(m)$ for some non-decreasing function $f(\cdot, \cdot)$ and $\hat{\phi}(m)$ such that $\sum_{m=1}^{\infty} m^{d-1}\hat{\phi}(m) < \infty$.

Theorem 5 Let $\{D_n\}$ be a sequence of arbitrary finite subsets of D such that $|D_n| \rightarrow \infty$ as $n \rightarrow \infty$, where $D \subset \mathbb{R}^d$, $d \geq 1$ is as in Assumption 1, and let T_n be a sequence of subsets of D such that $D_n \subseteq T_n$. Suppose that $Z = \{Z_{i,n}, i \in D_n, n \geq 1\}$ satisfies Assumption 15 and is L_1 -NED on $X = \{X_{i,n}, i \in T_n, n \geq 1\}$ with magnitude indices $d_{i,n}$. If $X_{i,n}$ is either

(a) α -mixing satisfying Assumption 16, or

(b) ϕ -mixing satisfying Assumption 17, then

$$a_n^{-1} \sum_{i \in D_n} (Z_{i,n} - EZ_{i,n}) \xrightarrow{L_1} 0,$$

where $a_n = M_n |D_n|$ and $M_n = \max_{i \in D_n} \max(c_{i,n}, d_{i,n})$.

Of course, L_1 -convergence implies convergence in probability, and hence the $Z_{i,n}$ also satisfy a weak law of large numbers. We also note that Z and X can be vector-valued random fields, possibly of different dimensions. In this case, $|Z_{i,n}|$ should be understood as the Euclidean norm in the respective vector-space.

Assumption 15 is a standard moment assumption employed in weak laws of large numbers for dependent processes. It requires existence of absolute moments of order slightly greater than 1, which in turn implies existence of first moments. Clearly, this moment assumption is weaker than that of the

CLT for NED random fields, Theorem 4. However, it is slightly stronger than its counterpart in the LLN for mixing random fields of Section 3.4.2. The latter theorem relies on uniform integrability, which is implied by uniform L_p -boundedness for some $p > 1$. Strengthening of the moment condition can be explained by weakening of the restrictions on the dependence structure of the random field: the NED condition is weaker than mixing conditions.

As in Theorem 4, $c_{i,n}$ and $d_{i,n}$ are the normalizing constants that reflect the magnitudes of potentially trending moments. They can be chosen as $\|Z_{i,n}\|_p$. The case of variables with uniformly bounded moments is covered by setting $c_{i,n} = d_{i,n} = 1$.

Assumptions 16 and 17 restrict the dependence structure of the mixing input field X . These conditions are weaker than the mixing assumptions maintained in the CLT for NED fields. First, the mixing coefficients in Assumptions 16 and 17 may decrease with the distance at a slower rate than those in Assumptions 9 and 10. Second, there is no loss associated with the cardinalities of index sets. Recall that to obtain the CLT, we had to impose specific structure on the functions $f(k, l)$ and to account for the growth of this function in k and l . It turns out that the cardinalities of index sets do not play the same role in the LLN. This is not surprising since LLNs are weaker results than CLTs. Furthermore, note that in contrast to the CLT for NED fields, the LLN does not require any assumptions with respect to the size of the NED coefficients.

In the time series literature, weak LLNs for NED processes have been obtained by Andrews (1988) and Davidson (1993b), among others. Andrews (1988) derives an L_1 -law for triangular arrays of L_1 -mixingales. He then

shows that NED processes are L_1 -mixingales, and hence, satisfy his LLN. Davidson (1993b) extends the latter result to processes with trending moments. The mixingale concept, which exploits the natural order of the time line, is weaker than that of mixing. It allows these authors to circumvent restrictions on the mixingale sizes, i.e., the rates at which dependence declines. 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 NED random fields with a mixing input process. Due to the higher dimensionality and unevenness of the index sets, we have to restrict the rates of decay of mixing coefficients with the distance.

5 Conclusion

The dissertation develops an asymptotic theory for spatial processes exhibiting considerable heterogeneity and dependence. More specifically, it derives new central limit theorems, uniform and pointwise laws of large numbers for arrays of weakly dependent random fields that can be readily used to establish the asymptotic properties of spatial estimators in many socioeconomic models. Relative to the existing literature, the contribution of the dissertation is threefold. First, the proposed limit theorems accommodate nonstationary random fields with asymptotically unbounded or trending moments. Second, they cover a larger class of weakly dependent random fields than mixing random fields. Third, they allow for arrays of fields located on unevenly spaced lattices in \mathbb{R}^d , and place minimal restrictions on the configuration and growth behavior of index sets.

All these features are critical for many econometric applications. Clearly, processes encountered in applications are often nonstationary, and in particular, heteroscedastic. Sometimes, their second moments may grow unboundedly or trend as the index set expands. This form of nonstationarity may lead to violation of the asymptotic negligibility condition essential for CLTs, and therefore, needs to be checked. Furthermore, some weakly dependent processes do not generally satisfy the mixing property, e.g., linear random fields and Cliff-Ord type spatial processes used widely in applications. Therefore, limit theorems that cover not only mixing random fields but also more general weakly dependent random fields are required. This goal is achieved by considering the class of near-epoch-dependent random fields which is richer than that of mixing random fields. The limit theorems for NED random fields generalize nicely their one-dimensional counterparts in the time series literature.

In contrast to the previous results, the proposed limit theorems allow for random fields located on unevenly spaced lattices and sampled over regions of arbitrary configuration, which significantly facilitates their application in socioeconomic models. In addition, each of the theorems is supplied with low-level sufficient conditions which are fairly easy to verify in applications.

Central limit theorems, uniform and pointwise laws of large numbers are the fundamental building blocks for the asymptotic theory of statistical estimators, which in turn serves as the basis for statistical inference. As such, our limit theorems can be used to establish consistency and asymptotic normality of estimators and tests statistics in a wide range of nonlinear spatial models with nonstationary dependent data-generating processes.

A Appendix: Cardinalities of Basic Sets on Irregular Lattices

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 $i = (i_1, \dots, i_d) \in \mathbb{R}^d$ let

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

denote, respectively, the half-open and closed unitary cubes with "south-west" corner i . Note that given the metric, $[i, i + 1] = B(j, 1/2)$, i.e., is the ball centered at j of radius $1/2$; where $j = (i_1 + 1/2, \dots, i_d + 1/2)$.

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 $h \geq 1$*

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

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

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

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

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

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

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

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 half-open unitary cubes. Hence, in light of Lemma A.1(i)

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

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

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

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

(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 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 $B(i, m + 1) \setminus B(i, m - 1)$ is

given by

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

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

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

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

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 $N_i(1, 1, m) \leq C m^{d-1}$, $C < \infty$. 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 $N_i(1, 1, m) |V| \leq C m^{d-1} |V|$. Furthermore, one can form at most $|U|$ different two-element subsets of U that include i . Hence, $N_i(2, 2, m) \leq N_i(1, 1, m) |V| |U| \leq C m^{d-1} |V| |U|$. Thus, $\sup_{i \in U} N_i(2, 2, m) \leq C m^{d-1} |U| |V|$, where C does not depend on i . \blacksquare

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 $N_i(1, 1, m) \leq C m^{d-1}$, $C < \infty$. 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: $N_i(1, 3, m) \leq N_i(1, 1, m) |V|^2 \leq C m^{d-1} |V|^2$. Since C does not depend on i furthermore $\sup_{i \in R^d} N_i(1, 3, m) \leq C m^{d-1} |V|^2$. ■

B Appendix: Proof of CLT for Mixing Processes

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) $\lim_{n \rightarrow \infty} \int (\mathbf{i}\lambda - y) \exp(\mathbf{i}\lambda y) \mu_n(dy) = 0$ 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 , $n = 1, 2, \dots$ and V_{nk} , $k = 1, 2, \dots$; $n = 1, 2, \dots$, be random vectors such that*

- (i) $V_{nk} \implies V_k$ as $n \rightarrow \infty$ for each $k = 1, 2, \dots$;
- (ii) $V_k \implies V$ as $k \rightarrow \infty$, and

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

Then $Y_n \implies V$ as $n \rightarrow \infty$.

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 $\sigma_{n,Z}^2 = \text{Var} [\sum_{i \in D_n} Z_{i,n}]$ and $\sigma_{n,X}^2 = \text{Var} [\sum_{i \in D_n} X_{i,n}] = M_n^{-2} \sigma_{n,Z}^2$. Since

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

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

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

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,

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

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:

$$\liminf_{n \rightarrow \infty} |D_n|^{-1} \sigma_n^2 > 0. \quad (\text{B.2})$$

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

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

and the corresponding variances as

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

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

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

Furthermore, by (B.1)

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

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

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

Since $X_{i,n}^k$ and $\tilde{X}_{i,n}^k$ 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:

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

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)

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

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 $n \geq N_*$:

$$B_1|D_n| \leq \sigma_n^2 \leq B_2|D_n|, \quad (\text{B.6})$$

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

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

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

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

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

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

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

and

$$\begin{aligned} & \lim_{k \rightarrow \infty} \sup_{n \geq N_*} \left| \frac{\sigma_n^2 - \sigma_{n,k}^2}{\sigma_n^2} \right| \tag{B.8} \\ & \leq \frac{2B_2''}{B_1} \left[\lim_{k \rightarrow \infty} \sup_n \sup_{i \in D} \|\tilde{X}_{i,n}^k\|_{2+\delta} \right] + \frac{B_2'}{B_1} \left[\lim_{k \rightarrow \infty} \sup_n \sup_{i \in D} \|\tilde{X}_{i,n}^k\|_{2+\delta} \right]^2 = 0. \end{aligned}$$

$$\begin{aligned} \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{\tilde{\sigma}_{n,k}^2}{\sigma_n^2} & \leq \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \left\{ \frac{B_2'}{B_1} \left[\sup_n \sup_{i \in D} \|\tilde{X}_{i,n}^k\|_{2+\delta} \right]^2 \right\} \tag{B.9} \\ & = \frac{B_2'}{B_1} \left[\lim_{k \rightarrow \infty} \sup_n \sup_{i \in D} \|\tilde{X}_{i,n}^k\|_{2+\delta} \right]^2 = 0. \end{aligned}$$

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

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

with

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

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

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

for each $k = 1, 2, \dots$. 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

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

In light of (B.8)

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

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 $r(k) = \lim_{n \rightarrow \infty} \sigma_{n,k}/\sigma_n$ exists, then $V_{nk} \implies V_k \sim N(0, r(k)^2)$ in light of (B.10). If furthermore, $\lim_{k \rightarrow \infty} r(k) \rightarrow 1$, then $V_k \implies V \sim N(0, 1)$, and the claim would follow by Lemma B.2. However, in the case of nonstationary variables $\lim_{n \rightarrow \infty} \sigma_{n,k}/\sigma_n$ need not exist, and therefore, we have to use a different argument to show that $Y_n \implies V \sim N(0, 1)$. 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 $d(\cdot, \cdot)$. Let μ_n and μ be the probability measures corresponding to Y_n and V , respectively, then $\mu_n \implies \mu$ iff $d(\mu_n, \mu) \rightarrow 0$ as $n \rightarrow \infty$. Now suppose that Y_n does not converge to V . Then for some $\varepsilon > 0$ there exists 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 $\sigma_{n(m(l_1)),1}/\sigma_{n(m(l_1))} \rightarrow r(1)$ as $l_1 \rightarrow \infty$. For $k = 2$ there exists a subsubsubsequence $\{n(m(l_1(l_2)))\}$ such that $\sigma_{n(m(l_1(l_2))),2}/\sigma_{n(m(l_1(l_2)))} \rightarrow r(2)$ as $l_2 \rightarrow \infty$. The argument can

be repeated for $k = 3, 4, \dots$. 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 $k = 1, 2, \dots$, we have:

$$\lim_{l \rightarrow \infty} \frac{\sigma_{n_l, k}}{\sigma_{n_l}} = r(k). \quad (\text{B.11})$$

Moreover, since by (B.9)

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

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

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

Given (B.12), it follows that $V_{n_l k} \implies V_k \sim N(0, r(k)^2)$. Then, by Lemma B.2, $Y_{n_l} \implies V \sim N(0, 1)$ as $l \rightarrow \infty$. 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 $Y_n \implies N(0, 1)$ if (B.10) holds. In light of this it suffices to prove the CLT for bounded variables. In the following, we assume that $|X_{i, n}| \leq C_X < \infty$.

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

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

and

$$m_n^d |D_n|^{-1/2} \rightarrow 0 \quad (\text{B.14})$$

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

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

so that

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

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

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

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

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

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 $|a_n| \leq \sum_{i,j \in D_n, \rho(i,j) \leq m_n} |\text{cov}(X_{i,n}, X_{j,n})| \leq B_2 |D_n|$. Hence, for sufficiently large n , say $n \geq N_{**} \geq N_*$:

$$0 < B_1 |D_n| \leq a_n \leq B_2 |D_n|, \quad 0 < B_1 \leq B_2 < \infty, \quad (\text{B.15})$$

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

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 $\bar{S}_n \Rightarrow N(0, 1)$, it suffices to show that

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

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

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

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 $E|A_{k,n}| \rightarrow 0$ as $n \rightarrow \infty$ for $k = 1, 2, 3$.

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

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

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

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

By (B.15), we have

$$\begin{aligned}
E|A_{1,n}|^2 &\leq C_*|D_n|^{-2} \sum_{\substack{i,j,i',j' \in D_n \\ \rho(i,j) \leq m_n, \rho(i',j') \leq m_n}} |\text{cov}(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})| & (B.16) \\
&= C_*|D_n|^{-2} \sum_{\substack{i,j,i',j' \in D_n \\ \rho(i,j) \leq m_n, \rho(i',j') \leq m_n, \rho(i,i') \geq 3m_n}} |\text{cov}(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})| \\
&\quad + C_*|D_n|^{-2} \sum_{\substack{i,j,i',j' \in D_n \\ \rho(i,j) \leq m_n, \rho(i',j') \leq m_n, \rho(i,i') < 3m_n}} |\text{cov}(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})|,
\end{aligned}$$

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.

$\gamma(a)$ 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 $k = l = 2$, and observing that $\bar{\alpha}(k, l, h)$ is nonincreasing in h yields

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

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.,

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

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

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

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 $i \in D_n$:

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

for some $C_1 < \infty$.

7(b) Next consider the case where $r = \rho(i, i') < 3m_n$. Let $V_i = \{x \in D_n : \rho(x, i) \leq 4m_n\}$ 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 $\rho(i, i') < 3m_n$, $\rho(i, j) \leq m_n$, and $\rho(i', j') \leq m_n$. 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 $k = 1, l = 3$, and then with $k = l = 1$ gives

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

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

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

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

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

$$\begin{aligned}
&\sum_{\substack{j, i', j' \in D_n \\ \rho(i, j) \leq m_n, \rho(i', j') \leq m_n, \rho(i, i') < 3m_n}} |cov(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})| & (B.22) \\
&\leq \sum_{j, i', j' \in V_i} |cov(X_{i,n}X_{j,n}; X_{i',n}X_{j',n})| \\
&\leq 8C_X^4 \sum_{j, i', j' \in V_i} \bar{\alpha}(1, 3, h(j, i', j')) = 8C_X^4 \sum_{l=1}^{4m_n} \sum_{A \in W_i(l)} \bar{\alpha}(1, 3, l) \\
&\leq 8C_X^4 \bar{M}_* m_n^{2d} \sum_{l=1}^{4m_n} l^{d-1} \bar{\alpha}(1, 3, l) \leq C_2 m_n^{2d}
\end{aligned}$$

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

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

$$E|A_{1,n}|^2 \leq C_* |D_n|^{-2} \sum_{i \in D_n} (C_1 + C_2) m_n^{2d} \leq \text{const} * |D_n|^{-1} m_n^{2d} \rightarrow 0$$

as $n \rightarrow \infty$.

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

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

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

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

Since $|X_{i,n}| \leq C_X < \infty$, using this inequality and the same arguments as before gives:

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

for some $C_5 < \infty$. The last inequality used Assumption 3. Hence, by (B.14), $E|A_{2,n}| \rightarrow 0$ as $n \rightarrow \infty$.

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

$$|EA_{3,n}| = \left| E a_n^{-1/2} \sum_{j \in D_n} X_{j,n} e^{i\lambda(\bar{S}_n - \bar{S}_{j,n})} \right| \leq \text{const} * |D_n|^{-1/2} \sum_{j \in D_n} \left| E X_{j,n} e^{i\lambda(\bar{S}_n - \bar{S}_{j,n})} \right|$$

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 $k = 1, l = |D_n|$ gives

$$\left| E X_{j,n} e^{i\lambda(\bar{S}_n - \bar{S}_{j,n})} \right| \leq 4C_X \bar{\alpha}(1, |D_n|, m_n)$$

and hence as $n \rightarrow \infty$ by (B.13). This completes the proof of the CLT. ■

C Appendix: Proofs of 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

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

Let

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

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

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

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

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

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

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

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

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

To see this, observe that

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

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 $\delta \rightarrow 0$, as demonstrated above. Moreover, by Theorem 2.1 in Pötscher and Prucha (1994a), $Q_n(\theta)$ is

also AL_0UEC , i.e., for every $\varepsilon > 0$

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

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

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

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

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

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 $a.s.ACUEC \implies ACL_0UEC$ 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 $\bar{Q}_n(\theta) = EQ_n(\theta)$ is *AUEC*, by Theorem 3.3 in Pötscher and Prucha (1994a), which proves part (b) of the theorem. \blacksquare

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

$$\begin{aligned} \lim_{k \rightarrow \infty} \sup_n \sup_{i \in D_n} E[|X_{i,n}| \mathbf{1}(|X_{i,n}| < k)] & \quad (C.4) \\ & \leq \lim_{k \rightarrow \infty} \sup_n \sup_{i \in D_n} E[|Z_{i,n}/c_{i,n}| \mathbf{1}(|Z_{i,n}/c_{i,n}| > k)] = 0, \end{aligned}$$

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

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

In light of (C.4)

$$\lim_{k \rightarrow \infty} \sup_n \sup_{i \in D_n} E \left| \tilde{X}_{i,n}^k \right| = 0. \quad (C.5)$$

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

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

and thus

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

where $\|\cdot\|_1$ 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 proof we now demonstrate that also the second term converges to zero. To that effect it suffices to show that $X_{i,n}^k$ satisfies an L_1 -norm LLN for fixed k .

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

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

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

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

with $C < \infty$, and $K = \sum_{m=1}^{\infty} m^{d-1} \bar{\alpha}_X(1, 1, m) < \infty$ by Assumption 3b. Consequently, the r.h.s. of (C.8) is seen to go to zero as $n \rightarrow \infty$, 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. \blacksquare

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

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

Further observe that

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

By Markov's inequality and the i.p. part of Condition 7, we have

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

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 7 we have a.s.

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

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

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

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

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 from equicontinuity Condition 8(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 $1 \leq k \leq K$:

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

where $B = \limsup_{n \rightarrow \infty} |D_n|^{-1} \sum_{i \in D_n} E |r_{ki,n}(Z_{i,n})|$, which is finite by Condition 8(a). Since η was arbitrary it follows that

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

Also, for each $1 \leq k \leq K$ it follows from by Condition 8(b) that

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

Hence

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

which completes the proof. ■

D Appendix: Proofs for Section 4.2

D.1 Proof of Lemma 3

Let $\mathfrak{F}_i(s) = \sigma(X_j; j \in \mathbb{Z}^d: \rho(i, j) \leq s)$. First note that

$$\|Y_i - E[Y_i | \mathfrak{F}_i(s)]\|_p = \left\| \sum_{j \in \mathbb{Z}^d, \rho(i, j) > s} g_{i, j} \{X_j - E[X_j | \mathfrak{F}_i(s)]\} \right\|_p.$$

By Minkowski's inequality for infinite sums and (9)-(10), we have

$$\begin{aligned} \|Y_i - E[Y_i | \mathfrak{F}_i(s)]\|_p &\leq \sum_{j \in \mathbb{Z}^d, \rho(i, j) > s} |g_{i, j}| \|X_j - E[X_j | \mathfrak{F}_i(s)]\|_p \\ &\leq 2 \sup_{j \in \mathbb{Z}^d} \|X_{j, n}\|_p \sup_{i \in \mathbb{Z}^d} \sum_{j \in \mathbb{Z}^d, \rho(i, j) > s} |g_{i, j}| \\ &= c\psi(s) \rightarrow 0 \text{ as } s \rightarrow \infty \end{aligned}$$

where $c = 2 \sup_{j \in \mathbb{Z}^d} \|X_{j, n}\|_p$ and

$$\psi(s) = \sup_{i \in \mathbb{Z}^d} \sum_{j \in \mathbb{Z}^d, \rho(i, j) > s} |g_{i, j}|,$$

by condition (10). ■

D.2 Proof of Lemma 4

Let $\mathfrak{F}_{i,n}(s) = \sigma(X_{j,n}; 1 \leq j \leq n: \rho(i, j) \leq s)$. By Minkowski's inequality, we have:

$$\begin{aligned}
& \|Y_{i,n} - E[Y_{i,n} | \mathfrak{F}_{i,n}(s)]\|_p \\
& \leq \beta \sum_{1 \leq j \leq n, \rho(i,j) > s} |a_{ij,n}| \|Z_{j,n} - E[Z_{j,n} | \mathfrak{F}_{i,n}(s)]\|_p \\
& \quad + \sum_{1 \leq j \leq n, \rho(i,j) > s} |b_{ij,n}| \|\varepsilon_{j,n} - E[\varepsilon_{j,n} | \mathfrak{F}_{i,n}(s)]\|_p \\
& \leq 2\beta \sup_n \sup_{1 \leq i \leq n} \|Z_{i,n}\|_p \sup_n \sup_{1 \leq i \leq n} \sum_{1 \leq j \leq n, \rho(i,j) > s} |a_{ij,n}| \\
& \quad + 2 \sup_n \sup_{1 \leq i \leq n} \|\varepsilon_{i,n}\|_p \sup_n \sup_{1 \leq i \leq n} \sum_{1 \leq j \leq n, \rho(i,j) > s} |b_{ij,n}| \leq c\psi(s)
\end{aligned}$$

where

$$\psi(s) = \sup_n \sup_{1 \leq i \leq n} \sum_{1 \leq j \leq n, \rho(i,j) > s} (|a_{ij,n}| + |b_{ij,n}|) \rightarrow 0 \text{ as } s \rightarrow \infty$$

by condition (12), and

$$c = 2 \max \left\{ \beta \sup_n \sup_{1 \leq j \leq n} \|Z_{j,n}\|_p, \sup_n \sup_{1 \leq j \leq n} \|\varepsilon_{j,n}\|_p \right\} < \infty$$

by condition (13). ■

D.3 Proof of Lemma 5

Let $\mathfrak{F}_i(s) = \sigma(X_{i+j}; j \in D : |j| \leq s)$. Define a $\mathfrak{F}_i(s)/\mathfrak{B}$ -measurable approximating function for $Y_i = H(X_{i+j}, j \in D)$ by replacing the arguments with spatial lags j outside the s -neighborhood of i by zeroes:

$$h^s(X_{i+j}; j \in D : \rho(i, j) \leq s) = H(X_{i+j} \mathbf{1}_{|j| \leq s}, j \in D)$$

where $\mathbf{1}_A$ is the indicator function of set A .

By the minimum mean-squared error property of the conditional expectation, we have:

$$\begin{aligned}
& \|Y_i - E[Y_i | \mathfrak{F}_i(s)]\|_2 \leq \|Y_i - h^s(X_{i+j}; j \in D : |j| \leq s)\|_2 \\
& = \|H(X_{i+j}, j \in D) - H(X_{i+j} \mathbf{1}_{|j| \leq s}, j \in D)\|_2 \\
& \leq \left\| \sum_{j \in D: |j| > s} w_{i+j} |X_{i+j}| \right\|_2 \leq \sum_{j \in D: |j| > s} w_{i+j} \|X_{i+j}\|_2 \\
& \leq \sup_{i \in D} \|X_i\|_2 \sum_{j \in D: |j| > s} w_{i+j} \leq c\psi(s).
\end{aligned}$$

with $c = \sup_{i \in D} \|X_i\|_2$ and $\psi(s) = \sup_{i \in D} \sum_{j \in D: |j| > s} w_{i+j}$. In deriving these inequalities, we used Lipschitz condition (16), Minkowski's inequality and moment condition (18). Finally, by condition (17), $\psi(s) \rightarrow 0$ as $s \rightarrow \infty$. ■

E Appendix: Proof of CLT for NED Processes

The proof of the CLT follows the approach of Ibragimov and Linnik (1971), pp. 352-355 and makes use of the Bernstein's Lemma (B.2) given in Appendix B. We prove the theorem for the α -mixing case. The proof for the ϕ -mixing case is analogous.

1. *Transition from $Z_{i,n}$ to Scaled variables $Y_{i,n} = Z_{i,n}/M_n$*

Throughout the proof, $\mathfrak{F}_{i,n}(s) = \sigma(X_{j,n}; j \in T_n : \rho(i, j) \leq s)$, $s \in \mathbb{N}$ denotes the σ -algebra generated by $X_{j,n}$ located in the s -neighborhood of point $i \in D$.

Let $M_n = \max_{i \in D_n} c_{i,n}$ and $Y_{i,n} = Z_{i,n}/M_n$. Also, let $\sigma_{Z,n}^2 = Var [\sum Z_{i,n}]$ and $\sigma_{Y,n}^2 = Var [\sum Y_{i,n}] = M_n^{-2} \sigma_{Z,n}^2$. Since

$$\sigma_{Y,n}^{-1} \sum_{i \in D_n} Y_{i,n} = \sigma_{Z,n}^{-1} \sum_{i \in D_n} Z_{i,n},$$

to prove the theorem, it suffices to show that $\sigma_{Y,n}^{-1} \sum_{i \in D_n} Y_{i,n} \implies N(0, 1)$. Therefore, it proves convenient to switch notation from the text and to define

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

That is, in the following, S_n denotes $\sum_{i \in D_n} Y_{i,n}$ rather than $\sum_{i \in D_n} Z_{i,n}$, and σ_n^2 denotes the variance of $\sum_{i \in D_n} Y_{i,n}$ rather than of $\sum_{i \in D_n} Z_{i,n}$. We now establish moment and mixing conditions for $Y_{i,n}$ from the assumptions of the theorem. Observe that by definition of M_n

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

and hence

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

so that Assumption 11 implies that

$$\lim_{k \rightarrow \infty} \sup_n \sup_{i \in D_n} E[|Y_{i,n}|^{2+\delta} \mathbf{1}(|Y_{i,n}| > k)] = 0. \quad (\text{E.1})$$

Since $Y_{i,n}$ is uniformly $L_{2+\delta}$ integrable, it is also uniformly $L_{2+\delta}$ bounded.

Let $\|Y\|_{2+\delta} = \sup_{i,n} \|Y_{i,n}\|_{2+\delta}$. Further, note that

$$\begin{aligned} \|Y_{i,n} - E(Y_{i,n}|\mathfrak{F}_{i,n}(s))\|_2 &= M_n^{-1} \|Z_{i,n} - E(Z_{i,n}|\mathfrak{F}_{i,n}(s))\|_2 \\ &\leq c_{i,n}^{-1} d_{i,n} \psi(s) \leq C \psi(s) \end{aligned} \quad (\text{E.2})$$

since $\sup_n \sup_{i \in D} c_{i,n}^{-1} d_{i,n} \leq C < \infty$, by assumption. Thus, $Y_{i,n}$ is L_2 -NED on X with the NED coefficients $\psi(s)$ and magnitude index C . Finally, observe that by Assumption 12:

$$\liminf_{n \rightarrow \infty} |D_n|^{-1} \sigma_n^2 > 0. \quad (\text{E.3})$$

Hence, there exists an N_* and $0 < B_1 < \infty$ such that for all $n \geq N_*$, we have

$$B_1 |D_n| \leq \sigma_n^2 \quad (\text{E.4})$$

In the following, without loss of generality we assume $N_* = 1$.

2. Decomposition of $Y_{i,n}$

The proof of the theorem will make use of the following two auxiliary random variables:

$$\xi_{i,n}^s = E(Y_{i,n}|\mathfrak{F}_{i,n}(s)), \quad \eta_{i,n}^s = Y_{i,n} - \xi_{i,n}^s \quad (\text{E.5})$$

for some $s > 0$. Note that $E\xi_{i,n}^s = 0$ and $E\eta_{i,n}^s = 0$. Let

$$\begin{aligned} S_{n,s} &= \sum_{i \in D_n} \xi_{i,n}^s; \quad \tilde{S}_{n,s} = \sum_{i \in D_n} \eta_{i,n}^s \\ \sigma_{n,s}^2 &= \text{Var}[S_{n,s}], \quad \tilde{\sigma}_{n,s}^2 = \text{Var}[\tilde{S}_{n,s}] \end{aligned}$$

Repeated use of Minkowski's inequality yields:

$$|\sigma_n - \sigma_{n,s}| \leq \tilde{\sigma}_{n,s} \quad (\text{E.6})$$

Similarly, one can show that

$$|\sigma_n - \tilde{\sigma}_{n,s}| \leq \sigma_{n,s} \quad (\text{E.7})$$

Furthermore, Jensen's conditional expectation and Lyapunov's inequalities give for all $s > 0$ and any $1 \leq q \leq 2 + \delta$:

$$\begin{aligned} E |\xi_{i,n}^s|^q &= E \{ |E(Y_{i,n} | \mathfrak{F}_{i,n}(s))|^q \} \\ &\leq E \{ E(|Y_{i,n}|^q | \mathfrak{F}_{i,n}(s)) \} \\ &= E |Y_{i,n}|^q \leq \sup_{n,i \in D_n} E |Y_{i,n}|^q = \|Y\|_{2+\delta}^q \end{aligned} \quad (\text{E.8})$$

By Minkowski's and Lyapunov's inequalities, we have for all $s > 0$ and any $1 \leq q \leq 2 + \delta$:

$$\|\eta_{n,i}^s\|_q = \|Y_{i,n} - \xi_{n,i}^s\| \leq 2 \|Y_{i,n}\|_q \leq 2 \|Y_{i,n}\|_{2+\delta}$$

Furthermore, note that for any $q \leq 2$:

$$\|\eta_{i,n}^s\|_q \leq \|\eta_{i,n}^s\|_2 \leq c\psi(s). \quad (\text{E.9})$$

Thus, both $\xi_{n,i}^s$ and $\eta_{n,i}^s$ are uniformly $L_{2+\delta}$ bounded.

Now, let $\sigma(\xi_{i,n}^s)$ be σ -field generated by $\xi_{i,n}^s$. Since $\sigma(\xi_{i,n}^s) \subseteq \mathfrak{F}_{i,n}(s)$, we

have the following bounds for mixing coefficients of $\xi_{i,n}^s$ any $s \in \mathbb{N}$:

$$\bar{\alpha}_\xi(1, 1, h) \leq \begin{cases} 1, & h \leq 2s \\ \bar{\alpha}_X((2s+2)^d, (2s+2)^d, h-2s), & h > 2s \end{cases}$$

since, as shown in the proof of Lemma A.1(ii), the s -neighborhood of any point on the lattice D contains at most $(2s + 2)^d$ points of D for any $s \in \mathbb{N}$.

To simplify notation, hereafter, we suppress the dependence of $\bar{\alpha}_X(1, 1, h)$ on X , i.e. write

$$\bar{\alpha}(1, 1, h) = \bar{\alpha}_X(1, 1, h)$$

Furthermore, we have

$$\begin{aligned} & E [E(Y_{i,n}|\mathfrak{F}_{i,n}(s))|\mathfrak{F}_{i,n}(m)] \\ = & \begin{cases} E(Y_{i,n}|\mathfrak{F}_{i,n}(s)), & m \geq s, \\ E(Y_{i,n}|\mathfrak{F}_{i,n}(m)), & m < s. \end{cases} \end{aligned}$$

Define the L_2 -approximation error of any field $U_{i,n}$ by the base field X as follows:

$$\varphi_U(m) \equiv \|U_{i,n} - E(U_{i,n}|\mathfrak{F}_{i,n}(m))\|_2$$

Then, we have

$$\begin{aligned} \varphi_\eta^2(m) & \equiv \|\eta_{i,n}^s - E(\eta_{i,n}^s|\mathfrak{F}_{i,n}(m))\|_2^2 \\ & = E\{Y_{i,n} - E[Y_{i,n}|\mathfrak{F}_{i,n}(s)] - E[Y_{i,n}|\mathfrak{F}_{i,n}(m)] \\ & \quad + E[(Y_{i,n}|\mathfrak{F}_{i,n}(s))|\mathfrak{F}_{i,n}(m)]\}^2 \\ & = \begin{cases} \varphi_Y^2(m), & m \geq s, \\ \varphi_Y^2(s), & m < s. \end{cases} = \varphi_Y^2(\max(m; s)) \end{aligned}$$

In other words, if $m \geq s$ we have

$$\|\eta_{i,n}^s - E(\eta_{i,n}^s|\mathfrak{F}_{i,n}(m))\|_2 = \|Y_{i,n} - E(Y_{i,n}|\mathfrak{F}_{i,n}(m))\|_2 \leq c\psi(m) \quad (\text{E.10})$$

If $m < s$, then

$$\|\eta_{i,n}^s - E(\eta_{i,n}^s|\mathfrak{F}_{i,n}(m))\|_2 = \|Y_{i,n} - E(Y_{i,n}|\mathfrak{F}_{i,n}(s))\|_2 \leq c\psi(s) \leq c\psi(m)$$

since $\psi(m)$ is non-increasing sequence. Hence, $\eta_{i,n}^s$ is L_2 -NED on X and we can, with no loss of generality, assume that the NED coefficients of $\eta_{i,n}^s$ for any fixed s are equal to those of $Y_{i,n}$, i.e., for all $m \in \mathbb{N}$

$$\psi_\eta(m) \leq \psi_Y(m) = \psi(m). \quad (\text{E.11})$$

Furthermore, $\eta_{i,n}^s$ is $L_{2+\delta}$ bounded. Thus, for all $s > 0$, $\eta_{i,n}^s$ has the same structure as $Y_{i,n}$. This observations will be exploited further in the proof of the theorem.

3. Bound for Variance of $\sum \eta_{i,n}^s$

Let $\tilde{\sigma}_{n,s}^2 = \text{Var} [\sum_{D_n} \eta_{n,i}^s]$. To simplify notation, in the following, we suppress dependence of $\eta_{i,n}^s$ on s and set

$$U_{i,n} := \eta_{i,n}^s$$

Since $U_{i,n}$ has the same structure as $Y_{i,n}$, i.e., L_2 -NED on X , we can similarly decompose $U_{i,n}$ as follows:

$$\begin{aligned} U_{i,n} &= \left(\tilde{\xi}_{i,n}^m + \tilde{\eta}_{i,n}^m \right) \\ \tilde{\xi}_{i,n}^m &= E(U_{i,n} | \mathfrak{F}_{i,n}(m)) \\ \tilde{\eta}_{i,n}^m &= U_{i,n} - \tilde{\xi}_{i,n}^m \end{aligned}$$

Then, for any i, j such that $\rho(i, j) = h \geq 3$, we have

$$\begin{aligned} E(\eta_{i,n}^s, \eta_{j,n}^s) &= E \left(\tilde{\xi}_{i,n}^{[h/3]} + \tilde{\eta}_{i,n}^{[h/3]} \right) \left(\tilde{\xi}_{j,n}^{[h/3]} + \tilde{\eta}_{j,n}^{[h/3]} \right) \\ &\leq \left| E \left(\tilde{\xi}_{i,n}^{[h/3]} \tilde{\xi}_{j,n}^{[h/3]} \right) \right| + \left| E \left(\tilde{\xi}_{i,n}^{[h/3]} \tilde{\eta}_{j,n}^{[h/3]} \right) \right| \\ &\quad + \left| E \left(\tilde{\xi}_{j,n}^{[h/3]} \tilde{\eta}_{i,n}^{[h/3]} \right) \right| + \left| E \left(\tilde{\eta}_{i,n}^{[h/3]} \tilde{\eta}_{j,n}^{[h/3]} \right) \right| \end{aligned}$$

By (E.8) and the definition of $\eta_{i,n}^s$:

$$\begin{aligned} \sup_{n,i \in D_n} \left\| \tilde{\xi}_{j,n}^{[h/3]} \right\|_2 &\leq \sup_{n,i \in D_n} \left\| \eta_{i,n}^s \right\|_2 \\ &= \sup_n \sup_{i \in D_n} \left\| Y_{i,n} - E(Y_{i,n} | \mathfrak{F}_{i,n}(s)) \right\|_2 \leq c\psi(s). \end{aligned} \quad (\text{E.12})$$

Furthermore, by (E.8),

$$\sup_{n,i \in D_n} \left\| \tilde{\xi}_{i,n}^{[h/3]} \right\|_{2+\delta} \leq \sup_{n,i \in D_n} \left\| \eta_{i,n}^s \right\|_{2+\delta} \leq 2 \|Y\|_{2+\delta} \quad (\text{E.13})$$

and by assumption

$$\bar{\alpha}(k, l, r) \leq (k + l)^\tau \hat{\alpha}(r).$$

Using (E.12), (E.13) and the covariance inequalities of Lemma 1 with $p = 2$; $q = 2 + \delta$, and $r = 2(2 + \delta)/\delta$ yields the following upper bound on the first term:

$$\begin{aligned} \left| E \left(\tilde{\xi}_{i,n}^{[h/3]} \tilde{\xi}_{j,n}^{[h/3]} \right) \right| &\leq 8 \left\| \tilde{\xi}_{i,n}^{[h/3]} \right\|_{2+\delta} \left\| \tilde{\xi}_{j,n}^{[h/3]} \right\|_2 \bar{\alpha}_\xi^{\frac{\delta}{2(2+\delta)}}(1, 1, [h/3]) \\ &\leq 8 \left\| \tilde{\xi}_{i,n}^{[h/3]} \right\|_{2+\delta} \left\| \tilde{\xi}_{j,n}^{[h/3]} \right\|_2 \times \\ &\quad \times \bar{\alpha}^{\frac{\delta}{2(2+\delta)}} \left((2[h/3] + 2)^d, (2[h/3] + 2)^d, h - 2[h/3] \right) \\ &\leq 2^{\frac{\delta\tau}{2(2+\delta)}} 16 \|Y\|_{2+\delta} (2[h/3] + 2)^{\frac{\delta\tau d}{2(2+\delta)}} \hat{\alpha}^{\frac{\delta}{2(2+\delta)}}([h/3]) \sup_{n,i \in D_n} \left\| \eta_{n,i}^s \right\|_2 \\ &\leq 2^{\frac{\delta\tau}{2(2+\delta)}(1+d)} 16 \|Y\|_{2+\delta} ([h/3] + 1)^{\frac{\delta\tau d}{2(2+\delta)}} \hat{\alpha}^{\frac{\delta}{2(2+\delta)}}([h/3]) c\psi(s) \\ &\leq c_1 ([h/3] + 1)^{d\tau_*} \hat{\alpha}^{\frac{\delta}{2(2+\delta)}}([h/3]) \psi(s), \end{aligned}$$

where $\tau_* = \frac{\delta\tau}{2(2+\delta)}$; $c_1 = 2^{4+(1+d)\tau_*} c \|Y\|_{2+\delta}$.

Using Cauchy-Schwartz inequality, (E.10) and (E.11) gives the following bound on the second and third terms:

$$\begin{aligned} \left| E \left(\tilde{\xi}_{i,n}^{[h/3]} \tilde{\eta}_{j,n}^{[h/3]} \right) \right| &\leq \left\| \tilde{\xi}_{i,n}^{[h/3]} \right\|_2 \left\| \tilde{\eta}_{j,n}^{[h/3]} \right\|_2 \\ &\leq \sup_{n,i \in D_n} \left\| \eta_{n,i}^s \right\|_2 c\psi([h/3]) \\ &\leq c^2 \psi(s) \psi([h/3]). \end{aligned}$$

Similarly, using Cauchy-Schwartz inequality, we have for the fourth term:

$$\begin{aligned}
\left| E \left(\tilde{\eta}_{i,n}^{[h/3]} \tilde{\eta}_{j,n}^{[h/3]} \right) \right| &\leq \left\| \tilde{\eta}_{i,n}^{[h/3]} \right\|_2 \left\| \tilde{\eta}_{j,n}^{[h/3]} \right\|_2 \\
&\leq 2c \sup_{n,i \in D_n} \left\| \eta_{n,i}^s \right\|_2 \psi([h/3]) \\
&\leq 2c^2 \psi(s) \psi([h/3])
\end{aligned}$$

Collecting terms, we have:

$$\left| E(\eta_{i,n}^s, \eta_{j,n}^s) \right| \leq \psi(s) \left\{ c_1 ([h/3] + 1)^{d\tau^*} \hat{\alpha}^{\frac{\delta}{2(2+\delta)}} ([h/3]) + c_2 \psi([h/3]) \right\}$$

where $c_2 = 4c^2$ and c_1 is as defined above.

Using the last inequality as well as Lemma A.1(iii) of Appendix A, we can now establish an upper bound for $\tilde{\sigma}_{n,s}^2$:

$$\begin{aligned}
\tilde{\sigma}_{n,s}^2 &\leq \sum_{i \in D_n} E \left| \eta_{i,n}^s \right|^2 + \sum_{\substack{i \in D_n, j \in D_n \\ i \neq j, \rho(i,j) < 3}} \left\| \eta_{i,n}^s \right\|_2 \left\| \eta_{j,n}^s \right\|_2 + \sum_{\substack{i \in D_n, j \in D_n \\ i \neq j, \rho(i,j) \geq 3}} \left| E(\eta_{i,n}^s \eta_{j,n}^s) \right| \\
&\leq |D_n| \sup_n \sup_{i \in D_n} \left\| \eta_{n,i}^s \right\|_2^2 + \sum_{i \in D_n} \sum_{m=1}^2 \sum_{j \in D_n: \rho(i,j) \in [m, m+1)} \left\| \eta_{i,n}^s \right\|_2 \left\| \eta_{j,n}^s \right\|_2 \\
&\quad + c_1 \left[\sum_{i \in D_n} \sum_{m=3}^{\infty} \sum_{j \in D_n: \rho(i,j) \in [m, m+1)} \left([\rho(i,j)/3] + 1 \right)^{d\tau^*} \hat{\alpha}^{\frac{\delta}{2(2+\delta)}} ([\rho(i,j)/3]) \right] \psi(s)
\end{aligned}$$

$$\begin{aligned}
& +c_2 \left[\sum_{i \in D_n} \sum_{m=3}^{\infty} \sum_{j \in D_n: \rho(i,j) \in [m, m+1)} \psi([\rho(i,j)/3]) \right] \psi(s) \\
\leq & |D_n| [1 + M(1 + 2^{d-1})] \sup_n \sup_{i \in D_n} \|\eta_{n,i}^s\|_2^2 + \\
& +c_1 \left[\sum_{i \in D_n} \sum_{m=3}^{\infty} N_i(1, 1, m) ([m/3] + 1)^{d\tau_*} \widehat{\alpha}^{\frac{\delta}{2(2+\delta)}}([m/3]) \right] \psi(s) \\
& +c_2 \left[\sum_{i \in D_n} \sum_{m=3}^{\infty} N_i(1, 1, m) \psi([m/3]) \right] \psi(s) \\
\leq & |D_n| [1 + M(1 + 2^{d-1})] c^2 \psi^2(s) + \\
& +c_1 M 4^{d\tau_*} \left[\sum_{i \in D_n} \sum_{m=3}^{\infty} m^{d-1} [m/3]^{d\tau_*} \widehat{\alpha}^{\frac{\delta}{2(2+\delta)}}([m/3]) \right] \psi(s) \\
& +c_2 M \left[\sum_{i \in D_n} \sum_{m=3}^{\infty} m^{d-1} \psi([m/3]) \right] \psi(s) \\
\leq & B_2 |D_n| \psi^2(s) + B_3 |D_n| \psi(s),
\end{aligned}$$

where

$$\begin{aligned}
B_2 &= c^2 [1 + M(1 + 2^{d-1})] \\
B_3 &= 3^d 4^{d\tau_*} c_1 M K_1 + 3^d c_2 M K_2
\end{aligned}$$

with

$$K_1 = \sum_{m=1}^{\infty} m^{d(\tau_*+1)-1} \widehat{\alpha}^{\frac{\delta}{2(2+\delta)}}(m) \leq \sum_{m=1}^{\infty} m^{d(\tau+1)-1} \widehat{\alpha}^{\frac{\delta}{2(2+\delta)}}(m) < \infty$$

since $\tau_* < \tau$, and

$$K_2 = \sum_{m=1}^{\infty} m^{d-1} \psi(m) < \infty.$$

These two series are convergent by Assumptions 9 and 13, respectively. We also used the following elementary inequality

$$([m+1]/3 + 1)^{d\tau_*} \leq ([m/3] + 2)^{d\tau_*} \leq 4^{d\tau_*} [m/3]^{d\tau_*}$$

Thus,

$$0 \leq \tilde{\sigma}_{n,s}^2 \leq (B_2 + B_3) |D_n| \psi(s) \quad (\text{E.14})$$

since $\psi^2(s) \leq 1$ and hence $\psi^2(s) \leq \psi(s)$. In light of (E.4), we have:

$$\frac{\tilde{\sigma}_{n,s}^2}{\sigma_n^2} \leq \frac{B_2 + B_3}{B_1} \psi(s)$$

and hence

$$\lim_{s \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{\tilde{\sigma}_{n,s}^2}{\sigma_n^2} \leq \frac{B_2 + B_3}{B_1} \lim_{s \rightarrow \infty} \psi(s) = 0 \quad (\text{E.15})$$

Furthermore, by (E.6) we have

$$\lim_{s \rightarrow \infty} \limsup_{n \rightarrow \infty} \left| 1 - \frac{\sigma_{n,s}}{\sigma_n} \right| \leq \lim_{s \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{\tilde{\sigma}_{n,s}}{\sigma_n} = 0 \quad (\text{E.16})$$

and hence for all $s \geq 1$ and $n \geq 1$

$$\frac{\sigma_{n,s}}{\sigma_n} \leq \tilde{B} < \infty. \quad (\text{E.17})$$

4. CLT for $\sigma_{n,s}^{-1} \sum_{i \in D_n} \xi_{i,n}^s$

We now show that for fixed $s > 0$, $\xi_{i,n}^s$ satisfies the CLT for α -mixing fields, Theorem 1. First, note that since $\xi_{i,n}^s$ is a measurable function of $X_{i,n}$,

$$\bar{\alpha}_\xi(k, l, m) \leq \bar{\alpha}(k(2s+2)^d, l(2s+2)^d, m-2s)$$

for $k+l \leq 4$ and $m > 2s$. Therefore, by Assumption 9 we have for fixed $s > 0$

$$\begin{aligned}
& \sum_{m=1}^{\infty} m^{d-1} \bar{\alpha}_{\xi}^{\delta/(2+\delta)}(1, 1, m) \\
\leq & \sum_{m=1}^{2s} m^{d-1} + \sum_{m=2s+1}^{\infty} m^{d-1} \bar{\alpha}^{\delta/(2+\delta)}((2s+2)^d, (2s+2)^d, m-2s) \\
\leq & \sum_{m=1}^{2s} m^{d-1} + (2s+2)^{d\tau_1} 2^{\tau_1} \sum_{m=2s+1}^{\infty} m^{d-1} \widehat{\alpha}^{\frac{\delta}{2+\delta}}(m-2s) \\
\leq & \sum_{m=1}^{2s} m^{d-1} + (2s+2)^{d\tau_1} 2^{\tau_1} \sum_{m=1}^{\infty} (m+2s)^{d-1} \widehat{\alpha}^{\frac{\delta}{2+\delta}}(m) < \infty
\end{aligned}$$

where $\tau_1 = \tau\delta/(2+\delta)$, since $\widehat{\alpha}^{\frac{\delta}{2+\delta}}(m) \leq \widehat{\alpha}^{\frac{\delta}{2(2+\delta)}}(m)$.

Similarly,

$$\begin{aligned}
& \sum_{m=1}^{\infty} m^{d-1} \bar{\alpha}_{\xi}(k, l, m) \\
\leq & \sum_{m=1}^{2s} m^{d-1} + \sum_{m=2s+1}^{\infty} m^{d-1} \bar{\alpha}(k(2s+2)^d, l(2s+2)^d, m-2s) \\
\leq & \sum_{m=1}^{2s} m^{d-1} + (2s+2)^{d\tau} (k+l)^{\tau} \sum_{m=1}^{\infty} (m+2s)^{d-1} \widehat{\alpha}(m) < \infty,
\end{aligned}$$

since $\widehat{\alpha}(m) \leq \widehat{\alpha}^{\frac{\delta}{2(2+\delta)}}(m)$. By Assumption 9 for each given $s \in \mathbb{N}$, we have

$$\begin{aligned}
\bar{\alpha}_{\xi}(1, \infty, m) & \leq \bar{\alpha}((2s+2)^d, \infty, m-2s) \\
& \leq C'(m-2s)^{-d-\varepsilon} \leq C_* m^{-d-\varepsilon}
\end{aligned}$$

for some $C' < \infty$, $C_* < \infty$ and $\varepsilon > 0$. Hence, $\bar{\alpha}_{\xi}(1, \infty, m) = O(m^{-d-\varepsilon})$.

Furthermore, since $|\xi_{i,n}^s|^{2+\delta} \leq |Y_{i,n}|^{2+\delta}$ for all $i \in D_n$, $n \geq 1$:

$$E[|\xi_{n,i}^s|^{2+\delta} \mathbf{1}(|\xi_{n,i}^s| > k)] \leq E[|Y_{n,i}|^{2+\delta} \mathbf{1}(|Y_{n,i}| > k)],$$

and hence by (E.1)

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

Thus, the moment Assumption 2 of that theorem is also fulfilled. Finally, we demonstrate that for sufficiently large s ,

$$0 < \liminf_{n \rightarrow \infty} |D_n|^{-1} \sigma_{n,s}^2.$$

First note that by (E.3),

$$0 < B = \liminf_{n \rightarrow \infty} |D_n|^{-1/2} \sigma_n.$$

Furthermore, it follows from (E.14) that

$$|D_n|^{-1/2} \tilde{\sigma}_{n,s} \leq (B_2 + B_3)^{1/2} \psi^{1/2}(s).$$

Since $\lim_{s \rightarrow \infty} \psi^{1/2}(s) = 0$, there exists s_* such that for all $s \geq s_*$

$$|D_n|^{-1/2} \tilde{\sigma}_{n,s} \leq \frac{B}{2}$$

Next observe that by (E.7)

$$|D_n|^{-1/2} (\sigma_n - \tilde{\sigma}_{n,s}) \leq |D_n|^{-1/2} \sigma_{n,s}$$

and hence for all $s \geq s_*$

$$\begin{aligned} \liminf_{n \rightarrow \infty} |D_n|^{-1/2} \sigma_{n,s} &\geq \liminf_{n \rightarrow \infty} |D_n|^{-1/2} \sigma_n + \liminf_{n \rightarrow \infty} [-|D_n|^{-1/2} \tilde{\sigma}_{n,s}] \\ &= \liminf_{n \rightarrow \infty} |D_n|^{-1/2} \sigma_n - \limsup_{n \rightarrow \infty} [|D_n|^{-1/2} \tilde{\sigma}_{n,s}] \\ &\geq B - \frac{B}{2} = \frac{B}{2} > 0 \end{aligned}$$

Thus, for all $s \geq s_*$,

$$\sigma_{n,s}^{-1} \sum_{i \in D_n} \xi_{i,n}^s \implies N(0, 1).$$

5. CLT for $\sigma_n^{-1} \sum_{i \in D_n} Y_{i,n}$

Finally, we show that

$$\sigma_n^{-1} \sum_{i \in D_n} Y_{i,n} \Longrightarrow N(0, 1).$$

Define

$$\begin{aligned} W_n &= \sigma_n^{-1} \sum_{D_n} Y_{i,n}; \quad V_{ns} = \sigma_n^{-1} \sum_{i \in D_n} \xi_{i,n}^s \\ W_n - V_{ns} &= \sigma_n^{-1} \sum_{i \in D_n} \eta_{i,n}^s; \quad V \sim N(0, 1) \end{aligned}$$

so that we can exploit Lemma B.2 to prove the above claim. In Step 4, we showed that for $s \geq s_*$

$$\sigma_{n,s}^{-1} \sum_{i \in D_n} \xi_{i,n}^s \Longrightarrow N(0, 1)$$

W.L.O.G., we can assume that $s_* = 1$. We first verify condition (iii) of

Lemma B.2. By Markov's inequality and (E.15), for every $\varepsilon > 0$ we have

$$\begin{aligned} \lim_{s \rightarrow \infty} \limsup_{n \rightarrow \infty} P(|W_n - V_{ns}| > \varepsilon) &= \lim_{s \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left(\left|\sigma_n^{-1} \sum_{i \in D_n} \eta_{i,n}^s\right| > \varepsilon\right) \\ &\leq \lim_{s \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{\tilde{\sigma}_{n,s}^2}{\varepsilon^2 \sigma_n^2} = 0. \end{aligned}$$

Next observe that

$$V_{ns} = \frac{\sigma_{n,s}}{\sigma_n} \left[\sigma_{n,s}^{-1} \sum_{i \in D_n} \xi_{i,n}^s \right].$$

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 $d(\cdot, \cdot)$. Let μ_n and μ be the probability measures corresponding to W_n and V , respectively, then $\mu_n \Longrightarrow \mu$

iff $d(\mu_n, \mu) \rightarrow 0$ as $n \rightarrow \infty$. Now suppose that W_n does not converge to V . Then for some $\varepsilon > 0$ there exists a subsequence $\{n(m)\}$ such that $d(\mu_{n(m)}, \mu) > \varepsilon$ for all $n(m)$. By (E.17) we have $0 \leq \sigma_{n,s}/\sigma_n \leq \tilde{B} < \infty$ for all $s \geq 1$ and $n \geq 1$. Hence, $0 \leq \sigma_{n(m),s}/\sigma_{n(m)} \leq \tilde{B} < \infty$ for all $n(m)$. Consequently, for $s = 1$ there exists a subsubsequence $\{n(m(l_1))\}$ such that $\sigma_{n(m(l_1)),1}/\sigma_{n(m(l_1))} \rightarrow r(1)$ as $l_1 \rightarrow \infty$. For $s = 2$, there exists a subsubsubsequence $\{n(m(l_1(l_2)))\}$ such that $\sigma_{n(m(l_1(l_2))),2}/\sigma_{n(m(l_1(l_2)))} \rightarrow r(2)$ as $l_2 \rightarrow \infty$. The argument can be repeated for $s = 3, 4, \dots$. 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

$$\lim_{l \rightarrow \infty} \frac{\sigma_{n_l, s}}{\sigma_{n_l}} = r(s) \quad (\text{E.18})$$

for $s = 1, 2, \dots$. Given (E.18), it follows that as $l \rightarrow \infty$

$$V_{n_l s} \implies V_s \sim N(0, r(s)^2).$$

Then, it follows from (E.16) that

$$\lim_{s \rightarrow \infty} |r(s) - 1| \leq \lim_{s \rightarrow \infty} \lim_{l \rightarrow \infty} \left| r(s) - \frac{\sigma_{n_l, s}}{\sigma_{n_l}} \right| + \lim_{s \rightarrow \infty} \sup_{n \geq 1} \left| \frac{\sigma_{n, s}}{\sigma_n} - 1 \right| = 0.$$

Now, by Lemma B.2 $W_{n_l} \implies V \sim N(0, 1)$ as $l \rightarrow \infty$. Since $\{n_l\} \subseteq \{n(m)\}$ this contradicts the assumption that $d(\mu_{n(m)}, \mu) > \varepsilon$ for all $n(m)$. This completes the proof of the CLT. ■

F Appendix: Proof of LLN for NED Processes

Define $Y_{i,n} = Z_{i,n}/M_n$, and observe that

$$\begin{aligned} a_n^{-1} \sum_{i \in D_n} (Z_{i,n} - EZ_{i,n}) &= |D_n|^{-1} \sum_{i \in D_n} M_n^{-1} (Z_{i,n} - EZ_{i,n}) \\ &= |D_n|^{-1} \sum_{i \in D_n} (Y_{i,n} - EY_{i,n}) \end{aligned}$$

since, by definition, $a_n = |D_n| M_n$. So, to prove the theorem, it suffices to show that

$$|D_n|^{-1} \sum_{i \in D_n} (Y_{i,n} - EY_{i,n}) \xrightarrow{L_1} 0.$$

We first establish mixing and moment conditions for $Y_{i,n}$ from those for $Z_{i,n}$.

Observe that by definition of M_n

$$\begin{aligned} \sup_n \sup_{i \in D_n} E |Y_{i,n}|^r &= \sup_n \sup_{i \in D_n} E |Z_{i,n}/M_n|^r \\ &\leq \sup_n \sup_{i \in D_n} E |Z_{i,n}/c_{i,n}|^r < \infty. \end{aligned}$$

Thus, $Y_{i,n}$ is uniformly L_p -bounded for some $p > 1$. Let $\mathfrak{F}_{i,n}(s) = \sigma(X_{j,n}; j \in T_n : \rho(i, j) \leq s)$, $s \in \mathbb{N}$ denote the σ -algebra generated by $X_{j,n}$ located in the s -neighborhood of point $i \in D$. If $Z_{i,n}$ is L_1 -NED on $X = \{X_{i,n}, i \in T_n, n \geq 1\}$, so is $Y_{i,n}$:

$$\begin{aligned} \sup_n \|Y_{i,n} - E(Y_{i,n} | \mathfrak{F}_{i,n}(s))\|_1 &= \sup_n M_n^{-1} \|Z_{i,n} - E(Z_{i,n} | \mathfrak{F}_{i,n}(s))\|_1 \text{ (F.1)} \\ &\leq \sup_n M_n^{-1} d_{i,n} \psi(s) \leq \psi(s), \end{aligned}$$

since $M_n = \max_{i \in D_n} \max(c_{i,n}, d_{i,n})$.

We first show that for each given $s > 0$, the conditional mean $V_{i,n}^s = E(Y_{i,n} | \mathfrak{F}_{i,n}(s))$ satisfies the L_1 -norm LLN of Section 3.4.2, Theorem 3. Note

that $EV_{i,n}^s = EE(Y_{i,n}|\mathfrak{F}_{i,n}(s)) = EY_{i,n} < \infty$. Using Jensen's conditional expectation and Lyapunov's inequalities gives for each $s > 0$:

$$\begin{aligned} E|V_{i,n}^s|^p &= E\{|E(Y_{i,n}|\mathfrak{F}_{i,n}(s))|^p\} \\ &\leq E\{E(|Y_{i,n}|^p|\mathfrak{F}_{i,n}(s))\} \\ &= E|Y_{i,n}|^p \leq \sup_n \sup_{i \in D_n} E|Y_{i,n}|^p < \infty. \end{aligned}$$

So, $V_{i,n}^s$ is uniformly L_p -bounded for $p > 1$ and hence uniformly integrable. For each fixed $s \in \mathbb{N}$, $V_{i,n}^s$ is a measurable function of $X_{i,n}$, and for $c = \{\alpha, \phi\}$:

$$\bar{c}_V(1, 1, h) \leq \begin{cases} 1, & h \leq 2s \\ \bar{c}_X((2s+2)^d, (2s+2)^d, h-2s), & h > 2s \end{cases}$$

since, as shown in the proof of Lemma A.1(ii), the s -neighborhood of any point on the lattice D contains at most $(2s+2)^d$ points of D for any $s \in \mathbb{N}$.

Then, we have

$$\begin{aligned} &\sum_{m=1}^{\infty} m^{d-1} \bar{c}_V(1, 1, m) \\ &= \sum_{m=1}^{2s} m^{d-1} \bar{c}_V(1, 1, m) + \sum_{m=2s+1}^{\infty} m^{d-1} \bar{c}_V(1, 1, m) \\ &\leq \sum_{m=1}^{2s} m^{d-1} + f((2s+2)^d, (2s+2)^d) \sum_{m=1}^{\infty} (m+2s)^{d-1} \hat{c}_X(m) < \infty, \end{aligned}$$

where $\hat{c} = \{\hat{\alpha}, \hat{\phi}\}$. Thus, for each fixed s , $V_{i,n}^s$ is uniformly integrable and α -mixing [ϕ -mixing] satisfying the mixing assumptions of Theorem 3. Therefore, for each s , we have

$$\left\| |D_n|^{-1} \sum_{i \in D_n} (E(Y_{i,n}|\mathfrak{F}_{i,n}(s)) - EY_{i,n}) \right\|_1 \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (\text{F.2})$$

By (F.1), we have for all $i \in D_n$ and $n \geq 1$:

$$\|Y_{i,n} - E(Y_{i,n}|\mathfrak{F}_{i,n}(s))\|_1 \leq \psi(s),$$

and hence by Minkowski's inequality:

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left\| |D_n|^{-1} \sum_{i \in D_n} (Y_{i,n} - E(Y_{i,n}|\mathfrak{F}_{i,n}(s))) \right\|_1 &\leq & \text{(F.3)} \\ &\leq \limsup_{n \rightarrow \infty} |D_n|^{-1} \sum_{i \in D_n} \|Y_{i,n} - E(Y_{i,n}|\mathfrak{F}_{i,n}(s))\|_1 \leq \psi(s). \end{aligned}$$

Hence, it follows from (F.2) and (F.3) that

$$\begin{aligned} &\lim_{n \rightarrow \infty} \left\| |D_n|^{-1} \sum_{i \in D_n} (Y_{i,n} - EY_{i,n}) \right\|_1 \\ &\leq \lim_{s \rightarrow \infty} \limsup_{n \rightarrow \infty} \left\| |D_n|^{-1} \sum_{i \in D_n} (Y_{i,n} - E(Y_{i,n}|\mathfrak{F}_{i,n}(s))) \right\|_1 \\ &\quad + \lim_{s \rightarrow \infty} \lim_{n \rightarrow \infty} \left\| |D_n|^{-1} \sum_{i \in D_n} (E(Y_{i,n}|\mathfrak{F}_{i,n}(s)) - EY_{i,n}) \right\|_1 = 0. \end{aligned}$$

The proof of the LLN is complete. ■

References

- [1] Andrews, D.W.K. (1984): "Non-strong mixing autoregressive processes," *Journal of Applied Probability*, 21, 930-934.
- [2] Andrews, D.W.K. (1987): "Consistency in Nonlinear Econometric Models: a Generic Uniform Law of Large Numbers," *Econometrica*, 55, 1465-1471.
- [3] Andrews, D.W.K. (1988): "Laws of Large Numbers for Dependent Non-identically Distributed Variables," *Econometric Theory* 4, 458-467.
- [4] Andrews, D.W.K. (1992): "Generic Uniform Convergence," *Econometric Theory*, 8, 241-257.
- [5] Baltagi, B.H., and D. Li, (2001a): "Double Length Artificial Regressions for Testing Spatial Dependence," *Econometric Reviews*, 20, 31-40.
- [6] Baltagi, B.H., and D. Li, (2001b): "LM Test for Functional Form and Spatial Error Correlation," *International Regional Science Review*, 24, 194-225.
- [7] 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.
- [8] 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*.

- [9] Bao, Y., and A. Ullah (2007): "Finite Sample Properties of Maximum Likelihood Estimators in Spatial Models," *Journal of Econometrics*, 137, 396-413.
- [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] Bierens, H.J. (1994): *Topics in Advanced Econometrics*. Cambridge: Cambridge University Press.
- [12] Billingsley, P. (1968): *Convergence of Probability Measures*. New York: John Wiley and Sons.
- [13] Billingsley, P. (1986): *Probability and Measure*. John Wiley, New York.
- [14] Bolthausen, E. (1982): "On the Central Limit Theorem for Stationary Mixing Random Fields," *The Annals of Probability*, 10, 1047-1050.
- [15] Bradley, R. (1992): "On the Spectral Density and Asymptotic Normality of Weakly Dependent Random Fields," *Journal of Theoretical Probability*, 5, 355-373.
- [16] Bradley, R. (1993): "Some Examples of Mixing Random Fields," *Rocky Mountain Journal of Mathematics*, Vol. 23, 2, 495-519.
- [17] Bradley, R. (2007): *Introduction to Strong Mixing Conditions*. Vol. 1-3, Kendrick Press.

- [18] Brock, W., and S. Durlauf (2001): "Interactions-based Models," *Handbook of Econometrics*, 5, J. Heckman and E. Leamer eds., Amsterdam: North Holland.
- [19] Brock, W., and S. Durlauf (2007): "Identification of Binary Choice Models with Social Interactions," forthcoming in *Journal of Econometrics*.
- [20] Brockwell, P., and R. Davis (1991): *Times Series: Theory and Methods*, Springer Verlag.
- [21] Bulinskii, A.V.(1986): "Limit Theorems under Weak Dependence Conditions," *Probability Theory and Math. Statistics, Proceedings of the Fourth Vilnius Conference*, Vol. 1, VNU Sci. Press, Utrecht, The Netherlands, pp. 307-326.
- [22] Bulinskii, A.V.(1988): "On Various Conditions for Mixing and Asymptotic Normality of Random Fields," *Soviet Math. Doklady*, 37, 215-226.
- [23] Bulinskii, A.V. (1989): *Limit Theorems under Weak Dependence Conditions*. Moscow University Press (in Russian).
- [24] Bulinskii, A.V., and P. Doukhan (1990): "Vitesse de convergence dans le theorem de limite centrale pour des champs melangeants des hypotheses de moment faibles," C.R. Academie Sci. Paris, Serie I, 801-105.
- [25] Carbon, M., Tran, L. and B. Wu (1997): "Kernel Density Estimation for Random Fields," *Statistics and Probability Letters*, 36, 115-125.

- [26] Cliff, A., and J. Ord (1973): *Spatial Autocorrelation*, London: Pion.
- [27] Cliff, A., and J. Ord (1981): *Spatial Processes, Models and Applications*. London: Pion.
- [28] Conley, T. (1999): "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 92, 1-45.
- [29] Conley, T., and B. Dupor (2003): "Spatial Analysis of Sectoral Complementarity," *Journal of Political Economy*, 111, 311-352.
- [30] Conley, T., and F. Molinari (2007): "Spatial Correlation Robust Inference with Errors in Location or Distance," forthcoming in *Journal of Econometrics*.
- [31] Comets, F., and M. Janžura (1998): " A Central Limit Theorem for Conditionally Centered Random Fields with an Application to Markov Fields," *Journal of Applied Probability*, 35, 608-621.
- [32] Cressie, N.A.C. (1993): *Statistics for Spatial Data*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. New York: Wiley.
- [33] 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.
- [34] Davidson, J. (1992): "A Central Limit Theorem for Globally Nonstationary Near-Epoch Dependent Functions of Mixing Processes," *Econometric Theory*, 8, 313-329.

- [35] Davidson, J. (1993a): "A Central Limit Theorem for Globally Non-stationary Near-Epoch Dependent Functions of Mixing Processes: Asymptotically Degenerate Case," *Econometric Theory*, 9, 402-412.
- [36] Davidson, J. (1993b): "An L_1 Convergence Theorem for Heterogenous Mixingale Arrays with Trending Moments," *Statistics and Probability Letters*, 8, 313-329.
- [37] Davidson, J. (1994): *Stochastic Limit Theory*. Oxford University Press.
- [38] De Jong, R.M. (1997): "Central Limit Theorems for Dependent Heterogeneous Random Variables," *Econometric Theory*, 13, 353-367.
- [39] Dedecker, J. (1998): "A Central Limit Theorem for Stationary Random Fields," *Probability Theory and Related Fields*, 110, 397-426.
- [40] Dedecker, J. (2001): "Exponential Inequalities and Functional Central Limit Theorem for Random Fields," *ESAIM Probability and Statistics*, 5, 77-104.
- [41] Driscoll, J. and A. Kraay (1998): "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *The Review of Economics and Statistics*, LXXX, 549-560.
- [42] 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.

- [43] Dobrushin, R. (1968b): "Gibbs Random Fields for Lattice Systems with Pairwise Interactions," *Functional Analysis and Applications*, 2, 292-301.
- [44] Doukhan, P. (1994): *Mixing. Properties and Examples*. Springer-Verlag.
- [45] Doukhan, P., and X. Guyon (1991): "Mixing for Linear Random Fields," *C.R. Academie Sciences Paris, Serie 1*, 313, 465-470.
- [46] Doukhan, P., and S. Louhichi (1999): "A New Weak Dependence Condition and Applications to Moment Inequalities," *Stochastic Processes and Their Applications*, 84, 313-342.
- [47] Eberlein, E., and A. Csenki (1979): "A Note on Strongly Mixing Lattices of Random Variables," *Z. Wahrsch. verw. Gebiete*, 50, 135-136.
- [48] El Machkouri, M. (2002): "Kahane-Khintchine Inequalities and Functional Central Limit Theorem for Stationary Real Random Fields," *Stochastic Processes and Their Applications*, 120, 285-299.
- [49] El Machkouri, M., and L. Ouchti (2006): "Invariance Principles for Standard-Normalized and Self-Normalized Random Fields," *Alea*, 2, 177-194.
- [50] El Machkouri, M., and D. Volný (2002): "Counter-example dans le Théorème Centrale Limite Fonctionnel pour les Champs Aléatoire Réels," *Annales de l'IHP*, 2, 325-337.

- [51] Fazekas, I. (1983): "Convergence of Vector Valued Martingales with Multidimensional Indices," *Publ. Math. Debrecen*, 30, 157-164.
- [52] Gallant, A.R., and H. White (1988): *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. New York: Basil Blackwell.
- [53] Gorodetskii, V. V. (1977): "On the Strong Mixing Property for Linear Sequences," *Theory of Probability and Applications* 22, 411-413.
- [54] Gut, A. (1978): "Marcinkiewicz Laws and Convergence Rates in the Law of Large Numbers for Random Variables with Multidimensional Indices," *Annals of Probability*, Vol. 6, No 3, 469-482.
- [55] Guyon, X., and S. Richardson (1984): "Vitesse de Convergence du Théorème de la Limite Centrale pour des Champs Faiblement Dependants," *Z. Wahrsch. verw. Gebiete*, 66, 297-314.
- [56] Guyon, X. (1995): *Random Fields on a Network. Modeling, Statistics, and Applications*. Springer-Verlag.
- [57] Hall, P., and C. Heyde (1980): *Martingale Limit Theory and its Applications*, Academic Press.
- [58] Hansen, B. (1991): "Strong Laws for Dependent Heterogeneous Processes," *Econometric Theory*, 7, 213-221.
- [59] Hegerfeldt, G.C., and C. R. Nappi (1977): "Mixing Properties in Lattice Systems," *Communications of Math. Physics*, 53, 1-7.

- [60] Ibragimov, I.A. (1962): "Some Limit Theorems for Stationary Processes," *Theory of Probability and Applications*, 7, 349-382.
- [61] Ibragimov, I.A., and Y. V. Linnik (1971): *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen.
- [62] Kapoor, M., H.H. Kelejian, and I.R. Prucha (2007): "Panel Data Models with Spatially Correlated Error Components," *Journal of Econometrics*, 140, 97-130
- [63] Kelejian, H.H., and I.R. Prucha (1997): "Estimation of the Spatial Autoregressive Parameter by Two-Stage Least Squares Procedures: A Serious Problem," *International Regional Science Review*, 20, 103-111.
- [64] 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.
- [65] 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.
- [66] 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.
- [67] 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.

- [68] Kelejian, H.H., and I.R. Prucha (2004): "Estimation of Simultaneous Systems of Spatially Interrelated Cross Sectional Equations," *Journal of Econometrics*, 118, 27-50.
- [69] Kelejian, H.H., and I.R. Prucha (2007a): "HAC Estimation in a Spatial Framework," *Journal of Econometrics*, 140, 131-154.
- [70] 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*.
- [71] 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.
- [72] Korniotis, G.M. (2005): "A Dynamic Panel Estimator with Both Fixed and Spatial Effects," Department of Finance, University of Notre Dame.
- [73] Lee, L.-F. (2002): "Consistency and Efficiency of Least Squares Estimation for Mixed Regressive Spatial Autoregressive Models," *Econometric Theory*, 18, 252-277.
- [74] Lee, L.-F. (2003): "Best Spatial Two-Stage Least Squares Estimator for a Spatial Autoregressive Model with Autoregressive Disturbances," *Econometric Reviews*, 22, 307-322.
- [75] Lee, L.-F. (2004): "Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models," *Econometrica*, 72, 1899-1925.

- [76] Lee, L.-F. (2007a): "GMM and 2SLS for Mixed Regressive, Spatial Autoregressive Models," *Journal of Econometrics*, 137, 489–514.
- [77] Lee, L.-F. (2007b): "Method of Elimination and Substitution in the GMM Estimation of Mixed Regressive Spatial Autoregressive Models," forthcoming in *Journal of Econometrics*.
- [78] Lee, L.-F. (2007c): "Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects," forthcoming in *Journal of Econometrics*.
- [79] LeSage, J.P., and R.K. Pace (2007): "A Matrix Exponential Spatial Specification," forthcoming in *Journal of Econometrics*.
- [80] 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.
- [81] McLeish, D. L. (1975a): "A Maximal Inequality and Dependent Strong Laws," *Annals of Probability*, 3, 826-36.
- [82] McLeish, D. L. (1975b): "Invariance Principles for Dependent Variables," *Z. Wahrsch. verw. Gebiete*, 32, 165-78.
- [83] Moricz, F., (1977): "Moment Inequalities for the Maximum of Partial Sums of Random Fields," *Acta Sci. Math. (Szeged)*, 39, 353-366.
- [84] Moricz, F. (1978): "Multiparameter Strong Laws of Large Numbers. (Second Order Moment Restrictions)," *Acta Sci. Math. (Szeged)*, 40, 143-156.

- [85] Nahapetian, B. (1980): "A Central Limit Theorem for Random Fields," *Multicomponent Random Systems* edited by Dobrushin and Sinai, 531-542.
- [86] Nahapetian, B. (1987): "An Approach to Limit Theorems for Dependent Random Variables," *Theory of Probability and its Applications*, 32, 589-594.
- [87] Nahapetian, B. (1991): *Limit Theorems and Some Applications in Statistical Physics*. Teubner-Texte zur Mathematik, Leipzig.
- [88] Nahapetian, B. (1995): "Billingsley-Ibragimov Theorem for Martingale-Difference Random Fields and Its Application to Some Models of Classical Statistical Physics," *C. R. Acad. Sci. Paris*, er. I, 320, 1539-1544.
- [89] Nahapetian, B., and A. Petrosian (1992): "Martingale-Difference Gibbs Random Fields and a Central Limit Theorem," *Ann. Acad. Sci. Fennicae*, A-I. 17, 105-110.
- [90] Neaderhouser, C. (1978a): "Limit Theorems for Multiply Indexed Mixing Random Variables," *Annals of Probability*, 6.
- [91] Neaderhouser, C. (1978b): "Some Limit Theorems for Random Variables," *Commun. Math. Phys.*, 61, 293-305.
- [92] Neaderhouser, C. (1981): "An Almost Sure Invariance Principle for a Partial Sums Associated with a Random Field," *Stochastic Processes and their Applications*, 11, 1-10.

- [93] Newey, W.K. (1991): "Uniform Convergence in Probability and Stochastic Equicontinuity," *Econometrica*, 59, 1161-1167.
- [94] 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.
- [95] Peligrad, M., and A. Gut (1999): "Almost-sure Results for a Class of Dependent Random Variables," *Journal of Theoretical Probability*, 12, 87-104.
- [96] Perera, G. (1997): "Geometry of \mathbb{Z}^d and the Central Limit Theorem for Weakly Dependent Random Fields," *Journal of Theoretical Probability*, 10, 581-603.
- [97] Pinkse, J., and M.E. Slade (1998): "Contracting in Space: Application of Spatial Statistics to Discrete-Choice Models," *Journal of Econometrics*, 85, 1111-1153.
- [98] Pinkse, J., M.E. Slade, and C. Brett (2002): "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, 70, 1111-1153.
- [99] Pinkse, J., L. Shen, and M.E. Slade (2006): "A Central Limit Theorem for Endogenous Locations and Complex Spatial Interactions," forthcoming in *Journal of Econometrics*.
- [100] Pollard, D. (1984): *Convergence of Stochastic Processes*. Springer-Verlag, New York.

- [101] Pötscher, B. M., and I. R. Prucha (1989): "A Uniform Law of Large Numbers for Dependent and Heterogeneous Data Processes," *Econometrica*, 57, 675-683.
- [102] Pötscher, B.M., and I. R. Prucha (1994a): "Generic Uniform Convergence and Equicontinuity Concepts for Random Functions," *Journal of Econometrics*, 60, 23-63.
- [103] Pötscher, B.M., and I. R. Prucha (1994b): "On the Formulation of Uniform Laws of Large Numbers: A Truncation Approach," *Statistics*, 25, 343-360.
- [104] Pötscher, B.M. and I. R. Prucha (1997): *Dynamic Nonlinear Econometric Models*. Springer-Verlag, New York.
- [105] Ripley, B.D. (1981): *Spatial Statistics*. Wiley, New York.
- [106] Robinson, P.M. (2007a): "Nonparametric Spectrum Estimation for Spatial Data," *Journal of Statistical Planning and Inference*, 137, 1024-1034.
- [107] Robinson, P.M. (2007b): "Efficient Estimation of the Semiparametric Spatial Autoregressive Model," forthcoming in *Journal of Econometrics*.
- [108] Rosenblatt, M. (1956): "A Central Limit Theorem and Strong Mixing Condition," *Proceedings of National Academie of Sciences of U.S.A.*, 42, 43-47.

- [109] Rosenblatt, M. (1980): "Linear Processes and Bispectra," *Journal of Applied Probability*, 17, 265-270.
- [110] Sain, S.R., and N. Cressie (2007): "A Spatial Model for Multivariate Lattice Data," forthcoming in *Journal of Econometrics*.
- [111] Smythe, R. (1973): "Strong Laws of Large Numbers for r -Dimensional Arrays of Random Variables," *Annals of Probability*, 1, 164-170.
- [112] Smythe, R. (1974): "Sums of Independent Random Variables on Partially Ordered Sets," *Annals of Probability*, 2, 906-917.
- [113] 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.
- [114] Su, L., and Z. Yang (2007): "QML Estimation of Dynamic Panel Data Models with Spatial Errors," Working Paper, Peking University and Singapore Management University.
- [115] Takahata, H. (1983): "On the Rates in the Central Limit Theorem for Weakly Dependent Random Fields," *Z. Wahrsch. verw. Gebiete*, 64, 445-456.
- [116] Tran, L. (1990): "Kernel Density Estimation on Random Fields," *Journal of Multivariate Analysis*, 34, 37-53.
- [117] 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.

- [118] 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.
- [119] Wald, A. (1949): "Note on the Consistency of Maximum Likelihood Estimate," *Annals of Mathematical Statistics*, 20, 595-601.
- [120] Withers, C. S. (1981): "Conditions for Linear Processes to be Strong Mixing," *Z. Wahrsch. verw. Gebiete*, 57, 477-480.
- [121] White, H. (1980): *Asymptotic Theory for Econometricians*. New York: Academic Press.
- [122] Whittle, P. (1954): "On Stationary Processes in the Plane," *Biometrika*, 41, 434-449.
- [123] Wooldridge, J. (1986): "Asymptotic Properties of Econometric Estimators," University of California San Diego, Department of Economics. Ph.D. Dissertation.
- [124] Wooldridge, J., and H. White (1988): "Some Invariance Principles and Central Limit Theorems for Dependent Heterogeneous Processes," *Econometric Theory*, 4, 210-230.