C. R. Acad. Sci. Paris, Ser. I 354 (2016) 205–210

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Statistics

# Nonparametric recursive density estimation for spatial data

## *Estimation nonparamétrique récursive de la densité pour données spatiales*

Aboubacar Amiri [a], Sophie Dabo-Niang [a,b], Mohamed Yahaya [a,c]

[a] *Université Lille 3, Laboratoire LEM CNRS 9221, France*
[b] *INRIA-MODAL Team, France*
[c] *FST, Université des Comores, Comoros*

ARTICLE INFO

*Article history:*
Received 26 February 2014
Accepted after revision 18 October 2015
Available online 23 December 2015

Presented by the Editorial Board

ABSTRACT

This paper deals with non-parametric density estimation for spatial data. We study the asymptotic properties of a new recursive version of the Parzen–Rozenblatt estimator. The mean square error and an almost sure convergence result with rate of such estimator are derived.



RÉSUMÉ

Ce papier traite de l'estimation de la densité spatiale dans le cas récursif. Nous étudions les propiétés asymptotiques d'une nouvelle version de l'estimateur de Parzen–Rozenblatt. Nous établissons les convergences en moyenne quadratique et presque sûre de cet estimateur ; des vitesses de convergence sont données.



## 1. Introduction

We consider recursive kernel density estimation for spatially dependent data. Spatial density estimation is an interesting and crucial problem in statistical inference for a number of applications, where the considered variables have spatial dependence. Spatial data are modeled as finite realizations of random fields collected from different spatial locations on the earth and they are widely used in a variety of fields, including soil science, geology, oceanography, econometrics, epidemiology, environmental science, forestry, and many others, see Cressie [9], Chiles and Delfiner [7]. Although potential applications of nonparametric spatial models are abundant, little theoretical work has been devoted to nonparametric space modeling compared to the parametric case.

This work concerns a nonparametric estimation of the probability density function from dependent spatial data, using a recursive kernel approach. For an overview on results and applications considering independent or dependent spatial data for density estimation by classical kernel method, we highlight the works by Biau and Cadre [3], Carbon et al. [5], Dabo-Niang and Yao [10], Tran [19]. The method and the results obtained here generalize the previous works in the recursive framework.

*E-mail addresses:* aboubacar.amiri@univ-lille3.fr (A. Amiri), sophie.dabo@univ-lille3.fr (S. Dabo-Niang), mohamed.yahaya@etu.univ-lille3.fr (M. Yahaya).

http://dx.doi.org/10.1016/j.crma.2015.10.010

With the advances in hardware technology and the sophistication of modern computing and data acquisition techniques, voluminous spatial data are collected in various applied areas. New challenges arise since in many applications the volume of the data is so large that it may be impossible to store them on desk and their modeling requires many time. Nonparametric estimators such as the kernel density estimator of Parzen–Rozenblatt can be used to estimate the density function. However, they suffer from a serious computational drawback in big data context. In fact, if the density is estimated by a non-recursive estimator, this last must be completely recalculated when new data items are received.

In situations where data items arrive sequentially and a large amount of data can be generated at a rapid rate, a recursive updating algorithms provide great help for the difficulties arising from the large volume of data. Then they are much preferred over non-recursive ones, since recursive estimators can be updated from their immediate past and the new observation. Therefore, the update can be computed instantly and does not require extensive storage of data. This arrangement is particular to recursive methods and is called the online or real-time updating property. This recursive property is particularly interesting when the sample data are continually captured over time.

Recursive kernel approaches were introduced and abundantly studied in the non-spatial case. We refer, for instance, to the pioneer works by Wolverton and Wagner [20], Deheuvels [12], Wegman and Davies [11], who have paid attention to nonparametric recursive density estimation for time processes. There are some recent contributions on asymptotic properties of recursive kernel density for dependent data. Amiri [1] generalized all the above-cited versions to a general family of recursive estimators, using different values of a parameter $\ell \in [0,1]$ associated with each estimator. For the fixed values $\ell = 1/2$ and $\ell = 1$, the asymptotic normality under negative association was previously treated by Liang and Baek [2], while the mean-square convergence and the asymptotic normality were studied by Masry [14] under $\alpha$-mixing condition. Mezhoud et al. [15] have extended the results obtained by Amiri [1] to $\eta$-dependent time process.

We extend the feature of time-varying sample recursive density estimate to the spatial case. Namely, we present a recursive version of the classical spatial kernel density estimator and establish some asymptotic results of such estimator. As far as we know, the investigation of recursive kernel methods for spatial data remain an open problem.

The rest of the paper is organized as follows. In section 2, we present the spatial recursive density estimator, while Section 3 gives general assumptions. In section 4, we establish asymptotic convergence, namely mean square error and almost-sure convergence with the rate of our estimator. Finally Section 5 is devoted to some indications to prove the theoretical results presented in this note.

## 2. The recursive spatial kernel density estimate

Let $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^N}$, $N \geq 1$ be a measurable strictly stationary spatial process defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ and valued in $\mathbb{R}^d$, where $d \geq 1$. We assume that the $X_{\mathbf{i}}$'s have the same distribution as a random variable $X$, with unknown probability density function $p$. Suppose that we observe the process on a spatial set of sites $\mathcal{I}_n$ of cardinal $n$, which is a finite subset of a potentially observable region $\mathcal{D} \subset \mathbb{R}^N$, where $\mathbb{R}^N$ is endowed with the uniform metric. For convenience, we treat the observations sites as an array that is $\mathcal{I}_n = \left\{s_j, j = 1, \ldots, n\right\} \subset \mathbb{N}^N$. The estimator of the probability density function based on $(X_{\mathbf{i}},\ \mathbf{i} \in \mathcal{I}_n)$ to be studied in the present work is of the form:

$$p_n^\ell(x_0) := \frac{1}{S_{n,\ell}} \sum_{\mathbf{i}\in\mathcal{I}_n} \frac{1}{h_{\mathbf{i}}^{d\ell}} K\left(\frac{X_{\mathbf{i}} - x_0}{h_{\mathbf{i}}}\right),\ x_0 \in \mathbb{R}^d\ (\ell \in [0,1]), \tag{1}$$

where $S_{n,\ell} := \sum_{\mathbf{i}\in\mathcal{I}_n} h_{\mathbf{i}}^{d(1-\ell)}$, $h_{\mathbf{i}}$ is a bandwidth corresponding to the observation $X_{\mathbf{i}}$. By enumerating the sites, one may rewrite $p_n^\ell(x_0)$ as

$$p_n^\ell(x_0) := \frac{1}{S_{n,\ell}} \sum_{j=1}^{n} \frac{1}{h_{s_j}^{d\ell}} K\left(\frac{X_{s_j} - x_0}{h_{s_j}}\right),\ x_0 \in \mathbb{R}^d\ (\ell \in [0,1]), \tag{2}$$

where $S_{n,\ell} := \sum_{j=1}^{n} h_{s_j}^{d(1-\ell)}$. This representation of the estimator presents a great interest from the computational point of view. In fact, if $X_{s_{n+1}}$ is a new observation of the process at a site $s_{n+1}$ added to $\mathcal{I}_n$, the estimator can be updated recursively by the following formula:

$$p_{n+1}^\ell(x_0) = \frac{S_{n,\ell}}{S_{n+1,\ell}} p_n^\ell(x_0) + K_{n+1}^\ell\left(X_{s_{n+1}} - x_0\right) \text{with} \quad K_i^\ell(.) := \frac{1}{h_{\mathbf{s}_i}^{d\ell} S_{i,\ell}} K\left(\frac{.}{h_{s_i}}\right). \tag{3}$$

In (3), $p_{n+1}^\ell(x_0)$ is the density estimator based on the domain $\mathcal{I}_n \cup \{s_{n+1}\}$. This recursive formula is useful when the number of the spatial sites increases sequentially on space.

To simplify the presentation, the parameter $\ell$ is throughout supposed to be equal to 0 and the corresponding estimator will be later noted $p_n$.

Basically, two asymptotic methods occur in the spatial literature: increasing domain and infill asymptotics, see [9], p. 480. The growth of the sample in increasing domain asymptotics is a consequence of an unbounded expansion of the sample region $\mathcal{I}_n$, whereas under infill asymptotics the sample region is fixed and the growth of the sample size is due to sampling that is dense in the region $\mathcal{D}$. In what follows we consider the increasing domain asymptotics and for simplicity the bivariate regular lattice ($N = 2$), described in the following assumption.

**H 2.1.** *$D$ is a regular lattice and $\mathcal{I}_n = \{\mathbf{i} = (i_1, i_2),\ 1 \le i_j \le n_j,\ j = 1, 2\}$ is rectangular. In this case, $n = n_1 \times n_2$ goes to infinity means that $\min n_{j\, j=1,2}$ goes to infinity. Therefore, we use the lexicographic order and renumber the observations as a triangular array in the following way. The observation site $(i, j)$ can be indexed by $t = n_2(i-1) + j$ in the triangular array setting, see [17]. In this case, the new observation site is $(n_1 + 1, 1)$, which is indexed by $n + 1$ in the array setting.*

Notice that to obtain our asymptotic results in the irregular lattice context, one can consider additional assumptions on the number of spatial unit on a closed ball of $\mathcal{D}$ with the following:

**H 2.2.** *The potential observable region $\mathcal{D}$ is an irregular lattice and infinite countable. All elements of $\mathcal{D}$ are located at distances of at least $\delta > 1$ from each other.*

*In this case (as in [8]), we consider a sequence of finite closed, convex regions $\{R_u\}$ of increasing area as $u \to +\infty$ and let the sample set consists of the intersection of one of the convex regions and $D$: $\mathcal{I}_n = \mathcal{D} \subset R_u$. We do not specify here any order and any other assumption on the configuration and growth of sample size except the assumption of a uniformly increasing area of $\{R_u\}$ in at least two non-opposing directions to increase the sample size $n$ (as $u \to +\infty$), see [8]. Thus, a new observation corresponds to any new site $s_{n+1}$ of $\mathcal{D}$ not already included in $\mathcal{I}_n$.*

## 3. General assumptions

In order to establish the asymptotic results, the following assumptions will be considered. Let us consider the representation of the recursive density estimator given in (2), with $\ell = 0$.

**H 3.1.** *(a) $h_{s_n} \downarrow 0$ and $nh_{s_n}^{d+2} \to \infty$ as $n \to \infty$; (b) For $r \in ]-\infty, 2+d]$, $B_{n,r} := \frac{1}{n}\sum_{i=1}^{n}\left(\frac{h_{s_i}}{h_{s_n}}\right)^r \to B_r > 0$.*

**H 3.2.** *$K$ is a bounded symmetric density such that $\int_{\mathbb{R}^d} uK(u)\mathrm{d}u = 0$ and $\int_{\mathbb{R}^d} \|u\|^2 K(u)\mathrm{d}u < \infty$.*

**H 3.3.** *The density $p$ is bounded and twice continuously differentiable on $\mathbb{R}^d$.*

**H 3.4.** *For any $i, j \in \{1, \dots, n\}$ such that $s_i \neq s_j$, the random vector $(X_{s_i}, X_{s_j})$ has a density $f_{s_i,s_j}$ such that:*

$$\sup_{s_i \neq s_j} \left\| g_{s_i,s_j} \right\|_\infty < \infty, \text{ where } g_{s_i,s_j} = f_{s_i,s_j} - p \otimes p, \quad i, j = 1, \dots, n.$$

**H 3.5.**

(i) *The field $(X_{s_i})_{1 \le i \le n}$ is $\alpha$-mixing: there exists a function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ with $\phi(t) \searrow 0$ as $t \to \infty$, such that whenever $E, E' \subset \mathbb{R}^2$ with finite cardinals $Card(E)$, $Card(E')$*

$$\alpha\left(\sigma\left(E\right), \sigma\left(E'\right)\right) := \sup_{A \in \sigma(E),\ B \in \sigma(E')} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\,\mathbb{P}(B)| \le \psi\left(Card(E), Card(E')\right)\phi\left(dist(E, E')\right),$$

*where $\sigma(E) = \{X_\mathbf{i}, \mathbf{i} \in E\}$ and $\sigma\left(E'\right) = \left\{X_\mathbf{i}, \mathbf{i} \in E'\right\}$, $dist(E, E')$ is the Euclidean distance between $E$ and $E'$ and $\psi(\cdot)$ is a positive symmetric function nondecreasing in each variable.*

*The functions $\phi$ and $\psi$ are such that $\phi(i) \le Ci^{-\theta}$ and $\psi(n, m) \le C\min(n, m)$.*

(ii) *For $\epsilon > 0$, let $u_n = \prod_{i=1}^{N}(\log n_i)(\log\log n_i)^{1+\epsilon}$ and assume that*

$$n \log n^{\frac{\theta - 2N}{4N - \theta}} h_n^{\frac{d\theta}{\theta - 4N}} u_n^{\frac{2N}{4N - \theta}} \to \infty, \text{ where } \theta > \max\left\{4N, \frac{Nd + 2}{2}\right\}.$$

The first condition of H 3.1 is usual in nonparametric recursive estimation, while the second one is crucial in our calculus. This last is particularly used in the recursive kernel estimation, for instance by Masry [14], Wegman and Davies [11], Yamato [21], Samanta and Mugisha [16] and Isogai [13]. If $h_{s_n} = A_n n^{-q}$ where $A_n \downarrow A > 0$, $0 < q < 1$, then the assumption H 3.1 is satisfied for all $q$ such that $rq < 1$ and $B_r = \frac{1}{1 - rq}$. Also the choice $h_{s_n} = A_n\left(\frac{\ln n}{n}\right)^q$, $0 < q < \frac{1}{2+d}$ satisfies assumption H 3.1. Assumptions H 3.2 and H 3.3 are classical in nonparametric estimation and easily verified by many kernels such as Epanechnikov, Gaussian kernels.

About the assumption H 3.5(i), basically, in the spatial literature it is supposed that $\phi(i)$ tends to zero with polynomial rate, or $\phi(i) \le C\exp(-si)$, for some $C, s > 0$, i.e. $\phi(i)$ tends to zero with exponential rate. Our results can be proved under the above condition. Finally let us mention that the definition of $u_n$ in H 3.5 leads to $\sum_{\mathbf{n}\in\mathbb{N}^N} \frac{1}{\widehat{\mathbf{n}}u_n} < \infty$ (recall that $\widehat{\mathbf{n}} = n$).

## 4. Consistency results

Theorem 1 below establishes the asymptotic mean square error of the estimator $p_n$ via its variance and bias.

**Theorem 1.** *Under the assumptions **H** 2.2 and **H** 3.1–3.5, we have:*

$$nh_{s_n}^d Var(p_n(x_0)) \to p(x_0)B_d^{-1}\int_{\mathbb{R}^d} K^2(u)\mathrm{d}u \tag{4}$$

*and*

$$h_{s_n}^{-2}(\mathbb{E}(p_n(x_0)) - p(x_0)) \to B_d^{-1}B_{d+2}\int_{\mathbb{R}^d} \frac{K(u)}{2}\sum_{j_1,j_2=1}^{d} u_{j_1}u_{j_2}\frac{\partial^2 p(x_0)}{\partial x_{j_1}\partial x_{j_2}}\mathrm{d}u \tag{5}$$

*as $n \to \infty$, for all $x_0$ such that $p(x_0) > 0$. In particular, if $h_{s_n} = n^{\frac{-1}{4+d}}$, we have the asymptotic behavior according to the mean square error:*

$$n^{\frac{4}{4+d}}\mathbb{E}(p_n(x_0) - p(x_0))^2 \to \left(\int_{\mathbb{R}^d} K(u)\sum_{j_1,j_2=1}^{d} u_{j_1}u_{j_2}\frac{\partial^2 p(x_0)}{\partial x_{j_1}\partial x_{j_2}}\mathrm{d}u\right)^2 + \frac{4}{4+d}p(x_0)\int_{\mathbb{R}^d} K^2(u)\mathrm{d}u \tag{6}$$

*as $n \to \infty$ for all $x_0$ such that $p(x_0) > 0$.*

Theorem 1 is an extension of the previous results obtained in the temporal case by Amiri [1] with $\alpha$-mixing condition. Also, a similar result were proved by Mezhoud et al. [15] for temporal $\eta$-dependence process. Now, we state the following theorem that establishes the almost convergence of $p_n$.

**Theorem 2.** *Under the assumptions **H** 2.1, **H** 3.1–3.5, we have:*

$$|p_n(x_0) - p(x_0)| = O\left(\sqrt{\frac{\log n}{nh_n^d}}\right), \quad a.s. \tag{7}$$

## 5. Brief outline of proofs

### 5.1. Theorem 1

For $x_0 \in \mathbb{R}^d$, let $I_1 = \frac{1}{B_{n,d}^2 n^2 h_{s_n}^{2d}}\sum_{i=1}^{n}\mathrm{Var}\left(Z_{s_i}\right)$ and $I_2 = \frac{1}{B_{n,d}^2 n^2 h_{s_n}^{2d}}\sum_{s_i\neq s_j}\mathrm{Cov}\left(Z_{s_i}, Z_{s_j}\right)$, with $Z_{s_i} = K\left(\frac{X_{s_i} - x_0}{h_{s_i}}\right)$.

On the one hand, under H 3.2 with the help of the Toeplitz Lemma (see [18]), we readily have

$$nh_{s_n}^d|I_1| \to p(x_0)B_d^{-1}\int_{\mathbb{R}^d} K^2(u)\mathrm{d}u \quad \text{as } n \to \infty.$$

On the other hand, for $x_0 \in \mathbb{R}^d$, let us split: $I_2 = J_1 + J_2$, where

$$J_1 = \frac{1}{B_{n,d}^2 n^2 h_{s_n}^{2d}}\sum_{\|s_i - s_j\|\le c_n}\mathrm{Cov}(Z_{s_i}, Z_{s_j}) \text{ and } J_2 = \frac{1}{B_{n,d}^2 n^2 h_{s_n}^{2d}}\sum_{c_n<\|s_i - s_j\|}\mathrm{Cov}(Z_{s_i}, Z_{s_j}),$$

$c_n$ being a sequence tending to infinity as $n$ tends to infinity. If $d \ge 3$, then $2d > d + 2$. In this case, we consider $\zeta \in \mathbb{R}$ such that $\frac{N}{\theta - 1} < \zeta \le \frac{2}{d}$, and using the decrease of $(h_{s_n})_n$, one may write:

$$\left|\mathrm{Cov}(Z_{s_i}, Z_{s_j})\right| \le h_{si}^d h_{s_j}^d \sup_{s_i\neq s_j}\left\|g_{s_i,s_j}\right\|_\infty \le \frac{h_{s_i}^{2d} + h_{s_j}^{2d}}{2}\sup_{s_i\neq s_j}\left\|g_{s_i,s_j}\right\|_\infty \le \frac{h_{s_i}^{d(\zeta+1)}h_{s_1}^{d(1-\zeta)} + h_{s_j}^{d(\zeta+1)}h_{s_1}^{d(1-\zeta)}}{2}\sup_{s_i\neq s_j}\left\|g_{s_i,s_j}\right\|_\infty.$$

If $1 \le d \le 3$, take $\zeta = 1$ and make as previously. As $\mathcal{I}_n$ is a regular design (see for instance [19])

$$\mathrm{Card}\left\{(\mathbf{u},\mathbf{v}) \in \mathcal{I}_n^2 : \|\mathbf{u}-\mathbf{v}\| \le c_n\right\} = O\left(nc_n^N\right).$$

It follows that $nh_{s_n}^d\,|J_1| \le \frac{B_{n,d(\zeta+1)}}{B_{n,d}^2} h_{s_1}^{d(1-\zeta)} h_{s_n}^{d\zeta} c_n^N \sup_{s_i \ne s_j} \left\|g_{s_i,s_j}\right\|_\infty = O\left(h_{s_n}^{d\zeta} c_n^N\right)$.

Turning to $J_2$, by the Billingsley inequality (see [4]), we have $\left|\mathrm{Cov}(Z_{s_i}, Z_{s_j})\right| \le 4\phi(\left\|s_i - s_j\right\|)\,\|K\|_\infty^2$. Consequently, we get

$$nh_{s_n}^d\,|J_2| \le \frac{4\,\|K\|_\infty^2}{B_{n,d}^2 h_{s_n}^d (\theta-1)} c_n^{1-\theta} = O\left(\frac{c_n^{1-\theta}}{h_{s_n}^d}\right).$$

The choice of $c_n = \left\lfloor h_{s_n}^{-\frac{d(\zeta+1)}{N+\theta-1}} \right\rfloor$ leads to $nh_{s_n}^d\,|I_2| = O\left(h_{s_n}^{\frac{d(\zeta(\theta-1)-N)}{N+\theta-1}}\right) \to 0$ as $n \to \infty$, as well as $\zeta > \frac{N}{\theta-1}$, and the result (4) follows. About the bias term, we apply the Taylor formula and the dominated convergence, which together with Toeplitz' lemma, leads to the result (5). □

### 5.2. Theorem 2

Let us set $G_n(x_0) = p_n(x_0) - \mathbb{E}p_n(x_0) = \sum_{i=1}^{n} \Delta_i$ with $\Delta_i = \frac{1}{S_{n,0}}\left(Z_{s_i} - \mathbb{E}Z_{s_i}\right)$. Next, let $a = a_n$ be an integer and split the random variables $\Delta_i$ into blocks as follows:

$$U(1,n,\mathbf{j},x_0) = \sum_{\substack{i_k=2j_k a_n+1\\ k=1,\dots,N}}^{(2j_k+1)a_n} \Delta_{\mathbf{i}}; \qquad U(2,n,\mathbf{j},x_0) = \sum_{\substack{i_k=2j_k a_n+1\\ k=1,\dots,N-1}}^{(2j_k+1)a_n} \sum_{i_N=(2j_N+1)a_n+1}^{2(j_N+1)a_n} \Delta_{\mathbf{i}};$$

$$U(3,n,\mathbf{j},x_0) = \sum_{\substack{i_k=2j_k a_n+1\\ k=1,\dots,N-2}}^{(2j_k+1)a_n} \sum_{i_{N-1}=(2j_{N-1}+1)a_n+1}^{2(j_{N-1}+1)a_n} \sum_{i_N=2j_N a_n+1}^{(2j_N+1)a_n} \Delta_{\mathbf{i}};$$

$$U(4,n,\mathbf{j},x_0) = \sum_{\substack{i_k=2j_k a_n+1\\ k=1,\dots,N-2}}^{(2j_k+1)a_n} \sum_{i_{N-1}=(2j_{N-1}+1)a_n+1}^{2(j_{N-1}+1)a_n} \sum_{i_N=(2j_N+1)a_n+1}^{2(j_N+1)a_n} \Delta_{\mathbf{i}}, \tag{8}$$

and so on. Finally note that $U(2^{N-1},n,\mathbf{j},x_0) = \sum_{\substack{i_k=2j_k a_n+1\\ k=1,\dots,N-1}}^{(2j_k+1)a_n} \sum_{i_N=2j_N a_n+1}^{2(j_N+1)a_n} \Delta_{\mathbf{i}}$; $\quad U(2^N,n,\mathbf{j},x_0) = \sum_{\substack{i_k=2j_k a_n+1\\ k=1,\dots,N}}^{(2j_k+1)a_n} \Delta_{\mathbf{i}}$.

For each integer $1 \le i \le 2^N$, define $E(n,i,x_0) = \sum_{\substack{j_k=0\\ k=1,\dots,N}}^{r_k-1} U(i,n,\mathbf{j},x_0)$, then one may write $G_n(x_0) = \sum_{i=1}^{2^{N+1}} E(n,i,x_0)$. Without loss of generality, we will show the result for $i = 1$.

Because $K$ is bounded, we get $U(1,n,\mathbf{j},x_0) \le \frac{a_n^N\|K\|_\infty}{B_{n,d}} \frac{1}{nh_n^d}$. Consider the sequences $\lambda_n = \left\lfloor \left(nh_n^d \log n\right)^{1/2} \right\rfloor$, and $r_i = n_i/(2a_n)$, $\quad i = 1,\dots,N$. We deduce that, for $n$ large enough,

$$\lambda_n\,|U(1,n,\mathbf{j},x_0)| = O\left(\frac{\log n}{\sqrt{nh_n^d}}\right). \tag{9}$$

Enumerate the random variables $U(1,n,\mathbf{j},x)$ in $E(n,1,x)$ in an arbitrary manner and refer to them as $\hat{U}_1,\dots,\hat{U}_M$, $M = r_1 \dots r_N$. By Markov's inequality and Lemma 4.5 of [6], there exist r.v.'s $\tilde{U}_1,\dots,\tilde{U}_M$ independent of $\hat{U}_1,\dots,\hat{U}_M$ and such that $\hat{U}_i = \hat{U}_i(x_0)$ has the same distribution as $\tilde{U}_i = \tilde{U}_i(x_0)$ and:

$$\mathbb{P}\left(\left|\hat{U}_i - \tilde{U}_i\right| > \epsilon\right) \le 18\left(\frac{\left\|\hat{U}_i\right\|_\infty}{\epsilon}\right)\psi(n, a_n^N)\phi(a_n). \tag{10}$$

Let $\epsilon_n = \eta\sqrt{\frac{\log n}{nh_n^d}}$, where $\eta$ is a positive number. Obviously

$$\mathbb{P}(|E(n,1,x_0)| > \epsilon_n) \le \mathbb{P}\left(\sum_{j=1}^{M}\left|\tilde{U}_i\right| > \frac{\epsilon_n}{2}\right) + \mathbb{P}\left(\sum_{j=1}^{M}\left|\hat{U}_i - \tilde{U}_i\right| > \frac{\epsilon_n}{2}\right). \quad (11)$$

Using the independence of the $\tilde{U}_i$'s and applying Bernstein inequality, we get

$$\mathbb{P}\left(\sum_{j=1}^{M}\left|\tilde{U}_i\right| > \frac{\epsilon_n}{2}\right) \le 2\exp\left(-\frac{\lambda_n\epsilon_n}{2} + \frac{\lambda_n^2}{4}\sum_{i=1}^{M}\mathbb{E}\left(\hat{U}_i^2(x_0)\right)\right) \le 2\exp\left(-\frac{\eta\log n}{2} + \frac{\lambda_n^2}{4}\sum_{i=1}^{M}\mathbb{E}\left(\hat{U}_i^2(x_0)\right)\right).$$

Clearly, $\frac{\lambda_n^2}{4}\sum_{i=1}^{M}\mathbb{E}\left(\hat{U}_i^2(x_0)\right) \le \frac{\lambda_n^2}{4}\mathrm{Var}(p_n(x_0)) \le \frac{\log(n)}{4}nh_n^d\mathrm{Var}(p_n(x_0))$.

From (4), we deduce that for $n$ large enough, $nh_n^d\mathrm{Var}(p_n(x_0)) \le p(x_0)B_d^{-1}\int_{\mathbb{R}^d} K^2(u)\mathrm{d}u$. It follows that

$$\frac{\lambda_n^2}{4}\sum_{i=1}^{M}\mathbb{E}\left(\hat{U}_i^2(x_0)\right) \le \log(n)p(x_0)\frac{B_d^{-1}}{4}\int_{\mathbb{R}^d} K^2(u)\mathrm{d}u.$$

Hence $\mathbb{P}\left(\sum_{j=1}^{M}\left|\tilde{U}_i\right| > \frac{\epsilon_n}{2}\right) \le n^{-\delta}$, where $\delta > 1$, for $\eta$ large enough. Concerning the second term in the right-hand-side of (11), from (10) one has $\left\|\hat{U}_i(x_0)\right\|_\infty \le \left(\frac{a_n^N\|K\|_\infty}{nh_n^d}\right)$. We get, by Markov's inequality:

$$\mathbb{P}\left(\sum_{i=1}^{M}\left|\hat{U}_i - \tilde{U}_i\right| > \frac{\epsilon_n}{2}\right) \le CM\left(\frac{a_n^N\|K\|_\infty}{nh_n^d}\right)\epsilon_n^{-1}\psi(n,a_n^N)\phi(a_n),$$

then $\mathbb{P}\left(\sum_{i=1}^{M}\left|\hat{U}_i - \tilde{U}_i\right| > \frac{\epsilon_n}{2}\right) \le C\left(\frac{1}{h_n^d}\right)\epsilon_n^{-1}\psi(n,a_n^N)\left(\phi(a_n)\right) \le C\left(\frac{1}{h_n^d}\right)\epsilon_n^{-1}a_n^N a_n^{-\theta}$.

Let $a_n = \left(\frac{\log n}{nh_n^d}\right)^{-1/(2N)}$, then $\mathbb{P}\left(\sum_{i=1}^{M}\left|\hat{U}_i - \tilde{U}_i\right| > \frac{\epsilon_n}{2}\right) \le Cn^{\frac{2N-\theta}{2N}}\log n^{\frac{\theta-2N}{2N}}h_n^{\frac{-d\theta}{2N}} = C\beta_n$.

By hypothesis, we have $nu_n\beta_n = \left(n\log n^{\frac{\theta-2N}{4N-\theta}}h_n^{\frac{d\theta}{\theta-4N}}u_n^{\frac{2N}{4N-\theta}}\right)^{\frac{4N-\theta}{2N}} \to 0$, then Theorem 2 follows by the Borel–Cantelli lemma. □

## References


[1] A. Amiri, Sur une famille paramétrique d'estimateurs séquentiels de la densité pour un processus fortement mélangeant, C. R. Acad. Sci. Paris, Sér. I 347 (2009) 309–314.
[2] J. Baek, H.-Y. Liang, Asymptotic normality of recursive density estimates under some dependence assumptions, Metrika 60 (2004) 155–166.
[3] G. Biau, B. Cadre, Nonparametric spatial prediction, Stat. Inference Stoch. Process. 7 (2004) 327–349.
[4] D. Blanke, D. Bosq, Inference and Prediction in Large Dimensions, Wiley Series in Probability and Statistics, Dunod, 2007.
[5] M. Carbon, L.-T. Tran, B. Wu, Kernel density estimation for random fields: the $L_1$ theory, J. Nonparametr. Stat. 6 (1997) 157–170.
[6] M. Carbon, L.-T. Tran, B. Wu, Kernel density estimation for random fields, Stat. Probab. Lett. 36 (1997) 115–125.
[7] J.-P. Chiles, P. Delfiner, Statistics for Spatial Data, Wiley, New York, 1999.
[8] T. Conley, GMM estimation with cross sectional dependence, J. Econom. 92 (1999) 1–45.
[9] N.-A.-C. Cressie, Statistics for Spatial Data, Wiley Series in Probability and Mathematical Statistics, Wiley-Interscience, New York, 1991.
[10] S. Dabo-Niang, A.-F. Yao, Kernel regression estimation for continuous spatial processes, Math. Methods Stat. 16 (2007) 298–317.
[11] H.-I. Davies, E.-J. Wegman, Remarks on some recursive estimators of a probability density, Ann. Stat. 7 (1979) 316–327.
[12] P. Deheuvels, Sur une famille d'estimateurs de la densité d'une variable aléatoire, C. R. Acad. Sci. Paris 276 (1974) 1013–1015.
[13] E. Isogai, A Berry–Esseen-type bound for recursive estimators of a density and its derivatives, J. Stat. Plan. Inference 40 (1994) 1–4.
[14] E. Masry, Recursive probability density estimation for weakly dependent stationary processes, IEEE Trans. Inf. Theory 32 (1986) 254–267.
[15] K.-A. Mezhoud, et al., Recursive kernel estimation of the density under $\eta$-dependence, J. Korean Stat. Soc. (2014) 1–12.
[16] R.-X. Mugisha, M. Samanta, On the class of estimates of the probability function and mode based on random number of observations, Calcutta Stat. Assoc. Bull. 117 (1981) 23–40.
[17] P.-M. Robinson, Asymptotic theory for nonparametric regression with spatial data, J. Econom. 165 (2011) 5–19.
[18] W.-F. Stout, Almost Sure Convergence, Academic Press, New York, 1974, pp. 120–121.
[19] L.-T. Tran, Kernel density estimation on random fields, J. Multivar. Anal. 34 (1990) 37–53.
[20] T. Wagner, C. Wolverton, Recursive estimates of probability density, IEEE Trans. Syst. Sci. Cybern. 5 (1969) 307–308.
[21] H. Yamato, Sequential estimation of a continuous probability density function and mode, Bull. Math. Stat. 14 (1972) 1–12.