Asymptotic normality for deconvolution kernel density estimators from random fields

Jiexiang Li

Department of Mathematics, College of Charleston, U.S.A.

Abstract

The paper discusses the estimation of a continuous density function of the target random field \( X_i, \ i \in \mathbb{Z}^N \) which is contaminated by measurement errors. In particular, the observed random field \( Y_i, \ i \in \mathbb{Z}^N \) is such that \( Y_i = X_i + \epsilon_i \), where the random error \( \epsilon_i \) is from a known distribution and independent of the target random field. Compared to the existing results, the paper is improved in two directions. First, the random vectors in contrast to univariate random variables are investigated. Second, a random field with a certain spatial interactions instead of i. i. d. random variables is studied. Asymptotic normality of the proposed estimator is established under appropriate conditions.

AMS 2000 subject classifications. Primary 62G07, 62G20; Secondary 62M40.

Key words and phrases. Random field, characteristic function, deconvolution, asymptotic normality.

Send Correspondence to:
Jiexiang Li
College of Charleston
Math Dept
Charleston, SC 29424
1 Introduction

Denote the integer lattice points in the $N$ dimensional Euclidean space by $\mathbb{Z}^N$ and consider a strictly stationary $\mathbb{R}^d$ valued random field $X_i$, $i \in \mathbb{Z}^N$. One tries to observe a random variable $X_i$ but the observations are contaminated with noise such as measurement errors. Hence, one can only observe $Y_i$, the sum of the true random variable $X_i$ and the error variable $\epsilon_i$, where the true random variable and the error are assumed to be independent. The observable random variable is given as $Y_i = X_i + \epsilon_i$ and one can only observe a sample from the convolved density $f_Y = f_X * f_\epsilon$. However, we are interested in the density $f_X$, so we have to solve a deconvolution problem. There are many different approaches in this setting that lead to consistent estimators of $f_X$. The most common method is the deconvolving kernel approach introduced by Stefanski and Carroll (1990). Fan (1991a) showed that for independent identically distributed (i.i.d.) random variables the convergence rate obtained by the deconvolving kernel density estimator is optimal for estimating density $f_X$ and its derivatives. Fan (1991b) showed asymptotic normality for the deconvolving kernel density estimators. Masry (1993a, 1993b) studied the convergence rate and asymptotic normality of the deconvolving kernel density estimators for stationary processes satisfying strongly mixing or $\rho$-mixing conditions. Suppose we observe $Y_i$ on a rectangular region $I_n$ defined by

$$I_n = \{i : i \in \mathbb{Z}^N, 1 \leq i_k \leq n_k, k = 1, ..., N\}.$$

The construction of the kernel density estimators is similar to that proposed by Stefanski and Carroll (1990) and Masry (1993b). Let $K$ be $\mathbb{R}^d$ valued smooth kernel density with characteristic function $\phi_K$ and $\phi_\epsilon$ be the characteristic function of the error variable $\epsilon_i$. The bandwidth $b_n$ is a sequence of positive numbers such that $b_n \to 0$ as $n \to \infty$, which means $\min(n_1, ..., n_N) \to \infty$. The deconvolving kernel density estimator of $f_X$ is

$$\hat{f}_n(x) = \frac{1}{nh_n^d} \sum_{i \in I_n} g_n\left(\frac{x - Y_i}{b_n}\right)$$

(1.1)

with

$$g_n(x) = \frac{1}{(2\pi)^d} \int e^{-i t \cdot x} \frac{\phi_K(t)}{\phi_\epsilon(t/b_n)} dt.$$
The kernel density estimator in (1.1) can also be written as
\[
\hat{f}_n(x) = \frac{1}{(2\pi)^d} \int e^{-it \cdot x} \hat{\phi}_n(t) \frac{\phi_K(b_n t)}{\phi_\epsilon(t)} dt
\]
with
\[
\hat{\phi}_n(t) = \frac{1}{n} \sum_{i \in I_n} e^{i t \cdot y_i}.
\]
The purpose of the paper is to establish the asymptotic normality of the deconvolving kernel density estimators from dependent random fields. Note that the extension is not trivial because of the difficulties coming from spatial ordering.

A point \( i \) in \( \mathbb{Z}^N \) will be referred to as a site and written as \( i = \langle i_1, i_2, \ldots, i_N \rangle \). Let \( S \) and \( S' \) be two sets of sites. The Borel \( \sigma \)-fields \( \mathcal{B}(S) = \mathcal{B}((X_i, \epsilon_i), i \in S) \) and \( \mathcal{B}(S') = \mathcal{B}((X_i, \epsilon_i), i \in S') \) are the \( \sigma \)-fields generated by the random variables \( (X_i, \epsilon_i) \) with \( i \) in, respectively, \( S \) and \( S' \). Define distance between \( S \) and \( S' \) as follows: \( d(S, S') = \min(d(i, j) : i \in S, j \in S') \) where \( d(i, j) = ||i - j|| = \sqrt{(i_1 - j_1)^2 + \ldots + (i_N - j_N)^2} \). We will assume that \( (X_i, \epsilon_i), i \in \mathbb{Z}^N \) satisfies the following mixing condition: There exists a function \( \psi(t) \downarrow 0 \) as \( t \to \infty \), such that whenever \( S, S' \subset \mathbb{Z}^N \),
\[
\alpha(\mathcal{B}(S), \mathcal{B}(S')) = \sup\{|P(A \cap B) - P(A)P(B)|, A \in \mathcal{B}(S), B \in \mathcal{B}(S')\} \leq g(\text{Card}(S), \text{Card}(S'))\psi(d(S, S')) \tag{1.2}
\]
where \( \text{Card}(S) \) denotes the cardinality of \( S \) and \( g \) is a symmetric positive function non-decreasing in each variable. We assume that \( g \) satisfies
\[
g(n, m) \leq C(n + m + 1)^\ell \tag{1.3}
\]
for some \( \ell \) with \( \ell > 1 \) and some \( C > 0 \). If \( g \equiv 1 \), then \( X_i, i \in \mathbb{Z}^N \) is called strongly mixing, which is one of the mixing conditions discussed in Masry (1993a, 1993b). Mixing conditions (1.2) and (1.3) are satisfied by many spatial models and examples can be found in Guyon (1987). In the following proof, \( C \) is generic whose values are unimportant and may vary from line to line. And if not explicitly stated, the limit in the following is taken as \( n \to \infty \).
2 Assumptions

Assumption 2.1 The density of $Y_i$, $f_y$, is continuous.

Assumption 2.2 The joint density $f_{ij}$ of $Y_i$ and $Y_j$ exists and satisfies

$$|f_{ij}(u,v) - f_Y(u)f_Y(v)| \leq C$$ (2.1)

for all $i \neq j$ and for all $(u,v)$ in some neighborhood of $(x,x)$.

Assumption 2.3 For some $\beta > 0$, $\|t\|^{\beta d} |\phi_\epsilon(t)| \rightarrow B$ as $t \rightarrow \infty$.

Remark 2.1 As stated in Fan (1991b), the limiting behavior of the estimator in (1.1) depends heavily on the tail of $\phi_\epsilon$. Here, we assume the ordinary smoothness of the characteristic function $\phi_\epsilon$. Double exponential, Gamma distribution, and their mixtures are examples of this smooth type.

Assumption 2.4 $\phi_K(t)$ vanishes outside $[-1, 1]^d$.

Remark 2.2 Let $K(x) = \frac{1}{\pi} \frac{1}{1 + \frac{x^2}{2}}$ for $-\infty < x < \infty$, then $\phi_K(t) = (1 - |t|)I_{(-1,1)}(t)$.

In the paper, we assume either algebraically or geometrically decreasing mixing rate of target random field and noise random field.

Assumption 2.5 i) Bandwidth goes to zero not too slow: for some $\alpha$ with $0 < \alpha < 1$,

$$\hat{n}^{\kappa d(2\beta + 1)} \rightarrow 0.$$

ii) The mixing rate decreases at a polynomial speed:

$$\psi(t) \leq |t|^{-\theta}$$

for some $\theta > N$, $|t| \rightarrow \infty$.

iii) Bandwidth goes to zero not too fast:

$$\hat{n}^{\kappa(\frac{d}{N} - 1)} \rightarrow \infty.$$
Remark 2.3 (i) and (iii) in assumption 2.5 implies \( \theta > N(1 + \frac{1}{2\beta+1}) \).

Assumption 2.6 i) Bandwidth goes to zero not too slow: for some \( \alpha \) with \( 0 < \alpha < 1 \),

\[
\hat{n}^\alpha b_n^{(2\beta+1)d} \to 0.
\]

ii) The mixing rate decreases exponentially:

\[
\psi(t) \leq e^{-\lambda|t|}
\]

for some \( \lambda > 0 \), \( |t| \to \infty \).

iii) Bandwidth goes to zero not too fast: for some \( \theta_1 > 0 \),

\[
\hat{n}^{1-2\theta_1} b_n^d \to \infty.
\]

Remark 2.4 (i) and (iii) in assumption 2.6 implies \( \theta_1 < \frac{1}{2} - \frac{\alpha}{4\beta+2} \).

In kernel density estimation, it is important to choose the most appropriate bandwidth. The estimation would not be useful if the bandwidth is too small or too large.

3 Preliminaries

Lemma 3.1 For smooth density \( K(x) \), we have

\[
E[\hat{f}_n(x)] \to f_X(x)
\]

as \( n \to \infty \).

Proof. Let \( \phi_Y(t) \), \( \phi_X(t) \) be the characteristic functions of \( Y_i \) and \( X_i \) respectively and note that \( \phi_Y = \phi_X * \phi_e \).

\[
E[\hat{f}_n(x)] = \frac{1}{(2\pi)^d} \int e^{-it \cdot x} \phi_Y(t) \frac{\phi_K(b_n t)}{\phi_e(t)} dt
\]

\[
= \frac{1}{(2\pi)^d} \int e^{-it \cdot x} \phi_X(t) \phi_K(b_n t) dt
\]

\[
\to \frac{1}{(2\pi)^d} \int e^{-it \cdot x} \phi_X(t) dt = f_X(x).
\]
We will need a lemma in Ibragimov and Linnik (1971) or Deo (1973), which is presented here for completeness.

**Lemma 3.2** (1) Suppose (1.2) holds. Let $\mathcal{L}_r(\mathcal{F})$ denote the class of $\mathcal{F}$-measurable r.v.’s $X$ satisfying $\|X\|_r = (E|X|^r)^{1/r} < \infty$. Let $X \in \mathcal{L}_r(\mathcal{B}(S))$ and $Y \in \mathcal{L}_s(\mathcal{B}(S'))$. Suppose $1 \leq r,s,h < \infty$ and $r^{-1} + s^{-1} + h^{-1} = 1$. Then

$$|EXY - EXEY| \leq C\|X\|_r \|Y\|_s \{g(\text{Card}(S), \text{Card}(S'))\psi(d(S, S'))\}^{1/h}. \quad (3.1)$$

(2) For r.v.’s bounded with probability 1, the right-hand side of (3.1) can be replaced by $Cg(\text{Card}(S), \text{Card}(S'))\psi(d(S, S'))$.

Denote

$$\mu_n = \frac{1}{b_n^d}Eg_n\left(\frac{x - Y_i}{b_n}\right)$$

and

$$\xi_i = \frac{1}{b_n^d}g_n\left(\frac{x - Y_i}{b_n}\right) - \mu_n.$$  

Hence,

$$\hat{f}_n(x) - E[\hat{f}_n(x)] = \frac{1}{n} \sum_{i \in I_n} \xi_i.$$  

**Lemma 3.3** Suppose the strictly stationary random field $(X_i, \epsilon_i)$, $i \in \mathbb{Z}^N$ satisfies (1.2) and (1.3). If assumptions 2.1–2.5 hold, then

$$\lim \hat{n}_n^{(2\beta+1)d} \text{Var}(\hat{f}_n(x)) = \frac{f_Y(x)}{(2\pi)^d B^2} \int \|t\|^{2\beta d} |\phi_K(t)|^2 dt := \sigma^2(x).$$

Proof.

$$\hat{n}_n^{(2\beta+1)d} \text{Var}(\hat{f}_n(x))$$

$$= b_n^{(2\beta+1)d} \text{Var}(\xi_1)$$

$$+ \hat{n}_n^{-1}b_n^{(2\beta+1)d} \sum_{i \neq j \in I_n} \text{Cov}(\xi_i, \xi_j)$$

$$:= V_1 + V_2. \quad (3.2)$$
Note
\[ E b_n^{-d} g_n \left( \frac{x - Y_i}{b_n} \right) = \int g_n(u) f_Y(x - b_n u) du \leq C \]

Therefore,
\[ b_n^{(2\beta+1)d} E^2 b_n^{-d} g_n \left( \frac{x - Y_i}{b_n} \right) \rightarrow 0 \tag{3.3} \]

Note
\[
\begin{align*}
b_n^{(2\beta+1)d} E b_n^{-2d} g_n^2 \left( \frac{x - Y_i}{b_n} \right) &= b_n^{2\beta d} \int g_n^2(u) f_Y(x - b_n u) du \\
&= b_n^{2\beta d} f_Y(x) \int g_n^2(u) du (1 + o(1)) \\
&= \sigma^2(x) (1 + o(1)) \tag{3.4}
\end{align*}
\]

Many difficulties arise since points in higher dimensional space \( N \geq 2 \) cannot be linearly ordered. To deal with this, we divide sites in the rectangular region, \( I_n \), into two groups.

Let \( c_n^N = \hat{n}^\alpha \) and define

\[
S_1 = \{ (i, j) \in I_n | 0 < d(i, j) \leq c_n \},
\]

\[
S_2 = \{ (i, j) \in I_n | d(i, j) > c_n \}.
\]

On one hand, for \( (i, j) \in S_1 \), we establish an upper bound for covariance between \( \xi_i \) and \( \xi_j \).

\[
| \text{Cov}(\xi_i, \xi_j) | = | \int \int g_n(u) g_n(v) [ f_{i,j}(x - b_n u, x - b_n v) - f_Y(x - b_n u) f_Y(x - b_n v)] du dv |
\]

\[
\leq C \left( \int g_n(u) du \right)^2 \leq C
\]

Hence,
\[
\hat{n}^{-1} b_n^{(2\beta+1)d} \sum_{(i,j) \in S_1} | \text{Cov}(\xi_i, \xi_j) | \leq C c_n^N b_n^{(2\beta+1)d} \rightarrow 0 \tag{3.5}
\]

On the other hand, for \( (i, j) \in S_2 \), we establish the covariance inequality with an application of lemma 3.2.

\[
| \text{Cov}(\xi_i, \xi_j) | \leq C b_n^{-2d} \| g_n \|_\infty^2 \psi(\| i - j \|)
\]

\[
\leq C b_n^{-2+2\beta d} \psi(\| i - j \|)
\]
Under assumption 2.5, for sufficiently large $n$

\[
\hat{n}^{-1} b_n^{2\beta+1} \sum_{(i,j) \in S_2} |\text{Cov}(\xi_i, \xi_j)| \leq C b_n^{-d} \sum_{i \geq cn} i^{N-1} e^{-\theta}
\leq C b_n^{-d} c_n^{N-\theta} \to 0. \tag{3.6}
\]

The proof of the lemma is complete by (3.2) through (3.3), (3.4) and (3.5), (3.6). \hfill \square

**Lemma 3.4** Suppose the strictly stationary random field $(X_i, \epsilon_i), i \in \mathbb{Z}^N$ satisfies (1.2) and (1.3). If assumptions 2.1–2.4 and assumption 2.6 hold, then

\[
\lim \hat{n} b_n^{2\beta+1} \text{Var}(\hat{f}_n(x)) = \frac{f_Y(x)}{(2\pi)^d B^2} \int \|t\|^{2\beta} |\phi_K(t)|^2 dt := \sigma^2(x).
\]

Proof. Under assumption 2.6 for sufficiently large $n$ and some $\lambda_0 < \lambda$,

\[
\hat{n}^{-1} b_n^{2\beta+1} \sum_{(i,j) \in S_2} |\text{Cov}(\xi_i, \xi_j)| \leq C b_n^{-d} \sum_{i \geq cn} i^{N-1} e^{-\lambda_i}
\leq C b_n^{-d} \hat{n}^{\lambda_0 C} \to 0 \tag{3.7}
\]

The convergence follows from an arbitrary choice of $C$ such that $\lambda_0 C \geq 1 - 2\theta_1$. Following (3.2) with (3.3), (3.4) and (3.5), (3.7), the proof of the lemma is complete. \hfill \square

### 4 Asymptotic normality

Decompose

\[
\hat{n}(\hat{f}_n(x) - \mathbb{E}[\hat{f}_n(x)]) = \sum_{i_k=1}^{n_k} \xi_i
\]
into smaller pieces involving “large” and “small” blocks. More specifically, consider all
sums are running over \( i := (i_1, \ldots, i_N) \).

\[
U(1, n, x, j) := \sum_{i_k = j_k(p+q)+1}^{j_k(p+q)+p} \xi_i,
\]

\[
U(2, n, x, j) := \sum_{i_k = j_k(p+q)+1}^{j_k(p+q)+p} \sum_{i_N = j_N(p+q)+p+1}^{(j_N+1)(p+q)} \xi_i,
\]

\[
U(3, n, x, j) := \sum_{i_k = j_k(p+q)+1}^{j_k(p+q)+p} \sum_{i_{N-1} = j_{N-1}(p+q)+p+1}^{(j_{N-1}+1)(p+q)} \sum_{i_N = j_N(p+q)+p+1}^{j_N(p+q)+p} \xi_i,
\]

\[
U(4, n, x, j) := \sum_{i_k = j_k(p+q)+1}^{j_k(p+q)+p} \sum_{i_{N-1} = j_{N-1}(p+q)+p+1}^{(j_{N-1}+1)(p+q)} \sum_{i_N = j_N(p+q)+p+1}^{(j_N+1)(p+q)} \sum_{i_{N-2} = j_{N-2}(p+q)+p+1}^{j_{N-2}(p+q)+p+1} \xi_i,
\]

and so on. Note that

\[
U(2^N - 1, n, x, j) := \sum_{i_k = j_k(p+q)+p+1}^{(j_k+1)(p+q)} \sum_{i_N = j_N(p+q)+p}^{j_N(p+q)+p+1} \xi_i,
\]

and

\[
U(2^N, n, x, j) := \sum_{i_k = j_k(p+q)+p+1}^{(j_k+1)(p+q)} \xi_i
\]

where \( p := p_n = \hat{n}^{\theta_n} \) and \( q := q_n = \hat{n}^{\theta_3} \) with \( 0 < \theta_3 < \frac{\theta_1}{N} \leq \frac{1-\alpha(\theta-N)}{2N} \) and \( \ell + 1 < \theta_1 + \theta \theta_3 \).

Note \( \frac{p}{q} \to 0 \). Without loss of generality, assume that, for some integers \( r_1, \ldots, r_N, n = (n_1, \ldots, n_N) \) is such that \( n_1 = r_1(p + q), \ldots, n_N = r_N(p + q) \), with \( r_k \to \infty \) for all \( k = 1, \ldots, N \). Let \( M = r_1 \cdots r_N \), then \( \hat{n} = M(p + q)^N \). For each integer \( 1 \leq i \leq 2^N \), define

\[
T(n, x, i) := \sum_{j_k = 0}^{r_k-1} U(i, n, x, j).
\]

Clearly \( S_n = \sum_{i=1}^{2^N} T(n, x, i) \). Note that \( T(n, x, 1) \) is the sum of the random variables \( \xi_i \) over “large” blocks, whereas \( T(n, x, i), 2 \leq i \leq 2^N \) are sums over “small” blocks. If it is not the case that \( n_1 = r_1(p + q), \ldots, n_N = r_N(p + q) \) for some integers \( r_1, \ldots, r_N \), then an additional term \( T(n, x, 2^N + 1) \), say, containing all the \( \xi_i \)'s that are not included in the
large or small blocks, can be considered. This term will not change the proof much. The general approach consists of showing that, as $n \to \infty$,

$$Q_1 := \left| \exp[iuT(n, x, 1)] - \prod_{j_k=0}^{r_k-1} \exp[iuU(1, n, x, j)] \right| \to 0, \quad (4.1)$$

$$Q_2 \equiv \hat{n}^{-1}b_n^{(2\beta+1)d} \mathbb{E} \left( \sum_{i=2}^{2N} T(n, x, i) \right)^2 \to 0, \quad (4.2)$$

$$Q_3 := \hat{n}^{-1}b_n^{(2\beta+1)d} \sum_{j_k=0}^{r_k-1} \sum_{k=1}^{N} \mathbb{E}U^2(1, n, x, j) \to \sigma^2(x), \quad (4.3)$$

$$Q_4 \equiv \hat{n}^{-1}b_n^{(2\beta+1)d} \sum_{j_k=0}^{r_k-1} \sum_{k=1}^{N} \mathbb{E}[U^2(1, n, x, j)]I\{|U(1, n, x, j)| > \varepsilon \sigma(x) n^{\frac{1}{2}} \} \to 0 \quad (4.4)$$

for every $\varepsilon > 0$. The term $\hat{n}^{-\frac{1}{2}}b_n^{(2\beta+1)d} \sum_{i=2}^{2N} T(n, x, i)$ is asymptotically negligible by (4.2). The random variables $U(1, n, x, j)$ (with $j_k = 0, ..., r_k - 1$ for $k = 1, ..., N$) are asymptotically mutually independent by (4.1). The asymptotic normality of $\hat{n}^{-\frac{1}{2}}b_n^{(2\beta+1)d} T(n, x, 1)$ follows from (4.3) and the Lindeberg-Feller condition (4.4). The arguments are reminiscent of those used by Masry (1986) and Nakhapetyan (1987).

Proof of (4.1). Following the proof of Lemma 5.3 in Hallin et al. (2004), we know (4.1) is bounded by $\frac{\hat{n}^{\ell+1}}{p_N} \phi(q)$ under (1.3).

Under assumption 2.5,

$$\frac{\hat{n}^{\ell+1}}{p_N} \phi(q) \leq C \hat{n}^{\ell+1-\theta_1-\theta_3} \to 0.$$  

Under assumption 2.6,

$$\frac{\hat{n}^{\ell+1}}{p_N} \phi(q) \leq C \hat{n}^{\ell+1-\theta_1-C\lambda} \to 0.$$  

Proof of (4.2) and (4.3) follows from the proof of Lemma 3.3 and our choice of $p$ and $q$.  

10
Proof of (4.4). Note

\[ \|g_n\|_\infty \leq C \int \frac{\phi_K(t)}{\phi(t/b_n)} \leq C b_n^{-\beta_d}, \]

which implies

\[ |\xi_i| \leq C b_n^{-(\beta+1)d}. \]

Therefore

\[ |U(1, n, x, j)| \leq C p^N b_n^{-(\beta+1)d}. \]

Note

\[ \frac{p^N}{b_n^d} \to 0. \]

For sufficiently large \( n \),

\[ Q_4 \leq C \hat{n}^{-1} p^{2N} b_n^{-d} \sum_{j_k=0}^{r_k-1} P[|U(1, n, x, j)| > \varepsilon \sigma(x) \hat{n}^{(2\beta+1)d}{1/2} b_n^{2/2}] \equiv 0. \]

Following the argument above and Lemma 3.1, we have our theorem on the asymptotic normality of the deconvolving kernel density estimator.

**Theorem 4.1** Suppose the strictly stationary random field \((X_i, \epsilon_i), i \in \mathbb{Z}^N\) satisfies (1.2) and (1.3). If assumptions 2.1–2.5 hold and \( \theta \) satisfies \( 2(\ell + 1 - \theta \theta_3) \leq 2\theta_1 \leq 1 - \alpha (\theta - N) \) for any \( \theta_1 \) and \( \theta_3 \) such that \( 0 < \theta_3 < \frac{\theta}{N} \), then

\[ (\hat{n}^{(2\beta+1)d}{1/2} b_n^{(2\beta+1)d}{1/2})(\hat{f}_n(x) - f_X(x)) \to N(0, \sigma^2(x)) \]

**Theorem 4.2** Suppose the strictly stationary random field \((X_i, \epsilon_i), i \in \mathbb{Z}^N\) satisfies (1.2) and (1.3). If assumptions 2.1, 2.4, and assumption 2.6 hold and \( \lambda > \ell + 1 - \theta_1 \) for some \( \theta_1 > 0 \), then,

\[ (\hat{n}^{(2\beta+1)d}{1/2} b_n^{(2\beta+1)d}{1/2})(\hat{f}_n(x) - f_X(x)) \to N(0, \sigma^2(x)) \]

**References**

[1] Deo, C. M. (1973). A note on empirical processes of strong mixing sequences. *Ann. Probab.* 1, 870-875.
[2] Fan, J., 1991. On the optimal rates of convergence for nonparametric deconvolution problems. *J. Ann. Statist.* **19**, 1257–1272.

[3] Fan, J., 1991. Asymptotic normality for deconvolution kernel density estimators. *Sankhya Ser. A* **53**, 97–110.

[4] Guyon, X. (1987). Estimation d’un champ par pseudo-vraisemblance conditionnelle: étude asymptotique et application au cas markovien. Proc. 6th Franco-Blegian Meeting of Statisticians.

[5] Hallin, M., Lu Z. and Tran, L.T. (2004) Local linear spatial regression. *Ann. Statist.* **32**, 2469-2500.

[6] Ibragimov, I. A. and Linnik, Y. V. (1971). Independent and Stationary Sequence of Random Variables. Wolters-Noordhoff publishing, Groningen. 443 pp

[7] Masry, E. (1986). Recursive probability density estimation for weakly dependent processes. *IEEE Trans. Inform. Theory* **32**, 254-267.

[8] Masry, E., 1993a. Strong consistency and rates for deconvolution of multivariate densities of stationary processes. *Stochastic Process. Appl.* **47**, no. 1, 53–74.

[9] Masry, E., 1993b. Asymptotic normality for deconvolution estimators of multivariate densities of stationary processes. *J. Multivariate Anal.* **44**, 47–68.

[10] Nakhapetyan, B. S. (1987). An approach to the proof of limit theorems for dependent random variables. *Teor. Veroyatnost. i Primenen.* **32**, 589–594.

[11] Stefanski, L., Carroll, R. J., 1990. Deconvolution kernel density estimators. *Statistics* **2**, 169–184.