Metrika
DOI 10.1007/s00184-009-0233-8

Local linear M-estimation for spatial processes in fixed-design models

Jia Chen · Li-Xin Zhang

Received: 19 May 2008

Abstract We investigate the local linear M-estimation for regression in a fixed-design model when the errors are from a strongly mixing random field. We establish the weak and strong consistency as well as the asymptotic normality of the local linear M-estimator. The conditions on $\rho(\cdot)$ used in this paper are mild and allow many important special cases such as the least square estimator and the least absolute distance estimator.

Keywords Asymptotic normality · Consistency · Fixed-design model · Local linear M-estimator · Strongly mixing random fields

1 Introduction

Spatial statistics has recently received increasing attention. Applications of spatial statistical models are extremely numerous in many fields, such as econometrics, epidemiology, environmental science and image analysis, among others. Many authors have studied parametric methods for statistical inference in spatial models. However, one often encounters situations where a particular parametric model can not be adopted with confidence and thus a nonparametric method is used as an alternative. In the last two decades, some developments in nonparametric estimation for spatial processes

J. Chen (✉) · L.-X. Zhang
Department of Mathematics, Zhejiang University, Yuquan Campus, 310027 Hangzhou, China
e-mail: chenjiazju@yahoo.com.cn

L.-X. Zhang
e-mail: stazlx@zju.edu.cn

J. Chen
School of Economics, The University of Adelaide, Adelaide, SA 5005, Australia

Published online: 30 January 2009

have been achieved. Among them, Tran (1990), Carbon et al. (1997), Hallin et al. (2001, 2004a) and Lee et al. (2004) discussed the problems of density estimation for spatial processes. Lu and Chen (2004) obtained the weak consistency for kernel regression estimation. Hallin et al. (2004b) recently established weak consistency and asymptotic normality for local linear regression estimation.

Let $\mathbf{Z}^2$ be the set of integer lattice points in the 2-dimensional Euclidean space. For $\mathbf{n} = (n_1, n_2)^\top \in \mathbf{Z}^2$, define the rectangle $\Lambda_\mathbf{n}$ as $\Lambda_\mathbf{n} := \{\mathbf{i} = (i_1, i_2)^\top \in \mathbf{Z}^2 : 1 \le i_k \le n_k, k = 1, 2\}$. The notation $\mathbf{n} \to \infty$ stands for $\min_{1\le k\le 2} n_k \to \infty$ and $|\frac{n_i}{n_j}| \le C$ for some $0 < C < \infty$, $1 \le i, j \le 2$. When $\mathbf{n} \to \infty$, all the component of $\mathbf{n}$ tend to infinity at the same rate.

For two integer-valued vectors $\mathbf{i}, \mathbf{j} \in \mathbf{Z}^2$, $\mathbf{i}/\mathbf{j}$ denotes for $(i_1/j_1, i_2/j_2)^\top$. In this paper, we are interested in the following fixed design regression model

$$Y_\mathbf{i} = m(\mathbf{i}/\mathbf{n}) + \varepsilon_\mathbf{i}, \quad \mathbf{i} \in \Lambda_\mathbf{n}, \tag{1.1}$$

where $m(\cdot)$ is an unknown smooth function and $\{\varepsilon_\mathbf{i}, \mathbf{i} \in \mathbf{Z}^2\}$ is a real valued, zero mean stationary random field. In fact, the results of this paper still hold for general $\mathbf{n} = (n_1, \ldots, n_d)$ and $\mathbf{i} = (i_1, \ldots, i_d)$, $d \ge 1$. However, when $d \ge 3$, the problem of the curse of dimensionality may occur. So we only consider the case $d = 2$. Hall and Hart (1990), Bosq (1993) studied model (1.1) for time series case (i.e. $\mathbf{n} = n \in \mathbf{Z}$). Machkouri (2005, 2007) considered the model when $\mathbf{n} = (n, \ldots, n)^T$, $n \in \mathbf{Z}$. In particular, Machkouri (2005) established the asymptotic normality of the Nadaraya–Watson type estimator by adapting Lindeberg's method. In Machkouri (2007), the Nadaraya–Watson type estimator of $m(\cdot)$ was studied and a uniform convergence rate was obtained by an exponential inequality.

It is well known that regression estimation is an important topic in nonparametric statistics. It has various applications in econometrics, finance, control systems as well as other areas. There is a vast of literature on the nonparametric estimation of regression functions and many nonparametric estimators have been proposed, such as Nadaraya–Watson type estimators and local polynomial estimators. It is known that local polynomial smoothers have significant advantages over the Nadaraya–Watson regression estimators, as they reduce the bias and cope well with the edge effects. So they have become popular because of their attractive properties. We refer to Fan and Gijbels (1996) and Fan (1992, 1993) for the details. However, local polynomial smoothers are not robust. In order to overcome this drawback, Fan and Gijbels (1996) proposed a local robust technique. And Fan and Jiang (1999) studied local linear M-estimators with variable bandwidth for independent data and showed that they carry over the desirable properties of local linear smoothers. Jiang and Mack (2001) studied robust local polynomial regression for processes satisfying some mixing conditions. Cai and Ould-Saïd (2003) studied local M-estimators for mixing time series and established weak and strong consistency as well as asymptotic normality of the estimators.

The purpose of our paper is to investigate the local linear M-estimator of the regression function in a fixed design model where the errors are from a stationary and strongly mixing random field. We adopt the method used in the proof of Theorem 2.2 in Peligrad and Utev (1997) to show the asymptotic normality of the proposed local linear M-estimator, which is completely different from the method used in Machkouri

(2005). Furthermore, we obtain strong consistency of the estimator with the help of an exponential inequality.

Let S and S' be two subsets of $\mathbf{Z}^2$. The Borel fields $\mathcal{B}(S) = \mathcal{B}(\varepsilon_{\mathbf{i}}, \mathbf{i} \in S)$ and $\mathcal{B}(S') = \mathcal{B}(\varepsilon_{\mathbf{i}}, \mathbf{i} \in S')$ are the $\sigma-$fields generated by $\varepsilon_{\mathbf{i}}$ with $\mathbf{i}$ being elements of S and S', respectively. Let $d(S, S')$ be the Euclidean distance between S and S', i.e. $d(S, S') = \min_{\mathbf{i}\in S\ \mathbf{j}\in S'} \|\mathbf{i} - \mathbf{j}\|$, where $\|\cdot\|$ is the Euclidean norm. We assume that $\{\varepsilon_{\mathbf{i}}\}$ is stationary and mixing, with spatial mixing coefficient satisfying the following condition as defined in Tran (1990): there exist a function $\varphi(t) \downarrow 0$ as $t \to \infty$, and a function $\phi : \mathbf{N}^2 \to \mathbf{R}^+$ symmetric and nondecreasing in each of its two arguments, such that

$$\begin{aligned}\alpha(\mathcal{B}(S), \mathcal{B}(S')) &:= \sup\left\{|P(AB) - P(A)P(B)|,\ A \in \mathcal{B}(S),\ B \in \mathcal{B}(S')\right\}\\ &\le \phi(\mathrm{Card}(S), \mathrm{Card}(S'))\varphi(d(S, S')),\end{aligned} \tag{1.2}$$

where $\mathrm{Card}(S)$ is the cardinality of S. In the case $\phi \equiv 1$, the random fields $\{\varepsilon_{\mathbf{i}}\}$ is called strongly mixing. In the serial case, many stochastic processes and time series are known to be strongly mixing. Withers (1981) obtained various conditions for linear processes to be strongly mixing. Note that autoregressive and more general nonlinear time series models are strongly mixing with exponential mixing rates under certain weak assumptions, see Pham and Tran (1985), Pham (1986), Lu (1998) for details.

In this paper, we study the local linear M-estimator of m when $\{\varepsilon_{\mathbf{i}}, \mathbf{i} \in Z^2\}$ is a strongly mixing random field, i.e., it satisfies (1.2) with $\phi \equiv 1$. Our paper is organized as follows. In Sect. 2, we present the local M-estimator and the assumptions. In Sect. 3, we give our main results. Proofs and technical lemmas are concentrated in Sect. 4.

2 Local linear M-estimator and assumptions

We first present the local linear M-smoother. Let $\mathbf{x}_0 = (x_{01}, x_{02})^\top$ be a fixed point in $(0, 1)^2$ and we assume that the regression function $m(\mathbf{x})$ is twice differentiable at $\mathbf{x}_0$, then $m(\mathbf{x})$ can be approximated by a linear function in a neighborhood of $\mathbf{x}_0$ as $m(\mathbf{x}) \approx m(\mathbf{x}_0) + (m'(\mathbf{x}_0))^\top(\mathbf{x} - \mathbf{x}_0)$, where $m'(\mathbf{x}_0)$ is the gradient of $m(\cdot)$ at $\mathbf{x}_0$. Let $\mathbf{x}_{\mathbf{i}} = (x_{\mathbf{i}1}, x_{\mathbf{i}2})^\top = \mathbf{i}/\mathbf{n}$, then the robust-type local linear estimators of $m(\cdot)$ and $m'(\cdot)$ are defined as $(\widehat{m}_{\mathbf{n}}(\mathbf{x}_0), \widehat{m}'_{\mathbf{n}}(\mathbf{x}_0)) = (\widehat{a}, \widehat{b})$, where

$$(\widehat{a}, \widehat{b}) = \arg \min_{a\in\mathbf{R},\mathbf{b}\in\mathbf{R}^2} \sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} \rho\left(Y_{\mathbf{i}} - a - \mathbf{b}^\top(\mathbf{x}_{\mathbf{i}} - \mathbf{x}_0)\right) K\left(\frac{\mathbf{x}_0 - \mathbf{x}_{\mathbf{i}}}{h_{\mathbf{n}}}\right) \tag{2.1}$$

or to satisfy the local estimation equations:

$$\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} \psi\left(Y_{\mathbf{i}} - a - \mathbf{b}^\top(\mathbf{x}_{\mathbf{i}} - \mathbf{x}_0)\right) K\left(\frac{\mathbf{x}_0 - \mathbf{x}_{\mathbf{i}}}{h_{\mathbf{n}}}\right)\begin{pmatrix}1\\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h_{\mathbf{n}}}\end{pmatrix} = \begin{pmatrix}0\\ \mathbf{0}\end{pmatrix}, \tag{2.2}$$

where $\rho(\cdot)$ is a given outlier-resistant loss function and $\psi(\cdot)$ is its derivative, $\mathbf{0} = (0, 0)^\top \in \mathbf{R}^2$, $K(\cdot)$ is a kernel function, and $h_{\mathbf{n}}$ is a sequence of positive numbers tending to zero. In the rest of the paper, we will write h instead of $h_{\mathbf{n}}$ for the sake of simplicity and denote by C a positive constant whose value may vary from place to place.

Let $\widehat{m}_{\mathbf{n}}(\mathbf{x}_0)$ and $\widehat{m}'_{\mathbf{n}}(\mathbf{x}_0)$ be the solutions of (2.2). They are the M-type estimators of $m(\cdot)$ and $m'(\cdot)$ at $\mathbf{x}_0$. For $\mathbf{n} \in \mathbf{Z}^2$, denote $|\mathbf{n}| = \prod_{k=1}^{2} n_k$. We now list the assumptions of our paper.

(A1) $m(\cdot)$ has continuous second partial derivatives in a neighborhood of $\mathbf{x}_0$.
(A2) $K(\cdot)$ is a bounded probability density function with bounded support, say $[-1, 1]^2$. Furthermore, $K(\cdot)$ is Lipschitz continuous, i.e. $|K(\mathbf{u}) - K(\mathbf{v})| \leq C\|\mathbf{u} - \mathbf{v}\|_\infty$, where $\|\mathbf{u} - \mathbf{v}\|_\infty = \max_{1\leq k\leq 2} |u_k - v_k|$.
(A3) $\rho(\cdot)$ is a convex function. Let $\psi(\cdot)$ be any choice of the derivative of $\rho(\cdot)$ and denote by $\mathcal{D}$ the set of discontinuity points of $\psi(\cdot)$, which is the same for all choices of $\psi(\cdot)$.
(A4) $F(\mathcal{D}) = 0$, where $F(\cdot)$ denotes the common distribution of $\varepsilon_{\mathbf{i}}$.
(A5) There exists a constant $A_1 > 0$, such that as $|z| \to 0$, $E[\psi(\varepsilon_{\mathbf{1}}+z)] = A_1 z + o(z)$.
(A6) $0 < E\psi^2(\varepsilon_{\mathbf{0}}) < \infty$, $\sum_{\mathbf{i}\in\mathbf{Z}^2} |E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{i}})| < \infty$ and $\sum_{\mathbf{i}\in Z^2} E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{i}}) > 0$. Furthermore, there exists some $\gamma > 0$ such that $E|\psi(\varepsilon_{\mathbf{0}})|^{2+\gamma} < \infty$.
(A7) $\max\{E(\psi(\varepsilon_{\mathbf{1}} + z) - \psi(\varepsilon_{\mathbf{1}}))^2,\ E|\psi(\varepsilon_{\mathbf{1}} + z) - \psi(\varepsilon_{\mathbf{1}})|^{2+\gamma}\} \leq A_2(z)$ holds for all sufficiently small $|z|$ and $A_2(\cdot)$ is continuous at $u = 0$ with $A_2(0) = 0$.
(A8) $\sum_{k=1}^{\infty} k[\varphi(k)]^{\frac{\gamma}{2+\gamma}} < \infty$, where γ is the same as in (A6) and $\varphi(\cdot)$ is the spatial mixing coefficient function defined as in (1.2).
(A9) There exist a positive constant β, $\frac{2}{2+\gamma} < \beta < 1$ and a sequence of $\mathbf{q}_{\mathbf{n}} = \mathbf{q} = (q_1, q_2)$, $1 \leq q_k \leq n_k$, $1 \leq k \leq 2$, such that $|\mathbf{n}|^{1+\beta}h^2/|\mathbf{q}| \to \infty$, $(|\mathbf{q}|h^2)/(|\mathbf{n}|^\beta \log|\mathbf{n}|) \to \infty$ and $\sum_{n_1,n_2\geq 1}(|\mathbf{n}|^\beta/h^2)\varphi(\min_{1\leq k\leq 2} \frac{n_k}{2q_k}) < \infty$.

Remark 2.1 (A1), (A2), (A6) and (A8) are mild conditions and they are common in literature about robust type local linear estimation under mixing dependence, such as Jiang and Mack (2001) and Cai and Ould-Saïd (2003). (A9) is imposed to derive the strong consistency of the local M-estimator and it is also assumed in Cai and Ould-Saïd (2003). The Lipschitz continuity of $K(\cdot)$ is employed a lot in the context of nonparametric estimation for fixed design models, see e.g. Machkouri (2007) and Peng and Yao (2004), among others. By Theorem 17.2.2 in Ibragimov and Linnik (1971), we know that,

$$|\mathrm{Cov}(\psi(\varepsilon_{\mathbf{0}}), \psi(\varepsilon_{\mathbf{i}}))| \leq C\,[\varphi(\|\mathbf{i}\|)]^{\gamma/(2+\gamma)}\,(E|\psi(\varepsilon_{\mathbf{0}})|^{2+\gamma})^{2/(2+\gamma)}.$$

Therefore, $\sum_{\mathbf{i}\in\mathbf{Z}^2} |E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{i}})| < \infty$ in (A6) is implied by $E|\psi(\varepsilon_{\mathbf{0}})|^{2+\gamma} < \infty$ and (A8).

Remark 2.2 The Conditions (A3), (A4),(A5) and (A7) imposed on $\psi(\cdot)$ are mild and they hold for some well known estimators such as the least square estimator (LSE), the least absolute distance estimator (LADE) and the mixed LSE and LADE. (A4) is imposed to provide unique values for certain functionals of $\psi(\cdot)$ which appear in the discussion and it automatically holds when $\rho(\cdot)$ is differentiable. Note that in

Fan and Jiang (1999), $\psi(\cdot)$ was assumed to be differentiable and satisfy

$$E \sup_{|z|\le\delta} |\psi(\varepsilon_{\mathbf{1}}+z)-\psi(\varepsilon_{\mathbf{1}})-\psi'(\varepsilon_{\mathbf{1}})z| = o(\delta),$$
$$E \sup_{|z|\le\delta} |\psi'(\varepsilon_{\mathbf{1}}+z)-\psi'(\varepsilon_{\mathbf{1}})| = o(1). \tag{2.3}$$

However, we do not need even the continuity of $\psi(\cdot)$ and we replace their condition (2.3) by (A5) and (A7). For more details about these conditions we refer to Bai et al. (1992).

Remark 2.3 By (A9), we know $|\mathbf{n}|h^4 \to \infty$. As an example, we suppose that $\varphi(k) = O(k^{-\theta})$ as $k \to \infty$, where $\theta > \frac{(5+3\beta)}{(1-\beta)}$ and $\frac{2}{2+\gamma} < \beta < 1$. Then we choose $h = |\mathbf{n}|^{-\frac{1-\beta}{8}}$ and $\mathbf{q} = (q,q)$ with $q = |\mathbf{n}|^{\frac{\beta+1}{4}}$. As $\beta > \frac{2}{2+\gamma}$, $\theta > \frac{(5+3\beta)}{(1-\beta)} > \frac{2(2+\gamma)}{\gamma}$. So (A8) is satisfied. Besides, $|\mathbf{n}|^{1+\beta}h^2/q^2 = |\mathbf{n}|^{\frac{1+3\beta}{4}} \to \infty$, $q^2h^2/(|\mathbf{n}|^{\beta}\log|\mathbf{n}|) = |\mathbf{n}|^{\frac{1-\beta}{4}}/\log|\mathbf{n}| \to \infty$. When $|\frac{n_i}{n_j}| \le C$, $i,j=1,2$,

$$(|\mathbf{n}|^{\beta}/h^2)\varphi\left(\min_{1\le k\le 2}\frac{n_k}{2q}\right) = O\left(|\mathbf{n}|^{\frac{3}{4}\beta+\frac{1}{4}}|\mathbf{n}|^{-\frac{(1-\beta)\theta}{4}}\right) = O\left(|\mathbf{n}|^{-(\frac{(1-\beta)\theta}{4}-\frac{3}{4}\beta-\frac{1}{4})}\right).$$

As $\frac{1-\beta}{4}\theta - \frac{3}{4}\beta - \frac{1}{4} > 1$,

$$\sum_{n_1,n_2\ge1}(|\mathbf{n}|^{\beta}/h^2)\varphi\left(\min_{1\le k\le 2}\frac{n_k}{q}\right) < \infty.$$

Therefore, (A9) is also satisfied.

3 Main results

Our first result is about the consistency of the local M-estimator.

Theorem 3.1 *Suppose that* (A1)–(A8) *hold and* $n_k h \to \infty$ *for* $k=1,2$, *then there exist solutions* $\widehat{m}_{\mathbf{n}}(\mathbf{x}_0)$ *and* $\widehat{m}'_{\mathbf{n}}(\mathbf{x}_0)$ *to* (2.2), *such that*

$$\begin{pmatrix} \widehat{m}_{\mathbf{n}}(\mathbf{x}_0)-m(\mathbf{x}_0) \\ h(\widehat{m}'_{\mathbf{n}}(\mathbf{x}_0)-m'(\mathbf{x}_0)) \end{pmatrix} \xrightarrow{P} 0. \tag{3.1}$$

If, in addition, (A9) *holds, then*

$$\begin{pmatrix} \widehat{m}_{\mathbf{n}}(\mathbf{x}_0)-m(\mathbf{x}_0) \\ h(\widehat{m}'_{\mathbf{n}}(\mathbf{x}_0)-m'(\mathbf{x}_0)) \end{pmatrix} \longrightarrow 0 \quad \text{a.s.} \tag{3.2}$$

The next result is about the asymptotic normality of the proposed estimator, from which the asymptotic bias and variance of the estimator can be derived.

Theorem 3.2 *Suppose that conditions* (A1)–(A8) *hold and* $n_k h \to \infty$ *for* $k = 1, 2$. *Then, we have*

$$\sqrt{|\mathbf{n}|h^2}\left[\begin{pmatrix} \widehat{m}_{\mathbf{n}}(\mathbf{x}_0) - m(\mathbf{x}_0) \\ h(\widehat{m}'_{\mathbf{n}}(\mathbf{x}_0) - m'(\mathbf{x}_0)) \end{pmatrix} - \frac{1}{2}h^2\mathbf{U}^{-1}\mathbf{M}\right] \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix}, \Sigma(\mathbf{x}_0)\right), \tag{3.3}$$

where

$$\mathbf{U} = \begin{pmatrix} \int K(\mathbf{u})d\mathbf{u} & -\int K(\mathbf{u})\mathbf{u}^\top d\mathbf{u} \\ -\int K(\mathbf{u})\mathbf{u}d\mathbf{u} & \int K(\mathbf{u})\mathbf{u}\mathbf{u}^\top d\mathbf{u} \end{pmatrix},$$
$$\mathbf{M} = \int \mathbf{u}^\top m''(\mathbf{x}_0)\mathbf{u}K(\mathbf{u}) \begin{pmatrix} 1 \\ -\mathbf{u} \end{pmatrix} d\mathbf{u},$$
$$\Sigma(\mathbf{x}_0) = A_1^{-2}\left(\sum_{\mathbf{j}\in Z^2} E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{j}})\right)\mathbf{U}^{-1}\mathbf{V}\mathbf{U}^{-1},$$

and

$$\mathbf{V} = \begin{pmatrix} \int K^2(\mathbf{u})d\mathbf{u} & -\int \mathbf{u}^\top K^2(\mathbf{u})d\mathbf{u} \\ -\int \mathbf{u}K^2(\mathbf{u})d\mathbf{u} & \int \mathbf{u}\mathbf{u}^\top K^2(\mathbf{u})d\mathbf{u} \end{pmatrix}.$$

Remark 3.1 Following the proofs in Sect. 4, we can show that the above theorems still hold for general $\mathbf{n} = (n_1, \dots, n_d)$ and $\mathbf{i} = (i_1, \dots, i_d)$, $d \geq 1$. However, when $d \geq 3$, the problem of the curse of dimensionality may occur. Hence, we only consider the special case of $d = 2$ in this paper.

Since the proofs of Theorems 3.1 and 3.2 are quite technical, we relegate them to the next section. Next, we shall give some special cases as corollaries of our results. Some of the following examples have been discussed by other authors.

Example 3.1 Let $\rho(z) = z^2$ and $\psi(z) = 2z$, which corresponds to least square estimation. Let U, M and V be defined as in Theorem 3.2 and

$$\Sigma_1(\mathbf{x}_0) = \frac{1}{4}\left(\sum_{\mathbf{j}\in Z^2} E\varepsilon_{\mathbf{0}}\varepsilon_{\mathbf{j}}\right)\mathbf{U}^{-1}\mathbf{V}\mathbf{U}^{-1}.$$

Then, under some mild assumptions, we have, by Theorem 3.2,

$$\sqrt{|\mathbf{n}|h^2}\left[\begin{pmatrix} \widehat{m}_1(\mathbf{x}_0) - m(\mathbf{x}_0) \\ h(\widehat{m}'_1(\mathbf{x}_0) - m'(\mathbf{x}_0)) \end{pmatrix} - \frac{1}{2}h^2\mathbf{U}^{-1}\mathbf{M}\right] \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix}, \Sigma_1(\mathbf{x}_0)\right), \tag{3.4}$$

where $\widehat{m}_1(\mathbf{x}_0)$ and $\widehat{m}_1'(\mathbf{x}_0)$ are the least square estimators of $m(\mathbf{x}_0)$ and $m'(\mathbf{x}_0)$, respectively.

Example 3.2 Consider $\rho_p(z) = z(p - I\{z \le 0\})$, then $\psi_p(z) = p - I\{z \le 0\}$, which corresponds to quantile regression estimation. Let $u_{\mathbf{i}} = \varepsilon_{\mathbf{i}} - z_p$, where z_p is the pth quantile of $\varepsilon_{\mathbf{1}}$. If $\psi(\cdot)$ and $\varepsilon_{\mathbf{i}}$ are substituted by $\psi_p(\cdot)$ and $u_{\mathbf{i}}$, then (A5) and (A7) are satisfied with $A_1 = f_\varepsilon(z_p)$ and $A_2(z) = f_\varepsilon(z_p)z$, where $f_\varepsilon(\cdot)$ is the density function of ε. Moreover, if

$$\sum_{\mathbf{i}\in\mathbf{Z}^2} \left| P\left(\varepsilon_{\mathbf{0}} \le z_p, \varepsilon_{\mathbf{i}} \le z_p\right) - p^2 \right| < \infty \quad \text{and} \quad \sum_{\mathbf{i}\in\mathbf{Z}^2} \left(P\left(\varepsilon_{\mathbf{0}} \le z_p, \varepsilon_{\mathbf{i}} \le z_p\right) - p^2 \right) > 0,$$

then under conditions (A1), (A2), (A8) and (A9), we have

$$\begin{aligned} &\sqrt{|\mathbf{n}|h^2} \left[\begin{pmatrix} \widehat{q}_p(\mathbf{x}_0) - q_p(\mathbf{x}_0) \\ h(\widehat{q}_p'(\mathbf{x}_0) - q_p'(\mathbf{x}_0)) \end{pmatrix} - \frac{1}{2}h^2\mathbf{U}^{-1}\mathbf{M}_2 \right] \\ &\quad \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix}, \Sigma_2(\mathbf{x}_0) \right), \end{aligned} \tag{3.5}$$

where

$$\begin{aligned} \Sigma_2(\mathbf{x}_0) &= \frac{1}{f^2(z_p)} \left[\sum_{\mathbf{i}\in\mathbf{Z}^2} \left(P\left(\varepsilon_{\mathbf{0}} \le z_p, \varepsilon_{\mathbf{i}} \le z_p\right) - p^2 \right) \right] \mathbf{U}^{-1}\mathbf{V}\mathbf{U}^{-1}, \\ \mathbf{M}_2 &= \int \mathbf{u}^\top q_p''(\mathbf{x}_0)\mathbf{u}K(\mathbf{u}) \begin{pmatrix} 1 \\ -\mathbf{u} \end{pmatrix} d\mathbf{u}, \end{aligned}$$

$q_p''(\mathbf{x}_0)$ is the matrix of the second derivatives of $q_p(\cdot)$ at $\mathbf{x}_0$, $\widehat{q}_p(\mathbf{x}_0)$ and $\widehat{q}\,'_p(\mathbf{x}_0)$ are the local linear quantile estimators of $q_p(\cdot)$ and $q_p'(\cdot)$ at $\mathbf{x}_0$. When the covariates $\mathbf{X}_{\mathbf{i}}$, $\mathbf{i} \in \Lambda_{\mathbf{n}}$, are random, Hallin et al. (2005) studied local linear spatial quantile regression estimation and obtained a Bahadur representation for the estimators of $q_p(\cdot)$ and its derivatives.

4 Proofs

In this section, we will present the detailed proofs of Theorems 3.1 and 3.2. Let $\mathbf{r} = (r_1, h\mathbf{r}_2^\top)^\top$, $\mathbf{r}_0 = (m(\mathbf{x}_0),\ hm'(\mathbf{x}_0)^\top)^\top$ and define

$$\begin{aligned} L_{\mathbf{n}}(\mathbf{r}) &= \frac{1}{|\mathbf{n}|h^2} \sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} \rho\left(Y_{\mathbf{i}} - \mathbf{r}^\top \begin{pmatrix} 1 \\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h} \end{pmatrix} \right) K\left(\frac{\mathbf{x}_0 - \mathbf{x}_{\mathbf{i}}}{h} \right), \\ \Psi_{\mathbf{n}}(\mathbf{r}) &= \frac{1}{|\mathbf{n}|h^2} \sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} \psi\left(Y_{\mathbf{i}} - \mathbf{r}^\top \begin{pmatrix} 1 \\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h} \end{pmatrix} \right) K\left(\frac{\mathbf{x}_0 - \mathbf{x}_{\mathbf{i}}}{h} \right) \begin{pmatrix} 1 \\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h} \end{pmatrix}, \\ \Psi_{\mathbf{n}0} &= \frac{1}{|\mathbf{n}|h^2} \sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} \psi(\varepsilon_{\mathbf{i}}) K\left(\frac{\mathbf{x}_0 - \mathbf{x}_{\mathbf{i}}}{h} \right) \begin{pmatrix} 1 \\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h} \end{pmatrix}. \end{aligned}$$

We first give some preliminary lemmas that are useful.

Lemma 4.1 *Under* (A1)–(A6) *and* $n_k h \to \infty$ *for* $k = 1, 2$, *we have*

$$\mathrm{Var}(\Psi_{\mathbf{n}0}) = \frac{1}{|\mathbf{n}|h^2}\left[\sum_{\mathbf{i}\in\mathbf{Z}^2} E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{i}})\right]\mathbf{V}(1+o(1)). \tag{4.1}$$

Proof We first calculate $\mathrm{Var}(|\mathbf{n}|h^2\mathbf{c}^\top\Psi_{\mathbf{n}0})$ for any $\mathbf{c}\in\mathbf{R}^3$. In view of (A5), we have $E\Psi_{\mathbf{n}0} = (0, \mathbf{0}^\top)^\top$. Define $b_{\mathbf{i}} = K(\frac{\mathbf{x}_0-\mathbf{x}_{\mathbf{i}}}{h})\mathbf{c}^\top\begin{pmatrix}1\\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h}\end{pmatrix}$. Let $\{d_{\mathbf{n}}\}$ be a sequence of integers that satisfies $d_{\mathbf{n}}\to\infty$ and $d_{\mathbf{n}} = o(n_k h)$, $k = 1, 2$. As $\mathbf{x}_0\in(0,1)^2$, $K(\frac{\mathbf{x}_0-\mathbf{x}_{\mathbf{i}}}{h}) = 0$ for $\mathbf{i}\in\{\mathbf{i}\in\Lambda_{\mathbf{n}}$: for some $1\le k\le 2,\ 1\le i_k\le d_{\mathbf{n}}$ or $n_k - d_{\mathbf{n}} + 1\le i_k\le n_k\}$ when $\mathbf{n}$ is large enough. As a result,

$$\begin{aligned}
&\mathrm{Var}\left(|\mathbf{n}|h^2\mathbf{c}^\top\Psi_{\mathbf{n}0}\right)\\
&= \sum_{\substack{i_k=d_n+1\\k=1,2}}^{n_k-d_{\mathbf{n}}}\ \sum_{\substack{j_k=i_k-d_{\mathbf{n}}\\k=1,2}}^{i_k+d_{\mathbf{n}}} b_{\mathbf{i}}b_{\mathbf{j}}E\psi(\varepsilon_{\mathbf{i}})\psi(\varepsilon_{\mathbf{j}}) + \sum_{\substack{i_k=d_n+1\\k=1,2}}^{n_k-d_{\mathbf{n}}}\ \sum_{\substack{|j_k-i_k|>d_{\mathbf{n}}\\k=1,2}} b_{\mathbf{i}}b_{\mathbf{j}}E\psi(\varepsilon_{\mathbf{i}})\psi(\varepsilon_{\mathbf{j}})\\
&=: I_{\mathbf{n}1} + I_{\mathbf{n}2}.
\end{aligned} \tag{4.2}$$

$$\begin{aligned}
&I_{\mathbf{n}1}\\
&= \sum_{\substack{i_k=d_n+1\\k=1,2}}^{n_k-d_{\mathbf{n}}}\ \sum_{\substack{j_k=i_k-d_{\mathbf{n}}\\k=1,2}}^{i_k+d_{\mathbf{n}}} E\psi(\varepsilon_{\mathbf{i}})\psi(\varepsilon_{\mathbf{j}})\left[b_{\mathbf{i}}^2 - b_{\mathbf{i}}(b_{\mathbf{i}}-b_{\mathbf{j}})\right]\\
&= \sum_{\substack{i_k=d_n+1\\k=1,2}}^{n_k-d_{\mathbf{n}}}\ \sum_{\substack{j_k=i_k-d_{\mathbf{n}}\\k=1,2}}^{i_k+d_{\mathbf{n}}} b_{\mathbf{i}}^2E\psi(\varepsilon_{\mathbf{i}})\psi(\varepsilon_{\mathbf{j}}) - \sum_{\substack{i_k=d_n+1\\k=1,2}}^{n_k-d_{\mathbf{n}}}\ \sum_{\substack{j_k=i_k-d_{\mathbf{n}}\\k=1,2}}^{i_k+d_{\mathbf{n}}} b_{\mathbf{i}}(b_{\mathbf{i}}-b_{\mathbf{j}})E\psi(\varepsilon_{\mathbf{i}})\psi(\varepsilon_{\mathbf{j}})\\
&=: I_{\mathbf{n}3} + I_{\mathbf{n}4}.
\end{aligned} \tag{4.3}$$

As $d_{\mathbf{n}} = o(n_k h)$, $k = 1, 2$, we have

$$\begin{aligned}
I_{\mathbf{n}3} &= \sum_{\substack{i_k=d_n+1\\k=1,2}}^{n_k-d_{\mathbf{n}}} b_{\mathbf{i}}^2 \sum_{\substack{j_k=-d_{\mathbf{n}}\\k=1,2}}^{d_{\mathbf{n}}} E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{j}})\\
&= \left[\sum_{\substack{j_k=-d_{\mathbf{n}}\\k=1,2}}^{d_{\mathbf{n}}} E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{j}})\right]|\mathbf{n}|\int\left(K^2\left(\frac{\mathbf{x}_0-\mathbf{x}}{h}\right)\left[\mathbf{c}^\top\begin{pmatrix}1\\ \frac{\mathbf{x}-\mathbf{x}_0}{h}\end{pmatrix}\right]^2\right)d\mathbf{x}(1+o(1))\\
&= \left[\sum_{\substack{j_k=-d_{\mathbf{n}}\\k=1,2}}^{d_{\mathbf{n}}} E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{j}})\right]|\mathbf{n}|h^2\mathbf{c}^\top\mathbf{V}\mathbf{c}(1+o(1)).
\end{aligned} \tag{4.4}$$

As $K(\cdot)$ is Lipschitz continuous, we have

$$|b_{\mathbf{i}} - b_{\mathbf{j}}| \leq \frac{C}{h} \|\mathbf{i}/\mathbf{n} - \mathbf{j}/\mathbf{n}\|_\infty \leq \frac{C}{h} \max_{1\leq k\leq 2} \left|\frac{d_{\mathbf{n}}}{n_k}\right| = o(1)$$

for all $d_{\mathbf{n}} + 1 \leq i_k \leq n_k - d_{\mathbf{n}}$, $|j_k - i_k| \leq d_{\mathbf{n}}$, $k = 1, 2$. Therefore, by (A6),

$$|I_{\mathbf{n}4}| = o\left(\sum_{\substack{i_k=d_n+1\\k=1,2}}^{n_k-d_{\mathbf{n}}} |b_{\mathbf{i}}| \sum_{\substack{j_k=-d_{\mathbf{n}}\\k=1,2}}^{d_{\mathbf{n}}} \left|E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{j}})\right|\right) = o(|\mathbf{n}|h^2). \tag{4.5}$$

By (4.3)–(4.5), we have

$$I_{\mathbf{n}1} = |\mathbf{n}|h^2 \left[\sum_{\mathbf{i}\in\mathbf{Z}^2} E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{i}})\right] \mathbf{c}^\top\mathbf{V}\mathbf{c}(1 + o(1)). \tag{4.6}$$

As $\max_{\mathbf{i}\in\Lambda_{\mathbf{n}}} |b_{\mathbf{i}}| \leq C$, by (A6) we have

$$|I_{\mathbf{n}2}| \leq C \sum_{\substack{i_k=d_n+1\\k=1,2}}^{n_k-d_{\mathbf{n}}} |b_{\mathbf{i}}| \sum_{\substack{|j_k|>d_{\mathbf{n}}\\k=1,2}} |E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{j}})| = o(|\mathbf{n}|h^2). \tag{4.7}$$

By (4.2), (4.4) and (4.7), we get

$$\mathrm{Var}\left(\mathbf{c}^\top\Psi_{\mathbf{n}0}\right) = \frac{1}{|\mathbf{n}|h^2} \left[\sum_{\mathbf{i}\in\mathbf{Z}^2} E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{i}})\right] \mathbf{c}^\top\mathbf{V}\mathbf{c}(1 + o(1)), \tag{4.8}$$

which implies (4.1). □

Lemma 4.2 *Under* (A1)–(A8), *we have, for small* δ,

$$\sup_{\|\mathbf{r}-\mathbf{r}_0\|\leq\delta} \left|L_{\mathbf{n}}(\mathbf{r}) - L_{\mathbf{n}}(\mathbf{r}_0) + (\Psi_{\mathbf{n}}(\mathbf{r}_0))^\top(\mathbf{r}-\mathbf{r}_0) - \frac{A_1}{2}(\mathbf{r}-\mathbf{r}_0)^\top\mathbf{U}(\mathbf{r}-\mathbf{r}_0)\right| = o_P(1). \tag{4.9}$$

Proof Let

$$\begin{aligned}&L_{\mathbf{n}}(\mathbf{r}) - L_{\mathbf{n}}(\mathbf{r}_0) + (\Psi_{\mathbf{n}}(\mathbf{r}_0))^\top(\mathbf{r}-\mathbf{r}_0) - E\left[L_{\mathbf{n}}(\mathbf{r}) - L_{\mathbf{n}}(\mathbf{r}_0) + (\Psi_{\mathbf{n}}(\mathbf{r}_0))^\top(\mathbf{r}-\mathbf{r}_0)\right]\\ &=: \frac{1}{|\mathbf{n}|h^2} \sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} (W_{\mathbf{ni}} - EW_{\mathbf{ni}}),\end{aligned}$$

where

$$W_{\mathbf{ni}} = K\left(\frac{\mathbf{x}_0 - \mathbf{x_i}}{h}\right)\left[\rho\left(Y_{\mathbf{i}} - \mathbf{r}^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)\right) - \rho\left(Y_{\mathbf{i}} - \mathbf{r}_0^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)\right)\right.$$
$$\left. + \psi\left(Y_{\mathbf{i}} - \mathbf{r}_0^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)\right)(\mathbf{r}-\mathbf{r}_0)^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)\right].$$

As $\rho(\cdot)$ is convex, we have

$$|W_{\mathbf{ni}}| \le \left|K\left(\frac{\mathbf{x}_0 - \mathbf{x_i}}{h}\right)(\mathbf{r}-\mathbf{r}_0)^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)\right|$$
$$\times\left|\psi\left(Y_{\mathbf{i}} - \mathbf{r}^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)\right) - \psi\left(Y_{\mathbf{i}} - \mathbf{r}_0^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)\right)\right|. \tag{4.10}$$

Let $z_{\mathbf{i}0} = m(\mathbf{x_i}) - \mathbf{r}_0^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)$ and $z_{\mathbf{i}1} = m(\mathbf{x_i}) - \mathbf{r}^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)$, then for $|x_{\mathbf{i}l} - x_{0l}| \le h$, $l = 1, 2$,

$$z_{\mathbf{i}0} = \frac{1}{2}(\mathbf{x_i}-\mathbf{x}_0)^\top m''(\mathbf{x}_0)(\mathbf{x_i}-\mathbf{x}_0) + o\left(\|\mathbf{x_i}-\mathbf{x}_0\|^2\right) = O(h^2) \tag{4.11}$$

and for $\|\mathbf{r}-\mathbf{r}_0\| \le \delta$,

$$|z_{\mathbf{i}1}| \le |z_{\mathbf{i}0}| + \left|(\mathbf{r}-\mathbf{r}_0)^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)\right| = O(h^2) + O(\delta). \tag{4.12}$$

Also by (4.10), we have

$$|W_{\mathbf{ni}}| \le \left|K\left(\frac{\mathbf{x}_0 - \mathbf{x_i}}{h}\right)(\mathbf{r}-\mathbf{r}_0)^\top\left(\begin{smallmatrix}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{smallmatrix}\right)\right|[|\psi(\varepsilon_{\mathbf{i}} + z_{\mathbf{i}1}) - \psi(\varepsilon_{\mathbf{i}})| + |\psi(\varepsilon_{\mathbf{i}} + z_{\mathbf{i}0}) - \psi(\varepsilon_{\mathbf{i}})|].$$

Let $\chi_{\mathbf{ni}} = \psi(\varepsilon_{\mathbf{i}} + z_{\mathbf{i}1}) - \psi(\varepsilon_{\mathbf{i}})$, then by (4.12) and (A7), we know that for small δ and large $\mathbf{n}$, $E\chi_{\mathbf{n0}}^2$ and $E|\chi_{\mathbf{n0}}|^{2+\gamma}$ exist. So by Theorem 17.2.2 in Ibragimov and Linnik (1971),

$$\begin{aligned}
\sum_{\mathbf{i}\in\mathbf{Z}^2} |\mathrm{Cov}(\chi_{\mathbf{n0}}, \chi_{\mathbf{ni}})| &\le E\chi_{\mathbf{n0}}^2 + \sum_{\mathbf{i}\ne\mathbf{0}} |\mathrm{Cov}(\chi_{\mathbf{n0}}, \chi_{\mathbf{ni}})| \\
&\le E\chi_{\mathbf{n0}}^2 + C\sum_{\mathbf{i}\ne\mathbf{0}} [\varphi(\|\mathbf{i}\|)]^{\gamma/(\gamma+2)}\left(E|\chi_{\mathbf{n0}}|^{2+\gamma}\right)^{2/2+\gamma} \\
&\le E\chi_{\mathbf{n0}}^2 + C\left(E|\chi_{\mathbf{n0}}|^{2+\gamma}\right)^{2/2+\gamma}\sum_{i=1}^{\infty} i[\varphi(i)]^{\frac{\gamma}{2+\gamma}} \\
&= O\left(A_2(h^2+\delta)\right),
\end{aligned}$$

where $A_2(\cdot)$ is the function defined in (A7).

Similarly, if we let $\chi'_{\mathbf{ni}}=\psi(\varepsilon_{\mathbf{i}}+z_{\mathbf{i}0})-\psi(\varepsilon_{\mathbf{i}})$, then by (A7), $\sum_{\mathbf{i}\in\mathbf{Z}^2}|\mathrm{Cov}(\chi'_{\mathbf{n0}},\ \chi'_{\mathbf{ni}})|=O(A_2(h^2))$. We can prove, by taking the same argument as the proof of (4.1), that $\mathrm{Var}(\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}W_{\mathbf{ni}})=o(|\mathbf{n}|h^2)$, when δ is sufficiently small and $\mathbf{n}$ is sufficiently large. Therefore

$$\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}(W_{\mathbf{ni}}-EW_{\mathbf{ni}})=o_P\left((|\mathbf{n}|h^2)^{-\frac{1}{2}}\right)=o_P(1),$$

i.e.

$$\begin{aligned}&L_{\mathbf{n}}(\mathbf{r})-L_{\mathbf{n}}(\mathbf{r}_0)+(\Psi_{\mathbf{n}}(\mathbf{r}_0))^\top(\mathbf{r}-\mathbf{r}_0)\\&-E\left[L_{\mathbf{n}}(\mathbf{r})-L_{\mathbf{n}}(\mathbf{r}_0)+(\Psi_{\mathbf{n}}(\mathbf{r}_0))^\top(\mathbf{r}-\mathbf{r}_0)\right]=o_P(1).\end{aligned}\tag{4.13}$$

On the other hand,

$$\begin{aligned}E(L_{\mathbf{n}}(\mathbf{r})-L_{\mathbf{n}}(\mathbf{r}_0))&=\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}EK\left(\frac{\mathbf{x}_0-\mathbf{x_i}}{h}\right)[\rho(\varepsilon_{\mathbf{i}}+z_{\mathbf{i}1})-\rho(\varepsilon_{\mathbf{i}}+z_{\mathbf{i}0})]\\&=\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}K\left(\frac{\mathbf{x}_0-\mathbf{x_i}}{h}\right)\left(\frac{1}{2}A_1z_{\mathbf{i}1}^2-\frac{1}{2}A_1z_{\mathbf{i}0}^2\right)(1+o(1))\\&=\frac{1}{2}A_1\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}\left(K\left(\frac{\mathbf{x}_0-\mathbf{x_i}}{h}\right)\left[(\mathbf{r}-\mathbf{r}_0)^\top\begin{pmatrix}1\\\frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{pmatrix}-2z_{\mathbf{i}0}\right]\right.\\&\qquad\left.\times(\mathbf{r}-\mathbf{r}_0)^\top\begin{pmatrix}1\\\frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{pmatrix}(1+o(1))\right).\end{aligned}\tag{4.14}$$

The second equality above holds because for any small $u\in\mathbf{R}$,

$$\begin{aligned}&\frac{u}{k}\sum_{i=1}^{k}\psi\left(\varepsilon_{\mathbf{1}}+\frac{i-1}{k}u\right)\le\rho(\varepsilon_{\mathbf{1}}+u)-\rho(\varepsilon_{\mathbf{1}})\\&=\sum_{i=1}^{k}\left[\rho\left(\varepsilon_{\mathbf{1}}+\frac{i}{k}u\right)-\rho(\varepsilon_{\mathbf{1}}+\frac{i-1}{k}u)\right]\le\frac{u}{k}\sum_{i=1}^{k}\psi\left(\varepsilon_{\mathbf{1}}+\frac{i}{k}u\right).\end{aligned}$$

By taking expectation and letting $k\to\infty$, we know, by (A5), that

$$E[\rho(\varepsilon_{\mathbf{1}}+u)-\rho(\varepsilon_{\mathbf{1}})]=\frac{1}{2}A_1u^2(1+o(1)).$$

Furthermore,

$$E(\Psi_{\mathbf{n}}(\mathbf{r}_0))^{\top}(\mathbf{r}-\mathbf{r}_0) = \frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} K\left(\frac{\mathbf{x}_0-\mathbf{x}_{\mathbf{i}}}{h}\right) E[\psi(\varepsilon_{\mathbf{i}}+z_{\mathbf{i}0})](\mathbf{r}-\mathbf{r}_0)^{\top}\begin{pmatrix}1\\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h}\end{pmatrix}$$
$$= \frac{A_1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} K\left(\frac{\mathbf{x}_0-\mathbf{x}_{\mathbf{i}}}{h}\right) z_{\mathbf{i}0}(\mathbf{r}-\mathbf{r}_0)^{\top}\begin{pmatrix}1\\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h}\end{pmatrix}(1+o(1)). \tag{4.15}$$

By (4.14) and (4.15), we know that

$$E\left[L_{\mathbf{n}}(\mathbf{r})-L_{\mathbf{n}}(\mathbf{r}_0)+(\Psi_{\mathbf{n}}(\mathbf{r}_0))^{\top}(\mathbf{r}-\mathbf{r}_0)\right]$$
$$= \frac{A_1}{2}\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} K\left(\frac{\mathbf{x}_0-\mathbf{x}_{\mathbf{i}}}{h}\right)\left[(\mathbf{r}-\mathbf{r}_0)^{\top}\begin{pmatrix}1\\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h}\end{pmatrix}\right]^2(1+o(1))$$
$$= \frac{A_1}{2}(\mathbf{r}-\mathbf{r}_0)^{\top}\mathbf{U}(\mathbf{r}-\mathbf{r}_0)(1+o(1)). \tag{4.16}$$

Combining this with (4.13), we know

$$\left|L_{\mathbf{n}}(\mathbf{r})-L_{\mathbf{n}}(\mathbf{r}_0)+(\Psi_{\mathbf{n}}(\mathbf{r}_0))^{\top}(\mathbf{r}-\mathbf{r}_0)-\frac{1}{2}A_1(\mathbf{r}-\mathbf{r}_0)^{\top}\mathbf{U}(\mathbf{r}-\mathbf{r}_0)\right| = o_P(1).$$

As $L_{\mathbf{n}}(\mathbf{r})-L_{\mathbf{n}}(\mathbf{r}_0)+(\Psi_{\mathbf{n}}(\mathbf{r}_0))^{\top}(\mathbf{r}-\mathbf{r}_0)$ is convex in $\mathbf{r}$ and $(\mathbf{r}-\mathbf{r}_0)^{\top}\mathbf{U}(\mathbf{r}-\mathbf{r}_0)$ is continuous and convex in $\mathbf{r}$, by Theorem 10.8 in Rockafellar (1970), we obtain (4.9). □

Lemma 4.3 *Under* (A1)–(A8), *the following asymptotic property holds*

$$\Psi_{\mathbf{n}}(\mathbf{r}_0) = O_P\left((|\mathbf{n}|h^2)^{-1/2}\right). \tag{4.17}$$

Proof By the definition of $\Psi_{\mathbf{n}}(\mathbf{r}_0)$ and $z_{\mathbf{i}0}$, we have

$$\Psi_{\mathbf{n}}(\mathbf{r}_0) = \frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}[\psi(\varepsilon_{\mathbf{i}}+z_{\mathbf{i}0})-\psi(\varepsilon_{\mathbf{i}})]K\left(\frac{\mathbf{x}_0-\mathbf{x}_{\mathbf{i}}}{h}\right)\begin{pmatrix}1\\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h}\end{pmatrix}$$
$$+\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}\psi(\varepsilon_{\mathbf{i}})K\left(\frac{\mathbf{x}_0-\mathbf{x}_{\mathbf{i}}}{h}\right)\begin{pmatrix}1\\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h}\end{pmatrix}$$
$$= \frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}\chi'_{\mathbf{n}\mathbf{i}}K\left(\frac{\mathbf{x}_0-\mathbf{x}_{\mathbf{i}}}{h}\right)\begin{pmatrix}1\\ \frac{\mathbf{x}_{\mathbf{i}}-\mathbf{x}_0}{h}\end{pmatrix}+\Psi_{\mathbf{n}0}. \tag{4.18}$$

By proof similar to that of $\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_\mathbf{n}}(W_{\mathbf{ni}}-EW_{\mathbf{ni}})=o_P(1)$ above, we get

$$\operatorname{Var}\left(\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_\mathbf{n}}\chi'_{\mathbf{ni}}K\left(\frac{\mathbf{x}_0-\mathbf{x_i}}{h}\right)\left(\begin{array}{c}1\\ \frac{\mathbf{x_i}-\mathbf{x}_0}{h}\end{array}\right)\right)=o\left(\frac{1}{|\mathbf{n}|h^2}\right). \tag{4.19}$$

Combing this with (4.1) and (4.18), we know that (4.17) holds. □

Proof of Theorem 3.1 We first prove the weak consistency of the local linear M-estimator. By Lemma 4.2 and Lemma 4.3, we have, for any $\|\mathbf{r}-\mathbf{r}_0\|\leq\delta$,

$$L_\mathbf{n}(\mathbf{r})-L_\mathbf{n}(\mathbf{r}_0)-\frac{A_1}{2}(\mathbf{r}-\mathbf{r}_0)^\top\mathbf{U}(\mathbf{r}-\mathbf{r}_0)=o_P(1).$$

As $\mathbf{U}$ is positive definite, let $\lambda_0>0$ be the smallest eigenvalue of $\mathbf{U}$, then for $\|\mathbf{r}-\mathbf{r}_0\|=\delta$,

$$(\mathbf{r}-\mathbf{r}_0)^\top\mathbf{U}(\mathbf{r}-\mathbf{r}_0)>\frac{1}{2}\lambda_0(\mathbf{r}-\mathbf{r}_0)^\top(\mathbf{r}-\mathbf{r}_0)=\frac{1}{2}\lambda_0\delta^2.$$

Therefore, as $\mathbf{n}\to\infty$,

$$P\left(\inf_{\|\mathbf{r}-\mathbf{r}_0\|=\delta}L_\mathbf{n}(\mathbf{r})-L_\mathbf{n}(\mathbf{r}_0)>\frac{1}{4}A_1\lambda_0\delta^2>0\right)\longrightarrow 1. \tag{4.20}$$

Hence, $L_\mathbf{n}(\mathbf{r})$ has a local minimum in the interior of the sphere $\{\mathbf{r}:\ \|\mathbf{r}-\mathbf{r}_0\|=\delta\}$. So there exists a solution to Eq. (2.2). Let $\widehat{\mathbf{r}}=\left(\begin{array}{c}\widehat{m}_\mathbf{n}(\mathbf{x}_0)\\ h\widehat{m}'_\mathbf{n}(\mathbf{x}_0)\end{array}\right)$ be the closest solution to $\mathbf{r}_0=\left(\begin{array}{c}m(\mathbf{x}_0)\\ hm'(\mathbf{x}_0)\end{array}\right)$, then

$$P\left\{(\widehat{m}_\mathbf{n}(\mathbf{x}_0)-m(\mathbf{x}_0))^2+h^2\left[\left(\widehat{m}'_\mathbf{n}(\mathbf{x}_0)-m'(\mathbf{x}_0)\right)^\top\left(\widehat{m}'_\mathbf{n}(\mathbf{x}_0)-m'(\mathbf{x}_0)\right)\right]<\delta^2\right\}\to 1,$$

which, in turn, implies (3.1).

Next, we will prove (3.2). From the proof of (3.1), we know that in order to prove (3.2), it suffices for us to show

$$\begin{aligned}&\sup_{\|\mathbf{r}-\mathbf{r}_0\|\leq\delta}\left|L_\mathbf{n}(\mathbf{r})-L_\mathbf{n}(\mathbf{r}_0)+(\Psi_\mathbf{n}(\mathbf{r}_0))^\top(\mathbf{r}-\mathbf{r}_0)-\frac{1}{2}A_1(\mathbf{r}-\mathbf{r}_0)^\top\mathbf{U}(\mathbf{r}-\mathbf{r}_0)\right|\\ &\quad=o(1)\quad\text{a.s.}\end{aligned}$$

and

$$\Psi_\mathbf{n}(\mathbf{r}_0)=o(1)\quad\text{a.s.}$$

We only derive the second assertion, noting that the proof of the first one is similar. From (4.18), we know that it's enough to show

$$\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_\mathbf{n}}\chi'_{\mathbf{ni}}K\left(\frac{\mathbf{x}_0-\mathbf{x}_\mathbf{i}}{h}\right)\left(\begin{array}{c}1\\ \frac{\mathbf{x}_\mathbf{i}-\mathbf{x}_0}{h}\end{array}\right)=o(1)\quad\text{a.s.}\tag{4.21}$$

and

$$\Psi_{\mathbf{n}0}=o(1)\quad\text{a.s.}\tag{4.22}$$

We only prove (4.22). Let $\psi_1(\varepsilon_\mathbf{i})=\psi(\varepsilon_\mathbf{i})I(|\psi(\varepsilon_\mathbf{i})|\leq|\mathbf{n}|^\beta)$ and $\bar\psi(\varepsilon_\mathbf{i})=\psi(\varepsilon_\mathbf{i})-\psi_1(\varepsilon_\mathbf{i})$, where β is defined in (A8). Define

$$T_{\mathbf{n}3}^{(1)}=\sum_{\mathbf{i}\in\Lambda_\mathbf{n}}\psi_1(\varepsilon_\mathbf{i})K\left(\frac{\mathbf{x}_0-\mathbf{x}_\mathbf{i}}{h}\right)\left(\begin{array}{c}1\\ \frac{\mathbf{x}_\mathbf{i}-\mathbf{x}_0}{h}\end{array}\right)$$

and

$$T_{\mathbf{n}3}^{(2)}=\sum_{\mathbf{i}\in\Lambda_\mathbf{n}}\bar\psi(\varepsilon_\mathbf{i})K\left(\frac{\mathbf{x}_0-\mathbf{x}_\mathbf{i}}{h}\right)\left(\begin{array}{c}1\\ \frac{\mathbf{x}_\mathbf{i}-\mathbf{x}_0}{h}\end{array}\right).$$

Then,

$$\Psi_{\mathbf{n}0}=T_{\mathbf{n}3}^{(1)}+T_{\mathbf{n}3}^{(2)}.\tag{4.23}$$

By (A6) and Markov inequality, we know that for $|x_{\mathbf{i}l}-x_{0l}|\leq h$, $l=1,2$, and arbitrary $\mathbf{c}\in\mathbf{R}^3$, $\mathbf{c}\neq\mathbf{0}$,

$$P\left(\left|\mathbf{c}^\top T_{\mathbf{n}3}^{(2)}\right|/(|\mathbf{n}|h^2)\neq 0\right)\leq\sum_{\mathbf{i}\in\Lambda_\mathbf{n}}P\left(|\psi(\varepsilon_\mathbf{i})|>|\mathbf{n}|^\beta\right)\leq C|\mathbf{n}|^{1-\beta(2+\gamma)}.$$

Hence, by $\beta>\frac{2}{2+\gamma}$,

$$\sum_{n_1,n_2\geq 1}P\left(\left|\mathbf{c}^\top T_{\mathbf{n}3}^{(2)}\right|/(|\mathbf{n}|h^2)\neq 0\right)\leq C\sum_{n_1,n_2\geq 1}|\mathbf{n}|^{1-\beta(2+\gamma)}<\infty.$$

So

$$T_{\mathbf{n}3}^{(2)}=o\left(|\mathbf{n}|h^2\right)\quad\text{a.s.}\tag{4.24}$$

Let $\eta_\mathbf{i}=\frac{1}{h^2}[\psi_1(\varepsilon_\mathbf{i})-E\psi_1(\varepsilon_\mathbf{i})]K\left(\frac{\mathbf{x}_0-\mathbf{x}_\mathbf{i}}{h}\right)\mathbf{c}^\top\left(\begin{array}{c}1\\ \frac{\mathbf{x}_\mathbf{i}-\mathbf{x}_0}{h}\end{array}\right)=\frac{1}{h^2}b_\mathbf{i}[\psi_1(\varepsilon_\mathbf{i})-E\psi_1(\varepsilon_\mathbf{i})]$, where $b_\mathbf{i}$ is defined in Lemma 4.1, then $\max_{\mathbf{i}\in\Lambda_\mathbf{n}}|\eta_\mathbf{i}|\leq C|\mathbf{n}|^\beta h^{-2}=:K_\mathbf{n}$. By applying

Lemma A in the Appendix to $\eta_{\mathbf{i}}$, we have, for any $\epsilon > 0$,

$$\begin{aligned} & P\left(\frac{1}{h^2}\left|\mathbf{c}^\top\left(T_{\mathbf{n}3}^{(1)} - ET_{\mathbf{n}3}^{(1)}\right)\right| \geq |\mathbf{n}|\epsilon\right) \\ & = P\left(\left|\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} \eta_{\mathbf{i}}\right| \geq |\mathbf{n}|\epsilon\right) \\ & \leq 2^2\left\{2\exp\left(-\frac{\epsilon^2|\mathbf{q}|}{2^3 v^2(\mathbf{q})}\right) + \frac{4K_{\mathbf{n}}}{\epsilon}\varphi\left(\left[\min_{k=1,2} p_k\right]\right)\right\}, \end{aligned} \tag{4.25}$$

where $\mathbf{q}$ is defined in (A9) and

$$v^2(|\mathbf{q}|) = 2^7(|\mathbf{q}|/|\mathbf{n}|)^2 \max_{\mathbf{i},\mathbf{j}} E\left(\sum_{\mathbf{k}\in A_{\mathbf{i},\mathbf{j}}} \eta_{\mathbf{k}}\right)^2 + K_{\mathbf{n}}\epsilon$$

with $A_{\mathbf{i},\mathbf{j}} = \prod_{k=1}^2 ((i_k + 2j_k)p_k,\ (i_k + 2j_k + 1)p_k]$ and $i_k = 0, 1$, $j_k = 0, \ldots, q_k - 1$, $k = 1, 2$. As $\max_{\mathbf{k}\in\mathbf{Z}^2} |b_{\mathbf{k}}| \leq C$ and $\sum_{\mathbf{k}\in\mathbf{Z}^2} |\mathrm{Cov}(\psi_1(\varepsilon_{\mathbf{0}}), \psi_1(\varepsilon_{\mathbf{k}}))| < \infty$, we have

$$\begin{aligned} \max_{\mathbf{i},\mathbf{j}} E\left(\sum_{\mathbf{k}\in A_{\mathbf{i},\mathbf{j}}} \eta_{\mathbf{k}}\right)^2 & \leq (C/h^4)\max_{\mathbf{i},\mathbf{j}} E\left(\sum_{\mathbf{k}\in A_{\mathbf{i},\mathbf{j}}} [\psi_1(\varepsilon_{\mathbf{k}}) - E\psi_1(\varepsilon_{\mathbf{k}})]\right)^2 \\ & = (C/h^4)E\left(\sum_{\mathbf{k}\in A_{\mathbf{0},\mathbf{0}}} [\psi_1(\varepsilon_{\mathbf{k}}) - E\psi_1(\varepsilon_{\mathbf{k}})]\right)^2 \\ & \leq C|\mathbf{p}|/h^4. \end{aligned}$$

Hence,

$$v^2(|\mathbf{q}|) \leq C|\mathbf{p}||\mathbf{q}|^2/(|\mathbf{n}|^2h^4) + K_{\mathbf{n}}\epsilon \leq C|\mathbf{q}|/(|\mathbf{n}|h^4) + K_{\mathbf{n}}\epsilon \leq CK_{\mathbf{n}}.$$

The last inequality above holds because $|\mathbf{n}|^{1+\beta}h^2/|\mathbf{q}| \to \infty$. By (4.25), we get

$$P\left(\frac{1}{h^2}\left|\mathbf{c}^\top\left(T_{\mathbf{n}3}^{(1)} - ET_{\mathbf{n}3}^{(1)}\right)\right| \geq |\mathbf{n}|\epsilon\right) \leq C\left\{\exp(-C|\mathbf{q}|/K_{\mathbf{n}})\right\} + CK_{\mathbf{n}}\varphi\left(\min_{1\leq k\leq 2} p_k\right). \tag{4.26}$$

By $(|\mathbf{q}|h^2)/(|\mathbf{n}|^\beta \log|\mathbf{n}|) \to \infty$ and $\sum_{n_1,n_2\geq 1}(|\mathbf{n}|^\beta/h^2)\varphi(\min_{1\leq k\leq 2}\frac{n_k}{q_k}) < \infty$, we know that

$$|\mathbf{q}|/(K_{\mathbf{n}}\log|\mathbf{n}|) \to \infty \quad \text{and} \quad \sum_{n_1,n_2\geq 1} K_{\mathbf{n}}\varphi\left(\min_{1\leq k\leq 2} p_k\right) < \infty.$$

These combined with (4.26) result in

$$\sum_{n_1,n_2\geq 1} P\left(\frac{1}{h^2}\left|\mathbf{c}^\top\left(T_{\mathbf{n}3}^{(1)}-ET_{\mathbf{n}3}^{(1)}\right)\right|\geq |\mathbf{n}|\epsilon\right)<\infty,$$

which implies

$$T_{\mathbf{n}3}^{(1)}=o(|\mathbf{n}|h^2)\quad \text{a.s.} \tag{4.27}$$

In view of (4.23), (4.24) and (4.27), (4.22) is proved. □

Before giving the proof of Theorem 3.2, we present the following lemma.

Lemma 4.4 *Under* (A1)–(A8), $\Psi_{\mathbf{n}}(\mathbf{r}_0)$ *has the following asymptotic distribution*

$$\sqrt{|\mathbf{n}|h^2}\left(\Psi_{\mathbf{n}}(\mathbf{r}_0)-\frac{A_1}{2}h^2\mathbf{M}\right)\xrightarrow{d} N\left(0,\ \mathbf{V}\sum_{\mathbf{i}\in\mathbf{Z}^2}E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{i}})\right). \tag{4.28}$$

Proof When deriving the asymptotic normality of the Nadaraya–Watson estimator of $m(\cdot)$ in model (1.1) in the case of $\mathbf{n}=(n,\ldots,n)^\top$, Machkouri (2005) adapted a method that was used in the proof of the central limit theorem in Dedecker (1998), which was based on Lindeberg's decomposition method. In this paper, we will employ a different method which has been used by Peligrad and Utev (1997) to prove the central limit theorem for partial sums of weighted mixing random variables.

Note that

$$\begin{aligned}
E\Psi_{\mathbf{n}}(\mathbf{r}_0)&=\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}A_1 z_{\mathbf{i}0}K\left(\frac{\mathbf{x}_0-\mathbf{x}_\mathbf{i}}{h}\right)\left(\frac{1}{\frac{\mathbf{x}_\mathbf{i}-\mathbf{x}_0}{h}}\right)(1+o(1))\\
&=\frac{A_1}{2}\frac{m''(\mathbf{x}_0)}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}(\mathbf{x}_\mathbf{i}-\mathbf{x}_0)^2K\left(\frac{\mathbf{x}_0-\mathbf{x}_\mathbf{i}}{h}\right)\left(\frac{1}{\frac{\mathbf{x}_\mathbf{i}-\mathbf{x}_0}{h}}\right)(1+o(1))\\
&=\frac{A_1}{2}h^2\mathbf{M}(1+o(1))
\end{aligned}$$

and

$$\begin{aligned}
&\Psi_{\mathbf{n}}(\mathbf{r}_0)-\Psi_{\mathbf{n}0}\\
&\quad=\frac{1}{|\mathbf{n}|h^2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}\left[\psi(\varepsilon_\mathbf{i}+z_{\mathbf{i}0})-\psi(\varepsilon_\mathbf{i})\right]K\left(\frac{\mathbf{x}_0-\mathbf{x}_\mathbf{i}}{h}\right)\left(\frac{1}{\frac{\mathbf{x}_\mathbf{i}-\mathbf{x}_0}{h}}\right)\\
&\quad=o_P\left((|\mathbf{n}|h^2)^{-\frac{1}{2}}\right).
\end{aligned}$$

So in order to prove Lemma 4.4, it suffices for us to prove the asymptotic normality of $(|\mathbf{n}|h^2)^{\frac{1}{2}}\Psi_{\mathbf{n}0}$. Remind that $\mathrm{Var}(\Psi_{\mathbf{n}0})=\frac{1}{|\mathbf{n}|h^2}[\sum_{\mathbf{i}\in\mathbf{Z}^2}E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_\mathbf{i})]\mathbf{V}(1+o(1))$. By the

Cramér–Wold device, it suffices to prove the asymptotic normality of $(|\mathbf{n}|h^2)^{\frac{1}{2}}\mathbf{c}^\top\Psi_{\mathbf{n}0}$ for any $\mathbf{c}\in\mathbf{R}^3$, $\mathbf{c}\neq\mathbf{0}$.

Let $c_{\mathbf{ni}}=(|\mathbf{n}|h^2)^{-\frac{1}{2}}K(\frac{\mathbf{x}_0-\mathbf{x}_\mathbf{i}}{h})\mathbf{c}^\top\begin{pmatrix}1\\ \frac{\mathbf{x}_\mathbf{i}-\mathbf{x}_0}{h}\end{pmatrix}$, then $(|\mathbf{n}|h^2)^{\frac{1}{2}}\mathbf{c}^\top\Psi_{\mathbf{n}0}=\sum_{\mathbf{i}\in\Lambda_\mathbf{n}}c_{\mathbf{ni}}\psi(\varepsilon_\mathbf{i})$. We now adopt the method in the proof of Theorem 2.2 in Peligrad and Utev (1997). By (A6), (A7) and the stationarity, we have

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{j_k=a_k,k=1,2}^{b_k}c_{\mathbf{nj}}\psi(\varepsilon_\mathbf{j})\right)\\
&=\sum_{j_k=a_k,k=1,2}^{b_k}c_{\mathbf{nj}}^2E\psi^2(\varepsilon_\mathbf{j})+\sum_{i_k=a_k,k=1,2}^{b_k}\ \sum_{j_k=a_k,k=1,2,\mathbf{i}\neq\mathbf{j}}^{b_k}c_{\mathbf{ni}}c_{\mathbf{nj}}E\psi(\varepsilon_\mathbf{i})\psi(\varepsilon_\mathbf{j})\\
&\leq\sum_{j_k=a_k,k=1,2}^{b_k}c_{\mathbf{nj}}^2E\psi^2(\varepsilon_\mathbf{0})+2\sum_{i_k=a_k,k=1,2}^{b_k}c_{\mathbf{ni}}^2\sum_{j_k=a_k,k=1,2,\mathbf{i}\neq\mathbf{j}}^{b_k}|E\psi(\varepsilon_\mathbf{i})\psi(\varepsilon_\mathbf{j})|\\
&\leq\sum_{j_k=a_k,k=1,2}^{b_k}c_{\mathbf{nj}}^2E\psi^2(\varepsilon_\mathbf{0})+C\sum_{i_k=a_k,k=1,2}^{b_k}c_{\mathbf{ni}}^2\left[\sum_{k=1}^{\infty}k\varphi^{\frac{\gamma}{2+\gamma}}(k)\right]\|\psi(\varepsilon_\mathbf{0})\|_{2+\gamma}^2\\
&\leq\left[E\psi^2(\varepsilon_\mathbf{0})+C\|\psi(\varepsilon_\mathbf{0})\|_{2+\gamma}^2\sum_{k=1}^{\infty}k\varphi^{\frac{\gamma}{2+\gamma}}(k)\right]\sum_{j_k=a_k,k=1,2}^{b_k}c_{\mathbf{nj}}^2\\
&\leq C\sum_{j_k=a_k,k=1,2}^{b_k}c_{\mathbf{nj}}^2.
\end{aligned}\tag{4.29}$$

Truncate the variable $\psi(\varepsilon_\mathbf{i})$ at the level $A>0$ and denote

$$\zeta_\mathbf{i}=\psi(\varepsilon_\mathbf{i})I(|\psi(\varepsilon_\mathbf{i})|\leq A)-E\left[\psi(\varepsilon_\mathbf{i})I(|\psi(\varepsilon_\mathbf{i})|\leq A)\right]$$

and

$$\zeta_\mathbf{i}'=\psi(\varepsilon_\mathbf{i})I(|\psi(\varepsilon_\mathbf{i})|>A)-E[\psi(\varepsilon_\mathbf{i})I(|\psi(\varepsilon_\mathbf{i})|>A)].$$

By the proof of (4.29), we know that

$$\mathrm{Var}\left(\sum_{\mathbf{i}\in\Lambda_\mathbf{n}}c_{\mathbf{ni}}\zeta_\mathbf{i}'\right)\leq\sum_{\mathbf{i}\in\Lambda_\mathbf{n}}c_{\mathbf{ni}}^2\left(E\left[\psi^2(\varepsilon_\mathbf{0})I(|\psi(\varepsilon_\mathbf{0})|>A)\right]+C\|\zeta_\mathbf{0}'\|_{2+\gamma}^2\right),$$

which tends to zero as first $\mathbf{n}\to\infty$ then $A\to\infty$. As $\psi(\varepsilon_\mathbf{i})=\zeta_\mathbf{i}+\zeta_\mathbf{i}'$ and

$$\mathrm{Var}\left(\sum_{\mathbf{i}\in\Lambda_\mathbf{n}}c_{\mathbf{ni}}\psi(\varepsilon_\mathbf{i})\right)\to\left[\sum_{\mathbf{i}\in Z^2}E\psi(\varepsilon_\mathbf{0})\psi(\varepsilon_\mathbf{i})\right]\mathbf{c}^\top\mathbf{V}\mathbf{c}\quad\text{as }\mathbf{n}\to\infty,$$

we have

$$\operatorname{Var}\left(\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} c_{\mathbf{ni}}\zeta_{\mathbf{i}}\right) \to \left[\sum_{\mathbf{i}\in Z^2} E\psi(\varepsilon_{\mathbf{0}})\psi(\varepsilon_{\mathbf{i}})\right]\mathbf{c}^{\top}\mathbf{V}\mathbf{c} =: \sigma^2 > 0 \tag{4.30}$$

as first $\mathbf{n}\to\infty$ then $A\to\infty$. Hence, it remains for us to prove

$$\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} c_{\mathbf{ni}}\zeta_{\mathbf{i}} \xrightarrow{d} N\left(0,\sigma^2\right)$$

as first $\mathbf{n}\to\infty$ then $A\to\infty$. To this end, we apply Lemma B in the Appendix. Notice that by a computation similar to (4.29), (a) holds for $\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} c_{\mathbf{ni}}\zeta_{\mathbf{i}}$. Since $E\psi^2(\varepsilon_{\mathbf{0}}) > 0$, $E\zeta_{\mathbf{0}}^2 > 0$ when A is large enough. Besides, $\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}}\operatorname{Var}(c_{\mathbf{ni}}\zeta_{\mathbf{i}}) = \mathbf{c}^{\top}\mathbf{V}\mathbf{c}E\zeta_{\mathbf{0}}^2(1+o(1))$ as $\mathbf{n}\to\infty$. This, combined with (4.30), implies that (b) holds. As $|c_{\mathbf{ni}}\zeta_{\mathbf{i}}| \le CA(|\mathbf{n}|h^2)^{-\frac{1}{2}} \to 0$ when $\mathbf{n}\to\infty$, we have $\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} c_{\mathbf{ni}}^2 E\zeta_{\mathbf{i}}^2 I(|c_{\mathbf{ni}}\zeta_{\mathbf{i}}| > \epsilon\sigma) \to 0$ for any $\epsilon > 0$, i.e. (d) is valid. Next, we will verify (c). Notice that

$$\begin{aligned}
&\left|\operatorname{Cov}\left(\exp\left\{it\sum_{i_k=a_k,k=1,2}^{b_k} c_{\mathbf{ni}}\zeta_{\mathbf{i}}\right\},\ \exp\left\{it\sum_{j_k=b_k+u_k,k=1,2}^{c_k} c_{\mathbf{nj}}\zeta_{\mathbf{j}}\right\}\right)\right| \\
&\quad = \left|\operatorname{Cov}\left(\left[\exp\left\{it\sum_{i_k=a_k,k=1,2}^{b_k} c_{\mathbf{ni}}\zeta_{\mathbf{i}}\right\}-1\right],\ \left[\exp\left\{it\sum_{j_k=b_k+u_k,k=1,2}^{c_k} c_{\mathbf{nj}}\zeta_{\mathbf{j}}\right\}-1\right]\right)\right| \\
&\quad \le Ct^2\left[\varphi^{\frac{\gamma}{2+\gamma}}(\|\mathbf{u}\|)\right]\left\|\sum_{i_k=a_k,k=1,2}^{b_k} c_{\mathbf{ni}}\zeta_{\mathbf{i}}\right\|_{2+\gamma}\left\|\sum_{j_k=b_k+u_k,k=1,2}^{c_k} c_{\mathbf{nj}}\zeta_{\mathbf{j}}\right\|_{2+\gamma}
\end{aligned} \tag{4.31}$$

and

$$\begin{aligned}
&E\left|\sum_{i_k=a_k,k=1,2}^{b_k} c_{\mathbf{ni}}\zeta_{\mathbf{i}}\right|^{2+\gamma} \\
&\quad \le (2A)^{\gamma}E\left(\sum_{i_k=a_k,k=1,2}^{b_k} c_{\mathbf{ni}}\zeta_{\mathbf{i}}\right)^2\left(\sum_{i_k=a_k,k=1,2}^{b_k} |c_{\mathbf{ni}}|\right)^{\gamma} \\
&\quad \le C(2A)^{\gamma}\left(\sum_{i_k=a_k,k=1,2}^{b_k} c_{\mathbf{ni}}^2\right)\left(\sum_{i_k=a_k,k=1,2}^{b_k} c_{\mathbf{ni}}^2\right)^{\gamma/2}\left[\prod_{k=1}^{2}(b_k-a_k+1)\right]^{\gamma/2} \\
&\quad = C(2A)^{\gamma}\left(\sum_{i_k=a_k,k=1,2}^{b_k} c_{\mathbf{ni}}^2\right)^{1+\frac{\gamma}{2}}\left[\prod_{k=1}^{2}(b_k-a_k+1)\right]^{\gamma/2}.
\end{aligned} \tag{4.32}$$

As $|b_k - a_k + 1| \le |c_k - a_k|$, $|c_k - b_k - u_k + 1| \le |c_k - a_k|$ and $u_k = (c_k - a_k)^{1-\iota}$, where ι will be specified later, we have, by (4.31) and (4.32),

$$\begin{aligned}
&\left|\mathrm{Cov}\left(\exp\left\{it\sum_{i_k=a_k,k=1,2}^{b_k} c_{\mathbf{ni}}\zeta_{\mathbf{i}}\right\}, \exp\left\{it\sum_{j_k=b_k+u_k,k=1,2}^{c_k} c_{\mathbf{nj}}\zeta_{\mathbf{j}}\right\}\right)\right| \\
&\le Ct^2\left[\varphi^{\frac{\gamma}{2+\gamma}}(\|\mathbf{u}\|)\right](2A)^{\frac{2\gamma}{2+\gamma}}\left(\sum_{i_k=a_k,k=1,2}^{c_k} c_{\mathbf{ni}}^2\right)\left[\prod_{k=1}^{2}(c_k - a_k)\right]^{\frac{\gamma}{2+\gamma}} \\
&= Ct^2\,(2A)^{\frac{2\gamma}{2+\gamma}}\left(\sum_{i_k=a_k,k=1,2}^{c_k} c_{\mathbf{ni}}^2\right)\left[\varphi(\|\mathbf{u}\|)\prod_{k=1}^{2} u_k^{\frac{1}{1-\iota}}\right]^{\gamma/(2+\gamma)} \\
&\le Ct^2\,(2A)^{\frac{2\gamma}{2+\gamma}}\left(\sum_{i_k=a_k,k=1,2}^{c_k} c_{\mathbf{ni}}^2\right)\left[\varphi(\|\mathbf{u}\|)\,\|\mathbf{u}\|^{\frac{2}{1-\iota}}\right]^{\gamma/(2+\gamma)}.
\end{aligned}$$

By (A7), we know that $\varphi(n) = o(n^{-2(2+\gamma)/\gamma})$. If we choose $\iota = \frac{1}{2+\gamma}$, then $\frac{1}{1-\iota} < \frac{2+\gamma}{\gamma}$. Therefore,

$$\sum_{i=1}^{\infty}\left[(2^i)^{\frac{2}{1-\iota}}\varphi(2^i)\right]^{\gamma/(2+\gamma)} \le C\sum_{i=1}^{\infty}\left[2^{-\frac{2+\gamma}{\gamma}+\frac{1}{1-\iota}}\right]^{2i\gamma/(2+\gamma)} < \infty.$$

Therefore, (c) holds. Thus we have proved Lemma 4.4. $\square$

Proof of Theorem 3.2 By proof similar to that of Lemma 4.2, we have, for any constant $c > 0$,

$$\begin{aligned}
\sup_{\sqrt{|\mathbf{n}|h^2}\|\mathbf{r}-\mathbf{r}_0\|\le c}\Big||\mathbf{n}|h^2&\left[L_{\mathbf{n}}(\mathbf{r}) - L_{\mathbf{n}}(\mathbf{r}_0) + (\Psi_{\mathbf{n}}(\mathbf{r}_0))^\top(\mathbf{r}-\mathbf{r}_0)\right] \\
&-\frac{1}{2}|\mathbf{n}|h^2 A_1(\mathbf{r}-\mathbf{r}_0)^\top\mathbf{U}(\mathbf{r}-\mathbf{r}_0)\Big| = o_P(1).
\end{aligned} \tag{4.33}$$

Next, we will prove

$$\sqrt{|\mathbf{n}|h^2}\|\widehat{\mathbf{r}} - \mathbf{r}_0\| = O_P(1).$$

It suffices for us to show that for any positive sequence $\{c_{\mathbf{n}}\}$ satisfying $c_{\mathbf{n}} \to \infty$, we have

$$P\left(\sqrt{|\mathbf{n}|h^2}\|\widehat{\mathbf{r}} - \mathbf{r}_0\| \ge c_{\mathbf{n}}\right) \to 0. \tag{4.34}$$

By (4.33), we know that there exists a sequence of positive constants $\{c'_{\mathbf{n}}\}$ such that $c'_{\mathbf{n}} \to \infty$, $c'_{\mathbf{n}} \le c_{\mathbf{n}}$ and

$$\sup_{\sqrt{|\mathbf{n}|h^2}\|\mathbf{r}-\mathbf{r}_0\|\le c'_{\mathbf{n}}}\Bigg||\mathbf{n}|h^2\left[L_{\mathbf{n}}(\mathbf{r})-L_{\mathbf{n}}(\mathbf{r}_0)+(\Psi_{\mathbf{n}}(\mathbf{r}_0))^\top(\mathbf{r}-\mathbf{r}_0)\right]$$
$$-\frac{1}{2}|\mathbf{n}|h^2A_1(\mathbf{r}-\mathbf{r}_0)^\top\mathbf{U}(\mathbf{r}-\mathbf{r}_0)\Bigg|=o_P(1). \tag{4.35}$$

As $\mathbf{U}$ is positive definite, let $\lambda_0>0$ be the smallest eigenvalue of $\mathbf{U}$. Then when $\sqrt{|\mathbf{n}|h^2}\|\mathbf{r}-\mathbf{r}_0\|=c'_{\mathbf{n}}$, we have

$$\frac{1}{2}|\mathbf{n}|h^2A_1(\mathbf{r}-\mathbf{r}_0)^\top\mathbf{U}(\mathbf{r}-\mathbf{r}_0)>\frac{1}{4}A_1\lambda_0c'^2_{\mathbf{n}}. \tag{4.36}$$

By Lemma 4.3, we know that

$$|\mathbf{n}|h^2\Psi_{\mathbf{n}}(\mathbf{r}_0)=O_P\left(\sqrt{|\mathbf{n}|h^2}\right).$$

Hence, when $\sqrt{|\mathbf{n}|h^2}\|\mathbf{r}-\mathbf{r}_0\|=c'_{\mathbf{n}}$, we have

$$\left||\mathbf{n}|h^2\left(\Psi_{\mathbf{n}}(\mathbf{r}_0)\right)^\top(\mathbf{r}-\mathbf{r}_0)\right|=O_P(c'_{\mathbf{n}}). \tag{4.37}$$

As $c'_{\mathbf{n}}\to\infty$, by (4.35)–(4.37), we know that

$$P\left(\inf_{\sqrt{|\mathbf{n}|h^2}\|\mathbf{r}-\mathbf{r}_0\|=c'_{\mathbf{n}}}|\mathbf{n}|h^2(L_{\mathbf{n}}(\mathbf{r})-L_{\mathbf{n}}(\mathbf{r}_0))<0\right)\to 0.$$

As $\rho(\cdot)$ is a convex function, we have

$$P\left(\inf_{\sqrt{|\mathbf{n}|h^2}\|\mathbf{r}-\mathbf{r}_0\|\ge c'_{\mathbf{n}}}|\mathbf{n}|h^2L_{\mathbf{n}}(\mathbf{r})\le|\mathbf{n}|h^2L_{\mathbf{n}}(\mathbf{r}_0)\right)\to 0.$$

By the definition of $\widehat{\mathbf{r}}$, we know that

$$P\left(\sqrt{|\mathbf{n}|h^2}\|\widehat{\mathbf{r}}-\mathbf{r}_0\|\ge c'_{\mathbf{n}}\right)\to 0,$$

which implies the validity of (4.34) by $c'_{\mathbf{n}}\le c_{\mathbf{n}}$.

On the other hand, by Theorem 2.5.7 in Rockafellar (1970) and arguments similar to the proof of Lemma 4.2, we have, for any $c>0$,

$$\sup_{\sqrt{|\mathbf{n}|h^2}\|\mathbf{r}-\mathbf{r}_0\|\le c}\left|\left(|\mathbf{n}|h^2\right)^{1/2}[\Psi_{\mathbf{n}}(\mathbf{r})-\Psi_{\mathbf{n}}(\mathbf{r}_0)]+\left(|\mathbf{n}|h^2\right)^{1/2}A_1\mathbf{U}(\mathbf{r}-\mathbf{r}_0)\right|=o_P(1). \tag{4.38}$$

By $\Psi_{\mathbf{n}}(\widehat{\mathbf{r}}) = 0$, (4.38) and (4.34), we have

$$\widehat{\mathbf{r}} - \mathbf{r}_0 = A_1^{-1}\mathbf{U}^{-1}\Psi_{\mathbf{n}}(\mathbf{r}_0) + o_P\left((|\mathbf{n}|h^2)^{-1/2}\right). \tag{4.39}$$

By (4.39) and Lemma 4.4, we know that Theorem 3.2 holds. □

Appendix

Lemma A below gives a exponential inequality for strongly mixing random fields. It can be found in Lee et al. (2004).

Lemma A *Let* $\{Z_{\mathbf{i}}, \mathbf{i} \in Z^d\}$ *be a zero-mean real valued and strongly mixing random field such that* $\sup_{\mathbf{i}\in\Lambda_{\mathbf{n}}} \|Z_{\mathbf{i}}\|_\infty \le K < \infty$. *Then for each* $\mathbf{q} = (q_1, \ldots, q_d)$ *with integer-valued coordinates* $q_k \in [1, n_k/2]$ *and for each* $\epsilon > 0$ *we have*

$$P\left(\left|\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} Z_{\mathbf{i}}\right| > |\mathbf{n}|\epsilon\right) \le 2^d \left\{2\exp\left(-\frac{\epsilon^2|\mathbf{q}|}{2^{d+1}v^2(\mathbf{q})}\right) + \frac{4K}{\epsilon}\varphi\left(\left[\min_{1\le k\le d} p_k\right]\right)\right\},$$

where $p_k = n_k/(2q_k)$, $v^2(\mathbf{q}) = 2^{d+1}\sigma^2(\mathbf{q})/|\mathbf{p}|^2 + K\epsilon$, $\sigma^2(\mathbf{q}) = \max_{\mathbf{i},\mathbf{j}} E\left(\sum_{k\in A_{\mathbf{i},\mathbf{j}}} Z_{\mathbf{k}}\right)^2$ *and* $A_{\mathbf{i},\mathbf{j}} = \prod_{k=1}^d ((i_k + 2j_k)p_k, (i_k + 2j_k + 1)p_k]$. *The maximization in the defining equation for* $\sigma^2(\mathbf{q})$ *is taken over all pairs of d-tuple indices* $\mathbf{i}$ *and* $\mathbf{j}$ *with* $i_k = 0, 1$ *and* $j_k = 0, 1, \ldots, q_k - 1$.

The next lemma is an extension of Theorem B in Peligrad and Utev (1997) to random fields. The proof is analogous to that of Theorem 4.1 in Utev (1990).

Lemma B *Let* $\{X_{\mathbf{ni}}, \mathbf{i} \in \Lambda_{\mathbf{n}}\}$ *be an array of random variables such that the following hold.*

(a) $\mathrm{Var}(\sum_{j_k=a_k,1\le k\le d}^{b_k} X_{\mathbf{nj}}) \le C\sum_{j_k=a_k,1\le k\le d}^{b_k} \mathrm{Var}(X_{\mathbf{nj}})$ *for every* $0 \le a_k \le b_k \le n_k$, $1 \le k \le d$;

(b) $\liminf_{\mathbf{n}\to\infty} \frac{\mathrm{Var}\left(\sum_{\mathbf{j}\in\Lambda_{\mathbf{n}}} X_{\mathbf{nj}}\right)}{\sum_{\mathbf{j}\in\Lambda_{\mathbf{n}}} \mathrm{Var}(X_{\mathbf{nj}})} > 0$;

(c) $|\mathrm{Cov}(\exp(it\sum_{j_k=a_k,1\le k\le d}^{b_k} X_{\mathbf{nj}}), \exp(it\sum_{j_k=b_k+u_k,1\le k\le d}^{c_k} X_{\mathbf{nj}}))| \le h_t(\|\mathbf{u}\|) \sum_{j_k=a_k,1\le k\le d}^{c_k} \mathrm{Var}(X_{\mathbf{nj}})$ *for every* $1 \le a_k \le b_k \le c_k \le n_k$, *where* $h_t(\|\mathbf{u}\|) \ge 0$, $\sum h_t(2^i) < \infty$ *and* $\mathbf{u}$ *is of the form* $u_k = [(c_k - a_k)^{1-\iota}]$ *for a certain* $0 < \iota < 1$;

(d) $\sigma_{\mathbf{n}}^{-2}\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} EX_{\mathbf{ni}}^2 I(|X_{\mathbf{ni}}| > \epsilon\sigma_{\mathbf{n}}) \to 0$ *as* $\mathbf{n} \to \infty$ *for every* $\epsilon > 0$, *where* $\sigma_{\mathbf{n}}^2$ *denotes* $\mathrm{Var}(\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} X_{\mathbf{ni}})$.

Then $\sum_{\mathbf{i}\in\Lambda_{\mathbf{n}}} X_{\mathbf{ni}}/\sigma_{\mathbf{n}} \xrightarrow{d} N(0, 1)$ *as* $\mathbf{n} \to \infty$.

Acknowledgments The authors are grateful to the editor and two anonymous referees for their valuable comments, which help to improve the presentation of a former version of this paper. We also would like to thank Dr. Degui Li for his constructive suggestions. This work was supported by the National Natural Science Foundation of China (Grant no. 10771192).

References

Bai ZD, Rao CR, Wu Y (1992) M-estimation of multivariate linear regression parameters under a convex discrepancy function. Stat Sin 2:237–254

Bosq D (1993) Bernstein-type large deviation inequalities for partial sums of strongly mixing processes. Statistics 24:59–70

Cai ZW, Ould-Saïd E (2003) Local M-estimator for nonparametric time series. Statist Probab Lett 65: 433–449

Carbon M, Tran LT, Wu B (1997) Kernel density estimation for random fields. Statist Probab Lett 36: 115–125

Dedecker J (1998) A central limit theorem for stationary random fields. Probab Theory Relat Fields 110:397–426

Fan JQ (1992) Design-adaptive nonparametric regression. J Am Statist Assoc 87:998–1004

Fan JQ (1993) Local linear regression smoothers and their minimax efficiency. Ann Statist 21:196–216

Fan JQ, Gijbels I (1996) Local polynomial modeling and its applications. Chapman and Hall, London

Fan JQ, Jiang J (1999) Variable bandwidth and one-step local M-estimator. Sci China Ser A 29:1–15

Hall P, Hart J (1990) Nonparametric regression with long-range dependence. Stoch Proc Appl 36:339–351

Hallin M, Lu ZD, Tran LT (2001) Density estimation for spatial linear processes. Bernoulli 7:657–668

Hallin M, Lu ZD, Tran LT (2004a) Density estimation for spatial processes: the L_1 theory. J Multivar Anal 88:61–75

Hallin M, Lu ZD, Tran LT (2004b) Local linear spatial regression. Ann Statist 32:2469–2500

Hallin M, Lu ZD, Yu K (2005) Local linear spatial quantile regression. Technical report available at www.stat.ucl.ac.be/ISpub/tr/2005/TR0552.pdf

Ibragimov IA, Linnik YuV (1971) Independent and stationary sequences of random variables. Wolters, Groningen

Jiang J, Mack YP (2001) Robust local polynomial regression for dependent data. Stat Sin 11:705–722

Lee YK, Choi H, Park BU, Yu KS (2004) Local likelihood density estimation on random fields. Statist Probab Lett 68:347–357

Lu ZD (1998) On geometric ergodicity of a non-linear autoregressive (AR) model with an autoregressive conditional heteroscedastic (ARCH) term. Stat Sin 8:1205–1217

Lu ZD, Chen X (2004) Spatial kernel regression estimation: weak consistency. Statist Probab Lett 68: 125–136

Machkouri MEl (2005) Asymptotic normality of kernel estimates in a regression model for random fields. Technical report available at http://math.univ-lille1.fr/~elmachko/PagepersoLille_fichiers/Asymptotic_normality.pdf

Machkouri MEl (2007) Nonparametric regression estimation for random fields in a fixed-design. Statist Infer Stochast Process 10:29–47

Peligrad M, Utev SA (1997) Central limit theorems for linear processes. Ann Probab 25:443–456

Peng L, Yao QW (2004) Nonparametric regression under dependent errors with infinite variance. Ann Inst Statist Math 56:73–86

Pham TD (1986) The mixing properties of bilinear and generalized random coefficient autoregressive models. Stoch Proc Appl 23:291–300

Pham TD, Tran LT (1985) Some strong mixing properties of time series models. Stoch Proc Appl 19: 297–303

Rockafellar RT (1970) Convex analysis. Princeton University Press, Princeton

Tran LT (1990) Kernel density estimation on random fields. J Multivar Anal 34:37–53

Utev SA (1990) Central limit theorem for dependent random variables. Probab Theory Math Statist 2:519–528

Withers CS (1981) Conditions for linear processes to be strong mixing. Z Wahrsch Verw Gebiete 57: 477–480