Large and moderate deviation principles for recursive kernel estimators of a regression function for spatial data defined by stochastic approximation method

Yousri Slaoui, Salim Bouzebda

Large and moderate deviation principles for recursive kernel estimators of a regression function for spatial data defined by stochastic approximation method

Salim Bouzebda[1*] and Yousri Slaoui[2†]
[1]LMAC, Université de Technologie de Compiègne, France
[2]Univ. Poitiers, Lab. Math. et Appl., Futuroscope Chasseneuil, France

March 12, 2019

Abstract

In the present paper, we are mainly concerned with a family of kernel type estimators based upon spatial data. More precisely, we establish large and moderate deviations principles for the recursive kernel estimators of a regression function for spatial data defined by the stochastic approximation algorithm.
AMS 2010 subject classifications: 62G08, 62L20, 60F10.
Keywords: Nonparametric regression; Stochastic approximation algorithm; Large and Moderate deviations principles.

1 Introduction

Spatial data, collected on measurement sites in a variety of fields and the statistical treatment, typically arise in various fields of research, including econometrics, epidemiology, environmental science, image analysis, oceanography, meteorology, geostatistics and many others. For good sources of references to research literature in this area along with statistical applications consult Ripley (1981), Rosenblatt (1985), Guyon (1995) and Cressie (2015) and the references therein. In the context of nonparametric estimation for spatial data, the existing papier are mainly concerned with the estimation of a probability density and regression functions, we cite some key references Tran (1990), Tran and Yakowitz (1993), Carbon *et al.* (1997), Biau and Cadre (2004), Dabo-Niang and Yao (2013), Dabo-Niang *et al.* (2016) and the references therein. In the works of Amiri *et al.* (2016) and Bouzebda and Slaoui (2019a,b), recursive versions of non-parametric density estimation for spatial data are investigated. These results are extended to a more general setting in Bouzebda and Slaoui (2018), which includes the previous works as particular cases, this generalization is far from being trivial. We start by giving some notation and definitions that are needed for the forthcoming sections. We consider a spatial process $(\mathbf{Z_i} = (\mathbf{X_i}, Y_\mathbf{i}) \in \mathbb{R}^d \times \mathbb{R} : \mathbf{i} \in \mathbb{Z}^N)$ defined over some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the same distribution as $(\mathbf{X}, Y)$ having unknown density $g_{\mathbf{X},Y}(\cdot)$ relatively to the Lebesgue measure on $\mathbb{R}^{d+1}$. The density function of $\mathbf{X}$ on $\mathbb{R}^d$ is $g_\mathbf{X}(\cdot)$. In this paper, we are interested in the following regression model

$$Y_\mathbf{i} = r(\mathbf{X_i}) + \varepsilon_\mathbf{i},$$

where $r(\mathbf{x}) = \mathbb{E}(Y|\mathbf{X} = \mathbf{x})$ is an unknown function, with real values, the noise $\varepsilon_\mathbf{i}$ is centered and independent of $\mathbf{X_i}$. The process is observed over the domain $\mathcal{I}_\mathbf{n} = \{\mathbf{i} = (i_1, \ldots, i_N) : 1 \le i_k \le n_k, k = 1, \ldots, N\}$. For simplicity we restrict ourselves to rectangles as domains for the observations. We denote $\mathbf{n} = (n_1, \ldots, n_N)$; let $\widehat{\mathbf{n}} := n_1 \times \cdots \times n_N$ be the sample size. We consider the following family of estimators investigated in Bouzebda and Slaoui (2018). Given any two continuous functions $c_f(\cdot)$ and $d_f(\cdot)$, define

$$\widehat{\Psi}_{\mathbf{n},h_\mathbf{n}}(\mathbf{x}, f, K) = \Pi_\mathbf{n} \sum_{\mathbf{i}\in\mathcal{I}_n} \Pi_\mathbf{i}^{-1} \gamma_\mathbf{i} h_\mathbf{i}^{-d} \left\{(c_f(\mathbf{x}) f(Y_\mathbf{i}) + d_f(\mathbf{x})) K(h_\mathbf{i}^{-1}(\mathbf{x} - \mathbf{X_i}))\right\}, \tag{1.1}$$

where $(\gamma_\mathbf{n})$ is a nonrandom positive sequence tending to zero as $\widehat{\mathbf{n}} \to \infty$, $(h_\mathbf{n})$ is a nonrandom positive sequence tending to zero as $\widehat{\mathbf{n}} \to \infty$, called bandwidth and $\Pi_\mathbf{n} = \prod_{\mathbf{i}\in\mathcal{I}_n} (1 - \gamma_\mathbf{i})$. Here and elsewhere, $f(\cdot)$ denotes a specified measurable function, which is assumed to be bounded on each compact subinterval of $\mathbb{R}$. The process in (1.1) was considered, in non recursive setting, by Einmahl and Mason (2000, 2005). For convenience, we treat the observations sites as an array that is $\mathcal{I}_n = \{s_j : j = 1, \ldots, n\}$. By enumerating the sites, one may rewrite $\widehat{\Psi}_{\mathbf{n},h_\mathbf{n}}(\mathbf{x}, f, K)$ in the following way

$$\widehat{\Psi}_{\mathbf{n},h_\mathbf{n}}(\mathbf{x}, f, K) = \Pi_\mathbf{n} \sum_{j=1}^{n} \Pi_j^{-1} \gamma_{s_j} h_{s_j}^{-d} \left\{(c_f(\mathbf{x}) f(Y_{s_j}) + d_f(\mathbf{x})) K\left(h_{s_j}^{-1}\left(\mathbf{x} - \mathbf{X}_{s_j}\right)\right)\right\}, \tag{1.2}$$

where $\Pi_\mathbf{n} = \prod_{j=1}^{n} \left(1 - \gamma_{s_j}\right)$. This recursive property is particularly useful when the number of the spatial sites increase on space since $\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K)$ can be easily updated with each additional observation. In fact, if X_{s_n} is a new observation of the process

*e-mail adress: salim.bouzebda@utc.fr
†e-mail adress: Yousri.Slaoui@math.univ-poitiers.fr

at a site s_n added to $\mathcal{I}_{n-1}$, the estimators $\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K)$ can be updated recursively by the relation (1.2). From a practical point of view, this arrangement provides important savings in computational time and storage memory which a consequence of the fact that the estimate updating is independent of the history of the data. The main drawback of the classical kernel kernel estimator is the use of all data at each step of estimation. From a theoretical point of view, the main advantage of the investigation of such processes is that we can prove almost sure consistency with exact rate for several kernel-type estimators simultaneously. It is worth noting that the quantity $\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K)$ includes as particular cases : the kernel type density estimator, the Nadaraya Watson estimator and the kernel type estimator of the conditional distribution, we may refer to Einmahl and Mason (2000, 2005) for more details. In this sense, the present paper extends, in non trivial, way some previous results by considering a general kernel-type estimators given in (1.2), see Remark 2.7, below.

Recently, large and moderate deviations results have been proved for the recursive density estimators defined by stochastic approximation method in Slaoui (2013) and for the averaged stochastic approximation method for the estimation of a regression function in Slaoui (2015a). The purpose of this paper is to establish large and moderate deviations principles for the recursive regression estimator for spatial data $\widehat{\Psi}_{\mathbf{n},h_\mathbf{n}}(\mathbf{x}, f, K)$, The present paper completes the investigation of Bouzebda and Slaoui (2018) and extends in non trivial way the works Slaoui (2013) and Slaoui (2015a). In Bouzebda and Slaoui (2018), we have obtained the central limit theorem and strong pointwise convergence rate for the nonparametric recursive general kernel-type estimators under some mild conditions. We have investigated the MISE of the proposed estimators and provide the optimal bandwidth. The aim of the present paper is quite different from the Bouzebda and Slaoui (2018), since we are interested in LDP and MDP type results. Mokkadem *et al.* (2006, 2008) and Slaoui (2015a) have established some results about LDP and MDP for some nonparametric estimators. Their results are not directly applicable here since we are considering more general setting in the present work. Their results are not only useful in their own right but essential in some steps of our proofs. To the best of our knowledge, the results presented here, respond to a problem that has not been studied systematically up to the present, which was the basic motivation of the paper.
Let us first recall that a $\mathbb{R}^m$-valued sequence $(Z_n)_{n\geq 1}$ satisfies a large deviations principle (LDP) with speed (ν_n) and good rate function I if :

1. (ν_n) is a positive sequence such that $\lim_{n\to\infty} \nu_n = \infty$;
2. $I : \mathbb{R}^m \to [0, \infty]$ has compact level sets;
3. for every borel set $B \subset \mathbb{R}^m$,

$$-\inf_{x\in \overset{\circ}{B}} I(x) \quad \leq \quad \liminf_{n\to\infty} \nu_n^{-1} \log \mathbb{P}[Z_n \in B] \leq \limsup_{n\to\infty} \nu_n^{-1} \log \mathbb{P}[Z_n \in B] \leq -\inf_{x\in \overline{B}} I(x),$$

   where $\overset{\circ}{B}$ and $\overline{B}$ denote the interior and the closure of B respectively. Moreover, let (v_n) be a nonrandom sequence that goes to infinity; if $(v_n Z_n)$ satisfies a LDP, then (Z_n) is said to satisfy a moderate deviations principle (MDP).

The first purpose of this paper is to establish pointwise LDP for the recursive kernel estimators of a regression function for spatial data (1.2). It turns out that the rate function depend on the choice of the stepsize (γ_{s_n}); We focus in the first part of this paper on the following two special cases : (1) $(\gamma_{s_n}) = \left(n^{-1}\right)$ and (2) $(\gamma_{s_n}) = \left(h_{s_n}^d \left(\sum_{k=1}^n h_{s_k}^d\right)^{-1}\right)$, the first stepsize belongs to the subclass of recursive kernel estimators which have a minimum $MISE$ and the second stepsize belongs to the subclass of recursive kernel estimators which have a minimum variance, we may refer to Bouzebda and Slaoui (2018).

We show that using the stepsize $(\gamma_{s_n}) = \left(n^{-1}\right)$ and the bandwidth $(h_{s_n}) \equiv (cn^{-a})$ with $c > 0$ and $a \in]0, 1/d[$, the sequence $\left(\widehat{\Psi}_{\mathbf{n},h_\mathbf{n}}(\mathbf{x}, f, K) - \Psi(\mathbf{x}, f)\right)$ satisfies a LDP with speed $\left(nh_{s_n}^d\right)$ and the rate function defined as follows:

$$I_{a,\mathbf{x}}(t) = \sup_{u\in\mathbb{R}} \{ut - \psi_{a,\mathbf{x}}(u)\}, \tag{1.3}$$

which is the Fenchel-Legendre transform of the function $\psi_{a,\mathbf{x}}$ defined as follows:

$$\psi_{a,\mathbf{x}}(u) = g_{\mathbf{X}}^{-1}(\mathbf{x}) \int_{\mathbb{R}^d\times\mathbb{R}\times[0,1]} \left[\exp\left(ut^{ad} K(\mathbf{z})(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x}))\right) - 1\right] g_{\mathbf{X},Y}(\mathbf{x}, y)\, d\mathbf{z}dydt. \tag{1.4}$$

Moreover, we show that using the stepsize $(\gamma_{s_n}) = \left(h_{s_n}^d \left(\sum_{k=1}^n h_{s_k}^d\right)^{-1}\right)$ and more general bandwidths defined as $h_{s_n} = h(s_n)$ for all n, where h is a regulary varing function with exponent $(-a)$, $a \in]0, 1/d[$. We prove that the sequence $\left(\widehat{\Psi}_{\mathbf{n},h_\mathbf{n}}(\mathbf{x}, f, K) - \Psi(\mathbf{x}, f)\right)$ satisfies a LDP with speed $\left(nh_{s_n}^d\right)$ and the following rate function:

$$I_{\mathbf{x}}(\mathbf{t}) = \sup_{u\in\mathbb{R}} \{ut - \psi_{\mathbf{x}}(u)\},$$

which is the Fenchel-Legendre transform of the function $\psi_{\mathbf{x}}$ defined as follows:

$$\psi_{\mathbf{x}}(u) = g_{\mathbf{X}}^{-1}(\mathbf{x}) \int_{\mathbb{R}^d\times\mathbb{R}} \left[\exp\left(uK(\mathbf{z})(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x}))\right) - 1\right] g_{\mathbf{x},y}(\mathbf{x}, y)\, d\mathbf{z}dy. \tag{1.5}$$

Our second purpose in this paper is to provide pointwise MDP for the proposed regression estimator for spatial data defined by the stochastic approximation algorithm (1.2). In this case, we consider more general stepsizes defined as $\gamma_{s_n} = \gamma(s_n)$ for all n, where γ is a regularly function with exponent $(-\alpha)$, $\alpha \in]1/2, 1]$. Throughout this paper we will use the following notation:

$$\xi = \lim_{n\to+\infty} (n\gamma_{s_n})^{-1}. \tag{1.6}$$

We show that using the stepsize defined as $\gamma_{s_n} = \gamma(n)$ for all n, where γ is a regularly function with exponent $(-\alpha)$, $\alpha \in (\frac{1}{2}, 1]$, and for any positive sequence (v_n) satisfying

$$\lim_{n\to\infty} v_n = \infty, \qquad \lim_{n\to\infty} \frac{\gamma_{s_n} v_n^2}{h_{s_n}} = 0 \quad \text{and} \quad \lim_{n\to\infty} v_{s_n} h_{s_n}^2 = 0$$

and the bandwidths defined as $h_{s_n} = h(n)$ for all n, where h is a regularly function with exponent $(-a)$, $a \in (0, \frac{\alpha}{d})$, we prove that the sequence $v_n \left(\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K) - \Psi(\mathbf{x}, f)\right)$ satisfies a LDP of speed $\left(h_{s_n}/\left(\gamma_{s_n} v_{s_n}^2\right)\right)$ and rate function $J_{a,\alpha,\mathbf{x}}(.)$ defined by

$$\begin{cases} if \quad g_{\mathbf{X}}(\mathbf{x}) \neq 0, \quad J_{a,\alpha,\mathbf{x}} : t \to \dfrac{t^2\left(2 - (\alpha - ad)\,\xi\right)}{2V(\mathbf{x}, f) \displaystyle\int_{\mathbb{R}^d} K^2(z)\, dz} \\ if \quad g_{\mathbf{X}}(\mathbf{x}) = 0, \quad J_{a,\alpha,\mathbf{x}}(0) = 0 \quad and \quad J_{a,\alpha,\mathbf{x}}(t) = +\infty \quad \text{for} \quad t \neq 0. \end{cases} \tag{1.7}$$

where $V(\mathbf{x}, f)$ is given in (2.2).

The layout of the present article is as follows. Section 2 is devoted to the main results of the present work. To avoid interrupting the flow of the presentation, all mathematical developments are relegated to Section 3.

2 Assumptions and main results

Let us first define the class of positive sequences that will be used in the statement of our assumptions.

Definition 2.1 *Let $\gamma \in \mathbb{R}$ and $(v_n)_{n\geq 1}$ be a nonrandom positive sequence. We say that $(v_n) \in \mathcal{GS}(\gamma)$ if*

$$\lim_{n\to\infty} n\left[1 - \frac{v_{n-1}}{v_n}\right] = \gamma. \tag{2.1}$$

Condition (2.1) was introduced by Galambos and Seneta (1973) to define regularly varying sequences (see also Bojanic and Seneta (1973)); it was used in Mokkadem and Pelletier (2007) in the context of stochastic approximation algorithms. Typical sequences in $\mathcal{GS}(\gamma)$ are, for $b \in \mathbb{R}$, $n^\gamma (\log n)^b$, $n^\gamma (\log\log n)^b$, and so on. To unburden our notation a bit, we let

$$\begin{aligned} L(\mathbf{x}, f) &= \mathbb{E}\left(f(Y) \mid \mathbf{X} = \mathbf{x}\right) g_{\mathbf{X}}(\mathbf{x}), \quad \Psi(\mathbf{x}, f) = c_f(\mathbf{x})\, L(\mathbf{x}, f) + d_f(\mathbf{x}) g_{\mathbf{X}}(\mathbf{x}), \\ V(\mathbf{x}, f) &= c_f^2(\mathbf{x})\, L\left(\mathbf{x}, f^2\right) + d_f^2 g_{\mathbf{x}}(\mathbf{x}) + 2c_f(\mathbf{x})\, d_f(\mathbf{x})\, L(\mathbf{x}, f), \\ g_{ij}^{(2)}(\mathbf{x}) &= \frac{\partial^2 g_{\mathbf{X}}}{\partial x_i \partial x_j}(\mathbf{x}), \quad L_{ij}^{(2)}(\mathbf{x}, f) = \frac{\partial^2 L}{\partial x_i \partial x_j}(\mathbf{x}, f), \quad R(K) = \int_{\mathbb{R}^d} K^2(\mathbf{z})\, d\mathbf{z}. \end{aligned} \tag{2.2}$$

2.1 Pointwise LDP for $\widehat{\Psi}_{n,h_n}$ in the case when $(\gamma_{s_n}) = (n^{-1})$

2.1.1 Choices of (γ_{s_n}) minimizing the MISE of $\widehat{\Psi}_{n,h_n}$

In order to establish pointwise LDP for $\widehat{\Psi}_{n,h_n}$ in the special case when $(\gamma_{s_n}) = \left(n^{-1}\right)$, we need the following assumptions. The assumptions to which we shall refer are the following:

$(L1)$ $K : \mathbb{R}^d \to \mathbb{R}$ is a continuous, bounded function satisfying $\int_{\mathbb{R}^d} K(\mathbf{z})\, d\mathbf{z} = 1$, and, for all $j \in \{1, \ldots, d\}$, $\int_{\mathbb{R}} z_j K(\mathbf{z})\, dz_j = 0$ and $\int_{\mathbb{R}^d} z_j^2 \|K(\mathbf{z})\|\, d\mathbf{z} < \infty$.

$(L2)$ (i) $(h_{s_n}) = (cn^{-a})$ with $a \in (0, 1/d)$ and $c > 0$.
(ii) $(\gamma_{s_n}) = \left(n^{-1}\right)$.

$(L3)$ (i) $g_{\mathbf{X},Y}(s, t)$ is twice continuously differentiable with respect to s.
(ii) For $q \in \{0, 1, 2\}$, $s \mapsto \int_{\mathbb{R}} t^q g_{\mathbf{X},Y}(\mathbf{s}, t)\, dt$ is a bounded function continuous at $\mathbf{s} = \mathbf{x}$.
For $q \in [2, 3]$, $s \mapsto \int_{\mathbb{R}} |t|^q g_{\mathbf{X},Y}(\mathbf{s}, t)\, dt$ is a bounded function.
(iii) For $q \in \{0, 1\}$, $\int_{\mathbb{R}} |t|^q \left|\frac{\partial g_{\mathbf{X},Y}}{\partial x}(\mathbf{x}, t)\right| dt < \infty$, and $\mathbf{s} \mapsto \int_{\mathbb{R}} t^q \frac{\partial^2 g_{\mathbf{X},Y}}{\partial s^2}(\mathbf{s}, t)\, dt$ is a bounded function continuous at $\mathbf{s} = \mathbf{x}$.
(iv) For any $i, j \in \{1, \ldots, n\}$ such that $s_i \neq s_j$, the random vector $\left(X_{s_i}, X_{s_j}\right)$ and $\left(Z_{s_i}, Z_{s_j}\right)$ admit a density $f_{s_i,s_j}(\cdot)$ and $g_{s_i,s_j}(\cdot)$ such that $\sup_{s_i\neq s_j} \left\|f_{s_i,s_j}\right\| < \infty$, and $\sup_{s_i\neq s_j} \left\|g_{s_i,s_j}\right\| < \infty$.

(L4) For any $u \in \mathbb{R}$, $\mathbf{t} \to \int_{\mathbb{R}} \exp\left(uf\left(y\right)\right) g_{\mathbf{X},Y}\left(\mathbf{t}, y\right) dy$ is continuous at $\mathbf{x}$ and bounded.

(L5) (i) The field $\left(Z_{s_i}\right)_{1\leq i\leq n}$ is α-mixing: there exists a function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ with $\phi\left(t\right)$ goes to zero as t goes to infinity, such that for E, $F \subset \mathbb{R}^2$ with finite cardinals $Card\left(E\right)$, $Card\left(F\right)$

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

where $\sigma\left(E\right) = \left\{Z_{\mathbf{i}}, \mathbf{i}\in E\right\}$ and $\sigma\left(F\right) = \left\{Z_{\mathbf{i}}, \mathbf{i}\in F\right\}$, $dist\left(E, F\right)$ is the Euclidean distance between E and F and $\psi\left(.\right)$ is a positive symmetric function nondecreasing in each variable. The functions ϕ and ψ are such that $\phi\left(i\right) \leq Ci^{-\theta}$ and $\psi\left(n, m\right) \leq C\min\left(m, n\right)$.

(ii) $\sum_{k=0}^{\infty}\left(k+1\right)^2 \alpha_n^{\frac{\delta}{4+\delta}}\left(k\right) < c$ for some $c, \delta > 0$ and all n, where

$$\alpha_n\left(k\right) = \alpha_n\left(\mathbf{Z}, k\right) = \sup_{A\in\mathcal{F}_{-\infty}^n, B\in\mathcal{F}_{n+k}^{\infty}} \left|\mathbb{P}\left(A\cap B\right) - \mathbb{P}\left(A\right)\mathbb{P}\left(B\right)\right|,$$

with $\mathbf{Z} = \left\{Z_{s_i}\right\}_{i=1}^n$, $\mathcal{F}_m^n$ denote the σ-algebra generated by $\left\{Z_{s_i}\right\}_{i=m}^n$.

Assumption $(\mathbf{L1})$ on the kernel is widely used in the recursive and the nonrecursive framework for the functional estimation. Assumption $(\mathbf{L2})$ on the stepsize and the bandwidth was used in the recursive framework for the estimation of the density function (see Mokkadem *et al.* (2009a) and Slaoui (2013, 2014a)) and for the estimation of the distribution function (see Slaoui (2014b)), and Slaoui (2015a,b, 2016)) for the estimation of the regression function. Assumption $(\mathbf{L3})$ on the density of Z was used in Mokkadem *et al.* (2009b) and Slaoui (2015a,b, 2016)). Assumption $(\mathbf{L5})$(i) are classical in nonparametric estimation in the spatial literature (see Amiri *et al.* (2016)). However, assumption $(\mathbf{L5})$(ii) was considered in Ekström (2014) to establish a general central limit theorem for strong mixing sequences, see Rosenblatt (1956). The proof of the following comment is given in Mokkadem *et al.* (2008).

Comment

Assumption $(\mathbf{L2})(iii)$ on the limit of $(n\gamma_{s_n})$ as n goes to infinity is standard in the framework of stochastic approximation algorithms. It implies in particular that the limit of $\left(\left[n\gamma_{s_n}\right]^{-1}\right)$ is finite.

Assumption $(\mathbf{L4})$ implies that $\forall m \geq 0, \forall \rho \geq 0$

$$\text{the function} \quad t \mapsto \int_{\mathbb{R}} \left|f\left(y\right)\right|^m \exp\left(\rho\left|f\left(y\right)\right|\right) g_{\mathbf{X},Y}\left(t, y\right) dy \quad \text{is bounded.} \tag{2.3}$$

Before stating our results, we set $S_+ = \left\{\mathbf{x}\in\mathbb{R}^d; K\left(\mathbf{x}\right) > 0\right\}$ and $S_- = \left\{x\in\mathbb{R}^d; K\left(\mathbf{x}\right) < 0\right\}$ and for fixed $x \in \mathbb{R}$

$$T_+ = \left\{y\in\mathbb{R}; c_f\left(\mathbf{x}\right)f\left(y\right) + d_f\left(\mathbf{x}\right) > 0\right\} \quad \text{and} \quad T_- = \left\{y\in\mathbb{R}; c_f\left(\mathbf{x}\right)f\left(y\right) + d_f\left(\mathbf{x}\right) < 0\right\}$$

Moreover, we set $O_+ = \left(S_+\cap T_+\right)\cup\left(S_-\cap T_-\right)$ and $O_- = \left(S_+\cap T_-\right)\cup\left(S_-\cap T_+\right)$

The following proposition gives the properties of the functions $\psi_{a,\mathbf{x}}$ and $I_{a,\mathbf{x}}$; in particular, the behavior of the rate function $I_{a,\mathbf{x}}$.

Proposition 2.2 (Properties of $\psi_{a,\mathbf{x}}$ and $I_{a,\mathbf{x}}$)

Let λ be the Lebesgue measure on $\mathbb{R}^d$ and let Assumptions $(L1)$ and $(L4)$ hold.

(i) *$\psi_{a,\mathbf{x}}$ is strictly convex, twice continuously differentiable on $\mathbb{R}$, and $I_{a,\mathbf{x}}$ is a good rate function on $\mathbb{R}$.*

(ii) *If $\lambda\left(O_-\right) = 0$, $I_{a,\mathbf{x}}\left(t\right) = +\infty$, when $t < 0$, and*

$$I_{a,\mathbf{x}}\left(0\right) = \begin{cases} \lambda\left(S_+\right) & \text{if} \quad \lambda\left(S_+\cap T_+\right) > 0 \\ \lambda\left(S_-\right) & \text{if} \quad \lambda\left(S_-\cap T_-\right) > 0 \end{cases}$$

$I_{a,\mathbf{x}}$ is strictly convex on $\mathbb{R}$ and continuous on $\left(0, +\infty\right)$, and for any $t > 0$

$$I_{a,\mathbf{x}}\left(t\right) = t\left(\psi'_{a,\mathbf{x}}\right)^{-1}\left(t\right) - \psi_{a,\mathbf{x}}\left(\left(\psi'_{a,\mathbf{x}}\right)^{-1}\left(t\right)\right), \tag{2.4}$$

(iii) If $\lambda\left(O_-\right) > 0$, then $I_{a,\mathbf{x}}$ is finite and strictly convex on $\mathbb{R}$ and (2.4) holds for any $t \in \mathbb{R}$.

The following Theorem gives the pointwise LDP for $\widehat{\Psi}_{n,h_n}$ in the case when $\left(\gamma_{s_n}\right) = \left(n^{-1}\right)$.

Theorem 2.3 (Pointwise LDP for $\widehat{\Psi}_{n,h_n}$ in the case when $\left(\gamma_{s_n}\right) = \left(n^{-1}\right)$)

Let Assumptions $(L1)$ and $(L4)$ hold and assume that $g_{X,Y}$ is continuous at $(\mathbf{x}, y)$. Then, the sequence $(\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K) - \Psi(\mathbf{x}, f))$ satisfies a LDP with speed $\left(nh_{s_n}^d\right)$ and rate function defined as follows:

$$I_{a,\mathbf{x}}\left(t\right) = t\left(\psi'_{a,\mathbf{x}}\right)^{-1}\left(t\right) - \psi_{a,\mathbf{x}}\left(\left(\psi'_{a,\mathbf{x}}\right)^{-1}\left(t\right)\right),$$

where $\psi_{a,\mathbf{x}}$ is defined in (1.4).

2.1.2 Choices of (γ_{s_n}) minimizing the Variance of $\widehat{\Psi}_{n,h_n}$

In order to establish pointwise LDP for $\widehat{\Psi}_{n,h_n}$ in this case, we assume that.

($L2'$) $i)$ $(h_{s_n}) \in \mathcal{GS}(-a)$ with $a \in (0, 1/d)$. $ii)$ $(\gamma_{s_n}) = \left(h_{s_n}^d \left(\sum_{k=1}^n h_{s_k}^d\right)^{-1}\right)$.

The following proposition gives the properties of the functions $\psi_{\mathbf{x}}$ and $I_{\mathbf{x}}$; in particular, the behavior of the rate function $I_{\mathbf{x}}$.

Proposition 2.4 (Properties of $\psi_{\mathbf{x}}$ and $I_{\mathbf{x}}$)
Let λ be the Lebesgue measure on $\mathbb{R}^d$ and let Assumptions Let Assumptions $(L1)$, $(L2')$, $(L3)$ and $(L4)$.

(i) *$\psi_{\mathbf{x}}$ is strictly convex, twice continuously differentiable on $\mathbb{R}$, and $I_{\mathbf{x}}$ is a good rate function on $\mathbb{R}$.*

(ii) *If $\lambda(O_-) = 0$, $I_{\mathbf{x}}(t) = +\infty$, when $t < 0$, and*

$$I_{\mathbf{x}}(0) = \begin{cases} \lambda(S_+) & \text{if} \quad \lambda(S_+ \cap T_+) > 0 \\ \lambda(S_-) & \text{if} \quad \lambda(S_- \cap T_-) > 0 \end{cases}$$

$I_{\mathbf{x}}$ is strictly convex on $\mathbb{R}$ and continuous on $(0, +\infty)$, and for any $t > 0$

$$I_{\mathbf{x}}(t) = t\left(\psi'_{\mathbf{x}}\right)^{-1}(t) - \psi_{\mathbf{x}}\left(\left(\psi'_{\mathbf{x}}\right)^{-1}(t)\right), \tag{2.5}$$

(iii) *If $\lambda(O_-) > 0$, then $I_{\mathbf{x}}$ is finite and strictly convex on $\mathbb{R}$ and (2.5) holds for any $t \in \mathbb{R}$.*

The following Theorem gives the pointwise LDP for $\widehat{\Psi}_{n,h_n}$ in this case.

Theorem 2.5 (Pointwise LDP for $\widehat{\Psi}_{n,h_n}$ in the case when $(\gamma_{s_n}) = \left(h_{s_n}^d \left(\sum_{k=1}^n h_{s_k}^d\right)^{-1}\right)$)
Let Assumptions $(L1)$, $(L2')$, $(L3)$ and $(L4)$ hold and assume that $g_{X,Y}$ is continuous at $(\mathbf{x}, y)$. Then, the sequence $\left(\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K) - \Psi(\mathbf{x}, f)\right)$ satisfies a LDP with speed $\left(nh_{s_n}^d\right)$ and rate function defined as follows:

$$I_{\mathbf{x}}(t) = t\left(\psi'_{\mathbf{x}}\right)^{-1}(t) - \psi_{\mathbf{x}}\left(\left(\psi'_{\mathbf{x}}\right)^{-1}(t)\right),$$

where $\psi_{\mathbf{x}}$ is defined in (1.5).

2.2 Pointwise MDP for $\widehat{\Psi}_{n,h_n}$

Let (v_n) be a positive sequence, we assume that

$(M1)$ $K : \mathbb{R}^d \to \mathbb{R}$ is a continuous, bounded function satisfying $\int_{\mathbb{R}^d} K(\mathbf{z})\, d\mathbf{z} = 1$, and, for all $j \in \{1, \ldots, d\}$, $\int_{\mathbb{R}} z_j K(\mathbf{z})\, dz_j = 0$ and $\int_{\mathbb{R}^d} z_j^2 \|K(\mathbf{z})\|\, d\mathbf{z} < \infty$.

$(M2)$ (i) $(h_{s_n}) = (cn^{-a})$ with $a \in (0, 1/d)$ and $c > 0$.
$ii)$ $(\gamma_{s_n}) = \left(n^{-1}\right)$.

$(M3)$ (i) $g_{\mathbf{X},Y}(s, t)$ is twice continuously differentiable with respect to s.
(ii) For $q \in \{0, 1, 2\}$, $s \mapsto \int_{\mathbb{R}} t^q g_{\mathbf{X},Y}(\mathbf{s}, t)\, dt$ is a bounded function continuous at $\mathbf{s} = \mathbf{x}$.
For $q \in [2, 3]$, $s \mapsto \int_{\mathbb{R}} |t|^q g_{\mathbf{X},Y}(\mathbf{s}, t)\, dt$ is a bounded function.
(iii) For $q \in \{0, 1\}$, $\int_{\mathbb{R}} |t|^q \left|\frac{\partial g_{\mathbf{X},Y}}{\partial x}(\mathbf{x}, t)\right| dt < \infty$, and $\mathbf{s} \mapsto \int_{\mathbb{R}} t^q \frac{\partial^2 g_{\mathbf{X},Y}}{\partial s^2}(\mathbf{s}, t)\, dt$ is a bounded function continuous at $\mathbf{s} = \mathbf{x}$.
(iv) For any $i, j \in \{1, \ldots, n\}$ such that $s_i \neq s_j$, the random vector $\left(X_{s_i}, X_{s_j}\right)$ and $\left(Z_{s_i}, Z_{s_j}\right)$ admit a density $f_{s_i,s_j}(\cdot)$ and $g_{s_i,s_j}(\cdot)$ such that $\sup_{s_i \neq s_j} \left\|f_{s_i,s_j}\right\| < \infty$, and $\sup_{s_i \neq s_j} \left\|g_{s_i,s_j}\right\| < \infty$.

(M4) (i) The field $(Z_{s_i})_{1 \leq i \leq n}$ is α-mixing: there exists a function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ with $\phi(t)$ goes to zero as t goes to infinity, such that for $E, F \subset \mathbb{R}^2$ with finite cardinals $Card(E)$, $Card(F)$

$$\begin{aligned} \alpha(\sigma(E), \sigma(F)) &:= \sup_{A \in \sigma(E), B \in \sigma(F)} \left|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)\right| \\ &\leq \phi(dist(E, F))\, \psi(Card(E), Card(F)), \end{aligned}$$

where $\sigma(E) = \{Z_{\mathbf{i}}, \mathbf{i} \in E\}$ and $\sigma(F) = \{Z_{\mathbf{i}}, \mathbf{i} \in F\}$, $dist(E, F)$ is the Euclidean distance between E and F and $\psi(.)$ is a positive symmetric function nondecreasing in each variable. The functions ϕ and ψ are such that $\phi(i) \leq Ci^{-\theta}$ and

$\psi(n,m) \le C \min(m,n)$.

(ii) $\sum_{k=0}^{\infty} (k+1)^2 \alpha_n^{\frac{\delta}{4+\delta}}(k) < c$ for some c, $\delta > 0$ and all n, where

$$\alpha_n(k) = \alpha_n(\mathbf{Z}, k) = \sup_{A \in \mathcal{F}_{-\infty}^n, B \in \mathcal{F}_{n+k}^{\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

with $\mathbf{Z} = \{Z_{s_i}\}_{i=1}^n$, $\mathcal{F}_m^n$ denote the σ-algebra generated by $\{Z_{s_i}\}_{i=m}^n$.

$(M5)$ For any $u \in \mathbb{R}$, $t \to \int_{\mathbb{R}} \exp(uf(y)) g_{\mathbf{X},Y}(t,y)\,dy$ is continuous at x and bounded.

$(M6)$ $\lim_{n\to\infty} v_{s_n} = \infty$ and $\lim_{n\to\infty} \gamma_{s_n} v_{s_n}^2 / h_{s_n}^d = 0$.

The following Theorem gives the pointwise MDP for $\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K)$.

Theorem 2.6 (Pointwise MDP for the recursive estimators defined by (1.2))
Let Assumptions $(M1)$–$(M5)$ hold. Then, the sequence $\left(\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K) - \Psi(\mathbf{x}, f)\right)$ satisfies a MDP with speed $\left(h_{s_n}^d / \left(\gamma_{s_n} v_{s_n}^2\right)\right)$ and rate function $J_{a,\alpha,\mathbf{x}}$ defined in (1.7).

Remark 2.7 *One can choose in (1.1), $c_f(\mathbf{x}) = 1/g_{\mathbf{X}}(\mathbf{x})$ and $d_f(\mathbf{x}) = -\mathbb{E}(f(Y) \mid \mathbf{X} = \mathbf{x})/g_{\mathbf{X}}(\mathbf{x})$ this corresponds to regression setting, see equation (3.1) in Einmahl and Mason (2000). An other trivial choice $c_f(\mathbf{x}) = 0$ and $d_f(\mathbf{x}) = 1$ correspond to the kernel density estimator. The introduction of the function $f(\cdot)$ in (1.1), is motivated by the following choices. By setting $f(y) = y$ (or $f(y) = y^k$, where k is a strictly positive integer) into (1.1) we get the recursive Nadaraya-Watson kernel regression function estimator of $m(\mathbf{x}) := \mathbb{E}(Y \mid \mathbf{X} = \mathbf{x})$. The choice $f(y) = f_t(y) = \mathbb{1}\{y \le t\}$ may be used to study the recursive kernel estimator of the conditional distribution function $F(t|\mathbf{x}) := \mathbb{P}(Y \le t | X = \mathbf{x})$. For more motivation on the use of the function $f(\cdot)$ in (1.1), one can see Remark 1.1 of Deheuvels (2011).*

Remark 2.8 *Notice that the most existing results for random fields which require certain regularity conditions on the boundary of $\mathcal{I}_{\mathbf{n}}$. Theorem 1 of El Machkouri* et al. *(2013) has the very interesting property that no condition on the boundary of $\mathcal{I}_{\mathbf{n}}$ is needed, and the central limit theorem holds under the minimal condition that the cardinal of $\mathcal{I}_{\mathbf{n}}$ tends to infinity. This is a very attractive property in spatial applications in which the underlying observation domains can be quite irregular, we can refer to El Machkouri (2014). In the paper by Lu and Tjøstheim (2014), the authors proposed nonparametric kernel estimators for both the marginal and in particular the joint probability density functions for nongridded spatial data. The asymptotic distribution of the proposed estimators are obtained under general conditions, and in particular, a new interesting framework of expanding-domain infill asymptotics is suggested to circumvent the shortcomings of spatial asymptotics in the existing literature, we refer to the last cited paper for more details and discussions. In the paper by Al-Sulami* et al. *(2017), the authors considered the semiparametric nonlinear regression allowing the sampling spatial grids can be irregular. It would be interesting to consider the extension of the present work to a more general setting for boundary of $\mathcal{I}_{\mathbf{n}}$, that requires non trivial mathematics. Another problem to be studied in the future is the characterization of the asymptotic properties of our estimators in the setting of stationary random fields without the mixing conditions.*

3 Proofs

This section is devoted to the proofs of our results. The previously defined notation continues to be used below. Throughout this section we will use the following notation:

$$\begin{aligned} Z_{s_n}(\mathbf{x}, f) &= h_{s_n}^{-d} f(Y_{s_n}) K\left(h_{s_n}^{-1}(\mathbf{x} - \mathbf{X}_{s_n})\right), \quad W_{s_n}(\mathbf{x}) = h_{s_n}^{-d} K\left(h_{s_n}^{-1}(\mathbf{x} - \mathbf{X}_{s_n})\right), \\ T_{s_n}(\mathbf{x}, f) &= h_{s_n}^{-d} \{c_f(\mathbf{x}) f(Y_{s_n}) + d_f(\mathbf{x})\} K\left(h_{s_n}^{-1}(\mathbf{x} - \mathbf{X}_{s_n})\right), \\ T_{s_n} &= \{c_f(\mathbf{x}) f(Y_{s_n}) + d_f(\mathbf{x})\} K\left(h_{s_n}^{-1}(\mathbf{x} - \mathbf{X}_{s_n})\right). \end{aligned} \tag{3.1}$$

Let us first state the following technical lemma.

Lemma 3.1 *Let $(v_{s_n}) \in \mathcal{GS}(v^*)$, $(\gamma_{s_n}) \in \mathcal{GS}(-\alpha)$, and $m > 0$ such that $m - v^*\xi > 0$ where ξ is defined in (1.6). We have*

$$\lim_{n\to+\infty} v_{s_n} \Pi_n^m \sum_{i=1}^n \Pi_i^{-m} \frac{\gamma_{s_i}}{v_{s_i}} = \frac{1}{m - v^*\xi}.$$

Moreover, for all positive sequence (α_{s_n}) such that $\lim_{n\to+\infty} \alpha_{s_n} = 0$, and all $\delta \in \mathbb{R}$,

$$\lim_{n\to+\infty} v_{s_n} \Pi_n^m \left[\sum_{i=1}^n \Pi_i^{-m} \frac{\gamma_{s_i}}{v_{s_i}} \alpha_{s_i} + \delta\right] = 0.$$

The proof is given in Mokkadem *et al.* (2009a). Lemma 3.1 is widely applied throughout the proofs. Let us underline that it is its application, which requires Assumption $(\mathbf{A2})(iii)$ on the limit of $(n\gamma_{s_n})$ as n goes to infinity. Now, we let (Φ_n) and (B_n) be the sequences defined as follows:

$$\Phi_n(\mathbf{x}, f, K) = \Pi_n \sum_{k=1}^{n} \Pi_k^{-1} \gamma_{s_k} h_{s_k}^{-d} \left(T_{s_k} - \mathbb{E}\left[T_{s_k}\right]\right), \quad B_n(\mathbf{x}, f, K) = \mathbb{E}\left[\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K)\right] - \Psi(x, f).$$

It is clear that, we have

$$\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K) - \Psi(x, f) = \Phi_n(\mathbf{x}, f, K) + B_n(\mathbf{x}, f, K). \tag{3.2}$$

We then deduce that, Theorems 2.3, 2.5 and 2.6 are consequences of (3.2) and the pointwise LDP and MDP for (Φ_n), which is given in the following propositions.

Proposition 3.2 (Pointwise LDP and MDP for (Φ_n))

1. *Under the assumptions $(L1)$ and $(L2)$, the sequence* $\left(\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K) - \mathbb{E}\left[\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K)\right]\right)$ *satisfies a LDP with speed* $\left(nh_{s_n}^d\right)$ *and rate function* $I_{a,\mathbf{x}}$.
2. *Under the assumptions $(L1)$ and $(L3)$, the sequence* $\left(\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K) - \mathbb{E}\left[\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K)\right]\right)$ *satisfies a LDP with speed* $\left(nh_{s_n}^d\right)$ *and rate function* $I_{\mathbf{x}}$.
3. *Under the assumptions $(M1) - (M6)$, the sequence $(v_n \Phi_n(\mathbf{x}, f, K))$ satisfies a LDP with speed* $\left(h_{s_n}^d / \left(\gamma_{s_n} v_n^2\right)\right)$ *and rate function* $J_{a,\alpha,\mathbf{x}}$.

The proof of the following proposition is given in Bouzebda and Slaoui (2018).

Proposition 3.3 (Pointwise convergence rate of (B_n))
Let Assumptions $(M1) - (M5)$ hold. We assume that, for all $i, j \in \{1, \ldots d\}$, $g_{ij}^{(2)}(\cdot)$ and $L_{ij}^{(2)}(\cdot)$ are continuous at $\mathbf{x}$. Then, we have

If $a \leq \alpha/(d+4)$, $B_n(\mathbf{x}, f, K) = O\left(h_{s_n}^2\right)$.

If $a > \alpha/(d+4)$, $B_n(\mathbf{x}, f, K) = o\left(\sqrt{\gamma_{s_n} h_{s_n}^{-d}}\right)$.

Set $\mathbf{x} \in \mathbb{R}^d$; since the assumptions of Theorems 2.3 and 2.5 gives that $\lim_{n\to\infty} B_n(\mathbf{x}, f, K) = 0$, Theorem 2.3 (respectively Theorem 2.5) is a consequence of the application of the first Part (respectively of the second Part) of Proposition 3.2. Moreover, under the assumptions of Theorem 2.6, the application of Proposition 3.3, $\lim_{n\to\infty} v_n B_n(\mathbf{x}, f, K) = 0$; Theorem 2.6 thus follows from the application of third Part of Proposition 3.2.

Remark 3.4 *It was shown in Bouzebda and Slaoui (2018) that under the assumptions of the previous proposition, that* $Var\left[\widehat{\Psi}_{n,h_n}(\mathbf{x}, f, K)\right] = O\left(\frac{\gamma_{s_n}}{h_{s_n}^d}\right)$. *Then, we are interested by the following quantity*

$$\Lambda_{n,\mathbf{x}}(u) = \frac{\gamma_{s_n} v_n^2}{h_{s_n}^d} \log \mathbb{E}\left[\exp\left(u \frac{h_{s_n}^d}{\gamma_{s_n} v_{s_n}} \Phi_n(\mathbf{x}, f, K)\right)\right], \quad \forall u \in \mathbb{R}.$$

Let us now state a two preliminary lemmas, which are the key of the proof of Proposition 3.2.

Lemma 3.5 *[Convergence of $\Lambda_{n,\mathbf{x}}$ in the case $v_n \equiv 1$]*
Let Assumption $(L1)$ holds and $v_n \equiv 1$. If $g_{\mathbf{X}}$ is continuous at $\mathbf{x}$, then for all $u \in \mathbb{R}$,

$$\lim_{n\to\infty} \Lambda_{n,\mathbf{x}}(u) = \begin{cases} g_{\mathbf{x}}(\mathbf{x})\left[\psi_{a,\mathbf{x}}(u) - u\left(c_f(\mathbf{x})\mathbb{E}\left(f(Y)|X = \mathbf{x}\right) + d_f(\mathbf{x})\right)\right] & \text{when} \quad (L2) \text{ holds} \\ g_{\mathbf{x}}(\mathbf{x})\left[\psi_{\mathbf{x}}(u) - u\left(c_f(\mathbf{x})\mathbb{E}\left(f(Y)|X = \mathbf{x}\right) + d_f(\mathbf{x})\right)\right] & \text{when} \quad (L2') \text{ holds} \end{cases}.$$

Lemma 3.6 *[Convergence of $\Lambda_{n,\mathbf{x}}$ in the case $v_n \to \infty$]*
Let Assumptions $(M1) - (M5)$ hold and $v_n \to \infty$. If $g_{\mathbf{X}}$ is continuous at $\mathbf{x}$, then for all $u \in \mathbb{R}$,

$$\lim_{n\to\infty} \Lambda_{n,\mathbf{x}}(u) = \frac{u^2}{2\left(2 - (\alpha - ad)\xi\right)} V(\mathbf{x}, f) R(K).$$

Our proofs are now organized as follows: Lemma 3.5 and 3.6 are proved in Section 3.1 and Proposition 3.2 in Section 3.4.

3.1 Proof of Lemmas 3.5 and 3.6.

Set $u \in \mathbb{R}$, $u_n = u/v_n$ and $a_{s_n} = h_{s_n}^d \gamma_{s_n}^{-1}$. We have:

$$\Lambda_{n,\mathbf{x}}(u) = \frac{v_n^2}{a_{s_n}} \log \mathbb{E}\left[\exp\left(u_n a_{s_n} \Phi_n(\mathbf{x}, f, K)\right)\right] = \frac{v_n^2}{a_{s_n}} \sum_{k=1}^n \log \mathbb{E}\left[\exp\left(u_n \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right)\right] - uv_n\Pi_n \sum_{k=1}^n \Pi_k^{-1} a_{s_k}^{-1} \mathbb{E}\left[T_{s_k}\right].$$

By Taylor expansion, there exists $c_{k,n}$ between 1 and $\mathbb{E}\left[\exp\left(u_n \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right)\right]$ in such a way that we have

$$\log \mathbb{E}\left[\exp\left(u_n \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_{s_k}} T_{s_k}\right)\right] = \mathbb{E}\left[\exp\left(u_n \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right) - 1\right] - \frac{1}{2c_{k,n}^2}\left(\mathbb{E}\left[\exp\left(u_n \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right) - 1\right]\right)^2,$$

and $\Lambda_{n,\mathbf{x}}$ can be rewriten as

$$\begin{aligned}\Lambda_{n,\mathbf{x}}(u) &= \frac{v_n^2}{a_{s_n}} \sum_{k=1}^n \mathbb{E}\left[\exp\left(u_n \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right) - 1\right] - \frac{v_n^2}{2a_{s_n}} \sum_{k=1}^n \frac{1}{c_{k,n}^2}\left(\mathbb{E}\left[\exp\left(u_n \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right) - 1\right]\right)^2 \\ &\quad - uv_n\Pi_n \sum_{k=1}^n \Pi_k^{-1} a_{s_k}^{-1} \mathbb{E}\left[T_{s_k}\right]. \end{aligned} \tag{3.3}$$

3.1.1 Proof of Lemma 3.5

We obtain from (3.3) that

$$\begin{aligned}\Lambda_{n,\mathbf{x}}(u) &= \frac{1}{a_{s_n}} \sum_{k=1}^n \mathbb{E}\left[\exp\left(u \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right) - 1\right] - \frac{1}{2a_{s_n}} \sum_{k=1}^n \frac{1}{c_{k,n}^2}\left(\mathbb{E}\left[\exp\left(u \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right) - 1\right]\right)^2 - u\Pi_n \sum_{k=1}^n \Pi_k^{-1} a_{s_k}^{-1} \mathbb{E}\left[T_{s_k}\right] \\ &= \frac{1}{a_{s_n}} \sum_{k=1}^n h_{s_k}^d \left[\int_{\mathbb{R}^d \times \mathbb{R}} \left(\exp\left(uV_{n,k}\left(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x})\right) K(\mathbf{z})\right) - 1\right)\right. \\ &\quad \left. - uV_{n,k}\left(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x})\right) K(\mathbf{z})\right] g_{\mathbf{X},Y}(\mathbf{x}, y)\, d\mathbf{z} dy - R_{n,\mathbf{x}}^{(1)}(u) + R_{n,\mathbf{x}}^{(1)}(u), \end{aligned} \tag{3.4}$$

with

$$\begin{aligned} V_{n,k} &= \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k}, \quad R_{n,\mathbf{x}}^{(1)}(u) = \frac{1}{2a_{s_n}} \sum_{k=1}^n \frac{1}{c_{k,n}^2}\left(\mathbb{E}\left[\exp\left(u \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right) - 1\right]\right)^2 \\ R_{n,\mathbf{x}}^{(2)}(u) &= \frac{1}{a_{s_n}} \sum_{k=1}^n h_{s_k}^d \int_{\mathbb{R}^d \times \mathbb{R}} \left[\exp\left(u \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k}\left(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x})\right) K(\mathbf{z})\right) - 1\right] \left[g_{\mathbf{X},Y}(\mathbf{x} - \mathbf{z}h_{s_k}, y) - g_{\mathbf{X},Y}(\mathbf{x}, y)\right] d\mathbf{z} dy \\ &\quad - u\Pi_n \sum_{k=1}^n \Pi_k^{-1}\gamma_{s_k} \int_{\mathbb{R}^d \times \mathbb{R}} \left(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x})\right) K(\mathbf{z}) \left[g_{\mathbf{X},Y}(\mathbf{x} - \mathbf{z}h_{s_k}, y) - g_{\mathbf{X},Y}(\mathbf{x}, y)\right] d\mathbf{z} dy. \end{aligned}$$

Moreover, it follows from (3.10), that $\lim_{n\to\infty} \left|R_{n,\mathbf{x}}^{(1)}(u)\right| = 0$. Now, since $|e^t - 1| \leq |t|\, e^{|t|}$, we have

$$\begin{aligned}\left|R_{n,\mathbf{x}}^{(2)}(u)\right| &\leq |u| \left(e^{|u||c_f(\mathbf{x}) f(y) + d_f(\mathbf{x})|\|K\|_\infty} + 1\right) \Pi_n \sum_{k=1}^n \Pi_k^{-1}\gamma_{s_k} \int_{\mathbb{R}^d} |K(\mathbf{z})| \left|c_f(\mathbf{x}) f(y) + d_f(\mathbf{x})\right| \\ &\quad \times \left|g_{\mathbf{X},Y}(\mathbf{x} - \mathbf{z}h_{s_k}, y) - g_{\mathbf{X},Y}(\mathbf{x}, y)\right| d\mathbf{z} dy. \end{aligned}$$

Since, Lemma 3.1 ensures that, the sequence $\left(\Pi_n \sum_{k=1}^n \Pi_k^{-1}\gamma_{s_k}\right)$ is bounded, then, the dominated convergence theorem ensures that $\lim_{n\to\infty} R_{n,\mathbf{x}}^{(2)}(u) = 0$. Then, it follows from (3.4), that

$$\begin{aligned}\lim_{n\to\infty} \Lambda_{n,\mathbf{x}}(u) &= \lim_{n\to\infty} \frac{\gamma_{s_n}}{h_{s_n}^d} \sum_{k=1}^n h_{s_k}^d \int_{\mathbb{R}^d \times \mathbb{R}} \left[\left(\exp\left(uV_{n,k}\left(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x})\right) K(\mathbf{z})\right) - 1\right)\right. \\ &\quad \left. - uV_{n,k}\left(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x})\right) K(\mathbf{z})\right] g_{\mathbf{X},Y}(\mathbf{x}, y)\, d\mathbf{z} dy. \end{aligned} \tag{3.5}$$

In the case when $(\gamma_{s_n}) = (n^{-1})$
We have, see Slaoui (2013, 2015a)

$$\frac{\Pi_n}{\Pi_k} = \prod_{j=k+1}^{n}\left(1-\gamma_{s_j}\right) = \frac{k}{n}, \quad \text{then, } V_{n,k} = \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} = \left(\frac{k}{n}\right)^{ad}.$$

Consequently, it follows from (3.5) and from some analysis considerations that

$$\begin{aligned}\lim_{n\to\infty}\Lambda_{n,\mathbf{x}}(u) &= \int_{\mathbb{R}^d\times\mathbb{R}}\left[\int_0^1 t^{-ad}\exp\left(ut^{ad}\left(c_f(\mathbf{x})f(y)+d_f(\mathbf{x})\right)K(\mathbf{z})\right)-1\right.\\&\quad\left.-ut^{ad}\left(c_f(\mathbf{x})f(y)+d_f(\mathbf{x})\right)K(\mathbf{z})\right]g_{\mathbf{X},Y}(\mathbf{x},y)\,dtd\mathbf{z}dy\\&= g_{\mathbf{x}}(\mathbf{x})\left[\psi_{a,\mathbf{x}}(u)-u\left(c_f(\mathbf{x})\mathbb{E}\left(f(Y)|X=\mathbf{x}\right)+d_f(\mathbf{x})\right)\right].\end{aligned}$$

In the case when $(\gamma_{s_n}) = \left(h_{s_n}^d\left(\sum_{k=1}^n h_{s_k}^d\right)^{-1}\right)$
We have, see Slaoui (2013, 2015a), $\frac{\Pi_n}{\Pi_k} = \prod_{j=k+1}^n\left(1-\gamma_{s_j}\right) = \frac{\gamma_{s_n}}{\gamma_{s_k}}\frac{h_{s_k}^d}{h_{s_n}^d}$, then, $V_{n,k}=1$. Consequently, it follows from (3.5) that

$$\begin{aligned}\lim_{n\to\infty}\Lambda_{n,\mathbf{x}}(u) &= \int_{\mathbb{R}^d\times\mathbb{R}}\left[\left(\exp\left(u\left(c_f(\mathbf{x})f(y)+d_f(\mathbf{x})\right)K(\mathbf{z})\right)-1\right)-u\left(c_f(\mathbf{x})f(y)+d_f(\mathbf{x})\right)K(\mathbf{z})\right]g_{\mathbf{X},Y}(\mathbf{x},y)\,dtd\mathbf{z}dy\\&= g_{\mathbf{x}}(\mathbf{x})\left[\psi_{\mathbf{x}}(u)-u\left(c_f(\mathbf{x})\mathbb{E}\left(f(Y)|X=\mathbf{x}\right)+d_f(\mathbf{x})\right)\right],\end{aligned}$$

and thus Lemma 3.5 is proved. □

3.1.2 Proof of Lemma 3.6

A Taylor's expansion implies the existence of $c'_{k,n}$ between 0 and $u_n\frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k}T_{s_k}$ such that

$$\mathbb{E}\left[\exp\left(u_n\frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k}T_{s_k}\right)-1\right] = u_n\frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k}\mathbb{E}\left[T_{s_k}\right]+\frac{1}{2}\left(u_n\frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k}\right)^2\mathbb{E}\left[T_{s_k}^2\right]+\frac{1}{6}\left(u_n\frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k}\right)^3\mathbb{E}\left[T_{s_k}^3e^{c'_{k,n}}\right].$$

Therefore, we infer that

$$\Lambda_{n,\mathbf{x}}(u) = \frac{1}{2}V(\mathbf{x},f)\,u^2a_{s_n}\Pi_n^2\sum_{k=1}^n\Pi_k^{-2}a_{s_k}^{-1}\gamma_{s_k}\int_{\mathbb{R}^d}K^2(\mathbf{z})\,d\mathbf{z}+R_{n,\mathbf{x}}^{(3)}(u)-R_{n,\mathbf{x}}^{(4)}(u)+R_{n,\mathbf{x}}^{(5)}(u), \tag{3.6}$$

with

$$\begin{aligned}R_{n,\mathbf{x}}^{(3)}(u) &= \frac{1}{2}u^2c_f^2(\mathbf{x})\,a_{s_n}\Pi_n^2\sum_{k=1}^n\Pi_k^{-2}a_{s_k}^{-1}\gamma_{s_k}\int_{\mathbb{R}^d}K^2(\mathbf{z})\left[L\left(\mathbf{x}-\mathbf{z}h_{s_k},f^2\right)-L\left(\mathbf{x},f^2\right)\right]d\mathbf{z}\\&\quad+\frac{1}{2}u^2d_f^2(\mathbf{x})\,a_{s_n}\Pi_n^2\sum_{k=1}^n\Pi_k^{-2}a_{s_k}^{-1}\gamma_{s_k}\int_{\mathbb{R}^d}K^2(\mathbf{z})\left[g_{\mathbf{x}}\left(\mathbf{x}-\mathbf{z}h_{s_k}\right)-g_{\mathbf{x}}(\mathbf{x})\right]d\mathbf{z}\\&\quad+c_f(\mathbf{x})\,d_f(\mathbf{x})\,u^2a_{s_n}\Pi_n^2\sum_{k=1}^n\Pi_k^{-2}a_{s_k}^{-1}\gamma_{s_k}\int_{\mathbb{R}^d}K^2(\mathbf{z})\left[L\left(\mathbf{x}-\mathbf{z}h_{s_k},f\right)-L(\mathbf{x},f)\right]d\mathbf{z},\\R_{n,\mathbf{x}}^{(4)}(u) &= \frac{1}{2}u^2a_{s_n}\Pi_n^2\sum_{k=1}^n\Pi_k^{-2}a_{s_k}^{-1}\gamma_{s_k}h_{s_k}^d\left(\int_{\mathbb{R}^d}K(\mathbf{z})\left(c_f(\mathbf{x})\,L\left(\mathbf{x}-\mathbf{z}h_{s_k},f\right)+d_f(\mathbf{x})\,g_{\mathbf{x}}\left(\mathbf{x}-\mathbf{z}h_{s_k}\right)\right)dz\right)^2.\\R_{n,\mathbf{x}}^{(5)}(u) &= \frac{1}{6}\frac{u^3}{v_n}a_{s_n}^2\Pi_n^3\sum_{k=1}^n\Pi_k^{-3}a_{s_k}^{-3}\mathbb{E}\left[T_{s_k}^3e^{c'_{k,n}}\right]-\frac{v_n^2}{2a_{s_n}}\sum_{k=1}^n\frac{1}{c_{k,n}^2}\left(\mathbb{E}\left[\exp\left(u_n\frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k}T_{s_k}\right)-1\right]\right)^2.\end{aligned}$$

In view of the assumption $(A3)$, we have

$$\lim_{k\to\infty}\left|L\left(\mathbf{x}-\mathbf{z}h_{s_k},f^2\right)-L\left(\mathbf{x},f^2\right)\right|=0,\quad \lim_{k\to\infty}\left|g_{\mathbf{x}}\left(\mathbf{x}-\mathbf{z}h_{s_k}\right)-g_{\mathbf{x}}(\mathbf{x})\right|=0,\quad \lim_{k\to\infty}\left|L\left(\mathbf{x}-\mathbf{z}h_{s_k},f\right)-L(\mathbf{x},f)\right|=0.$$

Thus a straightforward application of Lebesgue dominated convergence theorem in connection with condition $(M1)$ implies that

$$\begin{aligned}\lim_{k\to\infty}\int_{\mathbb{R}^d}K^2(\mathbf{z})\left|L\left(\mathbf{x}-\mathbf{z}h_{s_k},f^2\right)-L\left(\mathbf{x},f^2\right)\right|d\mathbf{z} &= 0,\quad \lim_{k\to\infty}\int_{\mathbb{R}^d}K^2(\mathbf{z})\left|g_{\mathbf{x}}\left(\mathbf{x}-\mathbf{z}h_{s_k}\right)-g_{\mathbf{x}}(\mathbf{x})\right|d\mathbf{z}=0,\\\lim_{k\to\infty}\int_{\mathbb{R}^d}K^2(\mathbf{z})\left|L\left(\mathbf{x}-\mathbf{z}h_{s_k},f\right)-L(\mathbf{x},f)\right|d\mathbf{z} &= 0.\end{aligned}$$

Morever, since $(a_{s_n}) \in \mathcal{GS}(\alpha - ad)$, and $\lim_{n\to\infty}(n\gamma_{s_n}) > (\alpha - ad)/2$. The application of Lemma 3.1 ensures that

$$a_{s_n}\Pi_n^2 \sum_{k=1}^n \Pi_k^{-2} a_{s_k}^{-1}\gamma_{s_k} = \frac{1}{(2-(\alpha-ad)\xi)} + o(1). \tag{3.7}$$

Then, we have $\lim_{n\to\infty}\left|R_{n,\mathbf{x}}^{(3)}(u)\right| = 0$. Moreover, since L and $g_{\mathbf{X}}$ are bounded and in view of Lemma 3.1, we have

$$a_{s_n}\Pi_n^2 \sum_{k=1}^n \Pi_k^{-2} a_{s_k}^{-1}\gamma_{s_k} h_{s_k}^d = O\left(h_{s_n}^d\right).$$

Then, we have $\lim_{n\to\infty}\left|R_{n,\mathbf{x}}^{(4)}(u)\right| = 0$. Moreover, in view of (3.1), we have $|T_{s_k}| \leq |c_f(\mathbf{x}) f(Y) + d_f(\mathbf{x})| \, \|K\|_\infty$, then

$$c'_{k,n} \leq \left|u_n \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right| \leq |u_n| \, |c_f(\mathbf{x}) f(Y) + d_f(\mathbf{x})| \, \|K\|_\infty . \tag{3.8}$$

Since, we have

$$\mathbb{E}|T_{s_k}|^3 \leq h_{s_k}^d \left|c_f^3(\mathbf{x}) L\left(\mathbf{x}, f^3\right) + 3c_f^2(\mathbf{x}) d_f(\mathbf{x}) L\left(\mathbf{x}, f^2\right) + 3c_f(\mathbf{x}) d_f^2(\mathbf{x}) L(\mathbf{x}, f) + d_f^3(\mathbf{x}) g_{\mathbf{X}}(\mathbf{x})\right| \int_{\mathbb{R}^d} \left|K^3(\mathbf{z})\right| d\mathbf{z}.$$

It follows from, Lemma 3.1 and (3.8), that, there exists a positive constant c_1 such that, for n large enough,

$$\begin{aligned} \left|\frac{u^3}{v_n} a_{s_n}^2 \Pi_n^3 \sum_{k=1}^n \Pi_k^{-3} a_{s_k}^{-3} \mathbb{E}\left[T_{s_k}^3 e^{c'_{k,n}}\right]\right| &\leq c_1 e^{|u_n||c_f(\mathbf{x})f(Y)+d_f(\mathbf{x})|\|K\|_\infty} \frac{u^3}{v_n} \\ & \left|c_f^3(\mathbf{x}) L\left(\mathbf{x}, f^3\right) + 3c_f^2(\mathbf{x}) d_f(\mathbf{x}) L\left(\mathbf{x}, f^2\right) + 3c_f(\mathbf{x}) d_f^2(\mathbf{x}) L(\mathbf{x}, f) + d_f^3(\mathbf{x}) g_{\mathbf{X}}(\mathbf{x})\right| \int_{\mathbb{R}^d} \left|K^3(\mathbf{z})\right| d\mathbf{z}, \end{aligned} \tag{3.9}$$

which goes to 0 as $n \to \infty$. Furthermore, Lemma 3.1 ensures that

$$\begin{aligned} \left|\frac{v_n^2}{2a_{s_n}} \sum_{k=1}^n \frac{1}{c_{k,n}^2}\left(\mathbb{E}\left[\exp\left(u_n \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right) - 1\right]\right)^2\right| &\leq \frac{v_n^2}{2a_{s_n}} \sum_{k=1}^n \left(\mathbb{E}\left[\exp\left(u_n \frac{a_{s_n}\Pi_n}{a_{s_k}\Pi_k} T_{s_k}\right) - 1\right]\right)^2 \\ &= \frac{u^2}{2} V(\mathbf{x}, f) R(K) 2a_{s_n}\Pi_n^2 \sum_{k=1}^n \Pi_k^{-2} a_{s_k}^{-1}\gamma_{s_k} h_{s_k}^d + o\left(a_{s_n}\Pi_n^2 \sum_{k=1}^n \Pi_k^{-2} a_{s_k}^{-1}\gamma_{s_k} h_{s_k}^d\right) = o(1). \end{aligned} \tag{3.10}$$

The combination of (3.9) and (3.10) ensures that $\lim_{n\to\infty}\left|R_{n,\mathbf{x}}^{(5)}(u)\right| = 0$. Then, it follows from (3.6) and (3.7), that $\lim_{n\to\infty}\Lambda_{n,\mathbf{x}}(u) = \frac{u^2}{2(2-(\alpha-ad)\xi)} V(\mathbf{x}, f) R(K)$. Lemma 3.6 is proved. □

3.2 Proof of Proposition 2.2

- Since $\left|e^t - 1\right| \leq |t| e^{|t|}$, it follows from (2.3) and $(L1)$, that

$$|\psi_{a,\mathbf{x}}(u)| \leq \int_{\mathbb{R}^d\times\mathbb{R}\times[0,1]} g_{\mathbf{X}}^{-1}(x) \left|\exp\left(ut^{ad} K(\mathbf{z})(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x}))\right) - 1\right| g_{\mathbf{X},Y}(x,y)\, d\mathbf{z}dydt < \infty$$

  which ensures the existence of $\psi_{a,\mathbf{x}}$. It is straightforward to check that $\psi_{a,\mathbf{x}}$ is twice differentiable, with

$$\begin{aligned} \psi'_{a,\mathbf{x}}(u) &= g_{\mathbf{X}}^{-1}(\mathbf{x}) \int_{[0,1]\times\mathbb{R}^d\times\mathbb{R}} t^{ad} K(\mathbf{z})(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x})) \\ &\quad \times \exp\left(ut^{ad} K(\mathbf{z})(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x}))\right) g_{\mathbf{X},Y}(\mathbf{x},y)\, dtd\mathbf{z}dy \\ \psi''_{a,\mathbf{x}}(u) &= g_{\mathbf{X}}^{-1}(\mathbf{x}) \int_{[0,1]\times\mathbb{R}^d\times\mathbb{R}} t^{2ad}(K(\mathbf{z}))^2 (c_f(\mathbf{x}) f(y) + d_f(\mathbf{x}))^2 \\ &\quad \times \exp\left(ut^{ad} K(\mathbf{z})(c_f(\mathbf{x}) f(y) + d_f(\mathbf{x}))\right) g_{\mathbf{X},Y}(\mathbf{x},y)\, dtd\mathbf{z}dy. \end{aligned}$$

  Since $\psi''_{a,\mathbf{x}}(u) > 0\, \forall u \in \mathbb{R}$, $\psi'_{a,\mathbf{x}}$ is increasing on $\mathbb{R}$, and $\psi_{a,\mathbf{x}}$ is strictly convex on $\mathbb{R}$. It follows that its Cramer transform $I_{a,\mathbf{x}}$ is a good rate function on $\mathbb{R}$ (see Dembo and Zeitouni (2010)) and (i) of Proposition 2.2 is proved.

- Let us now assume that $\lambda(O_-) = 0$. We then have $\lim_{u\to-\infty}\psi'_{a,\mathbf{x}}(u) = 0$ and $\lim_{u\to+\infty}\psi'_{a,\mathbf{x}}(u) = +\infty$ so that the range of $\psi'_{a,\mathbf{x}}$ is $(0,+\infty)$. Moreover

$$\lim_{u\to-\infty}\psi_{a,\mathbf{x}}(u) = \begin{cases} -\lambda(S_+) & \text{if} \quad \lambda(S_+\cap T_+) > 0 \\ -\lambda(S_-) & \text{if} \quad \lambda(S_-\cap T_-) > 0 \end{cases}$$

(which can be $-\infty$). This implies in particular that

$$I_{a,\mathbf{x}}(0) = \begin{cases} \lambda(S_+) & \text{if} \quad \lambda(S_+\cap T_+) > 0 \\ \lambda(S_-) & \text{if} \quad \lambda(S_-\cap T_-) > 0 \end{cases}$$

Now, when $t < 0$, $\lim_{u\to-\infty}(ut - \psi_{a,\mathbf{x}}(u)) = +\infty$ and $I_{a,\mathbf{x}}(t) = +\infty$. Since $\psi'_{a,\mathbf{x}}$ is increasing with range $(0,+\infty)$, when $t > 0$, $\sup_u(ut - \psi_{a,\mathbf{x}}(u))$ is reached for $u_0(t)$ such that $\psi_{a,\mathbf{x}}(u_0(t)) = t$, i.e., for $u_0(t) = (\psi'_{a,\mathbf{x}})^{-1}(t)$; this prove (2.4). (Note that, since $\psi''_{a,\mathbf{x}}(t) > 0$, the function $t \mapsto u_0(t)$ is differentiable on $(0,+\infty)$). Now, differentiating (2.4), we have$I'_{a,\mathbf{x}}(t) = u_0(t) + tu'_0(t) - \psi'_{a,\mathbf{x}}(u_0(t))u'_0(t) = (\psi'_{a,\mathbf{x}})^{-1}(t) + tu'_0(t) - tu'_0(t) = (\psi'_{a,\mathbf{x}})^{-1}(t)$. Since $(\psi'_{a,\mathbf{x}})^{-1}$ is an increasing function on $(0,+\infty)$, it follows that $I_{a,\mathbf{x}}$ is strictly convex on $(0,+\infty)$ (and differentiable). Thus (ii) is proved.

- We assume that $\lambda(O_-) > 0$. In this case, $\psi'_{a,\mathbf{x}}$ can be rewritten as

$$\begin{aligned} \psi'_{a,\mathbf{x}}(u) = & \int_{[0,1]\times(\mathbb{R}^{d+1}\cap O_+)} t^{ad}K(\mathbf{z})\left((c_f(\mathbf{x})f(y) + d_f(\mathbf{x})\right)) \\ & \times\exp\left(ut^{ad}K(z)\left(c_f(\mathbf{x})f(y) + d_f(\mathbf{x})\right)\right)g_{\mathbf{X},Y}(\mathbf{x},y)\,dtd\mathbf{z}dy \\ & + \int_{[0,1]\times(\mathbb{R}^{d+1}\cap O_-)} t^{ad}K(\mathbf{z})\left(c_f(\mathbf{x})f(y) + d_f(\mathbf{x})\right) \\ & \times\exp\left(ut^{ad}K(\mathbf{z})\left(c_f(\mathbf{x})f(y) + d_f(\mathbf{x})\right)\right)g_{\mathbf{X},Y}(\mathbf{x},y)\,dtd\mathbf{z}dy \end{aligned}$$

and we have $\lim_{u\to-\infty}\psi'_{a,\mathbf{x}}(u) = -\infty$ and $\lim_{u\to+\infty}\psi'_{a,\mathbf{x}}(u) = +\infty$ so that the range of $\psi'_{a,\mathbf{x}}$ is $\mathbb{R}$ in this case. The proof of (iii) follows the same lines as previously, except that, in the present case, $(\psi'_{a,\mathbf{x}})^{-1}$ is defined on $\mathbb{R}$, and not only on $(0,+\infty)$.

□

3.3 Proof of Proposition 2.4

The proof of this Proposition can be found in Bouzebda and Slaoui (2019c) and is available upon request. □

3.4 Proof of Proposition 3.2

To prove Proposition 3.2, we apply similar result as the one given by Proposition 1 Mokkadem *et al.* (2006) in the non spatial case, Lemmas 3.5 and 3.6 and the following result (see Puhalskii (1994)).

Lemma 3.7 *Let (W_n) be a sequence of real random variables, (ν_n) a positive sequence satisfying $\lim_{n\to\infty}\nu_n = +\infty$, and suppose that there exists some convex non-negative function Γ defined on $\mathbb{R}$ such that $\forall u \in \mathbb{R}$, $\lim_{n\to\infty}\frac{1}{\nu_n}\log\mathbb{E}\left[\exp(u\nu_n W_n)\right] = \Gamma(u)$. If the Legendre function Γ^* of Γ is a strictly convex function, then the sequence (W_n) satisfies a* LDP *of speed (ν_n) and good rate fonction Γ^*.*

In our framework, when $v_n \equiv 1$ and $\gamma_{s_n} = n^{-1}$, we take $W_n = \widehat{\Psi}_{n,h_n}(\mathbf{x},f,K) - \mathbb{E}\left(\widehat{\Psi}_{n,h_n}(\mathbf{x},f,K)\right)$, $\nu_n = nh^d_{s_n}$ with $h_{s_n} = cn^{-a}$ where $a \in\,]0,1/d[$ and $\Gamma = \Lambda^{L,1}_{\mathbf{x}}$. In this case, the Legendre transform of $\Gamma = \Lambda^{L,1}_{\mathbf{x}}$ is the rate function $I_{a,\mathbf{x}}$ defined in (1.3) which is strictly convex by Proposition 2.2. Farther, when $v_n \equiv 1$ and $\gamma_{s_n} = h^d_{s_n}\left(\sum_{k=1}^n h^d_{s_k}\right)^{-1}$, we take $W_n = \widehat{\Psi}_{n,h_n}(\mathbf{x},f,K) - \mathbb{E}\left(\widehat{\Psi}_{n,h_n}(\mathbf{x}),f,K\right)$, $\nu_n = nh^d_{s_n}$ with $h_{s_n} \in \mathcal{GS}(-a)$ where $a \in\,]0,1/d[$ and $\Gamma = \Lambda^{L,2}_{\mathbf{x}}$. In this case, the Legendre transform of $\Gamma = \Lambda^{L,2}_{\mathbf{x}}$ is the rate function $I_{\mathbf{x}}$ which is st rictly convex by Proposition 2.4. Otherwise, when, $v_n \to \infty$, we take $W_n = v_n\left(\widehat{\Psi}_{n,h_n}(\mathbf{x},f,K) - \mathbb{E}\left(\widehat{\Psi}_{n,h_n}(\mathbf{x},f,K)\right)\right)$, $\nu_n = h^d_{s_n}/\left(\gamma_{s_n}v_n^2\right)$ and $\Gamma = \Lambda^M_{\mathbf{x}}$; Γ^* is then the quadratic rate function $J_{a,\alpha,\mathbf{x}}$ defined in (1.7) and thus Proposition 3.2 follows. □

Funding

This work benefited from the financial support of the GDR 3477 GeoSto.

Acknowledgement

The authors are indebted to the Editor-in-Chief, the Associate Editor and the referee for their very valuable comments and suggestions which led to a considerable improvement of the manuscript.

References

Al-Sulami, D., Jiang, Z., Lu, Z., and Zhu, J. (2017). Estimation for semiparametric nonlinear regression of irregularly located spatial time-series data. *Econom. Stat.*, **2**, 22–35.

Amiri, A., Dabo-Niang, S., and Yahaya, M. (2016). Nonparametric recursive density estimation for spatial data. *C. R. Math. Acad. Sci. Paris*, **354**(2), 205–210.

Biau, G. and Cadre, B. (2004). Nonparametric spatial prediction. *Stat. Inference Stoch. Process.*, **7**(3), 327–349.

Bojanic, R. and Seneta, E. (1973). A unified theory of regularly varying sequences. *Math. Z.*, **134**, 91–106.

Bouzebda, S. and Slaoui, Y. (2018). Nonparametric recursive method for kernel-type function estimators for spatial data. *Statist. Probab. Lett.*, **139**, 103–114.

Bouzebda, S. and Slaoui, Y. (2019a). Bandwidth selector for nonparametric recursive density estimation for spatial data defined by stochastic approximation method. *Comm. Statist. Theory Methods*, **48, To appear**. DOI:10.1080/03610926.2019.1584313

Bouzebda, S. and Slaoui, Y. (2019b). Large and moderate deviation principles for recursive kernel estimators for spatial data defined by stochastic approximation method with application to testing symmetry. Submitted.

Bouzebda, S. and Slaoui, Y. (2019c). Large and moderate deviation principles for recursive kernel estimators of a regression function for spatial data defined by stochastic approximation method. Preprint

Carbon, M., Tran, L. T., and Wu, B. (1997). Kernel density estimation for random fields (density estimation for random fields). *Statist. Probab. Lett.*, **36**(2), 115–125.

Cressie, N. A. C. (2015). *Statistics for spatial data*. Wiley Classics Library. John Wiley & Sons, Inc., New York, revised edition. Paperback edition of the 1993 edition [MR1239641].

Dabo-Niang, S. and Yao, A.-F. (2013). Kernel spatial density estimation in infinite dimension space. *Metrika*, **76**(1), 19–52.

Dabo-Niang, S., Ternynck, C., and Yao, A.-F. (2016). Nonparametric prediction of spatial multivariate data. *J. Nonparametr. Stat.*, **28**(2), 428–458.

Deheuvels, P. (2011). One bootstrap suffices to generate sharp uniform bounds in functional estimation. *Kybernetika (Prague)*, **47**(6), 855–865.

Dembo, A. and Zeitouni, O. (2010). *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin. Corrected reprint of the second (1998) edition.

Einmahl, U. and Mason, D. M. (2000). An empirical process approach to the uniform consistency of kernel-type function estimators. *J. Theoret. Probab.*, **13**(1), 1–37.

Einmahl, U. and Mason, D. M. (2005). Uniform in bandwidth consistency of kernel-type function estimators. *Ann. Statist.*, **33**(3), 1380–1403.

Ekström, M. (2014). A general central limit theorem limit theorem for strong mixing sequences. *Statist. Probab. Lett.*, **46**, 195–214.

El Machkouri, M. (2014). Kernel density estimation for stationary random fields. *ALEA Lat. Am. J. Probab. Math. Stat.*, **11**(1), 259–279.

El Machkouri, M., Volný, D., and Wu, W. B. (2013). A central limit theorem for stationary random fields. *Stochastic Process. Appl.*, **123**(1), 1–14.

Galambos, J. and Seneta, E. (1973). Regularly varying sequences. *Proc. Amer. Math. Soc.*, **41**, 110–116.

Guyon, X. (1995). *Random fields on a network*. Probability and its Applications (New York). Springer-Verlag, New York. Modeling, statistics, and applications, Translated from the 1992 French original by Carenne Ludeña.

Lu, Z. and Tjøstheim, D. (2014). Nonparametric estimation of probability density functions for irregularly observed spatial data. *J. Amer. Statist. Assoc.*, **109**(508), 1546–1564.

Mokkadem, A. and Pelletier, M. (2007). A companion for the Kiefer-Wolfowitz-Blum stochastic approximation algorithm. *Ann. Statist.*, **35**(4), 1749–1772.

Mokkadem, A., Pelletier, M., and Thiam, B. (2006). Large and moderate deviations principles for recursive kernel estimator of a multivariate density and its partial derivatives. *Serdica Math. J.*, **32**(4), 323–354.

Mokkadem, A., Pelletier, M., and Thiam, B. (2008). Large and moderate deviations principles for kernel estimators of the multivariate regression. *Math. Methods Statist.*, **17**(2), 146–172.

Mokkadem, A. Pelletier, M. and Slaoui, Y. (2009a). The stochastic approximation method for the estimation of a multivariate probability density. *J. Statist. Plann. Inference*, **139**, 2459–2478.

Mokkadem, A. Pelletier, M. and Slaoui, Y. (2009b). Revisiting Révész's stochastic approximation method for the estimation of a regression function. *ALEA Lat. Am. J. Probab. Math. Stat.*, **6**, 63–114.

Puhalskii, A. (1994). The method of stochastic exponentials for large deviations. *Stochastic Process. Appl.*, **54**(1), 45–70.

Ripley, B. D. (1981). *Spatial statistics*. John Wiley & Sons, Inc., New York. Wiley Series in Probability and Mathematical Statistics.

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U. S. A.*, **42**, 43–47.

Rosenblatt, M. (1985). *Stationary sequences and random fields*. Birkhäuser Boston, Inc., Boston, MA.

Slaoui, Y. (2013). Large and moderate deviation principles for recursive kernel density estimators defined by stochastic approximation method. *Serdica Math. J.*, **39**(1), 53–82.

Slaoui, Y. (2014a). Bandwidth selection for recursive kernel density estimators defined by stochastic approximation method. *J. Probab. Stat.*, **2014**, ID 739640, doi:10.1155/2014/739640.

Slaoui, Y. (2014b). The stochastic approximation method for the estimation of a distribution function. *Math. Methods Statist.* **23**, 306–325.

Slaoui, Y. (2015a). Large and moderate deviation principles for averaged stochastic approximation method for the estimation of a regression function. *Serdica Math. J.*, **41**(2-3), 307–328.

Slaoui, Y. (2015b). Plug-In Bandwidth selector for recursive kernel regression estimators defined by stochastic approximation method. *Stat. Neerl.*, **69**, 483–509.

Slaoui, Y. (2016). Optimal bandwidth selection for semi-recursive kernel regression estimators. *Stat. Interface*, **9**, 375–388.

Tran, L. T. (1990). Kernel density estimation on random fields. *J. Multivariate Anal.*, **34**(1), 37–53.

Tran, L. T. and Yakowitz, S. (1993). Nearest neighbor estimators for random fields. *J. Multivariate Anal.*, **44**(1), 23–46.